OUKA
Osaka University Knowledge Archive

Title	Substrate Recognition Mechanism of Molecular Chaperone GroEL
Author(s)	星野, 大
Citation	大阪大学, 1998, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3143767
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

Substrate Recognition Mechanism of Molecular Chaperone GroEL

Masaru Hoshino

Department of Biology, Graduate School of Science,
Osaka University

1998

Contents

Abbreveations

α-MF	α-mating factor derived from *Saccharomyces cerevisiae*
AOX1	alcohol oxidase 1
β-LG	bovine β-lactoglobulin A
B-factor	temperature factor
BMGY	buffered glycerol-complex medium
BMMY	buffered methanol-complex medium
BrCN	cyanogen bromide
CD	circular dichroism
COSY	correlation spectroscopy
CSI	chemical shift index
cyt *c*	cytochrome *c*
DACM	*N*-(7-dimethylamino-4-methyl-3-coumarinyl)-maleimide
DACM-apo-cyt *c*	DACM-labeled apo-cytochrome *c*
ΔC_α	C_α chemical shift differences from random coil chemical shifts
ΔCO	CO chemical shift differences from random coil chemical shifts
ΔG_U	free energy change of unfolding
ΔH_α	H_α chemical shift differences from random coil chemical shifts
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
fragment 1-65	cytochrome *c* fragment from 1 to 65 residues
fragment 1-38	cytochrome *c* fragment from 1 to 38 residues
fragment 11-22	cytochrome *c* fragment from 11 to 22 residues
Gdn-HCl	guanidine hydrochloride
HPLC	high perfomance liquid chromatography
Hsp	heat shock protein
HSQC	hetelonuclear single quantum coherence
MD	minimal dextrose medium
MG	molten globule
MM	minimal methanol medium
Mut	methanol utilization
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
P. pastoris	*Pichia pastoris*
PCR	polymerase chain reaction
pI	isoelectric point

$[\theta]$	molar residue ellipticity
Rh-β-LG	tetramethylrhodamine-labeled β-lactoglobulin
Rh-β-LG.red	disulfied-reduced tetramethylrhodamine-labeled β-lactoglobulin
Rubisco	ribulose-bisphosphate carboxylase
S. cerevisiae	*Saccharomyces cerevisiae*
S^2	order parameter
SDS-PAGE	sodium lauryl sulfate polyacrylamide gel electrophoreces
T_1	longitudinal relaxation time
T_2	transverse relaxation time
τ_e	effective correlation times
TFE	2,2,2-trifluoloethanol
τ_M	overall tumbling motion
TOCSY	total correlation spectroscopy
Tris	tris(hydroxymethyl)aminomethane
TRiC	TCP-1 ring complex
UV	ultra violet

Abstract

Protein folding, the process by which a protein acquires its functionally active form (native structure), has been considered to be a spontaneous process. However, recent studies indicate that the protein folding in the living cell might be facilitated by a group of proteins called molecular chaperones. Molecular chaperones have been revealed to exist in everywhere in the cell, and are thought to recognize a nonnative form of other proteins to facilitate their folding.

The binding specificity of molecular chaperones for their substrates is extremely low, and the mechanism of their substrate recognition is not elucidated. In order to clarify this point, I examined the interaction between GroEL, a molecular chaperone derived form *Escherichia coli*, and horse cytochrome *c* and bovine β-lactoglobulin. By modifying the covalently bound heme group, I prepared several kinds of cytochrome *c* derivatives with different conformational properties. I also prepared fluorescence-labeled β-lactoglobulin as well as its disulfide-bond reduced derivative. By examining the interaction between these substrate proteins and GroEL, I found that the binding affinity is correlated with the extent of solvent-exposed fluctuating hydrophobic clusters. I also found that electrostatic interaction between GroEL and the substrates is also an important factor influencing the binding affinity. Kinetic analysis of interaction indicated that the refolding rate of the substrate protein is another important factor determining the interaction. As the first step for characterizing the conformation of the bound substrate at atomic resolution, I analyzed the solution structures of β-lactoglobulin in aqueous and 50% (v/v) trifluoroethanol conditions by heteronuclear multi dimensional NMR.

Chapter 1

General Introduction

1.1. Molecular chaperones

All proteins synthesized in the cell are transformed into well-defined three-dimensional biologically active form, the native conformation. The process by which a protein acquires its native structure is called "protein folding reaction". Since Anfinsen's pioneering study (Anfinsen, 1973), it has been shown that most of small globular proteins can reach their native state spontaneously *in vitro*. From these studies, it has been assumed that all the information necessary for specifying the sophisticated tertiary structure of a protein is completely included in its amino acid sequence, leading to a view that the protein folding *in vivo* also occurs in a spontaneous process. However, recent studies suggested that the protein folding in the living cell might be facilitated by a group of proteins called the molecular chaperones. Molecular chaperones have been revealed to exist in everywhere in the cell, and are thought to facilitate the folding reaction of other proteins, including newly synthesized proteins, denatured proteins which are damaged by various cellular stress, and organellar proteins which have to unfold to traverse the lipid bilayers (Gething & Sambrook, 1992; Hartl *et al.*, 1994; Rothman, 1989; Figure 1.1).

The term 'molecular chaperone' was originally coined to describe nucleoplasmin and then the chloroplast ribulose-bisphosphate carboxylase (Rubisco)-binding protein, which were observed to promote the oligomeric assembly of nucleosomes and Rubisco (Ellis, 1987; Ellis & van der Vies, 1991). The definition of the term 'molecular chaperone' was subsequently expanded, and is currently defined as the proteins that mediate the correct folding or assembly of other proteins, but are not themselves components of the final functional structures (Hendrick & Hartl, 1993). They are also classified as stress or heat-shock proteins because their expression can be induced by a variety of celluar stresses. It should be noted, however, most of stress proteins are constitutively expressed and have essential functions under normal cellular conditions. Moreover, molecular chaperones are ubiquitously distributed in a variety of

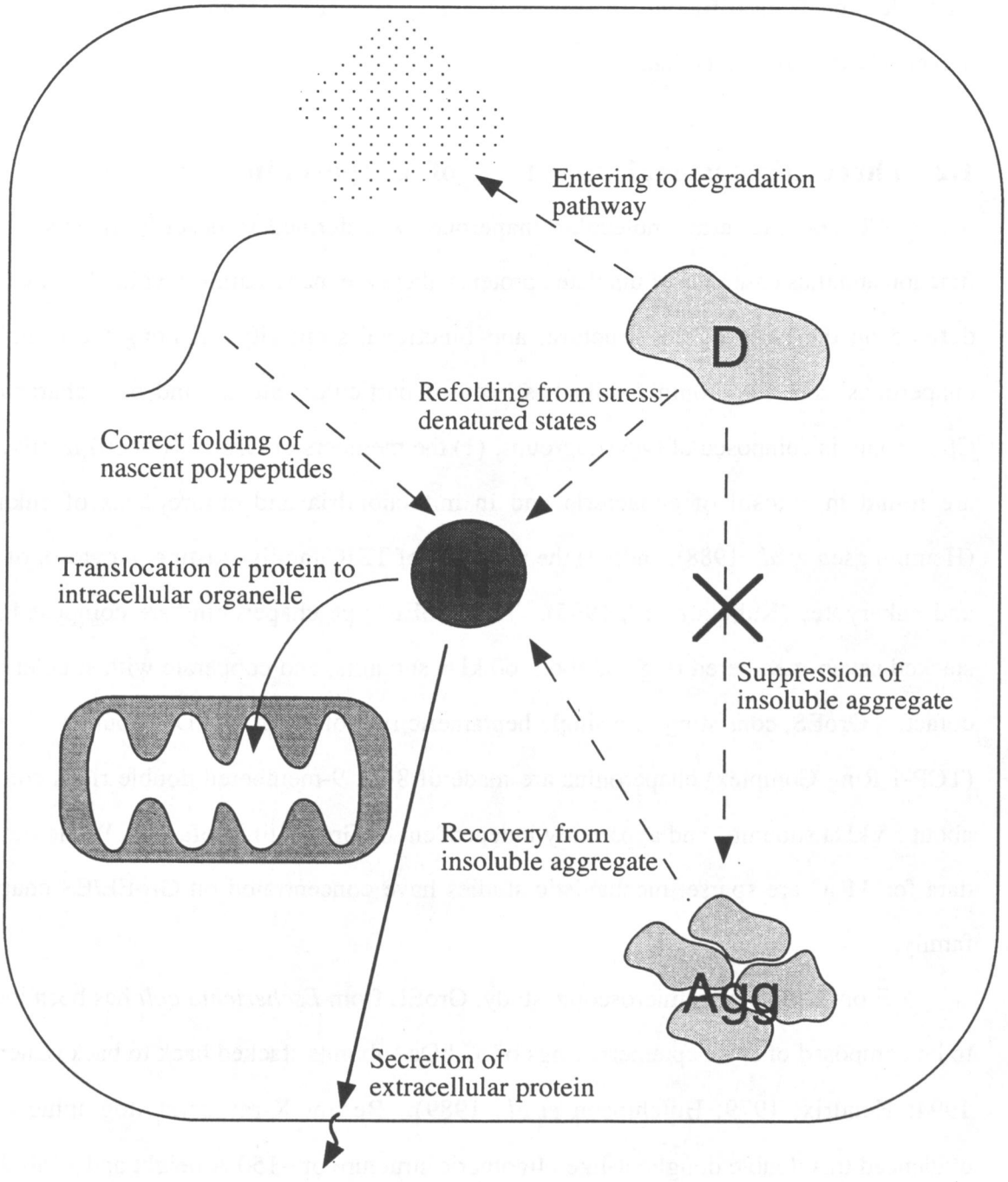


Figure 1.1. Significant roles of molecular chaperones in the cell. "N", "D" and "Agg" designate the native state, denatured state and insoluble aggregate, respectively.

organisms from bacteria to human, and, in the same family, the sequence homology between those derived from different organisms is highly conserved, indicating that they fulfill a universally important function.

1.2. Three dimensional structure of chaperonin

Whereas the term 'molecular chaperone' was defined to describe their specialized function and thus ensemble of unrelated proteins, there are many families within them which are defined on the basis of the structural and functional similarity. Among these molecular chaperones, the chaperonin family has been of particular interest and best characterized. Chaperonin is composed of two subgroups: (1) the members of GroEL (Hsp60) family, which are found in cytosol of eubacteria and in mitochondria and chloroplasts of eukaryotes (Hemmingsen *et al.*, 1988), and, (2) the members of TRiC family existing in cytosol of archae and eukaryotes (Kubota *et al.*, 1995). The GroEL-type chaperonins are composed of two stacked seven-membered rings of about 60 kDa subunits, and cooperate with smaller protein cofactor, GroES, consisting of a single heptameric ring of about 10 kDa subunits. The TRiC (TCP-1 Ring Complex) chaperonins are made of 8- or 9-membered double rings containing about 55 kDa subunits, and apparently independent of GroES-like cofactor. While functional data for TRiC are sparse, mechanistic studies have concentrated on GroEL/ES chaperonin family.

From early electromicroscopic study, GroEL from *Escherichia coli* has been indicated to be composed of two heptameric rings of 57 kDa subunits stacked back to back (Chen *et al.*, 1994; Hendrix, 1979; Hutchinson *et al.*, 1989). Recent X-ray crystallographic analysis evidenced this double doughnut-like oligomeric structure of ~150 Å height and ~140 Å width with a central hole of ~50 Å diameter (Boisvert *et al.*, 1996; Braig *et al.*, 1994, 1995; Figure 1.2). Each subunit is composed of an apical, an intermediate and an equatorial domain. The equatorial domain is the largest one, and provides most interring contact as well as all innerring interaction. This domain also contains an ATP-binding site. An apical domain forms an opening of the central channel and provides the potential binding site for the substrates and co-chaperonin GroES. In this domain, some of the segment facing the channel and the top surface

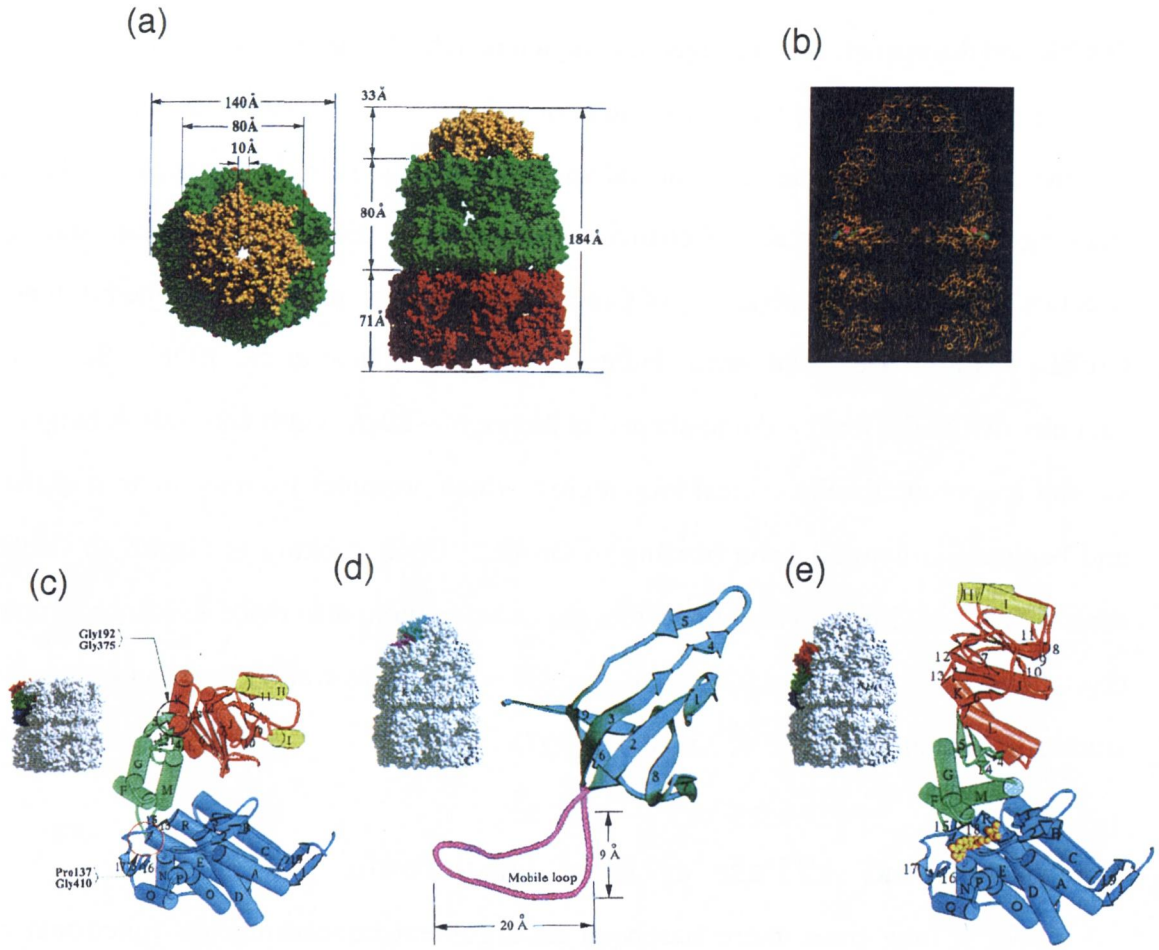


Figure 1.2. Oligomeric structure of $(GroEL)_{14}(GroES)_7$ complex (a), (b), and schematic drawing of monomer of unliganded GroEL (c), GroES (d) and liganded GroEL in the *cis*-ring. In (a), GroEL *cis*-ring, GroEL trans-ring and GroES are colored green, red and yellow, respectively. In (d) and (e), the apical, intermediate and equatorial domains are colored red, green and blue, respectively. The H8 and H9 helices, which are involved in substrate binding and also binding of the co-chaperonin GroES, are colored yellow. In (d), the mobile loop, which is the binding site for GroEL, is colored purple. In (c)-(e), the corresponding protomer in the $(GroEL)_{14}(GroES)_7$ complex are also indicated. Figures are taken from Xu *et al.* (1997).

have very high B-factor and are most poorly resolved, indicating that this region is highly flexible and disordered. Interestingly, this region is rich in hydrophobic and negatively charged residues protruding toward the cavity, mutations of which abolish substrate binding. The small intermediate domain connects the apical and equatorial domains and acts as flexible 'hinge' enabling the large conformational change upon binding of GroES. The X-ray crystallographic structure of smaller heptameric ring of GroES, a co-chaperonin required for the full function of GroEL, has also been elucidated (Hunt *et al.*, 1996; Mande *et al.*, 1996). Seven 10 kDa subunits of GroES form a dome-shaped complex of ~80 Å width and ~30 Å height. Each subunit has a functionally critical loop region, which protrudes from the bottom of the dome and becomes structured upon binding to GroEL. Upon docking of GroES to GroEL, the central cavity of GroEL cylinder becomes expanded to more than twice, evidenced from recent crystallographic structure of $(GroEL)_{14}$-$(GroES)_7$-$(ADP)_7$ as well as the electron microscopic studies (Roseman *et al.*, 1996; Xu *et al.*, 1997).

1.3. Functional ATPase cycle of chaperonin

For a long time, there has been an argument concerning the functional role of chaperonin: whether chaperonin actively assists the refolding reaction of its substrate proteins by using the energy of ATP hydrolysis, or just it traps and isolates the substrate proteins into its central cavity, preventing them from associating into insoluble aggregates. Recently, based on their sophisticated experiments, Horwich *et al.* proposed an attractive model for chaperonin ATPase cycle, relating a functional role of chaperonin and co-chaperonin with their three dimensional structures (Fenton & Horwich, 1997; Weissman *et al.*, 1995,1996; see also Hartl, 1996; Figure 1.3). The abstract of this model is: (1) In a physiological condition, a bullet-shaped asymmetric $(GroEL)_{14}$-$(GroES)_7$-$(ADP)_7$ complex might be the most populated state. (2) An unfolded protein first bind to the open side, the ring which does not associate with GroES, of this asymmetric complex to form a *trans* ternary complex. (3) By binding and hydrolysis of ATP in the opposite ring to GroES, this *trans* complex convert to the *cis* ternary complex through transient formation of GroES-free complex or football-shaped symmetric $(GroEL)_{14}$-$(GroES)_{14}$ complex. Upon formation of the *cis* complex, the substrate binding site

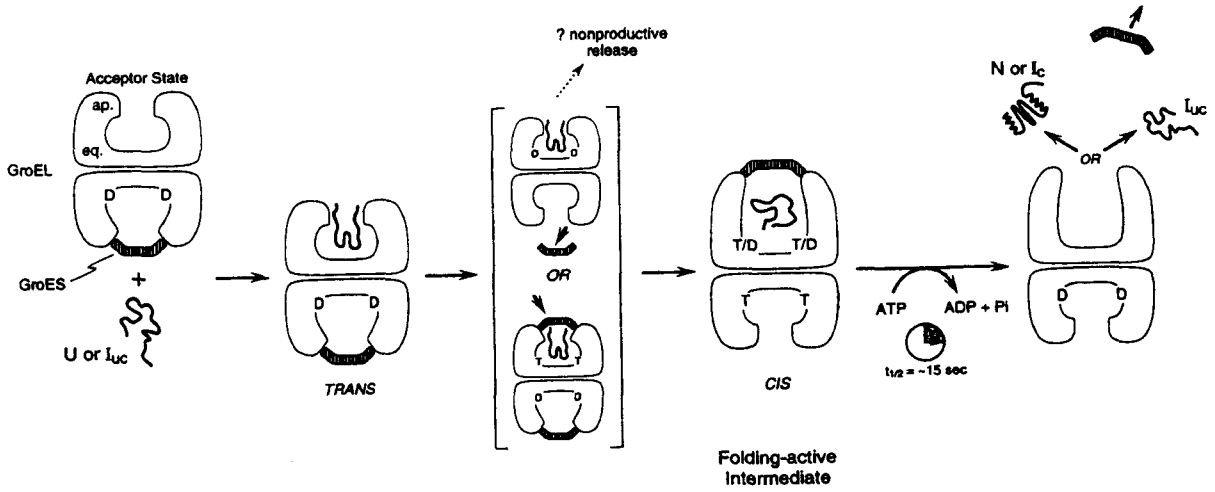


Figure 1.3. Model for a GroEL-GroES-mediated folding reaction. In the GroEL-GroES oligomer, ap. and eq. indicate apical and equatorial domains, respectively. T and D designates ATP and ADP, respectively. For the substrate polypeptides, U, I_{uc}, N represent unfolded state, kinetically trapped intermediates and the native state, respectively. Figure is taken from Fenton & Horwich (1997).

on GroEL changes a partner to GroES, releasing substrate protein into the central cavity, which expanded about twice in volume. (4) This *cis* ternary complex has the most significant role in the refolding cycle of chaperonin system, providing a hydrophilic environment suitable for refolding of a substrate protein, as well as isolating and preventing a substrate from forming aggregates. (5) By binding and hydrolysis of ATP in GroES-free ring, *cis* complex releases GroES and substrate. ATP hydrolysis in GroES-free ring takes several tens of seconds, serving as a 'molecular timer' for substrate refolding. If substrate protein succeeded in reaching its native state within this limit, it would be simply released from chaperonin. On the other hand, if this time-limit is too short for a substrate to fold to the native state, substrate protein would rebind to the same or other chaperonin molecule, entering the chaperonin cycle repeatedly.

1.4. Substrate recognition mechanism of chaperonin

The functional cycle of chaperonin can be divided into two parts, i.e., the recognition and binding of nonnative substrate protein and the release of substrate in ATP and co-chaperonin dependent manner. Although GroES is necessary for the full function of chaperonin system, it is known that GroEL alone has an ability to bind and facilitate the refolding of denatured protein in a ATP-dependent manner. Whereas the recent progress in understanding of the functional role of chaperonin in promoting the correct folding of substrate polypeptides is significant, as summarized above, it remains still ambiguous what conformation chaperonin recognizes and binds. It has been shown that a wide variety of proteins can bind to chaperonin in their kinetic folding intermediate states. They are different in primary structure, secondary structure, molecular weight, isoelectric point, with/without disulfide bridges, and so on. The only known structural preference of chaperonin is that it does not bind polypeptides in the native conformation.

Several models have been proposed, explaining the required structural features for substrate proteins. Previous NMR studies have indicated that some peptides bind to GroEL as α-helices, suggesting that the amphiphilic α-helix is a conformation required for the binding (Landry & Gierasch, 1991; Landry *et al.*, 1992). Subsequently, however, it was found that

other conformations could also bind to GroEL (Landry & Gierasch, 1994; Schmidt & Buchner, 1992). It was proposed that GroEL recognizes substrate proteins in the molten globule (MG) state (Hayer-Hartl *et al.*, 1994; Martin *et al.*, 1991; Mendoza *et al.*, 1992), a compact intermediate with a certain amount of native-like secondary structure but with significantly fluctuating side chains (Kuwajima, 1989; Ptitsyn, 1992). Although a full picture of the MG state has still to emerge, it is proposed to be similar to a major and common folding intermediate of small globular proteins (Kuwajima, 1989; Ptitsyn, 1992). One of the characteristics of the MG state is a high tendency to form aggregates. GroEL is suggested to prevent such aggregation by interacting with the MG state, consequently promoting its efficient folding to the native state.

On the other hand, a number of observations suggest that the conformation recognized by GroEL is more disordered. Cyclophilin bound to GroEL is substantially unfolded, being devoid of a stable secondary structure (Zahn *et al.*, 1994). Okazaki *et al.* (1994), using size-exclusion chromatography, indicated that the MG state of α-lactalbumin, i.e. apo-α-lactalbumin, does not interact with GroEL, whereas a more disordered conformational state, i.e., disulfide-reduced α-lactalbumin, can bind. Lilie & Buchner (1995) studied the interaction of GroEL with a Fab antibody fragment and proposed that the prerequisite for binding is not a certain folding state but the general surface properties of non-native proteins. Similar conclusion was reached by Katsumata *et al.* (1996), who studied kinetically the interaction of GroEL with the MG state of α-lactalbumin (i.e., apo-α-lactalbumin) and detected a weak but significant interaction between them. They suggested that the structure recognized by GroEL is not unique and that the binding strength varies depending on the conformation of the substrate protein.

Although it seems certain that nonspecific hydrophobic interaction is involved in the interaction, the exact structural features recognized by GroEL are ambiguous. One of the major reasons for the difficulty is that even the conformational properties of the substrate proteins themselves in solution are not well known. Therefore, in order to clarify this type of

interaction, it is important, first, to prepare substrates with different and distinct conformational properties, and then to analyze their interactions with GroEL. For this purpose, I chose two well-characterized proteins as substrates and examined their interaction with GroEL using various biophysical methods.

In **Chapter 2**, I examined the interaction of GroEL with cytochrome *c* (cyt *c*) and its derivatives. By modifying the covalently linked heme group, I prepared two kinds of cyt *c* derivatives which have different conformational properties under physiological conditions. I also prepared three cyt *c* fragments with different length. By examining the interaction between these cyt *c* derivatives and GroEL, I will show that the binding affinity is correlated with the extent of solvent-exposed fluctuating hydrophobic clusters. Kinetic analysis of interaction suggested that the refolding rate of the substrate protein is another important factor determining the interaction.

In **Chapter 3**, in order to examine the generality of the results obtained with cyt *c*, I studied the interaction of GroEL with β-lactoglobulin derivatives. β-LG is thought to assume a nonnative α-helical intermediate state on its folding pathway although the native conformation is predominantly β-sheet structure. Whereas cyt *c* is a basic protein (pI = 10), β-LG is an acidic protein (pI = 5). Therefore I tried to address the role of electrostatic interaction between substrate proteins and GroEL (pI = 5). It was also suggested that the conformational state of β-LG recognized by GroEL is the collapsed state which is stabilized mainly by local interaction, lacking the non-local interaction necessary for correct packing of the native conformation. The results suggested that the hydrophobic clusters on the collapsed state are responsible for the interaction with GroEL, consistent with the results obtained with cyt *c*.

In order to obtain further information of the conformational properties necessary for binding to GroEL, I tried to analyze the solution structures of the native and α-helical intermediate of β-LG by heteronuclear NMR. First, I established the high expression system of bovine β-LG using a methylotropic yeast, *Pichia pastoris* (**Chapter 4**). Taking advantage of the secretory pathway of this system, I succeeded in obtaining recombinant β-LG with correct fold at a yield exceeding 0.5 g/l.

In **Chapter 5**, I analyzed the solution structures of β-LG in aqueous and 50% trifluoroethanol (TFE) conditions by heteronuclear multi dimensional NMR, using ^{13}C,^{15}N-double labeled recombinant β-LG expressed in *P. pastoris*. The secondary structures in aqueous condition were agreed well with those of the crystal structure. TFE-induced helical state was composed of several α-helical segments. Analysis of backbone dynamics indicated the enhanced internal mobility of the TFE state. Although I could not examine by NMR the interaction of rhodamine-labeled β-LG with GroEL, further NMR studies with the labeled β-LG will be important to characterize the conformational properties necessary for GroEL binding at the atomic level.

Chapter 2

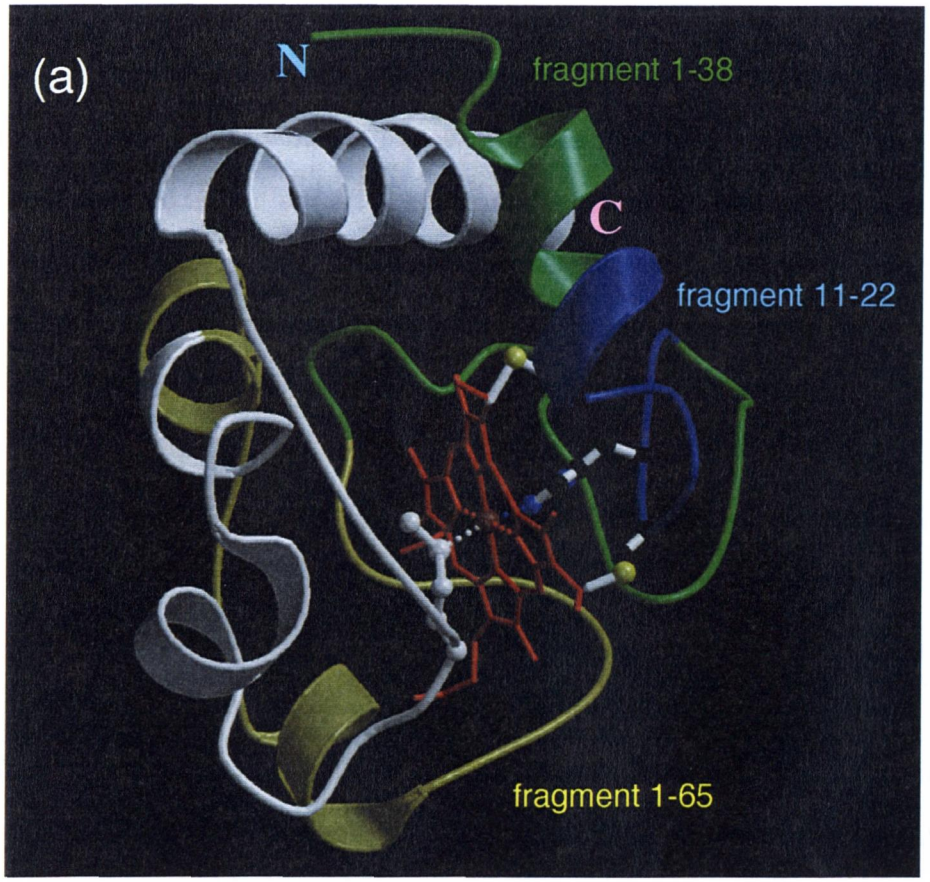
Interaction of GroEL with Cytochrome *c* Derivatives

2.1. Introduction

Although it seems certain that nonspecific hydrophobic interaction is involved in the interaction with substrate polypeptides (see Lin *et al.*, 1995), the exact structural features recognized by GroEL are ambiguous. One of the major reasons for the difficulty is that even the conformational properties of the substrate proteins themselves in solution are not well known. Therefore, in order to clarify this type of interaction, it is important, first, to prepare substrates with different and distinct conformational properties, and then to analyze their interactions with GroEL. I noticed that a series of conformational states of horse cytochrome *c* (cyt *c*) prepared by Hamada *et al.* (1993, 1996a) would be useful for this purpose.

Horse cyt *c* is a small globular protein of 104 amino acid residues containing a heme group (Figure 2.1). The heme is covalently attached by thioether bonds through Cys 14 and Cys 17 and, in the native state, a ring nitrogen of His 18 and the sulfur of Met 80 are coordinated to the heme iron. It is a basic protein with pI = 10.1. Hamada *et al.* (1993, 1996a) prepared apo-cytochrome *c* (apo-cyt *c*) and porphyrin-cytochrome *c* (porphyrin-cyt *c*) and studied the role of heme and its axial ligations in the conformation of the native and MG states. Apo-cyt *c*, lacking the heme group, is largely disordered in terms of secondary structure, but is relatively compact, suggesting the presence of fluctuating hydrophobic clusters. Iron-free porphyrin-cyt *c* has a native-like secondary structure, as measured by CD, but largely flexible side chains, being similar to the MG-like state of the holoprotein.

In the present study, in order to clarify the structural features recognized by GroEL, I investigated the interaction of GroEL with apo-cyt *c*, porphyrin-cyt *c* and three fragments of different sizes, i.e. fragments 1-65, 1-38, and 11-22 (Figure 2.1). When the affinity for GroEL was correlated with the degree of folding of the substrates, the maximum was observed for apo-cyt *c*, suggesting that the affinity is correlated with the extent of solvent-exposed



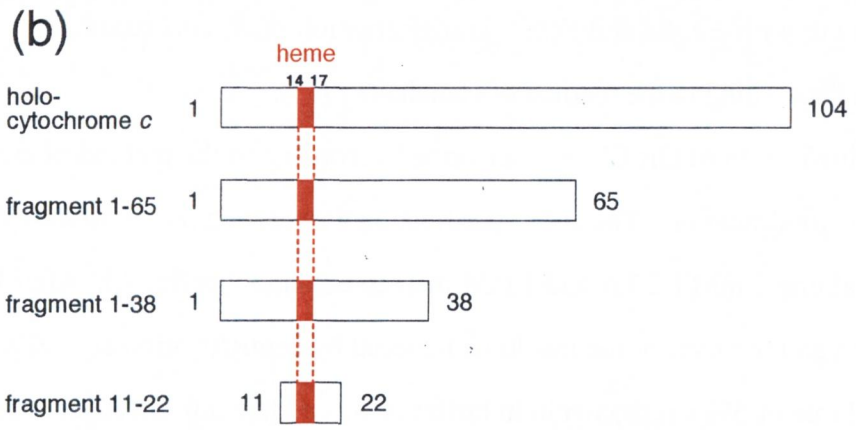


Figure 2.1. (a) Structure of horse cyt *c* illustrating the helices, the heme group and the location of the fragments used in the present study. The figure was created using the coordinates of Bushnell *et al.* (1990) and the program Molscript (Kraulis, 1991) and Raster3D (Merritt & Murphy, 1994). (b) Comparison of the cyt *c* fragments.

fluctuating hydrophobic clusters. The electrostatic interactions between the negatively charged GroEL and positively charged substrates were also shown to be important in increasing the affinity. I propose that this mode of interaction is similar to that between the cyt *c* derivatives and negatively charged phospholipid membranes (Jordi et al, 1989). I further characterized the kinetics of the interaction using fluorescence-labeled apo-cyt *c*. The results suggested that the refolding rate of the substrate protein is another important factor determining the interaction.

2.2. Materials and Methods

2.2.1. Materials

2.2.1.1. GroEL

GroEL was prepared from a GroE-overproducing strain, *E. coli* DH1/pKY206. The plasmid pKY206 was kindly provided by Dr. Y. Kawata (Tottori Univ.). The plasmid pKY206 is constructed by ligation of a partially digested, 3.5-kbp *Eco*RI-*Eco*RV (the *Eco*RI site had been artificially constructed) fragment of pNRK267 (designated pgroESgroEL in a previous study; Kusukawa and Yura, 1988) containing the GroE region with pACYC184 that had been cut with *Eco*RI and *Sca*I. Transformation of *E. coli* strain DH1 with pKY206 was performed according to the method of Hanahan (1983).

Purification of GroEL was performed according to the method of Buchner *et al.* (1991) with some modification. The entire purification was carried out in 50 mM Tris-HCl buffer, pH 7.8, containing 2 mM EDTA and 1 mM mercaptoethanol (buffer A). After lysis of the cells by sonication and removal of the insoluble material by centrifugation at 15,000 x g for 40 min, an equal volume of 5% streptomycin in buffer A was added at 0°C. After stirring for 1 hr on ice, the mixture was centrifuged at 15,000 x g for 40 min and the supernatant was collected. Ammonium sulfate was added to the solution to a concentration of 55% saturation and stirred for 1 hr on ice. The precipitated protein was collected by centrifugation for 30 min at 15,000 x g and dissolved in buffer A. This solution was dialyzed overnight at 4°C against 100 volumes of buffer and applied to a 4 x 15 cm Q-Sepharose (Pharmacia) column. The GroEL protein was eluted from the column with a 0-1M linear NaCl gradient. A 100 ml of fraction corresponding to GroEL was collected and concentrated to 3 ml by pressure filtration using YM10 membrane

(Amicon), and applied to a 2.5 x 150 cm Sephacryl S-300HR (Pharmacia) column. Purity of the protein was examined by sodium dodecylsulfate-polyacrylamide gel electrophoresis.

2.2.1.2. Cyt *c* heme-modified derivatives

Horse cyt *c* (type VI) and cyt *c* fragment 11-21 (microperoxidase, MP11) were purchased from Sigma and used without further purification.

Apo-cyt *c* was prepared by reaction with silver sulfate according to the method of Fisher *et al.* (1973) with some modification (Figure 2.2). Eighty milligrams of Ag_2SO_4 in 9 ml of 10% (w/w) acetic acid were added to 48 mg of cyt *c* dissolved in 1 ml of water. The solution was incubated in the dark for 4 hrs at 40°C and was then centrifuged to remove precipitated heme aggregates. The solution was applied to a 2.5 x 50 cm Sephadex G-25 (Pharmacia) column equilibrated with 0.1N acetic acid. The protein was eluted with the same buffer at 4°C and 10 ml fractions were collected and lyophilized. The silver mercaptide bond was dissociated with the use of dithiothreitol. The lyophilized protein was dissolved in 2 ml of 50 mM ammonium acetate (pH 5.0) which contained 6 M Gdn-HCl and 1 M dithiothreitol. The solution was incubated for 2 hrs at 25°C in the dark and was then centrifuged to remove a greenish yellow precipitate which was presumed to be the silver mercaptide of dithiothreitol. The clear supernatant solution was then applied to a 2.5 x 150 cm Sephadex G-50 (Pharmacia) column equilibrated with 50 mM ammonium acetate (pH 5.0) and eluted with the same buffer. The apoprotein fractions were collected and dialyzed against 20 mM HCl, and lyophilized. The removal of the heme group and the number of free thiol groups were confirmed by the absorption spectrum (see Figure 2.6b, pp. 25) and by titration with 5,5'-dithiobis-(2-nitrobenzoic acid).

DACM-apo-cyt *c* was prepared by treating the two thiol groups of apo-cyt *c* with the fluorescent reagent *N*-(7-dimethylamino-4-methyl-3-coumarinyl)-maleimide (DACM) (Figure 2.2; Yamamoto *et al.*, 1977). The reaction was started by adding 3 mg of DACM dissolved in 400 µl of acetone to 20 ml of apo-cyt *c* (0.5 mg ml^{-1}) dissolved in 6 M Gdn-HCl, 20 mM Tris-HCl (pH 7.0) at 4°C and continued for 4 hr. After the treatment, the succinimide ring of bound DACM was opened by incubating the protein solution at pH 9, 20°C for 12 hr. The solution

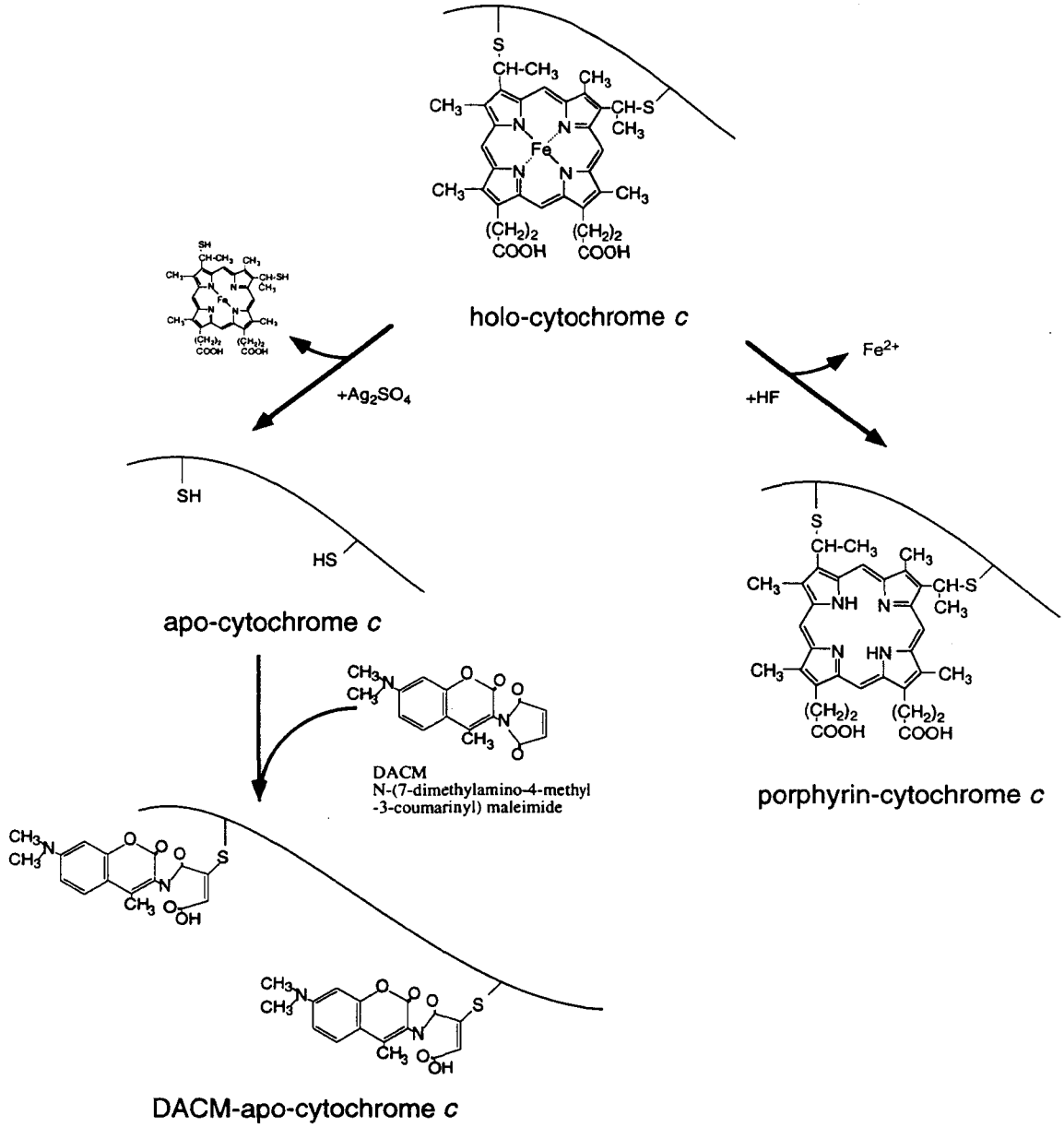


Figure 2.2. Preparation of porphyrin-cytochrome *c* and DACM-apo-cytochrome *c*.

was dialyzed against 20 mM HCl to remove the excess DACM. The number of bound DACM molecules per protein molecule was determined to be 1.8-2.0 from the absorption spectra, assuming the additivity of the spectra of nonlabeled apo-cyt *c* and the reagent reacted with cysteine (see Figure 2.6b, pp. 25). The absorption coefficient of DACM was assumed to be $\varepsilon_{380} = 19800\ M^{-1}\ cm^{-1}$ (Yamamoto *et al.*, 1977).

Porphyrin-cyt *c* was prepared by the reaction of cyt *c* with anhydrous hydrofluoride (Figure 2.2; Fisher *et al.*, 1973). One hundred milligrams of cyt *c* was dissolved in 2 ml of hydrofluoride in a Teflon tube dipped in dry ice/ethanol at -70°C. As soon as the protein dissolved, it was dried under reduced pressure on ice. The dried protein was gently dissolved in 50 mM ammonium acetate (pH 5.0), 6 M Gdn-HCl and purified by gel filtration through a column of Sephadex G-50 (2.5 x 150 cm) equilibrated with 50 mM ammonium acetate. The peak appearing slightly earlier than the position of holo-cyt *c* was collected and stored at -20°C.

2.2.1.3. Cyt *c* fragments

Cyt *c* fragment 1-65 was prepared by treatment of intact horse cyt *c* with cyanogen bromide (BrCN). One hundred milligrams of cyt *c* was dissolved in 10 ml of 70% formic acid and 3 mM BrCN, and the reaction was carried out at 20°C for 4 hr. The resulting mixture was lyophilized and redissolved in 2 ml of 0.1 M acetic acid and purified by gel filtration through a column of Sephadex G-50 (2.5 x 150 cm) equilibrated with 0.1 M acetic acid. The peak fraction with absorption at 400 nm was collected.

Cyt *c* fragment 1-38 was prepared by the method of Taniuchi (Juillerat *et al.*, 1980) with some modification. One hundred milligrams of horse cyt *c* was dissolved in 10 ml of 6.4 M Gdn-HCl. The pH was adjusted to 8.4 with 2N NaOH and 50 µl aliquot of citraconic anhydride were added every 20 min for 5 hrs. The pH was maintained at 8.4 with 2N NaOH. The excess of reagents was removed by 3 successive dialyses against dilute NH_4OH (pH 9.0) and once against 0.1 M ammonium bicarbonate (pH 8.3) at 4°C. After dialysis, trypsin digestion was carried out by adding 50 µl of 1% trypsin for 2 hrs at 20°C, and quenched by adding 50 µl of 2% soybean trypsin inhibitor. The protein was lyophilized, redissolved in 2 ml of 50 mM ammonium bicarbonate (pH 8.3) and applied to a column of Sephadex G-50 (2.5 x

150 cm) equilibrated with the same buffer at 4°C (Figure 2.3a). The peak fraction with absorption at 400 nm was collected and lyophilized. This citraconylated cyt *c* fragment 1-38 was dissolved in 4 ml of 30% acetic acid to deblock the lysine residues and purified by gel filtration on a Sephadex G-50 column equilibrated with 10% formic acid followed by ion-exchange chromatography with a 3 x 20 cm column of SP-Sephadex C-25 (Pharmacia) at pH 7.0 and 20°C (Figure 2.3b). The ion-exchange chromatography was carried out using a 20-300 mM linear gradient of ammonium bicarbonate.

The sizes and purity of the cyt *c* fragments were confirmed by SDS-PAGE (Figure 2.4).

2.2.2. Methods

All spectroscopic measurements in this work were carried out at 20°C with thermostatically controlled cell holders. The pH of the solutions was measured using a Radiometer pH meter, model PHM83, at 20°C.

2.2.2.1. Protein concentration

The concentrations of cyt *c* and their derivatives were determined spectrophotometrically with a Hitachi U3000 spectrophotometer. Molar absorption coefficients used for holo-cyt *c*, apo-cyt *c*, fragments 1-65, fragment 1-38 and fragment 11-22 were $\varepsilon_{409} = 1.10 \times 10^5\ M^{-1}cm^{-1}$, $\varepsilon_{276} = 1.20 \times 10^5\ M^{-1}cm^{-1}$, $\varepsilon_{406} = 1.27 \times 10^5\ M^{-1}cm^{-1}$, $\varepsilon_{406} = 1.14 \times 10^5\ M^{-1}cm^{-1}$, and $\varepsilon_{394} = 1.78 \times 105\ M^{-1}cm^{-1}$, respectively (Hantgan & Taniuchi, 1977; Parr *et al.*, 1978; Peterson *et al.*, 1983). The concentration of porphyrin-cyt *c* was determined using protein assay reagent (Bio-Rad), assuming that the standard curve is the same as that of holo-cyt *c*. The concentration of DACM-apo-cyt *c* was determined from the absorption at 276 nm after subtraction of the contribution of DACM groups. The concentration of GroEL was determined using $\varepsilon_{277} = 1.89 \times 10^6\ M^{-1}cm^{-1}$ (Mizobata *et al.*, 1992).

2.2.2.2. HPLC measurements

The binding of substrate proteins to GroEL was detected by size exclusion chromatography experiments using a Gilson HPLC system. Typically, 70 µl of GroEL

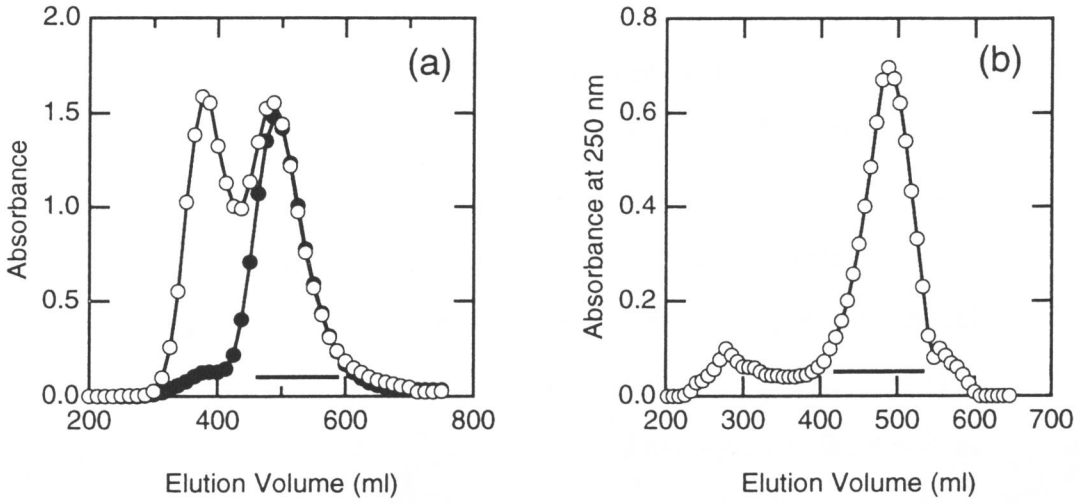


Figure 2.3. (a) Elution profile of the trypsin digest of citraconylated cytochrome *c* on Sephadex G-50. The flow rate was 1 ml/min. Absorbance was measured at 240 nm (○) and 390 nm (●). The fractions indicated by a bar was collected. (b) Ion exchange chromatography of deblocked cyt *c* fragment 1-38 on a SP-Sephadex C-25 column. Elution was performed with a linear gradient of 20-300 mM ammonium bicarbonate at pH 7.0.

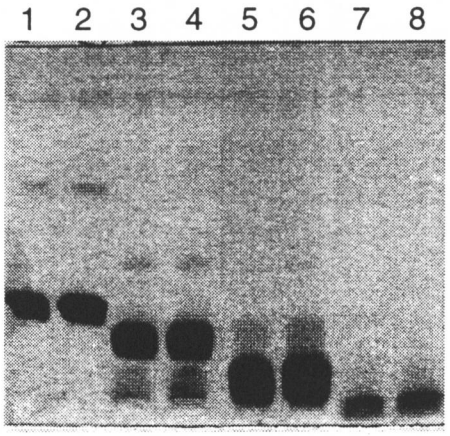


Figure 2.4. SDS-PAGE of cytochrome *c* and its fragments. 10 μg of proteins were applied to 15% polyacrylamide gel. lanes 1 and 3, intact cyt *c*; lanes 3 and 4, fragment 1-65; lanes 5 and 6, fragment 1-38; lanes 7 and 8, fragment 11-22

(2 mg ml^{-1}) was added to a 1/2 equimolar amount of substrate protein (e.g. 10 μl of 0.1 mg ml^{-1} of DACM-apo-cyt *c*), and the solution was incubated for 15 min at 20°C. The protein solution was applied to a Bio-Rad SEC 40XL column (300 x 7.8 mm) equilibrated with 20 mM Tris-HCl (pH 7.0) at a flow rate of 0.5 ml ml^{-1}. The experiments were carried out in the presence of three different concentrations of KCl, i.e., 50, 100, and 200 mM, in the sample and buffer solutions. Eluted proteins were detected by absorption at 220 nm for GroEL, absorption at 400 nm for holo-cyt *c*, porphyrin-cyt *c* and cyt *c* fragments, fluorescence at 480 nm with 380 nm of excitation for DACM-apo-cyt *c*.

2.2.2.3. CD measurements

CD measurements were carried out with a Jasco spectropolarimeter, Model J-720. The instrument was calibrated with ammonium *d*-10-camphorsulfonic acid. The results are expressed as the mean residue ellipticity, $[\theta]$, which is defined as $[\theta] = 100\theta_{obs}(lc)^{-1}$, where θ_{obs} is the observed ellipticity in degrees, *c* is the concentration in residue moles per liter, and *l* is the length of the light path in centimeters. Far- and near-UV CD spectra were measured using cells with light paths of 1 and 10 mm, respectively. The protein concentration was 0.1 mg ml^{-1}.

2.2.2.4. Fluorescence measurements

Fluorescence spectra were measured with a Hitachi F4500 fluorescence spectrophotometer and a cell with a 5-mm light path was used. The fluorescence of DACM was measured at a protein concentration of 0.01 mg ml^{-1} with the excitation at 380 nm.

2.2.2.5. Stopped-flow measurements

Kinetic experiments to investigate the interaction were carried out with an Applied Photophysics stopped-flow fluorescence spectrophotometer, SX-17MV, using a cell with a 2-mm light path. The fluorescence signal was obtained using excitation at 380 nm in conjunction with a GG435 cut-off filter to remove scattered light below 435 nm.

2.2.2.6. Equilibrium analysis

For analysis of the interaction of substrate with GroEL, I assumed a bimolecular two-state reaction.

$$[EL] + [S] \underset{k-1}{\overset{k1}{\rightleftarrows}} [EL{\cdot}S], \qquad \text{Scheme 1}$$

where [EL], [S] and [EL·S] are the concentrations of GroEL, substrate, and the complex, respectively, and k_1 and k_{-1} are the association and dissociation rate constants, respectively. From Scheme 1, the binding constant, K_b, is given by

$$K_b = \frac{[EL \cdot S]}{[EL][S]} = \frac{[EL \cdot S]}{([EL]_0 - [EL \cdot S])([S]_0 - [EL \cdot S])} \qquad (1)$$

where $[EL]_0$ and $[S]_0$ are the total concentrations of GroEL and substrate, respectively. In the present study, $[S]_0$ is constant, and $[EL]_0$ is variable. Therefore, the fraction of complex (f) is given as a function of $[EL]_0$ by

$$f([EL]_0) = \frac{\Delta F}{\Delta F_\infty} = \frac{[EL \cdot S]}{[S]_0} = \frac{([EL]_0 + [S]_0 + 1/K_b) - \sqrt{([EL]_0 + [S]_0 + 1/K_b)^2 - 4[EL]_0[S]_0}}{2[S]_0} \qquad (2)$$

where ΔF is the increase in fluorescence at a given concentration of GroEL, and ΔF_∞ is the limiting amplitude at saturating GroEL. Assuming an appropriate baseline for ΔF_∞, I can calculate the theoretical titration curve using equation (2). K_b was estimated by fitting the theoretical curve to the observed one.

2.2.2.7. Kinetic analysis

From Scheme 1, the rate of complex formation is given by:

$$\frac{d[\mathrm{EL}\cdot\mathrm{S}]}{dt} = k_1[\mathrm{EL}][\mathrm{S}] - k_{-1}[\mathrm{EL}\cdot\mathrm{S}] \tag{3}$$

By solving this differential equation, the fraction of complex is expressed as a function of t by

$$f(t) = \frac{[\mathrm{EL}\cdot\mathrm{S}]}{[\mathrm{S}]_0} = \frac{1}{[\mathrm{S}]_0}\left(\beta - \frac{\beta-\alpha}{1-(\alpha/\beta)\exp(-k_1(\beta-\alpha)t)}\right) \tag{4}$$

where α and β are the solutions of the following quadratic equation ($\alpha < \beta$).

$$x^2 - ([\mathrm{EL}]_0 + [\mathrm{S}]_0 + k_{-1}/k_1)x + [\mathrm{EL}]_0\cdot[\mathrm{S}]_0 = 0 \tag{5}$$

The values of k_1 and k_{-1} and K_b ($=k_1/k_{-1}$) were estimated by fitting equation (4) to each of the binding kinetics measured by fluorescence.

2.3. Results

2.3.1. Conformation of cyt *c* derivatives

Figure 2.5 shows the far-UV and near-UV CD spectra of various conformational states of cyt *c* at pH 7.0 and 20°C. Holo-cyt *c* has a minimum at 222 nm representing the presence of α-helical structures. According to the X-ray structure (see Figure 2.1, pp. 13), the N-terminal, C-terminal, and 60's helices constitute 30% helical content. In the near-UV region, there are two sharp peaks at 280 and 290 nm, which have been assigned to Trp 59 (Davies *et al.*, 1993), indicating that the side chains have unique tertiary structures.

The far-UV CD spectrum of porphyrin-cyt *c* was similar to that of the native state of holo-cyt *c*. On the other hand, in the near-UV region, the two sharp peaks observed for holo-cyt *c* were ambiguous. These results suggest that, at neutral pH, porphyrin-cyt *c* has native-like secondary structures, whereas its side chains are flexible. Hamada *et al.* (1996a) examined the

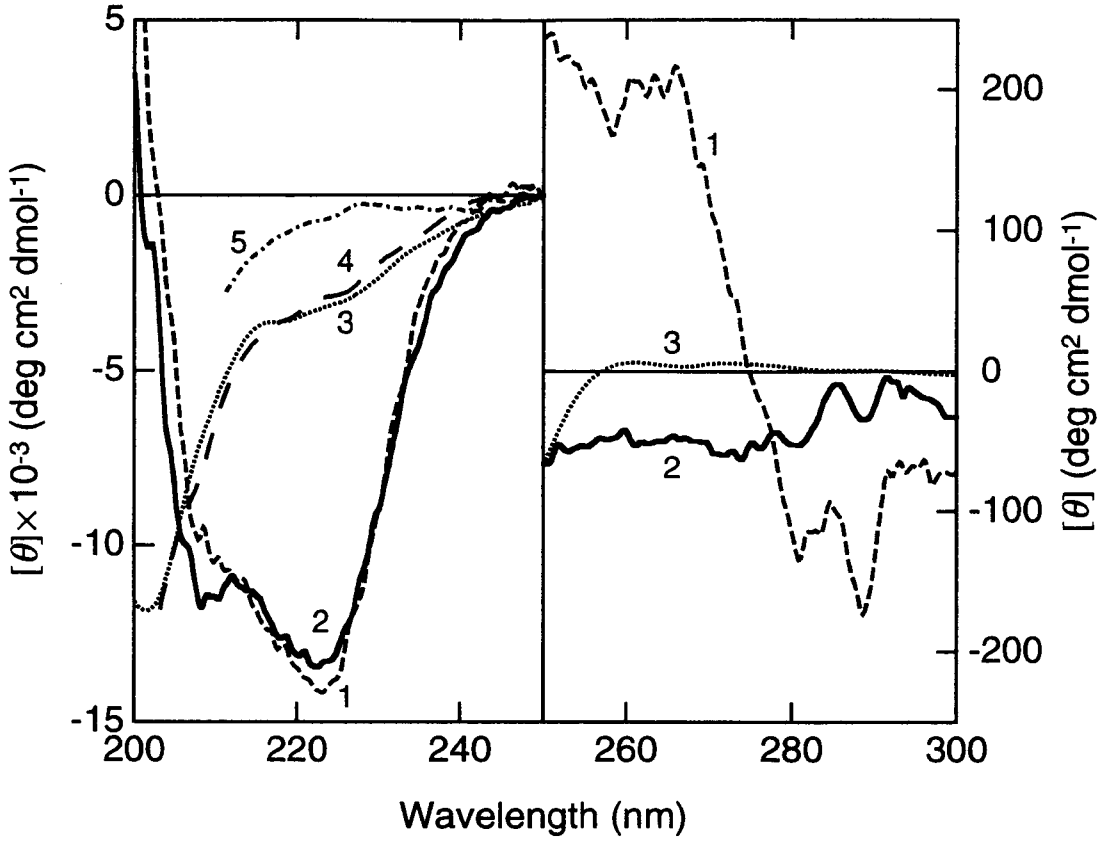


Figure 2.5. CD spectra of horse cyt *c* and its derivatives at pH 7 and 20°C. 1, cyt *c*; 2, porphyrin-cyt *c*;3, DACM-apo-cyt *c*; 4, fragment 1-65; 5, cyt *c* in 6.4 M Gdn-HCl.

conformation of porphyrin-cyt *c* and reported that it assumes a structure similar to the acidic MG state of holo-cyt *c*.

In order to detect the interaction with GroEL, I prepared DACM-apo-cyt *c* in which two thiol groups are labeled with DACM (Figure 2.2, pp. 16; Figure 2.6). The far-UV CD spectrum of DACM-apo-cyt *c* showed a minimum at 200 nm, indicating that the protein is largely disordered in terms of secondary structure (Figure 2.5). No significant near-UV CD signal was observed. The CD spectrum of DACM-apo-cyt *c* is similar to that of apo-cyt *c* and also to that of AEDANS-apo-cyt *c*, in which two thiol groups are labeled with the AEDANS group (Hamada *et al.*, 1993). These results indicate that the effects of the modification are negligible and that DACM-apo-cyt *c*, like apo-cyt *c*, is largely disordered. I consider that, as in the case of apo-cyt *c* and AEDANS-apo-cyt *c*, DACM-apo-cyt *c* assumes a relatively compact denatured state (C state) without a significant ordered secondary structure.

Figure 2.5 also shows the far-UV CD spectrum of fragment 1-65. The spectrum was similar to that of the acid-unfolded cyt *c*, indicating that the fragment is largely disordered. The far-UV CD spectra of fragments 1-38 and 11-21 were similar to that of fragment 1-65 (data not shown).

2.3.2. Size-exclusion chromatography

I monitored the interaction of cyt *c* derivatives with GroEL by size-exclusion chromatography. The absorption spectra of the derivatives are shown in Figure 2.6. Their elution positions were followed by absorption at 380 nm for DACM-apo-cyt *c* and at 400 nm for other derivatives. The elution position of GroEL was detected by absorption at 220 nm, where the contribution of GroEL is predominant. The interaction was examined at different KCl concentrations at pH 7. Figure 2.7 shows the typical elution profiles observed at 50 mM KCl.

When applied separately, the retention time of GroEL (~16 min) was much faster than those of the substrate proteins (21 - 33 min) because of the difference in size (800 kDa for GroEL oligomer, 12 kDa for cyt *c* and 2 - 6.5 kDa for cyt *c* fragments). As shown in Figure 2.7a, when cyt *c* incubated with GroEL was applied to the column, cyt *c* was eluted separately

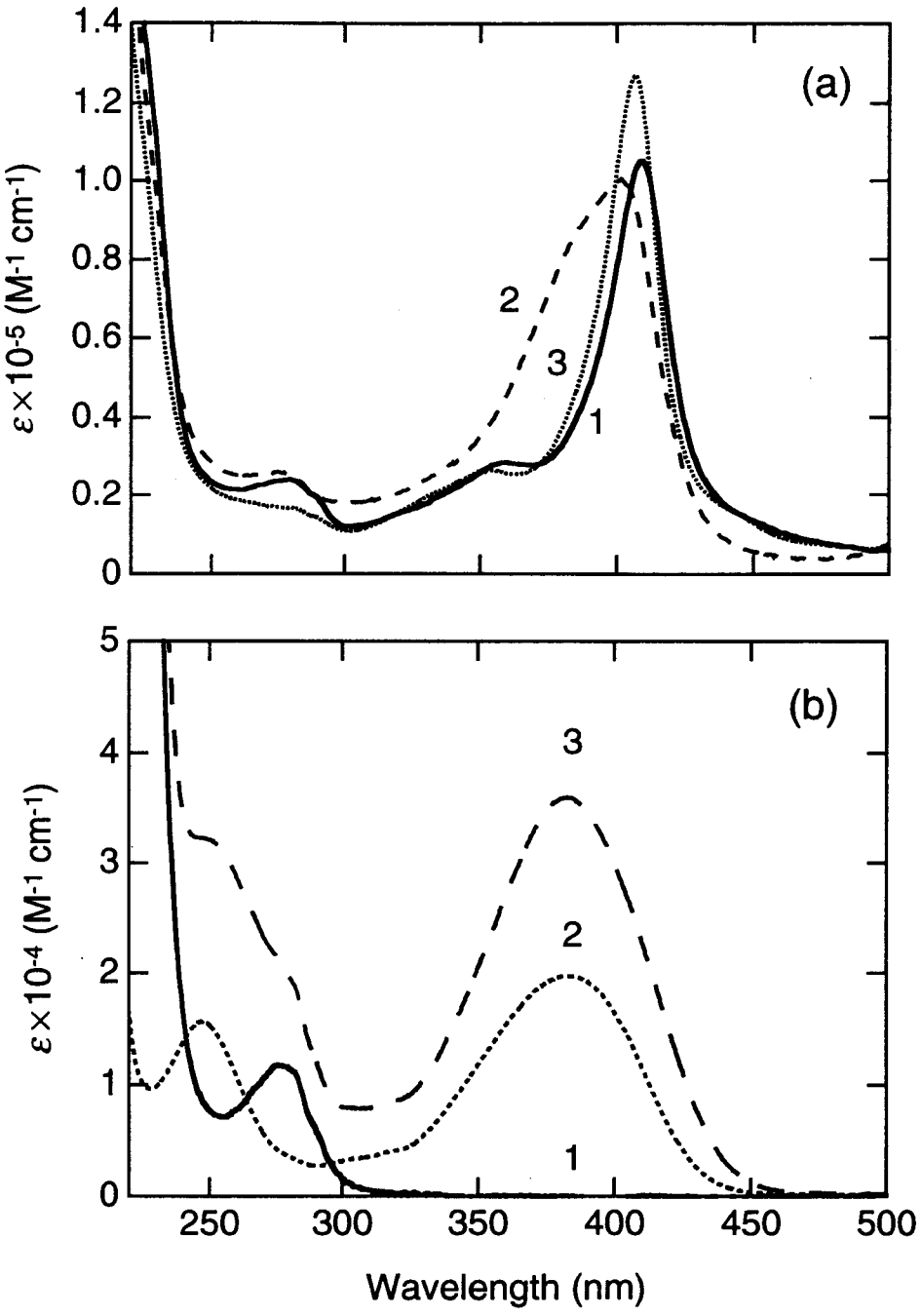


Figure 2.6. Absorption spectra of horse cytochrome *c* and the derivatives at pH 7.0 and 20°C. (a) 1, cyt *c*; 2, porphyrin-cyt *c*; 3, fragment 1-65. (b) 1, apo-cyt *c*; 2, DACM-cysteine; 3, DACM-apo-cyt *c*.

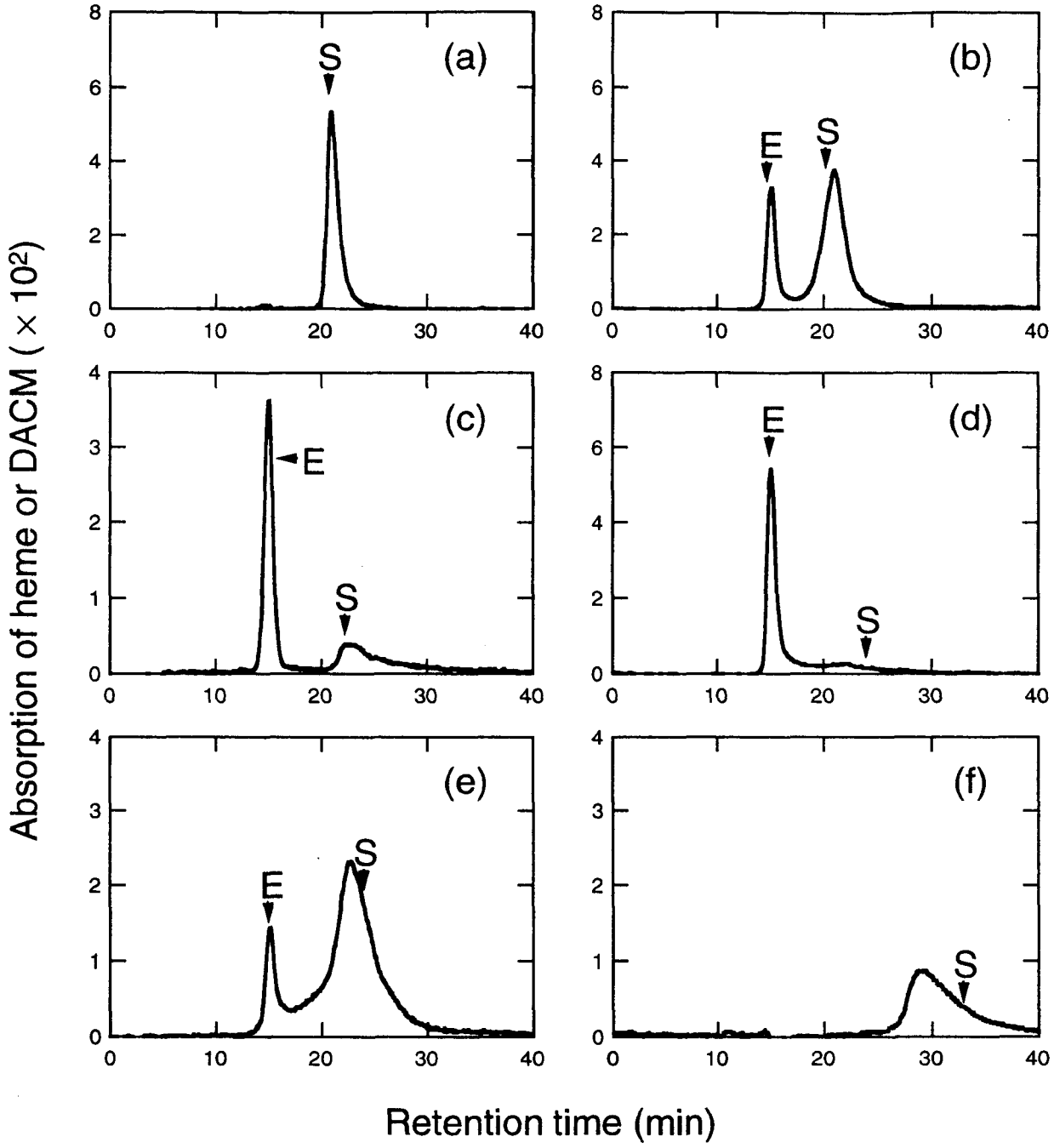


Figure 2.7. HPLC elution profiles of horse cyt *c* and its derivatives in the presence of GroEL at 50 mM KCl, pH 7.0. a, cyt *c*; b, porphyrin-cyt *c*; c, DACM-apo-cyt *c*; d, fragment 1-65; e, fragment 1-38; f, fragment 11-22. The elution profile of DACM-apo-cyt *c* was detected by the absorption at 380 nm and others were detected by the absorption at 400 nm. The small absorption due to GroEL in these wavelengths, detected by the reference elution profile in the absence of substrate, was subtracted. The elution positions of the GroEL and the free substrate are indicated by 'E' and 'S', respectively.

from GroEL. The elution position of cyt *c* was the same as that when the protein was applied without GroEL. These results indicate that GroEL does not interact with the native state of holo-cyt *c*, consistent with the general idea that GroEL does not recognize the substrate proteins in their native state (Landry & Gierasch, 1994).

By contrast, when porphyrin-cyt *c*, DACM-apo-cyt *c*, fragment 1-65, or fragment 1-38 was applied with GroEL, there were two peaks for the substrate proteins: one at the same position as that of GroEL oligomer and the other at a position close to that of the free substrate. It is obvious that GroEL bound these cyt *c* derivatives. The amount of bound substrate was shown to decrease in the order:

apo-cyt *c* > fragment 1-65 > porphyrin-cyt *c* > fragment 1-38 Series 1

Whereas holo-cyt *c* and fragment 11-21 showed negligible binding, several cyt *c* derivatives were bound to GroEL and their affinities varied with a maximum for apo-cyt *c*. It was noted that, for the cyt *c* fragments, similar binding profiles were obtained with the DACM-labeled derivatives (data not shown). Therefore, the tight binding of DACM-apo-cyt *c* is not attributable to the DACM label, but represents an intrinsic property of apo-cyt *c*. It is suggested from the comparison of the three fragments that a certain length is required for tight binding to GroEL.

2.3.3. Effects of salt concentration

Because cyt *c* (pI = 10.1) and GroEL (pI = 4.7 according to amino acid composition) are charged positively and negatively, respectively, at pH 7 (see Table 2.2, pp. 34), the contribution of electrostatic attraction was examined by changing the concentration of salt in the buffer. I carried out HPLC analysis in the presence of 100 and 200 mM KCl and compared the results with those in 50 mM KCl (Figure 2.7). As summarized in Figure 2.8, I found that the increase in KCl concentration substantially weakened the affinity of porphyrin-cyt *c*, fragment 1-65, and fragment 1-38 for GroEL. Interestingly, the interaction of DACM-apo-cyt *c* with GroEL was not notably affected by the salt concentration.

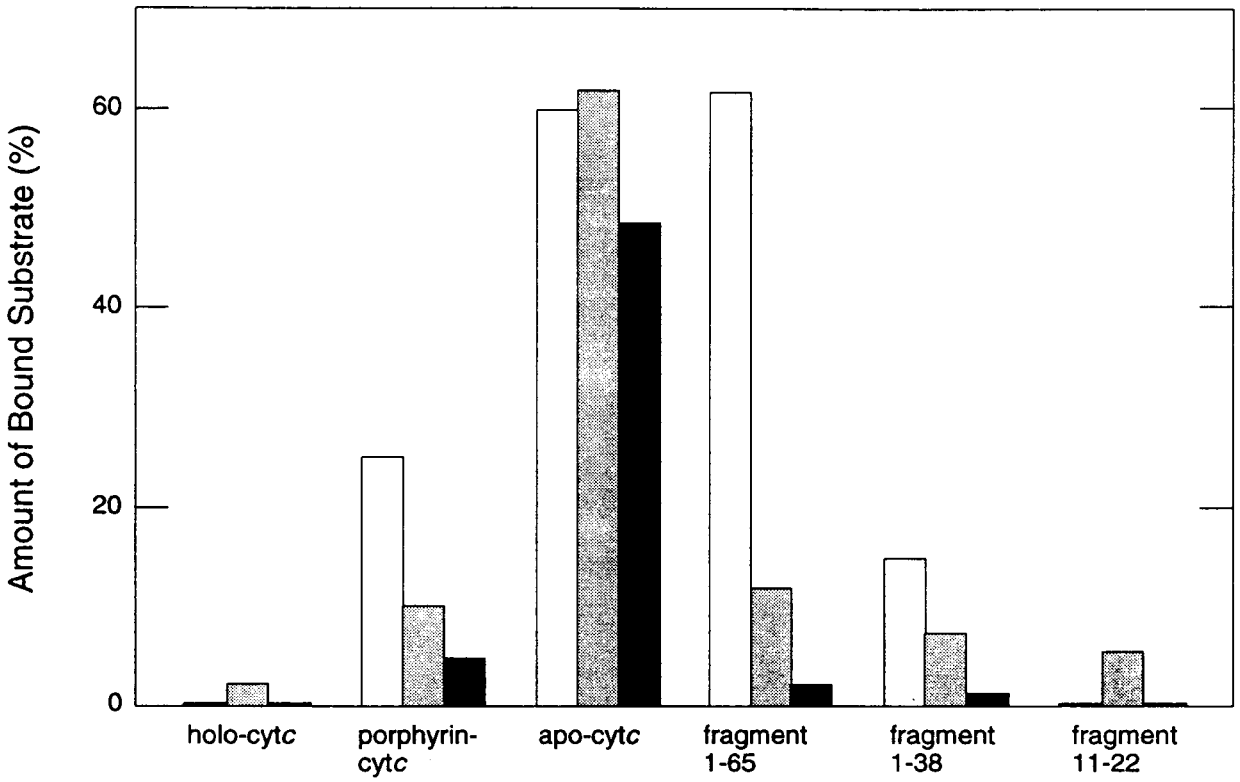


Figure 2.8. Interaction of horse cyt *c* and its derivatives with GroEL at pH 7.0 and various concentrations of KCl. The affinity for GroEL is the amount (%) of applied substrate appearing as a complex with GroEL. KCl concentrations are 50 mM (clear bars), 100 mM (shaded bars), and 200 mM (solid bars).

These results indicate that the electrostatic attractive forces between the positively charged cyt *c* and the negatively charged GroEL promote binding under low salt conditions and that an increase in salt concentration shields the attractive forces, consequently decreasing the affinity. However, electrostatic attraction should not be the major driving force of binding because intact cyt *c* cannot bind to GroEL even under low salt conditions and the interaction of DACM-apo-cyt *c* was not affected significantly by an increase in salt concentration.

2.3.4. Fluorescence spectrum of DACM-apo-cyt *c*

In order to analyze the mode of interaction in detail, I measured the fluorescence spectrum of DACM-apo-cyt *c* in the presence and absence of GroEL (Figure 2.9a). The fluorescence spectrum of 0.8 μM DACM-apo-cyt *c* had a maximum at 474 nm, similar to that of DACM-labeled cysteine. In the presence of 0.8 μM GroEL, the fluorescence intensity was increased two-fold and a shift of the maximum wavelength to 472 nm was seen. The significant increase in the maximum intensity and the slight blue shift indicate that the DACM chromophores are transferred to a hydrophobic environment, suggesting that hydrophobic interaction is involved in the binding.

Taking advantage of the increase in fluorescence intensity upon binding, titration of apo-cyt *c* with GroEL was carried out. Figure 2.9b shows the fluorescence intensity of 0.8 μM DACM-apo-cyt *c* at 480 nm in the presence of various concentrations of GroEL. The titration curve exhibited a typical saturation pattern: at low concentrations of GroEL, the fluorescence intensity increased linearly with increasing GroEL concentration, the slope gradually decreased, and then became saturated. I assumed that, at low concentrations of GroEL, all the added GroEL bound one mole of DACM-apo-cyt *c*. By extrapolating the initial slope of the titration curve, the stoichiometry of GroEL : DACM-apo-cyt *c* in the complex was estimated be 1 : 1.2, i.e. close to a 1 : 1 ratio (Figure 2.9b). Assuming a reversible binding reaction with a 1 : 1 ratio (see Materials and Methods), the binding constant was estimated to be $2.0 \times 10^7\ M^{-1}$.

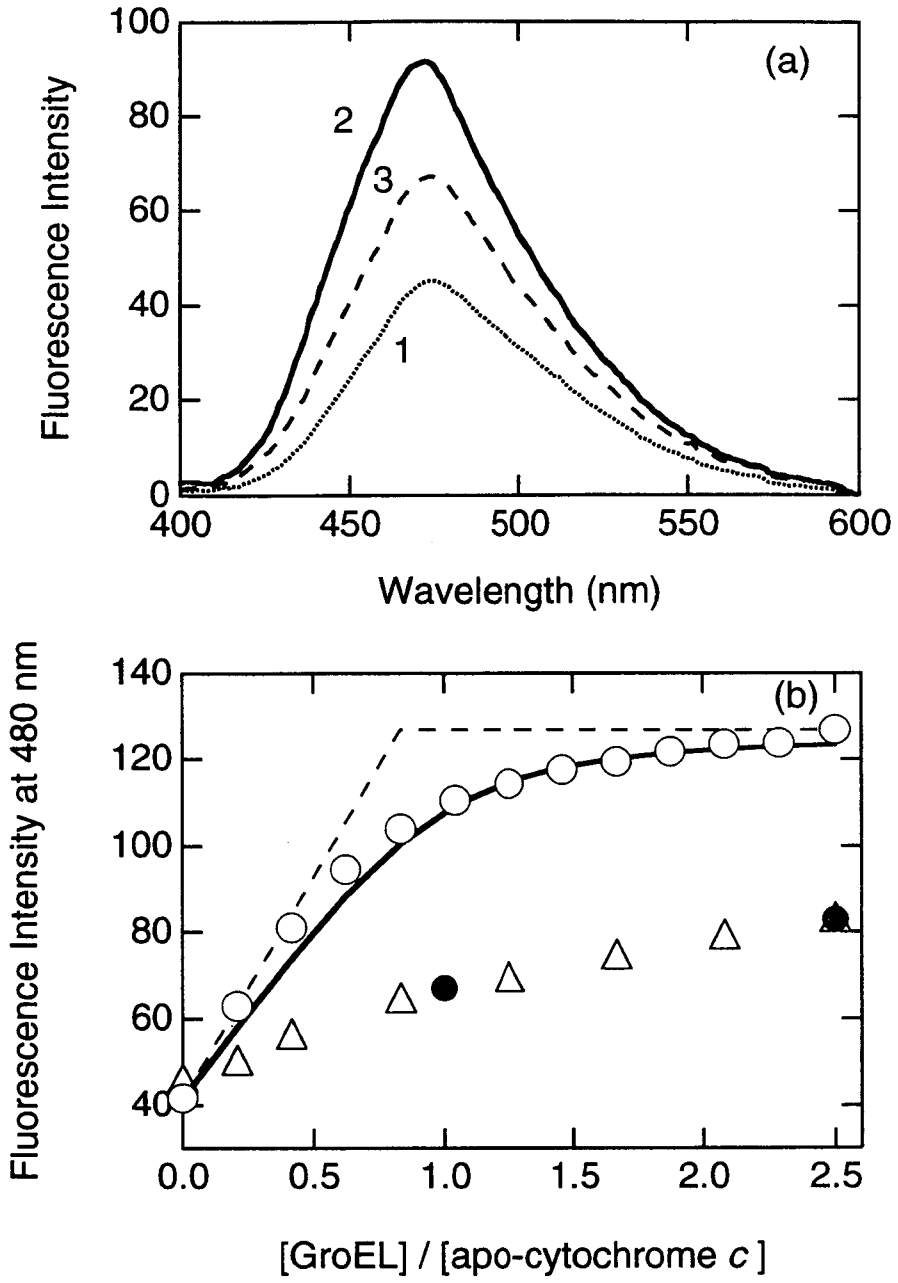


Figure 2.9. (a) Fluorescence spectra of DACM-apo-cyt *c* (0.8 mM) at pH 7.0 and 20°C. 1, in the absence of GroEL; 2, in the presence of 0.8 mM GroEL; 3, in the presence of 0.8 mM GroEL and 2 mM ATP. Excitation was at 380 nm. (b) Titration of DACM-apo-cyt *c* (0.8 mM) with GroEL measured by the fluorescence at 480 nm. Circles, in the absence of ATP; triangles, in the presence of 2 mM ATP. Solid circles indicates the intensity for the solution in which the complex was partially dissociated by addition of 2 mM ATP. The broken line indicates that the stoichiometry of interaction estimated by extrapolating the initial slope is 1: 1.2. The solid line indicates the theoretical binding curve with K_b = 2 x10^7 M^{-1} and a 1 : 1 stoichiometry (Materials and Methods, Scheme 1).

2.3.5. Effects of Mg-ATP

Interaction of substrate proteins with GroEL is known to be regulated by Mg-ATP and GroES (Jackson *et al.*, 1993; Kubo *et al.*, 1993; Landry & Gierasch, 1994; Mayhew & Hartl, 1996; Todd *et al.*, 1994; Weissman *et al.*, 1995). The binding of Mg-ATP to GroEL is proposed to transform GroEL from a form with high substrate affinity to one with low affinity. Here, I examined the effects of Mg-ATP on the interaction of DACM-apo-cyt *c* with GroEL.

By adding Mg-ATP at a final concentration of 2 mM, sufficient to induce the conformational change (Todd *et al.*, 1994), to the performed complex of DACM-apo-cyt *c* (0.8 μM) and GroEL (0.8 or 2.0 μM), the fluorescence intensity was decreased, indicating that the Mg-ATP induced dissociation of the complex (Figure 2.9a). However, it was noted that the fluorescence intensity did not return to a level equivalent to that of the free substrate in the absence of GroEL, suggesting that the dissociation was incomplete. This ATP-induced dissociation of the complex was specific, because ADP did not induce the decrease in fluorescence intensity (data not shown).

The titration of DACM-apo-cyt *c* with GroEL was carried out in the presence of 2 mM Mg-ATP (Figure 2.9b). Although a slight increase in DACM fluorescence was observed upon addition of GroEL, it was significantly lower in extent than that in the absence of Mg-ATP. These results are consistent with the effects of 2 mM Mg-ATP on the performed complex (solid circles in Figure 2.9b).

2.3.6. Measurements of the interaction kinetics

I measured the kinetics of the interaction of DACM-apo-cyt *c* with GroEL using the fluorescence stopped-flow technique. The binding reaction was started by mixing at a 1 : 1 ratio a constant concentration (final 0.4 μM) of DACM-apo-cyt *c* with various concentrations of GroEL (final 0 - 1.6 μM). Figure 2.10a shows the typical binding kinetics. It can be seen that the reactions were completed within 300 ms and that, with increasing GroEL concentration, the reaction became faster, suggesting that the intermolecular process is rate-limiting.

I assumed that the reaction can be approximated by a bimolecular two-state mechanism (see Materials and Methods). On the basis of manual curve fitting analysis, the association (k_1)

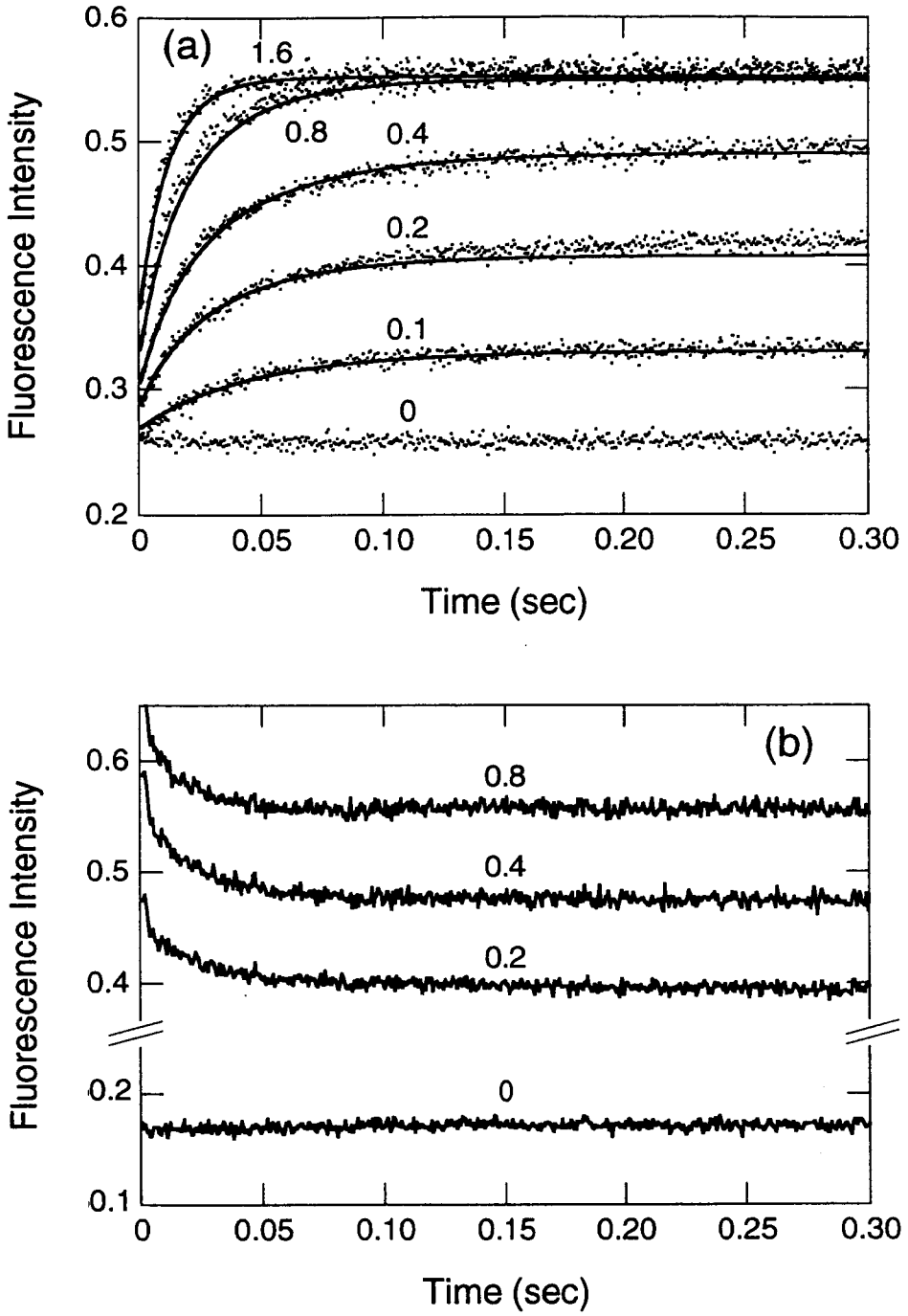


Figure 2.10. Stopped-flow kinetics of the interaction of DACM-apo-cyt *c* with GroEL at 20°C followed by the fluorescence of DACM groups. (a) Kinetics of binding between DACM-apo-cyt *c* (0.4 mM) and GroEL. The figures indicate the concentration of GroEL in mM. The solid lines indicate the theoretical curves on the basis of a bimolecular two state model (Materials and Methods, Scheme 1) and the rate constants shown in Table 2.1. (b) Dissociation kinetics of the complex between DACM-apo-cyt *c* (0.4 mM) and various concentrations of GroEL induced by addition of 2 mM ATP. The figures indicate the concentration of GroEL in mM.

and dissociation (k_{-1}) rate constants were estimated for each curve (Table 2.1). The average values of k_1 and k_{-1} were 7.8 x 10^7 $M^{-1}s^{-1}$ and 1.9 s^{-1}, respectively. The binding constant (K_b) determined from these rate constants ($K_b = k_1/k_{-1}$ = 4.1 x 10^7 M^{-1}) was similar to that (2.0 x 10^7 M^{-1}) estimated from the equilibrium titration experiment. Figure 2.10b shows the dissociation kinetics of the complex between DACM-apo-cyt *c* and GroEL initiated by addition of 2 mM Mg-ATP. The rapid decrease in DACM fluorescence occurred in less than 200 ms. However, the fluorescence intensity did not decrease to a level equivalent to that of the free substrate without GroEL, indicating that the dissociation is only partial. These results are consistent with those obtained from the equilibrium titration experiments (Figures 2.9). It is evident that, contrary to the binding kinetics, the rapid dissociation phase is independent of GroEL concentration, suggesting that the rate-limiting step is an intramolecular process.

2.4. Discussion

2.4.1. Conformational features recognized by GroEL

One of the most important aspects of the mechanism of GroEL action is the conformational features of the substrate proteins (Ewbank *et al.*, 1995; Hayer-Hartl *et al.*, 1994; Martin *et al.*, 1991; Okazaki *et al.*, 1994, 1995). To address this issue, I measured the interaction with GroEL of a series of cyt *c* derivatives. The most significant finding was that maximal binding was observed with apo-cyt *c* (Figure 2.8, Table 2.2). Apo-cyt *c* has little ordered secondary structure, and exists in a collapsed state with fluctuating hydrophobic clusters (Hamada *et al.*, 1993). Although porphyrin-cyt *c* has several features of the MG state (Hamada *et al.*, 1996a), including a high helical content, compactness, and flexible side chains, its affinity for GroEL is lower than that of apo-cyt *c*. These results indicate that the MG state with a significant amount of ordered secondary structure is not the required conformational state, although it often satisfies the required conditions, as described below.

The present results suggest that the most important feature responsible for the interaction is exposure of the hydrophobic clusters to the solvent. Protein folding is a process in which an extended polypeptide chain acquires compact packing through the formation of specific secondary and tertiary structures and hydrophobic interactions (Nishii *et al.*, 1994; Dill *et al.*,

Table 2.1. Kinetic parameters of the association reaction between GroEL and DACM-apocytochrome *c*.

$[S]_0$ (mM)	$[EL]_0$ (mM)	k_1 (M^{-1} s^{-1})	k_{-1} (s^{-1})	K_b (M^{-1})
0.4	0.1	5.5×10^7	1.6	3.5×10^7
0.4	0.2	1.0×10^8	0.7	1.5×10^8
0.4	0.4	9.1×10^7	2.8	3.3×10^7
0.4	0.8	7.1×10^7	1.5	5.1×10^7
0.4	1.6	6.1×10^7	2.8	2.2×10^7

Table 2.2. Structures and interaction with GroEL of various derivatives of horse cyt *c*.

	Ordered secondary structure	Ordered tertiary structure	Net charge at pH 7	Compact-ness	Exposed hydrophobic clusters	Interaction with GroEL
cyt *c*	+	+	+10	+	-	-
porphyrin-cyt *c*	+	-	+7	+	+	+
apo-cyt *c*	-	-	+9	+-	++	+++
fragment 1-65	-	-	+7	-	n.d.[a]	++
fragment 1-38	-	-	+6	n.d.[a]	n.d.[a]	+-
fragment 11-22	-	-	+1	n.d.[a]	n.d.[a]	-

[a]Not determined

1995). The formation of the compact state such as apo-cyt *c* and porphyrin-cyt *c* is driven by the hydrophobic interactions (Hamada *et al.*, 1996a). On the other hand, the amount of exposed hydrophobic clusters is determined by a balance of two processes occurring consecutively during protein folding: first, their formation and, second, their burial in the interior of the protein molecule. Because the formation of an ordered secondary structure in porphyrin-cyt *c* is accompanied by the burial of the hydrophobic surfaces, the amount of exposed hydrophobic clusters in porphyrin-cyt *c* should be smaller than that of apo-cyt *c*, explaining the difference in the affinity to GroEL (see Figure 2.11, pp. 37). The present results further suggest that the affinity of the ordered MG state for GroEL is lower than that of the disordered compact state. Indeed, such a situation has been observed for α-lactalbumin (Hayer-Hartl *et al.*, 1994; Okazaki *et al.*, 1994).

In order for such exposed hydrophobic clusters to be formed cooperatively, a certain size of polypeptide may be required. Because the fluctuating hydrophobic clusters at the early stage of protein folding may be stabilized by relatively nonspecific encounter of the groups, the increase in size i.e., the increase in the number of hydrophobic groups, cooperatively increases the chance of encounter, thus stabilizing the hydrophobic clusters. This would explain why affinity was reduced with the decrease in size of the cyt *c* fragments, even though these fragments are similarly disordered from their CD spectra. In the case of fragment 11-22, the absence of a positive charge further discourages the interaction (Table 2.2). Hlodan *et al.* (1995) examined the interaction of the fragments of rhodanese with GroEL and showed that there is a critical size for effective GroEL binding: a length of about 50 amino acid residues is required for a substrate polypeptide to interact with GroEL, consistent with the present results.

2.4.2. Role of electrostatic interactions

The present work clarified the role of electrostatic interaction in the binding: the electrostatic attractive forces stabilize the complex formed between the positively charged substrates and the negatively charged GroEL. Nevertheless, because no interaction was observed between intact cyt *c* and GroEL, it can be concluded that electrostatic interaction is not a major driving force of the interaction.

There is evidence to indicate that electrostatic interaction plays a role in the binding of substrates to GroEL. Itzhaki *et al.* (1995) studied the refolding of barley chymotrypsin inhibitor 2 and its mutants in the presence of GroEL and indicated that hydrophobic and positively charged side chains on the substrate promote the interaction with GroEL, whereas negatively charged side chains tend to repel, consistent with our results. Okazaki *et al.* (1994) examined the interaction between GroEL and α-lactalbumin in which the disulfide bonds are reduced. They found that the interaction was strengthened when the ionic strength of the buffer was increased. Because α-lactalbumin is an acidic protein with pI = 4, and charged negatively at neutral pH, the electrostatic repulsion contributes to weakening the interaction. Shielding of such repulsion by increasing the salt concentration thus favors the binding. Therefore, although the effects of salts are opposite in the two cases, they are consistent with each other, confirming the role of electrostatic interaction in modulating the binding.

2.4.3. Similarity to membrane binding

It is notable that the mode of interaction between cyt *c* derivatives and GroEL resembles that between cyt *c* derivatives and acidic phospholipid membranes (Figure 2.11). It has been shown that apo-cyt *c* has high affinity for acidic phospholipid membranes, whereas intact cyt *c* binds to them only slightly (Jain & Zakim, 1987; de Jongh & de Kruijff, 1990; Jordi *et al.*, 1989; Muga *et al.*, 1991). Hamada *et al.* (1996a) reported that porphyrin-cyt *c* and apo-cyt *c* showed binding to acidic phospholipid liposomes consisting of cardiolipin and phosphatidylcholine, but not to neutral liposomes consisting of phosphatidylcholine. Jordi *et al.* (1989) reported that various fragments of cyt *c*, including fragments 1-65 and 1-38, were bound to acidic phospholipid membranes. It is known that the interaction between cyt *c* derivatives and phospholipid membranes is modulated by salt conditions (Jain & Zakim, 1987). The affinity decreases significantly with an increase in ionic strength, indicating that electrostatic attraction promotes the binding.

The crystal structures of GroEL (Braig *et al.*, 1994, 1995) and GroES (Hunt *et al.*, 1996; Mande *et al.*, 1996) have been determined, and in combination with mutational and functional analyses (Fenton *et al.*, 1994; Weissman *et al.*, 1995), the detailed mechanism of

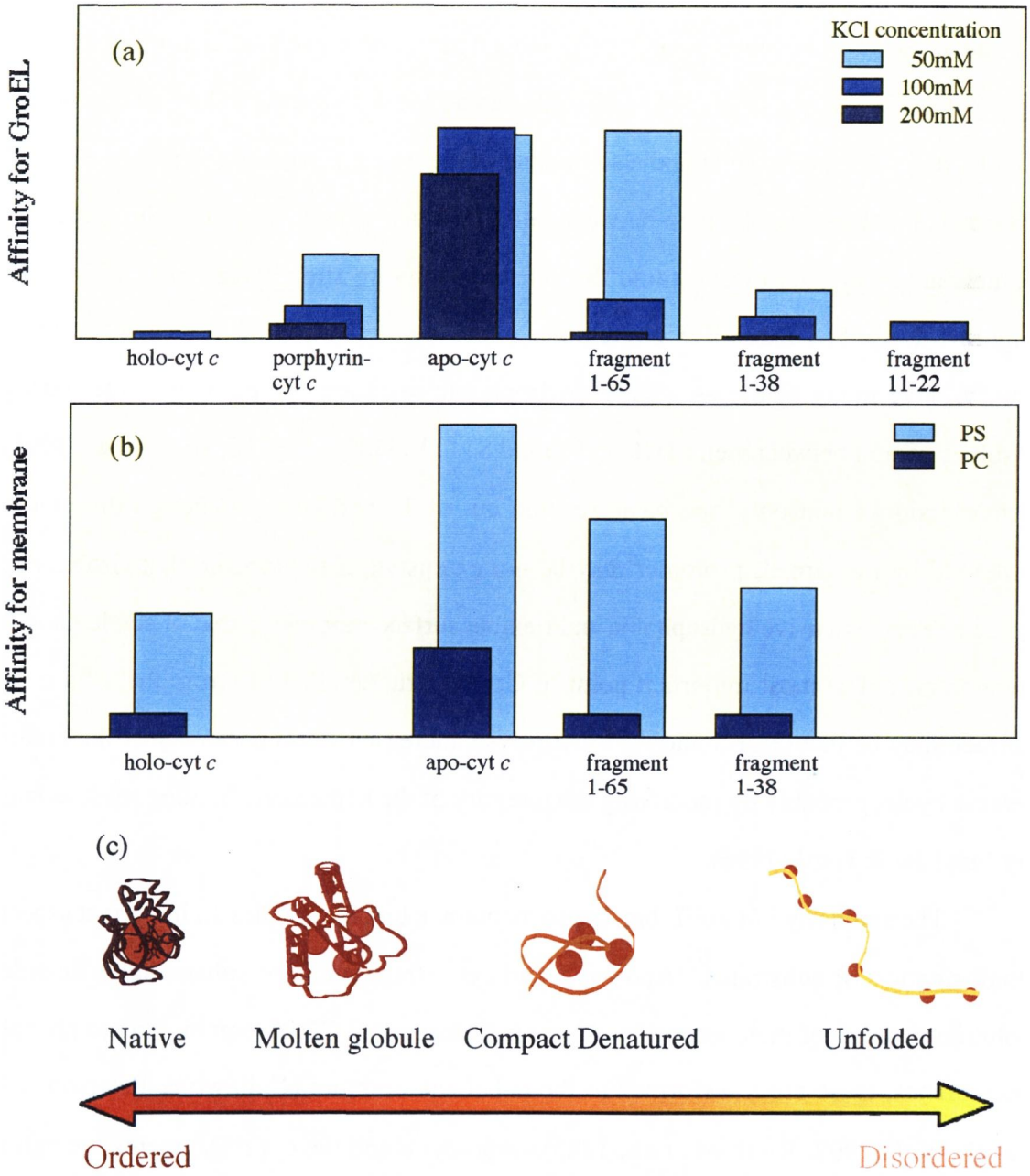


Figure 2.11. Comparison of the affinity for GroEL and for lipid monolayers of horse cytochrome *c* and its derivatives. (a), The affinity for GroEL at different concentrations of KCl, redrawn from Figure 2.8; (b), The affinity for lipid monolayer, phosphatidylseline (PS) and phosphatidylcholine (PC), taken from Jordi *et al.* (1989). (c), Schematic drawing of the conformational states of cytochrome *c* and its derivatives.

GroEL action is being elucidated (Mayhew & Hartl, 1996). The functional properties of GroEL can be assigned to different regions: a flexible hydrophobic patch on the inside surface of the apical domain (residues 191-374) of GroEL, facing the central cavity, has been indicated to be involved in substrate binding and also binding of the co-chaperonin GroES. More specifically, helices H8 and H9 and the loop between helix H10 and strand S11, which have the highest B values, are suggested to constitute the substrate binding site (Braig *et al.*, 1994, 1995; see Figure 1.2, pp. 5). These flexible regions, exposing the hydrophobic residues to the central cavity, have many acidic as well as hydrophobic residues. In particular, the most poorly resolved region between helix H10 and strand S11 (Val300 - Asp316; Braig *et al.*, 1995) has 6 acidic residues and only one basic residue out of 17 residues. Although the binding site provided by the GroEL protomer may be not extensive, it is probable that GroEL heptamer constitutes an extensive hydrophobic and flexible surface mimicking that of acidic phospholipid membranes. The most important point of GroEL structure distinct from that of a membrane surface may be that GroES and ATP-hydrolysis increase the cooperativity of the binding and release cycle, probably by modifying the integrity of the heptameric binding sites, as suggested by Mayhew & Hartl (1996).

The similarity of GroEL binding to membrane binding implies an important aspect of the conformation of substrates. Apo-cyt *c* and cyt *c* fragments are substantially disordered in solution in terms of their secondary structure measured by CD. Upon binding to phospholipid membranes, these are transformed to the α-helical structure (Collawn & Paterson, 1990; de Jongh *et al.*, 1992; Rietveld *et al.*, 1985). Apo-cyt *c* and the cyt *c* fragments are assumed to interact with the membrane through the hydrophobic surface of the amphiphilic α-helices. It is possible, as schematic drawn in Figure 2.12, that apo-cyt *c* and the cyt *c* fragments, which have no significant secondary structure in solution, undergo such conformational transition upon binding to GroEL.

On the other hand, it is reported that the binding of cyt *c* to cardiolipin membranes induces disordering of the rigid native structure (Spooner & Watts, 1991a, b). Therefore, it is also possible that a relatively ordered substrate such as porphyrin-cyt *c* becomes disordered upon binding to GroEL so as to give the exposed hydrophobic surfaces complementary to the

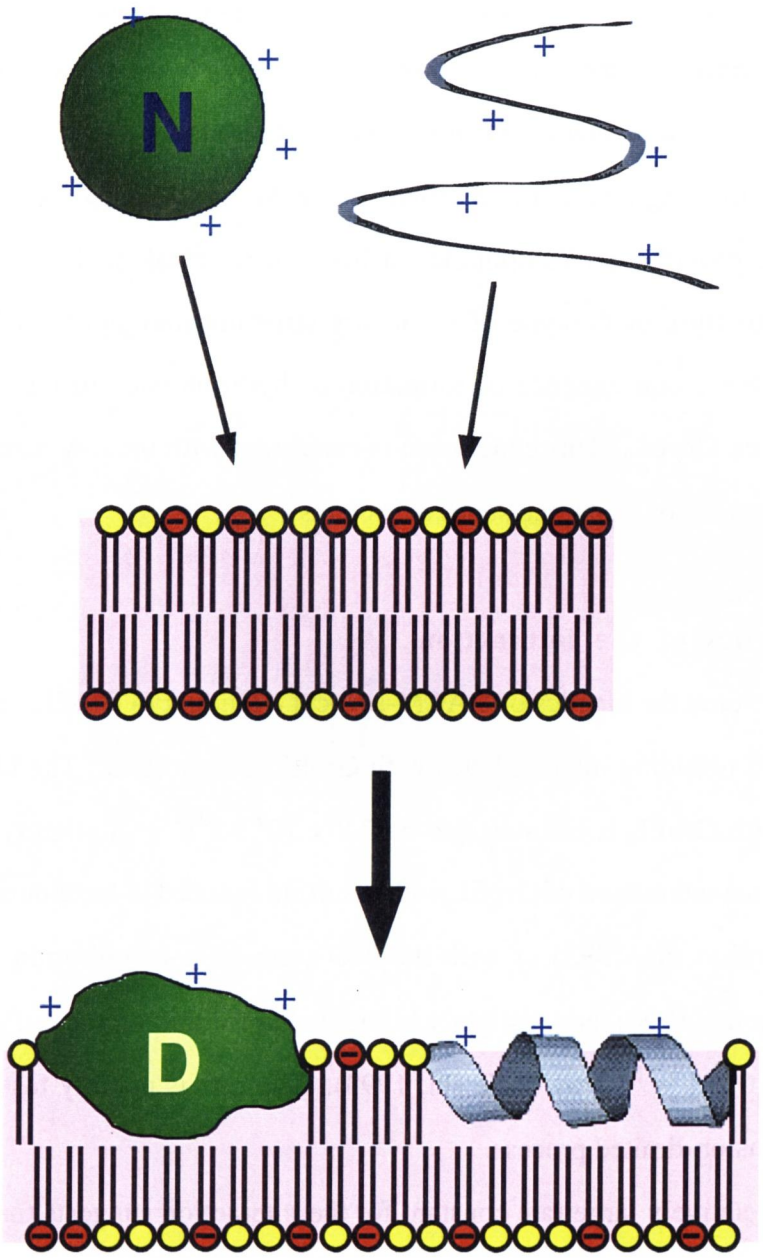


Figure 2.12. A model of possible conformational changes of proteins on association with membrane. N and D designate the native and denatured states, respectively. Crosses indicate the positive charges on the protein. Red and yellow circles represent acidic and neutral phospholipids, respectively.

binding site (Figure 2.12). Such structural rearrangement may be useful for restarting the folding of trapped intermediates (Ranson *et al.*, 1995). It is also expected for other substrates that they assume a β-sheet conformation when bound to GroEL, as suggested for the immunoglobulin fragment (Lilie & Buchner, 1995), or other disordered structures, as long as the structure provides a hydrophobic surface. In conclusion, the presence or absence of a secondary structure or the type of secondary structure may not be a factor determining the interaction, but a consequence of formation of hydrophobic clusters complementary to the binding site on GroEL. This conclusion is consistent with the low barriers to interconversion of the intermediate or denatured states of proteins.

2.4.4. Kinetics of the interaction

I measured the kinetics of DACM-apo-cyt *c* binding to GroEL. Their implication in the interaction of refolding intermediates with GroEL is important. The binding rate constant of apo-cyt *c* with GroEL is estimated to be 7.8×10^7 $M^{-1}s^{-1}$. A slightly lower value has been reported for the interaction of GroEL with disulfide-reduced α-lactalbumin ($k_1 = 1.95 \times 10^5$ $M^{-1}s^{-1}$; Murai *et al.*, 1995) or with the MG state of α-lactalbumin ($k_1 = 1 \times 10^6$ $M^{-1}s^{-1}$; Katsumata *et al.*, 1996), and a slightly higher value for the interaction of barnase with GroEL ($k_1 = 1.3 \times 10^8$ $M^{-1}s^{-1}$; Gray & Fersht, 1993). The rate is notably fast, although slower than that of a diffusion-limited process.

The relatively large rate constant for the interaction suggests that the interaction of the folding intermediate with GroEL is determined by competition between the interaction and the refolding, which is also suggested by Fisher & Yuan (1994). Assuming the concentration of GroEL to be 1 μM (0.8 mg ml^{-1}), the pseudo-first order rate constant of the association is 80 s^{-1} (relaxation time, 12.5 ms). For a protein with a large refolding rate constant, the interaction with GroEL should be marginal, because the intermediate which can interact with GroEL would disappear rapidly. On the other hand, the refolding rate of a large protein is slow in general, thus providing a chance for accumulation of the collapsed intermediate, which can interact with GroEL. Therefore, the refolding rate determines the accumulation of the collapsed intermediate, thus determining the interaction with GroEL. Indeed, it is known that, whereas the refolding *in*

vitro of several large proteins is completely trapped by GroEL (Kawata *et al.*, 1994), refolding of other proteins, in particular small rapid-folding proteins, is merely retarded by GroEL (Gray & Fersht, 1993, Itzhaki *et al.*, 1995; Katsumata *et al.*, 1996). However, it is probable that the substrate proteins, in particular smaller ones, can refold to the native state, maintaining the complex with GroEL. Therefore, the exact mechanism of the interaction may be more complicated than that approximated by a simple competition mechanism, in which a bound substrate cannot refold to the native state.

Chapter 3

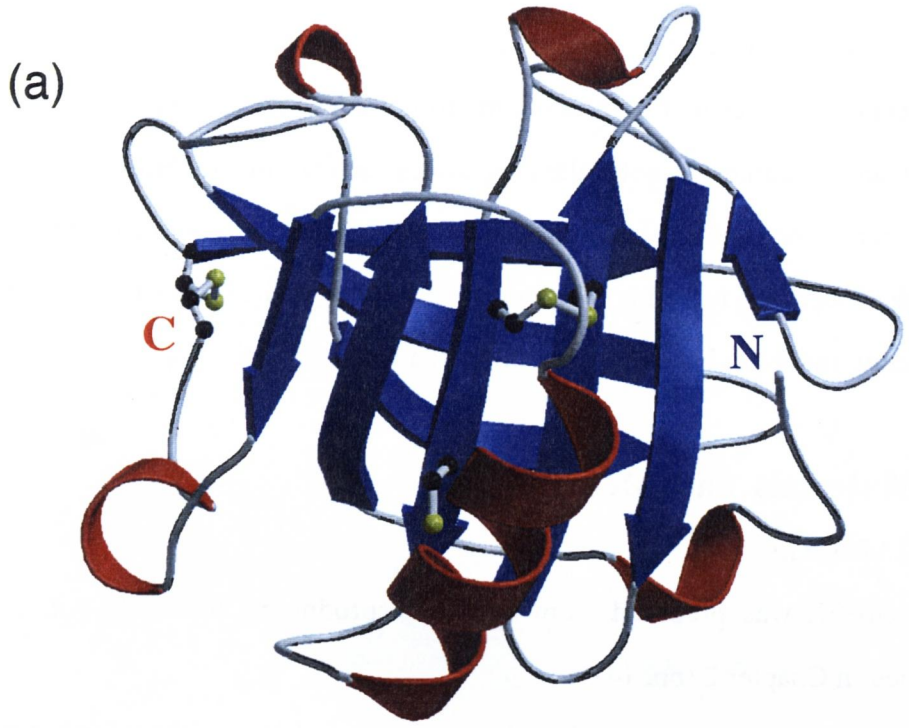
Interaction of GroEL with β-lactoglobulin Derivatives

3.1. Introduction

Although it is believed that, molecular chaperones bind proteins and promote their efficient folding, it is yet unclear whether they actively facilitate the folding of bound protein. In order to answer this question, β-lactoglobulin will be useful as a model substrate protein because the secondary structures of its folding intermediate state has been considered to be very different from those of the native state. Whereas cyt *c* is a basic protein (pI = 10) as described in Chapter 2, β-LG is an acidic protein (pI = 5). Therefore, by comparing the binding properties of the two proteins, I would also address the role of electrostatic interaction between substrate proteins and GroEL (pI = 5).

Bovine β-LG is a small globular protein of 162 amino acid residues with two disulfide bonds and one free thiol group (Figure 3.1a). From X-ray crystallography (Brownlow *et al.*, 1997), it has been assumed that Cys66-Cys160 and Cys106-Cys119 are connected by disulfide bonds, while remaining Cys121 exists as a free thiol. Although the native conformation of β-LG consists mostly of β-sheet structure, recent studies have revealed that this protein has a high tendency to form an α-helical structure in environment where short range interactions are dominant. In accordance with this high helical propensity, it has been demonstrated that the intermediate state with nonnative helical structure is accumulated during the folding process.

In the present study, I prepared tetramethylrhodamine-labeled β-LG in which tetramethylrhodamine was covalently linked to a single free thiol group of β-LG, as well as disulfide-bond reduced form of it. Then, I investigated the interaction between GroEL and these modified β-LGs. There is another important reason why I chose tetramethylrhodamine-labeled β-LG as a substrate protein. I chose tetramethylrhodamine among various fluorescent dyes as a fluorescent probe, because it emits strong fluorescence in the visible regions. Recently, visualization of single molecular event is progressing as a field for elucidating various



(b)

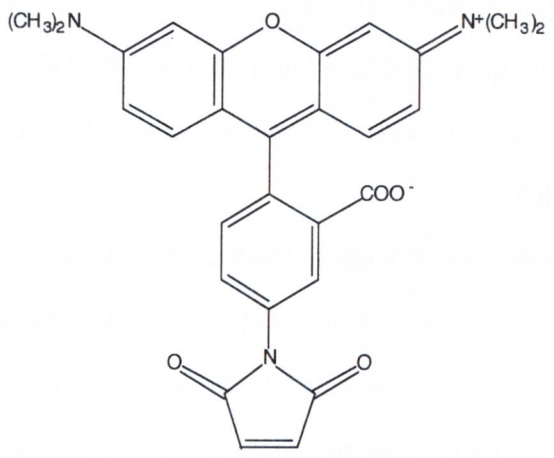


Figure 3.1. (a) Three dimensional structure of bovine β-lactoglobulin. The figure was created using the coordinates of Brownlow *et al.* (1997) and the program Molscript (Kraulis, 1991) and Raster3D (Merritt & Murphy, 1994). (b) Schematic drawing of tetramethylrhodamine-5-maleimide.

processes of biomolecules (Noji *et al.*, 1997; Funatsu *et al.*, (1995)). Imaging of single fluorescent molecules by fluorescence microscopy will be a promising approach to analyze the reaction and dynamics of molecular chaperones such as GroEL. However, the crucial technical requirement at present is that a fluorescent dye with strong fluorescence at visible regions must be used. Therefore, for using the fluorescence-labeled GroEL for future study imaging single fluorescent molecule, I chose tetramethylrhodamine.

3.2. Materials and Methods

3.2.1. Materials

GroEL was prepared from GroE-overproducing strain, *E. coli* DH1/pKY206, as described in Chapter 2 (pp. 14).

Bovine β-LG A was purchased from Sigma and used without further purification.

tetramethylhodamine-labeled β-LG (Rh-β-LG) was prepared by treatment of intact bovine β-LG with tetramethylrhodamine-5-maleimide (Figure 3.1b). Ten milligrams of β-LG was dissolved in 1 ml of 2 M urea and 20 mM Tris-HCl (pH 7.8). The reaction was started by adding 0.6 mg of tetramethylrhodamine dissolved in 5 μl of dimethylsulfoxide. The reaction was immediately quenched by adding 20 μl of 1 M HCl and subsequently the reaction mixture was applied to a PD-10 column (Pharmacia Biotech, equivalent to Sephadex G-25) equilibrated with 20 mM HCl.

Disulfide-reduced tetramethylhodamine-labeled β-LG (Rh-β-LG.red) was prepared by treatment of Rh-β-LG with dithiothreitol. One hundred microliter of Rh-β-LG solution (5 mg ml^{-1} in 20 mM HCl) was mixed with 900 μl of 2 M urea, 0.2 mg ml-1 of dithiothreitol and 20 mM Tris-HCl (pH 8.6), and incubated at room temperature for 15 min. After incubation, the reaction was quenched by adding of 10 μl of 1 M HCl and excess dithiothreitol was removed by gel-filtration chromatography through a column of PD-10 equilibrated with 20 mM HCl.

3.2.2. Methods

All spectroscopic measurements otherwise mentioned were carried out at 20°C with thermostatically controlled cell holders, as described in Chapter 2 (pp. 18).

3.2.2.1. Protein concentration

The concentration of tetramethylrhodamine-labeled β-LG was determined from the absorption at 276 nm after subtraction of the contribution of tetramethylrhodamine, using ε_{278} = 1.77 x 10^4 (Byler *et al.*, 1983) for β-LG and ε_{541} = 9.1 x 10^4 (in MeOH) for tetramethylrhodamine (see Figure 3.3, pp. 46).

3.2.2.2. HPLC measurements

The binding of substrate proteins to GroEL was detected by size exclusion chromatography experiments using a Gilson HPLC system, as described in Chapter 2 (pp. 18). Eluted proteins were detected by absorption at 220 nm for GroEL and fluorescence at 567 nm with excitation at 543 nm for Rh-β-LG and Rh-β-LG.red.

3.2.2.3. Fluorescence measurements

Fluorescence spectra were measured with a Hitachi F4500 fluorescence spectrophotometer. The fluorescence of tetramethylrhodamine was measured at a protein concentration of 0.01 mg ml^{-1} with excitation at 556 nm.

3.3. Results

3.3.1. Solution structure of β-lactoglobulin and its disulfide derivatives.

Figure 3.2 shows the far-UV CD spectra of bovine β-LG. To facilitate the detection of the complex formation between substrate protein and GroEL, I prepared tetramethylrhodamine-labeled β-lactoglobulin (Rh-β-LG) in which a single free thiol group of Cys121 is covalently labeled with tetramethylrhodamine-5-maleimide while the other four disulfide bridged cysteines are unmodified (Figures 3.1b, 3.3). I also prepared disulfide-reduced tetramethylrhodamine-labeled β-lactoglobulin (Rh-β-LG.red) in which two disulfide bonds of Rh-β-LG are reduced. Figure 3.2 includes the spectra of these derivatives. The spectrum of intact β-LG showed a broad minimum at ~215 nm, indicating that this species consists of predominantly β-sheet structure. By contrast, in the spectrum of Rh-β-LG, broad minimum at ~215 nm was

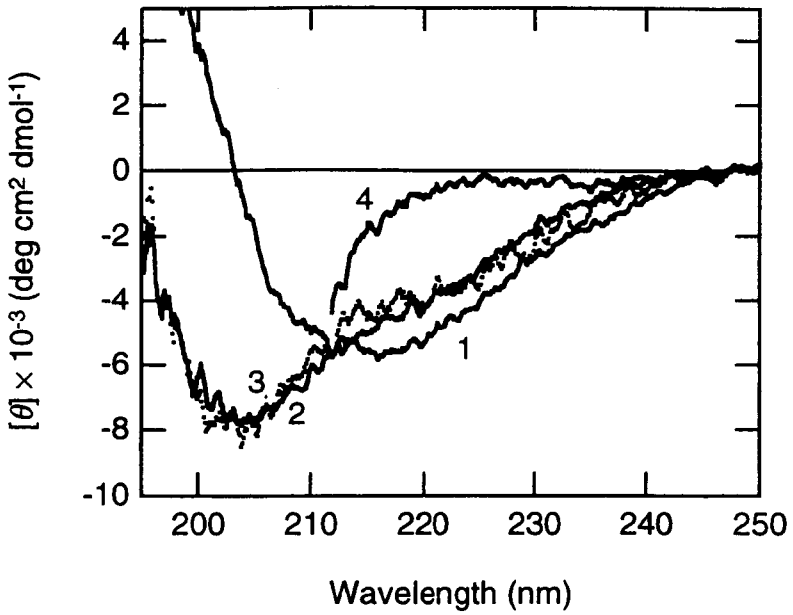


Figure 3.2. Far-UV CD spectra of bovine β-lactoglobulin and its derivatives at pH 7.0 and 20°C. 1. intact β-lactoglobulin; 2. Rh-β-LG (solid line); 3. Rh-β-LG.red (dotted line); 4. β-lactoglobulin in 6.4 M Gdn-HCl.

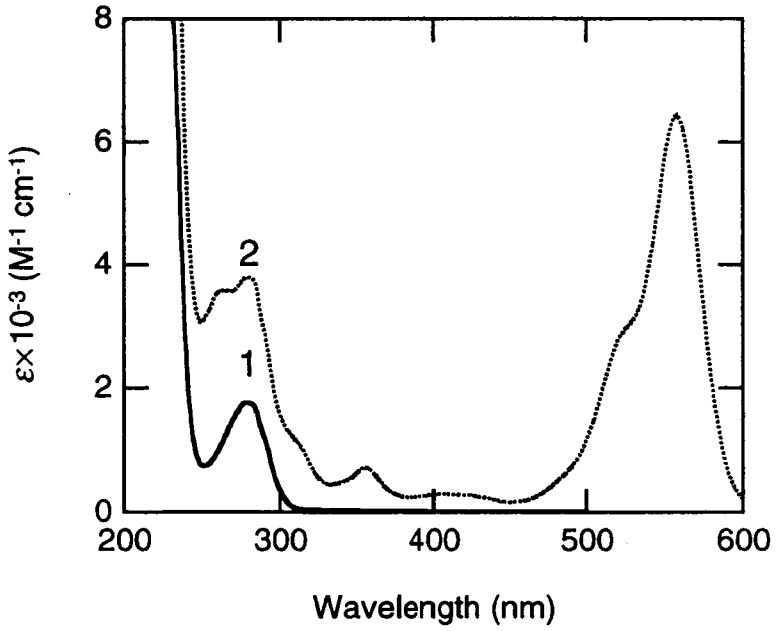


Figure 3.3. Absorption spectra of bovine β-lactoglobulin (1) and tetramethylrhodamine-β-lactoglobulin (2) at pH 7.0 and 20°C.

disappeared and, instead, rather sharp minimum at ~205 nm was marked. This spectral change indicates that the labeling of buried thiol group of Cys121 resulted in structural perturbation of the native conformation. This is also supported by the observation that, without destablization of the native structure by denaturant such as Gdn-HCl or urea, the free thiol group of β-LG was hardly reacted with thiol reactive reagents including tetramethylrhodamine-5-maleimide (data not shown).

On the other hand, the spectrum of Rh-β-LG.red was similar to that of Rh-β-LG, indicating that both species are disordered to a similar extent at least under this experimental condition. Judging from the CD spectra, although rhodamine-labeled species were largely disordered independent of the nature of the cysteine residues, they are much more structured than the fully unfolded state in a high concentration of Gdn-HCl. Interestingly, the spectra of Rh-β-LG and Rh-β-LG.red show some similarity to that of apo-cyt *c* (see Figure 2.5, pp. 23), suggesting the possibility that Rh-β-LG and Rh-β-LG.red also assume a compact denatured conformation as observed for apo-cyt *c* (see Hamada *et al.*, 1993)

3.3.2. The interaction between GroEL and disulfide-intact form of rhodamine-labeled β-lactoglobulin

I measured the interaction between GroEL and β-LG derivatives by size exclusion chromatography at various salt concentrations. The complex formation between GroEL and β-LG was monitored by the fluorescence of rhodamine (see Figure 3.8a, pp. 53). When Rh-β-LG or GroEL was applied separately to the column, the chromatogram showed a single peak at ~20 min or ~16 min, respectively (data not shown). On the other hand, when the mixture of Rh-β-LG and GroEL was applied to the column after pre-incubation of 15 min in the presence of 50 mM KCl, a small peak at ~16 min and a much larger peak at ~20 min were observed by fluorescence at 570 nm (Figure 3.4). Because almost the same small peak appeared in the chromatogram when GroEL alone was applied, this small peak at ~16 min does not derive from the complex but would be originated from the scattering of excitation light by large GroEL molecule. Thus, the results indicate that Rh-β-LG does not interact notably with GroEL at 50 mM KCl. However, by increasing salt concentration in the buffer up to 200 mM by KCl, the

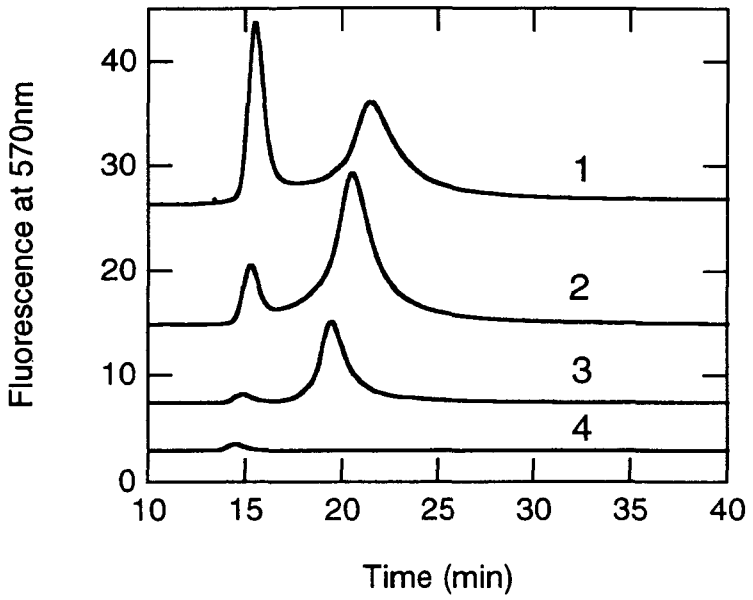


Figure 3.4. HPLC elution profiles of tetramethylrhodamine-β-lactoglobulin in the presence of GroEL at pH 7.0 and various concentrations of KCl. 1-3, mixture of GroEL and Rh-β-LG; 1, in the presence of 200 mM KCl; 2, 100 mM KCl; 3, 50 mM KCl; 4, GroEL.

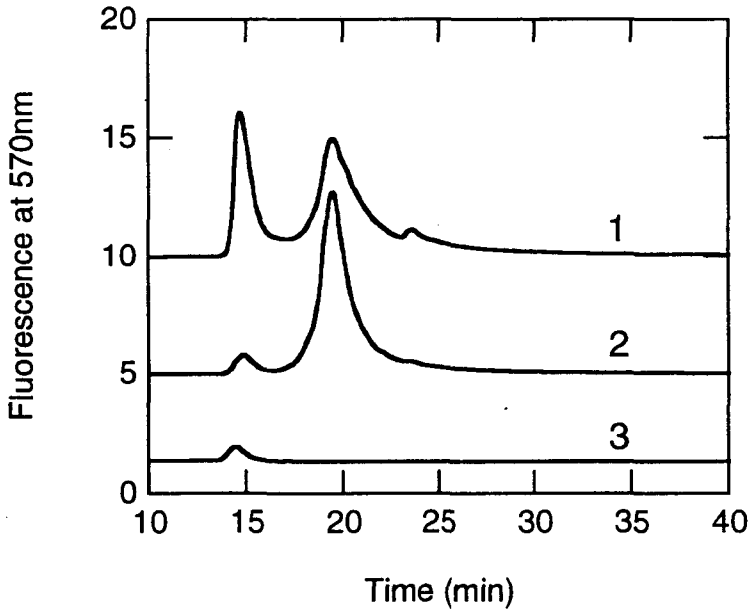


Figure 3.5. HPLC elution profiles of disulfide-intact and reduced form of tetramethylrhodamine-β-lactoglobulin in the presence of GroEL at pH 7.0 and 50 mM KCl. 1, Rh-β-LG.red; 2, Rh-β-LG; 3, GroEL.

peak at ~16 min increased significantly and, accordingly, the peak at ~20 min decreased. This indicates that, although the interaction between Rh-β-LG and GroEL is too weak to detect by HPLC when salt concentration is low, tight binding occurs when salt concentration is increased. Because both Rh-β-LG and GroEL are negatively charged under the conditions used (net charge of Rh-β-LG and GroEL at pH 7 is -6 and -18, respectively), the increase in salt concentration should shield the electrostatic repulsion between the negatively charged molecules, resulting in stabilization of the complex.

3.3.3. The interaction between Rh-β-LG.red and GroEL

Figure 3.5 shows chromatograms of the Rh-β-LG and Rh-β-LG.red after incubation with GroEL for ten minutes in the presence of 50 mM KCl. As can be seen in Figure 3.4, the chromatogram of the disulfide intact form, Rh-β-LG, did not show notable interaction with GroEL at this low salt concentration. In contrast, the elution profile of the disulfide reduced form, Rh-β-LG.red, showed a large peak at ~15 min, indicating that the interaction of Rh-β-LG.red with GroEL is stronger than that of the disulfide intact form of Rh-β-LG. Although Rh-β-LG.red could interact with GroEL even at this low salt concentration, the peak at ~15 min became higher with increase in salt concentration, as was the case of the disulfide intact form of Rh-β-LG, suggesting that the interaction between Rh-β-LG and GroEL is also affected by electrostatic interaction. The present results demonstrate a substantial difference in binding affinity for GroEL between Rh-β-LG and Rh-β-LG.red. However, the far-UV CD spectra of these molecules are quite similar, and both of these have little secondary structure, as shown in Figure 3.2. These results suggest that the secondary structure content is not a determinant for the interaction with GroEL. This view is consistent with the results obtained with cyt *c* derivatives: although holo-cyt *c* and porphyrin-cyt *c* have similar secondary structures, the affinity of porphyrin-cyt *c* toward GroEL is stronger than that of holo-cyt *c*.

3.3.4. Re-formation of the disulfide bonds

Although there was no obvious difference in the far-UV CD spectra between Rh-β-LG and Rh-β-LG.red, the affinity for GroEL of Rh-β-LG.red was much higher than that of the

disulfide intact form. To elucidate whether disulfide-bond formation does influence the affinity for GroEL, I observed the interaction between GroEL and the intermediates of disulfide-bond formation of Rh-β-LG by size exclusion chromatography. Figure 3.6a shows HPLC elution profiles of a mixture of GroEL and Rh-β-LG.red in the presence of glutathione. I chose the low salt conditions (i.e., 50 mM KCl) so that the disulfide-intact form of Rh-β-LG does not interact with GroEL. The fully disulfide reduced form of Rh-β-LG was first incubated in the presence of 5 mM oxidized glutathione and 2 mM reduced glutathione at pH 7.8 for various times in order to allow disulfide exchange reaction. After various incubation periods, aliquot was taken and disulfide exchange reaction was quenched by lowering pH and then GroEL was added. As shown in Figure 3.6, the prolonged incubation time in the glutathione redox buffer resulted in decrease in peak intensity at ~15 min and, correspondingly, peak at ~20 min has gradually increased. From the ratio of the peaks at ~15 min and at ~20 min, I evaluated the affinity change for GroEL during the disulfide exchange reaction (Figure 3.6b). Because the disulfide exchange reaction was carried out in an environment where the native conformation is stable, the molecule will reach to a species with correct disulfide bonds through several kinds of intermediates with nonnative disulfide pairs. Therefore, the results show that, while the cleavage of disulfide-bonds increases the affinity of Rh-β-LG for GroEL, re-formation of correct disulfide-bond reversibly decreases the affinity of Rh-β-LG for GroEL, suggesting that the complex formation is determined by a factor related to the extent of correct disulfide bond formation.

3.3.5. Interaction with GroEL of refolding intermediate of Rh-β-LG

It has been suggested that the non-native α-helical intermediate state accumulates in the folding pathway of bovine β-LG, through which native β-sheet structure is formed (Hamada *et al.*, 1996b; Kuwajima *et al.*, 1987, 1996). It has not yet been understood whether this nonnative helical intermediate state is a major on-pathway intermediate or off-pathway trapped state. Even if it is a kinetically trapped state, it might be appropriate to consider that this compact helical intermediate state is on a closest detour to the native state rather than it climbs up the energy funnel to the original unfolded state. It is intriguing to examine the interaction

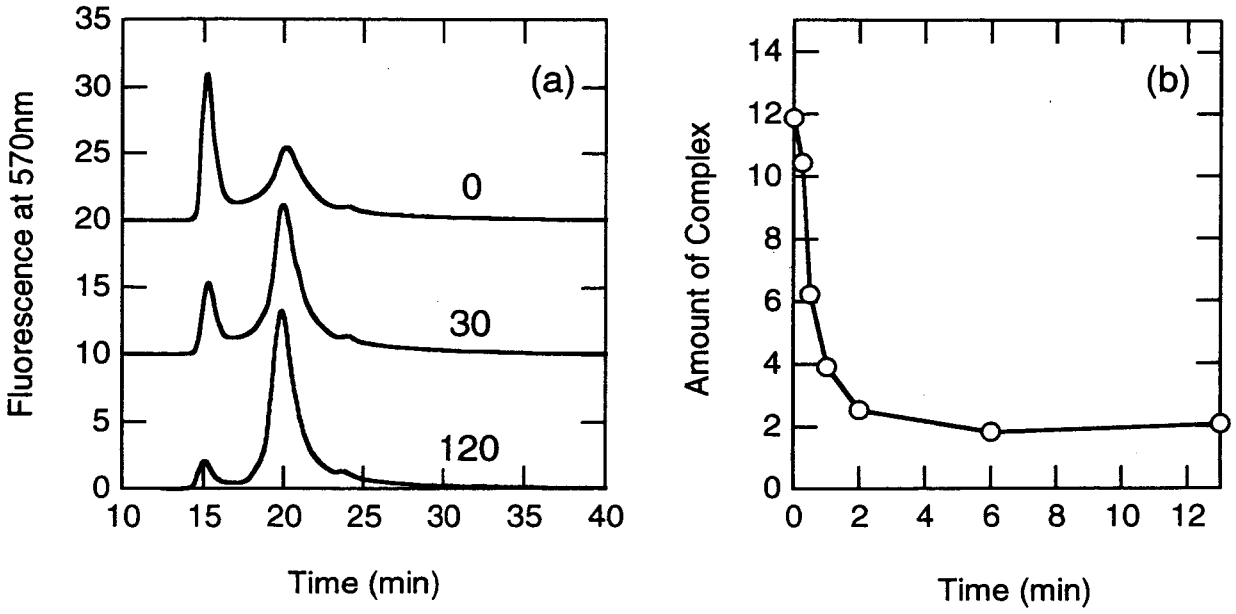


Figure 3.6. Decrease of the affinity for GroEL of disulfide-reduced tetramethylrhodamine-β-lactoglobulin induced by disulfide bond formation. (a) HPLC elution profiles of mixture of Rh-β-LG.red and GroEL in the presence of glutathione. The figures indicate the reaction time in min. (b) Decrease of the peak intensity of the complex between GroEL and Rh-β-LG.red during the reaction with glutathione.

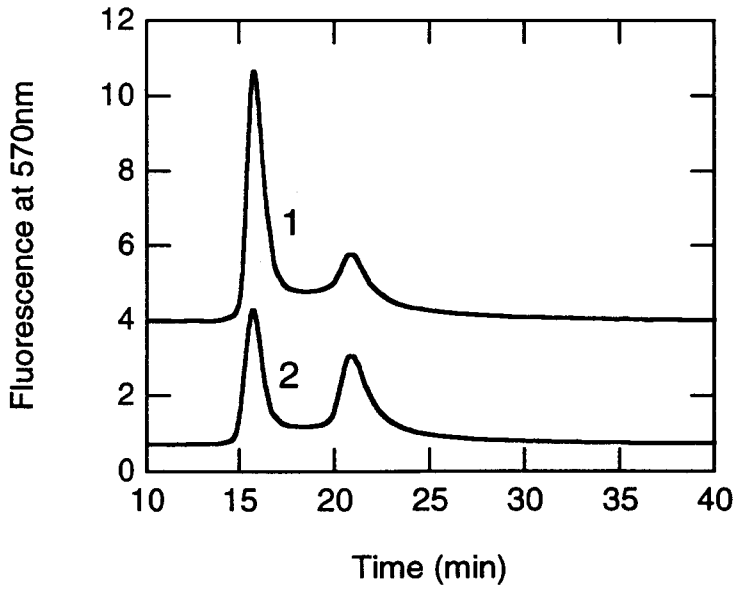


Figure 3.7. Interaction between GroEL and refolding intermediate from Gdn-HCl denatured state of tetramethylrhodamine-β-lactoglobulin in the presence of 200 mM KCl at pH 7.0 and 20°C. 1, Denatured Rh-β-LG in 4 M Gdn-HCl and GroEL; 2, Control: Non-denatured Rh-β-LG and GroEL.

between such a nonnative intermediate and GroEL in order to understand the functional role of molecular chaperone.

Figure 3.7 shows the interaction between GroEL and the refolding intermediate of Rh-β-LG in the presence of 200 mM KCl. Rh-β-LG unfolded by 4 M Gdn-HCl was rapidly diluted hundred times by a buffer (200 mM KCl and 20 mM Tris-HCl, pH 7.0) containing GroEL, and applied to a size-exclusion column. Figure also shows the result of control experiment where Rh-β-LG in non-denaturing condition was diluted and mixed with GroEL in the same manner. In the case of control experiment, the chromatogram shows two peaks at ~15 min and ~20 min, indicating a relatively weak interaction between GroEL and Rh-β-LG (see also Figure 3.3a, pp. 46). The HPLC elution profile of the refolding intermediate from Gdn-HCl also shows two peaks at ~15 min and ~20 min. Clearly, peak intensity at ~15 min was increased, and correspondingly, peak at ~20 min was decreased, showing that the refolding intermediate of Rh-β-LG interacts with GroEL more tightly than the non-denatured state of Rh-β-LG. The results suggest that the nonnative α-helical intermediate which accumulates in the folding pathway of β-LG can interact with GroEL, confirming the idea that the type of secondary structure is not a determinant for GroEL binding.

3.3.6. Interaction with GroEL studied by fluorescence analysis

As shown in Figure 3.4 (pp. 48), Rh-β-LG interacts with GroEL in the presence of 200 mM KCl. In order to evaluate quantitatively the interaction between them, I measured the change in fluorescence spectrum of rhodamine upon binding. Figure 3.8a shows fluorescence spectra of Rh-β-LG in the presence of various concentrations of GroEL at 200 mM KCl. In the absence of GroEL, the fluorescence spectra of Rh-β-LG showed a maximum at 574 nm. The peak intensity varied depending on the concentration of GroEL, and increased as much as twofold in the presence of excess amount of GroEL. I assumed that this spectral change reflects the binding between Rh-β-LG and GroEL. Then, I measured the fluorescence spectra in the presence of various concentrations of GroEL and plot its peak intensity against GroEL concentration to draw a titration curve. As shown in Figure 3.8b, the fluorescence intensity increased linearly at the beginning of the titration curve, and then the slope gradually decreased

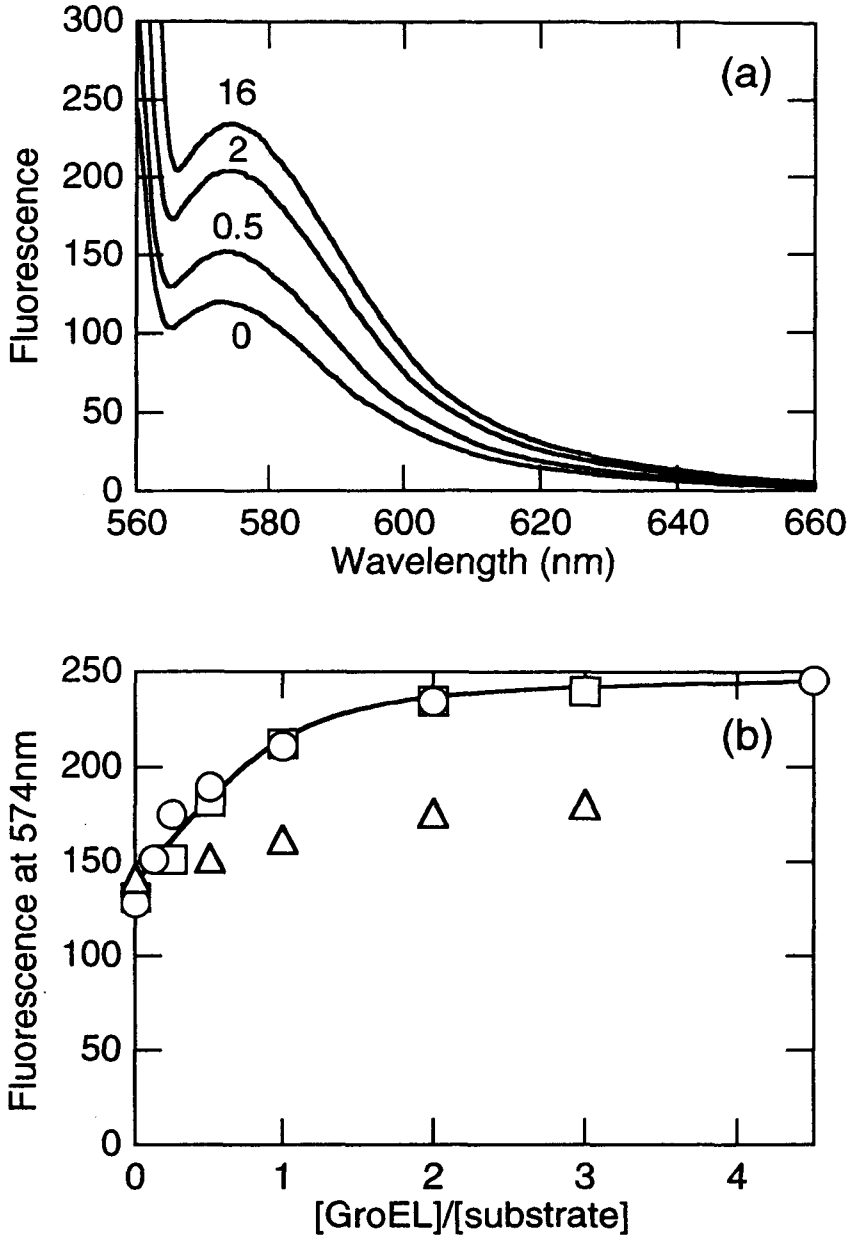


Figure 3.8. (a) Fluorescence spectra of tetramethylrhodamine-β-lactoglobulin in the presence of various concentrations of GroEL at pH 7.0 and 20°C. Figure indicates the ratio of [GroEL]/[substrate]. (b) Changes in fluorescence intensity at 574 nm in the presence of nucleotide. (○) in the absence of nucleotide; (△) in the presence of 2 mM ATP; (□) in the presence of 2 mM ADP. The line indicates the theoretical binding curve with K_b = 2 x 10^7 M^{-1} and a 1 : 1 stoichiometry (Materials and Methods, Scheme 1).

and finally saturated. By assuming that one GroEL molecule binds one molecule of Rh-β-LG (i.e., a 1:1 binding scheme; see Material and Methods in Chapter 2, pp. 21), I analyzed this titration curve. Consequently, the binding constant between them was estimated to be 2 x 10^7 M^{-1}. I also carried out the same experiment in the presence of ATP or ADP and the results are also included in Figure 3.6b. While the titration curve in the presence of 2 mM Mg-ADP was quite similar to that in the absence of nucleotide, little increase in fluorescence intensity was observed in the presence of 2 mM Mg-ATP. This result is consistent with the idea that ATP binding/hydrolysis induces the substrate release from GroEL, suggesting that the interaction between Rh-β-LG and GroEL must be specific and closely related to the GroEL function.

3.4. Discussion

3.4.1. Electrostatic interaction between GroEL and β-LG

In the present chapter, I showed that elevating salt concentration results in increase the affinity of Rh-β-LG for GroEL, in contrast to the case for cyt *c*, in which the affinity for GroEL is decreased when salt concentration is raised (Chapter 2.4.2, pp. 35). This apparently opposing effects of salt concentration are explained by electrostatic interaction between GroEL and substrate proteins. When salt concentration in the buffer is increased, electrostatic attractive or repulsive forces between charged molecules with the opposing or the same signs, respectively, will be weakened. Under the conditions examined (pH 7), both β-LG (pI = 5) and GroEL (pI = 5) is negatively charged while cyt *c* (pI = 10) is positively charged. Therefore, interaction between β-LG and GroEL will be apparently strengthened with increasing salt concentration because the electrostatically repulsive forces preventing complex formation is weakened. On the other hand, complex formation between cyt *c* and GroEL will be suppressed when salt concentration is increased, because electrostatically attractive forces between them is weakened (see also Chapter 2.4.2, pp. 35).

3.4.2. Conformational characteristics of β-LG

The present work showed that disulfide-reduced form of Rh-β-LG interacts with GroEL more tightly than the disulfide-intact form. This was evidenced by the results that, whereas Rh-

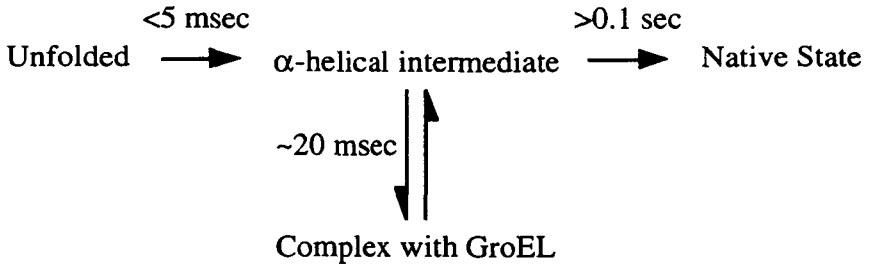
β-LG could interact with GroEL only in the presence of high concentration of salt, Rh-β-LG.red interacted with GroEL at any salt concentrations examined. Nevertheless, it seems that the two forms have very similar conformations at least examined by far-UV CD spectra. As mentioned in the Chapter 2.4.3, (pp. 36; see also Figure 2.12, pp. 39), it is possible to consider that the conformational state of a certain protein which binds to GroEL might be different from the structure in solution. In other words, the binding to GroEL might induce some structural change on the substrate proteins. Zahn *et al.* (1996) examined the conformational states of barnase bound to GroEL and SecB by using hydrogen-deuterium exchange combined with NMR and indicated that barnase binds to GroEL in a fully unfolded state. Similarly, Nieba-Axmann *et al.* (1997) examined the interaction between cyclophilin A and GroEL by using NMR spectroscopy and suggested that cyclophilin A with native-like structure undergoes multiple cycles of unfolding while bound to GroEL.

The disulfide bond, which covalently connects one region to another in a linear polypeptide chain, often restricts largely the conformational space of fully unfolded state, resulting in the stabilization of the native conformation or prevention of further unfolding. Indeed, it was suggested that the four disulfide bonds of hen egg white lysozyme, 14.5 kDa of globular protein, prevents fully unfolding even in the presence of 4 M Gdn-HCl or 40% (v/v) of trifluoroethanol (Hoshino *et al.*, 1997). If a substrate protein actually undergoes unfolding reaction upon binding to GroEL, it is possible to consider that the two disulfide bonds existing in Rh-β-LG prevent forming a fully unfolded state, and, as a result, the affinity for GroEL toward Rh-β-LG is not as strong as the disulfide reduced form. It is also possible to consider that the binding of a substrate protein to GroEL is an event in which already exposed hydrophobic regions on a substrate protein interact with GroEL, and, in concomitant with this, some structural rearrangements occur in the substrate, or simultaneously breaking intramolecular contacts to make new intermolecular contacts (Figure 2.11, pp. 37). The unfolding of a substrate protein might be also effective to increase the surface area that is available to contact with the flexible binding site on GroEL.

3.4.3. Interaction between GroEL and refolding intermediate of Rh-β-LG

Although the native structure of β-LG was disordered by labeling its free thiol group with tetramethylrhodamine, the refolding reaction from the fully denatured state in the presence of 4 M Gdn-HCl is considered to proceed through an α-helical intermediate, similarly to the non-labeled form. It is known that, although the native conformation of bovine β-LG is consisting of largely β-sheet structure, it has a high helical preference for α-helical structure (Hamada *et al.*, 1995; Hamada & Goto, 1997; Kuroda *et al.*, 1996; Shiraki *et al.*, 1995), and the refolding reaction of this protein is considered to proceed through a nonnative partly α-helical intermediate (Hamada *et al.*, 1996b; Kuwajima *et al.*, 1987, 1996). This helical intermediate accumulates within the dead-time of CD stopped flow mixing, and its conversion to the native β-sheet structure proceeds more slowly According to Hamada *et al.* (1996b), the refolding reaction of β-LG is consisted of at least three phases, and for the fastest phase, k_1 is 9.6 s^{-1}. In contrast, the binding reaction between GroEL and its substrate generally proceeds much faster, the binding rate constant of apo-cyt *c* and GroEL was estimated to be 7.8 x 10^7 $M^{-1}s^{-1}$, and similar value was reported for the disulfide reduced α-lactalbumin (k_1 = 1 x 10^6 $M^{-1}s^{-1}$; Murai *et al.* 1995), the MG state of α-lactalbumin (k_1 = 1 x 10^6 $M^{-1}s^{-1}$; Katsumata *et al.*, 1996) and for barnase (k_1 = 1.3 x 10^8 $M^{-1}s^{-1}$; Gray & Fersht). All of these reaction occur by the rate constant of the order of k_1 = 10^6 - 10^8 $M^{-1}s^{-1}$.

Providing that the binding reaction between GroEL and Rh-β-LG proceeds in the similar rate, a half time of the reaction (time needed for half of Rh-β-LG molecules bind to GroEL) was estimated to be $t_{1/2}$ = 20 msec using the following parameters; [Rh-β-LG] = 0.3 μM, [GroEL] = 0.6 μM, k_1 = 7.8 x 10^7 $M^{-1}s^{-1}$, k_{-1} = 1.9 s^{-1}. On the other hand, the conformational transition from α-helical intermediate to the native β-sheet structure of β-LG is consisted of at least three phases, and the fastest phase has a k_1 value of 9.6 s^{-1}. Therefore, the overall conversion must be slower than 0.1 s (= $1/k_1$). On the other hand, the α-helix formation from the fully unfolded state is a much faster process, completing within the dead-time of the stopped flow mixing (less than 5 msec). Taken together, the refolding process and binding reaction with GroEL of β-LG can be considered by the following scheme:



Scheme 2

The present work suggests that, although the non-denaturing form of Rh-β-LG shows some interaction with GroEL, the refolding intermediate from fully denaturing condition of Rh-β-LG possesses much higher affinity for GroEL.

According to the above scheme, in the refolding reaction from the fully unfolded state in the presence of 4 M of Gdn-HCl, Rh-β-LG rapidly adopts an α-helical intermediate, and then this α-helical intermediate preferentially binds to GroEL. It is suggested that such an α-helical intermediate of bovine β-LG is accumulated in an environment where the local interaction is predominant, including the early stage of the folding reaction, where non-local interaction is marginal. From their theoretical studies on the basis of simple exact models, Dill and coworkers (Chan & Dill, 1990; Dill *et al.*, 1995) suggested that the collapse of protein molecule drives the formation of secondary structure. From this point of view, the early α-helical intermediate of β-LG is also considered to be a compact intermediate. However, it would be a collapsed state without specific non-local interaction. In such a collapsed state, many hydrophobic residues would not be buried properly, and a correct specific packing of the molecule would be prevented. It is considered that these collapsed but not-correctly packed state with exposed hydrophobic surfaces may be preferentially bound by GroEL.

Chapter 4

High Level Expression of Bovine β-Lactoglobulin in *Pichia pastoris*

4.1. Introduction

In order to obtain further information of the conformational properties necessary for binding to GroEL, analysis of the bound and free substrate proteins by physicochemical methods with high structural resolution such as heteronuclear NMR will be useful. Therefore, I tried to analyze the solution structures of the native and α-helical intermediate of β-LG at atomic resolution by heteronuclear multidimensional NMR with a recombinant bovine β-LG expressed by methylotropic yeast *Pichia pastoris*.

It is noted that the researches described in Chapter 4 and Chapter 5 have been performed as a collaboration between Drs. K. Kuwata (Gifu Univ.), C. A. Batt (Cornell Univ.), Y. Goto (Osaka Univ.) and myself. Dr. Batt constructed a *P. pastoris* strain which secretes recombinant β-LG by fusion of the cDNA to the α-mating factor prepro-leader from *Saccharomyces cerevisiae*. Dr. Kuwata mainly performed the NMR measurements of recombinant β-LG. I carried out the fermentation of *P. pastoris*, preparation and physical characterization of the recombinant β-LG.

Recombinant bovine β-LG was first expressed in *Escherichia coli* (Batt *et al.*, 1990) and *Saccharomyces cerevisiae* (Totsuka *et al.*, 1990). More recently, it has also been expressed in *Kluyveromyces lactis* (Rocha *et al.*, 1996). We pursued an alternative host for high-level expression of recombinant bovine β-LG, since when produced in *E. coli*, it formed inclusion bodies, while the yields in other systems were not deemed sufficient for a number of biophysical studies. The methylotropic yeast, *P. pastoris* was selected as the expression host (Cregg & Higgins, 1995; Romanos, 1995) because of its ability to grow to very high cell density (Cregg & Higgins, 1995) while producing up alcohol oxidase up to 30% of its total soluble protein when fully induced (Couderc & Baratti, 1980; Cregg *et al.*, 1985). It also

secretes very little of its own protein, simplifying purification of any heterologous secreted protein (Barr *et al.*, 1992).

Using the secretion system of *P. pastoris* by connecting a prepro α-mating factor (α-MF prepro) secretion signal derived from *Saccharomyces cerevisiae* to the N-terminal of the protein of interest, we first established the high expression system of recombinant β-LG. The yield of the protein was reached at the order of exceeding several hundred milligrams. Then, we examined the solution structure of recombinant β-LG by CD and 2D-NMR. Both techniques revealed that the structure and stability of β-LG expressed in *P. pastoris* is practically identical to that derived from its natural source, bovine.

4.2. Materials and Methods

4.2.1. Materials

4.2.1.1. Strains and plasmids

E. coli TG1 was used as a host strain for constructing BLG/pPIC9. pPIC9 contains the alcohol oxidase (AOX) I promoter, His^+ selectable marker and prepro α-mating factor (α-MF prepro) secretion signal derived from *Saccharomyces cerevisiae* (Figure 4.1). Plasmid pTTQ18BLG, the source of the β-LG cDNA, was prepared from *E. coli* TG1 (Batt *et al.*, 1990). *P. pastoris* GS115 was selected as a host strain for expression (Cregg *et al.*, 1985).

4.2.1.2. DNA manipulation

Plasmid DNA was isolated using a Qiagen plasmid preparation kit (Santa Clarita). Restriction enzymes were purchased from New England Biolabs and ligase from GIBCO BRL. The purification and the ligation of digested plasmid or PCR products were performed by the procedures as described (Ausubel *et al.*, 1990).

4.2.1.3. Medium

E. coli TG1 was grown in Luria broth (Ausubel *et al.*, 1990) containing ampicillin (100 mg/l). BMGY (Buffered Glycerol-complex Medium; 1% yeast extract, 2% peptone, 0.1 M potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 4 x 10^{-5}% biotin and 1% glycerol)

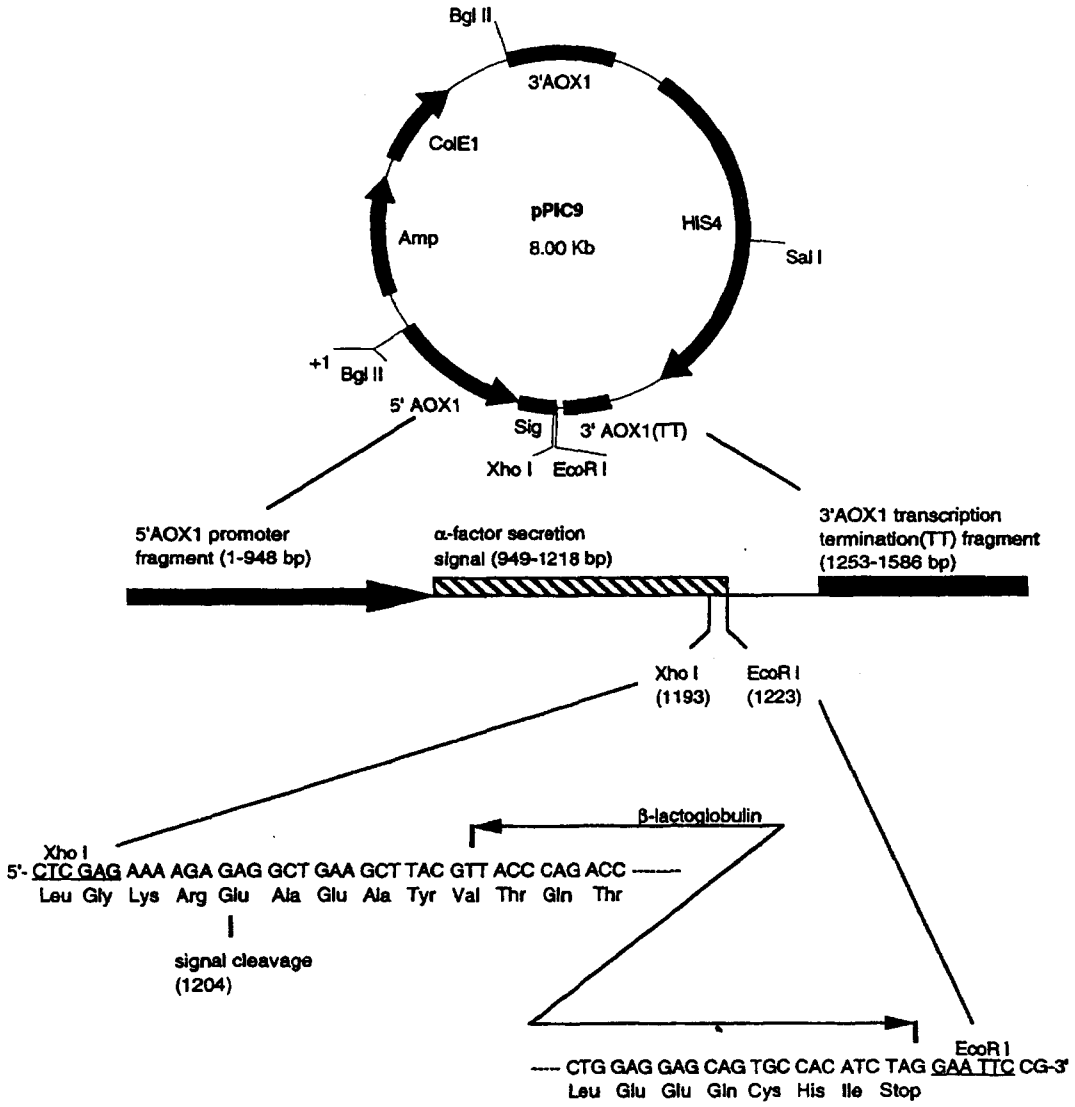


Figure 4.1. Restriction map of BLG/pPIC9.

and BMMY (Buffered Methanol-complex Medium; BMGY except 0.5% methanol was used instead of glycerol) were used for growing *P. patroris* and producing recombinant β-LG, respectively. MD (Minimal Dextrose Medium; 1.34% yeast nitrogen base, 4 x 10^{-5}% biotin, 1% dextrose) agar was used for screening of His^+ transformants and MM (Minimal Methanol Medium; MD except 0.5% methanol was added instead of dextrose) for Mut (Methanol Utilization) screening.

4.2.2. Methods

4.2.2.1. Construction of expression vector

A 480 bp DNA fragment encoding bovine β-LG was amplified by PCR from pTTQ18BLG (Batt *et al.*, 1990), with the following oligonucleotide primers: 5'CCGCTCGAGAAAAGAGAGGCTGAAGCTTACGTTACCCAGACC-3 (5'XhoI) and 5'CGGAATTCCTAGATGTGGCACTGCTCCTCCAG-3 (3'EcoRI). The PCR product purified using a Qiagen kit and after *Xho*I, *Eco*RI digestion, it was cloned into the *Xho*I and *Eco*RI sites of pPIC9, generating BLG/pPIC9 (Figure 4.1).

4.2.2.2. Transformation of *P. pastoris*

P. pastoris GS115(His^-) was transformed with BLG/pPIC9 digested with *Bgl*II. Approximately 0.3 μg of the linearized plasmid DNA was used for electroporation in 0.2 cm cuvettes at 1.5 kV, 25 mF, 400 W, using a Bio-Rad Gene Pulser (Becker *et al.*, 1991). Immediately after pulsing, 1 ml of cold 1 M sorbitol was added to the cuvette. Cells were plated onto MD agar for the selection of His^+ transformants. Transformants containing integrated BLG/pPIC9 were screened as follows: 4 ml of BMGY was inoculated a His^+ colony and incubated for 1 day at 30°C. A 1 μl aliquot from 20 μl of cell lysate was used for PCR with 5'XhoI and 3'EcoRI primers. To screen for methanol utilization, 2.5 μl of each culture was spotted onto MM agar plate.

4.2.2.3. Fermentation of *P. pastoris* [BLG/pPIC9]

The fermentation of *P. pastoris* [BLG/pPIC9] was carried out using a KMJ-2C fermenter (Mitsuwa Biosystems). The seed culture was prepared by inoculation from a frozen

glycerol stock of GS115/His$^+$/BLG/Mut$^+$ into 2 ml of BMGY. After overnight incubation at 30°C, the culture was transferred into 100 ml of the same medium and grown to an OD_{600} of 2 to 6. The seed culture was transferred into 1 liter of the same medium. The pH of medium was adjusted to 6.0 with potassium hydroxide. The culture was grown for approximately 15 hrs until an OD_{600} reached 10 to 15. To induce AOX1 transcription, the culture was centrifuged and the precipitated cell was suspended into 600 ml of BMMY. Fifteen milliliter of methanol was supplemented every 12 hours into the culture, and further incubated for 48-72 hrs. Samples were taken from the fermenter at different times and the cell density and protein level were measured.

4.2.2.4. Purification of β-LG

The culture supernatant was harvested 48-72 hrs after the start of methanol induction by centrifugation at 1,500 x g for 20 min and approximately 800 ml of supernatant was concentrated to 150 ml by pressure filtration using YM10 membrane (Amicon). The concentrated protein solution was dialyzed against 200 mM NaCl, 50 mM Gly-HCl (pH 3.0) for three times to decrease salt concentration in the solution. After dialysis, the protein was applied to a 2.5 x 25 cm CM-Sepharose CL-6B column (Pharmacia) equilibrated with the same buffer. β-LG was eluted with a linear gradient of 0.2-1.0 M NaCl in 50 mM Gly-HCl (pH 3.0). A 30 ml fraction of purified recombinant β-LG was dialyzed against 20 mM HCl and then lyophilized. Purity of the protein was examined by SDS-PAGE and reverse-phase HPLC using a μ-BONDASPHERE 5μC4-300Å column (Waters). The N-terminal sequence was determined by automated Edman degradation using ABI 437A protein sequencer (Applied Biosystems).

4.2.2.5. CD measurements

The CD spectra were measured using a Jasco spectropolarimeter, model J-720 at 20°C as described in Chapter 2 (pp. 20). The unfolding transition curves measured by the ellipticity at 218 nm were analyzed on the basis of a two-state transition mechanism between the native

state (N) and the unfolded state (U), N $\rightleftarrows$ U, and a linear dependence of the free energy change of unfolding (ΔG_U) upon Gdn-HCl concentration, [Gdn-HCl],

$$\Delta G_U = \Delta G_U(H_2O) - m[\text{Gdn-HCl}] \quad (6)$$

where $\Delta G_U(H_2O)$ is ΔG_U in the absence of denaturant and m provides a measure of the dependence of ΔG_U upon Gdn-HCl concentration (Myers *et al.*, 1995). Assuming manually the baselines for the native and unfolded states, the $\Delta G_U(H_2O)$ and m values were obtained by a least-squares curve fitting program.

4.2.2.6. NMR measurements

The ^{1}H-NMR spectra were measured using a Varian Inova 500 MHz spectrometers at 60°C. β-LG solutions were prepared by dissolving the lyophilized proteins in 10% (v/v) D_2O and 90% H_2O. The pH was adjusted to pH 2.0 at room temperature by adding 1.0 M HCl. No additional salt was present. Water suppression was accomplished using symmetrically shifted pulse (Smallcombe, 1993) with minimum presaturation. The 2D phase sensitive COSY spectra were recorded by the hypercomplex method (States *et al.*, 1982) with the following parameters: spectral width, 7500 Hz in each dimension and size, 4096 complex points in t_2 with 512 hypercomplex t_1 increments. Each dimension was processed with sine-bell and Gaussian functions. Acquisition time was 1 day for each spectrum.

4.3. Results

4.3.1. Construction of pPIC9 expression vector

The cDNA for β-LG was inserted in between the *Xho*I and *Sna*BI sites of pPIC9 3' to the secretion signal sequence using the 5'*Xho*I primer. The cDNA was placed in-frame with the 89 amino acids of *Saccharomyces cerevisiae* α-MF prepro (Kurjan & Herskowitz, 1982) producing a fusion between it and the α-MF prepro peptide. Heterologous proteins fused to the sequence are cleaved between Arg-85 and Glu-86 in the sequence Glu-Lys-Arg-Glu-Ala-Glu-Ala by the KEX2 endopeptidase which cleaves on the carboxyl side of dibasic residues (Julius

et al., 1984). The Glu-Ala spacer repeats at N-terminus that provide a hydrophilic environment at the KEX2 endopeptidase cleavage site, are then removed by the STE13 dipeptidyl aminopeptidase (Anna-Arriola & Herskowitz, 1994; Bussey, 1988; Julius *et al.*, 1983). The amino acid sequence (Tyr-Val) after Glu-Ala repeats was introduced by 5'*Xho*I primer slightly altering the N-terminal amino acid sequence of β-LG from ^{1}Leu-^{2}Ile-^{3}Val to ^{1}Ala-^{2}Tyr-^{3}Val. Nucleotide sequencing of BLG/pPIC9 with 5'AOX primer confirmed the junction between the secretion signal and β-LG cDNA.

4.3.2. Primary sequence

β-LG was purified using CM-Sepharose chromatography and a single band was observed on Coomassie Blue staining SDS-PAGE. Even with significant overloading of the SDS-PAGE, no extraneous proteins were observed. The purified β-LG migrated at the same position as purified bovine β-LG. A total of 6 amino acids were sequenced from the N-terminal of the secreted β-LG ($^{-3}$Glu-$^{-2}$Ala-$^{-1}$Glu-^{1}Ala-^{2}Tyr-^{3}Val) and they revealed that the α-MF prepro peptide was cleaved at the carboxyl side of the dibasic residue, Lys-Arg, but Glu-Ala spacer at the N-terminus remained.

4.3.3. CD spectra

Conformation of the recombinant β-LG was studied by CD and ^{1}H-NMR spectra. Whereas bovine β-LG exists as a dimer at neutral pH, it dissociate into monomers below pH 3, but still retains a native conformation even in an acidic environment as low as pH 2 (Fugate & Song, 1980). The far-UV CD spectrum of the recombinant β-LG at pH 3.0 and 20°C had a minimum at 217 nm, consistent with the abundance of β-sheet structures (Figure 4.2). The near UV-CD spectrum showed sharp peaks at 286 and 294 nm, indicating the unique tertiary structures. These spectral features of the recombinant β-LG were essentially the same as those of bovine β-LG A purchased from Sigma. Thus, the CD spectra indicate that the recombinant β-LG at pH 3.0 assumes unique native structure which is similar to that of bovine β-LG. I also measured the CD spectra of the recombinant and bovine β-LGs at pH 7.0. The spectra at pH 7.0 were similar to each other and were also similar to those at pH 3.0 (data not shown).

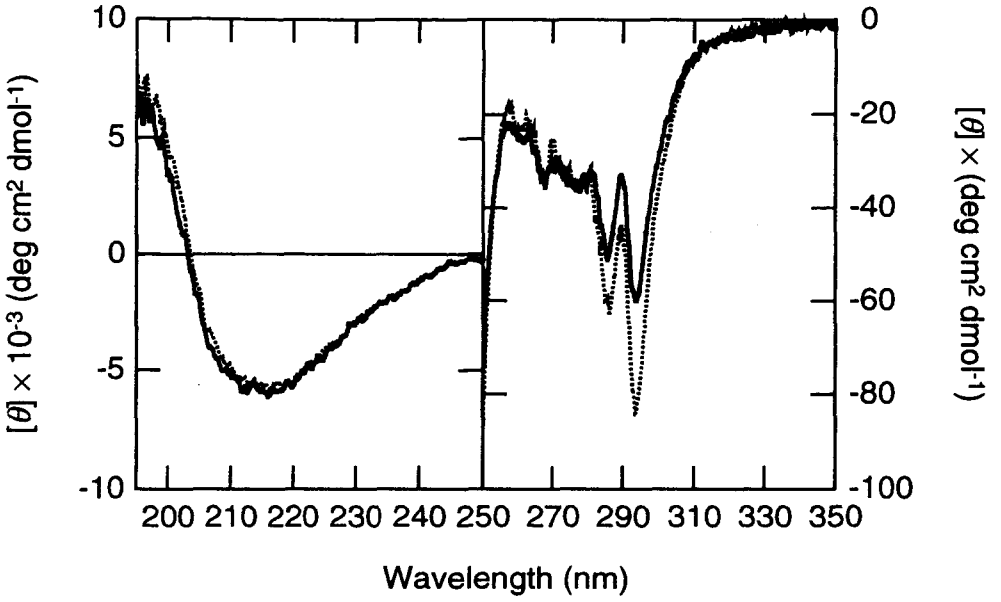


Figure 4.2. Far- and Near- UV CD spectra of β-lactogloblin derived from bovine (solid) and *P. pastoris* (broken) at pH 3.0 and 20°C.

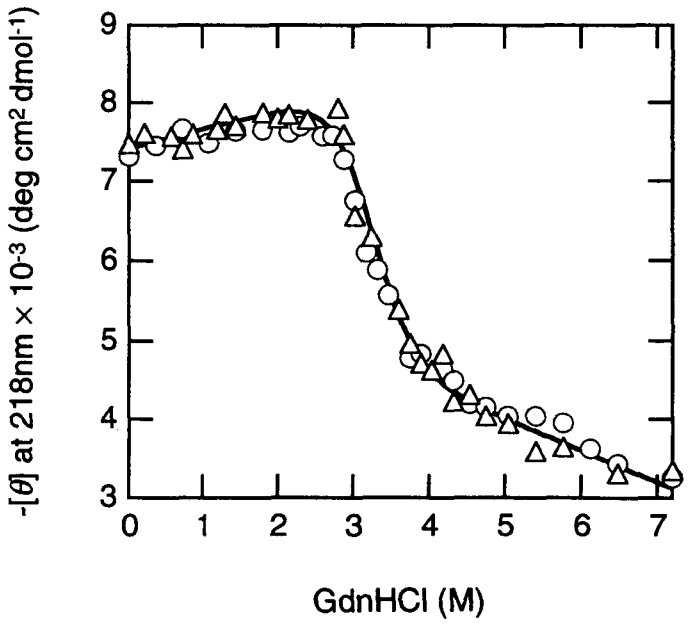


Figure 4.3. Unfolding transitions of the recombinant (○) and bovine (△) β-LGs at pH 3.0 and 20°C. The solid line indicates the theoretical curve on the basis of the two-state unfolding mechanism with the parameters described in the text.

4.3.4. Gdn-HCl unfolding profiles

In order to compare the conformational stability, the unfolding transitions of the recombinant and bovine β-LGs were measured by the ellipticity at 218 nm at pH 3.0 and 20°C (Figure 4.3). The unfolding transition curves of the two proteins were indistinguishable, again indicating the similarity of the two β-LGs. For both proteins, the unfolding started at about 2 M Gdn-HCl and major transition ended at 4 M Gdn-HCl, the apparent midpoint of the transition being 3.2 M Gdn-HCl. The unfolding transition of the recombinant β-LG at pH 3.0 was reversible (data not shown) as is the case of bovine β-LG (Hamada *et al.*, 1996b). The unfolding transition curves were analyzed on the basis two-state transition mechanism between the native state (N) and the unfolded state (U), and a linear dependence of the free energy change of unfolding (ΔG_U) upon Gdn-HCl concentration (see Materials and Methods). The $\Delta G_U(H_2O)$ and m values were estimated to be 28.4±0.1 kJ/mol and 8.8±0.1 kJ/mol/(mol of Gdn-HCl), respectively, for the recombinant β-LG and 33.4±0.1 kJ/mol and 10.6±0.1 kJ/mol/(mol of Gdn-HCl), respectively, for bovine β-LG.

4.3.5. NMR spectroscopy

To further compare the structural details, the 2D ^{1}H-NMR spectra of the recombinant and bovine β-LGs were measured at pH 2.0 and 60°C (Figure 4.4). The measurements at a high temperature was necessary to obtain a spectra with high peak resolution. The β-LGs were still native at 60°C and pH 2.0 on the basis of the CD spectra (data not shown). In fact, for both β-LGs, many αCH-NH cross peaks were clearly separated, consistent with the native structure with many β-sheets. The 162 amino acids residues of bovine β-LG A give rise to 156 αCH-NH cross peaks in the COSY spectrum in H_2O, i.e., 162 (total amino acid residues) - 8 (Pro) - 1 (N-terminal) + 3 (Gly), and up to 36 additional cross-peaks representing Arg and Lys side chain reasonances. A total of 128 cross-peaks were visible in Figure 4.4b. On the other hand, 165 amino acids of the recombinant β-LG give rise to 159 αCH-NH cross-peaks in the COSY spectrum, i.e., 165 (total amino acid residues) - 8 (Pro) - 1 (N-terminal) + 3 (Gly), and up to 36 additional cross-peaks representing Arg and Lys side chain reasonances. A total of 131 cross-peaks were visible in Figure 4.4a. Thus, we observed at least 80% of αCH-NH

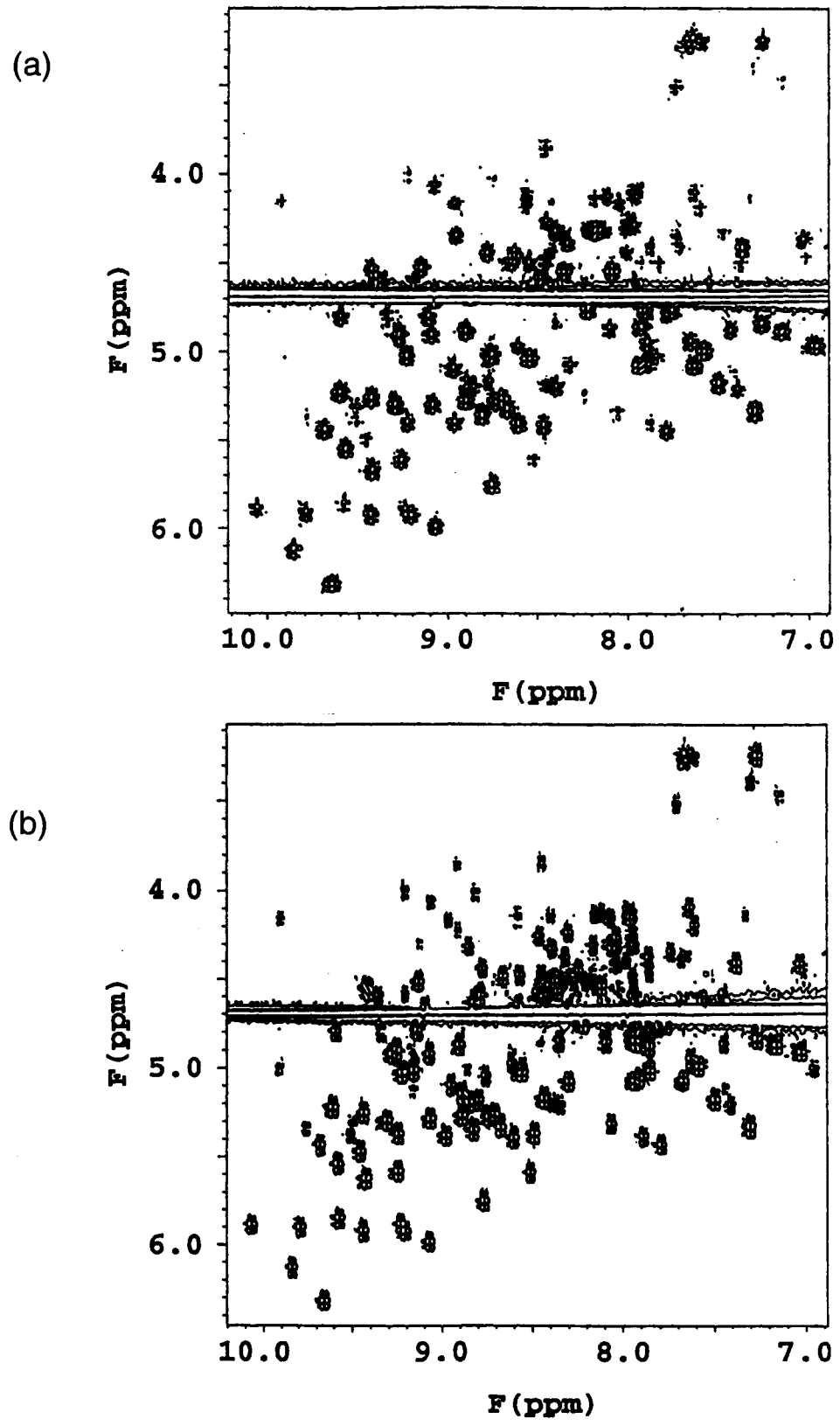


Figure 4.4. Finger print regions of phase sensitive COSY spectra of the recombinant (a) and bovine (b) β-lactoglobulin at pH 2.0 and 60°C.

cross-peaks in the finger print regions of COSY spectra of both β-LGs. Although the two spectra were very similar, several cross-peaks were evidently different, demonstrating the high sensitivity of the NMR methodology. Although the assignments of the peaks were not performed, these differences probably come from the difference in amino acid composition at the N-terminal region and the difference in three dimensional structure around the corresponding region. (See Chapter 5 for the assignments with heteronuclear NMR).

4.4. Discussion

β-LG has been expressed from both recombinant prokaryotic and eukaryotic hosts. Initial levels of β-LG expression in *E. coli* were low and the protein required solubilization followed by refolding. Typical yields were on the order of 8-10 mg/l (Batt *et al.*, 1990). Parallel efforts to express β-LG in *S. cerevisiae* using the glyceraldehyde 3 phosphate dehydrogenase promoter resulted in yields of 1.1 mg/l, but advantageously, this protein was secreted (Totsuka *et al.*, 1990). Finally β-LG has also been expressed in *K. lactis* using the phosphoglycerate kinase promoter and yields of 40-50 mg/l in secreted protein reported (Rocha *et al.*, 1996). In this latter effort, the native β-LG signal was used to direct secretion and proper processing observed. In the present study, recombinant β-LG was efficiently expressed from *P. pastoris* and copious amounts of fully folded protein on the order of several hundred milligrams could be recovered from the supernatant.

The prepro α-MF leader sequence derived from *S. cerevisiae* is effective at secreting and directing the processing of the recombinant β-LG in *P. pastoris*. Cleavage of the leader is apparently mediated through a KEX2-like activity in *P. pastoris*. On the other hand, the efficiency of the STE13 cleavage depends on the heterologous protein being expressed (Brake *et al.*, 1984). The Glu-Ala spacer residues, which provide a hydrophilic environment at the KEX2 cleavage site, are removed in certain cases, but not in other cases (Guisez *et al.*, 1991) suggesting a intrinsic role for the recombinant protein. Incomplete cleavage by the STE13-gene product observed in the present study might also be due to the occlusion of the cleavage size on the secreted protein from the peptidase (Hinchliffe & Kenny, 1993). β-LG exists as a dimer at physiological pH (Pessen *et al.*, 1985) and the negatively charged N-terminal may be involved

in dimer formation. Therefore, N-terminal of the partially folded recombinant β-LG may not be accessible to STE13.

CD spectra of β-LG produced by *P. pastoris* and bovine β-LG were almost the same except for a slight differences in the aromatic region (~295 nm; see Figure 4.2, pp. 65). This spectral change may be derived from a Tyr residue at the N-terminal of the recombinant β-LG, which was introduced by a mutation of ^{1}Leu-^{2}Ile-^{3}Val to ^{1}Ala-^{2}Tyr-^{3}Val. However, in the other aliphatic and aromatic regions, two spectra were indistinguishable, indicating that the differences in the N-terminal do not have a significant effect on the overall structure. In addition, it was shown that both spices have the same stability measured by Gdn-HCl denaturation experiments. This structural tolerance against the N-terminal difference is supported by the X-ray crystallographic structure, in which N-terminal four residues were not well defined (Brownlow *et al.*, 1997; Papiz *et al.*, 1986).

NMR results further confirm that the conformation of the recombinant and bovine β-LGs are practically the same in spite of the N-terminal differences. In particular, the COSY spectra (Figure 4.4, pp. 67) provide convincing evidence of structural similarity as well as the slight difference probably arising from the N-terminal region. These efforts comprise a very complete characterization of the recombinant β-LG as a prelude to using this system to study the structure-function of protein.

Protein structures are determined by the local and non-local interactions. β-LG is a unique example representing inconsistency of the local and non-local secondary structure preference (Hamada & Goto, 1997; Hamada *et al.*, 1995, 1996b; Kuroda *et al.*, 1996; Shiraki *et al.*, 1995). Detailed analysis of the folding mechanism of β-LG is of special importance for understanding the interplay between local and non-local interactions during protein folding. For such studies, heteronuclear NMR methods, which can provide the detailed structural information at the atomic level, will be essential. The COSY spectra with high peak resolution indicate that this protein is suitable for such NMR studies. The high level expression of the native β-LG in *P. pastoris* enables the efficient expression of the isotope-labeled proteins. In the next chapter, taking advantage of this high level expression system constructed in *P. pastoris*, I prepared the uniformly ^{15}N-labeled and, ^{15}N and ^{13}C double-labeled β-LGs, and

carried out further NMR characterization of the native state at pH 2 as well as the highly helical state in and 50% (v/v) trifluoroethanol solution.

Chapter 5

Heteronuclear NMR Characterization of the Native β-Sheet and TFE-induced α-Helical States of β-Lactoglobulin

5.1. Introduction

β-Lactoglobulin is a predominantly β-sheet protein consisting of nine anti-parallel β-strands and one α-helix (Brownlow *et al.*, 1997; Figure 3.1, pp. 43). β-LG has a high tendency to form α-helix, and it has been suggested to fold through α-helical intermediate, I(α): U → I(α) → N(β), where U and N(β) are the fully unfolded and native states, respectively (Kuroda *et al.*, 1995; Nishikawa & Noguchi, 1991; Shiraki *et al.*, 1995). However, role of the α-helical intermediate in the β-LG kinetic folding is unclear. Is it located in the folding pathway or an off-pathway intermediate? Is it a productive intermediate accelerating the folding reaction or a kinetic trap (Kiefhabor, 1995) decelerating the reaction?

In order to clarify the role of the helical intermediate in the folding of β-LG, it is essential to characterize the structure and dynamics of β-LG and its intermediate conformational states at the atomic level. The heteronuclear NMR is the best approach to perform such analysis. The heteronuclear NMR characterization of β-LG will be also important for characterizing the conformational states recognized by GroEL as described in Chapter 4. Therefore, I prepared the recombinant β-LG uniformly labeled with ^{15}N and ^{13}C using *Pichia pastoris* expression system, as described in Chapter 4. As a first step to elucidate the whole structural details of the refolding process of β-LG, I, in collaboration with Dr. K. Kuwata, assigned the resonance frequencies of backbone nuclei in aqueous solution (i.e., the native state) using heteronuclear 2D and 3D NMR spectra, and characterized both the secondary structures and dynamics. Because TFE stabilizes the α-helices of the kinetic intermediate (Hamada *et al.*, 1996b), TFE-induced structures would be useful to consider the kinetic intermediate in the absence of TFE. Thus, we then characterized the TFE-induced α-helical

state (i.e., the TFE state). The results confirmed the global conformations and dynamics of the TFE state and implicated the possible α-helical regions which exist in the kinetic intermediate.

Although I have not examined directly the interaction between GroEL and β-LG at atomic detail using NMR, the present results provide an scaffold for a further detailed analysis, such as pulse labeling H/D exchange experiment in order to elucidate the interaction between GroEL and kinetic α-helical intermediate of β-LG at atomic resolution.

5.2. Materials and Methods

5.2.1. Materials

Bovine β-LG was expressed in the methylotropic yeast *P. pastoris* with slight modifications from that described in Chapter 4 (pp. 59). That is, for the preparation of the uniformly doubly labeled β-LG, ^{15}N,^{13}C-labeled MD and MM, in which ^{15}N-labeled $(NH_4)_2SO_4$, ^{13}C-labeled glucose and ^{13}C-labeled methanol are contained in place of standard compounds, were used for growing *P. patroris* and producing recombinant β-LG, respectively. The secreted protein was purified as described, and I obtained 40 mg of doubly labeled β-LG.

One milligram of ^{13}C,^{15}N labeled β-LG was dissolved in 200 μl of 20 mM HCl (95% H_20/5% D_2O). pH was adjusted to 2.0 by adding HCl and sample was kept in 5 mm microcell (Shigemi). All the NMR spectra were recorded at 45°C and the acquisition time for each 3D spectrum was about four days with 2048 points in direct dimension.

5.2.2. Methods

NMR experiments were performed on a Varian Inova 500 spectrometer equipped with a triple resonance probe including shielded z gradients at the Instrument Analysis Center at Gifu University. Complete sequence-specific NMR assignments for recombinant β-LG were obtained based on scalar coupling connectivities. Assignments of backbone resonances were established using CBCA(CO)NH (Grzesiek & Bax, 1992) and HNCACB (Wittekind & Muller, 1993). HNCA and HN(CA)CO (Grzesiek & Bax, 1993) were acquired to obtain independent sequential linkage patterns to complement the above assignment. HCCH-TOCSY (Bax *et al.*,

1990), CCH-TOCSY (Fogh *et al.*, 1995), $(H^{\beta})C^{\beta}(C^{\gamma}C^{\delta})H^{\delta}$ and $H^{\beta}(C^{\beta}C^{\gamma}C^{\delta})H^{\delta}$ (Yamazaki *et al.*, 1993) were also acquired to refine the assignments. 3D ^{1}H-^{1}H-^{15}N NOESY experiments were performed varying the mixing time of 90, 120 and 200 msec. The order parameters and the effective correlation times are obtained using the programs coded by Dr. M. Akke at Columbia University. Data processing and peak picking were performed using the FELIX program version 95.0 (Biosym Technologies).

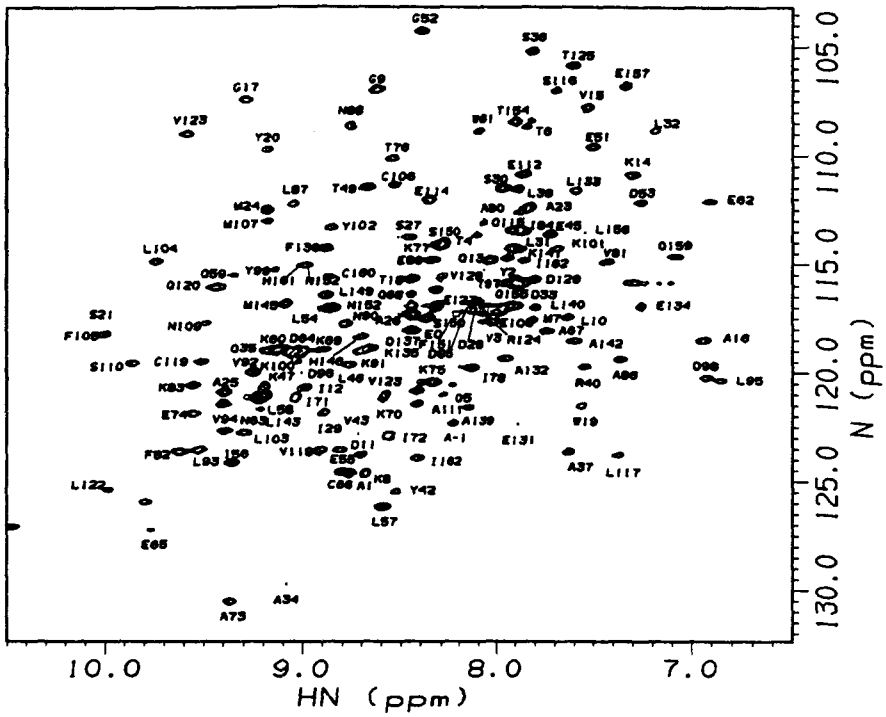
5.3. Results

5.3.1. Backbone resonance assignments

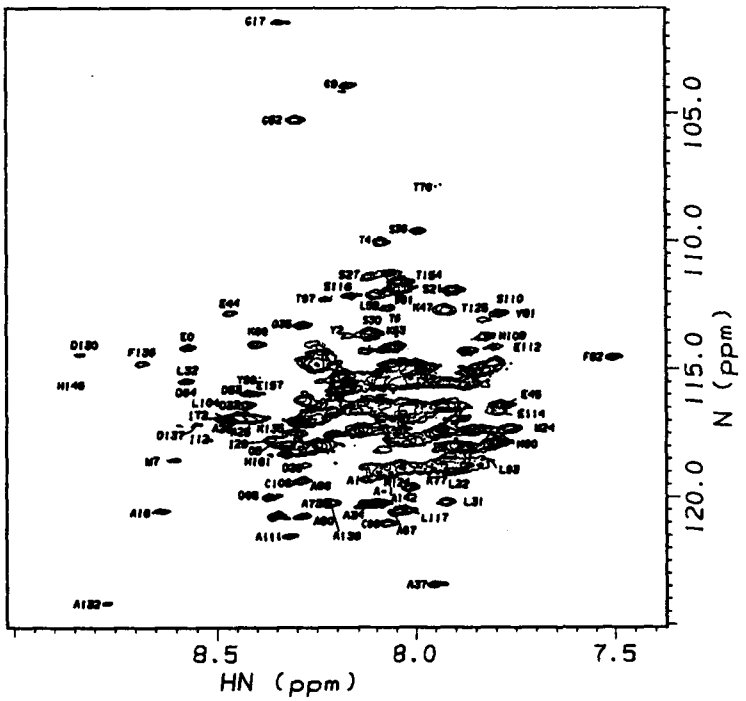
Amide ^{1}H and ^{15}N chemical shift range in the native state at pH 2.0 was 3.6 and 27 ppm (Figure 5.1a), respectively. High dispersion of amide ^{1}H suggests a predominantly β-sheet structure, consistent with the X-ray structure as well as that inferred from the CD spectrum. We could assign most of the backbone atoms in the native state using CBCACONH (Grzesiek & Bax, 1992) and HNCACB (Wittekind & Muller, 1993; spectra are not shown). Because chemical shifts of C_{α} resonances were significantly degenerated, we needed scalar connectivity via CO or C_{β} resonances to connect the neighboring spins uniquely. Especially, C_{β} scalar connectivities were quite essential for the assignment of the TFE state as will be described below. The residues whose backbone nitrogen atoms are not assigned at present are E-2, L22, Y102, D130, K138, I147, R148 and E158.

The TFE-induced $\beta \rightarrow \alpha$ transition of β-LG at pH 2 is highly cooperative, occurring between 15% and 20% (v/v) TFE (Shiraki *et al.*, 1995). The ellipticities at 222 nm before and after the transition are flat, and the maximal helical content is 80%. We chose 50% TFE, high enough to complete the $\beta \rightarrow \alpha$ transition, as a solvent condition stabilizing the TFE state. Amide ^{1}H and ^{15}N chemical shift range in the TFE state was 1.3 and 23 ppm, respectively (Figure 5.1b). The low dispersion of the amide ^{1}H chemical shift in the TFE state suggests a highly helical structure, consistent with the CD spectrum. Because of the less sensitivity of ^{15}N backbone resonance dispersion to the protein structure, we could assign most of them even in the TFE state by using CBCACONH and HNCACB (spectra are not shown). Consequently, although 2D ^{15}N HSQC spectrum shown in Figure 5.1c was remarkably crowded, peaks were

(a)


10.0
9.0
8.0
7.0
HN (ppm)
105.0
110.0
115.0
120.0
125.0
130.0
N (ppm)

(b)


8.5
8.0
7.5
HN (ppm)
105.0
110.0
115.0
120.0
N (ppm)

(c)

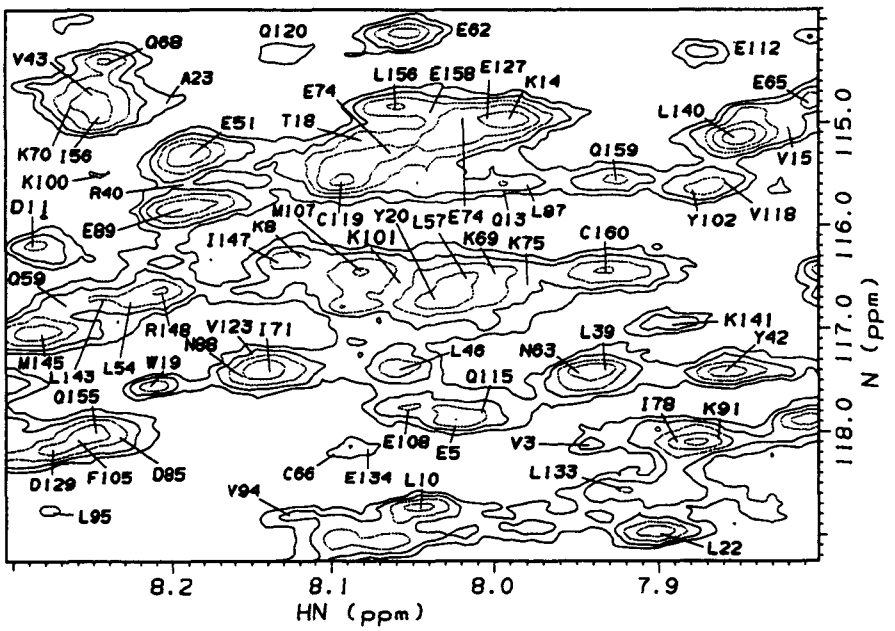


Figure 5.1. ^{1}H, ^{15}N HSQC spectra of β-lactoglobulin at 45°C and pH 2.0. Assignments of the backbone amide signals are indicated using the one-letter code and the residue number. (a) the native state in aqueous solution; (b) the TFE state in 50% TFE (v/v); (c) Expanded region of (b).

clearly separable in 3D CBCACONH and HNCACB. The residues in the TFE state whose backbone nitrogen atoms are not assigned at present are E2, V41, V92, D96, C121, L122, L149 and I162.

5.3.2. Structural features in the native state

H_α, C_α and CO chemical shift differences from random coil chemical shifts (i.e., secondary chemical shifts: ΔH_α, ΔC_α, and ΔCO, respectively), which are excellent indicators of secondary structures (Wishart *et al.*, 1992, 1995), are plotted in Figure 5.2a for the native state. Positive ΔH_α, negative ΔC_α, and negative ΔCO indicate that the residue assumes β-sheet, whereas the opposite signs indicate α-helix. Among the three secondary chemical shifts, ΔH_α is considered to provide the best measure of the secondary structures. It is evident that a substantial part of the molecule assumes β-sheet conformation. In order to compare the NMR results with the X-ray structure, the locations of the α-helix and β-sheet regions taken from the most recent X-ray structure (Brownlow *et al.*, 1997) and their locations estimated by the chemical shift index (CSI) of ΔH_α according to the method of Wishart *et al.* (1992) are shown in the upper and the lower strips, respectively. It is noted that, whereas the present NMR measurements were carried out for a monomeric state at pH 2.0 (Futage & Song, 1980), the X-ray coordinates were determined for a dimeric state with crystals prepared at pH 6.5. The X-ray structure contains nine β-sheets (β_A - β_I), one major α-helix, and four short helices. For the native state at pH 2, nine β-sheets, one major α-helix and three short helices are clearly discernible by CSI, although there are some shifts at the ends of several secondary structure units. ΔH_α of β_D-sheets (residues 65-75) was disrupted at the central part, which may explain the large temperature (B) factor at the corresponding region (see below). Although a short α-helix (residues 113-115) in the X-ray structure is not seen in the CSI profile, ΔC_α and ΔCO in the corresponding region clearly indicates the helical conformation.

Overall patterns of the ΔC_α and ΔCO were similar to that of ΔH_α. However, consistencies with the X-ray secondary structures seem less than that of ΔH_α, emphasizing that ΔH_α provides the best index of the secondary structures. Thus, the secondary structures of the

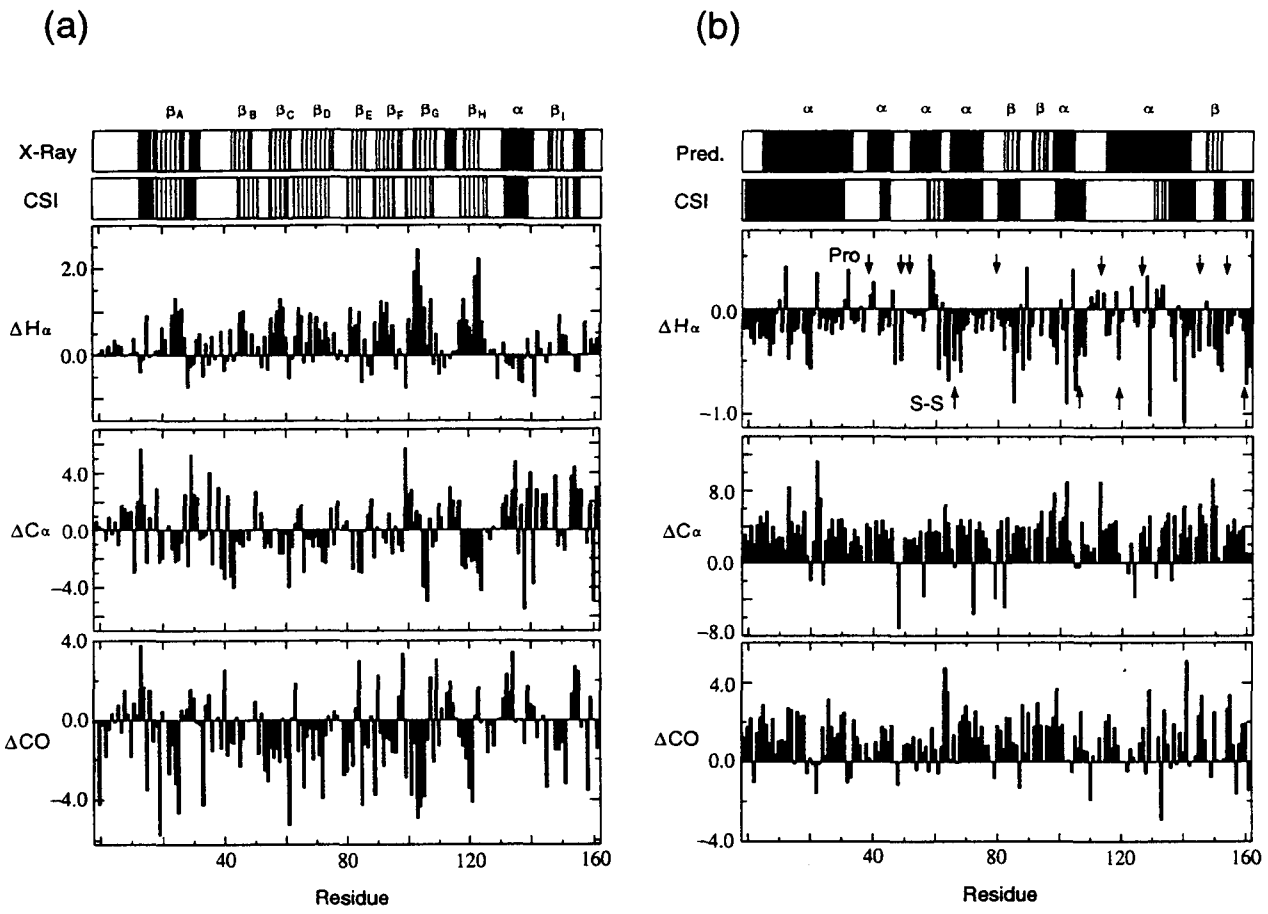


Figure 5.2. Secondary chemical shifts of backbone resonances vs. residue number of β-lactoglobulin. The chemical shift deviations of ^{1}H and ^{13}C resonances of A37, K47, T49, I78, E112, T125, L143 and N152 contain corrections for their sequential position following proline residues. Random coil shifts for other ^{1}H and ^{13}C resonances were taken from Wishart *et al.* (1995). (a) Native state in aqueous solution. Upper and lower strips indicate the secondary structures obtained from X-ray structure and those obtained by CSI (Wishart *et al.*, 1992), respectively. α-helix, solid; β-sheet, vertical stripes. (b) TFE state in 50% (v/v). Upper and lower strips indicate the secondary structures predicted by a PHD method (Rost et al., 1994) and those obtained by CSI, respectively. The symbols are as in panel a. The locations of proline and cysteine with disulfide bond are indicated by arrows.

native state estimated from the secondary chemical shifts are consistent with those of the X-ray structure.

In the native state, strong H^{α}_{i}-H^{N}_{i+1} ($d_{\alpha N}$) and weak or no medium range nuclear Overhauser enhancements (NOEs) (i.e., H^{N}_{i}-H^{N}_{j} ($d_{NN}(i,i+j)$, j=2,3,4; $d_{\alpha N}(i,i+j)$, j=2,3; $d_{\beta N}(i,i+j)$, j=2,3), characteristic for β-sheet, were observed in the corresponding β-sheet regions in the CSI profile, while strong medium range NOEs (j=3,4), characteristic for α-helix, were observed in the helical region in the CSI profile (data not shown). Interstrand NOEs between β-strands in the native state are shown in Figure 5.3. Ragona *et al.* (1997) studied by ^{1}H-NMR the native structure of β-LG at pH 2 and reported the interstrand NOE connectivities. Our results are very similar to theirs except some additional NOEs, which were marked by asterisks, were detected. It is noted that the residues 22, 44, and 95, which form a bulge in the X-ray β-sheets, are not involved in the NOE network. Thus, although the determination of three dimensional structure on the basis of the NMR methods awaits further study, we can fairly conclude that the overall structure of β-LG at pH 2.0 in aqueous solution is very similar to the X-ray structure obtained at pH 6.5.

5.3.3. Structural features in the TFE state

The ΔH_{α}, ΔC_{α}, and ΔCO values for the TFE state are plotted in Figure 5.2b. In a marked contrast with the native state, substantial residues have the negative ΔH_{α} (upper panel), positive ΔC_{α} (middle panel) and positive ΔCO (lower panel) values, demonstrating that the TFE state is a predominantly α-helical conformation. The positions of eight proline residues, which do not form α-helix, are shown in the upper panel of Figure 5.2b.

By the CSI method (Wishart *et al.*, 1992) on the basis of ΔH_{α}, the persistent helical regions were estimated, which are indicated in the lower strip of Figure 5.2b. There are eight helical segments and possibly two β-strands. However, the total helical content is 52%, much less than that (80%) estimated from CD spectrum (Hamada *et al.*, 1996b; Shiraki *et al.*, 1995). Therefore, in this case, the α-helical regions estimated from CSI provide the minimal helical regions and other parts of the molecule also assume the helical structure dynamically. Consistent with this, the ΔC_{α} and ΔCO profiles suggest a higher helical content than that

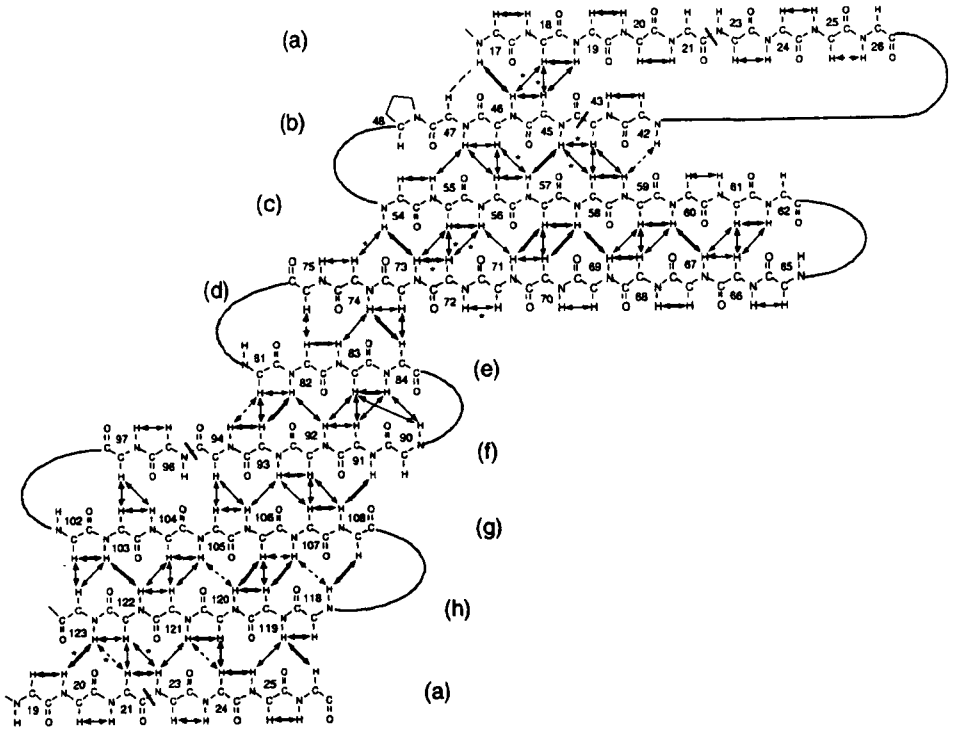


Figure 5.3. Interstrand NOEs obtained for the native state in 50% TFE (v/v) at pH 2.0. Weak, medium and strong NOE are indicated by (- - - -), (——), (▬▬), respectively. Assigned NOEs which were not reported by Ragona *et al.* (1997) are marked by asterisks. The picture was drawn on the basis of Figure 2 of Ragona *et al.* (1997).

estimated by CSI of ΔH_α. Many strong or medium range NOEs (j=3,4) further confirm that almost entire parts of the TFE state assume α-helical conformation (data are not shown).

For the TFE state, a sketch of a structural model consisting of eight persistent helices was illustrated (see Figure 5.5b, pp. 84). It is emphasized that the helical regions represent very stable ones defined by the CSI method and the loop regions can also assume helical structures dynamically. Although the native β_C-strand (residue 54 to 62) seemed to be present in the TFE state, β_B or β_D strands did not exist. Therefore the existence of a small amount of b sheets in the TFE state is uncertain at this stage.

5.3.4. Backbone dynamics

One of the advantages of the heteronuclear NMR approach is that we can analyze the backbone dynamics of psec to msec order. Here, the parameters in Lipari-Szabo spectral density function (Lipari & Szabo, 1982) were calculated, that is, correlation time of the overall tumbling motion (τ_M), order parameters (S^2) and effective correlation times (τ_e), which were obtained using longitudinal relaxation time (T_1), transverse relaxation time (T_2) and NOE of backbone ^{15}N nuclei (Farrow *et al.*, 1995). S^2 indicates the amplitude of the N-H vector motion in a molecular fixed coordinate from 0 to 1 scale. On the other hand, τ_e represents its rate of psec to nsec order. The S^2 and τ_e values of the native and TFE states derived from the NMR experiment are shown in Figure 5.4a and b, respectively. For comparison, B factors obtained from the X-ray structure (Brownlow *et al.*, 1997) are shown in the lower panel of Figure 5.4a. The S^2 values are also indicated in the schematic representation of the native and TFE states by gradually changing the residue color from blue ($S^2 = 1$) to red ($S^2 = 0$) (Figure 5.5).

In the native state, S^2 decreases at the N-terminal, C-terminal, loop and β_D sheet regions, while τ_e increases at the N-terminal, C-terminal, loop and especially β_C to β_D strand regions. The variations in τ_e are similar in general to those of S^2: the regions with low S^2 tend to exhibit high τ_e. These profiles of dynamic properties, in particular S^2, is resembling those of B-factors. However, the profile of S^2 is somewhat spiky compared to that of B-factor. The B-factor is affected by both the amplitude and the phase of the atom displacement as well as the

heterogeneity of the atom position in each cell, while S^2 includes only the contribution of the averaged amplitude of the atomic motions. This discrepancy may partly explain the differences between the S^2 and B-factor.

In the TFE state, as can be seen from difference in overall color of the schematic pictures (Figure 5.5), the S^2values decreased substantially compared with the native state, indicating the global increase in flexibility. There are many residues with the τ_e values higher than 200 psec, consistent with a view that the TFE state is rather expanded and fragile compared to the native state (Figure 5.4a). There is a tendency that $S^2$are high at the persistent helical regions and there seems to be a tendency that S^2around the disulfide bonds are high. The τ_e values increase sharply (> 500 psec) in several loop region, that is, between the persistent helical regions (Figure 5.4b).

5.4. Discussion

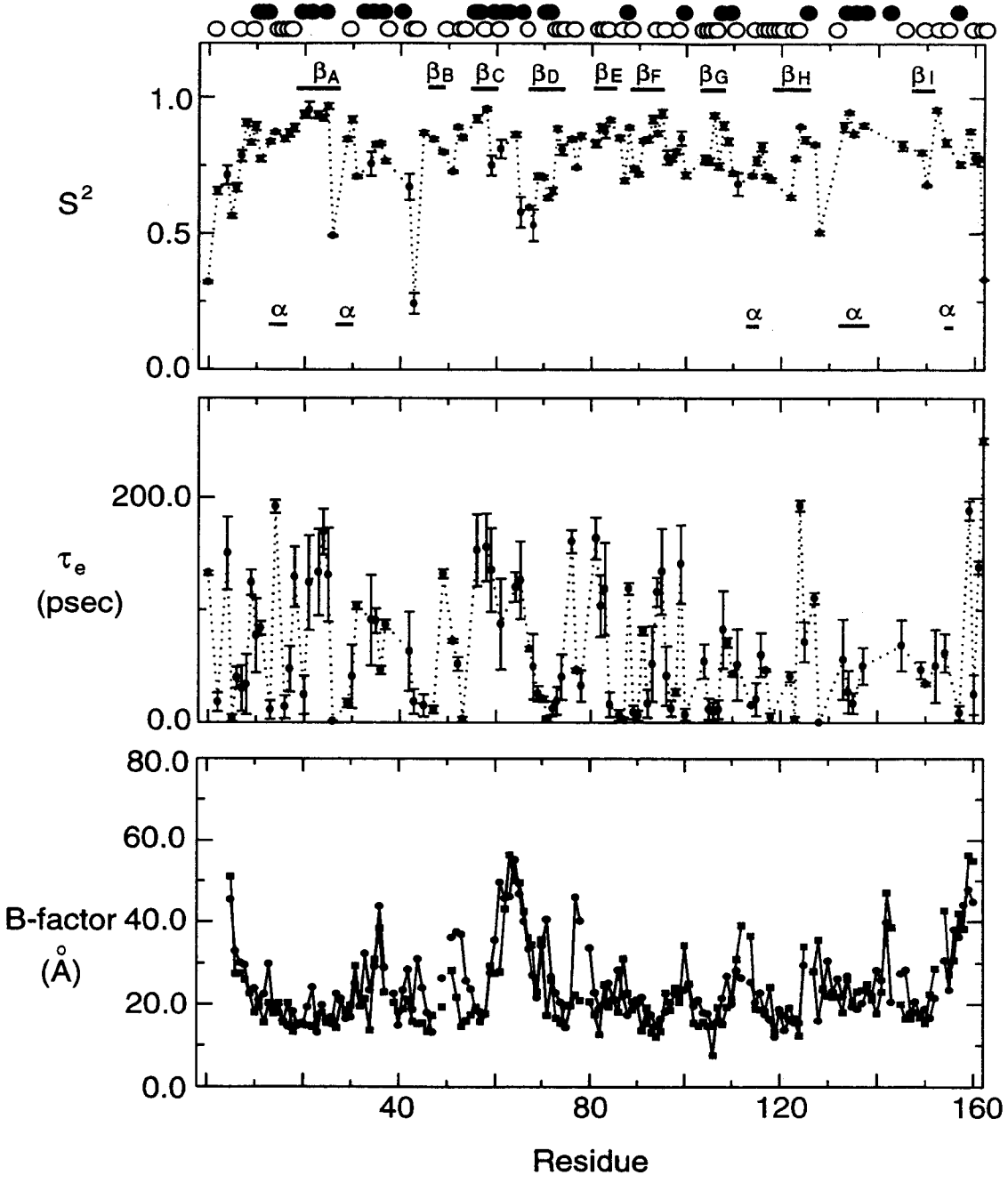
5.4.1. Possible structure of the kinetic intermediate

In the present study, the resonance frequencies of backbone nuclei of the recombinant bovine β-LG in both the native and TFE state as well as the backbone dynamics were analyzed. The results confirmed the high helical propensities of this protein at the residue level. In addtion, the results suggest a possible structure of the kinetic intermediate as will be discussed below.

Because TFE exaggerates the helical propensity of peptides and proteins, the TFE state should not be considered to reproduce the kinetic intermediate trapped under aqueous conditions (Hamada *et al.*, 1996b). In fact, the helical content of the burst-phase intermediate observed during the refolding kinetics in the absence of TFE is only 37%, much less than that of the TFE state. Nevertheless, it would be useful to discuss the possible helical regions in the kinetic intermediate on the basis of the persistent helical regions in the TFE state.

Secondary structure prediction indicates the regions of intrinsically high helical preference. Predicted secondary structures using PHD method (Rost *et al*, 1994) were shown in the upper strip in Figure 5.2b. There are six predicted helical regions which are also candidates for the helical regions in the kinetic intermediate. We provisionally assumed that the

(a)


βA
βB
βC
βD
βE
βF
βG
βH
βI
α
α
α
α
α
1.0
0.5
0.0
S²
200.0
0.0
τe
(psec)
80.0
60.0
40.0
20.0
0.0
B-factor
(Å)
40
80
120
160
Residue

(b)

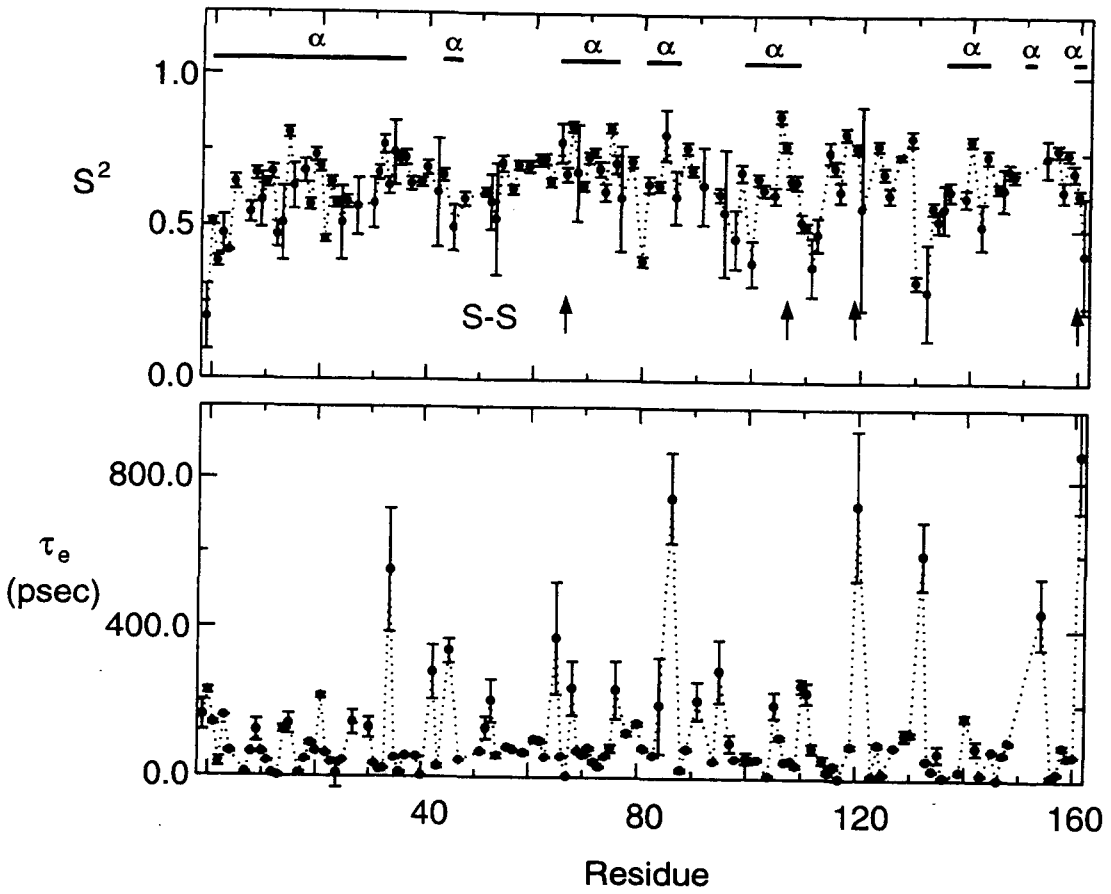


Figure 5.4. (a) (upper) Order parameters (S^2) and (center) effective correlation times (τ_e) of Lipari-Szabo spectral density function obtained from model-free simulations using T_1, T_2 and NOE data of the native state in aqueous solution vs. residue number. (lower) B-factors of backbone nitrogen vs. residue number. (●) and (■) indicate A and B subunit of dimer in the X-ray coordinates, respectively. Secondary structures obtained by CSI profile (Figure 5.2a) are also indicated by bars in the upper panel. Outside of the upper panel, locations of the 50 residues (○) with the native peak and the 28 residues (●) without the native peak in 2D ^{15}N-HSQC spectrum in 15% TFE (v/v) at 25°C are marked. (b) (upper) S^2 and (lower) τ_e of Lipari-Szabo spectral density function in 50% TFE (v/v) vs. residue number. Secondary structures obtained by CSI (Figure 5.2b) are also indicated by bars in the upper panel.

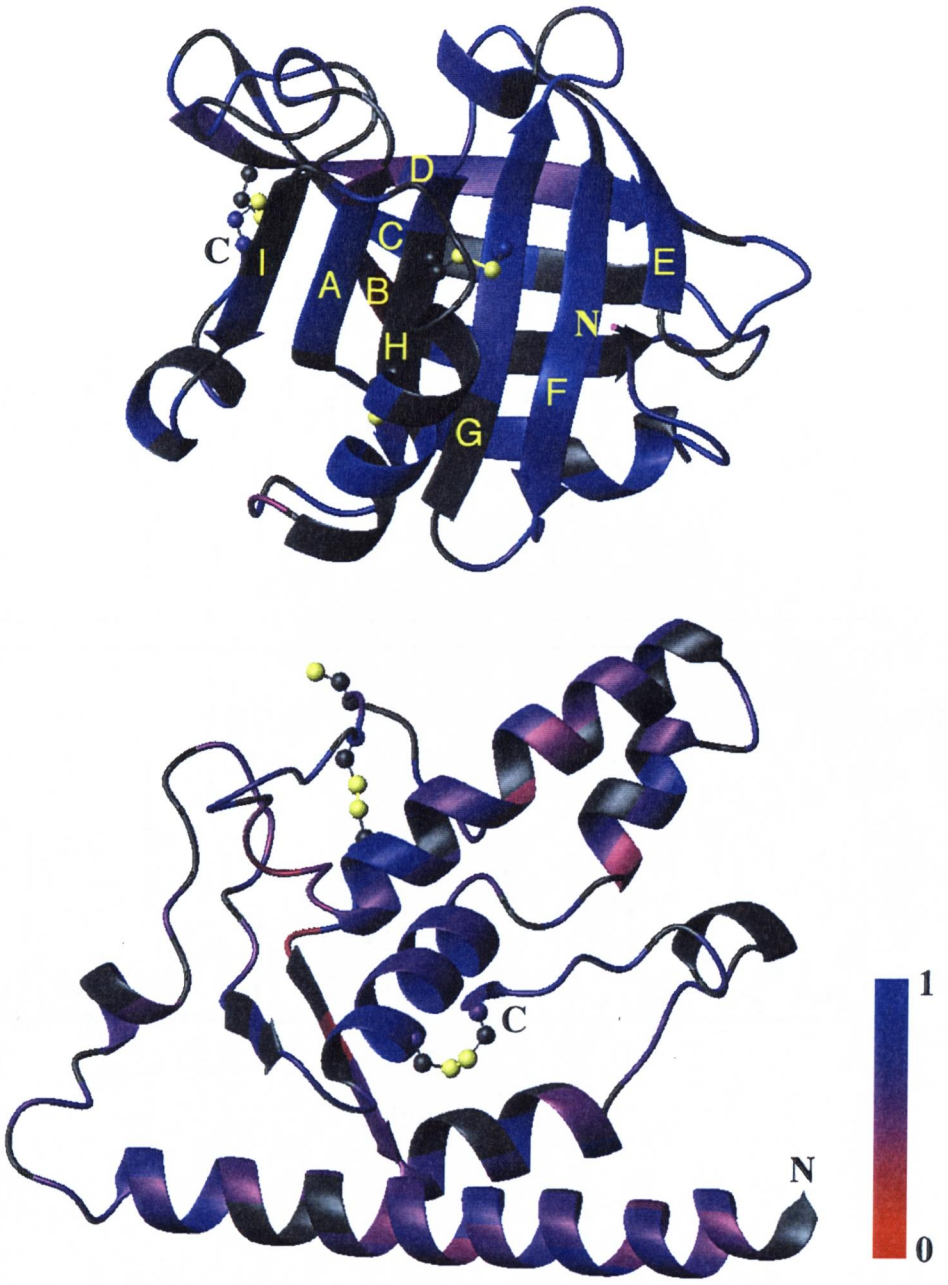


Figure 5.5. (a) X-ray structure of β-lactoglobulin in the native state. (b) Sketch of a structural model for the non-native α-helical conformation in 50% TFE (v/v). The relative arrangement of the helices should be regarded as on of the various possibilities. For both structures, the S^2 values are indicated by gradual change in color from blue (S^2 = 1) to red (S^2 = 0). The residues of which S^2 are not available are colored black. The locations of disulfide bonds are indicated. The diagrams were produced using MOLMOL (Koradi *et al.*, 1996). The native structure was constructed with the PDB coodinate file 1BEB (Brownlow *et al.*, 1997).

residues in both the persistent helices in TFE and the predicted helices are the helical residues in the kinetic intermediate. Such regions are the residues 4 to 30, 42 to 45, 66 to 75, 98 to 102 and 135 to 141, and total helical content is 32.7%, in agreement with that of the kinetic intermediate (Hamada *et al.*, 1996b). These regions are indicated in the native structure (Figure 5.6). A joint prediction method by Nishikawa and Noguchi (1991) also indicated the high helical propensity of these regions. It is noted that, except the C-terminal helix, the highly helical regions are located in the N-terminal half of the molecule, including the N-terminal loop, the first and second short helices, β_A, β_B, and β_D.

In order to clarify the structure of the kinetic intermediate, further studies with various NMR methods will be necessary. At the present stage of our research, it should be added that the NMR spectra at the midpoint of the $\beta \rightarrow \alpha$ transition suggest the validity of a model as shown in Figure 5.6. Although the TFE induced $\beta \rightarrow \alpha$ transition measured by CD is highly cooperative (Shiraki *et al.*, 1995), this was not true at the atomic level measured by the heteronuclear NMR. The 2D ^{15}N-HSQC spectra of β-LG was measured at 15% TFE at various temperatures between 5 and 45°C. At 15°C, the spectrum was close to that of the native state, while, at 45°C, it was close to that of the TFE state (data not shown). Intriguingly at 25°C, the midpoint of the TFE-induced transition, the spectrum was not a simple combination of those of the native state and TFE states, but only 48 peaks arising from the native state and 30 peaks from the TFE state were clearly observed. This suggests the presence of an intermediate for the $\alpha \rightarrow \beta$ transition, in which a part of the molecule remains native and others are helical. The locations of 48 native peaks and 30 peaks arising from the TFE state are indicated on the native secondary structures (the upper panel of Figure 5.4a, pp. 82). As can be seen, many of the native peaks are on the β-strands located on the C-terminal half of the molecule, i.e., β_E, β_F, β_G and β_H, whereas the peaks arising from the TFE state are located on the N-terminal half of the molecule, i.e., β_A, β_B, β_C, β_D, β_I and the C-terminal helix. These results implicate that the N-terminal half of the intermediate is helical and the C-terminal half remains native-like.

Taken together, a following model for the refolding process of β-LG under aqueous conditions was constructed. During the early stage of refolding, the formation of native-like

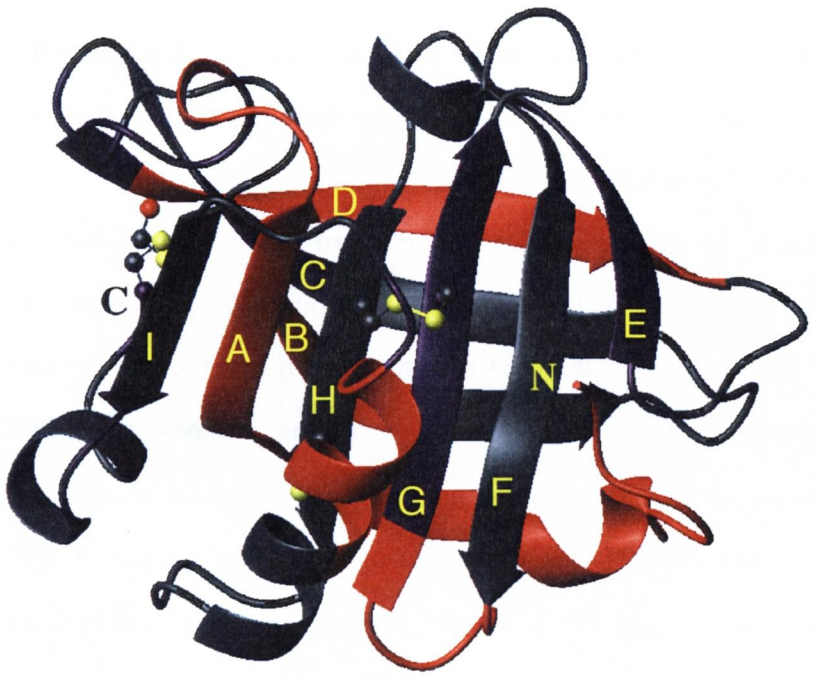


Figure 5.6. Schematic representation of the possible helical regions in the kinetic intermediate of β-lactoglobulin. The highly helical regions indicated from both the persistent helices in the TFE state and the secondary prediction by a PHD method are colored red. The helical regions observed in the TFE state but not in a PHD prediction are colored magenta. The diagram was produced using MOLMOL (Koradi *et al.*, 1996).

Acknowledgments

The author would like to specially express his profound gratitude to Dr. Yuji Goto (Department of Biology, Graduate School of Science, Osaka University) for his continuous guidance, valuable discussion, and assiduous encouragement throughout this work.

The auther also would like to sincerely acknowledge Prof. Seiki Kuramitsu (Department of Biology, Graduate School of Science, Osaka University), Dr. Yasushi Kawata (Department of Biotechnology, Faculty of Engineering, Tottori University), Dr. Kazuo Kuwata (Department of Physiology, School of Medicine, Gifu University), Dr. Carl A. Batt (Department of Food Science, Cornell University), Prof. Saburo Aimoto (Institute for Protein Research, Osaka University), Dr. Noboru Yumoto (Department of Organic Materials, Osaka Natural Research Institute) and Dr. Mikio Kataoka (Department of Earth and Space Science, Graduate School of Science, Osaka University) for their valuable discussions and collaborations.

The author wishes to express his deep appreciation to Dr. Yoshihisa Hagihara (Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology), Dr. Daizo Hamada (Oxford Centre for Molecular Science, University of Oxford), Mr. Kentaro Shiraki, Ms. Nami Hirota and Mr. Tsuyoshi Morishita, and the other members in Laboratory of Biophysical Chemistry, Department of Biology, Graduate School of Science, Osaka University for their continuous encouragement and valuable discussions.

January, 1998

M. Hoshino

Masaru Hoshino

β_E, β_F, β_G and β_H sheets occurs by hydrophobic interactions, while in the N-terminal half of the molecule and in the C-terminal helix region, several helices are formed due to the high helical preference. Subsequently, β_A, β_B, β_C, and β_D sheets are 'induced' through the interaction with the preexisting β sheets. Another indirect evidence for this mechanism is that the regions with high β-sheet preference precisely correspond to β_E, β_F and β_H sheets, while considerable inconsistencies exist in the N-terminal half of the molecule. Because there is almost no β-sheet preference for β_A, β_B, β_C and β_D (Hamada *et al.*, 1996b), this intramolecular induction process might be a rate-limiting step in the refolding reaction.

References

Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science*, **181**, 223-230.

Anna, A. S. & Herskowitz, I. (1994) Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase. *Yeast*, **10**, 801-810.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G. & Struht, K. (1990) Current Protocols in Molecular Biology. In *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, New York.

Barr, K. A., Hopkins, S. A. & Sreekrishna, K. (1992) Protocol for efficient secretion of HSA developed from *Pichia pastoris*. *Pharm. Eng.*, **12**, 48-51.

Batt, C. A., Rabson, L. D., Wong, D. W. & Kinsella, J. E. (1990) Expression of recombinant bovine β-lactoglobulin in *Escherichia coli*. *Agric. Biol. Chem.*, **54**, 949-955.

Bax, A., Clore, G. M. & Gronenborn, A. M. (1990) ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C enriched proteins. *J. Magn. Reson.*, **88**, 425-431.

Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenbach, G. T., Urdea, M. S., Valenzuela, P. & Barr, P. J. (1984) Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **81**, 4642-4646.

Boisvert, D. C., Wang, J., Otwinowski, Z., Horwich, A. L. & Singler, P. B. (1996) The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATPγS. *Nature Struct. Biol.*, **3**, 170-177.

Braig, K., Adams, P. D. & Brünger, A. T. (1995) Conformational variability in the refined structure of the chaperonin GroEL at 2.8Å resolution. *Nature Struct. Biol.*, **2**, 1083-1094.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature*, **371**, 578-586.

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

Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R., Rudolph, R., Schmid, F. X. & Kiefhaber, T. (1991) GroE facilitates refolding of citrate synthase by suppressing aggregation. *Biochemistry*, **30**, 1586-1591.

Bushnell, G. W., Louie, G. V. & Brayer, G. D. (1990) High-resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.*, **214**, 585-595.

Bussey, H. (1988) Proteases and the processing of precursors to secreted proteins in yeast. *Yeast*, **4**, 17-26.

Byler, D. M., Susi, H. & Farrell, H. J. (1983) Laser-Raman spectra, sulfhydryl groups, and conformation of the cystine linkages of β-lactoglobulin *Biopolymers*, **22**, 2507-2511.

Chan, H. S. & Dill, K. A. (1990) The effects of internal constraints of the configurations of chain molecules. *J. Chem. Phys.*, **92**, 3118-3135.

Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R. & Saibil, H. R. (1994) Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature*, **371**, 261-264.

Collawn, J. F. & Paterson, Y. (1990) Stabilization of helical structure in two 17-residue amphipathic analogues of the C-terminal peptide of cytochrome *c*. *Biopolymers*, **29**, 1289-1296.

Couderc, R. & Baratti, J. (1980) Oxidation of methanol by the yeast *Pichia pastoris*. Purification and properties of alcohol oxidase. *Agric. Biol. Chem.*, **44**, 2279-2289.

Cregg, J. M., Barringer, K. J., Hessler, A. Y. & Madden, K. R. (1985) *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.*, **5**, 3376-3385.

Davies, A. M., Guillemette, J. G., Smith, M., Greenwood, C., Thurgood, A. G. P., Mauk, A. G. & Moore, G. R. (1993) Redesign of the interior hydrophilic region of mitochondrial cytochrome *c* by site-directed mutagenesis. *Biochemistry*, **32**, 5431-5435.

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

Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) Principle of protein folding - A perspective from simple exact models. *Protein Sci.*, **4**, 561-602.

Ellis, R. J. (1987) Proteins as molecular chaperones. *Nature*, **328**, 378-379.

Ellis, R. J. & van der Vies, S. M. (1991) Molecular chaperones. *Annu. Rev. Biochem.*, **60**, 321-347.

Ewbank, J. J., Creighton, T. E., Hayer-Hartl, M. K. & Hartl, F. U. (1995) What is the molten globule? *Nature Struct. Biol.*, **2**, 10.

Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. (1995) Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry*, **34**, 868-878.

Fenton, W. A. & Horwich, A. L. (1997) GroEL-mediated protein folding. *Protein Sci.*, **6**, 743-760.

Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994) Residues in chaperonin GroEL required for polypeptide binding and release. *Nature*, **371**, 614-619.

Fisher, W. R., Taniuchi, H. & Anfinsen, C. B. (1973) On the role of heme in the formation of the structure of cytochrome *c*. *J. Biol. Chem.*, **248**, 3188-3195.

Fisher, M. T. & Yuan, X. (1994) The rates of commitment to renaturation of rhodanese and glutamine synthetase in the presence of the GroE chaperonins. *J. Biol. Chem.*, **269**, 29598-29601.

Fogh, R. H., Schipper, D., Boelens, R. & Kaptein, R. (1995) ^{1}H, ^{13}C & ^{15}N assignments and secondary structure of 269-residue serine protease PB92 from *Bacillus alcalophilus*. *J. Biomol. NMR*, **5**, 259-270.

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

Funatsu, T., Harada, Y., Tokunaga, M., Saito, K. & Yanagida, T. (1995) Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, **374**, 555-559.

Gething, M. J. & Sambrook, J. (1992) Protein folding in the cell. *Nature*, **355**, 33-45.

Gray, T. E. & Fersht, A. R. (1993) Refolding of barnase in the presence of GroE. *J. Mol. Biol.*, **232**, 1197-1207.

Grzesiek, S. & Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.*, **114**, 6291-6293.

Grzesiek, S. & Bax, A. (1993) Improved 3D triple resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.*, **96**, 432-430.

Guisez, Y., Tison, B., Vandekerckhove, J., Demolder, J., Bauw, G., Haegeman, G., Fiers, W. & Contreras, R. (1991) Production and purification of recombinant human interleukin-6 secreted by the yeast *Saccharomyces cerevisiae*. *Eur. J. Biochem.*, **198**, 217-222.

Hamada, D. & Goto, Y. (1997) The equilibrium intermediate of β-lactoglobulin with non-native α-helical structure. *J. Mol. Biol.*, **269**, 479-487.

Hamada, D., Hoshino, M., Kataoka, M., Fink, A. L. & Goto, Y. (1993) Intermediate conformational states of apocytochrome *c*. *Biochemistry*, **32**, 10351-10358.

Hamada, D., Kuroda, Y., Kataoka, M., Aimoto, S., Yoshimura, T. & Goto, Y. (1996a) Role of heme axial ligands in the conformational stability of the native and molten globule states of horse cytochrome *c*. *J. Mol. Biol.*, **256**, 172-186.

Hamada, D., Kuroda, Y., Tanaka, T. & Goto, Y. (1995) High helical propensity of the peptide fragments derived from β-lactoglobulin, a predominantly β-sheet protein. *J. Mol. Biol.*, **254**, 737-746.

Hamada, D., Segawa, S. & Goto, Y. (1996b) Non-native α-helical intermediate in the refolding of β-lactoglobulin, a predominantly β-sheet protein. *Nature Struct. Biol.*, **3**, 868-873.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.*, **166**, 557-580.

Hantgan, R. R. & Taniuchi, H. (1977) Formation of a biologically active, ordered complex from two overlapping fragments of cytochrome *c*. *J. Biol. Chem.*, **252**, 1367-1374.

Hartl, F. U. (1996) Molecular chaperones in cellular protein folding. *Nature*, **381**, 571-580.

Hartl, F. U., Hlodan, R. & Langer, T. (1994) Molecular chaperonins in protein folding: the art of avoiding sticky situations. *Trends Biochem. Sci.*, **19**, 20-25.

Hayer-Hartl, M. K., Ewbank, J. J., Creighton, T. E. & Hartl, F. U. (1994) Conformational specificity of the chaperonin GroEL for the compact folding intermediates of α-lactalbumin. *EMBO J.*, **13**, 3192-3202.

Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature*, **333**, 330-334.

Hendrick, J. P. & Hartl, F.-U. (1993) Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem.*, **62**, 349-384.

Hendrix, R. W. (1979) Purification and properties of groE, a host protein involved in bacteriophage assembly. *J. Mol. Biol.*, **129**, 375-392.

Hinchliffe, E. & Kenny, E. (1993) Yearsts as a vehicle for the expression of heterologous genes. In *The Yeasts* (Rose, A. H. & Harrison, J. S., Eds.), pp. 325-352, Academic Pres Inc., San Diego.

Hlodan, R., Tempst, P. & Hartl, F. U. (1995) Binding of defined regions of a polypeptide to GroEL and its implications for chaperonin-mediated protein folding. *Nature Struct. Biol.*, **2**, 587-595.

Hoshino, M., Hagihara, Y., Hamada, D., Kataoka, M. & Goto, Y. (1997) Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering. *FEBS Lett.*, **416**, 72-76.

Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L. & Deisenhofer, J. (1996) The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature*, **379**, 37-45.

Hutchinson, E. G., Tichelaar, W., Hofhaus, G., Weiss, H. & Leonard, K. R. (1989) Identification and electron microscopic analysis of a chaperonin oligomer from *Neurospora crassa* mitochondria. *EMBO J.*, **8**, 1485-1490.

Itzhaki, L. S., Daniel, E. O. & Fersht, A. R. (1995) Nature and consequences of GroEL-protein interactions. *Biochemistry*, **34**, 14581-14587.

Jackson, G. S., Staniforth, R. A., Halsall, D. J., Atkinson, T., Holbrook, J. J., Clarke, A. R. & Burston, S. G. (1993) Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: Implications for the mechanism of assisted protein folding. *Biochemistry*, **32**, 2554-2563.

Jain, M. K. & Zakim, D. (1987) The spontaneous incorporation of proteins into preformed bilayers. *Biochim. Biophys. Acta*, **906**, 33-68.

de Jongh, H. H. J., Killian, J. A. & de Kruijff, B. (1992) A water-lipid interface induces a highly dynamic folded state in apocytochrome *c* and cytochrome *c*, which may represent a common folding intermediate. *Biochemistry*, **31**, 1636-1643.

de Jongh, H. H. J. & de Kruijff, B. (1990) The conformational changes of apocytochrome *c* upon binding to phospholipid vesicles and micelles of phospholipid based detergents: a circular dichroism study. *Biochim. Biophys. Acta*, **1029**, 105-112.

Jordi, W., Li-Xin, Z., Pilon, M., Demel, R. A. & de Kruijff, B. (1989) The importance of the amino terminus of the mitochondrial precursor protein apocytochrome *c* for translocation across model membranes. *J. Biol. Chem.*, **264**, 2292-2301.

Juillerat, M., Parr, G. R. & Taniuchi, H. (1980) A biologically active, three-fragment complex of horse heart cytochrome *c*. *J. Biol. Chem.*, **255**, 845-853.

Julius, D., Brake, A., Blair, L., Kunisawa, R. & Thorner, J. (1984) Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-α-factor. *Cell*, **37**, 1075-1089.

Julius, D., Blair, L., Brake, A., Sprague, G. & Thorner, J. (1983) Yeast α-factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell*, **32**, 839-852.

Katsumata, K., Okazaki, A. & Kuwajima, K. (1996) Effect of GroEL on the re-folding kinetics of α-lactalbumin. *J. Mol. Biol.*, **258**, 827-838.

Kawata, Y., Nosaka, K., Hongo, K., Mizobata, T. & Nagai, J. (1994) Chaperonin GroE and ADP facilitate the folding of various proteins and protect against heat inactivation. *FEBS Lett.*, **345**, 229-232.

Kiefhabor, T. (1995) Kinetic traps in lysozyme folding. *Proc. Natl. Acad. Sci. USA*, **92**, 9029-9033.

Koradi, R., Billeter, M. & Wüthrich, K. (1996) MOLMOL – a program for display and analysis of macromolecular structures. *J. Molec. Graph.*, **14**, 51-55.

Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946-950.

Kubo, T., Mizobata, T. & Kawata, Y. (1993) Refolding of yeast enolase in the presence of the chaperonin GroE. *J. Biol. Chem.*, **268**, 19346-19351.

Kubota, H., Hynes, G. & Willison, K. (1995) The chaperonin containing *t*-complex polypeptide 1 (TCP-1). *Eur. J. Biochem.*, **230**, 3-16.

Kurjan, J. & Herskowitz, I. (1982) Structure of a yeast pheromone gene (MF α): a putative α-factor precursor contains four tandem copies of mature α-factor. *Cell*, **30**, 933-943.

Kuroda, Y., Hamada, D., Tanaka, T. & Goto, Y. (1996) High helicity of peptide fragments corresponding to β-strand regions of β-lactoglobulin observed by 2D-NMR spectroscopy. *Folding & Design*, **1**, 255-263.

Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins Struct. Funct. Genet.*, **6**, 87-103.

Kuwajima, K., Mitani, M. & Sugai, S. (1989) Characterization of the critical state in protein folding. Effects of guanidine hydrochloride and specific Ca^{2+} binding on the folding kinetics of α-lactalbumin. *J. Mol. Biol.*, **206**, 547-561.

Kuwajima, K., Yamaya, H. & Sugai, S. (1996) The burst-phase intermediate in the refolding of β-lactoglobulin studied by stopped-flow circular dichroism and absorption spectroscopy. *J. Mol. Biol.*, **264**, 806-822.

Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. & Nagamura, T. (1987) Rapid formation of secondary structure framework in protein folding studied by stopped-flow circular dichroism. *FEBS Lett.*, **221**, 115-118.

Landry, S. J. & Gierasch, L. M. (1991) The chaperonin GroEL binds a polypeptide in an α-helical conformation. *Biochemistry*, **30**, 7359-7362.

Landry, S. J. & Gierasch, L. M. (1994) Polypeptide interactions with molecular chaperones and their relationship to *in vivo* protein folding. *Ann. Rev. Biophys. Biomol. Struct.*, **23**, 645-669.

Landry, S. J., Jordan, R., McMacken, R. & Gierasch, L. M. (1992) Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. *Nature*, **355**, 455-457.

Lilie, H. & Buchner, J. (1995) Interaction of GroEL with a highly structured folding intermediate: Iterative binding cycles do not involving unfolding. *Proc. Natl. Acad. Sci. USA*, **92**, 8100-8104.

Lin, Z., Schwarz, F. P. & Eisenstein, E. (1995) The hydrophobic nature of GroEL-substrate binding. *J. Biol. Chem.*, **270**, 1011-1014.

Lipari, G. & Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, **104**, 4546-4559.

Mande, S. C., Mehra, V., Bloom, B. R. & Hol, W. G. J. (1996) Structure of the heat shock protein chaperonin-10 of *mycobacterium leprae*. *Science*, **271**, 203-207.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L. & Hartl, F. U. (1991) Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature*, **352**, 36-42.

Mayhew, M. & Hartl, F. U. (1996) Lord of the rings: GroES structure. *Science*, **271**, 161-162.

Mendoza, J. A., Butler, M. C. & Horowitz, P. M. (1992) Characterization of a stable, reactivatable complex between chaperonin 60 and mitochondrial rhodanese. *J. Biol. Chem.*, **267**, 24648-24654.

Merritt, E. A. & Murphy, M. E. P. (1994) *Raster3D* version 2.0. A program for photorealistic molecular graphics. *Acta. Cryst.*, **D50**, 869-873.

Mizobata, T., Akiyama, Y., Ito, K., Yumoto, N. & Kawata, Y. (1992) Effects of the chaperonin GroE on the refolding of tryptophanase from *Escherichia coli*. *J. Biol. Chem.*, **267**, 17773-17779.

Muga, A., Mantsch, H. H. & Surewicz, W. K. (1991) Apocytochrome *c* interaction with phospholipid membranes studied by Fourier-transform infrared spectroscopy. *Biochemistry*, **30**, 2629-2635.

Murai, N., Taguchi, H. & Yoshida, M. (1995) Kinetic analysis of interactions between GroEL and reduced α-lactalbumin. *J. Biol. Chem.*, **270**, 19957-19963.

Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.*, **4**, 2138-2148.

Nieba-Axmann, S. E., Ottiger, M., Wütrich, K. & Plückthun, A. (1997) Multiple cycles of global unfolding of GroEL-bound cyclophilin A evidenced by NMR. *J. Mol. Biol.*, **271**, 803-818.

Nishii, I., Kataoka, M., Tokunaga, F. & Goto, Y. (1994) Cold denaturation of the molten globule states of apomyoglobin and profile for protein folding. *Biochemistry*, **33**, 4903-4909.

Nishikawa, K. & Noguchi, T. (1991) Predicting protein secondary structure based on amino acid sequence. *Methods Enzymol.*, **202**, 21-24.

Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. J. (1997) Direct observation of the rotation of F1-ATPase. *Nature*, **386**, 299-302.

Okazaki, A., Ikura, T. & Kuwajima, K. (1995) In reply – What is the molten globule? *Nature Struct. Biol.*, **2**, 10-11.

Okazaki, A., Ikura, T., Nikaido, K. & Kuwajima, K. (1994) The chaperonin GroEL does not recognize apo-α-lactalbumin in the molten globule state. *Nature Struct. Biol.*, **1**, 439-446.

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

Parr, G. R., Hantgan, R. R. & Taniuchi, H. (1978) Formation of two alternative complementing structures from a cytochrome *c* heme fragment (residues 1 to 38) and the apoprotein. *J. Biol. Chem.*, **253**, 5381-5388.

Pessen, H., Purcell, J. M. & Farrell, H. M. J. (1985) Proton relaxation rates of water in dilute solution of β-lactoglobulin. Determination of cross relaxation and correlation with structural changes by the use of two generic variants of a self-associating globular protein. *Biochim. Biophys. Acta*, **828**, 1-12.

Peterson, J., Saleem, M. M. M., Silver, J. & Wilson, M. T. (1983) On the preparation and mössbauer properties of some heme peptides of cytochrome *c*. *J. Inorg. Biochem.*, **19**, 165-178.

Ptitsyn, O. B. (1992). The molten globule state. In *Protein Folding* (Creighton, T. E., ed.), pp. 243-300. W. H. Freeman and Co., New York.

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

Ranson, N. A., Dunster, N. J., Burston, S. G. & Clarke, A. R. (1995) Chaperonins can catalyze the reversal of early aggregation steps when a protein misfolds. *J. Mol. Biol.*, **250**, 581-586.

Rietveld, A., Ponjee, G. A., Schiffer, P., Jordi, W., van de Cookwijk, P. J. F. M., Demel, R. A., Marsh, D. & de Kruijiff, B. (1985) Investigations on the insertion of the mitochondrial precursor protein apocytochrome *c* into model membranes. *Biochim. Biophys. Acta*, **818**, 398-409.

Rocha, T. L., Paterson, G., Crimmins, K., Boyd, A., Sawyer, L. & Fothergill, G. L. (1996) Expression and secretion of recombinant ovine β-lactoglobulin in *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. *Biochem. J.*, **313**, 927-932.

Romanos, M. (1995) Advances in the use of *Pichia pastoris* for high-level gene expression. *Curr. Opin. Biotech.*, **6**, 527-533.

Roseman, A. M., Chen, S., White, H., Braig, K. & Saibil, H. R. (1996) The chaperonin ATPase cycle: Mechanism of allosteric switching and movements of substrate-binding domains in GroEL. *Cell*, **87**, 241-251.

Rost, B., Sander, C. & Schneider, R. (1994) PHD-an automatic mail server for protein secondary structure prediction. *CABIOS*, **10**, 53-60.

Rothman, J. E. (1989) Polypeptide chain binding proteins: Catalysts of protein folding and related processes in cells. *Cell*, **59**, 591-601.

Schmidt, M. & Buchner, J. (1992) Interaction of GroE with an all-β-protein. *J. Biol. Chem.*, **267**, 16829-16833.

Shiraki, K., Nishikawa, K. & Goto, Y. (1995) Trifluoroethanol-induced stabilization of the α-helical structure of β-lactoglobulin: Implication for non-hierarchical protein folding. *J. Mol. Biol.*, **245**, 180-194.

Smallcombe, S. H. (1993) Solvent supression with symmetrically shifted pulses. *J. Am. Chem. Soc.*, **115**, 4776-4785.

Spooner, P. J. R. & Watts, A. (1991a) Reversible unfolding of cytochrome *c* upon interaction with caldiolipin bilayers. 1. Evidence from deuterium NMR measurements. *Biochemistry*, **30**, 3871-3879.

Spooner, P. J. R. & Watts, A. (1991b) Reversible unfolding of cytochrome *c* upon interaction with caldiolipin bilayers. 2. Evidence from phosphorus-31 NMR measurements. *Biochemistry*, **30**, 3880-3885.

States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982) A two dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.*, **48**, 386-393.

Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science*, **265**, 659-665.

Totsuka, M., Katakura, Y., Shimizu, M., Kumagai, I., Miura, K. & Kaminogawa, S. (1990) Expression and secretion of bovine β-lactoglobulin in *Saccharomyces cerevisiae*. *Agric. Biol. Chem.*, **54**, 3111-3116.

Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H. R., Fenton, W. A. & Horwich, A. L. (1995) Mechanism of GroEL action: Productive release of polypeptide from a sequestered position under GroES. *Cell*, **83**, 577-587.

Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M. & Horwich, A. L. (1996) Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell*, **84**, 481-490.

Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shitfs of the common amino acids. I. Investigation of nearest-neighbor effects. *J. Biomol. NMR*, **5**, 67-81.

Wishart, D. S., Sykes, B. D. & Richards, F. M. (1992) The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, **31**, 1647-1651.

Wittekind, M. & Muller, L. (1993) HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha and beta carbon resonances in proteins. *J. Magn. Reson. B*, **101**, 201-205.

Xu, Z., L., H. A. & Sigler, P. B. (1997) The crystal structure of the asymmetric GroEL-GroES-$(ADP)_7$ chaperonin complex. *Nature*, **388**, 741-750.

Yamamoto, K., Sekine, T. & Kanaoka, Y. (1977) Fluorescent thiol reagents. *Anal. Biochem.*, **79**, 83-94.

Yamazaki, T., Kay, J. D. & Kay, L. E. (1993) Two-dimensional NMR experiments for correlating $^{13}C^{\beta}$ and $^{1}H^{\delta/\varepsilon}$ chemical shifts of aromatic residues in ^{13}C labeled proteins via scalar couplings. *J. Am. Chem. Soc.*, **115**, 11054-11055.

Zahn, R., Perrett, S. & Fersht, A. R. (1996) Conformational states bound by the molecular chaperones GroEL and SecB: A hidden unfolding (annealing) activity. *J. Mol. Biol.*, **261**, 43-61.

Zahn, R., Spitzfaden, C., Ottiger, M., Wüthrich, K. & Plückthun, A. (1994) Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature*, **368**, 261-265.

List of Publications

A part of this thesis has been or will be published in the following articles.

1) Hoshino, M., Kawata, Y. & Goto, Y. (1996) Interaction of GroEL with conformational states of horse cytochrome *c*. *J. Mol. Biol.*, **262**, 575-587.

2) Kuwata, K., Hoshino, M., Era, S., Batt, C. A. & Goto, Y. Heteronuclear NMR characterization of the native β-sheet and TFE-induced α-helical states of β-lactoglobulin. *in preparation.*

Other articles.

1) Hamada, D., Hoshino, M., Kataoka, M., Fink, A. L. & Goto, Y. (1993) Intermediate conformational states of apocytochrome *c*. *Biochemistry*, **32**, 10351-10358.

2) Goto, Y., Hagihara, Y., Hamada, D., Hoshino, M. & Nishii, I. (1993) Acid-induced unfolding and refolding transitions of cytochrome *c*: A three-state mechanism in H_2O and D_2O. *Biochemistry*, **32**, 11878-11885.

3) Hoshino, M. & Goto, Y. (1994) Perchlorate-induced formation of the α-helical structure of mastoparan. *J. Biochem.*, **116**, 910-915.

4) Hoshino, M., Yumoto, N., Yoshikawa, S. & Goto, Y. (1997) Design and characterization of the anion-sensitive coiled coil peptide. *Protein Sci.*, **6**, 1396-1404.

5) Hoshino, M., Hagihara, Y., Hamada, D., Kataoka, M. & Goto, Y. (1997) Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering. *FEBS Lett.*, **416**, 72-76.

6) Hagihara, Y., Hoshino, M., Hamada, D., Kataoka, M. & Goto, Y. Chain-like conformation of the heat-denatured ribonuclease A and cytochrome *c* evidenced by solution X-ray scattering. submitted to *Folding & Design.*

curriculum vitae

January 30, 1998

Name: Masaru Hoshino

Position: Graduate student, Graduate School of Science, Biochemistry, Osaka University.
Research Fellow of the Japan Society for the Promotion of Science.

Date of birth: January 9, 1970

Nationality: Japanese

Home Address: 1-12-7, Hotarugaike-Higashimachi, Toyonaka, Osaka, 560, Japan
Telephone: +81-6-841-5315

Office Address: 1-1, Machikaneyama-cho, Toyonaka, Osaka, 560, Japan
Telephone: +81-6-850-5435
Facsimile: +81-6-850-5442

Education:

Faculty of Science, Osaka University	April, 1989- March, 1993	Obtained a Bachelor of Science
Graduate School of Science, Osaka University	April, 1993- March, 1995	Obtained a Master of Science

Proceeded to Doctor Course, in Graduate School of Science, Osaka University in April, 1995.

Professional Specialty:
Biophysical Chemistry of Proteins

Professional Society Membership:
The Japanese Biochemical Society
The Biophysical Society of Japan