



Review

Structural dynamics and folding of β -lactoglobulin probed by heteronuclear NMR

Kazumasa Sakurai, Tsuyoshi Konuma, Masanori Yagi, Yuji Goto*

Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871, Japan

ARTICLE INFO

Article history:

Received 18 January 2009

Received in revised form 2 April 2009

Accepted 6 April 2009

Available online 10 April 2009

Keywords:

 β -Lactoglobulin

Tanford transition

Ligand binding

Conformational dynamics

Protein folding

Heteronuclear NMR

ABSTRACT

Bovine β -lactoglobulin (β LG) has been one of the most extensively studied proteins in the history of protein science mainly because its abundance in cow's milk makes it readily available to researchers. However, compared to other textbook proteins, progress in the study of β LG has been slow because of obstacles such as a low reversibility from denaturation linked with thiol–disulfide exchange or monomer–dimer equilibrium preventing a detailed NMR analysis. Recently, the expression of various types of recombinant β LGs combined with heteronuclear NMR analysis has significantly improved understanding of the physico-chemical properties of β LG. In this review, we address several topics including pH-dependent structural dynamics, ligand binding, and the complex folding mechanism with non-native intermediates. These unique properties might be brought about by conformational frustration of the β LG structure, partly attributed to the relatively large molecular size of β LG. We expect studies with β LG to continue to reveal various important findings, difficult to obtain with small globular proteins, leading to a more comprehensive understanding of the conformation, dynamics and folding of proteins.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Bovine β -lactoglobulin (β LG) has been one of the most extensively studied proteins in the history of protein science [1–4]. The major reason for this is simply its abundance in cow's milk: the concentration of β LG in milk is about 0.2 g/100 ml, next to casein (2.9 g/100 ml) [5], making β LG easily accessible to researchers. For the same reason, sperm whale myoglobin [6,7], hen egg lysozyme [8,9], and lactate dehydrogenase [10,11] are examples of proteins on which extensive studies were performed in the 1950s, leading to the determination of their 3D structure and an understanding of structure–function relationships. Extensive studies with β LG started in the 1950s and have provided important information regarding structure and function. These included the pH-dependent conformational transition known as the Tanford transition [12] and the binding of various hydrophobic ligands [13], making β LG an excellent model for the molecular study of proteins as reviewed previously [2–4,14–17].

However, compared to other textbook proteins, detailed structural studies on β LG were slower because of several obstacles. Medium resolution X-ray crystallographic structures of β LG were reported only in the late 1980s [18], followed by a series of high resolution structures, some including ligands [19–21]. Low reversibility from denaturation linked with thiol–disulfide exchange prevented the detailed study of protein folding with β LG [22], though such exchange is relevant to gelation [23,24] and advantageous for the food

applications. The X-ray structures suggested the mechanism of the Tanford transition [20] and ligand binding [19,21]. However, details in solution have remained unknown because heteronuclear NMR analysis, a powerful technique often used for other proteins, has been difficult to apply to β LG at neutral pH.

Recently, obstacles preventing the detailed molecular understanding of the structure and dynamics of β LG in solution have been cleared by expressing wild-type [25–31] and suitable mutants [32,33] or by choosing appropriate experimental conditions as well as employing advanced analytical techniques, in particular heteronuclear NMR. Consequently, we now understand the molecular basis of the unique properties of β LG indicated in earlier studies. Moreover, supported by heteronuclear NMR, we believe that β LG will continue to be one of the most important targets of protein science for several reasons. First of all, it is a typical globular protein of moderately large molecular size. So far, detailed studies of folding in solution have been performed with relatively small proteins, typically less than 100 amino acid residues, creating the concept of rapid folding without folding intermediate or complex folding funnel [34–36]. However, studies with larger proteins, e.g. apomyoglobin prepared by the removal of heme from myoglobin [37,38], suggest important roles for folding intermediates in productive folding [39]. It is likely that β LG will provide a wealth of information about the roles of such key intermediates. On the other hand, the binding of ligands to β LG continues to be an intriguing topic. From a more practical viewpoint, aggregation of β LG is closely related to food processing properties of milk. Taken together, it is increasingly important to obtain further understanding of the structure, dynamics, and ligand binding of β LG. In this review, we overview recent advances focusing on studies with

* Corresponding author.

E-mail address: ygoto@protein.osaka-u.ac.jp (Y. Goto).

NMR. For the physico-chemical properties of β LG, readers should also refer to other reviews [4,14,15].

2. Structure of β LG

β LG consists of 162 amino acid residues (18 kDa), containing two disulfide bonds (Cys 66–Cys 160 and Cys 106–Cys 119) and a free thiol (Cys 121). Structures of β LG have been reported by several groups with X-ray crystallography [19–21] and solution NMR [29,40,41] (Fig. 1A). It is a predominantly β -sheet protein. The β -barrel, or so-called calyx, is conical and is made of two β -sheets: the B–D strands and N-terminal half of the A strand (denoted A^N) form one sheet, and the E–H strands and C-terminal half of the A strand (denoted A^C) form the other (Fig. 1B). On the outer surface of the β -barrel, between the G and H strands, is the 3-turn α -helix. The loops that connect the β -strands at the closed end of the calyx, BC, DE, and FG, are generally quite short, whereas those at the open end, AB, CD, EF, and GH, are significantly longer and more flexible [19]. In the calyx, there is a large central cavity which is surrounded by hydrophobic residues and is accessible to solvent. This cavity provides the principal ligand-binding site. β LG contains two tryptophan residues, Trp 19 on the A strand and Trp 61 on the C strand. The former is buried in the hydrophobic core whereas the latter is exposed to the solvent in the native structure, making them useful probes for monitoring site-specific conformational changes. In addition, studies on the monomer–dimer equilibrium [30,32,42,43] and the reactivity of the thiol group of Cys121 deeply buried between the α -helix and H strand [44–48] revealed other important properties of β LG.

β LG belongs to the lipocalin family, a subclass of the calycin superfamily [49]. There are two other subfamilies in the calycin superfamily; the fatty acid-binding protein (FABP) family and the avidin family. The functions of most lipocalins, such as tear lipocalin [50], and FABPs, such as I-FABP [51], are in the transportation of ligands. However, some lipocalins, such as L-PGDS [52] and glycode-lins [53], etc., have other specific functions. It is suggested that these proteins originated from the same gene, but now have different genes and functions after gene duplication and evolution [16,53]. Indeed, the behavior of β LG described above is shared by other members of the family.

3. Molecular mechanism of the Tanford transition

3.1. Equilibrium transition

Although β LG exists in a native state over a wide range of pH values, it shows slight conformational changes during a change of pH [54]. Among the pH-dependent conformational changes of β LG, the Tanford transition is the most important because it is thought to be related to the function of β LG. Tanford et al. [55] observed a change in optical rotatory dispersion at pH 7.0 representing a certain conformational change. Subsequently, they found that this conformational change is accompanied by a deprotonation of a carboxyl group with an anomalous pK_a of 7.5 [20,55].

Qin et al. [20] reported an important observation about the Tanford transition from crystal structures solved at pH 6.2 and 7.1. By comparing these two crystal structures, they found that the EF loop,

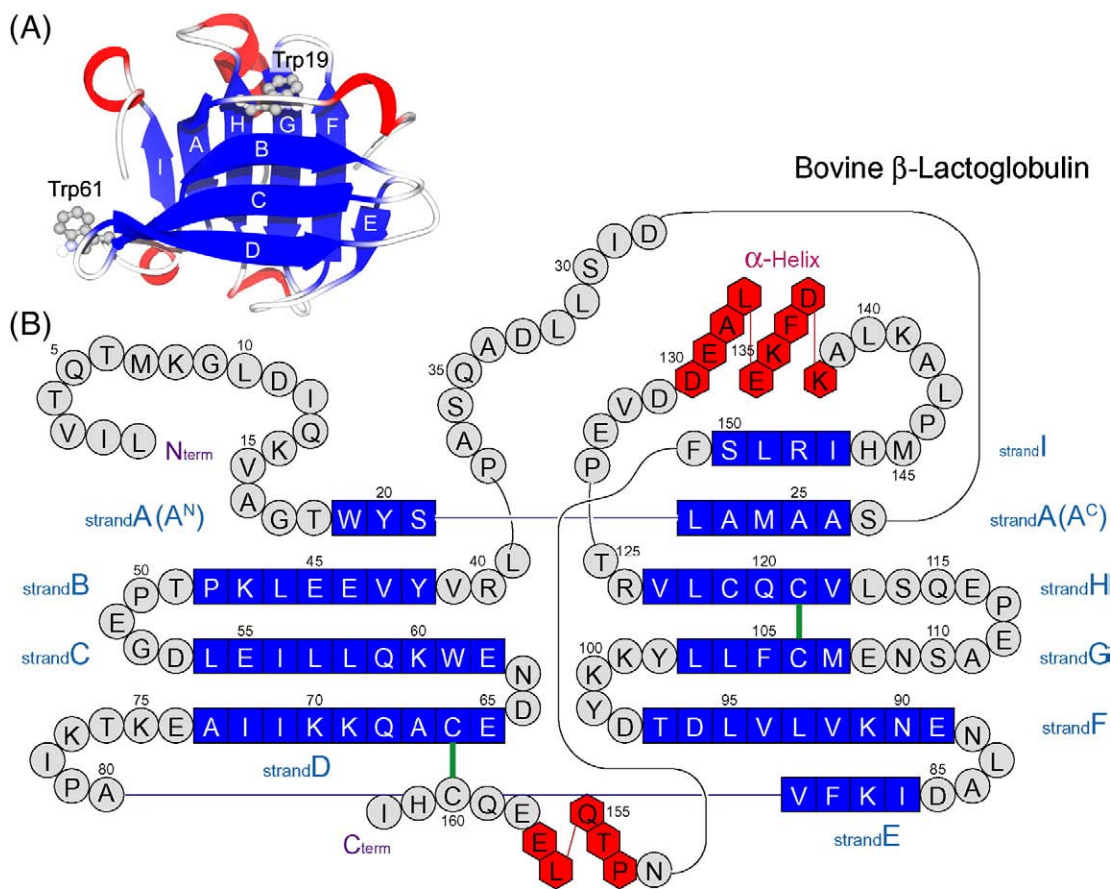


Fig. 1. The 3D structure and amino acid sequence of bovine β LG. (A) Ribbon diagram of a single subunit of bovine β LG lattice X, whose pdb code is 1BEB [19]. The β -strands are labeled. Trp residues are represented as balls and sticks. The diagram was produced using the program MolFeat (FiatLux, Tokyo, Japan). (B) A schematic representation of the amino acid residues of the β LG sequence. Residues making up the α -helix, β -sheet, and loop are represented by hexagons in red, squares in blue, and circles in grey, respectively. Green lines indicate the positions of disulfide bonds. It is seen that β LG has two β -sheets; The B–D strands and N-terminal half of the A strand (denoted A^N) consist of one and the E–H strands and C-terminal half of the A strand (denoted A^C) consist of the other.

one of the loops located at the open end of the central cavity, is oriented as if it “closes” the entrance to the cavity at pH 6.2, while it “opens” the entrance at pH 7.1. Furthermore, Glu 89, located on the EF loop, is buried and protonated in the closed form, while it is exposed and deprotonated in the open form. Glu 89 is attributed to the anomalous pK_a of 7.5 [20]. The conformational change seems to control the binding of ligands to β LG, since the EF loop is likely to regulate access to the central cavity [19,20,55]. In addition, Vijayalakshmi et al. [56] reported an interesting crystal structure of a β LG dimer formed under low-humidity conditions. In the dimer, one monomer assumes open conformation whereas the other does half-closed conformation. Their results suggest that the populations of the open and closed forms are determined by a subtle balance of stabilizing interactions, which are determined not only by pH but also by other parameters, such as humidity in this case.

We suggested an energetic mechanism for the Tanford transition based on experiments with fluorescence [57]. The Gdn-HCl denaturation of β LG was measured at various pH values by monitoring tryptophan fluorescence. The pH-dependence of the free energy of the N and I states with respect to the U state was obtained by a global fitting of the pH- and Gdn-HCl-dependent fluorescence data with theoretical equations, which were derived by combining the pH-dependence of unfolding equilibrium [58] with its Gdn-HCl dependence [59]. The fitting indicated a mechanism whereby, at acidic pH, the stability of the native (ΔG_{NU}) and intermediate (ΔG_{IU}) conformations varies in the same manner as a function of pH, while, at neutral pH, ΔG_{NU} decreases and ΔG_{IU} is constant (Fig. 2). This breaks the apparent two-state behavior at acidic pH, causing the I state to accumulate at neutral pH. As will be described below, from a detailed analysis of the pH-dependent change monitored using heteronuclear NMR measurements, we showed that regions including the CD and EF loops are locally unfolded at neutral pH. We suggest that this locally denatured state corresponds to the “open” state reported by Tanford et al. [55] and Qin et al. [20].

3.2. Molecular dynamics and NMR with a ^{13}C -enriched ligand

Not only the equilibrium properties but also the dynamic properties of the Tanford transition were addressed. One study involved molecular dynamics simulations on the basis of the crystal structure, performed by Fogolari et al. [60]. Their energy calculations indicated that the closed and open conformations are favored when Glu 89 is protonated and deprotonated, respectively, a result expected from the crystal structures. However, the conformational transition from an open to a closed state was not observed within the

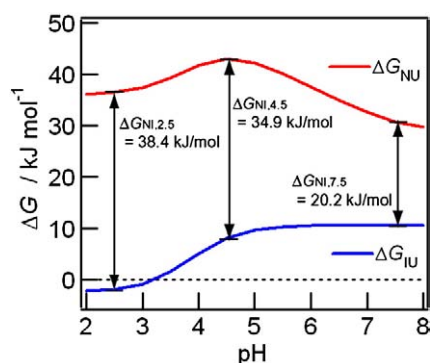


Fig. 2. pH-Dependence of folding free energies for native (red) and intermediate (blue) forms of wild-type β LG. This profile was obtained from the result of global fitting of the pH- and Gdn-HCl-dependent fluorescence spectra with the theoretical equations described in ref [62]. The energy difference at pH 7.5 between the native and intermediate states is smaller than that at pH 2.5 or pH 4.5, indicating that the relative population of the intermediate with respect to the native structure is larger at neutral pH than at acidic pH or at the isoelectric point (pI of β LG = 4.7).

period of their simulation (i.e. 3 ns), suggesting that the conformational transition occurs on a slower time scale. Ragona et al. [61] also investigated the dynamics of the Tanford transition by measuring NMR signals of a ^{13}C -enriched palmitic acid (PA) bound to β LG. They observed a split in the signal from the carboxyl carbon of the bound PA at around pH 7. This result also indicates that β LG showed the EF loop conformational change at this pH and the open and closed conformations are in a slow exchange, taking tens of milliseconds or more, with no discrepancy with the result of the simulation by Fogolari et al. [60].

3.3. Heteronuclear NMR approaches

We investigated the Tanford transition of β LG using heteronuclear NMR measurements [57,62]. Since severe signal broadening due to monomer–dimer equilibrium prevented NMR measurements of wild-type β LG at neutral pH, we searched for optimal sample conditions, finding that a disulfide bond linked-dimer of the mutant A34C gives a HSQC spectrum without signal broadening [57]. With the disulfide-linked dimeric A34C β LG as a model, we performed three types of NMR experiments to obtain information about the conformational dynamics of β LG in different time scales.

First, information on the dynamics over a time scale of the ps–ns was obtained from the change in chemical shifts. We measured pH-dependent HSQC spectral changes of β LG and applied a principal component analysis (PCA) to the chemical shift data [62]. Then, we found a cluster of residues with large chemical shift differences ($\Delta\delta$ s) at the G strand with a pK_a of 6.9 (Fig. 3A). Because these residues are located near the carboxyl group of Glu 89, the chemical shift changes seem due to the deprotonation of the carboxyl group.

Second, information of the motion on a time scale of μ s–ms was obtained from a Lipari–Szabo-type model-free analysis [63,64]. Among the parameters obtained from this analysis, we focused only on the excess transverse relaxation constant, R_{ex} , which is related to the degree of motion on this scale [57]. It was found that Ile 84, Asn 90, and Glu 108 showed notable differences in R_{ex} between the two pH values (Fig. 3B). These three residues are located at the hinge part of the EF loop flip motion and form hydrogen bonds with each other. Thus, we attributed the observed micro- to millisecond motions to fluctuations of hydrogen bonds in these residues. This motion becomes significant at pH higher than 7.0.

Third, for information about motion on a time scale of ms or slower, we inspected the change in signal intensity upon pH change, finding that residues at the D strand, EF loop and GH loop showed decreases in their signals without significant $\Delta\delta$ (Fig. 3C). Thus, we consider that the timescale of the conformational change to these regions is in the order of microseconds or slower.

By combining these events, a series of conformational changes giving rise to the Tanford transition has been suggested (Fig. 4) [57,62]. First, the carboxyl group of Glu 89 is deprotonated. Next, there is a fluctuation in the hydrogen bonds between the backbone atoms at residues Ile 84, Asn 90, and Glu 108. Finally, unfolding at the D strand and the EF and GH loops occurs. Although the initial deprotonation makes the open conformation more stable, the conversion to it cannot occur simultaneously because the second and third steps have different rate-limiting factors. As discussed above, a β LG molecule can assume various structural states, and the populations of each state are determined by a mechanism linking these states, which together describe the pH-dependent Tanford transition (Fig. 4).

It is interesting that the open–close behavior of β LG has not been reported for other calycin proteins. For example, there are no reports of pH-dependent structural changes for L-PGDS and I-FABP: only unique structures of these proteins were reported [65,66]. However, loops on the open side of the β -barrel show significant fluctuations. This is the same feature as that of the open form of β LG. Thus, it is tempting to consider that the open–close behavior is unique to β LG,

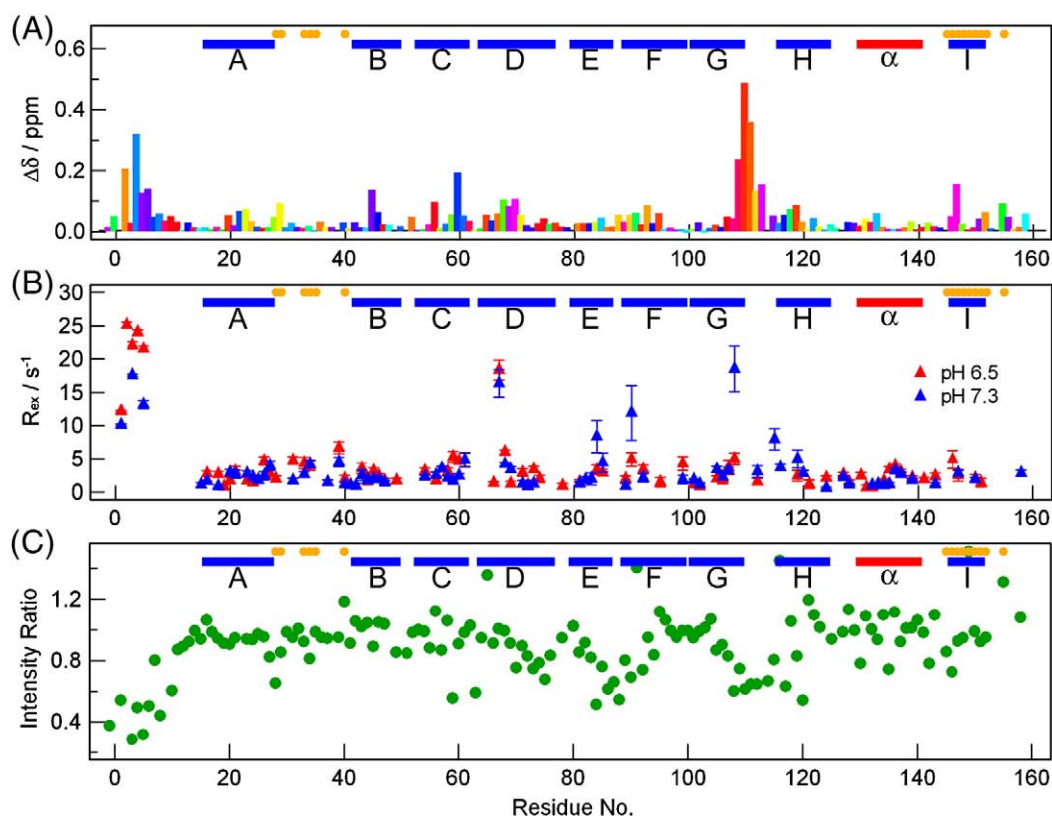


Fig. 3. Dynamics of β LG on various time scales. (A) Chemical shift differences for the transitions at pH 6.8 as information on motion over a time scale of ps–ns. This profile was obtained through the principal component analysis [62]. Heights and colorations of the bars indicate the distances and directions of the signal displacements, respectively. The blue and red bars at the top of each panel indicate the positions of β -strands and α -helices, respectively, and the orange circles are those of interfacial residues of the β LG dimer. (B) Comparison of R_{ex} values of β LG A34C obtained from relaxation analyses at pH 6.5 (red) and 7.3 (blue) as information on a time scale of μ s–ms. The arrows indicate the data on Ile 84, Asn 90, and Glu 108, which show a significant difference in R_{ex} values between pH 6.5 and pH 7.3. (C) Ratio of intensities of HSQC signals between pH 6.5 and pH 7.3 as information on motion over a time scale of ms–slower. Significant decreases in intensity were observed at the D strand and the EF and GH loops. Adopted from [57] and [62].

and the open form of β LG shares properties common with other calycin superfamily members.

4. Ligand binding

One major characteristic of β LG is its binding of various hydrophobic ligands. The binding of ligands has been studied extensively [16]. It has been proposed that β LG enhances the uptakes of fatty acids and vitamin D in vivo [67,68], although it is still controversial whether the binding property is relevant to the biological role of β LG. Specific binding of β LG had already been recognized in the 1940s [69]. Since

then, many ligands have been observed to bind β LG. The binding constants for the ligands were determined by a variety of methods [14,17,25,48]. Since the end of the 1990s, crystal structures of β LG in complexes with various ligands have been reported [16,21,70–72]. Although several possible binding sites were seen in the crystal structures, most ligands lie within the central cavity of the protein, indicating that the cavity is the primary binding site of ligands [16,21,70,71]. The central cavity seems to accommodate linear molecules like fatty acids and also retinol.

A β LG complex with palmitic acid (PA) reported by Wu et al. [71] provides important insight into the mechanism of binding. The

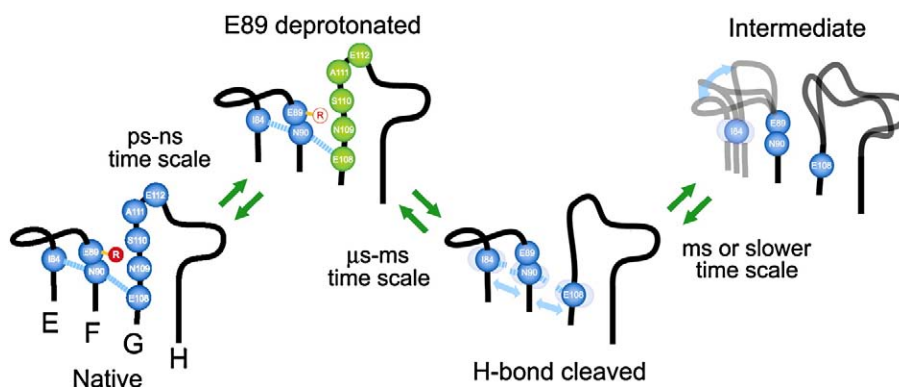


Fig. 4. The process of the closed-open conformational transition of β LG. This is an illustration of each step in the Tanford transition. The first step: the deprotonation of the carboxyl group of Glu 89. Next: the fluctuation or cleavage of hydrogen bonds between backbone atoms at Ile 84, Asn 90, and Glu 108. Third: unfolding at the D strand and the EF and GH loops occurs. The native, or closed, state, which is the leftmost form of the figure, is accessible to the intermediate, or open, state, the rightmost form, through deprotonation of the carboxyl group of Glu 89.

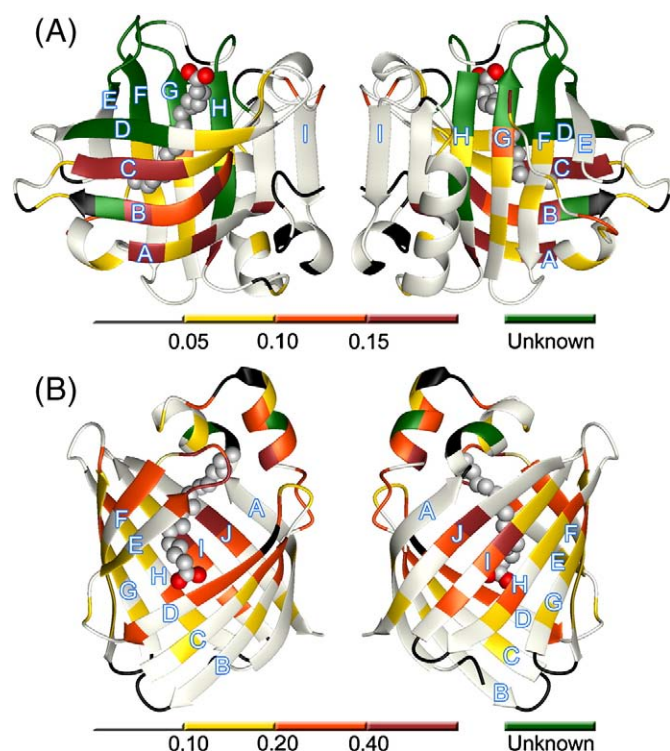


Fig. 5. Comparison of the spectral changes upon ligand binding of βLG and I-FABP. (A) Mapping of $\Delta\delta$ s of βLG. Residues whose $\Delta\delta$ is larger than 0.15, 0.10, and 0.05 ppm are colored red, orange, and yellow, respectively. Unassigned residues are colored green. (B) Mapping of $\Delta\delta$ s of I-FABP. Residues whose $\Delta\delta$ is larger than 0.4, 0.2, and 0.1 ppm are colored red, orange, and yellow, respectively. Residues that appeared upon the binding of PA are green. In both panels, left and right images are front and rear views of the crystal structures, respectively. Adopted from [73].

aliphatic tail of PA is deeply stuck into the barrel of the βLG molecule and fixed by hydrophobic interaction, while the carboxylate head group of PA electrostatically interacts with the amino groups of the Lys 60 and Lys 69 side chains. With this crystal structure, the βLG-PA complex seems rigid.

With the disulfide-linked dimeric A34C βLG as a model of βLG, we studied the mechanism of the binding of PA. Again, the advantage of using the dimeric A34C βLG was obvious: with highly resolved HSQC spectra at pH 6.5 and 40 °C, we could monitor the effects of binding at respective residues. Titration with PA monitored with HSQC spectra indicated that the fluctuation of the βLG molecule becomes more prominent upon the binding of PA [73]. HSQC spectra and transverse relaxation rates (R_2 s) in the presence and absence of PA were compared, revealing that some motions on a time scale of μ s–ms are enhanced by the binding of PA at residues at the open end of the barrel (Fig. 5A). On the other hand, residues encompassing the cavity were found to interact rigidly with the PA molecule [74,75]. The results of the R_2 measurements support this picture: significant enhancements of the R_2 upon the binding of PA were observed for the residues around the entrance of the barrel.

It should be noted that the pattern of change in signal intensity obtained from the PA-binding experiment (Fig. 5A) is quite similar to that from the pH titration experiment (Fig. 3C). This suggests that the binding of PA induces a conformational change similar to that from a closed to open form, as seen in the Tanford transition [4,19,55]. Thus, the binding of PA seems to make the residues located at the entrance of the barrel more mobile. The population of the open form, or the partially unfolded intermediate, is minor at neutral pH. However, when the intermediate encounters a ligand, they form a stable complex, making its population dominant. This interpretation is similar to the case of the “pre-existing model” or “population shift” schemes of ligand binding [76].

More evidence supporting the flexibility of the entrance region is provided by the crystal structures of βLG–ligand complexes. Six pdb codes are available for βLG molecules in complexes with ligands in the cavity [21,70–72]. After the C α traces of these structures were aligned, RMSD values for each residue were calculated. The residues with a high RMSD are seen at the D strand, EF loop and GH loop (Fig. 6), indicating that the entrance region is capable of adopting a variety of conformations in the open form.

NMR measurements of the binding of PA to I-FABP performed by Hodsdon et al. [65] showed contrasting behavior (Fig. 5B). In I-FABP, residues with a large $\Delta\delta$ upon the binding of the ligand are distributed around the barrel, similar to the case of βLG. However, the number of observable signals for I-FABP increased upon the binding. Furthermore, a relaxation analysis revealed that the “portal site”, which acts as a lid to the cavity and is mobile in the ligand-free conformation, becomes more rigid upon the binding. These results indicate that the binding of the ligand leads to the immobilization of some residues in I-FABP, in marked contrast to the case of βLG (Fig. 5A). On the other hand, the binding mechanism of mouse major urinary protein (MUP), which is one of calycin protein, to its ligand seems similar to that of βLG. Zidek et al. [77] estimated the entropic contribution to the free energy of the binding of MUP to the ligand arising from (NMR-derived) backbone motions. They concluded that, although the residues with increased mobility after ligand binding are distributed throughout the β-barrel structure of MUP, these residues are particularly prevalent in a helical turn and irregularly structured region (residues 27–40) proposed to form a ligand entry gate at one end of the binding cavity. The conclusions are similar to those of βLG.

L-PGDS, which is also one of calycin proteins, has a huge cavity larger than that of βLG. The binding site has specific residues complementary to the conformation of its inherent substrate, prostaglandin H₂ (PGH₂), tightly confining the ligand in the bound form. However, at the bottom of the cavity, L-PGDS can simultaneously bind another hydrophobic ligand [66]. Because there are no residues which could constrain the ligand's conformation by electrostatic interactions or hydrogen bonds, the binding of the ligand seems to be driven by hydrophobic interactions, and thus less specific.

The dominant force driving the binding of ligands to the calycin members is hydrophobic interaction, as the central cavity is almost completely lined by hydrophobic residues. Jameson et al. [78] calculated surface free energy of the hydrophobic cavity of βLG

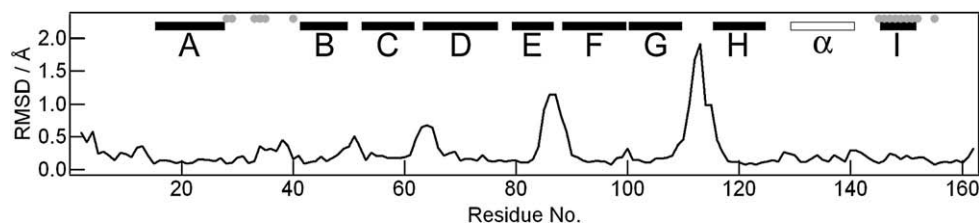


Fig. 6. RMSD of C α of six structures of ligand-bound βLGs. Structural variability of βLG in the ligand-bound state was examined as RMSD of main chain C α atom coordinates. As sample structures of βLG in the ligand-bound state, the coordinates with 2GJ5, 1B00, 1BS0, 1GX8, 1GX9 and 1GXA, whose ligands are vitamin D, PA, 12-bromododecanoic acid, retinol, retinoic acid and palmitate, respectively [21, 70–72], were used. The maxima of the RMSD are seen at the D strand, and the EF and GH loops.

considering the surface tension. They discussed that, although maintaining the hydrophobic cavity imposes entropic (or free energy) cost, the binding of ligands compensates for the energetic cost. In addition, the open end of the barrel has structural variability which confers low specificity but high affinity. This scheme resembles the recognition of substrates by GroEL [79–81]. GroEL recognized the exposed hydrophobic regions of fully or partly unfolded proteins. The substrate-binding site in the apical domain is hydrophobic and flexible and can indiscriminately catch various substrates. However,

additional specificity is provided by other geometric factors, that is, the shape and volume of the cavity and the arrangement of a few hydrophilic or charged residues in the cavity. Recently, binding promiscuity has been proposed as a basic mechanism increasing the complexity of networks, in which intrinsically unstructured proteins might play important roles [82–84]. With β LG, we have another example of binding with low specificity but high affinity, which may play an important role in protein-ligand and protein-protein networks.

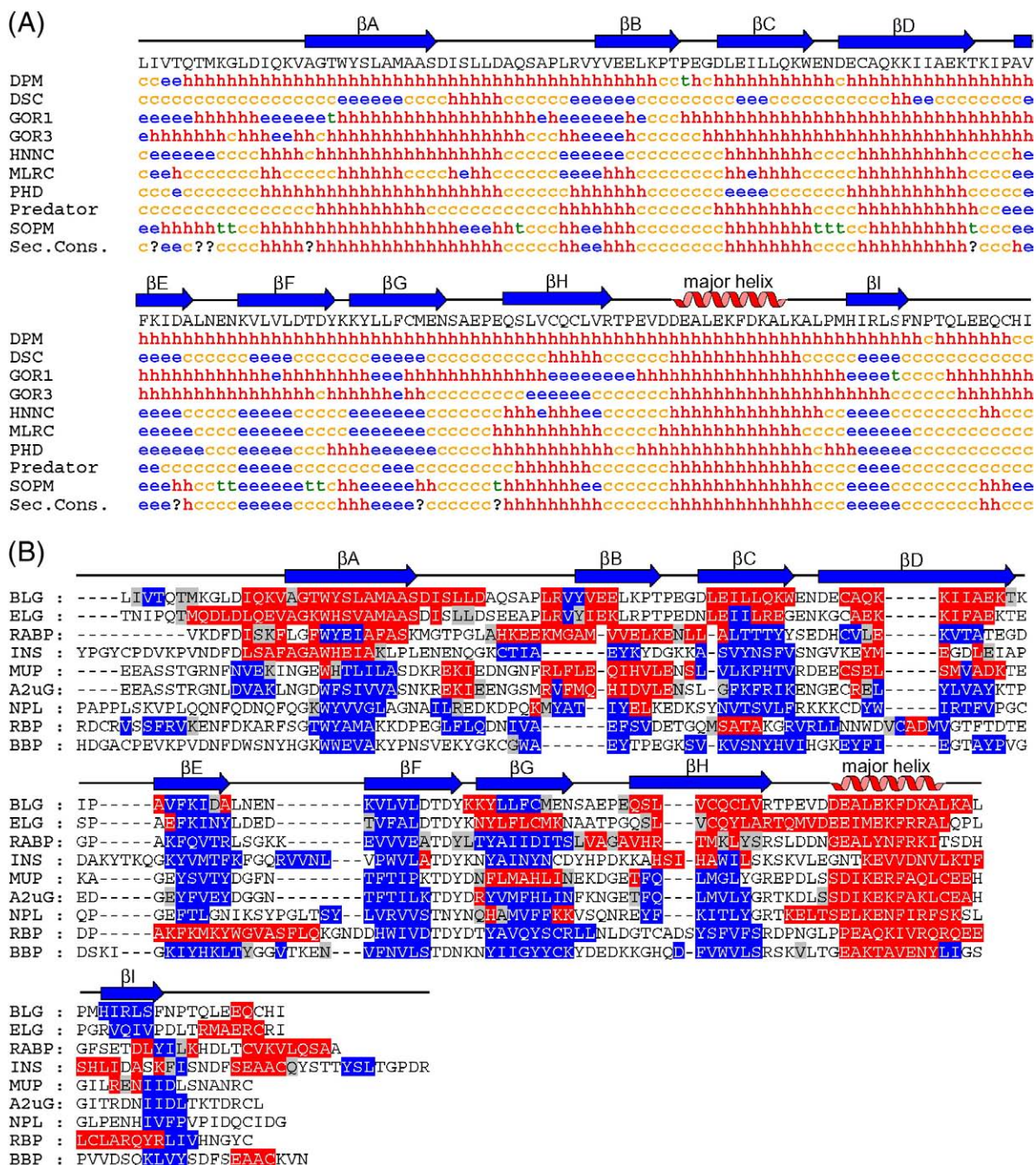


Fig. 7. High helical propensity of β LG in comparison with other lipocalin proteins. (A) Probabilities of secondary structure of β LG predicted by DPM [112], DSC [113], GOR1 [114], GOR3 [115], HNNC [116], MLRC [117], PHD [118], Predator [119], and SOPM [120] algorithms indicated by h (helix), e (sheet), t (turn), and c (coil). These algorithms were executed with ANTHEPROT [121]. The secondary structural elements determined on the basis of the crystal structure of bovine β LG (PDB entry 1BEB) is shown on the top. (B) Secondary structure predictions for lipocalin proteins. Each result shown is a consensus of the prediction algorithms used in (A). The proteins analyzed here are as follows; BLG, bovine β LG; ELG, equine β LG; RABP, epididymal retinoic acid-binding protein; INS, insecticystanin; MUP, mouse major urinary protein; A2uG, α_{2u} -globulin; NPL, neutrophil lipocalin; RBP, retinol-binding protein; BBP, bilin-binding protein. Sequences of these proteins were quoted from ref. [122] and articles therein. Residues indicated by blue and red are predicted to be in β -strand and α -helix, respectively. Residues indicated by grey are those which were not given a clear prediction.

5. Folding with non-native intermediates

β LG folds through a unique pathway with non-native intermediates [85]. On the basis of stopped-flow experiments with conventional spectrometric methods [28,31,85,86], the folding scheme for β LG has been summarized as follows;



where U and N are the unfolded and native states, respectively, and Is are different kinds of intermediate. I_{burst} , a burst-phase intermediate observed in a CD stopped-flow experiment, is unique to β LG because it contains a significant amount of a non-native α -helical component. In the subsequent I_{fold} state, the α -helical content is decreased and concomitantly the β -sheet content is increased. The I_{fold} state might include two or four distinct intermediates from CD and fluorescence stopped-flow experiments, respectively [31,85]. However, because these intermediates (I_{1-2} or I_{1-4}) have not been further characterized, here we combine them for convenience. The unique folding process for β LG is often referred to as the α - β transition, and considered to be a case of non-hierarchical folding which might play an important role in biological processes [87,88]. Related to this, it has been suggested that β LG in the partially unfolded state with a non-native α -helix forms amyloid fibrils [89].

The accumulation of the non-native α -helix is explained by a high α -helical propensity of its amino acid sequence according to secondary structure predictions [90,91] (Fig. 7A). However, the non-native α -helices are eventually converted to the native β -structure because of the non-local interactions formed in the subsequent refolding steps [85,92]. For clarification of the folding of these proteins, the characterization of folding intermediates using heteronuclear NMR combined with hydrogen/deuterium (H/D) exchange will be useful [38,93].

The I_{burst} state of β LG has been investigated by H/D exchange pulse labeling combined with NMR by Kuwata et al. [31]. They have characterized the structural and dynamic properties of I_{burst} . The folding of β LG was monitored over the 100 μ s to 10 s range using ultrarapid mixing techniques in conjunction with fluorescence detection and H/D exchange. Trp fluorescence indicated that the conversion to I_{burst} occurs with a time constant of 140 μ s. Within 2 ms of refolding, well protected amide protons indicative of a stable hydrogen bonded structure were found only in a domain comprising β -strands F, G and H, and the main α -helix. On the other hand, weak protection of amide protons in a segment spanning residues 12–21

was observed, which is consistent with the formation of marginally stable non-native α -helices in the A strand region (Fig. 8).

In addition, we should mention the result of Ragona et al. [94]. An analysis of the urea-concentration-dependent H/D exchange rate showed that most of the residues are in a unique cooperative folding unit, encompassing most strands. They suggested that the formation of the I_{burst} state is tightly coupled with the formation of the N state, which leads to a two-state transition in equilibrium unfolding.

It is quite interesting that equine and porcine β LGs showed a similar intermediate with a non-native α -helical content [95,96]. Among them, precise characterization of the folding intermediate of equine β LG was performed by Ikeguchi et al. [95]. They observed a significant change in CD intensity within the dead time (25 ms) of stopped-flow measurements, suggesting a transient accumulation of the folding intermediate. They further found that equilibrium molten globule intermediates occurring during pH- and urea-titration experiments are equivalent to the kinetic burst-phase intermediate [97]. The α -helical content of the acid-unfolded form was greater than that of the native conformation, similar to bovine β LG. They measured hydrogen exchange protection of the molten globule using NMR [98]. In the molten globule form at acidic pH, significant protection from exchange was observed for residues located in the A, F, G and H strands in the native structure. The pattern of protection is consistent with a native-like β -sheet formation by these strands. The residues located in the major helix of the native structure are also protected, suggesting that this helix is formed in the molten globule and is packed against the sheet as in the native structure. The intermediate of equine β LG, therefore, has quite similar characteristics to that of bovine β LG.

From these reports, an idea emerges; the amino acid sequences of β LGs are designed so as to fold in a stepwise manner, that is, the first step is the formation of the β -sheet consisting of A^C and E–H strands, and the second step is the formation of the other β -sheet consisting of A^N and B–D strands. During the first step, redundant donors and acceptors of the hydrogen bond in the main chain of A–D strands should be sequestered not to interrupt the E–H sheet's formation. Thus, Nature might give local α -helix propensity to the A strand to sequester excess donors and acceptors. This mechanism might apply to other lipocalins, and indeed to other protein families as well. Many of the amino acid sequences of lipocalin proteins also show high α -helical propensity in the A and B strand region (Fig. 7B) and, similar protection was observed in the molten globule intermediate of serum retinol-binding protein, a member of the lipocalin family [99]. The validity of this hypothesis should be examined in future work.

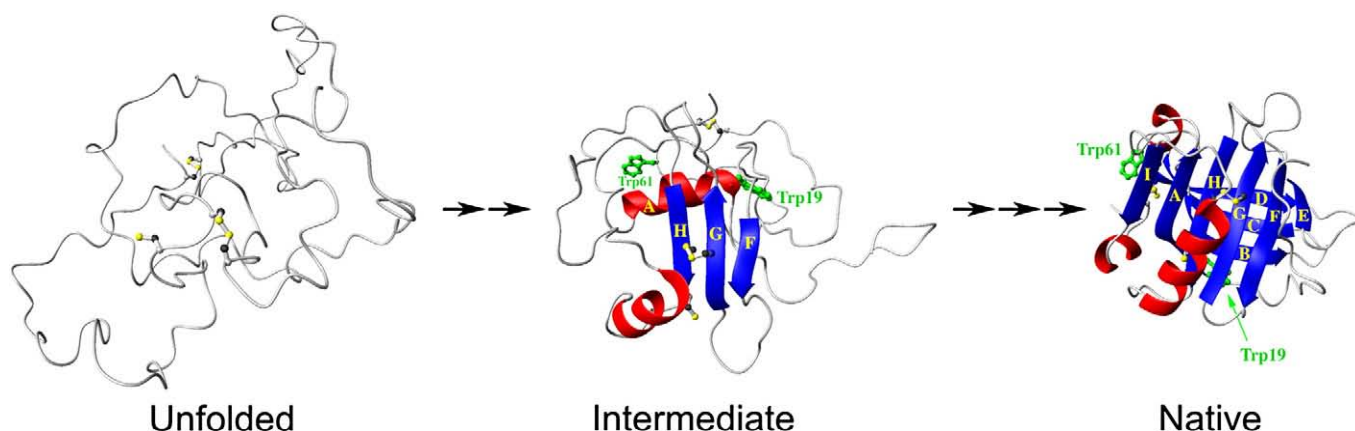


Fig. 8. Non-hierarchical folding of β LG. This is a schematic representation of the conformational states encountered during folding of β LG, including the unfolded ensemble, a partially structured intermediate populated on the millisecond time scale and the native state. The ribbon diagram was based on the X-ray structure with the PDB code, 3BLG [20]. The intermediate contains a hydrogen bonded structure as measured by burst-phase labeling in the core of the β -sheet (F, G and H) and the major α -helix, as well as some fluctuating helical structure near the N-terminus that is later converted into a β -sheet. Side chains are shown for the fluorescence probes, Trp 19 and Trp 61, and five cysteine residues, including two disulfides. Adopted from [31].

6. Folding of dimeric β LG

Finally, we would like to discuss a report by Yagi et al. [100]. In this paper, we addressed the relationship between the size of a protein and the complexity of its folding. Small proteins with a MW of less than 10,000 fold according to an ideal two-state scheme, and those with a MW of 10,000–20,000, such as β LG and apomyoglobin, fold through intermediate states. Proteins of more than 20,000 seldom unfold reversibly and are apt to aggregate. However, in the folding process in cells, such large proteins circumvent aggregative states through certain intermediates. For example, during domain-separated folding as observed for immunoglobulins [101,102] or the participation of special factors such as molecular chaperones [103–105], certain intermediates are seen. Improvements in our understanding of such mechanisms will allow applications for the expression, refolding, and purification of such larger proteins.

The C121A mutation was introduced to avoid disulfide shuffling in the unfolded state, improving the refolding yield from denaturation markedly [33]. It was confirmed that the folding kinetics of C121A β LG is the same as that of wild-type β LG (scheme 1 in the previous section), from stopped-flow CD and fluorescence measurements. Thus, we treated C121A β LG as the pseudo-wild type of β LG. Then we prepared a double mutant, A34C/C121A β LG. This is a dimeric mutant in which two β LG molecules are linked by a disulfide bridge through Cys 34 created at the dimer's interface.

Interestingly, the folding was perturbed significantly by the introduction of a dimer mutation, A34C/C121A. Although, A34C/C121A β LG showed a notable burst-phase amplitude corresponding to the non-native α -helical intermediate of the wild type within the dead time of the folding experiment, the mutant was trapped in a state with hydrophobic regions exposed and did not attain the native state. H/D exchange pulse labeling with a folding time of 10 ms combined with NMR analysis [37,106,107] indicated that the burst-phase intermediate of both C121A and A34C/C121A β LGs had several structured regions, including the F, G, and H strands and the loop region after the major helix, which corresponds to I_{burst} in scheme 1. Thus, both mutants pass through the same pathway until the I_{burst} state.

These observations revealed that the initial step in the folding of β LG includes the formation of core secondary and tertiary structures driven by local hydrophobic interactions. This step is independent of

molecular size. The subsequent processes include the acquisition of peripheral structures with a rearrangement of the non-native α -helices into the native β -strands and the concomitant formation of specific side chain interactions, resulting in accumulation of the I_{fold} intermediates. The latter processes are controlled by a balance between the stabilization of native interactions and destabilization of non-native interactions, so that they are apt to be perturbed by MW-doubling leading to the accumulation of quasi-stable (or kinetically trapped) non-native conformations under conditions where the non-native interactions are too strong. The ensemble of these trapped conformations represents a “glass”-like misfolded state, a consequence of structural inconsistency.

One interesting finding is that, although this mutant β LG cannot refold at low concentrations of urea, it refolds at moderate concentrations of urea. It is assumed that, in the presence of 3.0 M urea, A34C/C121A β LG manages to refold to the native structure, jumping over the kinetic high energy barriers. On the other hand, in 0.82 M urea, the concentration at which wild-type β LG refolds properly, the incorrect interactions are too stable to rearrange into the native forms, and so the molecules are trapped in the misfolded conformation with an exposed hydrophobic surface. This in turn leads to the low reversibility under strongly native conditions (Fig. 9).

For large proteins, kinetic trapping by non-native interactions would be inevitable. Nature should have created ways of circumventing such a trap. In the case of an artificial protein, A34C/C121A β LG, interactions that are too strong make the folding funnel rough producing many kinetic traps which prevent efficient and rapid folding [35,36]. It is assumed that many large proteins can refold by taking advantage of a mechanism similar to that observed for the refolding of A34C/C121A β LG. One approach to preventing such kinetic trapping is, as demonstrated here in 3.0 M urea, to destabilize the native form, which simultaneously has the effect of destabilizing the trapped intermediate. An alternative approach might be to make the folding pathways evading the kinetic traps by some modifications of the energy surface, instead of straight paths from each unfolded conformation to the native form running on a smooth folding funnel.

7. Conclusion

The ideal two-state folding reaction might be explained by the folding funnel or consistency principle, that is, if the interactions in

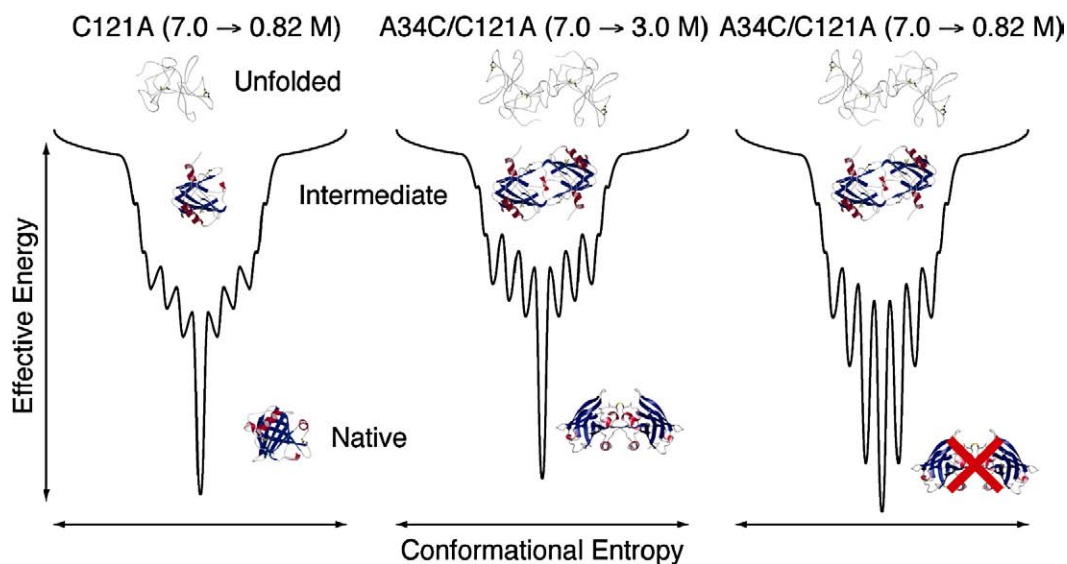


Fig. 9. Schematic representation of the folding mechanisms of the β LG mutants using the folding funnels. At the early stage of refolding, the β LG mutants fold similarly, driven by hydrophobic interactions. In the following steps, while the monomeric β LG folds as smoothly as wild-type β LG, the disulfide-linked dimeric β LG folds much more slowly due to many non-native interactions causing the rough funnel surface. Unless the non-native interactions are destabilized by moderate concentrations of denaturant, making the funnel surface smoother, the dimeric β LG remains in a glasslike trapped state. Adopted from [100] with permission.

the native conformation are consistent with each other, the energy landscape along the folding reaction coordinate becomes funnel-like with the smooth surface [108,109]. However, when the proteins become larger, their folding reactions become multi-state and/or irreversible because the consistency is lost and local minima are generated on the folding energy surface.

Recently, an interesting idea about a relation between a non-native folding intermediate and conformational frustration in the native structure has been reported: Ferreira et al. [110,111] proposed that, if there are localized frustrations in the native structure, a non-native folding intermediate will occur in the folding process. This idea seems to be consistent with the properties of β LG. Indeed, we have reported the existence of a non-native intermediate during the folding of β LG [85] and local frustration on the EF loop conformation in the native state [57]. Furthermore, such local frustration in the native state is probably related to the conformational dynamics [57]. Conformational frustration is likely to occur more frequently with an increase in the size of proteins. Because β LG is relatively large, ~20 kDa, it will show behavior unexpected for small proteins. Even though the mechanisms of such behavior are still enigmatic, studies of the folding and conformational dynamics of β LG will lead to a general understanding of the conformation, dynamics and folding of proteins.

8. Methods

8.1. RMSD calculation of the structures of ligand-bound β LGs

Structural variability of β LG in a ligand-bound state was examined as RMSD of main chain C α atom coordinates. The following calculations were performed using Igor Pro (WaveMetrics, Lake Oswego, OR). As a template of β LG in the open state, the coordinate with PDB ID 1BSY was used [20]. As sample structures of β LG in the ligand-bound state, the coordinates with 2GJ5, 1B00, 1GX8, 1GX9 and 1GXA, whose ligands are vitamin D, PA, 12-bromododecanoic acid, retinol, retinoic acid and palmitate, respectively [21,70–72], were used. First, each sample coordinate was aligned to the template coordinate by the least-squares curve fitting algorithm in the Igor software package with a self-written alignment program, in which Δx , Δy , Δz (translations along each axis), θ_x , θ_y , and θ_z (rotation about each axis) were adjustable parameters. Then deviations in C α of the six aligned coordinates from that of the template coordinate were calculated. The final RMSD values were calculated using these deviations from the six sampled coordinates together. Details of other materials and methods were described in the original papers.

Acknowledgements

We thank Carl A. Batt (Cornell University) for valuable comments, Hideo Akutsu, Takahisa Ikegami (Institute for Protein Research) and Masaru Hoshino (Kyoto University) for instructions regarding NMR measurements. This work was supported, in part, by Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Science, Culture and Sports.

References

- [1] S. Pervaiz, K. Brew, Homology of β -lactoglobulin, serum retinol-binding protein, and protein HC, *Science* 228 (1985) 335–337.
- [2] J. Godovac-Zimmermann, The structural motif of β -lactoglobulin and retinol-binding protein: a basic framework for binding and transport of small hydrophobic molecules? *Trends Biochem. Sci.* 13 (1988) 64–66.
- [3] S.G. Hambling, A.S. MacAlpine, L. Sawyer, β -lactoglobulin, in: P.F. Fox (Ed.), *Advanced Dairy Chemistry 1, Proteins*, Elsevier, Amsterdam, 1992, pp. 141–190.
- [4] L. Sawyer, G. Kontopidis, The core lipocalin, bovine β -lactoglobulin, *Biochim. Biophys. Acta* 1482 (2000) 136–148.
- [5] J. Cerbulis, H.M. Farrell Jr., Composition of milks of dairy cattle. I. Protein, lactose, and fat contents and distribution of protein fraction, *J. Dairy Sci.* 58 (1975) 817–827.
- [6] M. Brunori, Structural dynamics of myoglobin, *Biophys. Chem.* 86 (2000) 221–230.
- [7] I. Schlichting, K. Chu, Trapping intermediates in the crystal: ligand binding to myoglobin, *Curr. Opin. Struct. Biol.* 10 (2000) 744–752.
- [8] R.L. Baldwin, On-pathway versus off-pathway folding intermediates, *Fold. Des.* 1 (1996) R1–8.
- [9] T. Fukamizo, Chitinolytic enzymes: catalysis, substrate binding, and their application, *Curr. Protein Pept. Sci.* 1 (2000) 105–124.
- [10] M.J. Adams, G.C. Ford, R. Koekoek, P.J. Lentz, A. McPherson Jr., M.G. Rossmann, I.E. Smiley, R.W. Schevitz, A.J. Wonacott, Structure of lactate dehydrogenase at 2–8 Å resolution, *Nature* 227 (1970) 1098–1103.
- [11] A.R. Clarke, T. Atkinson, J.J. Holbrook, From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part II, *Trends Biochem. Sci.* 14 (1989) 145–148.
- [12] C. Tanford, Y. Nozaki, Physico-chemical comparison of β -lactoglobulins A and B, *J. Biol. Chem.* 234 (1959) 2874–2877.
- [13] R.D. Fugate, P.S. Song, Spectroscopic characterization of β -lactoglobulin–retinol complex, *Biochim. Biophys. Acta* 625 (1980) 28–42.
- [14] L. Sawyer, S. Brownlow, I. Polikarpov, S.Y. Wu, β -Lactoglobulin: structural studies, biological clues, *Int. Dairy J.* 8 (1998) 65–72.
- [15] J.E. Kinsella, D.M. Whitehead, Proteins in whey: chemical, physical, and functional properties, *Adv. Food Nutr. Res.* 33 (1989) 343–438.
- [16] G. Kontopidis, C. Holt, L. Sawyer, Invited review: β -lactoglobulin: binding properties, structure, and function, *J. Dairy Sci.* 87 (2004) 785–796.
- [17] L. Sawyer, β -lactoglobulin, in: P.F. Fox, P. McSweeney (Eds.), *Advanced Dairy Chemistry 1*, 3rd ed, Kluwer, Amsterdam and New York, 2003, pp. 319–386.
- [18] M.Z. Papiz, L. Sawyer, E.E. Eliopoulos, A.C. North, J.B. Findlay, R. Sivaprasadarao, T.A. Jones, M.E. Newcomer, P.J. Kraulis, The structure of β -lactoglobulin and its similarity to plasma retinol-binding protein, *Nature* 324 (1986) 383–385.
- [19] S. Brownlow, J.H. Morais Cabral, R. Cooper, D.R. Flower, S.J. Yewdall, I. Polikarpov, A.C. North, L. Sawyer, Bovine β -lactoglobulin at 1.8 Å resolution—still an enigmatic lipocalin, *Structure* 5 (1997) 481–495.
- [20] B.Y. Qin, M.C. Bewley, L.K. Creamer, H.M. Baker, E.N. Baker, G.B. Jameson, Structural basis of the Tanford transition of bovine β -lactoglobulin, *Biochemistry* 37 (1998) 14014–14023.
- [21] B.Y. Qin, L.K. Creamer, E.N. Baker, G.B. Jameson, 12-Bromododecanoic acid binds into the calyx of bovine β -lactoglobulin, *FEBS Lett.* 438 (1998) 272–278.
- [22] J.F. Cupo, C.N. Pace, Conformational stability of mixed disulfide derivatives of β -lactoglobulin B, *Biochemistry* 22 (1983) 2654–2658.
- [23] A.C. Altling, H.H. de Jongh, R.W. Visschers, J.W. Simons, Physical and chemical interactions in cold gelation of food proteins, *J. Agric. Food Chem.* 50 (2002) 4682–4689.
- [24] R. Mercade-Prieto, R.J. Falconer, W.R. Paterson, D.I. Wilson, Effect of gel structure on the dissolution of heat-induced β -lactoglobulin gels in alkali, *J. Agric. Food Chem.* 54 (2006) 5437–5444.
- [25] Y. Cho, C.A. Batt, L. Sawyer, Probing the retinol-binding site of bovine β -lactoglobulin, *J. Biol. Chem.* 269 (1994) 11102–11107.
- [26] Y. Cho, W. Gu, S. Watkins, S.P. Lee, T.R. Kim, J.W. Brady, C.A. Batt, Thermostable variants of bovine β -lactoglobulin, *Protein Eng.* 7 (1994) 263–270.
- [27] T.R. Kim, Y. Goto, N. Hirota, K. Kuwata, H. Denton, S.Y. Wu, L. Sawyer, C.A. Batt, High-level expression of bovine β -lactoglobulin in *Pichia pastoris* and characterization of its physical properties, *Protein Eng.* 10 (1997) 1339–1345.
- [28] K. Kuwata, M. Hoshino, S. Era, C.A. Batt, Y. Goto, $\alpha \rightarrow \beta$ transition of β -lactoglobulin as evidenced by heteronuclear NMR, *J. Mol. Biol.* 283 (1998) 731–739.
- [29] K. Kuwata, M. Hoshino, V. Forge, S. Era, C.A. Batt, Y. Goto, Solution structure and dynamics of bovine β -lactoglobulin A, *Protein Sci.* 8 (1999) 2541–2545.
- [30] S. Uhrínová, M.H. Smith, G.B. Jameson, D. Uhrín, L. Sawyer, P.N. Barlow, Structural changes accompanying pH-induced dissociation of the β -lactoglobulin dimer, *Biochemistry* 39 (2000) 3565–3574.
- [31] K. Kuwata, R. Shastri, H. Cheng, M. Hoshino, C.A. Batt, Y. Goto, H. Roder, Structural and kinetic characterization of early folding events in β -lactoglobulin, *Nat. Struct. Biol.* 8 (2001) 151–155.
- [32] K. Sakurai, Y. Goto, Manipulating monomer–dimer equilibrium of bovine β -lactoglobulin by amino acid substitution, *J. Biol. Chem.* 277 (2002) 25735–25740.
- [33] M. Yagi, K. Sakurai, C. Kalidas, C.A. Batt, Y. Goto, Reversible unfolding of bovine β -lactoglobulin mutants without a free thiol group, *J. Biol. Chem.* 278 (2003) 47009–47015.
- [34] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, *Nat. Struct. Biol.* 4 (1997) 10–19.
- [35] N.D. Socci, J.N. Onuchic, P.G. Wolynes, Protein folding mechanisms and the multidimensional folding funnel, *Proteins* 32 (1998) 136–158.
- [36] P.G. Wolynes, Energy landscapes and solved protein-folding problems, *Philos. Trans. A Math. Phys. Eng. Sci.* 363 (2005) 453–464 discussion 464–457.
- [37] C. Nishimura, H.J. Dyson, P.E. Wright, Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin, *J. Mol. Biol.* 355 (2006) 139–156.
- [38] C. Nishimura, H.J. Dyson, P.E. Wright, Enhanced picture of protein-folding intermediates using organic solvents in H/D exchange and quench-flow experiments, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 4765–4770.
- [39] C.M. Dobson, Protein folding and misfolding, *Nature* 426 (2003) 884–890.
- [40] F. Fogolari, L. Ragona, L. Zetta, S. Romagnoli, K.G. De Kruijff, H. Molinari, Monomeric bovine β -lactoglobulin adopts a β -barrel fold at pH 2, *FEBS Lett.* 436 (1998) 149–154.
- [41] S. Uhrínová, D. Uhrín, H. Denton, M. Smith, L. Sawyer, P.N. Barlow, Complete assignment of ^1H , ^{13}C and ^{15}N chemical shifts for bovine β -lactoglobulin:

- secondary structure and topology of the native state is retained in a partially unfolded form, *J. Biomol. NMR* 12 (1998) 89–107.
- [42] L.A. Joss, G.B. Ralston, β -Lactoglobulin B: a proposed standard for the study of reversible self-association reactions in the analytical ultracentrifuge? *Anal. Biochem.* 236 (1996) 20–26.
 - [43] K. Sakurai, M. Oobatake, Y. Goto, Salt-dependent monomer-dimer equilibrium of bovine β -lactoglobulin at pH 3, *Protein Sci.* 10 (2001) 2325–2335.
 - [44] R.K. Apenten, Protein stability function relations: β -lactoglobulin-A sulphhydryl group reactivity and its relationship to protein unfolding stability, *Int. J. Biol. Macromol.* 23 (1998) 19–25.
 - [45] K. Sakai, K. Sakurai, M. Sakai, M. Hoshino, Y. Goto, Conformation and stability of thiol-modified bovine β -lactoglobulin, *Protein Sci.* 9 (2000) 1719–1729.
 - [46] R. Yamasaki, M. Hoshino, T. Wazawa, Y. Ishii, T. Yanagida, Y. Kawata, T. Higurashi, K. Sakai, J. Nagai, Y. Goto, Single molecular observation of the interaction of GroEL with substrate proteins, *J. Mol. Biol.* 292 (1999) 965–972.
 - [47] T.V. Burova, Y. Choiset, V. Tran, T. Haertle, Role of free Cys121 in stabilization of bovine β -lactoglobulin B, *Protein Eng.* 11 (1998) 1065–1073.
 - [48] M. Narayan, L.J. Berliner, Mapping fatty acid binding to β -lactoglobulin: ligand binding is restricted by modification of Cys 121, *Protein Sci.* 7 (1998) 150–157.
 - [49] D.R. Flower, The lipocalin protein family: structure and function, *Biochem. J.* 318 (Pt 1) (1996) 1–14.
 - [50] D.A. Breustedt, I.P. Korndörfer, B. Redl, A. Skerra, The 1.8-Å crystal structure of human tear lipocalin reveals an extended branched cavity with capacity for multiple ligands, *J. Biol. Chem.* 280 (2005) 484–493.
 - [51] J.C. Sacchettini, J.L. Gordon, L.J. Banaszak, Crystal structure of rat intestinal fatty-acid-binding protein. Refinement and analysis of the *Escherichia coli*-derived protein with bound palmitate, *J. Mol. Biol.* 208 (1989) 327–339.
 - [52] Y. Urade, N. Fujimoto, O. Hayashi, Purification and characterization of rat brain prostaglandin D synthetase, *J. Biol. Chem.* 260 (1985) 12410–12415.
 - [53] M. Seppälä, R.N. Taylor, H. Koistinen, R. Koistinen, E. Milgrom, Glycodelin: a major lipocalin protein of the reproductive axis with diverse actions in cell recognition and differentiation, *Endocr. Rev.* 23 (2002) 401–430.
 - [54] N. Taulier, T.V. Chalikian, Characterization of pH-induced transitions of β -lactoglobulin: ultrasonic, densimetric, and spectroscopic studies, *J. Mol. Biol.* 314 (2001) 873–889.
 - [55] C. Tanford, L.G. Bunville, Y. Nozaki, The reversible transformation of β -lactoglobulin at pH 7.5, *J. Am. Chem. Soc.* 81 (1959) 4032–4036.
 - [56] L. Vijayalakshmi, R. Krishna, R. Sankaranarayanan, M. Vijayan, An asymmetric dimer of β -lactoglobulin in a low humidity crystal form-structural changes that accompany partial dehydration and protein action, *Proteins* 71 (2008) 241–249.
 - [57] K. Sakurai, Y. Goto, Dynamics and mechanism of the Tanford transition of bovine β -lactoglobulin studied using heteronuclear NMR spectroscopy, *J. Mol. Biol.* 356 (2006) 483–496.
 - [58] C. Tanford, Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Adv. Protein Chem.* 24 (1970) 1–95.
 - [59] D. Barrick, R.L. Baldwin, Three-state analysis of sperm whale apomyoglobin folding, *Biochemistry* 32 (1993) 3790–3796.
 - [60] F. Fogolari, E. Moroni, M. Wojciechowski, M. Baginski, L. Ragona, H. Molinari, MM/PBSA analysis of molecular dynamics simulations of bovine β -lactoglobulin: free energy gradients in conformational transitions? *Proteins* 59 (2005) 91–103.
 - [61] L. Ragona, F. Fogolari, L. Zetta, D.M. Peréz, P. Puyol, K. De Kruijff, F. Löhr, H. Rüterjans, H. Molinari, Bovine β -lactoglobulin: interaction studies with palmitic acid, *Protein Sci.* 9 (2000) 1347–1356.
 - [62] K. Sakurai, Y. Goto, Principal component analysis of the pH-dependent conformational transitions of bovine β -lactoglobulin monitored by heteronuclear NMR, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 15346–15351.
 - [63] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules, *J. Am. Chem. Soc.* 104 (1982) 4546–4559.
 - [64] P. Dosset, J.C. Hus, M. Blackledge, D. Marion, Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data, *J. Biomol. NMR* 16 (2000) 23–28.
 - [65] M.E. Hodsdon, D.P. Cistola, Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}N NMR relaxation and ^1H exchange, *Biochemistry* 36 (1997) 2278–2290.
 - [66] S. Shimamoto, T. Yoshida, T. Inui, K. Gohda, Y. Kobayashi, K. Fujimori, T. Tsurumura, K. Aritake, Y. Urade, T. Ohkubo, NMR solution structure of lipocalin-type prostaglandin D synthase: evidence for partial overlapping of catalytic pocket and retinoic acid-binding pocket within the central cavity, *J. Biol. Chem.* 282 (2007) 31373–31379.
 - [67] S. Kushibiki, K. Hodate, J. Kurisaki, H. Shingu, Y. Ueda, A. Watanabe, M. Shinoda, Effect of β -lactoglobulin on plasma retinol and triglyceride concentrations, and fatty acid composition in calves, *J. Dairy Res.* 68 (2001) 579–586.
 - [68] M.C. Yang, N.C. Chen, C.J. Chen, C.Y. Wu, S.J. Mao, Evidence for β -lactoglobulin involvement in vitamin D transport in vivo-role of the γ -turn (Leu-Pro-Met) of β -lactoglobulin in vitamin D binding, *FEBS J.* 276 (2009) 2251–2265.
 - [69] T.L. McMeekin, B.D. Polis, E.S. Della-Monica, J.H. Custer, A crystalline compound of β -lactoglobulin with dodecyl sulphate, *J. Am. Chem. Soc.* 71 (1949) 3606–3609.
 - [70] G. Kontopidis, C. Holt, L. Sawyer, The ligand-binding site of bovine β -lactoglobulin: evidence for a function? *J. Mol. Biol.* 318 (2002) 1043–1055.
 - [71] S.Y. Wu, M.D. Pérez, P. Puyol, L. Sawyer, β -Lactoglobulin binds palmitate within its central cavity, *J. Biol. Chem.* 274 (1999) 170–174.
 - [72] M.C. Yang, H.H. Guan, M.Y. Liu, Y.H. Lin, J.M. Yang, W.L. Chen, C.J. Chen, S.J. Mao, Crystal structure of a secondary vitamin D₃ binding site of milk β -lactoglobulin, *Proteins* 71 (2008) 1197–1210.
 - [73] T. Konuma, K. Sakurai, Y. Goto, Promiscuous binding of ligands by β -lactoglobulin involves hydrophobic interactions and plasticity, *J. Mol. Biol.* 368 (2007) 209–218.
 - [74] S.Y. Stevens, S. Sanker, C. Kent, E.R. Zuiderweg, Delineation of the allosteric mechanism of a cytidyltransferase exhibiting negative cooperativity, *Nat. Struct. Biol.* 8 (2001) 947–952.
 - [75] E.R. Zuiderweg, Mapping protein-protein interactions in solution by NMR spectroscopy, *Biochemistry* 41 (2002) 1–7.
 - [76] C.S. Goh, D. Milburn, M. Gerstein, Conformational changes associated with protein-protein interactions, *Curr. Opin. Struct. Biol.* 14 (2004) 104–109.
 - [77] L. Zidek, M.V. Novotny, M.J. Stone, Increased protein backbone conformational entropy upon hydrophobic ligand binding, *Nat. Struct. Biol.* 6 (1999) 1118–1121.
 - [78] G.B. Jameson, J.J. Adams, L.K. Creamer, Flexibility, functionality and hydrophobicity of bovine β -lactoglobulin, *Int. Dairy J.* 12 (2002) 319–329.
 - [79] K. Braig, Z. Otwinowski, R. Hegde, D.C. Boisvert, A. Joachimiak, A.L. Horwich, P.B. Sigler, The crystal structure of the bacterial chaperonin GroEL at 2.8 Å, *Nature* 371 (1994) 578–586.
 - [80] J. Chatellier, A.M. Buckle, A.R. Fersht, GroEL recognises sequential and non-sequential linear structural motifs compatible with extended β -strands and α -helices, *J. Mol. Biol.* 292 (1999) 163–172.
 - [81] P. Gómez-Puertas, J. Martín-Benito, J.L. Carrascosa, K.R. Willison, J.M. Valpuesta, The substrate recognition mechanisms in chaperonins, *J. Mol. Recognit.* 17 (2004) 85–94.
 - [82] P.E. Wright, H.J. Dyson, Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm, *J. Mol. Biol.* 293 (1999) 321–331.
 - [83] A.L. Fink, Natively unfolded proteins, *Curr. Opin. Struct. Biol.* 15 (2005) 35–41.
 - [84] P. Tompa, C. Szász, L. Buday, Structural disorder throws new light on moonlighting, *Trends Biochem. Sci.* 30 (2005) 484–489.
 - [85] D. Hamada, S. Segawa, Y. Goto, Non-native α -helical intermediate in the refolding of β -lactoglobulin, a predominantly β -sheet protein, *Nat. Struct. Biol.* 3 (1996) 868–873.
 - [86] V. Forge, M. Hoshino, K. Kuwata, M. Arai, K. Kuwajima, C.A. Batt, Y. Goto, Is folding of β -lactoglobulin non-hierarchical? Intermediate with native-like β -sheet and non-native α -helix, *J. Mol. Biol.* 296 (2000) 1039–1051.
 - [87] F.E. Cohen, Protein misfolding and prion diseases, *J. Mol. Biol.* 293 (1999) 313–320.
 - [88] V.N. Uversky, A.L. Fink, Conformational constraints for amyloid fibrillation: the importance of being unfolded, *Biochim. Biophys. Acta* 1698 (2004) 131–153.
 - [89] D. Hamada, T. Tanaka, G.G. Tartaglia, A. Pawar, M. Vendruscolo, M. Kawamura, A. Tamura, N. Tanaka, C.M. Dobson, Competition between folding, native-state dimerisation and amyloid aggregation in β -lactoglobulin, *J. Mol. Biol.* 386 (2009) 878–890.
 - [90] D. Hamada, Y. Kuroda, T. Tanaka, Y. Goto, High helical propensity of the peptide fragments derived from β -lactoglobulin, a predominantly β -sheet protein, *J. Mol. Biol.* 254 (1995) 737–746.
 - [91] Y. Kuroda, D. Hamada, T. Tanaka, Y. Goto, High helicity of peptide fragments corresponding to β -strand regions of β -lactoglobulin observed by 2D-NMR spectroscopy, *Fold. Des.* 1 (1996) 255–263.
 - [92] D. Hamada, Y. Goto, The equilibrium intermediate of β -lactoglobulin with non-native α -helical structure, *J. Mol. Biol.* 269 (1997) 479–487.
 - [93] H. Roder, G.A. Elöve, S.W. Englander, Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR, *Nature* 335 (1988) 700–704.
 - [94] L. Ragona, F. Fogolari, S. Romagnoli, L. Zetta, J.L. Maubois, H. Molinari, Unfolding and refolding of bovine β -lactoglobulin monitored by hydrogen exchange measurements, *J. Mol. Biol.* 293 (1999) 953–969.
 - [95] K. Fujiwara, M. Arai, A. Shimizu, K. Ikeguchi, K. Kuwajima, S. Sugai, Folding-unfolding equilibrium and kinetics of equine β -lactoglobulin: equivalence between the equilibrium molten globule state and a burst-phase folding intermediate, *Biochemistry* 38 (1999) 4455–4463.
 - [96] L. D'Alfonso, M. Collini, L. Ragona, R. Ugolini, G. Baldini, H. Molinari, Porcine β -lactoglobulin chemical unfolding: identification of a non-native α -helical intermediate, *Proteins* 58 (2005) 70–79.
 - [97] M. Ikeguchi, S. Kato, A. Shimizu, S. Sugai, Molten globule state of equine β -lactoglobulin, *Proteins* 27 (1997) 567–575.
 - [98] T. Kobayashi, M. Ikeguchi, S. Sugai, Molten globule structure of equine β -lactoglobulin probed by hydrogen exchange, *J. Mol. Biol.* 299 (2000) 757–770.
 - [99] L.H. Greene, R. Wijesinha-Bettoni, C. Redfield, Characterization of the molten globule of human serum retinol-binding protein using NMR spectroscopy, *Biochemistry* 45 (2006) 9475–9484.
 - [100] M. Yagi, A. Kameda, K. Sakurai, C. Nishimura, Y. Goto, Disulfide-linked bovine β -lactoglobulin dimers fold slowly, navigating a glassy folding landscape, *Biochemistry* 47 (2008) 5996–6006.
 - [101] W. Hoyer, K. Ramm, A. Plückthun, A kinetic trap is an intrinsic feature in the folding pathway of single-chain Fv fragments, *Biophys. Chem.* 96 (2002) 273–284.
 - [102] M. Tsunenaga, Y. Goto, Y. Kawata, K. Hamaguchi, Unfolding and refolding of a type κ immunoglobulin light chain and its variable and constant fragments, *Biochemistry* 26 (1987) 6044–6051.
 - [103] W.A. Fenton, A.L. Horwich, GroEL-mediated protein folding, *Protein Sci.* 6 (1997) 743–760.
 - [104] R. Horst, E.B. Bertelsen, J. Fiaux, G. Wider, A.L. Horwich, K. Wüthrich, Direct NMR observation of a substrate protein bound to the chaperonin GroEL, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 12748–12753.
 - [105] R. Horst, W.A. Fenton, S.W. Englander, K. Wüthrich, A.L. Horwich, Folding trajectories of human dihydrofolate reductase inside the GroEL GroES

- chaperonin cavity and free in solution, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 20788–20792.
- [106] C. Nishimura, H.J. Dyson, P.E. Wright, The apomyoglobin folding pathway revisited: structural heterogeneity in the kinetic burst phase intermediate, *J. Mol. Biol.* 322 (2002) 483–489.
- [107] C. Nishimura, P.E. Wright, H.J. Dyson, Role of the B helix in early folding events in apomyoglobin: evidence from site-directed mutagenesis for native-like long range interactions, *J. Mol. Biol.* 334 (2003) 293–307.
- [108] N. Go, The consistency principle in protein structure and pathways of folding, *Adv. Biophys.* 18 (1984) 149–164.
- [109] J.N. Onuchic, P.G. Wolynes, Theory of protein folding, *Curr. Opin. Struct. Biol.* 14 (2004) 70–75.
- [110] L. Sutton, J. Lätzer, J.A. Hegler, D.U. Ferreira, P.G. Wolynes, Consequences of localized frustration for the folding mechanism of the IM7 protein, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 19825–19830.
- [111] D.U. Ferreira, J.A. Hegler, E.A. Komives, P.G. Wolynes, Localizing frustration in native proteins and protein assemblies, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 19819–19824.
- [112] G. Deléage, B. Roux, An algorithm for protein secondary structure prediction based on class prediction, *Protein Eng.* 1 (1987) 289–294.
- [113] R.D. King, M.J. Sternberg, Identification and application of the concepts important for accurate and reliable protein secondary structure prediction, *Protein Sci.* 5 (1996) 2298–2310.
- [114] J. Garnier, D.J. Osguthorpe, B. Robson, Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins, *J. Mol. Biol.* 120 (1978) 97–120.
- [115] J.F. Gibrat, J. Garnier, B. Robson, Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs, *J. Mol. Biol.* 198 (1987) 425–443.
- [116] Y. Guermeur, *Combinaison de classifieurs statistiques, Application à la prédiction de structure secondaire des protéines.*, vol. Ph. D, Université Paris, France, 1997.
- [117] Y. Guermeur, C. Geourjon, P. Gallinari, G. Deléage, Improved performance in protein secondary structure prediction by inhomogeneous score combination, *Bioinformatics* 15 (1999) 413–421.
- [118] B. Rost, C. Sander, Prediction of protein secondary structure at better than 70% accuracy, *J. Mol. Biol.* 232 (1993) 584–599.
- [119] D. Frishman, P. Argos, Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence, *Protein Eng.* 9 (1996) 133–142.
- [120] C. Geourjon, G. Deléage, SOPM: a self-optimized method for protein secondary structure prediction, *Protein Eng.* 7 (1994) 157–164.
- [121] G. Deléage, C. Combet, C. Blanchet, C. Geourjon, ANTHEPROT: an integrated protein sequence analysis software with client/server capabilities, *Comput. Biol. Med.* 31 (2001) 259–267.
- [122] D.R. Flower, A.C. North, C.E. Sansom, The lipocalin protein family: structural and sequence overview, *Biochim. Biophys. Acta* 1482 (2000) 9–24.