doi:10.1016/j.jmb.2008.12.038 J. Mol. Biol. (2009) 386, 878–890

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Competition between Folding, Native-State Dimerisation and Amyloid Aggregation in β-Lactoglobulin

Daizo Hamada[1,2,3]*, Toshiki Tanaka[4], Gian Gaetano Tartaglia[5], Amol Pawar[5], Michele Vendruscolo[5], Mei Kawamura[6], Atsuo Tamura[6], Naoki Tanaka[7] and Christopher M. Dobson[5]*

[1]*Department of Developmental Infectious Diseases, Osaka Medical Center for Maternal and Child Health, 840 Murodo, Izumi, Osaka 594-1101, Japan*

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

[3]*Division of Structural Biology (G-COE), Department of Biochemistry and Molecular Biology, Graduate School of Medicine, Kobe University, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan*

[4]*Graduate School of Material Science, Nagoya Institute of Technology, Gokiso-cho, Nagoya 466-8555, Japan*

[5]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*

[6]*Department of Chemistry, Graduate School of Science, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan*

[7]*Department of Polymer Science and Engineering, Kyoto Institute of Technology, Matsugasaki, Sakyo, Kyoto 606-8585, Japan*

We show that a series of peptides corresponding to individual β-strands in native β-lactoglobulin readily form amyloid aggregates and that such aggregates are capable of seeding fibril formation by a full-length form of β-lactoglobulin in which the disulfide bonds are reduced. By contrast, preformed fibrils corresponding to only one of the β-strands that we considered, βA, were found to promote fibril formation by a full-length form of β-lactoglobulin in which the disulfide bonds are intact. These results indicate that regions of high intrinsic aggregation propensity do not give rise to aggregation unless at least partial unfolding takes place. Furthermore, we found that the high aggregation propensity of one of the edge strands, βI, promotes dimerisation of the native structure rather than misfolding and aggregation since the structure of βI is stabilised by the presence of a disulfide bond. These findings demonstrate that the interactions that promote folding and native-state oligomerisation can also result in high intrinsic amyloidogenicity. However, we show that the presence of the remainder of the sequence dramatically reduces the net overall aggregation propensity by negative design principles that we suggest are very common in biological systems as a result of evolutionary processes.

**Corresponding authors.* D. Hamada is to be contacted at Division of Structural Biology (G-COE), Department of Biochemistry and Molecular Biology, Graduate School of Medicine, Kobe University, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan; C. M. Dobson, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK. E-mail addresses: daizo@med.kobe-u.ac.jp; cmd44@cam.ac.uk.

Abbreviations used: RCM-β-lactoglobulin, reduced and carboxymethylated β-lactoglobulin; TEM, transmission electron microscopy.

Received 31 July 2008;
received in revised form
10 December 2008;
accepted 15 December 2008
Available online
24 December 2008

Edited by K. Kuwajima

Keywords: fragment seeding; transmission electron microscopy; thioflavin T; fluorescence; negative design principle

Introduction

Despite the natural tendency of proteins to fold into native structures unique to their amino acid sequences,[1,2] they are capable of aggregating both *in vitro* and *in vivo* into large assemblies with a variety of morphological features.[3–6] Of such aggregates, amyloid fibrils represent particularly well-ordered structures that can be micrometers in length while just a few nanometers in diameter and contain cross β-strands aligned perpendicularly to the fibril axis.[7–11] Such fibrils are the subject of intense interest at the present time since their deposition in tissue is associated with a range of increasingly common human diseases, including both systemic and organ-specific conditions, such as Alzheimer's disease and type II diabetes.[4,12,13] Remarkably, many peptides and proteins, and indeed homopolymers, such as polyalanine and polythreonine, without any known involvement in amyloid disease can also form amyloid fibrils *in vitro*,[14–21] suggesting that the ability to form amyloid fibrils is an intrinsic property of polypeptide chains.[22] Furthermore, such structures have the potential to generate new forms of nanoscale biomaterials based on the amyloid fibrillar architecture since they are characterised by well-ordered and stable structures that can be formed by self-association in aqueous solution.[23–26] Thus, elucidation of the mechanism by which polypeptide chains convert into amyloid fibrils is an important goal in medical, biological and material sciences.

Formation of amyloid fibrils appears to occur in a nucleation-dependent manner that is followed by a rapid extension process that gives rise to fibrillar architectures.[27] It has been observed by analysis of native-state structures that the existence of exposed β-strands at the edge of β-sheet structures tends to be avoided since they help in the formation of β-aggregates, such as amyloid fibrils.[28] Extensive surveys of structures in the Protein Data Bank suggest that a variety of mechanisms prevent the exposure of hydrophobic surfaces in such edge strands (e.g., by formation of intermolecular β-sheets or by placing inward-pointing charged side chains).[28]

One of the proteins exemplifying the former situation is bovine β-lactoglobulin, which has been shown to be capable of forming amyloid fibrils by prolonged incubation at pH 7.0 in the presence of 5.0 M urea.[18] However, the reaction is relatively slow and can take several weeks to reach completion. β-Lactoglobulin forms a homodimeric native state at neutral pH in the absence of denaturant, although it dissociates into monomeric but still natively folded species at pH values below 3.0.[29] In the native state, the protein has a predominantly β-sheet structure, consisting of a β-barrel of eight antiparallel β-strands (A–H) and an additional *edge* β-strand (I) that is part of the dimer interface located at the edge of the β-sheet of each individual molecule, although there is one major α-helix at the C-terminus (Fig. 1). The β-barrel structure has the shape of a flattened cone resembling a calyx involving the N- and C-terminal β-sheets; the N-terminal β-sheet consists of the βB, βC and βD strands, while the C-terminal one involves the βA, βE, βF, βG, βH and βI strands. The protein also contains two disulfide bonds (C66–C160 and C106–C119, linking strands βD–βI and βG–βH, respec-

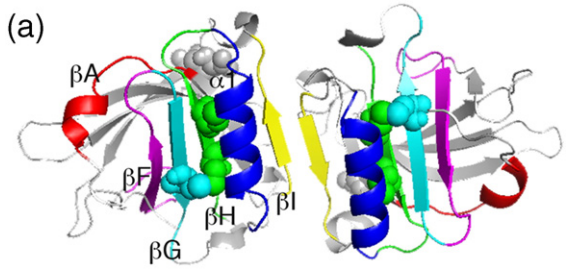

Fig. 1. Structure of bovine β-lactoglobulin. (a) Dimeric native-state structure from X-ray crystallography (Protein Data Bank ID 1BEB). The regions corresponding to the peptides studied in this work are represented by different colors (βA is shown in red; βF, magenta; βG, cyan; βH, green; βI, yellow; and α1, blue). (b) Amino acid sequences of the peptide fragments studied in this work. The aromatic residues shown in lower case were added to the original sequence in order to allow determination of the peptide concentration by means of UV absorption. The additional aromatic residues should not significantly influence the aggregation propensity of these peptides according to the Z_{agg} score from Zyggregator (Table 1). Residue numbers are relative to the sequence of the intact protein.

tively) and one free cysteine residue (C121). Despite the fact that the native structure of β-lactoglobulin contains a predominantly β-sheet structure, the sequence of β-lactoglobulin encodes a relatively high α-helical propensity; indeed, the major folding intermediate observed to be populated during the folding of the protein contains non-native α-helical motifs[30–37] located mainly in the vicinity of the βA-strand region in the native structure.[35]

In this work, we describe studies on the relative abilities of peptide fragments corresponding to the amino acid sequences of the β-strands and the α-helical regions of the hydrophobic core of native β-lactoglobulin to form amyloid fibrils and to seed the aggregation of full-length β-lactoglobulin. These results reveal new insights into the general principles underlying the competition between folding, native-state oligomerisation and multimolecular aggregation and provides specific information about the way these factors influence the behaviour of β-lactoglobulin itself.

Results and Discussion

Amyloid formation by peptide fragments of β-lactoglobulin

The peptide fragments that we have analysed here correspond to the βA, βF, βG, βH and βI strands and the C-terminal α-helical region (α1) of native β-lactoglobulin (Fig. 1); the amino acid sequences are shown in Fig. 1b. Except for βA, the peptides are all located at the C-terminal β-sheet or at the dimer interface of the native protein. These fragments were chosen because the corresponding regions of the sequence form the hydrophobic core of the native structure of the full-length protein.[36] Moreover, in the process of folding from the denatured state, the βF, βG and βH strands have been found to assume a native-like structure rapidly after the initiation of the reaction.[36] α1 (described as fragment 3 in Ref. 31 and as fragment 4 in Ref. 33) has previously been shown to possess a high intrinsic helical propensity. In this context, it is particularly interesting that, in the native state, βI is located at the edge of a C-terminal β-sheet that is also involved in the interface of the native homodimer of β-lactoglobulin.

The kinetics of fibril formation by the peptide fragments in solutions containing 10 mM sodium phosphate and 5.0 M urea at pH 7.0 were monitored by thioflavin T fluorescence (Fig. 2a and b). Since full-length β-lactoglobulin is known to form fibrillar aggregates under these experimental conditions,[18] the identification of peptides that are amyloidogenic enables "fragment seeding" studies to be performed. In these experiments, the degree to which the addition of fibrils preformed from the various peptide fragments is able to promote fibril formation by intact β-lactoglobulin can be examined under conditions in which fibrils have been shown to be formed by the full-length protein.

Of the peptide fragments examined here, βA, βG and βI show a clear increase in the intensity of

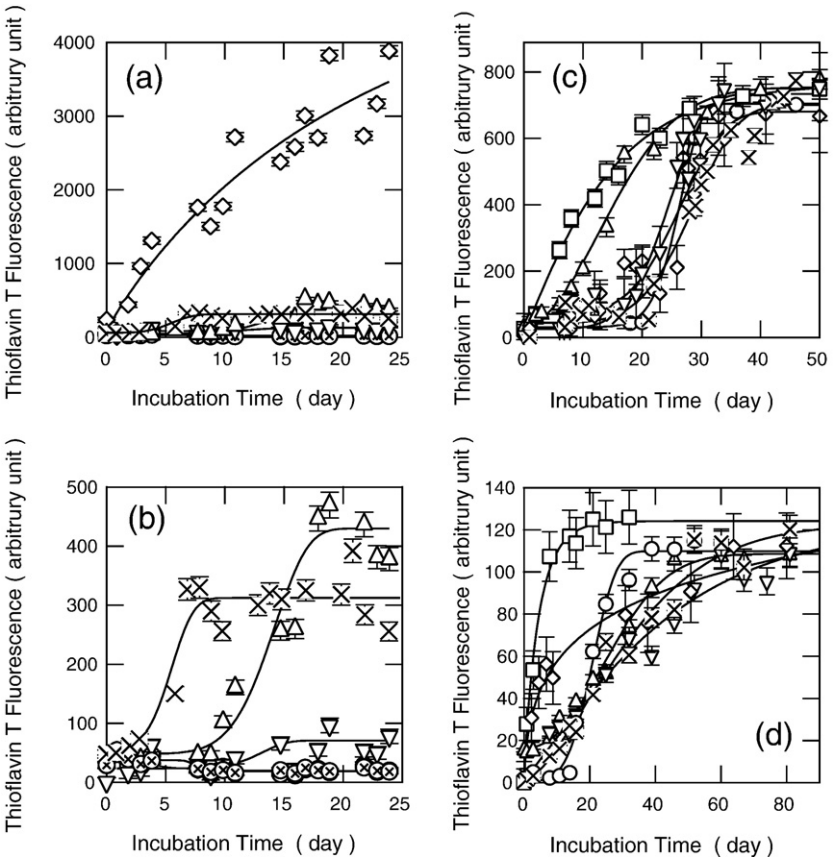

Fig. 2. Kinetics of amyloid fibril formation by β-lactoglobulin and its peptide fragments in the presence of 5.0 M urea at pH 7.0 and 37 °C followed by thioflavin T fluorescence. (a) Kinetic traces for spontaneous fibril formation by βA (triangle), βF (plus sign in circle), βG (diamond), βH (inverted triangle), βI ("X" sign) and α1 (multiplication sign in circle). (b) Same as (a), except that the ordinate is expanded to show more clearly the kinetics for βA, βF, βH, βI and α1. (c and d) Fragment seeding effects of the fibrils formed by the different peptide fragments on amyloid fibril formation by (c) intact β-lactoglobulin and (d) RCM-β-lactoglobulin. Spontaneous fibril formation is represented by circles. Fibril formation in the presence of preformed fibrils of the same peptide or protein (self-seeding) (whisker ends) and that in the presence of preformed fibrils of βA (triangle), βG (diamond), βH (inverted triangle) and βI ("X") (fragment seeding) are also shown. Error bars represent standard deviations of triplicate experiments.

thioflavin T fluorescence after prolonged incubation at pH 7.0 in the presence of 5.0 M urea (Fig. 2a and b). The kinetic traces of the increase in fluorescence in solutions containing βA and βI are sigmoidal curves, consistent with the idea that the formation of amyloid fibrils occurs in a nucleation-dependent manner. The kinetic profile for the solution containing βG, on the other hand, does not show the presence of a well-defined lag phase; the fluorescence intensity of thioflavin T was found to increase immediately after the preparation of samples. It should be noted, however, that βG is highly insoluble in water but soluble in the presence of denaturants. The lack of an observable lag phase is therefore likely to reflect the fact that the solutions contain preformed aggregates of βG prior to incubation and that the growth of fibrils does not require further nucleation events. The βG peptide, in addition, has a single free cysteine residue in the central region of its sequence that may permit intermolecular disulfide bonds to form under the conditions employed here, hence providing an additional mechanism for nucleation. The long-term incubation in solution containing urea may induce carbamylation, thus potentially affecting the aggregation propensity of the protein. However, a previous analysis of the molecular masses of intact β-lactoglobulin before and after amyloid formation suggested that no such modification is introduced to the protein during the formation of fibrils.[30] Thus, such modification, even if present, should not significantly affect the ability of this protein or probably the peptide fragments.

Transmission electron microscopy (TEM) images confirm the presence of fibrillar aggregates following incubation of solutions containing βA, βG or βI fragments (Fig. 3a). In addition, solutions of βH were unexpectedly found to contain well-defined fibrillar aggregates by TEM, although they do not show an increase in thioflavin T fluorescence. Interestingly, the fibrils from the βH peptide are highly intertwined but do not differ significantly in appearance from typical amyloid fibrils as observed for intact β-lactoglobulin (Fig. 3a). The inability of the fibrils formed by this peptide to increase thioflavin T fluorescence could be associated with this difference in morphology. More generally, it has been observed in other systems that the amyloid fibrils formed by short peptides tend to show lower thioflavin T fluorescence intensities, perhaps as a result of the rigidity

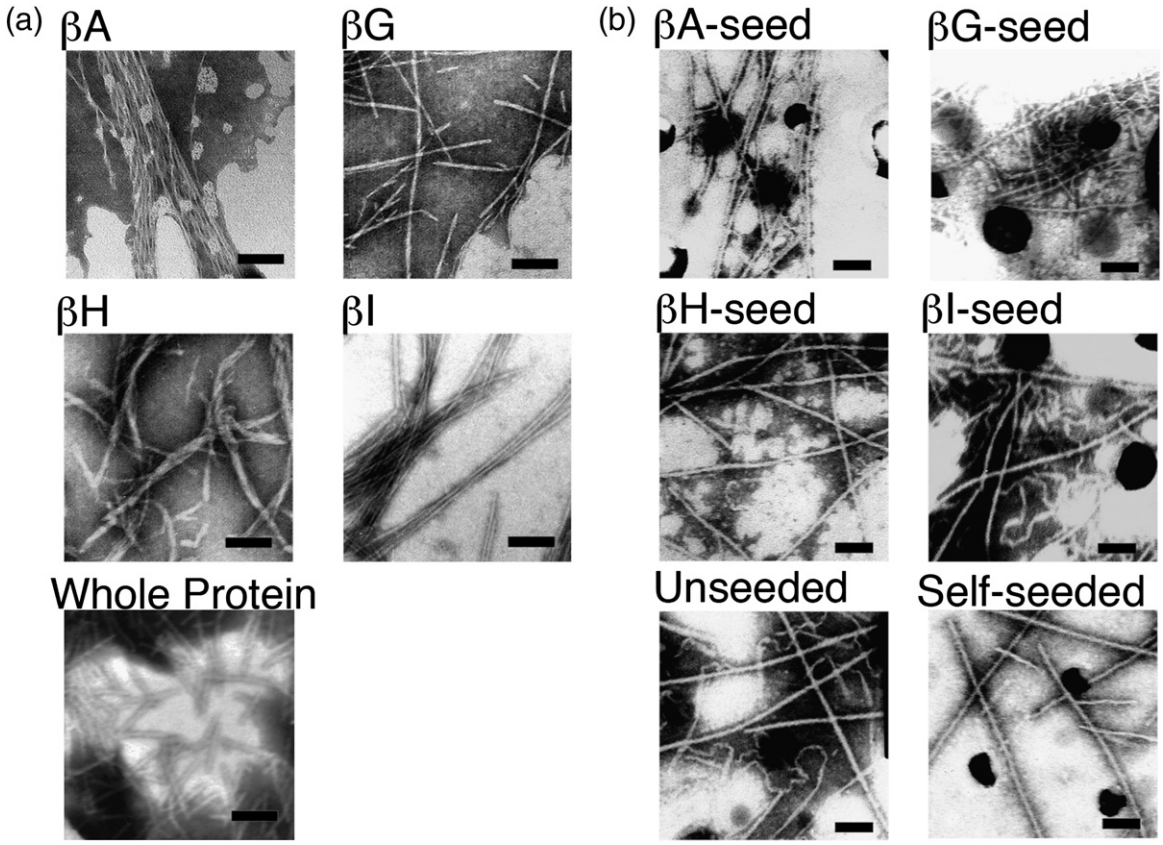

Fig. 3. TEM images of amyloid fibrils formed by β-lactoglobulin and its peptide fragments under different conditions. (a) Fibrils formed by βA, βG, βH, βI and intact β-lactoglobulin. The images were taken for the solution after 1 month of incubation at pH 7.0 and 37 °C in the presence of 5.0 M urea where the reaction curves in Fig. 2 obtained by thioflavin T fluorescence reach the plateau region. (b) Fibrils formed by RCM-β-lactoglobulin under unseeded or seeded conditions in the presence of preformed fibrils by RCM-β-lactoglobulin or βA, βG, βH and βI peptides. The scale bars correspond to 100 nm. The images were taken for the solution after 3 months of incubation at pH 7.0 and 37 °C in the presence of 5.0 M urea where the reaction curves in Fig. 2 obtained by thioflavin T fluorescence reach the plateau region.

of the cross-β strands in the fibril structures (Yuji Goto, personal communication).

Fragment seeding effects

Amyloid formation by large polypeptide chains appears to be promoted by specific regions of their amino acid sequences, sometimes known as "aggregation-prone" regions, which, in addition, are likely to be incorporated into the core of the fibrils.[37] A major aim of the present study was to understand the role played by such regions in the process of fibril formation by β-lactoglobulin. As part of this objective, we adopted the strategy of adding fibrils formed by the peptide fragments to solutions of the full-length protein to explore their efficacy in promoting the formation of fibrils by the latter. Such fragment seeding behaviour has been observed in studies on peptides and proteins, including insulin, hen lysozyme, islet amyloid polypeptide and β_2-microglobulin.[38–43]

Figure 2c indicates the degree to which the fibrillar aggregates formed by βA, βG, βH and βI fragments are capable of seeding the aggregation of the intact protein in 5.0 M urea at pH 7.0. Of the fibrils formed by the four peptides, only those formed by the βA fragment were observed to detectably accelerate the fibril formation by full-length β-lactoglobulin; the kinetics of fibril formation are essentially unchanged by the presence of preformed fibrils of the other fragments. This analysis suggests that the βA-strand region, but not the regions corresponding to the βG, βH and βI strands, is likely to be directly involved in promoting the formation of fibrils by full-length β-lactoglobulin.

However, these results do not rule out the possibility that other regions of the sequence (e.g., the βB, βC, βD and βE strands or loop regions) also contribute to the core region of the fibrillar structure. Indeed, the fact that several of the peptides studied here readily form amyloid fibrils supports the idea that the ability to form fibrillar architectures is a generic property of polypeptide chains.[13] For large polypeptide chains, in particular, different regions of the molecules may well assemble in a range of distinct ways to generate cross-β core regions and give rise to the well-established structural polymorphism or heterogeneity of these fibrils.[44–48] This observation is consistent with the fact that very different morphologies of fibrillar structures are observed for full-length β-lactoglobulin under different conditions.[18,49] It is also possible that the amyloid core is formed by various segments apart from one another. In this sense, further analysis of fragment seeding using several peptides at the same time to make seeds may clarify the regions required for stabilisation of the amyloid core by β-lactoglobulin.

Implications for amyloid formation by β-lactoglobulin

We used our previously published method[50] to calculate the intrinsic aggregation propensity of β-lactoglobulin as a function of its sequence (see Materials and Methods). The results show that the propensities of the regions corresponding to the βA, βG, βH and βI strands to aggregate are substantially higher than those of the other regions of the sequence (Fig. 4; Table 1). Consistent with these predictions, amyloid fibrils are formed very rapidly by the βA, βG, βH and βI peptides in solution, as shown in Fig. 2a and b. Addition of preformed fibrils of the βG, βH and βI peptides however does not detectably promote fibril formation by full-length β-lactoglobulin under the conditions used here, suggesting that the intrinsic propensity for aggregation of segments of a protein sequence is not the only factor that determines the aggregation process of the intact molecule. Indeed, only preformed fibrils of βA, which in fact has a slightly lower intrinsic propensity for aggregation than do βG, βH and βI (Fig. 4; Table 1), were found to accelerate the formation of fibrils by full-length β-lactoglobulin. It is well established that, for intact proteins, only part of the polypeptide chain needs to be involved in the formation of the cross-β core of amyloid fibrils.[51–54] The present results imply that the region corresponding to the βA strand could be specifically involved in stabilising the cross-β structure of full-length β-lactoglobulin amyloid fibrils.

In order to understand the differences between the fragment seeding effects of fibrils corresponding to the βA, βG, βH and βI peptides, despite the fact that they all have high intrinsic propensities for aggregation, we calculated two additional aggregation propensity profiles by taking into account the protection of the aggregation-prone regions that will be provided by the formation of the native structure of the protein. In this approach, the local stability of a region in a folded protein is predicted from knowledge of its structure.[55] When the stability of native-like structures in a localised region is very high, the effective propensity for aggregation around such region is strongly reduced as the region itself will tend to remain in a native-like conformation even under at least moderately destabilising conditions or in the presence of local conformational fluctuations. The presence of such localised native-like structure will reduce the accessibility for aggregation of the protein through the association of this region. We therefore calculated the aggregation profile taking into account the structure present in the *monomeric* form of β-lactoglobulin (Fig. 4). In this new profile, except for the terminal regions of the sequence, the only region with an aggregation propensity significantly above the average corresponds to βI in the native fold. The fact that dimerisation occurs through this β-strand shows the value of the method for predicting regions likely to form intermolecular β-sheets in functional and pathogenic assemblies. This idea is supported by the fact that equine β-lactoglobulin, for which the propensity to aggregate in the βI strand is much lower (Fig. 4b), forms a monomeric native state. A comparison of the amino acid sequences of bovine and equine β-lactoglobulin reveals a clear difference in the sequence in the

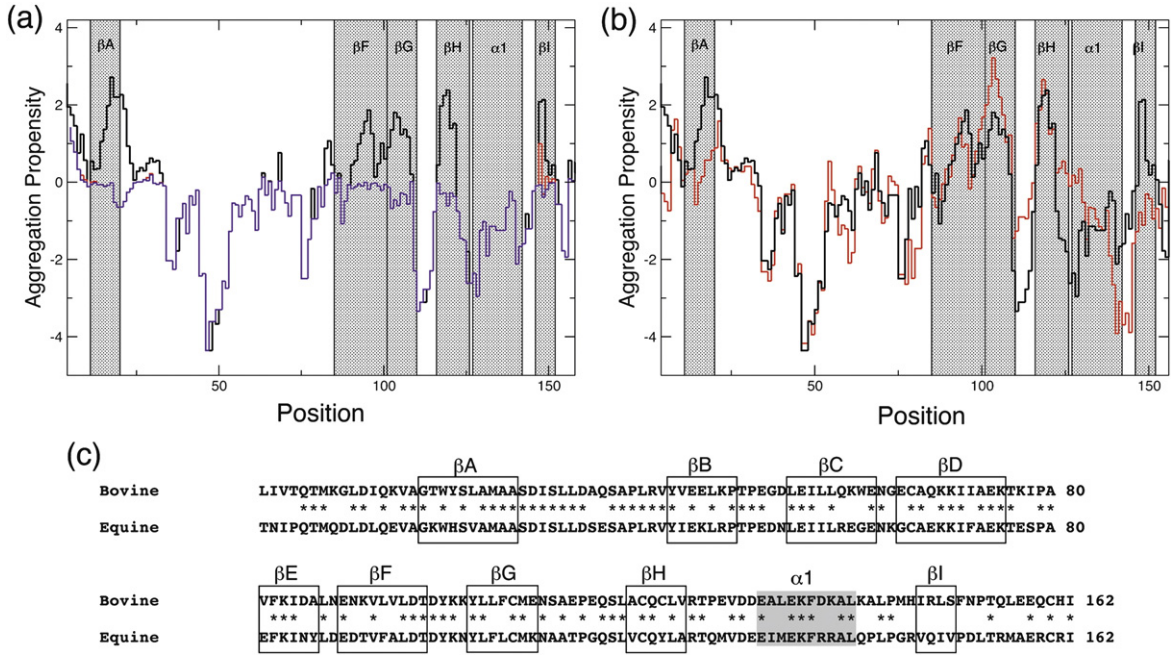

Fig. 4. Structural and physicochemical properties of β-lactoglobulin predicted from the amino acid sequence. (a) Predicted aggregation propensity profiles of bovine β-lactoglobulin at pH 7.0. The intrinsic aggregation profile (black) is compared with the profiles obtained by incorporating conformational factors in the monomeric (red) and dimeric (violet) forms. Gray shaded areas indicate the locations of the peptides analysed in this work. (b) Predicted intrinsic aggregation profile of equine β-lactoglobulin (black) at pH 7.0 in comparison with that of bovine β-lactoglobulin (red). (c) Sequence alignment of bovine and equine β-lactoglobulin showing a less conserved βI region.

region of the βI strand in contrast to the highly conserved sequences in the other regions (Fig. 4c).

We calculated the second profile by considering the structure of the *dimeric* native state of β-lactoglobulin (Fig. 4a). This profile is essentially identical with the previous one, except in the region of βI, where the aggregation propensity is much reduced as a result of the burial of this strand in the dimer interface. In this profile, there is no region of significant aggregation propensity, in agreement with the observation that the native dimer of β-lactoglobulin is highly soluble and must be destabilised in order to form amyloid aggregates efficiently and indicating that the native fold of the protein has a strongly protective function against aggregation. These results indicate that fibrils formed by the βA, βG, βH and βI strands are likely to be able to promote amyloid formation by full-length β-lactoglobulin if the structure of the latter becomes substantially unfolded.

Table 1. Aggregation propensities of peptides derived from β-lactoglobulin

Peptide fragment[a]	Z_{agg}[b]	Amyloidogenicity[c]
Full-length β-lactoglobulin	0.87	
βA	1.36	High
βF	0.64	Low
βG	1.48	High
βH	1.79 (1.73)[d]	High
βI	2.28 (1.66)	High
α1	0.00 (0.41)	Low

[a] Sequence position of fragments in intact β-lactoglobulin (Fig. 1).
[b] Intrinsic aggregation propensity corresponding to the amino acid sequence of the peptide at pH 7.0 calculated using the Zyggregator method (see Materials and Methods for details).
[c] Qualitative experimental rates of growth of fibrils at pH 7.0 in the presence of 5.0 M urea (Fig. 2a).
[d] Values calculated without considering the effect of artificially added aromatic residues (Fig. 1).

Role of the disulfide bonds in reducing amyloidogenicity

Further insight into the different roles played by the βA, βG, βH and βI strands in the process of formation of amyloid fibrils by β-lactoglobulin is provided by studies on the folding mechanism of this protein using amide hydrogen exchange and nuclear magnetic resonance spectroscopy.[35,36,56,57] These studies indicate that the βG and βH strands strongly favour the formation of a native-like β-sheet structure, both in intermediates formed early in the folding process[35,36] and in those formed in cold-denatured states.[56] This finding can be rationalised by noting the presence of a disulfide bond between the βG and βH strands (C106–C119) that restricts conformational freedom by preferentially stabilising native-like structures around the βG and βH strands. We have previously shown that the protein populates significantly not only the monomeric and the dimeric native states but also a partially unfolded state at pH 7.0 and 37 °C in the

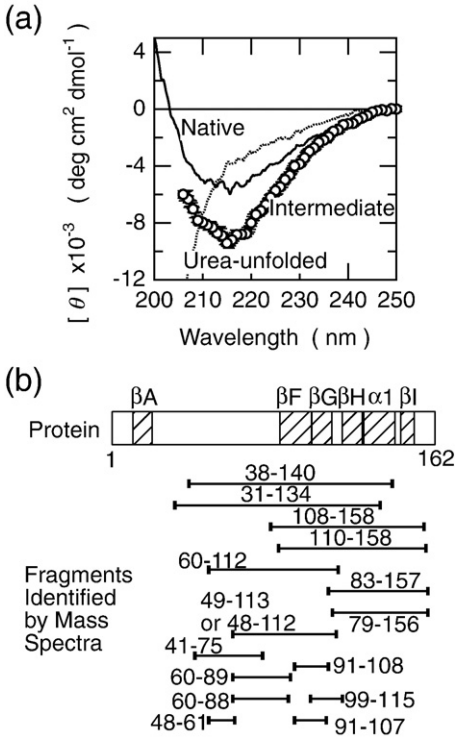

Fig. 5. Conformational properties of partially folded intermediate state of β-lactoglobulin accumulated at pH 7.0 and 37 °C in the presence of urea. (a) CD spectra of the native (continuous line), intermediate (circles) and fully unfolded (dotted line) states. Error bars are also shown for the spectrum of the intermediate state. (b) Sequence position of fragments resistant against proteinase K cleavage at pH 7.0 and 37 °C in the presence of 5.0 M urea.

presence of 5.0 M urea.[18] A similar result was recently obtained using an analysis of luminescence lifetime distributions[58] that showed a significant population (~100%) of β-lactoglobulin in the partially folded states in the presence of 5.0 M urea at pH 7.0 and 37 °C. Therefore, under such conditions, the partially folded conformation that is most readily accessible to the protein is expected to have stable native-like regions of β-sheet structure in the vicinity of the βG and βH strands. To investigate these possibilities, we characterised the structural properties of an intermediate state accumulated at pH 7.0 and 37 °C during an equilibrium unfolding induced by urea using circular dichroism (CD) according to our previously established method and limited proteolysis by proteinase K, a non-specific protease.

We previously demonstrated that a partially folded intermediate state with non-native α-helical structures accumulated during the guanidinium hydrochloride unfolding at pH 2.0 and 4 °C by global-fitting analysis of far-UV CD spectra.[30] A similar strategy has been applied here for urea unfolding of β-lactoglobulin at pH 7.0 and 37°C (Fig. 5a; see also Supplementary Data). Unlike the spectrum obtained at low pH and low temperature,[30] the CD spectrum of the intermediate state at pH 7.0 and 37 °C shown in Fig. 5a does not show a significant increase of absolute ellipticity at 222 nm, which is diagnostic of the accumulation of α-helical structure. The spectrum was further analysed by CDpro,[59] a software package that includes three independent programs (SELCON3, CDSSTR and CONTIN) to calculate the secondary structure content. The results are summarised in Table 2. According to this analysis, the secondary structure content of the intermediate state is highly consistent with that of the native state.

Limited proteolysis of β-lactoglobulin by proteinase K was performed at pH 7.0 and 37 °C in the presence of 5.0 M urea, and the resulting proteolysis-resistant fragments were examined by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (Fig. 5b). The observed masses of these fragments were compared with the calculated masses of peptide fragments of β-lactoglobulin, thus enabling the identification of the regions of fragments resistant against proteinase K cleavage. This analysis indicated that most fragments identified here are located downstream of the βA-strand region. According to Giurleo *et al.*[58] almost 100% of β-lactoglobulin molecules assume a partially folded structure at pH 7.0 and 37 °C in the presence of 5.0 M urea.

These results therefore suggest that the conformational state accumulated at pH 7.0 and 37 °C in the presence of 5.0 M urea is a partially folded intermediate in which the N-terminal βA-strand region in the native state is relatively flexible and accessible to the proteolytic cleavage by proteinase K, while the downstream of the protein assumes relatively rigid native-like structures resistant against proteinase K cleavage.

The native-like secondary structures in the partially folded state of β-lactoglobulin identified here should act to inhibit at least to some extent the formation of aggregates, despite the fact that isolated peptides corresponding to these regions are readily able to form amyloid fibrils. In addition, the C-terminal region close to the βI strand, which is involved in the dimer interface, contains another disulfide bond between residue C66 and residue C160. This bond is likely to generate a conformational restraint that will also inhibit the formation of

Table 2. Secondary structure contents of various conformational states of β-lactoglobulin accumulated at 37 °C and pH 7.0

State	α-Helix (%)	β-Sheet (%)	Others (%)
Native	13±5 (6.8)	37±2 (51.2)	50±3 (42.0)
Intermediate	11±4	31±6	58±3
Intermediate at 4 °C[a]	28±10	23±14	49±4
Fully unfolded	9±7	32±2	59±6

The reported data are the average of the values estimated from the SELCON3, CONTIN and CDSSTR software in CDpro package. Values in parentheses were calculated from the crystal structure of β-lactoglobulin (Protein Data Bank ID 1BEB).

[a] Spectral data were taken from Ref. 30.

misfolded aggregation of β-lactoglobulin by stabilising the dimeric native state through intermolecular association involving the βI-strand region. By contrast, the βA-strand region in the unfolded, partially folded and even native states should be relatively accessible, as shown by its limited protection against amide hydrogen/deuterium exchange with the solvent[35,36,57] for various conformational states and our limited proteolysis (Fig. 5b). Importantly, the aggregation propensity at the βA-strand region has a relatively high aggregation propensity (Table 1), possibly due to the hydrophobic nature of its amino acid sequence (Fig. 1). Thus, the hydrophobic character, together with the flexible nature of the βA-strand region, may strongly promote the formation of amyloid fibrils by β-lactoglobulin through intermolecular interactions that are likely to be readily formed around this region. These results suggest that the βG, βH and βI strands are hindered from becoming fully unfolded by the presence of disulfide bridges, while the βA strand does not have such structural constraints and so may more readily become accessible and promote amyloid formation.

The conclusions discussed above suggest that addition of fibrils formed from the various peptides into solutions of full-length β-lactoglobulin in which the disulfide bonds are reduced and the protein is destabilised could well accelerate the rate of fibril formation by the protein. To test this idea, we prepared fully disulfide-reduced and carboxymethylated β-lactoglobulin (RCM-β-lactoglobulin) and monitored the kinetics of amyloid fibril formation using thioflavin T fluorescence (Fig. 2d). As with the disulfide-containing intact protein, RCM-β-lactoglobulin shows a sigmoidal kinetic behaviour under unseeded conditions, although the lag time could be abolished by adding the preformed fibrils corresponding to the full-length protein. Importantly, the addition of preformed fibrils formed by any of the peptides studied here could also completely abolish the lag phase observed in the absence of added fibrils.

We then used TEM imaging to examine the morphology of the fibril structures formed by RCM-β-lactoglobulin when seeded with fibrils from the different peptides. While intact β-lactoglobulin shows uniformity in fibril morphology and unbranched straight fibrils with a diameter of 16.9±2.9 nm were observed by TEM (Fig. 3a), two types of fibrils can be seen for RCM-β-lactoglobulin under unseeded conditions (Fig. 3b): One type is very straight with a diameter of 14.7±1.8 Å, which is similar to the fibril diameter formed by intact protein, and the other has a curly appearance with a diameter of 10.2±1.9 Å. Interestingly, only straight fibrils with a diameter of 14.6±3 Å could be observed under self-seeding conditions. However, seeding with preformed fibrils of the β-lactoglobulin peptides tended to generate fibrils thinner than those formed under self-seeded or unseeded conditions. In the presence of βA, βG and βH seeds, straight fibrils with diameters of 11.6±2.1, 8.3±1.8 and 11.9±1.8 Å were formed, respectively, while seeding with βI fibrils resulted in both straight and curly fibrils, with diameters of 10.2±1.9 and 12.6±2.6 Å, respectively.

These changes in the fibril structures formed under different seeding conditions suggest that different aggregation-prone regions can act as nucleation sites (Fig. 2d). They also offer an explanation for the observed polymorphism in fibril structures under unseeded conditions as different regions of RCM-β-lactoglobulin could initiate the nucleation process and hence affect the fibril morphology. These results therefore strongly support the idea that fibril formation by intact β-lactoglobulin potentially occurs through regions of the sequence that have high aggregation propensities, such as βI, and can be efficiently inhibited by formation of the monomeric and dimeric native states and by conformational restrictions stabilised by the presence of disulfide bonds.

A unified view of the behaviour of the folding and association of β-lactoglobulin

The study presented in this article illustrates the power of combining experimental data and theoretical predictions to probe specific aspects of the principles of protein misfolding and aggregation. We summarise the findings reported in this work concerning β-lactoglobulin in the schematic diagram presented in Fig. 6. Initially, the protein at equilibrium in 5.0 M urea at pH 7.0 and 37 °C has significant populations of native, partially folded and unfolded states. In the partially folded state, however, much of the polypeptide chain, with the exception of the βA strand, assumes a native-like topology that decreases very considerably the aggregation propensity of even the most highly aggregation-prone regions. The βA-strand region within the partially folded state appears to be exposed to the solvent and to assume a native-like extended structure as suggested by the limited proteolysis and CD analysis demonstrated here (Fig. 5) or, with a lower probability, a non-native α-helical structure.[30–33,35–37] The association of the extended βA-strand region with other molecules can thus initiate aggregation into ordered fibrils that is followed by maturation processes, including the association of filaments and the elongation of fibrillar structures.[5,60]

The propensity to form a non-native α-helical structure in the vicinity of the βA-strand region in the intermediate state of β-lactoglobulin particularly at low pH and low temperature[30] could be an important factor in reducing its propensity to form amyloid fibrils because of the involvement of the main-chain amide and carboxyl groups in intermolecular hydrogen bonds. The intermediate state in the folding and unfolding of the protein can then directly or indirectly become a precursor for amyloid fibril formation, as has been found to be the case in lysozyme, for example.[61] When the folding intermediate state of the protein is a direct precursor of amyloid fibrils, the aggregation process will involve a large conformational change in the

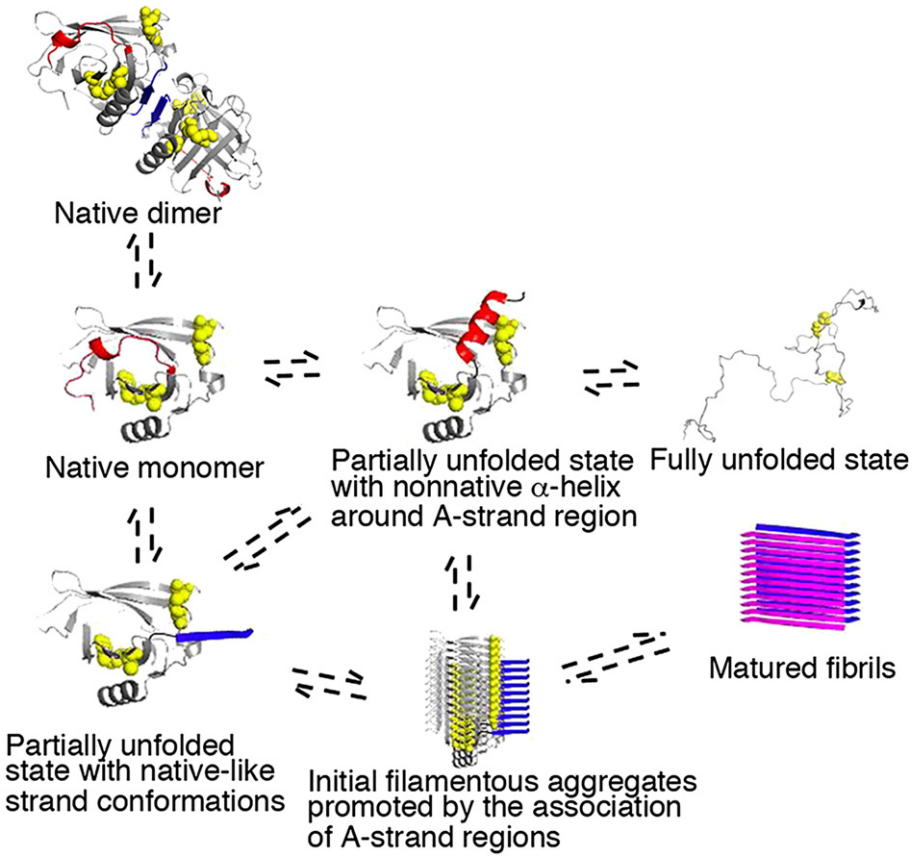

Fig. 6. Schematic view of a unified mechanism for amyloid formation by β-lactoglobulin.

region of the sequence corresponding to the βA strand from its non-native α-helical state to a native-like β-strand structure prior to the formation of the cross-β core of the fibril (Fig. 6). In the absence of disulfide bonds in intact β-lactoglobulin, amyloid aggregation can be initiated from various regions of the sequence due to the absence of structural constraints that stabilise the native structure. These results indicate that great care needs to be taken in using peptide fragments to probe the mechanism of aggregation of an intact protein or indeed of longer fragments of the protein.

The results obtained here suggest that the determinants for folding, native-state dimerisation and aggregation differ only in rather subtle ways, at least in the case of β-lactoglobulin. Regions with high intrinsic amyloidogenicity are also important for folding (the βG and βH strands) or for homodimerisation (the βI strand) but generate a risk of conversion into amyloid aggregates. Our results indicate that the presence of disulfide bonds represents a design principle that is additional to those already identified as promoting folding rather than aggregation of these regions that form β-sheets in the native state.[28] Our findings also indicate that the βA strand located at the N-terminal region of the structure is likely to be a key component of the amyloid core, because of its high intrinsic amyloidogenicity and because it has sufficient local flexibility to become exposed to the solvent and hence to be incorporated efficiently into the amyloid fibril structure. It is clear, however, that the amino acid sequence of β-lactoglobulin has evolved to inhibit efficiently the intrinsic tendency of a protein that forms extensive β-sheet structures in its native state to form similar structures in non-native and potentially pathogenic aggregates.

Materials and Methods

Materials

The A variant of bovine β-lactoglobulin, which differs from the B variant at positions 64 and 118, where an aspartate and a valine in the A variant are substituted by a glycine and an alanine in the B variant, was purchased from Sigma (St. Louis, MO). RCM-β-lactoglobulin was prepared as described in Ref. 32. Other chemicals were purchased either from Nacalai Tesque (Kyoto, Japan) or from Wako Pure Chemical Industries (Osaka, Japan). The polyvinyl formal–carbon-coated 200-mesh copper grids (200 Å) were obtained from Oken Shoji (Tokyo, Japan).

Peptide synthesis

Peptides were synthesised on an Applied Biosystems Model 433A automated synthesiser, using Rink amide resin, based on the standard Fastmoc 0.1 mmol protocol. The resulting crude peptides were purified by reversed-

phase HPLC, and the degree of purity was checked by analytical reversed-phase HPLC and mass spectrometry. Since all the peptides contain tryptophan or tyrosine residues, the peptide concentrations were determined by UV absorption spectroscopy using the extinction coefficients calculated according to Gill and von Hippel.[62]

Kinetics of fibril formation by peptides using thioflavin T fluorescence

Thioflavin T fluorescence spectra were obtained using a JASCO fluorimeter (Model FP-6500). Peptide solutions (1 mg mL^{-1}) in 10 mM sodium phosphate, pH 7.0, and 5.0 M urea were incubated at 37 °C. After given times of incubation, 5-μL aliquots of the peptide solutions were mixed with 400 μL of 5 μM thioflavin T dissolved in 10 mM sodium phosphate at pH 7.0, and the fluorescence at 460–660 nm was measured with an excitation wavelength of 450 nm. For seeding experiments using fragments, preformed fibrils of the various peptides (initial concentration of 1 mg mL^{-1}) were broken into smaller segments by sonication, and 10 μL of the resulting solutions containing fragmented fibrils was added to solutions of β-lactoglobulin at a concentration of 1 mg mL^{-1} and dissolved in 10 mM sodium phosphate at pH 7.0 and 5.0 M urea. The samples were incubated at 37 °C, and the presence of fibrils was monitored by thioflavin T fluorescence as described above.

Fibril imaging by TEM

TEM images of fibrils were acquired with a JEM-1200EX II transmission electron microscope (JEOL, Tokyo, Japan) with an acceleration voltage of 85 keV. The samples were negatively stained with 1.5%–2.0% phosphotungstate adjusted at pH 7.5 by adding sodium hydroxide. The reported diameters of fibrils are the average of more than 100 randomly selected sections of different fibrils.

Predictions of amyloid aggregation propensities

Predictions were made by using the Zyggregator† algorithm,[50,55] which we previously introduced to predict the intrinsic aggregation propensities of amino acid sequences by considering the physicochemical properties of amino acids. In these predictions, a region of a polypeptide sequence should meet two conditions in order to promote aggregation: (1) it should have a high intrinsic aggregation propensity Z^{P} and (2) it should be sufficiently unstable to have a significant propensity to form intermolecular interactions. In order to describe the latter, we used the CamP method for the prediction of structural flexibility (ln*P*) from the amino acid sequence.[63] For those values that have $Z^{P}>0$, we modified the aggregation propensity profile Z^{P} by modulating it with ln*P*,

$$Z_i^{PS} = Z_i^{P}\left(1 - \frac{\ln P_i}{15}\right),$$

thus resulting in the aggregation propensity profile Z^{PS} with structural corrections.

Limited proteolysis by proteinase K

β-Lactoglobulin (1 mg mL^{-1}) was dissolved in 1 mL of 10 mM sodium phosphate, pH 7.0, and 5.0 M urea. One microliter of 10 ng mL^{-1} of proteinase K dissolved in distilled water was added into the solution containing β-lactoglobulin, and the solution was incubated for 1 min at 37 °C. The cleavage reaction was quenched by addition of phenylmethylsulfonyl fluoride at room temperature. Disulfide bonds were then reduced by addition of dithiothreitol (final concentration of 10 mM) for 1 h at room temperature. The solution was treated by ZipTip with 0.6 μL of C_4 resin (Millipore, Billerica, MA). Masses of the peptides in the solution after ZipTip treatment were analysed by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry using AXIMA-CFR (SHIMADZU, Kyoto, Japan). α-Cyano-4-hydroxycinnamic acid was used as matrix. The fragments were identified by using FindPept tool in ExPaSy‡ server. Further details are provided as supplementary data.

CD spectra

Far-UV CD spectra were monitored at 37 °C using a J-720 spectropolarimeter from JASCO (Tokyo, Japan). Protein samples were prepared in 10 mM sodium phosphate and various concentrations of urea at pH 7.0. A protein concentration of 1 mg mL^{-1} and the cell of 0.1-mm path length were used. It was essential to use the cell with such short path length and relatively high concentration of protein in order to efficiently eliminate the high absorption of urea, particularly at shorter wavelength, below ~220 nm. Data were expressed as molar residue ellipticity [θ].

The global-fitting analysis was performed as described in Ref. 30. In order to obtain the CD spectrum of an intermediate state and various thermodynamic parameters to explain the urea-unfolding reactions, we introduced the following assumptions:

(1) Three states, including native dimer (N), intermediate (I) and fully unfolded (U), were responsible for the urea-unfolding transition. Thus, the observed ellipticity at wavelength x, $[\theta](x)$, in the presence of a given concentration of urea is expressed by a linear combination of the contribution from the three states:

$$[\theta](x) = f_N[\theta]_N(x) + f_I[\theta]_I(x) + f_U[\theta]_U(x)$$

where f_i and $[\theta]_i(x)$ are the fraction and ellipticity at wavelength x of the species i (U, N or I), respectively.

(2) The free-energy difference between states i and j (ΔG_{ij}) is a linear function of urea concentration, $\Delta G_{ij} = \Delta G_{ij}(H_2O) - m_{ij}$ [urea], where $\Delta G_{ij}(H_2O)$ and m_{ij} are the free-energy difference in the absence of urea and a measure of the cooperativity for urea-unfolding transition, respectively. In the case of urea unfolding of β-lactoglobulin at pH 7.0, "N" corresponds to the dimeric native state; therefore, ΔG_{NI} is also a function of total concentration of protein in solution and urea concentration.

† http://www-vendruscolo.ch.cam.ac.uk/zyggregator.php

‡ kr.expasy.org/tools/findpept.html

For the global-fitting analysis, the CD spectra of β-lactoglobulin in the absence and presence of 7.0 M urea were used as $[\theta]_N(x)$ and $[\theta]_U(x)$, respectively. $[\theta]_I(x)$ of the intermediate state and other thermodynamic parameters were included as fitting parameters. The data were analysed by IgorPro version 6.04 (Wavemetrics, OR).

We also analysed the data according to a two-state model including native and fully unfolded states. However, this approach could not provide reasonable convergence of the fitting result; we thus adopted the three-state assumption made above. Details of the results of global-fitting analysis are available as supplementary data.

Acknowledgements

This work was supported in part by Grants-in-Aid for Scientific Research on Priority Areas (no. 15076201) and the Global COE Program A08 from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to D.H.) and through grants from the Leverhulme and Wellcome Trusts (to M.V and C. M. D.). D.H. acknowledges Dr. Takashi Fukui for assistance in CDpro analysis. A.P. acknowledges with gratitude the receipt of a Cambridge Gates Scholarship; M.V., that of a Royal Society University Research Fellowship and an EMBO Young Investigator award.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.12.038

References

1. Dobson, C. M., Šali, A. & Karplus, M. (1998). Protein folding: a perspective from theory and experiment. *Angew. Chem.* **110**, 908–935.
2. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–230.
3. Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, **426**, 884–890.
4. Stefani, M. (2004). Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta*, **1739**, 5–25.
5. Stefani, M. & Dobson, C. M. (2003). Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **81**, 678–699.
6. Vendruscolo, M. & Dobson, C. M. (2005). Towards complete descriptions of the free-energy landscapes of proteins. *Philos. Trans. R. Soc. London, Ser. A*, **363**, 433–450.
7. Jaroniec, C. P., MacPhee, C. E., Bajaj, V. S., McMahon, M. T., Dobson, C. M. & Griffin, R. G. (2004). High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl Acad. Sci. USA*, **101**, 711–716.
8. Khurana, R., Souillac, P. O., Coats, A. C., Ionescu-Zanetti, C., Solomon, A. & Fink, A. L. (2003). A model for amyloid fibril formation in immunoglobulin light chains based on comparison of amyloidogenic and benign proteins and specific antibody binding. *Amyloid*, **10**, 97–109.
9. Makin, O. S. & Serpell, L. C. (2002). Examining the structure of the mature amyloid fibril. *Biochem. Soc. Trans.* **30**, 521–525.
10. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A.Ø., Riekel, C., Grothe, R. & Eisenberg, D. (2005). Structure of the cross-β spine of amyloid-like fibrils. *Nature*, **435**, 773–778.
11. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I. *et al.* (2007). Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature*, **447**, 453–457.
12. Pepys, M. B. (2001). Pathogenesis, diagnosis and treatment of systemic amyloidosis. *Philos. Trans. R. Soc. London, Ser. B*, **356**, 203–210.
13. Dobson, C. M. (2001). The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc. London, Ser. B*, **356**, 133–145.
14. Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G. & Dobson, C. M. (1999). Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. *Proc. Natl Acad. Sci. USA*, **96**, 3590–3594.
15. Chiti, F., Bucciantini, M., Capanni, C., Taddei, N., Dobson, C. M. & Stefani, M. (2001). Solution conditions can promote formation of either amyloid protofilaments or mature fibrils from the HypF N-terminal domain. *Protein Sci.* **10**, 2541–2547.
16. Fändrich, M. & Dobson, C. M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* **21**, 5682–5690.
17. Fändrich, M., Fletcher, M. A. & Dobson, C. M. (2001). Amyloid fibrils from muscle myoglobin. *Nature*, **410**, 165–166.
18. Hamada, D. & Dobson, C. M. (2002). A kinetic study of β-lactoglobulin amyloid fibril formation promoted by urea. *Protein Sci.* **11**, 2411–2426.
19. Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D. & Dobson, C. M. (1998). Amyloid fibril formation by an SH3 domain. *Proc. Natl Acad. Sci. USA*, **95**, 4224–4228.
20. López de la Paz, M., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C. M., Hoenger, A. & Serrano, L. (2002). *De novo* designed peptide-based amyloid fibrils. *Proc. Natl Acad. Sci. USA*, **99**, 16052–16057.
21. Wilkins, D. K., Dobson, C. M. & Gross, M. (2000). Biophysical studies of the development of amyloid fibrils from a peptide fragment of cold shock protein B. *Eur. J. Biochem.* **267**, 2609–2616.
22. Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24**, 329–332.
23. Hamada, D., Yanagihara, I. & Tsumoto, K. (2004). Engineering amyloidogenicity towards the development of nanofibrillar materials. *Trends Biotechnol.* **22**, 93–97.
24. Yeates, T. O. & Padilla, J. E. (2002). Designing supramolecular protein assemblies. *Curr. Opin. Struct. Biol.* **12**, 464–470.
25. Hamada, D., Tsumoto, K., Sawara, M., Tanaka, N., Nakahira, K., Shiraki, K. & Yanagihara, I. (2008). Effect of an amyloidogenic sequence attached to yellow fluorescent protein. *Proteins*, **72**, 811–821.
26. Knowles, T. P., Fitzpatrick, A. W., Meehan, S., Mott, H. R., Vendruscolo, M., Dobson, C. M. & Welland,

M. E. (2007). Role of intermolecular forces in defining material properties of protein nanofibrils. *Science*, **318**, 1900–1903.
27. Chiti, F. & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366.
28. Richardson, J. S. & Richardson, D. C. (2002). Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl Acad. Sci. USA*, **99**, 2754–2759.
29. Hambling, S., MacAlpine, A. S. & Sawyer, L. (1992). β-Lactoglobulin. In *Advanced Dairy Chemistry* (Fox, P. F., ed), pp. 141–190, Elsevier, Amsterdam, The Netherlands.
30. Hamada, D. & Goto, Y. (1997). The equilibrium intermediate of β-lactoglobulin with non-native α-helical structure. *J. Mol. Biol.* **269**, 479–487.
31. 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**, 737–740.
32. Hamada, D., Segawa, S. & Goto, Y. (1996). Non-native α-helical intermediate in the refolding of β-lactoglobulin, a predominantly β-sheet protein. *Nat. Struct. Biol.* **2**, 868–874.
33. 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. *Fold Des.* **1**, 243–251.
34. 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.
35. Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y. & Roder, H. (2001). Structural and kinetic characterization of early folding events in β-lactoglobulin. *Nat. Struct. Biol.* **8**, 151–155.
36. Forge, V., Hoshino, M., Kuwata, K., Arai, M., Kuwajima, K., Batt, C. A. & Goto, Y. (2000). Is folding of β-lactoglobulin non-hierarchic? Intermediate with native-like β-sheet and non-native α-helix. *J. Mol. Biol.* **296**, 1039–1051.
37. 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.
38. Devlin, G. L., Knowles, T. P., Squires, A., McCammon, M. G., Gras, S. L., Nilsson, M. R. *et al.* (2006). The component polypeptide chains of bovine insulin nucleate or inhibit aggregation of the parent protein in a conformation-dependent manner. *J. Mol. Biol.* **360**, 497–509.
39. Hong, D. P. & Fink, A. L. (2005). Independent heterologous fibrillation of insulin and its B-chain peptide. *Biochemistry*, **44**, 16701–16709.
40. Krebs, M. R., Wilkins, D. K., Chung, E. W., Pitkeathly, M. C., Chamberiain, A. K., Zurdo, J. *et al.* (2000). Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the β-domain. *J. Mol. Biol.* **300**, 541–549.
41. Scrocchi, L. A., Ha, K., Chen, Y., Wu, L., Wang, F. & Fraser, P. E. (2003). Identification of minimal peptide sequences in the (8–20) domain of human islet amyloid polypeptide involved in fibrillogenesis. *J. Struct. Biol.* **141**, 218–227.
42. Kozhukh, G. V., Hagihara, Y., Kawakami, T., Hasegawa, K., Naiki, H. & Goto, Y. (2002). Investigation of a peptide responsible for amyloid fibril formation of β_2-microglobulin by Achromobacter protease I. *J. Biol. Chem.* **277**, 1310–1315.
43. Jones, S., Manning, J., Kad, N. M. & Radford, S. E. (2003). Amyloid-forming peptides from β_2-microglobulin—insights into the mechanism of fibril formation *in vivo*. *J. Mol. Biol.* **325**, 249–257.
44. Ionescu-Zanetti, C., Khurana, R., Gillespie, J. R., Petrick, J. S., Trabachino, L. C., Minert, L. *et al.* (1999). Monitoring the assembly of Ig light-chain amyloid fibrils by atomic force microscopy. *Proc. Natl Acad. Sci. USA*, **96**, 13175–13179.
45. Chamberlain, A. K., MacPhee, C. E., Zurdo, J., Morozova-Roche, L. A., Hill, H. A., Dobson, C. M. & Davis, J. J. (2000). Ultrastructural organization of amyloid fibrils by atomic force microscopy. *Biophys. J.* **79**, 3282–3293.
46. Zurdo, J., Guijarro, J. I., Jimenez, J. L., Saibil, H. R. & Dobson, C. M. (2001). Dependence on solution conditions of aggregation and amyloid formation by SH3 domain. *J. Mol. Biol.* **311**, 325–340.
47. Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P. & Tycko, R. (2005). Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science*, **307**, 262–265.
48. Krishnan, R. & Lindquist, S. L. (2005). Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature*, **435**, 763–772.
49. Gosal, W. S., Clark, A. H. & Ross-Murphy, S. B. (2004). Fibrillar β-lactoglobulin gels: Part 1. Fibril formation and structure. *Biomacromolecules*, **5**, 2408–2419.
50. Pawar, A. P., Dubay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M. & Dobson, C. M. (2005). Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **350**, 379–392.
51. Glover, J. R., Kowal, A. S., Schirmer, E. C., Patino, M. M., Liu, J. J. & Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of *S. cerevisiae*. *Cell*, **89**, 811–819.
52. Esteras-Chopo, A., Serrano, L. & Lopez de la Paz, M. (2005). The amyloid stretch hypothesis: recruiting proteins toward the dark side. *Proc. Natl Acad. Sci. USA*, **102**, 16672–16677.
53. Ventura, S., Zurdo, J., Narayanan, S., Parreño, M., Mangues, R., Reif, B. *et al.* (2004). Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3) case. *Proc. Natl Acad. Sci. USA*, **101**, 7258–7263.
54. Frare, E., Mossuto, M. F., Polverino de Laureto, P., Dumoulin, M., Dobson, C. M. & Fonrana, A. (2006). Identification of the core structure of lysozyme amyloid fibrils by proteolysis. *J. Mol. Biol.* **361**, 551–561.
55. Tartaglia, G. G., Pawar, A. P., Campioni, S., Dobson, C. M., Chiti, F. & Vendruscolo, M. (2008). Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.* **380**, 425–436.
56. Katou, H., Hoshino, M., Kamikubo, H., Batt, C. A. & Goto, Y. (2001). Native-like β-hairpin retained in the cold-denatured state of bovine β-lactoglobulin. *J. Mol. Biol.* **310**, 471–484.
57. Sakurai, K. & Goto, Y. (2007). Principal component analysis of the pH-dependent conformational transitions of bovine β-lactoglobulin monitored by heteronuclear NMR. *Proc. Natl Acad. Sci. USA*, **104**, 15346–15351.
58. Giurleo, J. T., He, X. & Talaga, D. S. (2008). β-Lactoglobulin assembles into amyloid through sequential aggregated intermediates. *J. Mol. Biol.* **381**, 1332–1348.
59. Sreerama, N. & Woody, R. W. (2000). Estimation of

protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260.
60. Dumoulin, M. & Dobson, C. M. (2004). Probing the origins, diagnosis and treatment of amyloid diseases using antibodies. *Biochimie*, **86**, 589–600.
61. Canet, D., Last, A. M., Tito, P., Sunde, M., Spencer, A., Archer, D. B. *et al.* (2002). Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme. *Nat. Struct. Biol.* **9**, 308–315.
62. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **189**, 319–326.
63. Tartaglia, G. G., Cavalli, A. & Vendruscolo, M. (2007). Prediction of local structural stabilities of proteins from their amino acid sequences. *Structure*, **15**, 139–143.