

# Partially folded structure of monomeric bovine $\beta$ -lactoglobulin

Henriette Molinari<sup>a,\*</sup>, Laura Ragona<sup>b</sup>, Luca Varani<sup>b</sup>, Giovanna Musco<sup>c</sup>, Roberto Consonni<sup>b</sup>,  
Lucia Zetta<sup>b</sup>, Hugo L. Monaco<sup>d</sup>

<sup>a</sup>*Istituto Policattedra della Facoltà di Scienze, Università degli Studi di Verona, Ca Vignal I, Strada le Grazie, 37100 Verona, Italy*

<sup>b</sup>*Laboratorio NMR, ICM, CNR, Via Ampere 56, 20131 Milano, Italy*

<sup>c</sup>*EMBL, Meyerhofstrasse, Heidelberg, Germany*

<sup>d</sup>*Dipartimento di Genetica, Università degli Studi di Pavia, 27100 Pavia, Italy*

Received 18 September 1995; revised version received 8 January 1996

**Abstract** Bovine  $\beta$ -LG ( $\beta$ -lactoglobulin) has been studied under a variety of solution conditions by one- and two-dimensional NMR spectroscopy. At highly acidic pH (pH = 2) and low ionic strength the protein is present in a monomeric form, exhibiting a highly structured  $\beta$ -sheet core and less ordered regions as evidenced by both CD data and the NOESY spectra. Marginal protection was observed for most of the amide protons as a result of high conformational mobility. This structural state of  $\beta$ -LG may be considered as an attractive model for a partially folded structure occurring late in the folding process of the protein.

**Key words:** Locally unfolded structure;  $\beta$ -Lactoglobulin; CD; NMR

## 1. Introduction

The characterization of unfolded and partly folded states of proteins is central to the understanding of protein stability and folding. Most of the partially folded structures characterized so far appear to be highly disordered, and only recently, more persistent structures or more ordered molten globule states have been identified, thus providing insight into factors that stabilize protein structures [1].

In this paper we present data that indicate that at low pH the  $\beta$ -barrel protein  $\beta$ -LG retains a well-defined highly structured  $\beta$ -sheet core and loses some of its secondary structure. The presence of a high structural variability in a significant number of residues of the molecule, indicates local unfolding. Bovine  $\beta$ -LG is an 18,300 dalton molecular weight protein found in the milk of ruminants where it is often present as a population of more than one genetic variant of almost identical amino acid compositions [2–5]. The function of the molecule is unknown but since its three-dimensional structure is very similar to that of plasma retinol-binding protein and since it has been shown that it can bind many hydrophobic molecules, most notably retinol itself,  $\beta$ -LG has been included in the family of a broad class of protein transporters that are sometimes called either hydrophobic molecule-binding proteins or lipocalins [6,7].  $\beta$ -LG has been the subject of many extensive studies using very different physico-chemical techni-

ques [8]. The three-dimensional structure of  $\beta$ -LG has been revealed by X-ray diffraction studies of more than one crystal form:  $\beta$ -LG is a  $\beta$ -barrel made up of 8 strands of antiparallel  $\beta$ -sheets and a short  $\alpha$ -helical segment [9,10].

In spite of the wealth of data found in the literature on this molecule, no NMR structural studies have been reported which is probably due to the complex pattern of association and aggregation exhibited by  $\beta$ -LG under physiological conditions [11–13]. The association behaviour of  $\beta$ -LG is largely influenced by pH changes: the two most common variants of bovine  $\beta$ -LG,  $\beta$ -LG-A and -B (which differ in two amino acids), are in a monomer–dimer equilibrium at pH 3.0 and appear to form higher aggregates above pH 3.5 [14]. Since of these two forms it is only the monomer that can be amenable to a detailed NMR study, we undertook a search for the conditions under which the equilibrium is completely shifted towards the formation of the monomer.

The sequence of bovine  $\beta$ -LG is reported in Fig. 1, together with an indication of the secondary structure elements as defined by the non-refined X-ray structures. Out of a total of fifteen  $\beta$ -LGs from different species listed in the Swiss-Prot Databank, ten are reported to be monomers under physiological conditions, while the other five (cow, sheep, goat, water buffalo and pig) are reported to form dimers or higher aggregates [15]. It is worth noticing that these are highly homologous in contrast to the ten monomeric  $\beta$ -LGs which show high homology only in the N-terminal region.

## 2. Materials and methods

Bovine  $\beta$ -LG was prepared from freshly obtained milk following a procedure that avoids heating and the use of extreme values of pH [10]. The two genetic variants A and B were separated by the method of Piez et al. [16]. Following purification, potential endogenous hydrophobic ligands were eliminated by stirring the protein dissolved in 0.05 M acetate, pH = 5, at a concentration of 0.75 mg/ml with the same weight of activated charcoal for about 2 h at room temperature. The electrospray mass of the product of this purification procedure gave for the B form of bovine  $\beta$ -LG a molecular weight of  $18,275 \pm 1$  Da and no other spurious peaks. Capillary zone electrophoresis confirmed that the protein was highly pure and crystallization experiments yielded the canonical crystals indicating that during this process the three-dimensional structure of the protein had been conserved. The samples were buffered exchanged using concentration cells with 10-kDa cut-off membranes (Amicon, Gloucester, UK). The pH values quoted are uncorrected for isotope effects. Sample concentrations were between 0.9 and 1.5 mM.

### 2.1. CD

Spectra of  $\beta$ -LG were collected using a Jasco Model J600 spectropolarimeter. The concentration was 15  $\mu$ M with a cuvette pathlength

\*Corresponding author. Lab. NMR, Via Ampere 56, 20131 Milano, Italy. Fax: (39) (2) 266-3030; E-mail: henry@labnmr.icmmr.mi.cnr.it

**Abbreviations:** NMR, nuclear magnetic resonance; DQF-COSY, double quantum filtered correlated spectroscopy; TOCSY, total correlated spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; CD, circular dichroism;  $\beta$ -LG,  $\beta$ -lactoglobulin.

of 0.1 and 0.02 cm for the far-UV determination and 30–60  $\mu\text{M}$  with a cuvette pathlength of 1 cm for the near-UV. Each spectrum is the average of 8 scans; a solvent spectrum was subtracted from each protein spectrum, at each temperature. The molar ellipticity was calculated according to:

$$[\theta]_R^\lambda = \theta^\lambda 100 / [(N-1)lc]$$

where  $\theta^\lambda$  is the measured ellipticity (in degrees) at a wavelength  $\lambda$ ,  $N$  is the number of residues,  $l$  is the cuvette pathlength (in cm) and  $c$  is the protein concentration (in mol/l). The molar ellipticity  $[\theta]_R^\lambda$  was plotted vs. the wavelength in the 190–340 nm region. The denaturation curve was obtained by plotting  $[\theta]_R^{213}$  vs. the experimental temperatures (in  $^\circ\text{C}$ ). The deconvolution of the CD spectra was performed using a program kindly provided by S. Mammì, based on the method of Pribic et al. [17].

## 2.2. NMR

Measurements were carried out at  $^1\text{H}$  frequency of 500.13 and 600.13 MHz on DMX-Bruker spectrometers. 1D spectra were collected using 4K or 8K data points over a spectral width of 7002.8 or 7788.162 Hz, collecting 128 scans. DQF-COSY, TOCSY, and NOESY spectra were acquired over 4K data points and 500–1024  $t_1$  increments in the absorption mode with time-proportional phase incrementation (TPPI) for quadrature detection in the  $t_1$  dimension. Water saturation was achieved by low power irradiation during the relaxation delay introduced between scans. Gradients were used for water suppression for the NOESY at 600 MHz, in water. A total of 128 or 256 transients were collected for each  $t_1$  increment. Mixing times of 20, 50 and 80 ms were employed for each TOCSY experiment; NOESY experiments were acquired with 60, 100, 160 and 180 ms mixing. 2D spectra were processed on a X32 using the UXNMR program provided by Bruker. The data set was resolution enhanced using Lorentzian-Gaussian transformation prior to zero-filling in  $F_1$ .

## 3. Results and discussion

Searching for the conditions under which the  $\beta$ -LG monomer is predominant, we have checked different pH values from 1.9 to 3.4, different ionic strengths, down to 0.012 M sodium phosphate buffer in the presence and in the absence of NaCl and different temperatures, in the range 10–37 $^\circ\text{C}$ . We found, on the basis of both NMR and electrospray mass data (see below), that the monomer is obtained at pH 2–2.4 at 0.012 M sodium phosphate without NaCl (ionic strength  $I = 0.006$ ). It is likely that under these conditions aggregation is inhibited by the electrostatic repulsion of the positively charged protein at low pH values and low ionic strength where screening of the charges is minimal.



Fig. 1. Sequence of  $\beta$ -LG, containing the indication on the secondary structure elements, as derived from the unrefined X-ray structure [10].

## 3.1. CD studies

The CD spectra of  $\beta$ -LG were recorded at pH 2.1–2.3. In aqueous solution, the CD spectrum is typical of a folded structure containing predominantly antiparallel  $\beta$ -sheets, in agreement with X-ray data [9,10], with a negative band at 213 nm ( $[\theta] = -6780 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$ ) and a strong positive band at 189 nm. The CD spectrum is concentration independent in the range  $1.6 \times 10^{-5}$  to  $1.6 \times 10^{-3}$  M protein. The deconvolution of the CD spectra indicated the presence of ca. 40%  $\beta$ -structure. At 37 $^\circ\text{C}$  a comparison has been reported (Fig. 2b) between the far UV CD spectra of the low pH form of  $\beta$ -LG (pH 2.1) and the native protein (pH 7.3): it is clear from the figure that the two spectra are very similar. The far and near UV CD spectra of the low pH form of  $\beta$ -LG at variable temperatures are shown in Fig. 2a and c: the low pH form of  $\beta$ -LG exhibits very similar far UV CD spectra in the entire range 27–70 $^\circ\text{C}$  (Fig. 2a) and only at temperatures higher than 70 $^\circ\text{C}$  the minimum at 213 nm shows a blue shift, together with a more negative  $[\theta]$  value. The change observed in the far UV CD spectra at high temperatures can be attributed to partial denaturation that reduces the amount of  $\beta$ -structure present in the protein. The curve obtained for bovine  $\beta$ -LG thermal denaturation is reported in Fig. 2d. A sigmoidal curve, indicative of a cooperative mechanism, cannot be expected under these conditions to be due to an additional irreversible association that can take place at high temperature and involves intermolecular chemical reactions through the formation of new disulphide bonds among denaturated molecules [14,18,19]. This is confirmed by the sudden decrease of ellipticity on going from 70 to 90 $^\circ\text{C}$ . The existence of such an irreversible change at higher temperature is confirmed by the differences observed for the two CD spectra obtained at 37 $^\circ\text{C}$  before and after thermal treatment up to 80 $^\circ\text{C}$  (data not shown), suggesting that renaturation was not viable under these conditions. This observation is in contrast with data reporting that the CD spectra of  $\beta$ -LG renatured after heating up to 80 $^\circ\text{C}$  are either identical or very similar to those of the native protein [20]. In the near UV CD region the ellipticity  $[\theta]$  of the low pH form of  $\beta$ -LG is substantially reduced as compared to that of the neutral form of  $\beta$ -LG (pH 2.1,  $[\theta]^{293} = -50 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$  and pH 7.3,  $[\theta]^{293} = -80 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$ ) (Fig. 2c). This suggests that the environment of aromatic groups in the low pH form of  $\beta$ -LG is time-averaged in such a way that these groups have lost part of their rigid environment. The effect of temperature, at pH 2.1, is shown in Fig. 2c; it is worth noticing that a substantial loss of signal is already observed on going from 37 to 55 $^\circ\text{C}$  (differently from the behaviour observed in the far UV), while a small signal is left at 75 $^\circ\text{C}$ . Loss of near-UV CD spectrum, characteristic of the absence of well-defined structure in the vicinity of aromatic chromophores, is one of the characteristic of partially folded states.

## 3.2. NMR studies

The monodimensional NMR spectra of  $\beta$ -LG in  $\text{D}_2\text{O}$ , at 37 $^\circ\text{C}$ , at pH 2.3 and 3.4, are shown in Fig. 3A and B, respectively. Inspection of the first spectrum shows a large dispersion of chemical shifts, in particular a significant number of downfield shifted  $\text{H}_\alpha$  proton resonances between 5 and 6.2 ppm indicating the presence of a  $\beta$ -sheet structure. In addition, several methyl group resonances appear upfield of the main aliphatic region (between 0 and 0.7 ppm), and reso-

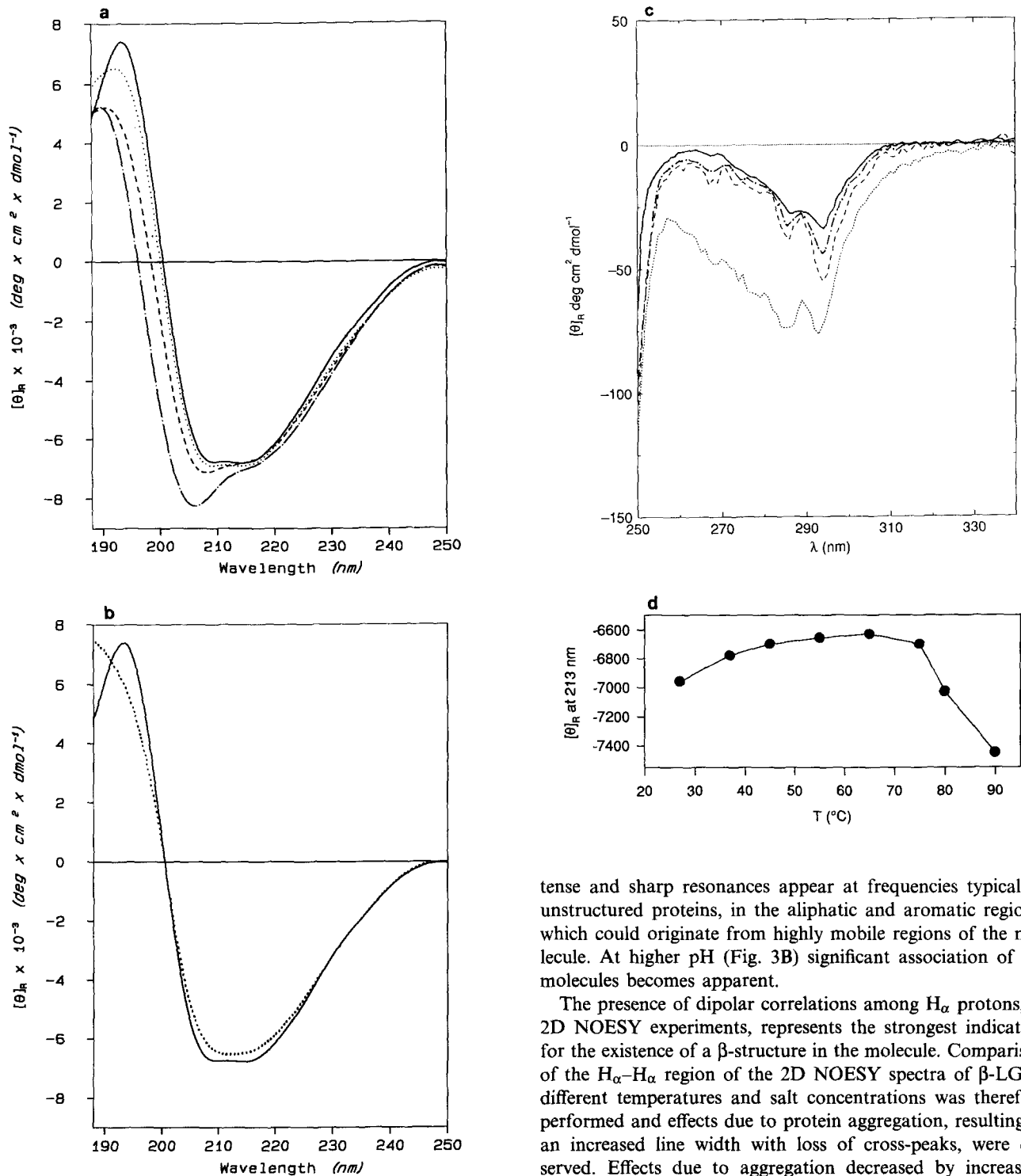


Fig. 2. (a) Far-UV CD spectra of  $\beta$ -LG, pH 2.1, 12 mM H<sub>3</sub>PO<sub>4</sub>/NaOH, at different temperatures: 37°C (solid line), 55°C (dotted line), 75°C (long-dashed line), 80°C (long-dashed/dotted line). (b) Far-UV CD spectra of  $\beta$ -LG at pH 2.1 (solid line) and pH 7.3 (dotted line). (c) Near-UV CD spectra of  $\beta$ -LG pH 7.3 37°C (dotted line), pH 2.1 37°C (long-dashed line), pH 2.1 55°C (long-dashed/dotted line), pH 2.1 75°C (solid line). (d) Thermal denaturation of  $\beta$ -LG:  $[\theta]_R$  at 213 nm is reported against temperature.

nances of a small number of aromatic ring protons are resolved in the region of 6.4–6.7 ppm, both indicating the presence of globular structure. At the same time some very in-

tense and sharp resonances appear at frequencies typical of unstructured proteins, in the aliphatic and aromatic regions, which could originate from highly mobile regions of the molecule. At higher pH (Fig. 3B) significant association of the molecules becomes apparent.

The presence of dipolar correlations among H <sub>$\alpha$</sub>  protons, in 2D NOESY experiments, represents the strongest indication for the existence of a  $\beta$ -structure in the molecule. Comparison of the H <sub>$\alpha$</sub> –H <sub>$\alpha$</sub>  region of the 2D NOESY spectra of  $\beta$ -LG at different temperatures and salt concentrations was therefore performed and effects due to protein aggregation, resulting in an increased line width with loss of cross-peaks, were observed. Effects due to aggregation decreased by increasing the temperature, as observed from the comparison of 2D TOCSY, DQF-COSY and NOESY performed in the range 10–37°C on a 1.5 mM sample at pH 2.1 (50 mM H<sub>3</sub>PO<sub>4</sub>/NaOH). The effect of the ionic strength is shown in Fig. 4A–C, reporting the H <sub>$\alpha$</sub> –H <sub>$\alpha$</sub>  region of the NOESY spectra of a 1.5 mM sample of  $\beta$ -LG at pH 2.1, 37°C in the presence of 150 mM Na<sub>2</sub>HPO<sub>4</sub>/HCl, 50 mM H<sub>3</sub>PO<sub>4</sub>/NaOH and 12 mM H<sub>3</sub>PO<sub>4</sub>/NaOH, respectively. It is apparent that the spectrum in Fig. 4C must be due to the protein in the monomeric form, in fact it contains the most intense and the highest number of H <sub>$\alpha$</sub> –H <sub>$\alpha$</sub>  cross-peaks. The chemical shifts of the observed H <sub>$\alpha$</sub> –H <sub>$\alpha$</sub>  correlations fall in the region between 4.7

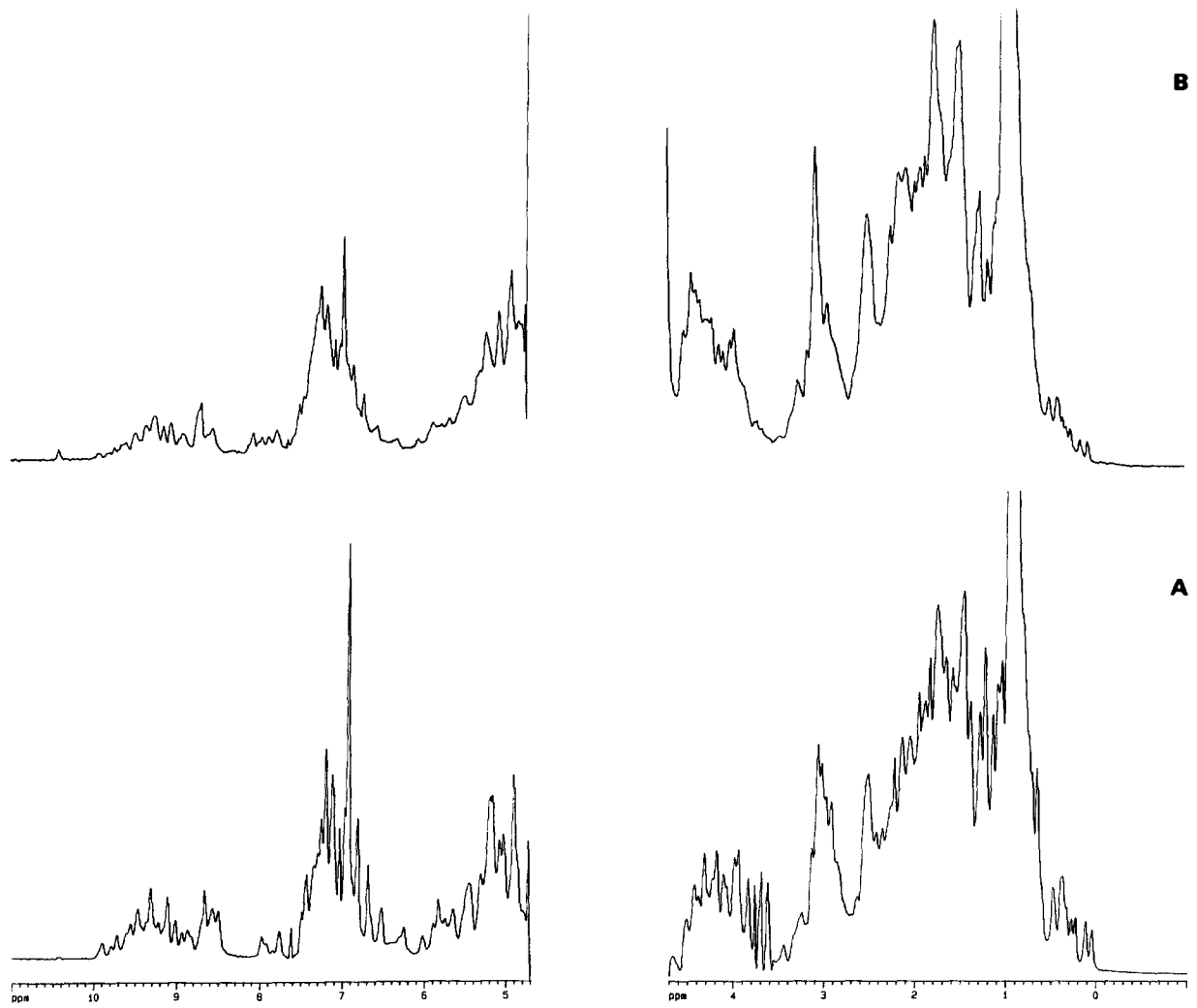


Fig. 3. 500 MHz 1D NMR spectra of  $\beta$ -LG at 37°C, 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ , at: (A) pH 2.3; (B) pH 3.4.

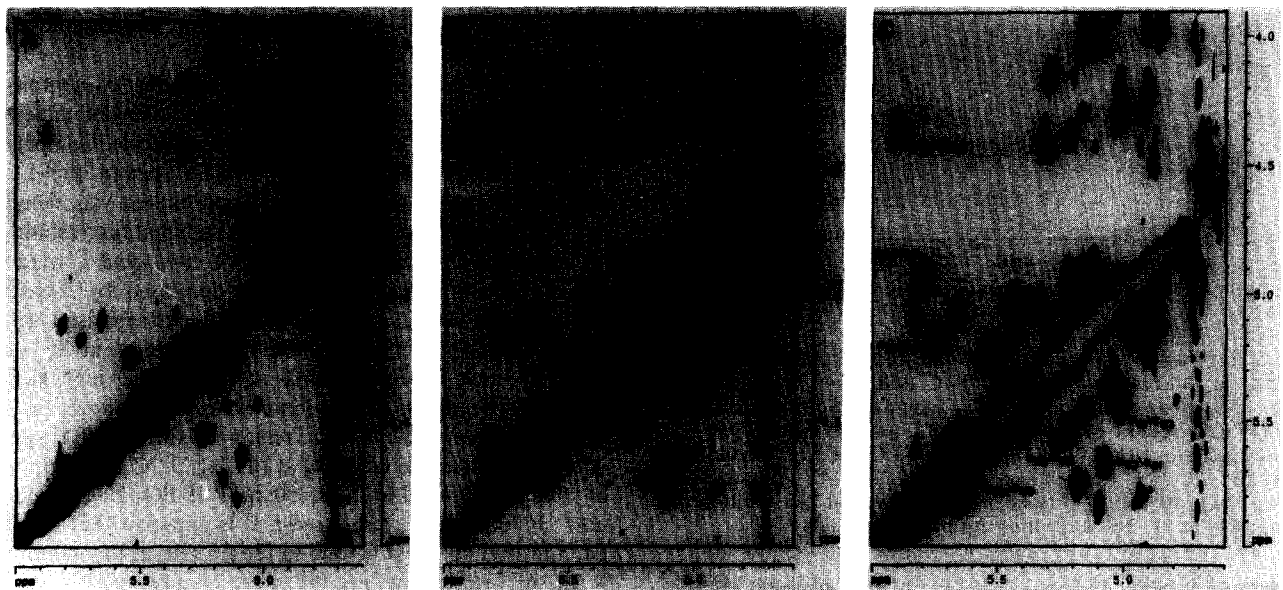


Fig. 4.  $\text{H}_\alpha$ - $\text{H}_\alpha$  region of 500 MHz NOESY spectra (mix-100 ms) of  $\beta$ -LG at different salt concentrations: (A) 150 mM  $\text{Na}_2\text{HPO}_4/\text{HCl}$ ; (B) 50 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ ; (C) 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ .

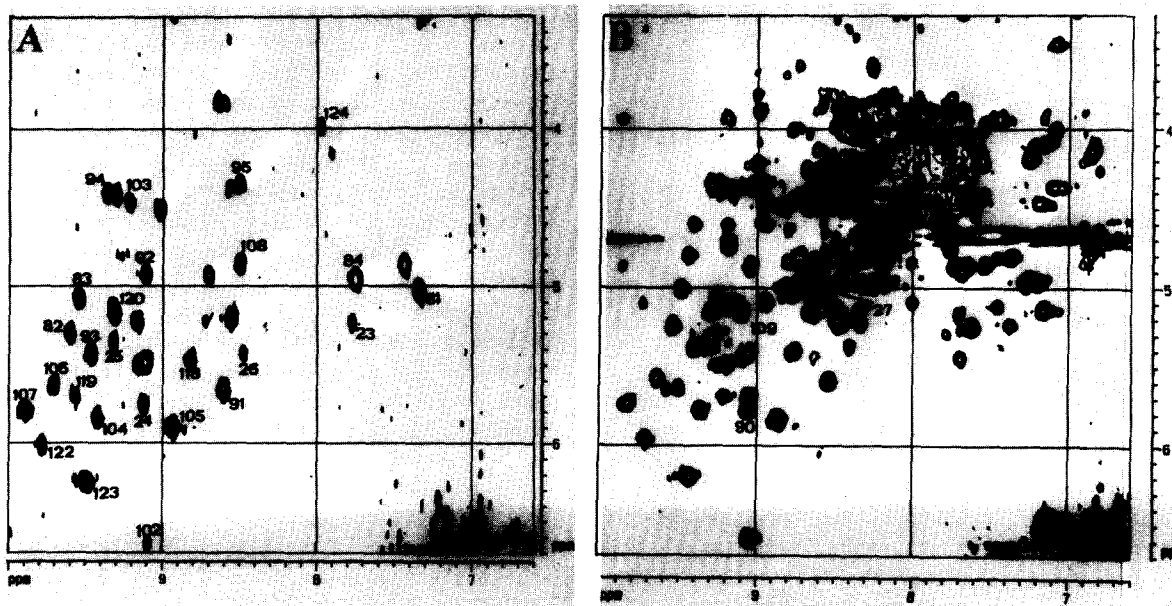


Fig. 5. Fingerprint region of the TOCSY spectra of  $\beta$ -LG at 37°C, 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ , pH 2.4 in: (A)  $\text{D}_2\text{O}$ ; (B)  $\text{H}_2\text{O}$ .

and 6.2 ppm, thus indicating the existence of a region of  $\beta$ -sheet structure, in agreement with the CD data.

The fact that the protein is a monomer under the experimental conditions of spectrum 4C (pH 2.1, 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ ,  $I = 0.006$ ) is in agreement with data previously reported in the literature [21]. Positive electrospray mass (EM) spectrum of an aqueous solution of this protein (1 mM sample, pH 2.1, 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ ,  $I = 0.006$ ) agreed with the presence of a pure monomeric protein having the molecular weight  $18,275 \pm 1$  Da, since: (i) the EM spectrum exhibited a single-charge status distribution containing a number of sharp peaks extending between +11 and +16; (ii) the same mass spectrum did not contain charge states which may derive from the dimeric structure (i.e. +23, +25, +27, etc.).

A rough estimate of  $\tau_c$ , employing the cross-diagonal peak ratio [22] led to  $\tau_c$  ranging from 6.6 to 7.2 ns, using, as reference distances, the ortho protons of Y 102 and the  $\text{H}_\alpha$ -HN intraresidue distances of 3 Å (from the X-ray structure) for Leu-122 and Val-123 (all the mentioned residues are in the  $\beta$ -sheet core). This  $\tau_c$  is consistent with a 18,300 Da protein [23].

The electrospray mass alone is not unequivocal in assessing the monomeric state of the protein but the combination of (a) prior work on sedimentation equilibria, (b) the electrospray mass data, and (c) the  $\tau_c$  estimated from the NMR data, together compel the conclusion that the protein is a monomer.

When the 2D experiments were performed at pH 3.4, where the aggregation into a dimer is rather important, the pattern of observable chemical shifts showed little change with respect to that at lower pH, but the resonances were so broad and the overlap so severe that any detailed NMR study became unfeasible.

Sequence specific assignments were undertaken in the experimental conditions reported for spectrum 4C, following both the classical methods described by Wüthrich [24] and the main-chain-directed strategy [25]. The NOE patterns observed for the amide protons present in  $\text{D}_2\text{O}$  are characteristic of  $\beta$ -sheet, with strong  $d_{\alpha\text{N}(i,i+1)}$  interactions, weak or absent intraresidue  $d_{\text{N}\alpha}$  connectivities, and few NOEs common to consecutive amide protons. Complete assignment of the spec-

trum has not so far been possible, because of the broad line widths and the limited resolution in some regions of the spectrum. Furthermore the assignment is seriously hindered by the lack of residue-type information from the TOCSY data for many of the amide protons. Since the expression of  $\beta$ -LG in yeast has so far produced the wild protein in very low yields [26], selectively labelled protein that could be of great importance in helping to interpret its very complex NMR spectra has not been available.

Of the 40–43 slowly exchanging amides giving detectable cross-peaks in the fingerprint region of the TOCSY spectrum of the protein freshly dissolved in  $\text{D}_2\text{O}$ , (Fig. 5A), 30 have been identified. They comprise the segments 23–27, 81–84, 90–95, 102–109 and 118–124 belonging to the  $\beta$ -sheet core, in particular to chains a, e, f, g, and h (Fig. 1). The interstrand dipolar correlations between these segments were all assigned. The starting point for these assignments was the individuation of the segment Phe-Cys-Met which could only belong to the g segment, at positions 105, 106 and 107. The fingerprint region of the TOCSY spectrum (Fig. 5A) remains unchanged after three weeks. Approximately the same number of  $\text{NH-H}_\alpha$  correlations (35–38) is visible in the DQF spectrum where, due to the intrinsically high line width of this 18 kDa protein and to the anti-phase nature of this type of experiment, the cross-peaks detected must have a high J coupling and therefore belong to  $\beta$ -sheet segments.

It is worth comparing this experimentally measured number of amides protected from exchange with that expected from structural data reported in the literature [9,10]. In spite of the fact that the crystal structure for this protein, at pH higher than 7, was determined nearly ten years ago, refined coordinates are not available. On the basis of the definition of the secondary structure elements reported by Monaco et al., and represented in Fig. 1 [27], the total number of amides involved in hydrogen bonds in the barrel structure should be roughly 60, of which approximately 30 should correspond to the assigned segments a, e, f, g and h.

The protein was then dissolved in  $\text{H}_2\text{O}$ , to help in the assignment procedure, since the lack of  $\text{NH-H}_\alpha$  correlations

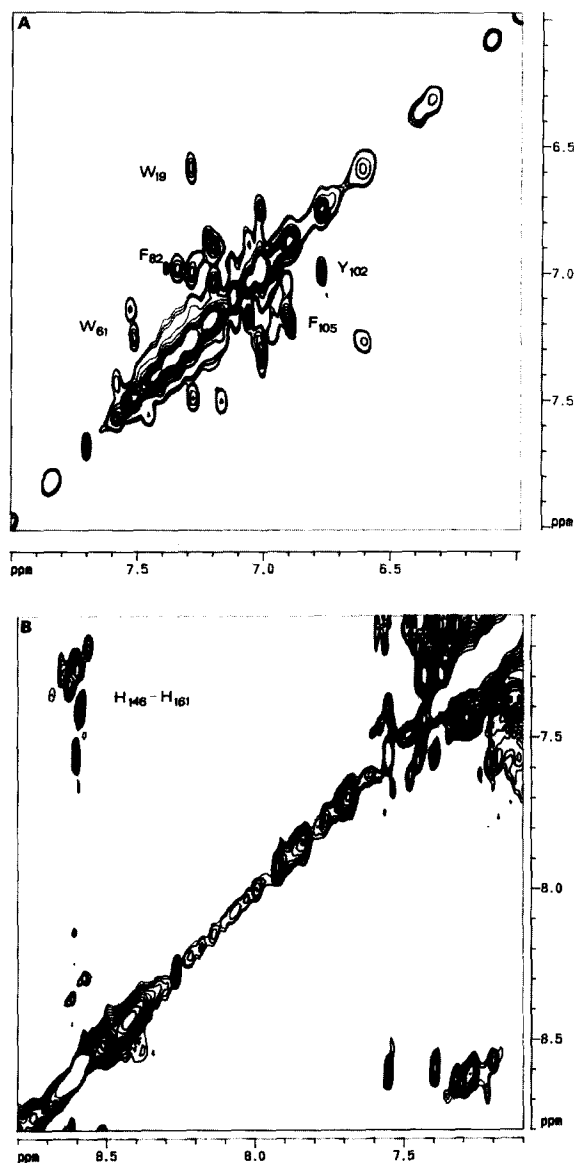


Fig. 6. (A,B) Different aromatic regions of the 500 MHz TOCSY spectrum of  $\beta$ -LG at 37°C, 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$ .

prevented further steps in the assignment. Fig. 5B shows the fingerprint region of the TOCSY experiment of  $\beta$ -LG dissolved in water. The noticeable feature of this spectrum is that in addition to the correlations observed in  $\text{D}_2\text{O}$ , a region of broadly overlapping peaks, corresponding to the rest of the backbone correlations is also present. These peaks clearly show a very small dispersion of secondary chemical shifts which appear centred at the random coil position [28], indicating the presence of high conformational mobility resulting in the averaging on the chemical shift time scale and in line broadening.

Even if only a few of the new peaks appearing in the resolved region of the protein spectrum in  $\text{H}_2\text{O}$  have been assigned (see Fig. 5B), we have no evidence that any of them belong to the strands b, c or d. These results altogether strongly support the hypothesis that under highly acidic conditions this protein contains a subdomain with a highly ordered resistant  $\beta$  sheet core, while the rest of the molecule exhibits an increased flexibility, giving rise to locally disor-

dered structures. It is worth noting that recent data on the cleavage of  $\beta$ -LG by pepsin in water/ethanol (where a cooperative conformational change has been reported) show the presence of many cleavage sites, which leave however unaffected the region 83–122, that roughly corresponds to the resistant core identified by these NMR experiments [29].

Extensive flexibility is also evident from the analysis of the aromatic region (Fig. 6A,B), showing a severe spectral overlap, consequence of the large number of cross-peaks close to the diagonal. This correlates with an increase in intensity between 6.8 and 7.3 ppm, as seen in the 1D spectra, and results from resonances of residues in the unstructured segment of the protein.  $\beta$ -LG contains 12 aromatic residues, 4 Tyr, 4 Phe, 2 Trp and 2 His. A preliminary assignment of the resonances of Phe-82, Phe-105, Trp-19, Trp-61, Tyr-102, His-146 and His-161 has been made. The aromatic chemical shifts do not show any noticeable change when the ionic strength is reduced from 150 mM  $\text{Na}_2\text{HPO}_4/\text{HCl}$  to 12 mM  $\text{H}_3\text{PO}_4/\text{NaOH}$  with the only exception of the cross-peak H2,6/H3,5 of Tyr-102, which moves downfield of about 0.2 ppm, and its amide proton peak which shows a significant downfield shift as well. Of the two Trp residues, Trp-61 occurring at lower fields, is located on the surface of the molecule, as judged from the behaviour of its indolic amide proton which is not protected from exchange. Resonances of ring protons 5H and 6H of Trp-19 appear at 6.44 ppm, and those of 7H and 4H at 7.14 ppm, far from the random coil values, indicating that this residue belongs to the folded core of the protein. This is confirmed by the fact that the indolic amide of Trp-19 is rather resistant to deuterium exchange, and is still present in the spectrum after the protein has been three weeks dissolved in  $\text{D}_2\text{O}$ . Fig. 6B shows the TOCSY correlations observed for His-146 and His-161, both located outside the  $\beta$ -strand region, in the C-terminal part of the molecule. In both cases a high number of correlations is observed, indicating the presence of many conformers. For His-146, the *cis-trans* isomerism of Pro-144 together with high flexibility could explain this behaviour.

In conclusion the CD and NMR data reported here indicate that under extremely acidic conditions  $\beta$ -LG contains a region of highly ordered  $\beta$ -sheet, which appears to form a compact core that does not encompass the whole of the eight  $\beta$ -sheet strands, detected by X-ray diffraction under very different conditions.

Fink et al. [30] have investigated the effect of acid conditions on twenty monomeric proteins showing that there can be different types of conformational behaviour depending on the protein, the acid, the presence or absence of salts and denaturants and the temperature. In their study,  $\beta$ -LG is classified in the protein group that shows native-like behaviour at pH=2.0, even if it behaves differently from the other members of the same class. In fact, urea titrations did not show unfolding transitions in the 2–4 M range, as observed for all the other proteins of the same class. We suggest that our results may explain some of the reported observations, although the main difference between the reported observations and our data may be due to differences in the starting protein, a commercial sample in Fink's study, and a protein purified by us from milk in our case.

The coexistence, in the low pH structural state of  $\beta$ -LG, of a subdomain containing a stable core together with an unstructured region, makes this structure an attractive model

for an intermediate occurring late in the folding process of the protein, since stable partly folded states have been shown to be related to kinetic intermediates in the folding pathways [1,31,32]. Although most partially folded states so far characterized appear to be highly disordered, there is a growing interest in the identification of those conformational states that retain a significant part of the native-like structure [33–36]. It has been recently suggested that subdomains with persistent and well-ordered structure are particularly likely to exist for well-defined supersecondary structural motifs such as the  $\beta$ -barrel topologies [37]. Furthermore this study confirms that NMR is the appropriate tool to study such partially ordered structures. NMR studies both on the intact protein and on synthetic peptides from different regions of the molecule are currently in progress.

*Acknowledgements:* The authors are indebted to Annalisa Pastore (EMBL) for stimulating discussions, to Stefano Mammi (University of Padova), to Giuseppe Vecchio (CNR, Milano), to Gianluca Fossati (Italfarmaco, Milano) and to Maddalena Collini (University of Milano) for providing to us the CD facilities, to Heinz Ruterjans and Frank Löhner for making available to us the NMR facility at the Biophysic Centre of Frankfurt University, to Marcella Chiari for electrophoresis measurements, to Mahmoud Hamdan (Glaxo, Verona) for providing all the electrospray mass measurements and to Fulvia Greco for technical assistance. This work was supported by the Italian National Research Council: "Comitato Biotecnologie" and by the Associazione "Antonio De Marco".

## References

- [1] Dobson, C.M. (1992) *Curr. Opin. Struct. Biol.* 2, 6–12.
- [2] Braunitzer, G. and Chen, R. (1972) *Hoppe-Seyler's Z. Physiol. Chem.* 353, 674–676.
- [3] Braunitzer, G., Chen, R., Schrank, B. and Stangl, A. (1972) *Hoppe-Seyler's Z. Physiol. Chem.* 353, 832–834.
- [4] Pervaiz, S. and Brew, K. (1985) *Science* 228, 335–337.
- [5] Godovac-Zimmerman, J. (1988) *Trends Biochem. Sci.* 13, 64–66.
- [6] Akerstroem, B. and Loegdberg, L. (1990) *Trends Biochem. Sci.* 15, 240–243.
- [7] Flower, D.R. (1995) *J. Mol. Recognition* 8, 185–195.
- [8] Hattori, M., Ametani, A., Katakura, Y., Shimizu, M. and Kaminogawa, S. (1993) *J. Biol. Chem.* 268, 22414–22419.
- [9] Papiz, M.Z., Sawyer, L., Eliopoulos, E.E., North, A.C.T., Findlay, J.B.C., Sivaprasadarao, R., Jones, T.A., Newcomes, M.E. and Kraulis, P.J. (1986) *Nature* 324, 383–385.
- [10] Monaco, H.L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L. and Eliopoulos, E.E. (1987) *J. Mol. Biol.* 197, 695–706.
- [11] Townend, R., Winterbottom, R.J. and Timasheff, S.N. (1960) *J. Am. Chem. Soc.* 82, 3161–3168.
- [12] Townend, R., Weinberger, L. and Timasheff, S.N. (1960) *J. Am. Chem. Soc.* 82, 3175–3179.
- [13] Townend, R. and Timasheff, S.N. (1960) *J. Am. Chem. Soc.* 82, 3168–3174.
- [14] Roefs, P.F.M. and De Kruijff, K.G. (1994) *Eur. J. Biochem.* 226, 883–889.
- [15] Bairoch, A. and Boeckmann, B. (1994) *Nucleic Acids Res.* 22, 3578–3580.
- [16] Piez, K.A., Davie, E.W., Folk, J.E. and Gladner, J.A. (1961) *J. Biol. Chem.* 236, 2912–2916.
- [17] Pribic, R., van Stokkum, I.H.M., Chapman, D., Haris, P.I. and Bloemendal, M. (1993) *Anal. Biochem.*, 214, 366–378.
- [18] Li, H., Hardin, C.C. and Foegeding, E.A. (1994) *J. Agric. Food Chem.* 42, 2411–2420.
- [19] Cairoli, S., Iametti, S. and Bonomi, F. (1994) *J. Protein Chem.* 13, 347–354.
- [20] Lapanje, S. and Poklar, N. (1989) *Biophys. Chem.* 34, 155–162.
- [21] Albright, D.A. and Williams, J.W. (1968) *Biochemistry* 7, 67–78.
- [22] Esposito, G. and Pastore, A. (1988) *J. Magn. Reson.* 76, 331–336.
- [23] Wagner, G. (1993) *J. Biomol. NMR* 3, 375–385.
- [24] Wüthrich, K. (1986) in: *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
- [25] Englander, S.W. and Wand, A.J. (1987) *Biochemistry* 26, 5953–5958.
- [26] Katakura, Y., Totsuka, M., Ametani, A. and Kaminogawa, S. (1994) *Biochem. Biophys. Acta* 1207, 58–67.
- [27] Monaco, H.L. and Zanotti, G. (1992) *Biopolymers* 32, 457–465.
- [28] Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) *J. Mol. Biol.* 222, 311–333.
- [29] Chobert, J.M., Briand, L., Grinberg, V. and Haertlé, T. (1995) *Bioch. and Biophys. Acta* 170–176.
- [30] Fink, A.L., Calciano, L.J., Goto, Y., Kurotsu, T. and Palleros, D. (1994) *Biochemistry* 33, 12504–12511.
- [31] Ptitsyn, O.B. (1992) in: *Protein Folding*, (Creighton, T.E. ed.) pp. 243–300, W.H. Freeman and Co., New York.
- [32] Peng, Z. and Kim, P.S. (1994) *Biochemistry* 33, 136–141.
- [33] Redfield, C., Smith, R.A.G. and Dobson, C.M. (1994) *Nature Struct. Biol.* 1, 23–29.
- [34] Feng, Y., Sligar, S.G. and Wand, A.J. (1994) *Nature Struct. Biol.* 1, 30–35.
- [35] Van Mierlo, C.P.M., Darby, N.J., Keeler, J., Neuhaus, D. and Creighton, T.E. (1993) *J. Mol. Biol.* 229, 1125–1146.
- [36] Staley, J.P. and Kim, P.S. (1990) *Nature* 344, 685–688.
- [37] Nowak, U.K., Cooper, A., Saunders, D., Smith, R.A.G. and Dobson, C. (1994) *Biochemistry* 33, 2951–2960.