Structural Changes Accompanying pH-Induced Dissociation of the β-Lactoglobulin Dimer[†,‡]

Stanislava Uhrínová,[§] Mark H. Smith,[§,||] Geoffrey B. Jameson,[⊥] Dusan Uhrín,[§] Lindsay Sawyer,[@] and Paul N. Barlow*[,§]

The Edinburgh Centre for Protein Technology, Joseph Black Chemistry Building, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, Scotland, and The Institute of Cell and Molecular Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, Scotland

Received November 15, 1999; Revised Manuscript Received January 13, 2000

ABSTRACT: We have used NMR spectroscopy to determine the three-dimensional (3D) structure, and to characterize the backbone dynamics, of a recombinant version of bovine β-lactoglobulin (variant A) at pH 2.6, where the protein is a monomer. The structure of this low-pH form of β-lactoglobulin is very similar to that of a subunit within the dimer at pH 6.2. The root-mean-square deviation from the pH 6.2 (crystal) structure, calculated for backbone atoms of residues 6−160, is ∼1.3 Å. Differences arise from the orientation, with respect to the calyx, of the A−B and C−D loops, and of the flanking three-turn α-helix. The hydrophobic cavity within the calyx is retained at low pH. The E−F loop (residues 85−90), which moves to occlude the opening of the cavity over the pH range 7.2−6.2, is in the "closed" position at pH 2.6, and the side chain of Glu89 is buried. We also carried out measurements of ^{15}N T_1s and T_2s and $^{1}H-^{15}N$ heteronuclear NOEs at pH 2.6 and 37 °C. Although the residues of the E−F loop (residues 86−89) have the highest crystallographic *B*-factors, the conformation of this loop is reasonably well defined by the NMR data, and its backbone is not especially mobile on the pico- to nanosecond time scale. Several residues (Ser21, Lys60, Ala67, Leu87, and Glu112) exhibit large ratios of T_1 to T_2, consistent with conformational exchange on a micro- to millisecond time scale. The positions of these residues in the 3D structure of β-lactoglobulin are consistent with a role in modulating access to the hydrophobic cavity.

β-Lactoglobulin (β-LG)[1] is a small (162 amino acid residues, 18 350 Da), soluble, acid-stable protein which is found in copious quantities in the whey fraction of the milk of ruminants and other species (*1*), but not in human milk. Many variants of bovine β-LG have been identified, the most common of which, termed A and B, differ by two amino acids (*2*). Despite a very wide range of physical studies performed upon β-LG over the last five decades (*3*, *4*) and the recent availability of several high-resolution crystal structures (*5*−*9*), its properties in solution are not fully understood. Of interest in this context are the stability and aggregation properties of the bovine variants of β-LG under various conditions. These affect the properties of dairy milk during processing (*10*). Furthermore, it is emerging that a putative function of β-LG as a transporter of small hydrophobic molecules (*11*) within the digestive system may depend on dynamic features of the protein (*7*). Intermolecular contacts and the mobilities of loops and turns differ according to the crystal form, and there is a clear need for solution structure information.

The structure determination of the triclinic crystal form of an A/B mixture of β-LG, at pH 6.7 (*5*), confirmed that the protein is a dimer, and has the typical lipocalin topology seen in other proteins such as odorant binding protein (*12*), epididymal retinoic acid binding protein (*13*), retinol binding protein (RBP) (*14*), bilin-binding protein (*15*), and major urinary protein from mouse (*16*). The structure of β-LG contains an eight-stranded, flattened β-barrel or calyx and a flanking three-turn α-helix (*5*). One curved face of the calyx is formed from a twisted antiparallel β-sheet containing the N-terminal half of strand A [i.e., $A^{(1)}$], together with strands B−D. The other curved face is a similarly twisted antiparallel β-sheet containing strands E−H and the C-terminal half of strand A [i.e., $A^{(2)}$]. Strand A (residues 16−27) has a 90° bend at its midpoint, Ser21, reflecting its participation in both sheets. A ninth strand (I) extends the $EFGHA^{(2)}$ sheet. A short 3_{10}-helix precedes strand A, and a second 3_{10}-helical turn lies in the long A−B loop (i.e., the loop connecting

[†] This work has been supported in part by a research grants from the New Zealand Foundation for Research, Science and Technology (Contracts DRI 403 and 801). Support was also provided by the U.K. Biotechnology and Biological Sciences Research Council (Grant D09446), the Department of Trade and Industry through a Foresight Challenge initiative, and the Wellcome Trust through funding of the 600 MHz spectrometer.

[‡] Atomic coordinates have been submitted to the Protein Data Bank as file name 1DV9.

* Corresponding author. Phone: 44-131-650-4727. Fax: 44-131-650-4717. E-mail: barlow@chem.ed.ac.uk.

[§] The Edinburgh Centre for Protein Technology, Joseph Black Chemistry Building, University of Edinburgh.

[||] Current address: Food Science Section, New Zealand Dairy Research Institute, Private Bag 11029, Palmerston North, New Zealand.

[⊥] Institute of Fundamental Sciences-Chemistry, Massey University, Palmerston North, New Zealand.

[@] The Institute of Cell and Molecular Biology, University of Edinburgh.

[1] Abbreviations: β-LG, β-lactoglobulin; RBP, retinol binding protein; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; HSQC, heteronuclear single-quantum coherence; FID, free induction decay; SD, standard deviation; rmsd, root-mean-square deviation; ANS, 8-anilino-1-naphthalenesulfonate.

10.1021/bi992629o CCC: $19.00
Published on Web 03/04/2000

strand A to strand B) which forms part of the dimer interface along with strand I. Short 3_{10}-helical turns are also found between strands G and H, and within the C-terminal region. The α-helix lies in the sequence between strands H and I. A disulfide bond (Cys66−Cys160) connects the C−D loop to the carboxyl-terminal region, and a second disulfide bond (Cys106−Cys119) links strands G and H. Cys121 is buried and does not normally participate in a disulfide linkage in the mature protein.

Subsequently, studies of the purified A and B variants of bovine β-LG revealed subtle differences in their tertiary structure (*9*). Working with trigonal crystal forms, Qin et al. showed that the B variant has a less well packed core than the A variant, arising from the V118A substitution. The other substitution (D64G) causes a significant difference in the main chain conformation in the C−D loop.

β-Lactoglobulin undergoes changes of physical, chemical, and spectroscopic properties over a pH range of 6−8 (*17*). The structural basis of this so-called Tanford transition (*18*) was determined by X-ray crystallographic studies at several pH values (*8*). At basic pH, the open end of the calyx provides an access route to a cavity at its center which is a binding site for small hydrophobic molecules; this may be considered the "open" conformation. At acidic pH, the E−F loop (residues 86−89) is arranged so as to occlude the open end of the calyx, giving the "closed" conformation. The physiological significance of this may lie in the putative role of β-LG as a transporter of small molecule ligands. In the closed conformation, bound ligand might be protected in the acidic stomach. Ligand could then be released upon adoption of the open conformation of β-LG within the basic small intestine.

Also of fundamental interest is a putative α → β transition which β-LG is reported to undergo during the folding process (*19*). This is purported to be a consequence of an inherent propensity for helix formation within the β-LG sequence as a consequence of local interactions. Nonlocal interactions favor the predominant β-structure observed under native conditions. Thus, β-LG may serve as a model for the α → β conformational change associated with prion disease, or with Alzheimer's disease, where the α-helical forms are normal and the β-sheet forms are amyloidogenic (*20*).

In solution, when the pH is reduced from 6.2 to 2.6, bovine β-LG undergoes a transition from a dimer to a monomer under low-salt conditions (*21*). Given its relatively small size, the monomeric form of the protein is accessible to modern multidimensional protein NMR spectrometry which is able to deliver a high level of structural and dynamical detail. Early studies utilized ^{1}H NMR to show that at pH 2.6, much of the secondary structure observed in the dimer already exists (*22*). Furthermore, much of the hydrophobic core appears intact (*23*). Binding studies with ANS on the other hand implied some partial unfolding of the protein under these conditions (*24*). Two subsequent studies utilizing isotopically labeled protein led to complete assignment of backbone and side chains (*19*, *25*). Using the chemical shift index, and a subset of readily assigned NOEs, these studies further demonstrated that the monomeric protein has a three-dimensional structure very similar to the dimeric one (*25*).

We now present the three-dimensional structure of bovine β-LG variant A at pH 2.6 and 37 °C. We have also performed measurements of its ^{15}N spin−lattice and spin−spin relaxation rates, and ^{1}H−^{15}N NOEs, and interpreted these in terms of the backbone dynamics of the protein. These data are compared with the various crystal structures of β-LG.

MATERIALS AND METHODS

Protein Expression and Purification. The bovine β-LG A variant was expressed as a recombinant protein from a genetically modified strain of *Pichia pastoris* (kindly supplied by C. Batt, Cornell University, Ithaca, NY) as described by Kim et al. (*26*). The preparation and purification of ^{15}N- and ^{15}N- and ^{13}C-labeled protein samples and preparation for NMR is fully described by Denton et al. (*27*) and Uhrinova et al. (*25*). The protein used in this study differs from wild-type β-LG variant A at the N-terminus where L1 and I2 have been replaced by E−3, A−2, E−1, A1, and Y2. The ^{15}N−^{1}H HSQC spectra of the expressed protein and of the wild-type protein (based on the natural abundance of ^{15}N) overlay sufficiently well to suggest that the substitution does not alter the structure of β-LG A (see ref *25*).

For structure determination, the NMR sample contained 1 mM protein in 50 mM potassium phosphate adjusted to pH 2.6, in 550 μL of $H_2O/^2H_2O$ (90/10 v/v). Under these conditions, β-LG is a monomeric protein. For investigation of dynamics, data were obtained using a sample containing 0.6 mM protein in 50 mM potassium phosphate adjusted to pH 2.6, in 300 μL of $H_2O/^2H_2O$ (90/10 v/v), in a Shigemi NMR tube (Shigemi Inc.). The measured value of pH was not corrected for isotope effects. All spectra were recorded at 37 °C on a Varian (Palo Alto, CA) "INOVA 600" spectrometer in a 5 mm, triple-resonance probe, equipped with triple-axis pulsed field gradients.

NMR Spectroscopy. NOE connectivities were established from a three-dimensional (3D) ^{15}N-edited NOESY-HSQC spectrum (*28*, *29*) with a mixing time of 70 ms, and a 3D ^{13}C- and ^{15}N-edited NOESY-HSQC spectrum with a mixing time of 80 ms. The spectral widths of the former experiment were 5000 Hz in F3 (N^1H), 2015 Hz in F2 (^{15}N), and 7200 Hz in F1 (^{1}H). The size of the acquired matrix was 512 × 48 × 96. The spectral widths for the 3D ^{13}C- and ^{15}N-edited NOESY-HSQC experiments were 8000 Hz in F3 (N^1H), 3017 Hz in F2 (^{13}C and ^{15}N), and 7200 Hz in F1 (^{1}H). The size of the acquired matrix was 1024 × 96 × 32 complex data points. The ^{1}H carrier was placed on the H_2O resonance at 4.644 ppm or in the middle of the NH region; the ^{15}N carrier was placed at 117 ppm, and the ^{13}C carrier at 82 ppm, utilizing multiple folding. The $J_{HNH\alpha}$ coupling constants were determined from an HNHA experiment (*30*). The spectral widths were the same as in the 3D ^{15}N-edited NOESY experiment with the exception of the acquisition dimension, which was set to 8000 Hz. The indirectly detected dimensions were referenced as described by Live et al. (*31*), and Bax and Subramanian (*32*). Data sets were processed using home-written macros within the package FELIX95 (Biosym Technologies/Molecular Simulations, San Diego, CA). The FIDs were zero-filled prior to Fourier transformation. Final 3D matrixes typically consisted of 1024 × 128 × 128 real data points.

The T_1 and T_2 ^{15}N relaxation times were measured according to the method of Kay et al. (*33*) with water flip-back pulses inserted when the magnetization was in the zz state. Soft pulse WATERGATE was used for water sup-

pression (*34*). The $^1H-^{15}N$ steady-state NOEs were recorded keeping the water along the *z*-axis (*35*). Proton and nitrogen sweep widths were 10 000 and 2015 Hz, respectively; the number of complex data points in the nitrogen dimension was 128 in all experiments. Spectra were recorded using 64 scans per increment, giving a total experimental acquisition time for each data point of 10 h for T_1 determinations and 7 h for T_2 determinations. Sampling of the exponential decay of the ^{15}N magnetization in T_1 and T_2 experiments was optimized according to Hore (*36*). Three data points were acquired. For measurements of T_1, relaxation delays of 6.25 ms (one data point) and 960.5 ms (two data points) were used; relaxation delays of 16.26 ms (one data point) and 162.56 ms (two data points) were used in the T_2 measurements. A reference NOE experiment was recorded with a 5 s relaxation delay. A second NOE spectrum was recorded with saturation of the ^1H signal for the last 3 s of the 5 s delay. Both spectra were recorded using 32 scans per increment, giving a total acquisition time of 12 h per increment.

NOE Assignments and Structure Calculations. A total of 1405 cross-peaks from the two NOESY spectra were assigned and integrated in the XEASY suite of software (*37*). These, along with a list of picked, integrated, cross-peaks and a table of chemical shifts, were used as input for the ARIA extension (*38*) of XPLOR (*39*). Cross-peak volumes were classified into categories as follows: strong (corresponding to interproton distances of <2.7 Å), medium (<3.4 Å), weak (<4.2 Å), or very weak (<5 Å). ARIA is able to deal with multiple assignment possibilities for cross-peaks by consideration of the contribution (C_k; see ref *38*) of each possible assignment. For a set of possible NOE assignments, values of C_k are summed starting with the highest value. Once the sum exceeds the "cutoff" (ρ), no more possibilities are included. After eight cycles of structure calculations (*38*), ARIA generated a total of 2752 nondegenerate NOEs that were unambiguously assigned to pairs of protons, using a final value for ρ of 0.8. These were subsequently subjected to a process of manual editing by a careful re-examination of the spectra, resulting in the net elimination of 224 NOEs. Cross-peaks were removed from the table, or re-assigned, on the grounds that (i) they were unreliable due to overlap, (ii) they were weak and could not be confirmed by the presence of a symmetry-related cross-peak or (iii) despite falling within the margin of error supplied to ARIA for the purposes of automated assignment (0.03 ppm in ^1H, 0.03 ppm in indirectly detected ^1H, and 0.3 ppm in ^{15}N and ^{13}C), they did not align within the digital resolution of the spectra with the other cross-peaks assigned to the same proton. The edited list of 2528 cross-peaks was then used as input for a simulated annealing protocol within the program "Crystallography and NMR systems" (*40*) that had been modified to include provision for floating stereochemistry at, and to include active swapping of, prochiral centers in random order using a metropolis style acceptance criterion (*41*). Also used as input were 110 $J_{HNH\alpha}$ coupling constants derived from an HNHA experiment, 62 distance restraints for sets of atoms implicated in H-bonds on the basis of slowly exchanging (>1 h) amide data and supporting NOEs, and a further 33 ambiguous assignments where the H-bonding partner of a slowly exchanging (>1 h) amide was not immediately identifiable. In 15 of these, a further round of structure calculations allowed identification of the likely partners and the appropriate H-bonds were entered as unambiguous restraints in the final structure calculation (the remaining 18 slowly exchanging amides were not included in the final calculation). A total of 50 structures were calculated, out of which 21 were selected on the basis of lowest total potential energy.

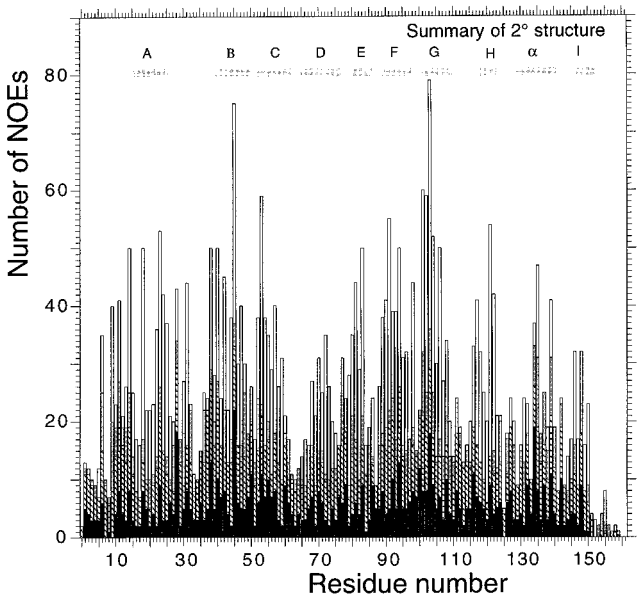

FIGURE 1: Distribution of NOEs used in the structure calculation as a function of residue number. Black bars represent intraresidue NOEs, hatched bars sequential NOEs, gray stippled bars medium-range NOEs [$i \rightarrow i + (2-4)$], and white bars long-range NOEs.

RESULTS

Structure Calculations. A plot of the number of NOEs identified per residue as a function of the position of the residue within the sequence exhibits a series of peaks and troughs (Figure 1). Peaks correspond approximately to the middle residues of the eight β-strands A−H as identified in crystal structures of the dimer. These comprise the core β-barrel or calyx feature. The troughs correlate with the positions of loops and turns. There are few detectable long-range NOEs which involve the five amino-terminal residues, or the four carboxy-terminal residues. The average number of long-range NOEs per residue excluding these nine residues is 6.8. The region of the protein best defined by NOE-derived distance restraints is β-strand G (residues 101−109) with an average of 23.1 long-range restraints per residue. The sequence of the α-helix exhibits a ~3.5 residue periodicity in the pattern of long-range NOEs, with significant numbers of connectivities assigned to Asp129, Ala132 and Ala133, Leu135 and Phe136, and Ala139 and Leu140. For 48 of the β-strand residues and 15 of the helix residues, the H-bond partner was identified on the basis of supporting NOEs or by preliminary structure calculations. All of the H-bonding potential of strands C, F, G, and H is exploited, while strand D participates in only six H-bonds. Of the 110 residues for which three-bond $J_{HNH\alpha}$ couplings could be measured reliably, 68 lie in the β-strands and 12 within the α-helix. The average value of $J_{HNH\alpha}$ for strand residues was 7.7 Hz, while for helix residues it was 4.8 Hz. The average value of this coupling constant for the remaining residues was 6.5 Hz.

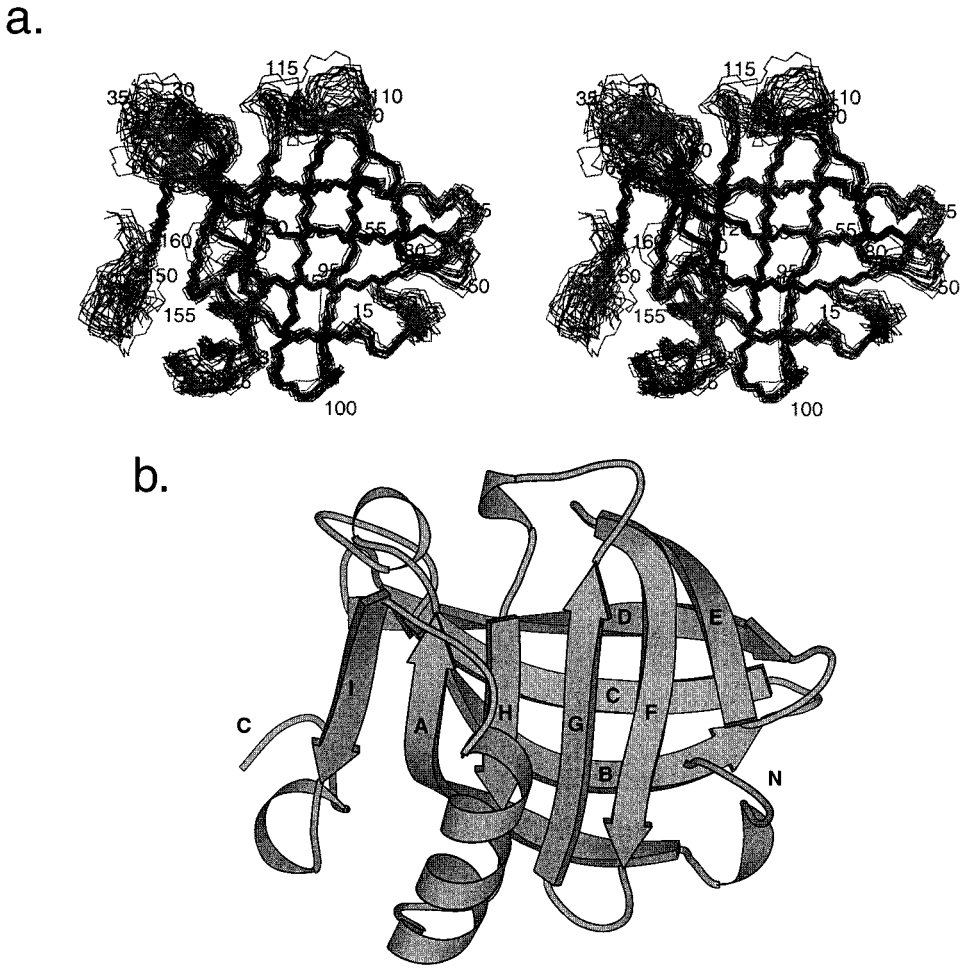

FIGURE 2: Structure of β-LG at pH 2.6. (a) An overlay, on the coordinate-averaged structure (using the Cα atoms of residues of strands A−H and the long α-helix), of the backbones of 21 structures calculated on the basis of the NMR-derived distance and angle restraints. Residues 1−5, 161, and 162, which are not structured, have been omitted from the figure for clarity. The structure is shown as a stereopair. Numerals refer to residue sequence numbers. (b) A cartoon representation of the structure (residues 6−160), in the same perspective, generated using Molscript (*47*).

The solution structure of β-LG at low pH and in a momomeric form has nine β-strands and an α-helix [as anticipated from the chemical shift index (*24*)]. These secondary structural elements are arranged into a tertiary structure that is similar to that of a subunit of β-LG in the dimeric form. Figure 2a shows a superposition (on the mean structure) of the backbone traces of 21 calculated structures of bovine β-LG variant A. Figure 2b is a schematic representation of the arrangement of secondary structure elements in the NMR-derived structure. Figure 3a is a plot of sequence number versus the rmsd of the Cα from the average structure. Table 1 summarizes some structural statistics.

Residues toward the amino and carboxy termini are not well-defined by the NMR data (Figures 2a and 3a) as expected from the lack of long-range NOEs. Between residues 6 and 160 (average rmsd for backbone atoms = 0.84 Å), the least well-defined regions (consecutive residues with an rmsd for Cα of >1.0 Å) involve residues 34−36 within the A−B loop, residues 49−52 in the B−C loop, residues 62−65 in the C−D loop, residues 125−127 just before the long α-helix, residues 142−144 after the helix, and residues 157−160. The best defined regions are the β-strands where consecutive residues occur with rmsds from the average structure in the region of 0.5 Å. Strand G is the best defined of the strands, and strand I is the least well-defined. The D−E, E−F, and F−G loops converge quite well and contain only isolated residues for which the rmsd exceeds 1 Å. Within this global superposition, the residues within the α-helix [as defined according to Qin et al. (*7*) as residues 129−141] converge with a mean rmsd (Cα atoms) of 0.66 Å. Overall, the left-hand side of the protein as viewed in Figure 2a with the open end of the calyx at the top, consisting of the C-terminus, the A−B and C−D loops, and strand I, appears less well converged than the right-hand side.

Analysis of Relaxation Experiments. Values of ^{15}N T_1 and T_2 and the steady-state ^{1}H−^{15}N NOE enhancement for each amide in a protein may be interpreted to furnish information about the motion of the backbone (*42*). T_1 values probe motion in the range of 10^8−10^{12} s^{-1}; T_2 values are sensitive to these fast motions in addition to dynamics on the micro- to millisecond time scale. The NOE enhancement is more sensitive to molecular motion on a nanosecond time scale than either T_1 or T_2. In general, rapidly fluctuating, highly mobile regions of the protein are associated with small (<0.7) or negative NOE enhancements, and larger than average values of T_1 and T_2. If T_1/T_2 exceeds the average by more than a standard deviation (i.e., the value is >8.0 in the case of β-LG) for a particular amide, this probably reflects a contribution to the line width by conformational exchange.

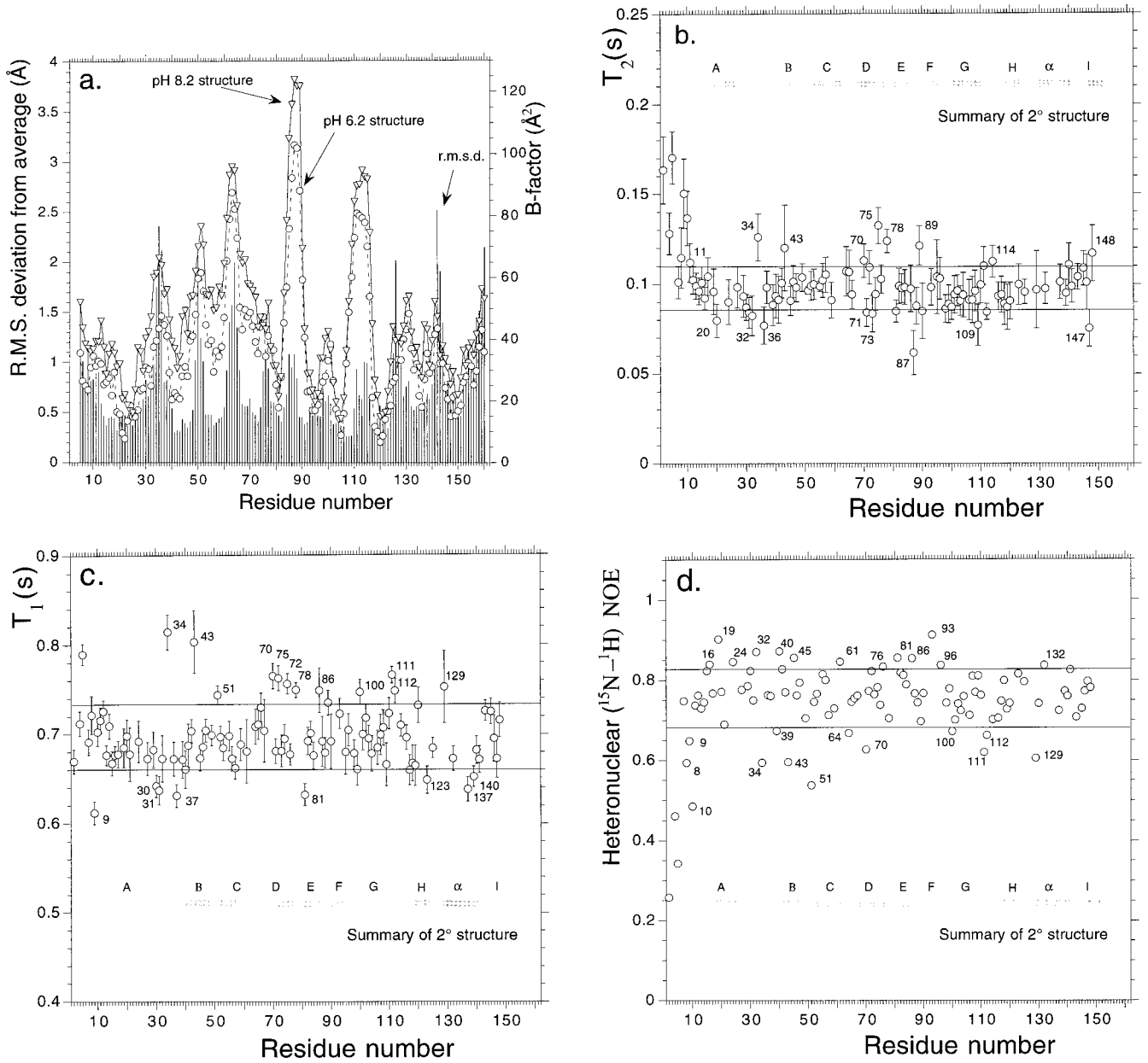

FIGURE 3: Definition and mobility of the solution structure of β-LG at 37 °C. (a) A plot (thin black bars) of the rmsd from the mean structure of Cα atoms among the ensemble of 21 calculated structures (for residues 6−160). Structures were superimposed using the backbone atoms of residues 6−160 on the coordinate-averaged structure prior to the calculation. Also shown are the crystallographic *B*-factors for Z lattice crystal structures at pH 6.2 (▽) and 8.2 (○). (b) A plot of the measured ^{15}N spin−lattice (T_1) relaxation times vs residue number. (c) A plot of ^{15}N spin−spin (T_2) relaxation times vs residue number. (d) A plot of heteronuclear (^{1}H−^{15}N) NOE enhancements as a function of residue number. The relative error in the NOE data, as determined from the ratio of the peak height and the background noise in saturated and reference spectra, was 5−10% depending on cross-peak intensity. In plots b−d, the positions of the nine β-strands (labeled A−I) and α-helix (defined according to ref *8*) are indicated. Parameters for ∼100 residues are shown. Horizontal lines indicate values corresponding to the mean ± one SD (calculated for residues 6−160), and residues lying outside of these limits are labeled with their residue sequence numbers.

More sophisticated analyses based on the derivation of order parameters (S^2) are available (*43*). Our data are not of sufficient quality to allow such a treatment for the majority of amides, but selected values, calculated using the spectral density function of Lipari and Szabo (*43*), are sufficient for the semiquantitative analysis presented here. The existence of very slowly exchanging amide protons involved in H-bonds also yields insight into the rigidity of a protein, and in the case of β-LG at pH 2.6, the amide protons of 70 residues were still detectable after 2 weeks.

The relaxation parameters of only 100 out of 162 residues in β-LG were considered reliable for analysis. The HSQC cross-peaks of other residues were too overlapped to be of use. Transverse relaxation times (T_2) were in the range of 0.062−0.170 s (average of 0.104 s), while longitudinal relaxation times (T_1) were in the range of 0.61−0.81 s (average of 0.70 s). Three residues (Ser21, Trp61, and Ala67) exhibited no signal in the T_2 spectrum acquired with a relaxation delay of 16.26 ms. This is presumably a consequence of a chemical exchange contribution to relaxation, and these residues might be in conformers exchanging on a time scale comparable with the difference in their chemical shifts. Ser21 lies in a critical position in the 90° bend at the midpoint of strand A [i.e., between A$^{(1)}$ and A$^{(2)}$], and

Table 1: NMR Structure Determination Statistics

total no. of NOEs used for CNS calculation	2528	
no. of intraresidue NOEs	844	
no. of sequential NOEs	597	
no. of short-range NOEs ($\|i - j\| < 5$)	63	
no. of long-range NOEs ($\|i - j\| > 4$)	1024	
no. of torsion angles ϕ	110	
no. of hydrogen bonds	77	
structure calculation statistics (21 structures out of 50)		
maximum violation (Å)	0.46	
no. of NOE violations[a] of >0.3 Å	3	
no. of torsion angle violations[a] of >1.5 Hz	1	
Procheck statistics		
residues in most favored region (%)	65.1	
residues in additional allowed regions (%)	27.5	
residues in generously allowed regions (%)	4.7	
residues in disallowed regions (%)	2.7	
rms deviations from the experimental restraints[b]		
NOEs (Å)	0.034 ± 0.009	
J restraints (deg)	0.79 ± 0.05	
rms deviation from idealized geometry[b]		
bond lengths (Å)	0.0030 ± 0.0001	
bond angles (deg)	0.36 ± 0.02	
rms deviation from the average structure (Å)		
backbone atoms (N, C^{α}, C′, and O) 5−160	0.845	
excluding loops of residues 34−38, 61−65, 111−115	0.767	
sheets only	0.520	
comparison[c] of NMR structure with X-ray		
residues 5−160	1.678	1.310
excluding loops of residues 34−38, 61−65, 111−115	1.710	1.248
sheets only	0.812	0.778

[a] The number given is the accumulated total for all 21 structures in the ensemble. [b] Number is mean ± SD. [c] Numbers are average deviations (angstroms) for backbone atoms N, C^{α}, C′, and O at pH 8.2 and 6.2 (columns 2 and 3, respectively).

conformational exchange here could conceivably result in a slight expansion or contraction of the calyx. Residues 61 and 67 share an H-bond, and conformational exchange of these two residues might indicate slow segmental motion of the C−D loop. Other residues that have high T_1/T_2 ratios (>8.0) are (in ascending order) Gln120, Ile71, Asn90, Ala73, Leu32, Lys100, Asn109, Ser36, Tyr20, Glu112, Ile147, and Leu87.

The overall rotational correlation time τ_c was calculated (*42*) from the ratio of longitudinal and transverse relaxation times. Twelve residues for which the heteronuclear NOE < 0.65 were excluded from these calculations together with residues which exhibited larger deviations from average T_2 and T_1 values according to a criterion proposed by Tjandra (*44*). This left 53 residues, out of the 100 analyzed, for inclusion in the calculation of τ_c. The calculated value of τ_c is 7.30 ± 0.05 ns which is in line with those of proteins of a comparable size and shape (*42*, *45*, *46*).

The $^{15}N-^{1}H$ heteronuclear NOEs (Figure 3d) were in the region of 0.26−0.912 (mean of 0.74, median of 0.76). With the exception of the N-terminus, the measured NOE enhancements fall within a relatively narrow range; thus, after Leu10, only 10 residues have NOE values that are more than one standard deviation (SD) below the mean (calculated on the basis of residue 6 onward). Of the 10 residues with low NOE enhancements, eight have relatively high T_1 values (Ala34, Val43, Glu51, Lys70, Lys100, Ala112, Glu112, and Asp129 all more than one SD higher than the mean), underlining their flexible nature on the rapid time scale. Ala34, Val43, and Lys70 also possess high T_2 values (>0.11 s) and probably therefore have the most mobile backbones apart from the N-terminal residues. The S^2 values for these three residues were calculated to be 0.69, 0.70, and 0.76, respectively. Of the residues with low NOE enhancements and high T_1 values, six are in loops, but two of these mobile residues (Val43 and Lys70) lie in β-strands B and D, respectively. In strand D, Ile72 and Lys75 also appear mobile as evidenced by relatively high T_1 values. A feature shared by these four mobile strand residues is that they do not participate in intrastrand H-bonds.

The 13 residues with NOE enhancements more than one SD above the mean mostly lie within β-strand or α-helix, and most have hydrophobic side chains that contribute to the hydrophobic core. The exception is Leu32 which is in an H-bond-stabilized helical portion of the A−B loop. The residues with large NOE enhancements have lower than average T_1s, but unexceptional T_2s, apart from Val81 (T_2 = 0.085) in strand E which probably corresponds to one of the most rigid parts of the backbone and has a calculated S^2 of 0.99. None of the five cysteine residues within β-LG appear to be in especially rigid parts of the backbone.

Comparison of NMR-Derived and Crystal Forms of the Structure. Panel b of Figure 4 shows an overlay of the ensemble of NMR-derived structures superimposed on ribbon representations of a subunit (from the Z lattice crystal structure) of dimeric β-LG at pH 6.2 and 8.2. Figure 4a shows the deviation in the position of the Cα atoms of the coordinate-averaged NMR structure from their equivalents in the X-ray structures at pH 6.2 (PDB file 3BLG) and 8.2 (PDB file 2BLG) following a (pairwise) superimposition of NMR and X-ray-derived structures using the backbone atoms of residues 5−160. The average deviations (backbone atoms) are 1.35 and 1.66 Å for the pH 6.2 and 8.2 structures, respectively.

(1) N-Terminal Region. Among the 10 N-terminal residues, Met7 is unusual in having an NOE of >0.7 and an average value of T_2; the others appear disordered as judged by small NOE enhancements and relatively large T_2 values. This is consistent with the lack of convergence for this region in the structure calculation and the lack of long-range NOEs involving these residues. It is noteworthy that crystal forms of dimeric β-LG in lattice X and in lattice Y also exhibit a lack of definition in this region (*5*, *6*). In contrast, in the lattice Z crystal structure, the N-terminus is defined and residues Thr4−Met7 adopt a β-strand conformation with a single H-bond from Met7 to Val94 of strand E (*8*). In solution, Met7 has an amide proton which does not exchange with solvent during the course of 1 h, but is no longer detectable after 9 h, indicating a weak H-bond. The H-bond partner was not identified directly; however, NOE evidence supported Val94 as a candidate, and an H-bond between the NH of Met7 and the CO of residue 94 was included in the final calculation of the structure. In the lattice Z crystal structure, the α-amino group of Leu1 forms a salt bond with Glu108 in strand G. Since the recombinant protein used in the current study has extra residues at the N-terminus, no such salt bridge could be formed.

Residues 11−16 exhibit reasonably uniform values of relaxation parameters and are about as rigid as the β-strands. Residues 14−16 all have very slowly exchanging amides. These form a helical structure in solution. In the lattice X

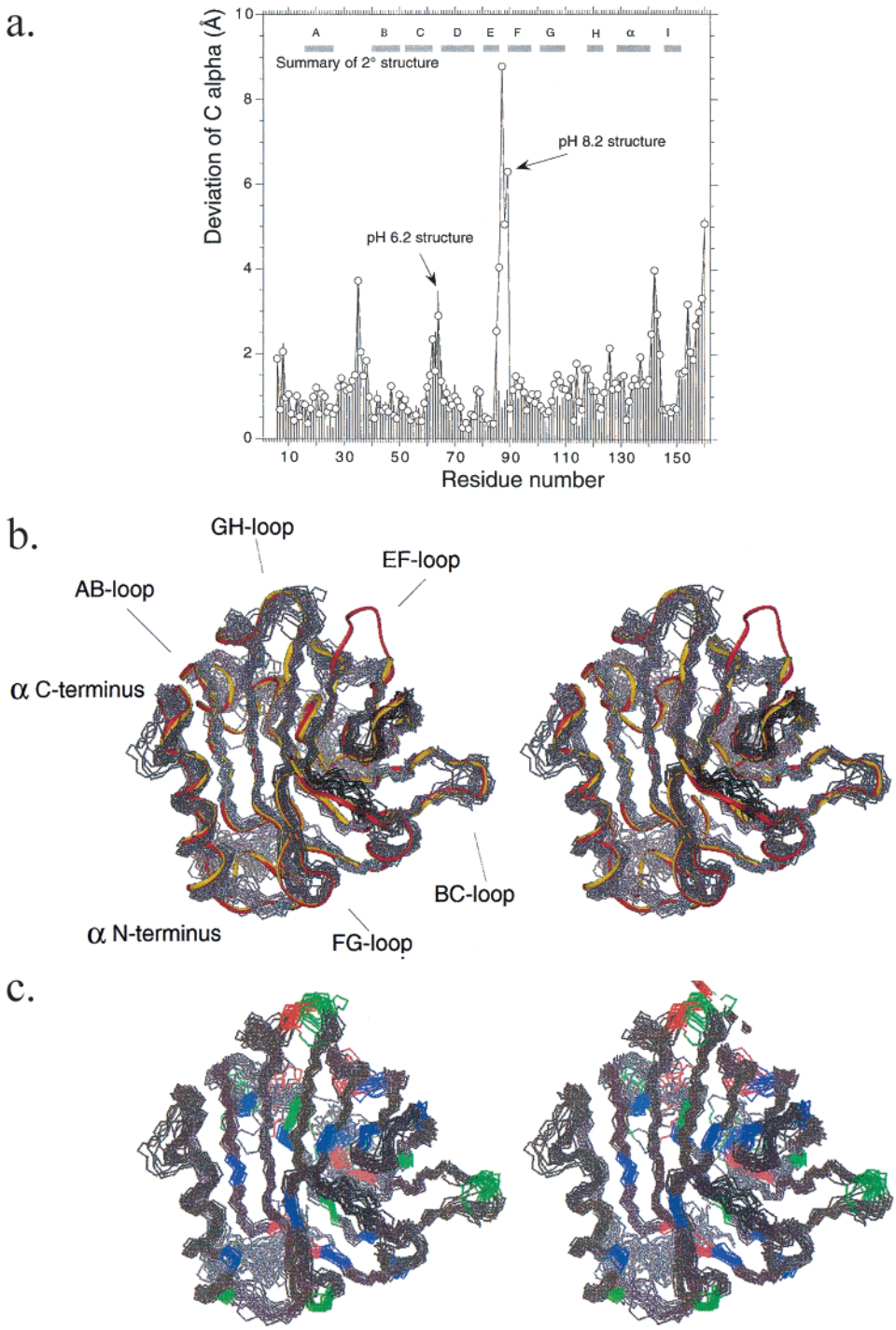

FIGURE 4: Comparison of NMR-derived and crystal (Z lattice) structures of β-LG. (a) A plot to show the deviation in the positions of Cα atoms between the NMR and pH 6.2 (thin black bars) and 8.2 (○) X-ray structures, as a function of residue numbers. The NMR structure was superimposed on the backbone atoms of residues 6−160 in the X-ray structure prior to the calculation. (b) NMR-derived ensemble of structures, shown as backbone traces (black), superimposed (residues 6−160, Cα atoms) on the pH 6.2 crystal structure (yellow backbone ribbon) and the pH 8.2 crystal structure (red backbone ribbon). The molecule has been rotated about the *y*-axis with respect to its orientation in Figure 2b so as to reveal more clearly the E−F loop. For clarification, several loops and the ends of the α-helix are labeled. (c) NMR-derived structures (backbone traces) in approximately the same orientation as in panel b. Residues with high heteronuclear NOE enhancements (i.e., relatively immobile residues on the pico- to nanosecond time scale, as defined in Figure 3d) are colored blue, while residues with low NOE enhancements are colored green. Residues implicated in conformational exchange (micro- to millisecond time scale) are shown in red.

and Z crystal structures, these residues form a 3_{10}-helix with low *B*-factors (*5*, *8*).

(2) *β-Strands, α-Helix, and the Core.* Analysis of chemical shifts for β-LG and the use of the chemical shift index had suggested disparities between solution and crystal structures in the extent of β-strands and α-helix (*24*). These are not borne out by the full structure determination. On the basis of a global superposition (Figure 4a), it can be seen that residues in strands A−E and I of both X-ray structures overlay well with equivalent residues in the solution structure (mean deviation for Cα atoms = 0.88 Å vs the pH 6.2 structure). Strands F−H and the long α-helix overlay less

well (mean deviation for Cα atoms = 1.1 Å vs the pH 6.2 structure). However, a local superimposition of the long α-helices of the X-ray and NMR-derived structures exhibits close agreement (0.44 Å for backbone heavy atoms, residues 129−141). There are significant differences in the path traced by the backbone at either end of the long α-helix. The segment joining the C-terminal end of the helix to strand I is not well defined by the NMR data. A distinct set of minor peaks corresponding to an alternative conformation of residues 138−143 were detected in the NMR spectra (*24*). Residues 143, 145, and 146 all have well above average T_2s (no data for residue 142; residue 144 is a proline), indicative of mobility, and Ile147 has a high T_1/T_2 and is probably in conformational exchange. However, the equivalent region of the crystal structure does not exhibit exceptionally high *B*-factors. The result of these differences in conformation is that the helix appears to adopt a different orientation with respect to the calyx, exhibiting a slightly greater tilt of its C-terminal end away from the N-terminal end of strand H.

With the exception of residues in strands B, D, and I whose H-bonding potential is not fully exploited, residues within the β-strands have relatively uniform values for their relaxation parameters and appear not to be mobile. Val43 in strand B is particularly mobile. It does not participate in interstrand H-bonds unlike residues 40 and 45 which are immobile. The backbone of strand D exhibit similar dynamic properties since ^{15}N relaxation parameters for residues 70, 72, and 75 are consistent with flexibility, while the parameters for residues 71, 73, and 74 indicate a lack of mobility. None of these three residues, Lys70, Ile72, and Lys75, are involved in H-bonds, while Ile71, Ala73, and Glu74 participate in two H-bonds each. This is a result of the exposed position of strand D at the open end of the calyx. Within short strand I, residues are more flexible than was observed for most of the strand residues in the calyx, but less flexible than residues within the longer loops. Residues of the three-turn α-helix (except residue 129 at the N-terminus) have average to above average NOE values, and shorter than average T_1s, so may be regarded as relatively immobile.

The cavity within the calyx, which is lined with hydrophobic side chains and is a binding site for fatty acids, is retained unaltered at pH 2.6. As at pH 6.2, Tyr42 is buried and unavailable for chemical modification. Tyr102 is almost completely buried, while Tyr20 and Tyr99 are substantially solvent-exposed. Cys121 is also buried, as is the Cys106−Cys119 disulfide. The Cys106 Cα−Cys119 Cα distance is 4.3 Å (cf. 3.9 Å in the crystal structure). The other disulfide, Cys66−Cys160, is largely exposed, and the Cys66 Cα−Cys160 Cα distance is 6.6 Å (cf. 5.9 Å in the crystal structure).

(3) Loops, Turns, and Connecting Segments. The occurrence of the short α-helix (*29*−*32*) beyond strand A is probably responsible for the relatively rigid nature of this segment. However, some of the residues (residues 34 and 39) connecting this helix to strand B are among the most flexible in the protein. The conformation of this loop differs between the NMR and crystal structures. It appears to be slightly less closely associated with the body of the protein in solution. Residue 35 exhibits a deviation of nearly 4 Å (Cα), but lies in a poorly converged region of the solution structure. Residues 34−36 do not have large crystallographic *B*-factors compared to other loops. At the pH used for crystallization (pH 6.2−8.2), residues 29−35 are involved in forming the dimer. Overall, it may be concluded that dimerization is associated with a distinct, relatively well-ordered, conformation of this loop, but that the loop is disordered at pH 2.6.

The average conformation of the B−C loop (residues 50−52) at pH 2.6 in solution is close to its conformation at higher pH values, despite the poor convergence in the structure calculation, and the high mobility of Glu51. Crystallographic *B*-factors in this region are similar to the A−B loop. It converges poorly in the NMR-based structure calculations, and closer inspection reveals that it exhibits two conformations; one is quite close to the crystal structure, and the other is flipped away from the body of the protein.

The C−D loop is likely to be undergoing conformational exchange on a slow time scale according to the large T_1/T_2 values of residues 60 and 67 (see above). The conformation of the short D−E loop does not differ markedly between solution and crystal structures, converges relatively well, and has low crystallographic *B*-factors and mobility in solution.

The E−F loop of the pH 2.6 structure overlays quite well with its equivalent in the structure at pH 6.2. It is in the closed position at pH 2.6 and has adopted a more flattened appearance than at pH 6.2 (see Figure 4b), forming a more intimate association with the rim of the aperture to the hydrophobic pocket. As in the crystal structure, the (protonated) side chain of Glu89 is buried and is within H-bonding distance of the carbonyl oxygen of Ser116. The E−F loop converges relatively well in the NMR calculation, and its backbone is not particularly mobile in solution on a nano- to picosecond time scale; yet residues in this region have the highest crystallographic *B*-factors in both open and closed conformations. It is possible that the closed conformation is more established at pH 2.6 than at pH 6.2; however, residue 87 appears to be undergoing conformational exchange according to the relaxation studies, so motion on an intermediate time scale could explain the *B*-factors for this residue.

The short F−G loop is a γ-turn, is relatively well defined by the NMR data, and is of average mobility for a turn and similar in conformation to the crystal structure. The backbone of residue 111 in the long G−H loop is flexible (small NOEs, long T_1 and T_2), while Glu112 has a high T_1/T_2 (8.8) and is probably in chemical exchange; this could explain the high crystallographic *B*-factor of this segment. The remaining residues of the G−H loop form an H-bond-stabilized helical structure very similar to the short helix observed in the crystal structure.

The segments of backbone on either side of the α-helix have already been discussed. The residues beyond strand I are poorly defined by the data and appear to be mobile (although relaxation data for this region are sparse) despite the disulfide connection to the C−D loop and the presence of two slowly exchanging amides which might be evidence for a helical turn as seen in the crystal structure. This region forms part of an intermolecular interface within the Z lattice and could be stabilized by crystal packing forces.

DISCUSSION

The three-dimensional structure of β-LG in solution, at an acidic pH similar to that in the stomach, resembles its

structure in a variety of crystal forms. The protein, however, is monomeric at pH 2.6 rather than dimeric. Significant differences in structure occur at the termini, but the recombinant form of β-LG used for the current study has a nonnative N-terminal sequence so any differences here might be artifactual; in addition, the C-terminus might be stabilized by crystal-packing forces. Differences in the conformation of the C−D loop could also be explained by crystal-packing interactions, although this appears to be a mobile region of the protein both in the crystal and in solution on several time scales. More significant might be changes in the conformation of the residues in the A−B loop involved in forming part of the dimer interface at higher pH. These are shifted by up to 3.5 Å compared to the crystal structure, and appear very mobile at low pH on the nano- to picosecond time scale.

Unexpectedly, the biggest structural difference lies away from the potential dimer interface and involves the orientation of the long α-helix with respect to the calyx. This surface of the protein exhibits patches of negative and positive charge at pH 6−8 (*8*). At pH 2.6, there will be a radical shift toward an overall positive charge and this could cause a disruption of tertiary as well as quaternary structure. The possibility that this difference is a consequence of the protein being in solution rather than in a crystalline form cannot however be excluded. The increased affinity of β-LG for ANS at pH 2−3 (*22*) is consistent with an improved solvent accessibility of the hydrophobic region under low-pH conditions. The bulk of the protein, however, is compactly folded at pH 2.6 as evidenced by its backbone ^{15}N relaxation characteristics and the persistence of 70 amide protons after incubation for 11 days in 99.9% D_2O.

As shown in Figure 4c, residues in the C−D, E−F, and G−H loops exhibit relaxation parameters that are consistent with conformational exchange on a milli- to microsecond time scale. Each of these loops is located at the open end of the calyx, and might be involved in modulating access to the hydrophobic cavity. The only other residue which has similar relaxation properties is Ser21. This residue occupies a key position between the two sheets [formed from strands $A^{(1)}$BCD and EFGH$A^{(2)}$] that form the calyx, and is part of the unusual 90° bend which allows strand A to participate in both sheets. Any small movement of the two sheets as might occur in a minor expansion or contraction of the hydrophobic cavity would probably involve changes in the conformation of Ser21. Therefore, it is possible that at pH 2.6 β-LG is undergoing conformational exchange between open and closed conformations, resulting in changes in the accessibility and capacity of the hydrophobic pocket. Judging by the conformation of the E−F loop in the NMR-derived structure, and assuming we have calculated an average of possible conformations, then the equilibrium is clearly biased toward the closed conformation at pH 2.6. Thus, this work is consistent with the hypothesis that β-LG encloses and protects certain hydrophobic molecules in the acidic stomach, but releases them upon reaching the basic small intestine.

Retinol binding protein, another lipocalin, has also been reported to undergo conformational exchange between open and closed conformations on this time scale (*47*). In the case of RBP, however, no conformational exchange was ascribed to Met11, the residue equivalent to Ser21.

Another possible consequence of the observation of conformational exchange of Ser21 is that this residue might be critical to the α−β transition that has been reported for β-LG during folding or unfolding. In support of this, it has been shown that in the presence of trifluoroethanol, the residues of strand A form an α-helix (*19*).

ACKNOWLEDGMENT

We thank Robert Wakefield and Brian Smith (University of Edinburgh) and Lawrie Creamer (New Zealand Dairy Research Institute) for valuable help and discussion.

REFERENCES

1. Hambling, S. G., McAlpine, A. S., and Sawyer, L. (1992) in *Advanced Dairy Chemistry: 1. Proteins* (Fox, P. F., Ed.) pp 141−190, Elsevier Applied Science, Amsterdam.
2. Brignon, G., and Ribadau-Dumas, B. (1973) *FEBS Lett. 33*, 73−76.
3. Qi, X. L., Brownlow, S., Holt, C., and Sellers, P. (1995) *Biochim. Biophys. Acta 1248*, 43−49.
4. Townend, R. J., Herskovits, T. T., Timasheff, S. N., and Gorbunoff, M. J. (1969) *Arch. Biochem. Biophys. 129*, 567−580.
5. Brownlow, S., Cabral, J. H. M., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., North, A. C. T., and Sawyer, L. (1997) *Structure 5*, 481−495.
6. Bewley, M. C., Qin, B. Y., Jameson, G. B., Sawyer, L., and Baker, E. N. (1997) Milk protein polymorphism, International Dairy Federation Bulletin, Special Issue 9702, pp 100−109, International Dairy Federation, Brussels.
7. Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., Baker, E. N., and Jameson, G. B. (1998) *Biochemistry 37*, 14014−14023.
8. Qin, B. Y., Creamer, L. K., Bewley, M. C., Baker, E. N., and Jameson, G. B. (1998) *FEBS Lett. 438*, 272−278.
9. Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., and Jameson, G. B. (1999) *Protein Sci. 8*, 75−83.
10. Hill, J. P., et al. (1996) in *Macromolecular interactions in food technology* (Parris, N., Kato, A., Creamer, L. K., and Pearch, J., Eds.) pp 281−294, American Chemical Society, Washington, DC.
11. Sawyer, L., Brownlow, S., Polikarpov, I., and Wu, S.-Y. (1998) *Int. Dairy J. 8*, 65−72.
12. Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S., and Cambilau, C. (1996) *Nat. Struct. Biol. 3*, 863−867.
13. Newcomer, M., Jones, T. A., Aquist, J., Sundelin, J., Erriksson, V., Rask, L., and Peterson, P. A. (1984) *EMBO J. 3*, 1451−1454.
14. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1990) *Proteins 8*, 44−61.
15. Huber, R., Schneider, M., Mayr, I., Muller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H., and Kayser, H. (1987) *J. Mol. Biol. 198*, 499−513.
16. Bocskei, Z., Groom, C. P., Flower, D. R., Wright, C. E., Cavaggioni, A., Findlay, J. B., and North, A. C. T. (1992) *Nature 360*, 186−188.
17. Pederson, K. O. (1936) *Biochem. J. 30*, 961−970.
18. Tanford, C., Bunville, L. G., and Nozaki, Y. (1959) *J. Am. Chem. Soc. 81*, 4032−4035.
19. Kuwata, K., Hoshino, M., Era, S., and Batt, C. A. (1998) *J. Mol. Biol. 283*, 731−739.
20. Prusiner, S. B. (1997) *Science 278*, 245−251.
21. Fugate, R. G., and Song, P.-S. (1980) *Biochim. Biophys. Acta 625*, 28−42.
22. Molinari, H., Ragona, L., Varani, L., Musco, G., Consonni, R., Zetta, L., and Monaco, H. L. (1996) *FEBS Lett. 381*, 237−243.
23. Ragona, L., Pusterla, F., Zetta, L., Monaco, H. L., and Molinari, H. (1997) *Folding Des. 2*, 281−290.
24. Fogolari, F., Ragona, L., Zetta, L., Romagnoli, S., Dekruif, K. G., and Molinari, H. (1998) *FEBS Lett. 436*, 149−154.
25. Uhrinova, S., Uhrin, D., Denton, H., Smith, M., Sawyer, L., and Barlow, P. N. (1998) *J. Biomol. NMR 12*, 89−107.

26. Kim, T. R., Goto, Y., Hirota, N., Kuwata, K., Denton, H., Wu, S. Y., Sawyer, L., and Batt, C. A. (1997) *Protein Eng. 10*, 1339−1345.
27. Denton, H., Smith, M., Husi, H., Uhrin, D., Barlow, P. N., Batt, C. A., and Sawyer, L. (1998) *Protein Expression Purif. 14*, 97−103.
28. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson., Ser. A 102*, 241−245.
29. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. M. (1995) *J. Magn. Reson., Ser. B 108*, 94−98.
30. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR 4*, 871−878.
31. Live, D. H., Davis, D. G., Agosta, W. C., and Cowburn, D. (1984) *J. Am. Chem. Soc. 106*, 1939−1941.
32. Bax, A., and Subramanian, S. (1986) *J. Magn. Reson. 67*, 565−569.
33. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., and Torchia, D. A. (1992) *J. Magn. Reson. 97*, 359−375.
34. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
35. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
36. Jones, J. A. (1997) *J. Magn. Reson. 126*, 283−286.
37. Bartels, C., Xia, T. H., Billeter, M., Guntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR 6*, 1−10.
38. Nilges, M., Macias, M. J., OíDonaghue, S. I., and Oschkinat, H. (1997) *J. Mol. Biol. 269* 408−422.
39. Brünger, A. (1992) *X-PLOR Version 3.1, A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
40. Brünger, A. T., Adams, P. D., Clore, G. M., et al. (1998) *Acta Crystallogr. D54*, 905−921.
41. Raine, A. R. C., Smith, B. O., and Domaille, P. J. Unpublished results.
42. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry 28*, 8972−8989.
43. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
44. Tjandra, N., Wingfield, P., Stahl, S., and Bax, A. (1996) *J. Biomol. NMR 8*, 273−284.
45. Clore, G. M., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) *Biochemistry 29*, 7387−7401.
46. Lu, J., Lin, C.-L., Tang, C., Ponder, J. W., Kao, J. L. F., Cistola, D. P., and Li, E. (1999) *J. Mol. Biol. 286*, 1179−1195.
47. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI992629O