

Crystal Structures of Apo- and Holo-bovine α -Lactalbumin at 2.2-Å Resolution Reveal an Effect of Calcium on Inter-lobe Interactions*

Received for publication, June 1, 2000, and in revised form, June 30, 2000
Published, JBC Papers in Press, July 14, 2000, DOI 10.1074/jbc.M004752200

Evangelia D. Chrysina[‡], Keith Brew[§], and K. Ravi Acharya^{‡¶}

From the [‡]Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom and the [§]Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, Miami, Florida 33101

High affinity binding of Ca^{2+} to α -lactalbumin (LA) stabilizes the native structure and is required for the efficient generation of native protein with correct disulfide bonds from the reduced denatured state. A progressive increase in affinity of LA conformers for Ca^{2+} as they develop increasingly native structures can account for the tendency of the apo form to assume a molten globule state and the large acceleration of folding by Ca^{2+} . To investigate the effect of calcium on structure of bovine LA, x-ray structures have been determined for crystals of the apo and holo forms at 2.2-Å resolution. In both crystal forms, which were grown at high ionic strength, the protein is in a similar global native conformation consisting of α -helical and β -subdomains separated by a cleft. Even though alternative cations and Ca^{2+} liganding solvent molecules are absent, removal of Ca^{2+} has only minor effects on the structure of the metal-binding site and a structural change was observed in the cleft on the opposite face of the molecule adjoining Tyr¹⁰³ of the helical lobe and Gln⁵⁴ of the β -lobe. Changes include increased separation of the lobes, loss of a buried solvent molecule near the Ca^{2+} -binding site, and the replacement of inter- and intra-lobe H-bonds of Tyr¹⁰³ by interactions with new immobilized water molecules. The more open cleft structure in the apo protein appears to be an effect of calcium binding transmitted via a change in orientation of helix H3 relative to the β -lobe to the inter-lobe interface. Calcium is well known to promote the folding of LA. The results from the comparison of apo and holo structures of LA provide high resolution structural evidence that the acceleration of folding by Ca^{2+} is mediated by an effect on interactions between the two subdomains.

α -Lactalbumin (LA)¹ is the regulatory protein of lactose synthase (1–3), modulating the affinity of the catalytic component,

UDP-galactose β -N-acetylglucosaminide β 1,4-galactosyltransferase I, for acceptor substrates through a reversible protein-protein interaction (4). LA is also secreted into milk and aggregates of a partially folded form of the human protein have been found to have selective apoptotic effects on tumor cells (5). LA is homologous with the c-type lysozymes (LZs) (6) and provides an example of extreme functional divergence in homologous proteins with closely similar structures (7–9). LA binds calcium strongly and specifically but the LZs include subgroups that bind Ca^{2+} at a site corresponding to that in LA and others (“conventional” LZs) that lack a Ca^{2+} -binding site (10). The K_d^{app} for Ca^{2+} binding to apo-LA under physiological conditions is of the order of 10^{-7} M (11–13). The Ca^{2+} ion in LA has a structural role, being required for folding and native disulfide bond formation in the reduced denatured protein (14, 15). At ambient temperature and low ionic strength, apo-LA undergoes a transconformation to a molten globule (MG) state, which lacks a fixed tertiary structure but retains much of the native secondary structure. The MG is also stabilized by mildly disruptive treatments such as low pH or intermediate concentrations of denaturants (16, 17); it is also identical to a kinetic MG that is highly populated during the refolding of chemically denatured LA (18–20).

Crystal structures of holo-LA from various species in different crystal forms and metal complexes (7, 8, 21–26) are similar to those of the homologous LZs, with two lobes of structure separated by a cleft (Fig. 1A). The larger α -helical lobe is formed by the amino- and carboxyl-terminal sections of the polypeptide chain (residues 1–34 and 86–123) while the smaller lobe, which encompasses a small three-stranded antiparallel β -sheet, a small 3_{10} helix and some irregular structure, is formed by the central section of the polypeptide chain (residues 35–85). We refer to the lobes as “subdomains” since they lack properties normally expected of true domains such as sequence contiguity and distinct functional properties and origins. The high affinity Ca^{2+} -binding site is located at the junction of the subdomains and is composed of a contiguous section of chain (residues 79–88) connecting the 3_{10} -helix of the β -subdomain (residues 77–80) and helix H3 of the helical subdomain (residues 86–98), and contains highly conserved Ca^{2+} liganding aspartates (7). A secondary Ca^{2+} -binding site (different from Zn^{2+} or SO_4^{2-} site) has also been identified 7.9 Å away from the primary binding site (26) (Fig. 1A). This site has features in common with a Mn^{2+} -binding site originally described by Gerken (27).

In the absence of Ca^{2+} , denatured LA, with intact disulfide bonds, refolds slowly to a native conformation and this process is accelerated by at least 2 orders of magnitude by the preferential binding of Ca^{2+} binding to high energy intermediate(s) (17). Structural differences between the apo and holo proteins are of considerable interest in providing high resolution struc-

* This work was supported by Leverhulme Trust (UK) Research Grant F351/U and Wellcome Trust (UK) equipment Grant 055505/98/Z (to K. R. A.), National Institutes of Health Research Grant GM58773 (to K. B.) and European Union Training Fellowship ERBFACRT965075 (to E. D. C.), work at EMBL, Hamburg, is supported by the European Union EU TMR/LSF Grant, Contract number ERBFMGECT 980134. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates for apo-bLA and bLA structures have been deposited with the RCSB Protein Data Bank accession numbers 1F6R and 1F6S, respectively.

¶ To whom correspondence should be addressed. Tel.: 44-1225-826238; Fax: 44-1225-826779; E-mail: K.R.Acharya@bath.ac.uk.

¹ The abbreviations used are: LA, α -lactalbumin; LZ, lysozyme; MG, molten globule; bLA, bovine α -lactalbumin; r.m.s., root mean square.

TABLE I
 Data collection and refinement statistics of apo-bLA and bLA

Space group details	Apo-bLA	bLA
Cell dimensions (a, b, c) (Å)	$a = b = 119.57, c = 152.74$	$a = 72.05, b = 104.65, c = 117.41$
α, β, γ (°)	90, 90, 90	90, 90, 90
Space group	$P4_12_12$	$P2_12_12$
No. of molecules per asymmetric unit	6	6
Data collection and processing statistics		
Station (synchrotron)	Daresbury PX95	EMBL, BW7B
Image plate	Mar. 30	Mar. 30
Wavelength (Å)	0.9	0.847
No. of observations	561254	371734 ^a
No. of unique reflections	57273	45363 ^a
Max. resolution (Å)	2.2	2.2 ^a
Completeness (outermost shell) (%) ^b	99.3 (99.5)	99.2 (99.2) ^a
$R_{\text{sym}}(I)$ (outermost shell) (%) ^c	8.6 (63.3)	8.1 (51.7) ^a
$\langle I/\sigma(I) \rangle$ (outermost shell)	8.8 (2.5)	11.5 (2.1) ^a
Outermost shell (Å)	2.28–2.20	2.32–2.20 ^a
Refinement statistics and model quality		
Resolution range (Å)	40.0–2.20	40.0–2.20
No. of reflections	56422	45348
No. of protein atoms	5844	5856
No. of solvent molecules	167	89
Ca ²⁺ ions		6
R_{free} (%) ^d	24.8	25.3
R_{conv} (%) ^e	19.1	21.6
R.m.s. deviations in		
Bond lengths (Å) ^f	0.011	0.006
Bond angles (°)	1.582	1.27
Temperature factors (Å ²)		
Overall	36.1	49.7
Average main chain	34.1	48.9
Average side chain	37.8	50.7
Ca ²⁺ ions		42.3
From Wilson plot	20.1	31.2

^a For bLA, includes reflections after merging with the in-house data collected on a Siemens area detector mounted on a Siemens X-ray generator (wavelength = 1.5418 Å, 0.25° per frame).

^b Completeness in the range ∞ -resmax, where resmax is the maximum resolution to which data were collected.

^c $R_{\text{sym}}(I) = \sum_{hkl} \sum_i |I_{hkl,i} - I_{\text{average},hkl}| / \sum_{hkl} \sum_i I_{hkl,i} \times 100$.

^d $R_{\text{free}} = \sum_{hkl} |F_{\text{obs}} - k|F_{\text{calc}}| / \sum_{hkl} |F_{\text{obs}}| \times 100$, where $hkl \subset T$ represents the test set (5% of the diffraction data for apo-bLA and 10% for bLA).

^e $R_{\text{conv}} = \sum_{hkl} |F_{\text{obs}} - k|F_{\text{calc}}| / \sum_{hkl} |F_{\text{obs}}| \times 100$.

^f Deviations from ideal values (59).

tural information relevant to late events in LA folding and early steps in unfolding to the MG but the detailed structure of the native conformation of the Ca²⁺-free protein has not been determined previously. We describe here x-ray crystallographic structures for both apo and holo forms of bovine LA that allow a detailed comparison of their structures at 2.2-Å resolution. The two forms were crystallized in tetragonal and orthorhombic space groups, respectively, with six molecules in the asymmetric unit. Throughout our discussion, we refer to the Ca²⁺-bound protein as bLA and the Ca²⁺-free form as apo-bLA.

EXPERIMENTAL PROCEDURES

Preparation of Proteins

Native bovine LA (bLA) was obtained from Sigma and used without further purification. Apo-bLA was prepared by gel filtration of holo protein in 0.01 M HCl (pH 2.0) in Ca²⁺-free water as described by Kronman *et al.* (11) and was subsequently lyophilized. Containers were treated with EDTA solutions and washed with Chelex-treated water.

Crystallization

Apo-bLA—Crystals were grown by the vapor diffusion method. Equal volumes (2.0 μ l) of protein (~20 mg/ml in water) and reservoir solution (2.0 M ammonium sulfate, 0.1 M Tris-HCl, pH 6.0) were mixed on siliconized coverslips and left to equilibrate at 16 °C against the reservoir solution. Crystals appeared in ~20 days as tetrahedrons. Care was taken to avoid contamination by adventitious metal ions during crystallization experiments. All glassware were acid washed using nitric

acid and rinsed with Chelex-treated water according to the protocol of Veillon and Vallee (28).

bLA—Native bovine LA was crystallized by the vapor diffusion method at 16 °C against a reservoir solution containing 15% PEG 8,000 and 0.05 M potassium dihydrogen orthophosphate. Crystals can also be grown using sodium acetate (pH 5.0). Equal volumes (2.0 μ l) of protein solution (~20 mg/ml) and reservoir solution were used. Lozenge-shaped crystals appeared after a day and continued to grow for about 5 days.

Diffraction Data Collection

Apo-bLA—Data were collected from one crystal at the Synchrotron Radiation Source, Daresbury on station PX 9.5 (oscillation range 1.0°, exposure time ~150 s/image, $\lambda = 0.90$ Å) at room temperature (19 °C) using a 30-cm MAR Research image plate. The data set was indexed in tetragonal space group $P4_12_12$, which indicated the possibility of 6 or 8 molecules in the asymmetric unit with a solvent content of 61.3% (Matthews coefficient of ~3.2 Å³/Da). Data integration and reduction were performed with the programs DENZO and SCALEPACK (29). Diffraction spots up to 2.0 Å were observed but completeness dropped beyond 2.2-Å resolution. The data collection statistics are presented in Table I.

bLA—Data from one bLA crystal were collected to 2.2 Å on station BW7B at EMBL-Hamburg Synchrotron Radiation Source, using a 30-cm MAR Research image plate (Table I). The oscillation range per image was 1.0° ($\lambda = 0.847$ Å) and the exposure time was 40–50 s/image. Both DENZO and SCALEPACK programs were used for data processing (29). To improve the completeness at low resolution shells, a medium resolution data set which was collected previously, using crystals exposed to graphite monochromated CuK α x-rays ($\lambda = 1.5418$ Å) pro-

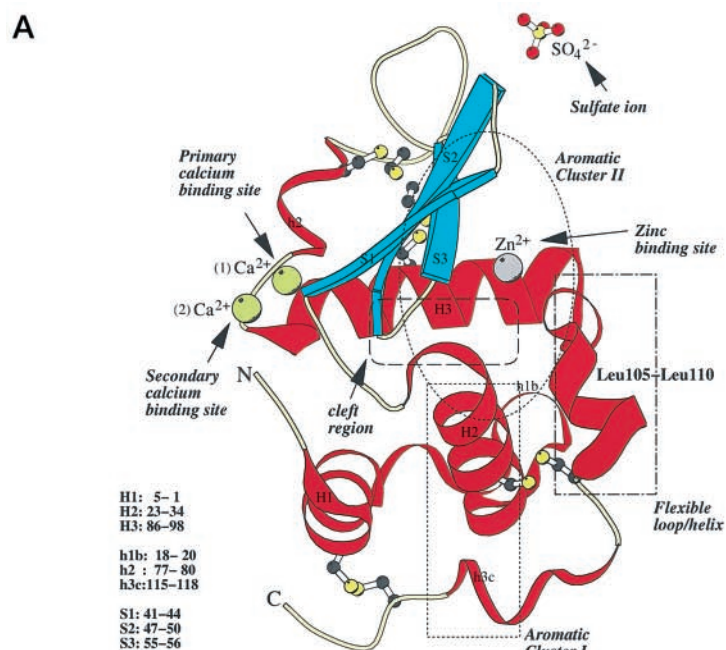


FIG. 1. A, structure of LA and the functional regions of the molecule showing the location of metal ions identified in LA crystal structures. Secondary structural elements (*S*, β -strand; *H*, α -helix; *h*, 3_{10} helix) are marked. The figure was produced using BOBSCRIPT (55). B, superposition of bovine LA structures. The Ca^{2+} -free (apo-bLA) structure is shown in red, the native bovine LA (bLA) structure is shown in yellow, and recombinant bLA (mLA) in green. The superposition of the structures was performed with program O (35). The insets show the details of packing interactions in aromatic clusters I and II (the hydrophobic box).

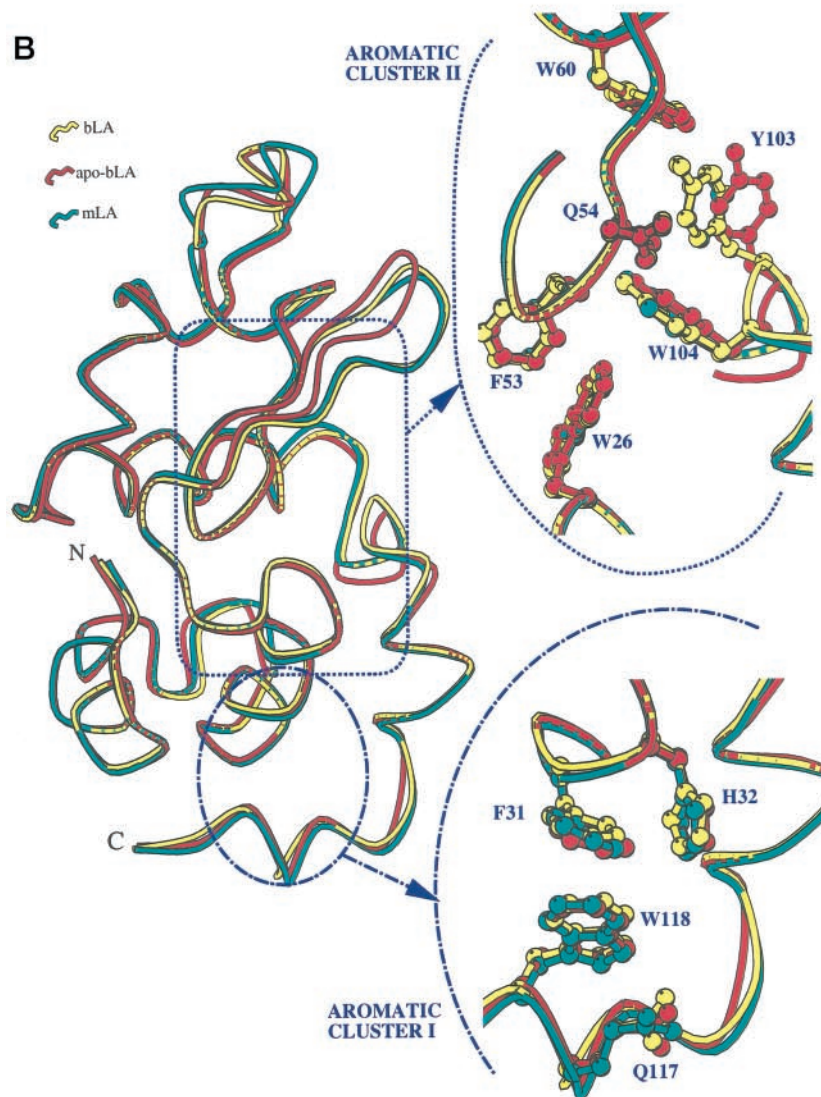


TABLE II
 Comparison of different LA structures

The r.m.s. deviations in Å are given after superposition using the least square fitting option in O applied to C α atoms of residues 1–120 (values in boldface) and to the core of the molecule (values in normal typeface). LA core: residues 5–11, 23–40, 50–61, and 71–104. mLA: recombinant bovine LA at 2.3 Å (PDB code; 1HFZ (24)). Baboon LA at 1.7 Å (PDB code; 1ALC (9)). Guinea pig LA at 1.9 Å (PDB code, 1HFX (24)). Goat LA at 2.3 Å (PDB code, 1HFY (24)). Buffalo LA at 2.3 Å (25). Human LA at 1.7 Å (21). Human LA in complex with Zn²⁺ at 1.7 Å (PDB code, 1HML (23)). Human LA with two Ca²⁺ ions at 1.8 Å (PDB code, 1A4V (26)).

	Apo-bLA	bLA	mLA (recom. bovine)	Baboon	Guinea pig	Goat	Buffalo	Human	Human in complex with Zn ²⁺	Human with two Ca ²⁺ ions
Apo-bLA		0.68	0.80	0.93	1.01	0.75	0.80	0.99	0.97	0.94
bLA	0.39		0.51	0.89	0.82	0.61	0.60	0.87	0.87	0.89
mLA (Bovine)	0.41	0.28		0.94	0.84	0.77	0.70	0.92	0.90	0.93
Baboon	0.65	0.65	0.65		1.06	1.06	1.04	0.41	0.36	0.47
Guinea pig	0.57	0.44	0.44	0.66		1.11	1.08	1.01	1.03	1.04
Goat	0.60	0.45	0.54	0.80	0.70		0.52	1.05	1.01	1.02
Buffalo	0.53	0.95	0.48	0.76	0.63	0.27		1.01	1.01	1.02
Human	0.66	1.00	0.65	0.27	0.63	0.81	0.78		0.35	0.35
Human in complex with Zn ²⁺ ion	0.62	0.94	0.58	0.21	0.61	0.71	0.69	0.26		0.35
Human with two Ca ²⁺ ions	0.63	1.37	0.60	0.23	0.59	0.74	0.71	0.23	0.16	

duced by a Siemens rotating anode generator operating at 50 kV and 80 mA at room temperature was merged using SCALA and subsequently converted to structure factor amplitudes with the program TRUNCATE as implemented by the CCP4 suite of programs (30). The systematic absences were found to agree with an orthorhombic P2₁2₁2 space group with either 6 or 8 molecules in the asymmetric unit and with a solvent content of 52.3% (Matthews coefficient of ~2.6 Å³/Da) (Table I).

Structure Determination

Apo-bLA—The structure of recombinant bLA (mLA, in the monoclinic form) determined previously at 2.3 Å (24) was used for molecular replacement with the program AMoRe (31). A combination of molecular replacement, rigid body refinement using X-PLOR (32), and visual inspection of the $2F_o - F_c$ electron density maps were helpful in identifying positions for five of the six apo-bLA molecules in the asymmetric unit. Further attempts to determine the sixth molecule using molecular replacement were not successful. However, visual inspection of the $F_o - F_c$ electron density map calculated using the refined model with five molecules, identified the position of the sixth molecule as a large piece of contiguous density sufficient to accommodate one more molecule. Careful examination of the packing features of five molecules and the symmetry information derived from the self-rotation function enabled accurate positioning of the sixth molecule in the unaccounted electron density region. These results confirmed the presence of only six molecules (three dimers) in the asymmetric unit.

bLA—The structure of bLA was determined by molecular replacement with the program AMoRe (31) using the structure of recombinant bovine LA (24) as a search model. As in the case of apo-bLA, packing considerations and visual inspection of electron density maps revealed the positions of the six molecules (three dimers) in the asymmetric unit.

Refinement

The resultant models of the apo-bLA and bLA structures (six molecules) were first subjected to rigid body refinement. Cycles of refinement were performed using simulated annealing protocol with the program X-PLOR (32) and non-crystallographic restraints for apo-bLA and with CNS (33) for bLA. The progress of refinement was monitored through both free and conventional *R*-factors (34). Alternate cycles of manual rebuilding with the graphics program "O" (35), and refinement cycles using the standard protocol improved the quality of the model. In the case of the bLA structure, simulated annealing omit maps calculated using CNS indicated density at the calcium site sufficient to accommodate a Ca²⁺ ion in all six molecules (no Ca²⁺ ion was detected for apo-bLA throughout refinement). Six Ca²⁺ ions were incorporated into the model (one per molecule) at final stages of the refinement. During the final stages of refinement water molecules were inserted into the model only if there were peaks in the $F_o - F_c$ electron density maps with heights greater than 3 σ and they were at hydrogen bond forming distances from appropriate atoms. $2F_o - F_c$ maps were also used to check the consistency in peaks. During the final cycles of refinement the non-crystallographic restraints were released and individual molecules were examined on the graphics with the aid of electron density maps. Analysis of the Ramachandran plot (calculated with the program PROCHECK, Ref. (30)) for both structures showed that all residues lay in the allowed regions. The details of refinement statistics

and model accuracy are listed in Table I. Structural superpositions were performed with program O (35) and SHP (36).

RESULTS

Overall Structures—The structures of both apo-bLA (tetragonal form) and bLA (orthorhombic form) were refined to 2.2-Å resolution (Table I). The six molecules in the asymmetric unit of each structure are very similar with r.m.s. deviations of ~0.39 Å in apo-bLA and ~0.52 Å for bLA. The two structures are similar to each other and to recombinant bovine LA, mLA (24) (Fig. 1B), and LA from other species (Table II), but bLA is more similar to each of the other known structures than is apo-bLA (average r.m.s. deviation of 0.60 Å versus 0.75 Å). Differences between various holo-LA structures are mainly confined to the flexible, solvent exposed loop regions between residues 43–47, 62–65, and the disordered C-terminal tail. The flexible loop region, residues 105–111, which is located adjacent to the lower end of the cleft, adopts a "helical" conformation in both apo-bLA and bLA, as in the previously determined structure of mLA crystallized at pH 8.0 (24) (Fig. 1A). Currently known LA structures indicate that at higher pH values (6.5–8.0) this loop adopts the helical conformation and at low pH (4.6), the "loop" conformation (22, 24).

Apo-bLA and bLA, have, respectively, ~6,512 Å² and ~6,820 Å² accessible surface area. 167 water molecules were identified in apo-bLA and 89 in bLA with temperature factors (*B*-factors) <45 Å² (Table I). This difference could be explained due to higher solvent content in apo-bLA crystals (see "Experimental Procedures") and different packing interactions. The formation of dimers in the crystal results in a loss of 463 Å² (~3.6%) of the accessible surface area in apo-bLA and 536 Å² (~3.9%) in bLA, consistent with other relatively low affinity protein-protein interfaces. Interactions in the dimers are mainly mediated through charged residues. A comparison of the *B*-factors in apo-bLA and bLA indicates that the absence of calcium results in an increase in mobility in main chain and side chain atoms in the Ca²⁺-binding site and C-terminal region but a decrease in mobility in the β -lobe (Fig. 2).

Calcium-binding Site—The Ca²⁺-binding site in bLA has an identical conformation (distorted pentagonal bipyramid coordination) to that found in all currently known LA structures (Fig. 3, Table III). In bLA, as well as mLA, the average *B*-factor for the Ca²⁺ ion is slightly higher than in the known structures of LAs from other species. However, the Ca²⁺ ion and the two water molecules that form its coordination sphere are clearly visible in the electron density map, except in molecule 5, where only one water molecule was identified. At low ionic strength and neutral pH, apo-bLA assumes a MG state in solution. This

FIG. 2. **B-factor plot (main chain) for apo-bLA and bLA.** In apo-bLA, the absence of calcium results in an increase in mobility in main chain and side chain atoms in the Ca^{2+} -binding site and COOH-terminal region but a decrease in mobility in the β -lobe.

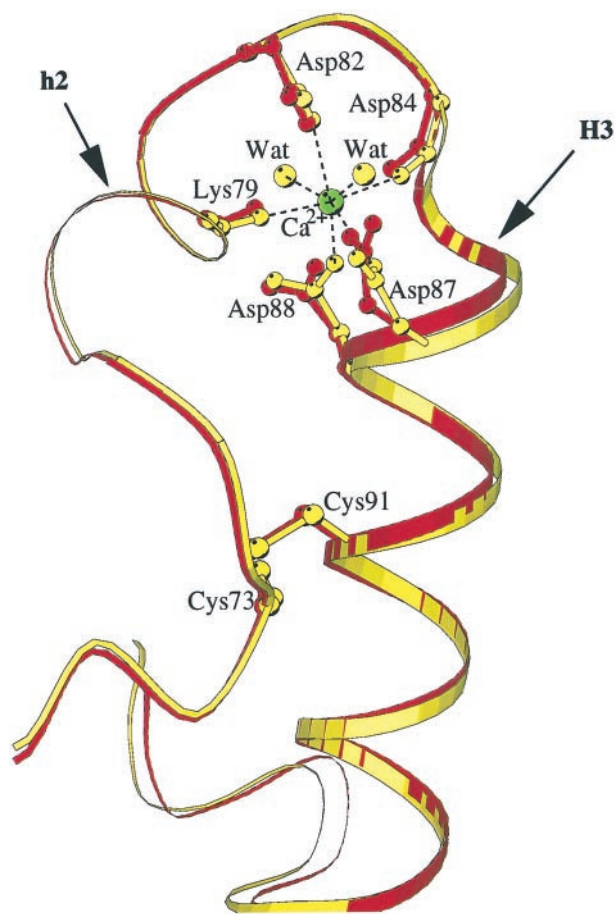
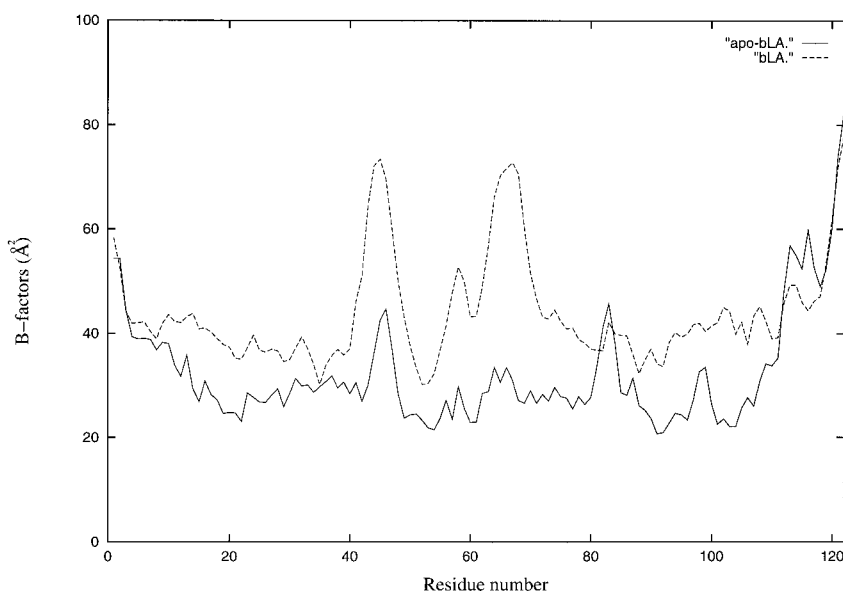


FIG. 3. **Comparison of apo-bLA (red) and bLA (yellow) structures at the Ca^{2+} -binding site and the flanking 3_{10} helix (h2) and α -helix (H3).**

is converted to a native conformation at higher concentration of NaCl or KCl (37). In the crystal structure, apo-bLA is probably stabilized in a native conformation by the high ionic strength of the crystallization medium since no ion was found replacing the absent calcium. In addition, the two conserved water molecules (part of the Ca^{2+} site), which are present in all known holo-LA structures, are absent in apo-bLA. In apo-bLA, the side

TABLE III
Thermal parameters at the calcium-binding site ($\text{\AA}^2$)

	bLA					
	Lys ⁷⁹ -O	Asp ⁸² -OD1	Asp ⁸⁴ -O	Asp ⁸⁷ -OD1	Asp ⁸⁸ -OD1	Wat (1) Wat (2)
mol1	37.68	35.58	38.16	38.23	32.34	35.71 31.82
mol2	38.45	52.59	50.39	40.58	33.80	46.88 42.57
mol3	43.24	49.74	54.50	44.35	43.04	40.12 40.61
mol4	40.96	43.17	46.41	34.57	36.63	32.68 39.32
mol5	48.31	62.25	54.45	53.05	51.10	40.97
mol6	39.11	41.37	38.25	39.84	30.87	33.86 35.17

	Apo-bLA ^a				
	Lys ⁷⁹ -O	Asp ⁸² -OD1	Asp ⁸⁴ -O	Asp ⁸⁷ -OD1	Asp ⁸⁸ -OD1
mol1	29.76	55.50	35.78	56.53	28.38
mol2	30.99	64.83	35.44	42.54	34.26
mol3	36.30	62.98	39.62	42.83	34.20
mol4	37.52	57.51	38.07	49.20	37.74
mol5	35.04	64.89	36.57	51.08	33.74
mol6	39.87	69.30	43.10	50.59	37.72

^a Equivalent residues in Apo-bLA.

chain of Asp⁸⁷ is reoriented but there is no other major structural change (r.m.s. deviation for C^α atoms between residues 78–90 for apo-bLA and bLA is 0.26 $\text{\AA}$). Overall, the Ca^{2+} binding loop in apo-bLA has a slightly higher solvent accessibility and significantly increased B-factors compared with the corresponding loop in bLA (Fig. 2).

Cleft and Subdomain Interface—In the c-type LZs, the cleft contains the active site comprising six monosaccharide-binding subsites (A–F) to which the hexasaccharide substrate binds. In all known LA structures, the side chain of Tyr¹⁰³, which is an alanine in most LZs (9) blocks subsites A and B. In bLA the phenolic hydroxyl group of Tyr¹⁰³ forms hydrogen bonds with the amide oxygen of Gln⁵⁴ (2.9 $\text{\AA}$) and the main chain amide of Asn⁵⁶; these bonds are part of the interaction interface between the two lobes, Tyr¹⁰³ being part of the helical lobe while Gln⁵⁴ and Asn⁵⁶ are in the β -lobe. The largest structural change observed in apo-bLA is in this region; Tyr¹⁰³ is shifted toward the interior of the cleft and water-mediated interactions with Gln⁵⁴ and Asn⁵⁶ replace the direct hydrogen bonds. Also, an H-bond between the carbonyl group of Tyr¹⁰³ and the amide nitrogen of Leu¹⁰⁴ is replaced by an interaction with a new immobilized solvent molecule. In apo-bLA, Tyr¹⁰³ forms contacts ($<4.5\text{\AA}$) with more residues in the α -subdomain than in bLA, particularly Val⁹⁹ and Gly¹⁰⁰, and fewer with residues in

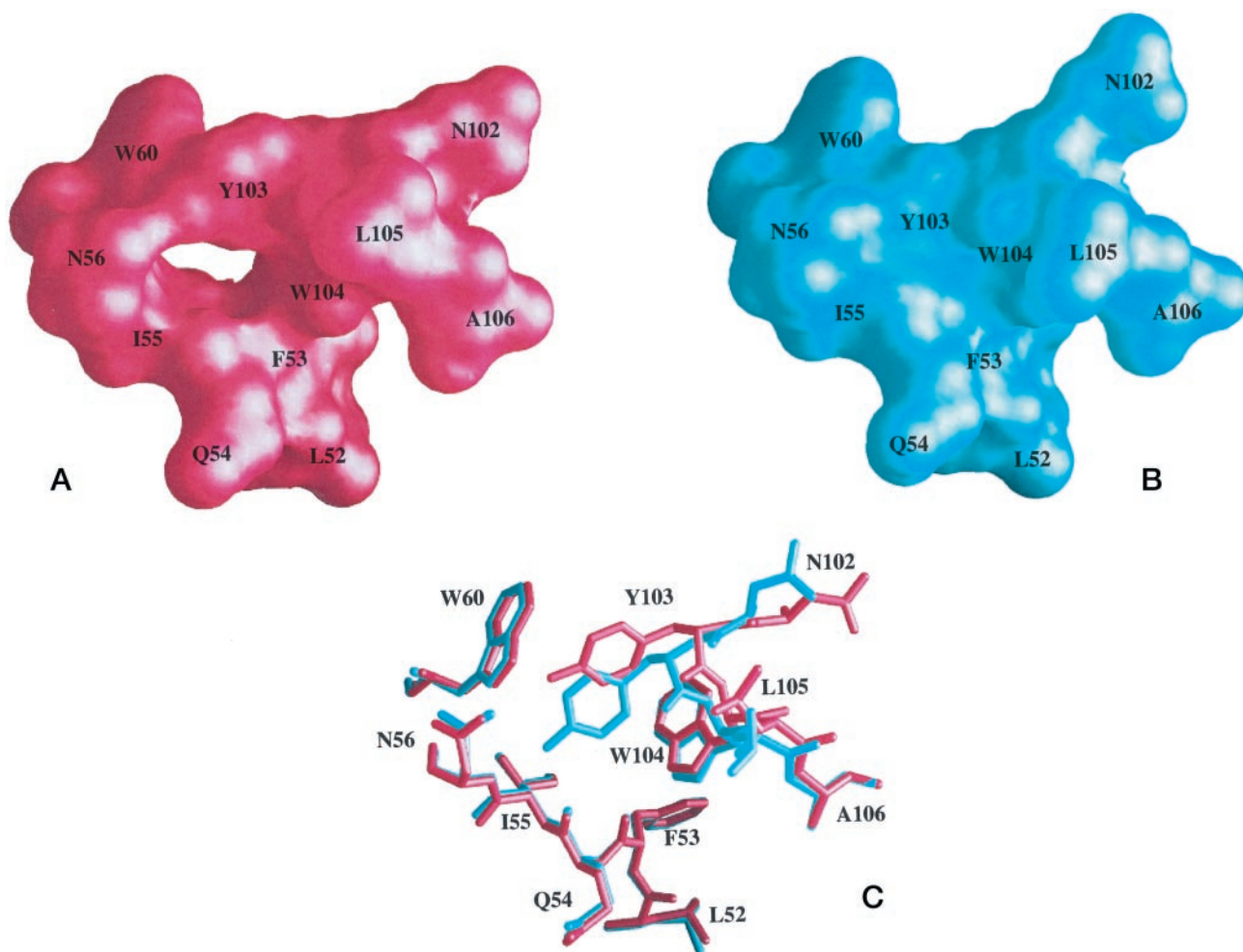


FIG. 4. A and B, molecular surfaces for apo-bLA and bLA at the entrance of the cleft. In the apo-bLA structure a water molecule makes interaction with the hydroxyl group of Tyr¹⁰³ and the main chain atoms of Gln⁵⁴ and Asn⁵⁶ (not shown in the figure) (C). Superposition of residues showing the location of Tyr¹⁰³ in apo-bLA (in pink) and bLA (in blue). The perturbation of Tyr¹⁰³ in apo-bLA results in a slight opening of the cleft.

the β -subdomain, notably Gln⁵⁴ and Ile⁵⁵. This change involving Tyr¹⁰³ is present in all 6 molecules of apo-bLA and results in a more open cleft that can be seen in the altered position of the β -lobe in Figs. 1B and 4. In the 1.7-Å resolution structure of baboon LA, an internal channel between Ile²⁷ and Asp⁸⁸ (from the Ca²⁺-binding site) is partly occupied by internal water molecules (9). The corresponding channel is present in both bLA and apo-bLA. Three buried water molecules are present in bLA (Wat⁹, Wat⁹², and Wat¹²) and these have counterparts in baboon LA. Of these, Wat¹² is a ligand of the Ca²⁺ ion which blocks its access to solvent. The buried solvent molecules are highly conserved in known LA and LZ structures. All three are present in the Ca²⁺ binding echidna LZ (38) while hen egg white LZ (39) has solvent molecules corresponding to the two that are not part of the calcium-binding site. Apo-bLA has buried solvent molecules corresponding to Wat⁸ and Wat⁴ but, like the non-Ca²⁺ binding hen egg white LZ, lacks the water associated with the Ca²⁺-binding site.

Model building using the carbohydrate of hexa-*N*-acetyl-chitohexaose from the human LZ structure (40) and a simple molecular dynamics exercise using the program CNS (33) indicates that the more open cleft of apo-bLA, unlike the closed cleft in bLA, could potentially accommodate carbohydrate residues in sites A-F (Fig. 5).

DISCUSSION

The effects of Ca²⁺ on folding and stability in LA appear to reflect differences in its affinity for different LA conformers.

The native state binds Ca²⁺ with high affinity ($K_d < 1 \mu\text{M}$) whereas the kinetic MG state has a relatively weak affinity ($K_d \sim 1 \text{ mM}$) for calcium (17). The removal of calcium from LA consequently increases the stability of the MG relative to the native state. Despite its low affinity for Ca²⁺, the presence of calcium affects the structure of the MG, resulting in more highly persistent structure (41). Calcium accelerates the rate of refolding in denatured bLA (disulfide bonds intact) by more than 2 orders of magnitude (17, 42); the dependence of folding kinetics on the Ca²⁺ concentration is thought to result from calcium binding to high energy intermediates in the folding process which have partially formed Ca²⁺-binding sites (17). The requirement for Ca²⁺ in the formation of native protein with correct disulfide bond arrangements from the reduced denatured protein (14, 15) suggests that it also binds selectively to similar rate-limiting intermediates in the oxidative folding process. The MG and the high energy intermediates in folding are arrays of conformations as opposed to discrete species and Ca²⁺ binding appears to be a common property of a range of partially folded states of LA in which affinity for the ion increases as their Ca²⁺-binding sites become more native.

The structural studies described here show that the overall fold is the same in the holo and apo-forms of bLA which is consistent with the fact that calcium is not required for activity in LA (11, 43). Interestingly, crystal structures of the apo and holo-forms of a mutant of the non-Ca²⁺ binding human LZ in which a Ca²⁺-binding site has been introduced by mutagenesis, indicate

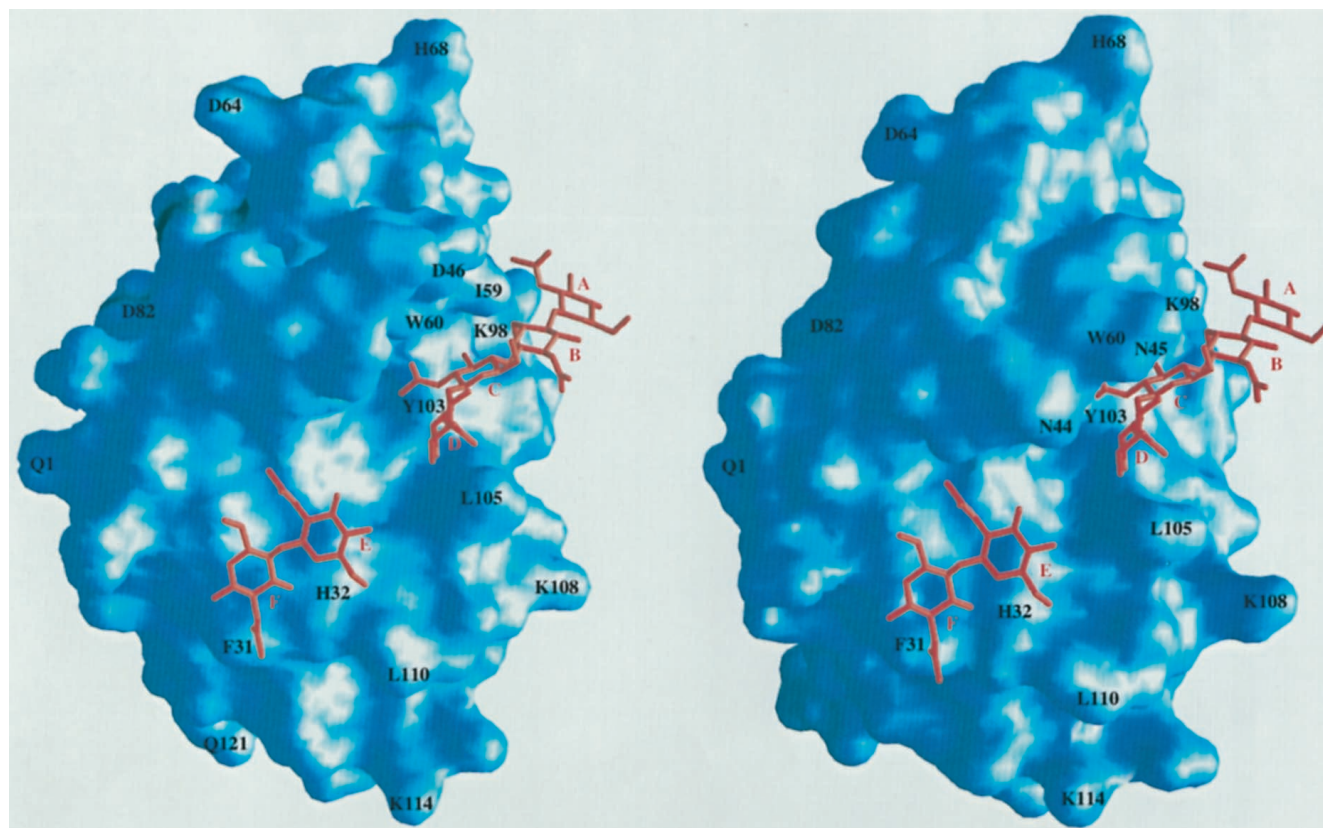


FIG. 5. **The molecular surfaces of apo-bLA (left) and bLA (right).** The structure of hexa-N-acetylchitohexaose (from the human LZ structure, Song *et al.* (see Ref. 40)) was modeled in the cleft of apo-bLA, followed by a molecular dynamics simulation using the program CNS (33) (left). For comparison, the position of the hexasaccharide modeled in the apo-bLA structure is also shown in bLA (right). The figure was prepared with the program GRASP (56).

that only subtle, local perturbations are required to accommodate the Ca^{2+} ion (44). Similarly, the Ca^{2+} binding loop in equine LZ (45) and the corresponding loop in hen egg white LZ, a prototype non- Ca^{2+} binding LZ (39), are closely similar in structure (r.m.s. deviation of 0.59 Å for C^{α} atoms between residues 81–93).

The absence of any cation from the crystallization medium in apo-bLA replacing the calcium of the holo protein, supports the view that stabilization of the native state of the apo protein by higher concentrations of monovalent cations (19, 16) results, in this case, from increased ionic strength and not specific binding. Water molecules that interact with the Ca^{2+} ion in the holo protein are also absent. Although changes in the immediate vicinity of the site are small, a larger structural change is found at the inter-lobe interface (which connects the Ca^{2+} -binding site and the cleft involving Tyr¹⁰³ through a channel containing trapped water molecules) located on the opposite face of the apo-bLA structure. As shown in Figs. 1B and 4, this change reflects a perturbation of interactions between the α - and β -lobes of LA at the edge of the hydrophobic box, resulting in a more open cleft structure in the apo protein. Despite the minor structural adjustments in the Ca^{2+} -binding site of the apo protein, it is reasonable to think that the larger change in the cleft is propagated from the Ca^{2+} binding loop. Charge repulsion between the carboxylates of the five aspartyl residues in this region is a probable source of this change; the reorientation of the side chain of Asp⁸⁷ and a slight expansion of the loop appear to be in response to this, reducing the negative charge density. The effects of charge repulsion in this region have been previously discussed by Kronman (16) and Forge *et al.* (42). Local destabilization of this region may also account for increased *B*-factors in the loop in apo-bLA (Fig. 2). In the native form of apo-bLA described here, the relatively low level of

repulsion between the aspartates resulting from shielding by the high ionic strength, produces a slight tilt of 3_{10} helix h2 (residues 76–82) relative to α -helix H3 (residues 86–99). Helix H3 is clamped to the region immediately NH_2 -terminal to h2 by the Cys⁷³–Cys⁹¹ disulfide bond, so that slight expansion of the loop is transmitted in the form of an increased separation of the helical and β -sheet subdomains (see Figs. 1B and 3) to disrupt native interactions in aromatic cluster II. In the helical subdomain, the structural change only affects the vicinity of Tyr¹⁰³ so that interactions within aromatic cluster I are not significantly changed (Fig. 1B, inset). We suggest that, at low ionic strength, increased charge repulsion will further expand the calcium binding elbow to produce, by the same mechanism, further separation of the subdomains and sufficient disruption of interactions in the hydrophobic box to result in a loss of net stability in the native structure. Interactions between the subdomains are necessary for the formation of fixed native structure as shown by previous protein dissection studies that indicate that the isolated α -helical subdomain forms a MG but not a native structure (46) while the isolated β -subdomain is unstructured (47). Thus, the changes observed in the apo-bLA structure appear to represent an initial step in the transition to the MG. The enhancement of the rate of folding in LA is thought to result from Ca^{2+} binding to rate-limiting intermediates containing Ca^{2+} -binding sites; this has been proposed to stabilize native interactions around the binding site (42). The apo- and holo-bLA structures indicate that Ca^{2+} binding affects packing at the subdomain interface as well as the immediate environment of the Ca^{2+} -binding site in folding intermediates. The importance of this interface may be reflected in the strong conservation of a region of sequence in very distant relatives of the LAs and c-type LZs including goose-type LZs and bacterial

PROTEIN	Residue
Bovine LA	26: WVCTTFHTSGYDTQAIQVQNN DSTEYGLFQINNKIW
Chicken LZ	28: WVCAAKFESNYNTQATNRD-----GSTDYGILQINSRW
Human LZH	28: WVCLAYTSGFNAAALYEA-----GSTNNGIFQINSRW
Pigeon LZ	27: WVCLVKHESYRTAFNNNG-----PNSRDYGIQINSRW
DD LZH	43: ISSISYVESWNPQAYNPSS-----ASGLWQLITLHC
EC SLTG	471: AMAIARQESAWNPVKVSPVG-----ASGLWQLIMPSTA
HI SLTG	446: AMAIARQESAWNPMAQSSAN-----ARGMLQLLPSTA
PI ORF404	222: YAAASWQESHWDPHATSPGT-----VRLGMLLTRETA
EC MLTE	95: ITAIAIESGGNPNAVSKN-----AIGLMQIKASTS
EC MLTC	210: ILAIMQTESFNPYAVSRSD-----ALGLMQVQVQSTA
EC MLTD	118: LVLLPIVESAFDPHATSGAN-----AAGIWLQIIPST-
Phage P1 STLG	763: LYAIAATESGGNPNAKSTLTRSPNGKLSGGALGMFQTSVAR
Plasmid R64 p1T	56: LKAISAGESSLRPGAININKDKKTGKASSTDYGLMLQINSSTH
BS LTH	78: ITRAVIKQESGPNKAVSGAG-----AMGLMLQMLPSTA
Phage PRD1 P7 protein	25: VAGVQTESGPNRPTTSVG-----AMGLMLQMLPSTA
AS TBA	42: LALMPLIESDFNASAYSHKH-----AMGLMLQMLPSTA
HP Dnir	96: FLFLAMAKESFSSRAYSRKK-----AVGIWQFMFSTA
EC ORF P19	37: LRAISWKESSYRVNAIGINPTGY-----GSLGMLQVDSCHF
EC rOrf3	35: LKAIAWESKESKSGIKSKTNGT-----YDIGLMQINSSTH
BB ORF BB0259	59: VFLGLKASSEFKNVSKPG-----AVGLMQVMPSTA
ST IAGB	36: LYAIAQQESAMKPGAIGHNRD-----GSTDLGMLQIMHKR
SF Ipgf protein	35: LKAIAEKESGFKSAVNVNN-----NSGKDYGMQINDPHS
AA HP	232: IYAVMRQESLFDYAVSSSG-----AKGLMLQIDSTA
Goose LZ	91: IAGTISRESHAGKVLKNGWD-----RGNGFGLMQVDKRS
Chicken g-type LZ	66: IAGTISRESHAGKVLKNGWD-----RGNGFGLMQVDKRYH

FIG. 6. Conservation of a region of sequence within the β -lobe in distant relatives of LAs and type-c lysozymes, identified by searches using PSIBLAST (57) and also by structural correlations (58). The proteins are: LZH, lysozyme homologue; SLTG, soluble lytic transglycosylase; MLTE, membrane-bound lytic transglycosylase E; MLTC, membrane-bound lytic transglycosylase C; MLTD, membrane-bound lytic transglycosylase D; STLG, structural lytic transglycosylase; LTH, lytic transglycosylase homologue; TBA, tributyltin chloride resistance protein; Dnir, Dnir regulatory protein; IAGB, invasion protein IAGB; HP, hypotheic protein. Organisms are abbreviated as: DD, *Dictyostelium discoideum*; EC, *Escherichia coli*; BS, *Bacillus subtilis*; AS, *Alteromonas* sp. M1; HI, *Hemophilus influenzae*; PM, *Proteus mirabilis*; HP, *Helicobacter pylori*; BB, *Borrelia burgdorferi*; ST, *Salmonella typhimurium*; SF, *Shigella flexneri*; AA, *Aquifex aeolicus*.

transglycosylases and invasion proteins that represents the β -subdomain component of the interface (Fig. 6). A mutation at one of these conserved sites in human LZ is the cause of hereditary systemic amyloidosis (48).

The structure of the acid-stabilized MG state of LA has been characterized by NMR (42, 49–52) but probably cannot be investigated by x-ray crystallography. bLA is a useful model for studies of protein folding because of its lower cooperativity, reflected in alternative conformations in different crystal forms (24), its equilibrium stable MG state, and the unique structure of its apo-form revealed here. A pre-MG state that is stable at equilibrium under mildly acidic conditions has been also described (53, 54), suggesting that additional near-native partially unfolded conformations may be accessible to structural analysis. Studies are in progress to determine if apo-bLA can be crystallized under more destabilizing conditions to allow the structural analysis of other conformers that illuminate additional steps in the unfolding process.

Conclusions—LA, the regulatory protein of lactose synthase, is a calcium-binding protein. The binding of Ca^{2+} stabilizes the native structure and calcium is required for the generation of native LA from the reduced denatured state. Previous studies have shown that on removal of Ca^{2+} from native LA, the protein adopts a MG structure, a compact conformer with a large proportion of native secondary structure but little fixed tertiary structure. There is considerable interest in the MG state of LA since it is identical to an early intermediate in the folding pathway.

In order to elucidate the influence of Ca^{2+} binding on the structure of LA, structural studies were performed on apo- and holo-bovine LA. The results reveal subtle, but significant conformational differences between the two forms of the protein. In the apo protein the major perturbation is on the opposite face from the Ca^{2+} -binding site in a cleft region at the interlobe interface. This change appears to be result from a slightly expanded structure in the Ca^{2+} -binding site in the apo protein that is magnified by a change in the relative orientation of two helices; the perturbed structure of the apo protein provides

high resolution information about a late step in the transition from the MG to native state in LA.

Acknowledgments—We are grateful to the staff at the Synchrotron Radiation Sources at Daresbury (UK) and EMBL outstation at Hamburg (HASYLAB, c/o DESY, Germany), Demetres Leonidas, and Tassos Papageorgiou for their help with x-ray data collection. We thank the members of the Structural Biology Group for the constructive criticisms of the manuscript; and Ashley Pike, Naveen Chandra, and Songlin Li for their contributions to the LA project.

REFERENCES

- Brew, K., Vanaman, T. C., and Hill, R. L. (1968) *Proc. Natl. Acad. Sci. U. S. A.* **59**, 491–497
- Hill, R. L., and Brew, K. (1975) *Adv. Enzymol. Rel. Areas Mol. Biol.* **43**, 411–489
- Grobler, J. A., and Brew, K. (1992) in *Advanced Dairy Chemistry, Proteins* (Fox, P., ed) Vol. I, pp. 191–229, Elsevier, New York
- Wang, M., Scott, W. A., Rao, K. R., Uday, J., Conner, G. E., and Brew, K. (1989) *J. Biol. Chem.* **264**, 21116–21121
- Svensson, M., Sabharwal, H., Hakansson, A., Mossberg, A.-K., Lipniunas, P., Lefler, H., Svanborg, C., and Linse, S. (1999) *J. Biol. Chem.* **274**, 6388–6396
- Brew, K., Vanaman, T. C., and Hill, R. L. (1967) *Nature* **242**, 3747–3749
- Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M., and Phillips, D. C. (1986) *Nature* **324**, 84–87
- Smith, S. G., Lewis, M., Aschaffenburg, R., Fenna, R. E., Wilson, I. A., Sundaralingam, M., Stuart, D. I., and Phillips, D. C. (1987) *Biochem. J.* **242**, 353–360
- Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M., and Phillips, D. C. (1989) *J. Mol. Biol.* **208**, 99–127
- Grobler, J. A., Rao, K. R., Pervaiz, S., and Brew, K. (1994) *Arch. Biochem. Biophys.* **313**, 360–366
- Kronman, M. J., Sinha, S. K., and Brew, K. (1981) *J. Biol. Chem.* **256**, 8582–8587
- Permyakov, E. A., Yarmolenko, V. V., Kalinichenko, L. P., Morozova, L. A., and Burstein, E. A. (1981) *Biochem. Biophys. Res. Commun.* **100**, 191–197
- Schaefer, J. J., Milos, M., and Cox, J. A. (1985) *FEBS Lett.* **190**, 77–81
- Rao, K. R., and Brew, K. (1989) *Biochem. Biophys. Res. Commun.* **163**, 1390–1396
- Ewbank, J. J., and Creighton, T. E. (1993) *Biochemistry* **32**, 3677–3693
- Kronman, M. J. (1989) *Crit. Rev. Biochem. Mol. Biol.* **24**, 565–667
- Kuwajima, K., Mitani, M., and Sugai, S. (1989) *J. Mol. Biol.* **206**, 547–561
- Kuwajima, K. (1989) *Proteins* **6**, 87–103
- Ikeguchi, M., Kuwajima, K., and Sugai, S. (1986) *J. Biochem. (Tokyo)* **99**, 1191–2001
- Arai, M., and Kuwajima, K. (1996) *Fold. Design* **1**, 275–287
- Acharya, K. R., Ren, J. S., Stuart, D. I., Phillips, D. C., and Fenna, R. E. (1991) *J. Mol. Biol.* **221**, 571–581
- Harata, K., and Muraki, M. (1992) *J. Biol. Chem.* **267**, 1419–1421
- Ren, J. S., Stuart, D. I., and Acharya, K. R. (1993) *J. Biol. Chem.* **268**, 19292–19298
- Pike, A. C. W., Brew, K., and Acharya, K. R. (1996) *Structure* **4**, 691–703
- Calderone, V., Giuffrida, M. G., Viterbo, D., Napolitano, L., Fortunato, D., Conti, A., and Acharya, K. R. (1996) *FEBS Lett.* **394**, 91–95
- Chandra, N., Brew, K., and Acharya, K. R. (1998) *Biochemistry* **37**, 4767–4772
- Gerken, T. A. (1984) *Biochemistry* **23**, 4688–4697
- Veillon, C., and Vallee, B. L. (1978) *Methods Enzymol.* **54**, 446–484
- Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
- CCP4 (1994) *Acta Crystallogr. Sect. D* **50**, 760–763
- Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50**, 157–163
- Brünger, A. T. (1992) *X-PLOR Version 3.1 Manual: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54**, 905–921
- Brünger, A. T. (1992) *Nature* **355**, 472–475
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
- Stuart, D. I., Levine, M., Muirhead, H., and Stammers, D. K. (1979) *J. Mol. Biol.* **134**, 109–142
- Hiraoka, Y., and Sugai, S. (1985) *Int. J. Peptide Protein Res.* **26**, 252–261
- Guss, J. M., Messer, M., Costello, M., Hardy, K., and Kumar, V. (1997) *Acta Crystallogr. Sect. D* **53**, 355–363
- Steinrauf, L. K. (1998) *Acta Crystallogr. Sect. D* **54**, 767–779
- Song, H., Inaka, K., Maenaka, K., and Matsushima, M. (1994) *J. Mol. Biol.* **244**, 522–540
- Vanderheeren, G., and Hanssens, I. (1994) *J. Biol. Chem.* **269**, 7090–7094
- Forge, V., Wijesinha, R. T., Balbach, J., Brew, K., Robinson, C. V., Redfield, C., and Dobson, C. M. (1999) *J. Mol. Biol.* **288**, 673–688
- Berliner, L. J., Ellis, P. D., and Murakami, K. (1983) *Biochemistry* **22**, 5061–5063
- Inaka, K., Kuroki, R., Kikuchi, M., and Matsushima, M. (1991) *J. Biol. Chem.* **266**, 20666–20671
- Tsuge, H., Ago, H., Noma, M., Nitta, K., Sugai, S., and Miyano, M. (1992) *J. Biochem. (Tokyo)* **111**, 141–143
- Peng, Z., and Kim, P. S. (1994) *Biochemistry* **33**, 2136–2141
- Wu, L., Peng, Z., and Kim, P. S. (1995) *Nat. Struct. Biol.* **2**, 281–286
- Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A.,

- Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C. F., Terry, C. J., Feast, T. G., Zalin, A. M., and Hsuan, J. J. (1993) *Nature* **362**, 553–557
49. Baum, J., Dobson, C. M., Evans, P. A., and Hanley, C. (1989) *Biochemistry* **28**, 7–13
50. Ewbank, J. J., and Creighton, T. E. (1991) *Nature* **350**, 518–520
51. Alexandrescu, A. T., Broadhurst, R. W., Wormald, C., Chyan, C. L., Baum, J., and Dobson, C. M. (1992) *Eur. J. Biochem.* **210**, 699–709
52. Alexandrescu, A. T., Evans, P. A., Pitkeathly, M., Baum, J., and Dobson, C. M. (1993) *Biochemistry* **32**, 1707–1718
53. Lala, A. K., and Kaul, P. (1992) *J. Biol. Chem.* **267**, 19914–19918
54. Gussakovsky, E. E., and Haas, E. (1995) *Protein Sci.* **4**, 2319–2326
55. Esnouf, R. M. (1997) *J. Mol. Graph.* **15**, 132–134
56. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins* **11**, 281–296
57. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
58. Thunnissen, A. M., Isaacs, N. W., and Dijkstra, B. W. (1995) *Proteins* **22**, 245–258
59. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. Sect. A* **47**, 100–119