Amino acid sequence and crystal structure of buffalo α-lactalbumin

V. Calderone[a,b,c], M.G. Giuffrida[b], D. Viterbo[c], L. Napolitano[b], D. Fortunato[b], A. Conti[b,*], K.R. Acharya[a]

[a]*School of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK*
[b]*Centro Studi Alimentazione Animali CNR, Turin, Italy*
[c]*Department of Inorganic, Physical and Materials Chemistry, University of Turin, Turin, Italy*

Received 18 June 1996; revised version received 12 August 1996

Abstract Isolation, purification, amino acid sequence determination and X-ray crystal structure of buffalo α-lactalbumin were performed in order to gain further knowledge of the molecular basis of α-lactalbumin in the lactose synthase complex. The deduced amino acid sequence differs at one position from the bovine α-lactalbumin sequence (at position 17). The refined crystal structure at 2.3 Å is very similar to those previously reported for human and baboon α-lactalbumins. However, a portion of the molecule (residues 105–109) exhibits different conformation. It forms a 'flexible loop', and appears to be a functionally important region in forming the lactose synthase complex.

Key words: Amino acid sequence; Crystal structure; α-Lactalbumin; Buffalo milk

1. Introduction

The biosynthesis of lactose in the lactating mammary gland is controlled by the activity of the lactose synthase enzyme system (LS complex) [1]. This enzyme consists of two protein components: β-1,4-galactosyltransferase (GT), the catalytic component, and α-lactalbumin (α-LA), the regulatory component. In tissues other than the lactating mammary gland, GT catalyses the transfer of β-galactose from UDP-galactose to the 4-position of β-*N*-acetylglucosamine residues in the N-linked glycans of glycoproteins. In the presence of α-LA, the K_m of GT for glucose is reduced by up to 3 orders of magnitude so that lactose synthesis proceeds effectively at physiological levels of glucose. α-LA, a major protein of the whey fraction of milk, possesses a number of biological features in addition to its role as a 'specifier' protein in the LS complex. Even though functionally divergent, it is homologous to C-type lysozymes [2]. It is a calcium binding protein [3] and binds different metal ions [4,5]. The precise role of calcium ion in α-LA is not known, it appears that the role of calcium binding in vivo is to mediate the release and folding of nascent LA at the site of translation [6] and is required for the stability of the protein.

Two crystal structures of α-LA (baboon [7], human [8,9]) indicate that the conformation of α-LA is very similar to that of C-type lysozyme. The Ca^{2+} ion is positioned in a helix-turn-helix motif bridging two domains and forming one of the most rigid parts of the molecule. In an effort towards understanding the molecular mechanism of LA action in the LS complex, we have set about looking at sequence variation on LA's modulatory properties. Here we report the amino acid sequence and the 3D structure of buffalo milk α-LA at 2.3 Å resolution determined using X-ray crystallography.

2. Materials and methods

2.1. Purification of buffalo milk α-LA

A pool of buffalo milk was defatted by centrifugation at 3000 rpm for 30 min at 10°C; the skimmed milk was then ultracentrifuged at 16000 rpm for 16 h at 0°C to precipitate the casein fraction. The supernatant was treated with $(NH_4)_2SO_4$ up to 50% saturation and the precipitate was resuspended in 0.06 M $NaHCO_3$, at pH 8.3. The protein fraction obtained was chromatographed on a Sephadex G-75 column (32×1000 mm, Pharmacia, Uppsala, Sweden) using the same buffer as eluent. The fractions collected were tested by SDS-PAGE with a 250 Midget Electrophoresis Unit (LKB, Bromma, Sweden), following Laemmli [10], except that the stacking gel buffer was at pH 8.0 instead of pH 6.8. The resolving gel was 12% T, 3% C and the stacking gel 4% T, 3% C and the gels were 0.75 mm thick. The fractions containing α-LA were pooled and rechromatographed on an anion exchange Sephadex A-50 column (32×500 mm, Pharmacia, Uppsala, Sweden): eluent A was 20 mM Tris-HCl, pH 7.2 and eluent B was 20 mM Tris-HCl, 0.5 M NaCl, pH 7.2 with a linear gradient from 0 to 100% of B. To check the purity of α-LA, the fractions were tested by SDS-PAGE followed by Western blotting using a Midget Multi Blot Electrophoresis Transfer Unit (LKB). The minigels were blotted onto Problott membrane (Applied Biosystems, Foster City, CA, USA), as described by Towbin [11], at 10°C and 120 mA with constant current for 60 min (methanol and SDS were omitted in the transfer buffer). The membrane was stained in 0.2% Coomassie brilliant blue R250, dissolved in the destaining solution (5:1:4 methanol:acetic acid:water). The bands showing an apparent molecular weight of about 14 kDa were cut out from the Problott membrane and directly microsequenced on an Applied Biosystems 470A gas-phase sequencer [12] equipped with a 120A phenylthiohydantoin-amino acid derivative analyser [13].

2.2. Amino acid analysis and sequence

The amino acid composition of buffalo α-LA and of its peptides was determined after hydrolysis with 6 N HCl at 110°C for 24 h under vacuum on a Beckman 118CL (Palo Alto, CA, USA) amino acid analyser. For sequence analysis the protein was modified by reduction in 6 M guanidinium chloride, 0.001 M EDTA, 0.1 M Tris-HCl, pH 8.3; 0.002 M dithiothreitol was added and the reduction proceeded under nitrogen, in the dark, at 37°C for 1 h. *S*-Carboxamidomethylation was performed with 0.1 M iodoacetamide in the same conditions as for the reduction of the protein. The reduced, carboxamidomethylated α-LA was digested using different enzymes: 0.5 mg of protein were incubated with 3% (w/w) of endoprotease Glu-C (Boehringer, Mannheim, Germany) in 0.1 M NH_4HCO_3 at 37°C for 6 h and the same amount with 1% trypsin (TPCK) (Worthington Biochemical Corp., Freehold, NJ, USA) in 0.2 M *N*-ethylmorpholine acetate, pH 8.3, at 37°C for 1 h; both reactions were stopped by freezing at −40°C. The peptides obtained were purified by RP-HPLC (Varian LC 5000, Palo Alto, CA, USA) on a LiChrospher 100 RP-18 column (5 mm, Merck, Darmstadt, Germany) in 0.1% trifluoroacetic acid (TFA) and eluted with a linear gradient of 10–55% acetonitrile in 0.1% TFA for 60 min at a flow rate of 0.7 ml/min. Some fractions were rechromatographed on the same column in 0.05 M ammonium acetate, pH 6.0 using a gradient of 0.0–75% acetonitrile at a flow rate of 0.7 ml/min. A limited proteolysis on the native protein was per-

*Corresponding author or K.R. Acharya.

PII S0014-5793(96)00933-7

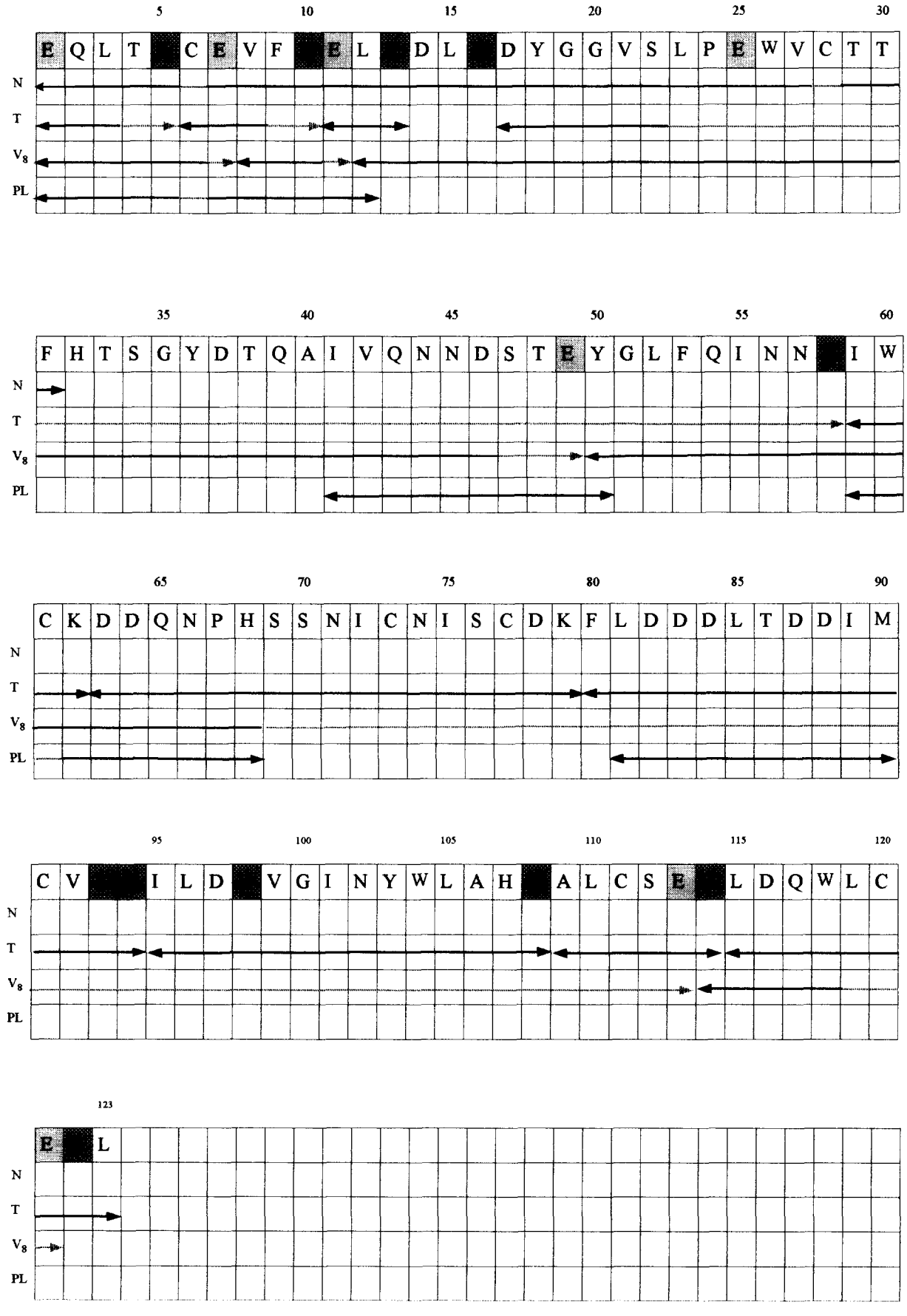

5 10 15 20 25 30
E Q L T C E V F E L D L D Y G G V S L P E W V C T T
N
T
V8
PL
35 40 45 50 55 60
F H T S G Y D T Q A I V Q N N D S T E Y G L F Q I N N I W
N
T
V8
PL
65 70 75 80 85 90
C K D D Q N P H S S N I C N I S C D K F L D D D L T D D I M
N
T
V8
PL
95 100 105 110 115 120
C V I L D V G I N Y W L A H A L C S E L D Q W L C
N
T
V8
PL
123
E L
N
T
V8
PL

Fig. 1. The complete amino acid sequence of buffalo milk α-LA. The only position which differs from bovine α-LA is position 17 (Asp (buffalo α-LA), Gly (bovine α-LA)). N = N-terminal sequence; T = trypsin; V_8 = Glu-C; PL = limited proteolysis; black line corresponds to amino acid residues identified by sequencing; grey line corresponds to amino acid residues deduced from amino acid analysis.
←

formed using 5% thermolysin (Sigma Chemical Co., St. Louis, MO, USA), in 0.05 M Tris-HCl, 0.005 M $CaCl_2$, pH 7.0 in 50% trifluoroethanol at 40°C for 45 min [14]. The complete amino acid sequence derived is shown in Fig. 1.

2.3. Crystallisation

Crystals of buffalo α-LA were obtained at 16°C using the hanging drop vapour diffusion method from a reservoir solution containing 0.2 M $(NH_4)_2SO_4$, 0.1 M CH_3COONa, 25% PEG 4000 and 10 mM $CaCl_2$, pH 4.5. The protein was first dissolved in water and the hanging drops were placed on a siliconised glass cover slip with 3 μl of the reservoir solution mixed with 3 μl of the protein solution to a final protein concentration of approximately 10 mg/ml; the volume of the reservoir was 800 μl. Triclinic crystals appeared within a few days: they belong to space group P1 (a = 36.72 Å, b = 39.55 Å, c = 42.88 Å, α = 106.78°, β = 90.98°, γ = 104.72°) with two molecules in the unit cell.

2.4. Data collection, structure determination and refinement

The X-ray data were collected (maximum resolution 2.3 Å) on a Siemens area detector mounted on a Siemens rotating anode operating at 45 kV and 70 mA. CuKα radiation (λ = 1.5418 Å) was selected using a graphite monochromator. The data were processed using the XDS package of programs [15] to give a scaled set of intensities (Table 1). The structure was determined using the molecular replacement technique with the program AMoRe [16] and the refined human α-LA structure (1.7 Å resolution; [8]) as the starting model (75% sequence identity). After the initial rigid body refinement (R_{cryst} = 0.507), the buffalo α-LA amino acid sequence was incorporated into the model using the program O [17]. The resultant model was subjected to positional, simulated annealing and B-factor refinement using X-PLOR 3.1 [18] with the Engh and Huber force and geometric constants [19]. Non-crystallographic restraints were used at the beginning but were released towards the end of the refinement. SIGMAA [20] weighted $2F_o$-F_c and F_o-F_c electron density maps were used for assessing the quality of the structure and for identifying the positions of water molecules respectively. Attempts were made to define the conformation of the tripeptide at the C-terminus (residues 121–123) which were partially disordered in human and buffalo α-LA. The final refined model (R_{cryst} = 0.19 and R_{free} = 0.29 between 8.0 and 2.3 Å using all reflections) includes 1954 protein atoms, 2 Ca^{2+} ions and 96 water molecules. The atomic coordinates of buffalo α-LA will be deposited in the Brookhaven Protein Data Bank.

2.5. Electrospray mass spectrometric (ESMS) analysis

An aliquot (10 μl) of the HPLC purified α-LA was directly subjected to ESMS analysis using a VG BIO-Q triple quadrupole mass spectrometer (VG Biotech) equipped with an electrospray ion source at a flow rate of 10 ml/min. The quadrupole was scanned from 400 to 2000 *m/z* at 10 s/scan and the spectra were acquired using the MASS LINK software. Calibration was performed by a separate injection of lysozyme (M_r 14306.6 Da). All mass values were reported as average masses.

3. Results and discussion

The amino acid sequence (Fig. 1) obtained by chemical methods clearly showed one difference at position 17 between buffalo (Asp) and bovine (Gly) α-LAs. It has been confirmed by the X-ray structure of buffalo α-LA (Fig. 2A). Further, the crystal structure of buffalo α-LA showed another difference at position 27 (Ile) (Ile in human α-LA) (Fig. 2B). However, the Edman degradation results showed a Val both from the N-terminal sequence of the native (Fig. 3A) and of the digested protein (Fig. 3B). This disagreement has been explained using mass spectrometry analysis. The ES spectrum showed a bell-shaped distribution of multiple charged ions in the region between *m/z* 1100 and 1800. Two major components could be distinguished in the spectra whose molecular weight was measured as 14236.4 ± 0.9 Da and 14251 ± 1.5 Da, respectively. On the basis of the accurate molecular mass value, the first component was identified as the buffalo α-LA containing four disulphide bridges (theoretical molecular weight 14236 Da). The molecular weight of the second component was 14–15 Da higher than that of buffalo α-LA. The relative abundance of the two proteins was estimated from the ionic current as 69% for the first component and 31% for the second. This confirms the existence of a mixture of two genetic variants in vivo; the one containing Ile at position 27 is present in a slightly lower percentage than the second variant containing Val and thus it would not be detectable with Edman degradation. However, it appears that in the crystals of buffalo α-LA the Ile variant is more predominant as observed in the electron density map.

The overall features of the structure are similar to those reported for baboon and human α-LA [7–9] (Fig. 4). The root mean square (rms) deviation for the 120 equivalent C^{α} atoms for human and buffalo α-LA is 1.36 Å (average deviation 1.07 Å). The most rigid part of the molecule is the Ca^{2+} binding site which has similar conformation and ligand coordination as observed in other α-LA structures. The only significant deviation in the peptide backbone occurs in the region between residues 105 and 109 adjacent to the bottom portion of the cleft region (Fig. 4). In human and baboon α-LA structures this portion of the polypeptide adopts a distorted helical conformation whereas in buffalo α-LA structure it

Table 1
Data processing and refinement statistics for buffalo milk α-LA

Cell dimensions: a = 36.72, b = 39.55, c = 42.88 Å, α = 106.78, β = 90.98, γ = 104.72°	
Resolution	20.0–2.3 Å
N_m[a]	27282
N_u[b]	8940
Overall completeness	88.2%
$I/\sigma(I)$ (average)	19.3
R_{sym}[c]	7.3%
No. of reflections used in refinement (8.0–2.3 Å)	8743
R_{cryst}[d], ($F > 0\sigma$) 19.1%; R_{free}[e]	29.7%
No. of protein atoms (two molecules)	1954
No. of water molecules	96
rms deviation in bond lengths (Å)	0.012
bond angles (°)	1.66
Average B-factors (Å^2)	
backbone	14.8
all atoms	18.7
water molecules	32.1
calcium ion	14.5

[a]Number of measurements. [b]Number of unique reflections. [c]$R_{sym} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - < I(hkl) > |/\Sigma_{hkl}\Sigma_i I_i(hkl)$ where $< I >$ is the averaged intensity of the i observations of reflection hkl. [d]$R_{cryst} = \Sigma ||F_o| - |F_c||/\Sigma|F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. [e]$R_{free} = R_{cryst}$ for a randomly selected 5% subset of reflections not used in the refinement [24].

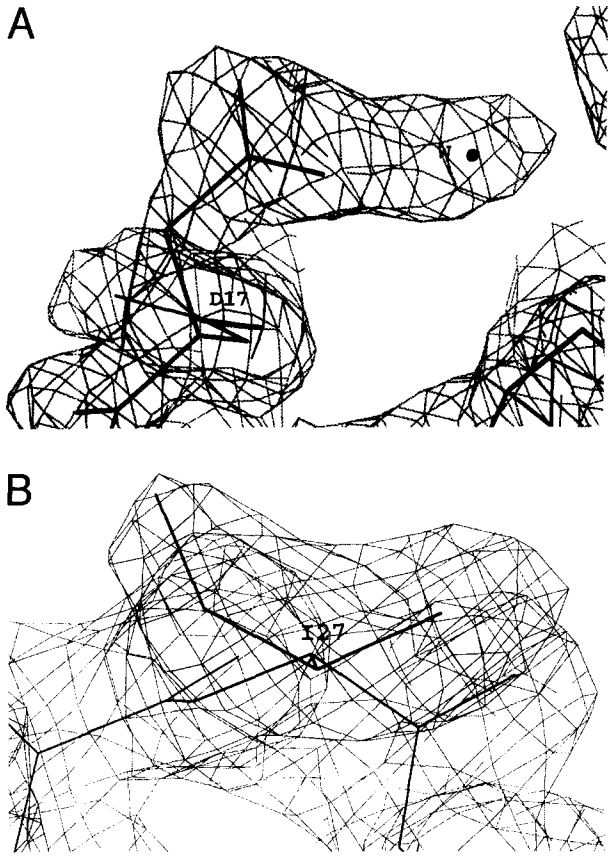
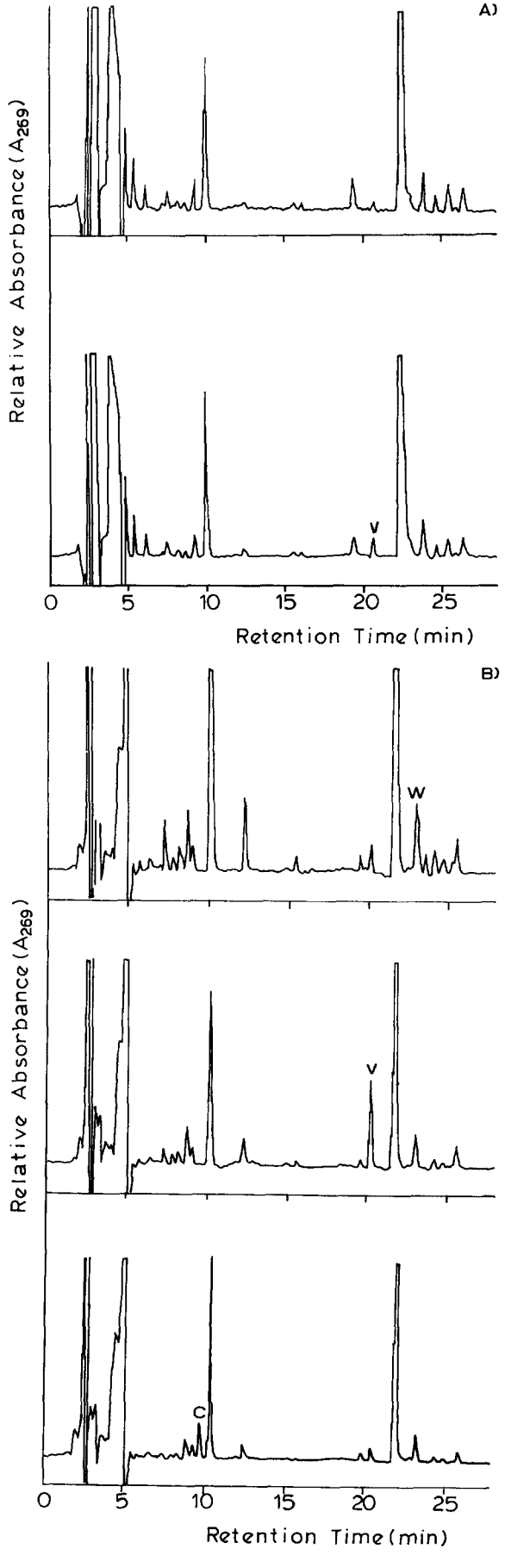


Fig. 2. Portions of the electron density map from buffalo α-LA refined structure at (A) Asp-17 and (B) Ile-27.

forms a 'flexible loop'. This local structural flexibility appears to be dependent on the crystallisation conditions as reported for human α-LA [9] – at high pH (6.5–7.5) the helical conformation is predominant (human and baboon α-LA) whereas at low pH (4.6), as in the case of buffalo α-LA, the loop structure is present. It has been observed before that at low pH, the hydrogen bond between the buried His-107 residue and the carboxylate group of Glu-25 (both are conserved in buffalo α-LA) is disrupted, making the fully protonated His solvent accessible and hence the loop structure.

The conformation of the flexible loop described here has particular interest in understanding α-LA function since it comprises part of the GT interaction site. Since the amino acid sequence of buffalo milk α-LA as described in this report is almost identical to the bovine sequence, we should be able to use the substantial amount of functional data available for the bovine protein. Based on mutagenesis data on bovine α-LA, amino acid substitutions at position 106, 107 and 110 have considerable effect on α-LA's ability to bind GT [21]. Furthermore, the helix to loop transition appears to invoke subtle changes in the environment of His-32 (increased accessibility in buffalo α-LA), an important residue implicated in



→

Fig. 3. A: Chromatogram of PTH-aa identification after automatic Edman degradation of native α-LA (residues 26 and 27). B: Chromatogram of PTH-aa identification of part of the peptide from V_8 protease digestion (residues 26, 27 and 28) at wavelength of 269 nm.

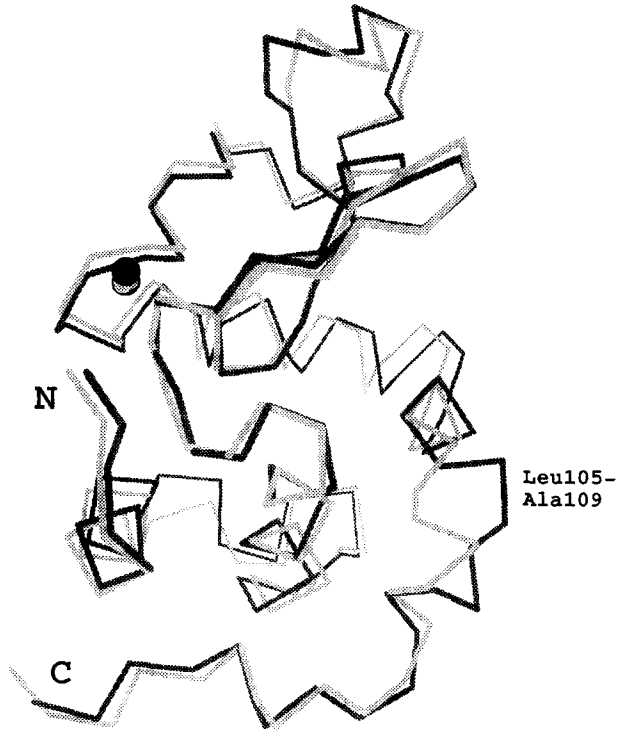


Fig. 4. A view of the C^{α} backbone of human α-LA (grey) and buffalo α-LA (black) is shown after least squares superposition. A portion of the molecule which differs in the two structures (residues 105–109), known to possess conformational flexibility (helix in human α-LA, [8]; coil in buffalo α-LA, present study) is highlighted. The Ca^{2+} ion is shown with large sphere. The figure was produced by MOLSCRIPT [25].

providing stabilising interactions for the bound monosaccharide when α-LA binds to GT. It is possible that conformational flexibility of this region might allow structural adjustments upon complex formation. The enhanced conformational flexibility of this particular region has also been observed in recent crystallographic analysis of the crystal structures of guinea pig, goat and bovine α-LA [22]. Interestingly, a structurally equivalent region of lysozyme has some flexibility (as judged from the crystallographic analyses of hen and human lysozymes [23]), but it does not undergo major rearrangements that are seen in α-LA. However, detailed structural information is required on the α-LA-GT complex to establish a definitive functional role for the 'flexible loop'.

Acknowledgements: This work was supported by a Leverhulme Trust research grant to K.R.A. and the Italian MURST and the British Council (UK) travel grant to K.R.A. and A.C. V.C. is supported by the Italian MURST through a protein chemistry Ph.D. course grant.

References

[1] Hill, R.L. and Brew, K. (1975) Adv. Enzymol. Rel. Areas Mol. Biol. 43, 411–489.
[2] Brew, K., Vanaman, T.C. and Hill, R.L. (1967) J. Biol. Chem. 242, 3747–3749.
[3] Hiroaka, Y., Segawa, Y., Kuwajima, K., Sugai, S. and Murai, N. (1980) Biochem. Biophys. Res. Commun. 93, 1098–1104.
[4] Kronman, M.J. (1989) CRC Crit. Rev. Biochem. Mol. Biol. 24, 565–667.
[5] Ren, J., Stuart, D.I. and Acharya, K.R. (1993) J. Biol. Chem. 268, 19292–19298.
[6] Brew, K. and Grobler, J.A. (1992) In: Advances in Dairy Chemistry-1 (P. Fox, Ed.), pp. 191–229. Elsevier, Amsterdam.
[7] Acharya, K.R., Stuart, D.I., Walker, N.P.C., Lewis, M. and Phillips, D.C. (1989) J. Mol. Biol. 208, 99–127.
[8] Acharya, K.R., Ren, J., Stuart, D.I., Phillips, D.C. and Fenna, R.E. (1991) J. Mol. Biol. 221, 571–581.
[9] Harata, K. and Muraki, M. (1992) J. Biol. Chem. 267, 1419–1421.
[10] Laemmli, U.K. (1970) Nature 227, 680–685.
[11] Towbin, H. and Gordon, J. (1984) J. Immunol. Methods 72, 313–340.
[12] Hewick, R.M., Hunkapillar, M.W., Hood, L.E. and Dreyer, W.J. (1981) J. Biol. Chem. 256, 7990–7997.
[13] Hunkapillar, M.W. and Hood, L.E. (1983) Methods Enzymol. 91, 486–493.
[14] De Laureto, P., De Filippis, V., Di Bello, M., Zambonin, M. and Fontana, A. (1995) Biochemistry 34, 12596–12604.
[15] Kabsch, W. (1988) J. Appl. Crystallogr. 21, 67–71.
[16] Navaza, J. (1994) Acta Crystallogr. D50, 760–763.
[17] Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Acta Crystallogr. A47, 110–119.
[18] Brünger, A.T. (1992) X-PLOR Manual – Version 3.1. A System for Crystallography and NMR. Yale University Press, New Haven, CT.
[19] Engh, R.A. and Huber, R. (1991) Acta Crystallogr. A47, 392–401.
[20] Read, R.J. (1986) Acta Crystallogr. A42, 140–149.
[21] Malinovskii, V.A., Tian, J., Grobler, J.A. and Brew, K. (1996) Biochemistry (in press).
[22] Pike, A.C.W., Brew, K. and Acharya, K.R. (1996) Structure 4, 691–703.
[23] Artymiuk, P.J., Blake, C.C.F., Grace, D.E.P., Oatley, S.J., Phillips, D.C. and Sternberg, M.J.E. (1979) Nature 280, 563–568.
[24] Brunger, A.T. (1992) Nature 355, 472–474.
[25] Kraulis, P.J. (1991) J. Appl. Crystallogr. 24, 946–950.