RESEARCH PAPERS

Acta Cryst. (1997). D**53**, 355–363

Structure of the Calcium-Binding Echidna Milk Lysozyme at 1.9 Å Resolution

J. MITCHELL GUSS,* MICHAEL MESSER, MICHAEL COSTELLO, KRISTINE HARDY AND VINAY KUMAR†

Department of Biochemistry – G08, University of Sydney, NSW 2006, Australia.
E-mail: m.guss@biochem.usyd.edu.au

(Received 24 October 1996; accepted 23 December 1996)

Abstract

A lysozyme isolated from the milk of a monotreme, the echidna, *Tachyglossus aculeatus multiaculeatus*, has been crystallized (space group $P2_1$, with unit-cell dimensions $a = 37.1$, $b = 42.0$, $c = 38.1$ Å, $\beta = 91°$ and $Z = 2$) and the structure refined to an R value of 0.167 for all measured data in the resolution range 7.0–1.9 Å. It had previously been inferred from sequence homology with α-lactalbumins that echidna milk lysozyme (EML) would bind one calcium ion per molecule. This has been confirmed in the present study in which the largest peak in a difference Fourier synthesis is associated with a calcium ion. The calcium binding site of EML is very similar to that observed in baboon and human α-lactalbumins, and in a human lysozyme engineered to contain a calcium-binding site. The overall fold of the protein is similar to that of chick-type lysozymes. EML, like pigeon lysozyme, has only 125 residues terminating at a cysteine but in EML this forms a disulfide with a cysteine at residue 9 whereas the equivalent cysteine residue in all other lysozymes of known sequence occurs at position 6. These changes cause some minor structural rearrangements. The binding of calcium appears to have had little effect on the polypeptide backbone conformation and caused only small changes in the conformation of side chains coordinating the calcium ion. A homology modelling study [Acharya, Stuart, Phillips, McKenzie & Teahan (1994). *J. Protein Chem.* **13**(6), 569–584] correctly predicted the overall structure of EML and the nature of its calcium binding site but generally failed to model some more subtle differences observed in the EML structure as evidenced by the fact that the homology model more closely resembles the starting structure from which the model was derived than it does the crystal structure.

1. Introduction

Hen egg-white lysozyme (HEWL), the first enzyme to have its structure determined by X-ray diffraction (Blake *et al.*, 1965), catalyzes the hydrolysis of the glycosidic bonds between *N*-acetylmuramic acid [2-acetamido-2-deoxy-3-*O*-(D-1-carboxyethyl)-D-glucopyranose] and *N*-acetylglucosamine (2-acetamido-2-deoxy-β-D-glucose) residues in bacterial cell wall polysaccharides. Lysozymes occur in a variety of species including birds and mammals and are found in a wide range of tissues and secretions. In contrast, α-lactalbumins are found only in the colostrum and milk of lactating mammals and their genes are expressed only in the mammary gland. The role of α-lactalbumin is to make glucose rather than *N*-acetylglucosamine the preferred substrate for galactosyltransferase, the other component of lactose synthase. In accordance with earlier predictions based on amino-acid sequence homology (Blake *et al.*, 1965), the structure of baboon α-lactalbumin was shown to be very similar to that of HEWL (Acharya, Stuart, Walker, Lewis & Phillips, 1989; Smith *et al.*, 1987; Stuart *et al.*, 1986).

α-Lactalbumin differs from lysozyme in a number of ways, including the absence of the acidic residues which are implicated in lysozyme's enzyme mechanism and in having a tightly bound calcium ion (Hiraoka, Segawa, Kuwajima, Sugai & Murai, 1980). The crystal structure of baboon α-lactalbumin (Acharya *et al.*, 1989; Stuart *et al.*, 1986) showed that the calcium binding site of this protein involves at least three conserved aspartate residues on a short loop of the polypeptide between the α and β domains. This calcium binding site is also seen in the human (Acharya, Ren, Stuart & Phillips, 1991), guinea-pig, caprine and bovine α-lactalbumins (Pike, Brew & Acharya, 1996).

Sequence homology of equine and pigeon lysozymes with α-lactalbumins suggested that these enzymes should also bind calcium (Stuart *et al.*, 1986) since they contained all of the aspartate residues necessary for calcium binding. It was subsequently confirmed that equine and pigeon lysozymes do bind one atom of calcium per molecule whereas human and chicken lysozyme and several other lysozymes, which lack these aspartate residues, do not (Nitta, Tsuge, Sugai & Shimazaki, 1987). Since then a number of other naturally occurring c-type lysozymes have been proposed to have a calcium binding site on the basis of their amino-acid sequences. These include the lysozymes of donkey

† Permanent address: Solid State Physics Division, BARC, Bombay 400085, India.

© 1997 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 © 1997

(Godovac-Zimmermann, Conti & Napolitano, 1988), dog, cat and hooded seal, (Grobler, Rao, Pervaiz & Brew, 1994) as well as echidna (Teahan, McKenzie, Shaw & Griffiths, 1991). It is generally assumed that all these lysozymes bind calcium, but so far this has been unequivocally demonstrated only for the equine, pigeon and canine enzymes.

What is the role of calcium in α-lactalbumin and in some lysozymes? Calcium is not located in the region of the protein thought to be involved in galactosyltransferase binding (Malinovskii, Tian, Grobler & Brew, 1996; Pike *et al.*, 1996), but may modify the kinetics of the reaction (Musci & Berliner, 1985). Calcium may protect α-lactalbumin from thermal denaturation (Hiraoka *et al.*, 1980) and act as a regulator of folding and disulfide bond formation (Rao & Brew, 1989). Since the α-lactalbumins of all species investigated have all the amino-acid residues required for calcium binding, it would be difficult to argue that calcium is not important for this protein, but the same argument cannot be applied to lysozymes.

There is agreement in the literature that α-lactalbumins and c-type lysozymes are derived from a common ancestor (Nitta & Sugai, 1989; Shewale, Sinha & Brew, 1984; McKenzie & White, 1991) and that α-lactalbumins arose by gene duplication from an ancestral lysozyme. The precise time of the gene duplication event is, however, still disputed. More recent studies, including the discovery of calcium-binding lysozymes and the sequencing of marsupial and monotreme proteins, have provided new evidence for the debate but have not resolved the question of whether the calcium-binding lysozymes are the genuine ancestors of the α-lactalbumins (Grobler *et al.*, 1994; Nitta & Sugai, 1989).

A number of structural studies have been undertaken with the aim of characterizing the calcium binding site in calcium-binding lysozymes. Horse and pigeon lysozymes have been crystallized and their structures have been determined at 2.5 Å (Tsuge *et al.*, 1992) and 3.0 Å resolution (Yao, Tanaka, Hikichi & Nitta, 1991), respectively, but so far the only high-resolution structural analysis of a calcium-binding lysozyme is that of a human lysozyme with an engineered calcium binding site. The double mutant Q86D/A92D human enzyme was studied in the calcium-bound and calcium-free form and compared with the structure of the wild-type enzyme (Inaka, Kuroki, Kikuchi & Matsushima, 1991). The calcium binding site was very similar to that found in α-lactalbumin. The effect of calcium binding was to cause a minor rearrangement of the side chains in the calcium-binding loop and the displacement of a water molecule from the site. No other significant changes were observed.

Echidna milk lysozyme (EML) is an attractive candidate for structural study. Echidnas are members of the Order Monotremata and, as egg-laying mammals, occupy a unique position in the evolution of the Class Mammalia. EML is the first monotreme or marsupial (*i.e.* non-eutherian) protein to have its three-dimensional structure determined. The amino-acid sequence of EML contains all of the residues which are found in the calcium-binding c-type lysozymes (Teahan *et al.*, 1991) and EML is the first naturally occurring lysozyme to have its structure determined at high resolution with a bound atom of calcium. EML is unusual in that its carboxyl terminal is a cysteine residue; this feature is not found in any other mammalian lysozyme. The most N-terminal cysteine residue is thus far uniquely found at position 9 in EML instead of at position 6 as in all other c-type lysozymes. The consequences of calcium binding and these differences in primary and tertiary structures of EML provide points of interest for a crystal structure analysis. There are two variants of EML, designated as I and II, which are present in the milk of the Kangaroo Island, South Australia, echidna, *T. a. multiaculeatus* and the New South Wales (mainland) echidna, *T. a. aculeatus*, respectively (Teahan *et al.*, 1991). EML I and II differ at three locations: 13V/A, 37S/G and 41R/Q. In this paper we present the structure of EML-I with a bound calcium ion refined at 1.9 Å resolution.

2. Materials and methods

2.1. *Isolation of echidna lysozyme I*

Milk samples from free-living lactating echidnas (*Tachyglossus aculeatus multiaculeatus*) were collected on Kangaroo Island, South Australia as described previously (Griffiths, Green, Leckie, Messer & Newgrain, 1984) and stored at 253 K for up to 6 months and then at 193 K for 2 years. Three samples collected from two echidnas in 1991 were pooled (20 g), diluted with 20 ml of 50 m*M* Tris–HCl buffer (pH 7.4) containing 0.1 *M* KCl and centrifuged at 40 000*g* (60 min, 277 K). The fat was discarded and the supernatant (whey) was concentrated to 16 ml by centrifugation using micro-concentrators with 3000 molecular weight cutoff (Microsep^{TM}, Filtron Technology Corp, MA, USA).

Lysozyme has been previously isolated from echidna whey by a one-step procedure involving Heparin–Sepharose chromatography (Teahan *et al.*, 1991). In this work, in which maximum purity of the lysozyme was considered desirable, the following four-step procedure was adopted. The echidna whey was applied to a column (2.5 × 118 cm) of Sephadex G-100 (Pharmacia) and eluted with 50 m*M* Tris–HCl/0.1 *M* KCl (pH 7.4). The whey proteins resolved into four peaks of which only the fourth, containing the low molecular weight proteins, exhibited lysozyme activity as determined by a turbidometric method (Weisner, 1984). The column fractions constituting peak 4 were combined, concentrated to 6 ml as described above and applied to a column (1.6 × 40 cm) of DEAE–Sephadex

Table 1. *Data processing and merging statistics*

		Crystal 1			Crystal 2			Combined data		
D_{min} (Å)	D_{min} (Å)	% Complete	Mean $I/\sigma(I)$	R_{sym}	% Complete	Mean $I/\sigma(I)$	R_{sym}*	% Complete	Mean $I/\sigma(I)$	R_{merge}*
17.0	4.08	76.3	10.8	0.062	74.7	17.5	0.046	87.4	9.2	0.061
4.08	3.24	80.5	10.2	0.078	79.6	15.8	0.044	90.9	8.7	0.070
3.24	2.84	83.8	8.3	0.070	81.0	9.4	0.071	94.0	6.6	0.077
2.84	2.58	85.2	6.5	0.080	82.2	6.6	0.102	94.3	5.0	0.086
2.58	2.39	86.4	5.6	0.091	83.3	5.1	0.127	95.9	4.2	0.096
2.39	2.25	86.7	5.2	0.088	82.8	4.2	0.142	95.8	3.6	0.108
2.25	2.14	86.2	4.2	0.100	83.7	3.2	0.162	90.2	2.7	0.155
2.14	2.05	87.7	3.4	0.119	—	—	—	87.7	2.1	—
2.05	1.97	87.7	2.6	0.131	—	—	—	87.7	1.5	—
1.97	1.90	75.7	1.9	0.142	—	—	—	75.7	1.3	—
Overall		84.3	7.2	0.077	80.9†	10.3	0.064	90.0‡	5.6	0.074

* R_{sym} or $R_{merge} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$; † Overall values for crystal 2 refer to the data with $17.0 \geq d \geq 2.2$ Å; ‡ The completeness and mean $I/\sigma(I)$ values for the combined data include all measurements to the 1.9 Å limit of the resolution. The merging R value refers to the overlapping data between the two crystals in the range $17.0 \geq d \geq 2.2$ Å.

(Pharmacia) equilibrated with 50 m*M* Tris–HCl pH 7.4. All of the lysozyme activity eluted with the initial, unretarded protein peak (*c.f.* Shaw, Messer, Scrivener, Nicholas & Griffiths, 1993), the contents of which were combined and filtered through a small column (1.0 × 6.0 cm) of Heparin–Sepharose CL-6B (Pharmacia) which had been equilibrated with 5 m*M* barbital sodium buffer (pH 7.4) containing 50 m*M* NaCl. The column was washed with buffer solution to remove unbound proteins, after which the bound proteins were eluted using a linear salt gradient to 1.0 *M* NaCl (Teahan *et al.*, 1991). This yielded a single unsymmetrical protein peak eluting between 0.15 and 0.30 *M* NaCl. Examination by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) (Shaw *et al.*, 1993) showed that although lysozyme was present throughout this peak, the leading-edge fractions were contaminated with proteins of higher molecular weight. These fractions were discarded while the others were pooled, concentrated, diluted with starting buffer and re-chromatographed on Heparin–Sepharose as above. The column fractions which contained only lysozyme protein (SDS–PAGE) were pooled and dialysed for 24 h against a slight molar excess of $CaCl_2$ using dialysis tubing with 6000–8000 molecular weight cutoff; this was followed by dialysis (24 h) against three changes of ultrapure water (18 MΩ) and freeze-drying. The final yield of freeze-dried EML I was 24.5 mg.

2.2. *Crystallization and data collection*

Crystals of EML were grown in 6 µl hanging drops initially containing a 1:1 mixture of protein solution and well solution (protein solution: 10 mg ml^{-1} protein, 10 m*M* citrate, pH 4.5; well solution: 0.05 *M* KH_2PO_4/K_2HPO_4, pH 7.4, 20% PEG 8000). These conditions are one of the set represented in a modification of a published sparse-matrix screen (Jancarik & Kim, 1991). Elongated flat-plate crystals grew up to 0.5 × 0.2 × 0.05 mm in one week at room temperature. The pH in the drop after the crystals had grown was measured to be 6.34. All data were recorded on an R-AXIS IIc imaging-plate detector mounted on a Rigaku RU-200 rotating-anode generator with a Cu $K\alpha$ target, operating at 50 kV 100 mA with a 0.3 mm cathode. The X-ray beam was monochromated and focused with a mirror system incorporating a 0.00038 cm Ni filter originally developed by Zbyszek Otwinowski and Gerald Johnson at Yale University, supplied by MSC Corporation, Texas. Diffraction data were recorded from two crystals to a resolution of 1.9 Å for one crystal and to 2.2 Å for the other using a crystal-to-image-plate distance of 100 mm. The data were processed with the *HKL* suite of programs, *DENZO* and *SCALEPACK* (Otwinowski, 1993). No reflections were omitted from the data because they were weak. The crystals are monoclinic, space group $P2_1$, with unit-cell dimensions. $a = 37.1$, $b = 42.0$, $c = 38.1$ Å, $\beta = 91°$ and $Z = 2$, obtained from the average values found from the two crystals. The scaling and merging statistics are given in Table 1. The final data set comprises 8418 unique reflections corresponding to 90% of the possible data to 1.9 Å resolution.

2.3. *Structure solution and refinement*

The structure of EML was determined using molecular replacement with a model for echidna lysozyme derived from the structure of hen egg-white lysozyme [HEWL; Protein Data Bank entry, 4LYZ; (Diamond, 1974)] by making the necessary amino-acid changes using *INSIGHT* software (MSI/Biosym Inc, 1993). The sequences were aligned as shown in Fig. 1 following the alignment published by Acharya *et al.* (Acharya, Stuart, Phillips, McKenzie & Teahan, 1994). Atoms in the EML model with no direct equivalents in HEWL were omitted from the initial calculations of structure factors. This model was then subjected to a molecular dynamics/energy minimization protocol with the program *X-PLOR* (Brünger, 1992) prior to the molec-

ular replacement searches which were also made with *X-PLOR*. The correct rotation-function solution was 15% larger than the largest noise peak at 3.0 and 3.5 Å resolution. The translation in the x and z directions was determined by both translation- and packing-function searches which yielded a consistent solution. Rigid-body refinement reduced the conventional R value from 0.47 to 0.42 ($6 \geq d \geq 3.5$ Å) and R_{free}, calculated for 10% of the data, from 0.47 to 0.46. Several cycles of simulated annealing with *X-PLOR*, difference Fourier syntheses with the program package *CCP*4 (Collaborative Computational Project, Number 4, 1994) and model refitting with the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991) were followed by restrained refinement *PROLSQ* (Hendrickson & Konnert, 1980) and model refitting (Table 2).

The calcium ion was identified with the largest peak in a difference Fourier synthesis (Fig. 2) at an early stage in the structure analysis, more than 10σ and twice the height of the next largest peak. The calcium ion was not included in the model until later in the refinement when solvent was being included. In the absence of chemical analysis, the identity of the peak as a calcium ion could only be inferred indirectly, but refinement as a calcium ion resulted in a B value of 16.2 Å^2, similar to that of the coordinating atoms, and the hepta-coordinate geometry is consistent with calcium and not favoured by other cations of similar atomic number.

Table 2. *Refinement statistics*

Number of residues	125
Number of atoms	1068
Solvent atoms	65
$\langle B \rangle$ main-chain atoms (Å^2)	18.0
$\langle B \rangle$ side-chain atoms (Å^2)	22.8
$\langle B \rangle$ solvent atoms (Å^2)	34.6
$\langle B \rangle$ Ca^{2+} ion (Å^2)	16.2
Resolution range (Å)	7.0–1.9
R value $= \sum \lvert \lvert F_o \rvert - \lvert F_c \rvert \rvert / \sum \lvert F_o \rvert$ (No. of reflections)	0.169 (8308)
R_{free} (No. of reflections)	0.229 (852)

3. Results and discussion

3.1. *Quality of the structure*

The quality of the structure was assessed using the program *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993). In a Ramachandran plot, not shown, (Ramachandran, Ramakrishnan & Sasisekharan, 1963), 90.1% of residues have backbone torsion angles in most favoured regions. No residues lie in disallowed or generously allowed regions of the plot. Other measures of quality applied by *PROCHECK* indicate that the refined parameters, including ω-angle standard deviation, bad contacts, chirality, hydrogen-bond energy and side-chain torsion angles are better or inside the limits derived from well refined structures of comparable resolution. The theoretical estimate of coordinate errors using a Luzzati plot (Luzzati, 1952) is 0.2 Å, and 0.18 Å from *SIGMAA* (Read, 1990).

3.2. *Overall structure and the catalytic active site*

The structure of EML (Fig. 3) shows the fold which is known to be archetypal for c-type lysozymes and for α-lactalbumins. The structure is comprised of two domains. The larger domain, predominantly α-helical, is composed of polypeptide segments from the N- and C-terminal regions, residues 1–40 and 80–125, respectively. An analysis of the structure using *PROMOTIF* (Hutchinson & Thornton, 1996) indicates that this domain contains four α-helices: Lys5–Ala14; Leu25–Ser36; Ile89–Ala102 and Val109–Lys114. There are also two shorter 3_{10} helices: Cys80–Leu83 and Ser121–Phe123. The smaller domain contains a three-stranded antiparallel β-sheet formed from the central portion of the sequence, residues 41–79 (Fig. 3). In all c-type lysozymes the active site lies in a deep cleft between the two domains. The two essential conserved acidic residues, Glu35 and Asp53 in EML, are at the base of the cleft (Fig. 3).

There are four disulfide bonds in EML: Cys9—Cys125; Cys30—Cys115; Cys64—Cys80 and Cys76—Cys94. The first of these is the only one not spatially or sequentially equivalent to those in other c-type lysozymes. The first cysteine from the N-terminus normally occurs at position 6, approximately one turn along the first α-helix.

3.3. *The calcium ion binding site*

The calcium ion is centred in a loop of the polypeptide formed by residues 82–91 (Fig. 4) and is coordinated by the side-chain carboxyl groups of three aspartic acid residues, two main-chain carbonyl groups and two water molecules (Fig. 4; Table 3). The atoms which coordinate the calcium ion form a remarkably regular pentagonal bipyramid (Table 3). The atoms forming the central plane of the bipyramid, Asp85 $O^{\delta 1}$, Asp90 $O^{\delta 2}$, Asp91 $O^{\delta 2}$, Wat128 O and Wat130 O, deviate from a least-squares plane drawn through the atoms by –0.010, 0.017, –0.022, 0.021 and –0.005 Å, respectively. The calcium ion lies 0.04 Å from this plane. The angles at the calcium ion formed by adjacent ligand atoms in the ring range from 66 to 77° compared with the ideal value of 72° for a regular pentagon (Table 3). The apical calcium-ligand bonds are nearly perpendicular to the central plane.

3.4. *Comparison with other lysozyme structures*

While the overall structure of EML closely resembles those of other members of the c-type lysozyme family there are some notable differences (Figs. 5 and 6). There are displacements of C^{α} positions of up to 5 Å when the structure of EML is superimposed on that of HEWL. The largest differences occur at the carboxy terminus

Table 3. *Geometry of the calcium-ion binding site*

(*a*) Distances from Ca^{2+}

Ligand	EML		3LHM*		1HML*		1ALC*	
group	Residue	Distance (Å)	Residue	Distance (Å)	Residue	Distance (Å)	Residue	Distance (Å)
Carbonyl	K82	2.32	A83	2.24	K79	2.26	K79	2.22
Carboxylate	D85	2.29	D86	2.41	D82	2.44	D82	2.36
Carbonyl	D87	2.31	N88	2.39	D84	2.28	D84	2.27
Carboxylate	D90	2.37	D91	2.44	D87	2.43	D87	2.32
Carboxylate	D91	2.49	D92	2.40	D88	2.42	D88	2.29
Water	128H†	2.32	58H	2.44	2H	2.57	77H	2.42
Water	130H	2.56	69H	2.90	1H	2.58	63H	2.60

(*b*) Angles at Ca^{2+} in EML

Atom 1	Atom 2	Angle (°)	Atom 1	Atom 2	Angle (°)
D85 $O^{\delta 1}$	128H	73	D90 $O^{\delta 2}$	D91 $O^{\delta 2}$	77
D85 $O^{\delta 1}$	130H	74	D91 $O^{\delta 2}$	128H	70
130H	D90 $O^{\delta 2}$	66	K82 O	D87 O	165

* 3LHM, 1HML and 1ALC are the Protein Data Bank references for the coordinates for holo mutant (Q86D/A92D) human lysozyme (Inaka *et al.*, 1991), human milk α-lactalbumin (Ren, Stuart & Acharya, 1993) and baboon milk α-lactalbumin (Acharya *et al.*, 1989). † Residues *n*H denote solvent molecules in the Protein Data Bank entry.

based on sequence homology with other lysozymes and α-lactalbumins (Acharya *et al.*, 1994). They used HEWL as a starting structure for their echidna lysozyme model. The calcium-binding loop and the loops 19–26 and 31–33 were based on the structure of human α-lactalbumin. A comparison of their model for echidna lysozyme and the structure reported here (Fig. 6*c*) reveals that the model captured the important features of the structure, including the movement at the carboxyl terminus. While the model failed to position correctly residue 103 the authors noted in their paper that region 101–103 forms an external loop and in spite of poor amino-acid sequence agreement was built as in hen egg-white lysozyme. These gross features of the model structure were correctly predicted because of the implications of changes in the sequence, notably the position of the disulfide bonds.

The next step in the modelling procedure which involved energy minimization might be said to have been less successful. Fig. 6 reveals that the model of echidna lysozyme more closely resembles HEWL (r.m.s.d. for C^{α} atoms is 0.82 Å) than it does the actual structure of echidna lysozyme EML (r.m.s.d. for C^{α} atoms is 1.18 Å). The structure which most closely resembles EML is PEWL (r.m.s.d. for C^{α} atoms is 1.03 Å). This is not surprising since the lysozyme sequence which is closest to echidna is pigeon (Grobler *et al.*, 1994), despite their distant evolutionary relationship. In fact PEWL has most of the eccentric characteristics of EML, including the short C-terminal strand terminating in a cysteine residue and the calcium-binding loop, and only lacks the shift in the sequence of the most N-terminal cysteine.

4. Concluding remarks

This study has confirmed that echidna milk lysozyme is a calcium-binding protein. Perhaps the most important finding is that calcium binding has almost no effect on its structure, confirming the observations made for the calcium-bound and calcium-free forms of a double mutant form of human lysozyme in which residues 86 and 92 were replaced by aspartates (Kuroki *et al.*, 1989). Other than changes to side-chain orientations directly associated with calcium binding, there were no significant changes to the backbone of the calcium-binding loop (Inaka *et al.*, 1991). The demonstration that human lysozyme, which has only one of the calcium-binding aspartates (at position 91) could be transformed to bind calcium simply by changing two residues to aspartates suggests that the necessary and sufficient condition for the ability of echidna lysozyme to bind calcium seems to be to have aspartate residues at positions 85, 90 and 91 (EML numbering).

It is of interest that the human D86/92 lysozyme was more stable against thermal denaturation and protease digestion than the native lysozyme (Kuroki *et al.*, 1989). In addition, the removal of calcium from the mutant enzyme decreased its stability, possibly because it resulted in electrostatic repulsion between Asp86 and Asp91 (Inaka *et al.*, 1991). A possible implication of these observations is that during the evolution of the non-stomach c-type lysozymes, loss of calcium binding through a mutation involving replacement of one of the three essential aspartates would be hindered by the fact that the product would be less stable. Conversely, acquisition of calcium binding by a non-calcium binding lysozyme might be favoured because the product would be more stable. It has been proposed (Grobler *et al.*, 1994) that calcium-binding activity is an ancient feature of the lysozyme-α-lactalbumin superfamily which has been lost from the non-calcium-binding gene line; however, this seems improbable if the loss of calcium significantly decreases the stability of these proteins under physiological conditions. In that case it would seem more likely that calcium-binding lysozymes such

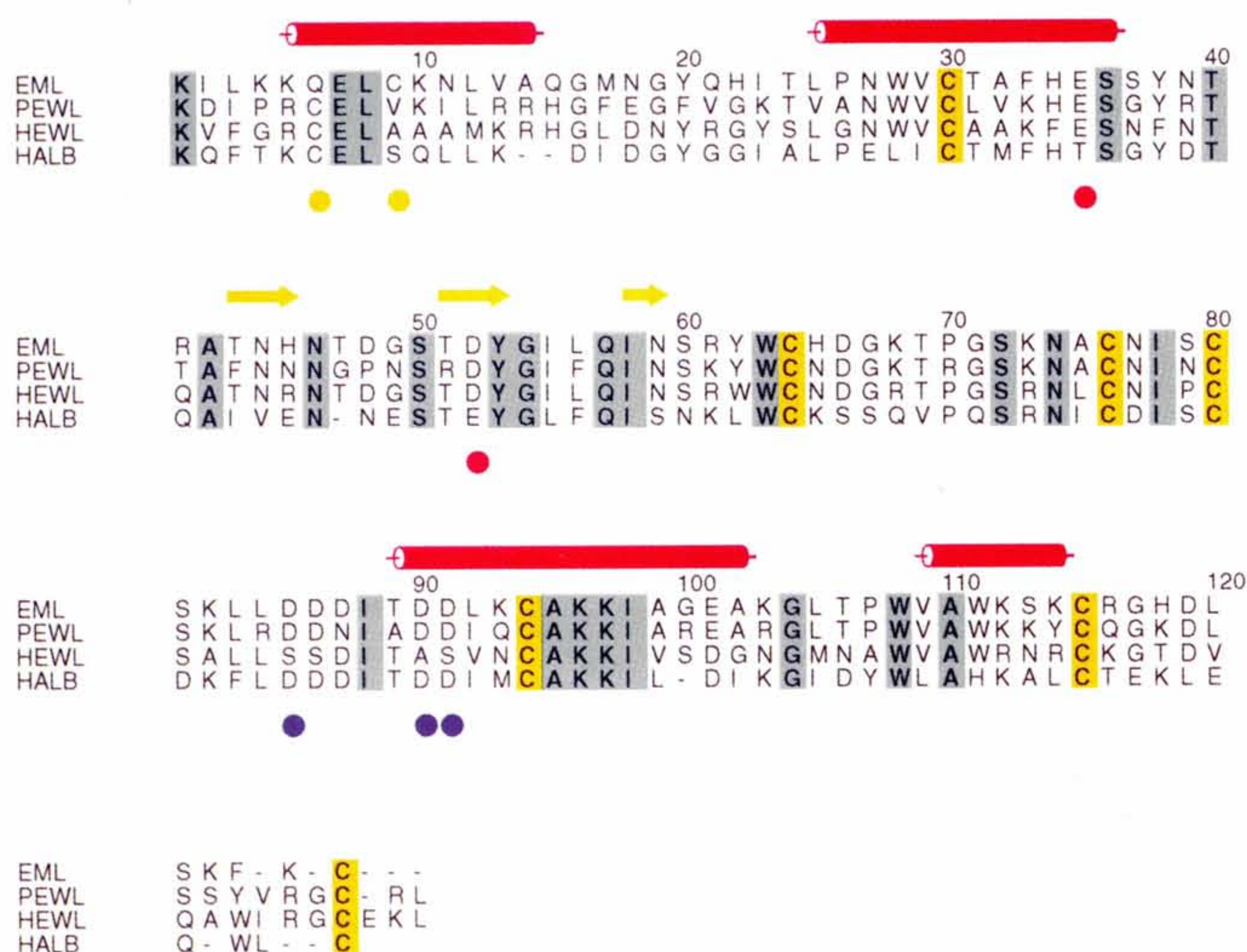

Fig. 1. Aligned amino-acid sequences for echidna milk lysozyme (EML), pigeon egg-white lysozyme (PEWL), hen egg-white lysozyme (HEWL) and human α-lactalbumin (HALB). Sequence identity in the four sequences is indicated by cyan shading and yellow shading for cysteine residues. The red cylinders and yellow arrows indicate the location of α-helices and β-sheets found in EML. The red spheres indicate the catalytically essential acidic residues found in all c-type lysozymes. The blue circles indicate the residues contributing side-chain atoms to the calcium coordination in EML, PEWL and HALB. The yellow circles indicate the remaining cysteine residue which is not aligned. This diagram was prepared with the program *ALSCRIPT* (Barton, 1993).

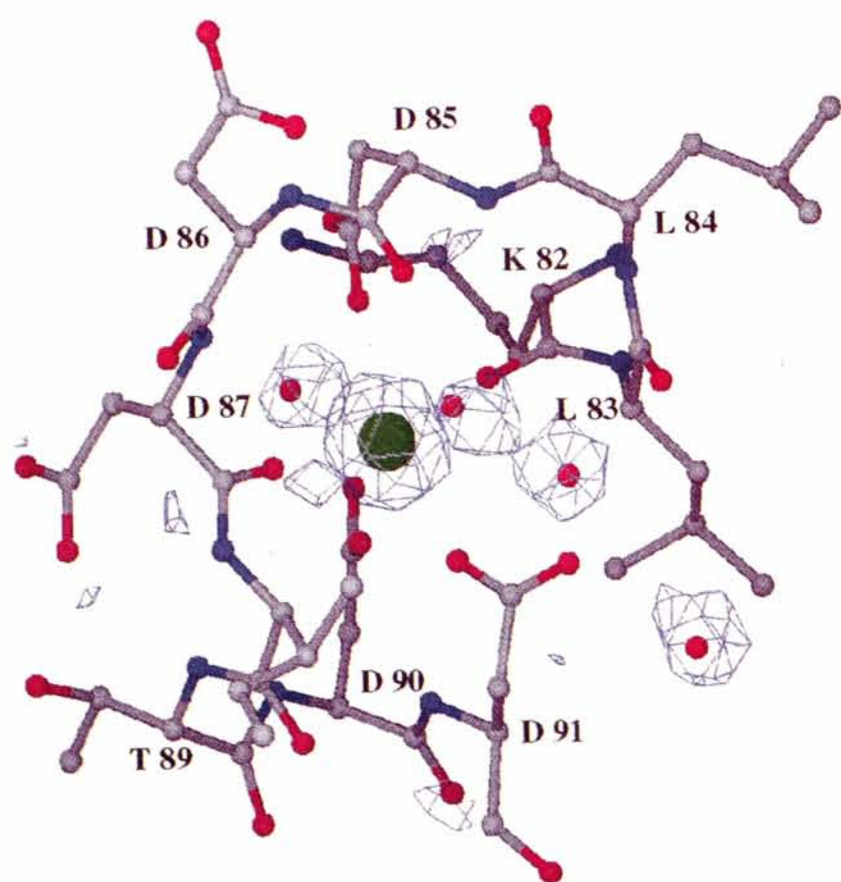

Fig. 2. Refined structure superimposed on the $(F_o - F_c)$ electron density. The calcium ion and water molecules had never been included at this stage of the refinement. The calcium ion is represented by a green sphere and the water molecules by red spheres.

and in the loop centred on residue 103. The former is a result of the shortening of the polypeptide chain and the relative movement of the Cys9—Cys125 disulfide bond and the associated polypeptide backbone (Fig. 5). It is of interest that while the S atoms of this disulfide bond in HEWL and HALB superimpose almost exactly, the C^{α} atoms of the cysteine residues are quite far apart (Fig. 5). In general, the carboxy terminal strand of EML more closely resembles that of HALB than of HEWL, despite a greater sequence similarity to the latter. Another significant difference between EML and HEWL occurs at Lys103 which lies on the link between the third and fourth α-helices at the open end of the active-site cleft. The equivalent region in pigeon egg-white lysozyme (PEWL) is very similar to that in EML. Another difference between HEWL and EML which is significant but not as large as those already mentioned involves the region from residues 4 to 20 (Fig. 6*a*). In this region EML is also different from PEWL (Fig. 6*b*) and PEWL and HEWL are also different from each other, implying that this whole region is free to move with respect to the remainder of the structure.

Calcium binding is accommodated with little or no effect on the structure. Specifically, the structure of the calcium-binding loop in EML and the equivalent residues in HEWL are very similar (Fig. 4), implying that calcium binding does not significantly alter the backbone conformation of the enzyme. Calcium binding also appears to have little or no effect on the mobility of the protein as measured by the B values. The loop which binds the calcium shows only a slightly elevated average B value in common with other loops.

3.5. *Molecular modelling of the calcium-binding lysozymes*

Acharya and colleagues have modelled the calcium-binding lysozymes from echidna, horse and pigeon

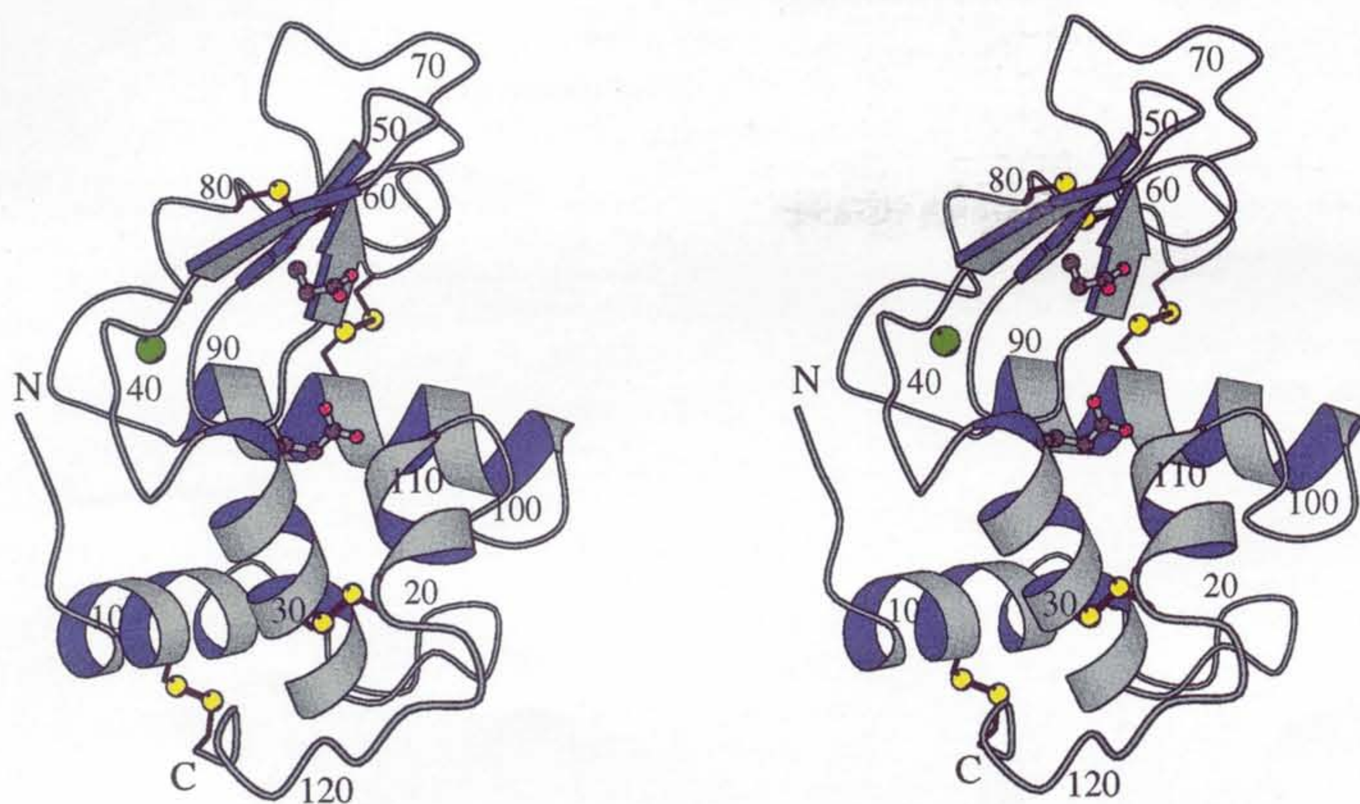

Fig. 3. Stereoview of the structure of echidna lysozyme. The disulfide bonds are represented by yellow balls, the calcium ion by a green ball and the catalytically essential acidic residues, Glu35 and Asp52, as their side chains in ball-and-stick representation with the O atoms in red. N and C represent the amino and carboxyl termini of the polypeptide chain, respectively. Figs. 3, 4 and 5 were prepared with the program *MOLSCRIPT* (Kraulis, 1991).

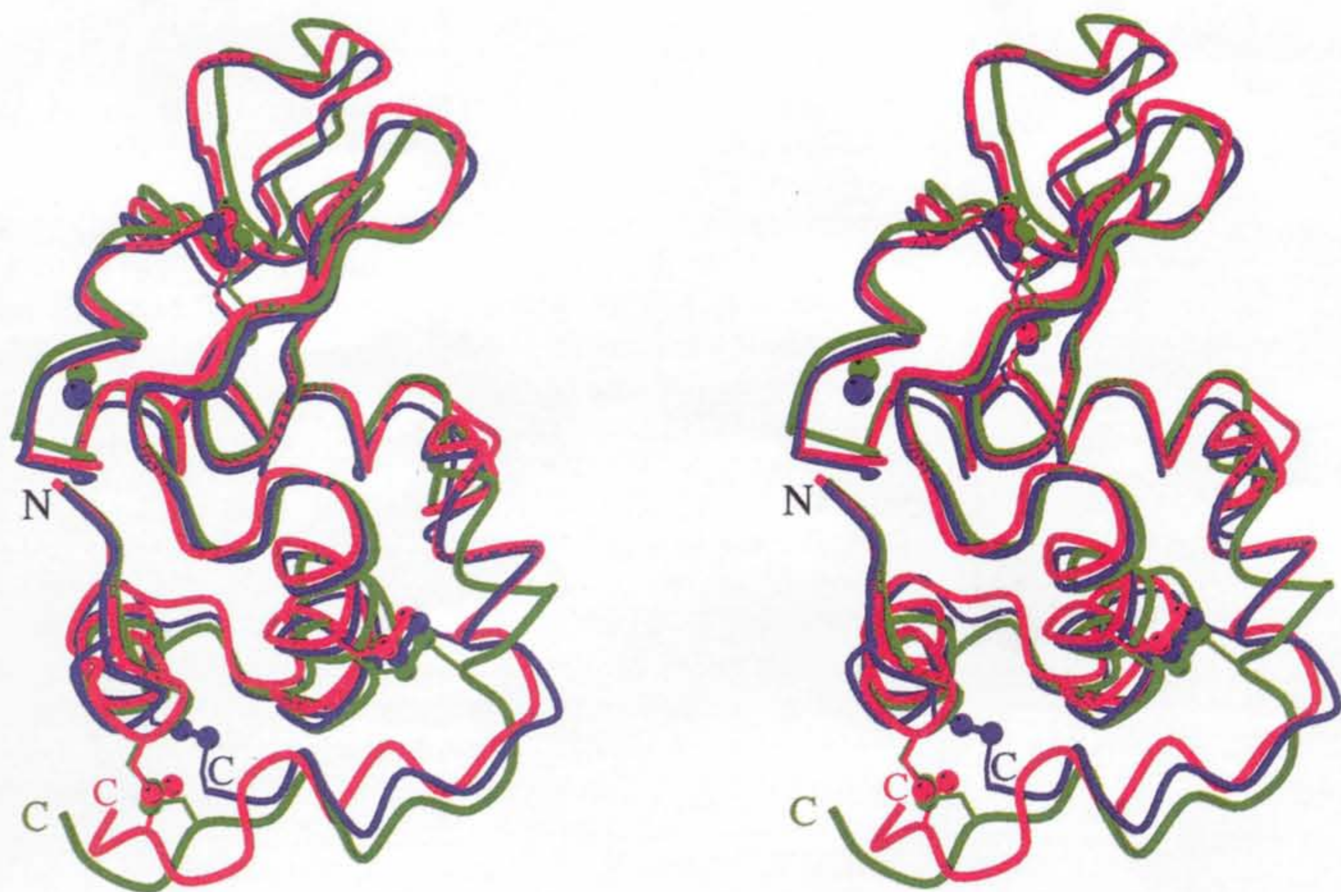

Fig. 4. Stereoview showing a superposition of the calcium-binding loops in echidna lysozyme (blue) and human α-lactalbumin (green) and the corresponding residues of hen egg-white lysozyme (red). The residues contributing to the calcium-ion coordination in echidna lysozyme are labelled.

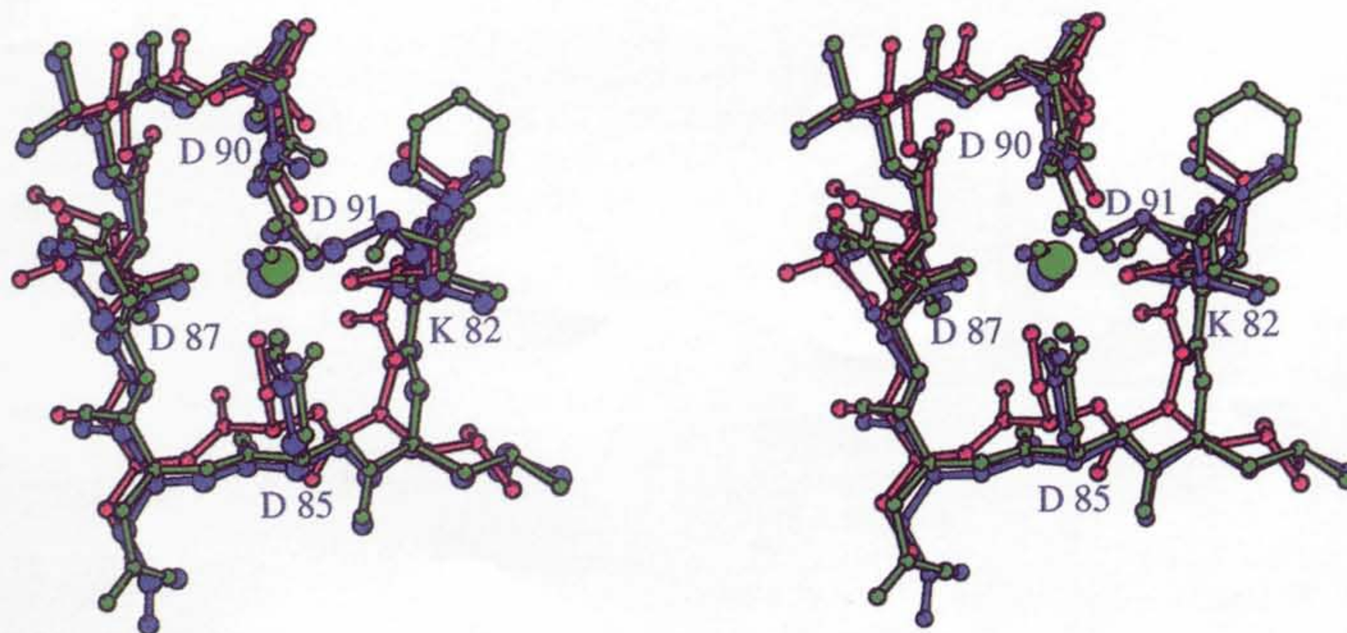

Fig. 5. Stereoview showing a superposition of the backbone traces for echidna lysozyme (blue), hen egg-white lysozyme (red) and human α-lactalbumin (green). The disulfide bonds are represented by spheres indicating the S-atom positions.

as EML acquired this property *via* appropriate mutations in one or more non-calcium-binding lysozymes.

We are greatly indebted to Dr Peggy Rismiller (Pelican Lagoon Research and Wildlife Centre, Kangaroo Island, South Australia) for the samples of echidna milk, to Professor K. Nitta, Hokkaido University, Japan for the coordinates of pigeon egg-white lysozyme, Dr R. Acharya for the coordinates of his model of echidna milk lysozyme and to Drs D. C. Shaw and G. B. Ralston for helpful comments on the manuscript. VK was supported by an ARC grant to H. C. Freeman and JMG.*

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1JUG, R1JUGSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: HE0188).

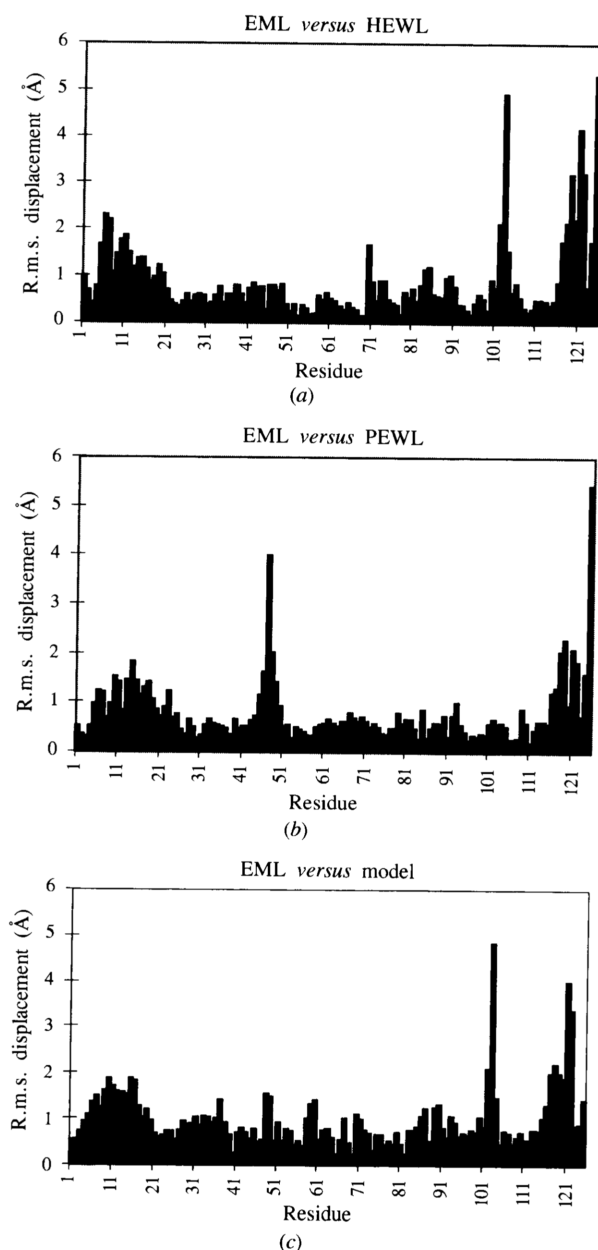

Fig. 6. Plots of r.m.s.Δ (Å) of C^{α} atoms for pairwise comparisons of structures. (*a*) EML *versus* HEWL. (*b*) EML *versus* PEWL. (*c*) EML *versus* model of EML (Acharya *et al.*, 1994).

References

Acharya, K. R., Ren, J., Stuart, D. I. & Phillips, D. C. (1991). *J. Mol. Biol.* **221**, 571–581.
Acharya, K. R., Stuart, D. I., Phillips, D. C., McKenzie, H. A. & Teahan, C. G. (1994). *J. Protein Chem.* **13**, 569–584.
Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M. & Phillips, D. C. (1989). *J. Mol. Biol.* **208**, 99–127.
Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.
Biosym Inc (1993). *Insight.* Biosym Inc, San Diego, California, USA.
Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965). *Nature (London)*, **206**, 757–761.
Brünger, A. T. (1992). *X-PLOR*, Version 3.1, *A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT, USA.
Collaborative Computational Project, Number 4. (1994). *Acta Cryst.* D**50**, 760–763.
Diamond, R. (1974). *J. Mol. Biol.* **82**, 371–391.
Godovac-Zimmermann, J., Conti, A. & Napolitano, L. (1988). *Hoppe-Seyler's Biol. Chem.* **369**, 1109–1115.
Griffiths, M., Green, B., Leckie, R. M. C., Messer, M. & Newgrain, K. W. (1984). *Aust. J. Biol. Sci.* **37**, 323–329.
Grobler, J. A., Rao, K. R., Pervaiz, S. & Brew, K. (1994). *Arch. Biochem. Biophys.* **313**, 360–366.
Hendrickson, W. A. & Konnert, J. H. (1980). *Biomolecular Structure, Function, Conformation and Evolution*, edited by R. Srinivasan, pp. 43–75. Oxford: Pergamon Press.
Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S. & Murai, N. (1980). *Biochem. Biophys. Res. Commun.* **95**, 1098–1104.
Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.
Inaka, K., Kuroki, R., Kikuchi, M. & Matsushima, M. (1991). *J. Biol. Chem.* **266**, 20666–20671.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kuroki, R., Taniyama, Y., Seko, C., Nakamura, H., Kikuchi, M. & Ikehara, M. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 6903–6907.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
McKenzie, H. A. & White, F. H. J. (1991). *Adv. Protein Chem.* **41**, 173–315.
Malinovskii, V. A., Tian, J., Grobler, J. A. & Brew, K. (1996). *Biochemistry*, **35**, 9710–9715.
Musci, G. & Berliner, L. J. (1985). *Biochemistry*, **24**, 6945–6948.
Nitta, K. & Sugai, S. (1989). *Eur. J. Biochem.* **182**, 111–118.

Nitta, K., Tsuge, H., Sugai, S. & Shimazaki, K. (1987). *FEBS Lett.* **223**, 405–408.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Pike, A. C. W., Brew, K. & Acharya, K. R. (1996). *Structure*, **4**, 691–703.

Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.

Rao, K. R. & Brew, K. (1989). *Biochem. Biophys. Res. Commun.* **163**, 1390–1395.

Read, R. J. (1990). *Acta Cryst.* A**46**, 140–149.

Ren, J., Stuart, D. I. & Acharya, K. R. (1993). *J. Biol. Chem.* **268**, 19292–19298.

Shaw, D. C., Messer, M., Scrivener, A. M., Nicholas, K. R. & Griffiths, M. (1993). *Biochim. Biophys. Acta*, **1161**, 177–186.

Shewale, J. K., Sinha, S. K. & Brew, K. (1984). *J. Biol. Chem.* **259**, 4947–4956.

Smith, S. G., Lewis, M., Aschaffenburg, R., Fenna, R. E., Wilson, I. A., Sundaralingham, M., Stuart, D. I. & Phillips, D. C. (1987). *Biochem. J.* **242**, 353–360.

Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M. & Phillips, D. C. (1986). *Nature (London)*, **324**, 84–87.

Teahan, C. G., McKenzie, H. A., Shaw, D. C. & Griffiths, M. (1991). *Biochem. Int.* **24**, 85–95.

Tsuge, H., Ago, H., Noma, M., Nitta, K., Sugai, S. & Miyano, M. (1992). *J. Biochem. (Tokyo)*, **111**, 141–143.

Weisner, B. (1984). *Methods in Enzymatic Analysis*, Vol. IV, edited by H. V. Bergmeyer, 3rd. ed., pp. 189–193. Weinheim: Verlag-Chemie.

Yao, M., Tanaka, I., Hikichi, K. & Nitta, K. (1991). *J. Biochem. (Tokyo)*, **111**, 1–3.