ELSEVIER

FEBS Letters 342 (1994) 154–158

FEBS 13803

FEBS LETTERS

Crystal structure of momordin, a type I Ribosome Inactivating Protein from the seeds of *Momordica charantia*

Jasmine Husain*, Ian J. Tickle, Stephen P. Wood

Department of Crystallography, Birkbeck College, Malet Street, London, WC1E 7HX, UK

Received 24 January 1994; revised version received 20 February 1994

Abstract

A type I ribosome-inactivating protein, extracted and purified from *M. charantia* seeds, was crystallised by vapour diffusion with polyethylene glycol at pH 7.2. X-ray data were collected to 2.1 Å resolution and the structure solved by molecular replacement using the A-chain coordinates of ricin. The overall fold of the protein is similar to ricin but there are differences in secondary structure, on the surface and in the active site cleft. These differences are probably due in part to the evolution of the protein without a B-chain partner. The most extensive reorganisation occurs at the C-terminus whereas Tyr^{70} shows the greatest change in the active site cleft.

Key words: *N*-Glycosidase; Enzyme; Ribosome-inactivating protein; Ricin; Trichosanthin; Pokeweed antiviral protein

1. Introduction

A number of higher plants produce proteins capable of inactivating ribosomes. Type II RIPs, such as ricin from castor bean, comprise two protein subunits of similar size (approx. M_r 30,000) linked by a disulphide bridge. The A-chain component is a highly specific *N*-glycosidase which excises a single base from a loop structure within the 28 S RNA component of the eukaryotic ribosome [1,2]. The B-chain component is a galactose-specific lectin which recognises cell surface glycoproteins and glycolipids and is mainly responsible for the endocytic internalisation of ricin and delivery of the A-chain component to the cytoplasm. Both the A- and B-chains of ricin are glycosylated and recognised by endocytic mannose receptors, providing a second route of cell entry [3]. Type II RIPs are thus extremely cytotoxic. Type I RIPs comprise only an A-chain component and exhibit much lower cytotoxicity. Indeed a number of these materials are components of oriental herbal medicines. Much current interest exists in targeting the A-chain cell-killing function of RIPs to particular cell types by conjugation to antibodies and hormones [4]. Furthermore evidence exists that trichosanthin and momordin (the type I RIPs from *Trichosanthes kirilowii* Maximowicz and *M. charantia*) show selective toxicity to T-cells and macrophages infected by HIV [5,6].

There is therefore considerable interest in the structure and function of these proteins. Significant progress has been made by Robertus and co-workers in defining the 3D structure of intact ricin [7], recombinant ricin A-chain [8], pokeweed antiviral protein [9], and various substrate analogue complexes, and in proposing a mechanism for the glycosidase activity [10,11]. In this paper we describe the 3D structure of momordin, a type I RIP derived from the seeds of the bitter gourd, *M. charantia*, which shows only 34% sequence similarity with ricin A-chain. The structural basis for the as yet incompletely defined variability in properties of type I RIPs may be an important factor in understanding their function and in engineering more selective cytotoxins.

2. Experimental

Momordin is an abundant constituent of *M. charantia* seeds and was purified using elements of previously published methods [13–15]. Dried seeds (10 g), obtained from Dhaka, Bangladesh, were powdered and dispersed in phosphate-buffered saline (60 ml). Solids and oils were removed by filtration over Hyflo Supercel filter aid. Proteins were precipitated from the filtrate with ammonium sulphate (60% saturation at 4°C). Following centrifugation (48,000 × *g* for 30 min), the protein pellet was dissolved in 50 mM phosphate buffer pH 6.3 and gel-filtered over Sephadex G50 (2.6 × 97 cm) equilibrated in the same buffer. Fractions from a well resolved peak (M_r 30,000, absorbance at 280 nm) were pooled, concentrated by ultrafiltration (M_r 10,000 cut-off membrane) and the buffer changed to 20 mM sodium acetate, pH 4.5. Final purification was achieved by cation exchange chromatography on a 10 ml MonoS column at pH 4.5 where the major constituent eluted by a sodium chloride gradient (120 ml, 0–0.5 M salt) was centred at 240 mM

* Corresponding author. Fax: (44) (71) 436 8918.

Abbreviations: RIP, ribosome-inactivating protein; RTA, ricin A-chain; RTC, recombinant ricin; MOM, momordin; PAP, pokeweed antiviral protein; FMP, formycin monophosphate; ApG, adenyl(3′-5′)guanosine; PAGE, polyacrylamide gel electrophoresis. $R_{merge} = \Sigma_h \Sigma_i | I_{hi} - \langle I_h \rangle | / \Sigma_h \Sigma_i I_{hi}$; $R = \Sigma_h (| | F_{obs} | - | F_{calc} | |) / \Sigma_h | F_{obs} |$

SSDI 0014-5793(94)00188-2

salt (yield 100 mg). Progress of the purification was monitored by SDS-PAGE.

The purified protein was desalted and concentrated to 20 mg/ml (absorbance at 280 nm = 1.3; 1 mg/ml; 1 cm path) by Centricon filtration in 20 mM phosphate buffer, pH 7.4. Crystallisation conditions were screened by hanging-drop vapour diffusion methods. X-ray data were collected using a MARResearch image plate mounted on a Siemens XP-18 rotating anode generator operating at 40 kV and 100 mA. Molecular replacement procedures were performed using the coordinates of the ricin A-chain as the search model. Normalised structure-factor amplitudes were employed in a fast cross-rotation function [16]. The positioned ricin model in the momordin unit cell was subjected to rigid body refinement. The atomic model was refined with geometrical restraints against the observed amplitudes using RESTRAIN [17]. Only common side chains were included initially, with others set as alanine. The main chain density in early difference maps was fragmented into 17 sections. Multiple rounds of refinement using both XPLOR [18] and PROLSQ [19] were interspersed with computer-graphics rebuilds using FRODO [20], gradually building in the side chains of the momordin sequence [21]. Once a satisfactory description of the protein density was complete, water molecules were added. The refined model was compared with ricin A-chain, recombinant ricin A-chain, pokeweed antiviral protein and their complexes with formycin monophosphate and dinucleotides [10].

3. Results and discussion

Well-formed rhombohedral crystals grew over 2 weeks equilibration against 20–40% w/v polyethylene glycol 4000 at pH 7.4 ± 0.2. Crystal constants and X-ray data collection statistics are recorded in Table 1. This form is similar to that reported previously [22]. Inclusion of lactose as a crystallisation additive (10 mM) produced isomorphous crystals with an altered morphology and improved diffraction characteristics although no evidence of bound disaccharide was eventually found. The refined electron density clearly matched the amino acid sequence of momordin 1 [21], except at residue 64 where it is a leucine instead of a valine.

The cross-rotation function using the coordinates of ricin A-chain, data from 8 to 3 Å resolution and a radius of integration of 25 Å provided a solution 6.5 σ above the mean and 1.6 σ above the highest random-noise peak. Fourier coefficients of the T_2-function [23] were computed and the translation function map showed a clear solution 4.1 σ above the highest random-noise peak. The rotation and translation parameters provided an acceptable packing arrangement in the momordin cell. Rigid body refinement resulted in $R = 0.496$, and correlation coefficient = 0.285 for data between 10.0 and

Table 1
X-ray diffraction data for momordin

Space group	*R3*
Cell constants	$a = 131.17$ Å, $c = 40.75$ Å
Number of molecules in asymmetric unit	1
Average redundancy	4.5
Number of unique reflections used	13591
Maximum resolution	2.16 Å
R_{merge}	0.100
Completeness	96.9%

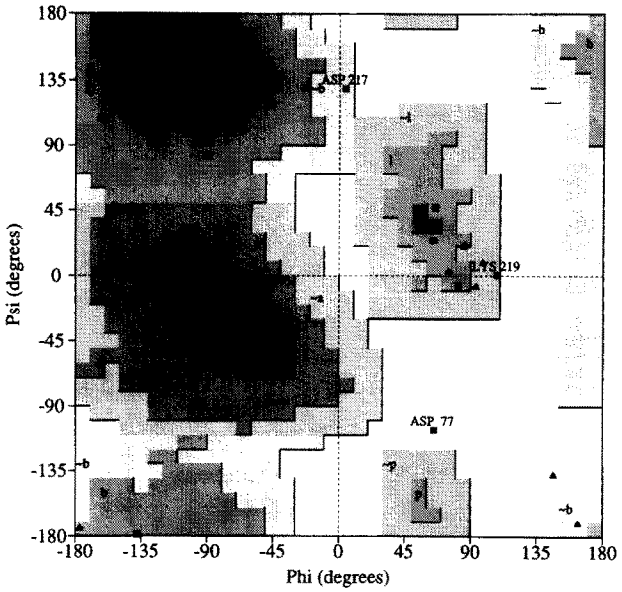

Fig. 1. A Ramachandran plot showing main chain (ϕ,φ) conformational angles of momordin. The glycine residues are shown as triangles. All residues but one (Asp^{77}) are within the allowed regions of the Ramachandran plot.

2.8 Å resolution. The final momordin model, which included 197 water molecules, gave $R = 0.187$ for all data between 8.0 and 2.16 Å resolution. Table 2 lists least-squares refinement statistics and Fig. 1 is a Ramachandran plot [24] (produced by PROCHECK [25]) showing (ϕ,φ) values of momordin. Fig. 2 shows Asp^{77} in a very tight type II β-turn normally occupied by glycine (as is the equivalent residue in ricin), with (ϕ,φ) = (65°,−100°) in a normally disallowed region. Similar occurrences have been observed elsewhere [26,27].

Ribbon diagrams [28] of the secondary structures of momordin and ricin are shown in Fig. 3a and b. Ricin A-chain has a greater accessible surface area than momordin (by 1,369 Å^2); the ribbon diagram illustrates that this is a consequence of sequence insertions in surface loops and the presence of extensive secondary struc-

Table 2
Least squares refinement statistics of momordin

Number of protein atoms	1958
Number of solvent atoms	194
Number of carbohydrate atoms	28
Resolution range	8.0–2.16 Å
Number of reflections used in refinement	13591
rms deviation from target values:	
Bond lengths	0.016 Å
Bond angles	0.047 Å
Peptide planarity	0.060 Å
Mean B-values:	
All protein atoms	28.6 Å^2
Main chain atoms	26.7 Å^2
Side chain atoms	30.5 Å^2
Solvent atoms	53.4 Å^2
R-factor	0.187

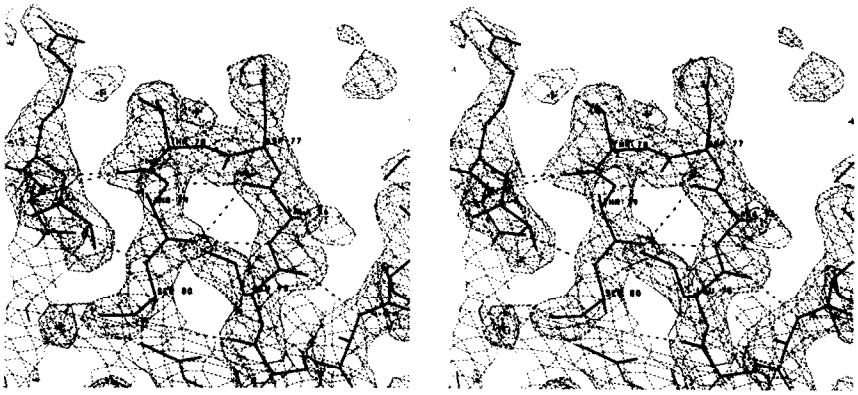

Fig. 2. A stereo view of Asp^{77} and the $(2 \mid F_{obs} \mid - \mid F_{calc} \mid)$ electron density contoured at 1 σ.

ture elements near the surface. The published sequence [21] for momordin contains 263 amino acids, but there is electron density for only 246 of them. The missing residues are located at the C-terminus. This may either be due to disorder or to proteolytic processing. A similar truncation has been reported for trichosanthin [29]. Of the 15 available sequences for RIPs from plant and bacterial origins, momordin shows the closest similarity (60%) with trichosanthin. Unlike trichosanthin, however, momordin is glycosylated at Asn^{227} [30]. Electron density is visible for the first *N*-acetyl glucosamine on strand **h** but there is weaker density for the second ring which, being adjacent to an extensive solvent channel on the crystal 3_1 axis, is probably disordered.

Robertus and co-workers have described the ricin fold in three regions [7]. A 6-stranded (**a**–**f**) twisted sheet with four central antiparallel strands and parallel outer strands is flanked by two helices, **A** and **B**. In momordin there are two additional antiparallel strands which, in keeping with the prime nomenclature of ricin we refer to as **a′** and **a″**. The central region of the molecule is made up of the eight helices, **A**–**H**, which enclose the active site cleft. Helices **A**, **B**, **D**, and **G** are shorter in momordin, whereas in the C-terminal region additional irregular helical regions occur. Residues 216–221 of the loop connecting strands **g** and **h** and residues 242–246 have hydrogen bond patterns and dihedral angles approaching that of α-helix. The most regular region 230–235 is shown in the ribbon diagram in Fig. 3a.

A number of residues in the active site cleft are identical across the RIP family and mutagenesis experiments [10,31] have confirmed their importance to the enzyme mechanism [11]. Structural studies of ricin and PAP complexed with formycin 5′-monophosphate, and ricin with the dinucleotide ApG show the adenine base sandwiched between Tyr^{80} and Tyr^{123} (ricin numbering). A mechanism has been proposed which focuses on Glu^{177} and Arg^{180} as central catalytic residues. The organisation of these residues is very similar in momordin with the exception of Tyr^{70} (Tyr^{80} in ricin). This side chain occupies

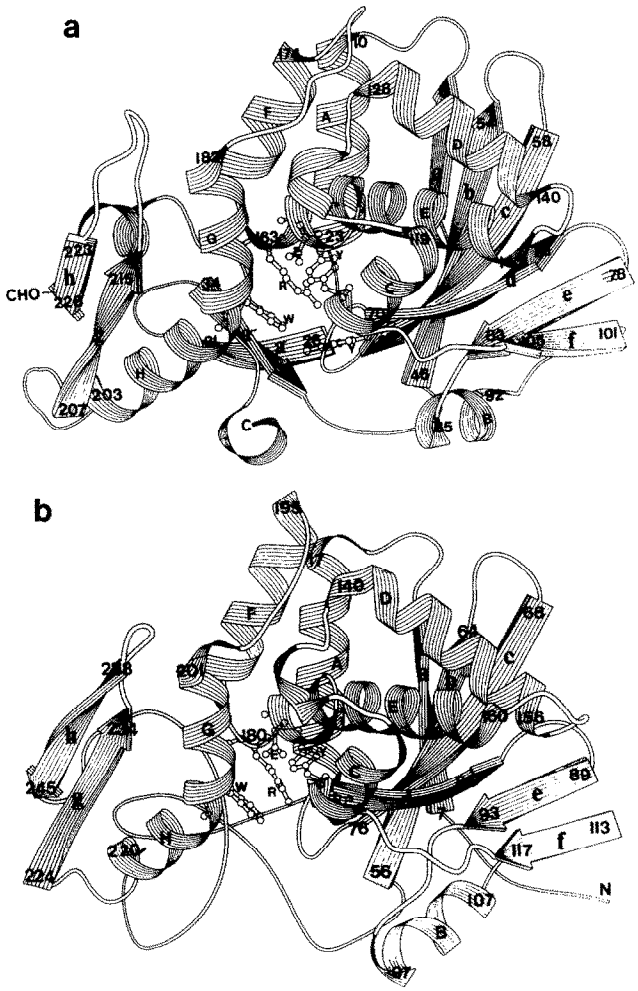

Fig. 3. A ribbon diagram showing the very similar secondary structures of (a) momordin and (b) recombinant ricin A-chain despite their low sequence similarity (34%). The carbohydrate attachment point in momordin is indicated as CHO. Side chains of active site residues Tyr^{70} = Y, Glu^{160} = E, Arg^{163} = R and Trp^{192} = W in momordin (and their equivalents in ricin) are also labelled.

a range of positions [8–10,12] in intact ricin, recombinant ricin A-chain, PAP and their complexes. Although the

rms fit of 215 Cα of the above mentioned RIPs (shown in Fig. 4) is only 0.9 Å, the rms fit of Tyr^{70} and its equivalents alone is 0.6 Å. This suggests that some of these differences may be related to an activation event on release of the ricin B-chain. The ring position in momordin most closely resembles that in PAP, which is also a type I RIP. The tyrosine ring effectively screens the active site Trp^{192} from solvent in momordin whereas it is exposed in other unliganded structures. The electron density for Trp^{192} also indicates that this side chain may adopt two different conformations. The ring plane is retained in both conformations but the ring is flipped over to provide H-bonding from NE2 either to $Leu^{239}CO$ or Tyr^{70} via a water molecule. When we model FMP into the momordin cleft it is necessary to disturb the *syn* conformation observed in ricin complexes to accommodate Tyr^{70}. The significance of this observation is not clear as it has been previously noted that FMP is not an effective inhibitor of ricin, although it is an excellent transition state mimic inhibitor for AMP nucleosidase, an enzyme that catalyses an analogous reaction. Furthermore, crystals of ricin, recombinant ricin A-chain, PAP and momordin grow at pH values between 4.8 and 8.9. Both PAP and momordin crystallise at about the same pH and have the most similar organisation for Tyr^{70}, suggesting that pH might be a significant and not unexpected determinant of the detailed active site geometry. Although the central active site residues are identical across the family, adjacent residues involved in defining their positions do vary to some degree. Glu^{160} and Arg^{163} in momordin (equivalents of 177 and 180 in ricin) protrude into the cleft from a distorted C-terminal segment of helix **E**. This segment is defined by a break in the continuous hydrogen bonding pattern of the long central helix **E**; a stereo view of helix **E** is shown in Fig. 5. The pattern break is clearest at Thr^{157} the carbonyl of which forms a hydrogen bond with the conserved Tyr^{14} side chain from nearby helix **A**. This arrangement and the degree of deviation of the dihedral angles is very close to that observed in ricin. The first residue beyond this kink

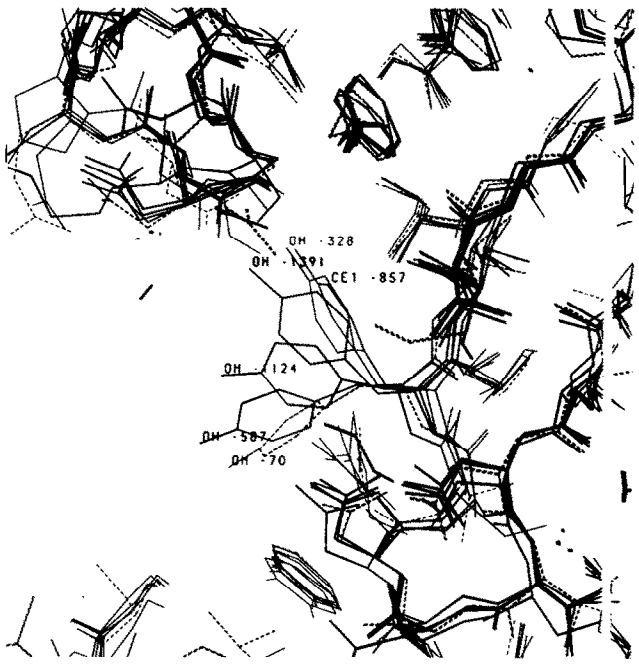

Fig. 4. Superimpositions of the 3D structures of MOM, PAP, RTC, FMP, APG and RICIN are shown. 215 Cαs of the RIPs were least-squares fitted by MNYFIT [32]; the rms fit is 0.9 Å. Tyr^{70} in the active site of momordin is compared with those of PAP (numbered 587), ricin complexes APG (1,124) and FMP (1,391), recombinant ricin A-chain (328) and ricin A-chain (857). Momordin is most similar to PAP.

in helix E in ricin, Ser^{176}, does not cap the continuing helix but forms a buried hydrogen bond with $Ala^{79}CO$ of the **c/d** loop which visits the cleft. The equivalent residue Ala 159 in momordin cannot make this stabilising hydrogen bond and this region of the helix may as a result enjoy more flexibility. Furthermore, the substitution of buried hydrophobic residues Met^{174} and Ile^{175} in ricin by buried hydrogen-bonded threonines in momordin may also lead to stability differences in the C terminus of helix **E**. Thr^{157} and Thr^{158} side chains and NHs share hydrogen bonds with $Val^{153}CO$ and $Leu^{154}CO$. It remains to be established what variability in catalytic

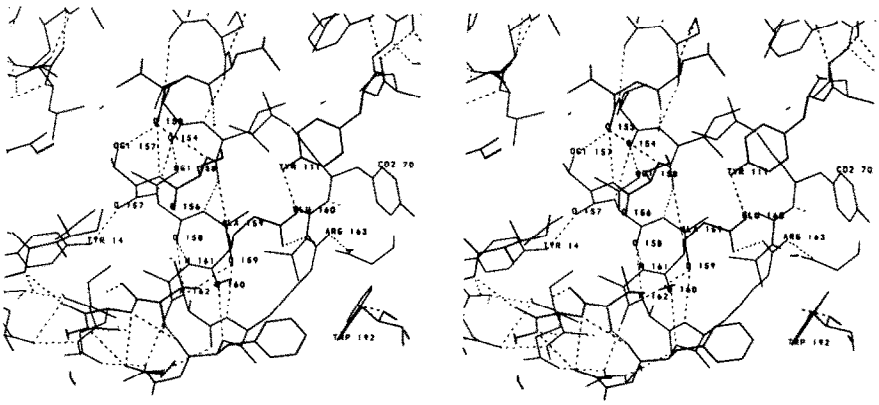

Fig. 5. A stereo view of the C-terminal region of helix **E** showing the distinctive hydrogen-bonding pattern of momordin adjacent to the active site residues Arg^{163} and Glu^{160}.

constants, if any, is shown by different RIPs, which may be related to these differences in active site organisation.

The interchain disulphide bond in the ricin heterodimer is formed through the A-chain Cys^{259}. This is located in the extended chain of the C-terminal region (Fig. 3b). Preceding this cysteine is a stretch of 10 hydrophobic amino acid residues, some of which are buried in the isolated A-chain and others which become buried on interaction with the B-chain. It has been suggested that this region might be involved in the translocation of the A-chain through membranes. In momordin, residues at these positions corresponding to exposed hydrophobics in the ricin A-chain are either buried by modifications to the fold or substituted by hydrophilic side chains. Equivalent modifications are also observed in PAP. Both proteins have evolved to a more stable solvent interaction surface and have lost the cysteine residue.

There are some indications that in spite of the strong conservation of the catalytic residues among members of the RIP family, there exists a degree of specificity in the action of RIPs on ribosomes from different species. Ricin A-chain, for instance, does not attack *E. coli* ribosomes whereas PAP does. This is not unexpected in that ribosome binding probably involves regions of the RIPs remote from the active site. Details of the 3D structure in surface regions are likely to become of increasing interest to those engineering toxins of defined target specificity, particularly with respect to recognition of virally infected cells.

Acknowledgements: We thank J.D. Robertus for supplying coordinates in advance of publication; M. Moshiuzzaman of Chemistry Department of Dhaka University for procuring the *Momordica* seeds; and colleagues at Birkbeck, especially D. Venugopal and G. Louie, for discussions on model building and refinement. J.H. acknowledges a grant from MRC.

References

[1] Roberts, W.K. and Selitrennikoff, C.P. (1985) Biosc. Rep. 6, 19–29.
[2] Stripe, F. and Barbieri, L. (1986) FEBS Lett. 195, 1–8.
[3] Magnusson, S. and Berg, T. (1993) Biochem. J. 291, 749–755.
[4] Bolognesi, P.C., Gromo, G., Lento, P., Mistza, G., Sciumbata, T., Stripe, F. and Modena, D. (1993) Cancer Immunol. Immunother. 36, 346–250.
[5] McGrath, M.S., Hwang, K.M., Caldwell, S.E., Gaston, I., Luk, K-C., Wu, P., Ng, V.L., Crowe, S., Daniels, J., Marsh, J., Deinhart, T., Lekas, P.V., Vennari, J.C., Yeung, H.W. and Lifson, J.D. (1989) Proc. Natl. Acad. Sci. USA 86, 2844–2848.
[6] Huang, S.L., Huang, P.L., Nara, P.L., Chen, H-C., Kung, H-F., Huang, P., Huang, H.I. and Huang, P.L. (1990) FEBS Lett. 272, 12–18.
[7] Montfort, W., Villafranca, J.E., Monzingo, A.F., Ernst, S.R., Katzin, B., Rutenber, E., Xuong, N.H., Hamlin, R., and Robertus, J.D. (1987) J. Biol. Chem. 262, 5398–5403.
[8] Mlsna, D., Monzingo, A.F., Katzin, B., Ernst, S.R. and Robertus, J.D. (1993) Protein Sci. 2, 429–435.
[9] Monzingo, A.F., Collins, E.J., Ernst, S.R., Irvin, J.D. and Robertus, J.D. (1993) J. Mol. Biol. 233, 705–715.
[10] Monzingo, A.F. and Robertus, J.D. (1992) J. Mol. Biol. 227, 1136–1145.
[11] Ready, M.P., Kim Youngsoo, K. and Robertus, J.D. (1991) Proteins: Struct. Funct. Genet. 10, 270–278.
[12] Kim, Y. and Robertus, J.D. (1992) Protein Eng. 5, 775–779.
[13] Barbieri, L., Stoppa, C. and Bolognesi, A. (1987) J. Chromatogr. 408, 235–243.
[14] Cunnick, J.E., Sakamoto, K., Chapes, S.K., Fortner, G.W. and Takemoto, D.J. (1990) Cell. Immunol. 126, 278–289.
[15] Ng, T.B., Chan, W.Y. and Yeung, H.W. (1992) Gen. Pharmacol. 23, 575–590.
[16] Crowther, R.A. (1972) in: The Molecular Replacement Method (Rossmann, M.G., ed.) pp. 173–178, Gordon & Breach.
[17] Haneef, I., Moss, D.S., Stanford, M.J. and Borkakoti, N. (1985) Acta Crystallogr., A41, 426–433.
[18] Brunger, A.T., Krukowski, A. and Erickson, J. (1990) Acta Crystallogr., A46, 585–593.
[19] Hendrickson, W.A. and Konnert, J.H. (1980) in: Computing in Crystallograllography (Diamond, R., Ramasheshan, S. and Venkatesan, K., eds.) pp. 13.01–13.23, Indian Acad. Sci., Bangalore.
[20] Jones, T.A. (1978) J. Appl. Crystallogr. 11, 268–272.
[21] Feng, Z., Li, W.W. and Yeung, H.W. (1990) J. Mol. Biol. 214, 525–626.
[22] Ho, W.K.K., Liu, S.C., Shaw, P.C., Yeung, H.W., Ng, T.B. and Chan, W.Y. (1991) Biochem. Biophys. Acta 1088, 311–314.
[23] Driessen, H.P.C., Bax, B., Slingsby, C., Lindley, P.F., Mahadevan, D., Moss, D.S., Blundell, T.L. and Tickle, I.J., (1991) Acta Crystallogr. B47, 987–997.
[24] Ramakrishnan, C. and Ramachandran, G.N. (1965) Biophys. J. 5, 909–933.
[25] Laskowski, R., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–291.
[26] Verschueren, K.H.G., Franken, S.M., Rozeboom, H.J., Kalk, K.H., and Dijkstra, B.W. (1993) J. Mol. Biol. 232, 856–872.
[27] Morris, A.L., MacArthur, M.W., Hutchinson, E.G. and Thornton, J.M. (1992) Proteins: Struct. Funct. Genet. 12, 345–364.
[28] Priestle, J.P. (1988) J. Appl. Crystallogr. 21, 572–576.
[29] Chow, T.P., Feldman, R.A., Lovett, M. and Piatak, M. (1990) J. Biol. Chem. 265, 8670–8674.
[30] Kimura, Y.K., Minami, Y., Tokuda, T., Nakajima, S., Takagi, S. and Funatsu, G. (1991) Agri. Biol. Chem. 55, 2031–2036.
[31] Endo, Y., Gluck, A. and Wool, I.G. (1991) J. Mol. Biol. 221, 193–207.
[32] Sutcliffe, M.J., Haneef, I., Carney, D. and Blundell, T.L. (1987) Protein Eng. 1, 377–384.