Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystal structure of β-luffin, a ribosome-inactivating protein, at 2.0 Å resolution

Qing-jun Ma, Jian-hui Li, Hui-guang Li, Shen Wu and Yi-cheng Dong*

Center for Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China

Correspondence e-mail: dyc@pedyc.ibp.ac.cn

Received 30 December 2002
Accepted 19 May 2003

PDB Reference: β-luffin, 1nio, r1niosf.

The crystal structure of β-luffin at 2.0 Å resolution was solved by the molecular-replacement method using polyalanyl trichosanthin as the search model. The structure was refined with *CNS*1.1, giving R_{work} = 0.162 and R_{free} = 0.204. The r.m.s.d.s of the bond lengths and bond angles are 0.008 Å and 1.3°, respectively. The overall structure is similar to those of other type I RIPs. Three *N*-acetylglucosamine (Nag) molecules are linked to residues Asn2, Asn78 and Asn85 of the protein.

1. Introduction

Ribosome-inactivating proteins (RIPs) are widely distributed in plant seeds and leaves. Many plants contain at least one type of ribosome-inactivating protein (Barbieri & Stirpe, 1982; Stirpe & Barbieri, 1986; Roberts & Claude, 1986). RIPs are RNA N-glycosidases that inhibit the protein synthesis of eukaryotic cells by cleaving a single adenine base from a highly specific site on the 28S RNA of the 60S ribosomal subunit (Endo & Tsurugi, 1987). RIPs have attracted attention as having potential application in the treatment of diseases such as cancer and AIDS. There are two types of RIPs (Stirpe & Barbieri, 1986). Type I RIPs are single-chained, whereas type II RIPs are double-chained. The *A* chain of type II RIPs possesses the ribosome-inactivating property; the *B* chain is responsible for attaching the protein molecule to the target-cell surface in order to assist the *A* chain in crossing the cell membrane. Type II RIPs are therefore among the most toxic cytotoxins. Trichosanthin and momorcharin belong to the type I RIPs, while ricin and abrin belong to type II. The crystal structures of trichosanthin (Gao *et al.*, 1993), α-momorcharin (Ren *et al.*, 1994), β-momorcharin (Yuan *et al.*, 1999) and ricin (Montfort *et al.*, 1987) have been elucidated. Trichosanthin and the ricin *A* chain are not only homologous in amino-acid sequence (Zhang & Wang, 1986), but are also similar in three-dimensional structure. Since trichosanthin and ricin are from taxonomically distant species, *Trichosanthes kirilowii* from the Cucurbitaceae family and *Ricinus communies* from the Euphorbiaceae family, respectively, it appears that the widely distributed RIPs of both types must originate from the same ancestor and assume the same 'RIP fold'.

β-Luffin is classified as a type I RIP. Despite the fact that β-luffin is a glycoprotein, whereas trichosanthin contains no carbohydrates, they share many common features. Like almost all type I RIPs (Stirpe & Barbieri, 1986), they have a comparable molecular mass (26–31 kDa) and a strongly basic pI (pH 9). They induce mid-term abortion in pregnant mice

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

Table 1
Data-collection and refinement statistics.

Values in parentheses correspond to the highest resolution shell.

Data collection	
Observed reflections	60539
Unique reflections	17034
Resolution range	30–2.0 (2.07–2.0)
Redundancy	3.54
Mosaicity	0.58
Completeness	99.6 (98.9)
Average $I/\sigma(I)$	12.8 (2.9)
$R_{\rm merge}$	11.1 (42.2)
Refinement and final model statistics	
Resolution range	29.21–2.0 (2.07–2.0)
No. of reflections in working set	14786 (1352)
No. of reflections in test set	1639 (157)
Data cutoff $\sigma(F)$	0
$R_{\rm work}$	0.162 (0.227)
$R_{\rm free}$	0.203 (0.271)
No. of non-H protein atoms	1911
No. of water molecules	128
No. of heteroatoms	42 (3 Nag)
R.m.s.d.s from ideal geometry	
Bond lengths (Å)	0.008
Bond angles (°)	1.3

and inhibit cell-free protein synthesis with similar potency (Yueng *et al.*, 1991).

Here, we report the crystal structure of β-luffin containing three Nag molecules at 2.0 Å resolution.

2. Experimental

2.1. Purification and crystallization

Proteins were purified as described in Wu *et al.* (1995). Crystals suitable for X-ray diffraction were grown using the hanging-drop vapour-diffusion method at room temperature. Crystallization drops were prepared by mixing 5 µl of reservoir solution [0.05 *M* Tris–HCl pH 7.5, 40%(*w*/*v*) $(NH_4)_2SO_4$] with 5 µl of 40 mg ml^{-1} protein solution [0.15 *M* NaCl, 0.1%(*w*/*v*) NaN_3].

2.2. Data collection and processing

X-ray diffraction data were collected to 2.0 Å resolution at room temperature on a MAR Research image plate (300 mm) with a Rugaku RU-200 rotating copper-anode generator operating at 40 kV and 100 mA. The data were processed using the *DENZO* and *SCALEPACK* programs (Otwinowski, 1993). The crystals belong to space group *C*2, with unit-cell parameters a = 89.902, b = 59.823, c = 55.184 Å, β = 120.81°. There is one molecule in the asymmetric unit, with an estimated $V_{\rm M}$ = 2.35 Å^3 Da^{-1} and a solvent content of 45.5% (Matthews, 1968). The statistics for the data set are summarized in Table 1.

2.3. Structure determination

The structure was solved by the molecular-replacement method using *CNS* (Brünger *et al.*, 1998) with polyalanyl trichosanthin as the initial model against 15–4 Å diffraction data. One obvious high peak was found by a cross-rotation function calculation with the direct rotation method with RF = 0.236 (the second high peak has RF = 0.107). The first ten peaks were used for subsequent PC refinement, translation-function calculation and rigid-body refinement. The best model was automatically output after rotating by α = 78.48, β = 84.76, γ = 27.047° and translating by x = −2.95, y = −0.42, z = 23.16 Å, with R = 0.517.

Table 2
Secondary structure of β-luffin.

(*a*) α-Helices.

Code	Starting/ending residues
α1	Asp11–Leu26
α2	Glu86–Gln94
α3	Asn110–Gly119
α4	Val121–His141
α5	Ala146–Phe164
α6	Phe164–Arg174
α7	Ser182–Gln203
α8	Ser230–Asn235
α9	Asn242–Ile246

(*b*) β-Sheets.

Code	β-Strand No. within the sheet	Starting/ending residues	Relationship between strands n and $n-1$
βa	1	Val3–Ser6	0
	2	Tyr48–Ser54	−1
	3	Ala60–Asp66	−1
	4	Ile72–Val77	−1
	5	Thr80–Phe83	−1
	6	Thr101–Thr104	1
βb	1	Ser28–Val32	0
	2	Ile35–Leu38	−1
βc	1	Ala208–Ile216	0
	2	Arg222–Asp227	−1

Table 3
Hydrogen bonds between *N*-acetylglucosamines and protein.

Symmetry-related molecules in the crystal are marked #.

Sugar atom	Protein atom	Distance (Å)
Nag248 O6	Asn235 OD1#	2.58
Nag249 O3	His141 ND1	3.05
Nag249 N2	Phe140 O	2.79
Nag250 O6	Ser43 OG#	3.38

2.4. Refinement

The refinement was performed with the program *CNS* using 29.9–2.0 Å diffraction data with a maximum-likelihood target on amplitudes. 10% of the data were randomly selected as a cross-validation data set to calculate $R_{\rm free}$ throughout the whole refinement. Torsion-angle molecular dynamics were applied in the earlier stages, while energy minimization and restrained individual *B*-factor refinement were included in the latter stages. Bulk-solvent correction and overall anisotropic *B*-factor correction were applied throughout the refinement.

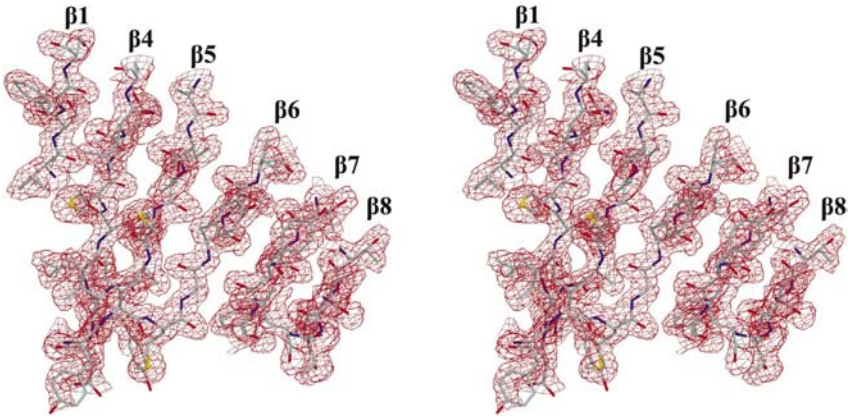

Figure 1
Stereo diagram of the $2F_o - F_c$ electron-density map around the six-stranded β-sheet contoured at 1.0σ. (All figures were prepared with *MOLSCRIPT* and *Raster3D*).

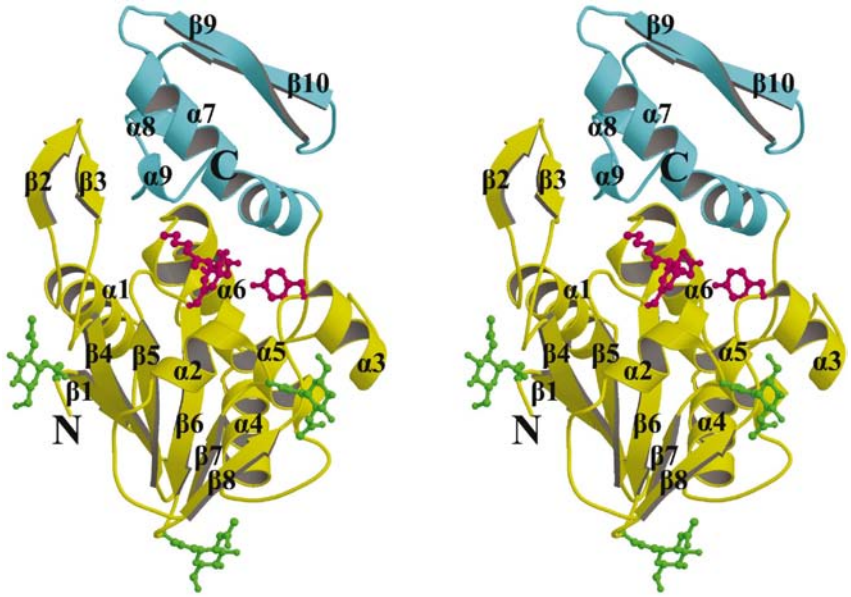

Figure 2
Stereo ribbon diagram of β-luffin. Stick models represent residues Tyr70, Tyr111, Glu160 and Arg163 in the active-site pocket and the three Nags at Asn2, Asn78 and Asn85.

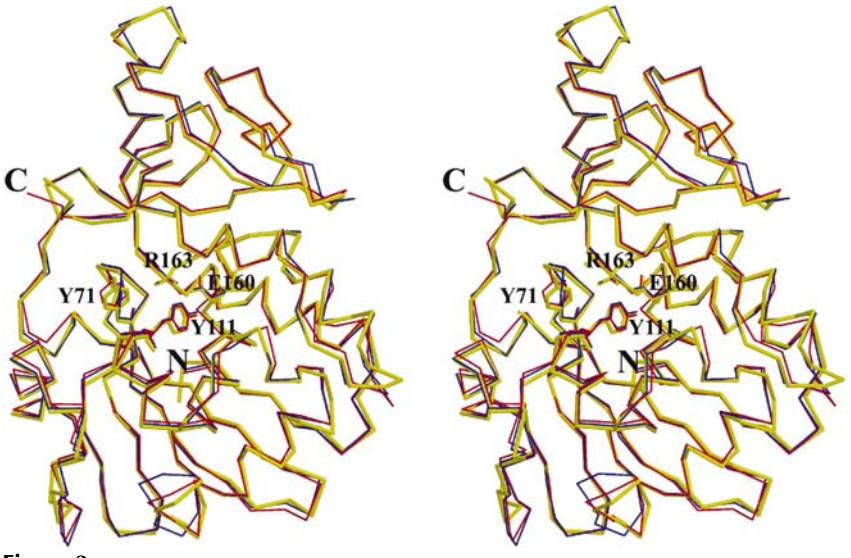

Figure 3
Stereo diagram of superposition of C^{α} atoms of β-luffin (yellow), trichosanthin (red) and α-momorcharin (blue).

Between each refinement round, the model was manually rebuilt using *XTALVIEW* (McRee, 1999), by inspecting σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ electron-density maps with reference to the β-luffin sequence. Except for two water molecules on special positions that were built manually, water molecules were placed automatically by *CNS*, meeting the criteria of both having electron density $\geq 3.5\sigma$ in a σ_A-weighted $F_o - F_c$ map and being within hydrogen-bonding distance (2.6–3.2 Å) of existing O or N atoms. The coordinates and restraints for the Nags were built automatically by *CNS*1.1. The final R factor and $R_{\rm free}$ of the model are 0.162 and 0.204, respectively. Refinement statistics are shown in Table 1.

3. Results and discussion

3.1. Quality of the model

Most residues fit the electron-density map quite well, except for some residues in loop regions and long side-chain residues on the protein surface that show smeared electron density. The electron-density map of the β-sheet is shown in Fig. 1. The geometry of the model was analyzed by *PROCHECK* (Laskowski *et al.*, 1993). The r.m.s.d.s of the bond length and bond angles are 0.008 Å and 1.3°, respectively. In the Ramachandran map (Ramachandran & Sasasekharan, 1968), 90% of the residues lie in the core region and 8.9% lie in the additionally allowed region. Residue Asn78, which is linked to a Nag, lies just outside the generally allowed region. Residue Asp236, which is located in a turn, lies in the generally allowed region. Data and refinement statistics are summarized in Table 1.

3.2. Overall structure and active site

β-Luffin has a large N-terminal (Ala1–Asn180) domain and a smaller C-terminal domain (Pro181–Ala247). The N-terminal domain is composed of six α-helices and two β-sheets, while the C-terminal domain consists of three α-helices and one β-sheet. Fig. 2 shows a ribbon representation of β-luffin. The secondary structure of β-luffin is very similar to those of other ribosome-inactivating proteins. The initial and terminal residues of the α-helices and β-strands are listed in Table 2. The tertiary structure of β-luffin shows a common 'RIP fold'. As superposition of the C^{α} atoms of β-luffin with trichosanthin and α-momorcharin is shown in

Fig. 3. The r.m.s. deviations of the C^{α} atoms between β-luffin and trichosanthin, and between β-luffin and α-momorcharin are 0.76 and 0.68 Å, respectively. However, the structure of one loop (Lys98–Thr101 in β-luffin) is obviously different from trichosanthin and α-momorcharin. The amino-acid sequence alignment indicates that there is a residue deletion in this loop of β-luffin. The C-terminal tail is also different owing to its flexibility.

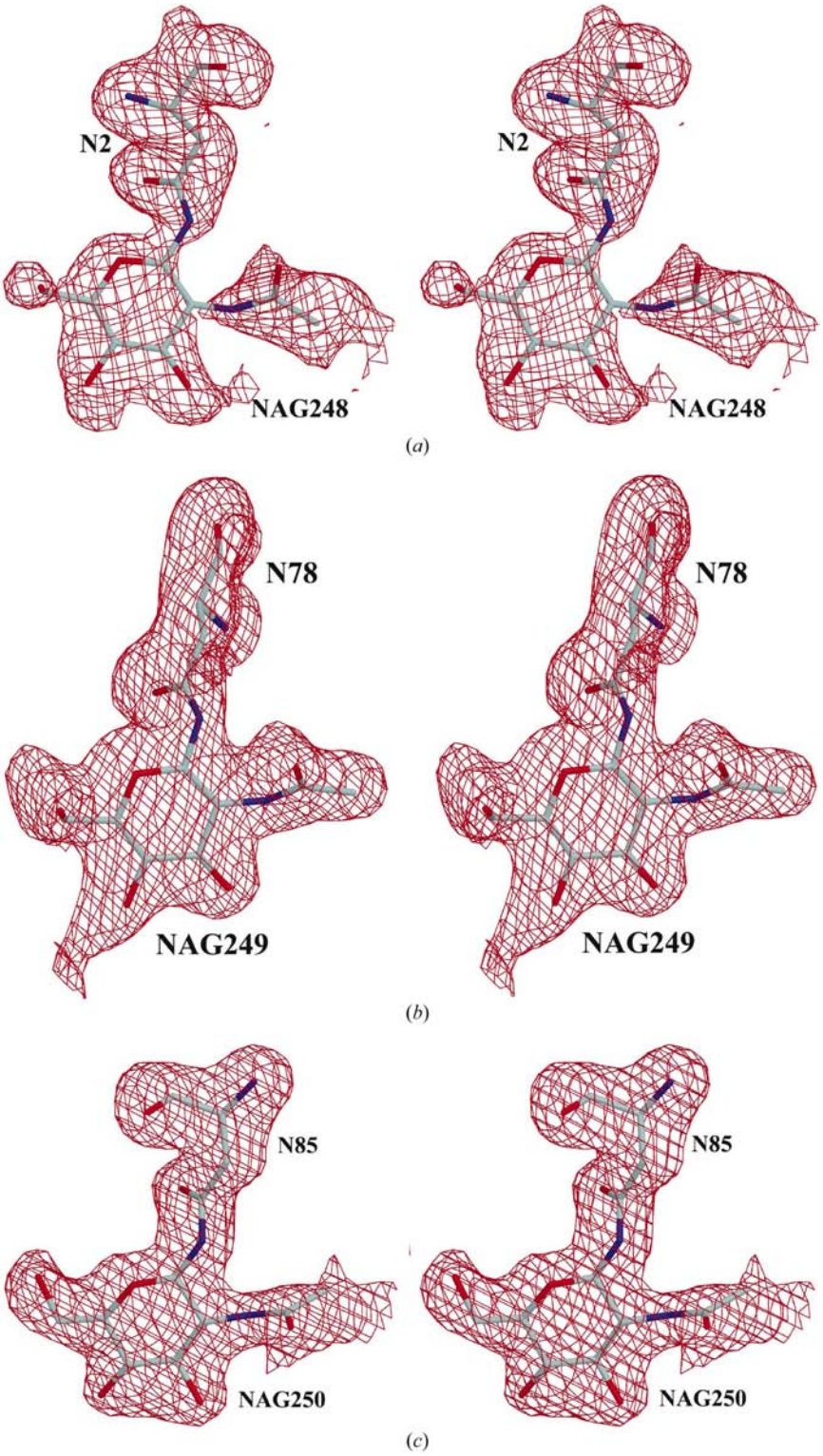

Figure 4
Stereo diagram of the simulated-annealing omit $2F_o - F_c$ electron-density map of three Nags contoured at 2σ. (*a*) Nag248 linked to Asn2; (*b*) Nag249 linked to Asn78; (*c*) Nag250 linked to Asn85.

In the cleft between the N-terminal domain and C-terminal domain lies the active pocket, which was proved to be the substrate-binding and catalysis site in ribosome-inactivating proteins. The positions of the key residues Arg163, Glu160, Tyr71 and Tyr111 are conserved except for the orientation of the aromatic ring of Tyr71. Tyr71 has a high *B* factor. In addition, the positions of some water molecules are well conserved.

3.3. Saccharides

Residues Asn2, Asn78 and Asn85 in the N-terminal domain were found to be linked to *N*-acetylglucosamines (Nags). This is consistent with the glycosylation sites reported in β-luffin (Islam *et al.*, 1991). These three Nags are named Nag248, Nag249 and Nag250, respectively. These Nags all protrude out from the protein molecule surface and extend into the solvent region. These saccharide moieties seem to be distant from the active-site pocket. Fig. 4 shows the electron-density map of the three Nags. Of these sugars, Nag249 forms more hydrogen bonds with atoms of the protein molecule. It therefore has a lower average *B* factor and shows a clearer electron density than that of Nag248 and Nag250. The hydrogen bonds formed by the three Nags are listed in Table 3. α-Momorcharin and β-momorcharin also have saccharide moieties in their three-dimensional structures (Ren *et al.*, 1994; Yuan *et al.*, 1999). However, these sites are different.

4. Recognition of one ambiguous residue

Residue 173 is reported as Glu in the β-luffin sequence. However, this residue is recognized as Lys in the σ_A-weighted $2F_o - F_c$ electron-density map, which was further confirmed by a simulated-annealing omit map (Fig. 5). Moreover, the NZ atom of Lys173 forms a hydrogen bond with Arg174 NH1.

This work was supported by the National Natural Science Foundation of China (No. 39900025). We give our heartfelt thanks to

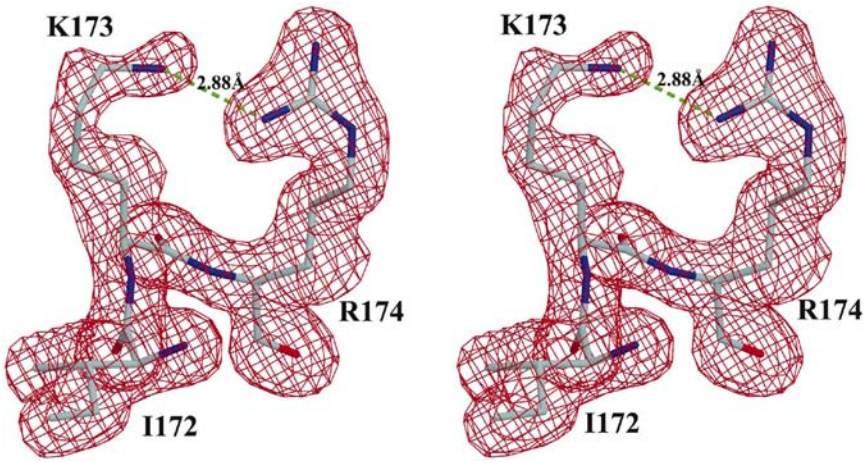

Figure 5
Stereo diagram of the simulated-annealing $2F_o - F_c$ omit map of residue 173 contoured at 3σ. Residue 173 is Lys.

Mr Hu Yugang, Dr Chai Jijie, Dr Zhang Hailong and Dr Guang Rongjin for their help.

References

Barbieri, L. & Stirpe, F. (1982). *Cancer Surv.* **10**, 489–520.
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. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Endo, Y. & Tsurugi, K. (1987). *J. Biol. Chem.* **262**, 8128–8130.
Gao, B., Ma, X.-Q., Wang, Y.-P., Wu, S., Chen, S.-Z. & Dong, Y. C. (1993). *Sci. China (Ser. B)*, **23**, 272–282.
Islam, M. R., Kung, S. S., Kimura, Y, & Funatsu, G. (1991). *Agric. Biol. Chem.* **55**, 1375–1381.
Laskowski, R. A., MacArthur, M. W., Moss, D. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Montfort, W., Villafranca, J. E., Monzingo, A. F., Ernst, S. R., Katzin, B., Rutenber, E., Xuong, N. H., Hamlin, R. & Robertus, J. D. (1987). *J. Biol. Chem.* **262**, 5398–5403.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). *Proteins Struct. Funct. Genet.* **12**, 345–364.
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.
Ramachandran, G. N. & Sasasekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–437.
Ren, J., Wang, Y., Dong, Y. & Stuart, D. (1994). *Structure*, **2**, 7–16.
Roberts, W. K. & Claude, P. S. E. (1986). *Biosci. Rep.* **6**, 19–29.
Stirpe, F. & Barbieri, L. (1986). *FEBS Lett.* **195**, 1–8.
Wu, S., Zhu, Y. R., Guo, F. & Liu, D. H. (1995). *Prog. Biochem. Biophys.* **22**, 464–468.
Yuan, Y.-R., He, Y.-N., Xiong, J.-P. & Xia, Z.-X. (1999). *Acta Cryst.* D**55**, 1144–1151.
Yueng, H. W., Li, W. W. & Ng, T. B (1991). *Int. J. Pept. Protein Res.* **38**, 15–19.
Zhang, X.-J. & Wang, J.-H. (1986). *Nature (London)*, **321**, 477–478.