Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Three-dimensional structure of β-momorcharin at 2.55 Å resolution

Yu-Ren Yuan,† Yong-Ning He,‡ Jian-Ping Xiong§ and Zong-Xiang Xia*

Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, People's Republic of China

† Present address: Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, Manhattan, NY 10032, USA.
‡ Present address: Department of Biological Science, Purdue University, West Lafayette, IN 47907, USA.
§ Present address: Massachussetts General Hospital/Harvard Medical School, USA.

Correspondence e-mail: xiazx@pub.sioc.ac.cn

Received 6 June 1998
Accepted 24 February 1999

PDB Reference:
β-momorcharin, 1cf5.

β-Momorcharin ($M_r \simeq 29$ kDa) is a single-chained ribosome-inactivating protein (RIP) with a branched hexasaccharide bound to Asn51. The crystal structure of β-momorcharin has been determined using the molecular-replacement method and refined to 2.55 Å resolution. The final structural model gave an R factor of 17.2% and root-mean-square deviations of 0.016 Å and 1.76° from ideal bond lengths and bond angles, respectively. β-Momorcharin contains nine α-helices, two 3_{10} helices and three β-sheets, and its overall structure is similar to those of other single-chained RIPs. Residues Tyr70, Tyr109, Glu158 and Arg161 are expected to define the active site of β-momorcharin as an rRNA N-glycosidase. The oligosaccharide is linked to the protein through an N-glycosidic bond, β-GlcNAc–(1-*N*)-Asn51, and stretches from the surface of the N-terminal domain far from the active site, which suggests that it should not play a role in enzymatic function. The oligosaccharide of each β-momorcharin molecule interacts with the protein through hydrogen bonds, although in the crystals most of these are intermolecular interactions with the protein atoms in an adjacent unit cell. This is the first example of an RIP structure which provides information about the three-dimensional structure and binding site of the oligosaccharide in the active chains of RIPs.

1. Introduction

β-Momorcharin (β-MMC) is a single-chained ribosome-inactivating protein (RIP; Barbieri *et al.*, 1993) isolated from the seeds of the bitter gourd *Momordica charantia* Linn of the Cucurbitaceae family (Yeng *et al.*, 1986). The bitter gourd has been used as a traditional Chinese medicine. β-MMC has been identified as possessing an abortifacient function (Yeng *et al.*, 1988) and it has been reported that the protein extracted from the seeds and fruit of *M. charantia* is a new inhibitor of human immunodeficiency virus infection and replication (Lee-Huang *et al.*, 1990).

RIPs from plants are RNA N-glycosidases which depurinate the major rRNA, thus damaging ribosomes and arresting protein synthesis (Barbieri *et al.*, 1993). Both trichosanthin (TCS) and ricin, the well known single-chained and double-chained RIPs, respectively, inactivate eukaryotic ribosomes by removing a single adenine base from a specific site (A^{4324}) of 28S rRNA in the 60S ribosomal subunit (Zhang & Liu, 1992; Endo & Tsurugi, 1987).

Several three-dimensional structures of single-chained RIPs, such as TCS, α-momorcharin (α-MMC) and pokeweed antiviral protein, have been reported (Xia *et al.*, 1993; Gao *et al.*, 1994; Ren *et al.*, 1994; Husain *et al.*, 1994; Monzingo *et al.*, 1993; Ago *et al.*, 1994) and shown to share a common fold. The

Table 1
Molecular-replacement solution of β-MMC.

The initial solution of the Euler angles was obtained from the two highest peaks of the Crowther fast-rotation function (Crowther, 1972), which was improved by applying the Lattman & Love rotation function (Lattman & Love, 1972) in the local area. The initial solution of the translation vector was obtained from the highest peak of the translation function (Crowther & Blow, 1967). Rigid-body refinement then leads to the final solution listed.

Molecule	Euler angle (°)			Translation vector		
	α	β	γ	F_A	F_B	F_C
A	299.20	71.11	298.71	0.0000	0.0000	0.0000
B	77.02	98.48	96.53	0.4993	0.5270	0.4377

A chains of double-chained RIPs are responsible for the enzymatic activity and their structures, such as that of the ricin *A* chain (Katzin *et al.*, 1991), are similar to single-chained RIPs. However, RIPs usually differ from each other in the structure of some of their surface loops; this leads to different toxicities and antigenicities, which are important for the application of these proteins as drugs.

The molecular weight of β-MMC is 29 kDa and the 249 amino-acid residue sequence has been determined by both chemical and gene-sequencing methods (Jin, 1993), showing identities of 60% and 54% to TCS and α-MMC, respectively (Collins *et al.*, 1990; Ho *et al.*, 1991). Trichosanthin does not contain carbohydrate, whereas β-MMC is a glycoprotein containing 3.5% sugar, and the primary structure of the oligosaccharide and its conformation in solution have been determined; the oligosaccharide is a branched hexasaccharide bound to the protein through a GlcNAc–(1-*N*)-Asn51 linkage (Wu & Jin, 1993). Many single-chain RIPs and the *A* chains of double-chain RIPs are glycoproteins. However, none of the reported three-dimensional structures contains information on carbohydrate binding.

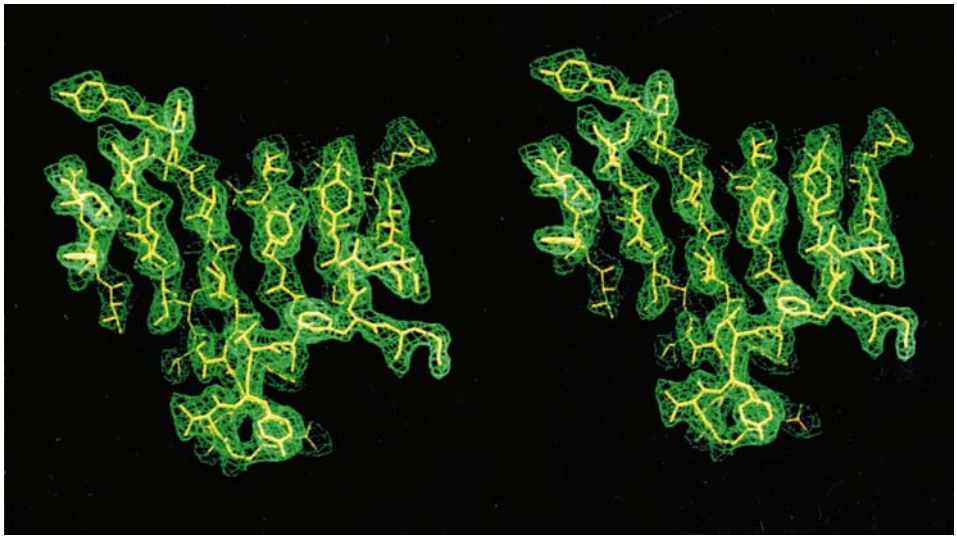


Figure 1
Stereoscopic view of the $(2F_o - F_c)$ electron density of the six-stranded β-sheet (molecule *A*) contoured at 1.0σ. This diagram was prepared using the molecular-graphics program *TURBO-FRODO*.

We have reported the crystallization and preliminary crystallographic study of β-MMC (Xiong, Xia, Zhang *et al.*, 1994). The crystals belong to space group *P*1 with unit-cell parameters a = 49.09, b = 50.58, c = 61.12 Å, α = 72.98, β = 78.39, γ = 76.97°. There are two molecules in the unit cell. The crystal structure of β-MMC was determined using the molecular-replacement method, and the statistics from the preliminary refinement have been reported without describing the structure (Yuan *et al.*, 1995). In this paper, we present the three-dimensional structure of β-MMC refined at 2.55 Å resolution. The structure displays some features which differ from other single-chained RIP structures and our refined model includes coordinates for the carbohydrate moiety.

2. Experimental

2.1. X-ray data collection and processing

Single crystals of β-MMC with sufficient dimensions for data collection were very difficult to grow and only one crystal suitable for X-ray data collection was obtained. Two sets of X-ray diffraction data were collected to 2.41 Å resolution from the crystal in two orientations, differing in χ by 30°, on an X-200B area detector using a rotating-anode X-ray generator as the radiation source. The image data were processed and reduced using *XENGEN* (Howard *et al.*, 1987). The two data sets were merged giving an $R_{\rm merge}$ ($\sum |I - \langle I \rangle| / \sum I$) of 7.8% and an average redundancy (number of observations/number of reflections) of 2.3. The data are 86.8 and 74.1% complete to 2.55 and 2.41 Å resolution, respectively.

2.2. Molecular-replacement solution

The crystal structure of β-MMC was determined using the molecular-replacement method (Rossmann, 1972) by applying the program package *MERLOT* (Fitzgerald, 1988) on an MV3300 computer. The X-ray data with $I > 2\sigma(I)$ in the resolution range 8.0–4.0 Å were used for the molecular-replacement calculation. Molecule *B* of TCS in the monoclinic crystal structure at 2.7 Å resolution (Xia *et al.*, 1993; Zhang *et al.*, 1994) was used as the search model; no insertions or deletions were made in the model, but side-chain atoms were omitted for residues differing from the amino-acid sequence of β-MMC. The procedure and the structure solution by molecular replacement are presented in Table 1.

2.3. Initial model building

The initial structural model of β-MMC was obtained by applying the molecular-replacement solution to the search model, refining the scale

Table 2
Refinement statistics.

(*a*) *R* factor

R factor (10.0–2.55 Å)	R_{free} (10.0–2.55 Å)	R_{shell} (2.66–2.55 Å)
0.172	0.278	0.267

(*b*) R.m.s. deviations of various geometric parameters from ideal values

Bond length (Å)	Bond angle (°)	Dihedral angle (°)	Improper angle (°)
0.016	1.756	27.408	2.661

(*c*) Average temperature factors (Å^2)

	Molecule *A*		Molecule *B*	
	Main chain	Side chain	Main chain	Side chain
Residues 1–244	19.43	21.30	18.88	20.34
Residues 245–249	78.99	84.40	81.09	90.51
Sugar	91.02		85.03	

(*d*) Transformation matrix and vector from molecule *A* to *B*

Transformation matrix	0.9999	0.0157	−0.0013
	0.0157	−0.9996	0.0247
	−0.0010	−0.0247	−0.9997
Vector	35.9657	32.8158	25.2830

factor using the Hendrickson–Konnert stereochemically restrained least-squares refinement method (Hendrickson, 1985) by applying the *PROLSQ* procedure (Finzel, 1987) at 2.7 Å resolution, computing the $(2F_o - F_c)$ electron-density map, roughly rebuilding the segments with appropriate insertions or deletions and fitting the omitted side chains to the map. The model fitting was carried out using the graphics software *TOM-FRODO* (Jones, 1985; Cambillau & Horjales, 1987) on an IRIS graphics workstation. A few C-terminal residues were omitted at this stage. The *R* factor calculated from this initial model was 45% at 2.7 Å resolution.

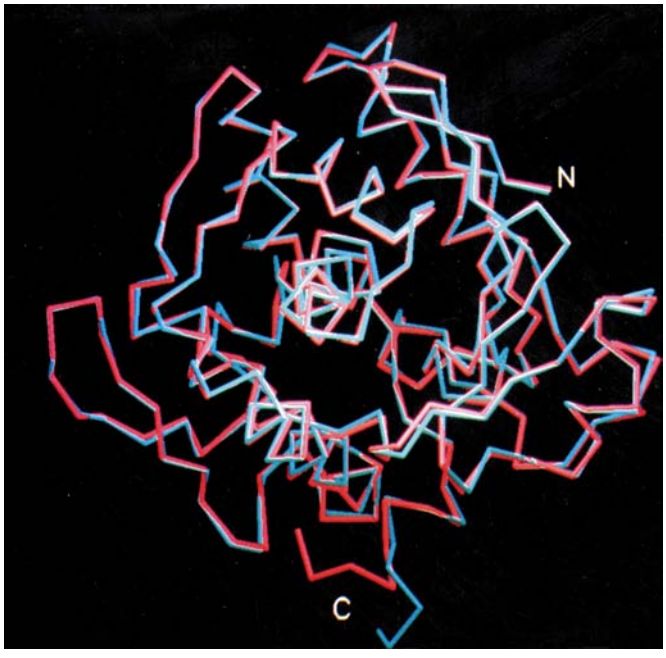


Figure 2
Superimposed polypeptide chains of molecules *A* and *B* (C^{α} drawing), shown in red and blue, respectively. N- and C-termini are indicated. This diagram was prepared using the molecular-graphics program *SETOR* (Evans, 1993).

Table 3
Secondary structure of β-MMC.

The secondary-structure assignment was based on the secondary-structure search using the molecular-graphics software *SETOR*, combined with manually examining the hydrogen-bonding patterns of the model.

(*a*) α-Helices†.

Code	Starting/ending residues	Irregularity (broken hydrogen bonds listed)
α1	Thr10–Thr24	
α2	Pro88–Leu94	
α3	Asn108–His117	
α4	Gly126–Tyr139	Ser131 O–Thr135 N‡
α5	Ser144–Phe162	Gln154 O–Glu158 N, Thr155 O–Ala159 N
α6	Phe162–Lys171	
α7	Leu181–Ala200	Glu187 O–Ser191 N, Asn188 O–Ala192 N
α8	Ser228–Asn234	
α9	Asn240–Thr245	

(*b*) β-Sheets§¶.

Code	β-Strand number within the sheet	Starting/ending residues	Relationship between strands *n* and *n* − 1††
βa	1	Val2–Leu6	0
	2	Arg46–Ser54	+1
	3	Tyr57–Val66	−1
	4	Asn68–Thr76	−1
	5	Val79–Lys84	−1
	6	Arg99–Leu103	+1
βb	1	Ser28–Val31	0
	2	Ile34–Leu36	−1
βc	1	Val211–Lys215	0
	2	Gly218–Val223	−1

† There are two additional 3_{10} helices: Ser42–Arg46 and Ile119–Ile123. ‡ The side chain of Ser131 is hydrogen bonded to the main-chain N atom of Tyr55, which makes the carbonyl group of Ser131 orient irregularly to form a hydrogen bond to the side chain of Thr135 instead of to its main-chain N atom. § The irregularity of the hydrogen-bonding patterns of the β-sheets are basically the same as that in TCS structure. ¶ The following pairs of adjacent segments forming only one pair of hydrogen bonds are not referred to as β-strands: Pro35 O···Leu238 N and Leu37 N···Leu238 O, Leu125 O···Phe177 N and Leu125 N···Phe177 O, Arg161 O···Lys236 N and Lys163 N···Asn234 O. †† 1, strand *n* parallel to strand *n* − 1; −1, strand *n* antiparallel to strand *n* − 1.

2.4. Crystallographic refinement

The crystallographic refinement was carried out in two stages. Firstly, the initial structural model of β-MMC was refined using the program package *TNT* (Tronrud *et al.*, 1987) on an MV3300 computer; no carbohydrates were included in the model throughout this stage. A total of 12 rounds of *TNT* refinement were carried out using the data in the resolution ranges 10–3.0 Å for the first five rounds and 10–2.41 Å for the remaining rounds. Each round consisted of three cycles of positional refinement employing the gradient-over-curvature method with loose stereochemical restraints, followed by one

cycle of temperature-factor refinement, as well as some additional cycles of positional refinement. The $(2F_o - F_c)$ and $(F_o - F_c)$ electron-density maps were computed after each round of the refinement, and the manual rebuilding of the model was carried out using *TOM-FRODO* software on an INDIGO graphics workstation. Solvent fitting commenced after the ninth round. The *R* factor dropped to 20.4% at the end of this stage.

In the second stage of refinement, the structure was further refined, for both positions and temperature factors, using the program package *X-PLOR* (Brünger, 1992*a*) on an INDIGO 2 graphics workstation. The data with $|F_o| > 2\sigma(|F_o|)$ in the resolution range 10–2.41 Å were used to refine the structure. During the refinement, the structure was further improved by manually adjusting the model and fitting more water molecules using *TURBO-FRODO* (Roussel & Cambillau, 1991). Simulated-annealing 'omit' maps were computed when necessary by omitting some particular residues from the model. At this point, the $(2F_o - F_c)$ map showed some additional areas of electron density near the molecular surface region, of which one was connected to the side chain of Asn51 and was identified to be a carbohydrate. The preliminary model of the oligosaccharide which was used for the initial carbohydrate fitting was produced using the software *QUANTA* on the IRIS graphics workstation by building the individual sugar units one by one and linking them according to the primary structure of the oligosaccharide. The fucose was in the L-configuration and the other sugars in the D-configuration, as they usually occur in nature. The fitting of the oligosaccharide was first carried out (also using *TURBO-FRODO* software) in molecule *B*, since the electron density of the oligosaccharide is stronger in molecule *B* than in molecule *A*. The sugar moiety of molecule *B* was then included in the model for further refinement. The refined oligosaccharide was then transformed into molecule *A* and manually shifted to fit the electron density of molecule *A*. Further runs of *X-PLOR* refinement and a number of rounds of model rebuilding of the oligosaccharide were carried out. Late in the *X-PLOR* refinement stage, 5% of the X-ray data were randomly selected as the test data set used for cross validation, and data in the 2.55–2.41 Å shell were omitted because the X-ray data in this shell have a completeness of only 30.2%. In order to evaluate objectively the free *R* factor and to remove the possible model bias which might be introduced during the model fitting, a simulated-annealing procedure from 3000 to 300 K was applied, during which the 5% test data set was set aside. Examination of electron-density maps led to minor rebuilding of four of the 12 sugars and two residues of one of the C-terminal tails, and this was followed by two rounds of the position and temperature-factor refinement. An additional positional-refinement run led to the final structure model of *β*-MMC.

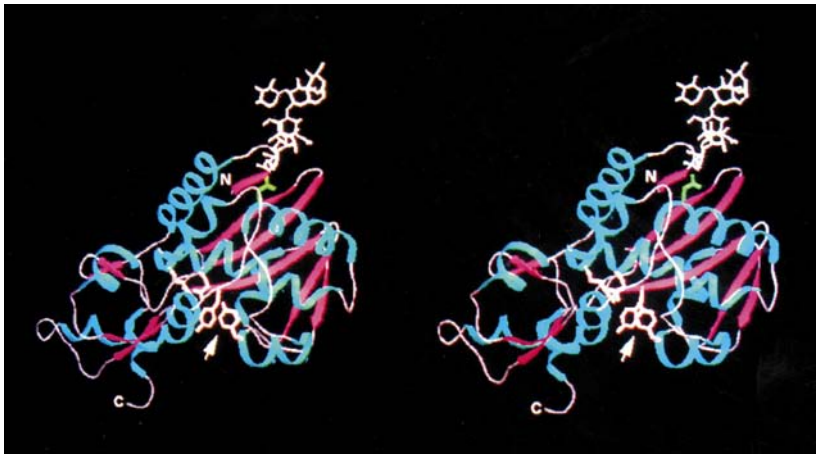


Figure 3
Stereoscopic ribbon diagram showing secondary structure of *β*-MMC (molecule *B*). *α*-Helices are shown in blue and *β*-strands are shown as red arrows. N- and C-termini are indicated. The oligosaccharide is shown in white and the side chain of Asn51 in green. A white arrow points to the active-site cleft; the side chains of the four active-site residues, Tyr70, Tyr109, Glu158 and Arg161, are shown in pink. This diagram was prepared using the molecular-graphics program *SETOR*.

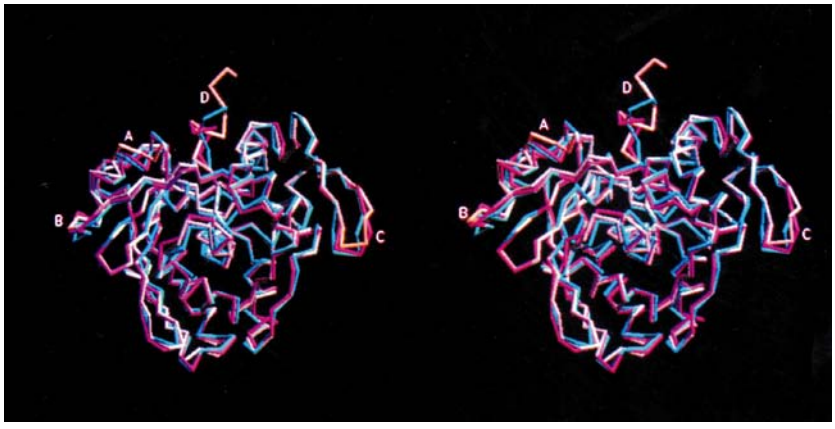


Figure 4
Stereoscopic view of the C^{α} backbone of *β*-MMC (molecule *B*, in grey) superimposed with monoclinic TCS (molecule *B*, in blue) and *α*-MMC (in purple). The three surface loops, Glu85–Glu89, Lys96–Arg99 and Lys215–Gln219, and the C-terminal tail of *β*-MMC are shown in orange and indicated as *A*, *B*, *C* and *D*, respectively. This diagram was prepared using the molecular-graphics program *SETOR*.

3. Results and discussion

3.1. Quality of the structure

The *R* factor and $R_{\rm free}$ (Brünger, 1992*b*) of the final model are 17.2 and 27.8%, respectively. The r.m.s. deviations of the bond lengths and bond angles from the ideal values are 0.016 Å and 1.76°, respectively. Tables 2(*a*) and 2(*b*) show the refinement statistics.

The Ramachandran plots (Ramachandran & Sasasekharan, 1968) of the polypeptide chains

of the two molecules show that the conformational angles φ_s and ψ_s are reasonable, with about 99% of the non-glycine residues within the acceptable regions as defined by *PROCHECK* (Morris *et al.*, 1992). Fig. 1 shows the representative electron density of the final $(2F_o - F_c)$ map, which is the six-stranded β-sheet region. As can be seen in Table 2(*c*), the average temperature factors of the C-terminal segments and sugar moieties are significantly higher than those of other regions of the structure, suggesting that they are more flexible than the remainder of the protein.

3.2. Overall structure

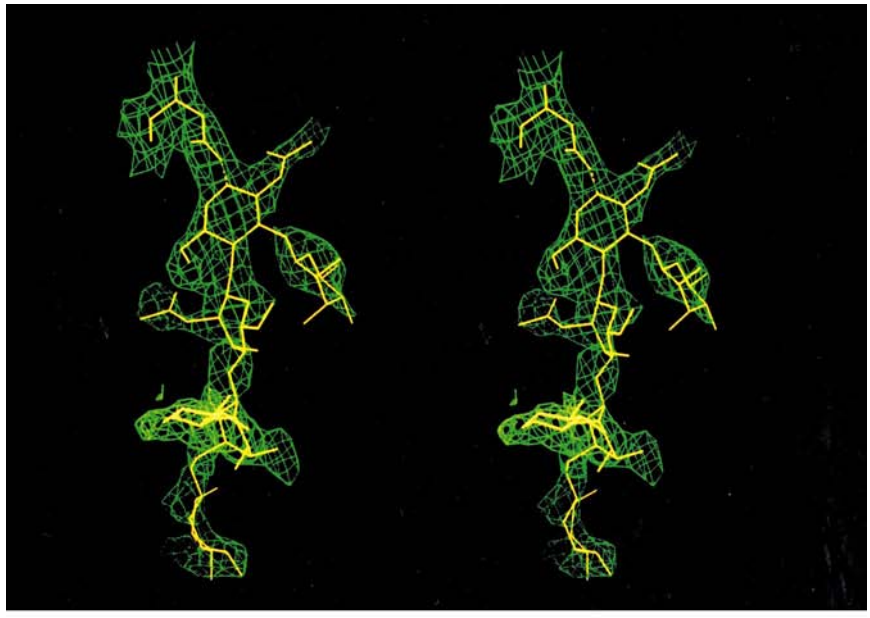
The two molecules in the unit cell are very similar to each other. After refinement without non-crystallographic symmetry constraints, the r.m.s. deviation of the 245 N-terminal C^{α} atoms between the two molecules is 0.36 Å; the transformation matrix and vector between them are shown in Table 2(*d*). However, the conformations of the flexible C-terminal tails in the two molecules are completely different. Fig. 2 shows the superposition of the C^{α} backbone of the two molecules.

The β-MMC molecule contains 249 amino-acid residues and a hexasaccharide. 68 water molecules are included in the refined structure. The polypeptide chain folds into two domains, a large N-terminal domain (Asp1–Asn180) and a smaller C-terminal domain (Leu181–Asn249). The oligosaccharide stretches from a surface region of the N-terminal domain. Fig. 3 shows the structure of the molecule *B*.

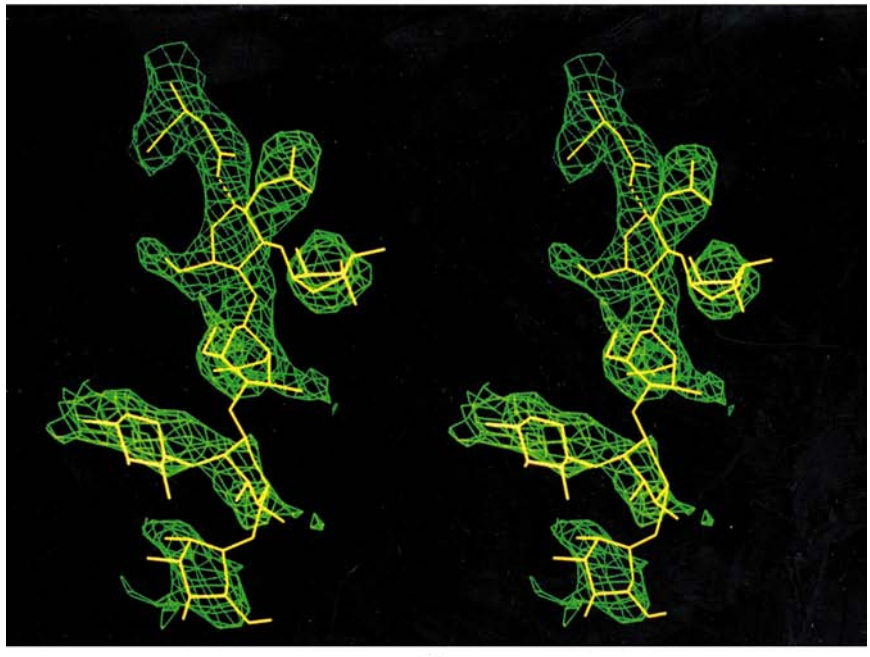
3.3. Secondary structure

The β-MMC molecule contains nine α-helices, two 3_{10} helices and three β-sheets with a total of ten β-strands. The secondary structure of β-MMC is shown in Fig. 3, and Table 3 shows the detailed secondary-structural elements of β-MMC.

The secondary structure of β-MMC is very similar to those of other single-chained RIPs. However, α2 (Pro88–Leu94) is obviously shorter in β-MMC, four residues shorter than those in TCS (Xia *et al.*, 1993) and in α-MMC (coordinates from the Protein Data Bank, entry 1ahc). The two consecutive prolines Pro87 and Pro88 in β-MMC hinder the formation of hydrogen bonds at its N-terminal end and a deletion exists at its C-terminal end, so that only three hydrogen bonds remain. The secondary structure of β-MMC also shows irregular hydrogen-bonding patterns similar

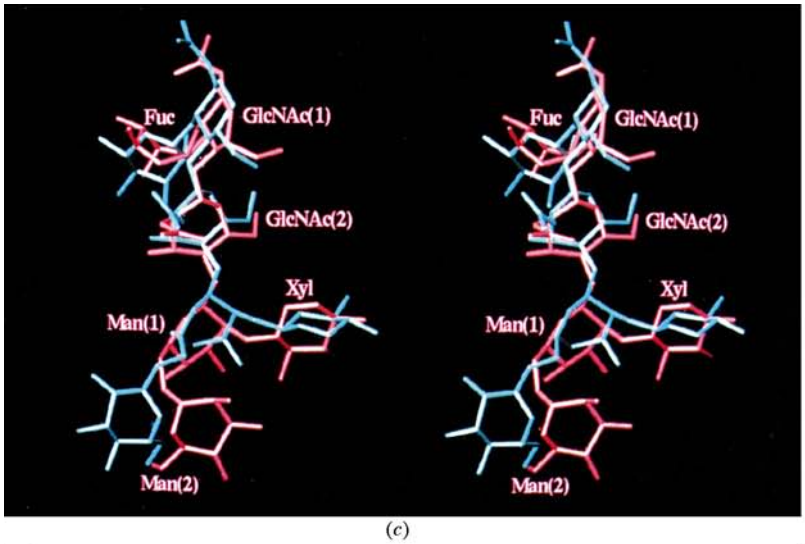


(*a*)

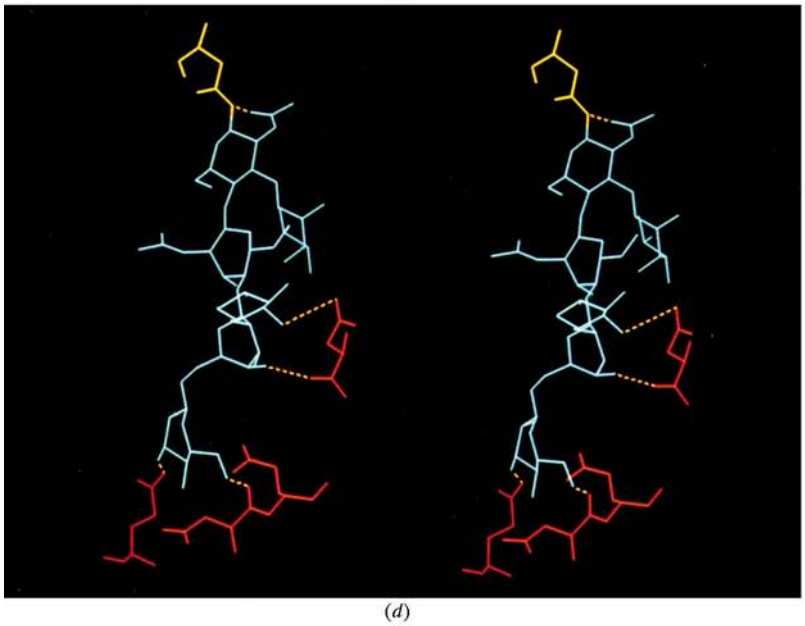


(*b*)

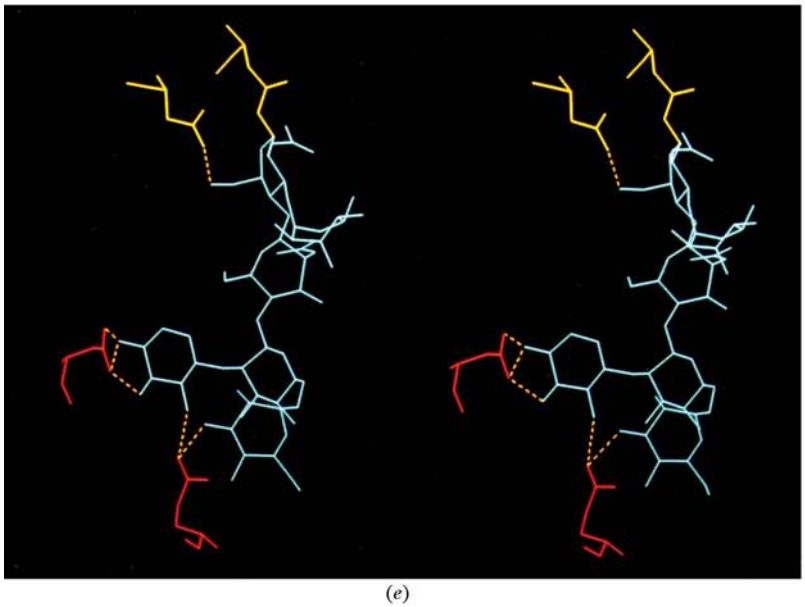
Figure 5
Stereoscopic diagrams of the oligosaccharide structure in β-MMC, prepared using the molecular-graphics program *TURBO-FRODO*. (*a*) The $(2F_o - F_c)$ electron density of the oligosaccharide of molecule *A*, contoured at 0.8σ. Asn51 is also shown. (*b*) The $(2F_o - F_c)$ electron density of the oligosaccharide of molecule *B*, contoured at 0.8σ. Asn51 is also shown. (*c*) The superimposed oligosaccharides of molecule *A* (in blue) and molecule *B* (in red). (*d*) Interactions of the oligosaccharide (in blue) of molecule *A* with the protein. Asn51 of molecule *A* is shown in yellow; Glu110, Asn188–Glu189 and the C-terminus (Asn249) of molecule #*A* are shown in red. The N-glycosidic bond is shown as a solid line and the hydrogen bonds are shown as dashed lines. (*e*) Interactions of the oligosaccharide (in blue) of molecule *B* with the protein. Asn51 and Asn3 of molecule *B* are shown in yellow; Gln189 and Asn209 of molecule #*B* are shown in red. The N-glycosidic bond is shown as a solid line and the hydrogen bonds are shown as dashed lines.



(*c*)



(*d*)



(*e*)

Figure 5 (continued)

to that of TCS (Xia *et al.*, 1993; Xiong *et al.*, 1997), with an additional irregularity in α4, as indicated in Table 3.

3.4. Tertiary structure

The tertiary structure of β-MMC shows that this protein shares a common 'RIP' fold with other single-chained RIPs. Fig. 4 shows the superposition of the C^{α} backbone of β-MMC with those of α-MMC and monoclinic TCS. They are very similar to each other up to Leu239 of β-MMC, but the flexible C-terminal tails in the three proteins are very different. The r.m.s. deviations of C^{α} atoms between β-MMC and monoclinic TCS and between β-MMC and α-MMC are 0.86 and 0.75 Å, respectively (calculated from the 239 N-terminal C^{α} atoms of molecule *B* of β-MMC and the corresponding C^{α} atoms of TCS and α-MMC).

However, the structure of three surface loops Glu85–Glu89, Lys96–Arg99 and Lys215–Gln219 in β-MMC obviously differ from the corresponding segments in TCS and α-MMC, as shown in Fig. 4. The amino-acid sequence alignment of the relevant segments indicates that the insertion, deletion and the replacement of non-proline residues by prolines lead to the different folding patterns of the polypeptide chain for these segments. Some side chains of these segments differ greatly between the protein structures; for example, the side chain of Glu85, a residue thought to be in the active site (Huang *et al.*, 1995), points towards the active-site cleft in TCS structure, but away from the active center towards the solvent area in the β-MMC structure.

3.5. Active-site structure

There is an active-site cleft located near the interface of the two domains, as indicated in Fig. 3; such clefts were also found in other RIPs and have proven to be the substrate-binding pocket (Monzingo & Robertus, 1992; Xiong, Xia & Wang, 1994; Ren *et al.*, 1994).

We have previously reported the active-site structure of TCS derived from the crystal structure of the TCS–NADPH complex at 1.7 Å resolution, and have demonstrated that the side chains of the four conserved residues, Tyr70, Tyr111, Glu160 and Arg163, interact with the substrate analogue NADPH (Xiong, Xia & Wang, 1994). Comparing the structure of β-MMC with the TCS–NADPH complex structure indicates that Tyr70, Tyr109, Glu158 and Arg161 are the active-site residues of β-MMC. The positions and conformations of the four active-site residues are similar in the two protein structures. Fig. 3 shows

Table 4
Sugar–protein interactions in β-MMC crystal structure.

# denotes the molecule in the adjacent unit cell $x + 1$, y, z.

Molecule *A*			Molecule *B*		
Atom 1 (in sugar)	Atom 2 (in protein)	Distance (Å)	Atom 1 (in sugar)	Atom 2 (in protein)	Distance (Å)
GlcNAc(1) O7	*A* Asp51 ND2	3.26	GlcNAc (1) O6	*B* Asn3 ND2	3.10
Man(1) O4	#*A* Asn249 OT1	2.88			
Man(2) O3	#*A* Glu110 OE1	2.80	Man(2) O3	#*B* Gln189 OE1	3.28
Man(2) O6	#*A* Asn188 O	2.80			
Xyl O2	#*A* Asn249 OD1	3.29	Xyl O2	#*B* Gln189 OE1	3.04
			Xyl O3	#*B* Asn209 ND2	2.50
			Xyl O4	#*B* Asn209 OD1	2.92
			Xyl O4	#*B* Asn209 ND2	3.38

the side chains of the four active-site residues in the β-MMC molecule, all of which point towards the active-site cleft.

3.6. Identification of two ambiguous residues

The amino-acid sequence determined by chemical methods agrees well with that derived from the gene sequence (the five C-terminal residues were not determined by gene-sequencing methods), with the exception of three residues 131, 157 and 219, which are Lys, Ser and Gln, respectively, from chemical methods and Ser, Ala and Glu, respectively, from gene sequencing. The electron density suggests that two of the three ambiguous residues should be Ser131 and Ala157; however, for residue 219 it is impossible to distinguish between Glu and Gln from the electron density. In addition to the electron density, the examination of the environment of residue 131 also supports Ser131 rather than Lys131: the side chain of Ser131 forms a hydrogen bond to the amide N atom of Tyr55, and the long side chain of lysine would make unreasonably close contacts with the surrounding residues.

3.7. The oligosaccharide structure

The β-MMC molecule contains a branched hexasaccharide which is composed of two β-*N*-acetylglucosamines [GlcNAc(1) and GlcNAc(2)], one α-fucose (Fuc), one β-mannose [Man(1)], one α-mannose [Man(2)] and one β-xylose (Xyl).

The oligosaccharide is bound to Asn51, stretches from the second strand of the six-stranded β-sheet and extends into the solvent area, as shown in Fig. 3. Most of the sugars show distorted chair conformations.

The average temperature factors of the oligosaccharides are much higher than those of the protein atoms in both molecules, as described above, and the electron density of the oligosaccharides in β-MMC is generally weaker than that of the protein parts, which implies that the conformation of the oligosaccharide is somewhat flexible.

The electron density of most of the sugar rings is above 1.0σ in the $(2F_o - F_c)$ electron-density map, where σ is the r.m.s. electron-density value of the map, but the rings of Man(1), Man(2), Xyl and GlcNAc(2) in molecule *A* and the Man(1) ring in molecule *B*, as well as most of the linkages between the rings, show poor electron density. Most of the acetylamide groups and hydroxymethyl groups as well as some of the hydroxyl groups are defined quite well, showing $(2F_o - F_c)$ electron densities above 1.0σ. Strong electron density is observed in both molecules between the atoms C1 of GlcNAc(1) and the amide N atoms of the side chain of Asn51 in the two molecules. The electron density, contoured at 0.8σ, of the oligosaccharides in molecules *A* and *B* are shown in Figs. 5(*a*) and 5(*b*), respectively.

Fig. 5(*c*) shows the superimposed oligosaccharide structures of the two molecules. The three carbohydrates constituting the stem of the branched oligosaccharide, GlcNAc(1), GlcNAc(2) and Man(1), of the two molecules are very similar to each other. However, the conformations of the other three carbohydrates, Fuc, Man(2) and Xyl, differ significantly between the two molecules. The differences can be accounted for by considering the intermolecular interactions.

The oligosaccharides in both molecules interact with the protein through hydrogen bonds as shown in Table 4, Figs. 5(*d*) and 5(*e*). Most of them are intermolecular interactions of the sugar atoms of molecule *A* or *B* with the protein atoms in an adjacent unit cell, molecules #*A* and #*B* at $(X + 1, Y, Z)$. These sugar–protein interactions differ between molecules *A* and *B*. For example, Man(1) and Xyl of molecule *A* are hydrogen bonded to the C-terminus of an adjacent molecule #*A*, but the oligosaccharide in molecule *B* is not in contact with the C-terminal tail of the corresponding molecule #*B*, in part because the conformation of the C-terminal tail of #*B* is different from that of molecule #*A*, as described above. These different sugar–protein interactions probably lead to the differences in the conformations of three of the six sugars in molecule *A* from those in molecule *B*. The different conformations of the sugars in turn result in some additional differences in sugar–protein interactions: for example, in molecule *A*, both atoms O3 and O6 of Man(2) form hydrogen bonds to the side chain of Glu110 and the main chain of Asn188 of molecule #*A*, while in molecule *B*, both Man(2) O3 and Xyl O2 hydrogen bond to the side chain of Gln189 of molecule #*B*. Furthermore, O3 and O4 of Xyl of molecule *B* interact with the side chain of Asn209 in molecule #*B*, and this accounts for the very different conformations of the side chains of Asn209 in the two molecules. In addition, there are some local differences in the oligosaccharide structure: for example, the hydroxymethyl and acetylamide groups of GlcNAc (1) of the two molecules orient differently, with different intramolecular sugar–protein interactions: O7 of GlcNAc(1) interacts with Asn51 in molecule *A* while in molecule *B*, O6 of GlcNAc (1) interacts with Asn3.

Several plant glycoproteins have been reported to contain xylose and share a common core structure (Kimura *et al.*, 1987). β-MMC consists of this core structure as well as an additional α-mannose, equivalent to one of the oligosaccharides in α-MMC and in stem bromelain isolated from pineapple stem (Kimura *et al.*, 1991; Bouwstra *et al.*, 1990). The oligosaccharide of β-MMC is far from the active site, suggesting that it should not play a role in the enzymatic function. This agrees with the fact that the recombinant ricin *A* chain, which is not glycosylated, has full biological activity (O'Hare *et al.*, 1987). The existence of the oligosaccharide in β-MMC also does not influence the fold of the polypeptide chain. The exact function of the oligosaccharides in RIPs, therefore, remains unclear and needs to be studied. The three-dimensional structure of β-MMC is the first example of an RIP structure which provides information about the three-dimensional structure and the binding site of the oligosaccharide as well as its interactions with the protein. It may shed light on the structure–function relationship of the carbohydrates in RIPs.

We wish to thank Professor S.-W. Jin for kindly providing the protein and to thank Professors Y. Wang, S.-W. Jin and H.-M. Wu for providing the amino-acid sequence as well as the primary structure and the conformation in solution of the oligosaccharide of β-MMC. We are also grateful to Professors D.-C. Liang and Z.-L. Wan for their help and to Mr L. Zhang for his assistance during the X-ray data collection.

References

Ago, H., Kataoka, J., Tsuge, H., Habuka, N., Inagaki, E., Noma, M. & Miyano, M. (1994). *Eur. J. Biochem.* **225**, 369–374.
Barbieri, L., Battelli, M. G. & Stirpe, F. (1993). *Biochim. Biophys. Acta*, **1154**, 237–282.
Bouwstra, J. B., Spoelstra, E. C., De Waard, P., Leeflang, B. R., Kamerling, J. P., & Vliegenthart, J. F. G. (1990). *Eur. J. Biochem.* **190**, 113–122.
Brünger, A. T. (1992*a*). *X-PLOR Version* 3.1. *A System for X-ray Crystallography and NMR.* New Haven: Yale University Press.
Brünger, A. T. (1992*b*). *Nature (London)*, **335**, 472–475.
Cambillau, C. & Horjales, E. (1987). *J. Mol. Graph.* **5**, 174–178.
Collins, E. J., Robertus, J. D., Lopresti, M., Stone, K. L., Williams, K. R., Wu, P., Hwang, K. & Piatak, M. (1990). *J. Biol. Chem.* **265**, 8665–8669.
Crowther, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. Rossmann, pp. 173–178. New York: Gordon & Breach.
Crowther, R. A. & Blow, D. M. (1967). *Acta Cryst.* **23**, 544–548.
Endo, Y. & Tsurugi, K. (1987). *J. Biol. Chem.* **262**, 8128–8130.
Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.
Finzel, B. C. (1987). *J. Appl. Cryst.* **20**, 53–55.
Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.
Gao, B., Ma, X.-Q., Wang, Y.-P., Chen, S.-Z., Wu, S., Dong, Y.-C. (1994). *Sci. China Ser. B*, **37**, 59–73.
Hendrickson, W. A. (1985). *Methods Enzymol.* **115**, 252–270.
Ho, W. K. K., Liu, S. C., Shaw, P. C., Yeng, H. W., Ng, T. B. & Chan, W. Y. (1991). *Biochim. Biophys. Acta*, **1088**, 311–314.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulus, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.
Huang, Q., Liu, S., Tang, Y., Jin, S. & Wang, Y. (1995). *Biochem. J.* **309**, 285–298.
Husain, J., Tickle, I. J. & Wood, S. P. (1994). *FEBS Lett.* **342**, 154–158.
Jin, S.-W. (1993). Personal communication.
Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.
Katzin, B. J., Collins, E. J., & Robertus, J. D. (1991). *Proteins Struct. Funct. Genet.* **10**, 251 -259.
Kimura, Y., Hase, S., Kobayashi, Y., Kyogoku, Y., Funatsu, G. & Ikenaka, T. (1987). *J. Biochem.* **101**, 1051–1054.
Kimura, Y., Minami, Y., Tokuda, T., Nakayma, S., Takagi, S. & Funatsu, G. (1991). *Agric. Biol. Chem.* **55**, 2031–2036.
Lattman, E. E. & Love, W. E. (1972). *Acta Cryst.* B**26**, 1854–1857.
Lee-Huang, S., Huang, P. L., Nara, P. L., Chen, H.-C., Kung, H. F., Huang, P., Huang, H. I. & Huang, P. L. (1990). *FEBS Lett.* **272**, 12–18.
Monzingo, A. F., Collins, E. J., Earst, S. R., Irvin, J. D. & Robertus, J. D. (1993). *J. Mol. Biol.* **233**, 705–715.
Monzingo, A. F. & Robertus, J. D. (1992). *J. Mol. Biol.* **227**, 1136–1145.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). *Proteins Struct. Funct. Genet.* **12**, 345–364.
O'Hare, M., Reberts, L. M., Thorpe, P. E., Watson, G. J., Prior, B. & Lord, J. M. (1987). *FEBS Lett.* **216**, 73–78.
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.
Rossmann, M. G. (1972). *The Molecular Replacement Method*, edited by M. G. Rossmann, pp. 4–24. New York: Gordon & Breach.
Roussel, A. & Cambillau, C. (1991). *TURBO-FRODO.* Silicon Graphics Geometry Partners Directory.
Tronrud, D. E., Ten Eyck, L. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.
Wu, H.-M. & Jin, S.-W. (1993). Personal communication.
Xia, Z.-X., Zhang, L., Zhang, Z.-M., Wu, S. & Dong, Y.-C. (1993). *Chin. J. Chem.* **11**, 280–288.
Xiong, J.-P., Xia, Z.-X. & Wang, Y. (1994). *Nature Struct. Biol.* **1**, 695–700.
Xiong, J.-P., Xia, Z.-X. & Wang, Y. (1997). *Chin. J. Chem.* **15**, 265–277.
Xiong, J.-P., Xia, Z.-X., Zhang, L., Ye, G.-J., Jin, S.-W. & Wang, Y. (1994). *J. Mol. Biol.* **238**, 284–285.
Yeng, H. W., Li, W. W., Law, L. K., Chan, W. Y. & Ng, T. B. (1986). *Int. J. Pept. Protein Res.* **28**, 518–524.
Yeng, H. W., Li, W. W., Zhang, F., Barbieri, L. & Stirpe, F. (1988). *Int. J. Pept. Protein Res.* **31**, 265–268.
Yuan, Y.-R., Xiong, J.-P. & Xia, Z.-X. (1995). *Chin. Chem. Lett.* **6**, 1053–1054.
Zhang, J. S. & Liu, W. Y. (1992). *Nucleic Acids Res.* **20**, 1271–1275.
Zhang, L., Xia, Z.-X., Wu, S. & Dong, Y.-C. (1994). *Chin. J. Chem.* **12**, 223–230.