Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a novel ribosome-inactivating protein from a hemi-parasitic plant inhabiting the northwestern Himalayas

Vandana Mishra,[a,b] Abdul S. Ethayathulla,[a] Radhey S. Sharma,[b] Savita Yadav,[a] Ruth Krauspenhaar,[c] Christian Betzel,[c] Cherukuri R. Babu[b] and Tej P. Singh[a]*

[a]Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India, [b]Centre for Environmental Management of Degraded Ecosystems, School of Environmental Studies, and Department of Botany, University of Delhi, Delhi, India, and [c]Institute of Medical Biochemistry and Molecular Biology, Hamburg, Germany

Correspondence e-mail: tps@aiims.aiims.ac.in

Received 29 July 2004
Accepted 21 September 2004

PDB Reference: HmRip, 1pc8, r1pc8sf.

This is the first report of the structural studies of a novel ribosome-inactivating protein (RIP) obtained from the Himalayan mistletoe (*Viscum album*) (HmRip). HmRip is a type II heterodimeric protein consisting of a toxic enzyme (A-chain) with an active site for ribosome inactivation and a lectin subunit (B-chain) with well defined sugar-binding sites. The crystal structure of HmRip has been determined at 3.8 Å resolution and refined to a crystallographic R factor of 0.228 ($R_{\rm free}$ = 0.271). A comparison of this structure with other type II RIPs reveals the presence of distinct structural features in the active site of the A-chain and in the 2γ sugar-binding site of the B-chain. The conformation of the side chain of Tyr110, which is a conserved active-site residue in the A subunit, is strikingly different from those observed in other mistletoe RIPs, indicating its unique substrate-binding preference. The deletion of two important residues from the kink region after Ala231 in the 2γ subdomain of the B-chain results in a significantly different conformation of the sugar-binding pocket. A ribosome-recognition site has also been identified in HmRip. The site is a shallow cavity, with the conserved residues Arg51, Asp70, Thr72 and Asn73 involved in the binding. The conformations of the antigenic epitopes of residues 1–20, 85–103 and 206–223 differ from those observed in other type II RIPs, resulting in the distinct antigenicity and pharmacological properties of HmRip.

1. Introduction

Ribosome-inactivating proteins (RIPs) are translation inhibitors that are present mainly in plants (Barbieri *et al.*, 1993). RIPs can be broadly classified into two groups: type I and type II. Type I RIPs are monomeric enzymes with a molecular weight of ~30 kDa, while type II RIPs are heterodimers with a molecular weight of ~60 kDa and are composed of two chains: an enzyme homologous to type I RIPs (A-chain) and a lectin subunit (B-chain). The A-chain is responsible for the N-glycosidase activity that is the characteristic property of RIPs. It removes the adenine base from a universally conserved GAGA hairpin loop in 28S rRNA and abolishes the binding of elongation factor in protein synthesis. The lectin subunit contributes by recognizing the receptors on the cell membranes and facilitates the entry of toxic enzyme into the cytosol. The RIPs have a great pharmacological significance because of their potential usefulness in the treatment of cancer (Hajto *et al.*, 1989; Heiny *et al.*, 1998) and AIDS (Elhaggar, 1993; Zhao *et al.*, 1999). Several RIPs are currently in the advanced stages of clinical trials (Chan *et al.*, 2000; French *et al.*, 1996).

RIPs are multifunctional proteins and contain distinct structural motifs that are responsible for individual functional properties. The residues in the binding sites of various RIPs are highly conserved. However, the relative orientations of the side chains of these residues may differ greatly, resulting in variation in their ribosome-inactivating properties. Although details of the active site and sugar-binding sites have generally been obtained, less is known about the ribosome-recognition site. In fact, the ribosome-binding site of mistletoe RIPs has not yet been clearly identified. Although RIPs are highly antigenic in nature and their antigenic epitopes have been biochemically identified (Chan *et al.*, 2000; Agapov *et al.*, 1999; Leung *et al.*, 1997), the structural features of the epitopes have not yet been reported.

Mistletoe RIPs constitute the active principle of anticancer mistletoe preparations used in traditional tumour therapy in some parts of the world (Bussing, 2000). The biological properties of RIPs isolated from mistletoe inhabiting different hosts and geographical regions vary considerably (Franz, 1989). A number of type I and type II RIPs have been purified and biochemically characterized, but only a few crystal structures have been reported so far (Krauspenhaar *et al.*, 1999; Rutenber & Robertus, 1991; Pascal *et al.*, 2001; Tahirov *et al.*, 1995). In order to understand the variations in structural motifs associated with varying biological properties, structural determinations of mistletoe RIPs are essential. We have purified a novel type II RIP from mistletoe parasitizing *Pyrus pashia* (wild apple) from the northwestern Himalayas (HmRip) (Mishra, Sharma, Yadav *et al.*, 2004). Unlike other type II RIPs, which are specific for one or two sugars only, HmRip shows high affinity for a number of sugars such as L-rhamnose, *meso*-inositol, L-arabinose, *N*-acetyl galactosamine and galactose. Since the anticancer activity of mistletoe RIPs is associated with its capability to recognize sugars on the surface of immunogenic cells, the multisugar recognition of HmRip is important. In view of its unique properties, we have determined the three-dimensional structure of HmRip. The structure has revealed unusual features of the active site in subunit A and a sugar-binding motif in subunit B.

2. Materials and methods

2.1. Purification of HmRip

Field surveys of the northwestern Himalayas were carried out for the collection of plant material. Samples of *V. album* parasitizing *P. pashia* (wild apple) were collected. Frozen green plant parts were cut into small pieces and ground into a fine powder in liquid nitrogen. The total soluble protein was extracted in 100 m*M* sodium phosphate buffer containing 500 m*M* NaCl and 2.5% polyvinyl polypyrrolidone (PVPP) pH 7.6. The crude extract was centrifuged at 10 000*g* for 30 min and the supernatant was subjected to ammonium sulfate precipitation. The total protein was separated into two fractions: fraction I, which was 0–25% ammonium sulfate saturated, and fraction II, which was 25–80% ammonium sulfate saturated. The protein-rich fraction II was extensively dialyzed against phosphate-buffered saline (PBS) and loaded onto an affinity chromatography column packed with acid-hydrolyzed Sepharose 4B. The unbound proteins were washed off with the same buffer and HmRip was eluted with 0.2 *M* lactose, dialyzed and lyophilized. These samples showed a single band on SDS–PAGE. The protein was further checked by determining the sequence of the first 20 N-terminal amino acids using a PPSQ-21A protein sequencer (Shimadzu, Japan).

2.2. Complete amino-acid sequence

The complete amino-acid sequence of HmRip was determined by the cDNA cloning method. Total RNA was isolated

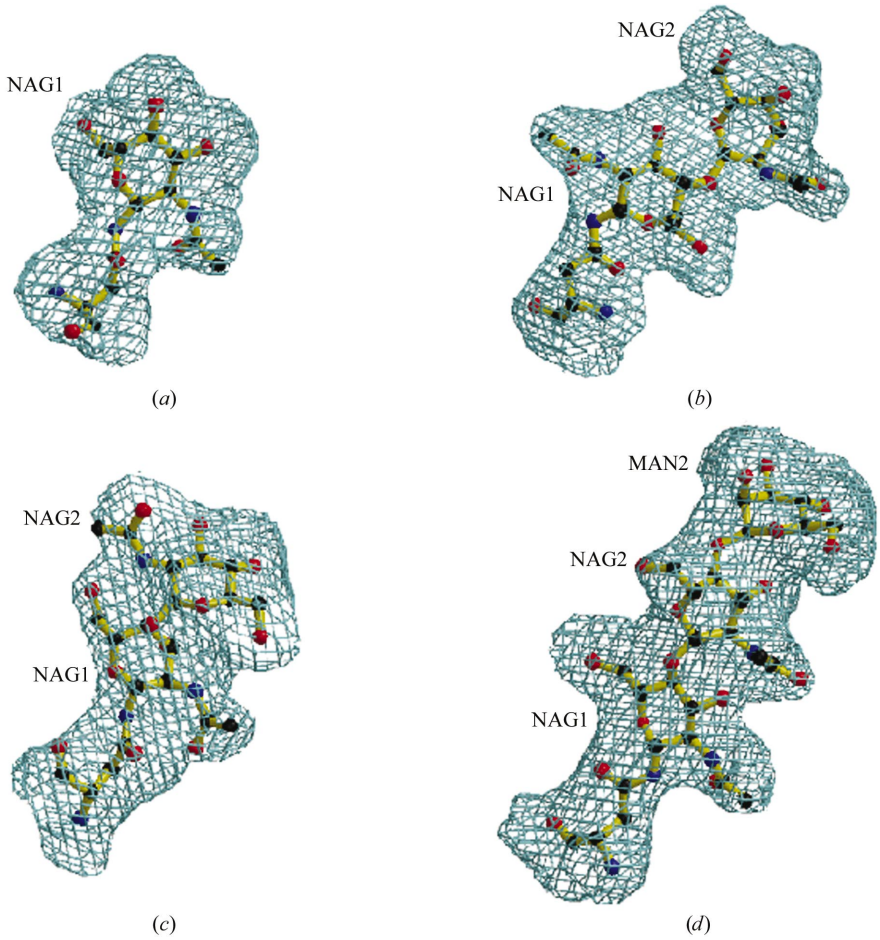


Figure 1
$|2F_o - F_c|$ electron density of the glycosylation sites at (*a*) Asn107 in the A-chain and (*b*) Asn57, (*c*) Asn92, (*d*) Asn132 in the B-chain of HmRip. The figure was drawn with *BOBSCRIPT* (Esnouf, 1997) and *RASTER3D* (Kraulis, 1991).

Table 1
Data-collection statistics.

Values in parentheses are for the last shell.

Wavelength (Å)	1.54
Space group	$P6_522$
Unit-cell parameters (Å)	
$a = b$	109.2
c	309.4
V_M (Å^3 Da^{-1})	4.3
Solvent content (%)	71.4
Z	12
Resolution range (Å)	32.0–3.8 (3.9–3.8)
No. of observed reflections	91625
No. of unique reflections	11420
Completeness (%)	99.5 (99.5)
R_{sym} (%)	17.9 (51.7)
$I/\sigma(I)$	6.1 (2.1)

Table 2
Summary of crystallographic refinement.

PDB code	1pc8
Resolution limits (Å)	32.0–3.8
No. reflections	11420
R_{cryst} (%)	22.8
R_{free} (%)	27.1
Protein atoms	3820
Carbohydrates (7 GlnAc and 1 Man)	109
Water molecules	13
R.m.s. deviations†	
Bond lengths (Å)	0.02
Bond angles (°)	2.3
Dihedral angles (°)	20.0
Overall G factor‡	−0.4
Average B factors (Å^2)	
Main-chain atoms	56.2
Side-chain atoms and waters	64.1
All atoms	59.2
Residues in most favoured regions (%)	80.1
Residues in additionally allowed regions (%)	19.9
Estimated coordinate error (after Luzzati, 1952) (Å)	0.39
Estimated coordinate error (from σ_A; Read, 1986) (Å)	0.36

† Target stereochemistry from Engh & Huber (1991) ‡ As reported by *PROCHECK* (Laskowski *et al.*, 1993)

from young frozen leaves and poly-(A)-rich mRNA and cDNA were synthesized as described by Mishra, Sharma, Paramiasivam *et al.*, (2004). PCR primers were designed by combination of the N-terminal amino-acid sequence of HmRip and the gene sequence of European ML-I. PCR was performed with Taq polymerase (Promega) using an MJ Research thermal cycler (model PTC-100). Nucleotide sequencing was performed on cloned double-stranded DNA (pGEM-T) using a DSQ-200L automated DNA sequencer (Shimadzu). The nucleotide sequence has been submitted to GenBank (AY625281).

2.3. Crystallization

Crystallization of HmRip was carried out using the hanging-drop vapour-diffusion method at 293 K. The protein was dissolved in 0.2 *M* phosphate buffer pH 7.6 to a final concentration of 10 mg ml^{-1}. 10 µl hanging drops were prepared and equilibrated against reservoir solution containing 35% $(NH_4)_2SO_4$, 20 m*M* NaCl pH 7.6. Hexagonal bipyramidal crystals of dimensions up to 0.3 × 0.2 × 0.2 mm were obtained in 3–4 weeks.

2.4. Data collection and processing

Attempts to cryoprotect these crystals for flash-cooling with a range of known cryoprotectants were unsuccessful and the crystals were therefore mounted in capillaries at 283 K for X-ray experiments. Data were collected from a single crystal on a MAR 345 image-plate scanner mounted on a Rigaku RU-300 X-ray generator with Osmic focusing mirrors. The data were processed and scaled using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The crystals belonged to space group $P6_522$, with unit-cell parameters $a = b = 109.2$, $c = 309.4$ Å. Details of the data-collection statistics are given in Table 1.

2.5. Structure determination and refinement

The structure has been determined by molecular replacement using the program *AMoRe* (Navaza, 1994). The coordinates of European ML-I (PDB code 1ce7) were used as a model (Krauspenhaar *et al.*, 1999). Both the rotation- and translation-function searches resulted in a clear single solution. Rigid-body refinement of the resulting model gave a correlation coefficient of 61.4% and an R factor of 32.7% in the resolution range 10.0–3.5 Å. The model was refined using *REFMAC* (Murshudov *et al.*, 1997). The initial steps of manual model building with the program *O* (Jones *et al.*, 1991) using $|2F_o - F_c|$ and $|F_o - F_c|$ Fourier maps on a Silicon Graphics O2 workstation gave R and R_{free} factors of 0.302 and 0.347, respectively. Densities for carbohydrate residues were clearly evident at Asn107 in the A-chain and Asn57, Asn92 and Asn132 in the B-chain (Fig. 1). These consisted of N-linked *N*-acetyl glucosamine at all the four sites. These molecules were included in further cycles of refinement with several rounds of manual model building. Although the positions of water molecules were not sought at this resolution, there were good densities in both $|2F_o - F_c|$ and $|F_o - F_c|$ for 13 water molecules and these made perfect hydrogen bonds with protein atoms; these were therefore included in the further refinement, to which they responded well. A bulk-solvent correction was also performed. The R and R_{free} factors finally converged to 0.228 and 0.271, respectively. The refinement statistics are given in Table 2. The program *PROCHECK* was used to validate the quality of the final structure (Laskowski *et al.*, 1993). 80% of the residues were found in the most favoured regions of the Ramachandran plot (Ramachandran & Sasisekharan, 1968) and the remaining residues were in the additionally allowed regions.

3. Results and discussion

3.1. Sequence analysis of HmRip

The complete amino-acid sequence determination of HmRip shows that the A-chain consists of 240 amino-acid

residues, while the B-chain contains 255 amino-acid residues. Both the A- and B-chains in HmRip are shorter than in ML-I and other type II RIPs (Fig. 2). Both chains show nearly 89% sequence identity with the respective chain of ML-I (Niwa *et al.*, 2003; Krauspenhaar *et al.*, 1999). However, the sequence identities with the A- and B-chains of other type II RIPs are considerably lower. The A-chain of HmRip shows 38% sequence identity with the A-chain of ricin and 62% with the B-chain. Similarly, the corresponding values for abrin and ebulin are 41 and 37% for the A-chain and 54 and 40% for the B-chain, respectively. Although the sequence identities between the respective chains of HmRip and ML-I are high, the observed deletions and variations in their sequences are of great structural significance. The absence of seven residues between Ser89 and Gly90 in the A-chain of HmRip compared with three deletions in ML-I should be of considerable interest (Fig. 2). There are several other differences in the A-chain compared with the A-chain of ML-I that appear to be responsible for important functional differences between the two proteins. Similarly, critical deletions and variations in the B-chains of HmRip and ML-I are responsible for the significant observed structural and functional variations. That two proteins have been found to show rather striking differences in their sugar specificities and hence in their responses towards cancer has wide implications. The sequences of the A-chain and B-chain reveal one and three potential N-glycosylation sites, respectively, with Asn-*X*-Ser/Thr motifs; these are at Asn107 in the A-chain and Asn57, Asn92 and Asn132 in the B-chain. Chemical studies in the similar protein ML-I have shown all four sites to be glycosylated (Niwa *et al.*, 2003). Our crystallographic analysis of HmRip clearly confirms attachment at all four sites and the remarkably good quality of the electron density has enabled us to model no fewer than eight sugar residues. The electron density at all the sites clearly corresponded to *N*-acetylglucosamine (NAG) residues, which also tallied well with those reported for ML-I (Niwa *et al.*, 2003; Krauspenhaar *et al.*, 1999).

3.2. Quality of the model

The final coordinate set consists of 3820 protein atoms from 495 amino-acid residues and eight sugar-moiety units (seven GlnAc and one Man) and 13 water molecules. The protein structure has geometry close to ideal values, with r.m.s. deviations of 0.02 Å and 2.3° from standard values of bond lengths and bond angles, respectively. The overall mean *B* factor was found to be 59.2 Å^2. The high values of the *B* factors and the high solvent content of 71% were probably the two main causes for the poor quality of the HmRip crystals. Despite this, the electron density was very well defined in almost all the regions of protein.

3.3. Overall structure

The general organization of HmRip is similar to that of other type II RIPs (Krauspenhaar *et al.*, 2002; Pascal *et al.*, 2001; Tahirov *et al.*, 1995; Rutenber & Robertus, 1991). It may be noted that HmRip, with 495 amino-acid residues, is the smallest protein of the known type II RIP structures ML-I (511 residues), ricin (529 residues), abrin (518 residues) and ebulin (520 residues) (Fig. 2). HmRip forms a heterodimer in which the two subunits are covalently linked through a disulfide bridge (Fig. 3). The two subunits associate in a T-like arrangement. The A-chain (toxin subunit) of 240 amino-acid

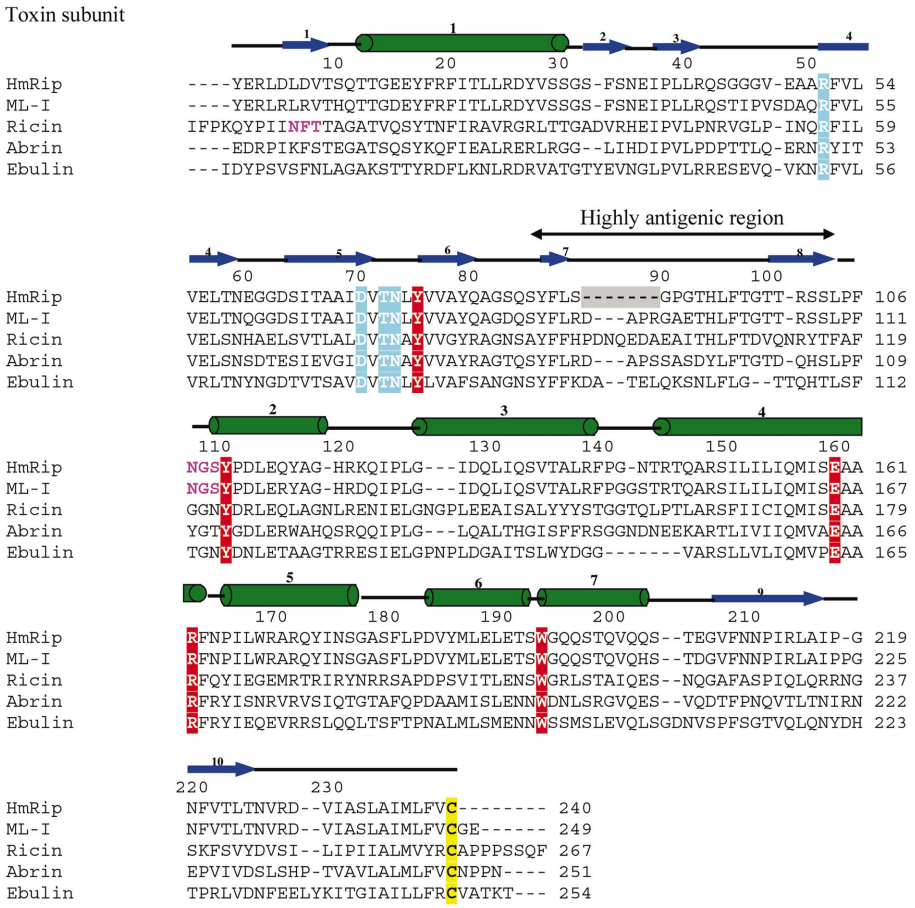


Figure 2
Sequence comparison of type II RIPs from Himalayan mistletoe (HmRip), European mistletoe (ML-I), *Ricinus communis* (ricin), *Abrus precatorius* (abrin) and *Sambucus ebulus* (ebulin). The secondary structure of HmRip is shown by representing α-helices as cylinders (green), β-sheets as arrows (blue) and loops as lines (black). Potential glycosylation sites are shown in magenta and conserved cysteines in yellow. In the A-chain (toxin subunit), the key active-site residues (N-glycosidase activity) are highlighted in red and conserved residues involved in ribosome recognition are highlighted in blue. The characteristic deletion of seven residues in the highly antigenic site is highlighted in grey.

residues folds into two non-homologous domains A1 and A2. The B subunit is a lectin with 255 amino-acid residues and folds into two well separated homologous domains B1 and B2. Domain A1 is comprised of residues 1–13 and 44–155 and contains both β-strands and α-helices in a nearly equal distribution. The A2 domain contains residues 14–43 and 156–240 and folds with a predominantly α-helical conformation. A1 contains six β-strands, β1 and β4–β8, and two helices, α2 and α3. Helix α4 crosses over to the A2 domain and is shared by domains A1 and A2 (Fig. 3). The A2 domain contains α1, β2–β3, α5–α7 and β9–β10, with its α-helical content being nearly twice its β-strand content. In contrast, the A1 domain of other type II RIPs such as ML-I (Krauspenhaar *et al.*, 1999), ricin (Rutenber & Robertus, 1991) and abrin (Tahirov *et al.*, 1995) contains four helices. One of the most striking deletions in HmRip corresponds to the seven helix-forming residues between β-strands β7 and β8 (Fig. 2). Even though three residues are deleted, in ML-I it still forms a well defined α-helix which connects the β-strands β7 and β8. In the A1 domain, α-helix α2 (112–116) is slightly shorter in HmRip than usually found in RIPs, while α3 (126–138) and α4 (143–155) are well conserved. Although most of the α-helices and β-strands are well conserved in these proteins, the loops between β3 and β4 and α3 and α4 are conformationally different from those in other RIPs. The loop between α3 and α4 is significantly shorter in HmRip in comparison with the type II RIP family including ML-I.

The basic structure of the B-chain is highly conserved. It folds into two distinct domains, B1 (1–132) and B2 (133–255). The domains are further divided into subdomains called λ, α, β and γ. The λ subdomains 1λ (1–13) and 2λ (133–144) link chains B and A and domains B1 and B2, respectively. The α and β subdomains are represented by a pair of β-sheets formed by two antiparallel strands each. The γ subdomains are represented by one β-sheet and a single strand. Sugar binding takes place in two non-homologous sites 1α (14–56) and 2γ (224–255). The 1β (57–97), 2α (145–179) and 2β (188–221) subdomains contain each an S—S bond between the conserved cysteine residues Cys60–Cys77, Cys148–Cys161 and Cys186–Cys202, respectively. However, the substitutions for conserved cysteines by Asn17 and Ser36 (Fig. 2) do not allow the formation of a disulfide bond in the 1α subdomain. The absence of this conserved S—S bond is a characteristic feature of mistletoe RIPs. The A1 and A2 domains are predominantly associated by hydrophobic interactions, while the B1 and B2 domains interact with each other primarily through hydrogen bonds. The formation of a hydrogen bond between Ser36 and Thr134 of the B1 and B2 domains, respectively, of HmRip is one of its characteristic features. The corresponding interaction is absent in other RIPs, where Ser36 is substituted by Cys36.

Overall, only a few polar interactions are present between the two subunits, although an extended network of van der Waals forces is observed. At the interface of the two subunits a shallow groove is formed by the B-chain, in which part of the A-chain is buried. Arg227 of the A-chain is inserted into the interdomain spacing formed by the B1 and B2 domains and makes a hydrogen bond with the carbonyl oxygen atoms of Trp128 and Gly91 of the B-chain. The interaction between Asp228 of the A-chain and Arg137 of the B-chain represents a characteristic

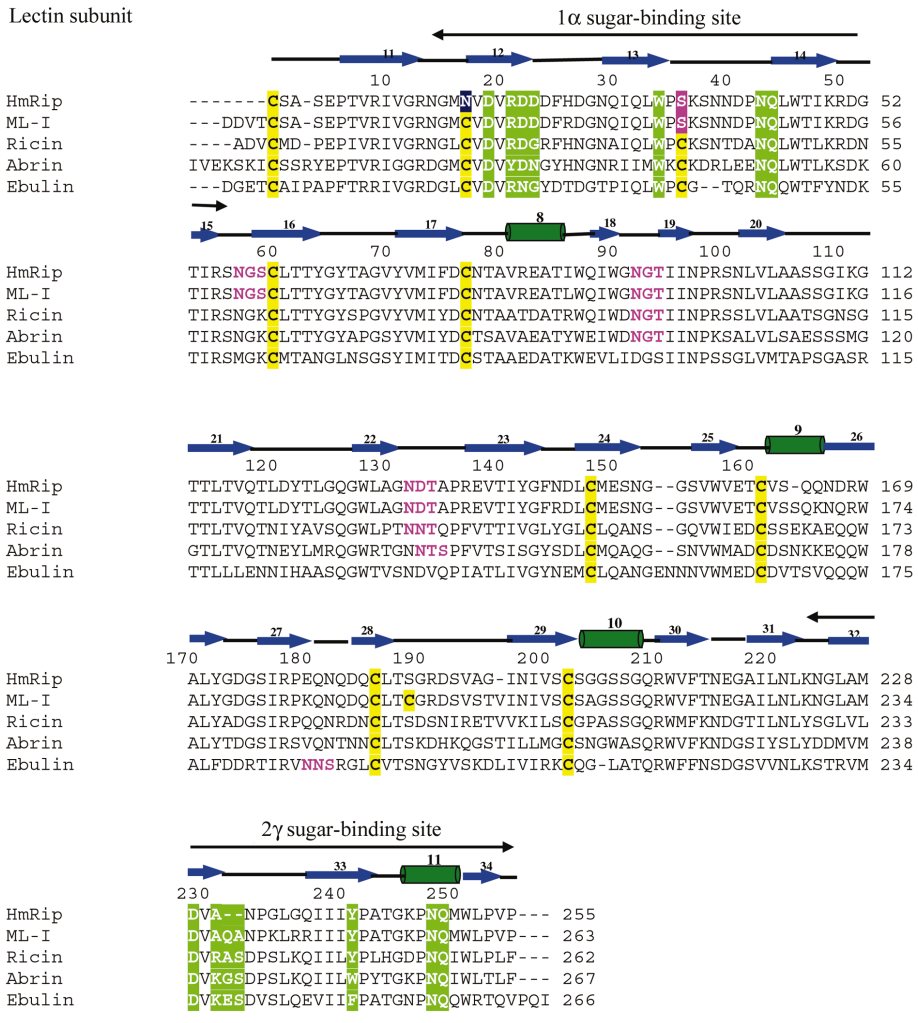


Figure 2 (continued)
In the B-chain (lectin subunit), the kink region forming the base of the sugar-binding pocket, Arg21–Asp23 (1α site) and Ala231 (2γ site), the conserved aromatic residue forming top of the pocket, Trp33 (1α site) and Tyr241 (2γ site), and the sugar-binding residues (Asp19, Asn43, Gln44 in 1α and Asp229, Asn248, Gln249 in 2γ) are highlighted in green. Two conserved cysteines have been substituted by Asn17 (highlighted in purple) and Ser36 (highlighted in pink) in HmRip.

feature of mistletoe RIPs. Four potential glycosylation sites exist inthe molecule at Asn107A, Asn57B, Asn92B and Asn132B and all were found to be glycosylated. The B1 domain is highly glycosylated, while the A2 and B2 domains have no glycosylation sites. In case of the A1 domain the glycosylation site is located very close to the active site.

3.4. Functionally important sites

Although the overall structure of HmRip is fairly similar to those of the other type II RIPs, striking differences were observed in the functionally significant regions, including the active site, sugar-binding site, antigenic epitopes and ribosome-recognition site.

3.4.1. Active site. The N-glycosidase activity site is located in a cleft formed by the association of domains A1 and A2 in the toxin subunit (Fig. 3). The C-terminal end of helix $\alpha4$ that crosses over from domain A1 to domain A2 forms the base of the cleft, while the loops between $\beta5$–$\beta6$ and $\beta8$–$\alpha2$ constitute

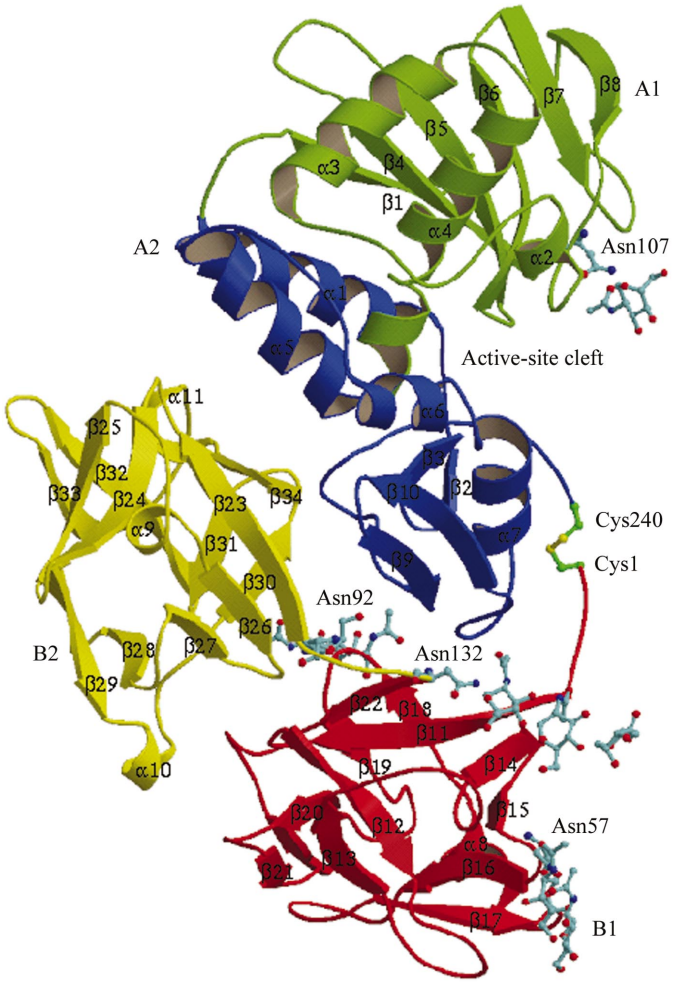


Figure 3
Overall structure of HmRip. The four domains are shown in different colours: A1 in green, A2 in blue, B1 in red and B2 in yellow. At the interface of the A1 and A2 domains, the active-site cleft is shown. The helices are indicated as $\alpha1$–$\alpha11$ and the sheets $\beta1$–$\beta34$. The disulfide bridge is shown between Cys240A and Cys1B. The carbohydrates are shown in grey. The figure was drawn with *MOLSCRIPT* (Merritt & Murphy, 1994) and *RASTER3D* (Kraulis, 1991).

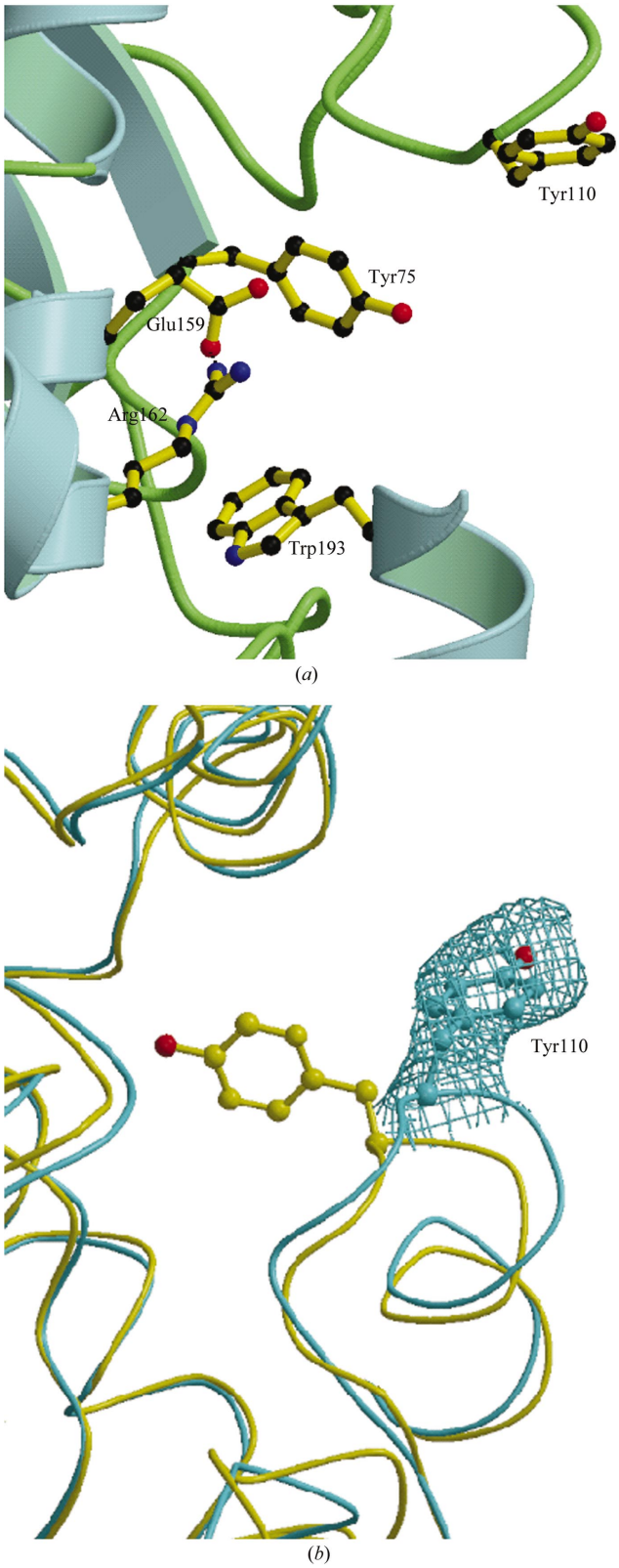


Figure 4
(*a*) Positioning of the key active-site residues (Tyr75, Tyr110, Glu159, Arg162 and Trp193) of the N-glycosidase activity site in the A-chain. (*b*) Overlap of the segments containing Tyr110 from the structure of HmRip (cyan) and ML-I (yellow). $|F_o - F_c|$ omit map at 1.5σ cutoff for the side chain of Tyr110 of HmRip. The side chain of Tyr110 in ML-I is also shown. The figure was drawn with *BOBSCRIPT* (Esnouf, 1997) and *RASTER3D* (Kraulis, 1991).

one of the walls. Helices α6 and α7 form the opposite wall. The residues Tyr75, Tyr110, Glu159, Arg162 and Trp193 represent the key active-site residues (Fig. 4*a*). These are conserved in the RIP family (Fig. 2). Glu159 and Arg162 are the principal residues that participate in catalysis. Both of them belong to the C-terminal part of the α-helix α4 and their side chains are aligned side-by-side in an antiparallel fashion and interact electrostatically. Arg162 also forms a hydrogen bond to Asn73. Trp193 belongs to helix α7 and protrudes into the active site. Trp193 NE2 forms a hydrogen bond with Ile235 O. This is a conserved interaction in type II RIPs. The tyrosines Tyr75 and Tyr110 belonging to β6 and α2, respectively, are placed on either side of the substrate-binding cleft. The most striking observation corresponds to the orientation of the Tyr110 side chain in HmRip (Fig. 4*b*); it adopts an entirely different conformation by flipping outward, resulting in a wider opening of the binding site. The corresponding tyrosine residues in ML-I (Krauspenhaar *et al.*, 1999; Niwa *et al.*, 2003) are parallel to each other. The unique arrangement of these two Tyr residues in HmRip indicates that it may be involved in accommodating a wider range of substrates.

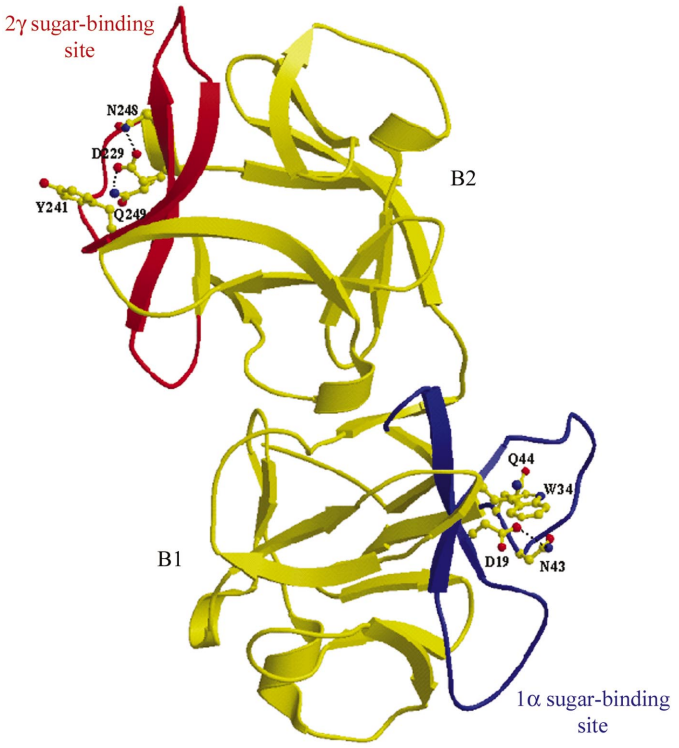


Figure 5
Overall fold of the B-chain (yellow) showing the location of the 2γ sugar-binding site (red) in the B2 domain and the 1α sugar-binding site (blue) in the B1 domain. The sugar-binding residues are also shown as a ball-and-stick model. The figure was drawn with *MOLSCRIPT* (Merritt & Murphy, 1994) and *RASTER3D* (Kraulis, 1991).

3.4.2. Sugar-binding sites. Two characteristic sugar-binding sites have been observed in the B-chain (lectin subunit) of type II RIPs. That corresponding to high affinity is located in the 2γ subdomain, while the low-affinity site is found in the 1α subdomain (Fig. 5). Generally, sugar-binding sites have a shallow architecture. The bottom of the sugar-binding site in RIPs is characterized by a three-residue kink formed in the protein main chain. Conserved aromatic residues form the top of the pockets in these proteins. A comparison of the sugar-binding sites of RIPs has revealed that the overall architecture of the 2γ sugar-binding site in HmRip is strikingly different from those found in ML-I and ricin, while the 1α site is very similar in all of them (Fig. 6). The unique structural features of the 2γ sugar-binding pocket in HmRip seem to have been introduced by the deletion of two critical residues after Ala232 from the kink region that is present in the lectin subunits of other type II RIPs (Fig. 2). This also causes the narrowing of the loop at the kink, which is 4 Å wide in HmRip but 8 Å wide in ML-I, ricin and abrin (Fig. 6). The presence of a kink in the backbone of ML-I, ricin and abrin makes the sugar-binding site narrow and deeper compared with that in HmRip, which is wider and shallow. Biochemical studies have shown that HmRip has a similar affinity for *N*-acetyl galactosamine and galactose, as the widening of the pocket provides sufficient space to accommodate the *N*-acetyl group. In contrast, ML-I lacks affinity for *N*-acetyl galactosamine owing to steric hindrance (Ziska *et al.*, 1993). It may be mentioned here that the ML-I, ricin and abrin are galactose-specific RIPs, whereas HmRip is shown to have a

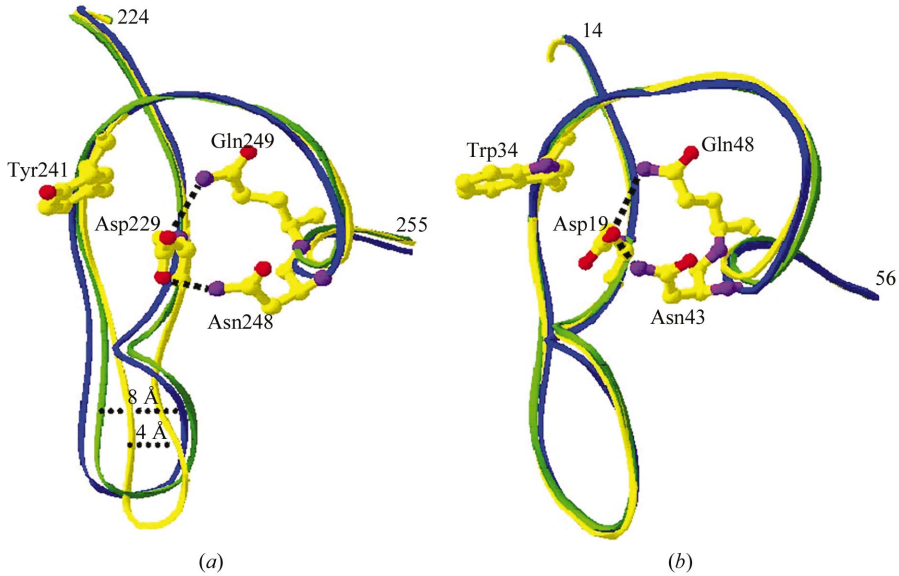


Figure 6
Superposition of (*a*) the 2γ and (*b*) the 1α sugar-binding sites of HmRip (yellow), ML-I (blue) and ricin (green). The 2γ site shows a strikingly different architecture in HmRip in comparision to ML-I and ricin. Absence of the kink in HmRip results in the narrowing of the sugar-binding loop by 4 Å. The 1α site shows a conserved conformation. The figure was drawn with *Swiss PDB Viewer* (Guex & Peitsch, 1997).

novel affinity for L-rhamnose, L-arabinose and *meso*-inositol in addition to galactose and *N*-acetyl galactosamine. On the other hand, the 1α sugar-binding sites in HmRip, ML-I, ricin and abrin are structurally similar and hence show similar preferences for sugar binding. It may be emphasized that owing to significant conformational differences of the 2γ site in HmRip, the sugar is expected to interact in a different manner, as has been suggested by sugar-binding studies (Mishra, Sharma, Yadav *et al.*, 2004). Since the recognition of the sugar chains present on the surface of the immunogenic cells is the first step in biological activity, the unique sugar affinity of HmRip has a great significance in determining its pharmacological properties.

3.4.3. Antigenic epitopes. RIPs are known to be among the most toxic natural substances so far known (Barbieri *et al.*, 2004). A great number of studies have been carried out to attempt to combat their toxic effects. One of the most successful methods is the synthesis of neutralizing antibodies against the A-chain. Peptide-binding studies have identified the highly antigenic regions in RIPs (Lebeda & Olson, 1999). In HmRip, it is the segment 85–103 (Fig. 7), while in ML-I and ricin it corresponds to the 86–108 and 90–116 segments, respectively. In HmRip, it is represented by β-strands β7 and β8 and the loop in between. Two other antigenic epitopes identified in these molecules include N-terminal (1–20) and C-terminal (206–223) fragments (Fig. 7). The N-terminal site is represented by a β-strand (β1), loop (α1–β1) and an α-helix (α1) and the organizations of these elements have been found to adopt slightly different arrangements in various RIPs. The C-terminal sites, which are represented by an antiparallel β-sheet (β9–β10) protruding out like a β-wing in the A2 domain, also adopt unique features in different RIPs. Thus, the crystal structure studies of various RIPs showed that the antigenic sites adopt distinct conformations, leading to the loss of cross-reactivity observed among antibodies raised against different RIPs.

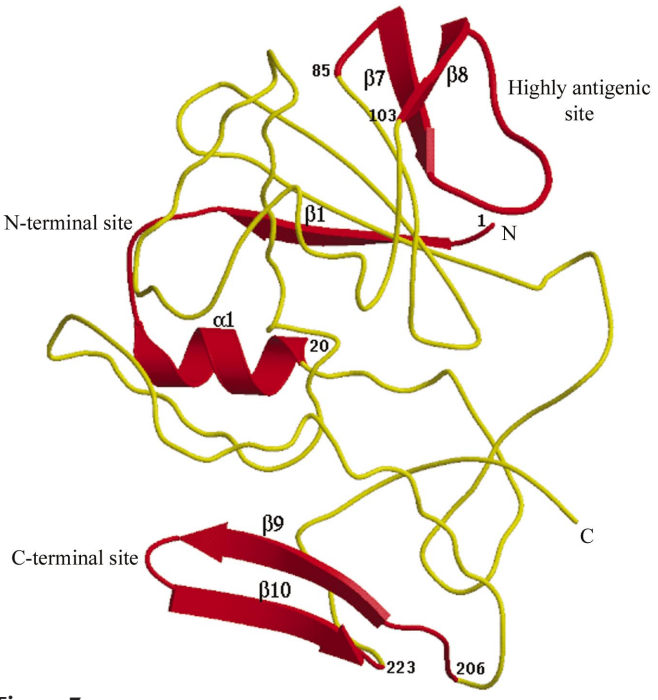


Figure 7
The three antigenic epitopes (red) present in the A-chain (yellow). The figure was drawn with *MOLSCRIPT* (Merritt & Murphy, 1994) and *RASTER3D* (Kraulis, 1991).

3.4.4. Ribosome-recognition site. One of the major focuses of current investigations on RIPs concerns the identification of the ribosome-recognition site. So far, however, it has not been clearly established (Savino *et al.*, 2000; Vater *et al.*, 1995). As mentioned previously, a highly antigenic peptide from the A1 domain binds strongly to neutralizing antibodies. Antibodies binding to the site can neutralize the toxicity, but the highly antigenic peptide is distinct from the N-glycosidase activity site, suggesting that the antibodies might be binding at or near the site involved in ribosome recognition. The ribosome-recognition site of mistletoe RIPs has not been reported so far. Analysis of the three-dimensional structure of HmRip revealed the presence of a well defined cleft near the antigenic peptide (Fig. 8*a*). The site is represented by a 22 Å wide shallow cleft in the A1 domain. The conserved β-strands β5 and β6 form the base of the cleft, while the loops between β3–β4 and β8–α2 form its two walls. The loop between β7 and β8, which also forms a part of the antigenic peptide, constitutes the roof of the cleft (Fig. 8*a*). The side chains of Arg51, Asp70, Thr72 and Asn73 point inside the cleft. These residues form a network of polar contacts available for interaction with the substrate. It may be noted that all the four residues are highly conserved in RIPs (Fig. 2).

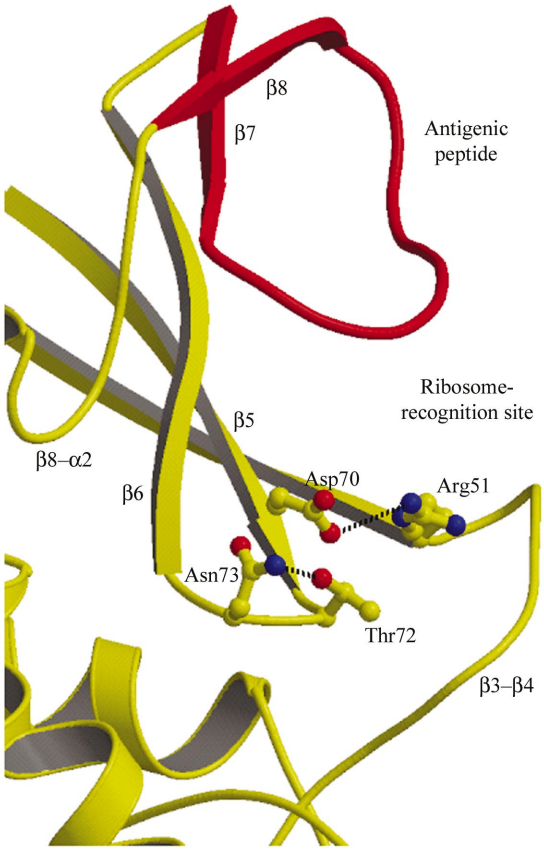
The ribosome-recognition site is formed as a shallow structure on the surface, while the N-glycosidase site corresponds to a well designed deep pocket (Fig. 8*b*). The sites are connected by a shallow channel. The binding of RIP and ribosome may occur as a two-step mechanism because of the large size of the ribosome. In this, the ribosome-recognition site binds to the ribosome first and holds it in a favourable position for attack on the rRNA in the second step. This mechanism is also supported by the fact that RIPs do not act on prokaryotic ribosomes, although they can attack the naked rRNA of prokaryotes as well as that of eukaryotes with equal affinity (Robertus *et al.*, 1991; Lord *et al.*, 1994). The distinct demarcation of the ribosome-recognition site and the N-glycosidase activity also support the double-step mechanism.

The comparison of the ribosome-recognition sites in various RIPs has shown that the region is represented by a well defined cleft. It is important to note that the highly antigenic peptide forming the top of the cleft is highly variable (Fig. 8*c*). Therefore, the overall structure of the ribosome-recognition site also varies in different RIPs. In ricin, the roof extends forward, whereas in HmRip, owing to the absence of a complete helix, the region has become slightly broader. The cavity is much wider in HmRip compared with other RIPs and hence differs in ribosome specificity. It should also be mentioned here that the RIPs act in a species-specific manner and different RIPs are known to recognize different ribosomal proteins (Savino *et al.*, 2000; Vater *et al.*, 1995; Hudak *et al.*,

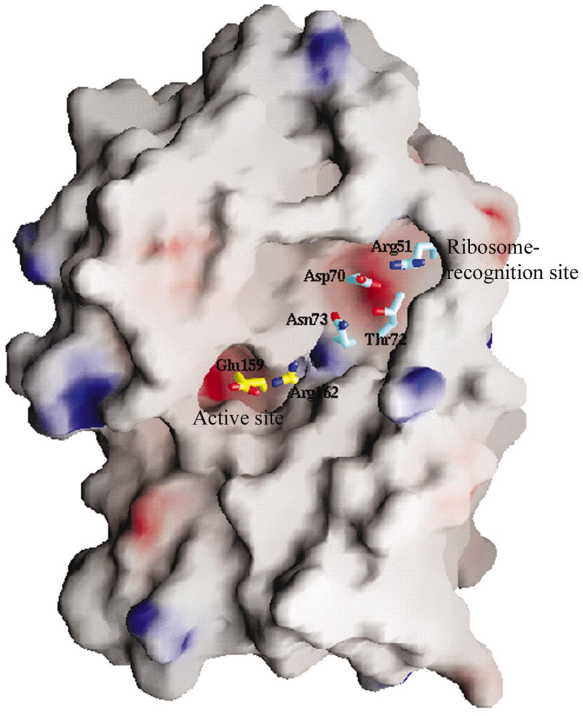
1999). The conformational differences observed in the overall structures of the ribosome-recognition site might be responsible for the species specificity among the RIPs.

4. Conclusions

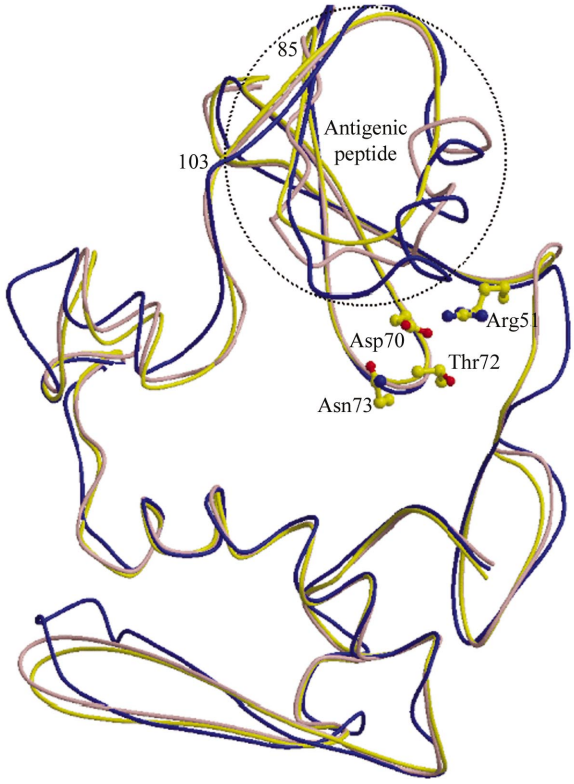
The three-dimensional structure of HmRip has revealed unique structural features of the functionally important sites. The active-site residue Tyr110 adopts a very different conformation in HmRip in which the side chain of Tyr110 is flipped in the opposite direction at the $C^{\alpha}-C^{\beta}$ bond, which has strong functional implications. The 2γ sugar-binding site in HmRip adopts a novel architecture which is very different from those reported so far, resulting in the unique sugar affinity. The structure analysis has also indicated the presence of a ribosome-binding site with highly characteristic features. Knowledge of these sites is important for pharmacological applications and the variations observed in the case of HmRip have strong potential for medical applications against a wide range of cancers. These studies also encourage the extension of the structural investigations of mistletoe RIPs purified from *V. album* parasitizing on diverse hosts and habitat so that other critical differences can be determined and new and novel therapeutically important RIPs can be identified.



(*a*)



(*b*)



(*c*)

Figure 8
The ribosome-recognition site has its roof represented by the highly antigenic peptide (red). The ribosome-binding residues are shown as a ball-and-stick model. (*b*) *GRASP* (Nicholls *et al.*, 1991) figure showing a shallow cavity representing the ribosome-recognition site and a well defined cleft corresponding to the N-glycosidase activity site, connected by a shallow channel. The ribosome-binding residues are shown in cyan and key active-site residues are shown in yellow. (*c*) Superposition of the C^{α} traces of HmRip (yellow), ML-I (pink) and ricin (blue), showing the highly variable nature of the antigenic peptide that results in different architectures of the ribosome-recognition sites.

The authors acknowledge the financial assistance from the Department of Biotechnology, Government of India. VM thanks CSIR, New Delhi, India for granting a fellowship. The authors thank Professor W. Voelter for useful discussions.

References

Agapov, I. I., Tonevitsky, A. G., Maluchenko, N. V., Moisenovich, M. M., Bulah, Y. S. & Kirpichnikov, M. P. (1999). *FEBS Lett.* **464**, 63–66.

Barbieri, L., Battelli, M. G. & Stirpe, F. (1993). *Biochim. Biophys. Acta*, **1154**, 237–287.

Barbieri, L., Ciani, M., Girbes, T., Liu, W. Y., Van Damme, E. J., Peumans, W. J. & Stirpe, F. (2004). *FEBS Lett.* **563**, 219–222.

Bussing, A. (2000). *Mistletoe: The Genus Viscum.* The Netherlands: Harwood Academic Publisher.

Chan, S. H., Shaw, P. C., Mulot, S. F. C., Xu, L. H., Chan, W. L., Tam, S. C. & Wong, K. B. (2000). *Biochem. Biophys. Res. Commun.* **270**, 279–285.

Elhaggar, S. (1993). *Arch. AIDS Res.* **7**, 120–121.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.

Franz, H. (1989). *Advances in Lectin Research.* New York: Springer.

French, R. R., Bell, A. J., Hamblin, T. J., Tutt, A. L. & Glennie, M. J. (1996). *Leuk. Res.* **20**, 607–617.

Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.

Hajto, T., Hostanska, K. & Gabius, H. J. (1989). *Cancer Res.* **49**, 4803–4808.

Heiny, B. N., Albrecht, V. & Beuth, J. (1998). *Anticancer Res.* **18**, 583–586.

Hudak, K. A., Dinman, J. D. & Tumer, N. E. (1999). *J. Biol. Chem.* **274**, 3859–3864.

Jones, T. A., Zou, Z., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Krauspenhaar, R., Eschenburg, S., Perbandt, M., Kornilov, V., Konareva, N., Mikailova, I., Stoeva, S., Wacker, R., Maier, T., Singh, T. P., Mikhailov, A., Voelter, W. & Betzel, C. (1999). *Biochem. Biophys. Res. Commun.* **257**, 418–424.

Krauspenhaar, R., Rypniewski, W., Kalkura, N., Moore, K., DeLucas, L., Stoeva, S., Kikhailov, A., Voelter, W. & Betzel, C. (2002). *Acta Cryst.* D**58**, 1704–1707.

Laskowski, R., MacArthur, M., Moss, D. & Thornton, J. (1993). *J. Appl. Cryst.* **26**, 283–290.

Lebeda, F. J. & Olson, M. A. (1999). *Int. J. Biol. Macromol.* **24**, 19–26.

Leung, K. C., Meng, Z. Q. & Ho, K. K. (1997). *Biochim. Biophys. Acta*, **1336**, 419–424.

Lord, J. M., Roberts, L. M. & Robertus, J. D. (1994). *FASEB J.* **8**, 201–208.

Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Mishra, V., Sharma, R. S., Paramasivam, M., Bilgrami, S., Yadav, S., Srinivasan, A., Betzel, C., Babu, C. R. & Singh, T. P. (2004). In the press.

Mishra, V., Sharma, R. S., Yadav, S., Babu, C. R. & Singh, T. P. (2004). *Arch. Biochem. Biophys.* **423**, 288–301.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.

Niwa, H., Tonevitsky, A. G., Agapov, I. I., Saward, S., Pfüller, U. & Palmer, R. A. (2003). *Eur. J. Biochem.* **270**, 2739–2749.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **176**, 307–326.

Pascal, J. M., Day, P. J., Monzingo, A. F., Ernst, S. R., Robertus, J. D., Iglesias, R., Perez, Y., Ferreras, J. M., Citores, L. & Girbes, T. (2001). *Proteins*, **43**, 319–326.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.

Robertus, J. D., Kim, Y. & Ready, M. P. (1991). *Proteins*, **10**, 270–278.

Rutenber, E. & Robertus, J. D. (1991). *Proteins*, **10**, 260–269.

Savino, C., Federici, L., Ippoliti, R., Lendaro, E. & Tsernoglou, D. (2000). *FEBS Lett.* **470**, 239–243.

Tahirov, T. H., Lu, T. H., Liaw, Y. C., Chen, Y. L. & Lin, J. Y. (1995). *J. Mol. Biol.* **250**, 354–367.

Vater, C. A., Bartle, L. M., Leszyk, J. D. & Lambert, J. M. (1995). *J. Biol. Chem.* **270**, 12933–12940.

Zhao, J., Ben, L. H., Wu, Y. L., Hu, W., Ling, K., Xin, S. M., Nie, H. L., Ma, L. & Pei, G. J. (1999). *Exp. Med.* **190**, 101–111.

Ziska, P., Gelbin, M. & Franz, H. (1993). *Lectin Biol. Biochem. Clin. Biochem.* **8**, 10–13.