Mistletoe lectin I in complex with galactose and lactose reveals distinct sugar-binding properties

**Ruth Mikeska,[a] Roland Wacker,[b]
Raghuvir Arni,[c] Tej P. Singh,[d]
Albert Mikhailov,[e] Azat
Gabdoulkhakov,[e] Wolfgang
Voelter[b] and Christian Betzel[a]***

[a]Institute of Biochemistry and Food Chemistry,
University of Hamburg, c/o DESY,
Notkestrasse 85, Building 22a, 22603 Hamburg,
Germany, [b]Institute of Physiological Chemistry,
University of Tübingen, Hoppe-Seyler-Strasse 4,
72076 Tübingen, Germany, [c]Department of
Physics, IBILCE/UNESP, São Jose do Rio Preto,
São Paul, Brazil, [d]Department of Biophysics, All
India Institute of Medical Sciences, New Delhi,
India, and [e]Institute of Crystallography of
Russian Academy of Sciences, Leninsky
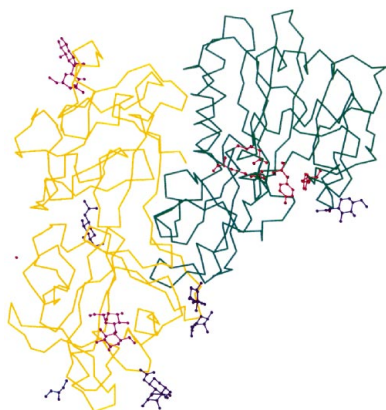Prospect 59, 117333 Moscow, Russia

Correspondence e-mail: betzel@unisgi1.desy.de

The structures of mistletoe lectin I (ML-I) from *Viscum album* complexed with lactose and galactose have been determined at 2.3 Å resolution and refined to R factors of 20.9% (R_{free} = 23.6%) and 20.9 (R_{free} = 24.6%), respectively. ML-I is a heterodimer and belongs to the class of ribosome-inactivating proteins of type II, which consist of two chains. The A-chain has rRNA N-glycosidase activity and irreversibly inhibits eukaryotic ribosomes. The B-chain is a lectin and preferentially binds to galactose-terminated glycolipids and glycoproteins on cell membranes. Saccharide binding is performed by two binding sites in subdomains α1 and γ2 of the ML-I B-chain separated by ~62 Å from each other. The favoured binding of galactose in subdomain α1 is achieved *via* hydrogen bonds connecting the 4-hydroxyl and 3-hydroxyl groups of the sugar moiety with the side chains of Asp23B, Gln36B and Lys41B and the main chain of 26B. The aromatic ring of Trp38B on top of the preferred binding pocket supports van der Waals packing of the apolar face of galactose and stabilizes the sugar–lectin complex. In the galactose-binding site II of subdomain γ2, Tyr249B provides the hydrophobic stacking and the side chains of Asp235B, Gln238B and Asn256B are hydrogen-bonding partners for galactose. In the case of the galactose-binding site I, the 2-hydroxyl group also stabilizes the sugar–protein complex, an interaction thus far rarely detected in galactose-specific lectins. Finally, a potential third low-affinity galactose-binding site in subunit β1 was identified in the present ML-I structures, in which a glycerol molecule from the cryoprotectant buffer has bound, mimicking the sugar compound.

1. Introduction

Lectins are known to provide molecular recognition and play key roles in diverse biological processes. As such, they provide tools for deciphering the biological information stored in the sugar code (Reuter & Gabius, 1999). Today, lectins are defined as proteins that possess at least one non-catalytic domain that specifically and reversibly binds a monosaccharide or oligosaccharide (Peumans & Van Damme, 1995). The lectin structures investigated so far are simply built up of β-sheet arrangements and folds such as the β-prism, β-barrel, β-sandwich and β-trefoil structures (Wright, 1997). Lectins are also categorized into families according to their sugar specificity. Monosaccharides are bound with dissociation constants in the 0.1–1.0 mM range. In order to increase the overall affinity, distinct multivalences are introduced at lectin subunits or subsites (Rini, 1995).

Mistletoe lectin I (ML-I) is a galactose-specific lectin from *Viscum album* consisting of two chains. While the lectin activity and specificity are located in the B-chain, the A-chain has rRNA N-glycosidase activity and irreversibly inhibits protein biosynthesis by cleavage at A$_{4324}$ located in the conserved GAGA loop of 28S rRNA (Endo *et al.*, 1988). By virtue of this activity, ML-I belongs to the class of ribosome-inactivating proteins (RIPs). The fusion of the toxic component with a galactose-specific lectin subunit supports efficient uptake into the target cells. In this way, the lectin subunit magnifies the RIP toxicity, as judged by comparison of IC$_{50}$ values of RIPs of type II with those of type I, which lack lectin activity (Barbieri *et al.*, 1993).

protein structure communications

ML-I and ricin are homologous RIPs and share similar structure and mechanism. However, the toxins differ in their specificity towards galactose-terminated complex oligosaccharides on the cell surface (Franz, 1986). As a result, the two RIPs occupy different binding sites on the cell surface, which partly explains the difference in toxicity, as the activity of the A-chain is approximately the same in both RIPs (Stirpe *et al.*, 1980). A crucial step in toxicity is the effectiveness of translocation through the cell membrane. The uptake of ricin occurs in the form of the inactive heterodimer (Mohanraj & Ramakrishnan, 1995), whereas the ML-I subunits are separated on the cell membrane prior to translocation (Agapov *et al.*, 1999). In spite of its toxicity, ML-I is today applied in the treatment of human cancer in doses of ~1 ng per kilogram body weight (Gabius *et al.*, 1991), which cause immunomodulatory, antimetastatic and antitumoral effects in animal studies and in humans. The lectin component of ML-I is responsible for the immunomodulation, as chemical labelling of the lectin localizes it on the surface of peripheral and intratumoral monocytes (Gabius *et al.*, 1991). Furthermore, it has been shown that MLB causes Ca^{2+} influx in Jurkat cells through interaction with surface glycoprotein receptors and is responsible for early T-cell activation (Walzel *et al.*, 2001).

The available X-ray structures of galactose-specific lectins in complex with their ligands, those of *Erythrina corallodendron* lectin (Elgavish & Shaanan, 1998) and peanut lectin (Banerjee *et al.*, 1996), have established the mode of sugar binding and emphasized the importance of recognition and binding through the 4′- and 3′-hydroxyl groups of galactose. Affinity studies against galactose and lactose and the recently published 3.0 Å low-resolution structure of ML-I in complex with galactose (Niwa *et al.*, 2003) confirmed the importance of galactose monosaccharide binding through the 4′- and 3′-hydroxyl groups (Lee *et al.*, 1994). Additional structural elements enhancing galactose specificity are achieved through van der Waals interactions of the sugar ring with aromatic side chains.
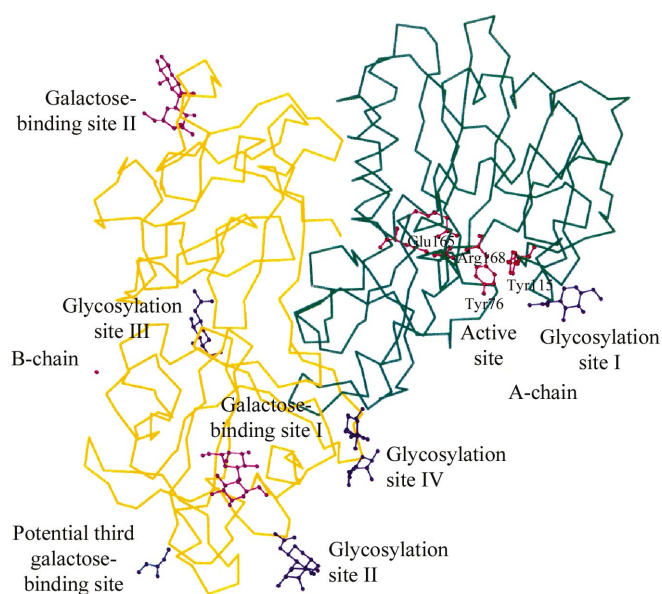
Table 1
Details of data-collection and refinement parameters.

	ML-I–lactose complex	ML-I–galactose complex
X-ray source	X13	X13
Temperature (K)	100	100
Resolution range (Å)	26–2.3	30–2.3
Wavelength (Å)	0.9073	0.9073
No. observations	701222	525723
No. unique reflections	47510	47779
R_{merge} (%)	4.9	4.8
Completeness (%)	99.6	96.3
Completeness in last resolution bin (%)	99.3	94.0
$I/\sigma(I)$ in high-resolution bin	3.72	2.57
Resolution limits (Å)	10–2.3	10–2.3
Final R factor (R_{free}) (%)	20.9 (23.6)	20.9 (24.6)
No. protein atoms	3921	3918
No. solvent molecules	328	361
No. carbohydrate atoms	140	118
Other atoms	120	108
Average B value, protein atoms (Å^2)	36.3	38.1
Average B value, solvent (Å^2)	49.8	54.2
Average B value, carbohydrate (Å^2)	59.4	59.3
Average B value, total (Å^2)	39.0	40.7
R.m.s. deviations from ideal bond lengths (Å)	0.010	0.009
R.m.s. deviations from ideal bond angles (°)	1.56	1.57
R.m.s. deviations from ideal dihedral angles (°)	25.05	25.1
R.m.s. deviations from ideal improper angles (°)	0.97	0.87

In order to gain a deeper and more detailed understanding of sugar specificity, we present and compare the modes of galactose and lactose binding to ML-I at 2.3 Å and compare the binding geometries to related RIPs of type II with higher and lower toxic potential than ML-I. Furthermore, we outline distinct structural features that are relevant for galactose binding and thereby for modulating toxicity. Finally, a potential third galactose-binding site is described and discussed relative to the architectures of galactose-binding sites I and II.

2. Material and methods

2.1. Crystallization

Mistletoe lectin I was purified for crystallization experiments following previously described procedures (Franz *et al.*, 1981; Krauspenhaar *et al.*, 1999, 2002) and concentrated to 10 mg ml^{-1} in 0.2 *M* glycine–HCl buffer pH 2.5 for crystallization setups. To obtain ML-I crystals in complex with lactose and galactose, the sugar compound was added to the above-mentioned buffer at a concentration of 0.2 *M*. Crystals of ML-I complexed with lactose and galactose were obtained by hanging-drop vapour diffusion equilibrated against a reservoir consisting of saturated ammonium sulfate of varying concentrations between 30 and 40% in 0.1 *M* glycine–HCl buffer pH 2.5 and 0.1 *M* lactose or 0.2 *M* galactose to which 40 µl dioxane had been added. Hexagonal crystals that diffracted to 2.3 Å resolution were grown after ~20 d at room temperature. Because of the relative high solvent content and rather large unit-cell parameters, attempts to improve the crystal quality to obtain higher resolution data and to lower the mosaic spread included an experiment in microgravity, as there is evidence that growth in microgravity can narrow mosaic spreads in protein and RNA crystals (Snell *et al.*, 1995; Borgstahl *et al.*, 2001) and a previous crystallization experiment of ML-I in complex with adenine monophosphate resulted in crystals that diffracted to 1.9 Å resolution. 16 crystallization trials were performed on the International Space Station (ISS) in the High Density Protein Crystal Growth (HDPCG) apparatus (previously described by Krauspenhaar *et al.*, 2002) during the mission ISS-8A for

Figure 1
C^α trace of mistletoe lectin I. The A-chain in the upper right region is shown in dark green; the active-site residues Tyr76A, Tyr115A, Glu165A and Arg168A are shown in red in ball-and-stick representation. The B-chain in the lower left part is coloured yellow. Lactose is shown in ball-and-stick representation in purple in galactose-binding sites I and II. Glycosylation sites I in chain A and sites II to IV in chain B are indicated by their oligosaccharide structure visible in the electron density and are shown in ball-and-stick representation. Glycerol bound in the potential third galactose-binding site is shown in blue.

a period of 67 d from 10 April to 17 June 2002. This lone experiment did not provide the improvements we sought and the data collected for structure determinations to 2.3 Å resolution were from the best crystals we obtained in the laboratory. The cryoprotectant solution contained 30–35% glycerol, 40% saturated ammonium sulfate, 0.1 M lactose/galactose in 0.1 M glycine–HCl buffer pH 2.5.

Figure 2
Seqence alignment of subdomains of mistletoe lectin I B-chain (MLB), ricin B-chain (RTB) and ebulin B-chain (EBB). Galactose-binding residues in subdomains α1 and γ2 are marked in red. The aromatic side chains of the galactose-binding sites are coloured green. Amino acids of the potential third galactose-binding site are shown in orange and the aromatic side chain is shown in light green. Amino acids that build up the hydrophobic centre stabilizing the β-trefoil are marked in magenta. The sequence numbering refers to MLB.
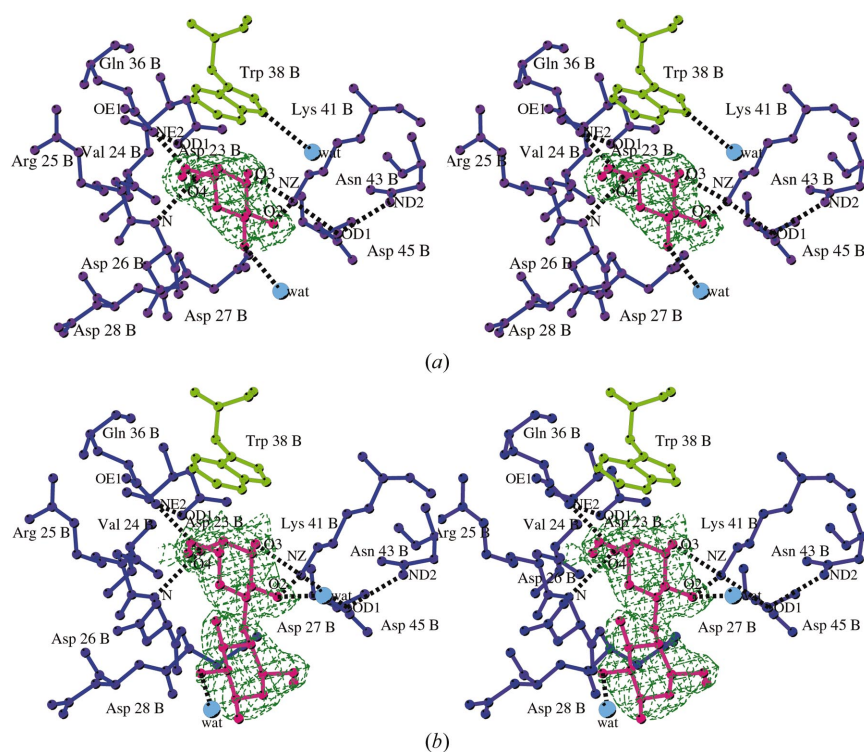

Figure 3
(a) Stereo figure showing the galactose-binding site I of the ML-I B-chain. Residues forming the sugar-binding pocket are indicated in blue in ball-and-stick mode and Trp38B providing the hydrophobic stack is shown in green. The galactose is superimposed on the electron density and is shown in red. Hydrogen bonds are shown as black dashed lines. Solvent water molecules are indicated as light blue circles. (b) As in (a) for the galactose-binding site I of ML-I in complex with lactose. All resides involved in hydrogen bonding of lactose and those indirectly stabilizing the complex are shown in blue, lactose is shown in red and the aromatic residue building up the top part is in green. Hydrogen bonds are shown as black dashed lines. Solvent waters are indicated as light blue circles.

2.2. Data collection

Intensity data from crystals of ML-I in complex with lactose and galactose were collected under cryogenic conditions at 100 K using synchrotron radiation at the consortium beamline X13 (HASYLAB/DESY), Hamburg equipped with a MAR CCD camera. The X-ray wavelength was set to 0.9073 Å. Both ML-I complexes belonged to the hexagonal space group $P6_522$ and diffracted to a resolution of 2.3 Å. The ML-I crystal in complex with lactose had unit-cell parameters $a = b = 106.28$, $c = 312.33$ Å, $\alpha = \beta = 90$, $\gamma = 120°$; in the case of the galactose-complex crystal the unit-cell parameters were almost identical, with $a = b = 106.83$, $c = 310.95$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. Owing to the long cell axis, the oscillation range per image was set to 0.5°. The data sets were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The data set of the ML-I crystal with lactose was 99.6% complete, with an R_{merge} of 4.9%; the ML-I–galactose complex showed a completeness of 96.3%, with an R_{merge} of 4.8%. The $I/\sigma(I)$ ratio in the last resolution bin was 3.7 and 2.6 for the lactose and galactose complexes, respectively. Statistics of data-collection parameters are summarized in Table 1.

2.3. Structure solution and refinement

For the initial structure calculations and refinement of the ML-I structures in complex with lactose and galactose, the native ML-I structure, previously refined to 2.3 Å (PDB code 1onk), was used as a starting model using the program *CNS* (Brünger *et al.*, 1998). Rigid-body minimization was performed first, followed by positional minimization and a gradual increase of the high-resolution limit from 3.2 to 2.3 Å. The lower limit of the data for the positional refinement was 10 Å. All the reflections measured within these ranges were included in refinement, with 5% of the data set reserved for R_{free} (Brünger, 1992). According to the fit of $2F_o - F_c$ and $F_o - F_c$ maps, the models were modified using *TURBO-FRODO* (Roussel & Cambillau, 1991) running on Silicon Graphics O2 workstations. As the electron density improved at high resolution, the ligands lactose and galactose and the sugar residues of the oligosaccharide structures in the glycosylation sites were introduced into the model. As refinement progressed, individual *B* factors were minimized and water molecules were added according to $F_o - F_c$ maps. The final refinement statistics and the overall quality parameters of the model are presented in Table 1.

3. Results

3.1. Overall fold

Mistletoe lectin I belongs to the group II carbohydrate-binding proteins, which bind their ligands in shallow preformed pockets at the protein surface (Rini, 1995). In the case of ML-I, the galactose-binding sites form a planar array distal to the toxic A-chain (Fig. 1). The toxic subunit has a globular fold divided into three domains according to the ricin nomenclature (Montfort *et al.*, 1987). Domain I
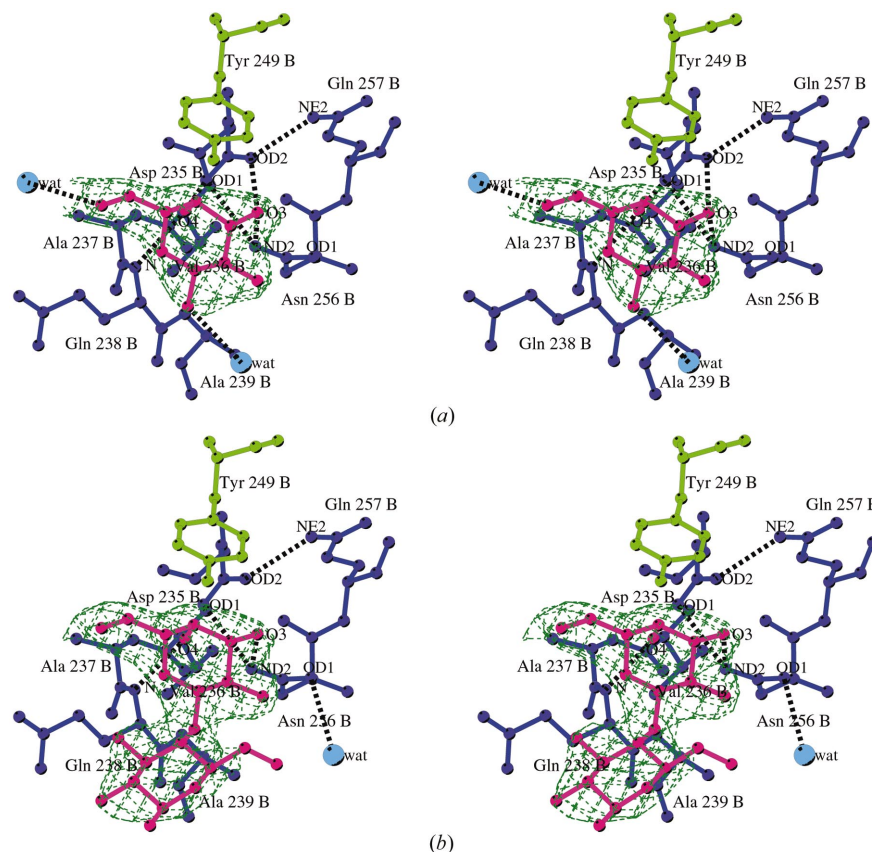

Figure 4
(*a*) Stereoview into the galactose-binding site II of the ML-I B-chain. Residues forming the sugar-binding pocket are shown in ball-and-stick representation and coloured blue; galactose is shown in red. Hydrogen bonds are indicated as black dashed lines. Water molecules are shown as blue circles. (*b*) As in (*a*) for the galactose-binding site II of ML-I in complex with lactose. Residues involved in hydrogen bonding of lactose and those indirectly stabilizing the complex are shown in blue, lactose is shown in red and the aromatic residue building the top part is in green. The electron density around lactose is coloured dark green. Hydrogen bonds are again shown as black dashed lines. Waters are indicated as blue circles.

compromises amino acids 1–109, which contain a six-stranded parallel β-sheet, an α-helix and a small antiparallel β-sheet. Domain II contains only α-helices, five of them, as secondary-structure elements, spanning amino acids 110–198. The last domain III of the A-subunit is built up of one α-helix and one antiparallel β-sheet through residues 199–249. The last five amino acids of the A-chain are flexible and were not visible in the electron density.

The B-chain appears to be a gene-duplication product from an ancient 40-residue galactose-binding unit (Robertus & Ready, 1984; Fig. 2). This chain consists of two main domains each divided in four subdomains, of which three are homologous to each other. The ML-I B-chain forms a so-called β-trefoil structure in which the β-sheet portions of the subdomains arrange themselves around a pseudo-threefold axis (Fig. 1). In the following the architectures of galactose-binding sites I and II will be described in detail.

3.2. Galactose-binding site I

The galactose-binding site I of ML-I is formed by Asp23B, a kinked loop Asp23B-Val24B-Arg25B and by Gln36B and Lys41B (Fig. 3a). Four key residues are involved in binding galactose via hydrogen bonds. In the centre of the binding pocket, the side chain of Asp23B is positioned via a hydrogen bond to Gln36B NE2. This orientation of Asp23B ensures establishment of hydrogen bonds to the 4′-hydroxyl group of galactose by Asp23B OD1 as well as to the 3′-hydroxyl group by Asp23B OD2, which form the shortest hydrogen bonds to the sugar. In addition, the 4′-hydroxyl group of the galactose moiety hydrogen bonds to of Gln36B NE2 and 26B N. In spatial proximity to the 3′-hydroxyl and 2′-hydroxyl groups of galactose is Lys41B NZ, which forms hydrogen bonds to both OH groups. The latter represents a rarely noted interaction with O2 of galactose (Drickamer, 1997). The aromatic ring of Trp38B forms the top part of the sugar-
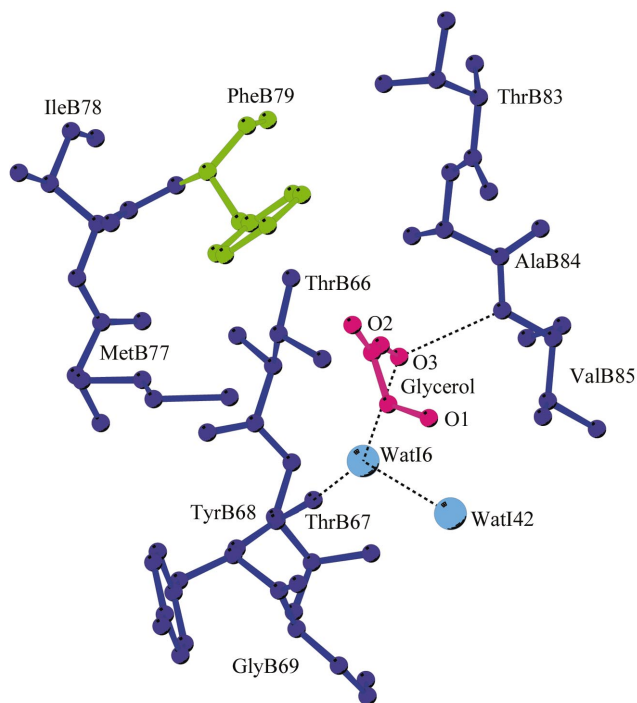
Table 2
Hydrogen bonds.

(a) Hydrogen bonds formed between galactose and lactose and the sugar-binding residues in ML-I for binding sites I and II.

	Protein and solvent atoms	Carbohydrate atoms	ML-I + galactose (Å)	ML-I + lactose (Å)
Galactose-binding site I	Asp23B OD1	O4	2.6	2.6
	Asp23B OD2	O3	2.6	2.6
	Asp26B N	O4	2.9	3.0
	Gln36B NE2	O4	3.1	3.1
	Lys41B NZ	O3	3.0	2.9
	Lys41B NZ	O2	3.0	2.7
	Wat273I OW	O1	3.1	
	Wat135I OW	O2		3.1
	Wat176I OW	O2		3.2
Galactose-binding site II	Asp235B OD1	O4	2.6	2.7
	Asp235B OD2	O3	2.7	2.7
	Gln238B N	O4	3.0	2.9
	Asn256B ND2	O3	3.1	3.1
	Wat157I OW	O6	3.4	
	Wat354I OW	O1	3.0	

(b) Hydrogen bonds formed between the galactose-binding residues in the galactose-binding site I and II in ricin and to lactose in ebulin.

	Protein atoms		Carbohydrate atoms		Ricin + galactose (Å)	Ebulin + lactose (Å)
	Ricin	Ebulin	Ricin	Ebulin		
Galactose-binding site I	Asp22B OD1	Asp24B OD1	O4	O4	2.9	3.1
	Asp22B OD2	Asp24B OD1	O3	O4	3.1	3.1
	Asp25B N	Asn27B N	O6	O4	3.1	2.7
	Gln35B NE2	Gln37B NE2	O6	O6	3.1	2.9
	Gln35B NE2		O4		3.2	
	Asn46B ND2	Asn46B ND2	O3	O3	2.9	2.7
Galactose-binding site II	Asp234B OD1		O4		2.9	
	Asp234B OD2		O3		2.9	
	Ala237B N	Glu238B N	O4	O4	2.8	2.7
	His251B NE2		O3		3.0	
		Asn254B ND2		O1		3.0
	Asn255B ND2	Asn256B ND2	O3	O3	2.8	2.8
		Asn256B ND2		O4		3.1
		Gln257B NE2		O2		3.1

binding pocket and provides a stacking interaction with the planar galactose ring. The orientation of the tryptophan side chain is further stabilized by a hydrogen bond to a solvent molecule. The ring planes form an angle of 19.4° in the case of the ML-I–galactose complex.

In the case of the ML-I–lactose complex, the galactose moiety of lactose is bound in the galactose-binding site I in the same manner as described previously for the galactose complex (Fig. 3b). However, the angle between the plane of the aromatic ring of Trp38B and the sugar ring of lactose is 33.4°. The galactosyl part of lactose is the primary partner of interaction with the lectin; the glucosyl part remains exposed to the solvent, forming hydrogen bonds with the solvent molecules.

3.3. Galactose-binding site II

The overall architecture of galactose-binding site II (Fig. 4a) is almost identical to galactose-binding site I, although not all amino acids involved in galactose binding are conserved. The centre of the binding pocket is formed by Asp235B, which is positioned via a hydrogen bond from OD2 to of Gln257B NE2. The kinked loop consisting of the tripeptide Asp235B-Val236B-Ala237B builds up the bottom of the sugar-binding site II. The side chain of Asp235B forms a hydrogen bond from its OD1 to the 4′-hydroxyl group of galactose and OD2 to the 3′-hydroxyl group of the sugar. Further stabilization of the sugar–lectin complex is achieved by a hydrogen bond between

Figure 5
Potential third galactose-binding site in ML-I located in subdomain β1. The glycerol molecule is marked in red, the aromatic amino acid Tyr79B in green, remaining residues of the pocket in blue and the water molecules in light blue. Hydrogen bonds are shown accordingly.

the 4′-hydroxyl group of galactose and Gln238B N and the 3′-hydroxyl group and Asn256B ND2. The top is covered by Tyr249B, which provides the stacking partner for the sugar ring. In this case, the planar galactose ring is tilted towards the aromatic ring of Tyr249B and forms an angle of 29.2° for the ML-I–galactose complex and an almost identical angle of 29.6° for the ML-I–lactose complex (Fig. 4b). In contrast to galactose-binding site I of ML-I, in site II no interactions are noted between the 2′-hydroxyl group of the sugar moiety and the lectin.

3.4. The potential third galactose-binding site

The B-chain of type II RIP proteins is a gene-duplication product of an ancient galactose-binding subunit (Robertus & Ready, 1984), but only two of the four subdomains retain galactose-binding capacity. The sequence alignment of galactose-binding residues and the architectures of subdomains α1 and γ2 are mimicked by a similar fold in subdomain β1. However, the polar side chains required for galactose binding are changed: the central Asp is replaced by Thr66B, Met77B is in equivalent position to Gln36B of subdomain α1 and Val85B corresponds to Asn256B in γ2 (Fig. 5). The replacement with more hydrophobic residues prevents sugar binding, although the aromatic top part of the binding pocket is conserved in that Phe79B is in an equivalent position to Trp38B in subdomain α1 and Tyr249B in γ2. Interestingly, a glycerol molecule is bound in subdomain β1, mimicking the sugar compound. Glycerol is bound to the lectin via hydrogen bonds between O3 and Val85B N and Thr67B O, the latter through a bridging water molecule. Met77B, on the other hand, is not capable of participation in hydrogen bonding to the glycerol and O1 and O2 of glycerol are oriented away from this residue toward the protein surface. Obviously, this pocket lacks the ability to bind galactose as it has nonpolar residues at key positions, especially 66B (Asp to Thr) and 77B (Gln to Met), but the ability to bind the hydroxyl group effectively remains.

4. Discussion

ML-I has two preformed galactose-binding pockets in subdomains α1 and γ2 of the B-chain; other subdomains have either lost the ability to bind galactose owing to mutations or deletion of sugar-binding residues or never developed the ability in the first place. A detailed analysis of the architectures of galactose-binding sites I and II shows a strikingly similar method of sugar binding. The basic features for galactose-specific binding are present in both galactose-binding sites of ML-I: a central Asp, further hydrogen-bonding donors in proximity to the 4′- and 3′-hydroxyl group of galactose and an aromatic side chain as hydrophobic stacking partner against the sugar. Not all sugar-binding residues are conserved in subdomains α1 and γ2 and the hydrogen-bonding network in binding site I is more complex than in site II, so that ML-I is an example of a lectin in which sugar-binding sites differ within one single chain.

Focusing on the central Asp, previous galactose-binding lectin structures such as those of *E. corallodendron* lectin (Shaanan et al., 1991), peanut lectin (Banerjee et al., 1996) and *Griffonia simplicifolia* (Debaere et al., 1993) reveal that this charged amino acid is conserved
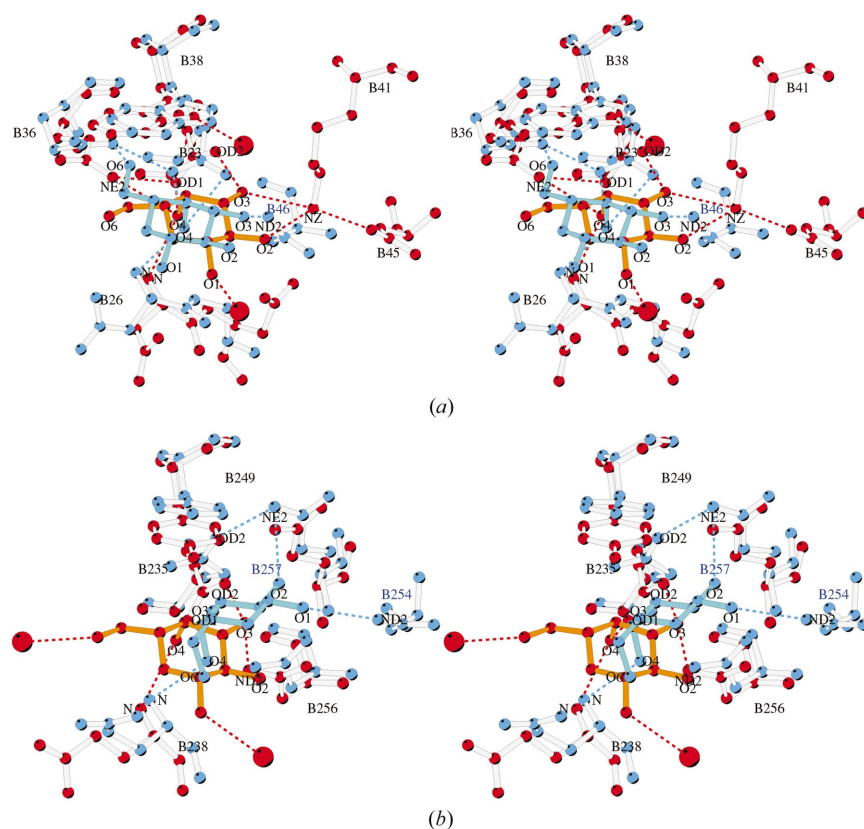

Figure 6
(a) Stereo representation of the superposition of the galactose-binding site I of ML-I in red and ebulin in blue as a ball-and-stick model. The galactose molecule is coloured red–orange in ML-I and blue in ebulin. Accordingly, hydrogen bonds are shown as red dashed lines for the ML-I and in blue for the ebulin complex. Water molecules in the ML-I complex are marked as red circles. (b) Superposition of the galactose-binding site II of ML-I in red and of ebulin in blue; stereoview in ball-and-stick mode. The galactose molecule is coloured red–orange in ML-I and in blue in ebulin. Accordingly, the hydrogen bonds are shown as red dashed lines for the ML-I and in blue for the ebulin complex. Water molecules in the ML-I complex are indicated as red circles.

and establishes the polar environment that is required for sugar binding (Rutenber & Robertus, 1991). Its planar carbohydrate side chain achieves excellent hydrogen-bond geometry to the vicinal sugar OHs (Vyas, 1991). The role and conformation of the Asp side chain is underlined when superimposing the galactose-binding sites I and II of ML-I with the site in ebulin (Figs. 6a and 6b), a type II RIP with low toxic potential (Pascal *et al.*, 2001). The superposition highlights a different conformation for the central Asp (Asp23B in ML-I and Asp24B in ebulin). There is a minor change in orientation of galactose in ebulin compared with ML-I, but the hydrogen bonding required for galactose binding in ebulin is conserved compared with ML-I. There is a minor difference: O3 of galactose forms a hydrogen bond with Asn46B ND2 in ebulin, whereas Lys41B provides this interaction in ML-I. In addition, a bond between O6 and Gln37B NE2, which further stabilizes the ebulin–galactose complex, is not detected in ML-I. Although nearly all amino acids in galactose-binding site II are conserved in ML-I and ebulin, the exception being that Tyr249B is changed to Phe at the top of the binding pocket, the mode of sugar binding is entirely different. The superposition of site II in ML-I and ebulin underlines the role of the Asp side chain (Fig. 6b); it is obvious that the side chain of Asp235B takes a different conformation in ebulin than in ML-I and is no longer the primary partner of interaction with galactose. This altered mode of galactose binding by ebulin correlates with its function as a type II RIP of low toxicity (Pascal *et al.*, 2001).

As lectin activity is a prerequisite for toxicity, we now compare the architecture of sugar-binding sites of ML-I and the related ricin, for which the isolated A-chains have approximately the same toxicity

(Stirpe *et al.*, 1980). Ricin shows a similar mode of galactose binding and key features are conserved with respect to ML-I (Table 2). However, owing to a slightly different orientation of the galactosyl part of lactose in the binding pocket, we can detect some deviations in the hydrogen-bonding pattern even though the amino acids within the pocket are conserved (Fig. 7a). O3 and O2 of the galactosyl moiety form hydrogen bonds to Lys41B in the case of ML-I, whereas only Asn46B ND2 binds to O3 in the ricin complex. Furthermore, O6 of the galactosyl part in the ricin complex is oriented towards Gln35B NE2 and Asp25B N and makes hydrogen bonds with these residues. In contrast, O6 of the galactosyl part of lactose in ML-I is oriented toward the protein surface without contributing to hydrogen bonds in the sugar–lectin complex. The plane through the galactosyl ring of lactose forms an angle of 33° with the planar aromatic ring of Trp38B in the case of ML-I and 45° for the ricin–lactose complex. Despite uncertainties inherent in these structure determinations, it seems unlikely that angular discrepancies this large are not significant. Still, based on conservation of amino-acid sequences we would not expect a different hydrogen-bonding pattern to bind galactose in binding site I in ML-I compared with ricin. Focusing on galactose-binding site II in ML-I, the galactose-binding residues 235B, 238B and 256B are equivalent to Asp234B, Gln237B and Asn255B in ricin and the hydrogen-bonding pattern to the galactose moiety of lactose is conserved (Fig. 7b). Orientation of the galactosyl part of lactose in both pockets and the angle formed between the planar ring of Tyr249B (248B in ricin) and the planarity through the sugar ring are nearly identical: 29° for ML-I and 26° for ricin. In contrast, in the ricin–lactose complex an additional hydrogen bond from the
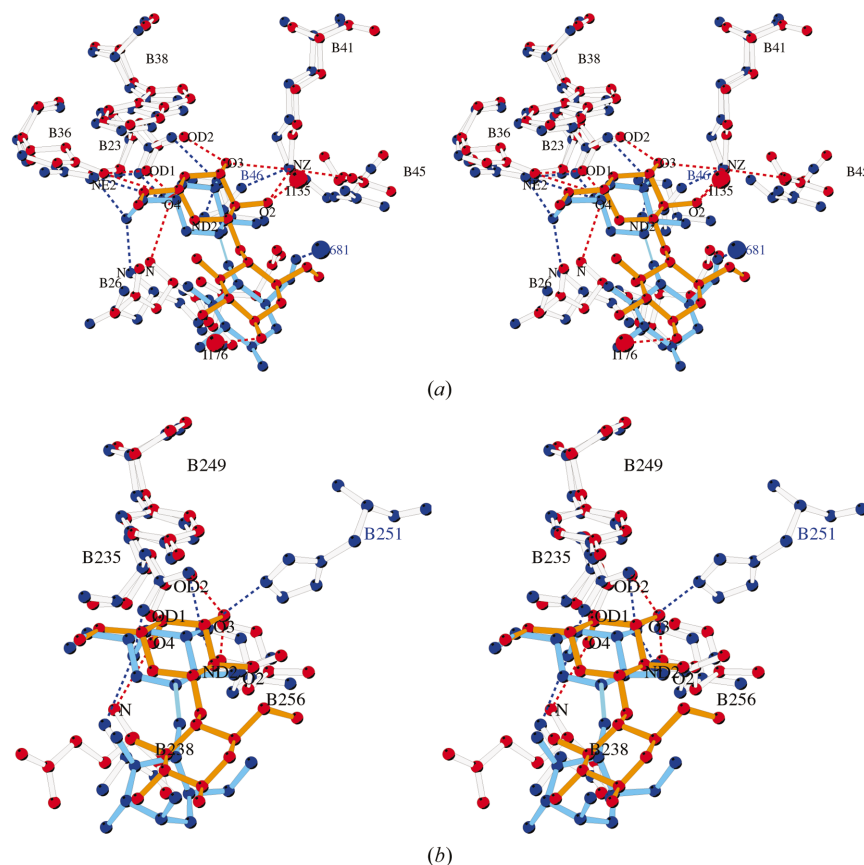

Figure 7
(a) Stereo representation of the superposition of the galactose-binding site I of ML-I in red and of ricin in blue as a ball-and-stick model. The bound lactose molecule is coloured red–orange in ML-I and blue in ricin. Hydrogen bonds are shown as red dashed lines for ML-I and in blue for ricin. Water molecules are marked as red circles in the ML-I complex and in blue in the ricin complex. (b) Superposition of galactose-binding site II of ML-I and ricin. The colour code is similar to (a).

$3'$-hydroxy group to of His251B NE2 stabilizes the sugar–lectin complex, whereas an equivalent interaction in ML-I cannot be performed by Thr252B. A further difference was shown in the tripeptide residues Asp234B-Val235B-Arg236B at the bottom of the ricin-binding pocket, which has the amino-acid sequence Asp-Val-Ala in ML-I. The impact of the loss of key hydrogen-bonding partners on the capacity for galactose binding of ricin was shown by site-directed mutagenesis. Galactose-binding site II was identified to be the high-affinity sugar-binding site in the case of ricin (Zentz *et al.*, 1978). A nearly complete loss of lectin activity and thereby cyto-toxicity was observed when Asn255B (Asn256B in ML-I) was mutated to alanine in galactose-binding site II (Vitetta & Yen, 1990). Interestingly, the angles between the sugar rings and the aromatic side chains are nearly identical in galactose-binding site II in ML-I and ricin. The average angle found between the planar sugar ring and the aromatic side chain in galactose-specific lectins investigated so far is $32°$ (Kolatkar & Weiss, 1996).

Affinity studies of ML-I and ricin show difference in affinity for N-acetyl-D-galactosamine and D-galactose (Lee *et al.*, 1994). The IC_{50} of ML-I against N-acetyl-D-galactosamine is 120 mM compared with 1.7 mM in the case of ricin, whereas the affinity towards D-galactose is

nearly identical: 0.9 mM for ML-I and 1.0 mM for ricin. Based on the structural determinants required for galactose binding in site I of ML-I, the N-acetyl group has a steric conflict with the side chain of Asp27B. For the ricin–lactose complex the N-acetyl group clashes with the side chain of Asp44B (Rutenber & Robertus, 1991). It can be concluded that the orientation of the sugar ring in the pocket indicates a strong preference for binding galactose in ML-I site I. In contrast, the $2'$-hydroxyl group in galactose-binding site II is directed towards the protein surface and N-acetylgalactosamine can also be bound. There is also structural evidence that at the bottom of site II in ricin the polar amino acid Ser238B can provide an additional polar interaction to O7 of the N-acetyl substituent (Fig. 8a). In contrast to ricin, ML-I has an alanine in this position, which is not capable of making a contribution in order to stabilize the complex (Fig. 8b). Of course, there may still be unrecognized determinants involved in binding N-acetyl-D-galactosamine that are needed to explain fully the significant difference in affinities of ML-I and ricin.

The affinity of ML-I towards galactose and lactose is not only conditioned by the architecture of the binding sites, because recombinant MLB possesses reduced affinity against monosaccharides (Eck *et al.*, 1999). Thus, it becomes clear that not only the mode of
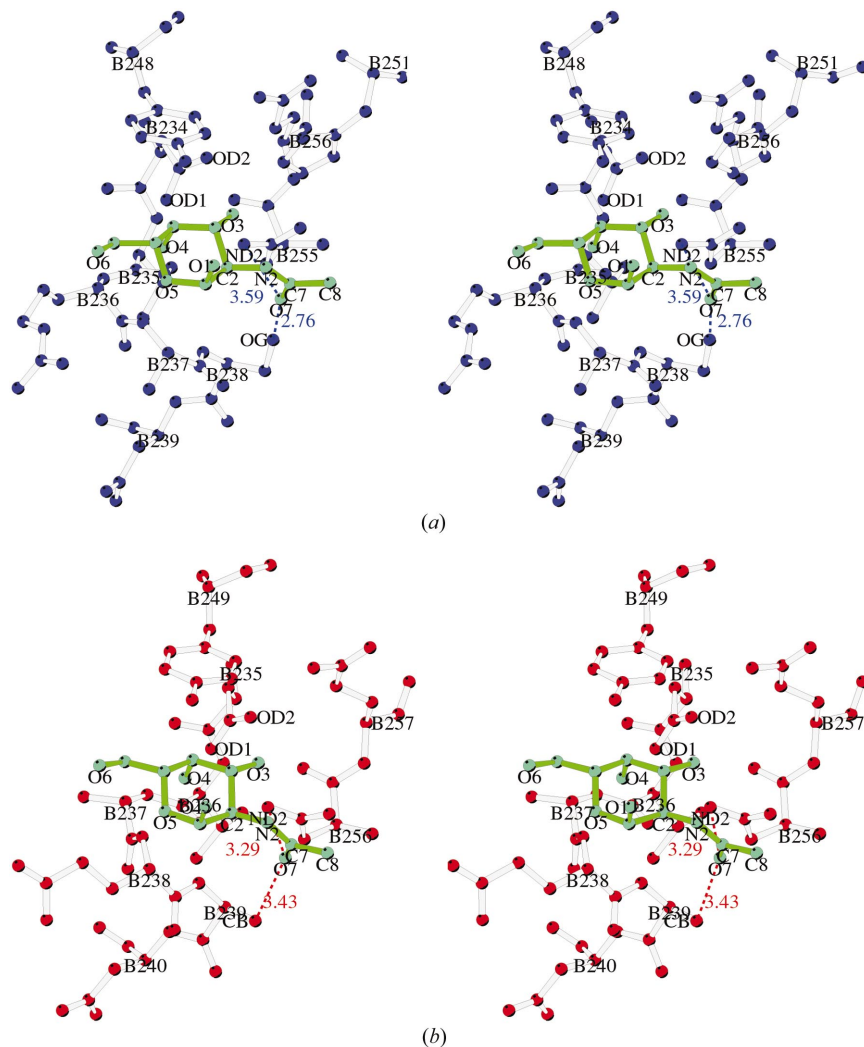

(a)

(b)

Figure 8
(a) Galactose-binding site II of ricin as a ball-and-stick representation in blue. The N-acetylgalactosamine, shown as a turquoise ball-and-stick representation, was modelled into the binding pocket according to the ricin–lactose complex. Potential hydrogen bonds of the substituent on C2 are marked in blue dashed lines and distances are given in Å. (b) The galactose-binding site II of ML-I is shown in red as a ball-and-stick model. The N-acetylgalactosamine was modelled into the binding pocket according to the ML-I–galactose complex and shown as a blue ball-and-stick model. Potential hydrogen bonds and distances is are indicated as in (a).

galactose binding but also the type of glycosylation has an influence on affinity and consequently on the toxic potential, as oligosaccharide side chains are known to be essential for stability and recognition. This phenomenon was also shown for the recombinant ricin B-chain (Wales *et al.*, 1991).

The toxicity of RIPs is based on the efficiency of uptake into target cells. Differences in this efficiency confer differences between ML-I and ricin in specificity against complex galactose-terminated oligosaccharide structures on cell membranes and in specificity against the type and quantity of cell receptors on target cells. As discussed and shown before, the specificity towards monosacharides is nearly identical. While the detailed comparison of galactose binding of ML-I and other sugar-complexed RIP structures with high and low toxic potential underline the overall impact of lectin activity on toxicity, RIP toxicity remains an activity influenced by many variables.

Based on the mode of galactose binding of ML-I in subdomains α1 and γ2, essential and conserved residues responsible for galactose binding as well as variable residues are outlined. These observations show the evidence that site-directed mutagenesis experiments replacing Thr66B by Asp and Met77B by Gln would most probably succeed in restoring the ability to bind galactose in subdomain β1. A third galactose-binding site would be expected to increase the sugar-binding capacity. This aspect is of importance in terms of the design studies of the antitumoral drug rViscumin, which is currently under investigation in phase I clinical studies and for which phase II studies are being considered. Further attempts may also succeed to engineer novel sugar specificities into ML-I that will bring MLB in focus as a potential transport molecule for pharmaceuticals and for efficient uptake in targeted cell types.

References

Agapov, I. I., Tonevitsky, A. G., Moysenovich, M. M., Maluchenko, N. V., Weyhenmeyer, R. & Kirpichnikov, M. P. (1999). *FEBS Lett.* **452**, 211–214.

Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1996). *J. Mol. Biol.* **259**, 281–296.

Barbieri, L., Battelli, M. G. & Stirpe, F. (1993). *Biochem. Biophys. Acta*, **1154**, 237–282.

Borgstahl, G. E. O., Vahedi-Fardi, A., Lovelace, J., Bellamy, H. & Snell, E. H. (2001). *Acta Cryst.* D**57**, 1204–1207.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Gross-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Debaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Wilson, K. S. & Dauter, Z. (1993). *J. Mol. Biol.* **230**, 950–965.

Drickamer, K. (1997). *Structure*, **5**, 465–468.

Eck, J., Langer, M., Möckel, B., Witthohn, K., Yinke, H. & Lentzen, H. (1999). *Eur. J. Biochem.* **265**, 788–797.

Elgavish, S. & Shaanan, B. (1998). *J. Mol. Biol.* **277**, 917–932.

Endo, Y., Tsurugi, K. & Franz, H. (1988). *FEBS Lett.* **231**, 378–380.

Franz, H. (1986). *Oncology*, **43**, 23–34.

Franz, H., Ziska, P. & Kindt, A. (1981). *Biochem. J.* **195**, 481–484.

Gabius, S., Kayser, K. & Gabius, H.-J. (1991). *Dtsch Zschr. Onkol.* **23**, 113–119.

Kolatkar, A. & Weiss, W. I. (1996). *J. Biol. Chem.* **271**, 6679–6685.

Krauspenhaar, R., Eschenburg, S., Perbandt, M., Kornilov, N., Konareva, N., Mikhailova, I., Stoeva, St., 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., Voelter, W. & Betzel, C. (2002). *Acta Cryst.* D**58**, 1704–1707.

Lee, R. T., Gabius, H. J. & Lee, Y. C. (1994). *Carbohydr. Res.* **254**, 269–276.

Mohanraj, D. & Ramakrishnan, S. (1995). *Biochim. Biophys. Acta*, **1243**, 399–406.

Montfort, W., Villafranca, J. E., Monzingo, A. F., Ernst, S. R., Katzin, B., Rutenber, E., Xoung, N. H., Hamlin, R. & Robertus, J. D. (1987). *J. Biol. Chem.* **262**, 5398–5403.

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

Otwinowski, Z. & Minor, W. (1997). *Methods Enyzmol.* **276**, 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 Struct. Funct. Genet.* **43**, 319–326.

Peumans, W. J. & Van Damme, E. J. M. (1995). *Plant Physiol.* **109**, 347–352.

Reuter, G. & Gabius, H.-J. (1999). *Cell. Mol. Life Sci.* **55**, 368–422.

Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577.

Robertus, J. D. & Ready, M. P. (1984). *J. Biol. Chem.* **259**, 13953–13956.

Roussel, A. & Cambillau, C. (1991). *Silicon Graphics Geometry Partners Directory*, p. 86. Mountain View, CA, USA: Silicon Graphics.

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

Shaanan, B., Lis, H. & Sharon, N. (1991). *Science*, **254**, 862–866.

Snell, E. H., Weisberger, S., Helliwell, J. R., Weckert, E., Hölzer, K. & Schroer, K. (1995). *Acta Cryst.* D**51**, 1099–1102.

Stirpe, F., Legg, R., Onyon, L., Ziska, P. & Franz, H. (1980). *Biochem. J.* **190**, 843–845.

Vitetta, E. S. & Yen, N. (1990). *Biochem. Biophys. Acta*, **1049**, 151–157.

Vyas, N. K. (1991). *Curr. Opin. Struct. Biol.* **1**, 732–740.

Wales, R., Richardson, P. T., Roberts, L. M., Woodland, H. R. & Lord, J. M. (1991). *J. Biol. Chem.* **266**, 19172–19179.

Walzel, H., Blach, M., Neels, P., Schulz, U., Wollenhaupt, K. & Brock, J. (2001). *Immunol. Lett.* **78**, 57–66.

Wright, C. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 631–636.

Zentz, C., Frenoy, J.-P. & Bourrillon, R. (1978). *Biochem. Biophys. Acta*, **536**, 18–26.