
Structural Basis for the Carbohydrate Specificities of Artocarpin: Variation in the Length of a Loop as a Strategy for Generating Ligand Specificity

A. Arockia Jeyaprakash, Anand Srivastav, A. Surolia and M. Vijayan*

*Molecular Biophysics Unit
Indian Institute of Science
Bangalore 560 012, India*

Artocarpin, a tetrameric lectin of molecular mass 65 kDa, is one of the two lectins extracted from the seeds of jackfruit. The structures of the complexes of artocarpin with mannotriose and mannopentose reported here, together with the structures of artocarpin and its complex with Me- α -mannose reported earlier, show that the lectin possesses a deep-seated binding site formed by three loops. The binding site can be considered as composed of two subsites; the primary site and the secondary site. Interactions at the primary site composed of two of the loops involve mainly hydrogen bonds, while those at the secondary site comprising the third loop are primarily van der Waals in nature. Mannotriose in its complex with the lectin interacts through all the three mannopyranosyl residues; mannopentose interacts with the protein using at least three of the five mannose residues. The complexes provide a structural explanation for the carbohydrate specificities of artocarpin. A detailed comparison with the sugar complexes of *heltuba*, the only other mannose-specific jacalin-like lectin with known three-dimensional structure in sugar-bound form, establishes the role of the sugar-binding loop constituting the secondary site, in conferring different specificities at the oligosaccharide level. This loop is four residues longer in artocarpin than in *heltuba*, providing an instance where variation in loop length is used as a strategy for generating carbohydrate specificity.

Keywords: moraceae lectin; carbohydrate specificity; crystal packing; weak interactions; carbohydrate conformation

*Corresponding author

Introduction

Lectins are multivalent carbohydrate-binding proteins of non-immune origin. They have received considerable attention on account of their ability to recognize diverse sugar structures with a high degree of specificity.^{1–5} They are involved in various biological processes such as cell–cell recognition, host–pathogen interactions, targeting of cells, cancer metastasis and differentiation. Lectins exist in almost all forms of life but those from plants are the most thoroughly characterized among the members of the family.^{6–8} Plant lectins account for about half the lectins of known three-dimensional structure†. The structures of plant

lectins encompass five different folds. One of them, the β -prism I fold, was first identified in this laboratory as a lectin fold in jacalin, one of the two lectins from jackfruit (*Artocarpus integrifolia*) seeds.⁹ Each subunit of the tetrameric lectin, 66,000 Da, consists of two chains generated by a post-translational modification involving proteolysis. The heavy chain α and the light chain β are composed of 133 and 20 amino acid residues, respectively.^{10–12} Jacalin is galactose-specific and the structure of its complex with Me- α -galactose demonstrated the role of post-translational modification in generating sugar specificity. Subsequent extensive X-ray and modelling studies carried out on various jacalin–sugar complexes revealed the combining site to be composed of a primary site and two secondary sites.^{13,14} These studies provided a structural rationale for the binding affinities of the lectin to different sugars and insights into its interactions with glycoproteins.

Abbreviations used: Gn₂, GlcNAc₂; Mn₃, mannotriose.
E-mail address of the corresponding author:
mv@mbu.iisc.ernet.in

Artocarpin (65,000 Da), the second lectin from jackfruit seeds is, unlike jacalin, non-glycosylated and specific to mannose at the monosaccharide level.^{15–19} It, along with calsepa (*Calystegia sepium* agglutinin) and heltuba (*Helianthus tuberosus* agglutinin), form a sub-group of jacalin-related mannose-binding lectins.^{20–23} Extensive thermodynamic studies have shown that mannopentose and mannotriose binds with much better affinity to artocarpin than di- and monosaccharides do. The order of the binding affinity of artocarpin is as follows: mannopentose \approx mannotriose $>$ Man α 1–3Man $>$ GlcNAc₂Man₃ $>$ Me- α -Man $>$ Man $>$ Man α 1–6Man $>$ Man α 1–2Man $>$ Me- α -Glc $>$ Glc. It has the highest affinity for the heptasaccharide from horseradish peroxidase.^{17–19} Although artocarpin differs from jacalin in its carbohydrate specificities and its biological action, their overall structure is the same.^{22,23} The crystal structures of artocarpin in complex with mannotriose and mannopentose reported here, together with the crystal structure of its complex with Me- α -Man reported by us earlier, reveal the involvement of three loops in carbohydrate recognition. The structures, together with simple modelling, provide an understanding of the structural basis for the carbohydrate specificities of artocarpin. A comparison of these structures with those involving heltuba brings to light the importance of the length of a loop in determining ligand specificity.

Results and Discussion

General features

Unlike jacalin, artocarpin is a single-chain protein and each of its subunits is 149 amino acid residues long. The molecule exists as a tetramer possessing 222 symmetry. Each subunit has a 3-fold symmetric β -prism fold formed by three four-stranded β -sheets (Figure 1). The β -sheets have greek-key topology. One of them is made up of a non-contiguous set of residues. We have reported the structure of the complex of artocarpin with Me- α -Man.²³ The structures of artocarpin bound to mannotriose and mannopentose reported here reveal that the overall structure remains essentially the same in all the complexes. The r.m.s. deviations in C $^{\alpha}$ positions when pairs of subunits from the same complex are superposed do not exceed 0.41 Å. The subunits in one structure superpose on those in the other with a maximum r.m.s. deviation of 0.51 Å.

Artocarpin–sugar interactions

The location of the carbohydrate-binding site of artocarpin in its subunits is the same as in that of jacalin. The binding-site cleft is formed by the three loops 14–17, 86–95 and 137–141. Among these, loops 14–17 and 137–141 connect the inner

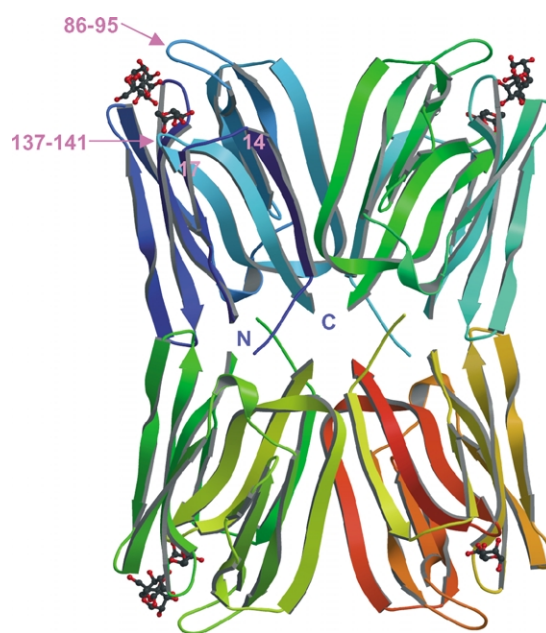


Figure 1. Structure of tetrameric artocarpin with bound mannotriose. The N and C termini, and the loops involved in sugar binding are indicated in one subunit.

strands of greek key I, while loop 86–95 connects those of greek key III. The crystal structure of artocarpin in complex with Me- α -Man revealed the involvement of only two loops (14–17 and 137–141) in monosaccharide recognition. The structures of the complexes of artocarpin with mannotriose (I) and mannopentose (II) show the involvement of loop 86–95, which is the longest loop in the molecule. The bound mannotriose has unambiguous electron density in six out of eight subunits (Figure 2(a)). Only one mannose residue is defined in the remaining subunits because of steric constraints in the crystal packing, as described in detail in Materials and Methods. In the mannopentose complex, a mannotriose stretch has clear density in one subunit, while four of the five sugar residues are defined in the other subunit (Figure 2(b)). The sugars modelled in the respective electron densities allow a thorough characterization of lectin–sugar interactions. On the basis of these interactions, the sugar-binding region can be considered as composed of a primary site and a secondary site. The primary site, which is formed by the amino acid residues in loops 14–17 and 137–141, involves numerous direct hydrogen bonding interactions with sugar hydroxyl groups. The amino acid residues in loop 86–95 form the secondary site. They are involved primarily in van der Waals interactions with the sugar molecule, although a few hydrogen bonds are present. The interactions at the primary and the secondary binding sites are described below.

Interactions at the primary site

The interactions of the bound Man residue at the primary site are the same in all the subunits

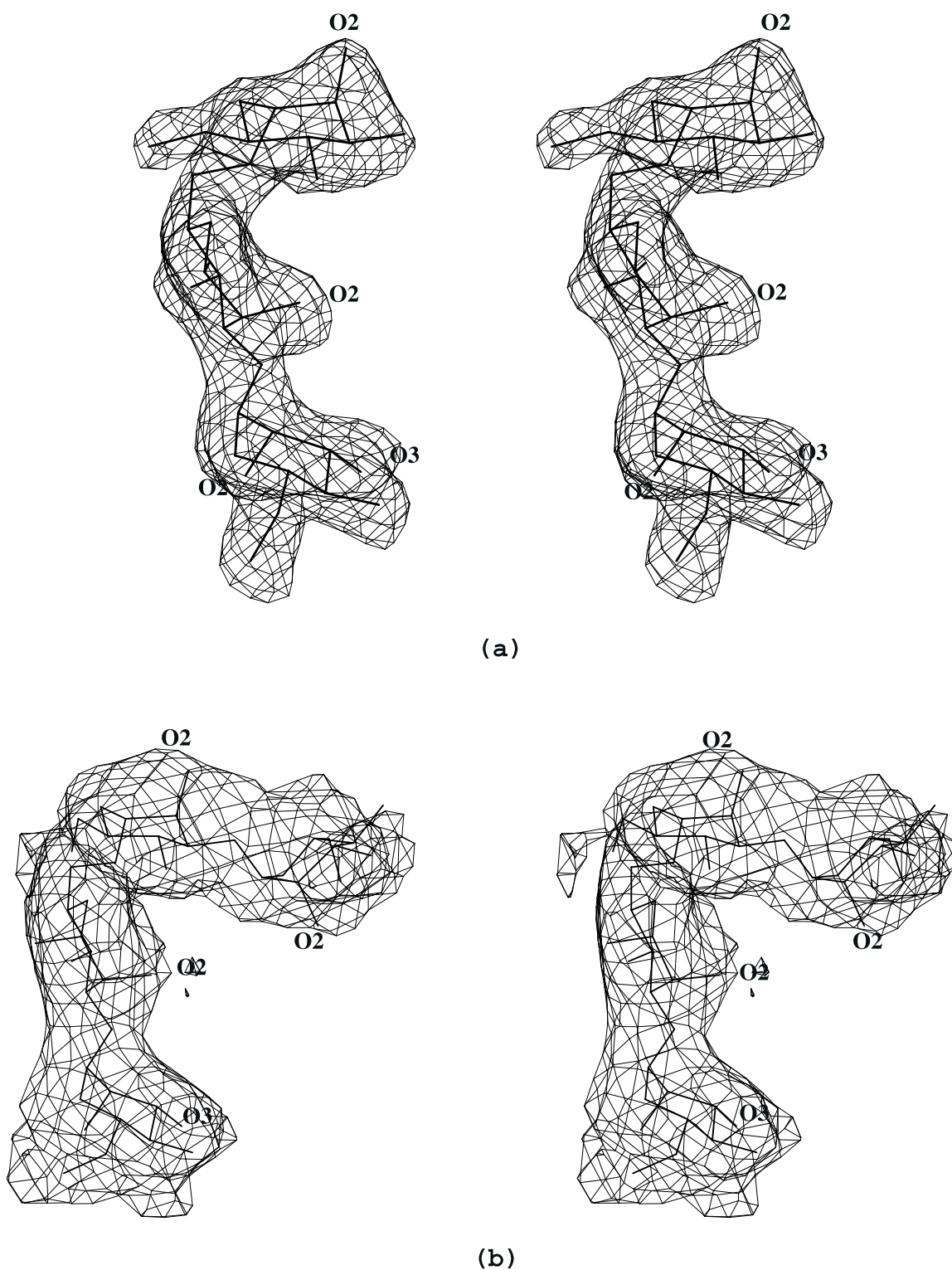


Figure 2. Stereo views of the electron densities in the $2F_o - F_c$ map contoured at 1σ for (a) mannotriose and (b) mannopentose (four of the five mannopyranosyl residues).

(Table 1). The amino acid residues involved in hydrogen bonding interactions at this site are Gly15, Gly137, Asp138 and Asp141. As observed in the Me- α -Man complex, Gly15 N hydrogen bonds to O3, Asp138 N to O5 and O6, Leu139 N to O6 and Asp141 OD1 to O4 and O6 of Man

(Figure 3). However, the distances between Gly15 N and O4 and Asp141 OD2 and O6 are now somewhat greater than 3.6 \AA , at 3.8 \AA and 3.7 \AA , respectively; the corresponding values in the Me- α -Man complex are 3.1 \AA and 3.5 \AA . The C^α superposition of the Me- α -Man complex on the

Table 1. Artocarpin–carbohydrate interactions (lengths in Å)

Sugar atom	Protein atom	Mannotriose						Mannopentose	
		A	B	C	D	E	F	A	B
Man1									
O3	Gly15 N	2.9	2.9	2.7	3.3	2.9	3.2	3.0	2.8
O4	Asp141 OD1	2.6	2.6	2.5	2.8	2.7	2.8	2.5	2.5
O5	Asp138 N	3.1	3.1	3.2	3.2	3.2	3.2	2.9	3.0
O6	Asp138 N	2.9	2.9	3.1	2.8	2.6	2.5	2.5	2.7
	Leu139 N	3.0	3.0	3.0	3.0	2.9	2.9	2.6	2.7
	Asp141 OD1	2.6	2.6	2.8	2.7	2.9	2.9	2.9	2.9
Man2									
O1	Ala90 N	2.6	2.8	3.0	–	–	–	–	–
	Asp138 OD1	–	–	–	–	3.2	–	3.1	3.5
O2	Ala90 N	3.0	2.9	2.9	–	3.2	–	3.1	3.0
	Thr91 N	3.0	3.3	2.8	–	3.1	–	3.2	2.9
	Thr91 OG1	2.8	2.7	3.0	–	2.8	–	2.5	2.8

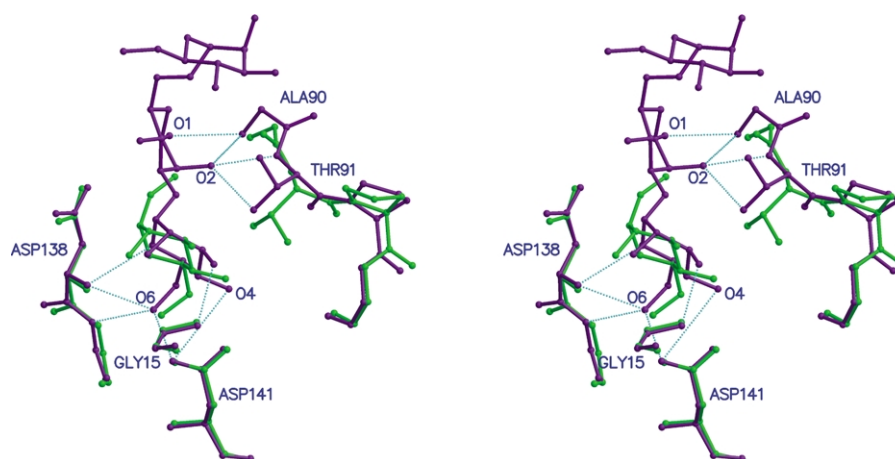
mannotriose complex revealed a small difference in the orientation of the bound Me- α -Man (Figure 3) between the two complexes. This reorientation involves a rotation of 21° with no translation. This change in the orientation could be related to the orientation of the methyl group in the Me- α -Man complex, which is not present in the trisaccharide. Notably, the binding of Me- α -Man involves an additional burial of 39 Å² of non-polar area compared to that of Man.

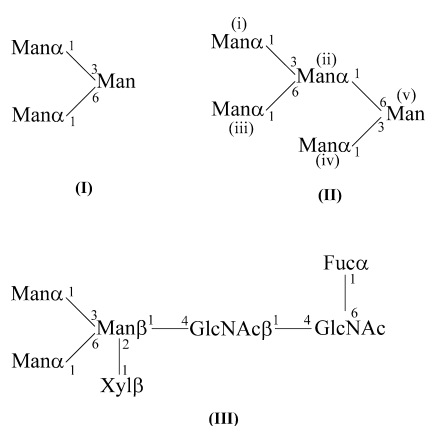
Interactions at the secondary site

All except one residue involved in carbohydrate binding at this site come from loop 86–95. The exception is Asp138 from loop 137–141. Among these, only Ala90, Thr91 and Asp138 form hydrogen bonds with the sugar molecule. All of them interact with O2 of the reducing Man (Table 1). O1 of this residue is axial in subunits A, B and C, and equatorial in subunits E, G and H. O1 form a hydrogen bond with Ala90 N when it is axial and with Asp138 OD1 when it is equatorial. The atoms in the reducing Man also make several favourable van der Waals contacts with protein atoms. The

mannose residue at the non-reducing end α 1–6 linked to the reducing Man has only van der Waals interactions with the lectin. In particular, the ring atoms C1, C2 and C3 in it are at distances less than 4.5 Å from the side-chain methyl groups of Ala90 and Thr91. Compared to the binding of Me- α -Man, the mannotriose binding involves the burial of nearly 180 Å² (82 Å² of non-polar) more surface area of which 86 Å² (69 Å² of non-polar) belongs to the 88–95 loop.

The structural formula of the mannopentose (see Scheme 1) is such that the mannotriose can be embedded into it in two different ways. Residues i, ii and iii in the mannopentose (II) constitute the mannotriose; so do iv, v and ii. In subunit A, residues iv, v, ii and i with iv at the primary site, are defined. On the other hand, in subunit B, residues i, ii and iii are defined, with i at the primary site. The interactions of residues iv, v and ii in the former and those of residues i, ii and iii in the latter, are the same as those observed in the mannotriose complex. The additional residue found in subunit A has van der Waals interactions particularly with Ala90 and Thr91. This results in the additional burial of 58 Å² (39 Å² of non-polar)

**Figure 3.** Stereo view of the superposition of the sugar-binding region of mannotriose complex (violet) on that in the Me- α -Man complex (green). Interactions between the mannotriose and the lectin are shown.



Scheme 1.

of the surface of the protein, of which 39 \AA^2 (34 \AA^2 of non-polar) is accounted for by loop 88–95. In subunit B, the linkage from sugar residue ii is such that residues iv and v point to solution, accounting for the absence of electron density for them.

Carbohydrate conformation

All the sugar residues in the two complexes are in the chair conformation, the most commonly observed conformation in hexoses. The glycosidic torsion angles in the carbohydrates bound to the combining sites are listed in Table 2. The angles ϕ and ψ at the $\alpha 1-3$ linkage do not show much variation and are in agreement with those ($\phi = 72.5(\pm 11.0)^\circ$, $\psi = -112.3(\pm 22.5)^\circ$) observed in 84 out of 91 crystal structures.²⁴ The values at the $\alpha 1-6$ linkage do show some variation but they can be considered as belonging to one ($\phi = 65.4(\pm 9.0)^\circ$, $\psi = 182.6(\pm 5.1)^\circ$, $\omega = 66.4(\pm 10.2)^\circ$) of the three ($\phi = 65.4(\pm 9.0)^\circ$, $\psi = 182.6(\pm 5.1)^\circ$, $\omega = 66.4(\pm 10.2)^\circ$; $\phi = 66.5(\pm 10.8)^\circ$, $\psi = 180.7(\pm 15.1)^\circ$, $\omega = 185.0(\pm 11.2)^\circ$; $\phi = 67.4(\pm 14.4)^\circ$, $\psi = 109.1(\pm 13.7)^\circ$, $\omega = 203.0(\pm 22.7)^\circ$) major conformational families seen in 69 crystal structures. The observed variations can be attributed to the flexibility of the sugars at the secondary site. Notably, the sugar does not make any direct hydrogen bond with the protein residues in the crystal structure and are involved only in molecular contacts.

Molecular assembly in the crystal

An interesting aggregation pattern is observed in the crystal structure of the mannopentose complex (Figure 4(a)). The crystals contain as much as 90% (v/v) solvent. A careful examination of the crystal packing shows a scaffolding composed of protein molecules supporting solvent channels as large as 200 Å in diameter. The arrangement is reminiscent of that of lectin molecules cross-linked by multivalent sugar molecules.^{25–27} The mannopentose, however, does not cross-link protein molecules directly in the crystals; it participates in cross-linking through a sulfate ion and water oxygen atoms as illustrated in Figure 4(b). The SO_4 ion, which is located on a 3-fold axis interconnects the terminal mannose molecules from three subunits related by the crystallographic 3-fold through water-mediated interactions. In each subunit, a water oxygen atom bridges an oxygen atom of the sulfate ion to O3 and O4 of the mannosyl residue. The distances between the sulfate oxygen atom and the water oxygen atom, and the water oxygen atom and O3 and O4 of the mannose are 2.5 Å, and 2.7 Å and 3.1 Å, respectively. Interestingly, the protein molecules assemble as trimers of two subunits, which may be referred to, for convenience, as dimers (Figure 4(c)) in the crystal. In all, 1327 \AA^2 of surface area per subunit is buried on the formation of a trimer. In the biological tetramer of artocarpin, 1508 \AA^2 area per subunit is buried at the dimer–dimer interface. Furthermore, the dimer–dimer interface of the biological tetramer involves more non-polar interactions (67% of the total area buried is non-polar) as compared to the trimer interface, where the non-polar surface area buried is only a little over half (57%) of the total area buried.

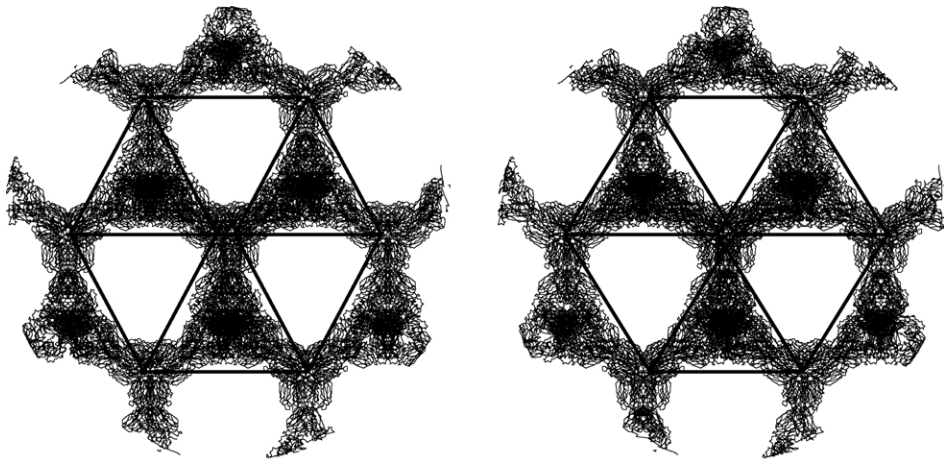
Structural rationale for carbohydrate specificity

The thermodynamic and inhibition studies carried out on artocarpin provide a vast amount of data on its binding to various carbohydrates. The crystal structures of the complexes of artocarpin with mannotriose and mannopentose together with the earlier reported structure of Me- α -Man complex provide the structural rationale for the protein's relative binding affinities for different

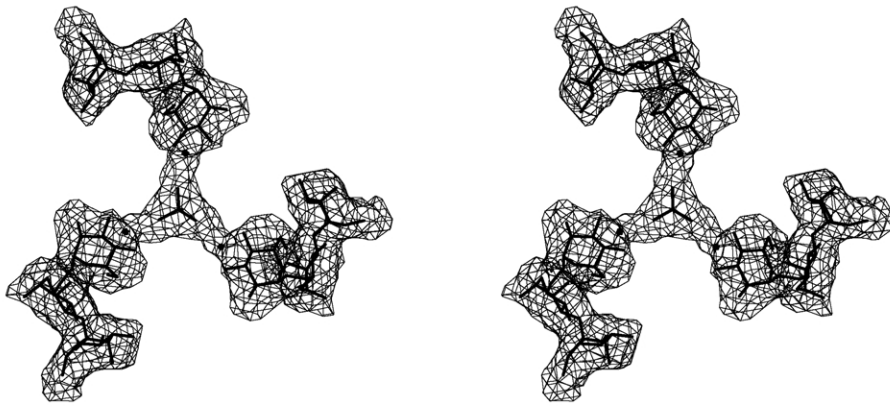
Table 2. Conformation of the bound mannotriose and mannopentose as observed in the crystal structures

	A	B	C	E	G	H	A'	B'
Man $\alpha 1-3$ Man								
ϕ (deg.) (O5–C1–O–C3')	80	87	86	86	85	87	89	88
ψ (deg.) (C1–O–C3'–C2')	-123	-120	-123	-127	-120	-129	-123	-138
Man $\alpha 1-6$ Man								
ϕ (deg.) (O5–C1–O–C6')	72	72	64	92	63	48	70	88
ψ (deg.) (C2–O–C6'–C5')	163	-177	150	-163	155	157	163	175
ω (deg.) (O–C6'–C5'–C4')	48	60	52	46	58	58	41	59

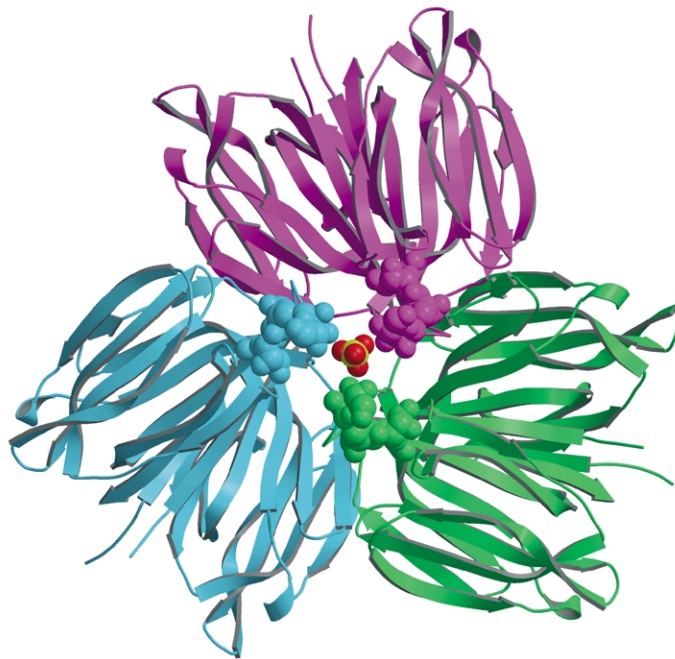
A' and B' refer to the two crystallographically independent subunits in the mannopentose complex.



(a)



(b)



(c)

Figure 4 (*legend opposite*)

Table 3. Correlation with the binding data

	Binding constants K_b (M^{-1})	Enthalpy $-\Delta H_b$ ($kJ\ mol^{-1}$)	Remarks
Glucose	150 ± 20	16.2 ± 2.0	Interaction of O2 is lost
Mannose^a	1640 ± 36	24.6 ± 1.4	348(156)^b
GlcNAc	NB	–	Interaction of O2 is lost; the introduction of additional flexibility of the acetyl arm
ManNAc	NB	–	Steric clash between acetyl arm and Gly15, Gly137 and Gly138
Me- α -Glc	341 ± 11	21.0 ± 1.0	Me has favorable interaction with Thr91 but with no interaction of O2
Me-α-Man^a	2500 ± 91	28.1 ± 0.8	367(195)^b
Me- β -Man	NB	–	Me has steric clash with Asp138
Man α 1,2Man	490 ± 10	18.6 ± 1.2	Interaction of the reducing Man O2 is lost
Man α 1,6Man	790 ± 34	25.9 ± 0.8	Interaction of the reducing Man O2 is lost
Man α 1,3Man	9600 ± 390	27.0 ± 0.8	O1 and O2 of the reducing Man interact with Ala90 and Thr91
Mannotriose^a	$20,100 \pm 900$	44.8 ± 0.4	547(277)^b
(GlcNAc) ₂ Man ₃	4317 ± 303	9.8 ± 0.7	Both the GlcNacs are exposed to solvent
Mannopentose^a	$21,200 \pm 600$	45.9 ± 0.9	610(316)/552(290)^b

^a Sugars in crystallographically characterized complexes are given in bold.

^b The surface areas (\AA^2) buried on complexation, with the hydrophobic components in parentheses.

carbohydrates. Table 3 lists the binding affinity and enthalpy (wherever it is available) of artocarpin to various sugars. At the monosaccharide level, Me- α -Man is the best ligand followed by Man, Me- α -Glc and Glc. The crystal structures of the complexes of artocarpin, provide the details of the basic set of interactions necessary for the recognition of a simple Man residue at the primary site. In addition to this set of interactions, the Me group of Me- α -Man, makes additional interactions with Thr91CB, Thr91CG2 and Asp138CG in the crystal structure of the Me- α -Man complex. Notably, binding to Me- α -Man involves 39 \AA^2 excess burial of non-polar accessible surface area compared to the binding of simple Man. The other monosaccharide complexes for which the crystal structures are not available include Glc and Me- α -Glc. Replacement of Man by Glc at the primary site results in the loss of the favourable interactions of O2 with Gly15 CA, Gly137 C, Gly137 CA and Asp138 N. This explains the poor affinity for Glc compared to that for Man. The addition of a Me group at the anomeric oxygen atom in axial configuration would establish additional interactions between the Me group and the side-chains of the residues Thr91 and Asp138 as observed in the Me- α -Man complex. This explains the better affinity for Me- α -Glc over that for Glc. However, the loss of the favourable interactions seen between O2 of Man and the protein residues leave Me- α -Glc as a weak binder compared to Me- α -Man and Man. ManNAc, GlcNAc and Me- β -Man do not show any binding to the lectin. Modelling involving the

replacement of Man by ManNAc resulted in severe steric clash between the acetyl arm of the sugar and the backbone atoms of protein residues Gly15, Gly137 and Asp138. Interestingly, the docking of GlcNAc did not lead to any steric clash, but the absence of the favourable interactions of O2 observed in Man and Me- α -Man, could be the reason for its poor binding. A β -substitution, as in Me- β -Man, leads to steric clashes with the side-chain atoms of Asp138, explaining the very poor affinity of the sugar for the protein.

At the disaccharide level, the order of binding affinity is as follows: Man α 1-3Man > Man α 1-6Man > Man α 1-2Man. The crystal structures of the complexes involving mannose and mannopentose provide the details of the interactions of Man α 1-3Man with the lectin. To understand the poor affinity exhibited by the lectin towards Man α 1-6Man as compared to that towards Man α 1-3Man, the former disaccharide was modelled at the carbohydrate-combining site of the lectin, with the torsion angle values corresponding to the three major conformers observed in 69 crystal structures: $\phi = 65.4(\pm 9.0)^\circ$, $\psi = 182.6(\pm 5.1)^\circ$, $\omega = 66.4(\pm 10.2)^\circ$ (conformer I), $\phi = 66.5(\pm 10.8)^\circ$, $\psi = 180.7(\pm 15.1)^\circ$, $\omega = 185.0(\pm 11.2)^\circ$ (conformer II) and $\phi = 67.4(\pm 14.4)^\circ$, $\psi = 109.1(\pm 13.7)^\circ$, $\omega = 203.0(\pm 22.7)^\circ$ (conformer III). The disaccharide models belong to conformer families I and II, when placed at the binding site resulted in steric clashes between the reducing mannose and the side-chain atoms of the protein residues Ala90 and Thr91, which are important for the

Figure 4. Stereo views of (a) the molecular assembly in the crystals of mannopentose complex, (b) the electron density in the $2F_o - F_c$ map contoured at 1σ together with the model indicating the role of SO_4 ion and water oxygen atoms in crosslinking the sugars. (c) Higher oligomeric state of artocarpin, as seen in the mannopentose complex. The SO_4 ion that sits on the 3-fold axis and the bound carbohydrates are shown in space-filling representation.

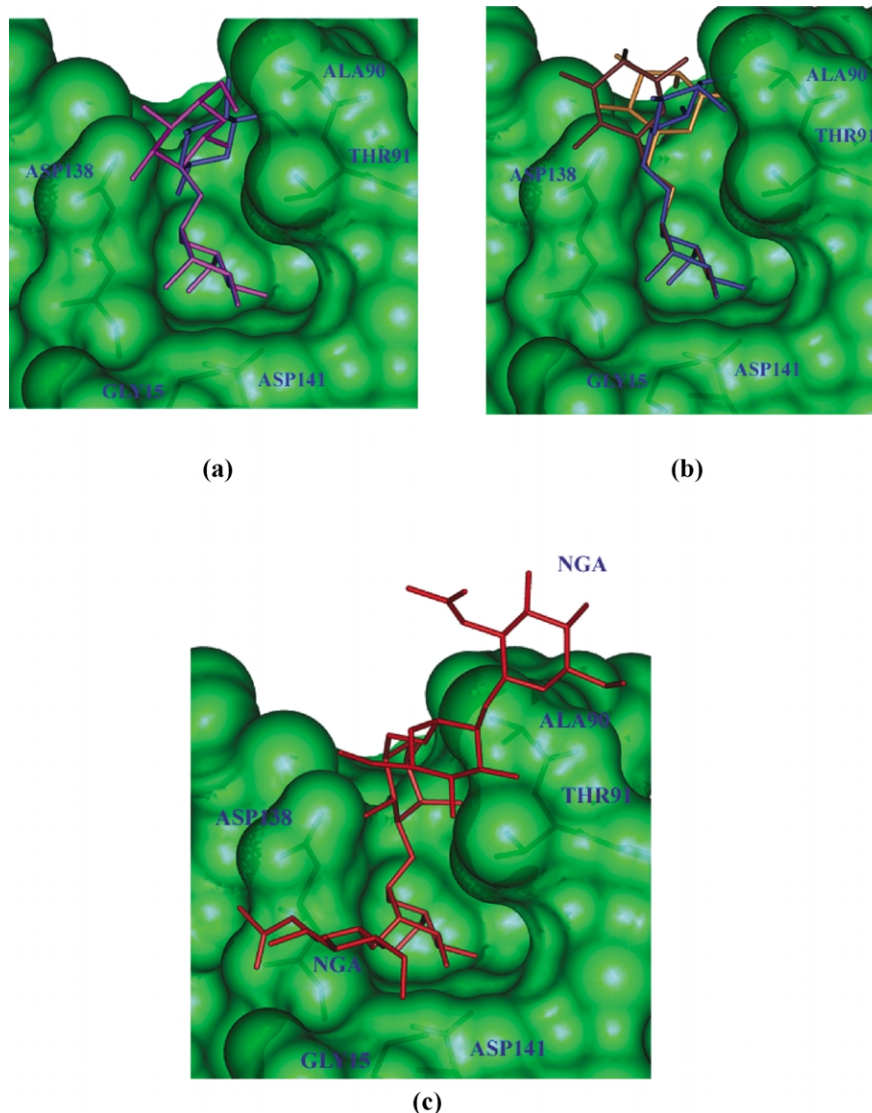
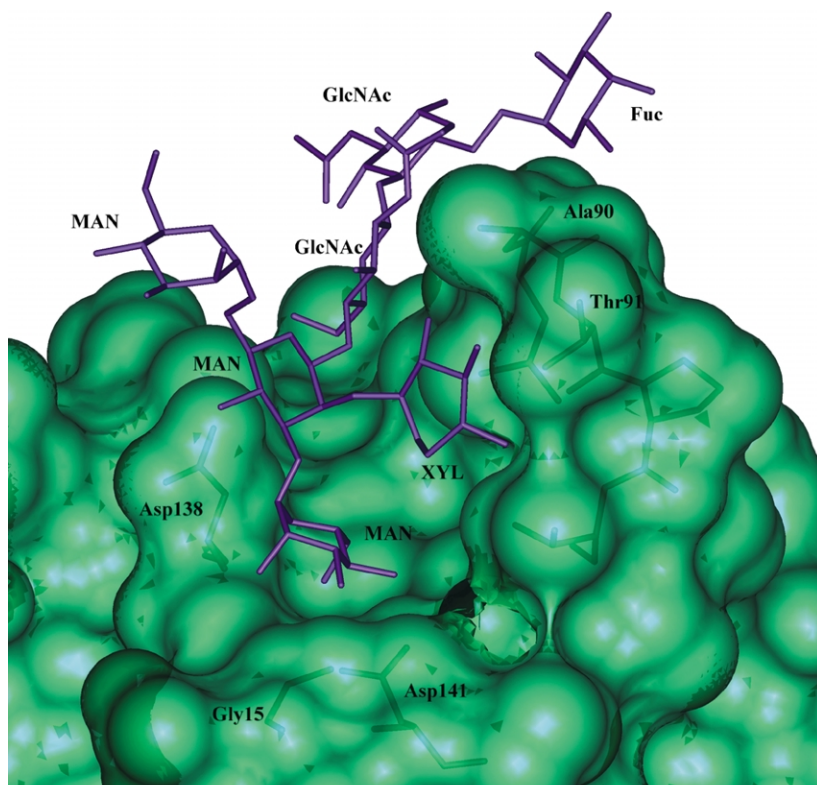


Figure 5 (legend opposite)

carbohydrate recognition at this site (Figure 5(a)). The disaccharides modelled with torsion angles corresponding to conformer III when bound to the lectin have fewer steric clashes with the protein than those in complexes involving conformers I and II. A permissible change in side-chain conformation, however, eliminates this clash, and leads to a hydrogen bond between O4 and Asp138OD1. Thus, only one of the three possible conformers of Man α 1-6Man can bind to the lectin. The two major conformers corresponding to Man α 1-2Man observed in 48 crystal structures are $\phi = 71.9(\pm 13.1)^\circ$, $\psi = -104.4(\pm 15.4)^\circ$ (conformer I) and $\phi = 62.2(\pm 8.3)^\circ$, $\psi = -175.0(\pm 10.3)^\circ$ (conformer II). The disaccharide modelled with torsion angles corresponding to conformer I resulted in severe steric clashes between O6 of the reducing Man with the main-chain atoms of the residues in loop 90-93 (Figure 5(b)). When the values of the torsion angles were changed to those of conformer II, there were no steric clashes, but the hydrogen bond involving O2 is lost. Thus, although

Man α 1-2Man in the second conformation can bind to the lectin, the interactions between them are likely to be weak.

Mannotriose, mannopentose and Gn₂Mn₃ (GlcNAc₂ mannotriose) are the only oligosaccharides that have been studied using isothermal titration calorimetry. Among these, the mannotriose and the mannopentose bind to artocarpin with almost the same affinity, while Gn₂Mn₃ binds weakly, even more weakly than the disaccharide Man α 1-3Man. As mentioned earlier, in the crystal structure of the mannotriose complex, mannotriose interacts with the protein through all of its sugar residues. Essentially the same interactions occur when the lectin binds the mannopentose, although a few additional van der Waals interactions exist in one of the subunits. The binding of Gn₂Mn₃ was modelled by adding two GlcNAc residues to the terminal Man of the mannotriose in its complex with artocarpin (Figure 5(c)). The addition of these two residues does not lead to any steric clash; nor does it result



(d)

Figure 5. Surface representation of the sugar-binding site with modelled (a) Man α 1–2Man (conformers I; blue, II; pink), (b) Man α 1–6Man (conformers I; blue, II; gold, III; brown), (c) Gn₂Mn₃ and (d) horseradish peroxidase oligosaccharide.

in any additional lectin–sugar interaction. The reason for the poor affinity of Gn₂Mn₃ compared to that of the mannotriose and the mannopentose, is not immediately obvious.

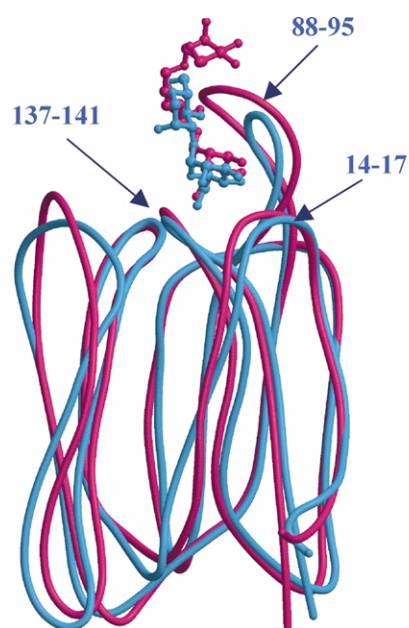
The binding affinity of artocarpin is the highest for horseradish peroxidase oligosaccharide (III), which is 1633-fold better compared to that for mannose. The branched mannosyl end of this oligosaccharide corresponds to mannotriose, the complex of which with artocarpin is one of the structures reported here. The superposition of the mannosyl end on the trisaccharide in the complex leads to severe steric clashes involving the xylose residue and loop 86–95 of the lectin. The unacceptable contacts can be removed by a movement of the loop (Figure 5(d)). This is precisely what happens when the modelled structure of the complex is energy minimized. In the process, the hydrogen bonds of O1 and O2 in the middle mannosyl residue with the lectin are lost. However, the hydroxyl groups of the xylose residue now make eight hydrogen bonds with the lectin. One of the non-reducing end residues has no interaction with the lectin in the energy-minimized model. However, the GlcNAc and fucose residues have many van der Waals contacts with the loop residues 87 to 93. The number of these protein–carbohydrate interactions is greater in the model of the complex of the lectin with the oligosaccharide than in the artocarpin–mannotriose complex. Also, the surface area buried in the latter

(852 Å², of which 454 Å² is non-polar) is greater than in the former (547 Å², of which 277 Å² is non-polar). That perhaps explains the high affinity of artocarpin for the horseradish peroxidase oligosaccharide.

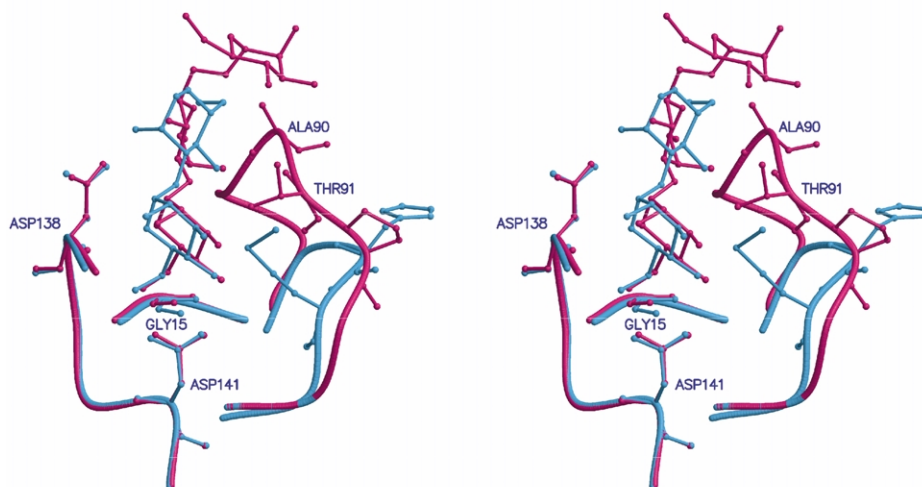
Specificity generation in jacalin-related mannose lectins

Artocarpin and heltuba lectin have 41% sequence identity and their subunits superpose with an r. m. s deviation of 1.3 Å in C α positions. They are specific to mannose at the monosaccharide level with comparable affinity. However, they differ considerably in their ability to bind di-, tri- and higher oligosaccharides.²⁰ The crystal structures reported here and those of the heltuba complexes with the same mannotriose and with the disaccharide Man α 1–2Man, provide the structural data to explain this difference.

In the complexes of heltuba with mannotriose, only the first two residues, namely, Man α 1–3Man, are crystallographically defined. In the corresponding complexes of artocarpin, at least three residues, namely, Man α 1–3Man α 1–6Man, are defined in both the structures. Therefore, a subunit in the heltuba–trisaccharide complex and one in the artocarpin–trisaccharide complex are used here for comparison. A superposition of these two subunits is shown in Figure 6(a). As far as the binding site is concerned, loops 14–17 and 137–141 (artocarpin



(a)



(b)

Figure 6. (a) Structural superposition of heltuba lectin (cyan) on artocarpin (red) with bound mannotriose. (b) Stereo view of the sugar-binding site (see the text for details).

numbering) have the same length in the two lectins. Loop 88–96 (artocarpin numbering) is considerably longer in artocarpin. Superposition of the binding sites and bound sugars, based only on the first two loops, is illustrated in Figure 6(b). The first sugar residue, that at the primary site, has nearly the same location and orientation in the two complexes; their interactions with the respective lectins are also the same. This is not true, however, about the second residue. Their

orientations are clearly different. This residue is involved in four hydrogen bonds in artocarpin. The corresponding number is only one in heltuba. The third sugar residue in the artocarpin complex has van der Waals interactions with loop 88–96. This residue is not seen in the crystal structure of the heltuba complex. However, its position could be modelled in more than one way. None of the models indicated interaction with the lectin. Thus, the crystal structures provide a clear rationale for

the higher affinity of artocarpin for the manno- triose compared to that of heltuba for the same sugar.

Man α 1–2Man binds to artocarpin only weakly. However, it binds with higher affinity to heltuba through interactions revealed by its complex with the lectin. A model of the artocarpin complex of the dimannose with the first residue anchored at the primary site in the same way as observed in crystal structures led to severe steric clashes of the hydroxyl groups of the second mannose residue with Leu89, Ala90 and Thr91. Models were constructed using the two independent sets of average values of torsion angles observed in crystal structures. As indicated earlier, one had severe steric clashes, while in the other the second mannose residue did not have any significant interaction with the lectin.

Lectins are known to employ several different strategies for generating carbohydrate specificity.^{9,28,29} One of them, responsible for the different blood group specificities of the two homologous lectins from winged beans, involve variation in loop length.²⁹ The same strategy appears to be responsible for the difference in the sugar specificities of artocarpin and heltuba. As illustrated in Figure 6, the main difference between the combining sites of artocarpin and heltuba is in the length of the third loop. The loop is longer in artocarpin. This is achieved through a four-residue insertion in the sequence. A mannose molecule bound at the primary site does not interact with this loop, and hence the two lectins have similar affinity at the monosaccharide level. The interactions with the loop become important in the binding of the disaccharide and higher oligo- saccharides and, at that level, the affinities of the two lectins for carbohydrates become different.

Materials and Methods

Crystallization

Artocarpin was extracted and purified using man- nose-Sepharose affinity chromatography as described.¹⁹ Crystals of artocarpin were grown using the vapour- diffusion technique by equilibrating a 10 μ l drop of 10 mg ml⁻¹ protein in 0.02 M phosphate buffer (pH 7.4), containing 0.1 M NaCl and 0.025% (w/v) sodium azide and 3–5% PEG1450 against a reservoir solution of 40% (w/v) PEG4000 in the same buffer. Crystals of size 1.0 mm \times 0.4 mm \times 0.4 mm grew in a week. In the presence of either manno- triose or mannopentose, crystallization trials failed in the same condition. New conditions were searched employing the vapour- diffusion technique using Hampton screening kits I and II.³⁰ This resulted in a condition where crystals of the complexes of artocarpin with manno- triose and manno- pentose could be grown by equilibrating a 10 μ l drop of 20 mg ml⁻¹ protein in 0.02 M phosphate buffer (pH 7.4), 0.1 M NaCl, 0.025% (w/v) sodium azide and 2–4 μ l of 2 M ammonium sulfate against a reservoir solution of 2 M ammonium sulfate. Crystals of size 0.6 mm \times 0.6 mm \times 0.2 mm grew in two weeks. However, the crystals

grown in the presence of manno- triose in this manner were highly mosaic and diffracted very poorly. So, the crystals of native artocarpin were soaked in the mother liquor containing 37 mM manno- triose and allowed to equilibrate for 48 hours. These crystals were used for further work. However, crystals of the complex involv- ing the mannopentose obtained through the use of ammonium sulfate as the precipitant diffracted moderately and were used in the work.

Data collection and processing

Data from the crystal of the manno- triose complex were collected at room temperature using a Mar research MAR300 imaging plate mounted on a Rigaku RU-200 X-ray generator. The mannopentose complex crystals did not diffract at room temperature and hence the crystals were soaked in a cryo-protectant solution of 25% (v/v) glycerol and the data were collected at 100 K (OXFORD cryo-systems) using a MAR345 imaging plate mounted on a Rigaku ULTRAX-18 X-ray generator. The data were processed using DENZO and SCALEPACK of the HKLSuite of programs.³¹ The structure factors from the processed data were calculated using program TRUNCATE of CCP4.³² Data collection statistics along with the cell parameters are given in Table 4. The Matthews coefficient³³ indicated the presence of two

Table 4. Data collection and refinement statistics

	Manno- triose	Mannopentose
Space group	$P2_1$	$P4_132$
<i>Unit cell dimensions</i>		
<i>a</i> (Å)	138.22	212.64
<i>b</i> (Å)	72.30	212.64
<i>c</i> (Å)	59.37	212.64
β (deg.)	94.5	–
<i>Z</i>	4	12
Resolution (Å)	2.35	3.5
Last shell (Å)	2.43–2.35	3.62–3.5
No. observations	134,666	67,845
No. unique reflections	48,516(4824)	19,233(1816)
Reflections with $I = 0$	3559(801)	1977(255)
Completeness (%)	99.4(99.3)	90.7(87.7)
R_{merge}^a (%)	10.4(49.8)	16.1(42.7)
Multiplicity	2.8	3.5
Protein atoms	9128	2280
Sugar atoms	182	79
Solvent atoms	490	1
R -factor ^b (%)	20.9	26.3
R_{free}^b (%)	22.6	29.0
Resolution range (Å)	20.0–2.35	20.0–3.5
No. reflections used in refine- ment	48504	19221
<i>RMS deviations from ideal values</i>		
Bond length (Å)	0.018	0.008
Bond angle (deg.)	3.4	1.8
Dihedral angle (deg.)	27.6	25.7
Improper (deg.)	2.75	1.43
<i>Residues (%) in Ramachandran plot</i>		
Core region (%)	90.4	87.7
Additionally allowed region (%)	8.5	12.3
Generously allowed region (%)	1.1	0.0
Disallowed region (%)	0.0	0.0

Values within parentheses refer to the last resolution shell.

^a $R_{\text{merge}} = \sum |I_i - \langle I \rangle| / \sum I_i$.

^b $R = \sum ||F_o| - |F_c|| / \sum |F_o|$; R_{free} is calculated in the same way but for a subset of reflections that is not used in the refinement.

tetrameric molecules in the asymmetric unit in the crystals of the mannotriose complex and two subunits in that of the mannopentose complex.

Structure solution and refinement

The structures of the complexes were solved using the molecular replacement program AMoRe.³⁴ The structure of artocarpin in its native form (PDB code 1J4S), which was solved earlier in this laboratory, was used as the search model. The binding of mannotriose to the native crystals upon soaking results in a unit cell transformation. The cell dimensions of the native crystal and the soaked crystal are the same along the *b* and *c* directions but the axis *a* is now doubled. Consequently, the number of the tetrameric molecules in the asymmetric unit doubles from one to two. Unique solutions for the two molecules were obtained with correlation coefficient (CC) and *R*-factor of 0.58 and 0.32, respectively. The transformation referred to above results in the disappearance of a 2₁ screw axis (that located at the midpoint of the *a* dimension of the native crystals). The two molecules related by this screw axis form the contents of the asymmetric unit of the crystals of the complex. These two molecules are now related by a pseudo 2₁ screw axis with a rotational component of 177.2° and a translational component of 35.6 Å. In the case of the mannopentose complex, a unique solution was obtained for two subunits in the asymmetric unit with CC and *R*-factor of 0.62 and 0.35, respectively. Both the structures were refined in a similar manner. To start with, 40 cycles of rigid body refinement followed by 100 cycles of positional refinement using CNS³⁵ with "mlf" target were carried out. At this stage, clear unambiguous density for the carbohydrates appeared in $F_o - F_c$ and $2F_o - F_c$ maps. Electron densities for mannotriose in six (A, B, C, E, G and H) of the eight subunits and for mannose in the remaining two subunits (D and F) were seen in the mannotriose complex. As reported elsewhere,¹⁹ the protein is eluted from a mannose-Sepharose column, and the protein used for crystallization could have been in the mannose bound form. A closer look at the binding-site region of the D and F subunits reveals no space for mannotriose to bind in the crystal. Any replacement of mannose by mannotriose will lead to severe steric clash with neighbouring protein molecules. In the mannopentose complex, densities for mannotriose in one subunit and for four of the five sugar residues of the mannopentose in the other subunit were seen. The coordinates of the sugars were generated using the web based program SWEET† and were fitted into the electron density using FRODO.³⁶ Subsequent cycles of positional refinement revealed clear densities for ordered solvent molecules and for four sulfate ions in the mannopentose complex structure. Water oxygen atoms were added successively to the model, using peaks with heights greater than 2.5σ in $F_o - F_c$ maps and 0.8σ in $2F_o - F_c$ maps. Although 54 water molecules could be identified in the mannopentose complex, only one critically involved in crystal packing was included in refinement, in view of the limited resolution of the structure. Omit maps were used in the course of refinement to remove model bias. Bulk solvent corrections and overall anisotropic *B*-factor corrections were used throughout the refinement. Iterative cycles of model

building and refinement were carried out until *R* and *R*_{free} converged. *B*-factors of the atoms were refined individually in the mannotriose complex structure. In the case of the mannopentose complex, group *B*-factor refinement was carried out. The final values of *R* and *R*_{free}, and other relevant refinement statistics are given in Table 4. The refined models were checked using PROCHECK.³⁷

Analysis and modelling

Possible hydrogen bonds were identified using program HBPLUS.³⁸ Contacts involving oxygen and nitrogen atoms with distances less than 3.6 Å and with donor–hydrogen–acceptor angle greater than 90° were treated as hydrogen bonds. Molecular superpositions were performed using program ALIGN.³⁹ Program NACCESS was employed for calculating accessible surface areas‡. Binding of various sugars to artocarpin was modelled using INSIGHT II and the refined coordinates of the mannotriose complex structure were used for the modelling. Distance-dependant dielectric constant was used throughout the minimization. During minimization, the loop residues 85–95 and the sugar residues except the mannose bound to the primary site were allowed to move. Then 200 cycles of conjugate gradient minimization followed by 1000 cycles of steepest descent minimization were carried out. Figures 1, 3, 4(c) and 6 were prepared using BOBSCRIPT,⁴⁰ Figures 2, and 4(a) and (b) were prepared using FRODO, and Figure 5 was prepared using INSIGHT II.

Protein Data Bank accession codes

The atomic coordinates and the structure factors of both the artocarpin–oligosaccharide complexes were deposited in the RCSB Protein Data Bank (PDB accession codes 1VBO and IVBP).

Acknowledgements

The data sets were collected at the X-ray Facility for Structural Biology, at the Institute, supported by the Department of Science & Technology (DST) and the Department of Biotechnology (DBT) of the Government of India. Computations were performed at the Supercomputer Education and Research Centre of the Institute, and the Bioinformatics Centre and the Graphics facility, both supported by DBT. Financial support from DST is acknowledged.

References

1. Sharon, N. & Lis, H. (1989). Lectins as cell recognition molecules. *Science*, **246**, 227–246.
2. Lis, H. & Sharon, N. (1998). Lectins: carbohydrate specific proteins that mediate cellular recognition. *Chem. Rev.* **98**, 637–674.
3. Drickamer, K. (1999). C-type lectin-like domains. *Curr. Opin. Struct. Biol.* **9**, 585–590.

† <http://www.dkfzheidelberg.de/spec/sweet2/doc/index.html>

‡ <http://wolf.bms.umist.ac.uk/naccess/>

4. Rini, J. (1999). New animal lectin structures. *Curr. Opin. Struct. Biol.* **9**, 578–584.
5. Vijayan, M. & Chandra, N. R. (1999). Lectins. *Curr. Opin. Struct. Biol.* **9**, 707–714.
6. Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). Legume lectin structure. *Biochim. Biophys. Acta*, **1383**, 9–36.
7. Rudiger, H. & Gabius, H. J. (2001). Plant lectins: occurrence, biochemistry, functions and applications. *Glycoconj. J.* **18**, 589–613.
8. Barre, A., Bourner, Y., Van Damme, E. J., Peumans, W. J. & Rouge, P. (2001). Mannose-binding plant lectins: different structural scaffolds for a common sugar-recognition process. *Biochimie*, **83**, 645–651.
9. Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A. & Vijayan, M. (1996). A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a beta-prism fold. *Nature Struct. Biol.* **3**, 596–603.
10. Aucouturier, P., Mihaesco, E., Mihaesco, C. & Preud'homme, J. L. (1987). Characterization of jacalin, the human IgA and IgD binding lectin from jackfruit. *Mol. Immunol.* **24**, 503–511.
11. Mahanta, S. K., Sanker, S., Rao, N. V., Swamy, M. J. & Surolia, A. (1992). Primary structure of a Thomsen–Friedenreich-antigen-specific lectin, jacalin [*Artocarpus integrifolia* (jackfruit) agglutinin]. Evidence for the presence of an internal repeat. *Biochem. J.* **284**, 95–101.
12. Yang, H. & Czapla, T. H. (1993). Isolation and characterization of cDNA clones encoding jacalin isolectins. *J. Biol. Chem.* **268**, 5905–5910.
13. Jeyaprakash, A. A., Rani, P. G., Reddy, G. B., Banumathi, S., Betzel, C., Sekar, K. *et al.* (2002). Crystal structure of the jacalin-T-antigen complex and a comparative study of lectin-T-antigen complexes. *J. Mol. Biol.* **321**, 637–645.
14. Jeyaprakash, A. A., Katiyar, S., Swaminathan, C. P., Sekar, K., Surolia, A. & Vijayan, M. (2003). Structural basis of the carbohydrate specificities of jacalin: an X-ray and modelling study. *J. Mol. Biol.* **332**, 217–228.
15. Bunn-Moreno, M. M. & Campos-Neto, A. (1981). Lectin(s) extracted from seeds of *Artocarpus integrifolia* (jackfruit): potent and selective stimulator(s) of distinct human T and B cell functions. *J. Immunol.* **127**, 427–429.
16. De Miranda-Santos, I. K. F., Mengel, J. O., Jr, Bunn-Moreno, M. M. & Campos-Neto, A. (1991). Activation of T and B cells by crude extract of *Artocarpus integrifolia* is mediated by a lectin distinct from jacalin. *J. Immunol. Methods*, **140**, 197–203.
17. Misquith, S., Rani, P. G. & Surolia, A. (1994). Carbohydrate binding specificity of the B-cell maturation mitogen from *Artocarpus integrifolia* seeds. *J. Biol. Chem.* **269**, 30393–30401.
18. Rani, P. G., Bachhawat, K., Misquith, S. & Surolia, A. (1999). Thermodynamic studies of saccharide binding to artocarpin, a B-cell mitogen, reveals the extended nature of its interaction with mannotriose [3,6-di-O-(α -D-mannopyranosyl)-D-mannose]. *J. Biol. Chem.* **274**, 29694–29698.
19. Rani, P. G., Bachhawat, K., Reddy, G. B., Oscarson, S. & Surolia, A. (2000). Isothermal titration calorimetric studies on the binding of deoxytrimannose derivatives with artocarpin: Implications for a deep-seated combining site in lectins. *Biochemistry*, **39**, 10755–10760.
20. Bourne, Y., Zamboni, V., Barre, A., Peumans, W. J., Van Damme, E. J. & Rouge, P. (1999). *Helianthus tuberosus* lectin reveals a widespread scaffold for mannose-binding lectins. *Struct. Fold. Des.* **7**, 1473–1482.
21. Bourne, Y., Roig-Zamboni, V., Barre, A., Peumans, W. J., Houles Astoul, C., Van Damme, E. J. & Rouge, P. (2004). The crystal structure of the *Calystegia sepium* agglutinin reveals a novel quaternary arrangement of lectin subunits with a beta-prism fold. *J. Biol. Chem.* **279**, 527–533.
22. Suresh, A. S., Rani, P. G., Pratap, J. V., Sankaranarayanan, R., Surolia, A. & Vijayan, M. (1997). Homology between jacalin and artocarpin from jack fruit (*Artocarpus integrifolia*) seeds. Partial sequence and preliminary crystallographic studies of artocarpin. *Acta Crystallog. sect. D*, **53**, 469–471.
23. Pratap, J. V., Jeyaprakash, A. A., Rani, P. G., Sekar, K., Surolia, A. & Vijayan, M. (2002). Crystal structures of artocarpin, a moraceae lectin with mannose specificity, and its complex with methyl- α -D-mannose: implications to the generation of carbohydrate specificity. *J. Mol. Biol.* **317**, 237–247.
24. Petrescu, A. J., Petrescu, S. M., Dwek, R. A. & Wormald, M. R. (1999). A statistical analysis of N and O-glycan linkage conformations from crystallographic data. *Glycobiology*, **9**, 343–352.
25. Brewer, C. F. (2001). Lectin cross-linking interactions with multivalent carbohydrates. *Advan. Expt. Med. Biol.* **491**, 17–25.
26. Brewer, C. F., Miceli, M. C. & Baum, L. G. (2002). Clusters, bundles, arrays and lattices: novel mechanisms for lectin-saccharide-mediated cellular interactions. *Curr. Opin. Struct. Biol.* **12**, 616–623.
27. Dam, T. K. & Brewer, C. F. (2003). Carbohydrate-lectin cross-linking interactions: structural, thermodynamic, and biological studies. *Methods Enzymol.* **362**, 455–486.
28. Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A. & Vijayan, M. (1997). Crystal structure of the peanut lectin-T-antigen complex. Carbohydrate specificity generated by water bridges. *Curr. Sci.* **72**, 855–861.
29. Manoj, N., Srinivas, V. R., Surolia, A., Vijayan, M. & Suguna, K. (2000). Carbohydrate specificity and salt-bridge mediated conformational change in acidic winged bean agglutinin. *J. Mol. Biol.* **302**, 1129–1137.
30. Jancarik, J. & Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallog.* **24**, 409–411.
31. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography, Part A* In Methods of Enzymology (Carter, C. W. Jr & Sweet, R. M., eds), vol. 276, pp. 307–326, Academic Press, New York.
32. Collaborative Computational Project No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
33. Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.
34. Navaza, J. (1994). AmoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 445–449.
35. Brunger, A. T., Adams, P. D. & Rice, L. M. (1998). Recent developments for the efficient crystallographic refinement of macromolecular structures. *Curr. Opin. Struct. Biol.* **8**, 606–611.
36. Jones, T. A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallog.* **11**, 268–272.

-
37. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067.
 38. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
 39. Cohen, G. E. (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallog.* **30**, 1160–1161.
 40. Esnouf, R. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 132–134.