Glycobiology vol. 15 no. 10 pp. 1033–1042, 2005
doi:10.1093/glycob/cwi088
Advance Access publication on June 8, 2005

Crystal structure of banana lectin reveals a novel second sugar binding site

Jennifer L. Meagher[2], Harry C. Winter[3], Porscha Ezell[2], Irwin J. Goldstein[3], and Jeanne A. Stuckey[1,2,3]

[2]Life Sciences Institute, University of Michigan, Ann Arbor, MI; and [3]Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI

Received on May 4, 2005; revised on May 17, 2005; accepted on June 1, 2005

Banana lectin (Banlec) is a dimeric plant lectin from the jacalin-related lectin family. Banlec belongs to a subgroup of this family that binds to glucose/mannose, but is unique in recognizing internal α1,3 linkages as well as β1,3 linkages at the reducing termini. Here we present the crystal structures of Banlec alone and with laminaribiose (LAM) (Glcβ1, 3Glc) and Xyl-β1,3-Man-α-*O*-Methyl. The structure of Banlec has a β-prism-I fold, similar to other family members, but differs from them in its mode of sugar binding. The reducing unit of the sugar is inserted into the binding site causing the second saccharide unit to be placed in the opposite orientation compared with the other ligand-bound structures of family members. More importantly, our structures reveal the presence of a second sugar binding site that has not been previously reported in the literature. The residues involved in the second site are common to other lectins in this family, potentially signaling a new group of mannose-specific jacalin-related lectins (mJRL) with two sugar binding sites.

Key words: banana lectin/laminaribiose/mannose-specific jacalin-related lectin

Introduction

Lectins are a group of proteins of nonimmune origin that recognize and bind to carbohydrates without modifying them. They have been enormously useful as reagents for typing blood, characterizing carbohydrate-containing molecules on cell surfaces and in solution, and establishing and confirming metabolic pathways (Sharon and Lis, 2004). Plant lectins are the most well characterized group of lectins.

The banana lectin (Banlec) was first isolated from *Musa paradisiac* by Koshte and colleagues (1990). They reported that Banlec is a homodimeric protein that binds mannose and mannose-containing oligosaccharides and functions as a potent T-cell mitogen. More recently, Peumans and colleagues cloned and sequenced Banlec (*Musa acuminata*) and confirmed that it is composed of two identical subunits of 15 kDa showing specificity toward mannose. Based on sequence homology and secondary structure predictions, they classified Banlec as a member of the jacalin-related superfamily of lectins (JRL) (Peumans *et al.*, 2000). The JRL family is subdivided into two groups based on carbohydrate specificity: galactose-specific jacalin-related lectins (gJRL) and mannose-specific jacalin-related lectins (mJRL) (Raval *et al.*, 2004). In addition to carbohydrate binding specificity, the gJRL and mJRL groups differ in subunit organization. Jacalin, the prototype gJRL, is isolated from the seeds of jackfruit (*Artocarpus integrifolia*) and is composed of two chains, a long α-chain and a short β-chain, which arise from proteolytic cleavage of a precursor protein (Yang and Czapla, 1993). Members of mJRL are composed of a single protein chain.

The structure of jacalin was solved in 1996, revealing a novel fold for lectins, termed a β-prism-I fold (Figure 1A), which is a three-sided β-prism with each side made up of a four-stranded Greek key (Sankaranarayanan *et al.*, 1996). Since that time, additional crystal structures of JRL family members have been solved: *Maclura pomifera* agglutinin (MPA) (1JOT; Lee *et al.*, 1998) from the gJRL group and *Helianthus tuberosus* lectin (heltuba) (1C3K; Bourne *et al.* 1999), artocarpin (also from *Artocarpus integrifolia*) (1J4U; Partap *et al.*, 2002), and *Calystegia sepium* lectin (calsepa) (1OUW; Bourne *et al.*, 2004) from the mJRL group. Of the mJRL structures examined, heltuba and artocarpin have been solved bound to ligand molecules. These structures show a common sugar binding site composed primarily of residues in the loop regions on the top of the first face of the β-prism (Figure 1) (Sankaranarayanan *et al.*, 1996).

Banlec, a member of the mJRL group, exhibits some extraordinary carbohydrate binding properties and much biochemical work has been done to explore its carbohydrate binding specificity (Mo *et al.*, 2001; Goldstein *et al.*, 2001). In contrast to other mannose/glucose binding lectins, which recognize α-linked gluco- and manno-pyranosyl groups of polysaccharide chain ends, Banlec was shown to bind to internal 3-*O*-α-D-glucopyranosyl units (Mo *et al.*, 2001). It differs further from typical mannose/glucose-binding lectins in its ability to bind to the reducing ends of β1,3-linked glucosyl oligosaccharides [e.g., laminaribiose (LAM) and its higher homologs, laminaridextrins] (Goldstein *et al.*, 2001). Banlec is analogous to animal receptors present in the plasma and blood cells of crayfish (Davic and Soderhall, 1990) and insect species (Chen *et al.*, 1998) which recognize and bind to β1,3-glucans. A recent report showed the x-ray crystallographic structure of a β1,3-glucanase (laminarinase), which binds the nonreducing terminus of the laminarin oligosaccharides, with Lam_6 being a 10-fold better ligand than Lam_2 (Van Bueren *et al.*, 2005). In contrast, Banlec binds Lam_2 or higher oligosaccharides with nearly identical affinity (Goldstein *et al.*, 2001).

[1]To whom correspondence should be addressed; e-mail: jass@umich.edu

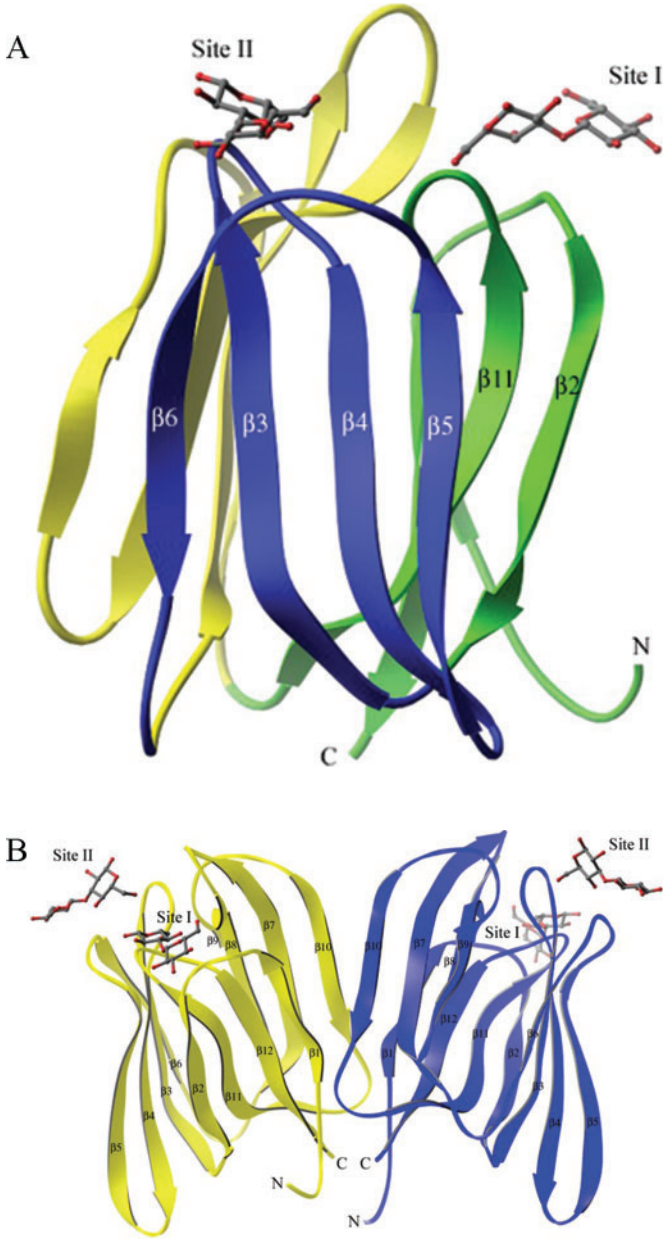

Fig. 1. Structure of Banlec. (**A**) Ribbon diagram of the Banlec monomer showing laminaribiose (LAM) bound in two distinct sugar binding sites. The three faces of the β-prism are colored differently, face 1 is green, face 2 is blue and face 3 is yellow. (**B**) Ribbon diagram of the biological dimer of banana lectin (Banlec) depicting the two sugar binding sites on each monomer. In the dimer, Banlec monomers are shown in blue and yellow. In both pictures, LAM is shown in gray (C atoms) and red (O atoms).

In view of the unusual variety of carbohydrates that bind to Banlec, it was highly desirable to obtain the X-ray crystal structures with and without carbohydrates bound to determine the precise mechanism of sugar binding. Previously, an attempt to solve the crystal structure of Banlec was only able to trace the Cα positions of the proposed biological dimer (Singh *et al.*, 2004). Here we provide the complete, refined structure of the *M. acuminata* Banlec, both free and complexed with two distinctly different disaccharides, LAM and Xyl-β1,3 Man-α-*O*-Methyl (XM). These structures show Banlec to be a dimeric protein with a similar fold to other family members. The surprising discovery was the presence of two distinct sugar binding sites per molecule (Figure 1). Site I, which was previously predicted based on sequence homology to heltuba and the newly found site II. With the information gained from this new site, we discovered sequence motifs that can be used to identify additional carbohydrate binding sites in other members of the JRL family.

Results and discussion

Overall structure of Banlec. A monomer of Banlec contains a single polypeptide chain that forms twelve β-strands in a β-prism-I fold (Figure 1A). This fold contains three four-stranded antiparallel β-sheets shaped like a prism with pseudo three-fold symmetry. The first face of the prism is a pseudo Greek key motif composed of strands 1, 12, 11, and 2. The second and third face of the prism are true Greek key motifs (Richardson, 1981) containing strands 5, 4, 3, and 6 and strands 9, 8, 7, and 10, respectively.

Banlec has been shown to be a dimer in solution (Peumans *et al.*, 2000). It crystallizes with two molecules in the asymmetric unit with the biological dimer forming about a crystallographic 2-fold axis (Figure 1B). The assignment of the biological dimer was based on the nature of the dimeric interactions. The interface of the biological dimer is formed primarily from the first face of the β-prism of each molecule. A predominately hydrophobic interface is formed between β1 (residues 4–10), β10 (residues 110–118), and the two C-terminal residues, E140 and P141 of each monomer encompassing a buried surface area of 1500 $Å^2$. β-strands 1 and 10 from different subunits pack in an antiparallel fashion forming backbone hydrogen bonds that produces an extended eight-stranded β-sheet structure. This dimeric interface is similar to some other family members, lending support that this dimer is the biological form of Banlec.

Additional support that the biological dimer forms about the crystallographic 2-fold axis can be found by examining the interface of the dimer in the asymmetric unit. This interface is mainly hydrophilic, composed of residues 50–58 on β5 of monomer A interacting with the loop regions between strands β7 and β8 (residues 106–109) and strands β9 and β10 (residues 82–85) of monomer B encompassing only 630 Å2 of buried surface area. This interaction is facilitated by the presence of a cadmium ion at the interface. Residues Asp41 and His54 from monomer A and residue His84 from monomer B form a cadmium binding site similar to the one found in the crystal structure of *Salmonella typhimurium* histidine-binding protein (Trakhanov *et al.*, 1998). The presence of this ion is a result of the crystallization process and not contained in the biological protein.

Comparison of Oligomeric State of Family Members. Previously, JRL family members were classified into three categories based on their quaternary structures (Raval *et al.*, 2004). The first group contains jacalin and artocarpin, which exist as a tetramer in solution and in crystal form (Sankaranarayanan *et al.*, 1996). The tetramer has been described as a dimer of dimers. Each dimer interface is

equivalent to that of the biological Banlec dimer. The second group contains calsepa whose quaternary structure is a dimer (Bourne *et al*., 1999). In this dimer, however, the two monomers form a parallel β-structure interface that positions the N- and C-termini of the monomers on opposing faces of the molecule, which is opposite to the configuration seen in the Banlec dimer interface (Figure 1B). The third category contains heltuba, which forms a tetramer in solution, but crystallizes as an octamer consisting of both Banlec- and calsepa–like dimer interfaces (Bourne *et al*., 1999). The structure of Banlec falls into a fourth structural category because its biological quaternary structure is a unique dimer.

Description of the carbohydrate binding sites. The initial difference electron density maps for the Banlec : LAM and Banlec : XM complexes each showed density for two distinct disaccharide binding sites per Banlec monomer (Figure 2). This is the first time a second binding site has been noted for a JRL family member. The binding of the two carbohydrate molecules per monomer of Banlec is consistent with recent calorimetric titration data presented by Winter and colleagues (2005).

The location of the first carbohydrate binding site (site I) is common to all lectins in this family. It is located on the top of the first face of the β-prism, involving two loop regions (Figure 1). The first loop, termed the GG loop, is between strands β1 and β2 and contains a Gly–Gly motif. The second loop, termed the ligand binding loop, lies between strands β11 and β12 and contains a G-X_3-D motif. The second carbohydrate binding site (site II) is located on top of the second face of the β-prism, formed from residues in the loop regions between strands β3 and β4 (ligand binding loop) and strands β5 and β6 (GG loop). The interactions between Banlec and carbohydrate are very similar in binding site I and II.

Not only is the discovery of a second carbohydrate binding site unique to Banlec, the mode of sugar binding is also unique. Structures of oligosaccharides bound to other jacalin family members show the nonreducing sugar to be inserted into binding site I, whereas the two Banlec carbohydrate-bound structures presented here have the reducing sugar unit bound to both sites.

In the Banlec : LAM complex, Glc1, the reducing glucose unit of LAM, is inserted into both binding sites and makes numerous interactions with the protein, whereas the nonreducing glucose unit (Glc2) has no specific subsite for binding. The majority of protein–LAM interactions are formed through main chain amide nitrogens of the protein with the

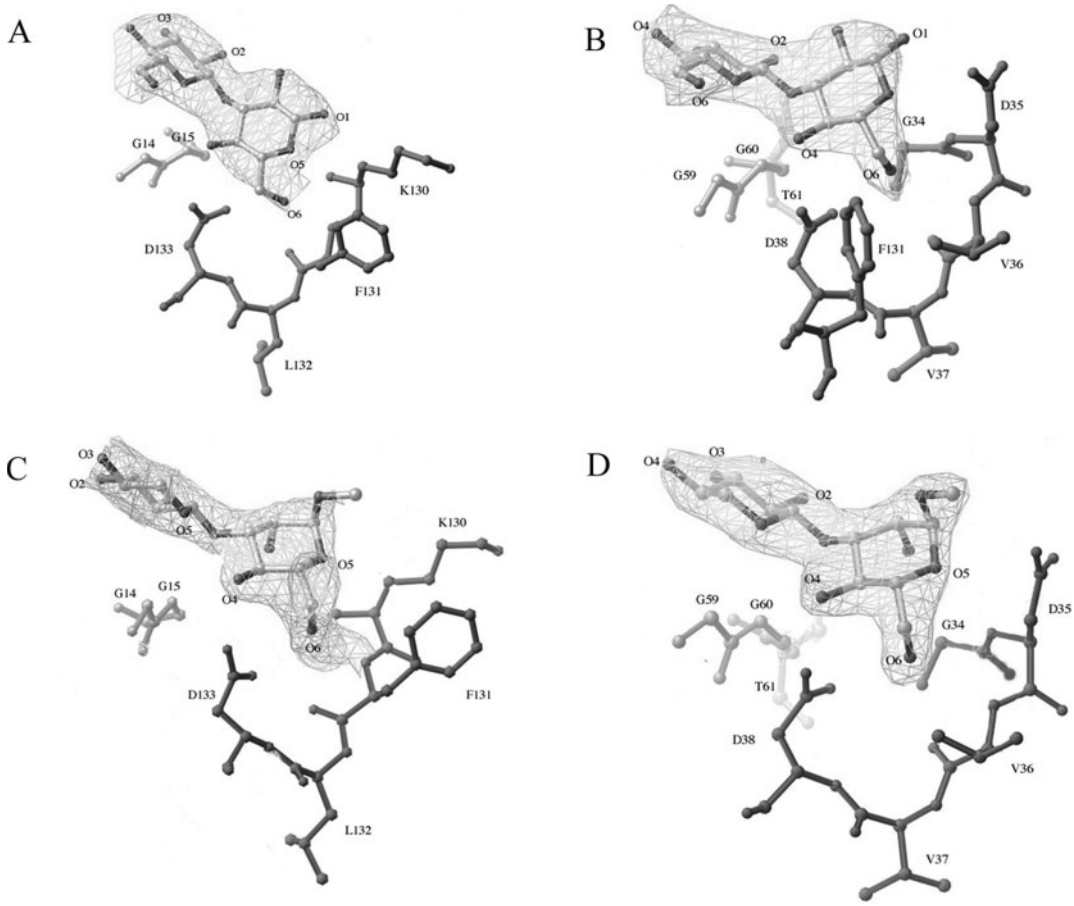

Fig. 2. Initial difference electron density for bound carbohydrates. (**A**) Monodiagram of laminaribiose bound to Banlec in binding site I, (**B**) binding site II, (**C**) Monodiagram of Xyl-β1,3 Man-α-*O*-methyl bound to banana lectin (Banlec) in binding site I, and (**D**) binding site II. *Fo – Fc* electron density at 3σ is shown for the disaccharide (large ball-n-sticks). Residues important in carbohydrate binding are labeled.

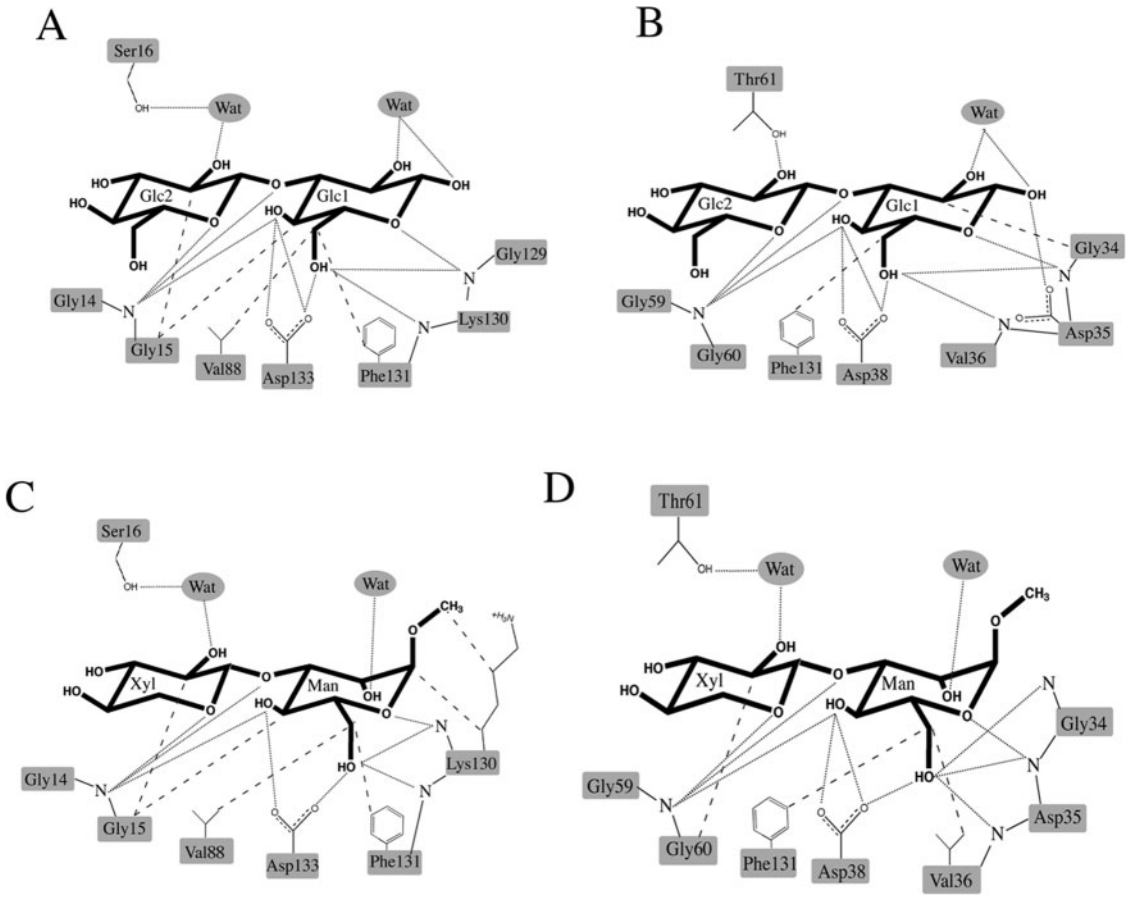

Fig. 3. Schematic diagram of binding interactions. (**A**) Schematic diagrams of Banlec bound to laminaribiose (LAM) in binding site I, (**B**) LAM in binding site II, (**C**) Xyl-β1,3 Man-α-*O*-Methyl (XM) in binding site I, and (**D**) XM in binding site II. The sugar residues are in bold. Hydrogen bonds are shown in small dashed lines and hydrophobic interactions are shown in large dashed lines.

O4, O6 hydroxyls and the O3 glycosidic oxygen of the reducing unit, which is consistent with the binding studies performed by Mo and colleagues (Figures 3A, B and Table I) (Mo *et al.*, 2001). Binding site I is composed of main chain atoms from residues Gly15 (a conserved glycine in the GG loop, Figure 4), Lys130, Phe131, and the side chain carboxylate of the conserved Asp133. The equivalent residues in the second binding site are Gly60, Asp35, Val36, and the side chain of Asp38. Although the other conserved glycine in the GG loop (Gly14 in binding site I; Gly59 in binding site II) makes no direct interactions with the sugar, the lack of side chain atoms on the GG loop allow binding of the nonreducing unit of a disaccharide having an internal β1–3 linkage.

In the first binding site, there are many interactions between Banlec and the reducing glucosyl residue (Glc1) of LAM (Figure 3A and Table I). The amide of Gly15 is hydrogen bonded to both the O4 atom and the oxygen atom of the β1,3 glycosidic bond, and the Gly15 Cα atom forms van der Waals interactions with the C4 atom of Glc1. The amide backbone of Lys130 is hydrogen bonded to both the O5 and O6. The amide backbone of Phe131 is hydrogen bonded to O6 and the Cδ1 atom is within van der Waals distance from the C6 atom. The carboxylate oxygen atoms of the conserved Asp133 are hydrogen bonded to the O4 and O6 atoms of Glc1. In addition, waters mediate interactions between O1 and O2 of Glc1 and O of the conserved Gly129 and Nζ of Lys130. In contrast, there are few interactions between the nonreducing Glc2 of LAM and Banlec. A water molecule hydrogen bonded to the O2 of Glc2 glucosyl residue mediates an interaction with Oγ1 atom of Ser16 and the Cα and amide group of the conserved Gly15 are within van der Waals distance of the C2 and O5 atoms of Glc2, respectively.

The interactions between the second binding site of Banlec and LAM differ only slightly from binding site I for equivalent residues (Figure 3B and Table I). The Oγ1 of Thr61 is hydrogen bonded directly to the O2 atom of Glc2 and the amide of Gly60 is hydrogen bonded to the O5 of Glc2. The Glc1 unit has several additional interactions, the carboxylate oxygen of Asp35 is hydrogen bonded to O2 atom of Glc1 and the Cα atom of Gly34 interacts with the C2 atom of Glc1. In the second binding site, there is a similar network of water molecules bonded to O2 of Glc2, O2 of Glc1 and the glycosidic bond oxygen atom (not shown). The only protein side chain to interact with the sugar residues in both binding sites is the phenyl ring of Phe131.

The binding of Banlec to XM is similar to LAM. Despite the differences in the two molecules, the atoms that are

Table I. Nonbonded interactions between laminaribiose atoms and banana lectin residues

Sugar	Binding site I		Binding site II	
	Protein atom	Distance (Å)[a]	Protein atoms	Distance (Å)[a]
Glc1 O1			D35 Oδ1	3.36
Glc1 O1	WAT	3.56	WAT	3.48
Glc1 O2	WAT	3.80	WAT	3.46
Glc1 C2			G34 Cα	4.13
Glc1 O4	G15 NH	2.97	G60 NH	3.19
Glc1 O4	D133 Oδ1	2.77	D38 Oδ1	3.02
Glc1 O4	D133 Oδ2	3.67[b]	D38 Oδ2	3.68
Glc1 C4	G15 Cα	3.72		
Gcl1 O5	K130 NH	3.06	D35 NH	3.24
Glc1 O6	K130 NH	2.81	D35 NH	2.93
Glc1 O6	F131 NH	2.65	V36 NH	3.30
Glc1 O6	D133 Oδ2	2.95	D38 Oδ2	3.14
Glc1 C6	F131 Cδ1	4.24	F131 Cε2	3.24
Glc1 C6	V88 Cγ1	3.77		
O3 (β1,3)	G15 NH	2.90	G60 NH	3.11
Glc2 O2			T61 Oγ1	3.42
Glc2 O2	WAT	2.83		
Glc2 C2	G15 Cα	3.98		
Glc2 O5	G15 NH	3.48	G60 NH	3.54

[a]Distances represent the average between the binding sites in the A chain and the B chain.
[b]Hydrogen bond is only seen in the B chain.

responsible for the majority of interactions, namely the O4 and O6 hydroxyls on the reducing sugar, are common to both sugars. In binding site I, XM has the same interactions with Banlec as LAM, except only one of the carboxylate atoms of Asp133 interacts with O4 of the reducing sugar (Figure 3C and Table II). In binding site II, there is an additional interaction between the amide backbone of Gly34 and the O6 of the mannose residue that is not seen in site I; however, the interaction with Thr61 seen with LAM in binding site II is not present (Figure 3D and Table II).

The K_d of Banlec binding to LAM is 1 mM and to XM is 0.2 mM (Winter *et al*., 2005). The 5-fold tighter binding to XM is interesting in that there are no additional hydrogen bonds to account for the difference. However, studying an overlay of the binding sites provides an explanation. In both binding sites, the O6 hydroxyl on the nonreducing sugar (Glc2) of LAM interferes with the hydrophobic interaction of the GG loop. This shifts the nonreducing Glc2 of LAM up to 1 Å away from Gly14 compared to the position of the Xyl residue of XM. In addition, in binding site I, the methyl group attached to the O1 of the mannose ring in XM is in the axial position allowing the C1 atom and the methyl group to make hydrophobic interactions with the aliphatic carbon atoms of the Lys130 side chain, thus adding strength to XM binding (Figure 3). In LAM, the O1 hydroxyl is in the equatorial position, preventing this interaction.

Comparison of Banlec, Heltuba, and Artocarpin. Previously, the structures of heltuba bound to man-α1,3man and man-α1,2man (Bourne et al., 1999) and artocarpin bound to methyl α-D-mannopyranoside, mannotriose, and mannopentaose (Partap *et al*., 2002; Jeyaprakash *et al*., 2004) were solved. A search for structures similar to Banlec in the protein database using distance matrix alignment (DALI) (Holm and Sander, 1995) failed to bring up these structures; however, pairwise comparison showed the structures to be very similar. The root mean square deviation (rmsd) between the Cα atoms of Banlec and heltuba is 1.6 Å with a *z*-score of 22.6 and of Banlec and artocarpin is 1.7 Å with a *z*-score of 22.5.

The carbohydrate binding sites in heltuba and artocarpin are equivalent to binding site I of Banlec located on the apex of the first face of the β-prism (Figure 1). A major difference between Banlec, heltuba, and artocarpin is the orientation of the bound sugars. Banlec binds to the reducing unit of the disaccharide with the nonreducing sugar extending out over the conserved GG containing loop of the binding site (Figure 4). Both heltuba and artocarpin bind to the nonreducing end of oligosaccharide with the reducing end extending out of the binding pocket and rotated greater than 90° with respect to the disaccharide bound to Banlec (Figure 4). Despite these differences, the binding of all three proteins to the first sugar residue is remarkably similar.

A detailed comparison of these structures explains the unique carbohydrate binding properties of Banlec. The ligand recognition loop, connecting strands β7 and β8 on the third face of the β-prism is one of the most important determinants of specificity for sugar binding (Figure 4C) (Jeyaprakash *et al*., 2004). Its length and composition may be responsible for ligand orientation. In artocarpin, this loop is four amino acids longer than Banlec and heltuba allowing Ala90 and Thr91 to interact with the second α-linked sugar residue via both hydrogen bonds and van der Waals interactions. In heltuba, this loop can still hydrophobically interact with the second residue through the large side chain of Met92. In Banlec, the ligand recognition loop contains amino acids with smaller side chains (GAV) and thus cannot interact with the sugar nor attract it into the orientation seen for heltuba and artocarpin. In addition, the second Asp residue in the ligand binding loop of heltuba (Asp136) and artocarpin (Asp138) helps stabilize the second sugar residue, whereas there is a lysine in the same position of Banlec that helps stabilize the binding of the reducing end of the sugar by hydrogen bonding via water molecules. In the absence of β7–β8 loop interactions, the nonreducing end of the sugar lies across the conserved GG loop stabilized by a hydrogen bond formed between the oxygen atom of β-linked glycosidic bond and the amide nitrogen of Gly15. This orientation appears to be the most energetically favorable conformation for β-linked sugars. A model with the nonreducing unit of LAM (β-1,3-linked glucose) inserted into the binding pocket of either binding site produced steric clashes with the side chain of the residue adjacent to the glycine in the G-X_3-D motif. However, models of internal α-1,3- or α-1,4-linked glucosyl residues, with the nonreducing end in the binding pocket, show that no steric hindrance would occur if the sugar binds in the

Table II. Nonbonded interactions between Xyl-β1,3 Man-α-*O*-Methyl atoms and banana lectin residues

Sugar	Binding site I		Binding site II	
	Protein atom	Distance (Å)[a]	Protein atoms	Distance (Å)[a]
Man O1-methyl	K130 Cδ	4.05		
Man C1	K130 Cβ	3.72		
Man O2	WAT	3.47	WAT	2.75
Man O4	G15 NH	3.06	G60 NH	2.95
Man O4	D133 Oδ1	3.70	D38 Oδ1	2.66
Man O4			D38 Oδ2	3.05
Man C4	G15 Cα	3.94		
Man O5	K130 NH	3.06	D35 NH	3.17
Man O6			G 34 NH	3.21
Man O6	K130 NH	2.96	D35 NH	2.97
Man O6	F131 NH	2.80	V36 NH	3.17
Man O6	D133 Oδ2	2.84	D38 Oδ2	3.00
Man C6	V88 Cγ1	4.03	V36 Cγ1	3.95
Man C6	F131 Cδ1	3.88	F131 Cε2	3.56
O (β1,3)	G15 NH	3.20	G60 NH	3.30
Xyl O2	WAT	2.95	WAT	3.12
Xyl C2	G15 Cα	4.19	G60 Cα	4.10
Xyl O5	G15 NH	3.66	G60 NH	3.27

[a]Distances represent the average between the binding sites in the A chain and the B chain.

same orientation as seen in heltuba and artocarpin (data not shown).

Binding site II of Banlec is composed of the GG loop and G-X_3-D ligand binding loop on the second face of the β-prism. Neither heltuba nor artocarpin contain these required motifs on this face of the structure. Binding site II of Banlec (GDVVD) is more similar in sequence to the residues in the binding sites of both heltuba (GDVLD) and artocarpin (GDLLD) than binding site I of Banlec (GKFLD). The third loop of binding site I, the ligand recognition loop, is absent in the second binding site of Banlec.

Identification of secondary carbohydrate binding sites in family members. A BLAST search of the database reveals that there are several other mJRL family members and hypothetical proteins from a variety of plants that could potentially contain a second carbohydrate binding site as seen in the structure of Banlec (Figure 5) (Terwilliger and Berendzen, 1996). The amino acid sequence of these proteins show the presence of the conserved GG motif on the predicted β5–β6 loop and the G-X_3-D on the predicted β3–β4 loop that make up the second binding site in Banlec (Figure 5). These proteins include the C-terminal domain of a putative 32.7 kDa jasomate-induced protein from *Hordeum vulgare* (AAA87042) (Lee *et al*., 1996), a root-specific rice gene GOS9 from *Oryza sativa* (CAA36189) (de Pater and Schilperoort, 1992), and a hypothetical protein from *Arabidopsis thaliana* (AAN15429).

The structural determination of Banlec complexed with two different carbohydrate molecules (LAM and XM) has provided new insights into the mode of carbohydrate binding for Banlec and the jacalin family of lectins. First, the structure of Banlec bound to reducing 3-*O*-β-linked sugar molecules led to the discovery of a second carbohydrate binding site on Banlec and the definition of signature motifs (Gly–Gly and G-X_3-D) that can be used to locate additional binding sites on other lectins. Secondly, the Banlec structures showed that carbohydrates with 3-*O*-β-linked reducing sugar molecules are recognized by the conserved Gly–Gly motif and finally, the Banlec : carbohydrate structures showed that although the bound carbohydrates may bind with alternate orientations with respect to their length and composition, the saccharide unit that binds to the conserved Asp in the binding pocket forms similar protein interactions whether it is the reducing or nonreducing end of the sugar.

Materials and methods

Materials

Banlec was prepared as previously described (Peumans *et al*., 2000) and dissolved in 10 mM HEPES, pH 7.4 and 150 mM NaCl. LAM (Glcβ1,3Glc) was purchased from Sigma (St. Louis, MO). XM was synthesized for a previous study (Winter *et al*., in press).

Crystallization and data collection

Large, data quality crystals of Banlec were grown from hanging drop vapor diffusion experiments by mixing 2 μL of protein (6 mg/mL) in 10 mM HEPES pH 7.5, 150 mM

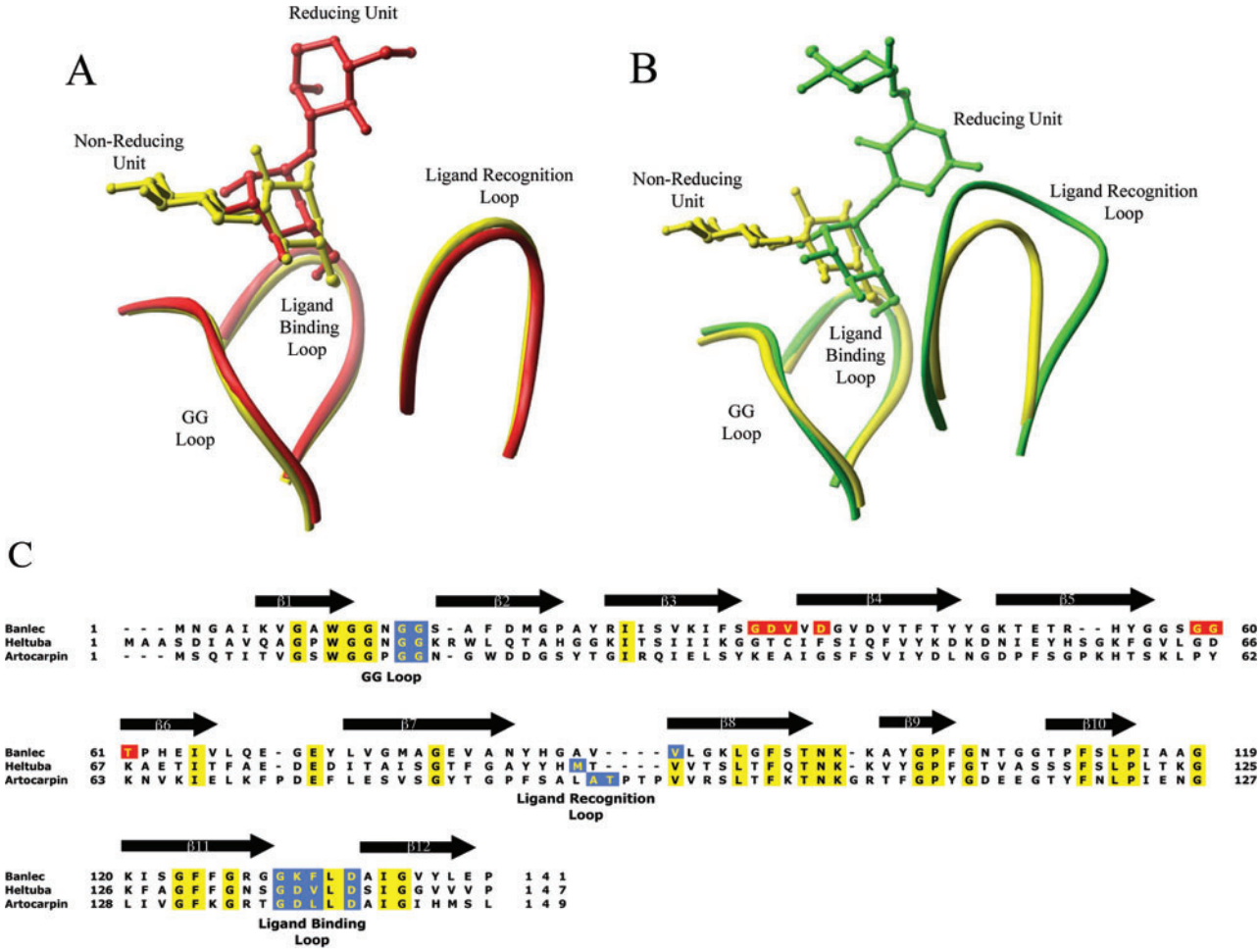

Fig. 4. Carbohydrate binding comparison of banana lectin (Banlec) to heltuba and artocarpin. Ribbon diagram of Banlec bound to laminaribiose (LAM) shown in yellow, heltuba bound to Manα1, 3Man (**A**) shown in red and artocarpin bound to trimannose (**B**) shown in green. The GG loop, ligand binding loop, and the ligand recognition loop are labeled. (**C**) Sequence alignment of banana lectin (Banlec), heltuba, and artocarpin. Banlec, heltuba, and artocarpin are three mannose-specific jacalin-related lectins (mJRL) whose structures have been solved bound to ligands. The residues in the first sugar binding site (I) are shaded blue. The residues in the second sugar binding site (II) are shaded red. The residues that are identical shaded yellow. Sequences were aligned with Clustal W(Thompson *et al*., 1994).

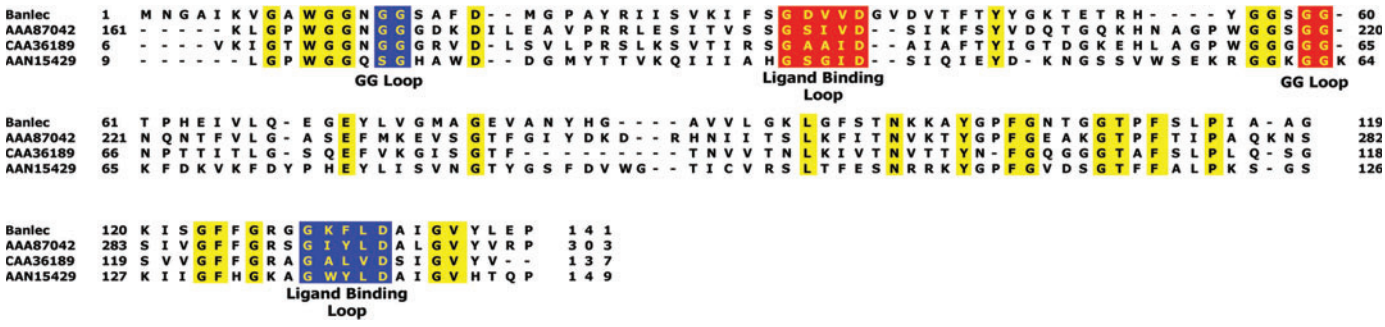

Fig. 5. Sequence alignment of mannose-specific jacalin-related lectins (mJRL) family members with a second potential carbohydrate binding site. Sequence of banana lectin (Banlec), AAA87042 from *Hordeum vulgare*, CAA36189 from *Oryza sativa*, and AAM46813 from *Arabidopsis thaliana*. Binding site one (I) is shaded blue, binding site two (II) is shaded red, and residues that are identical are shaded yellow. Sequences were aligned with Clustal W(Thompson *et al*., 1994).

Table III. Data collection, phasing and refinement statistics

Data collection				
Data set	$K_2O_sO_4$	Native	LAM	XM
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Unit cell	a = b = 81.15 Å, c = 147.28 Å	a = b = 81.54 Å, c = 147.15 Å	a = b = 81.8 Å, c = 147.19 Å	a = b = 81.623 Å, c = 146.815 Å
Wavelength (Å)	1.14	1.0332	1.0	1.2834
Resolution (Å)[a]	2.4 (2.4–2.5)	2.5 (2.50–2.59)	2.8 (2.8–2.9)	2.4 (2.40–2.49)
R_{sym} (%)[b]	6.2 (27.8)	3.8 (11.9)	9.4 (34.6)	6.3 (23.0)
<I/I>[c]	20 (5)	20 (10)	10 (3)	20 (5)
Completeness (%)[d]	97.8 (99.3)	99.7 (100)	99.9 (100)	99.8 (100)
Redundancy	12 (12)	9 (9)	5 (5)	9 (9)
Phase Resolution (Å)[e]	3.5			
Rcullis Anomalous[f]	0.682			
Ppano[g]	1.6			
Sites[h]	6			
Refinement				
Resolution (Å)		2.5	2.8	2.4
R-Factor (%)[i]		24.8	20.5	22.8
Rfree (%)[j]		28.3	26.2	25.7
Protein atoms		2046	2046	2042
Carbohydrate		0	92	66
Water molecules		130	196	222
Unique reflections		19529	25176	21764
R.m.s.d.[k]				
Bonds (Å)		0.007	0.007	0.007
Angles (°)		1.5	1.5	1.4

LAM, laminaribiose; XM, Xyl-β1,3 Man-α-*O*-Methyl.
[a]Statistics for highest resolution bin of reflections in parentheses.
[b]$R_{sym} = \Sigma_h\Sigma_j | I_{hj} - \langle I_h \rangle | / \Sigma_h\Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $\langle I_h \rangle$ is the mean intensity for multiply recorded reflections.
[c]Intensity signal-to-noise ratio.
[d]Completeness of the unique diffraction data.
[e]Resolution cut-off used during heavy-atom refinement and phase calculations.
[f]$R_{cullis} = \Sigma_h | I\, IF_{ph} - F_p I - F_h I | / \Sigma_h IF_{ph} - F_p I$ for centric data, where F_p and F_{ph} are the native and heavy-atom structure for factor amplitudes, respectively. The MLPHARE program from CCP4 slightly modifies this equation for acentric and anomalous data.
[g]Phasing power for isomorphic derivatives (PP_{iso}) = $\langle IF_h I \rangle / \langle E \rangle$, where F_h is the heavy-atom structure factor amplitude and E is the residual lack of closure error. Phasing power for anomalous-scattering derivatives (PP_{ano}) = $\langle 2IF_H' I \rangle / \langle E \rangle$, where F_H' is the calculated anomalous-scattering structure factor.
[h]Number of heavy-atom sites per asymmetric unit.
[i]*R*-factor = $\Sigma_h\, I\, IF_o I - IF_c I\, I / \Sigma_h IF_o I$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.
[j]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement.
[k]Root mean square deviation of bond lengths and bond angles

NaCl with 2 μL of precipitant (1.173–1.275 M ammonium sulfate, 30 mM cadmium chloride, 0.1 M Tris–HCl pH 8.25). Increasing the concentration of $CdCl_2$ in the crystallization buffer triggered the conversion of thin rod-like crystals into large well-diffracting bipyramidal crystals. The crystals grew to an average size of $0.2 \times 0.2 \times 0.3$ mm^3 over the course of 14 days at 20°C in the trigonal space group P3221 with 2 molecules in the asymmetric unit resulting in a large solvent content of 74% (Matthews, 1968).

The large solvent channels allowed ample room for diffusing in of carbohydrate ligands negating the need for cocrystallization. Carbohydrate bound protein crystals were produced by soaking native crystals overnight in a stabilizing solution of 1.275 M ammonium sulfate, 100 mM Tris–HCl pH 8.25, 30 mM cadmium chloride, and 500 mM lithium sulfate containing either 8.4 mM LAM or 3.6 mM XM. Prior to data collection, native and ligand bound crystals were equilibrated with cryoprotectant containing 2.5 M lithium sulfate, 0.1 M Tris-HCl pH 8.25 and 30 mM cadmium chloride then flash frozen in liquid nitrogen.

Diffraction data for the native crystals were collected at the BioCARS 14-ID-B station at the Advance Photon Source equipped with a Mar165 CCD detector and processed with HKL2000 (Table III) (Otwinowski and Minor,

1997). Banlec : carbohydrate complexes and single-wavelength anomalous diffraction data were collected at COMCAT 32-ID-B at the Advanced Photon Source equipped with a MAR165 CCD and processed with DENZO/SCALEPACK (Otwinowski and Minor, 1997).

Structure determination and refinement

Single-wavelength anomalous diffraction methods (Hendrickson and Ogata, 1997) were employed to solve the Banlec structure. Attempts to solve the structure by molecular replacement methods (Rossman and Blow, 1962) using the monomeric structure of Heltuba (pdb code:1C3M) as a search model were unsuccessful. Prior to data collection, native crystals of Banlec were soaked for 12–24 h in stabilizing solution (1.275 M ammonium sulfate, 30 mM cadmium chloride, 0.1 M Tris pH 8.25, and 0.5 M lithium sulfate) containing 2 mM of the heavy atom compound, potassium osmiate, then briefly back soaked in cryosolvent and cooled to 100 K in a liquid nitrogen stream. Two initial heavy atom sites were found in SOLVE (Terwilliger, 2004) using data collected at the osmium x-ray absorption edge (λ = 1.14 Å). The two heavy atom sites were refined in SHARP (La Fortelle and Bricogne, 1997). Difference fourier maps calculated in SHARP showed the existence of four additional heavy atom sites. Heavy atom refinement and density modification using DM with solvent flipping in SHARP produced electron density maps with a figure of merit of 0.8. Although the heavy atom phasing power fell off at 3.5 Å resolution, 90% of the protein model could be fit into the initial electron density map using O (Jones *et al*., 1991). Phase combination and further structural refinement to 2.5 Å was completed using CNS (Table III) (Brunger *et al*., 1998).

Difference fourier maps (*Fo – Fc*) calculated after CNS rigid-body refinement of the native protein against Banlec : LAM and Banlec : XM diffraction data showed the presence of two disaccharide molecules per monomer in each case (Figure 2). Models of the disaccharides were built using O and refined in CNS. Multiple rounds of CNS refinement and model building resulted in a Banlec : LAM structure refined to 2.8 Å resolution and a Banlec : XM structure refined to 2.4 Å resolution (Table III). Simulated annealing omit maps calculated in CNS were used to check for the correct placement of atoms in the structure. The final structures were analyzed in PROCHECK (Laskowski *et al*., 1993). All residues are in the allowed regions of the Ramanchandran plot with the exception of Ala9, which is found in a disallowed region. This distortion is a result of the side chain of neighboring Trp10 forming stacking interactions with Phe112 from an adjacent β-strand. Residues 1–3 on the B chain were disordered in all three structures, but the A chain however, showed clear density for these residues as a result of crystal packing.

Acknowledgments

The work described in the article was supported by a grant from the Michigan Economic Development Corporation for the Michigan Life Sciences Corridor Initiative (J.A.S) and NIH grant GM 29477 (IJG). We thank K. Pattridge for careful reading of the manuscript. We thank the Advanced Photon Source at Argonne National Laboratories, Illinois and the staff at BioCARS-14-ID-B and J. Brunzelle at COMCAT-32-ID-B for their help with data collection. Coordinates were submitted to the PDB with the following codes: 2BMY (Banlec), 2BMZ (Banlec : LAM), and 2BN0 (Banlec : XM).

Abbreviations

Banlec, banana lectin; gJRL, galactose-specific jacalin-related lectins; JRL, jacalin-related lectin; LAM, laminaribiose (Glcβ1,3Glc); mJRL, mannose-specific jacalin-related lectins; XM, Xyl-β1,3 Man-α-O-Methyl.

Note added in proof

The structure of the banana lectin from *M. Paradisiaca* is reported in the preceding paper in this issue of *Glycobiology* [Singh, D.D., Saikrishnan, K., Kumar, P., Surolia, A., Sekar K., and Vijayan, M. (2005). Unusual sugar specificity of banana lectin from *Musa paradisiaca* and its probable evolutionary origin. Crystallographic and modelling studies. *Glycobiology*, **15**, 1025–1032].

References

Bourne, Y., Roig-Zamboni, V., Barre, A., Peumans, W.J., Astoul, C.H., Van Dammer, E.J.M., and 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.

Bourne, Y., Zamboni, V., Barre, A., Peumans, W.J., Van Damme, E.J.M., and Rouge, P. (1999) Helianthus tuberosus lectin reveals a widespread scaffold for mannose-binding lectins. *Structure*, **7**, 1473–1482.

Brunger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., and others. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54,** 905–921.

Chen, C., Rowley, A.F., and Ratcliffe, N.A. (1998) Detection, purification by immunoaffinity chromatography and properties of β-1,3-glucan-specific lectins from the sera of several insect species. *Insect Biochem. Mol. Biol*., **28**, 721–731.

Davic, B. and Soderhall, K. (1990) Purification and characterization of a beta-1,3-glucan binding protein from plasma of the crayfish Pacifastacus leniusculus. *J. Biol. Chem*., **265**, 9327–9332.

Goldstein, I.J., Winter, H.C., Mo, H., Misaki, A., Van Damme, E.J.M., and Peumans, W.J. (2001) Carbohydrate-binding properties of the banana (Musa acuminata) lectin. II. Binding of laminaribiose oligosaccharide and β-glucans containing β1,6-linked glucosyl end groups. *Eur. J. Biochem*., **268**, 2616–2619.

Hendrickson, W.A. and Ogata, C.M. (1997) Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol*., **276**, 494–523.

Holm, L. and Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem. Sci*., **20**, 478–480.

Jeyaprakash, A.A., Srivastav, A., Surolia, A., and Viyanan, M. (2004) Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. *J. Mol. Biol*., **338**, 757–770.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density

maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Koshte, V.L., van Dijk, W., van der Stelt, M.E., and Aalberse, R.C. (1990) Isolation and characterization of BanLec-I, a mannoside-binding lectin from Musa paradisiac (banana). *Biochem. J.*, **272**, 721–726.

La Fortelle, E. and de and Bricogne, G. (1997). In Sweet, R.M. and Carter, C.W., Jr (eds.), *Methods in Enzymology, Macromolecular Crystallography*. Academic Press, New York, Vol. 276, pp. 472–494.

Lammerts van Bueren, A., Morland, C., Gilberts, H.J., and Boraston, A.B. (2005) Family 6 carbohydrate binding modules recognize the non-reducing end of β-1,3-linked glucans by presenting a unique ligand binding surface. *J. Biol. Chem.*, **280**, 530–537.

Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 91–97.

Lee, J., Parthier, B., and Lobler, M. (1996) Jasmonate signaling can be uncoupled from abscisic acid signaling in barley: identification of jasmonate-regulated transcripts which are not induced by abscisic acid. *Planta*, **199**, 625–632.

Lee, X., Thompson, A., Zhang, Z., Ton-that, H., Biesterfeldt, J., Ogata, C., Xu, L., Johnson, R.A.Z., and Martin Young, N. (1998) Structure of the complex of Maclura pomifera agglutinin and the T-antigen disaccharide, Galbeta1,3GalNAc. *J. Biol. Chem.*, **273**, 6312–6318.

Matthews, B. (1968) Solvent content of protein crystals. *J. Mol. Biol.*, **33**, 491–497.

Mo, H., Winter, H.C., Van Damme, E.J.M., Peumans, W.J., Misaki, A., and Goldstein, I.J. (2001) Carbohydrate-binding properties of the banana (Musa acuminata) lectin. I. Novel recognition of internal α1,3-linked glucosyl residues. *Eur. J. Biochem.*, **268**, 2609–2615.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Partap, J.V., Jeyaprakash, A.A., Rani, P.G., Sekar, K., Surolia, A., and Vijayan, M. (2002) Crystal structures of artocarpin, a Moraceae lectin with mannose specificity, and its complex with methyl-alpha-D-mannose: implications to the generation of carbohydrate specificity. *J. Mol. Biol.*, **317**, 237–247.

de Pater B.S., and Schilperoort, R.A. (1992) Structure and expression of a root-specific rice gene. *Plant Mol. Biol.*, **18**, 161–164.

Peumans, W.J., Zhang, W., Barre, A., Houles-Astoul, C., Balint-Kurti, P.L., Rovira, P., Rouge, P., May, G.D., Van Leuven, F., Truffa-Bachi, P., and Van Damme, E.J. (2000) Fruit-specific lectins from banana and plantain. *Planta*, **211**, 546–554.

Raval, S., Gowda, S.B., Singh, D.D., and Chandra, N.R. (2004) A database analysis of jacalin-like lectins: sequence-structure-function relationships. *Glycobiology*, **14**, 1247–1263.

Richardson, J.S. (1981) The anatomy and taxonomy of protein structure. *Adv. Protein Chem.*, **34**, 167–339.

Rossman, M.G. and Blow, D.M. (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.*, **15**, 24–31.

Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A., and Vijayan, M. (1996) A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a beta-prism fold. *Nat. Struct. Biol.*, **3**, 596–603.

Sharon, N. and Lis, H. (2004) History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology*, **14**, 53R–62R.

Singh, D.D., Saikrishnan, K., Kumar, P., Dauter, Z., Sekar, K., Surolia, A., and Vijayan, M. (2004) Purification, crystallization and preliminary X-ray structure analysis of the banana lectin from Musa paradisiaca. *Acta Crystallogr. D*, **60**, 2104–2106.

Terwilliger, T.C. (2004) SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Rad.*, **11**, 49–52.

Terwilliger, T.C. and Berendzen, J. (1996) Correlated phasing of multiple isomorphous replacement data. *Acta Crystallagr. D*, **52,** 749–757.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.*, **22**, 4673–4680.

Trakhanov, S., Kreimer, D.I., Parkin, S., Ames, G.F.-L., and Rupp, B. (1998) Cadmium-induced crystallization of proteins. II. Crystallization of the Salmonella typhimurium histidine-binding protein in complex with L-histidine, L-arginine, or L-lysine. *Protein Sci.*, **7**, 600–604.

Van Bueren A.L., Morland, C., Gilbert, H.J., and Boraston, A.B. (2005) Family 6 carbohydrate binding modules recognize the non-reducing end of β-1,3-linked glucans by presenting a unique ligand binding surface. *J. Biol. Chem.*, **280**, 530–537.

Winter, H., Oscarson, S., Slattegard, R., Tian, M., and Goldstein, I.G. (2005) Banana lectin recognizes the reducing unit of, 3-*O*-β-Glucosyl/Mannosyl Disaccharides. A Calorimetric Study. *Glycobiology*, **15**, 1043–1050.

Yang, H. and Czapla, T.H. (1993) Isolation and characterization of cDNA clones encoding jacalin isolectins. *J. Biol. Chem.*, **268**, 5905–5910.