Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure and sugar-specificity of basic winged-bean lectin: structures of new disaccharide complexes and a comparative study with other known disaccharide complexes of the lectin

Kiran A. Kulkarni,[a] Samiksha Katiyar,[a] Avadhesha Surolia,[a,b] Mamannamana Vijayan[a]* and Kaza Suguna[a]*

[a]Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India, and [b]National Institute of Immunology, New Delhi, India

Correspondence e-mail: mv@mbu.iisc.ernet.in, suguna@mbu.iisc.ernet.in

Received 29 October 2007
Accepted 21 April 2008

PDB References: complex of basic winged-bean agglutinin with Galα1-4Gal, 2zml, r2zmlsf; with Galα1-6Glc, 2zmn, r2zmnsf; with Galα1-4Galβ-Et, 2zmk, r2zmksf.

Crystal structures of the complexes of basic winged-bean agglutinin with the disaccharides Galα1-4Gal (galabiose), Galα1-6Glc (mellibiose) and Galα1-4Galβ-Et have been determined and the complex with Galα1-2Gal has been modelled. The interactions of the nonreducing Gal with the lectin at the primary site are the same as those in the known complexes with disaccharides having the $\alpha1\rightarrow3$ linkage. The second residue in Galα1-4Gal and Galα1-6Glc forms a water bridge to the lectin, while the ethyl group in Galα1-4Galβ-Et makes nonpolar interactions. In complexes involving disaccharides with α1-3 linkages, which form part of the A and B blood-group substances, the second sugar residue forms a direct hydrogen bond to the variable loop in the binding site of the lectin. This in part explains the specificity of the lectin for the blood-group substances and also the higher affinity of $\alpha1\rightarrow3$-linked disaccharides for the lectin compared with disaccharides involving other linkages. Including those reported here, 14 crystal structures involving the lectin, accounting for 54 crystallographically independent subunits, are available. A comparative study of these structures shows that the region involving the curved β-sheet which nestles the metal ions is relatively rigid. The carbohydrate-binding region is perched on this region. The flat β-sheet, which is involved in oligomerization and exhibits considerable variability in legume lectins, is relatively flexible. Indeed, the structures of basic winged-bean lectin have been of critical importance in establishing legume lectins as a family of proteins in which small alterations in essentially the same tertiary structure lead to large variations in quaternary association. They have also provided a structural explanation of the blood-group specificity of the lectin.

1. Introduction

The recognitive interaction between lectins, which are multivalent carbohydrate-binding proteins, and sugar ligands is known to play a crucial role in a variety of biological processes such as cell–cell communication, host–pathogen interaction, cell proliferation and signalling (Vijayan & Chandra, 1999; Lis & Sharon, 1998; Loris *et al.*, 1998; Rini, 1995). Lectins are found in various kinds of organisms such as plants, animals, bacteria and viruses. The best studied among these are from leguminous plants (http://www.cermav.cnrs.fr/lectines). The homodimeric N-glycosylated basic lectin from winged bean (*Psophocarpus tetragonolobus*; WBAI) is specific for the A blood-group substance and also binds to the B blood-group substance with lower affinity. It does not bind to the O

substance. It has also been shown that the lectin only binds to oligosaccharides in which the second residue is α-linked to the nonreducing Gal at the primary site; a β-linkage leads to unacceptable steric clashes (Prabu *et al.*, 1998). Previously, crystal structures of complexes of WBAI with A and B blood-group disaccharides and trisaccharides (A-disaccharide, GalNAcα1-3Gal; B-disaccharide, Galα1-3Gal; A-trisaccharide, GalNAcα1-3Galβ1-4Glc; B-trisaccharide, Galα1-3Galβ1-4Glc) have been reported from this laboratory (Kulkarni *et al.*, 2007). All of them have α1→3 linkage between the first and the second sugar residues.

The studies mentioned above have provided a clear picture of the interactions of WBAI with α1→3-linked saccharides. The work reported here is aimed at elucidating the interactions of the lectin with disaccharides containing αGal with linkages other than 1→3. With this objective in view, we have determined the structures of WBAI complexed with Galα1-4Gal (galabiose), Galα1-6Glc (mellibiose) and Galα1-4Galβ-Et at 2.65, 2.90 and 2.50 Å resolution, respectively. In addition, the structure of the complex with Galα1-2Gal has been modelled. Furthermore, these structures and those of other WBAI–sugar complexes reported previously (Prabu *et al.*, 1998; Kulkarni *et al.*, 2005, 2006, 2007) provide a rationale for the available thermodynamic data (Puri & Surolia, 1994*a*) on the interaction of WBAI with α-linked disaccharides containing Gal/GalNAc. A comparative study involving them also reveals two different modes of disaccharide binding to the lectin: one involving a direct carbohydrate–hypervariable loop interaction, as observed in the case of sugars with α1→3 linkages, and the other involving a water-mediated interaction between the carbohydrate and the lectin, as found in the current study.

Adhesion of lectins to cell-surface glycans, especially those containing Galα1-*X*Gal (where *X* can be 2, 3, 4 or 6), have attracted considerable attention (Galili, 2005; Mann, 2002) because of their implications in various infectious diseases and immunological functions of cells. Therefore, structural studies on complexes of lectins with Galα1-*X*Gal are also of interest in relation to pathophysiological processes.

Table 1
Data-collection statistics.

Values in parentheses are for the last shell.

	Galα1-4Gal	Galα1-6Glc	Galα1-4Gal-βEt
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters			
a (Å)	157.93	158.06	156.92
b (Å)	91.52	91.76	89.96
c (Å)	73.65	73.57	73.32
Resolution (Å)	30–2.65 (2.74–2.65)	30–2.9 (3.0–2.9)	30–2.50 (2.59–2.50)
No. of observed reflections	174153	164929	395869
No. of unique reflections	30795 (3084)	23768 (2335)	36542 (3195)
Data completeness (%)	97.0 (99.0)	97.6 (98.2)	98.1 (87.1)
R_{merge} (%)	11.9 (46.7)	12.4 (48.9)	13.0 (46.4)
$\langle I/\sigma(I)\rangle$	15.8 (2.4)	7.4 (2.8)	14.8 (1.9)
Solvent content (%)	47.3	47.4	45.8

Table 2
Refinement statistics.

Values in parentheses are for the last shell.

	Galα1-4Gal	Galα1-6Glc	Galα1-4Gal-βEt
No. of reflections used			
Working set	29278	22585	34379
Test set	1468	1154	1693
No. of non-H atoms			
Protein	7251	7284	7235
Sugar	311	355	401
Solvent	184	104	171
Average *B* factors (Å^2)			
Protein	25.7	31.8	43.9
Bound sugar	39.7	49.6	60.9
Solvent	28.6	32.3	49.1
R_{cryst} (%)	20.5 (27.9)	19.3 (28.8)	21.6 (31.6)
R_{free} (%)	24.8 (29.5)	23.7 (30.6)	25.5 (35.6)
R.m.s. deviations from ideal values			
Bond lengths (Å)	0.007	0.011	0.007
Bond angles (°)	1.4	1.5	1.4
Dihedral angles (°)	26.3	26.1	26.1
Improper angles (°)	0.92	0.96	0.9
Ramachandran plot, residues in (%)			
Most favoured regions	87.2	85.3	85.9
Additionally allowed regions	11.3	14.1	12.7
Generously allowed regions	1.5	0.6	1.4
Disallowed regions	0.0	0.0	0.0

2. Materials and methods

2.1. Crystallization and data collection

WBAI was isolated and purified following the previously reported method (Khan *et al.*, 1986; Sankaranarayanan *et al.*, 1993). The sugars were obtained from Sigma Chemicals. Crystals of the complexes were grown using vapour diffusion at room temperature employing the hanging-drop method. A 3 µl drop of 6.5 mg ml^{-1} protein containing a 20–50-fold molar excess of sugar, 6–7%(*w*/*v*) PEG 4000 and 5% 2-propanol in 0.02 *M* phosphate buffer pH 7.4 [with 0.15 *M* NaCl, 0.025%(*w*/*v*) sodium azide] was equilibrated against 500 µl reservoir solution containing 20%(*w*/*v*) PEG 4000 and 15%(*v*/*v*) 2-propanol in the same buffer. The diffraction data from the complexes were recorded at room temperature (293 K) using a MAR345 imaging-plate detector mounted on a Rigaku RU-300 generator. The crystal-to-detector distance was kept at 200 mm. The data were processed using *DENZO* and *SCALEPACK* from the *HKL* suite of programs (Otwinowski & Minor, 1997). *TRUNCATE* from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994) was used to convert the intensities to structure factors.

2.2. Refinement and structure analysis

The crystals obtained were isomorphous to those of the other WBAI–sugar complexes reported previously. Therefore, the coordinates of protein atoms of the WBAI–Me-α-galactose (Me-α-Gal) complex (Prabu *et al.*, 1998; PDB code 1wl1) were used as the starting model for refinement. After 35–50 cycles of rigid-body refinement followed by 130–150 cycles of positional refinement using the program *CNS* (Brünger *et al.*, 1998), *R* and R_{free} were in the ranges 0.25–0.27 and 0.26–0.29, respectively. Subsequently, model building was carried out for

bound sugars, metal ions and a few residues of the N-linked glycans based on $F_o - F_c$ and $2F_o - F_c$ maps. In subsequent steps of refinement, water O atoms were added to the model using peaks in electron density with heights greater than 3.0σ in $F_o - F_c$ maps and 1.0σ in $2F_o - F_c$ maps. Composite OMIT maps (Bhat, 1988) were calculated at various steps to remove the model bias. Bulk-solvent correction and NCS restraints were employed throughout the refinement. Manual model building was performed using the program *O* (Jones *et al.*, 1991). To maintain consistency, the R_{free} set of reflections were selected by copying the set used for the starting model and adding reflections where necessary on account of the better completeness of the present data sets. The structures were validated using *PROCHECK* (Laskowski *et al.*, 1993). The data-collection and refinement statistics are given in Tables 1 and 2. The rigid and flexible regions of the protein were delineated using the *Error-inclusive Structure Comparison and Evaluation Tool* (*ESCET*; Schneider, 2002). Structure superpositions were performed using *ALIGN* (Cohen, 1997).

2.3. Modelling study

Protein atoms, the atoms of the nonreducing Gal moiety and the water O atoms present in the vicinity of the sugar-binding pocket of subunit II of the WBAI–Galα1-4Gal complex were used for docking studies. The structure of Galα1-2Gal was constructed by changing the linkage from $\alpha1\rightarrow4$ to $\alpha1\rightarrow2$. The two glycosidic torsion angles (φ and ψ) about this new linkage were varied independently of each other, using *InsightII*, in steps of 20° to generate 324 conformers. Subsequently, a 5 Å water shell was generated around each model of the complex using *InsightII*. After generating H atoms, all 324 modelled complexes were subjected to conjugate-gradient energy minimization using *CNS* (Brünger *et al.*, 1998). The force field given in the program was used with a dielectric constant of unity. All residues, except those constituting the carbohydrate-binding site, crystallographically observed parts of the sugar and the crystallographically observed water O atoms were harmonically restrained with a force constant of 419 kJ mol^{-1} Å^{-1}. The residues constituting the carbohydrate-binding site were similarly restrained with a force constant of 82 kJ mol^{-1} Å^{-1}. Protein atoms which lie within 10 Å from the sugar atoms were used in the interaction-energy calculations.

3. Results and discussion

3.1. Overall structure of the complexes

Clear and unambiguous electron density for residues 6–241 and up to three sugar residues in each of the N-linked glycans at Asn44 and Asn219 and metal ions (Mn^{2+} and Ca^{2+}) is observed in all the four subunits (I, II, III and IV) constituting two dimers (I/II and III/IV) present in the asymmetric unit of the crystal of the complexes. The tertiary structure of the lectin has the characteristic legume-lectin fold consisting of three β-sheets and several loops arranged in a jelly-roll-like motif. The lectin forms the 'handshake' mode of dimerization (Fig. 1), similar to that observed in *Erythrina corallodendron* lectin (ECorL; Shaanan *et al.*, 1991). The lectin has the same tertiary and the quaternary structure in its saccharide-free (Manoj *et al.*, 1999) and bound forms, indicating that lectin–carbohydrate interaction does not influence the structure of the lectin. The different structures reported here have root-mean-square deviations in the range 0.16–0.28 Å when the C^{α} atoms are superposed.

3.2. Lectin–sugar interactions

The clear electron density for sugars (Fig. 2) found in all the complexes permits a detailed discussion of the lectin–sugar interactions. The known interactions of similar sugars with WBAI and other legume lectins, studied in this laboratory and elsewhere, lend additional confidence to the interpretation of the density illustrated in Fig. 2. As found in other legume lectins, the carbohydrate-binding pocket is a shallow depression on the surface formed by the four loops referred to as *A*, *B*, *C* and *D* (Sharma & Surolia, 1997). It is substantially preformed and largely unaffected by carbohydrate binding. In this paper, these loops have been designated L1, L2, L3 and L4 in order to avoid confusion between the names of the loops and carbohydrates. In WBAI, these four loops are formed by residues 78–87 (L1), 95–117 (L2), 124–134 (L3) and 211–222 (L4). Of these, loop L4 is hypervariable in terms of length and sequence (Sharma & Surolia, 1997). WBAI has a primary and a secondary site. Residues 214–216 of loop L4 contribute to the secondary site, whereas residues form all the four loops make up the primary site (Kulkarni *et al.*, 2006).

3.3. Interactions at the primary site

In all three complexes reported here, the nonreducing Gal moiety occupies the primary site with an identical orientation. The four hydrogen bonds, Asp87 OD1$\cdots$Gal O3, Asp87 OD2$\cdots$Gal O4, Gly105 N$\cdots$Gal O3 and Asn128 ND2$\cdots$Gal O3 (Fig. 3), between sugar and the protein provide the essential geometric framework for proper orientation for this Gal moiety. Furthermore, the side chain of Phe126 stacks against the sugar ring. Apart from these interactions, which are common to all legume lectins, O4 and O6 of Gal form hydrogen bonds to Asp212 N and His84 NE2, respectively. O2 of the Gal moiety forms a direct hydrogen bond to Asn128 ND2 (Fig. 3). These lectin–nonreducing Gal interactions at the primary site occur in all the three complexes reported here as well as in the structures involving galactose at the primary site (Prabu *et al.*, 1998; Kulkarni *et al.*, 2006, 2007) reported previously. A water bridge between O6 and Gln217 occurs in a majority of the subunits.

3.4. Interactions of the second sugar residue

Atom O5 of the reducing Gal of Galα1-4Gal interacts with Asn128 ND2 of the lectin through a water bridge (Fig. 3). The reducing Glc in Galα1-6Glc interacts with the lectin in a similar manner except that O4 is now the atom involved. In Galα1-4Galβ-Et this water bridge is absent. However, loss of this particular interaction appears to be compensated by the

nonbonded interactions of the ethyl group of the sugar with Thr129 CG2. The water molecule which bridges the second sugar and Asn128 ND2 is observed in 45 of 54 subunits in the 14 structures involving the lectin determined so far. The model of the WBAI–Galα1-2Gal complex further affirms the role of this water molecule in sugar binding. O3 of the second sugar of the disccharide could form the water bridge with Asn128 ND2 in the model.

Despite having different type of linkages at C1, the three disaccharides in the present work exhibit similar affinity for the lectin. Compared with the A (GalNAcα1-3Gal) and B (Galα1-3Gal) blood-group disaccharides these sugars have considerably lower affinity for the lectin, but bind slightly better than Gal (Puri & Surolia, 1994*a*). In the A and B blood-group disaccharides the second sugar (reducing Gal) moiety is α1$\rightarrow$3-linked to the nonreducing sugar with an orientation facilitating the formation of a direct interaction between O2 of the second Gal and Ser214 OG (Kulkarni *et al.*, 2007). The same sugar moiety (reducing Gal/Glc) of the complexes reported here has an orientation that is markedly different from those in the A and B blood-group disaccharides and therefore fails to form the Gal O2$\cdots$Ser214 OG direct interaction. Presumably, this is responsible for the higher affinity of WBAI for blood-group disaccharides with α1$\rightarrow$3 linkages than for the sugars presented here. However, the slightly higher affinity of these sugars compared with that of Gal could be attributed to the additional water bridge observed between them and the lectin. Water bridges between protein and sugar ligands are prevalent in many lectin–sugar complexes involving comparable affinity (Christensen & Toone, 2003; Weis & Drickamer, 1996). In at least one instance specificity is generated through a water-mediated interaction (Ravishankar *et al.*, 1997) in legume lectins.

Based on thermodynamic data on WBAI–carbohydrate interactions (Puri & Surolia, 1994*a*), the disaccharide ligands for WBAI can be classified into two categories: one involving α1$\rightarrow$2, α1$\rightarrow$4 or α1$\rightarrow$6 linkages and having nearly the same affinity as that of Gal and the other involving the α1$\rightarrow$3 linkage with affinity greater than or nearly equal to five times that of Gal (Puri & Surolia, 1994*a*). Galα1-4Gal, Galα1-6Glc and Galα1-4Galβ-Et fall into the first category, whereas GalNAcα1-3Gal, Galα1-3Gal and Galα1-3GalαMe belong to the second category. In both categories, binding is largely enthalpy-driven. It is interesting to note that among the legume lectins, loop L4 is the largest in WBAI and has a significant role in generating specificity for the sugars belonging to the second category (Kulkarni *et al.*, 2007). The length and the sequence of this loop play a specificity-conferring role in this case. The ligands involved in the crystal structures reported here belong to the first category. The second sugar residue in them interacts with the lectin through a water

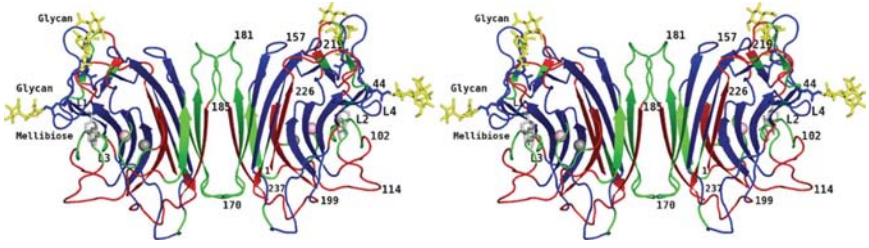

Figure 1
A stereoview of the dimer of WBAI complexed with mellibiose. Ca^{2+} and Mn^{2+} are shown as spheres. Mellibiose and N-linked glycans are shown in sticks. Rigid, flexible and most flexible regions of the molecule are shown in blue, green and red, respectively. See text for details. The carbohydrate-binding loops have been labelled. Figs. 1 and 2 were generated using *PyMOL* (DeLano, 2002).

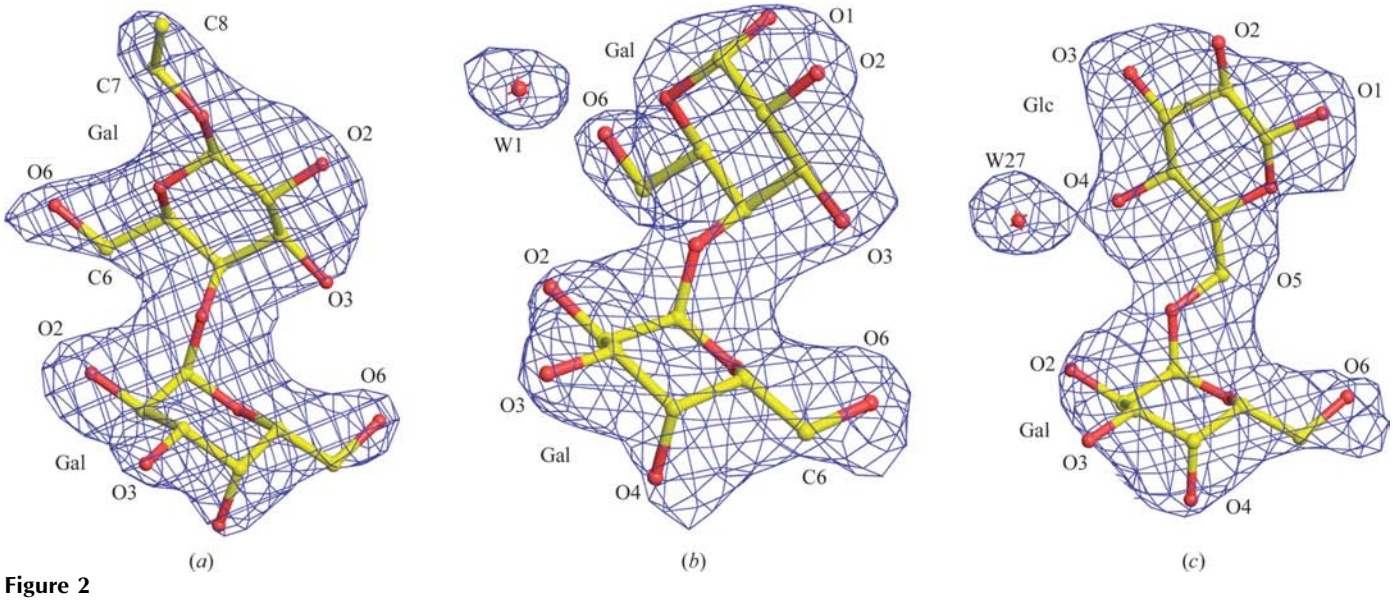

Figure 2
Electron density ($2F_o - F_c$ map) for (*a*) Galα1-4Gal, (*b*) Galα1-6Glc, (*c*) Galα1-4Galβ-Et and bridging water O atom contoured at 1σ. Some sugar atoms are numbered.

bridge. Clearly, WBAI exhibits two distinct ways of disaccharide binding: direct interaction between sugar and loop L4 or a water bridge between the sugar and Asn128.

3.5. Lectin–glycoprotein interactions

Apart from the use of lectins as serological agents for blood typing, they have recently also been used as effective probes to explore many cell-surface glycoproteins (Goldstein *et al.*, 1997; Sharon & Lis, 2003). With the availability of a large amount of structural and biochemical data on WBAI, it is worthwhile investigating the prospects for WBAI–glycoprotein interaction. Carbohydrate structural units containing Gal/GalNAc have been classified into 13 different groups. They are A, GalNAcα1-3Gal; B, Galα1-3Gal; F, GalNAcα1-3GalNAc; Tn, GalNAcα1-Ser(Thr); E, Galα1-4Gal; M, Galα1-6Glc; X, Galα1-2Gal; T, Galβ1-3GalNAc; I, Galβ1-3GlcNAc; II, Galβ1-4GlcNAc; L, Galβ1-4Glc; S, GalNAcβ1-4Gal; P, Galβ1-3Gal (Wu, 2003; Ahuja, 1985). Of these, WBAI does not bind to T-, I-, II-, L-, S- and P-type substances as all of them have a β-linkage involving sugar at the primary site which leads to severe steric contacts with the lectin (Prabu *et al.*, 1998). Previous structural studies on WBAI have indicated that the lectin binds to A, B and Tn with significantly higher affinity than to Gal/GalNAc (Kulkarni *et al.*, 2005, 2007). The present crystallographic and modelling study indicates the possibility of interaction of the lectin with α Gal containing structural units of type E, M and X through a water bridge. Furthermore, in order to explore the possibility of binding of F-type reactive substances, an acetamido group was attached by modelling to C2 of the Gal moiety of GalNAcα1-3Gal of subunit II of the crystal structure of the WBAI–GalNAcα1-3Gal complex (Kulkarni *et al.*, 2007). The attached acetamido group was rotated around the N2—C7 bond by changing the torsion angle χ(C2—N2—C7—O7) in steps of 20°. The resulting 18 conformations were subjected to energy minimization following the procedure used for modelling the WBAI—Galα1-2Gal complex. The galactopyranose ring of the second sugar in the model with minimum energy has a slightly different orientation from that in other α1→3-linked complexes. In the model, the Gal O2···Ser214 OG interaction observed in the crystal structure is replaced by a GalNAc O7···Ser214 OG interaction. This exercise hints at the possibility of WBAI binding Fossman disaccharide (F).

3.6. Structural plasticity of WBAI

The plasticity of protein molecules has been explored extensively in this laboratory (Madhusudan & Vijayan, 1991; Radha Kishan *et al.*, 1995; Sadasivan *et al.*, 1998; Biswal *et al.*, 2000; Natchiar *et al.*, 2006). The number of available WBAI structures elucidated from crystals grown under different conditions and in the presence of different sugars and in their absence is large enough to attempt a similar exercise in the present case. The data set used for this purpose had 54 subunits corresponding to 14 different crystal structures involving WBAI.

The parameter σ in the *Error-inclusive Structure Comparison and Evaluation Tool* (*ESCET*; Schneider, 2002) is used to delineate rigid and flexible regions of the subunit. This parameter is calculated from the error estimate in the coordinates of the structures being compared. At 0.59σ as the lower limit, nearly half (115) of the total of 236 defined residues in the subunit fall into rigid regions and the remainder fall into flexible regions. At the 0.69σ level, only 46 residues fall into flexible regions. These 46 may be designated as belonging to the most flexible regions. The rigid, flexible and most flexible regions defined in this manner are indicated in Fig. 1. As expected, the *B* values conform to this classification. For example, the *B* values in the rigid, flexible and most flexible regions are 25.6, 28.8 and 36.3 Å^2, respectively, in the Galα1-6Glc complex.

The general architecture of the WBAI monomer consists of three β-sheets at the core: a six-stranded back β-sheet made up of residues 1–7, 67–73, 161–167, 172–178, 186–192 and 228–235, a seven-stranded curved β-sheet consisting of residues 17–22, 45–52, 88–94, 118–125, 135–142, 148–152 and 203–210, and a smaller third β-sheet involving residues 24–26, 30–33, 74–77, 159–160 and 223–226 which connects the first two sheets. The remainder of the residues, accounting for nearly half the chain, belong to loops. The present analysis shows that 33 residues of the 46

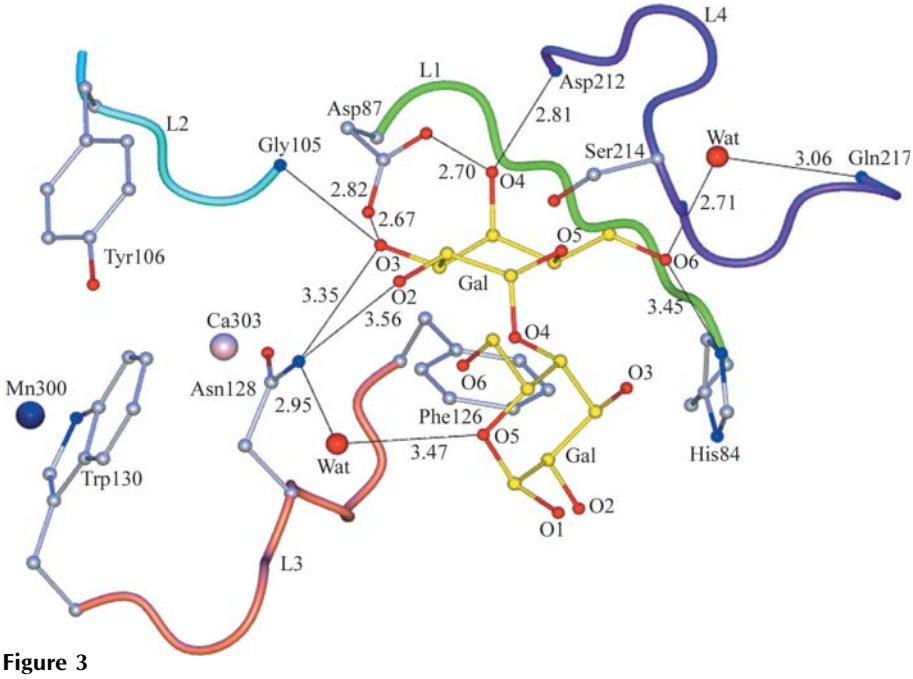

Figure 3
WBAI–Galα1-4Gal hydrogen-bonding interactions. Loops at the binding site are shown as coils. This figure was generated using *MOLMOL* (Koradi *et al.*, 1996).

of the front curved β-sheet, which nestles the metal ion-binding sites, are rigid. The fifth (172–178) and the sixth (186–192) strands in the flat β-sheet which are involved in the dimeric interface are found to be most flexible. Legume lectins are known to exhibit considerable variability in quaternary association (Prabu *et al.*, 1998; Vijayan & Chandra, 1999) and the flexibility of these strands is perhaps essential to accommodate this variability. The situation is more or less similar to what was observed in the case of peanut lectin (PNA; Natchiar *et al.*, 2006).

At the carbohydrate-binding site, out of four loops, loop L1 was found to be rigid and loops L2 and L3 fall into the flexible region of the lectin. Surprisingly, in contrast to the situation in PNA, loop L4, which is variable in length and composition among the legume lectins, is found to be rigid. In WBAI, the residues from this loop interact with symmetry-related molecules in the crystal structures of the complexes. Also, these residues make direct interactions with blood-group oligosaccharides (Kulkarni et al., 2007). These intramolecular and intermolecular interactions might be responsible for the rigidity of this loop.

4. Summary and conclusions

Including the results presented here, 14 crystal structures of WBAI in the free form and in complex with different sugars have been reported, all from this laboratory. This structural work has also been supported to a substantial degree by thermodynamic measurements. It is now perhaps appropriate to examine the new information and insights gained from these investigations on the structure and function of the lectin.

An important motivation for initiating work on WBAI was to do with investigating the quaternary structure of legume lectins. In the early 1990s, it was believed that legume lectins have an intrinsic propensity to form canonical dimers of the type found in ConA (Hardman & Anisworth, 1972; Reeke *et al.*, 1975) in which the two six-stranded back β-sheets from the two subunits come together to form a contiguous 12-membered antiparallel sheet. However, ECorL exhibited a different 'handshake' mode of dimerization in its structure reported in 1991 (Shaanan *et al.*, 1991). ECorL has a glycosylation site at the interface between subunits in the canonical dimer. Thus, it appeared that departure from the canonical mode of dimerization in legume lectins occurs on account of interference by other factors such as steric interactions caused by covalently linked sugar. The tetrameric peanut lectin molecule, the structure of which was solved in this laboratory in 1994, is not glycosylated (Banerjee *et al.*, 1994). However, the dimer in it is not canonical; the subunits in it have a back-to-back arrangement of back β-sheets. Thus, it appeared that the variability in quaternary association is caused by factors intrinsic to the protein and not by others such as interactions involving covalently linked sugar. WBAI is highly homologous to ECorL in sequence (63% identity; Puri & Surolia, 1994*b*). However, unlike ECorL, it does not have a glycosylation site in the region corresponding to the interface in the canonical dimer. Therefore, if oligomerization is influenced only by glycosylation, it is possible that WBAI can form a canonical dimer. However, if the mode of dimerization is primarily dictated by factors intrinsic to the protein itself, WBAI is likely to form an ECorL-type dimer. In the event, structure analysis showed that WBAI invariably formed an ECorL-type dimer (Prabu *et al.*, 1998; Manoj *et al.*, 1999). Thus, X-ray studies of WBAI contributed substantially in establishing that legume lectins are a family of proteins in which small alterations in essentially the same tertiary structure lead to large variations in quaternary association.

The role of glycosylation in WBAI has not been investigated structurally. In the lectin, the plant-specific heptasaccharide [Manα6(Manα3)(Xylβ2)Manβ4GlcNAcβ4-(LFucα3)GlcNAcβ] is N-linked to Asn44 and Asn219. The entire heptasccharide is not defined in any of the structures. The trisccharide GlcNAcβ4(LFucα3)GlcNAcβ is defined in a majority of them. The largest defined fragment, which occurs in a few, is the tetrasaccharide Manβ4GlcNAcβ4(LFucα3)-GlcNAcβ. The observed N-linked sugar fragments do not suggest any specific structural role. It may be that glycosylation affects the folding and stability of WBAI but not the structure, as in the case of its close homologue ECorL (Srinivas *et al.*, 2001; Kulkarni *et al.*, 2004).

Structural studies on WBAI also contributed to the elucidation of the nuances of carbohydrate recognition by lectins. The lectin exhibits specificity for Gal/GalNAc at the monosaccharide level and binds to GalNAc with nearly five times higher affinity than Gal. The structures of WBAI in complex with Gal, Me-α-Gal, 2Me-O-Gal, GalNAc and Me-α-GalNAc demonstrate that all of them bind to the lectin at the primary site with an identical orientation. The additional hydrophobic interactions of the acetamido group of GalNAc with the lectin are probably responsible for the higher affinity of GalNAc compared with that of Gal (Kulkarni *et al.*, 2006). Attachment of an α-methyl group to Gal or GalNAc further enhances the affinity of the ligand by generating nonbonded interactions with Gly215 of the hypervariable loop (L4). Among these monosaccharide ligands, Me-α-GalNAc makes interactions involving both acetamido group as well as the α-linked methyl group. Therefore, this sugar has the highest affinity among the monosaccharides. The monosaccharide complexes also showed that β-linkage in sugars at the primary site leads to a steric clash with the hypervariable loop (L4) of the lectin. Hence, the protein is specific for α-linked sugars. Compared with the structures of other Gal/GalNAc-specific legume lectins, this particular loop is the longest in WBAI and the presence of a proline residue (Pro213) at the beginning of the loop makes it a rigid lid-like structure. In most of the Gal/GalNAc-specific legume lectins the affinities for α and β anomeric sugars are comparable. However, in WBAI the steric contacts posed by the lid-like hypervariable loop makes it specific for α-linked sugars. Thus, WBAI provides an instance in which the carbohydrate specificity is generated by loop length (Prabu *et al.*, 1998).

Knowing that WBAI binds to GalNAc with higher affinity than to Gal, interactions of WBAI with Tn-determinant (GalNAc-α-O-Ser/Thr) were explored (Kulkarni *et al.*, 2005),

in view of its role in the diagnosis and prognosis of human malignancy. Tn-determinant is a human specific tumour-associated carbohydrate antigen and is suggested to be a potential diagnostic and therapeutic tool for various cancers. In the complex of the antigen with WBAI, apart from preserving the observed lectin–GalNAc interactions, the antigen makes additional water-mediated interactions involving the carboxyl O atoms of the seryl moiety of the antigen and Asn128 of the lectin (Kulkarni *et al.*, 2005), accounting for its higher affinity compared with that of GalNAc.

The interactions at the primary site itself provide a partial rationale for the higher affinity of WBAI for the A blood-group substance, in terms of the additional interactions introduced by the substitution of the acetamido group at the galactose residue at the primary site. Structural studies on the complexes of WBAI with blood-group disaccharide and trisaccharide complexes further showed that the interactions at the secondary site, made up of loop L4, also contribute to the differential affinity of the lectin for A and B substances. In these structures, part of the interactions is common in A- and B-reactive trisaccharides (A-tri and B-tri, respectively). However, the change in the orientation of the second (Gal) and the third (Glc) residues of B-tri caused by an internal hydrogen bond prevents B-tri but not A-tri from interacting with Lys216. Thus, the movement of part of B-tri on account of an internal hydrogen bond, supplemented by an interaction-dependent change in the rotamer conformation of Lys216, explains the differential affinity of the lectin for A and B blood-group substances. Structures of complexes of WBAI–blood-group sugars reconfirm the role of the hypervariable loop, both in terms of its length and sequence, in conferring the specificity of the lectin (Kulkarni *et al.*, 2007).

The results presented in the present paper provide an explanation for the preferential binding of the lectin to $\alpha1\rightarrow3$-linked Gal/GalNAc-containing oligosaccharides in terms of two distinct modes of binding of sugars to the lectin. The reducing sugar of the α-linked disaccharides with linkages other than $1\rightarrow3$ binds to the lectin through a water bridge, whereas the same sugar moiety with an $\alpha1\rightarrow3$ linkage makes direct interactions with loop L4 of the protein.

The availability of different structures involving WBAI permitted a study of the plasticity of the lectin. The front curved β-sheet, which nestles the metal-binding region and on which the carbohydrate-binding loops are perched, is relatively rigid. In contrast, the flat back β-sheet involved in the quaternary association in legume lectins is flexible. This flexibility is probably necessary to account for the variation in quaternary structure.

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

References

Ahuja, K. K. (1985). *Am. J. Anat.* **174**, 207–223.
Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanaraj, V., Mahanta, S. K., Suguna, K., Surolia, A. & Vijayan, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 227–231.
Bhat, T. N. (1988). *J. Appl. Cryst.* **21**, 279–281.
Biswal, B. K., Sukumar, N. & Vijayan, M. (2000). *Acta Cryst.* D**56**, 1110–1119.
Brünger, 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., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Christensen, T. & Toone, E. J. (2003). *Methods Enzymol.* **362**, 486–504.
Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
Galili, U. (2005). *Immunol. Cell. Biol.* **83**, 674–686.
Goldstein, I. J., Winter, H. C. & Poretz, R. D. (1997). *Glycoproteins II*, edited by J. Montreuil, J. F. G. Vliegenthart & H. Schachter, pp. 403–474. Amsterdam: Elsevier.
Hardman, K. D. & Ainsworth, C. F. (1972). *Biochemistry*, **11**, 4910–4919.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Khan, M. I., Sastry, M. V. K. & Surolia, A. (1986). *J. Biol. Chem.* **261**, 3013–3019.
Koradi, R., Billeter, M. & Wüthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.
Kulkarni, K. A., Katiyar, S., Surolia, A., Vijayan, M. & Suguna, K. (2006). *Acta Cryst.* D**62**, 1319–1324.
Kulkarni, K. A., Katiyar, S., Surolia, A., Vijayan, M. & Suguna, K. (2007). *Proteins*, **68**, 762–769.
Kulkarni, K. A., Sinha, S., Katiyar, S., Surolia, A., Vijayan, M. & Suguna, K. (2005). *FEBS Lett.* **579**, 6775–6780.
Kulkarni, K. A., Srivastava, A., Mitra, N., Sharon, N., Surolia, A., Vijayan, M. & Suguna, K. (2004). *Proteins*, **56**, 821–827.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lis, H. & Sharon, N. (1998). *Chem. Rev.* **98**, 637–674.
Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). *Biochem. Biophys. Acta*, **1383**, 9–36.
Madhusudan & Vijayan, M. (1991). *Curr. Sci.* **60**, 165–170.
Mann, B. J. (2002). *Int. Rev. Cytol.* **216**, 59–80.
Manoj, N., Srinivas, V. R. & Suguna, K. (1999). *Acta Cryst.* D**55**, 794–800.
Natchiar, S. K., Srinivas, O., Mitra, N., Surolia, A., Jayaraman, N. & Vijayan, M. (2006). *Acta Cryst.* D**62**, 1413–1421.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Prabu, M. M., Sankaranarayanan, R., Puri, K. D., Sharma, V., Surolia, A., Vijayan, M. & Suguna, K. (1998). *J. Mol. Biol.* **276**, 787–796.
Puri, K. D. & Surolia, A. (1994*a*). *Pure Appl. Chem.* **66**, 497–502.
Puri, K. D. & Surolia, A. (1994*b*). *J. Biol. Chem.* **269**, 30917–30926.
Radha Kishan, K. V., Chandra, N. R., Sudarsanakumar, C., Suguna, K. & Vijayan, M. (1995). *Acta Cryst.* D**51**, 703–710.
Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A. & Vijayan, M. (1997). *Curr. Sci.* **72**, 855–861.
Reeke, G. N. Jr, Becker, J. W. & Edelman, G. M. (1975). *J. Biol. Chem.* **250**, 1525–1547.
Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577.
Sadasivan, C., Nagendra, H. G. & Vijayan, M. (1998). *Acta Cryst.* D**54**, 1343–1352.

Sankaranarayanan, R., Puri, K. D., Ganesh, V., Banerjee, R., Surolia, A. & Vijayan, M. (1993). *J. Mol. Biol.* **229**, 558–560.
Schneider, T. R. (2002). *Acta Cryst.* D**58**, 195–208.
Shaanan, B., Lis, H. & Sharon, N. (1991). *Science*, **254**, 862–866.
Sharma, V. & Surolia, A. (1997). *J. Mol. Biol.* **267**, 433–445.
Sharon, N. & Lis, H. (2003). *Lectins*, 2nd ed., pp. 1–454. Boston: Kluwer Academic Publishers.
Srinivas, V. R., Bachhawat-Sikder, K., Habib, S., Hasnain, S. E. & Surolia, A. (2001). *Biosci. Rep.* **21**, 361–367.
Vijayan, M. & Chandra, N. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.
Weis, W. I. & Drickamer, K. (1996). *Annu. Rev. Biochem.* **65**, 441–473.
Wu, A. M. (2003). *J. Biomed. Sci.* **10**, 676–688.