Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural basis for the carbohydrate-specificity of basic winged-bean lectin and its differential affinity for Gal and GalNAc

Kiran A. Kulkarni, Samiksha Katiyar, Avadhesha Surolia, Mamannamana Vijayan and Kaza Suguna*

Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

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

Received 29 May 2006
Accepted 20 July 2006

PDB References: WBAI, Gal complex, 2du0, r2du0sf; 2Me-O-Gal complex, 2dtw, r2dtwsf; GalNAc complex, 2dty, r2dtysf; Me-α-GalNAc complex, 2du1, r2du1sf.

The crystal structure of the complexes of basic winged-bean lectin with galactose, 2-methoxygalactose, *N*-acetylgalactosamine and methyl-α-*N*-acetylgalactosamine have been determined. Lectin–sugar interactions involve four hydrogen bonds and a stacking interaction in all of the complexes. In addition, an N—H···O hydrogen bond involving the hydroxyl group at C2 exists in the galactose and 2-methoxygalactose complexes. An additional hydrophobic interaction involving the methyl group in the latter leads to the higher affinity of the methyl derivative. In the lectin–*N*-acetylgalactosamine complex the N—H···O hydrogen bond is lost, but a compensatory hydrogen bond is formed involving the O atom of the acetamido group. In addition, the CH_3 moiety of the acetamido group is involved in hydrophobic interactions. Consequently, the 2-methyl and acetamido derivatives of galactose have nearly the same affinity for the lectin. The methyl group α-linked to the galactose takes part in additional hydrophobic interactions. Therefore, methyl-α-*N*-acetylgalactosamine has a higher affinity than *N*-acetylgalactosamine for the lectin. The structures of basic winged-bean lectin–sugar complexes provide a framework for examining the relative affinity of galactose and galactosamine for the lectins that bind to them. The complexes also lead to a structural explanation for the blood-group specificity of basic winged-bean lectin.

1. Introduction

Recognition of carbohydrates by proteins is known to be crucial in myriad biological processes such as cell proliferation, regulation and signalling. Frequently, these interactions involve lectins, multivalent carbohydrate-binding proteins with a high degree of specificity (Vijayan & Chandra, 1999; Loris *et al.*, 1998). At the molecular level, protein–carbohydrate interactions involve an interplay of various types of forces. Although the nature of these forces is reasonably well understood, their relative contributions and magnitude in specific cases remain equivocal (Fersht *et al.*, 1985; Laederach & Reilly, 2005). A major objective of X-ray studies on lectin–carbohydrate complexes is the delineation of the structural basis of the known sugar-specificities of lectins. Based on their affinity at the monosaccharide level, lectins are classified as specific for mannose/glucose (Man/Glc), *N*-acetylglucosamine (GlcNAc), fucose (Fuc), galactose/*N*-acetylgalactosamine (Gal/GalNAc) and *N*-acetylneuraminic acid (NeuNAc). Many of the lectins tolerate variations at the anomeric (C1) and C2

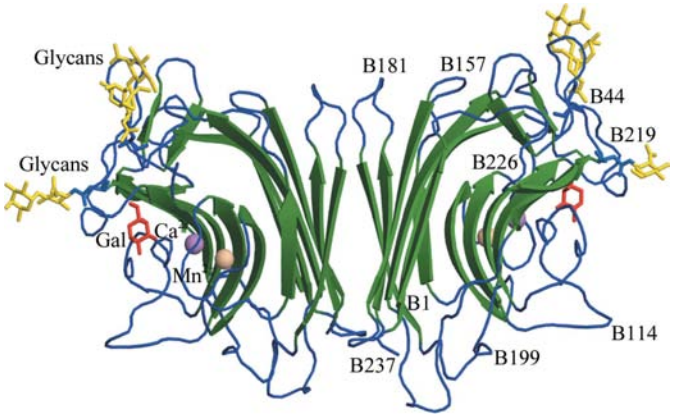

Figure 1
Dimeric structure of WBAI complexed with galactose. Ca^{2+} and Mn^{2+} are shown in spheres. Galactose and N-linked glycans are shown in stick representation. This figure was produced using *PyMOL* (http://www.pymol.org).

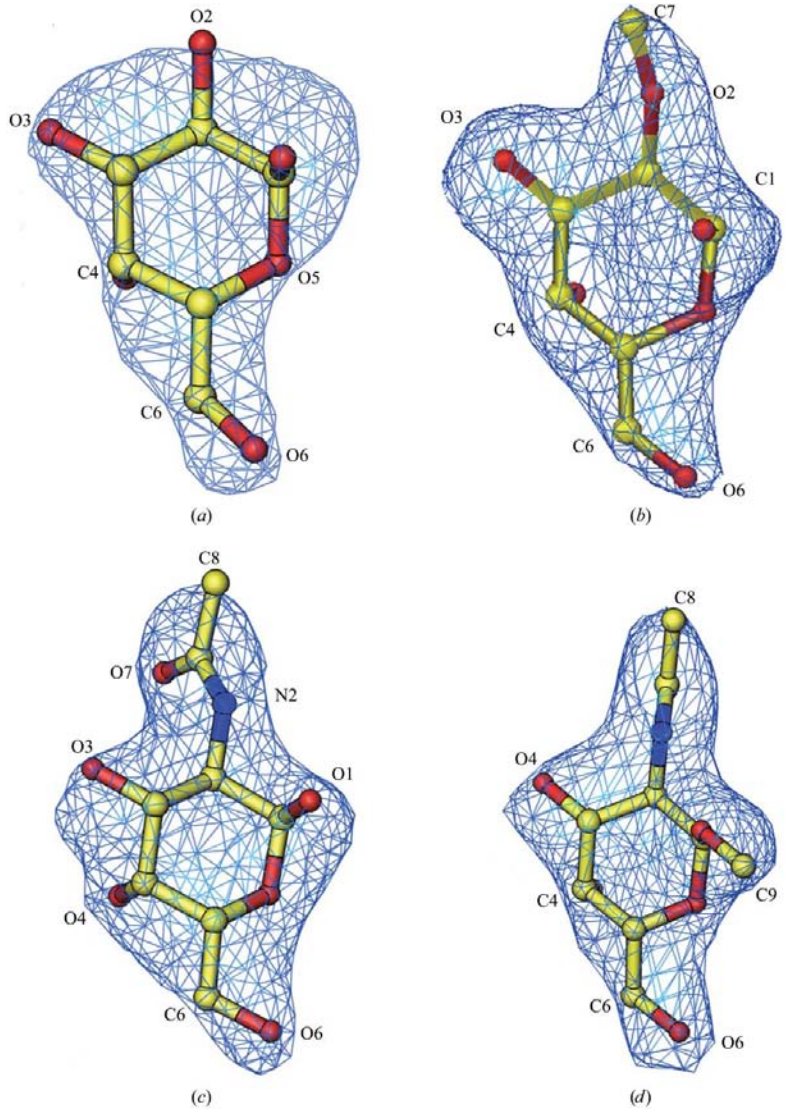

Figure 2
Electron density ($2F_o - F_c$) map of (*a*) Gal, (*b*) 2Me-O-Gal, (*c*) GalNAc and (*d*) Me-α-GalNAc contoured at 1σ. Some of the sugar atoms are labelled. This figure was generated with the program *RIBBONS* (Carson, 1997).

positions of the pyranose ring of their monosaccharide ligands more than variations at C3 and C4 (Rini, 1995; Lis & Sharon, 1998). For example, some lectins are specific for Man and Glc, while some others are specific for Gal as well as GalNAc in which the acetamido group is at the C2 position. There are of course exceptions, represented by peanut agglutinin (PNA), which binds Gal but not GalNAc, and the lectin from *Dolichos biflorus* (DBL), which binds GalNAc with considerably higher affinity than Gal (Banerjee *et al.*, 1996; Ravishankar *et al.*, 1999; Hamelryck *et al.*, 1999; Dam & Brewer, 2002).

The homodimeric N-glycosylated basic lectin from winged beans (*Psophocarpus tetragonolobus*; WBAI) exhibits a nearly fivefold higher affinity for *N*-acetylgalactosamine than galactose (Schwarz *et al.*, 1991). Here, we report the crystal structures of WBAI complexed with Gal, 2-methoxygalactose (2Me-O-Gal), GalNAc and methyl-α-GalNAc (Me-α-GalNAc) elucidated at 2.70, 2.40, 2.65 and 2.60 Å, respectively. A comparative study of these structures with the previously reported structure of the WBAI–methyl-α-galactose complex (Prabu *et al.*, 1998) provides a structural rationale for the higher affinity of WBAI for C1- and C2-substituted galactose derivatives. Furthermore, a comparative study of these complexes and the structure of other Gal/GalNAc-specific lectins leads to insights into the geometrical basis of the relative affinity of these lectins for Gal and GalNAc. Also, the structures presented here provide an explanation for the previously reported thermodynamic data on the basic winged-bean lectin. In addition, they shed light on the structural basis of the blood-group specificity of the lectin, which arises substantially through interactions at the monosaccharide level.

2. Materials and methods

2.1. Crystallization and data collection

WBAI was isolated as reported previously (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 solution containing a 20–50 molar excess of the sugar and 6–7%(*w*/*v*) polyethylene glycol (PEG) 4000 in 0.02 *M* phosphate buffer [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 for the Gal, 2Me-O-Gal and Me-α-GalNAc complexes were recorded at room temperature (293 K) using a MAR 345 imaging-plate detector mounted on a Rigaku RU-300 generator. A MAR 300 imaging-plate detector

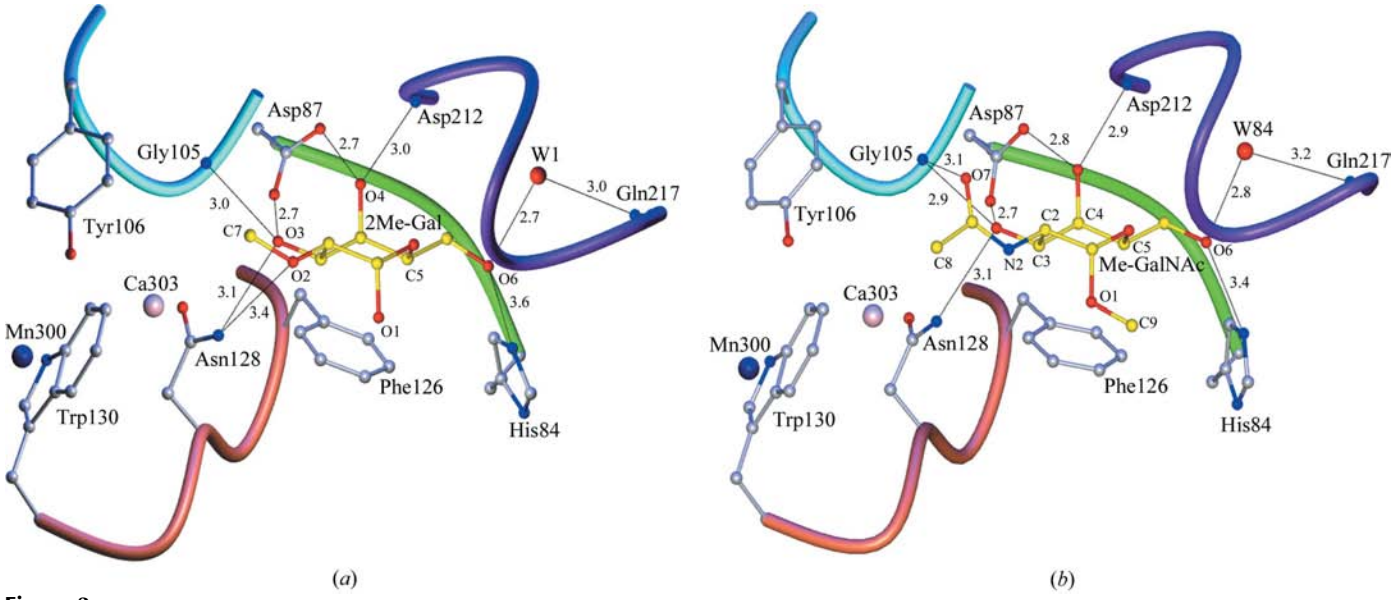

Figure 3
(*a*) WBAI–2Me-O-Gal hydrogen-bonding interactions. Four loops at the binding site are shown as coils. (*b*) WBAI–Me-α-GalNAc hydrogen-bonding interactions. Four loops at the binding site are shown as coils. This figure was produced using the program *MOLMOL* (Koradi *et al.*, 1996).

was instead used to record data for GalNAc complex. The crystal-to-detector distance was kept at 200 mm. The data were processed using *DENZO* and *SCALEPACK* from the *HKL* suite (Otwinowski & Minor, 1997). The intensities were converted to structure factors using *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

2.2. Structure refinement, validation and analysis

The crystals obtained were isomorphous to those of WBAI complexed with Me-α-galactose (Me-α-Gal; PDB code 1wl1; Prabu *et al.*, 1998). The lectin molecules in this complex (after removal of non-protein atoms) were used as the starting model for refinement. After a total of 35 cycles of rigid-body refinement followed by 150 cycles of positional refinement using *CNS* (Brünger *et al.*, 1998), R and $R_{\rm free}$ were in the ranges 0.25–0.27 and 0.26–0.28, respectively. Subsequently, bound sugar, metal ions and a few residues of N-linked glycans were added to the structure based on the $F_{\rm o} - F_{\rm c}$ and $2F_{\rm o} - F_{\rm c}$ maps. In subsequent steps of refinement, water O atoms were added to the model using peaks in electron density with heights greater than 2.5σ in $F_{\rm o} - F_{\rm c}$ and 0.8σ in $2F_{\rm o} - F_{\rm c}$ maps. Composite OMIT maps were used to remove the model bias (Bhat, 1988). Bulk-solvent correction and NCS restraints were employed throughout the refinement. The data-collection and refinement statistics are given in Table 1.

Manual model building was performed using *O* (Jones *et al.*, 1991). The structures were validated using *PROCHECK* (Laskowski *et al.*, 1993). *CONTACT* from *CCP*4 (Collaborative Computational Project, Number 4, 1994) was used to calculate the bond distances. Accessible surface area was calculated using *NACESS* (Hubbard & Thornton, 1993).

3. Results and discussion

3.1. Overall structure of the complexes

In all four structures, clear and unambiguous electron density for 236 out of 241 residues, bound sugars, up to three sugar residues in each of the N-linked glycans at Asn44 and Asn219 and metal ions is seen in all four subunits, constituting two dimers, present in the asymmetric unit of the crystal. The tertiary structure of the lectin exhibits the characteristic legume lectin fold consisting of three β-sheets and several loops arranged in a jelly-roll motif. WBAI forms the 'hand-shake' mode of dimerization (Fig. 1) first observed in *Erythrina corallodendron* lectin (ECorL; Shaanan *et al.*, 1991). The tertiary and the quaternary structure of WBAI is the same in its saccharide-free and saccharide-bound form, indicating that the lectin–carbohydrate interaction does not influence the structure of the lectin (Prabu *et al.*, 1998; Manoj *et al.*, 1999). The root-mean-square deviations between the C^{α} positions of lectin dimers in the various structures range from 0.14 to 0.26 Å.

3.2. Lectin–sugar interactions

The clear electron density for sugars (Fig. 2) found in all the structures permits a detailed discussion of the lectin–sugar interaction. As in other legume lectins, the carbohydrate-binding pocket is a shallow depression on the surface formed by four loops often referred to as *A*, *B*, *C* and *D* (Sharma & Surolia, 1997). In WBAI these four loops are formed by residues 78–87 (*A*), 95–117 (*B*), 124–134 (*C*) and 211–222 (*D*). Loop *D* is hypervariable in terms of length and sequence (Sharma & Surolia, 1997). In all four complexes the Gal moiety has an identical orientation, aided by a constellation of four hydrogen bonds characteristic of all Gal-specific legume lectins. These hydrogen bonds are Asp87 OD2—Gal O3, Asp87 OD1—Gal O4, Gly105 N—Gal O3 and Asn128 ND2—

Table 1
Data-collection and refinement statistics.

Values indicated in parentheses correspond to the last shell.

	Gal	2Me-Gal	GalNAc	Me-α-GalNAc
Data-collection statistics				
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters				
a (Å)	157.55	157.43	157.72	157.62
b (Å)	89.95	91.06	91.86	90.96
c (Å)	73.40	73.38	73.55	73.43
Resolution (Å)	30.0–2.70 (2.80–2.70)	30.0–2.40 (2.49–2.40)	25.0–2.65 (2.74–2.65)	30.0–2.60 (2.69–2.60)
No. of observed reflections	305351	287281	206515	310899
No. of unique reflections	29164 (2865)	41427 (3800)	30930 (2977)	31848 (3104)
Data completeness (%)	99.2 (99.3)	98.0 (91.4)	98.3 (96.2)	95.8 (95.1)
R_{merge}(%)	11.4 (31.6)	11.9 (49.8)	9.1 (45.3)	13.5 (43.8)
$\langle I/\sigma(I)\rangle$	16.8 (4.2)	15.8 (3.9)	11.1 (3.8)	24.3 (5.8)
Solvent content (%)	48.9	49.5	50.1	49.5
Refinement statistics				
No. of reflections used	29142	41180	30840	31799
No. of non-H atoms				
Protein	7302	7305	7330	7301
Sugar	228	333	375	286
Solvent	239	317	276	293
R_{cryst} (%)	19.4	20.4	19.6	19.4
R_{free} (%)	24.5	24.3	24.0	25.0
R.m.s. deviation from ideal values				
Bond lengths (Å)	0.007	0.007	0.007	0.008
Bond angles (°)	1.40	1.40	1.40	1.40
Dihedral angles (°)	26.3	26.4	26.3	26.4
Improper angles (°)	0.97	0.95	0.95	0.98
Ramachandran plot, residues in				
Most favoured regions (%)	85.6	88.0	87.1	85.8
Additionally allowed regions (%)	13.4	10.8	11.4	12.9
Generously allowed regions (%)	1.0	1.3	1.5	1.3
Disallowed regions (%)	0.0	0.0	0.0	0.0

Gal O3. In addition to these interactions common to all legume lectins, O4 and O6 of Gal form hydrogen bonds with Asp212 N and His84 NE2, respectively. Furthermore, the side chain of Phe126 stacks against the sugar ring. The O2 atom of the Gal moiety forms a direct hydrogen bond with Asn128 ND2. The binding affinity of galactose to WBAI results from these interactions common to all the complexes.

The lectin–sugar interactions mentioned above occur in all four complexes reported here as well as in the Me-α-Gal and Tn-antigen complexes reported previously (Prabu *et al.*, 1998; Kulkarni *et al.*, 2005), except for a small difference when there is a substitution of a methyl or acetamido group at C2. When a methyl group is substituted at this position, as in the case of 2Me-O-Gal, in addition to the above-mentioned interactions the methyl group is stabilized in the environment of Gly104, Gly105, Tyr106 and Trp130 (Fig. 3*a*), with an increase in surface area buried on complexation from 124 to 148 Å^2. Substitution of the acetamido group at C2 (as in GalNAc) on one hand abolishes interaction between O2 of the Gal moiety and Asn128 ND2, but on the other hand leads to a new hydrogen bond between O7 of the sugar and Gly105 N (Fig. 3*b*). Furthermore, CH_3 of the acetamido group is accommodated in a hydrophobic pocket formed by Tyr106 and Trp130. The surface area buried on complexation is now 162 Å^2. Both types of substitutions provide additional hydrophobic/non-polar interactions, while maintaining the same number of hydrogen-bonded interactions of the sugar with the lectin. This might account for the similar affinity of WBAI for 2Me-O-Gal and GalNAc and also for their increased affinity in comparison to that of Gal (Khan *et al.*, 1986; Swaminathan *et al.*, 1997). In the Me-α-Gal complex (Prabu *et al.*, 1998), the methyl group attached to O1, although substantially exposed to solvent, makes a close non-bonded interaction with C$^\alpha$ of Gly215 in the *D* loop (Fig. 3*b*), with a surface area buried on complexation of 134 Å^2. This may account for the higher affinity of Me-α-Gal than Gal for WBAI (Prabu *et al.*, 1998; Swaminathan *et al.*, 1997). In the complex with Me-α-GalNAc, additional interactions involving the acetamido group as well as involving the methyl group attached to O1 exist. Additionally, the surface area buried by the ligand (175 Å^2) is the highest in the series. Therefore, the affinity of Me-α-GalNAc for WBAI is higher than those of GalNAc and Me-α-Gal. In the four structures reported here, a water bridge connecting O6 with Gln217 N exists in one or more subunits (Fig. 3*a*). The structures do not provide evidence for any noticeable reorganization of water molecules in the binding site resulting from changes in the ligand.

In terms of haemagglutination, WBAI is specific for the blood group A substance, but also binds to the blood group B substance with lower affinity. It does not bind to the blood group O substance at all. As can be seen from Fig. 4, the O-group substance involves a β-linkage with the Gal. As shown previously (Prabu *et al.*, 1998), β-substitution at Gal bound to WBAI leads to unacceptable steric clashes. This explains the inability of the lectin to agglutinate O-group erythrocytes. The difference between A and B antigens is primarily in the terminal sugar, which is GalNAc in the former and Gal in the latter. WBAI has higher affinity for GalNAc than Gal. The structural basis for this difference brought to light by the results reported here forms the rationale for the differential affinity of WBAI for A and B blood-group substances. Indeed, WBAI presents an interesting case where the lectin's affinity for different blood groups can be substantially explained on the basis of the interactions of a monosaccharide at the primary binding site.

3.3. Differential specificity of Gal and GalNAc in lectins

The interaction of Gal and GalNAc with the lectin in the WBAI complexes provides a framework for exploring the structural basis of the differential affinity of these sugars for

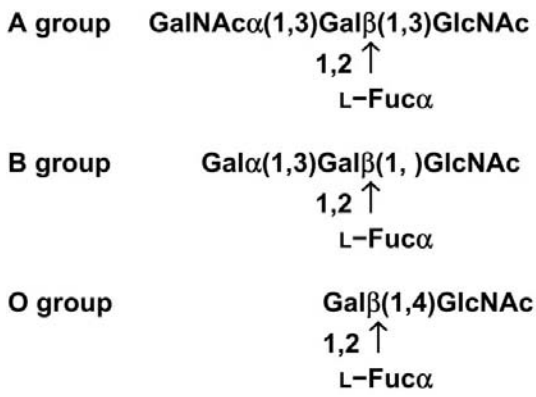

Figure 4
Schematic representation of blood-group determinants.

legume lectins. PNA, DBL and ECorL are the other Gal/GalNAc-specific lectins for which relevant crystallographic and thermodynamic data are available (Ravishankar *et al.*, 1999; Hamelryck *et al.*, 1999; Sharma *et al.*, 1998; Surolia *et al.*, 1996; Elgavish & Shaanan, 1998). Of these, PNA binds Gal but not GalNAc. This can be understood in terms of the substitution of a threonine by a glutamic acid at position 129, which is close to C2 in galactose. The bulkier side chain appears to offer steric hindrance to the acetamido group at C2 in GalNAc. Affinity for GalNAc can be generated in PNA by mutating Glu129 to Asp (Sharma *et al.*, 1998). On the other hand, DBL has poor affinity for Gal but binds GalNAc. In this lectin, the aromatic residue (Phe126 in WBAI) which stacks with the pyranose ring is replaced by an aliphatic residue Leu, thus abolishing the stacking interaction. This appears to lead to the poor affinity for Gal. However, apparently on account of the additional interactions involving the acetamido group, the lectin still binds GalNAc (Hamelryck *et al.*, 1999). ECorL presents an interesting case in which the lectin has nearly the same affinity for Gal and GalNAc (Surolia *et al.*, 1996; Elgavish & Shaanan, 1998), unlike in the case of WBAI, which has much greater affinity for GalNAc. In ECorL, the glycine at position 104 of WBAI, which is a part of conserved Gly-Gly stretch of GalNAc-binding legume lectins, is replaced by the bulkier tyrosine which displaces the acetamido group from its position in the WBAI–GalNAc complex. Consequently, O7, which has now moved by about 0.8 Å, no longer makes a hydrogen bond with Gly105 N. This results in the reduction of the affinity of EcorL for GalNAc (Hamelryck *et al.*, 1999; Elgavish & Shaanan, 1998).

Although not directly relevant to the question of the discrimination between Gal and GalNAc, a comparison between the relative affinities of WBAI and the acidic lectin from winged bean, WBAII, for Gal and Me-α-Gal is interesting. The higher affinity of WBAI for Me-α-Gal results from the interactions of the methyl group with Gly215 of the variable *D* loop. The *D* loop is much shorter in WBAII and the corresponding residue is nearly 6 Å away from the methyl group. Consequently, WBAII has nearly same affinity for Gal and Me-α-Gal (Acharya *et al.*, 1990; Manoj *et al.*, 2000).

The data were collected at the X-ray Facility for Structural Biology at 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 from the DBT. Financial support from the DST is acknowledged.

References

Acharya, S., Patanjali, S. R., Sajjan, U., Gopalakrishnan, B. & Surolia, A. (1990). *J. Biol. Chem.* **265**, 11586–11594.
Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanraj, V., Mahanta, S. K., Suguna, K., Surolia, A. & Vijayan, M. (1996). *Proc. Natl Acad. Sci. USA*, **91**, 227–231.
Bhat, T. N. (1988). *J. Appl. Cryst.* **21**, 279–281.
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.
Carson, M. (1997). *Methods Enzymol.* **277**, 493–505.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dam, T. K. & Brewer, F. C. (2002). *Chem. Rev.* **102**, 387–429.
Elgavish, S. & Shaanan, B. (1998). *J. Mol. Biol.* **277**, 917–932.
Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. & Winter, G. (1985). *Nature (London)*, **314**, 235–238.
Hamelryck, T. W., Loris, R., Bouckaert, J., Dao-Thi, M. H., Strecker, G., Imberty, A., Fernandez, E., Wyns, L. & Etzler, M. E. (1999). *J. Mol. Biol.* **286**, 1161–1177.
Hubbard, S. J. & Thornton, J. M. (1993). *NACCESS Computer Program.* Department of Biochemistry and Molecular Biology, University College, London.
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., Sinha, S., Katiyar, S., Suroila, A., Vijayan, M. & Suguna, K. (2005). *FEBS Lett.* **579**, 6775–6780.
Laederach, A. & Reilly, P. J. (2005). *Proteins*, **60**, 591–597.
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.
Manoj, N., Srinivas, V. R. & Suguna, K. (1999). *Acta Cryst.* D**55**, 794–800.
Manoj, N., Srinivas, V. R., Surolia, A., Vijayan, M. & Suguna, K. (2000). *J. Mol. Biol.* **302**, 1129–1137.
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.
Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1999). *Acta Cryst.* D**55**, 1375–1382.
Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577.
Sankaranarayanan, R., Puri, K. D., Ganesh, V., Banerjee, R., Surolia, A. & Vijayan, M. (1993). *J. Mol. Biol.* **229**, 558–560.

Schwarz, F. P., Puri, K. & Surolia, A. (1991). *J. Biol. Chem.* **266**, 344–24350.
Shaanan, B., Lis, H. & Sharon, N. (1991). *Science*, **254**, 862–866.
Sharma, V., Srinivas, V. R., Adhikari, P., Vijayan, M. & Surolia, A. (1998). *Glycobiology*, **8**, 1007–1012.
Sharma, V. & Surolia, A. (1997). *J. Mol. Biol.* **267**, 433–45.
Surolia, A., Sharon, N. & Schwarz, F. P. (1996). *J. Biol. Chem.* **271**, 17697–1703.
Swaminathan, C. P., Gupta, D., Sharma, V. & Surolia, A. (1997). *Biochemistry*, **36**, 13428–13434.
Vijayan, M. & Chandra, N. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.