Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Modification of the sugar specificity of a plant lectin: structural studies on a point mutant of *Erythrina corallodendron* lectin

Subbiah Thamotharan, Tangavelou Karthikeyan, Kiran A. Kulkarni, Kartika N. Shetty, Avadhesha Surolia, Mamannamana Vijayan and Kaza Suguna*

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

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

Received 8 December 2010
Accepted 7 February 2011

PDB References: Y106G mutant of EcorL, complex with Gal, 3n36; complex with GalNAc, 3n35; complex with citrate, 3n3h.

A mutant of *Erythrina corallodendron* lectin was generated with the aim of enhancing its affinity for *N*-acetylgalactosamine. A tyrosine residue close to the binding site of the lectin was mutated to a glycine in order to facilitate stronger interactions between the acetamido group of the sugar and the lectin which were prevented by the side chain of the tyrosine in the wild-type lectin. The crystal structures of this Y106G mutant lectin in complex with galactose and *N*-acetylgalactosamine have been determined. A structural rationale has been provided for the differences in the relative binding affinities of the wild-type and mutant lectins towards the two sugars based on the structures. A hydrogen bond between the O6 atom of the sugars and the variable loop of the carbohydrate-binding site of the lectin is lost in the mutant complexes owing to a conformational change in the loop. This loss is compensated by an additional hydrogen bond that is formed between the acetamido group of the sugar and the mutant lectin in the complex with *N*-acetylgalactosamine, resulting in a higher affinity of the mutant lectin for *N*-acetylgalactosamine compared with that for galactose, in contrast to the almost equal affinity of the wild-type lectin for the two sugars. The structure of a complex of the mutant with a citrate ion bound at the carbohydrate-binding site that was obtained while attempting to crystallize the complexes with sugars is also presented.

1. Introduction

Specific recognition and binding of cell-surface carbohydrates by lectins is an essential step in numerous processes involving cell–cell interactions in nature. Lectins, which are found in all kingdoms of life, are multimeric proteins of non-immune origin with diverse biological functions (Sharon & Lis, 1998, 2004; Rini, 1995; Vijayan & Chandra, 1999). Lectins are extensively used in biological and clinical applications and research because of their exquisite specificities for a wide variety of carbohydrates. The most common and the best characterized lectins in terms of biochemical properties and structure are those from legume plants (Sharon & Lis, 1990; Loris *et al.*, 1998). In spite of having similar sequences and identical protomer structures, legume lectins exhibit significant variations in their oligomeric status and affinity for monocarbohydrates to complex carbohydrates. Owing to their ability to discriminate between different carbohydrates, legume lectins are considered to be the best available model systems to explore various modes of carbohydrate recognition by proteins.

One well characterized legume lectin is *Erythrina corallodendron* lectin (EcorL), which has affinity for galactose (Gal) and *N*-acetylgalactosamine (GalNAc). The structural basis of galactose recognition by legume lectins was obtained for the first time from the crystal structure of this lectin in complex with lactose (Shaanan *et al.*, 1991). Many GalNAc-binding legume lectins such as soybean agglutinin (SBA), *Dolichos biflorus* lectin (DBL), *Robinia pseudoacacia* lectin (RBL), isolectin B4 from *Viscia villosa* (VVLB4) and winged bean basic and acidic agglutinins (WBAI and WBAII, respectively) have a greater affinity for GalNAc than for Gal. In contrast, in the case of EcorL the affinities for these two sugars are about the same. A comparison of the crystal structures of the reported legume lectin complexes with GalNAc at the primary binding site revealed that a conserved hydrogen bond between the *N*-acetamido group of GalNAc and the backbone N atom of a glycine residue of the binding site is absent in EcorL. This hydrogen-bonding interaction is prevented by the bulky side chain of a tyrosine residue at position 106 in EcorL which replaces the structurally equivalent glycine residue found in all GalNAc legume lectins mentioned above, except for SBA, which has an alanine at this position. With the objective of enhancing the affinity of EcorL for GalNAc by promoting the formation of this apparently critical hydrogen bond, the Y106G mutant of EcorL was generated by site-directed mutagenesis. The interactions between the mutant lectin and the sugars Gal and GalNAc were studied by crystallographic analysis and correlated with the observed thermodynamic data. The crystal structures of the Y106G mutant of EcorL in complex with Gal, GalNAc and a citrate ion (abbreviated as MEcGal, MEcNGal and MEc-Cit, respectively) presented here are compared with the reported structures of wild-type EcorL in complex with carbohydrates (EcGal, EcNGal *etc.*). The wild-type lectin that was isolated from the seeds of *E. corallodendron* is N-glycosylated at Asn17 and Asn113, whereas recombinant forms, including the mutants, are non-glycosylated. This study has provided a structural perspective of the altered sugar specificity of EcorL that results from the point mutation.

As a consequence of the available and rapidly increasing wealth of biochemical and structural information on lectins and the success in generating recombinant forms of many of them, it has become possible to modify the binding sites of lectins by site-directed mutagenesis in order to achieve selective carbohydrate binding, thereby converting them into more effective biomarkers or better receptors for preferential recognition of cell types in biological reactions. While many of the early studies on mutant legume lectins focused on exploring the role of individual residues on their properties, such as carbohydrate recognition, biological activity, stability and nutritional value (van Eijsden *et al.*, 1992, 1994; Hoedemaeker *et al.*, 1993; Jordan & Goldstein, 1995; Nishiguchi *et al.*, 1997; Arango *et al.*, 1993; Adar & Sharon, 1996; Moreno *et al.*, 1997; Zhu *et al.*, 1996), recent work on peanut lectin (PNA) has been directed towards altering the carbohydrate specificity (Sharma *et al.*, 1996, 1998; Adhikari *et al.*, 2001). However, no crystallographic studies of any of the mutants have been reported to date. Here, we present the crystal structure of a mutant legume lectin which has provided the first structural view of an engineered carbohydrate-binding site.

2. Materials and methods

2.1. Construction of the Y106G EcorL mutant

Site-directed mutagenesis (Y106G) was performed by PCR overlap extension using the pET-3d-EcorL wild type as the template (Arango *et al.*, 1990). The two PCR products with overlapping ends were generated by employing the EcorL wild-type For (5′-GATATA**CCATGG**TGGAAA-3′) primer together with EcorL Y106G Rev (5′-TATTCCGAGGTATC-CGCCACCT-3′) and EcorL wild-type Rev (5′-CGC**GGAT-CC**CACCATTGCAATCATTATT-3′) together with EcorL Y106G For (5′-AGGTGGCGGATACCTCGGAATA-3′) to introduce *Nco*I and *Bam*HI restriction sites (bold). The combined PCR product was digested with *Nco*I and *Bam*HI and cloned into pET-3d-EcorL digested with the same two enzymes. The recombinant DNAs were fully sequenced to confirm the planned mutation and to ascertain that no spurious mutations had occurred.

2.2. Overexpression and purification of the mutant EcorL protein

The pET-3d EcorL mutant clone was overexpressed in BL21 cells. The mutant protein that accumulated in inclusion bodies was isolated by denaturation and refolding as described previously for wild-type EcorL (Rodriguez-Arango *et al.*, 1992). The mutant protein was solubilized with 6 *M* urea in 10 m*M* 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS) buffer pH 10.5. The denatured protein was refolded by tenfold dilution and the pH was adjusted to 7.5 using 1 *M* Tris–HCl buffer followed by ultrafiltration and the solution was dialyzed against 10 m*M* Tris–HCl buffer at pH 7.5 containing 150 m*M* NaCl, 50 m*M* $CaCl_2$ and 50 m*M* $MnCl_2$. The soluble mutant protein was purified by gel-filtration chromatography on a Bio-Gel P100 column. The protein concentration was determined from its specific extinction coefficient ($A_{280}^{1\%}$ = 14.99). The protein purity and molecular weight were assessed by SDS–PAGE with standard molecular-weight markers.

2.3. Crystallization

Crystals suitable for X-ray diffaction were obtained after 4 d by vapour diffusion at room temperature employing the hanging-drop method. A 3 µl drop consisting of 20 mg ml^{-1} protein solution (150 m*M* NaCl, 50 m*M* $CaCl_2$, 50 m*M* $MnCl_2$ and a 20-fold molar excess of sugar in 10 m*M* Tris–HCl buffer pH 7.5) was mixed with 1 µl reservoir solution consisting of 20% PEG 3350 and 0.25 *M* diammonium hydrogen citrate in water and equilibrated against 500 µl reservoir solution. Crystals of the lectin in complex with *N*-acetyl-D-galactosamine (GalNAc) were obtained using these conditions, but a 150-fold molar excess of sugar was required in the case of the α-D-galactose (Gal) complex. Galactose did not bind to the lectin when a 20-fold molar excess of galactose was added to

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last shell.

	GalNAc	Gal	Citrate
Data-collection statistics			
Space group	$P3_121$	$P3_121$	$P3_121$
Unit-cell parameters (Å)			
$a = b$	102.51	102.58	102.78
c	57.02	57.11	56.98
Resolution (Å)	30.0–2.00 (2.07–2.00)	30.0–2.30 (2.38–2.30)	30.0–2.00 (2.07–2.00)
No. of observed reflections	214589	128509	209083
No. of unique reflections	23767 (2362)	15677 (1534)	23779 (2351)
Data completeness (%)	100 (100)	99.8 (100)	100 (100)
$R_{\rm merge}$† (%)	10.3 (42.3)	13.7 (44.4)	11.1 (40.4)
$\langle I/\sigma(I)\rangle$	22.6 (6.5)	16.7 (5.5)	21.0 (6.7)
Multiplicity	9.0	8.2	8.8
Solvent content (%)	62.3	62.4	62.5
Refinement statistics			
No. of reflections used	23647	15661	23739
No. of non-H atoms			
Protein atoms	1873	1873	1873
Sugar/citrate atoms	15	12	13
Solvent atoms	151	146	158
Ions	2	2	2
$R_{\rm cryst}$‡ (%)	17.7	17.7	18.4
$R_{\rm free}$‡ (%)	18.9	20.8	20.2
R.m.s. deviation from ideal values			
Bond lengths (Å)	0.007	0.008	0.013
Bond angles (°)	1.60	1.50	1.50
Dihedral angles (°)	27.10	26.30	26.30
Ramachandran plot, residues in			
Most favoured regions (%)	91.3	89.8	91.3
Additional allowed regions (%)	8.7	10.2	8.7
Generously allowed regions (%)	0.0	0.0	0.0
Disallowed regions (%)	0.0	0.0	0.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{\rm cryst}$ and $R_{\rm free} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$.

the protein solution. Instead, a citrate ion from an additive solution was found to be present in the binding site of the lectin.

2.4. Data collection and processing

The intensity data were collected with 1° oscillation at room temperature using a MAR345 image-plate detector mounted on a Rigaku RU-300 generator. The crystal-to-detector distance was kept at 110 mm. The incident X-ray beam was focused with an Osmic mirror system. The data were processed using *DENZO* and *SCALEPACK* as implemented in the *HKL* suite of programs (Otwinowski & Minor, 1997). The processed data were converted to structure factors using the program *TRUNCATE* from *CCP*4 (Collaborative Computational Project, Number 4, 1994). The Matthews coefficient (Matthews, 1968) indicated the presence of one protomer in the asymmetric unit in the crystals of all three of the mutant complexes reported here. The details of data collection are given in Table 1.

2.5. Structure determination and refinement

The crystal structure of the Y106G mutant of EcorL complexed with GalNAc was determined by the molecular-replacement technique using *Phaser* (Read, 2001), with one of the protomers of the structure of recombinant EcorL complexed with lactose (rEcorL–lactose; PDB entry 1sfy) solved previously in this laboratory (Kulkarni *et al.*, 2004) as the search model. The initial model was subjected to rigid-body refinement, treating the complete protomer as a rigid body. This was followed by several cycles of positional refinement using *CNS* (Brünger *et al.*, 1998) with a maximum-likelihood-formalism target function. At this stage, bound sugar (GalNAc), metal ions and three C-terminal residues were added to the model based on $F_o - F_c$ and $2F_o - F_c$ maps. In subsequent steps of refinement, water O atoms were added successively to the model at positions with peak heights greater than 2.5σ in $F_o - F_c$ and 0.8σ in $2F_o - F_c$ maps. Iterative cycles of model building using *FRODO* (Jones, 1978) and refinement were carried out until R and $R_{\rm free}$ converged. After removal of nonprotein atoms, the protomer of the GalNAc complex was used as the starting model for refinement of both the Gal and citrate complexes, as the crystals of all three complexes were found to be isomorphous. A total of 30 cycles of rigid-body refinement followed by 100 cycles of positional refinement were performed using *CNS*. Subsequently, bound galactose/citrate, metal ions and water molecules coordinated to the metal ions were added to the model based on the $F_o - F_c$ and $2F_o - F_c$ maps. OMIT maps were calculated for all the complexes and were used to remove model bias in the course of refinement. Bulk-solvent corrections and overall anisotropic B-factor corrections were used throughout the refinement. The final values of R and $R_{\rm free}$ and other relevant refinement statistics are listed in Table 1. The refined models were checked using *PROCHECK* (Laskowski *et al.*, 1993) and *SFCHECK* (Vaguine *et al.*, 1999) as implemented in the *CCP*4 program suite.

2.6. Analysis of the model

The superposition of structures was performed using the program *ALIGN* (Cohen, 1997) and accessible surface-area calculations were performed using *NACCESS* (Hubbard & Thornton, 1993). The shape complementarity for various dimers was calculated using the program *SC* from the *CCP*4 program suite. Hydrogen bonds were calculated using *CONTACT* from *CCP*4. Structurally invariant water molecules were identified using the web-based program *3DS* (Sumathi *et al.*, 2006). Figures of the structures were prepared and rendered with *PyMOL* (DeLano, 2002).

2.7. Isothermal titration calorimetry

To determine the binding of Gal to the mutant EcorL, 0.607 m*M* protein in Tris buffer pH 7.5 containing 1 m*M* $CaCl_2$, 1 m*M* $MnCl_2$ and 150 m*M* NaCl in a sample cell was titrated with a 150-fold molar excess of Gal in the same buffer from a 250 µl rotating-stirrer syringe. To estimate GalNAc binding, 0.624 m*M* of the protein was titrated with a 20-fold molar excess of GalNAc. The titrations were performed while the samples were being stirred at 300 rev min^{-1} at 286 K. The heats of dilution of the sugars were subtracted from the titration data. The data were fitted using a nonlinear least-squares minimization procedure to determine the binding stoichiometry (n), the binding constant (K_b) and the change in enthalpy of binding (ΔH_b^o) using *Origin* software (MicroCal). The change in the free energy of binding (ΔG_b^o) is calculated from $\Delta G_b^o = -nRT\ln K_b$, where R is the gas constant and T is the temperature in kelvin. These thermodynamic quantities were used to determine the change in entropy (ΔS) from $\Delta G_b^o = \Delta H_b^o - T\Delta S$. The experiments were performed using a VP-ITC titration microcalorimeter from MicroCal.

3. Results

3.1. General features of the structures

EcorL is a homodimeric legume lectin with 242 amino-acid residues and a protomer molecular weight of 26 309. Its structures in the ligand-free form and in complex with Gal (EcGal), GalNAc (EcNGal), lactose (EcLac) and *N*-acetyllactosamine (EcNLac) have been determined previously (Shaanan *et al.*, 1991; Elgavish & Shaanan, 1998). Subsequently, the crystal structure of the recombinant form of EcorL (rEcorL) in complex with lactose was reported from this laboratory (Kulkarni *et al.*, 2004). The structures of the Y106G mutant of EcorL complexed with Gal (MEcGal), GalNAc (MEcNGal) and citrate (MEc-Cit) presented here revealed that the overall fold of the lectin (Fig. 1) remains essentially the same in all of the reported structures. The carbohydrate ligands in two of the complexes, MEcGal and MEcNGal, and the citrate ion in the third complex, MEc-Cit, are clearly seen in the electron-density maps (Fig. 2). In the structures of EcorL reported previously, electron density could only be seen for the first 239 of the expected 242 residues. For the first time, clear and continuous electron density appeared for all 242 residues of the protomers in all three complexes presented here. The location of the two metal ions, Ca^{2+} and Mn^{2+}, and their interactions with the protein atoms in the present structures are similar to those in other legume lectins. The tertiary structure of EcorL adopts the jelly-roll motif characteristic of legume lectins. It comprises several loops that connect the various strands of the β-sheets, which include a six-stranded flat back β-sheet, a seven-stranded curved front β-sheet and a small three-stranded β-sheet bridging the front and back β-sheets. Glycosylated and non-glycosylated EcorL both form noncanonical dimers, with interfaces involving the six-stranded flat back β-sheet of the protomers in a handshake fashion. The asymmetric unit of the $P3_121$ cell contains one protomer of the mutant EcorL and the dimer is generated by the crystallographic twofold axis.

The r.m.s. deviations in C^α positions when protomers of the mutant lectins are superposed do not exceed 0.08 Å, indicating that the protomer structures are identical in the three complexes. To explore the influence of the glycan on the structure of the lectin, a comparative analysis was carried out by superimposing all of the crystal structures of the wild-type and recombinant forms of EcorL. The r.m.s. deviations in C^α positions for the protomers and the dimers are within 0.24 and 0.52 Å, respectively, indicating that the overall tertiary and quaternary structures of the lectin are mostly unaffected by glycosylation. However, noticeable deviations in the mutant lectin are observed in the stretches of residues 62–66 and 81–84, with a maximum shift of about 3 Å at Thr83 (Fig. 1). Surrounding symmetry-related protein molecules appear to cause these variations owing to differences in crystal packing. Another region which undergoes a significant conformational change is the stretch of residues 218–220 of a loop located in the vicinity of the mutated residue (Fig. 3). The residues Ala218 and Gln219 of this loop interact with carbohydrates in the wild-type EcorL complexes. The C^β atom of Ala218 makes favourable hydrophobic interactions with the side chain of Tyr106 in the wild-type protein. These interactions are lost in the mutant owing to the absence of the tyrosine side chain, leading to small rearrangements in the positions of the backbone atoms of Ala218 and changes in the side-chain conformations of Gln219 and Arg220.

3.2. Lectin–ligand interactions

In legume lectins, the carbohydrate-binding pocket is a shallow depression on the surface formed by four loops referred to as *A*, *B*, *C* and *D* (Sharma & Surolia, 1997). Loops *A*, *B* and *C* with conserved residues constitute the primary or monosaccharide-binding site, whereas the highly variable *D* loop determines the monosaccharide specificity. In EcorL, these loops are formed by residues 80–88 (*A*), 97–123 (*B*), 130–141 (*C*) and 217–224 (*D*). Not surprisingly, the galactose

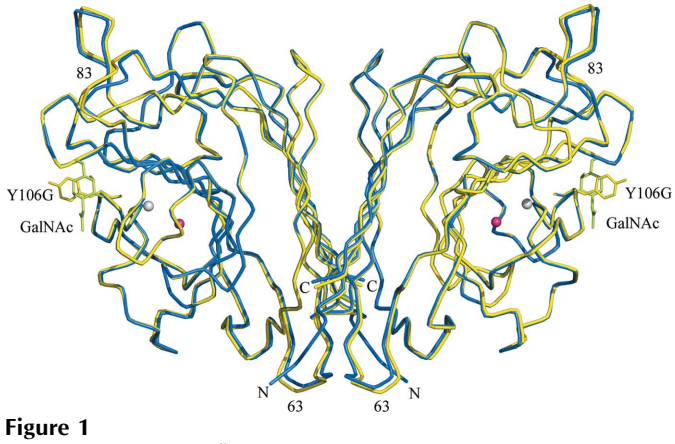

Figure 1
Superposition of the C^α traces of wild-type and mutant (Y106G) EcorL dimers are shown in yellow and blue, respectively. Bound GalNAc sugar and Tyr106 in the wild-type structure are shown as sticks and the metal ions Ca^{2+} and Mn^{2+} are shown as white and red spheres, respectively.

moiety in the MEcGal and MEcNGal complexes occupies the shallow depression on the surface of the protein between the four ligand-binding loops. However, there is a small shift of about 0.8 Å of the galactose moiety towards the lectin in both complexes of the mutant lectin (Fig. 3). Interestingly, the citrate ion is located in the binding pocket, with a significant overlap with the space occupied by the sugar molecule (Fig. 4). The positions of the amino-acid side chains within the binding-site region are the same in all three complexes. The surface areas of the ligands buried in the complexes are 225 Å^2 for galactose and citrate and 272 Å^2 for GalNAc, accounting for about 75% of the total surface areas of the ligands.

Small but significant structural changes occur at two ligand-binding residues, Ala218 and Gln219, as a consequence of the mutation. The C^{β} atom of Ala218 moves away from the binding site by about 1 Å and the conformation of Gln219 differs in the wild-type and mutant structures. The χ_1, χ_2 and χ_3 angles of Gln219 are −62.0, −64.0 and −61.6° in MEcGal, −61.1, −70.0 and −60.1° in MEcNGal and −58.9, −64.7 and −56.9° in MEc-Cit, respectively, whereas the corresponding values are −68.3, 175.2 and −98.8° in EcGal and −66.7, 172.2 and −81.3° in EcNGal, respectively. These changes have implications for the carbohydrate-binding affinity of the lectin, as discussed in the following sections.

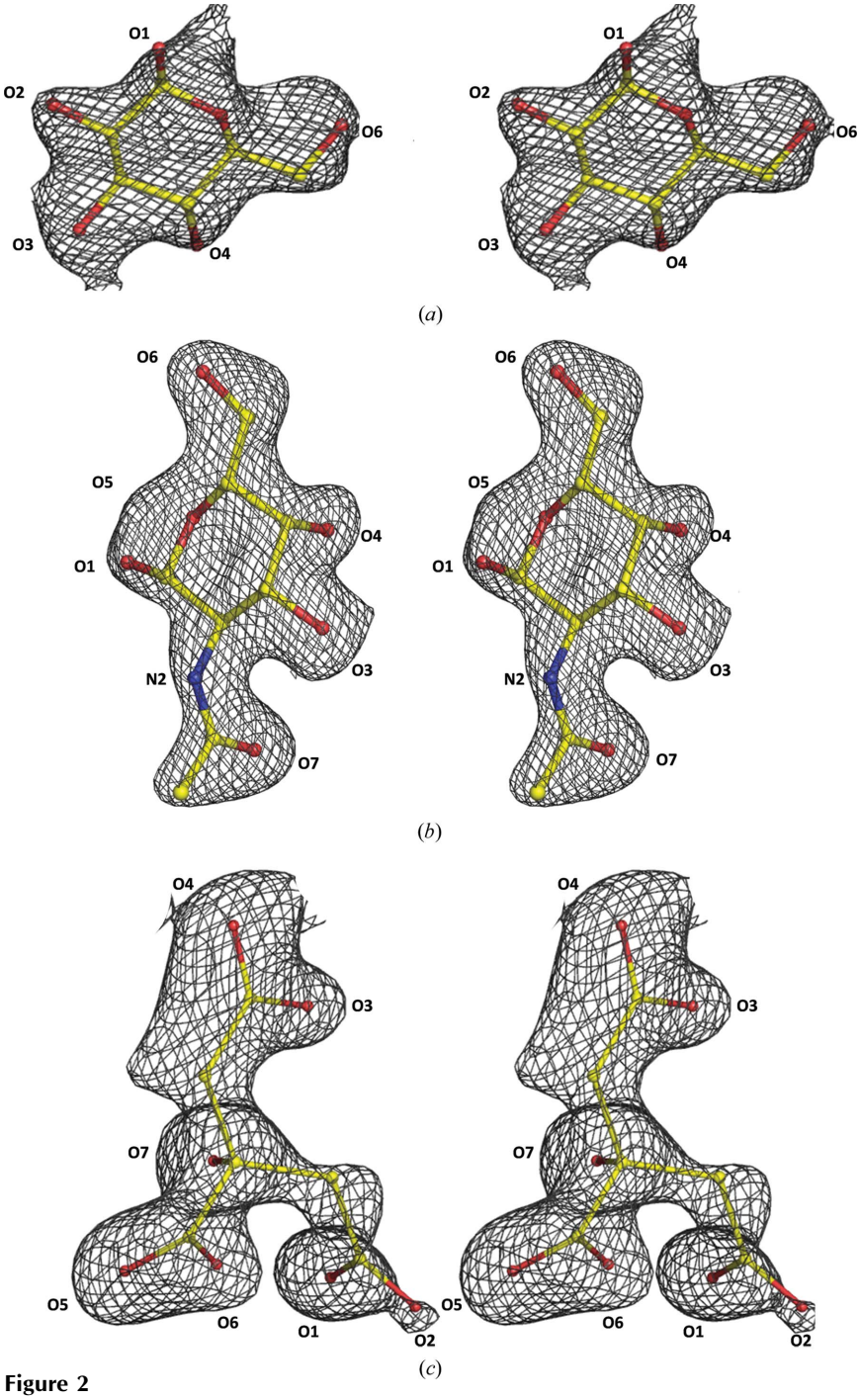

Figure 2
Stereoviews of the $2F_o - F_c$ electron density contoured at the 1σ level at the bound ligands: (*a*) galactose, (*b*) GalNAc and (*c*) citrate.

3.2.1. Mutant EcorL–Gal interactions. As observed in the reported complexes of EcorL, the hydroxyl O3 of the galactose moiety is involved in three hydrogen bonds to Asp89 OD1, Gly107 N and Asn133 ND2 and the O4 atom makes a hydrogen bond to Asp89 OD2 (Fig. 5*a*, Table 2). These four hydrogen bonds are conserved in all Gal-binding legume lectins. Also, the sugar ring makes additional van der Waals interactions with the aromatic ring of Phe131 as in most legume lectins.

Other interactions that are present in this complex as well as in all other EcorL complexes are a hydrogen bond between O4 and the N atom of Ala218 of the *D* loop and a water-mediated interaction between O6 and the carbonyl group of Leu86. In MEcGal, atoms O2 and O3 of the sugar make hydrogen bonds to water molecule W390, which in turn forms a hydrogen bond to the N atom of Gly107 of loop *B*. This water is not present in EcGal because of the presence of the Tyr106 side chain. Atom O6 of Gal in the EcGal complex makes a hydrogen bond to the side-chain atom NE2 of Gln219, which belongs to the carbohydrate-specificity loop *D*. This hydrogen bond is lost in the mutant complex owing to the change in the side-chain conformation of Gln219 as discussed above. The lack of this hydrogen bond substantially reduces the affinity of Gal for the mutant lectin. Because of this reduction in the affinity, larger amounts of the sugar were required to prepare this complex compared with that with GalNAc, as mentioned in §2.

3.2.2. Mutant EcorL–GalNAc interactions. In addition to the interactions made by the galactose moiety, the O7 atom of the acetamido group of GalNAc makes a

direct hydrogen bond to the N atom of Gly107 in MEcNGal by replacing water molecule W390 (Fig. 5*b*, Table 2). This interaction, which is observed in other GalNaAc-binding lectins but not in wild-type EcorL, enhances the affinity of the mutant lectin for GalNAc compared with that for Gal. In the MEcNGal complex, owing to the absence of the Tyr106 side chain, the O7 atom moves towards the lectin by about 1 Å when compared with the structure of EcNGal to interact with the backbone N atom of Gly107. The O7···N Gly107 distance of 2.95 Å in MEcNGal indicates a strong hydrogen bond, as opposed to a distance of 3.99 Å in EcNGal. Thus, there is clear structural evidence that the designed mutant, Y106G, of EcorL indeed has the predicted hydrogen-bonding interaction with GalNAc which could potentially affect the differential affinity of EcorL for Gal and GalNAc, as anticipated.

The O7 atom of the acetamido group also makes a water-mediated hydrogen bond to the main-chain O atom of Gly105, which precedes the mutated Gly at position 106. This water-mediated lectin–sugar interaction is absent in EcNGal. However, O7 of the GalNAc moiety interacts with the hydroxyl group of Tyr108 through two water molecules in EcNGal. In most of the legume lectins that are specific for Gal/GalNAc monosaccharides this aromatic amino-acid residue is conserved, except in PNA, which has a threonine at this position, and in some of the Man/Glc-specific legume lectins, which have a leucine residue instead.

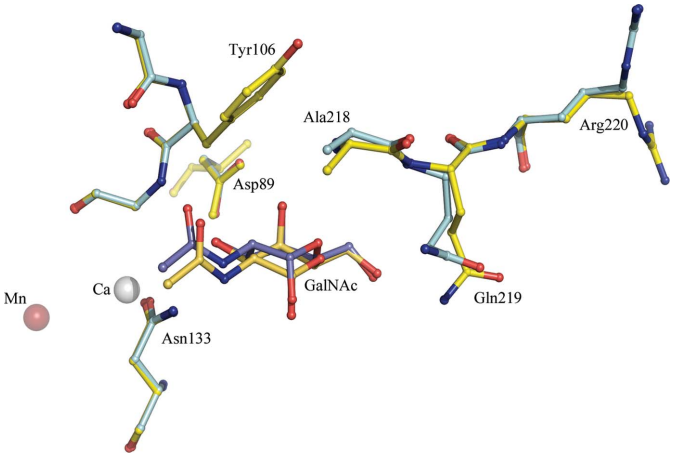

Figure 3
Superposition of the carbohydrate-binding sites of EcorL (blue) and mutant EcorL (yellow).

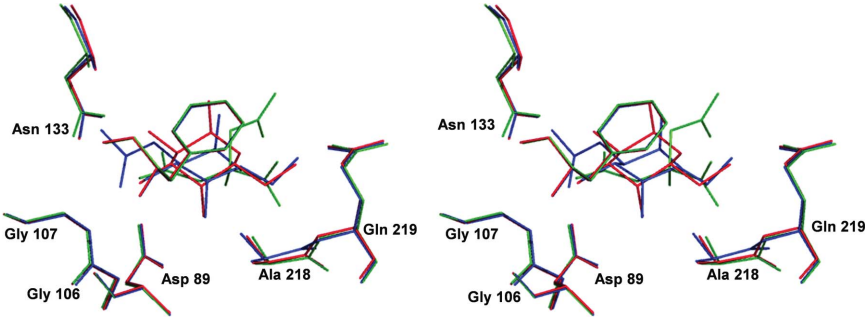

Figure 4
Superposition of the carbohydrate-binding sites of mutant EcorL complexes with galactose (red), GalNAc (blue) and citrate (green).

The methyl-group atom C8 of the acetamido moiety in EcNGal and MEcNGal is accommodated in the hydrophobic pockets formed by the aromatic residues Tyr108 and Trp135 [the atoms of both residues within 4.5 Å of C8 are Tyr108 CE2, 3.95 Å (4.26 Å); Tyr108 CZ, 4.20 Å (4.35 Å); Trp135 CH2, 3.80 Å (3.90 Å); Trp135 CZ3, 4.26 Å (4.30 Å); the values in parentheses correspond to the MEcNGal complex]. This hydrophobic interaction buries 20 Å^2 of the nonpolar area of both Tyr108 and Trp135 in the MEcNGal complex and 18 Å^2 in the EcNGal complex.

In addition to interacting with the protein atoms within the protomer, the O1 and O6 atoms of both Gal and GalNAc are involved in water-mediated interactions with the NE2 and OE1 atoms, respectively, of Gln122 of a symmetry-related molecule; these interactions arise owing to crystal packing and may not have any implications in lectin–sugar binding.

3.2.3. Mutant EcorL–citrate interactions. While attempting to cocrystallize the mutant EcorL in complex with galactose, up to a 20-fold molar excess of the sugar was added to the protein solution. However, on solving the structure it was noticed that galactose did not bind to EcorL as anticipated. Instead, a citrate ion from the additive solution appeared in the binding site bound to the lectin with clear and unambiguous electron density at the location usually occupied by bound sugars.

Interestingly, as observed in other complexes reported here and elsewhere, the framework residues belonging to all four carbohydrate-binding loops (*A*, *B*, *C* and *D*) are involved in hydrogen-bonding interactions with citrate (Fig. 5*c*, Table 2). The positions of the residues that are involved in hydrogen bonding to citrate are essentially similar to those of GalNAc as well as Gal complexes. Four atoms, C5, C4, C3 and C6, of the citrate ion occupy approximately the same positions as C3, C4, C5 and C6 of galactose, respectively. Atom O4 of the citrate ion occupies the same site as O3 of galactose and is involved in hydrogen-bonding interactions with the side-chain atoms Asp89 OD1 and Asn133 ND2 and the main-chain atom Gly107 N. Similar types of hydrogen bonds are observed in Gal and GalNAc complexed with the mutant lectin. The additional hydrogen bonds in the MEc-Cit complex are O1···NE2 Gln219, O3···ND2 Asn133, O5···N Ala218, O5···N Gln219 and O7···N Ala218. The conformation of Gln219 is the same as in MEcGal and MEcNGal and interestingly there is a direct hydrogen bond between the O1 atom of the citrate and the side-chain atom Gln219 NE2. The citrate ion makes three additional direct hydrogen bonds compared with those made by Gal in the MEcGal complex.

In addition to the above interactions as observed in other EcorL complexes, three water-mediated hydrogen bonds are also present in the citrate complex. These are between O3 and O4 of citrate and Gly107 N

Table 2
Mutant EcorL–carbohydrate/citrate interactions including water bridges.

Values in parentheses are the corresponding distances in the wild-type EcorL complexes (PDB codes 1axz and 1ax0).

Ligand atom	Ligand–water distance (Å)	Water	Ligand/water–protein distance (Å)	Protein atom
Gal				
O2	2.71	W390	2.91	Gly107 N
O3			2.77 (2.73)	Asp89 OD1
			2.88 (2.95)	Gly107 N
			3.38 (2.94)	Asn133 ND2
	2.99	W390	2.91	Gly107 N
O4			2.57 (2.76)	Asp89 OD2
			3.19 (3.03)	Ala218 N
O6			— (2.96)	Gln219 NE2
	2.78 (2.72)	W247	2.78 (2.78)	Leu86 O
GalNAc				
O1	3.03	W381		
O3			2.56 (2.67)	Asp89 OD1
			3.03 (2.94)	Gly107 N
			2.94 (2.93)	Asn133 ND2
O4			2.56 (2.76)	Asp89 OD2
			3.46 (3.04)	Ala218 N
O6	3.09	W381		
			— (2.81)	Gln219 NE2
	2.63 (2.71)	W247	2.83 (2.76)	Leu86 O
O7			2.95	Gly107 N
	2.84	W373	2.83	Gly105 O
N2	— (2.93)	W272	3.25 (2.80)	Asn133 ND2
Citrate				
O1			2.86	Gln219 NE2
O3	2.68	W401	3.07	Gly107 N
			3.19	Asn133 ND2
O4			2.57	Asp89 OD1
			3.06	Gly107 N
			2.73	Asn133 ND2
	3.17	W401	3.07	Gly107 N
O5			3.13	Ala218 N
			3.00	Gln219 N
O6	2.89	W247	2.80	Leu86 O
O7			3.25	Ala218 N

and between O6 of citrate and Leu86 O. The water molecules involved in the water-mediated hydrogen bonds in the MEc-Cit complex occupy the same positions as in the MEcGal complex, while in the MEcNGal complex atom O7 of GalNAc shares an approximately similar position with one of the water molecules that is involved in the water-mediated hydrogen bonds to Gly107 N. The O6 atom of the citrate also interacts with the OE1 atom of a symmetry-related Gln122 residue in the same manner as the O6 atoms of Gal and GalNAc. In addition to the hydrogen bonds, van der Waals interactions with Phe131 contribute to binding. Binding of the citrate ion buries 30 Å^2 of the nonpolar area of Phe131 as in the MEcGal and MEcNGal complexes.

3.3. Correlation of the thermodynamic data for the binding of Gal and GalNAc to wild-type EcorL with those for mutant EcorL

Isothermal calorimetric experiments were carried out to find the affinity of the mutant EcorL for Gal and GalNAc (Fig. 6). The thermodynamic parameters ΔG_b^o, ΔH_b^o and $T\Delta S$ at 286 K are −12.3, −10.9 and 1.4 kJ mol^{-1} for MEcGal and −18.0, −25.0 and −7.0 kJ mol^{-1} for MEcNGal, respectively. The corresponding values reported for wild-type EcorL are −18.5, −14.9 and 3.6 kJ mol^{-1} for Gal binding and −18.5, −23.4 and −4.9 kJ mol^{-1} for GalNAc binding, respectively (Surolia *et al.*, 1996). It has previously been demonstrated that Gal and GalNAc bind to wild-type EcorL and to recombinant EcorL with equal affinity (Adar & Sharon, 1996). The change in the enthalpy (ΔH_b^o) is nearly the same in the complexes of GalNAc with wild-type and mutant EcorL protein owing to the conservation of the number of hydrogen bonds. Interestingly, while five of these direct hydrogen bonds are conserved, the sixth differs between the wild-type and mutant EcorL complexes. The hydrogen bond between O6 and Gln219 NE2 observed in EcNGal is lost, but a new hydrogen bond between O7 and Gly107 N is observed in MEcNGal. The difference in the enthalpy change between the complex of Gal with the mutant lectin and that with the wild-type lectin is 4 kJ mol^{-1}. This difference is mainly a consequence of the absence of a direct hydrogen bond between O6 of Gal and Gln219 NE2 in the mutant caused by differences in the rotamer conformation of the side chain of Gln219. The entropy decrease observed in the mutant complexes could be attributed to the additional ordered water molecules at the binding site and the conformational changes observed in the mutant lectin when compared with the wild-type EcorL structure. The change in the binding free energy (ΔG_b^o) is about the same for the EcNGal and MEcNGal complexes, while it is better for Gal in the wild-type lectin, leading to a 13-fold decrease in the Gal-binding affinity to the mutant EcorL (the K_b values are 2.41 and $0.18 \times 10^3\ M^{-1}$ for EcGal and MEcGal and 2.32 and $1.94 \times 10^3\ M^{-1}$ for EcGalNAc and MEcGalNAc, respectively).

4. Discussion

The formulation of the present problem on the modification of the preferential/differential carbohydrate binding of EcorL by mutagenesis originates from our previous and ongoing studies on the structure, activity and mutagenesis of legume lectins. Many significant results related to the carbohydrate binding and quaternary association of peanut lectin (Banerjee *et al.*, 1994, 1996; Ravishankar *et al.*, 1999, 2001; Kundhavai Natchiar *et al.*, 2004; Natchiar *et al.*, 2006; Vijayan, 2007), winged bean lectins (Prabu *et al.*, 1998; Manoj *et al.*, 1999, 2000; Kulkarni *et al.*, 2005, 2006, 2007, 2008) and EcorL (Kulkarni *et al.*, 2004) emerged from these studies, which provided structural insights into the factors governing these features of legume lectins in general (Prabu *et al.*, 1999; Chandra *et al.*, 2001; Manoj & Suguna, 2001). These lectins are Gal/GalNAc-specific, with varying differential affinities for the two sugars. While PNA binds to Gal but not to GalNAc, the winged bean lectins have higher affinity for GalNAc and EcorL has almost equal affinities for the two sugars. The crystal structures have revealed the basis for the observed carbohydrate preferences at the atomic level, leading to further investigations.

The analysis that we have carried out on the winged bean agglutinins WBAI and WBAII is of direct relevance to the present study. Both of these lectins form handshake types of dimers similar to EcorL and have a higher affinity for GalNAc

than Gal. Of particular interest is the extensive study on the structure of WBAI in complex with several monosaccharides to trisaccharides, including those from the A and B blood group antigens (Prabu *et al.*, 1998; Manoj *et al.*, 2000; Kulkarni *et al.*, 2005, 2006, 2007, 2008), and the subsequent analysis resulting in structural details of Gal/GalNAc-binding modes in legume lectins, a brief account of which is given here.

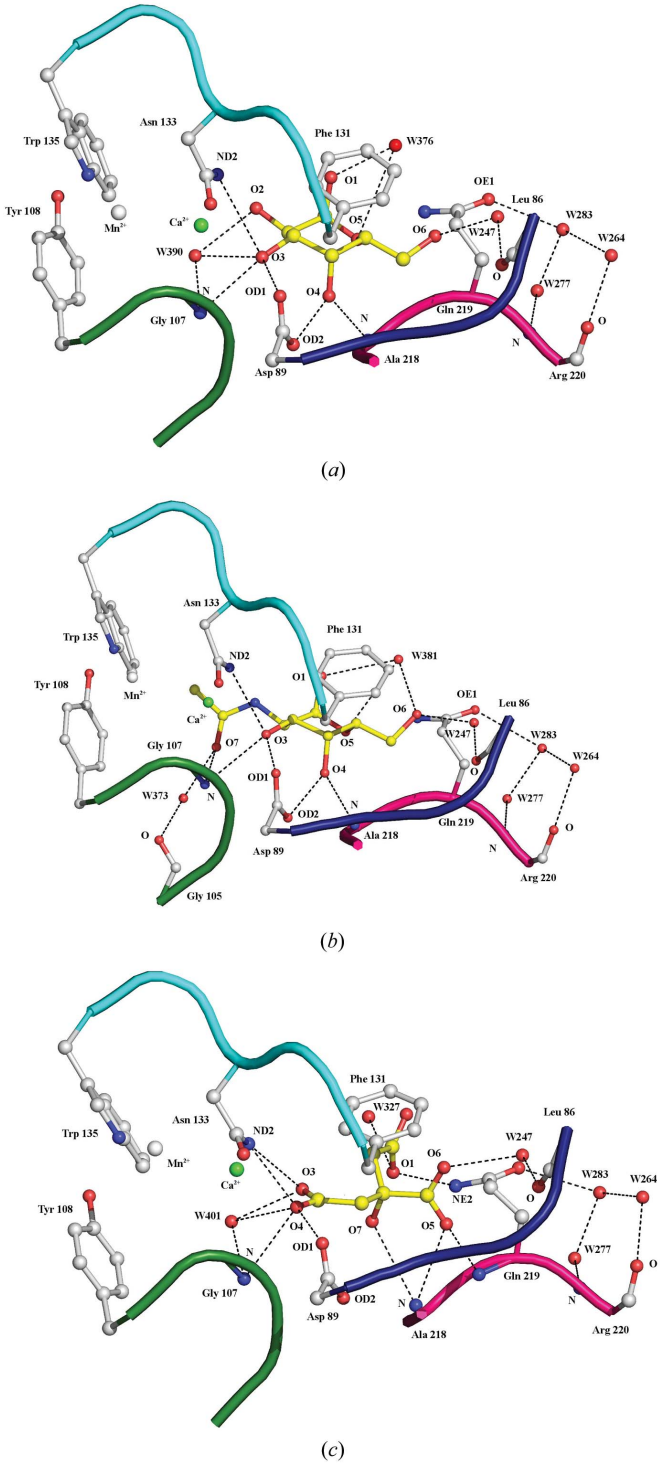

Figure 5
Interactions of mutant EcorL with (*a*) galactose, (*b*) GalNAc and (*c*) citrate. The ligands are shown in yellow and the carbohydrate-binding loops are shown as coils.

The diminished affinity of PNA for GalNAc has been attributed to the presence of Glu129, the side chain of which prevents substitutions at the C2 atom of the Gal moiety. However, GalNAc binding by PNA was achieved by mutating Gln129 to the shorter aspartate residue (Sharma *et al.*, 1998). Other successful mutagenesis studies on PNA include replacing the residues Leu212 and Asn41 to enhance its preferential binding to the tumour-associated disaccharide T-antigen (Sharma *et al.*, 1996; Adhikari *et al.*, 2001) in order to improve its efficiency in the detection of malignant cells.

Of the other galactose-binding lectins, PHAL and GSIV are specific for polysaccharides. SBA also has a high affinity for complex sugars, but at the monosaccharide level it has a higher affinity towards GalNAc compared with Gal. In addition to EcorL and WBAII, crystal structures of DBL, VVLB4 and RBL are available with GalNAc bound in the primary binding site. In DBL, which binds to both Gal and GalNAc, Gal binding is reduced because of the lack of the aromatic residue, which is replaced by Leu. However, DBL has a higher affinity for GalNAc than for Gal, as in the cases of VVL, RBP and WBAI. EcorL is distinct from the others as it binds to Gal and GalNAc with almost equal affinity, which is attributed to the presence of a tyrosine residue at 106, which is a unique feature of EcorL. Mutating this tyrosine to a glycine appeared to be the obvious choice in order to enhance its affinity for GalNAc. The result of this mutation, as revealed by the present work, is that although its affinity for Gal is reduced compared with that of the wild-type lectin owing to the change in the conformation of the *D* loop, the mutant EcorL binds GalNAc more efficiently than Gal. A modified strategy to retain the native conformation of the *D* loop upon mutagenesis should further increase the binding affinity of EcorL for both sugars.

The structures of wild-type and mutant EcorL are the same except at two loops involved in crystal contacts and at the *D* loop of the carbohydrate-binding site. The structural differences observed in the *D* loop are of particular significance as they influence the sugar-binding affinities of the lectin. Moreno *et al.* (1997) demonstrated by molecular-dynamics simulations that in the wild-type EcorL structure one of carbohydrate-binding residues, Gln219, can adopt different side-chain conformations. They also showed that the *D* loop undergoes a conformational change in the Y106G mutant, whereas in the wild-type EcorL structure the loop remains fixed during simulations, suggesting that the tyrosine residue is essential to hold this loop in position. However, the effect of these structural changes on Gal/GalNAc binding of the lectin was not explored. The present crystal structures of the mutant EcorL not only confirmed the theoretically predicted flexible nature of these regions but also provided precise details of the conformational changes. In addition to the structural change in the loop, two further factors contribute to the altered sugar specificities of the mutant. One is the movement of the Gal moiety towards the protein and the other is the addition of new water molecules in the mutant structure that occupy the space created owing to the absence of the tyrosine residue.

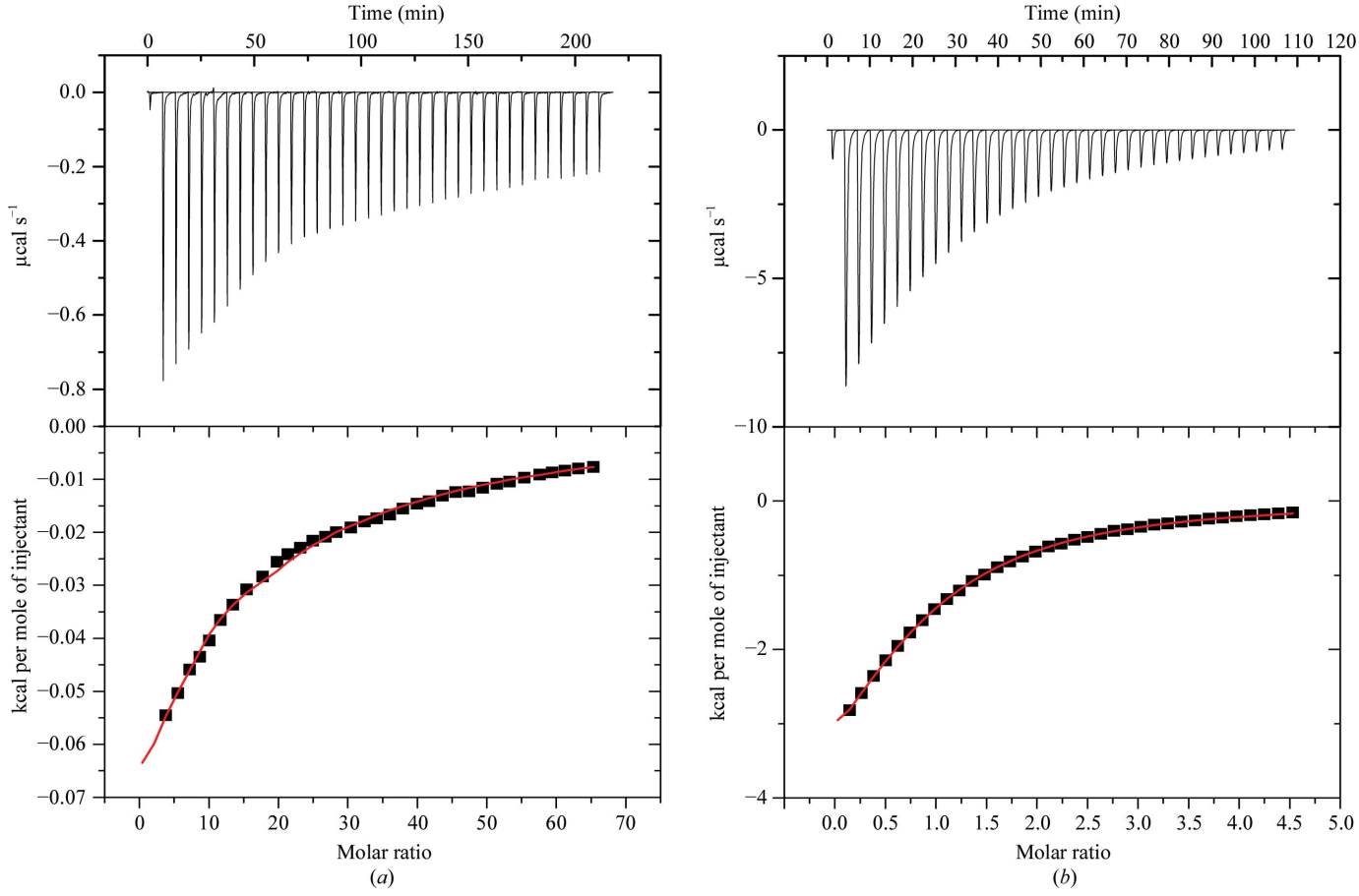

Figure 6
Isothermal calorimetric titration of EcorL Y106G with (*a*) α-D-galactose and (*b*) *N*-acetyl-D-galactosamine. Raw data were obtained after 8 µl injections of 13 and 90 m*M* *N*-acetyl-D-galactosamine and α-D-galactose into 0.624 and 0.607 m*M* EcorL Y106G, respectively (top). A nonlinear least-squares fit (lines) of the heat released as a function of the added ligand (filled squares) for the titration (bottom) is also shown. 1 cal = 4.186 J.

One of these waters, W390, makes a water-mediated hydrogen bond between the sugar and the N atom of Gly107 in MEcGal, whereas in MEcNGal it is replaced by the O7 atom, which makes a direct hydrogen bond to the same atom of the protein. The formation of this crucial O7···N Gly107 hydrogen bond in the mutant is possible because of the movement of the Gal moiety as the O7 atom moves about 1 Å closer to the N atom of Gly107. This interaction is mainly responsible for the higher affinity of the mutant lectin for GalNAc compared with Gal. This change in the differential affinities for the two sugars has been measured in solution by ITC experiments. While the solution studies confirm that the aim of taking up this particular work has been achieved, the structures provide a direct visualization of the atomic basis of the observed results.

To determine the effect of the mutation on the binding of higher saccharides, structures of wild-type EcorL in complex with the disaccharides lactose and *N*-acetyllactosamine and of rEcorL complexed with lactose were compared with those of the complexes of mutant EcorL. A superposition of these structures revealed that the second sugar moieties cannot be accommodated in the combining site owing to steric hindrance from the side chain of Gln219, thus preventing the binding of lactose and its derivatives to mutant EcorL in its present conformation.

5. Conclusions

A point mutation of EcorL was generated in order to alter its differential binding affinity for Gal and GalNAc. Its crystal structure has been determined in complex with these sugars and its carbohydrate-binding properties were studied by ITC experiments. Crystallographic analysis of this mutant, Y106G, of EcorL revealed that a concerted reorganization of the protein, sugar and water molecules at the binding site resulting from the mutation enhanced the affinity of the lectin for GalNAc compared with that for Gal. The present study has not only provided conclusive evidence for the modification of the differential affinity for Gal and GalNAc of EcorL, but has also presented the first three-dimensional image of the altered binding site of a legume lectin. The results of this study provide useful guidelines for the formulation of future work of a similar nature.

We thank N. Sharon for providing the wild-type EcorL clone. 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) 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 Department of Biotechnology (DBT). MV has been successively supported by the Distinguished Biotechnology Professorship and the Homi Bhabha Professorship of the Department of Atomic Energy. Financial support from the DST is acknowledged.

References

Adar, R. & Sharon, N. (1996). *Eur. J. Biochem.* **239**, 668–674.
Adhikari, P., Bachhawat-Sikder, K., Thomas, C. J., Ravishankar, R., Jeyaprakash, A. A., Sharma, V., Vijayan, M. & Surolia, A. (2001). *J. Biol. Chem.* **276**, 40734–40739.
Arango, R., Rodriguez-Arango, E., Adar, R., Belenky, D., Loontiens, F. G., Rozenblatt, S. & Sharon, N. (1993). *FEBS Lett.* **330**, 133–136.
Arango, R., Rozenblatt, S. & Sharon, N. (1990). *FEBS Lett.* **264**, 109–111.
Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1996). *J. Mol. Biol.* **259**, 281–296.
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.
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.
Chandra, N. R., Prabu, M. M., Suguna, K. & Vijayan, M. (2001). *Protein Eng.* **14**, 857–866.
Cohen, G. H. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *PyMOL.* http://www.pymol.org.
Elgavish, S. & Shaanan, B. (1998). *J. Mol. Biol.* **277**, 917–932.
Hoedemaeker, F. J., Van Eijsden, R. R., Díaz, C. L., De Pater, B. S. & Kijne, J. W. (1993). *Plant Mol. Biol.* **22**, 1039–1046.
Hubbard, S. J. & Thornton, J. M. (1993). *NACCESS.* Department of Biochemistry and Molecular Biology, University College London.
Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
Jordan, E. T. & Goldstein, I. J. (1995). *Eur. J. Biochem.* **230**, 958–964.
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., Katiyar, S., Surolia, A., Vijayan, M. & Suguna, K. (2008). *Acta Cryst.* D**64**, 730–737.
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.
Kundhavai Natchiar, S., Arockia Jeyaprakash, A., Ramya, T. N. C., Thomas, C. J., Suguna, K., Surolia, A. & Vijayan, M. (2004). *Acta Cryst.* D**60**, 211–219.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). *Biochim. 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.
Manoj, N. & Suguna, K. (2001). *Protein Eng.* **14**, 735–745.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Moreno, E., Teneberg, S., Adar, R., Sharon, N., Karlsson, K. A. & Angström, J. (1997). *Biochemistry*, **36**, 4429–4437.
Natchiar, S. K., Srinivas, O., Mitra, N., Surolia, A., Jayaraman, N. & Vijayan, M. (2006). *Acta Cryst.* D**62**, 1413–1421.
Nishiguchi, M., Yoshida, K., Sumizono, T. & Tazaki, K. (1997). *FEBS Lett.* **403**, 294–298.
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.
Prabu, M. M., Suguna, K. & Vijayan, M. (1999). *Proteins*, **35**, 58–69.
Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1999). *Acta Cryst.* D**55**, 1375–1382.
Ravishankar, R., Thomas, C. J., Suguna, K., Surolia, A. & Vijayan, M. (2001). *Proteins*, **43**, 260–270.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577.
Rodriguez-Arango, E., Arango, R., Adar, R., Galili, G. & Sharon, N. (1992). *FEBS Lett.* **307**, 185–189.
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**, 433–445.
Sharma, V., Vijayan, M. & Surolia, A. (1996). *J. Biol. Chem.* **271**, 21209–21213.
Sharon, N. & Lis, H. (1990). *FASEB J.* **4**, 3198–3208.
Sharon, N. & Lis, H. (1998). *Chem. Rev.* **98**, 637–674.
Sharon, N. & Lis, H. (2004). *Glycobiology*, **14**, 53–62.
Sumathi, K., Ananthalakshmi, P., Roshan, M. N. & Sekar, K. (2006). *Nucleic Acids Res.* **34**, W128–W132.
Surolia, A., Sharon, N. & Schwarz, F. P. (1996). *J. Biol. Chem.* **271**, 17697–17703.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Van Eijsden, R. R., De Pater, B. S. & Kijne, J. W. (1994). *Glycoconj. J.* **11**, 375–380.
Van Eijsden, R. R., Hoedemaeker, F. J., Díaz, C. L., Lugtenberg, B. J., De Pater, B. S. & Kijne, J. W. (1992). *Plant Mol. Biol.* **20**, 1049–1058.
Vijayan, M. (2007). *J. Biosci.* **32**, 1059–1066.
Vijayan, M. & Chandra, N. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.
Zhu, K., Bressan, R. A., Hasegawa, P. M. & Murdock, L. L. (1996). *FEBS Lett.* **390**, 271–274.