

# Carbohydrate Binding, Quaternary Structure and a Novel Hydrophobic Binding Site in Two Legume Lectin Oligomers from *Dolichos biflorus*

Thomas W. Hamelryck<sup>1\*</sup>, Remy Loris<sup>1</sup>, Julie Bouckaert<sup>1</sup>  
Minh-Hoa Dao-Thi<sup>1</sup>, Gerard Strecker<sup>2</sup>, Anne Imberty<sup>3</sup>, Elias Fernandez<sup>4</sup>  
Lode Wyns<sup>1</sup> and Marilynn E. Etzler<sup>4</sup>

<sup>1</sup>Laboratorium voor  
Ultrastructuur, Vlaams  
Interuniversitair Instituut voor  
Biotechnologie, Vrije  
Universiteit Brussel  
Paardenstraat 65, B-1640  
Sint-Genesius-Rode, Belgium

<sup>2</sup>Université des Sciences et  
Technologies de Lille, Bâtiment  
C9, 59655, Villeneuve D'Ascq  
CEDEX, France

<sup>3</sup>CERMAV-CNRS (affiliated to  
the University Joseph Fourier)  
BP 53, F-38041 Grenoble  
Cedex 9, France

<sup>4</sup>Section of Molecular and  
Cellular Biology, University of  
California, Davis  
CA 95616, USA

The seed lectin (DBL) from the leguminous plant *Dolichos biflorus* has a unique specificity among the members of the legume lectin family because of its high preference for GalNAc over Gal. In addition, precipitation of blood group A + H substance by DBL is slightly better inhibited by a blood group A trisaccharide (GalNAc( $\alpha$ 1-3)[Fuc( $\alpha$ 1-2)]Gal) containing pentasaccharide, and about 40 times better by the Forssman disaccharide (GalNAc( $\alpha$ 1-3)GalNAc) than by GalNAc. We report the crystal structures of the DBL-blood group A trisaccharide complex and the DBL-Forssman disaccharide complex.

A comparison with the binding sites of Gal-binding legume lectins indicates that the low affinity of DBL for Gal is due to the substitution of a conserved aromatic residue by an aliphatic residue (Leu127). Binding studies with a Leu127Phe mutant corroborate these conclusions. DBL has a higher affinity for GalNAc because the *N*-acetyl group compensates for the loss of aromatic stacking in DBL by making a hydrogen bond with the backbone amide group of Gly103 and a hydrophobic contact with the side-chains of Trp132 and Tyr104.

Some legume lectins possess a hydrophobic binding site that binds adenine and adenine-derived plant hormones, i.e. cytokinins. The exact function of this binding site is unknown, but adenine/cytokinin-binding legume lectins might be involved in storage of plant hormones or plant growth regulation. The structures of DBL in complex with adenine and of the dimeric stem and leaf lectin (DB58) from the same plant provide the first structural data on these binding sites. Both oligomers possess an unusual architecture, featuring an  $\alpha$ -helix sandwiched between two monomers. In both oligomers, this  $\alpha$ -helix is directly involved in the formation of the hydrophobic binding site. DB58 adopts a novel quaternary structure, related to the quaternary structure of the DBL heterotetramer, and brings the number of known legume lectin dimer types to four.

© 1999 Academic Press

**Keywords:** protein-carbohydrate interactions; quaternary structure; legume lectins; cytokinins; X-ray crystallography

\*Corresponding author

Abbreviations used: DB58, horse gram (*Dolichos biflorus*) stem and leaf lectin; DBL, horse gram (*Dolichos biflorus*) seed lectin; Con A, Jack bean (*Canavalia ensiformis*) lectin; EcorL, West Indian coral tree (*Erythrina corallodendron*) lectin; Fuc, L-fucose; Gal, D-galactose; GalNAc, *N*-acetyl-D-galactosamine; Glc, D-glucose; GlcNAc, *N*-acetyl-D-glucosamine; GS4, *Griffonia simplicifolia* (now *Bandeirea simplicifolia*) lectin IV; PHA-L, leucoagglutinating common bean (*Phaseolus vulgaris*) agglutinin; PNA, peanut (*Arachis hypogaea*) agglutinin; r.m.s.d., root-mean-square deviation; SBA, soybean (*Glycine max*) agglutinin; WBAI, basic winged bean (*Psophocarpus tetragonolobus*) agglutinin.

E-mail address of the corresponding author: [thamelry@vub.ac.be](mailto:thamelry@vub.ac.be)

## Introduction

Lectins are a structurally diverse class of proteins that bind carbohydrates in a reversible fashion, without showing enzymatic activity towards these carbohydrates. Lectins are ubiquitous in nature and can be found in animals and plants, but also in bacteria and even in viruses (for excellent reviews with an emphasis on structural features, see Rini, 1995; Weis & Drickamer, 1996; Lis & Sharon, 1998). Lectins are mainly involved in biological recognition functions and play important roles in embryogenesis, cancer, inflammation, immune response and fertilization (Gabiuss & Gabiuss, 1997). Plant lectins are valuable tools in different areas of biological and medical research (Van Damme *et al.*, 1997). A detailed understanding of lectin-carbohydrate recognition mechanisms is therefore of great practical as well as fundamental value.

One of the best-studied lectin families is the legume lectin family (Sharon & Lis, 1990; Loris *et al.*, 1998). Members of the legume lectin family are present in the seeds and the vegetative tissues of leguminous plants. The exact function of the legume lectins *in vivo* has not been established, although considerable attention has been devoted to the possibilities that they may be involved in the defense of plants against predators or in the interaction of the plant with *Rhizobium* symbionts (Chrispeels & Raikhel, 1991; Etzler, 1992; Brewin & Kardailsky, 1997). More recently, a legume lectin domain was found in a receptor-like serine/threonine kinase gene from *Arabidopsis thaliana* (*Brassicaceae*; Hervé *et al.*, 1996). In animals, type I integral membrane proteins with a legume lectin domain are involved in glycoprotein sorting (Fiedler & Simons, 1994, 1995; Itin *et al.*, 1996). Despite the lack of knowledge about their function *in vivo*, the legume lectins are widely used as a model system for studying protein-carbohydrate interactions.

Legume lectins are tetrameric or dimeric proteins whose subunits show high levels of sequential and structural identity. The legume lectins can be roughly divided into Man/Glc, Gal/GalNAc, Fuc, GlcNAc and complex specificity groups (Sharon & Lis, 1990). The latter bind only specific oligosaccharides.

At present, the crystal structures of 12 native or sugar complexed legume lectins have been determined (for a complete list of references up to 1997, see Loris *et al.*, 1998). Five belong to the Man/Glc specificity group: favin from *Vicia faba* (broad bean), concanavalin A from *Canavalia ensiformis* (Jack bean) and a closely related lectin from *Canavalia brasiliensis*, lentil lectin from *Lens culinaris*, pea lectin from *Pisum sativum* and two isolectins from *Lathyrus ochrus* (yellow-flowered pea). Five belong to the Gal/GalNAc specificity group: *Erythrina corallodendron* (coral tree) lectin (Elgavish & Boaz, 1998), soybean agglutinin (*Glycine max*), peanut agglutinin (*Arachis hypogea*) and, more recently, *Vicia villosa* (hairy winter vetch) isolectin B4

(Osinaga *et al.*, 1997) and winged bean agglutinin (*Psophocarpus tetragonolobus*; Prabu *et al.*, 1998). Only two crystal structures of lectins from the complex specificity group have been published: phytohemagglutinin-L (PHA-L) from *Phaseolus vulgaris* (kidney bean) and *Griffonia simplicifolia* lectin IV (GS4). In addition, the crystal structures of arcelin-5 and an  $\alpha$ -amylase inhibitor from the kidney bean, two plant defense proteins that are truncated and hence non-sugar-binding legume lectin homologues, have been determined. Structural information about the Fuc and GlcNAc-specific legume lectins is still lacking, with the exception of a recent molecular modeling study of Fuc binding by *Ulex europaeus* lectin I (Gohier *et al.*, 1996). Despite the abundance of crystal structures, structural information on specific oligosaccharide binding by legume lectins is relatively limited. Structures of only four legume lectins in complex with a specifically recognised oligosaccharide have been published. These include Con A in complex with a trisaccharide and a pentasaccharide (Moothoo & Naismith, 1998), *L. ochrus* lectin with a fucosylated and a non-fucosylated complex type sugar, *G. simplicifolia* lectin IV with the Le<sup>b</sup> and Le<sup>y</sup> tetrasaccharides and peanut agglutinin with the T-antigen disaccharide.

The *Dolichos biflorus* seed lectin (DBL) is one of at least four blood group A + H substance-binding lectins present in this plant (Etzler, 1996). It is a tetrameric glycoprotein with a molecular mass of 110 kDa that is composed of two types of subunits, designated subunit I and II (Carter & Etzler, 1975). Subunit II (241 amino acid residues) is post-translationally formed from subunit I (253 amino acid residues) by the removal of 12 residues from its C terminus (Roberts *et al.*, 1982; Young *et al.*, 1995). DBL agglutinates only epitopes with terminal non-reducing GalNAc residues and is unique among the GalNAc-binding legume lectins in its extreme preference for GalNAc over Gal (Etzler & Kabat, 1970; Hammarström *et al.*, 1977). Me- $\alpha$ -D-GalNAc is a twofold better inhibitor of binding in a solid phase assay than GalNAc (Etzler, 1994b) and binds to DBL with an association constant of  $4.2 \times 10^3$  M<sup>-1</sup> (Etzler *et al.*, 1981). Me- $\alpha$ -D-GalNAc, GalNAc( $\alpha$ 1-3)Gal and GalNAc( $\alpha$ 1-3)Gal( $\beta$ 1-3)GlcNAc are equally active as inhibitors in precipitation assays (Etzler & Kabat, 1970). However, a GalNAc( $\alpha$ 1-3)[Fuc( $\alpha$ 1-2)]Gal-containing pentasaccharide is a slightly better inhibitor, presumably due to the presence of the ( $\alpha$ 1-2)-linked fucose residue. Subsequent studies showed that the Forssman pentasaccharide GalNAc( $\alpha$ 1-3)GalNAc( $\beta$ 1-3)Gal( $\alpha$ 1-4)Gal( $\beta$ 1-4)Glc and the Forssman disaccharide GalNAc( $\alpha$ 1-3)GalNAc are 60 and 40-fold better inhibitors than GalNAc, respectively (Baker *et al.*, 1983). The Forssman disaccharide is present in tumor-associated Forssman glycolipid (Hakomori, 1984) and in lipopolysaccharides from certain pathogenic bacteria, e.g. *Vibrio mimicus* (Landersjö *et al.*, 1998).

Some legume lectins possess a hydrophobic binding site that binds adenine and adenine-derived plant hormones, i.e. cytokinins (Roberts & Goldstein, 1983). Binding of adenine does not interfere with carbohydrate binding (Gegg *et al.*, 1992) and has been described mainly for tetrameric legume lectins, including PHA-L and PHA-E from common bean, soybean agglutinin (SBA) and lima bean lectin (Roberts & Goldstein, 1983; Maliarik & Goldstein, 1988; Maliarik *et al.*, 1989), hog peanut lectin (Maliarik *et al.*, 1987) and *D. biflorus* seed lectin (Gegg *et al.*, 1992). The dimeric stem and leaf lectin from *D. biflorus* (DB58) also binds adenine (Gegg *et al.*, 1992). Binding studies indicate that the mentioned lectins bind adenine and cytokinins with an affinity ( $K_a = 10^5 - 10^6 \text{ M}^{-1}$ ) that is two to three orders of magnitude higher than their typical affinity for monosaccharides. The stoichiometry of binding is two adenine molecules per tetramer or one per dimer (Gegg *et al.*, 1992). The preservation of this binding site among legume lectins with different carbohydrate specificities isolated from different plant species suggests that it might be of crucial importance for understanding the role of the legume lectins *in vivo*.

We present the structures of DBL in complex with the blood group A trisaccharide (GalNAc( $\alpha$ 1-3)[Fuc( $\alpha$ 1-2)]Gal) and in complex with the Forssman disaccharide (GalNAc( $\alpha$ 1-3)GalNAc). We present the complex of DBL with adenine, and

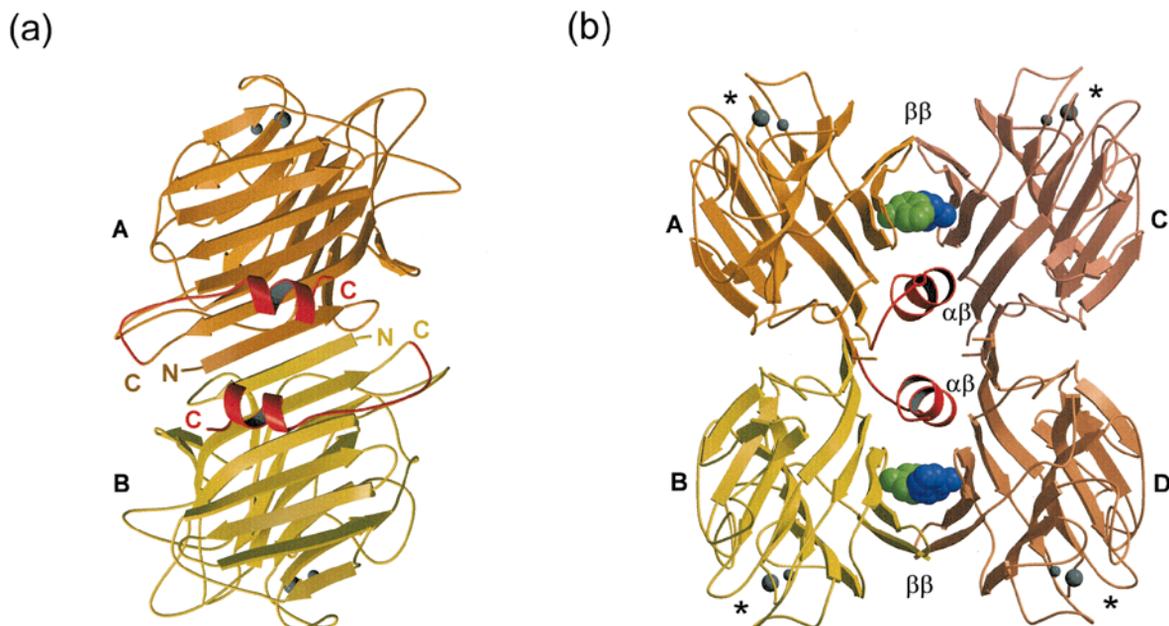
show that there is an intricate relationship between quaternary structure and adenine binding. In addition, we have determined the structure of the adenine-binding stem and leaf lectin DB58 (Etzler, 1994a) from the same plant. This heterodimeric lectin adopts a novel quaternary structure, related to the structure of the DBL tetramer, in which the hydrophobic binding site is preserved.

## Results and Discussion

### Overall structure of DBL

The overall structure of DBL in complex with adenine is shown in Figure 1(b). The DBL heterotetramer consists of two intact subunits (253 residues, subunits A and B in Figure 1) and two subunits with a post-translationally truncated C-terminal end (241 residues, subunits C and D in Figure 1: Schnell & Etzler, 1987; Young *et al.*, 1995). The C-terminal residues of the intact subunits are visible in the electron density, while the last six C-terminal residues of the truncated subunits are not visible.

This quaternary structure was first described for PHA-L (Hamelryck *et al.*, 1996) and is also adopted by SBA (Dessen *et al.*, 1995) and *V. villosa* isolectin B4 (Osinaga *et al.*, 1997). Like DBL, the SBA tetramer consists of C-terminally truncated subunits and intact subunits (Mandal *et al.*, 1994). This tetra-



**Figure 1.** (a) The canonical lectin dimer present in the DBL tetramer. The two subunits are shown in different colors. The sandwiched  $\alpha$ -helices (Leu244-Asn251), the linker region (Asp236-Asp243) and the two C-terminal residues (Val252 and Leu253) are shown in red. The part shown in red is present only for the intact subunits. The C-terminal ends of the truncated and the intact subunits are indicated in yellow and red, respectively.  $\text{Ca}^{2+}$  and  $\text{Mn}^{2+}$  are shown as large and small gray spheres, respectively. (b) The overall structure of DBL in complex with adenine. Each subunit is shown in a different color. The four observed adenine molecules are shown as space-filling models in green and blue. The two types of dimer-dimer interfaces (the  $\beta\beta$ -interface and the  $\alpha\beta$ -interface) are indicated. Dimer AB is shown in the same colors as in (a). Metal ions as in (a). The locations of the four sugar-binding sites are indicated with an asterisk (\*).

mer type consists of two canonical legume lectin dimers (dimers AB and CD in Figure 1(b)) that pack against each other in a parallel fashion. A first interface between the two canonical dimers is present at the outer ends of the tetramer (Figure 1(b)). This interface, which we will call the  $\beta\beta$ -interface, consists of two  $\beta$ -strands that pack together by the zipper-like intercalation of their side-chains (Ser187, Ile189, Ser191). The  $\beta\beta$ -interface decreases the accessible surface area of the two monomers involved by 1400 Å<sup>2</sup>. Essentially the same interface, involving the same Ser-X-Ile-X-Ser motif, is present in SBA and PHA-L (Hamelryck *et al.*, 1996).

However, we have observed a second, previously undescribed dimer-dimer interface. The above-described mode of association creates a large channel running through the center of the tetramer. In the case of SBA and PHA-L, two stretches of uninterpretable electron density were reported to be present in this channel, presumably due to the C-terminal regions of two of the four subunits (Dessen *et al.*, 1995; Hamelryck *et al.*, 1996). The corresponding densities in the channel of the DBL tetramer could be interpreted as two  $\alpha$ -helices (Leu244-Asn251), each formed by the C-terminal part of an intact subunit (Figure 1(a) and (b)). Clear density is present for the linker region between Asp235 and Leu244, and the two C-terminal residues (Val252 and Leu253) following the  $\alpha$ -helix. The presence of the  $\alpha$ -helices in the central channel breaks the apparent 222 symmetry of the tetramer, and changes its point group to 2. The two  $\alpha$ -helices are sandwiched between the  $\beta$ -sheets of two pairs of facing monomers (monomer pairs AC and BD in Figure 1(b)) and form an important stabilization of the tetramer. We will refer to this novel second interface as the  $\alpha\beta$ -interface. A comparable architecture has been observed in a thiolase from yeast (Mathieu *et al.*, 1997), but in that case the sandwiched  $\alpha$ -helices belong to an interdomain interface. The unique architecture of the DBL tetramer explains why two truncated and two intact subunits are present per tetramer: the two subunits that have their C-terminal region buried in the central channel remain intact, while the C-terminal regions of the two remaining subunits are proteolytically processed *in vivo* (Schnell & Etzler, 1987; Young *et al.*, 1995). There are some indications that C-terminal processing of half of the subunits is necessary before the assembly of the tetramer. Purified truncated subunits form aggregates that do not bind to blood group A + H substance (Etzler *et al.*, 1981). The intact subunits expressed in *E. coli* form oligomers with an anomalous molecular mass (Chao *et al.*, 1994). Crystals grown from recombinant DBL containing only intact subunits contain a mixture of intact and truncated subunits (Dao-Thi *et al.*, 1998).

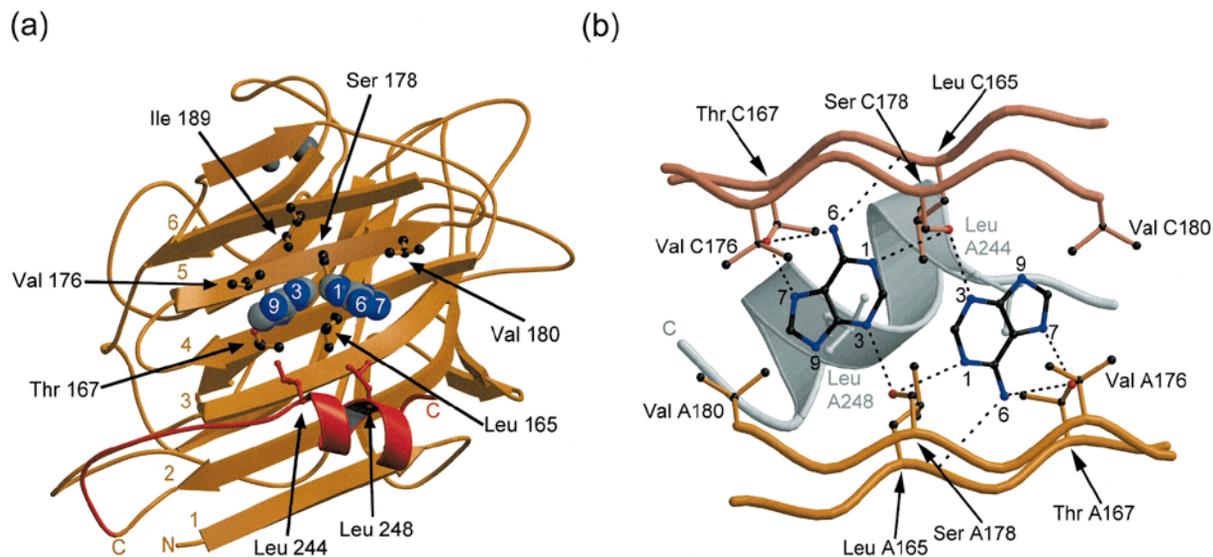
The side-chains of residues Ser246 and Arg250 of both helices protrude into a large, water-accessible cavity situated in the center of the tetramer. The side-chain of Arg250 is completely disordered and

is not visible in the electron density. On the opposite side of the  $\alpha$ -helices, the side-chains of residues Leu244 and Leu248 form the bottom of a large, hydrophobic cavity that contains density for two neighboring adenine residues (see below). For both tetramers in the asymmetric unit, the two intact subunits belong to a single canonical dimer (dimer AB in Figure 1(b)) and the two truncated subunits belong to a second canonical dimer (dimer CD in Figure 1(b)). A possible explanation for this fact is that an alternative arrangement (i.e. two canonical dimers each consisting of an intact and a truncated subunit) would bring the two Arg250 side-chains into the vicinity of each other, thereby creating an unfavorable interaction.

For the blood group A trisaccharide complex, one canonical dimer is present in the asymmetric unit and the complete tetramer is generated by the C2 symmetry operations. Similarly, for the Forssman disaccharide complex, the complete tetramer is generated by applying the  $I4_122$  symmetry operations to the single monomer in the asymmetric unit. This means that in both cases the electron density in the center of the tetramer is averaged because of statistical disorder in the crystal. Therefore, two C-terminal stretches (Asp236-Leu253) with half occupancy are present in the former case, while one C-terminal stretch with half occupancy is present in the latter case.

### Structure of the adenine-binding site

Two identical adenine-binding cavities with a hydrophobic character are found at opposite ends of the tetramer (Figure 1(b)). The top of this cavity is formed by the  $\beta\beta$ -interface (Ile189), while the bottom is formed by side-chains that protrude from the sandwiched  $\alpha$ -helices (Leu244 and Leu248; Figure 2(a)). The side-chains of the two C-terminal residues (Val252 and Leu253) also add to the hydrophobic character of the cavity. The walls are formed by the protruding side-chains from the back sheets of a pair of facing monomers (monomer pairs AC and BD in Figure 1(b)). All the side-chains that form the cavity are aliphatic (Leu, Val and Ile) or contain a hydroxyl group (Ser, Thr). The two cavities each have internal pseudo 2-fold symmetry (broken by the sandwiched  $\alpha$ -helices), and they are related to each other by a non-crystallographic 2-fold axis. Each cavity contains electron density for two neighboring adenine molecules. The adenine molecules are bound by hydrogen bonds (Leu165, Thr167, and Ser178; see Table 1) and hydrophobic interactions (Leu165, Val176, Val180, Ile189; see Table 2) with side-chains from  $\beta$ -strands 4, 5 and 6 of the back sheet (Figure 2(a) and (b)). In addition, two residues (Leu244 and Leu248) belonging to the sandwiched  $\alpha$ -helix are directly below the adenine rings. All nitrogen atoms of the adenine molecules are involved in hydrogen bonds, except N-9. Weak density in the neighborhood of the N-9 atom indicates that it is probably involved in a water bridge with the



**Figure 2.** (a) A view on a pair of neighboring adenine-binding sites. The adenine molecules are shown as space-fill models. Nitrogen atoms are shown in blue and labeled, carbon atoms in gray and oxygen atoms in red. The six  $\beta$ -strands that form the back sheet are numbered. The complete adenine-binding site is formed by corresponding residues from two facing monomers (monomers pairs AC and BD in Figure 1(b)). In this Figure only the side-chains involved in adenine binding from one monomer (monomer A in Figure 1(b)) and the sandwiched  $\alpha$ -helix (belonging to monomer A) are shown as ball-and-stick representations. Metal ions as in Figure 1(a). (b) A detailed view from the  $\beta\beta$ -interface of the DBL tetramer on the two neighboring adenine-binding sites in dimer AC from Figure 1(b). The upper two strands belong to monomer C, the lower two and the  $\alpha$ -helix belong to monomer A. All residues involved in hydrogen bonds or hydrophobic interactions with an adenine molecule are shown as ball-and-stick models and labeled, except IleA189 and IleC189, which are omitted for clarity. The two adenine molecules are shown as ball-and-stick representations. Carbon atoms are shown in gray, oxygen atoms in red and nitrogen atoms in blue. The nitrogen atoms of the adenine molecules are labeled. Hydrogen bonds are shown as broken lines. The sandwiched  $\alpha$ -helix that forms the bottom of the adenine-binding site is shown in gray. Note that the side-chain of Ser178 can be involved in only one of the two hydrogen bonds shown at the same time.

main-chain oxygen atom of Glu136. The position of the adenine-binding site is in accordance with the results of photoaffinity labeling experiments (Maliarik & Goldstein, 1988; Gegg & Etzler, 1994) and the putative location of the adenine-binding site in PHA-L and SBA as previously suggested by us (Hamelryck *et al.*, 1996).

We observe four adenine molecules with half occupancy, in accordance with binding studies indicating that DBL binds two adenine molecules per tetramer (Gegg *et al.*, 1992). In each cavity, two facing Ser178 residues are in the vicinity of the two neighboring adenine molecules. The posi-

tion of each Ser178 side-chain oxygen atom is compatible with a role as hydrogen bond donor to either N-1 of one adenine residue, or N-3 of the neighboring adenine residue (Figure 2(b)). The average distance in both cases is 3.3 Å. Hydrogen bond donor and acceptor sites in adenine are well defined (Jeffrey & Saenger, 1991). N-1 and N-3 act only as hydrogen bond acceptors, and since the Ser residue can serve only as a donor in one hydrogen bond, it cannot be involved in these two hydrogen bonds at the same time. This explains why only two adenine molecules bind per tetramer, despite the presence of four (two in each cavity) potential adenine-binding sites. Upon binding of one adenine molecule in a cavity, one Ser178 hydroxyl group hydrogen bonds to N-1, while the other Ser178 residue hydrogen bonds to N-3. In the same cavity, a second adenine molecule then necessarily binds with a much lower affinity, because two hydrogen bond donors are unavailable. Hence, each cavity will bind only one adenine molecule, bringing the number of bound adenine molecules per tetramer to two. Due to statistical disorder in the crystal, electron density is present for four adenine molecules, instead of two.

The two neighboring adenine-binding sites in a cavity are not completely equivalent. The two Leu

**Table 1.** Hydrogen bonds and average distances between hydrogen bond donors and acceptors between DBL and adenine for the eight equivalent adenine binding sites in the asymmetric unit

Adenine atoms	DBL	Distance (Å)
N-1	*Ser178 O $\gamma$	3.3
N-3	Ser178 O $\gamma$	3.3
N-6	Leu165 O	3.5
N-6	Thr167 O $\gamma$	2.8
N-7	Thr167 O $\gamma$	2.9

Three of the four hydrogen bond partners belong to the same subunit; the asterisk in front of Ser178 indicates it belongs to the facing subunit (Figure 2(b)).

**Table 2.** List of DBL residues with atoms within 4.6 Å of an adenine molecule in each of the eight equivalent adenine-binding sites in the asymmetric unit

Adenine atom	DBL residue (number of atoms within 4.6 Å)
N-1	Leu165 (5), Val176 (4), Ser178 (3), Ala177 (1)
C-2	Leu165 (2), Val176 (1), Ser178 (2), Ile189 (1), *Leu165 (1), *Ser178 (1)
N-3	Leu165 (1), Val176 (1), Ser178 (1), Ile189 (1), *Leu165 (1), *Ser178 (2)
C-4	Val176 (1), *Leu165 (1), *Ser178 (1)
C-5	Leu165 (1), Thr167 (3), Val176 (3)
C-6	Leu165 (3), Thr167 (4), Val176 (5)
N-7	Thr167 (3), Val176 (3)
C-8	Thr167 (1)
N-9	*Leu165 (2), *Ser178 (1), *Val180 (1)

Most residues belong to one subunit; an asterisk in front of a residue indicates that it belongs to the facing subunit. The number of atoms within 4.6 Å is shown between parentheses.

side-chains that protrude from the sandwiched  $\alpha$ -helix (Leu244 and Leu248), and form the bottom of the adenine-binding sites, are each in the vicinity of one adenine molecule (Figure 2(b)). However, the side-chain of Leu244 is considerably closer to the plane of the five-membered ring of the adenine molecule than the side-chain of Leu248 (4.5 Å and 5.5 Å, respectively).

### Structure of DB58, a vegetative legume lectin from *Dolichos biflorus*

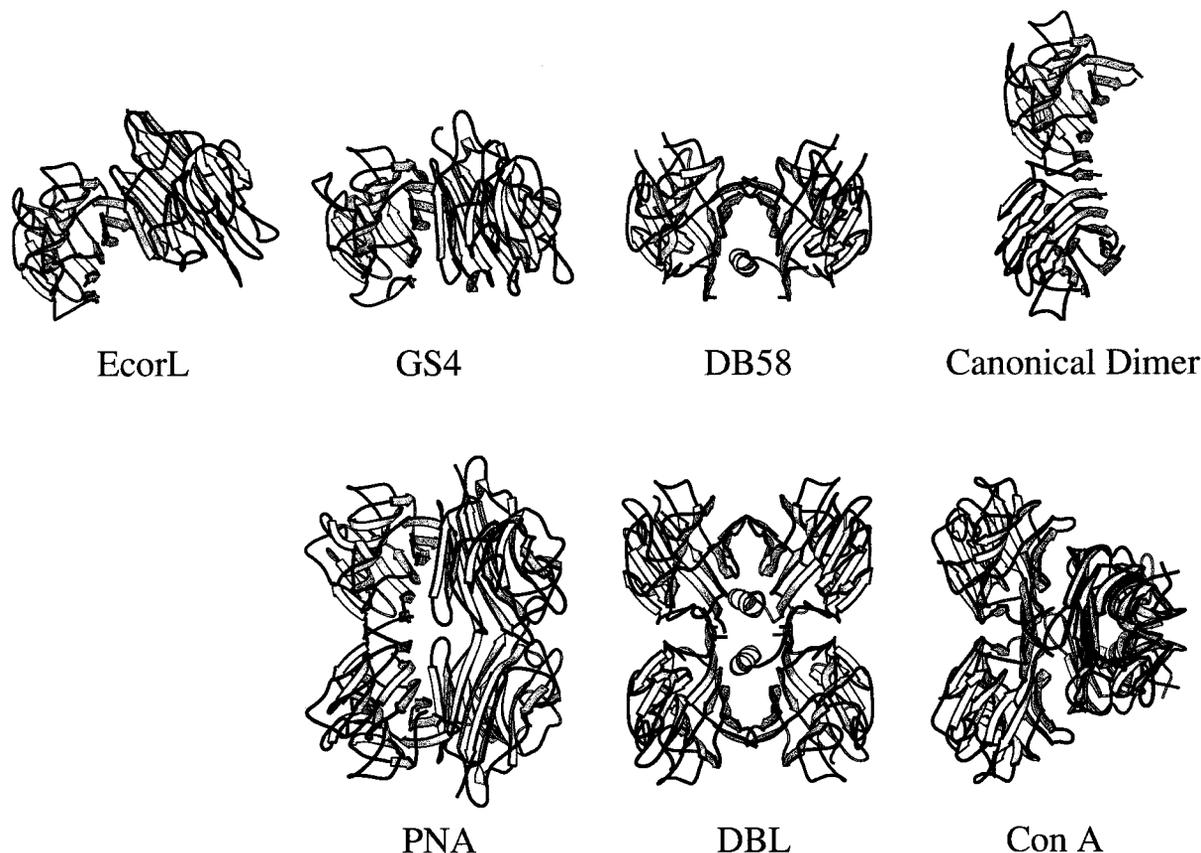
Our results indicate that adenine binding by a legume lectin depends on the tetrameric structure shared by PHA-L, SBA and DBL. However, the stem and leaves of *D. biflorus* also contain a 58 kDa dimeric adenine-binding legume lectin (DB58; Etzler, 1994a). DB58 is a heterodimer of an intact (253 residues) and a truncated subunit (241 or 242 residues), and has 87% sequence identity with DBL. DB58 proved to be very difficult to crystallize, but a 3.3 Å dataset of uncomplexed DB58 (Dao-Thi *et al.*, 1998) was collected at the Daresbury synchrotron facility, allowing us to describe the quaternary structure of DB58. Three DB58 dimers are present in the asymmetric unit. Surprisingly, two of these dimers associate in the crystal to form a tetramer that resembles a DBL tetramer. However, the third DB58 dimer does not form a tetramer in the crystal and thus represents the true quaternary structure of the dimer in solution. DB58 forms a novel dimer that is similar to the top half of a DBL tetramer (monomer pairs AC and BD in Figure 1(b)), i.e. it corresponds to two facing monomers that form the adenine-binding site (Figure 3). In DB58, the  $\alpha$ -helix breaks the potential 2-fold symmetry of the dimer. Tetramer formation by DB58 in solution is probably hindered by the substitution of Pro14 in DBL by Ser14 in DB58. In the canonical dimers present in DBL, the Pro14 ring is involved in a hydrophobic inter-subunit contact with Tyr203. At present, three different tetramer types and three different dimer types have been described for the legume lectins (Figure 4). The DB58 dimer brings the number of known legume

lectin dimer types to four. The DBL tetramer can thus be considered as a dimer of DB58 dimers, that associate with each other *via* the formation of two canonical dimers in the tetramer. Similarly, the peanut agglutinin dimer can be considered as a dimer of two *G. simplicifolia* lectin IV dimers that associate with each other *via* the formation of one canonical dimer. The quaternary structure of a legume lectin becomes important in binding multivalent ligands, which can lead to the formation of ordered, cross-linked lattices (Brewer, 1996). Our results indicate that the quaternary structure of certain legume lectins is also important for the formation of a hydrophobic binding site at a subunit interface.

All the residues involved in adenine binding in DBL are fully conserved in DB58, with the exception of the conservative substitution of Leu244 by Ile. Formation of a DB58 dimer buries an area of 1800 Å<sup>2</sup>, which is comparable to the buried area in a canonical dimer (1650 Å<sup>2</sup>). The fact that both DBL and DB58 bind adenine and adenine-derived plant hormones, despite their different quaternary structure, carbohydrate binding specificity (Etzler, 1994b) and locus of expression *in vivo* (Roberts & Etzler, 1984), illustrates the high potential importance of this binding site.



**Figure 3.** The overall structure of the DB58 dimer. The dimer corresponds to dimer AC in the DBL tetramer (Figure 1(b)). Metal ions as in Figure 1(a).



**Figure 4.** The different legume lectin oligomer types, including the novel DB58 dimer type (EcorL, *Erythrina corallo-dendron* lectin; GS4, *Griffonia simplicifolia* lectin IV; PNA, Peanut agglutinin; DB58, *Dolichos biflorus* stem and leaf lectin; DBL, *Dolichos biflorus* seed lectin; Con A, concanavalin A). The canonical dimer type is represented by the lentil lectin dimer. Each tetramer type can be considered as a dimer of dimers. In the Figure, the dimers and the corresponding tetramers are vertically aligned. No tetramer is known that contains an EcorL dimer. One of the subunits is in the same orientation for all oligomers (upper left monomer for the three tetramers).  $\beta$ -Strands are shown as arrows.

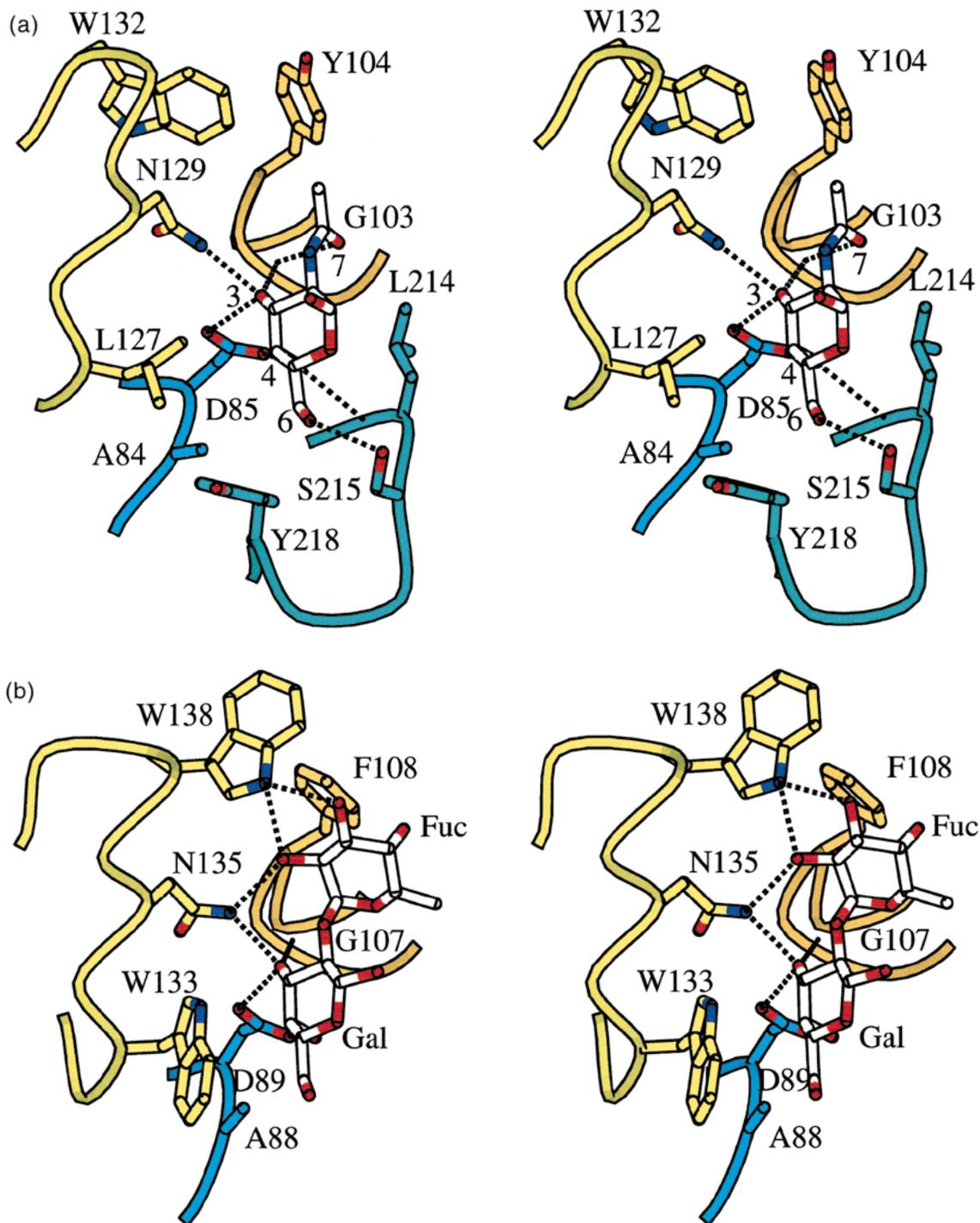
### Blood group A trisaccharide binding

For the blood group A trisaccharide complex, strong electron density is found only for the GalNAc moiety of the trisaccharide, which indicates that the Fuc( $\alpha$ 1-2)Gal moiety is not tightly bound by DBL. The visible GalNAc is bound in the conserved monosaccharide binding site, next to the equally conserved metal-binding site (Figure 5(a)). Three conserved sugar-binding residues (Asp85, Gly103 and Asn129) occupy positions similar to those in the other legume lectin structures. The Asp85 residue, which is involved in the *cis*-peptide bond with the preceding Ala residue, hydrogen bonds to the hydroxyl groups at positions 3 and 4 of the GalNAc residue. Asn129, which interacts with the bound  $\text{Ca}^{2+}$  via its  $\text{O}^{\delta 1}$  atom, hydrogen bonds to the hydroxyl group at position 4 via its  $\text{N}^{\delta 2}$  atom. Gly103 plays a crucial role in the GalNAc-DBL complex. The backbone amide group of Gly103 is involved in a bifurcated hydrogen bond with O-3 and the oxygen atom of the *N*-acetyl group. Further hydrogen bonds are formed between Ser215  $\text{O}^{\gamma}$  and O-6, and between Leu214 N and O-4. Leu214 is responsible for the  $\alpha$ -anomeric preference of DBL. Binding of  $\beta$ -GalNAc in the

monosaccharide-binding site of DBL would bring the  $\beta$ -anomeric oxygen atom close to the side-chain of Leu214.

Leu214 and Ser215 belong to the so-called specificity loop (Sharma & Surolia, 1997; Loris *et al.*, 1998). While the conserved Asn-Gly-Asp triad and a conserved aromatic residue (see below) confer affinity, the specificity loop is thought to confer specificity as well as affinity by excluding certain monosaccharides via sterical hindrance, while interacting favorably with others (Sharma & Surolia, 1997; Loris *et al.*, 1998). Hydrogen bonds between protein and sugar are listed in Table 3.

In addition to hydrogen bonds, hydrophobic interactions play an important role: the side-chains of Tyr104 and Trp132 form a hydrophobic pocket that is in the vicinity of the methyl group of GalNAc (atoms of both residues within 4.6 Å from the carbon atom are: Tyr104  $\text{C}^{\delta 2}$ : 4.3 Å; Tyr104  $\text{C}^{\epsilon 2}$ : 3.8 Å; Trp132  $\text{C}^{\eta 2}$ : 4.1 Å). Binding of GalNAc buries 23 Å<sup>2</sup> of the non-polar area of both residues. Tyr104 belongs to a large  $\Omega$ -loop, together with the conserved Gly103 residue, while Trp132 is part of the metal-binding loop. Tyr218, which belongs to the specificity loop, makes a favorable hydro-



**Figure 5.** (a) A view on the monosaccharide-binding site of the DBL-blood group A trisaccharide complex. The visible GalNAc residue and the side-chains of sugar-binding residues are shown as ball-and-stick models. The sugar-binding residues belong to four different stretches, each shown in a different color (metal loop, yellow; *cis*-peptide bond region, blue; specificity loop, green;  $\Omega$ -loop, orange). Nitrogen atoms are shown in blue; oxygen atoms in red. Hydrogen bonds between sugar and protein residues are shown as broken lines. The sugar hydroxyl groups that are involved in hydrogen bonds with the protein are labeled. (b) GS4 in complex with the Le<sup>b</sup> tetrasaccharide. For clarity, only the Fuc( $\alpha$ 1-2)Gal moiety is shown. The view on the sugar-binding site of GS4 corresponds to the view on the sugar-binding site of DBL in Figure 1(a). Sugar-binding residues are shown as ball-and-stick models. Atom color coding, hydrogen bonds and loop colors as in (a).

phobic interaction with C-6. The interaction buries 11 Å<sup>2</sup> of the non-polar area of Tyr218. Furthermore, Leu127 packs against the hydrophobic patch

of the B face of the GalNAc moiety, formed by C-3, C-4, C-5 and C-6. This buries 16 Å<sup>2</sup> of the non-polar area of Leu127. In most legume lectins, an

**Table 3.** Hydrogen bonds and average distances between hydrogen bond donors and acceptors between the non-reducing GalNAc in the monosaccharide binding site and DBL for the blood group A trisaccharide complex and the Forssman disaccharide complex

Hydroxyl group	Protein atom	Distance (Å)
O-3	Asp85 O <sup>δ1</sup>	2.9
	Gly103 N	2.9
	Asn129 N <sup>δ2</sup>	3.1
O-4	Asp85 O <sup>δ2</sup>	2.6
	Leu214 N	3.1
O-6	Ser215 O <sup>γ</sup>	2.7
O-7	Gly103 N	3.0

aromatic residue is present in the position of Leu127, and the significance of this substitution for the specificity of DBL is discussed below. All DBL residues with atoms within 4.6 Å of the bound sugar are listed in Table 4.

Imberty *et al.* (1994) presented a model of the DBL-blood group A trisaccharide complex that is in agreement with the results of NMR studies (Casset *et al.*, 1996). The structure of the conserved sugar-binding site and the position of the GalNAc residue in this site are in perfect agreement with the crystal structure presented here. In the model, the Fuc(α1-2)Gal disaccharide wraps around Leu127 and the Fuc moiety hydrogen bonds to the carbonyl group of Ser128. In the crystal structure, some weak density is indeed seen for the Fuc moiety in the neighborhood of Ser128, indicating that it probably transiently binds to the lectin in the proposed Fuc-binding subsite.

### GalNAc specificity: the role of Leu127

DBL is a strict GalNAc-specific lectin ( $K_a = 4 \times 10^3 \text{ M}^{-1}$  for Me-α-D-GalNAc; Etzler *et al.*, 1981), in the sense that precipitation of blood group A + H substance by DBL is inhibited by GalNAc but not by Gal (Etzler & Kabat, 1970). To determine the molecular basis underlying the specificity of DBL, we compared the sugar-binding site of DBL with the binding sites of four Gal-binding legume lectins (EcorL, SBA, PNA and WBAI). In EcorL, SBA, PNA and WBAI there is an aromatic residue (Phe128 in SBA, Phe131 in EcorL, Tyr125 in PNA and Phe126 in WBAI) that stacks

against the C-3, C-4, C-5 and C-6 patch of Gal. In DBL, this residue is replaced by the aliphatic residue Leu127. Strikingly, in all known crystal structures of lectins complexed with Gal, the apolar C-3, C-4, C-5, C-6 patch of the B face of the sugar packs against an aromatic residue (Weis & Drickamer, 1996), clearly demonstrating the importance of this interaction. Interactions of Man with aromatic residues show much more variability and seem less critical. This leads to the conclusion that the favorable interactions between the N-acetyl group and the Gly103/Tyr104/Trp134 triad compensate for the substitution of an aromatic residue by an aliphatic residue (Leu127). Since only GalNAc is able to compensate for the replacement of an aromatic residue with Leu, Gal binds with a much lower affinity.

To assess the role of the Leu127 residue, we constructed the mutant Leu127Phe and performed binding studies. The Leu127Phe mutant binds to Sepharose beads better than the native seed lectin (Figure 6): 50% inhibition of binding for the native lectin occurred at 120.2 mM Gal and 1.1 mM GalNAc. For the Leu127Phe mutant, 50% inhibition of binding occurred at only 28.4 mM Gal and 0.27 mM GalNAc. These results confirm that the Leu127 residue is one of the determining factors of the low affinity of DBL for Gal.

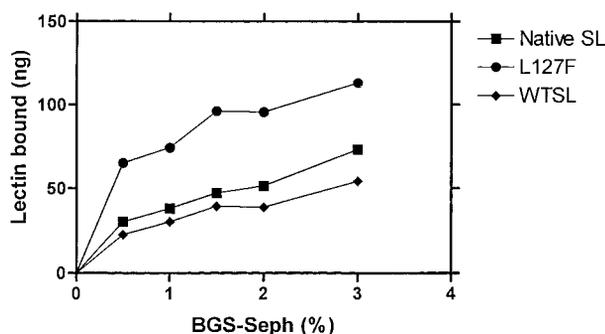
The results from two other studies on the effect of mutations on sugar binding by two Gal-binding legume lectins are in full accord with this view. In EcorL, two mutants in which the conserved aromatic residue is mutated to an aliphatic residue (Phe131Ala and Phe131Val) cannot induce hemagglutination (Adar & Sharon, 1996). In the *Robinia pseudoacacia* lectin, the Phe130Ala and the Phe130-Leu mutations impair hemagglutination as well (Nishiguchi *et al.*, 1997). This again demonstrates that an aliphatic residue (Val or Leu) cannot compensate for the loss of a sugar-stacking aromatic residue.

DBL is not the only legume lectin in which the conserved aromatic residue that stacks against the sugar ring is replaced by an aliphatic residue. The structure of (uncomplexed) PHA-L (Hamelryck *et al.*, 1996) from the common bean reveals that this lectin has a Leu residue (Leu126) at this pos-

**Table 4.** DBL residues with atoms within 4.6 Å of a carbohydrate atom

Non-reducing GalNAc (blood group A trisaccharide and Forssman disaccharide)	
Sugar ring (C-1, C-2, C-3, C-4, C-5, O-5)	Asp85 (6), Gly103 (2), Asn129 (2), Leu214 (10), Leu217 (2)
N-Acetyl group (N-2, C-7, O-7)	Asn101 (2), Gly102 (4), Gly103 (5), Asn129 (2), Leu214 (3)
N-Acetyl group (C-8)	Gly103 (2), Tyr104 (2), Asn129 (1), Trp132 (1)
O-3	Asp85 (3), Gly102 (2), Gly103 (2), Leu127 (1), Asn129 (3)
O-4	Asp85 (3), Gly102 (1), Leu127 (1), Gly213 (3), Leu214 (5)
C-6 and O-6	Gly213 (2), Leu214 (5), Ser215 (8), Tyr218 (11)
Reducing GalNAc (Forssman disaccharide)	
N-Acetyl group (C-8)	Leu127 (2), Tyr218 (3), *Asp109 (6), *Asn114 (1)

An asterisk in front of the residue name indicates it belongs to a symmetry mate. The number of atoms within 4.6 Å is shown between parentheses.

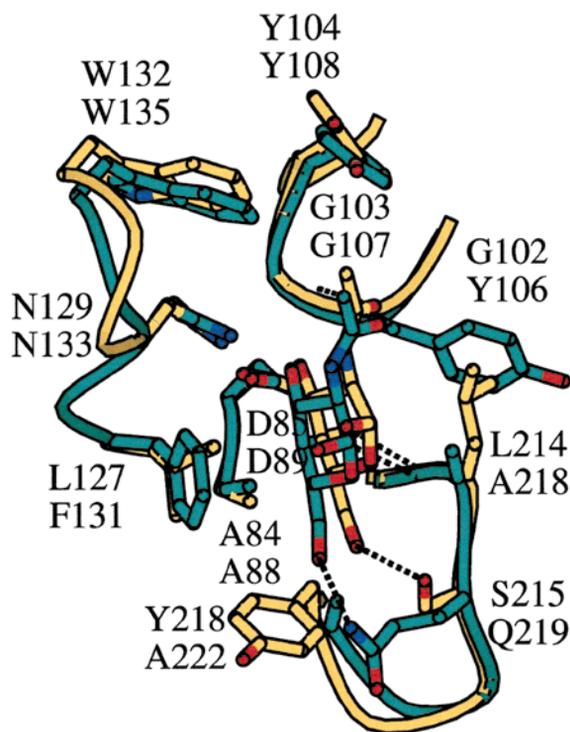
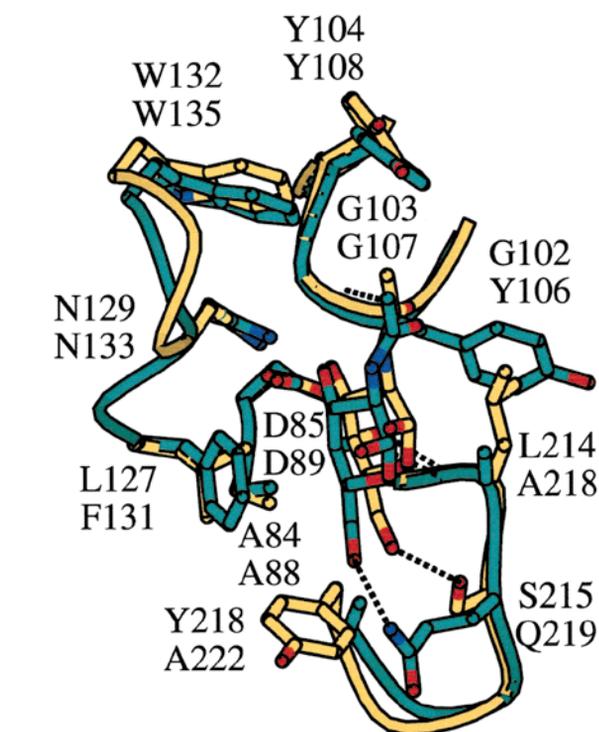


**Figure 6.** Binding of iodinated lectins (wild-type DBL, WTSL; native recombinant DBL, Native SL; and the Leu127Phe mutant, L127F) to hog blood group A + H-Sepharose (BGS-Seph).

ition. This may explain why precipitation by PHA-L is not inhibited by monosaccharides (Hammarström *et al.*, 1982).

### GalNAc specificity: the role of the specificity loop

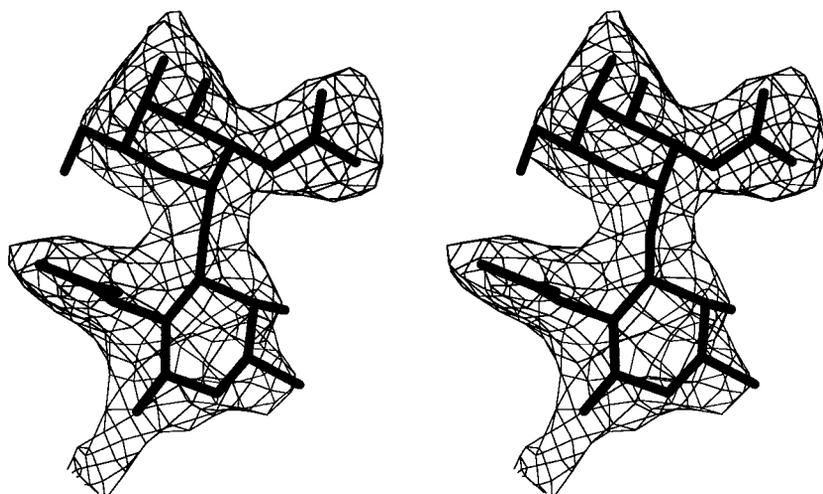
Recently, the crystal structure of EcorL in complex with GalNAc was published (Elgavish & Boaz, 1998). Unlike DBL, EcorL binds Gal and GalNAc with equal affinity ( $K_a = 2.4 \text{ M}^{-1}$  for Gal;  $K_a = 2.3 \text{ M}^{-1}$  for GalNAc; Surolia *et al.*, 1996). To study the molecular basis of the difference in specificity between EcorL and DBL, the monosaccharide-binding sites of DBL and EcorL were superimposed using the  $C^\alpha$  atoms of the four conserved sugar-binding residues (Asp85, Leu127, Gly103 and Asn129 for DBL; Asp89, Phe131, Gly107 and Asn133 for EcorL; Figure 7). The four residues occupy virtually identical positions in both lectins (r.m.s.d. 0.1 Å). Surprisingly, the three residues that interact with the *N*-acetyl group of GalNAc in DBL (Gly103, Tyr104 and Trp132) are present in EcorL in the same position (Gly107, Tyr108 and Trp135). The absence of a higher affinity for GalNAc in EcorL is thus not due to the simple absence of the Gly-Tyr-Trp triad.



**Figure 7.** Stereo Figure of the superimposed monosaccharide-binding sites of DBL (yellow) and EcorL (green). All residues involved in sugar-binding are shown as ball-and-stick models. The sugar-binding residues are labeled for DBL (upper) and EcorL (lower). Atom color coding and hydrogen bonds as in Figure 5.

ide-binding sites of DBL and EcorL were superimposed using the  $C^\alpha$  atoms of the four conserved sugar-binding residues (Asp85, Leu127, Gly103 and Asn129 for DBL; Asp89, Phe131, Gly107 and Asn133 for EcorL; Figure 7). The four residues occupy virtually identical positions in both lectins (r.m.s.d. 0.1 Å). Surprisingly, the three residues that interact with the *N*-acetyl group of GalNAc in DBL (Gly103, Tyr104 and Trp132) are present in EcorL in the same position (Gly107, Tyr108 and Trp135). The absence of a higher affinity for GalNAc in EcorL is thus not due to the simple absence of the Gly-Tyr-Trp triad.

The position of the GalNAc residue in the monosaccharide-binding sites of DBL and EcorL is similar, but not identical. In EcorL, the sugar ring is about 0.8 Å closer to the metal-binding loop than in DBL. Again, the monosaccharide-specificity loop plays an important role in positioning the monosaccharide in the monosaccharide-binding site. In EcorL, Ser215 and Tyr218 are replaced by Gln219 and Ala222. These two substitutions position the sugar ring closer to the metal-binding site in EcorL than in DBL, and push the sugar ring against the aromatic ring of Phe131. In DBL, the side-chains of Ser215 and Tyr218 position the sugar closer to the specificity loop. A second difference is the presence of Tyr106 in EcorL, which corresponds to Gly102 in DBL. The  $C^\beta$  atom of this residue might push the *N*-acetyl group away from the conserved Gly residue in EcorL. As a consequence of the above discussed structural features, the distance between



**Figure 8.** Stereo Figure of the  $F_o - F_c$  density around the Forsman disaccharide, calculated after simulated annealing refinement of the DBL structure in the absence of the sugar. The density was visualized at a level of  $1.0 \sigma$ .

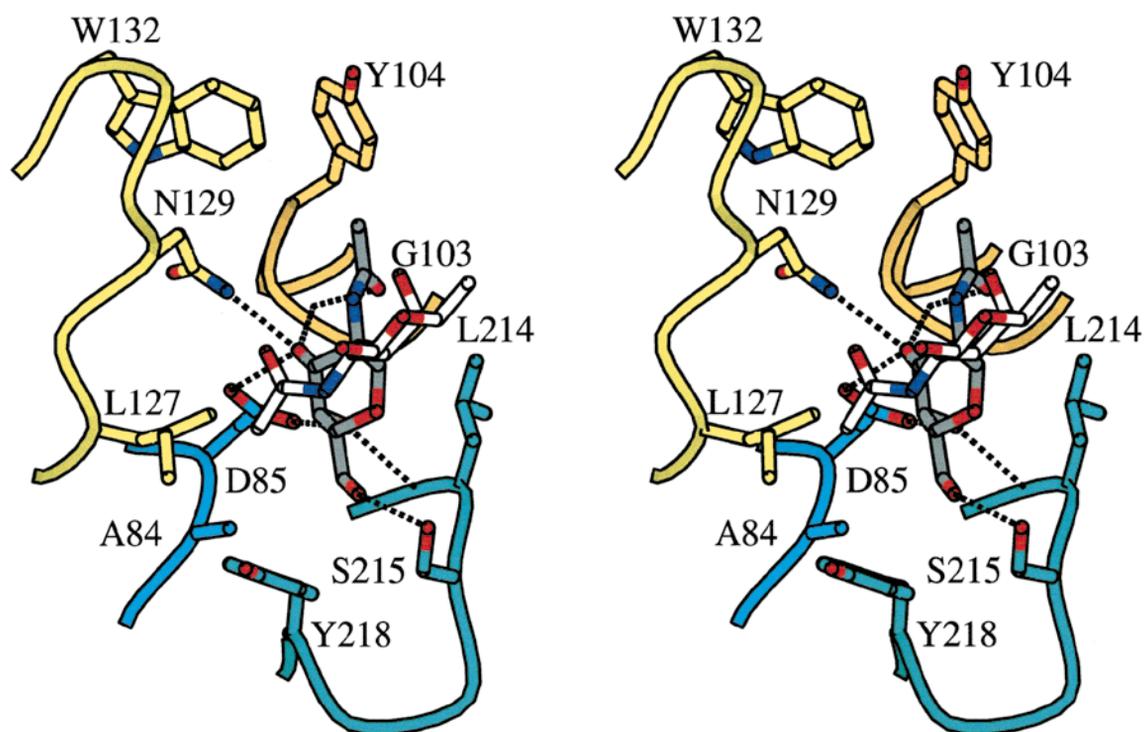
the amide group of Gly107 and the oxygen atom of the *N*-acetyl group of GalNAc in EcorL ( $4.0 \text{ \AA}$ ) is  $1.1 \text{ \AA}$  longer than the corresponding distance in DBL ( $2.9 \text{ \AA}$ ). Thus, EcorL does not bind GalNAc with a higher affinity than Gal because the sugar ring is not optimally positioned in the monosaccharide binding site.

#### Forsman disaccharide binding

The quality of the electron density of the Forsman disaccharide was high for both residues

(Figure 8), in contrast to the absence of electron density for two of the three residues of the blood group A trisaccharide. No density is seen for the O-6 hydroxyl group of the reducing GalNAc residue: this hydroxyl group is not in contact with the protein and seems to be completely flexible. The group attached to O-1 of the reducing GalNAc extends into the solvent and is not visible in the electron density (see Materials and Methods).

The non-reducing GalNAc residue of the GalNAc( $\alpha$ 1-3)GalNAc disaccharide is bound in the monosaccharide-binding site in exactly the same



**Figure 9.** Stereo Figure of the Forsman disaccharide in complex with DBL. The non-reducing GalNAc residue bound in the conserved monosaccharide binding site is shown in gray; the reducing GalNAc residue is shown in white. Sugar-binding residues are shown as ball-and-stick models. The view shown in this Figure is the same as in Figure 5. Atom color coding, hydrogen bonds and loop colors as in Figure 5.

way as the GalNAc residue of the blood group A trisaccharide (Figure 9, Tables 3 and 4). The reducing GalNAc residue extends from the monosaccharide-binding site and only interacts with the protein *via* its *N*-acetyl group (Table 4). The methyl group of the *N*-acetyl moiety of the reducing GalNAc residue points to a hydrophobic cavity formed by the side-chains of residues Leu127 and Tyr218. Both residues are involved in binding the non-reducing GalNAc in the monosaccharide-binding site. It should be noted that the methyl group of the reducing GalNAc is located directly above the hydrophobic patch (distances are 4.0 Å from C-5; 4.5 Å from C-6) of the non-reducing GalNAc. Binding of the Forssman disaccharide therefore creates a small, solvent-shielded hydrophobic cavity consisting of the side-chains of Leu127 and Tyr218, the methyl group of the reducing GalNAc and the hydrophobic patch of the non-reducing GalNAc. The presence of Leu at position 127 in DBL, instead of an aromatic residue as in most legume lectins, plays an important role in Forssman disaccharide binding: the larger Phe, Tyr or Trp side-chains would clash with the *N*-acetyl group.

Binding of GalNAc in the monosaccharide-binding site buries 27 Å<sup>2</sup> of the non-polar area of Leu127 and Tyr218, and the reducing GalNAc residue buries another 18 Å<sup>2</sup>. No additional hydrogen bond is made with the protein by the reducing GalNAc residue, and the shielding of the hydrophobic cavity from the solvent appears to be the crucial specificity-determining interaction between the Forssman disaccharide and DBL. The crucial role of the *N*-acetyl group in enhancing the affinity is further confirmed by the fact that Me- $\alpha$ -D-GalNAc and GalNAc( $\alpha$ 1-3)Gal are equally active as inhibitors of precipitation of blood group A substance (Etzler & Kabat, 1970).

The conformation of the Forssman disaccharide in the crystal structure ( $\Psi = 88^\circ$ ;  $\theta = 70^\circ$ ) broadly resembles the conformation in the modeled complex ( $\Psi = 77^\circ$ ;  $\theta = 80^\circ$ ; Imberty *et al.*, 1994). This conformation corresponds to the lowest-energy conformation in molecular mechanics simulations and is in accordance with NMR measurements of the disaccharide in solution (Grönberg *et al.*, 1994; Casset *et al.*, 1996) and bound to DBL (Casset *et al.*, 1997; Rinnbauer *et al.*, 1998). These observations support the view that lectins generally bind sugars in a conformation that is close to a minimum-energy conformation in solution.

### A multi-purpose secondary binding site in the legume lectin family

The *N*-acetyl group of the GalNAc residue bound in the monosaccharide-binding site of DBL interacts with a hydrophobic cavity formed by Tyr104 and Trp132. Comparison of this *N*-acetyl-binding region in DBL with the corresponding region in GS4 reveals that they possess a similar hydrophobic pocket, formed by Phe108 and Trp138 in GS4 (Figure 5(b)). GS4 belongs to the

complex specificity group and does not bind monosaccharides. However, the complex of GS4 with the Le<sup>b</sup> oligosaccharide (Fuc( $\alpha$ 1-2)Gal( $\beta$ 1-3)[Fuc( $\alpha$ 1-4)]GlcNAc) shows that the Gal residue is bound in a truncated monosaccharide-binding site, in a way similar to Gal in Gal-specific legume lectins. In the GS4-Le<sup>b</sup> complex, the Fuc residue ( $\alpha$ 1-2)-linked to the Gal residue in the monosaccharide-binding site is bound in the region that corresponds to the *N*-acetyl-binding pocket in DBL. The Fuc residue is involved in a stacking interaction with Phe108 and its O-2 and O-3 hydrogen bond to Trp138 N<sup>e1</sup>. The latter residue corresponds to Trp132 in DBL, but adopts a different conformation in GS4.

In the modeled structure of lentil lectin complexed with 3-*O*-nitrophenylmannose, the same subsite harbors the nitrophenyl group (Loris *et al.*, 1994) that interacts with the side-chains of Tyr100 $\beta$  and Trp128 $\beta$ , corresponding to Tyr104 and Trp132 in DBL. In the modeled structure of EcorL with Fuc( $\alpha$ 1-2)Gal( $\beta$ 1-4)GlcNAc $\beta$  (Moreno *et al.*, 1997), again the same subsite interacts with the fucose moiety, involving the side-chains of Phe108 and Trp135. In both structures, the backbone amide groups of the Gly residues equivalent to Gly103 in DBL are involved in a bifurcated hydrogen bond (with O-3 of the mannose moiety in lentil lectin, and O-2 of fucose in EcorL). Mutagenesis experiments and molecular modeling have indicated that the same subsite is involved in the high-affinity binding of methyl- $\alpha$ -*N*-dansylgalactosaminide by EcorL (Arango *et al.*, 1993; Adar *et al.*, 1998). The site formed by the Gly-Tyr-Trp triad thus seems to be a conserved multi-purpose binding site, involved in binding both hydrophobic groups and monosaccharide residues.

## Conclusions

The structure of DBL in complex with adenine shows that certain legume lectins possess a hydrophobic binding site that depends on their quaternary structure. The biological relevance of this binding site is unclear, but the fact that it is conserved in (at least) three tetrameric legume lectins (SBA, PHA-L and DBL) from different plant species indicates that it might play an important biological role. The DB58 dimer, which is the first structure of a legume lectin from the vegetative tissues, adopts a novel quaternary structure in which this site is strictly conserved. The DBL sugar complexes illustrate the fact that specific binding of an *N*-acetylated sugar is achieved through a give-and-take mechanism: a lack of aromatic stacking against the sugar ring is compensated by a favorable interaction of the *N*-acetyl group with the lectin. The latter interaction depends on the presence of a specific subsite, and on the correct positioning of the sugar ring by the specificity loop. High-affinity binding of the Forssman disaccharide is mediated solely *via* the formation of a small, sol-

vent-shielded hydrophobic cavity and does not involve any additional hydrogen bond.

## Materials and Methods

### Preparation of the L127F seed lectin mutant

Site-directed mutagenesis was performed using a two-step PCR method (Landt *et al.*, 1990), employing the oligonucleotide 5'-GCAGTCGAGTTCGACACGTTCTC-CAACAGCGGCTGG-3' as the mutagenic primer and using seed lectin cDNA (Schnell & Etzler, 1987) as template. Following sequence confirmation, a *KpnI/BamHI* fragment of the PCR product, containing the Leu127Phe mutation, was cloned into a *KpnI/BamHI* digested pTrc99A-based expression vector containing the coding sequence for methionine plus the entire mature seed lectin sequence (Chao *et al.*, 1994). The mutant protein was expressed in the protease-deficient SG21173 *E. coli* strain by induction with 0.4 mM IPTG for three hours at 37°C. The mutant protein was present in the soluble fraction obtained after sonication and centrifugation of the *E. coli*. This protein was purified by affinity chromatography on hog blood group A + H-Sepharose, followed by specific elution with 0.01 M GalNAc as described (Chao *et al.*, 1994).

### Binding assays

A solid phase assay (Etzler, 1994b), inhibition of binding of iodinated lectin to bind to hog blood group A + H-Sepharose, was employed for the determination of carbohydrate binding. Various amounts of inhibitors (Gal and GalNAc) were mixed with 72.5 ng of Leu127Phe mutant recombinant seed lectin or native recombinant seed lectin in a final volume of 200 µl of 2% (v/v) hog blood group A + H-Sepharose. After incubation at room temperature overnight, binding was measured as described (Etzler, 1994b).

### Crystallization

All DBL complexes were crystallized using recombinant DBL. DBL subunit I was cloned in *E. coli* and purified as described (Chao *et al.*, 1994). Crystallization of the complex of DBL with the blood group A trisaccharide (space group C2), the complex of DBL with adenine (space group  $P2_12_12_1$ ) and uncomplexed DB58 (space group  $P2_12_12_1$ ) is described elsewhere (Dao-Thi *et al.*, 1998). The crystal of DBL in complex with the Forssman disaccharide was made by soaking an uncomplexed I4<sub>22</sub> DBL crystal (Dao-Thi *et al.*, 1998) with a 5 mM sugar, 3 mM adenine solution. The Forssman disaccharide containing an additional group attached to the anomeric oxygen atom of the reducing GalNAc (GalNAc(α1-3)GalNAcβ1-OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>) was purchased from Syntesome, München.

### Data collection

Data collection for the DBL sugar complexes was done with an MAR image plate. CuKα X-rays were generated with a RIGAKU rotating anode X-ray generator operating at 40 kV, 100 mA. The crystals of the Forssman disaccharide complex diffract anisotropically to 2.4 Å in the  $c^*$  direction and to 2.8 Å perpendicular to  $c^*$ . Refinement was done using data to 2.6 Å. The crystals of the blood group A trisaccharide complex diffract to 2.8 Å.

Data collection for the complex of DBL with adenine (2.65 Å) and for uncomplexed DB58 (3.3 Å) was done with an MAR image plate at the Daresbury synchrotron facility, UK, operating at a wavelength of 1.488 Å.

In all cases, the datasets were indexed, scaled and merged using the programs DENZO and SCALEPACK (Otwinowski & Minor, 1997). Intensities were transformed to structure factors with the CCP4 (1994) program TRUNCATE. Details of data set quality are shown in Table 5.

### Molecular replacement, model building and refinement

The phase problem was solved for all structures with the molecular replacement technique, using the CCP4 (1994) program AMoRe (Navaza, 1994). The PHA-L coordinates (Hamelryck *et al.*, 1996) were used for molecular replacement of the DBL-adenine complex. The final model of the DBL-adenine complex was used for molecular replacement of the other three structures (the DBL-Forssman disaccharide complex, the DBL-blood group A trisaccharide complex and uncomplexed DB58).

Molecular replacement of the DB58 dataset was first attempted with a canonical dimer (subunits A and B of DBL, without the helices, see Figure 1). Two solutions were readily found (CC = 43.1%), and inspection of the crystal packing revealed that a DBL-like tetramer is formed in the crystal. However, inspection of an electron density map calculated from this model revealed density for a third non-canonical dimer. A subsequent molecular replacement attempt with an alternative dimer (subunits A and C of DBL, see Figure 1) yielded three clear solutions and a considerable better correlation coefficient (CC = 61.7%).

Positional refinement was done using the simulated annealing, torsion angle refinement and conjugate gradient least-squares refinement protocols of the program X-PLOR (Brünger, 1992). Throughout the refinement, a real space bulk solvent model was used. In addition, an overall anisotropic  $B$  factor was applied to  $F_{\text{obs}}$  for both carbohydrate complexes. This improved the quality of the electron maps dramatically in the case of the Forssman disaccharide complex. Non-crystallographic symmetry restraints ( $\sigma_b = 0.5 \text{ \AA}^2$ , weight = 600 kcal/mol $\text{\AA}^2$ ) were applied between the subunits in the asymmetric unit for the blood group A trisaccharide complex, the DBL-adenine complex and uncomplexed DB58. The Forssman disaccharide complex contains only one subunit in the asymmetric unit. In order not to overfit the data, only two  $B$  factors (one for side-chain atoms and one for main-chain atoms) per residue were refined for the latter structure. For both sugar complexes, the complete tetramer is generated by the crystal symmetry. The quality parameters of the structures are shown in Table 5.

Visualization and model building was done with the program O, version 6.2 (Jones *et al.*, 1991), starting from a modeled DBL subunit (Imberty *et al.*, 1994) for the DBL-adenine complex, and from the resulting DBL coordinates for the other three structures. The amino acid sequences were derived from the cDNA and genomic DNA sequences of DBL and DB58 (Schnell & Etzler, 1987, 1988; Harada *et al.*, 1990). The DB58 structure was built using an electron density map that was averaged over all subunits in the asymmetric unit with the program DM (Cowtan & Main, 1993). Weak density is present for regions Asn12-Ser13, Ser102-Gly103 and Asp131 for the DB58 structure. For the blood group A trisaccha-

**Table 5.** Statistics of crystallographic data and quality of the structures

	DBL-adenine complex	Uncomplexed DB58	DBL-blood group A trisaccharide complex	DBL-Forsman disaccharide complex
Space group	$P2_12_12_1$	$P2_12_12_1$	C2	$I4_122$
Unit cell parameters				
$a, b, c$ (Å)	81.38, 116.14, 224.62	101.39, 130.95, 138.23	96.73, 108.00, 80.96	79.05, 79.05, 260.11
$\alpha, \beta, \gamma$ (deg.)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 124.67, 90.00	90.00, 90.00, 90.00
Contents of asymmetric unit	Two tetramers, eight adenine residues	Three dimers	Two monomers, two GalNAc molecules	One monomer, one Forsman disaccharide one adenine residue
Resolution (Å)	20.0-2.65	20.0-3.3	15.0-2.8	20.0-2.6
$I/\sigma I$ ( $I/\sigma I$ last shell)	12.5 (2.5, 2.67-2.65 Å)	9.1 (2.7, 3.33-3.3)	10.7 (3.1; 2.82-2.80 Å)	15.2 (2.3; 2.62-2.60 Å)
$R_{\text{merge}}$ (%)	11.6	13.5	11.1	14.6
Completeness (%)	92.4	80.6	85.2	99.9
Number of reflections measured	162,686	47,794	39,403	91,892
Number of unique reflections	57,851	22,745	14,331	13,175
$R_{\text{cryst}}/R_{\text{free}}$ (%)	21.9, 24.6	22.7, 26.6	19.1/21.9	19.4/23.4
r.m.s.d. ideal bond length (Å)	0.008	0.011	0.008	0.008
r.m.s.d. impropers (deg.)	0.67	0.75	0.73	0.63
r.m.s.d. dihedrals (deg.)	28.45	28.62	28.9	28.7
r.m.s.d. ideal bond angles (deg.)	1.46	1.58	1.53	1.46
Ramachandran plot quality (%)				
Most favored	89.5	83.1	83.1	84.8
Additionally allowed	10.1	16.9	16.2	15.2
Generously allowed	0.4	0.0	0.7	0.0
Forbidden	0.0	0.0	0.0	0.0

$$R_{\text{merge}} = \frac{\sum_{hkl} \sum_I (|I_{hkl}| - I_{hkl})}{\sum_{hkl} \sum_I I_{hkl}}$$

$$R_{\text{cryst}} = \frac{\sum_{hkl} |F_{hkl,o} - F_{hkl,c}|}{\sum F_{hkl,o}}$$

where  $F_{hkl,o}$  and  $F_{hkl,c}$  are the amplitudes of the observed and calculated structure factors, respectively.

$$R_{\text{free}} = R_{\text{cryst}}$$

value calculated with 10% of the reflections, not included in the refinement.

ride structure, weak density was present for Ser79 and Lys80 in all subunits, except subunits B and H. Consequently, these regions were not built in these structures. Figures were made with the programs BOBSCRIPT (Kraulis, 1991; Esnouf, 1997) and RASTER3D (Merrit & Murphy, 1994). The quality of the structure was analyzed with PROCHECK (Laskowski *et al.*, 1993). Accessible surface areas of individual amino acid residues were calculated with NACCESS (Williams *et al.*, 1994). Accessible surface areas of monomers, dimers and tetramers were calculated with the program MSROLL (Connolly, 1993), using a probe size of 1.4 Å.

Torsion angles for the Forssman disaccharide (GalNAc( $\alpha$ 1-3)GalNAc) are defined as  $\Psi = C1-O1-C3'-C4'$  and  $\theta = O5-C1-O1-C3'$ . PDB entries 1LED and 1AX0 were used for comparison of GS4 and EcorL with DBL, respectively.

### Protein Data Bank

Coordinates and structure factors have been submitted to the Protein Data Bank: 1LU1 and r1LU1sf for the DBL-Forssman disaccharide complex; 1LU2 and r1LU2sf for the DBL-blood group A trisaccharide complex; 1BJQ and r1BJQsf for the DBL-adenine complex; 1LUL/r1LULsf for DB58.

### Acknowledgments

The authors thank Dr Edward Re and Yves Geunes for excellent technical assistance, and Dr J. Steyaert for interesting discussions. This work was supported by the Vlaams interuniversitair instituut voor Biotechnologie (V.I.B.) and by NIGMS-NIH grant GM21882 (M.E.E.). R.L. and J.B. are postdoctoral fellows of the Fonds voor Wetenschappelijk Onderzoek (F.W.O.). T.W.H. received financial support from the Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie (I.W.T).

### References

- Adar, R. & Sharon, N. (1996). Mutational studies of the amino acid residues in the combining site of *Erythrina corallodendron* lectin. *Eur. J. Biochem.* **239**, 668-674.
- Adar, R., Moreno, E., Streicher, H., Karlsson, K. A., Ångström, J. & Sharon, N. (1998). Structural features of the combining site region of *Erythrina corallodendron* lectin: role of tryptophan 135. *Protein Sci.* **7**, 52-63.
- Arango, R., Rodriguez-Arango, E., Adar, R., Belenky, D., Loontjens, F. G., Rozenblatt, S. & Sharon, N. (1993). Modification by site-directed mutagenesis of the specificity of *Erythrina corallodendron* lectin for galactose derivatives with bulky substituents at C-2. *FEBS Letters*, **330**, 133-136.
- Baker, D. A., Sugii, S., Kabat, E. A., Ratcliffe, R. M., Hermentin, P. & Lemieux, R. U. (1983). Immunological studies on the combining sites of Forssman haptens reactive hemagglutinins from *Dolichos biflorus*, *Helix pomatia*, and *Wistaria floribunda*. *Biochemistry*, **22**, 2741-2750.
- Brewer, C. F. (1996). Multivalent lectin-carbohydrate cross-linking interactions. *Chemtracts Biochem. Mol. Biol.* **6**, 165-179.
- Brewin, N. J. & Kardailsky, I. V. (1997). Legume lectins and nodulation by *Rhizobium*. *Trends Plant Sci.* **2**, 93-98.
- Brünger, A. (1992). *X-PLOR Version 3.1: A System for Crystallography and NMR*, Yale University, New Haven, CT.
- Carter, W. G. & Etzler, M. E. (1975). Isolation, characterization and subunit structures of multiple forms of *Dolichos biflorus* lectin. *J. Biol. Chem.* **250**, 2756-2762.
- Casset, F., Peters, T., Etzler, M., Korchagina, E., Nifant'ev, N., Pérez, S. & Imberty, A. (1996). Conformational analysis of blood group A trisaccharide in solution and in the binding site of *Dolichos biflorus* lectin using transient and transferred nuclear Overhauser enhancement (NOE) and rotating-frame NOE experiments. *Eur. J. Biochem.* **239**, 710-719.
- Casset, F., Imberty, A., Pérez, S., Etzler, M. E., Paulsen, H. & Peters, T. (1997). Transferred NOEs and ROEs reflect the size of the bound segment of the Forssman pentasaccharide in the binding site of *Dolichos biflorus* lectin. *Eur. J. Biochem.* **244**, 242-250.
- Chao, Q., Casalongue, C., Quinn, J. M. & Etzler, M. E. (1994). Expression and partial purification of *Dolichos biflorus* seed lectin in *Escherichia coli*. *Arch. Biochem. Biophys.* **313**, 346-350.
- Chrispeels, M. J. & Raikhel, N. V. (1991). Lectins, lectin genes and their role in plant defense. *Plant Cell*, **3**, 1-9.
- Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760-763.
- Connolly, M. L. (1993). The molecular surface package. *J. Mol. Graph.* **11**, 139-141.
- Cowtan, K. D. & Main, P. (1993). Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallog. sect. D*, **49**, 148-157.
- Dao-Thi, M., Hamelryck, T. W., Bouckaert, J., Körber, F., Burkow, V., Poortmans, F., Etzler, M., Strecker, G., Wyns, L. & Loris, R. (1998). Crystallization of two related lectins from the legume plant *Dolichos biflorus*. *Acta Crystallog. sect. D*, **54**, 1446-1449.
- Dessen, A., Gupta, D., Sabesan, S., Brewer, F. & Sacchettini, J. C. (1995). X-ray crystal structure of the soybean agglutinin cross-linked with a biantennary analog of the blood group I carbohydrate antigen. *Biochemistry*, **34**, 4933-4942.
- Elgavish, S. & Boaz, S. (1998). Structures of the *Erythrina corallodendron* lectin and of its complexes with mono- and disaccharides. *J. Mol. Biol.* **277**, 917-932.
- Esnouf, R. M. (1997). An extensively modified version of Molscrip that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 132-134.
- Etzler, M. E. (1992). Plant lectins: molecular biology, syntheses, and function. In *Glycoconjugates: Composition, Structure and Function* (Allan, H. J. & Kisailus, E. C., eds), pp. 521-539, Marcel Dekker, Inc., New York.
- Etzler, M. E. (1994a). Isolation and characterization of subunits of DB58, a lectin from the stems and leaves of *Dolichos biflorus*. *Biochemistry*, **33**, 9778-9783.
- Etzler, M. E. (1994b). A comparison of the carbohydrate binding properties of two *Dolichos biflorus* lectins. *Glycoconjugate J.* **11**, 395-399.
- Etzler, M. E. (1996). The *Dolichos biflorus* lectin family: a model system for studying legume lectin structure and function. In *Lectins, Biology-Biochemistry-Clinical*

- Biochemistry (Van Driessche, E., Rougé, P., Beeckmans, S. & Bøg-Hansen, T. C., eds), pp. 3-9, Textop, Hellerup, Denmark.
- Etzler, M. E. & Kabat, E. A. (1970). Purification and characterization of a lectin (plant hemagglutinin) with blood group A specificity from *Dolichos biflorus*. *Biochemistry*, **9**, 869-877.
- Etzler, M. E., Gupta, S. & Borrebaeck, C. (1981). Carbohydrate binding properties of the *Dolichos biflorus* lectin and its subunits. *J. Biol. Chem.* **256**, 2367-2370.
- Fiedler, K. & Simons, K. (1994). A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. *Cell*, **77**, 625-626.
- Fiedler, K. & Simons, K. (1995). Characterization of VIP36, an animal lectin homologous to leguminous lectins. *J. Cell Sci.* **109**, 271-276.
- Gabius, H. J. & Gabius, S. (1997). Editors of *Glycosciences, Status and Perspectives*, Chapman & Hal, Weinheim.
- Gegg, C. V. & Etzler, M. E. (1994). Photoaffinity labeling of the adenine binding sites of two *Dolichos biflorus* lectins. *J. Biol. Chem.* **269**, 5687-5692.
- Gegg, C. V., Roberts, D. D., Segel, I. H. & Etzler, M. E. (1992). Characterization of the adenine binding sites of two *Dolichos biflorus* lectins. *Biochemistry*, **31**, 6938-6942.
- Gohier, A., Espinosa, J. F., Jimenez-Barbero, J., Carrupt, P. A., Pérez, S. & Imberty, A. (1996). Knowledge-based modeling of a legume lectin and docking of the carbohydrate ligand: the *Ulex europaeus* lectin I and its interactions with fucose. *J. Mol. Graph.* **14**, 322-327.
- Grönberg, G., Nilsson, U., Bock, K. & Magnusson, G. (1994). Nuclear magnetic resonance and conformational investigations of the pentasaccharide of the Forssman antigen and overlapping di-, tri-, and tetra-saccharide sequences. *Carbohydrate Res.* **257**, 35-54.
- Hakomori, S. (1984). Tumor-associated carbohydrate antigens. *Annu. Rev. Immunol.* **2**, 103-126.
- Hamelryck, T. W., Dao-Thi, M., Poortmans, F., Chrispeels, M. J., Wyns, L. & Loris, R. (1996). The crystallographic structure of phytohemagglutinin-L. *J. Biol. Chem.* **271**, 20479-20485.
- Hammarström, S., Murphy, L. A., Goldstein, I. J. & Etzler, M. E. (1977). Carbohydrate binding specificity of four *N*-acetyl-D-galactosamine "specific" lectins: *Helix pomatia* hemagglutinin, soybean agglutinin, lima bean lectin and *Dolichos biflorus* lectin. *Biochemistry*, **16**, 2750-2755.
- Hammarström, S., Hammarström, M. L., Sundblad, G., Arnarp, J. & Lönngren, J. (1982). Mitogenic leucoagglutinin from *Phaseolus vulgaris* binds to a pentasaccharide unit in *N*-acetylglucosamine glycoprotein glycans. *Proc. Natl Acad. Sci. USA*, **79**, 1611-1615.
- Harada, J. J., Spadaro-Tank, J., Maxwell, J. C., Schnell, D. J. & Etzler, M. E. (1990). Two lectin genes differentially expressed in *Dolichos biflorus* differ primarily by a 116-base pair sequence in their 5' flanking regions. *J. Biol. Chem.* **265**, 4997-5002.
- Hervé, C., Dabos, P., Galaud, J. P., Rougé, P. & Lescure, B. (1996). Characterization of an *Arabidopsis thaliana* gene that defines a new class of putative plant receptor kinases with an extracellular lectin-like domain. *J. Mol. Biol.* **258**, 778-788.
- Imberty, A., Casset, F., Gegg, C. V., Etzler, M. E. & Pérez, S. (1994). Molecular modelling of the *Dolichos biflorus* seed lectin and its specific interactions with carbohydrates:  $\alpha$ -D-*N*-acetyl-galactosamine, Forssman disaccharide and blood group A trisaccharide. *Glycoconjugate J.* **11**, 400-413.
- Itin, C., Roche, A. C., Monsigny, M. & Hauri, H. P. (1996). ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. *Mol. Biol. Cell*, **7**, 483-493.
- Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Berlin-Heidelberg.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110-119.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
- Landersjö, C., Weintraub, A., Ansaruzzaman, M., Albert, M. J. & Widmalm, G. (1998). Structural analysis of the O-antigenic polysaccharide from *Vibrio mimicus* N-1990. *Eur. J. Biochem.* **251**, 986-990.
- Landt, O., Grunert, H. P. & Hahn, U. (1990). A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene*, **96**, 125-128.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
- Lis, H. & Sharon, N. (1998). Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.* **98**, 637-674.
- Loris, R., Casset, F., Bouckaert, J., Pletinckx, J., Dao-Thi, M., Poortmans, F., Imberty, A., Pérez, S. & Wyns, L. (1994). The monosaccharide binding site of lentil lectin: an X-ray and molecular modelling study. *Glycoconjugate J.* **11**, 507-517.
- Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). Legume lectin structure. *Biochim. Biophys. Acta*, **1383**, 9-36.
- Maliarik, M. J. & Goldstein, I. J. (1988). Photoaffinity labeling of the adenine binding site of the lectins from Lima bean, *Phaseolus lunatus*, and the kidney bean, *Phaseolus vulgaris*. *J. Biol. Chem.* **263**, 11274-11279.
- Maliarik, M. J., Roberts, D. D. & Goldstein, I. J. (1987). Properties of the lectin from the hog peanut (*Amphicarpaea bracteata*). *Arch. Biochem. Biophys.* **255**, 194-200.
- Maliarik, M., Plessas, N. R., Goldstein, I. J., Musci, G. & Berliner, L. J. (1989). ESR and fluorescence studies on the adenine binding site of lectins using a spin-labeled analogue. *Biochemistry*, **28**, 912-917.
- Mandal, D. K., Nieves, E., Bhattacharyya, L., Orr, G. A., Roboz, J., Yu, Q. & Brewer, C. F. (1994). Purification and characterization of three isolectins of soybean agglutinin. *Eur. J. Biochem.* **221**, 547-553.
- Mathieu, M., Modis, Y., Zeelen, J. P., Engel, C. K., Abagyan, R. A., Ahlberg, A., Rasmussen, B., Lamzin, V. S., Kunau, W. H. & Wieringa, R. K. (1997). The 1.8 Å crystal structure of the dimeric peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: implications for substrate binding and reaction mechanism. *J. Mol. Biol.* **273**, 714-728.
- Merrit, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869-873.
- Moothoo, D. N. & Naismith, J. H. (1998). Concanavalin A distorts the  $\beta$ -GlcNAc-(1-2)-Man linkage of

- $\beta$ -GlcNAc-(1-2)- $\alpha$ -Man-(1-3)-[ $\beta$ -GlcNAc-(1-2)- $\alpha$ -Man-(1-6)]-Man upon binding. *Glycobiology*, **8**, 173-181.
- Moreno, E., Teneberg, S., Adar, R., Sharon, N., Karlsson, K. & Angström, J. (1997). Redefinition of the carbohydrate specificity of *Erythrina corallodendron* lectin based on solid-phase binding assays and molecular modelling of native and recombinant forms obtained by site-directed mutagenesis. *Biochemistry*, **36**, 4429-4437.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157-163.
- Nishiguchi, M., Yoshida, K., Sumizono, T. & Tazaki, K. (1997). Studies by site-directed mutagenesis of the carbohydrate-binding properties of a bark lectin from *Robinia pseudoacacia*. *FEBS Letters*, **403**, 294-298.
- Osinaga, E., Tello, D., Batthyany, C., Bianchet, M., Tavares, G., Durán, R., Cerveñansky, C., Camoin, L., Roseto, A. & Alzari, P. M. (1997). Amino acid sequence and three-dimensional structure of the Tn-specific isolectin B4 from *Vicia villosa*. *FEBS Letters*, **412**, 190-196.
- Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
- Prabu, M. M., Sankaranarayanan, R., Puri, K. D., Sharma, V., Surolia, A., Vijayan, M. & Suguna, K. (1998). Carbohydrate specificity and quaternary association in basic winged bean lectin: X-ray analysis of the lectin at 2.5 Å resolution. *J. Mol. Biol.* **276**, 787-796.
- Rini, J. M. (1995). Lectin structure. *Annu. Rev. Biophys.* **24**, 551-577.
- Rinnbauer, M., Mikros, E. & Peters, T. (1998). Conformational analysis of a complex between *Dolichos biflorus* lectin and the Forssman pentasaccharide using transferred NOE build-up curves. *J. Carbohydr. Chem.* **17**, 217-230.
- Roberts, D. D. & Goldstein, I. J. (1983). Adenine binding sites of the lectin from lima beans (*Phaseolus lunatus*). *J. Biol. Chem.* **258**, 13820-13824.
- Roberts, D. M. & Etzler, M. E. (1984). Development and distribution of a lectin from the stems and leaves of *Dolichos biflorus*. *Plant Physiol.* **76**, 879-884.
- Roberts, D. M., Walker, J. & Etzler, M. E. (1982). A structural comparison of the subunits of the *Dolichos biflorus* seed lectin by peptide mapping and carboxy terminal amino acid analysis. *Arch. Biochem. Biophys.* **218**, 213-219.
- Schnell, D. J. & Etzler, M. E. (1987). Primary structure of the *Dolichos biflorus* seed lectin. *J. Biol. Chem.* **262**, 7220-7225.
- Schnell, D. J. & Etzler, M. E. (1988). cDNA cloning, primary structure, and in vitro biosynthesis of the DB58 lectin from *Dolichos biflorus*. *J. Biol. Chem.* **263**, 14648-14653.
- Sharma, V. & Surolia, A. (1997). Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J. Mol. Biol.* **267**, 433-445.
- Sharon, N. & Lis, H. (1990). Legume lectins—a large family of homologous proteins. *FASEB J.* **4**, 3198-3208.
- Surolia, A., Sharon, N. & Schwartz, F. P. (1996). Thermodynamics of monosaccharide and disaccharide binding to *Erythrina corallodendron* lectin. *J. Biol. Chem.* **271**, 17697-17703.
- Van Damme, E. J. M., Peumans, W. J., Pusztai, A. & Bardocz, S. (1997). *Handbook of Plant Lectins: Properties and Biomedical Applications*, John Wiley & Sons, New York.
- Weis, W. I. & Drickamer, K. (1996). Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* **65**, 441-473.
- Williams, M. A., Goodfellow, J. M. & Thornton, J. M. (1994). Buried waters and internal cavities in monomeric proteins. *Protein Sci.* **3**, 1224-1235.
- Young, M. N., Watson, D. C., Yaguchi, M., Adar, R., Arango, R., Rodriguez-Arango, E., Sharon, N., Blay, P. K. S. & Thibault, P. (1995). C-terminal post-translational proteolysis of plant lectins and their recombinant forms expressed in *Escherichia coli*. *J. Biol. Chem.* **270**, 2563-2570.

Edited by R. Huber

(Received 1 September 1998; received in revised form 30 December 1998; accepted 30 December 1998)