doi:10.1006/jmbi.2000.4214 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **304**, 423–433

JMB

The 2.2 Å Resolution Structure of the O(H) Blood-group-specific Lectin I from *Ulex europaeus*†

Gerald F. Audette, Margaret Vandonselaar and Louis T. J. Delbaere*

Department of Biochemistry
University of Saskatchewan
107 Wiggins Road, Saskatoon
Saskatchewan, Canada
S7N 5E5

The tertiary and quaternary structure of the lectin I from *Ulex europaeus* (UE-I) has been determined to 2.2 Å resolution. UE-I is a dimeric metalloglycoprotein that binds the H-type 2 human blood group determinant [α-L-Fucα(1 → 2)-β-D-Galβ(1 → 4)-β-D-GlcNAcα-]. Nine changes from the published amino acid sequence were necessary to account for the electron density. The quaternary structural organization of UE-I is that of the most commonly occurring legume lectin dimer. The tertiary structure of the monomeric subunits is similar to that in the conventional lectin subunit; however, some structural differences are noted. These differences include a four-stranded anti-parallel ''S'' sheet in UE-I *versus* the five-stranded S sheet in other lectin monomers. The Ala residue of the Ala-Asp *cis*-peptide bond present in the carbohydrate-binding site of the conventional lectin monomer is replaced with a Thr in the UE-I structure. Also, a novel disulfide bridge linking Cys115 and Cys150 is present. There are two metallic ions, one calcium and the other manganese, per subunit. *N*-linked oligosaccharides are at residues 23 and 111 of each subunit. One molecule of R-2-methyl-2,4-pentanediol (R-MPD) is present in a shallow depression on the surface of each subunit. In order to examine the binding of the H-type 2 blood group determinant by UE-I, its β-methyl glycoside (H-type 2-OMe) was docked into the binding site of R-MPD. The epitope previously identified for H-type 2-OMe by chemical mapping proved, with only minor adjustment of amino acid residues, to be complementary to the shallow cavity occupied by R-MPD in the structure. Several key interactions have been proposed between the H-type 2-OMe and UE-I.

Keywords: Ulex europaeus lectin I; H-type 2 human blood group determinant; protein-carbohydrate interaction; X-ray crystallography; chemical mapping

**Corresponding author*

Introduction

Lectins are a class of proteins which display a polyvalent affinity for oligosaccharides. Their polyvalency arises from the association of their monovalent subunits into mainly dimeric or tetrameric molecules.[1] The association of the individual lectin subunits within the molecule is normally by specific non-covalent interactions of complementary oligopeptide regions on their surfaces. A wide variety of different lectins are present in living cells and tissues. Their carbohydrate specificity is akin to antigenic recognition by antibodies[2] and enzymatic substrate specificity,[3] yet lectins are neither products of the immune system nor do they exhibit enzyme-like catalytic properties. The role of lectins appears limited to specific cross-linking interactions, which allow lectins to serve as recognition molecules for a wide variety of biological events.[4] Lectins from leguminous plants have come under considerable study in recent years. Though it has been suggested that legume lectins play roles in

†This manuscript is dedicated to the memory of Professor Raymond U. Lemieux (1920-2000)

Abbreviations used: UE-I, lectin I from *Ulex europaeus*; UE-II, lectin II from *Ulex europaeus*; GS-IV, lectin IV of *Griffonia simplicifolia*; R-MPD, R-2-methyl-2,4-pentanediol; HSEA, hard sphere exo-anomeric calculation; H-type 2-OMe, the β-methyl glycoside of H-type 2 blood group determinant; NCS, non-crystallographic symmetry; GlcNAc, *N*-acetylglucosamine; Gal, galactose; GalNAc, *N*-acetylgalactosamine..

E-mail address of the corresponding author: louis.delbaere@usask.ca

0022-2836/00/030423–11 $35.00/0

plant cell defense[5] and in bacterial-plant symbiosis,[6] the *in vivo* role of the legume lectin remains unclear. However, the specificity of carbohydrate recognition by legume lectins, coupled with their abundance in legume seeds, ease of purification, and non-enzymatic action have made lectins attractive biochemical tools.

The first lectin from the seeds of *Ulex europaeus* (common gorse) (UE-I) has been shown to agglutinate type O erythrocytes.[7] These cells, being devoid of A or B activity, were originally designated as having zero (O) activity.[8] In fact, the terminal trisaccharide unit on the surface of O-type cells proved to be the precursor of the A and B blood group determinants. This terminal trisaccharide unit is the H-type 2 determinant, which is recognized by UE-I.[9] The agglutination of O-type erythrocytes by UE-I is inhibited by an H-type 2 blood group substance, a glycoprotein present in the saliva of about 80 % of normal individuals who are termed secretors.[10,11] Another H activity may be present on the red cells of normal individuals; namely that expressed by H-type 1 antigens, which is only very weakly recognized by UE-I.[12,13] Indeed, UE-I is used in assays for H-type 2 determinants in the presence of H-type 1 substances;[11] in clinics, the O designation is reserved for the H-type 2 activity. The H-type 1 determinant [α-L-Fucα(1 → 2)-β-D-Gal((1 → 3)-β-D-GlcNAcα-] differs from the H-type 2 determinant (Figure 1) only in the position of the linkage of the β-D-Gal to the β-D-GlcNAc unit of the trisaccharide. The H-type 1 substances are formed in epithelial cells by way of a fucosyl transferase that can complete the synthesis of both H-type 1 and H-type 2 trisaccharides.[14] Conversely, the H-type 2 determinant originates in the endothelia from a specific fucosyl transferase that does not synthesize the H-type 1 determinant.[11]

UE-I is a homodimeric metalloglycoprotein; each subunit is approximately 27 kDa in molecular weight based on the published primary sequence.[15] As a result of the present X-ray analysis, modifications to the primary sequence and the locations of the *N*-linked oligosaccharides were necessary to account for the electron density observed.

Studies into the specificity of carbohydrate recognition by lectins have been greatly assisted by the development of methodologies for the synthesis of complex oligosaccharides.[16] Inhibition studies into the binding of H-type 2 and several congeners by UE-I (Table 1) have revealed several key interactions between the lectin and carbohydrate.[13,17] A method was applied for the systematic chemical mapping for the detection of the H-type 2 epitope that is bound by UE-I.[18] The chemical mapping[19] is based on the crystal structure of the fourth lectin from *Griffonia simplicifolia* (GS-IV) in complex with the Lewis b blood group determinant, a tetrasaccharide.[20] This method has proven successful for the delineation of the interactions between methyl α-isomaltoside and the glucoamylase enzyme as could be judged from the crystal structure of the inhibitor acarbose/enzyme complex.[3] The present project was initiated, in part, to provide further validation of the method by examination of UE-I in complex with the β-methyl glycoside of the H-type 2 determinant (H-type 2-OMe). Although the complex has been difficult to crystallize, useful crystals of the native lectin were produced.[21] This report provides the details of the X-ray structure of UE-I to 2.2 Å resolution. It has been found that the H-type 2-OMe epitope docked into the crystal structure of UE-I is in good agreement with that predicted by chemical mapping.

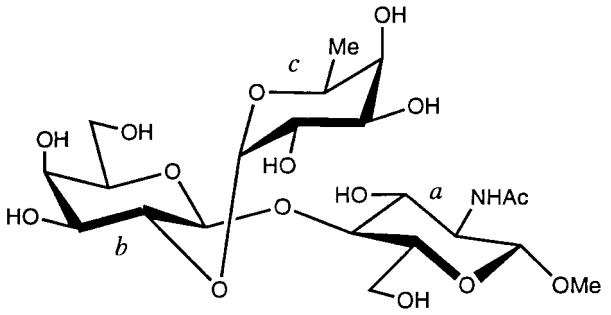

Figure 1. The structure of the β-methyl glycoside of the H-type 2 determinant [α-L-Fuc((1 → 2)-β-D-Galβ(1 → 4)-β-D-GlcNAcα-] (H-type 2-OMe) in a conformational representation which roughly depicts the much preferred conformer. Structural units a, b and c are β-D-GlcNAc-OMe, β-D-Gal and α-L-Fuc, respectively.

Results and Discussion

Quality of the final model

The structure of UE-I reported herein was obtained by refining the model with an initial data set at 2.8 Å resolution and subsequently with a separate data set to 2.2 Å resolution (diffraction data in Table 2). In the electron density maps, no electron density was observed for the final two

Table 1. Chemical mapping of the epitopes of H-type 2-OMe recognized by UE-I

	Relative potencies[b]	
Position altered[a]	Monodeoxy	Mono-O-methyl
3a	46	40
6a	250	150
3b	3	70
4b	37	44
6b	60	45
2c	0.6	Inactive
3c	0.11	Inactive
4c	Inactive	Inactive
5c (*nor*)[13]	3	-

[a]Structural unit designations are as described in Figure 1.
[b]By radioimmunoassay, setting the potency of H-type 2-OMe at 100.[61]
Source: Du *et al.* (1994).[17]

Table 2. Summary of data collection and refinement statistics

A. *Data collection*		
Data set	1.[a]	2
Resolution range	8.5-2.8 Å	40-2.2 Å
(Last shell)	(3.0-2.8 Å)	(2.32-2.2 Å)
Unique reflections	15,829	29,087
% Completeness	96.9	89.2
(Last shell)	(92.1)	(63.9)
R_{Symm}[b]	0.097	0.056
B. *Refinement statistics*		
Diffraction agreement		
Resolution range (Å)	40 - 2.2	
R-value.[c,d]	0.171	
R_{Free}[e]	0.217	
No. of unique reflections	26,189	
Quality of structure		
Model parameters		
Total protein non-H atoms	3695	
Solvent molecules	192	
B-factor model	Individual isotropic	
Mean *B*-factors (Å^2)[f]		
Main-chain	34.2	
Side-chain	35.8	
N-linked sugars	68.7	
Water molecules	41.3	
r.m.s.d from stereochemical ideality		
Bonds (Å)	0.011	
Angles (deg.)	1.738	
Dihedrals (deg.)	26.50	
Impropers (deg.)	1.527	

[a]Data set 1 was obtained by merging data from two separate crystals.

[b]$R_{Symm} = \Sigma_j|I_j - \langle I\rangle|/\Sigma_j\ I_j$, where j = a set of observations of equivalent reflections and $\langle I\rangle$ = average intensity of the reflection.

[c]R-value $= \Sigma_{hkl}\ ||F_O| - |F_C||/\Sigma_{hkl}\ |F_O|$.

[d]All data, final refinement cycle.

[e]Randomly chosen test set of 10% of unique reflections (2898 reflections).

[f]Mean *B*-factors are averaged over both monomers.

C-terminal amino acid residues in subunit A nor for the final three C-terminal amino acid residues in subunit B. An *N*-linked heptasaccharide was found attached to Asn23 of subunit A, whereas a trisaccharide was located at Asn111 of both subunits. A GlcNAc residue may be present at Asn23 of subunit B, however, the electron density was not well resolved. There are 192 solvent (water) molecules present in the final model, and a molecule of R-MPD was found bound in a shallow cavity on the surface of both subunits. The stereochemical quality of the final model was analyzed using the program PROCHECK;[22] 89.0% of the residues are within the most favored regions of the Ramachandran plot, and the remaining 11.0% are in the additional allowed regions. Mean coordinate error from Luzatti plot analysis[23] is 0.1 Å. The final refinement statistics are summarized in Table 2.

The Structure of UE-I

Nine amino acid residues in each subunit appear in the electron density maps to be different from those found in the published sequence.[15] Residues 23 and 111, sequenced as Asp and Glu residues, respectively, were changed to an Asn residue because clearly carbohydrate was attached at those positions. Residues 115 and 150, sequenced as Asp and Arg residues, respectively, were changed to a Cys residue due to a disulfide bridge connecting these two residues in both subunits. The other changes, Val65 to a Leu residue, Asp137 to a Trp residue, Gly218 to a Ser residue, Asn237 to a Thr residue, and Leu238 to a Phe residue were made on the basis of the shape of the electron density for the side-chains at a 2σ contour level. Asn116 was deleted from the model in both subunits because it could not be accommodated in the electron density map. The sequence from residue 117 to 243 was retained with the same numbering as the published sequence. Including the total 242 amino acid residues with the changes in the primary sequence, the resultant molecular weight of each polypeptide subunit is 26,606 Da.

Subunit A is more ordered in the crystal than subunit B. The electron density map for the backbone atoms in subunit A is continuous at a 2σ contour level while the map for the second subunit has many gaps in backbone density at 2σ but is continuous at a 1σ contour level. Subunit A has four side-chains: Glu28, Lys83, Asn240, and Thr241 for which there is little electron density, while subunit B has six side-chains: Glu28, Asp39, Asn40, Lys83, Lys114, and Asn240. Residues at these positions were replaced with alanine during structure refinement. The average *B*-factors for subunit A are 22.5 Å^2 and 25.0 Å^2 for main-chain and side-chain atoms, respectively, while the corresponding *B*-factors for subunit B are 45.8 Å^2 and 46.5 Å^2, respectively. The lower *B*-factors of subunit A suggest stronger crystal packing interactions in the unit cell for this subunit. Indeed, examination of symmetry interactions indicates that there are 97 contacts (<4.1 Å) between subunit A and symmetry-related molecules, whereas subunit B only makes 40 contacts with other molecules in the crystal.

Legume lectins form several different quaternary structures, which are either dimeric or tetrameric in organization. Of the dimeric lectins, the most common quaternary structure is the dimer initially described for the concanavalin A lectin (Con A)[24] and subsequently described in several other lectins (reviewed by Loris *et al.* (1995)).[25] Three other dimeric forms are described in the literature, these being the *Erythrina corallodendron* lectin (EcorL)[26] the lectin IV of *Griffonia simplicifolia* (GS-IV),[20] and the stem/leaf lectin of *Dolichos biflorus* (DB58).[32] The tetrameric organization has been described for several lectins: the peanut lectin (PNA),[27,28] soybean agglutinin (SBA),[29] phytohemagglutinin-L (PHA-L),[30] lectin II from *Ulex europaeus*(UE-II),[31] and the seed lectin from *D. biflorus* (DBL).[32] Although the quaternary structures of reported legume lectins differ widely, the secondary and tertiary structures of the lectin subunit are relatively constant. The legume lectin fold was initially described as consisting of two anti-parallel

β-sheets,[24,33] a six-stranded back sheet and a 12-stranded front sheet. More recently, this has evolved into a more detailed description of three β-sheets: a six-stranded back sheet, a seven-stranded front sheet, and a five-stranded ''S'' sheet which connects the front and back sheets.[25,28]

The overall quaternary structure of UE-I (Figure 2(a)) is that of the most commonly occurring legume lectin dimer.[25] The six-stranded anti-parallel back sheet of each monomeric subunit form a 12-stranded anti-parallel β-sheet scaffold in the dimeric molecule. Upon this scaffold, the front and S-sheets of the monomeric subunits are located with the carbohydrate binding sites present at both ends of the dimer. Both subunits of the UE-I dimer are seen to have the expected legume lectin fold, with some noticeable differences (Figure 2(b)). The six-stranded back sheet is present, consisting of residues Asp3 to Phe8, Asn231 to Thr237, Ala66 to Ser73, Ala169 to Tyr175, Thr181 to Thr187, and Thr193 to Ser199. The seven-stranded front sheet consists of residues Asp17 to Gly22, Thr45 to Tyr51, Val213 to Thr219, Asp87 to Ala94, Thr122 to Thr129, His143 to Asn149, and Ile154 to Arg158. The five-stranded S sheet[25] is a four-stranded sheet in the UE-I structure (highlighted in Figure 2(b)), comprised of residues Phe74 to Gln77, Glu224 to Leu230, Val31 to Val37, and Asn23 to Ile27. A second difference between the UE-I structure and that of the conventional legume lectin subunit is the presence of a β-strand connecting the front and back sheets in each subunit. This strand, comprised of residues Arg158 to Gly164, is an extension of the seventh strand of the front sheet, and connects to the fourth strand of the back sheet. Also, the amino acid Ala of the Ala-Asp *cis*-peptide bond seen in the binding site of many lectins for carbohydrate recognition[25] is a Thr in UE-I. Finally, the presence of a disulfide bridge (between residues 115 and 150) in UE-I is novel for the legume lectin.

N-linked carbohydrate

Covalently attached oligosaccharides were found at Asn23 and Asn111 of subunit A (Figure 2(a)). At position 23 on subunit A, a heptasaccharide commonly found in plant lectins,[34] namely α-D-Man-(1 ← 3)-[α-D-Man-(1 → 6)]-[β-D-Xyl-(1 → 2)]-β-D-Man-(1 → 4)-β-D-GlcNAc-(1 → 4)-[α-L-Fuc-(1→ 3)]-D-GlcNAc, could be fitted into electron density which stretched across a solvent channel to interact with a symmetry-related molecule. In the case of subunit B, some electron density for *N*-linked oligosaccharide was detected at Asn23; however, the

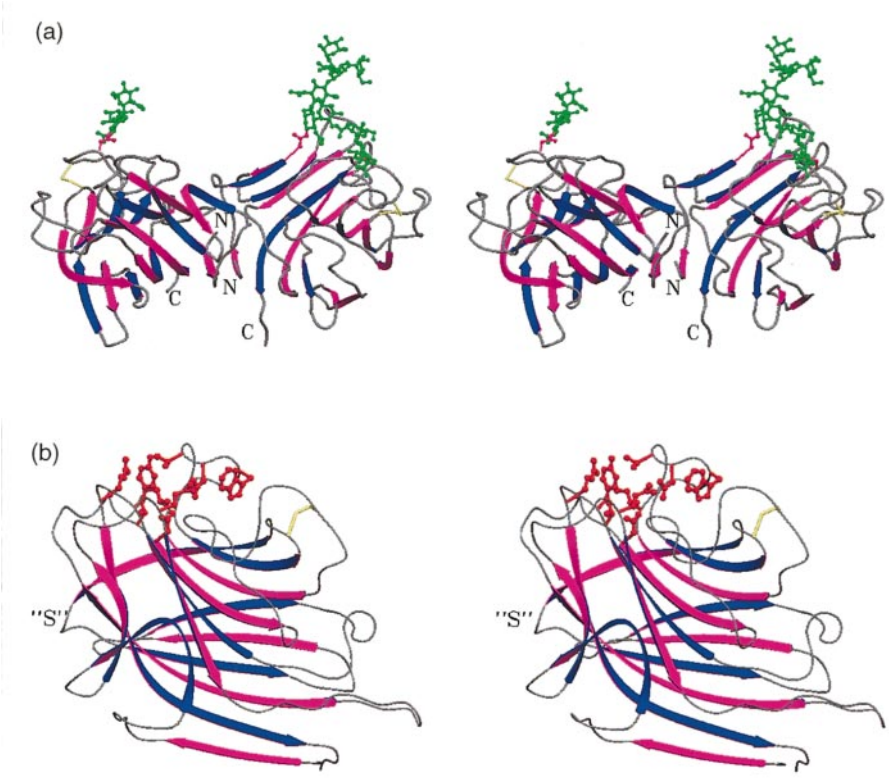

Figure 2. The structure of UE-I. (a) Stereographic representation of the C^{α} trace of the UE-I dimer. Beta strands are shown as blue and magenta arrows, and random coil in gray. The disulfide bridge between residues 115 and 150 is in yellow. *N*-linked carbohydrate in green, with Asn23 and Asn111 in magenta. (b) Stereographic representation of the C^{α} trace of the UE-I monomer. Colors are as in (a), with the residues of the carbohydrate-binding site in red. The ''S''-Sheet[25] is highlighted in the figure with an ''S''. Figures 2 through 5 were produced using the program SETOR.[62]

electron density was not sufficiently resolved to allow for modeling of a GlcNAc residue. A trisaccharide, α-L-Fuc-(1 → 3)-β-D-GlcNAc-(1 → 4)-D-GlcNAc, could be fitted into electron density at Asn111 of both subunits. The average *B*-factors for the attached oligosaccharides were quite high; 59.7 Å^2 for subunit A and 77.6 Å^2 for subunit B.

The carbohydrate-binding site

Though leguminous lectins exhibit both sequence and structural similarities, their carbohydrate binding specificities differ widely. With regards to carbohydrate specificity, the legume lectins are separated into five classes based upon the monosaccharide which exhibits the highest binding affinity, even though the binding of the monosaccharide is much weaker than that of the complete carbohydrate moiety recognized by the lectin. These classes are: Glc/Man; Gal/GalNAc; Fuc; chitobiose (GlcNAc); and Complex, in which the lectin is specific for a complex oligosaccharide rather than a single monosaccharide moiety. X-ray crystallographic analysis of lectins from the Glc/Man, Gal/GalNAc and complex specificities encompass the majority of knowledge about lectins and their interactions with carbohydrate to date.[25] Recently, the quaternary structure of the second lectin from *U. europaeus* (UE-II) has been reported.[31] UE-II is a chitobiose-specific lectin, and a detailed X-ray analysis should shed light into the nature of its carbohydrate specificity. Very little structural information is known about the fucose-specific lectins, though the molecular modeling of UE-I has been reported.[35] Although the isolation of a second α-L-fucose-specific legume lectin has been reported, the lectin from *Lotus tetragonolobus*,[36-38] to date the X-ray structure of a fucose-specific legume lectin has not been available.

Konami *et al*.[39] found that the peptide Asp128 to Trp137 of UE-I was retained on a Fuc-Gel affinity column and concluded that these residues constitute the carbohydrate binding site. Other legume lectin-carbohydrate complexes[25,40,41] indicate that residues homologous with Asp87, Asn135 and Gly104 of UE-I form hydrogen bonds with sugars. Residues involved in carbohydrate binding in the GS-IV lectin are Arg48, Ser49, Asp89, Tyr105, Gly106, Gly107, Phe108, His114, Asn135, Trp138, Tyr223.[20] Figure 3, showing the carbohydrate binding site of GS-IV superimposed upon that of UE-I using the program LSQKAB of the CCP4[42] suite of programs, illustrates the homology between the binding sites of both lectins. The UE-I binding site is therefore taken to be a depression made up of residues Glu44, Thr86, Asp87, Gly104-Gly105, Ile130, Val134, Asn135, Trp137, Tyr220, and Arg223.

The binding of R-MPD

A molecule of the crystallization precipitant, R-2-methyl-2,4-pentanediol (R-MPD), was found to occupy a binding site in each subunit of UE-I. This observation proved to be of great assistance in the docking of the H-type 2-OMe, especially since the structure of the epitope was known.[17] The binding of R-MPD by UE-I is very weak (D.R. Bundle, personel communication), and is a result of the large excess of the racemic diol being present during crystallization. Hydrogen bonds were observed between O2 of R-MPD and Glu44 OE1, O2 of R-MPD and Gly105 N, O4 of R-MPD and Gly105 N, and O4 of R-MPD and Asn135 ND2 and OD1. There are also distances of < 4.1 Å between non-hydrogen atoms of R-MPD and non-hydrogen atoms of residues Thr86, Gly104, Val134, Asn135, Trp137, and Tyr220 of UE-I.

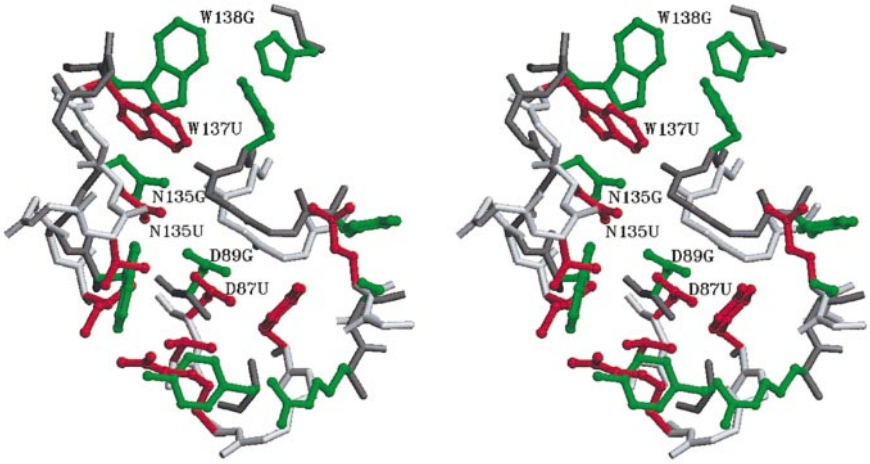

Figure 3. Overlay of the carbohydrate binding sites of UE-I (red) and the GS-IV lectin (green). The letters U and G in residue labels denote UE-I and GS-IV, respectively.

Cis-peptide bonds of the binding site

There are two *cis*-peptide bonds within the UE-I structure, and both are involved in the carbohydrate binding of UE-I. These two *cis*-peptide bonds are between residues Thr86 and Asp87 and between residues Ser132 and Pro133. The Thr86-Asp87 *cis*-peptide bond is analogous to the Ala-Asp *cis*-peptide bond observed in other lectin structures.[25] This *cis*-peptide is present in order to position the side-chain of the aspartate residue in the binding site of the protein. It is possible that the hydroxyl group of the Thr86 side-chain is also involved in carbohydrate binding. In the present structure, the hydroxyl group of the Thr86 side-chain points away from the carbohydrate binding site and hydrogen bonds with the Asp87 side-chain. If this group is indeed involved in carbohydrate binding, it would require a rotation about the C^{α}-C^{β} bond to position it within the binding site; however, in the present structure, this rotation about the Thr86 C^{α}-C^{β} bond is not in agreement with the observed electron density. The second *cis*-peptide bond in the UE-I structure involves proline, more commonly observed in protein structures, and is between Ser132 and Pro133. This *cis*-peptide bond positions Val134 into the carbohydrate-binding site, and therefore contributes to the hydrophobic nature of the binding site.

Calcium and manganese ions

Positively charged metal ions form part of a building unit in legume lectins. Both a calcium ion and a transition metal ion, generally manganese, are present and are adjacent to the binding site. The presence of these ions is considered to help lock the tertiary structure of the lectin into a favorable conformation for carbohydrate binding. As previously suggested by Delbaere *et al.*,[43] the positive charges may induce electron delocalization favorable to the stability of the complex. Both a calcium and manganese ion are present in each subunit of UE-I (Figure 4). The manganese ion is in a octahedral coordination; as is commonly found in proteins,[44] the calcium ion is coordinated by seven oxygen atoms in a pentagonal bipyramidal configuration. The ions are 4.17 Å apart and are 7.8 Å (Ca^{2+}) and 11.6 Å (Mn^{2+}) from the carbohydrate binding site, using C3 of R-MPD as an approximate center of the binding site.

The comparison of the UE-I structure to the previously predicted model

Gohier and co-workers[35] predicted the structure of the UE-I binding site by sequence alignment of several lectins with the published sequence of Konami *et al.*[15] and a structural superposition of previously determined lectin structures. In order to compare the previously predicted and actual observed binding sites, the previously modeled UE-I was superimposed onto the binding site of the present structure using the program LSQKAB of the CCP4[42] suite. The Asp87, Ile130 and Tyr220 residues were observed to be in similar locations in both structures. The Phe136 residue of the previously predicted UE-I structure was located in a position equivalent to that occupied by Trp137, previously sequenced as an aspartate.[15] Also, the location of the Ser132 side-chain was observed to be very different between the two structures. These differences can mainly be attributed to anomalies between the previously published primary sequence of UE-I[15] and the observed electron density.

The docking of H-type 2-OMe into the carbohydrate-binding site of UE-I

Studies into the interactions involved in the specificity of carbohydrate recognition by lectins have been greatly assisted by the development of methodologies for the synthesis of complex oligosaccharides.[16] The ability to chemically synthesize complex oligosaccharides provided access to congeners such as all the monodeoxy and mono-
O-methyl derivatives. By determining the effects

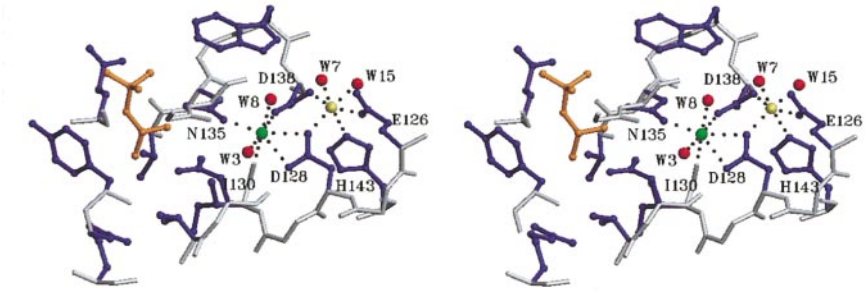

Figure 4. Stereographic representation of the carbohydrate-binding site and metal ion coordination in the UE-I subunit. Main-chain atoms are in gray, the calcium ion in green, the manganese ion in yellow, the R-MPD molecule in goldenrod, water molecules are in red, and side-chains involved in carbohydrate binding and/or ion coordination are in purple. Amino acid residues and water molecules involved in metal ion coordination are labeled.

of these slight structural changes on binding, a chemical map of the epitope, that is the groups which interact between the two molecules upon complexation, can be predicted.[19]

It is known that UE-I specifically binds the H-type 2 blood group determinant [α-L-Fucα (1 → 2)-β-D-Galβ(1 → 4)-β-D-GlcNAcα-]. Therefore, when both the deoxygenation and the *O*-methylation of the H-type 2-OMe essentially abolish binding (Table 1), the conclusion can be made that these hydroxyl groups provide critical hydrogen bonds with the protein. Also, when deoxygenation significantly decreases binding but the *O*-methylation causes little change in (and on occasion strengthens) binding, the indication is taken that the parent hydroxyl group resides at the periphery of the binding site in the complex where it will interact with peripheral amino acid residues. Finally, when small changes in binding are seen as a result of the structural modifications, it may be interpreted as arising from changes in hydration of polar groups which remain in the aqueous phase.

Noting the above interpretations, a chemical map of the H-type 2-OMe (Figure 1) can be predicted. The inhibitory studies[17] indicate that the three hydroxyls (O2c, O3c, O4c) of the α-L-Fuc unit provide critical hydrogen bonds to UE-I, the 3b-hydroxyl group of the β-D-Gal unit is interacting at the periphery of the UE-I binding site, and that the 3a-hydroxyl group of the β-D-GlcNAc unit remains within the aqueous phase. In addition to these polar interactions, replacement of the 6a-OH group of the β-D-GlcNAc moiety by a hydrogen atom more than doubled the binding potency, indicating a hydrophobic interaction. Also, the hydrophobic *C*-5-CH_3 of the fucose is close enough to be a part of this hydrophobic region, and its replacement by hydrogen has a significant deleterious effect on binding (Table 1).[13] It should be noted, however, that the replacement of the methyl group by a hydrogen atom may have a profound effect on the conformational preference of a building unit that provides three key hydroxyl groups to the binding region.

In order to examine the interactions between UE-I and the H-type 2 blood group determinant, the H-type 2-OMe was docked into the UE-I carbohydrate binding site and energy minimization was performed. It has been established that chemical alterations of the oligosaccharide cause no significant change in the conformational preference.[17] It has also been shown that in the crystal structures of several complexes of Le^b-OMe congeners with GS-IV, the occupancies of the congeners were found to be identical to that of the parent Le^b-OMe (M. Vandonselaar, personal communication). Therefore, the model used for the present studies was provided by the chemical map of H-type 2-OMe predicted from the inhibitory studies, with coordinates obtained from hard sphere exo-anomeric (HSEA) calculations.[45] Several potential hydrogen bonds and potential van der Waals interactions between the H-type 2-OMe and the binding site of UE-I were found in the model of the complex (Figure 5).

The potential hydrogen bonding pattern of the fucopyranose unit of the modeled H-type 2-OMe would explain the inhibitory studies into the binding of the H-type 2-OMe by UE-I.[13,17] All three hydroxyl groups of the fucopyranose unit are within the binding site and provide potentially critical hydrogen bonds with the protein; O3c and O4c hydroxyls would provide direct hydrogen bonds to the protein, while O2c is likely involved in a water-bridged interaction with UE-I. Disruption of these potential hydrogen bonds by alteration of these hydroxyl groups would lead to greatly diminished or abolished binding. The side-chains of Val134, Trp137 and Tyr220 in the binding site of UE-I form a surface which helps to surround the polar interactions of the key hydroxyl groups.

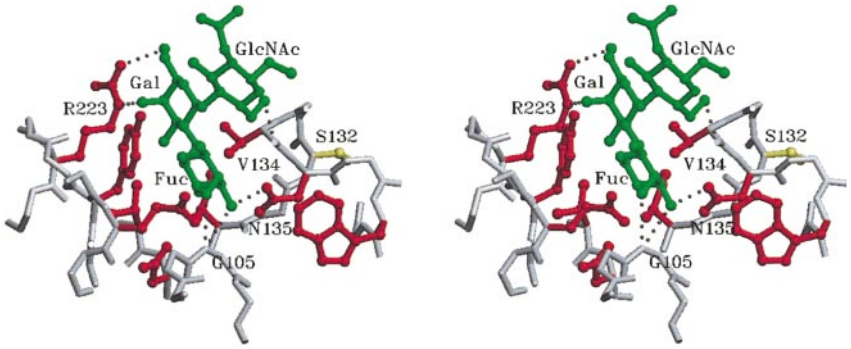

Figure 5. Stereographic representation of the docked H-type 2-OMe within the carbohydrate-binding site. All three hydroxyls of the fucopyranose unit are deep within the binding site and potentially provide critical interactions with UE-I. O3c and O4c would provide direct hydrogen bonds to G105-N and N135-ND2. O2c likely interacts with Asp87-OD2 and Asn135-OD1 through a bridging water molecule. The H-type 2-OMe is in green, main-chain atoms in gray, binding site residues in red. Hydrogen bonds between the H-type 2-OMe and binding site are represented as dark gray broken lines. Ser132, previously thought to be present in the binding site,[35] is in yellow.

An ordered water structure in the combining site of lectins is thought to be required for carbohydrate binding.[46-49] This knowledge, as well as the unusually high entropy barrier to H-type 2-OMe binding by UE-I ($T\Delta S = -20.5$ kcal/mol), allows for the inference that ordered water molecules will be present within the UE-I-binding site and interact with both the protein and the bound H-type 2 trisaccharide. The ordered water molecules are expected to be liberated upon complex formation and thus contribute significantly to the high enthalpy of binding (−29 kcal/mol). The term ''hydraphobic effect'' has been proposed[48] to identify this water-related contribution to the binding of strongly hydrophilic surfaces in the presence of very large molar excesses of water, such as found at the surface of proteins.

No ordered water molecules are seen in the carbohydrate-binding site of the present structure. It is believed that the bound R-MPD molecule is sufficiently hydrophobic in nature to extrude water from the binding site. It is likely that the O2c hydroxyl of the fucopyranosyl moiety of the H-type 2-OMe hydrogen bonds to the protein through ordered water molecules within the binding site. This potential water bridging is seen by the interatomic distances between O2c and Asp87 $O^{\delta 2}$ (3.87 Å), and between O2c and Asn135 $O^{\delta 1}$ (3.94 Å) and $N^{\delta 2}$ (3.98 Å) of UE-I. Other possible water-mediated interactions are observed between O4c and Glu44 $O^{\varepsilon 1}$ (3.91 Å), O5c and Tyr220 OH (4.34 Å), and between O3b and Arg223 N^{ε} (3.85 Å) and Arg223 NH2 (3.88 Å).

Materials and Methods

Ulex europaeus lectin I was purchased from Boehringer Mannheim (lot # 13178220-07) in pure form, and no further purification was performed. Crystals were grown by a sitting drop vapor diffusion technique in which 20 μl drops consisting of 5 mg/ml protein, 0.05 mM calcium chloride, 0.05 mM manganese chloride, 22.5 % R,S-2-methyl-2,4-pentanediol (R,S-MPD), and 0.05 M sodium acetate buffer (pH 4.0) were placed in sitting drop rods and equilibrated over reservoirs containing 45 % R,S-MPD, 0.1 M sodium acetate buffer (pH 4.0) for one week. Washed seed crystals were then introduced into the protein drops and equilibration was continued until the seed crystals stopped growing (5-6 months). The crystals are in the monoclinic space group *C*2, with unit cell dimensions $a = 79.61$ Å, $b = 70.67$ Å, $c = 122.15$ Å, and $\beta = 108.6^\circ$. This technique is a modification of the one originally published.[21]

Data from three crystals were collected on an Enraf Nonius FAST area detector. The X-ray source was an FR571 rotating anode generator operated at 45 kV, 95 mA, with copper radiation and a graphite monochromator. The data from two crystals were merged to produce a data set which was 96.9 % complete to 2.8 Å resolution. A summary of the diffraction data is given in Table 2. The structure was solved by molecular replacement using as a starting model the *G. simplicifolia* lectin IV coordinates (PDB accession code 1LEC)[20] edited to remove all side-chains which were not homologous in the two sequences. The program AMoRe from the CCP4 suite[42,50] was used for the molecular replacement search. The rotation search, carried out with data in the 10-4 Å resolution range, produced two major peaks, 9σ and 7σ above background. A translation search on the first peak alone yielded a solution with a correlation coefficient 0.314 and *R*-value 0.526. A translation search for the second rotation peak was then carried out with the first peak fixed. The correlation coefficient for the two peaks together was 0.436 and the *R*-value was 0.490. An AMoRe rigid body refinement resulted in a solution with correlation coefficient 0.596 and *R*-value 0.452. The program LSQKAB from the CCP4[42] suite was used to apply the two rotational and translational solutions to the original model coordinates. The resulting model turned out to be the 12-stranded beta-sheet dimer found in several other legume lectins[25] and was therefore accepted as the correct molecular replacement solution.

X-PLOR[50,51] rigid body refinement did not improve the *R*-value but gave a more interpretable $2\mathbf{F}_O - \mathbf{F}_C$ map than the one calculated from the AMoRe model. The X-PLOR map and a SIGMAA[53,54] map were used to begin manual rebuilding of the model with TURBO-FRODO.[55] Manual fitting was applied to only one subunit; the other was generated using a matrix found by superimposing one subunit onto the other, calculated with the LSQKAB program using coordinates from X-PLOR rigid body refinement. New $3\mathbf{F}_O - 2\mathbf{F}_C$, SIGMAA, and non-crystallographic symmetry (NCS) averaged maps were calculated using programs in the CCP4 package. Three rounds of manual model building, alternated with new map calculations, resulted in a model which had 219 correct side-chains built into the electron density.

At this stage, the model was subjected to X-PLOR simulated annealing refinement consisting of a slow-cooling step from an initial temperature of 4000 K to a final temperature of 300 K using steps of 50 K with a time step of 0.2 fs. This was followed by 120 cycles of conjugate gradient minimization and individual *B*-factor refinement. Coordinates of only one subunit were given and strict NCS symmetry was imposed using the refined matrix from the NCS-averaged program DM.[63] A randomly selected test set of 10 % of the reflections in the 10-2.8 Å range was excluded from the refinement to evaluate R_{Free}[56] After further modeling, the resolution was extended to 2.2 Å by refining with data from a third crystal. Simulated annealing with strict NCS symmetry gave an R_{Free} of 0.415 and *R*-value of 0.350. Simulated annealing of the dimer with NCS restraints gave an R_{Free} of 0.331 and *R*-value of 0.284. In subsequent cycles of manual fitting, each subunit was rebuilt individually and NCS restraints were no longer used in refinement. Three more rounds of model building and simulated annealing refinement were carried out, during which covalently linked carbohydrate, solvent (water) molecules, and one molecule of the precipitant, R-2-methyl-2,4-pentanediol (R-MPD) seen to be present within each binding site (one R-MPD molecule per subunit), were added to the model. The program ARP[57] was used to locate water molecules in $3\mathbf{F}_O - 2\mathbf{F}_C$ and $2\mathbf{F}_O - \mathbf{F}_C$ electron density maps. Near the end of the model building process, only X-PLOR positional refinement was carried out. The bond and angle parameters, reported by Engh & Huber[58] were employed throughout the refinement. The final R_{free} and *R*-value were 0.242 and 0.185, respectively, using reflections in the 10.0-2.2 Å range. A final refinement which included all data and a bulk solvent correction gave an *R*-value of 0.171 and an R_{Free} of 0.217. Examination of interactions between the final model and

symmetry-related molecules within the crystal was performed using the program CONTACSYM.[59,60] A summary of the final refinement statistics is shown in Table 2.

Modeling of the H-type 2-OMe into the UE-I-binding site proceeded in the following fashion. Coordinates for the β-methyl glycoside of the H-type 2 trisaccharide (H-type 2-OMe) molecule were obtained from hard sphere exo-anomeric (HSEA) calculations.[45] The H-type 2-OMe was manually docked into the UE-I-binding site using TURBO-FRODO. The first monomer of the UE-I dimer was used in modeling studies due to its lower *B*-factors in the structure. O4 of the α-L-fucose moiety was overlaid onto O2 of the R-MPD molecule, and the remainder of the H-type 2-OMe was modeled into the binding site making the most favorable contacts. The protein (minus R-MPD) and docked carbohydrate were then subjected to 600 cycles of conjugant gradient minimization using X-PLOR. During this minimization, the X-ray term was turned off. Examination of the final docked/minimized structure was performed using TURBO-FRODO, and interactions between UE-I and the H-type 2-OMe were identified using the program CONTACSYM and manual inspection using TURBO-FRODO.

Protein Data Bank accession code

Coordinates of the final model and structure amplitudes have been deposited into the RCSB Protein Data Bank (accession code 1FX5).

Acknowledgments

The authors express their appreciation to Professor Raymond U. Lemieux, who has recently passed away. Professor Lemieux provided many valuable discussions on this work. The authors also wish to thank Dr Lata Prasad for insightful discussions and critical review of the manuscript, and Miss Athena Sudom for helpful discussions. G.F.A. gratefully acknowledges financial support from the Colleges of Graduate Studies and Medicine at the University of Saskatchewan. This research was supported by a Medical Research Council of Canada operating grant to L.T.J.D. (MT-10162).

References

1. Lis, H. & Sharon, N. (1998). Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chemical Rev.* **98**, 637-674.
2. Lemieux, R. U., Cromer, R. & Spohr, U. (1988). Molecular Recognition VIII. The binding of the β-D-galactopyranosyl residue of the Lewis b human blood group determinant by the lectin IV of *Griffonia simplicifolia* and by a monoclonal anti-Lewis b antibody. Evidence for intramolecular hydrogen bonding. *Can. J. Chem.* **66**, 3083-3098.
3. Lemieux, R. U., Spohr, U., Bach, M., Cameron, D. R., Frandsen, T. P., Stoffer, B. B., *et al*. (1996). Chemical mapping of the active site of the glucoamylase of Aspergillus niger. *Can. J. Chem.* **74**, 319-335.
4. Sharon, N. & Lis, H. (1989). Lectins as cell recognition molecules. *Science,* **246**, 227-234.
5. Chrispeels, M. J. & Raikhel, N. V. (1991). Lectins, lectin genes, and their role in plant defense. *Plant Cell,* **3**, 1-9.
6. Diaz, C. L., Melchers, L. S., Hooykaas, P. J. J., Lutenberg, B. J. J. & Kijne, J. W. (1989). Root lectins as a determinant of host-plant specificity in the Rhizobium-legume symbiosis. *Nature,* **338**, 579-581.
7. Matsumoto, I. & Osawa, T. (1969). Purification and characterization of an anti-H(O) phytohemagglutinin of *Ulex europaeus*. *Biochim. Biophys. Acta,* **194**, 180-189.
8. Landsteiner, K. (1920). Specifishe Serumreaktionen mit einfach zusammengestzten Substanzen von bekannter Konstitution (organischen Säuren). *Biochem. Z.* **104**, 280-299.
9. Watkins, W. M. (1972). Blood-group specific substances. In *Glycoproteins: Their Composition, Structure and Function* (Gottschalk, A., ed.), pp. 830-891, Elsevier, Amsterdam.
10. Boyd, W. C. & Shapleigh, E. (1954). Separation of individuals of any blood group into secretors and non-secretors by use of a plant agglutinin (lectin). *Blood,* **9**, 1195-1198.
11. Le Pendu, J., Lemieux, R. U., Lambert, F., Dalix, A.-M. & Oriol, R. (1982). Distribution of H type 1 and H type 2 antigenic determinants in human sera and saliva. *Am. J. Hum. Genet.* **34**, 402-415.
12. Pereira, M. E. A., Kisailus, E. C., Gruezo, F. & Kabat, E. A. (1978). Immunochemical studies on the combining site of the blood group H-specific lectin I from *Ulex europaeus* seeds. *Arch. Biochem. Biophys.* **185**, 108-115.
13. Hindsgaul, O., Norberg, T., Le Pendu, J. & Lemiuex, R. U. (1982). Synthesis of type 2 human blood-group antigenic determinants. The H, X, and Y haptens and variations of the H type 2 determinant as probes for the combining site of the lectin I from *Ulex europaeus*. *Carbohyd. Res.* **109**, 109-142.
14. Le Pendu, J., Cartron, J. P., Lemiuex, R. U. & Oriol, R. (1985). The presence of at least two different H-blood-group-related β-D-Gal α-2-L-fucosyltransferases in human serum and the genetics of blood group H substances. *Am. J. Hum. Genet.* **37**, 749-760.
15. Konami, Y., Yamamoto, K. & Osawa, T. (1991). The primary structures of two types of the *Ulex europaeus* seed lectin. *J. Biochem.* **109**, 650-658.
16. Lemieux, R. U. (1978). Human blood groups and carbohydrate chemistry. *Chem. Soc. Rev.* **7**, 423-452.
17. Du, M., Sphor, U. & Lemieux, R. U. (1994). The recognition of three different epitopes for the H-type 2 human blood group determinant by lectins of *Ulex europaeus*, *Galactia tenuiflora* and *Psophocarpus tetragonolobus* (Winged Bean). *Glycoconj. J.* **11**, 443-461.
18. Spohr, U., Paszkiewicz-Hnatiw, E., Morishima, N. & Lemieux, R. U. (1992). Molecular recognition XI. The synthesis of extensively deoxygenated derivatives of the H-type 2 human blood group determinant and their binding by an anti-H-type 2 monoclonal antibody and the lectin 1 of *Ulex europaeus*. *Can. J. Chem.* **70**, 254-271.
19. Nikrad, P. V., Beiberbeck, H. & Lemieux, R. U. (1992). Molecular recognition. X. A novel procedure for the detection of the intermolecular hydrogen bonds present in a protein·oligosaccharide complex. *Can. J. Chem.* **70**, 241-253.
20. Delbaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Wilson, K. & Dauter, Z. (1993). Structures of the Lectin IV of *Griffonia simplicifolia* and its complex with the Lewis b human blood group

determinant at 2.0 Å resolution. *J. Mol. Biol.* **230**, 950-965.
21. Vandonselaar, M. & Delbaere, L. T. J. (1994). Crystallization of *Ulex europaeus* lectin I. *J. Mol. Biol.* **243**, 345-346.
22. 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.
23. Luzatti, V. (1952). Statistical treatment of errors in the determination of crystal structures. *Acta Crystallog. sect. A,* **5**, 802-810.
24. Hardman, K. D. & Ainsworth, C. F. (1972). Structure of concanavalin A at 2.4 Å resolution. *Biochemistry,* **11**, 4910-4919.
25. Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). Legume lectin structure. *Biochim. Biophys. Acta,* **1383**, 9-36.
26. Shaanan, B., Lis, H. & Sharon, N. (1991). Structure of a legume lectin with an ordered *N*-linked carbohydrate in complex with lactose. *Science,* **254**, 862-866.
27. Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanaraj, V., Mahanta, S. K., *et al*. (1994). Crystal structure of peanut lectin, a protein with an unusual quaternary structure. *Proc. Natl Acad. Sci. USA,* **91**, 227-231.
28. Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1996). Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex. *J. Mol. Biol.* **259**, 281-296.
29. Dessen, A., Gupta, D., Sabesan, S., Brewer, C. F. & Sacchettini, J. C. (1995). X-ray crystal structure of the soybean agglutinin cross-linked with a biatennary analog of the blood group I carbohydrate antigen. *Biochemistry,* **34**, 4933-4942.
30. Hamelryk, T. W., Dao-Thi, M.-H., Poortmans, F., Chrispeels, M. J., Wyns, L. & Loris, R. (1996). The crystallographic structure of phytohemagglutinin-L. *J. Biol. Chem.* **271**, 20479-20485.
31. Dao-Thi, M.-H., Rizkallah, P., Wyns, L., Poortmans, F. & Loris, R. (1998). Quaternary structure of UEA-II, the chitobiose specific lectin from gorse. *Acta Crystallog. sect. D,* **54**, 844-847.
32. Hamelryk, T. W., Loris, R., Bouckaert, J., Dao-Thi, M.-H., Strecker, G. & Imberty, A., *et al*. (1999). Carbohydrate binding, quaternary structure and a novel hydrophobic binding site in two legume lectin oligomers from *Dolichos biflorus*. *J. Mol. Biol.* **286**, 1161-1177.
33. Edelman, G. M., Cunningham, B. A., Reeke, G. N., Jr, Becker, J. W., Waxdal, M. J. & Wang, J. L. (1972). The covalent and three-dimensional structure of concanavalin A. *Proc. Natl Acad. Sci. USA,* **69**, 2580-2584.
34. Ashford, D. A., Dwek, R. A., Rademacher, T. W., Lis, H. & Sharon, N. (1991). The glycosylation of glycoprotein lectins. Intra and inter-genus variation in *N*-linked oligosaccharide expression. *Carbohyd. Res.* **213**, 215-227.
35. Gohier, A., Espinosa, J. F., Jimenez-Barbero, J., Carrupt, P., Perez, S. & Imberty, A. (1996). Knowledge based modeling of a legume lectin and docking of the carbohydrate ligand: the *Ulex europaeus* lectin I and its interaction with fucose. *J. Mol. Graph.* **14**, 322-327.
36. Yariv, J., Kalb, A. J. & Katchalski, E. (1967). Isolation of an L-fucose binding protein from *Lotus tetragonolobus* seed. *Nature,* **215**, 890-891.
37. Pereira, M. E. A. & Kabat, E. A. (1974). Specificity of purified hemagglutinin (Lectin) from *Lotus tetragonolobus*. *Biochemistry,* **13**, 3184-3192.
38. Konami, Y., Yamamoto, K. & Osawa, T. (1990). The primary structure of the Lotus tetragonolobus seed lectin. *FEBS Letters,* **268**, 281-286.
39. Konami, Y., Yamamoto, K. & Osawa, T. (1992). Purification and characterization of carbohydrate-binding peptides from *Lotus tetragonolobus* and *Ulex europaeus* seed lectins using affinity chromatography. *J. Chromatog.* **597**, 213-219.
40. Young, N. M. & Oomen, R. P. (1992). Analysis of sequence variation among legume lectins. A ring of hypervariable residues forms the perimeter of the carbohydrate-binding site. *J. Mol. Biol.* **228**, 924-934.
41. Sharma, V. & Surolia, A. (1997). Analysis of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J. Mol. Biol.* **267**, 433-445.
42. Collaborative Computing Project No. 4 (CCP4) (1994). A suite of programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.
43. Delbaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Pearlstone, J. R. & Carpenter, M. R., *et al*. (1990). Molecular recognition of a human blood group determinant by a plant lectin. *Can. J. Chem.* **68**, 1116-1121.
44. Strynadka, N. C. J. & James, M. N. G. (1989). Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* **58**, 951-998.
45. Lemieux, R. U., Bock, K., Delbaere, L. T. J., Koto, S. & Rao, V. S. (1980). The conformations of oligosaccharides related to the ABH and Lewis blood group determinants. *Can. J. Chem.* **58**, 631-653.
46. Bourne, Y., Rouge, P. & Cambillau, C. (1990). X-ray Structure of a (α-Man(1 $\rightarrow$ 3)β-Man(1 $\rightarrow$ 4)GlcNAc)-lectin complex at 2.1 Å resolution: the role of water in sugar-lectin interaction. *J. Biol. Chem.* **265**, 18161-18165.
47. Loris, R., Stas, P. P. G. & Wyns, L. (1994). Conserved waters in legume lectin crystal structures: the importance of bound water for the sequence-structure relationship within the legume lectin family. *J. Biol. Chem.* **269**, 26722-26733.
48. Lemieux, R. U. (1996). How water provides the impetus for molecular recognition. *Acc. Chem. Res.* **29**, 373-380.
49. Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1999). Structures of the complexes of peanut lectin with methyl-β-galactose and N-acetyllactosamine and a comparative study of carbohydrate binding in Gal/GalNAc-specific legume lectins. *Acta Crystallog. sect. D,* **55**, 1375-1382.
50. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157-163.
51. Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic *R*-factor refinement by molecular dynamics. *Science,* **235**, 458-460.
52. Brünger, A. T. (1993). *X-PLOR,* version 3.1. a system for X-ray crystallography and NMR, Yale University Press, New Haven and London.
53. Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A,* **42**, 140-149.

54. Read, R. J. (1990). Structure factor probabilities for related structures. *Acta Crystallog. sect. A,* **46**, 900-912.
55. Roussel, A. & Cambillau, C. (1992). *TURBO-FRODO*, Biographics, Marseilles, France.
56. Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature,* **355**, 472-475.
57. Lamzin, V. S. & Wilson, K. S. (1993). Automated refinement of protein models. *Acta Crystallog. sect. D,* **49**, 127-149.
58. Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein-structure refinement. *Acta Crystallog. sect. A,* **47**, 392-400.
59. Sheriff, S., Hendrickson, W. A. & Smith, J. L. (1987). Structure of myohemerethrin in the azidomet state at 1.7/1.3 Å resolution. *J. Mol. Biol.* **197**, 273-296.
60. Sheriff, S. (1993). Some methods for examining the interactions between two molecules. *Immunomethods,* **3**, 191-196.
61. Hindsgaul, O., Khare, D. P., Bach, M. & Lemieux, R. U. (1985). Molecular recognition. III. The binding of the H-type 2 human blood group determinant to the lectin I of *Ulex europaeus*. *Can. J. Chem.* **63**, 2653-2658.
62. Evans, S. V. (1993). SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134-138.
63. Cowtan, K. D. (1994). DM: an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography,* **31**, 34-38.

Edited by I. Wilson

(Received 18 August 2000; received in revised form 2 October 2000; accepted 8 October 2000)