doi:10.1006/jmbi.2001.4806 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **310**, 875–884

JMB

The 1.2 Å Resolution Structure of the Con A-Dimannose Complex

David A. R. Sanders[1], Davina N. Moothoo[1], James Raftery[2] Andrew J. Howard[3], John R. Helliwell[2] and James H. Naismith[1]*

[1]*Biomolecular Sciences, The University, St. Andrews Scotland, KY16 9ST, UK*

[2]*Department of Chemistry University of Manchester Oxford Rd., Manchester M13 9PL, UK*

[3]*IMCA, APS, Argonne National Laboratory, IL USA*

**Corresponding author*

The complex between concanavalin A (Con A) and α1-2 mannobiose (mannose α1-2 mannose) has been refined to 1.2 Å resolution. This is the highest resolution structure reported for any sugar-lectin complex. As the native structure of Con A to 0.94 Å resolution is already in the database, this gives us a unique opportunity to examine sugar-protein binding at high resolution. These data have allowed us to model a number of hydrogen atoms involved in the binding of the sugar to Con A, using the difference density map to place the hydrogen atoms. This map reveals the presence of the protonated form of Asp208 involved in binding. Asp208 is not protonated in the 0.94 Å native structure. Our results clearly show that this residue is protonated and hydrogen bonds to the sugar. The structure accounts for the higher affinity of the α1-2 linked sugar when compared to other disaccharides. This structure identifies different interactions to those predicted by previous modelling studies. We believe that the additional data presented here will enable significant improvements to be made to the sugar-protein modelling algorithms.

Keywords: carbohydrate conformation; Con A-saccharide complex; crystal structure; molecular recognition; thermodynamics

Introduction

The interactions between proteins and oligosaccharides play important roles in a number of biological processes, including cell differentiation, the immune response, neuronal development and infection. Understanding the specific ways in which proteins recognize, and differentiate between their carbohydrate ligands is critical to the understanding of how these systems function and to rationally design biologically active saccharide analogues.

Plant lectins are a group of highly homologous proteins, and although their function is unknown, their interaction with saccharides has proved to be a valuable source of fundamental information. These proteins bind monosaccharides with a relatively low affinity ($K_a \sim 1 \times 10^3$ M^{-1}), but bind oligosaccharides much more tightly ($K_a \sim 1 \times 10^6$ M^{-1}). The interest in these lectins stems from their high degree of specificity for complex oligosaccharides,[1] allowing them to serve as useful model systems for therapeutically relevant systems.

Concanavalin A (Con A) from jack bean (*Canavalia ensiformis*) is the most extensively studied member of the lectin family, with a wealth of structural and thermodynamic data already published.[1–10] There are also several reports of modelling studies on the Con A-oligosaccharide interaction.[11–13] Extensive thermodynamic studies have delineated the poly- and monosaccharide specificities for Con A.[5] A trisaccharide (α-Man-(1-3)-[α-Man-(1-6)]-Man) is central to the binding specificity for Con A.[4] A pentasaccharide (β-GlcNAc-(1-2)-α-Man-(1-3)-[β-GlcNAc-(1-2)-α-Man-(1-6)]-Man) appears to be the largest sugar specifically recognized. The structures of both the tri- and penta-complexes of Con A have been determined,[4,6] these show that the α1-6Man sits at the so-called monosaccharide binding site. The other two mannose residues span an extended trisaccharide ''core'' binding site. The central sugar sits in the reducing sugar site while the α1-3-linked mannose sits in the α1-3Man site.[4,6] The studies highlighted the role of a water molecule (denoted the ''structural water'') in anchoring the reducing sugar to the protein.[4]

Abbreviations used: Con A, concanavalin A; Man-(α1-2)Man, α1-2 mannobiose; r.m.s., root-mean-square.

E-mail address of the corresponding author: naismith@st-andrews.ac.uk

0022-2836/01/040875–10 $35.00/0

The tri- (and penta-) saccharide can be visually decomposed into disaccharide units. Thermodynamic studies show, however, that all of the possible disaccharides bind with approximately the same affinity as the Man-αOMe.[5] The crystal structure of the Man-(α1-6)Man(α1-O)Me[2] shows that as expected, the (methylated) reducing sugar binds in the same manner as the reducing sugar in the trisaccharide complex. However, as pointed out for the trisaccharide complex, the reducing sugar makes few direct interactions with the protein and pays a high entropy penalty in freezing the α1-6 linkage. Interestingly, Man-(α1-3)Man(α1-O)Me binds its non-reducing sugar in the monosaccharide site (as does Man-(α1-6)Man(α1-O)Me),[2] not in the α1-3Man site as would be predicted from the trisaccharide structure, confirming that it is the monosaccharide site which dominates binding energetics. In the survey of disaccharides, the α1-2 linked sugar stands out as having higher affinity (three to five times) than any other disaccharide.[5] In particular, Man-(α1-2)Man(α1-O)Me is particularly well recognised.

In the Man-(α1-2)Man(α1-O)Me-Con A complex;[7] the disaccharide adopts two different modes for binding to the protein, as predicted by Williams *et al*.[14,15] The two modes differ with respect to which sugar moiety binds to the monosaccharide binding site, one with the non-reducing sugar in the monosaccharide binding site and one with the reducing sugar in the monosaccharide binding site.[7]

A listing of this increasing wealth of knowledge on lectin-carbohydrate interactions is available at the Lectin Database (Bettler, Loris & Imberty, http://www.cermav.cnrs.fr/databank/lectin/).

The importance of high resolution in determining protein carbohydrate complexes is evident, the requirement for precise distance and spatial arrangements of atoms is important if accurate theoretical models are to be developed. Lectins are in a powerful position to lend themselves to parameterization of protein-carbohydrate interactions. As pointed out in a recent review,[16] the galactose and mannose/glucose binding proteins are the only group which have sufficient structural information available to begin to identify common recognition features. Identifying changes in protein conformation brought about by carbohydrate recognition is vital if proper assessments of the forces governing recognition are to be quantified. For lectins in general and Con A in particular this has been complicated as the complexes are co-crystals and the binding sites are almost always at points of crystal contact. This has made deconvoluting the changes wrought by binding and those as a result of crystal contacts problematic. In addition, subtle changes require accurate structures to be properly identified, thus both native and complex must be at high resolution.

This paper describes the structure of Con A complexed with α1-2 mannobiose (Man-(α1-2)Man) to 1.2 Å resolution. This is one of only two protein-sugar complexes determined to this high resolution[17] and the only lectin-carbohydrate complex. Fortuitously, the complex crystallizes in the same space group, with a very similar crystal packing arrangement and similar cell dimensions to the 0.94 Å native structure.[18] This has allowed us to examine changes in the protein structure and water structure surrounding the binding site in detail. The structure of the complex reveals how the protein utilizes the same residues to specifically recognize the α1-2 linkage as it uses to recognize the α1-6 linkage. Using electron density, we have, for the first time experimentally, identified many of the hydrogen atoms involved in forming hydrogen bonds between the lectin and the carbohydrate in an attempt to add this information to modelling parameters. The structure explains the higher affinity of this disaccharide.

Results and Discussion

Overall structure

The protein fold, with or without saccharide bound, remains the same.[3,8,19,20] The Man-(α1-2)Man-Con A complex crystallizes in the same space group as the native structure, and has dimensions $a = 90.9$ Å, $b = 86.4$ Å, $c = 65.4$ Å; (cf. native: $a = 89.6$ Å, $b = 86.5$ Å, $c = 62.1$ Å). The regions of poor density (the N-terminal residue, and the surface loops at residues 118-122, 149-151, 160-163) are consistent with the disordered regions in previous Con A structures.[2–4,6–9,18–22] Eleven residues (Ile27, His51, Arg60, Val64, Ser72, Ser96, Ser113, Ser117, Met129 Leu198 and Arg228) were modelled in two conformations. In addition, the side-chain Lys135 was stereochemically modelled and set to zero occupancy during refinement, as it was not located in the electron density. The disaccharide molecule was located in well-defined density prior to inclusion in the refinement. A total of 259 water molecules were located, including the ''structural'' water first described by Naismith & Field,[4] and later seen in the Man-(α1-2)Man(α1-O)Me, Man-(α1-3)Man(α1-O)Me, and Man-(α1-6)Man(α1-O)Me structures.[2,7] The peptide bond between Ala207 and Asp208 is in the unusual *cis* conformation in agreement with the previous Con A structures. In the final model, 86.8 % of residues are in the most favoured regions and the remaining 13.2 % in additionally allowed regions.

Effect of sugar binding: water structure

When making a comparison of the water structure between the Con A structures, it should be noted that the number of water molecules observed in each structure is highly dependent on temperature and the resolution of the structure. Thus, 319 water molecules were included in the final model of the 0.94 Å native cryostructure,[18] 170 more than were included in the earlier 1.6 Å native room temperature structure.[21]

The disaccharide molecule displaces a total of six water molecules from the binding site. One of these (Wat314 from pdb file 1NLS) is displaced by the 1.4 Å movement of Wat279 into the position occupied by the ''structural water''. Three of the remaining five displaced water molecules (252, 269 and 264, respectively) correspond with the positions of O4, O5 and O6 of the non-reducing sugar and were observed in the lower resolution studies. The fifth displaced water molecule (Wat375) sits close to the position of C6 of the non-reducing sugar. The sixth displaced water (Wat348) sits 1.5 Å from the position of C1 of the reducing sugar. An additional water molecule is partially occupied in the active site depending on the movement of the Arg228 side-chain. This additional water molecule makes a hydrogen bond with the ''structural water'' and Arg228 in one conformation, linking the reducing sugar to the protein through this network. When Arg228 is in its second conformation, a hydrogen bond is formed between Arg228 and the sugar directly, displacing the water.

Binding of α1-2 mannobiose

The monosaccharide-binding site is formed by residues 12-14, 98-100, 207-208 and 227-228 as described, previously.[3,8] The non-reducing mannose of the Man-(α1-2)Man complex binds in the monosaccharide site making a total of six hydrogen bonds (Figure 1), 44 van der Waals interactions and ten polar contacts with the protein, the alternate conformation of Arg228 increases the van der Waals interactions to 45 and the polar contacts to 12, with additional hydrogen bonds to water molecules (from O2) and to a symmetry related molecule (from O3).

The reducing sugar sits in an extended binding site formed by Tyr12, Asp16 and Arg228, first identified from the trisaccharide complex[4] (Figure 1). This sugar residue makes two hydrogen bonds (from O6), four van der Waals (including two from the alternate Arg228 conformation), and four polar contacts (including two from Arg228). The binding site is surrounded by crystal contacts, the closest being 2.65 Å from O1 of the reducing

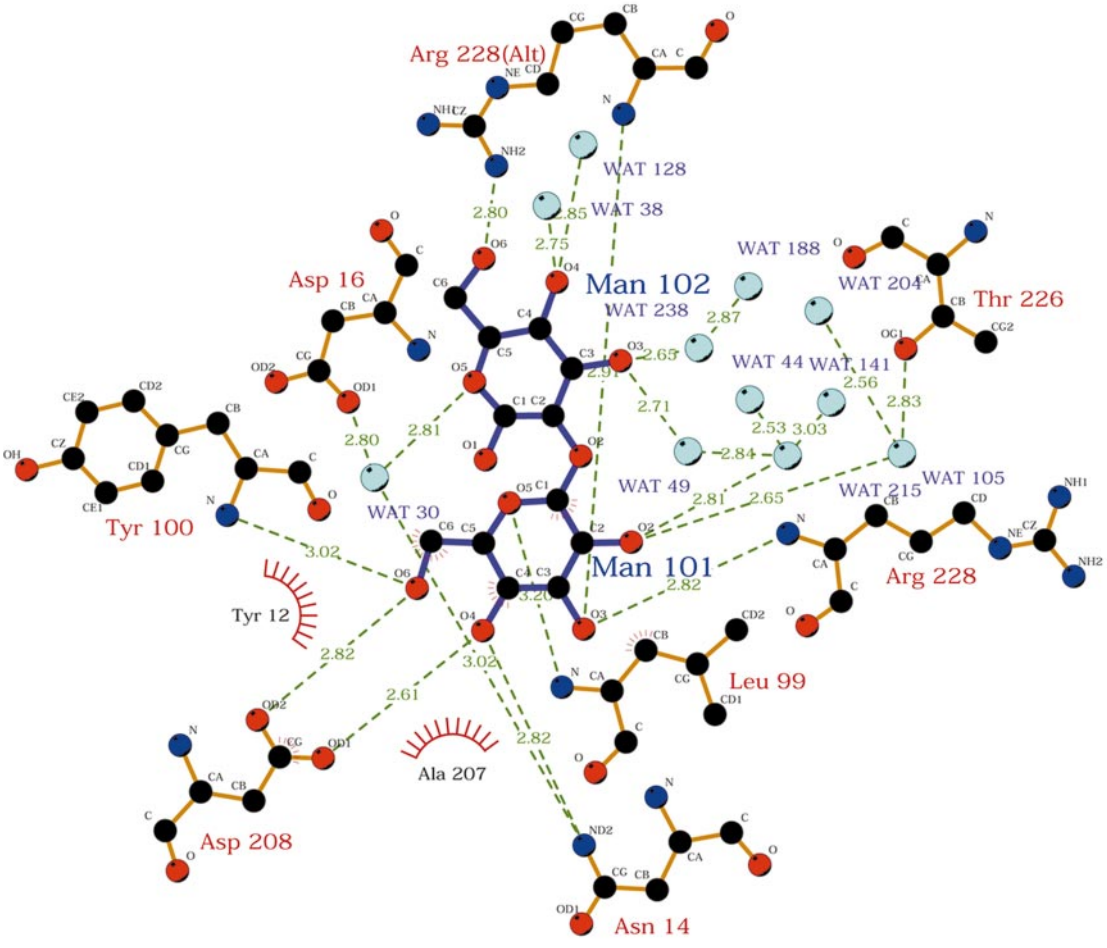

Figure 1. Schematic representation of binding of Man-(α1-2)Man to Con A. Structural water[4] labelled as Wat30. The two conformations for Arg228 (see text and Figure 2) are displayed, with the second conformation designated (Alt). Figure generated using LIGPLOT.[36]

sugar to Ser184. The loop at Ser185 from a symmetry related molecule occupies the a1-3Man site.

The reducing mannose makes hydrogen bonds with Asp16 and a number of water molecules, including the ''structural water'' (Figure 1). However, the Asp228 residue that binds to this water in the other dimannose structures is apparently in two conformations in the Man-(α1-2)Man structure (Figure 2). One conformation has the Arg228 side-chain in a position where the side-chain makes hydrogen bonds to the O6 of the reducing sugar and to Asp16, rather than the ''structural water''. The second conformation has Arg228 shifted away from the ligand-binding site, making a hydrogen bond to a symmetry related molecule and an additional water molecule is recruited into ligand binding. Additional hydrogen bonds are made from O4 and O1 of the reducing sugar to symmetry related molecules.

Water molecules bind to the dimannose through O2 of the non-reducing sugar and O3, O4 and O5 of the reducing sugar (Figure 1). Two of these water molecules play an important role in ligand binding; the water molecule bound to O5 of the reducing sugar is the ''structural water'' and makes hydrogen bonds to Asp16 and Asn14, whereas one of the two water molecules bonded to O2 of the non-reducing sugar makes a further hydrogen bond to Thr226. The remaining water molecules bind either to other water molecules or to symmetry related protein molecules.

Formation of the α1-2 mannobiose-Con A complex buries 148 Å^2 of protein surface area (an additional 47 Å^2 of protein surface buried if Arg228 adopts its second conformation) and 297 Å^2 (315 Å^2 for second Arg228 conformation) of sugar surface area, composed of 207(215) Å^2 from the non-reducing sugar and 90(100) Å^2 of the reducing sugar. The ϕ, ψ glycosidic linkage conformation angles (Table 1) of the disaccharide in this structure are close to those observed in the isolated crystal structure,[22] the solution structure of Man-(α1-2)Man(α1-O)Me[23] and as calculated by the relaxed residue approach with MM3 software.[24] The conformation lies in an energy minimum.[24,25] These values are also very similar to values reported for the Man-(α1-2)Man(α1-O)Me-Con A crystal structure[7] (Table 1). The location of the non-reducing sugar in the principal binding site mirrors the arrangement found in only one subunit of the Man-(α1-2)Man(α1-O)Me complex (Figure 3). As noted in the Introduction, the Man-(α1-2)Man(α1-

Table 1. Dihedral angles around the inter-sugar glycosidic linkage of Man-(α1-2)Man and Man-(α1-2)Man (α1-O)Me

	ϕ[a]	ψ[a]
Man-(α1-2)Man complex	72	−150
Man-(α1-2)Man isolated crystal[22]	65	−106
Man-(α1-2)Man relaxed residue model[24]	79	−145
Man-(α1-2)Man protein-ligand model[13]	143	−93
Man-(α1-2)Man(α1-O)Me-NMR (major conformation)[23]	60	−150
Man-(α1-2)Man(α1-O)Me:D subunit[7]	70	−147

[a] ϕ is O-5:C-1:O-X:C-X, ψ is C-1:O-X:C-X:C-(X-1) (IUPAC convention).

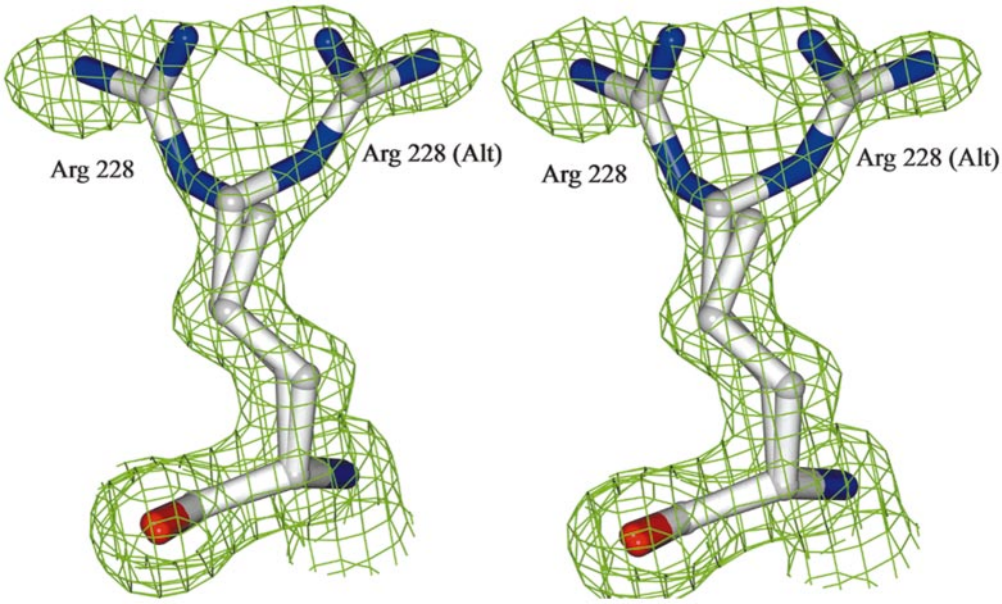

Figure 2. Stereo view of final $2F_o - F_c$ map showing the alternate conformations adopted by Arg228. The density was modelled at 2 r.m.s. Oxygen atoms are coloured red, nitrogen atoms blue, sulphur atoms yellow and carbon atoms grey. All molecular representations generated using BOBSCRIPT[37] utilizing the GL_RENDER interface (L. Esser & J. Deisenhofer, unpublished data) and POV-Ray™.

Table 2. Hydrogen bonds made by the second sugar in dimannose structures

Structure	Sugar hydroxyl	Protein residue	Distance
Man-(α1-2)Man(α1-O)Me (A)[a]	O6	Tyr-12 OH	3.0
Man-(α1-2)Man(α1-O)Me (D)[a]	O3	Thr-226 OG1	3.3
	O4	Ser-168 OG	3.3
Man-(α1-3)Man(α1-O)Me[b]	None		
Man-(α1-6)Man(α1-O)Me[b]	O4	Tyr-12 OH	2.7

[a] The second sugar is found in two conformations; from Moothoo *et al.*[7]
[b] From Bouckaert *et al.*[2]

O)Me is unique in that two binding modes are clearly observed.[7]

Structural basis of affinity for the 1-2 linked disaccharide

The thermodynamics of Con A binding to dimannose sugars has been reported.[5] The K_a value observed for the association between Con A-Man-(α1-2)Man displays a fivefold increase over MeMan and a threefold increase over Man-(α1-3)Man and Man-(α1-6)Man disaccharides. Man-(α1-2)Man(α1-O)Me binds with a 3.4-fold higher affinity than its non-methylated counterpart.

The Man-(α1-2)Man buries a larger accessible surface area (297 or 315 Å^2 depending on Arg228 conformation) than the Man-(α1-6)Man(α1-O)Me (290 Å^2) or Man-(α1-3)Man(α1-O)Me (276 Å^2) and makes a different hydrogen-bonding network than the other dimannose sugars (Figure 1 and Table 2). The additional hydrogen bonds compared to the Man-(α1-3)Man(α1-O)Me and Man-(α1-6)Man(α1-O)Me as well as the greater buried surface area are consistent with Man-(α1-2)Man binding with higher affinity than either the Man-(α1-3)Man(α1-O)Me or Man-(α1-6)Man(α1-O)Me dimannose.[5]

The 3.4-fold greater affinity displayed by Man-(α1-2)Man(α1-O)Me may be partially accounted for by the greater accessible surface area buried by the Man-(α1-2)Man(α1-O)Me complex compared with the Man-(α1-2)Man complex. However, methylation of the reducing sugar does not cause such an increase for any other disaccharide. This suggests that in the Man-(α1-2)Man(α1-O)Me complex an additional binding contribution arises from the existence of two mutually exclusive binding modes.[7] We see no evidence of disorder in the α1-2 complex reported here; however, crystallization is a kinetic process and an equilibrium between the two binding modes (seen in the Man-(α1-2)Man(α1-O)Me structure[7]) may exist in solution. The absence of two binding modes provides a simple explanation for the higher affinity of the Man-(α1-2)Man(α1-O)Me complex compared to Man-(α1-2)Man.

Hydrogen bonding patterns

Hydrogen atoms involved in the sugar-Con A interaction were modelled into the structure by removing them from the anisotropic refinement calculations and calculating a $F_o - F_c$ map. The hydrogen atoms were removed from all of the hydroxyl hydrogen atoms of the dimannose, and the appropriate hydrogen atoms from residues Tyr12, Asn14, Leu99, Tyr100, Asp208, Thr226 and Arg228, a number of which (Tyr OH groups, Thr OH, Asn side-chain NH, Arg side-chain NHs),

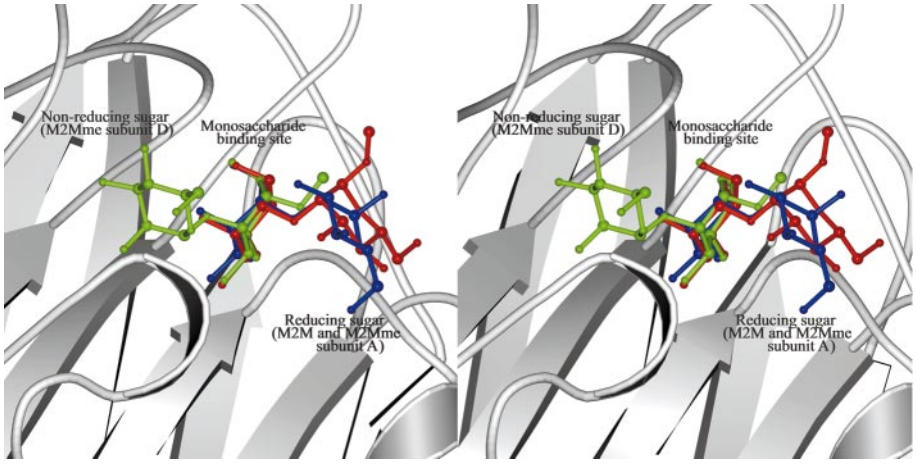

Figure 3. Stereo-view of superimposed disaccharides from Man-(α1-2)Man(α1-O)Me[7] (labelled as M2Mme), both subunit A (red) and subunit D (green) and the disaccharide from the Man-(α1-2)Man structure (blue) (labelled as M2 M), showing the different orientation of the binding of the three ligands.

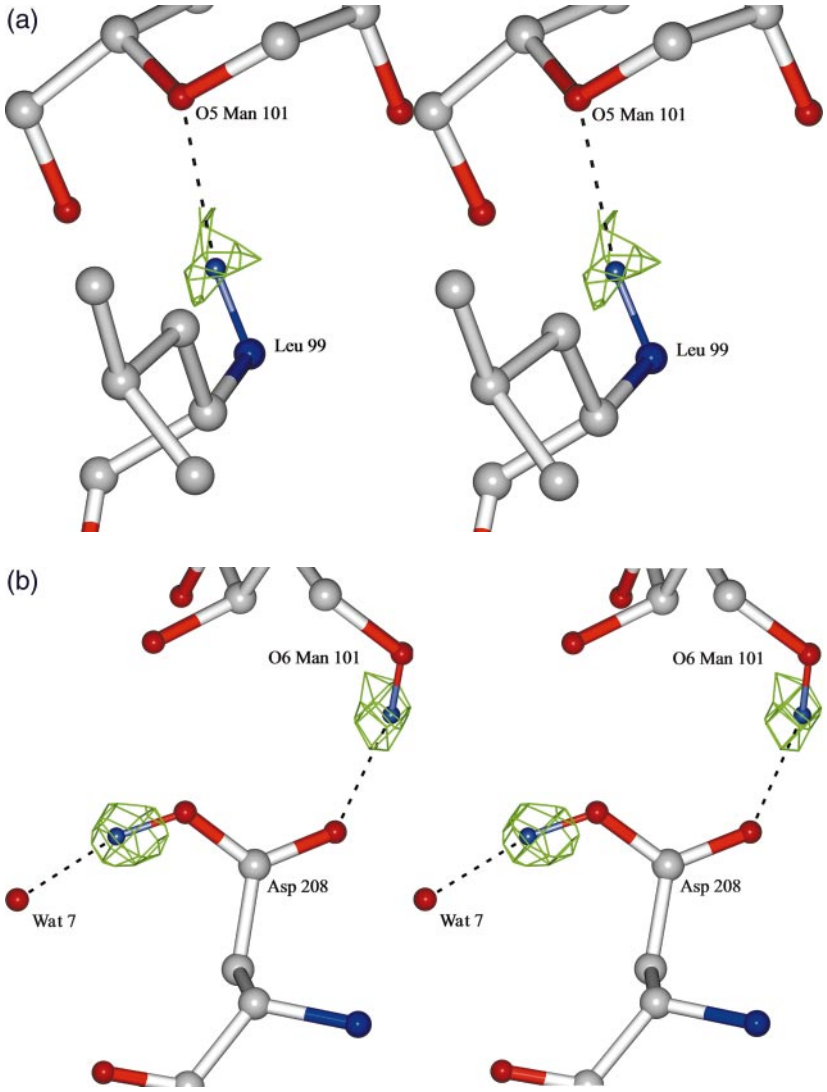

Figure 4. Difference electron density maps for various residues showing the likely positions of the hydrogen atoms. (a) Density at Leu99, showing density for hydrogen bonds between protein and sugar. (b) Density at Asp208 suggesting that the residue is in the protonated form. Hydrogen atoms are coloured purple, all other atoms coloured as for Figure 2.

were not included in the anisotropic refinement protocol. The hydrogen atoms were modelled into electron density (Figure 4(a) and (b)) and Table 3 summarizes the angles and distances of the hydrogen atoms to their bonded partner and to their putative hydrogen bond partner. The expected angles for hydrogen bonds (C(donor)-donor-H) for side-chains and main-chains have been taken from Baker *et al.*[26]). The hydroxyl groups of the sugar groups have been assumed to have a similar angle to those found for Ser-OH groups (118(±16)°[26]). All of the donor-H angles in Table 3 satisfy these criteria. The optimum angle between donor-H-acceptor is 180°. Deviations from the expected donor-H-acceptor angle of 180° are due to the constraints imposed by the arrangement of atoms within a structure.

The hydrogen atoms from Man101 O4, Man102 O3, Tyr12 and Asn14 were not located. Those from Tyr12 and Asn14 were positioned based on the requirements of chemical knowledge. The hydrogen atoms for the water molecules were not included at any stage in the refinement procedure. An extensive water-bridged hydrogen bond network was modelled from the difference map, guided by chemical knowledge (Figure 5).

In the case of O6 of Man102, the position of the hydrogen atom located from the difference density map is such that the hydroxyl hydrogen is bound to OD1 of a symmetry related Asn118 side-chain. Of course this interaction would not occur in solution. It is most likely that Asp16 makes this bond in solution. One conformation of Arg228 is in a position to make a hydrogen bond with the O6 atom; however, the Arg228 proton is not located in density. This may be the result of crystallographic disorder of this side-chain. Resolving the bonding pattern at O6 highlights the complexity imposed by crystal packing.

Thr226 hydrogen binds a water molecule that hydrogen bonds to the dimannose (Figure 1). Inspection of the difference density around the Thr226 does not allow placement of the hydrogen from the Thr226 oxygen.

Table 3. Summary of hydrogen placements

Donor	Acceptor	Distances[a]		Angles[a]	
		Donor:H	Accept:H	Donor-H[b]	Don-H-Acc[c]
O2 101	Wat105	1.04	1.70	110.3	150.0
Wat215	O2 101	2.79		99.6	
O3 101	SYM 71 Asp OD1	1.02	2.16	110.7	165.7
228 Arg N	O3 101	2.92/2.91		124.9	
O4 101	208 Asp OD1	2.61		112	
14 Asn ND2	O4 101	2.82		129.4	
99 Lys N	O5 101	1.12	2.09	99.6	169.7
O6 101	208 Asp OD2	1.00	1.87	122.7	158.7
100 Tyr N	O6 101	3.01		109.6	
O1 102	SYM 184 Ser O	2.65		109.8	
O3 102	Wat238	2.65		106.9	
Wat49	O3 102	2.71		121.3	
O4 102	Wat38	1.17	1.57	126.8	179.6
Wat128	O4 102	2.83		110.5	
O5 102	Wat30 (Structural)	2.81		130.0	
O6 102[d]	SYM Asn 118 OD1	1.31	2.05	93.9	176.5
	16 Asp OD2	2.76		109.7	
	228 Arg (2nd) NH2	2.80		146.2	
208 Asp OD1[e]	Wat7	1.02	1.75	123.7	159.8

[a] For hydrogen bonds where the hydrogen can not be placed in density, the distance refers to the distance between the acceptor and donator (sum of Donor:H and Acceptor:H distances), while the angle refers to the angle from C-x(sugar)-O-x-(O/N) (similar to Donor-H angle).
[b] Donor-H angle is defined as C(donor)-Donor-H angle (expected value = 118(±16)° for C-O-H (sugar), 121(±18)° for Asn, 125(±14)° for Asp,146(±19)° for Arg and 124(±16)° for Thr[26]).
[c] Don-H-Acc angle is defined as the angle between the donor and acceptor through the hydrogen.
[d] Hydrogen bonding pattern for O6 102 is unable to be resolved for Asp16 and Arg228.
[e] Discussed in text.

Asp208 makes hydrogen bonds to the sugar from both of the oxygens of the side-chain (to O4 101 and O6 101; Figure 1). The position of the hydrogen from O6 101 was located based on the difference map, while there is no density found for the O4 101 hydrogen atom. Density at Asp208 confirms that this residue is protonated. The protonated hydroxyl of Asp208 is bonded to Wat7, which bonds to the Ca^{2+} (Figure 4(b)).

Whilst this is an unexpected result, it is not unknown for Asp residues to remain protonated at elevated pH values (between pH 5 and 7 in this case). This protonated form of the Asp residue may be aided by more severe electron displacement from the water molecule caused by binding to the Ca^{2+} . Also, higher than usual p*K*a values for Asp and Glu have been correlated with reduced solvent exposure in a number of structures and the binding of Man-(α1-2)Man to Con A results in a 10 $Å^2$ less solvent accessible surface for Asp208 when compared to the native structure. In the 0.94 Å native map Con A[18] Asp208 is unprotonated, although the C-O bond lengths were noted to be unequal (1.25 and 1.30 Å) and the possibility

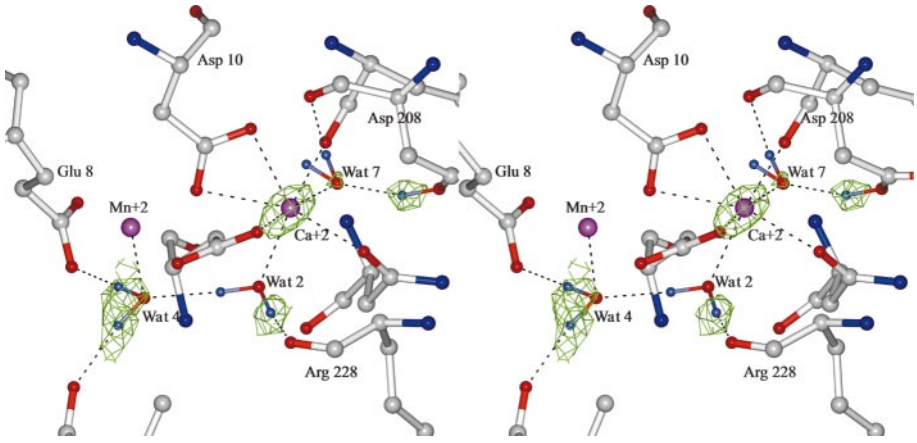

Figure 5. Difference map showing an example of water-mediated hydrogen bonding network between the protein and bound metals. One hydrogen atom from Wat7 could not be assigned a bonding partner. Metals are coloured pink, hydrogen atoms are coloured purple, all other atoms are colour coded as described for Figure 2.

of an equilibrium mixture between COO^- and COOH) was considered. Interestingly, in a 0.92 Å Con A study where Con A was demetallized, then re-populated with pure Mn^{2+} and Ca^{2+} , Asp208 is apparently protonated, suggesting the pK_a of this residue is indeed high (H.J. Price, J. Raftery & J.R. Helliwell, unpublished results).

Comparison of actual *versus* modelled complex structure

Prediction of the modes of binding of α1-2 mannobiose to Con A using energy minimization methods[13] docked the disaccharide to the protein with the non-reducing sugar in the monosaccharide binding site with a more favourable energy than the reducing sugar. In the model, the sugar conformation is very different to the crystal structure[13] and the reducing sugar makes a hydrogen bond as well as several van der Waals interactions with the protein (see Reddy *et al*.[13]). Our structure confirms that it is the non-reducing sugar which is recognized by the monosaccharide site and that the reducing sugar makes further interactions with the protein. The carbohydrate-protein hydrogen bonding network, and the conformation of the disaccharide, are very different between the model study and the crystallographic structure; most notably in the shift in binding of O5(Man101) from Tyr100 (model)[13] to Lys99 (crystal) (Figure 1). The binding of the reducing sugar in the minimization study is completely different to that of the crystal structure (Table 1). This reflects the insufficient parameters available for protein-saccharide modelling. The ''structural water'' was not identified in the modelling study. In peanut lectin bridging structural water molecules to the disaccharide have also been reported by Ravishankar *et al*.,[27] confirming the importance of such water molecules. In addition, the native Con A structure was used in the modelling studies and changes in the protein structure upon carbohydrate binding were not taken into account.

Conclusions

The majority of the hydrogen atoms modelled from the difference density map appear in locations close to the expected positions of the hydrogen atoms. The protonation of OD1 of Asp208 is the notable exception, highlighting the importance of high resolution data for proper determination of hydrogen bonding. As the other hydrogen atoms were modelled into realistic positions, we believe that the hydrogen of Asp208 is properly located by the difference electron density map, however, unexpected this position appears. This case is an excellent candidate for further study by neutron protein crystallography studies. This is an area undergoing substantial improvements in technology.[28]

The additional information, such as accurate geometry and proton location, that can be gained from examining high resolution protein-carbohydrate complexes will prove critical to improving the parameters used for modelling of sugar-protein interactions.

Materials and Methods

Crystallization of Con A-α1-2 mannobiose complex

Co-crystals were obtained from 10 µl of a solution containing 18 mM α1-2 mannobiose (Dextra Laboratories, Reading, UK), 0.6 mM jack bean concanavalin A (Sigma, Poole, UK), 1 mM $CaCl_2$, 1 mM $MnCl_2$, 20 mM Tris (pH 7.0) and 0.1 M NaCl, mixed with 10 µl of a reservoir containing 10 % (w/v) PEG 6000, 0.1 M citric acid (pH 5.0) in a sitting drop tray (Charles Supper Co.) and left to equilibrate against 1 ml of the reservoir. Crystals took several weeks to reach optimum size for data collection.

Data collection

A room-temperature data set to 1.8 Å was collected using a Nonius DIP2000 dual image plate. Data were recorded as 141 non-overlapping 25 minute 0.75° oscillations.

Data were indexed and merged using DENZO and SCALEPACK.[29] Data were indexed in the lattice group *I*222, $a = 91.7$ Å, $b = 86.8$ Å, $c = 66.6$ Å, with a monomer in the asymmetric unit.

For the higher resolution data sets, data were collected from two frozen crystals on the IMCA beamline at the APS. Small crystals (<0.2 mm) were chosen to avoid spot splitting. Data were collected on a Bruker 2 × 2 CCD detector (1165 images). The first crystal was subjected to a fast pass (to collect low resolution data) and a slow pass (for high resolution). The second crystal was subjected only to a slow pass for high resolution data. The detector was offset for both of the high resolution passes. Data were indexed in the lattice group *I*222, with cell dimensions $a = 90.9$ Å, $b = 86.4$ Å, and $c = 65.4$ Å, with a monomer in the asymmetric unit.

In the merge of the three passes, it was evident that there was a systematic error for common reflections between the tilted and non-tilted detector settings. We suspect this was a residual detector non-uniformity effect or a more subtle non-linearity of response effect. Hence each data pass was checked *via* rigid body and *B*-factor refinement starting from the room temperature 1.8 Å protein model; in each case the coordinates for the three cryo models agreed extremely closely but the atomic

Table 4. Data collection statistics for Con A:α1-2 mannobiose to 1.8 Å and 1.2 Å

	1.8 Å data	1.2 Å data
Resolution (Å)	25.0-1.8	8-1.16
Space group	*I*222	*I*222
Unique reflections	27,201	76,521
Completeness (%)	94.9 (80.3)	87.2 (89.0)
R_{merge} (*I*) (%)[a]	6.1 (23.9)	11.2 (45.3[b])

[a] $R_{merge}(I) = \Sigma_{hkl}\Sigma_i |I_i - I(hkl)| / \Sigma_{hkl}\Sigma_i I_i(hkl)$.

[b] R_{merge} taken from the crystal 2 slow pass SCALEPACK merge. All other stats are from the final merge of the three data sets from the two crystals. For details see Materials and Methods.

Table 5. Final refinement statistics for Con A-α1-2 mannobiose complex to 1.2 Å resolution

Number of protein atoms	1896
Number of sugar atoms	23
Number of water molecules	254
Number of metal ions	2
Number of restraints	30,485
Root-mean-square deviations from	
Ideal bond lengths[a]	0.020
Ideal bond angles[a]	1.82
Mean temperature factor	
Main-chain	18.32
Side-chain	20.27
Sugar atoms	25.42
Water molecules	35.30
Crystallographic *R*-factor (%)	16.97
Free crystallographic *R*-factor[b] (%)	19.09

[a] r.m.s. deviation from Engh and Huber ideal values.[35]
[b] R_{free} is calculated on 5% of data excluded during refinement.

B-factors were systematically shifted one with respect to the other. This confirmed that the fall off of the overall intensities was different between the three sets of data, consistent with a systematic error in data collection hardware. The three protein models were averaged and F_{calcs} value estimated with SFALL.[30] The three data sets were scaled against these F_{calcs} values and then merged together whereby all the quick pass resolution ranges were used (infinity to 1.48 Å), the slow pass from crystal one was used for 1.8 to 1.42 Å and the second crystal slow pass for the range 1.8 to 1.16 Å. Final data collection statistics for both the 1.8 Å data and the 1.2 Å datasets are listed in Table 4.

Structure solution

The initial structure was determined using the data set to 1.8 Å. The structure was determined using the molecular replacement method as implemented in the CCP4[30] program AMoRe[31] with data from 12.0 Å to 3.5 Å. One monomer of the methyl α-D-mannopyranoside-Con A complex (PDB code 5CNA) was used as the search model, with all metal ions, sugar molecules and water molecules removed, and with all atoms set to full occupancy.

Refinement

The initial structure to 1.8 Å was refined using CNS.[32] Rigid body refinement was followed by model building using O[33] and the metal ions, the two mannose residues and conserved water molecules were added at this point. Further refinement was carried out by alternating cycles of automated refinement (CNS restrained positional and temperature factor refinement, plus water additions) and manual rebuilding using O.[33] The refinement was deemed to have converged at an R_{free} value of 25.1%. The cryo data set to 1.2 Å resolution used the 1.8 Å model as a starting point and the resolution was gradually increased to 1.2 Å through the refinement, at this stage the refinement had an R_{free} value of 23.8%. Alternate conformations for residues Ile27, His51, Arg60, Val64, Ser72, Ser96, Ser113, Ser117, Met129, Leu198 and Arg228 were added to the model at this point and the occupancies were initially refined by CNS and then manually adjusted to satisfy the electron density. The structure was then refined using the CCP4 program REFMAC5[34] to refine by anisotropic temperature factors. The refinement converged at an R_{free} value of 19.09%. Final refinement values are listed in Table 5. Maps were then generated with the predicted positions of all hydrogen atoms, except those involved in ligand binding, included in the calculations.

Coordinate deposition

The coordinates for the structure described here have been deposited in the RCSB Protein Data Bank under accession code 1I3H.

Acknowledgements

J.H.N. thanks the BBSRC for funding. We thank Eleanor Dodson and Garib Murshudov for advice and discussion on the various Refmac versions used. J.R.H. thanks The Wellcome Trust for support and the APS IMCA consortium for the invitation to utilise beamtime on their undulator beamline; Dr R. Nuttall and Mr P. Faulder from Manchester accompanied J.R.H. on the beamtime run and they are thanked for their work and support.

References

1. Rini, J. M. (1995). Lectin-carbohydrate interactions. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551-577.
2. Bouckaert, J., Hamelryck, T. W., Wyns, L. & Loris, R. (1999). The crystal structures of Man(α1-3)Man(α1-O)Me and Man(α1-6)Man(α1-O)Me in complex with concanavalin A. *J. Biol. Chem.* **274**, 29188-29195.
3. Naismith, J. H., Emmerich, C., Habash, J., Harrop, S. J., Helliwell, J. R. & Hunter, W. N. *et al*. (1994). Refined structure of concanavalin-A complexed with methyl alpha-D-mannopyranoside at 2.0 Ångstrom resolution and comparison with the saccharide-free structure. *Acta Crystallog. sect. D,* **50**, 847-858.
4. Naismith, J. H. & Field, R. A. (1996). Structural basis of trimannoside recognition by concanavalin A. *J. Biol. Chem.* **271**, 972-976.
5. Mandal, D. K., Kishore, N. & Brewer, C. F. (1994). Thermodynamics of lectin-carbohydrate interactions. Titration microcalorimetry measurements of the binding of N-linked carbohydrates and ovalbumin to concanavalin A. *Biochemistry,* **33**, 1149-1156.
6. Moothoo, D. N. & Naismith, J. (1998). Concanavalin A distorts the B-GlcNAc-(1-2)-Man linkage of β-GlcNAc-(1-2)-α-Man-(1-3)-[β-GlcNAc-(1-2)-α-Man-(1-6)]-Man upon binding. *Glycobiol.* **8**, 173-181.
7. Moothoo, D. N., Canan, B., Field, R. A. & Naismith, J. H. (1999). Man α1-2 Man α-OMe-concanavalin A complex reveals a balance of forces involved in carbohydrate recognition. *Glycobiol.* **9**, 539-545.
8. Derewenda, Z., Tariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J. & Papiz, M. Z. *et al*. (1989). The structure of the saccharide-binding site of concanavalin A. *EMBO J.* **8**, 2189-2193.
9. Bradbrook, G. M., Gleichmann, T., Harrop, S. J., Habash, J., Raftery, J., J., K. G. & Yariv, J. *et al*. (1998). X-ray and molecular dynamics studies of concanavalin-A glucoside and mannoside com-

plexes: relating structure to thermodynamics of binding. *J. Chem. Soc. Faraday Trans.* **94**, 1603-1611.
10. Kalb, A. J., Habash, J., Hunter, N. S., Price, H. J., Raftery, J. & Helliwell, J. R. (2000). Manganese(II) in concanavalin A and other lectin proteins. *Met. Ions Biol. Syst.* **37**, 279-304.
11. Imberty, A. & Perez, S. (1994). Molecular modelling of protein-carbohydrate interactions. Understanding the specificities of two legume lectins towards oligosaccharides. *Glycobiol.* **4**, 351-366.
12. Carver, J. P., Mackenzie, A. E. & Hardman, K. D. (1985). Molecular model for the complex between concanavalin A and a biantennary-complex class glycopeptide. *Biopolymers,* **24**, 49-63.
13. Reddy, V. S. & Rao, V. S. R. (1992). Modes of binding of a (1-2) linked manno-oligosaccharides to concanavalin A. *Int. J. Biol. Macromol.* **14**, 185-192.
14. Williams, T. J., Shafer, J. A., Goldstein, I. J. & Adamson, T. (1978). Biphasic association of *p*-nitrophenyl 2-*O*-α-D-mannopyranosyl-α-D-mannopyranoside and concanavalin A as detected by stopped flow spectroscopy. *J. Biol. Chem.* **253**, 8538-8544.
15. Williams, T. J., Homer, L. D., Shafer, J. A., Goldstein, I. J., Garegg, P. J. & Hultberg, H. *et al.* (1981). Characterization of the extended carbohydrate binding site of concanavalin A: specificity for interaction with the nonreducing termini of α-(1-2)-linked disaccharides. *Arch. Biochem. Biophys.* **209**, 555-564.
16. Elgavish, S. & Shaanan, B. (1997). Lectin-carbohydrate interactions: different folds, common recognition principles. *Trends Biol. Sci.* **22**, 462-467.
17. Sulzenbacher, G., Mackenzie, L. F., Wilson, K. S., Withers, S. G., Dupont, C. & Davies, G. J. (1999). The crystal structure of a 2-fluorocellotriosyl complex of the *Streptomyces lividans* endoglucanase CelB2 at 1.2 Å resolution. *Biochemistry,* **38**, 4826-4833.
18. Deacon, A., Gleichmann, T., Kalb, J., Price, H., Raftery, J. & Bradbrook, G., *et al.* (1997). The structure of concanavalin A and its bound solvent structure determined with small molecule accuracy at 0.94 Å resolution. *J. Chem. Soc., Faraday Trans.* **53**, 4305-4312.
19. Hardman, K. D. & Ainsworth, C. F. (1972). Structure of concanavalin A at 2.4 Å resolution. *Biochemistry,* **11**, 4910-4919.
20. Naismith, J. H., Habash, J., Harrop, S., Helliwell, J. R., Hunter, W. N. & Wan, T. C. M., *et al.* (1993). Refined structure of cadmium-substituted concanavalin A at 2.0 Å resolution. *Acta Crystallog. sect. D,* **49**, 561-571.
21. Emmerich, C., Helliwell, J. R., Redshaw, M., Naismith, J. H., Harrop, S. & Raftery, J. *et al.* (1994). High-resolution structures of single-metal-substituted concanavalin A: the Co,Ca-protein at 1.6 Å and the Ni,Ca-protein at 2.0 Å. *Acta Crystallog. sect. D,* **50**, 749-756.
22. Srikrishnan, T., Chowdhary, M. S. & Matta, K. L. (1989). Crystal and molecular structure of methyl *O*-α-D-manno-pyranosyl-(1-2)-α-D-mannopyranoside. *Carbohydr. Res.* **186**, 167-175.
23. Peters, T. (1991). Synthesis and conformational-analysis of methyl 2-o-(alpha-D-mannopyranosyl)-alpha-D-mannopyranoside. *Liebigs Ann. Chem.* **2**, 135-141.
24. Dowd, M. K., French, A. D. & Reilly, P. J. (1995). Molecular mechanics modeling of α(1-2), α(1-3), and α(1-6) linked monnosyl disaccharides with MM3(92). *J. Carbohydr. Chem.* **14**, 589-600.
25. Imberty, A., Gerber, S., Tran, V. & Perez, S. (1990). Data bank of three-dimensional structures of disaccharides, a tool to build 3-D structures of oligosaccharides. Part I. Oligomannose type *N*-glycans. *Glycoconj. J.* **7**, 27-54.
26. Baker, E. N. & Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Progr. Biophys. Mol. Biol.* **44**, 27.
27. Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1999). Structures of the complexes of peanut lectin with methyl-beta-galactose and N-acetyllactosamine and a comparative study of carbohydrate binding in Gal/GalNAc-specific legume lectins. *Acta Crystallog. sect. D,* **55**, 1375-82.
28. Helliwell, J. R. (1997). Neutron Laue diffraction does it faster. *Nature Struct. Biol.* **4**, 874-876.
29. Otwinowski, Z. (1993). Oscilliation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, Daresbury Laboratory, Warrington, UK.
30. CCP4, (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.
31. Nazava, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157-163.
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P. & Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.
33. 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.
34. Murshudov, G. N., Lebedev, A., Vagin, A. A., Wilson, K. S. & Dodson, E. J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallog. sect. D,* **55**, 247-255.
35. Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A,* **47**, 392-400.
36. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127-134.
37. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 132-134.

Edited by I. A. Wilson

(Received 15 February 2001; received in revised form 29 May 2001; accepted 29 May 2001)