The 2.0 Å structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose: evidence for two unique binding modes

Christine Schubert Wright* and Gerko Hester†

Background: *Galanthus nivalis* agglutinin (GNA), a mannose-specific lectin from snowdrop bulbs, is a tetrameric member of the family of *Amaryllidaceae* lectins that exhibit antiviral activity towards HIV. Its subunits are composed of three pseudo-symmetrically related β sheet domains, each with a conserved mannose-binding site. Crystal structures of monosaccharide and disaccharide complexes of GNA have revealed that all 12 binding sites of the tetramer are functional, and that the degree of occupancy is dependent on the availability of subsidiary interactions from neighboring subunits. The complex of GNA with a branched mannopentaose ((Manα1,6-(α1,3-Man)Man-α1,6-(α1,3-Man)Man) described here simulates a more biologically relevant complex.

Results: Two unique mannopentaose binding modes co-exist in the tetragonal structure (1 subunit/asymmetric unit) of the complex. In one, the conserved monosaccharide-binding pocket in domain 1 (CRD1) is utilized for cross-linkage of twofold related GNA dimers by the outer 3,6 tri-Man arm, which alternates between two orientations consistent with crystal symmetry. Inter-linked dimers assemble helically along the 4_1 crystal axis forming a pore-like structure. In the second binding mode, the complete 3,6 tri-Man arm binds to an extended binding region in domain 3 (CRD3) with subsites for each terminal Man and the internal Man positioned in the conserved mono-saccharide pocket. The two remaining mannose residues are not visible in either binding mode.

Conclusions: This structure provides insights into possible mechanisms of the cross-linkage that is known to occur when lectins interact with specific multivalent cell surface receptors during events such as agglutination and mitogenic stimulation. By virtue of the large number of sites available for mannose binding, GNA has multiple possibilities of forming unique lattice structures. The two distinctly different binding modes observed in this study confirm that high affinity mannose binding occurs only at the two domain sites located near dimer interfaces.

Address: Departments of Biochemistry and Molecular Biophysics, and Medicinal Chemistry, Virginia Commonwealth University, Richmond, Virginia 23298, USA.

†Present address: Department of Structural Biology, Biocentre of the University of Basel, CH-4056 Basel, Switzerland.

*Corresponding author.
E-mail: cswright@gems.vcu.edu

Key words: *Amaryllidaceae*, mannose binding, snowdrop lectin, sugar-mediated cross-linkage

Received: 31 July 1996
Revisions requested: 28 August 1996
Revisions received: 23 September 1996
Accepted: 26 September 1996

Structure 15 November 1996, 4:1339–1352

Introduction

Lectins are oligomeric proteins with carbohydrate recognition functions that are widespread in plants and vertebrates. They play important biological roles in receptor-mediated cellular processes and inter-cellular recognition [1–4]. Although the biological functions of plant lectins have remained poorly understood [5], numerous well defined functions have been identified for animal and bacterial lectins [2,3,6]. Interesting biological effects induced by lectin binding to animal cells have been widely documented, and are thought to involve multi-site interactions that lead to cluster formation or cross-linkage [4,7].

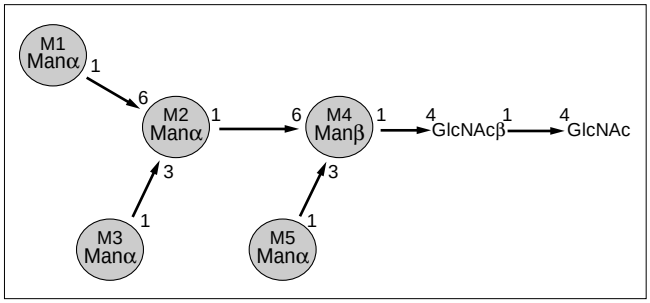
Galanthus nivalis (snowdrop) lectin (GNA), a tetramer of identical subunits (A,B,C,D, each of 12.5 kDa), belongs to a recently characterized non-seed family of mannose-specific lectins (from *Amaryllidaceae*), which constitutes a third structurally unique major plant lectin family [8]. Evidence suggests that GNA plays a protective role against plant predators [9]. As a result of its absolute specificity for mannose, GNA has been used as a tool for glycoprotein isolation and purification [10,11]. Of particular interest is its effective inhibition of HIV infectivity (50% inhibition at 0.04–0.15 $\mu g\,ml^{-1}$) through its specific binding to high-mannose glycans on the HIV envelope glycoprotein GP120 [12], and its ability to precipitate highly branched mannans (from *Saccharomyces cerevisiae*) with α1,3 linkages [13]. A series of hapten inhibition studies, using different mannose containing oligosaccharides, have shown a preference of GNA for α1,3-linked terminal mannosides [10,14]. These data were corroborated by binding studies of Chervenak and Toone [15], who reported association constants in the

range of 1.2–3.3 mM, which are, however, 100 times lower than the association constants observed for concanavalin A (ConA).

The recent elucidation of the crystal structure of GNA [16] revealed an interesting case of diversification of function within a single polypeptide that differs in structure from lectins of other well characterized families [17,18]. The polypeptide chain exhibits local threefold symmetry generated by three antiparallel four-stranded β sheets. These independently folded β sheets are termed carbohydrate recognition domains (CRD1, CRD2, CRD3), as they share strong sequence homology [19] and a conserved mannose-binding site [16]. They are structurally unique compared with domain structures of other mannose binding-proteins [17,20]. The tetramer, displaying *222* symmetry, consists of two tetramer consists of two tightly associated dimers. These will be referred to as A–D and B–C dimers throughout the text in keeping with the terminology established for the monosaccharide crystal complex (tetramer/asymmetric unit) [16]. These dimers are stabilized through C-terminal strand exchange and β sheet association. The structural determinants for monomannoside and dimannoside binding have been deduced from crystal structures in which methyl-α-D-mannose (MeMan) and Man-α1,3-D-Man-OMe (Man-2) were complexed with GNA at 2.3 Å and 3.0 Å resolution, respectively [21]. In the monosaccharide complex, binding was observed in each independent subunit at all three of the conserved monosaccharide binding sites (12 sites/tetramer), although occupancies varied depending on secondary interactions coming either from intra- or inter-tetramer subunit interactions. The dimannoside complex revealed that α1,3-linked dimannosides bind strongly only at the binding site in the CRD3 domain, where a secondary binding region is present for the non-reducing mannose residue, while the reducing mannose occupies the conserved monosaccharide pocket [21]. Strict linkage specificity of GNA was explained by the fact that the C2 hydroxyl groups of both mannose residues are inaccessible in the bound state and strongly involved in hydrogen bonding.

In the present study we report the 2 Å resolution structure of a complex of GNA with a branched 3,6 core mannopentaose (Manα1,6-(α1,3-Man)Man-α1,6-(α1,3-Man)Man) (Man-5) (Fig. 1), chosen as an extension of our studies on monomannoside and dimannoside complexes and to simulate a more complex and biologically relevant situation. The results show that this oligosaccharide binds strongly only to two of the domains (CRD1 and CRD3), and uses two distinct binding modes. One is a cross-linking mode in which the outer 3,6 tri-Man arm (M1–M2–M3) bridges between twofold related GNA dimers utilizing the conserved monosaccharide pocket in the CRD1 domain; only the terminal sugars are in contact with the protein surface. The other is a binding mode in which the same trimannoside (M1–M2–M3) is bound to an extended binding region in the CRD3 domain, which is composed of the conserved monosaccharide pocket and specific subsites for the terminal non-reducing mannose residues. The latter binding mode explains the linkage specificity requirements of this lectin. In agreement with previous studies, bound carbohydrate mediates crystal contacts determining the specific lattice symmetry.

Figure 1



Structure of a typical core oligosaccharide of high-mannose glycans. The portion shown as shaded circles and labeled M1–M5 represents the oligosaccharide which was co-crystallized with GNA.

Results

Structure refinement and model quality

As summarized in Table 1, the GNA–Man-5 complex refined with an R-factor of 16.0 % (R_{free} = 19.4 %), including reflections with $F_{obs} > 75.0$ and $> 2\sigma(F_{obs})$. In the final cycle

Table 1

Refinement statistics of the final cycle.

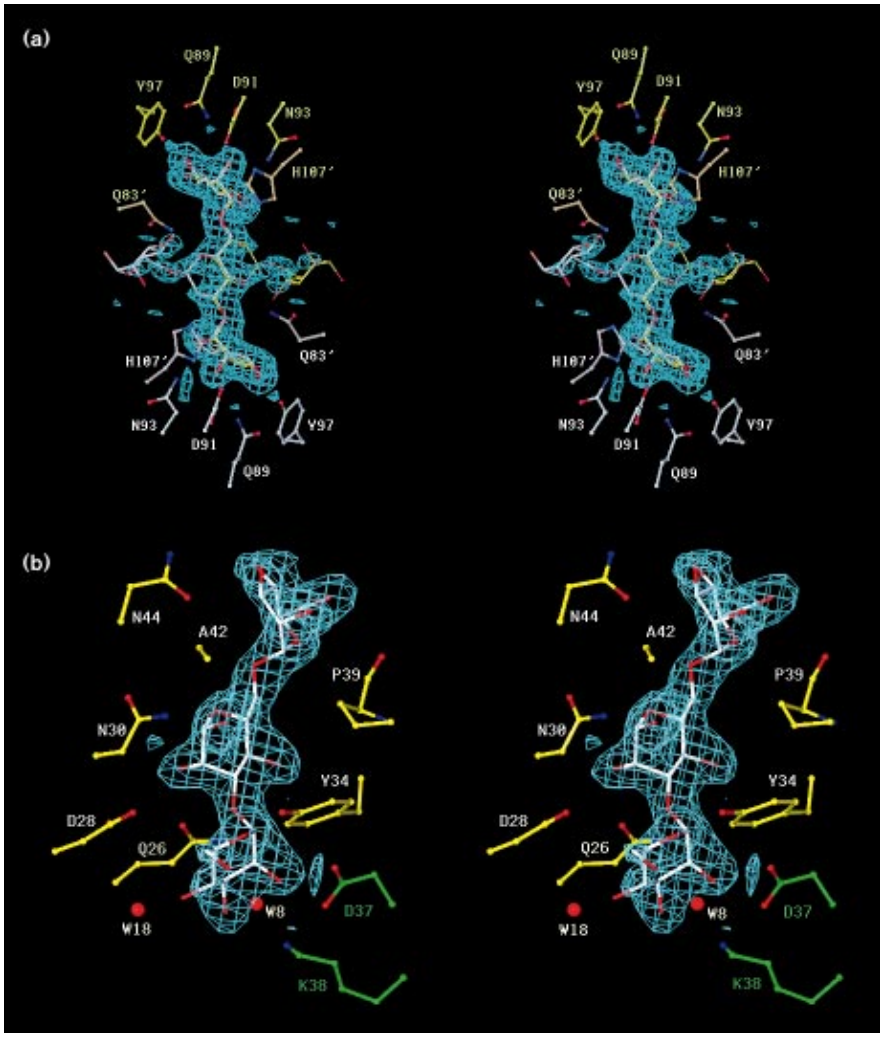
Resolution range (Å)	50.0–2.0
Reflections used* (total No. reflections)	8918 (11166)
Crystallographic R factor† (for all data)	0.16 (0.184)
R_{free}‡ (for all data)	0.194 (0.222)
No. of protein atoms§	860
No. of sugar atoms	91
No. of water molecules	55
Overall average temperature factor (Å^2)	32.7
Average temperature factor for protein atoms (Å^2)	32.1
Average temperature factor for sugar atoms (Å^2)	29.2
Average temperature factor for solvent atoms (Å^2)	47.1
Estimated coordinate error# (Å^2)	0.25/0.28
Rms deviation from ideal bond lengths (Å^2)**	0.007
Rms deviation from ideal bond angles (°)**	1.8
Rms deviation from ideal dihedral angles (°)**	24.8
Rms deviation from ideal improper angles (°)**	1.2

*$F_{obs} < 75.0$ and $F_{obs} < 2\sigma(F_{obs})$ excluded from refinement.
†Crystallographic R factor = $\Sigma(||F_{obs}| - |F_{calc}||)/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitude, respectively.
‡R_{free} is the crystallographic R factor calculated for a subset of randomly selected reflections (typically 5–10 %) not used in the phasing process [55]. §Residues Thr9, Lys38, Ser55 and His107 have alternative conformations. #Values estimated with methods of Luzzati [59] and Read [56], respectively. **The parameters of Engh & Huber [54] were used.

a considerable reduction in the value of the R-factor was observed when a correction was applied that takes into account the contribution of bulk solvent to the calculated structure factors. The final model included one GNA subunit plus 55 water molecules and eight bound mannose residues. The structure of the GNA monomer compares closely with that of the previously determined orthorhombic structure of the tetramer, with the exception of a number of surface residues. Superpositions yielded RMSD values of 0.68 Å for all atoms and 0.48 Å for Cα atoms only. The Ramachandran plot demonstrates that the final model has excellent geometry although a strained conformation is again observed for Val36 which is located in a tight β turn. It is noteworthy, that the side chains of Thr9, Lys38, Ser55 and His107 were found to have two discrete conformations which were refined separately with variable occupancies over several refinement cycles. In the final cycle, fixed partial occupancy values were assigned to these alternate side chain positions such that the combined occupancy for each residue totals 1.0. In the case of His107 and Lys38, side chain movements are caused by sugar binding. The side chain of His107 changes position as a result of steric crowding by the M4–M5 arm in one of the binding orientations of the Man-5 molecule. In the new position the imidazole ring (ND1 atom) is stabilized through a hydrogen bond with the side chain (OD1 atom) of N91 (2.75 Å distance). Occupancies of these two conformers refined to similar values, which is as expected because the two binding orientations of Man-5 are twofold related by the crystal symmetry and therefore occupied 50%. In the case of Lys38, the side chain moves closer to bound Man-5 to allow a stabilizing hydrogen bond contact to be made with the 2-OH of the non-reducing α1,3 Man (M3). The fact that this position is not fully occupied correlates well with incomplete Man-5 occupancy at this site (q = 0.7 in final cycle).

The C-terminal residue Gly109 was omitted from the model due to lack of density and the penultimate Thr108

Figure 2



Stereo views of the bound 3,6 tri-Man fitted to the final $2F_{obs}–F_{calc}$ electron-density map at two different binding locations. The map was contoured at 0.9 σ level. (a) 3,6 tri-Man bridging the CRD1 binding sites of twofold related molecules. The two alternative 3,6 tri-Man orientations are distinguished by color (white and yellow). Only four of the five mannose residues are shown. Binding-site residues that belong to different tetramers are colored either in white or yellow; side chains located on a twofold related subunit are labeled with a prime. (b) 3,6 tri-Man bound in the extended binding mode at the CRD3 site. Side chains located on different subunits are distinguished by color (yellow or green). Two water molecules are shown as red balls and labeled 'W'. (Figure generated using the program O [57].)

was also found to be flexible (B > 50.0 Å). Later refinement rounds concentrated mainly on determination of different pentasaccharide binding modes and on identification of water sites. The complete pentasaccharide was not visible at either of the three domain sites in the refined $2F_o–F_c$ map, although weak density for M4 and M5 was observed at the two major binding locations in the initial map based on phases of the molecular replacement model. Thus, the M4–M5 moiety appears to be conformationally disordered due to lack of binding interactions. It was possible to orient the tri-Man arm (M1–M2–M3) unambiguously in clearly defined density at these two binding locations (Figs 2a,b). Neither site is fully occupied as suggested from the refined occupancy values.

The number of water molecules included into the model in later cycles of refinement were found to be fewer than those refined for the tetrameric orthorhombic MeMan complex (55 versus 80 sites/subunit). Many of these waters refined with B-factors > 5 Å^2. Comparison of their positions by superposition demonstrated that a large number coincide in both crystal structures and are likely to be of structural importance.

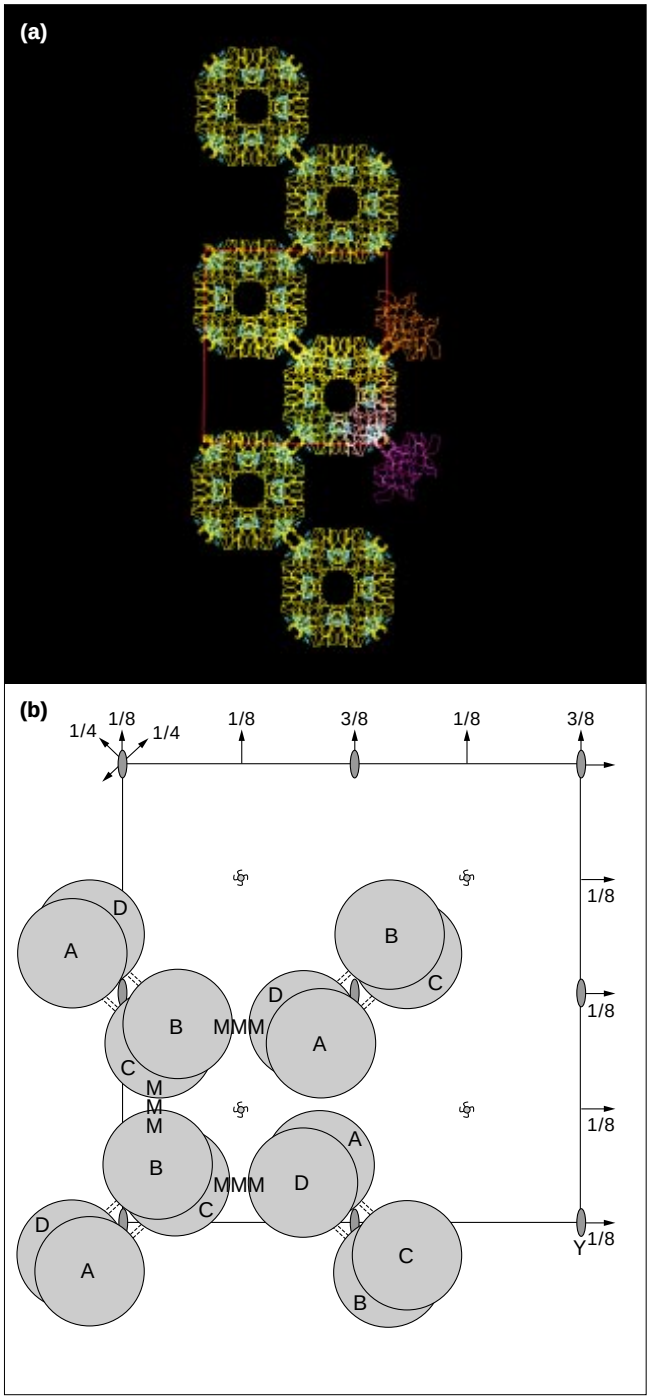
Overall structure of the cross-linked complex

Compared with the earlier determined structures of the orthorhombic MeMan and the cubic Man-2 complexes, the GNA–Man-5 complex described here has a uniquely different lattice structure. The branched 3,6 mannopentaose is incorporated into the tetragonal crystal lattice mediating intermolecular contacts, thereby creating a pore-like super structure as schematically illustrated in Figure 3. The outer tri-Man arm cross-links GNA dimers (A–D or B–C type) [16], which form the links of the helical assembly along the crystallographic 4_1-axis giving the appearance of defined pores when viewed along the crystal c-axis (Fig. 3a). A complete turn of this quasi-helical structure consists of four dimers (Fig. 3b). These helical columns occupy ¼ of the unit cell and are alternately packed with completely vacant 30 Å wide spaces, presumably filled with disordered water. They make contact with one another through the weaker of the subunit/subunit interfaces of the GNA tetramer (A with B and C with D), and involve only the A–D and B–C dimers [21] (Fig. 3b). The interior region of each pore consists of a 10–15 Å wide empty channel. Residues Arg101 and Gln80 are protruding furthest into the center. Bound saccharide is responsible for the majority of the molecular packing interactions either through cross-linkage between CRD1 sites, or clustering of oligosaccharides bound to CRD2- and CRD3-type sites (Figs 4a,b). In these regions, which center on crystallographic twofold axes, no defined positions are observed for the flexible inner two mannose residues (M4–M5).

Mannopentaose binding modes

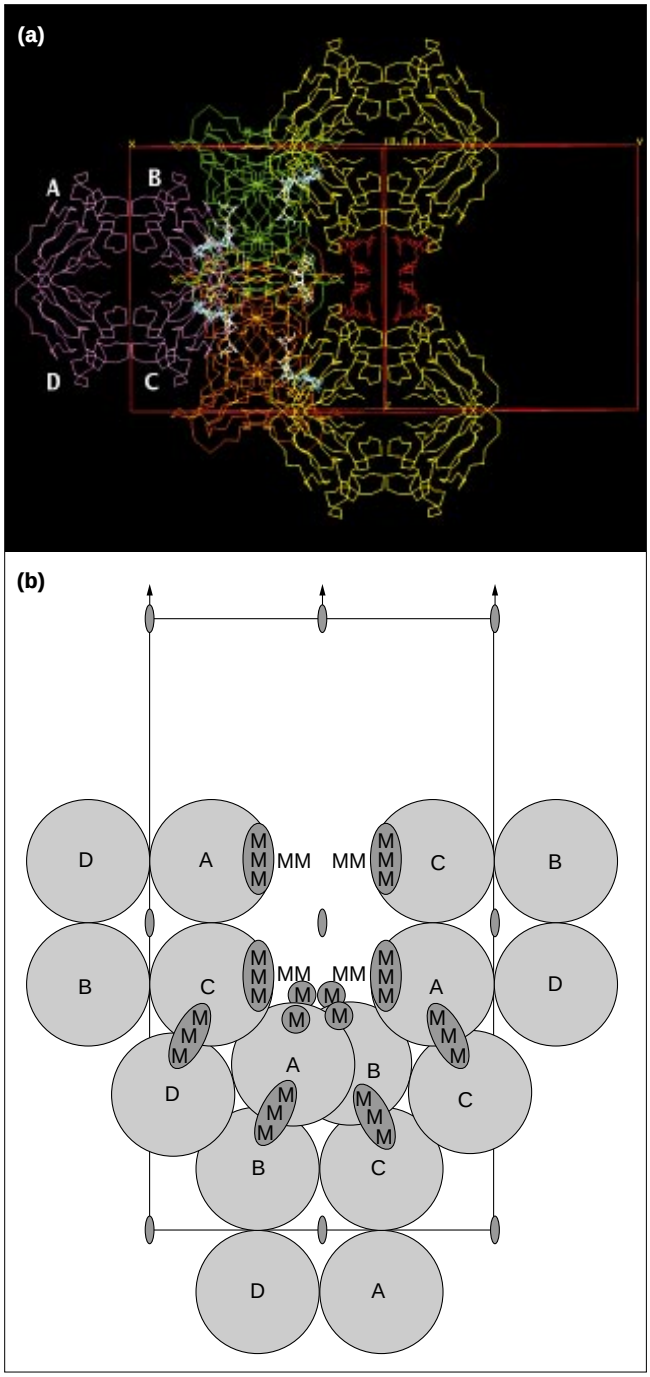
The mannopentaose molecule binds in distinctly different binding modes at the highly occupied CRD1 and the

Figure 3



Packing of GNA–Man-5 complexes along the 4_1 crystal axis. **(a)** Graphics display of the packing of six pore-like assemblies of GNA dimers (A–D or B–C). Subunits are shown as Cα traces in yellow. Bound carbohydrate is colored in cyan. The outline of the tetragonal cell is shown in red. Four complete tetramers are depicted in the center of the figure in gold, yellow, magenta and green. (Figure generated using the program O [57].) **(b)** Schematic diagram illustrating how the four cross-linked GNA dimer units assemble to form one helical turn. A, B, C and D designate individual GNA subunits and MMM represents the tri-Man molecule.

Figure 4



Packing of GNA–Man-5 complexes viewed along crystallographic twofold axes at a 45° angle between axes a and b. The diagonal between axes a and b lies in the horizontal direction, and the c-axis lies in the vertical direction. (a) Graphics display of GNA tetramers (yellow, magenta, light-green and gold), shown as Cα traces . Mannose oligosaccharides bound at the CRD1, CRD2 and CRD3 binding sites are shown in cyan, white and red, respectively. The outline of the unit is shown in red. (Figure generated using the program O [57].) (b) Schematic illustration of the packing of cross-linked GNA tetramers as viewed in Figure 4a. A, B, C and D designate individual GNA subunits and MMM represents the tri-Man molecule.

CRD3 binding sites, consistent with previous structural results from the Man-*2* crystal complex [21]. In both these binding modes only the outer trimannoside branch (M1–M2–M3) is involved in binding interactions. At the CRD*2* binding site, the observed electron density was weak and sufficient for fitting only two mannose residues (M1 and M2), which were refined despite high B-factors (>50 Å^2). The disaccharide was positioned with M2 bound in the conserved monosaccharide pocket and the α1,6-linked M1 residue oriented toward a non-polar subsite analogous to that seen in the CRD3 site. The α1,3-linked M3 residue is not visible, as it does not have a special binding region as present in the CRD3 binding site. Evidence from the Man-*2* complex, where the complete disaccharide (corresponding to M3–M*2*) was visible as a result of favorable lattice interactions [21] suggests that 3,6 tri-Man binds to an extended binding region in this domain similar to that observed in the CRD3 domain. However, much weaker binding affinities are expected in the absence of subunit/subunit interactions and favorable lattice contacts.

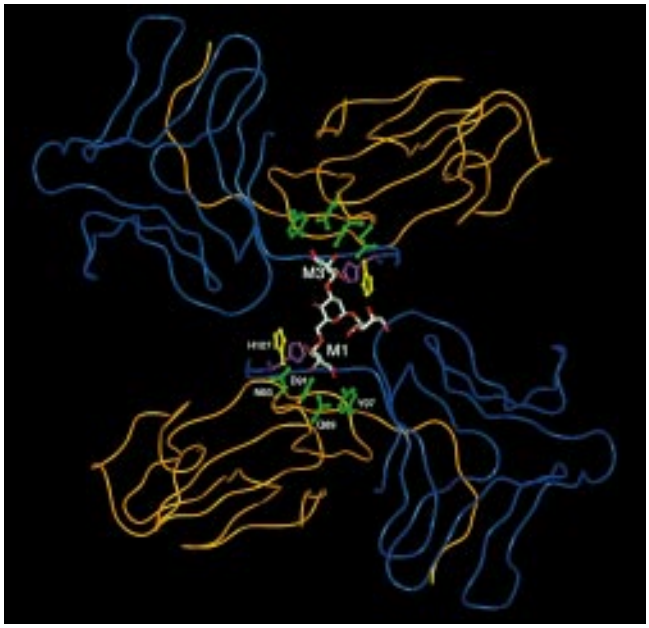
Cross-linking binding mode

The CRD1 binding site, identified as the highest affinity monosaccharide binding site, is utilized for cross-linkage of GNA dimers that assemble as helical columns. The outer tri-Man arm bridges between GNA dimers with either M1 or M3 bound to the conserved monosaccharide pocket (Fig. 5). Their binding positions are identical to that observed for MeMan, including the presence of the invariant water molecule [16]. Since the bridging electron density is centered on the crystallographic twofold axis (Fig. *2*a), and since there is no center of symmetry present in this 3,6 tri-Man arm (Fig. 1), it was concluded that two equally likely orientations must exist in this lattice, with either M1 or M3 positioned in the CRD1 site of either of the cross-linked dimers. Special cases of lattice symmetry of a similar nature have recently also been encountered in other ligand–protein complexes [*22*,*23*]. As shown in Figure *2*a, the electron density accommodates the two orientations reasonably well. It is broadened at the region of the central mannose (M*2*), as this residue occupies different positions in the two orientations generated by the twofold axis. No additional contacts are observed for M2 and the M4–M5 arm, which is exposed to solvent. Lack of visibility of this arm has been attributed to flexibility about the α1,6-linkage as suggested by the existence of two distinct conformers for this linkage in solution [24,25].

Extended tri-Man binding mode

The presence of an extended binding site in the CRD3 domain had already been postulated based on the Man-*2* complex [21], and is confirmed here by the higher resolution Man-5 complex. The electron density is consistent with a binding mode in which the internal Man (M*2*) of the outer 3,6 tri-Man branch sits in the conserved monosaccharide pocket (Fig. 2b). The terminal M3 occupies

Figure 5



Structure of 3,6 tri-Man cross-linked GNA dimers. Subunits that form the CRD1 monosaccharide-binding site (A–D or B–C pairs) are shown as Cα traces in blue and gold. The conserved binding-site residues (Gln86, Asp89, Asn91 and Tyr97) are colored green. The two refined side-chain positions of His107 are color-coded yellow and magenta; 3,6 tri-Man is shown in stick representation. Only one binding orientation is displayed for the bridging trimannoside. The bound non-reducing terminal mannose residues are labeled M1 and M3. (Figure prepared with InsightII).

the same subsite as observed in the Man-2 complex, with its axial C2-OH immobilized through H-bond interactions with residues Asp37 and Lys38 reaching across from the twofold related subunit (Fig. 6). This interface formed between subunits A and B (or C and D) has lower stability than that of the A–D or B–C dimer interface that forms part of the CRD1 binding sites. There is also a subsite for M1, which is located on the same subunit and has non-polar character (Pro39, Ala42). As depicted schematically in Figure 7, the binding interactions in this subsite consist primarily of van der Waals contacts. However, one long H-bond between C4–OH and the amide group of Asn44 is also present. An alternative binding mode in which the inner tri-Man arm (M2–M4–M5) occupies this extended site has been ruled out on the basis of the electron density map. In such a binding mode M4 would be positioned in the specificity pocket and M2, substituted on both its C3 and C6 hydroxyl groups, would occupy the 1,6 subsite. The observed electron density at this site is insufficient for such a large moiety.

Bound tri-Man conformations

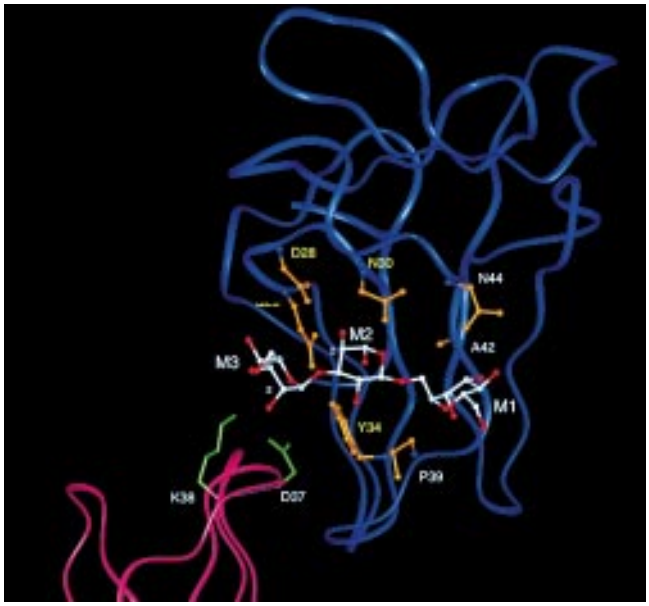
Analysis of the adopted conformations of bound 3,6 tri-Man in the two distinct binding environments revealed differences resulting from flexibility about the glycosidic torsion angles ϕ and ψ (Fig. 8). The observed values for the torsion angles of both the α1,3 and α1,6 linkages are, however, in good agreement with low energy solution conformers identified on potential energy surfaces and by NMR [25,26] (Table 2). Moreover, they are similar to values observed recently in other cross-linked crystal complexes [20,27]. The conformation selected in the cross-linking tri-Man is energetically somewhat more favorable than that adopted in the extended binding mode [26]. The two distinct bound conformers are shown superimposed in Figure 8 with the central M2 in register. The differences in the positions of the terminal mannose residues are primarily due to flexibility about the ψ angles, which differ by 60–70° in the two linkages. Unexpectedly, the ω angle for the terminal α1,6-linkage are in close agreement in all three binding environments (Table 2). The less favorable gt conformer is selected ($\varpi = 50–60°$) in this crystal complex, while in solution the gg conformer ($\varpi = 180°$) was shown to be preferable [26]. Changes in this angle are known to critically affect the overall structure of an oligosaccharide and its value may differ depending on whether the C1–OH of the reducing mannose is involved in an α- or β-linkage. For example, the ω-angle for the Man-α1,6-Manβ linkage appears to be sequence dependent, while that for the Man-α1,6-Manα linkage is not [25]. Thus, it is not surprising to find that the ω-values observed in this complex for the Man-α1,6-Manα linkage differ from those observed in other lectin-bound oligosaccharides, where a Man-α1,6-Manβ linkage is present [20,27]. In fact, flexibility in this angle appears to be the primary reason why an S-lectin–octasaccharide complex crystallizes in three different types of crystal lattice under the same crystallization conditions [27]. Our results further support the view that oligosaccharides structurally adapt themselves to the spatial requirements of the protein surface while maintaining low energy conformations that exist in solution.

Discussion

Relationship between structure and specificity among members of the *Amaryllidaceae* family

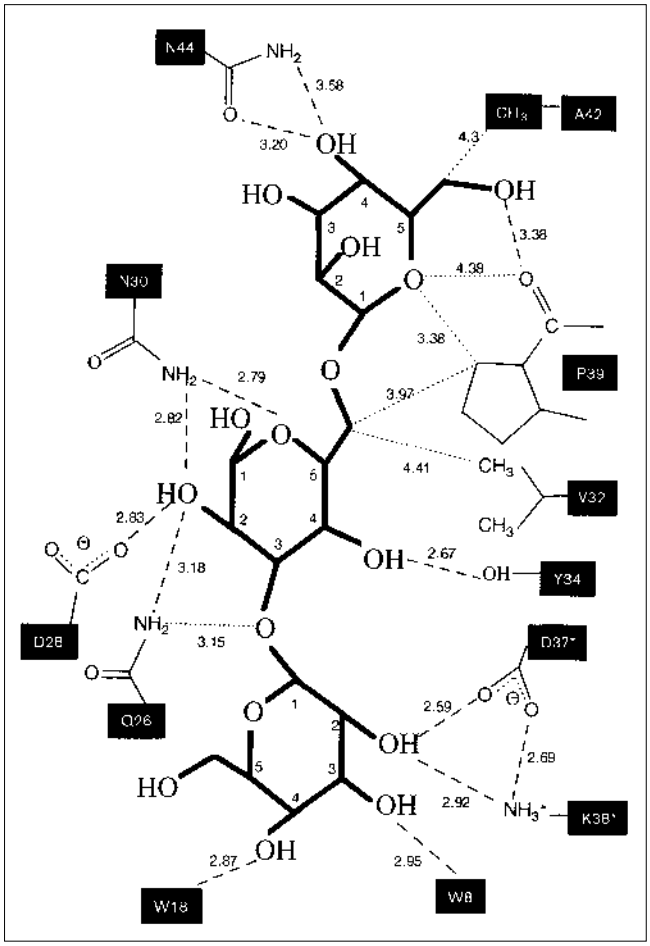
Knowledge of the three-dimensional structure of GNA and the binding locations of three types of oligo-mannosides now provides a basis for understanding the specificity differences among some members of the *Amaryllidaceae* lectins reported by Kaku and Goldstein [13,14]. In the case of GNA, hapten inhibition studies have shown specific requirements for the α anomer of mannose, an axial OH at C2, equatorial OH groups at C3 and C4, and the presence of a terminal α1,3-linked mannose for high affinity binding [10]. The two observed binding modes in the crystal explain the specificity requirements for both mono- and oligo-mannoside binding [16,21]. The fact that oligo-mannosides that possess terminal α1,2-linked mannose do not inhibit mannan precipitation by GNA, is consistent with

Figure 6



Extended 3,6 tri-Man binding mode at the CRD3 binding region. Subunits are depicted as Cα traces in blue and magenta. Mannose residues are labeled M1, M2 and M3 (Fig. 1). The C2–OH position in M2 and M3 is labeled with a '2'. Side chains that interact with bound tri-Man are displayed in gold; those with yellow labels belong to the conserved monosaccharide site, and those with white labels belong to the 1,6 subsite for M1. The two side chains (Asp37 and Lys38) that interact with terminal M3 across the dimer interface are shown in green.

Figure 7



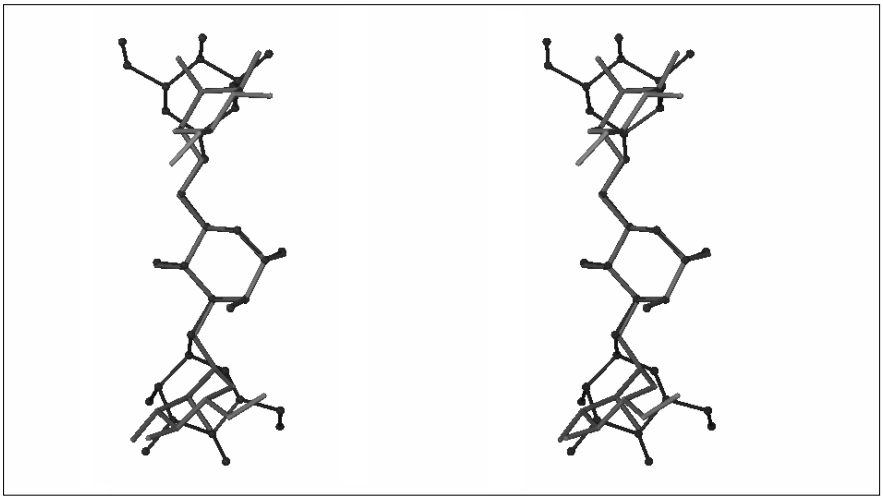
Schematic illustration of all ligand–protein contacts observed for 3,6 tri-Man bound at the extended CRD3 binding site. Dashed lines indicate hydrogen bonds and dotted lines van der Waals contacts. Contact distances are given in ångströms (Å). Amino acid residues are labeled by the one letter code and sugar carbon atoms are numbered. Residues labeled with an asterisk belong to a twofold related subunit.

our finding that the 2-OH is unavailable for substitution (Fig. 7).

Specificity differences among the *Amaryllidaceae* lectins involving oligo-mannosides of different linkages may be related to the oligomeric state (dimer or tetramer) [28] as well as variability in the amino acid sequence. Since the specific monosaccharide binding site is conserved among all lectins of this superfamily (including the *Alliaceae* lectins), the binding interactions are expected to be identical. However, substitution of other residues that contribute van der Waals contacts to bound MeMan, such as the C-terminal arm (Fig. 9), may explain small affinity differences among members. For example, stronger binding of MeMan to Narcissus pseudonarcissus agglutinin (NPA) could be the result of more extensive stacking contact with His83 and Tyr107 relative to GNA, where these residues are Asn and His (Fig. 9).

Specificity of GNA for terminal α1,3-linked mannosides is explained by the extended binding mode observed at the CRD3 site, which requires mannose residues in adjacent positions. When either of these residues is replaced by another sugar, the binding affinities are drastically reduced [13]. Although the polar dipeptide (Asp37, Lys38) that interacts across the dimer interface with the C2-OH of the terminal Man is conserved in all *Amaryllidaceae* lectins (Fig. 9), it is only used in the tetrameric proteins, GNA and Hippeastrum hydr. agglutinin (HHA). Thus, one would expect that specificity for α1,3-linked Man-2 at this site is much weaker in the case of a dimeric lectin (A–D dimer) where this interfacial contact is absent. However, the results of Kaku *et al.* [29] show comparable inhibitory potencies of Man-2 for dimeric NPA and tetrameric HHA, although the association constants indicate that HHA binds Man-2 4.5 times more strongly than does NPA [14]. One may speculate from these data that perhaps other interactions contribute to binding of the non-reducing mannose in the case of NPA. For instance, it is conceivable that Glu27 could provide a good H-bond contact to bound Man, as it is free to rotate and not involved in an ion pair interaction with residue 50 (Ser), which is an Arg in GNA.

Figure 8



Stereo view of the conformational differences of 3,6 tri-Man in the two distinct binding modes. The central residues are matched as reference. The oligosaccharide structure in the cross-linking binding mode (at CRD1 site) is modeled in stick representation and that in the extended binding mode (CRD3 site) in ball-and-stick representation.

The available binding data for GNA are even more difficult to rationalize. Although estimates of the relative potency of this disaccharide versus MeMan in inhibiting lectin precipitation show a much larger value for GNA than for NPA (14.2 versus 3.1), the magnitudes of association constants suggest that GNA binds Man-2 with lower affinity by a factor of 2–12 compared with NPA and HHA. These data cannot be rationalized only on the basis of the extended 3,6 tri-Man binding mode at the CRD3 site, which is assumed to be similar in GNA and HHA. Mutations of amino acid residues in the binding regions of the CRD1 and CRD2 domains may contribute to favorable subsites in these lectins not present in GNA. For instance, model building suggests that a tri-Man binding mode similar to that observed at the CRD3 binding site, is possible at the CRD1 site in NPA, considering mutations at several residues which could interact with the non-reducing terminal residues (M1 and M3). Replacement of Tyr84 (GNA) by Phe84 allows a favorable contact to the terminal α1,3-Man together with the hydroxyl group of Ser6. The terminal α1,6 Man would be situated close to the C-terminal region Ala103–Gly105, which is conserved in all four *Amaryllidaceae* lectins. It is uncertain, however, to what extent sequence variability beyond residue 105 (His107 to Tyr107), and the 3 residue extension in HHA, may contribute to some of the observed affinity differences.

While there is a strict requirement for mannose (axial 2-OH) at the α1,3-terminal position, sugars other than Man could be tolerated in the α1,6 subsite. Here, the C2-OH of bound mannose is exposed to solvent and the opposite side of the sugar ring faces the protein surface (C4–C5–C6–OH). There are numerous van der Waals contacts (Fig. 7) of non-polar nature that contribute to binding. This terminal sugar is also represented by strong density (Fig. 2b) suggesting that this arm is tightly bound. These findings confirm predictions based on hapten inhibition data of various branched mannose-containing trisaccharides [13]. As apparent from Figure 9, this binding region (Pro39, Ala42, Asn44) is also present in the CRD2 site with no significant sequence differences in the residues that contact α1,6-Man among the four lectins.

Differences among members of this family have also been observed in quantitative precipitation studies with linear α1,3- and α1,6-linked mannose oligomers. Even though both HHA and GNA are tetrameric, HHA is precipitated by tetrasaccharides, while GNA and also dimeric NPA require oligomers of at least five residues in length for precipitation [14]. Assuming that only the terminal non-reducing mannose residues are involved in the binding interactions, the ability to successfully generate 2- or 3-dimensional lattices may depend on subtle differences in

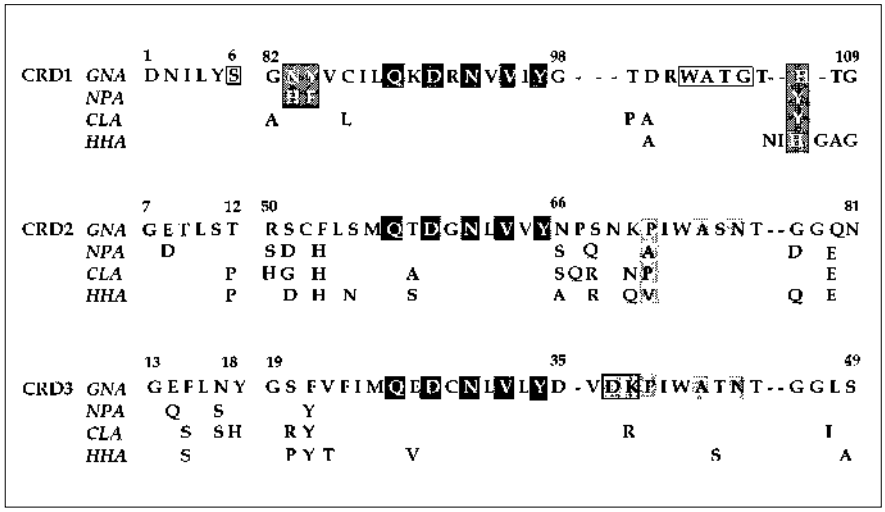
Table 2

Glycosidic bond torsion angles.

	Linkage*	ϕ	ψ	ω†
CRD1 site	M1 α1–6 M2	74.0	–177.0	60.0 (–57.0)†
	M3 α1–3 M2	87.0	175.0	–
	M2 α1–6 M4	77.0	–158.0	–76.0 (169.0)†
CRD2 site	M1 α1–6 M2	75.0	107.0	52.0 (–64.0)†
CRD3 site	M1 α1–6 M2	81.44	102.1	57.0 (–60.0)†
	M3 α1–3 M2	68.58	119.0	–

*Residues are abbreviated as shown in Figure 1. The following convention applies for the torsion angles: $\phi = O_5–C_1–O_1–C_{x'}$ (x is 3 or 6); $\psi = C_1–O_1–C_{x'}–C'_{(x+1)}$; $\omega = O_{1-}C_{6'}–C_{5'}–O_{5'}$; errors are estimated as ±15°. †Values in parentheses were measured for $\omega(H) = O_1–C_{6'}–C_{5'}–H_{5'}$.

Figure 9



Alignment of the amino acid sequences of four *Amaryllidaceae* lectins. Residue numbering is sequential from 1 to 109; dashes represent deletions. Each line represents one of the three β sheet domains (named CRD1, CRD2 and CRD3). Sugar-binding residues, conserved in all three domains in the monosaccharide-binding site, are shown in black boxes with white letters. In the top row (CRD1 domain) residues that could be involved in additional monosaccharide-binding interactions are shown within shaded, bordered boxes in white letters; residues that could potentially interact with the terminal residues of 3,6 tri-Man (M1 and M3) are shown boxed and in black letters. Residues that form the α-1,6 Man subsite in CRD2 and CRD3 are shaded in light gray (black letters). In CRD3, residues that form the α-1,3 Man subsite across the dimer interface are shown in a darkly shaded box in black letters.

surface properties of these lectins. Moreover, one cannot rule out the possibility that a different CRD binding site is used for cross-linkage in some of these lectins.

Comparison with other mannose-binding lectins

In comparison to the binding site topologies of other Man-specific lectins (from *Leguminoseae*), GNA clearly presents different structural determinants for mannose binding. Strict specificity for mannose is due to the fact that the axial C2-OH is completely buried in the bound complex and immobilized through hydrogen bonds with three different side chains. The binding interactions observed in both the legume and C-type mannose-binding animal lectins (MBP) involve the C3, C4 and C6 hydroxyl groups, while the C2–OH is exposed and available for substitution, explaining their secondary specificities for Glc, GlcNAc, and Gal. Among the diverse group of legume lectins, architecturally similar binding sites can distinguish different sugars [1,17,30,31], while the MBP lectins appear to be less sensitive to sugar orientation [32] as long as the correct side of the sugar ring (C3–C4 hydroxyls) is aligned for optimum contact with Ca^{2+}. In the case of the *Lathyrus ochrus* lectins (LOL1), 100-fold higher affinities were reported for fucosylated oligosaccharides [33], suggesting the presence of a specific subsite.

The general character of the sugar binding sites in lectins of all three families is polar, allowing optimum contact with the sugar hydroxyl groups through H-bonding. Bound metal ions (Ca^{2+} and Mg^{2+}) may also be involved in sugar stabilization either directly, as in the case of the MBP lectins [20,32], or indirectly, as observed in the legume lectins. In contrast, there is no evidence available to suggest metal ion involvement in mannose binding to GNA. Its mannose binding pocket is composed of five conserved residues (Asp, Asn, Gln, Tyr, Val) positioned on a tight β turn. Moreover, the absence of sugar ring-stacking contact with an aromatic side chain further distinguishes GNA from numerous other lectins, including mannose binding protein (MBP) and lectins from *Leguminoseae* and *Gramineae*.

Comparisons have also been carried out to assess the differences in binding affinities of mannose oligosaccharides for the *Amaryllidaceae* and *Leguminoseae* lectins. Inhibition studies have shown that ConA binds α1,3- and 1,6-linked oligo-mannosides much more tightly (10–50 times), and with 160–500 times higher association constants for α-D-MeMan, than the *Amaryllidaceae* lectins [13]. On the other hand, the *Amaryllidaceae* lectins have a much greater propensity to form precipitates with linear α1,3- or α1,6-linked mannans than do, for instance, ConA and pea lectin (PSA). Similarly, in the case of the branched 3,6 tri-Man, no cross-linkage is observed in co-crystals of this trimannoside with ConA [34]. Tri-Man binds to an extended binding region such that, in contrast with the binding mode observed at the CRD3 site in GNA, the terminal α1,6-linked mannose residue resides in the conserved monosaccharide pocket.

Three other properties of the *Amaryllidaceae* lectins contrast interestingly with those of the legume and MBP lectins. First is their threefold redundancy of a conserved mannose binding site within the polypeptide chain. Second is the modulation of their mannose affinities through inter-subunit contacts (a similar situation exists in the *Gramineae* lectins where four independent sugar-binding sites are composed of binding regions on both monomers [35]). Thirdly, GNA has flexibility in that its conserved monosaccharide pocket can accommodate non-reducing terminal mannose residues, internal mannose residues and reducing mannose residues.

These examples illustrate that nature has found multiple solutions in creating recognition sites for identical ligands that fine-tune specificity requirements through subtle variations in structure.

Oligosaccharide mediated super-structures

Since the first cross-linked complexes were characterized [20,36], others have followed, revealing in atomic detail ordered lattices each unique with respect to the particular lectin and oligosaccharide co-crystallized. In each case the lectin represents a highly conserved family with different molecular properties and carbohydrate binding sites [18]. The GNA–Man-5 complex described here provides yet another interesting example of a distinctly different cross-linked lattice. In analysing and comparing the properties of these lattices one can now begin to define common structural determinants and relate structure to function.

One question of interest is whether the size and structure of a particular multivalent oligosaccharide determines the symmetry of the crystal lattice. An affirmative answer can be given in view of the structural detail of five cross-linked crystal complexes and extensive evidence based on quantitative precipitation studies and electron microscopy [37–40] showing that unique lattices are formed for a particular lectin depending on the composition of the bound oligosaccharide. The combined results of three different GNA–mannoside complexes demonstrate that the symmetry of the lattice increases as a function of the size of the oligosaccharides. In the orthorhombic monosaccharide complex, the complete tetramer is contained in the asymmetric unit, while in the cubic Man-2 complex only half-tetramers constitute the asymmetric unit. In the tetragonal structure of the Man5 complex all subunits are equivalent and the three molecular twofold axes are incorporated into the crystal symmetry. Moreover, molecular assembly resulting from tri-Man cross-linkage obeys 4_1 symmetry. Other examples of high symmetry lattices are the hexagonal cross-linked pentasaccharide complex of Gal-specific soybean agglutinin (SBA) [40], and one of three octasaccharide-S-lectin complexes [27]. In the case of SBA, the unliganded protein crystallizes in a lower symmetry monoclinic space group (C2). In contrast, both low and high symmetry lattices (hexagonal, trigonal and monoclinic) are observed for the octasaccharide complex of S-lectin [27,41,42]. Lower symmetry lattices were also observed in other lectin co-crystals, as for instance, in the monoclinic ($P2_1$) lattice of MBP complexed with an N-linked octasaccharide, and in the asymmetric orthorhombic ($P2_12_12$) complex of wheat germ agglutinin (WGA) with a bivalent sialoglycopeptide. Based on these examples, it may be concluded that no consistent pattern has emerged by which specific crystal symmetry relationships can be predicted.

Although cross-linkage is readily observed with multivalent internally symmetric oligosaccharides [20, 27, 37–40], structural symmetry does not appear to be a requirement. This is demonstrated by two examples: the asymmetric O-linked tetrasaccharide that cross-links WGA molecules [36], and the asymmetric 3,6 tri-Man that constitutes the cross-link in the GNA–Man-5 structure, and sits on a twofold axis and alternates between two equally likely orientations. Analysis of the molecular packing in the five available crystal complexes reveals two types of preferred super structure motifs. Molecular assembly along a fourfold screw axis, resulting in a helicoidal motif, was found in the pore-like GNA–Man-5 structure of cross-linked dimers as well as in the trigonal S-lectin-octasaccharide crystal form. A second type of motif consists of parallel elongated arrays in which molecules that are translated with respect to one another are cross-linked [27,36]. In all complexes only terminal non-reducing sugars mediate binding interactions of the cross-links and identical binding sites are used in the cross-linked partner molecules. This is true in particular for those lectins that possess one binding site/subunit. In cases with multiple binding sites/subunit (WGA and GNA) not all binding sites are utilized equally. For instance, in the WGA sialoglycopeptide complex two non-identical sites participate in the cross-link, each on a different monomer of the cross-linked molecules [36]. This asymmetry extends to the fact that only three of the eight available sites were occupied in this complex and two types of binding modes, as found in the GNA–Man-5 complex, characterize the lattice.

The three residue cross-link present in the GNA–Man-5 complex is the shortest one thus far observed and may in fact constitute a minimum distance in terms of close-packing criteria. Four to seven residue cross-links have been observed in the other four complexes. In all cases the bridging oligosaccharides are represented by strong electron density and have adopted roughly extended immobilized conformation, which are consistent with low energy conformers observed in solution.

Another common feature is the fact that a high solvent content (>55%) appears to be characteristic of such cross-linked lattices. A crystal solvent content of 70% was reported for the SBA-pentasaccharide structure determined at 2.6Å resolution. Absence of electron density in the 30Å wide channels that alternate with the columns of tightly packed protein (Fig. 3a) also suggests a high content of disordered solvent. In addition, it is not surprising to find that lattice contacts consist primarily of sugar mediated protein interactions and direct sugar-sugar contacts.

Biological implications

Lectin-induced cell surface phenomena, such as mitogenic stimulation, agglutination, and signal transduction [4], are believed to involve cross-linkage or clustering of lectin–receptor complexes. Moreover, the widely observed cooperative binding behavior of lectins may also be

explained in terms of lectin–receptor cross-linkage and utilization of different classes of binding sites [36]. Despite their weak affinities towards simple oligosaccharides, lectins bind with significantly higher levels of specificity to their cell surface receptors [43]. Although a clear understanding of the structural basis of these binding phenomena is still lacking, the multivalent nature of both lectin and cell surface oligosaccharide plays a crucial role, increasing affinities by several magnitudes through 'subsite multivalency' [18]. Evidence in support of the cross-linkage theory has emerged in recent years from X-ray diffraction and electron microscopy studies [20,27, 36,38–40,44]. High resolution crystal structures of lectin–oligosaccharide co-crystals have revealed lattice structures unique to each type of lectin complex. Conformational restrictions of the branched oligosaccharide and relative orientations of sugar-binding sites on the surface of the oligomeric lectin are determining factors in the type of super structure motif (helical or sheet) generated.

The structure described here represents yet another example of a cross-linked lattice of unique symmetry and provides a glimpse of a possible molecular assembly that could exist *in vivo* when GNA binds to high-mannose glycans on cell membranes. The 3,6 core mannopentaose chosen for this study is a common component of biantennary and triantennary sialylated and high-mannose-type N-linked complex carbohydrate chains present, for instance, in glycans of the HIV surface envelope glycoprotein GP120 [45]. The effectiveness of GNA and other *Amaryllidaceae* lectins in inhibiting HIV infectivity has been attributed to their absolute specificity for **α**-D-mannose. The structural basis for specificity of GNA for **α**1,3-linked terminal dimannosides has now been explained in terms of the multi-subsite binding region in the CRD3 domain and involvement of the axial C2-OH groups of both mannose residues in hydrogen-bond interactions. Although the mechanism of HIV inhibition by these lectins is not yet understood, a number of binding studies have suggested that lectin binding may induce conformational disturbances or destabilization of the GP120–GP41 complex [46]. Cross-linkage between lectin and branched oligo-mannose moieties, as observed in this structure, would offer one possible mechanism by which binding of GP120 to specific receptors [47] could be blocked in subsequent viral fusion events.

The GNA tetramer can be considered a special case among the group of mannose-specific lectins in that it has more than one binding site per subunit. With 12 potential binding sites, it represents the highest density lectin characterized so far [48]. Reminiscent of the situation in wheat germ agglutinin [36], the three independent sites are not utilized equally and oligosaccharide binding modes and occupancies may differ depending on subsidiary interactions from neighboring or symmetrically related molecules. Availability of such a large number of sites allows formation of a wide range of multi-dimensional lattices of distinct symmetries with a high degree of specificity, as illustrated by the three types of crystal complexes examined so far. In each of these a different set of binding sites participates in lattice formation, and oligosaccharides of higher complexity tend to form higher symmetry lattices that exhibit a higher degree of sugar mediated packing contacts.

Materials and methods

Crystallization

GNA was isolated and purified as reported previously [8,49] and Man-5 was purchased from Dextra Laboratories (Reading, England). The GNA–Man-5 complex was crystallized by the hanging drop vapor diffusion technique using drop volumes of 7 μl and reservoir volumes of 0.5 ml. The reservoir solutions were composed of 33% saturated ammonium sulfate (0.05 M ammonium bicarbonate buffer, pH 8.2). The hanging drop contained 3.5 mg ml^{-1} protein, 0.28 mM Man-5 and 9% ammonium sulfate. The molar ratio of Man-5 to GNA tetramers was 4:1. Crystals formed within 1–2 weeks. They crystallized in the tetragonal space group $I4_122$ with unit cell dimensions a = b = 96.34 Å, c = 68.4 Å. A partial molar volume of 3.17 Å^3 Da^{-1} was consistent with the presence of 1 subunit/asymmetric unit and a crystal solvent content of 61%.

Data collection and processing

2.0 Å resolution X-ray intensity data were measured at room temperature on one crystal using a Rigaku R-AXIS IIC imaging plate system mounted on a Rigaku rotating anode X-ray generator. The data were processed with MOSFLM (Version 5.23) [50], and scaled using programs of the CCP4 program suite (SCALA, AGROVATA, TRUNCATE)[51]. Bijvoet-related reflections were averaged and Wilson scaling was performed (Wilson-B-factor = 25.0 Å^2). The final dataset, consisting of 11 166 reflections, is 99.7% complete. Relevant statistics are listed in Table 3.

Molecular replacement

The crystal structure of the tetragonal Man-5 complex was determined by the molecular replacement method using the GNA structure (subunit A) refined to 2.3 Å resolution [16] as the search model. Rotation and translation function searches were carried out with the program package AMORE [52]. Initially a self rotation function based on data in the resolution range 8–4 Å was calculated with an integration radius of 21 Å and a sampling step of 1°. Peak analysis did not reveal the presence of any significant local symmetry axis suggesting that a single subunit is contained in the asymmetric unit. This was consistent with estimated values of the crystal solvent content and V_m value. Trial cross rotation functions were calculated in several resolution ranges using a

Table 3

Statistics of data collection and processing.

Space group	$I4_122$
Unit cell dimensions (Å)	a=b=96.34 c=68.64
R_{merge}* (%)	7.4
No. of observations	54 560
Multiplicity	4.9 (4.8)†
No. of unique reflections	11 166
Highest resolution shell (Å)	2.05–2.00
Completeness (%)	
all data	99.7 (99.6)†
F_{obs}> 75.0 (%)	80.0 (53.0)†
$F_{obs} > 3\sigma$ (F_{obs}) (%)	83.2 (62.8)†

*$R_{merge} = \Sigma_{hkl}\Sigma_i|I_i - \langle I\rangle|/\Sigma\langle I\rangle$, where I_i is the intensity measurement of a symmetry-related reflection. †Results for highest resolution shell.

sampling step of 1°. These calculations were carried out with integration radii of 15 Å, 18 Å and 21 Å for resolution ranges 8–4 Å, 6–3 Å and 8–3 Å, respectively. The highest signal-to-noise ratio was observed for data in the 8–3 Å resolution range. Subsequent translation function computations, based on data in the same resolution range, resolved the space group ambiguity between $I4\,22$ and $I4_122$. Correct solutions were selected based on the magnitude of the R-factor, correlation coefficient and reasonable packing interactions. Values of 31.1 % for the R-factor and 77.3 % for the correlation coefficient yielded an unambiguous molecular replacement solution with favorable packing interactions for space group $I4_122$. The resulting model was optimized by rigid body refinement with AMORE yielding values for the R-factor and correlation coefficient of 28.8 % and 80.3 %, respectively.

Refinement

The molecular replacement model was further refined by simulated annealing with X-PLOR [53] (Version 3.1.f). A total of 13 rounds of refinement were performed using the parameters of Engh & Huber [54]. Initially several cycles of rigid body refinement were carried out in order to optimize the overall positions of the subunit molecules. After two rounds of rigid body refinement, using an overall B-factor of 15.0 Å and increasing the resolution from 3.0 to 2.5 Å the R-factors decreased to 28.4 % (R_{free} = 32.0 %). Refinement by simulated annealing was started without manual inspection of this model. However, refitting was necessary after round two. Only protein atoms were included in these two initial refinement rounds. The overall protein temperature factor refined to a value of 24.7 Å^2 , which compared closely to the Wilson B-factor (25 Å^2).

The following 'slow cooling' protocol was employed in rounds 1–3 and 5–7: T_{init} = 4000° K; T_{end} = 300° K; Δt = 0.5 fs; ΔT = –25° K for every 50 steps. In all other rounds the model was refined by conventional energy minimization [53]. A set of reflections comprising approximately 5 % of the data were randomly selected to compute the 'free R-factor', as a means of cross-validating the model [55]. Refinement was carried out against data with $F_{obs} > 2\sigma(F_{obs})$ in resolution range 8.0–2.0 Å in rounds 2 to 12. The pre-determined overall weight W_a, applied to the crystallographic energy term E_x, was reduced to obtain a more favorable balance between optimization of the stereochemistry of the model and minimization of the F_{obs}–F_{calc} difference. The specific values used in each round were based on the behavior of the free R-factor during test refinement runs. In order to weight bond lengths and bond angles stronger and to improve the stereochemistry of the atomic model the weights for the bond length and bond angle energy terms were increased from 1.0 to 1.1, except in the last refinement round. For refinement of the bound oligosaccharide molecules available parameter and topology files were found to be satisfactory.

Initial temperature factors (B-factors) for protein and sugar atoms were set to 15 Å^2 and those for water molecules to 20 Å^2. In the first two refinement rounds two B-factors were refined per residue. In later rounds individual atomic B-factors were refined at the end of each refinement round in three to four sets , each of which consisted of 20 to 30 steps of B-factor refinement alternated with 20 to 30 steps of positional refinement by energy minimization. In rounds 1–9 B-factors were reset before refinement, while in the last three rounds only values larger than 50 Å^2 were reset. In rounds 4–12 occupancies were also refined in the case of the Man-5 ligand (one per molecule) and for side chains with alternative conformations (Thr9, Lys38, Ser55 and His107). Typically 20 to 30 steps were performed after B-factor refinement and subsequently followed by positional refinement. Solvent molecules were included into the model and edited during the final refinement rounds. Since only a small number of well-ordered waters with B-values < 50 Å^2 could be identified, the refinement protocol was modified in round 13 experimenting with several parameters. First, a bulk solvent correction was applied (solvent density = 0.34; solvent radius = 0.25; B-factor = 50 Å^2) over the whole resolution range (100–2.0 Å) using all data, which had the effect of increasing atomic B-factors (B_{av} = 34.8 Å^2 versus 30.4 Å^2 in cycle 12) and reducing the R-factor from 0.219 to 0.187 (R_{free} from 0.236 to 0.220). Electron density maps indicated only marginally better definition of solvent regions, which allowed identification of an additional 32 water molecules. However, most of these had high B-factors (> 50 Å^2). In view of the relatively high protein B-factor (25 Å^2), indicating a sharp drop-off in the intensities at high resolution, we repeated the energy minimization run of round 13 using a trimmed reflection dataset which excluded weak or unobserved data ($F_{obs} > 75.0$) (Tables 1,3). Of the 2243 reflections eliminated 80 % were located in the 2.5–2.0 Å resolution range. In addition, occupancies for partially occupied residues (alternative side chain conformers and sugar residues) were fixed at previously established values and not refined in this last round. This reduced the R-factor further to 16.0 % , but did not change the model nor the electron density map significantly.

Electron density maps ($2F_{obs}$–F_{calc} and F_{obs}–F_{calc}) were calculated after each refinement round using phase information from the refined atomic model. Coefficients were calculated as suggested by Read [56] using the CCP4 program package (SFALL, SIGMAA, FFT).

The modeling program O [57] was utilized for inspection of the electron density map and for interactive modeling at the various stages during refinement. Water molecules were selected from electron density peaks > 3σ observed in F_{obs}–F_{calc} electron density maps. As acceptance criteria distance limits of 1.5 Å and 5 Å were applied utilizing the program WATPEAK [50]. Each site was checked graphically with program O and the solvent model was inspected and improved constantly during rounds 8–13. The final model included 55 water molecules, 24 of which with B-factors > 50 Å^2.

Refinement of bound Man-5

After the first two refinement rounds unambiguous electron density was observed for three mannose residues of the complexed mannopentaose at the binding sites in the CRD1 and CRD3 domains. The electron density at the third binding location (CRD2 site) sufficed only for a single mannose residue bound in the specificity pocket (M1). At the CRD1 site continuous difference density (F_o–F_c), centered on a crystallographic twofold axis (c-dimension) between neighboring tetramers, suggested the presence of either a cross-link or a close sugar–sugar lattice contact. The first hypothesis was rejected initially as Man-5 lacks internal twofold symmetry. Thus, in rounds 3–5 Man-5 was built into the difference density such that the central mannose (M2) occupies the conserved monosaccharide site and the terminal residues M1 and M3 are oriented close to the protein surface analogously to the binding mode observed at the CRD3 binding region. In this binding mode the 1-linked disaccharide (M4, M5) is directed away from the protein surface and was expected to occupy at least part of the bridging density. However, in view of weak electron density for M1 and M3, high B-factors and serious steric problems, this binding mode was rejected. Moreover, during refinement strong electron density remained centered on the twofold axis. Thus, after round 5 it was concluded that the density could accommodate only three covalently linked mannose residues, suggesting a cross-linking binding mode between twofold related dimers. The outer tri-Man arm (M1–M2–M3) could be fitted to the density with either M1 or M3 positioned in the CRD1 binding site of the reference molecule. Since the density is centered on a twofold symmetry axis, which relates the inter-linked dimers, the second orientation for M1–M2–M3 is generated as a symmetry mate. This interpretation of the map accounted best for the bridging difference electron density broadened at the center to accommodate the off-set positions of the central mannose (M2) in the two binding orientations (Fig. 2a). This model was successfully refined disabling packing interactions. The occupancy of the tri-Man molecule as a group refined to values close to 0.5, as expected, and was fixed in round 13. Weak density for the two remaining mannose residues (M4 and M5) connected to the central M2 was present in the initial map based on molecular replacement phases. M4 was included in the refinement until round 12 despite its low occupancy (0.2), high B-factors and slowly disappearing density. It was removed from the model in the final round 13.

Modeling of Man-5 at the CRD2 binding site initially included only one mannose residue, which was fitted to density in the conserved monosaccharide site. In the last two refinement rounds a second mannose residue (M1) was added with some uncertainty and the M1-M2 dissaccharide refined in a binding mode analogous to that observed in the CRD3 domain for M1–M2.

In the CRD3 binding site the complete 3,6 tri-Man structure was fitted without any difficulties into good density and refined satisfactorily with an occupancy near 0.7. Here again, the C1-linked dimannoside (M4 and M5) was initially visible, but electron density gradually vanished in successive refinement rounds. Thus, these two residues were also removed from the model in later cycles.

Accession numbers

The atomic coordinates for the GNA–Man-5 complex have been deposited at the Brookhaven data bank [58] under the code name 1JPC.

Acknowledgements

We thank the staff of the data collection facility in the Department of Medicinal Chemistry for their help with data measurement on the imaging plate detector, and Karen Kenny for excellent technical help with the illustrations. In addition, we are indebted to Dr Irwin Goldstein for initially providing us with protein material and supporting our efforts with great interest. This research was supported by a National Institutes of Health Grant, AI-17992.

References

1. Goldstein, I.J. & Poretz, R.D. (1986). Isolation, physiochemical characterization, and carbohydrate-binding specificity of lectins. In *The Lectins: Properties, Function and Application in Biology and Medicine*. (I.E. Liener, N. Sharon and I.J. Goldstein, eds) pp. 43–247, Academic Press, Orlando, USA.
2. Lis, H. & Sharon, N. (1981). Lectins in higher plants. *The Biochemistry of Plants* 6, 371–447.
3. Sharon, N. & Lis, H. (1989). Lectins as cell recognition molecules. *Science* 246, 227–234.
4. Nicolson, G.L. (1974) The interaction of lectins with animal cells. *Int. Rev. Cytology* 39, 89–190.
5. Etzler, M.E. (1985). Plant lectins: molecules and biological aspects. *Annu. Rev. Plant Physiol*. 36, 209–234.
6. Drickamer, K. & Taylor, M.E. (1993). Biology of animal lectins. *Annu. Rev. Cell Biol* 9, 237–264.
7. Nicolson, G.L. (1976). Transmembrane control of the receptors on normal and tumor cells. *Biochim. Biophys. Acta* 457, 57–108.
8. Van Damme, E.J.M., Allen, A.K. & Peumans, W.J. (1987) Isolation and characterization of a lectin with exclusive specificity towards mannose from snowdrop (*Galanthus nivalis*) bulbs. *FEBS Lett*. 215, 140–144.
9. Hilder, V.A., *et al*., & Boulder, D. (1994). Expression of snowdrop lectin in transgenic tobacco plants results in added protection against aphids. *Transgenic Res*. 4, 18–25.
10. Shibuya, N., Goldstein, I.J., Van Damme, E.J.M. & Peumans, W.J. (1988). Binding properties of a mannose-specific lectin from snowdrop (*Galanthus nivalis*) bulb. *J. Biol. Chem*. 263, 728–734.
11. Gilljam, G. (1993). Envelope glycoproteins of HIV-1, HIV-2, and SIV purified with *Galanthus nivalis* agglutinin induce strong immune responses. *Aids Research and Human Retroviruses* 9, 431–438.
12. Balzarini, J., Schols, D., Neyts, J., Van Damme, E.J.M., Peumans, W. & DeClerq, E. (1991). α-(1,3) and α-(1,6)-D-mannose-specific plant lectins are markedly inhibitory to human immunodeficiency virus and cytomegalovirus infections *in vitro*. *Antimicrobial Agents and Chemotherapy* 35, 410–416.
13. Kaku, H. & Goldstein, I.J. (1991). Interactions of five D-mannose-specific lectins with a series of synthetic branched trisaccharides. *Carbohydrate Res.* 213, 109–116.
14. Kaku, H. & Goldstein, I.J. (1992). Interaction of linear manno-oligosaccharides with three mannose-specific bulb lectins. Comparison with mannose/glucose binding lectins. *Carbohydrate Res.* 229, 337–346.
15. Chervenak, M.C. & Toone E.J. (1995). Calorimetric analysis of the binding of lectins with overlapping carbohydrate-binding ligand specificities. *Biochemistry* 34, 5685–5696.
16. Hester, G., Kaku, H., Goldstein, I. J. & Wright, C.S. (1995). Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family. *Nat. Struct. Biol.* 2, 472–479.
17. Sharon N. & Lis H. (1990). Legume lectins - a large family of homologous proteins. *FASEB J.* 4, 3198–3208.
18. Rini, J.M. Lectin Structure. (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 551–577.
19. Van Damme, E.J.M., *et al*., & Peumans, W.J. (1991). Biosynthesis, primary structure and molecular cloning of snowdrop (*Galanthus nivalis L*.) lectin. *Eur. J. Biochem.* 202, 23–30.
20. Weis, W.I., Drickamer, K. & Hendrickson, W.A. (1992). Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* 360, 127–134.
21. Hester, G. & Wright, C.S. The mannose-specific bulb lectin from *Galanthus nivalis* (snowdrop) binds mono- and dimannosides at distinct sites. structure analysis of refined complexes at 2.3 Å and 3.0 Å Resolution *J. Mol. Biol.* 262, 516–531 (1996).
22. Musacchio, A., Saraste, M. & Wilmanns, M. (1995) High-resolution crystal structure of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat. Struct. Biol*. 1, 546–551.
23. Lin., G., Bode, W., Huber, R., Chi, C. & Engh, R.A. (1993). The 0.25 nm X-ray structure of the Bowman-Birk-type inhibitor from mung bean in ternary complex with porcine trypsin. *Eur. J. Biochem*. 212, 549–555.
24. Woods, R. J., Edge, C.J. & Dwek, R.A. (1994). Protein surface oligosaccharides and protein function. *Nat. Struct. Biol*. 1, 499–501.
25. Wooten, E. W., Bazzo, R., Edge, C.J., Zamze, S., Dwek, R.A. & Rademacher, T.W. (1990). Primary sequence dependence of conformation in oligomannose oligosaccharides. *Eur. Biophys. J.* 18, 139–148.
26. Imberty, A., Gerber, S., Tran, V. & Perez, S. (1990). Data bank of three-dimensional structures of disaccharides. A tool to build 3-D structure of oligosaccharides. *Glycoconjugate J.* 7, 27–54.
27. Bourne, Y., *et al*., & Cambillau, C. (1994). Cross-linking of mammalian lectin (galectin 1) by complex biantennary saccharides. *Nat. Struct. Biol.* 1, 863–870.
28. Van Damme, E.J.M., Goldstein, I.J. & Peumans, W.J. (1991). A comparative study of mannose-binding lectins from the *Amaryllidaceae* and *Alliaceae*. *Phytochemistry* 30, 509–514.
29. Kaku, H., Goldstein, I.J. & Oscarson, S. (1991). Interaction of five D-mannose specific lectins with a series of synthetic branched trisaccharides. *Carbohydrate Res*. 213, 109–116.
30. Sharon, N. (1993). Lectin-carbohydrate complexes of plants and animals. An atomic view. *Trends in Biochem. Sci*. 18, 221–226.
31. 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.
32. Ng, K.K.-S., Drickamer, K. & Weis, W.I. (1996). Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. *J. Biol. Chem.* 271, 663–667.
33. Bourne Y., *et al*., & Cambillau C. (1994). Structure of a legume lectin complexed with the human lactotransferrin N2 fragment, and with an isolated biantennary glycopeptide: role of the fucose moiety. *Structure* 2, 209–219.
34. Naismith, J. H. & Field, R. A. (1996). Structural basis of trimannoside recognition by concanavalin A. *J. Biol. Chem*. 271, 972–976.
35. Wright, C.S. & Kellogg G. (1996). Differences in hydropathic properties of ligand binding at four independent sites in wheat germ agglutinin-oligosaccharide complexes. *Protein Science* 5, 1466–1476.
36. Wright, C.S. (1992). Crystal structure of a wheat germ agglutinin/glycophorin-sialoglycopeptide receptor complex. *J. Biol. Chem.* 267, 14345–14352.
37. Bhattacharyya, L., Fant J., Lonn, H. & Brewer, C.F. (1990). Binding and precipitating activities of lotus tetragonolobus isolectins with L-fucosyl oligosaccharides. Formation of unique homogeneous cross-linked lattices observed by electron microscopy. *Biochemistry* 29, 7523–7530.
38. Bhattacharyya, L., & Brewer, C.F. (1989). Interactions of concanavalin A with asparagine linked glycopeptides. *Eur. J. Biochem*. 178, 721–726.
39. Bhattacharyya, L., Haraldsson, M., & Brewer, C.F. (1988). Precipitation of Gal-specific lectins by complex-type oligosaccharides and glycopeptides. Studies with lectins from *Ricinus communis* (Agglutinin 1), *Erythrina indica, Erythrina arborescens, Abrus precatorius* (Agglutinin) and *Glycine max* (soybean). *Biochemistry* 27, 11034–11041.
40. 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 biantennary analog of the blood group I carbohydrate antigen. *Biochemistry* 34, 4933–4942.

41. Liao, D-I, Kapadia, G., Ahmed, H., Vasta, G.R. & Herzberg, O. (1994). Structure of S-lectin, a developmentally regulated vertebrate β-galactoside-binding protein. *Proc. Natl. Acad. Sci. USA* 91, 1428–1432.
42. Lobsanov, Y.D., Gitt, M.A., Leffler, H., Barondes, S.H. & Rini J.M. (1993). X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II in complex with lactose at 2.9 Å resolution. *J. Biol. Chem.* 268, 27034–27038.
43. Lis, H. & Sharon, N. Lectins. (1977). Their chemistry and application to immunology. *The Antigens* 4, 463–529.
44. Gupta, D., Arango, R., Sharon, N. & Brewer, C.F. (1994). Differences in the cross-linking activities of native and recombinant *Erythrina corallodendron* lectin with asialofetuin. Evidence for carbohydrate-carbohydrate interactions in lectin-glycoprotein complexes. *Biochemistry* 33, 2503–2508.
45. Mizuochi, T., Spellman, M.W., Larkin, M., Solomon, J., Basa, L.J. & Feizi, T. (1988). Carbohydrate structures of the human-immunodeficiency-virus (HIV) recombinant envelope glycoprotein gp120 produced in Chinese hamster ovary cells. *Biochem. J.* 254, 599–603.
46. Favero, J. (1994). Lectins in aids research. *Glycobiology* 4, 387–396.
47. Dracic, T., *et al*., & Paxton W.A. (1996). HIV-1 entry into CD^{4+} cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 381, 667–673.
48. Drickamer, K. (1995). Multiplicity of lectin-carbohydrate interactions. *Nat. Struct. Biol.* 2, 437–439.
49. Wright, C.S., Kaku, H. & Goldstein, I.J. (1990). Crystallization and preliminary X-ray diffraction results of snowdrop (*Galanthus nivalis*) lectin. *J. Biol. Chem.* 265, 1676–1677.
50. Leslie, A.G.W. (1987). Profile fitting. In *Proceedings of the CCP4 study weekend*. (Helliwell, J.R., Machin, P.A. & Papiz, M.Z., eds.), pp. 39–50, SERC Daresbury Laboratory, Warrington, UK.
51. Collaborative computational project, No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* 50, 760–763.
52. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* 50, 157–163.
53. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* 235, 458–460.
54. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A* 47, 392–400.
55. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472–475.
56. Read, R.J. (1986). Improved fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst. A* 42, 140–149.
57. 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 Cryst. A* 47, 110–119.
58. Bernstein, F.C., *et al*., & Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structure. *J. Mol. Biol.* 112, 535–542
59. Luzzati, P.V. (1952). Traitements statistique des erreurs dans la determination des structures cristallines. *Acta Cryst.* 5, 802–810.