JOURNAL OF STRUCTURAL BIOLOGY **116,** 345–355 (1996)
ARTICLE NO. 0052

The Crystal Structure of the Complexes of Concanavalin A with 4′-Nitrophenyl-α-D-mannopyranoside and 4′-Nitrophenyl-α-D-glucopyranoside

PANAGIOTIS N. KANELLOPOULOS

European Molecular Biology Laboratory, D-69012 Heidelberg, Germany

KYRIAKI PAVLOU

Department of Biochemistry, Cell and Molecular Biology and Genetics, University of Athens, Panepistimiopolis, Athens 157 01, Greece

ANASTASSIS PERRAKIS

European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

BOGOS AGIANIAN

Department of Biochemistry, Cell and Molecular Biology and Genetics, University of Athens, Panepistimiopolis, Athens 157 01, Greece

CONSTANTIN E. VORGIAS

European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

COSTAS MAVROMMATIS AND MARIA SOUFI

Department of Biochemistry, Cell and Molecular Biology and Genetics, University of Athens, Panepistimiopolis, Athens 157 01, Greece

PAUL A. TUCKER[1]

European Molecular Biology Laboratory, D-69012 Heidelberg, Germany

AND

STAVROS J. HAMODRAKAS

Department of Biochemistry, Cell and Molecular Biology and Genetics, University of Athens, Panepistimiopolis, Athens 157 01, Greece

Received November 9, 1995, and in revised form January 10, 1996

Concanavalin A (Con A) is the best-known plant lectin and has important *in vitro* biological activities arising from its specific saccharide-binding ability. Its exact biological role still remains unknown. The complexes of Con A with 4′-nitrophenyl-α-D-mannopyranoside (α-PNM) and 4′-nitrophenyl-α-D-glucopyranoside (α-PNG) have been crystallized in space group $P2_12_12$ with cell dimensions *a* = 135.19 Å, *b* = 155.38 Å, *c* = 71.25 Å and *a* = 134.66 Å, *b* = 155.67 Å, and *c* = 71.42 Å, respectively. X-ray diffraction intensities to 2.75 Å for the α-PNM and to 3.0 Å resolution for the α-PNG complex have been collected. The structures of the complexes were solved by molecular replacement and refined by simulated annealing methods to crystallographic R-factor values of 0.185/0.186 and free-R-factor values of 0.260/0.274, respectively. In both structures, the asymmetric unit contains four molecules arranged as a tetramer, with approximate 222 symmetry. A saccharide molecule is bound in the sugar-binding site near the surface of each monomer. The nonsugar (aglycon) portion of the

[1] To whom correspondence should be addressed.

1047-8477/96 $18.00

compounds used helps to identify the exact orientation of the saccharide in the sugar-binding pocket and is involved in major interactions between tetramers. The hydrogen bonding network in the region of the binding site has been analyzed, and only minor differences with the previously reported Con A–methyl-α-D-mannopyranoside complex structure have been observed. Structural differences that may contribute to the slight preference of the lectin for mannosides over glucosides are discussed. Calculations indicate a negative electrostatic surface potential for the saccharide binding site of Con A, which may be important for its biological activity. It is also shown in detail how a particular class of hydrophobic ligands interact with one of the three so-called characteristic hydrophobic sites of the lectins.

INTRODUCTION

Concanavalin A (Con A) is a representative member of the lectin class of plant proteins which are valuable tools for the investigation of cell surface structural organization and dynamics (Goldstein and Poretz, 1986; Lis and Sharon, 1986; Bittiger and Schnebli, 1976).

Con A generally binds to saccharides containing α-D-mannose or α-D-glucose residues but it may also recognize oligosaccharide sequences lacking these units. Con A has specific biological activities which depend on its binding to cell surface receptors. It preferentially agglutinates certain cells transformed by oncogenic viruses more than their untransformed counterparts, inhibits growth of malignant cells in animals, and exhibits mitogenic activity. It has also been used in studies on the number and mobility of cell-surface receptors associated with cell–cell interactions. Although the exact biological role of Con A remains unknown, its specific saccharide-binding properties make it an ideal object for the study of protein–saccharide interactions (Goldstein and Poretz, 1986; Lis and Sharon, 1986; Bittiger and Schnebli, 1976).

In solution it forms monomers, dimers, or tetramers, depending on pH and temperature conditions (Agrawal and Goldstein, 1986). At physiological pH, Con A exists as a tetramer. Each subunit consists of 237 residues and contains two metal sites: one (S1) that binds transition metal ions and another (S2) that binds preferentially calcium ions. Evidence suggests that both sites must be occupied for saccharide binding to occur (Goldstein and Poretz, 1986).

Con A was first isolated from the Jack bean (*Canavalia ensiformis*) and crystallized almost 60 years ago (Sumner and Howell, 1936). Its structure was initially determined by two independent groups (Edelman *et al.,* 1972; Hardman and Ainsworth, 1972) and subsequently refined to 2.0 Å (Becker *et al.,* 1976) and 1.75 Å resolution (Hardman *et al.,* 1982). The main structural feature of the protein is a β-sandwich, formed by two anti-parallel β-pleated sheets. Since most biological properties of Con A depend on its saccharide-binding ability, it is important to find the modes of binding of carbohydrates to this protein. Attempts to solve the structure of a complex of Con A with a saccharide at high resolution were unsuccessful for several years. The saccharide-binding site was independently discovered both by Becker *et al.* (1976) and by Hardman and Ainsworth (1976) and more recently, the crystal structure of Con A has been reinvestigated. The three-dimensional structure has been elucidated by X-ray crystallographic studies at 1.95 Å resolution with the S1 site occupied by a Mn ion (Weisberger and Helliwell, 1993), at 2.0 Å with the S1 site occupied by a Cd or Ni ion (Naismith *et al.,* 1993; Emmerich *et al.,* 1994), and at 1.6 Å with the S1 site occupied by a Co ion (Emmerich *et al.,* 1994). Helliwell and co-workers have also solved the structure of a Con A–methyl-α-D-mannopyranoside complex and described in detail the Con A–mannoside interaction (Derewenda *et al.,* 1989; Naismith *et al.,* 1994).

Our studies are focused on the modes of binding to Con A to a series of saccharides which consist of a mannoside or a glucoside group (saccharide residue) bonded to a second hydrophobic group (aglycon residue). We are trying to examine the role of the aglycon prosthetic group in binding. Experimentally determined binding constants (Farina and Wilkins, 1980, and references therein; Troganis and Stassinopoulou, 1994) show aglycon-dependent changes, for example, 4-methylumbelliferyl-α-D-mannopyranoside binds more strongly than 4′-nitrophenyl-α-D-mannopyranoside, and we wish to investigate whether there are corresponding structural changes. Suitable crystals of Con A–saccharide complexes of good quality have been produced for a series of such compounds, which diffract at medium resolution, and X-ray diffraction data have been collected for several complexes. In this report we present the crystal structure of complexes of ConA with 4′-nitrophenyl-αD-mannopyranoside (α-PNM) and 4′-nitrophenyl-α-D-glucopyranoside (α-PNG) at 2.75 and 3.0 Å resolution, respectively.

MATERIALS AND METHODS

Purification and Crystallization

Concanavalin A was obtained from Sigma. A novel method for a purification of the protein has been developed, which apparently removes fragments of Con A presumably resulting from incomplete posttranslational modification (Sharon and Lis, 1990). This step was essential in obtaining good quality crystals of Con A complexed with various saccharides. Con A was initially separated from the fragments by applying the protein on a DEAE-Sepharose column (Pharmacia) in a 50 m*M* Tris (pH 8.0) solution. The elution was performed using a 50 m*M* Tris (pH 8.0),

200 m*M* NaCl solution, and under these conditions the protein appears to be fragment-free as judged by SDS–PAGE. This protein solution was subsequently made 1 *M* in ammonium sulfate concentration and applied to an octyl-Sepharose CL-4B column (Sigma). The protein was eluted using a 20 m*M* Tris (pH 8.0) buffer and then was dialyzed against a 20 m*M* Tris (pH 8.0), 20 m*M* $MnCl_2$, 20 m*M* $CaCl_2$ solution to ensure enrichment of the protein with the metal ions. Finally, the protein was concentrated to 70 mg/ml. This Con A solution was used for cocrystallization experiments with the saccharides α-PNM and α-PNG obtained commercially from Sigma. A 10× molar excess of each of the saccharides was used to ensure high occupancy in crystals of the complex. The crystals were grown using the hanging drop vapor diffusion method. The reservoir solution contained a 10 m*M* Tris, pH 8.5, 1 *M* $(NH_4)_2SO_4$ solution. The drop was composed of a 1:1 mixture of the well solution with the 70 mg/ml ConA solution and a 10× molar excess of saccharide. The final volume of the drop was approximately 10 μl. The crystals are orthogonal prisms, with maximum dimensions 0.8 × 0.3 × 0.3 mm and grow after 2 to 3 weeks. They are orthorhombic and belong to space group $P2_12_12$. The unit cell dimensions for the α-PNM complex crystals are a = 135.19 Å, b = 155.38 Å, and c = 71.25 Å and for the α-PNG complex a = 134.66 Å, b = 155.67 Å, and c = 71.42 Å. The asymmetric unit for both crystals contains four molecules and the solvent content is 65.7% by volume.

Data Collection and Structure Solution

X-ray diffraction data from a single crystal of the Con A–α-PNM complex were collected at room temperature to 2.75 Å on a 30-cm MAR Research image plate detector. Pyrolytic graphite monochromatised CuK_α radiation (λ = 1.5418 Å) was used, produced by a rotating anode generator running at 40 kV, 90 mA. The crystals were kept at room temperature during data collection. The rotation method was employed and 1° oscillation data frames were collected. The X-ray data were processed using the program XDS (Kabsch, 1988). Under identical conditions a data set for a single Con A–α-PNG complex crystal was collected to 3.0 Å. The data were processed, as for the α-PNM complex data set. A summary of the data collection statistics is given in Table I.

The 2.75-Å Con A–α-PNM structure was solved by molecular replacement, using the 2.9-Å model of the concanavalin A–methyl-α-D-mannopyranoside complex (Protein Data Bank entry 4CNA) available at the time of calculations (Derewenda *et al.,* 1989). The methyl-α-D- mannopyranoside and all water molecules were omitted from the initial search model. All rotational and translational searches were carried out with the program AMoRe (Navaza, 1994). The fast rotation function calculation was used to determine the orientation of the Con A tetramer in the asymmetric unit. Using reflections in the resolution range 15 to 3.5 Å, and an integration radius of 45 Å, a set of four dominant peaks appeared. The highest of the four peaks was 18.8 σ and the remaining three had slightly lower values, while the next peak in the list was only 3.5 σ. The set of four peaks are related by the approximate 222 symmetry of the Con A tetramer; therefore, only the highest peak was chosen. The translation function calculations were performed using reflections from 15 to 2.75 Å. A single peak, corresponding to a crystallographic R factor of 0.349, was identified for the rotation function solution. The rigid body refinement procedure, implemented in AMoRe, was applied. An initial crystallographic R factor of 0.339 was calculated in the resolution range 15–2.75 Å, for a tetramer per asymmetric unit.

TABLE I
Summary of Data Collection

Data set		Con A–α-PNM	Con A–α-PNG
Resolution range (Å)		48.0–2.75	48.0–3.0
Observed reflections		181088	116891
Unique reflections		38837	30575
Percentage of unique data measured		97.4	99.1
R_{sym}(%)[a]	48.0–10.0Å	3.8	3.6
	10.0–6.0 Å	5.2	6.2
	6.0–4.5 Å	6.4	8.4
	4.5–3.8 Å	10.5	15.8
	3.8–3.3 Å	18.9	26.4
	3.3–3.0 Å	36.9	43.9
	3.0–2.75Å	58.4	
	Total	12.5	14.3

[a] $R_{sym} = \sum_h \sum_i |F_i^2(h) - \langle F^2(h) \rangle| / \sum_h \sum_i F_i^2(h)$

Refinement and Model Building of the 2.75-Å Con A–α-PNM Complex Structure

At this stage, the amino acid sequence was replaced by the most reliable sequence currently available (Min *et al.,* 1992). The model was further refined using the program X-PLOR (Brunger, 1992a). During all stages of refinement and model building, the four non-crystallographically related molecules in the asymmetric unit were treated independently. Initially the model was subjected to rigid body refinement. Each of the monomers was treated as a separate rigid group and the refinement was performed for all data between 6.0 and 2.75 Å. After this stage, $2F_o$-F_c and F_o-F_c electron density maps were calculated and electron density for the saccharide was clearly identified for all four subunits of the tetramer. Electron density for the saccharide was present for both maps even at a contouring level of five times the rms electron density. The model for α-PNM was built into the density, using the program O (Jones *et al.,* 1991) and the accurate geometry obtained from the crystal structure (Agianian *et al.,* 1995). The model was then subjected to further refinement with X-PLOR. The topology and parameter files were adjusted to accommodate the nonproline *cis* peptide (Ala207–Asp208) and the α-PNM in each monomer. No restraints were placed on metal–ligand distances and no charge was attributed to the metal ions. A step of refinement included: simulated annealing refinement with a starting temperature of 3000 K, 120 cycles of conjugate gradient minimisation, 15 cycles of overall B-factor refinement, 20 cycles of restrained individual B-factor refinement, and, finally, a further 160 cycles of conjugate gradient minimization (positional refinement). No σ cutoff was applied and 90% of the data in the resolution range 6.0 to 2.75 Å were used. In order to cross-validate in reciprocal space, the free-R value (Brunger, 1992b) was monitored during all steps of refinement using the remaining 10% fraction of the data. Model improvement was then performed for the polypeptide chain and saccharide molecules. Water molecules in the coordination sphere of Mn and Ca ions were added to the model as well as water molecules in the neighborhood of the saccharides. A total of 32 water molecules were included and fitted in peaks of the F_o-F_c electron density map higher than 3.0× the rms electron density level. At this resolution, no further attempt to find water molecules was regarded as justifiable. The final model was obtained after further refinement. The refinement statistics are shown in Table II.

Structure Determination and Refinement of the Con A–α-PNG Structure

The 2.75-Å structure of the Con A–α-PNM complex was used as a starting model for the Con A–α-PNG structure. From the initial coordinate set, the α-PNM molecule was omitted and only the water molecules in the coordination sphere of the metal ions were

included. The $2F_o$-F_c and F_o-F_c electron density maps were calculated and the model of α-PNG was fitted into clearly defined electron density in the saccharide-binding sites of all subunits of the tetramer. The model of α-PNG was derived from the accurate α-PNM geometry after transforming the mannose ring into a glucose ring. The model was then subjected to full refinement with X-PLOR, in a similar way to the Con A–α-PNM complex crystal structure, including simulated annealing, individual B-factor refinement, and conjugate gradient minimization. The final structure has a crystallographic R factor of 0.186 and a free-R factor of 0.274 in the resolution range 6.0 to 3.0 Å. Sixteen water molecules, identified in the F_o-F_c map, in the vicinity of the metal ions were included in the model.

RESULTS AND DISCUSSION

Accuracy and Quality of the Models

The chemical formula of α-PNM is shown in Fig. 1 (left) for clarity. The basic topology of the Con A–α-PNM, Con A–α-PNG tetramers is unchanged from previously reported Con A structures (e.g., Naismith *et al.,* 1994) and will not be discussed here. Each tetramer contains two dimers: one dimer is formed by subunits A and B and the other by subunits C and D (Fig. 1, right). The secondary structure elements and folding architecture of each monomer have been discussed in detail and in both Con A–α-PNM and Con A–α-PNG there are no significant deviations from the reported structures of Con A (e.g., Naismith *et al.,* 1993, 1994) and other legume lectins (Reeke and Becker, 1988; Delbaere *et al.,* 1989; Bourne *et al.,* 1990; Shaanan *et al.,* 1991; Rini *et al.,* 1993; Loris *et al.,* 1994; Bourne *et al.,* 1994). Dimer and tetramer formation and stabilization have also been analyzed in detail (Naismith *et al.,* 1993, 1994) and will not be discussed further. Suffice it to say that the tetramer in both complexes exhibits approximate 222 symmetry.

A representative portion of the 2.75 Å, final $2F_o$-F_c electron density map for the Con A–α-PNM complex is shown in Fig. 2. This portion of the map shows the saccharide-binding site of the B subunit in the Con A–α-PNM complex and how the aglycon part of the bound saccharide interacts in a characteristic way with the symmetry-related saccharide of the A subunit by a stacking of planar aromatic groups.

At this point, it is worth mentioning that the aglycon portion of the saccharides of the D subunit, in both crystal structures, appears to be less well ordered than that in the other three subunits and only the electron density for this part of the molecules is not very well defined. We cannot offer a satisfactory explanation for this disorder other than a rotational disorder of the aglycon part of subunit D. An analogous anomalous behavior for the saccharide of subunit C has been observed in the Con A–methyl-α-D-mannopyranoside complex at 2.0 Å resolution and there also, no explanation is apparent (Naismith *et al.,* 1994).

In the Con A–α-PNM complex, the refined model exhibits good fit in the $2F_o$-F_c map calculated with reflections from 15.0 to 2.75 Å. The per residue real space fit correlation coefficient plot for all four subunits is shown in Fig. 3. The coefficients have been calculated for complete residues as implemented in O (Jones *et al.,* 1991). The fit pattern is similar between the four independently refined subunits. Exceptions from good fit are observed in loop regions. The crystallographic R factor for the final model is 0.185 in the resolution range 6.0–2.75 Å, and the free-R factor is 0.26.

The structure has good stereochemistry with rms deviations from ideality of 0.009 Å for bond lengths, 1.6° for bond angles, and 27° for dihedral angles. The maximum expected error in atomic positions is estimated to be between 0.30 and 0.35 Å, at 2.75 Å resolution, based on a Luzatti plot (Luzatti, 1952). No residues lie outside allowed regions of the Ramachandran plot as calculated with the program PROCHECK (Laskowski *et al.,* 1993). The pattern of temperature factors (data not shown) are consistent for main chain and side chain atoms in the same subunit and also between the different subunits. They are also consistent with the RS-fit patterns of Fig. 3. A rather interesting observation is the poor electron density in the region of His 121, in all four subunits, which lies in the center of the hole formed by the Con A tetramer. This is close to the region where the unusual posttranslational modification which involves a ligation at residues 118 and 119 occurs (Bowles, 1990). The quality of the Con A–α-PNG structure is similar to the α-PNM complex structure. It has been assessed using all methods described above (data not shown).

TABLE II
Summary of Refinement of Con A–α-PNM Complex

Stage	Resolution range (Å)	R factor	R-free
Molecular replacement solution (AMoRe)	15.0–2.75	0.339	—
Rigid body refinement (XPLOR)	6.0–2.75	0.305	0.335
Addition of four saccharide molecules and refinement (XPLOR)	6.0–2.75	0.194	0.269
Addition of 32 water molecules and refinement (XPLOR)	6.0–2.75	0.185	0.260

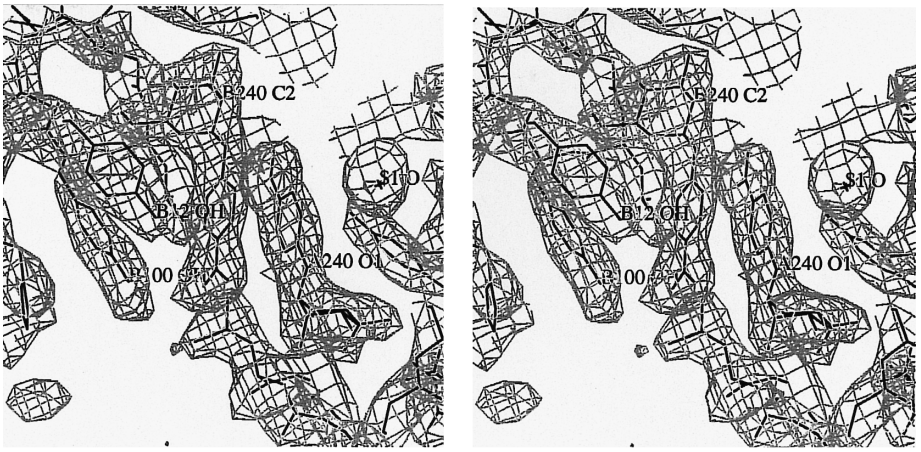


FIG. 2. A stereo pair showing the final $2F_O$-F_C electron density in the region of a sugar-binding site. The region shows where two tetramers contact each other in crystal lattice. The contact is between an A subunit and a B subunit. Contouring is at 1.0× the rms value of the electron density.

Metal Binding Sites

Each concanavalin A monomer contains a tightly bound manganese and calcium ion in the vicinity of the saccharide binding site. The two sites are denoted, respectively, S1 and S2. Each of the metal ions is coordinated by four amino acid side chains and two water ligands (data not shown). In all four subunits the coordination of the manganese ion is octahedral, whereas the calcium ion has pseudo-octahedral geometry, with Asp10 binding in a bidentate manner capping the sixth vertex of the octahedron. In the case of the calcium ion one of the water ligands forms a bridge between the metal and the main chain carbonyl of Asp208, thus stabilizing the unusual Ala207–Asp208 *cis*-peptide bond that is conserved in all known legume lectin crystal structures. The structure of both sites shows profound similarities with those determined in other legume lectins, which further confirms that this region is very well conserved within the legume lectin family.

The Saccharide Binding Site

Each monomer has a saccharide binding site on the outer surface of the tetramer as shown schematically in Fig. 1 (right).

An examination of the electrostatic surface potential of each monomer (calculated with GRASP (Nicholls *et al.,* 1991; Nicholls, 1993); Fig. 4) reveals that the saccharide binding site exhibits a negative electrostatic potential. This property of the saccharide binding site of Con A has not previously been noticed. It might be related to interactions with other, perhaps positively charged, molecules but is, in any case, likely to be relevant to the biological activity of Con A.

A view of the surface distribution of certain classes of residues with defined properties around the saccharide binding site reveals unexpected regularities (Fig. 4, inset). Thus, at the "top" of the binding site, toward Tyr12 and Tyr100, residues with aromatic side chains (Tyr, Phe, His, Trp) are clustered. On the "left," in the vicinity of Leu99, hydrophobic residues (other than aromatic) reside and on the "right," close to Arg228, charged residues. At the "bottom," close to Thr226, only polar residues are found. Therefore, the only regions close to the binding site with properties suitable to interact with hydrophobic ligands appear to be the "top" and "left" sites.

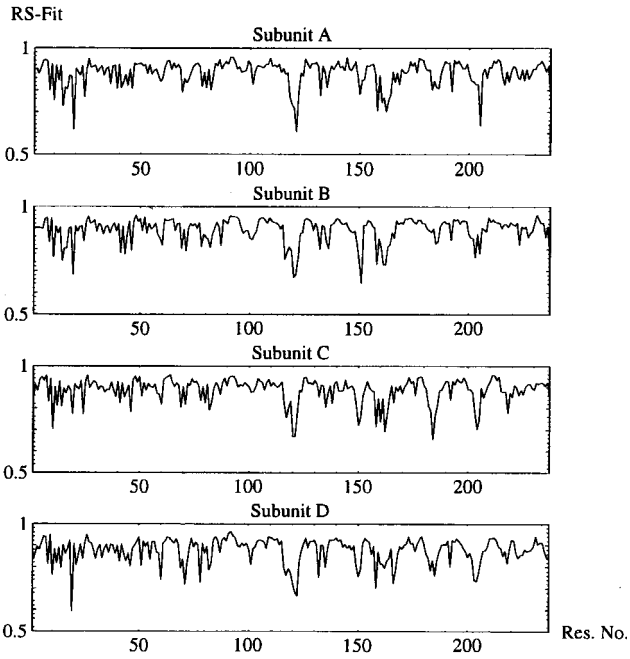
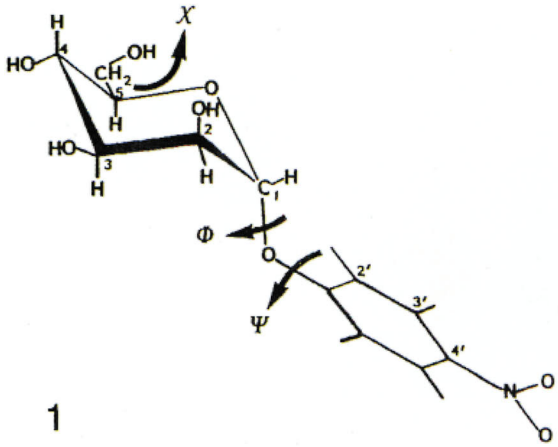
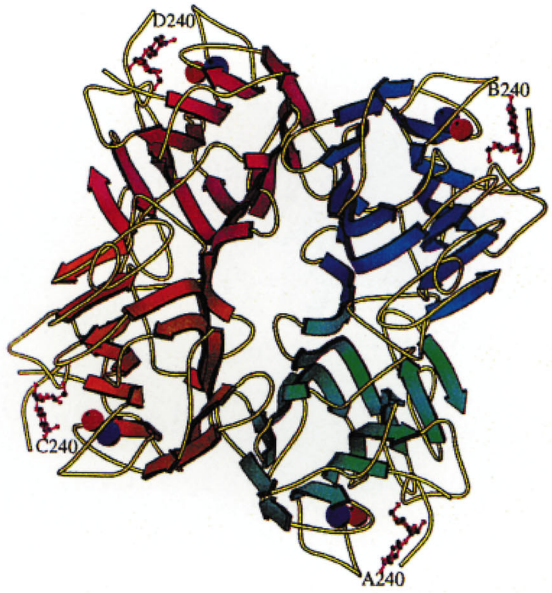


FIG. 3. Real space correlation coefficients as function of residue number for subunits A through D, with the final coordinate set and the corresponding $2F_O$-F_C electron density map. The program used for the calculations was "O" (Jones *et al.,* 1991).





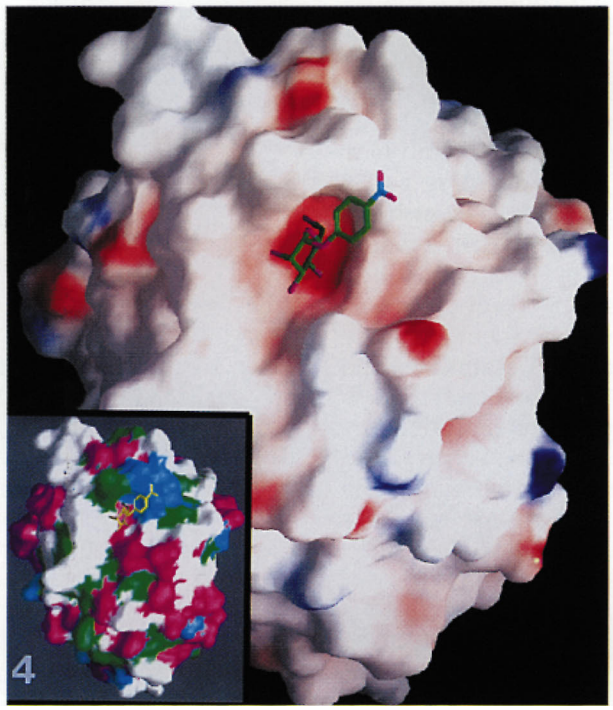
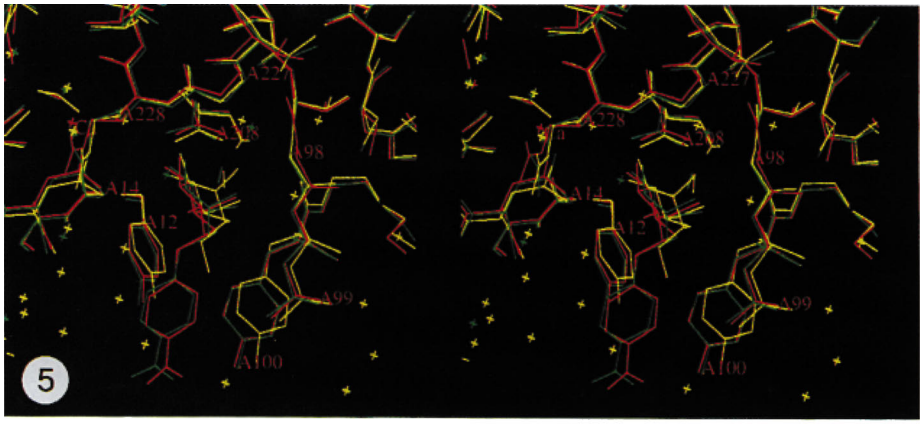


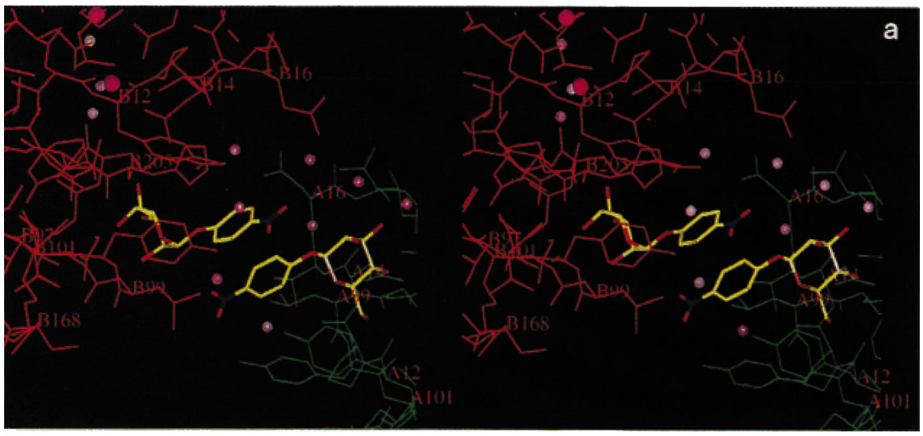
FIG. 1. (left) Chemical formula of 4′-nitrophenyl-α-D-mannopyranoside (α-PNM) (right) A ribbon representation (Kraulis, 1991) of the ConA–α-PNM tetramer viewed approximately along one of the noncrystallographic twofold axes and showing the subunit labeling. The sugar moieties are in ball and stick representation.

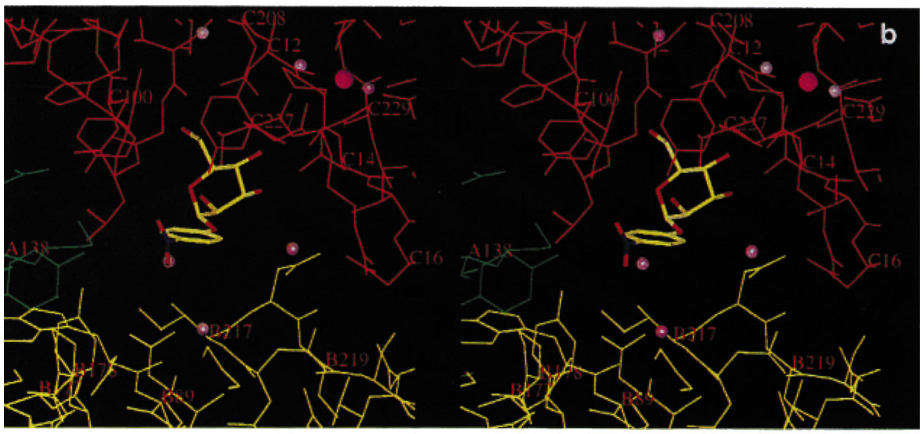
FIG. 4. A representation of the electrostatic surface potential as calculated for one monomer by the program DELPHI and displayed using GRASP (Nicholls *et al.,* 1991; Nicholls, 1993). The calculations were performed in water with the default charges file. Electrostatic potential is shown from –30kT (red) to +30kT (blue). The bound saccharide is shown as a stick model. The inset shows the same view with cyan for aromatic residues, green for other hydrophobic residues, and magenta for charged residues.

FIG. 5. Stereo pair showing the sugar binding sites after superposition of the Cα atoms for monomers of methyl-α-D-mannopyranoside (yellow, Naismith *et al.,* 1994), α-PNM (green), and α-PNG (red) complexes. The program used to produce the figure was "O" (Jones *et al.,* 1991).

FIG. 6. Stereo pairs showing (a) the interaction between tetramers packing along the crystallographic *b* axis. The A subunit is colored red and the B subunit of a second tetramer is colored green. (b) The interaction between the C subunit (red) and a second tetramer of ConA where the A and B subunits are colored yellow and green, respectively. (c) The interaction between the D subunit (red) and another tetramer for which A, B, and C subunits are colored green, yellow, and magenta, respectively. The α-PNM is shown in a stick representation colored yellow (carbon), red (oxygen), and blue (nitrogen). Water molecules and metal ions are represented as spheres and are colored pink and magenta, respectively.


A227
A228
A98
A14
A12
A99
A100
5


B16
B12
B14
A16
B99
B168
A12
A101
a


C208
C12
C100
C227
C229
C14
C16
A138
B217
B219
b

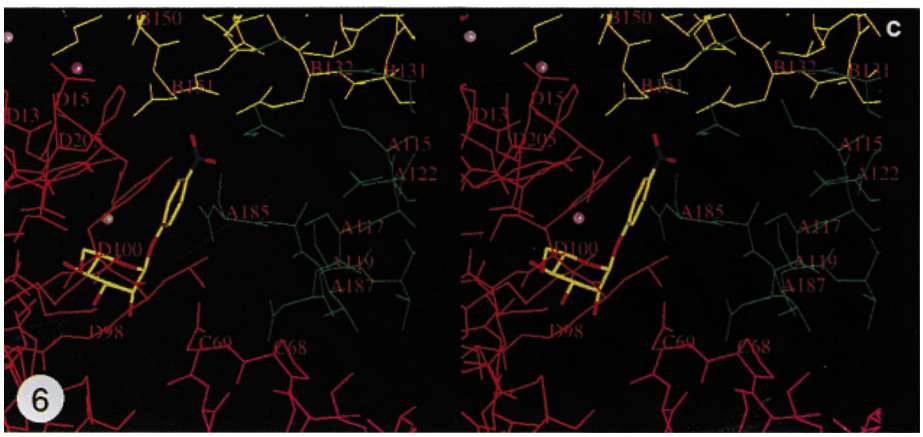

B150
B132
B131
D15
D13
D205
A115
A122
A185
A117
D100
A119
A187
D98
C69
C68
6
c

TABLE III
Conformation of α-PNM and α-PNG (in Parentheses)

Subunit	ϕ^a (O5–C1–01–C7)°	ψ (C1–O1–C7–C8)°	χ (O5–C5–C6–O6)°
A	69 (66)	165 (166)	–43 (–57)
B	65 (60)	–178 (–167)	–49 (–51)
C	52 (60)	–162 (–175)	–55 (–66)
D	59 (58)	–177 (165)	–59 (–59)

[a] The angles ϕ, ψ, and χ are defined in Fig. 1a.

In both complexes, the saccharides bind to each monomer adopting a conformation which is basically similar in all monomers, as shown in Table III, and is also almost identical to the conformation which α-PNM itself adopts in the crystalline state (Agianian *et al.,* 1995) excluding the paranitro group (see below). The sugar parts of the molecules have a $^4C_1$ chair conformation. Semiempirical classical energy calculations, using the program HYPERCHEM (E. Mikros, personal communication), showed that this conformation corresponds to a free-energy minimum, most probably the global minimum for the isolated α-PNM (Agianian *et al.,* 1995). Our calculations (Mavrommatis, Troganis, and Hamodrakas, in preparation), based on NMR measurements of aglycon and saccharide proton distances from the manganese ion, indicate that both saccharides, α-PNM and α-PNG, are bound to Con A in solution with the same conformation found in the crystal structures. Except for the orientation of the nitro group, the same saccharide conformation is observed in the different subunits of each complex, despite their having different environments in the crystal lattice.

The saccharide parts of α-PNM and also of α-PNG are bound to each Con A monomer via a rather complex network of hydrogen bonds and are also involved in a number of van der Waals interactions (Fig. 5 and Table IV). This network is only slightly different from the network of hydrogen bonds stabilizing the interaction of methyl-α-D-mannopyranoside with Con A (Naismith *et al.,* 1994). It was found that, in the Con A–α-PNM complex, the main chain NH of Leu99 interacts with O5, O6, and/or perhaps O2 of α-PNM in three subunits, A, B, and C, but not in subunit D, whereas, in the Con A–methyl-α-D-mannopyranoside complex only one possible weak hydrogen bond of this group with O2 (in subunit C) was detected. In addition, in the α-PNM structure, O2 of the saccharide of subunit A, at least, is linked to Thr226 (both carbonyl and OG1 oxygens) of the same subunit by a bridging water molecule (W3). These interactions are absent in the Con A–α-PNG complex. It has been estimated (utilizing the program WHATIF; Vriend, 1990) that 153 Å^2 of water-accessible surface area of the protein is "buried" when it interacts with α-PNM, whereas, upon binding of α-PNG, 149 Å^2 is buried. Major contributors to these accessibility losses are, in descending order, the side chains of Leu99, Tyr12, Tyr100, Arg228, and Asn14. The difference between α-PNM and α-PNG is mainly due to the difference in accessibility of the side chain of Leu99.

This difference might contribute to the higher affinity of Con A for para-substituted phenyl mannosides over paraphenyl glucosides (Loontiens *et al.,* 1973; Poretz and Goldstein, 1971; Troganis and

TABLE IV
Protein–Saccharide Hydrogen Bonds for α-PNM and α-PNG (in Parentheses)

		Hydrogen bond distance (Å)			
Sugar atom	Protein atom	A	B	C	D
O2	Leu99 N	3.5	3.3	3.3	
O2	W3	3.4			
O3	Arg228 N	3.2 (3.3)	3.0 (2.7)	3.0 (3.0)	3.2 (3.9)
O3	W4	2.7			
O4	Asp208 OD2	2.8	2.8 (2.6)		
O4	Asp208 OD1	3.3 2.6)	(3.5)	2.7 (2.7)	2.7 (2.7)
O4	Asn14 ND2	3.2 (3.0)	2.8 (2.8)	2.7 (3.3)	3.0 (3.3)
O4	Arg228 N	2.8 (3.0)	3.1 (3.4)	3.0 (3.1)	3.1 (3.1)
O5	Leu99 N	3.5 (3.8)	3.2 (3.3)	3.4 (3.5)	3.6 (3.5)
O6	Asp208 OD1	3.3	3.1 (2.9)	(3.4)	(3.7)
O6	Asp208 OD2	(3.1)		3.2 (3.2)	2.9 (3.0)
O6	Tyr100 N	2.9 (3.4)	3.0 (3.2)	3.0 (3.2)	3.3 (2.9)

Stassinopoulou, 1994) and is probably related to the higher affinity of Con A to mannosides rather than glucosides in general (Goldstein and Poretz, 1986).

The differences of the hydrogen bonding network may be attributed to a slight tilt of the sugar ring of approximately 5° toward Tyr100 and Tyr12 with pivot point the C5 of the sugar ring, with respect to the orientation of the sugar ring found in the Con A–methyl-α-D-mannopyranoside complex (Fig. 5). The tilt is such that C3, C4, and C5 of the ring and the hydroxyls O3, O4, and O6 retain their spatial positions as in the Con A–methyl-α-D-mannopyranoside complex. Presumably, this is due to the interactions of the aglycon part of α-PNM and α-PNG with this part of Con A (see also below).

Since no significant distortions occur on either the structure of the protein or the structure of the saccharides upon binding, it would appear that the energy of interaction is used primarily to promote binding and not to distort either molecule.

Hydrophobic Site

It is well known that in addition to carbohydrate binding sites, legume lectins frequently possess hydrophobic sites of three distinct types based on different ligand affinities (Sharon and Lis, 1990). One of these is believed to be adjacent to the carbohydrate binding site, as evidenced by the finding that hydrophobic glucosides and mannosides or other hydrophobic derivatives of monosaccharides bind more strongly (up to 10- to 50-fold) to the lectin than the analogous nonhydrophobic compounds (Sharon and Lis, 1990; Poretz and Goldstein, 1986; Farina and Wilkins, 1980; Loontiens *et al.,* 1973; Troganis and Stassinopoulou, 1994).

In this context, it is interesting to observe the modes of interaction of the aglycon parts of the saccharides α-PNM and α-PNG with Con A. Figures 2 and 6a show that Tyr12 stacks against one face of the sugar ring. This type of interaction is common in protein–saccharide complexes (Vyas, 1991), as has been noted previously (Naismith *et al.,* 1994). Tyr12 also interacts with one face of the paranitrophenyl group of the aglycon portion of the sugars. Furthermore, the phenyl ring of Tyr100 is stacked against the same face of the paranitrophenyl group of both α-PNM and α-PNG.

Hydrogen bonds between the hydroxyls of both Tyr12 and Tyr100 and the oxygens O7 and O8 of the paranitro group are possible but vary in number and strength from subunit to subunit, presumably due to different packing constraints. Therefore, it appears that this hydrophobic site is not a hydrophobic cavity capable of binding specifically hydrophobic ligands independently, but, rather, a portion of the surface of the protein with hydrophobic properties (cf. Fig. 4, inset) assisting in a sense sugar binding.

It remains to be shown which portion of the protein surface is utilized for interactions with the hydrophobic groups of β-substituted sugars, which are known to interact less strongly than α-substituted anomers (Farina and Wilkins, 1980, and references therein; Loontiens *et al.,* 1973).

Packing Interactions of the Subunits

The crystal packing is such that along the *b* axis, Con A tetramers related by the twofold screw axis of symmetry form infinite chains in which contacts occur between the A and B monomers of symmetry-related tetramers. It appears to be generated by a π-stacking interaction of the paranitrophenyl rings of sugars bound to the A and B monomers, respectively. This is shown in Fig. 2 and in more detail in Fig. 6a. This interaction between the bound saccharides apparently dictates how tetramers pack along the *b* axis and does not distort, but rather favors, the coplanarity of the paranitro group with the phenyl ring of the aglycon part of the saccharides of both subunits A and B, in both complexes.

The infinite chains of the tetramers of Con A (data not shown) are reminiscent of the structure of infinite networks of the carbohydrate protein galectin-1 (a mammalian lectin) in complex with branched oligosaccharides, which may be important in understanding of how this class of lectins act at the atomic level on cells (Sharon, 1994; Bourne *et al.,* 1994).

The packing interactions for subunits C and D are very different from those for subunits A and B. Thus, the tetramers are packed along the other twofold screw axis *a,* so that the saccharide binding site of subunit C is in contact with subunits A and B of symmetry-related molecules. This results in the burial of the bound saccharide of the C subunit into a shallow pocket involving two other Con A monomers (Fig. 6b). The strongest interaction between the C-subunit bound saccharide and its neighbors appears to be a hydrogen bond formed between the terminal NZ of the side chain of Lys A138 and the paranitro oxygen atom O7 of the saccharide. The same nitrogen atom may form a hydrogen bond with the hydroxyl of the side chain of Tyr C100. The hydrogen bond between NZ of Lys A138 and O7 coupled with hydrogen bonds from the hydroxyl groups of Tyr C12 and Tyr C100 results in a deviation of 67° in the α-PNM complex and 67° in the α-PNG complex of the paranitro group from coplanarity with the phenyl ring of the aglycon. Such rotations of a nitro group attached to a phenyl ring from coplanarity with the phenyl ring have frequently been observed in small molecule crystal structures (e.g., Yang *et al.,* 1994) and usually arise in order to optimize local packing geometry.

Along the twofold axis *c,* Con A tetramers related

by this symmetry element do not interact. However, through simple unit cell translations they interact so that the saccharide bound in subunit D is in contact with portions of the three other subunits A, B, and C (Fig. 6c). Although the density for the paranitro group of subunit D is not good there appear to be hydrogen bonds formed between NE2 of Gln B132 and both oxygens of the paranitro group and weaker interactions between ND2 of Asn C69 and O1, O2, in both α-PNM and α-PNG complexes. The former, together with hydrogen bonds formed between the Tyr D12, Tyr D100, and the oxygens of the paranitro group result in a distortion of 40° for the Con A–α-PNM complex and 72° for the Con A–α-PNG complex from coplanarity with the phenyl ring of the aglycon.

Unfortunately, the medium resolution in which both structures were refined leaves some uncertainty about the important water structure around each tetramer and in particular in the saccharide vicinity. Therefore, only those water molecules clearly identified in these regions were included. Higher resolution is needed to clarify these details.

S.J.H. thanks the EMBO for a short-term fellowship.

REFERENCES

Agianian, B., Perrakis, A., Kanellopoulos, P. N., Sheldrick, B., and Hamodrakas, S. J. (1995) The crystal and molecular structure of 4′-nitrophenyl-α-D-mannopyranoside ethanol solvate, submitted for publication.

Agrawal, B. B. L., and Goldstein, I. J. (1986) Protein–carbohydrate interaction. VII. Physical and chemical studies on concanavalin A, the hemagglutinin of the jack bean, *Arch. Biochem. Biophys.* **124,** 218–229.

Becker, J. W., Reeke, G. N., Jr., Cunningham, B. A., and Edelman, G. M. (1976) New evidence on the location of the saccharide-binding site of concanavalin A, *Nature (London)* **259,** 406–409.

Bittiger, H. S., and Schnebli, H. P. (1976) Concanavalin A as a Tool, Wiley–Interscience, London.

Bowles, D. J. (1990) Defense-related proteins in higher plants, *Annu. Rev. Biochem.* **59,** 873–907.

Bourne, Y., Roussel, A., Frey, M., Rouge, P., Fontecilla-Camps, J. C., and Cambillau, C. (1990) Three-dimensional structures of complexes of Lathyrus ochrus isolectin I with glucose and mannose: Fine specificity of the monosaccharide-binding site, *Proteins* **8,** 365–376.

Bourne, Y., Bolgiano, B., Liao, D. I., Strecker, G., Canteau, P., Herzberg, O., Feizi, T., and Cambillau, C. (1994) Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides, *Nature Struct. Biol.* **1,** 863–870.

Bourne, Y., Mazurier, J. B., Legrand, D., Rouge, P., Montreuil, J., Spik, G., and Cambillau, C. (1994) Structures 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.

Brunger, A. T. (1992a) X-PLOR Version 3.1: A system for x-ray crystallography and NMR, Yale Univ. New Haven, CT.

Brunger, A. T. (1992b) Free R value: A novel statistical quantity for assessing the accuracy of crystal structures, *Nature* **335,** 472–475.

Delbaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Pearlstone, J. R., Carpenter, M. R., Smillie, L. R., Nikrad, P. V., Spohr, U., and Lemieux, R. U. (1990) Molecular recognition of a human blood group determinant by a plant lectin, *Can. J. Chem.* **68,** 1116–1121.

Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb (Gilboa), A. J., Dodson, E. J., Papiz, M. Z., Wan, T., and Cambell, J. W. (1989) The structure of the saccharide-binding site of concanavalin A, *EMBO J.* **8,** 2189–2193.

Edelman, G. M., Cunningham, B. A., Reeke, G. N., Becker, J. W., Waxdal, M. J., and Wang, J. L. (1972) The covalent and three-dimensional structure of concanavalin A, *Proc. Natl. Acad. Sci. USA* **69,** 2580–2584.

Emmerich, C., Helliwell, J. R., Redshaw, M., Naismith, J. H., Harrop, S. J., Raftery, J., Kalb (Gilboa), A. J., Yariv, J., Dauter, Z., and Wilson, K. S. (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 Cyrstallogr.* **D50,** 749–756.

Farina, R. D., and Wilkins, R. G. (1980) Kinetics of interaction of some α- and β-D-monosaccharides with Concanavalin A, *Biochim. Biophys. Acta* **631,** 428–438.

Goldstein, I. J., and Poretz, R. P. (1986) *in* Liener, I. E., Sharon, N., and Goldstein, I. J., (Eds.), The Lectins: Properties, Functions and Applications in Biology and Medicine, pp. 33–247, Academic Press, New York.

Hardman, K. D., and Ainsworth, C. F. (1972) Structure of concanavalin A at 2.4 Å resolution, *Biochemistry* **11,** 4910–4919.

Hardman, K. D., and Ainsworth, C. F. (1976) Structure of the concanavalin A-methyl alpha-D-mannopyranoside complex at 6 Å resolution, *Biochemistry* **15,** 1120–1128.

Hardman, K. D., Agarwal, R. C., and Freiser, M. J. (1982) Manganese and calcium binding sites of concanavalin A, *J. Mol. Biol.* **157,** 69–86.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr.* **A47,** 110–119.

Kabsch, W. (1988) Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector, *J. Appl. Crystallogr.* **21,** 916–924.

Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr.* **24,** 946–950.

Laskowski, R. A., MacArthur, M. W., Moss, D., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* **26,** 283–291.

Lis, H., and Sharon, N. (1986) *in* Liener, I. E., Sharon, N., and Goldstein, I. J., (Eds.), The Lectins: Properties, Functions and Applications in Biology and Medicine, pp. 265–291, Academic Press, New York.

Loontiens, F. G., Van Wauwe, J. P., De Gussem, R., and De Bruyne, C. K. (1973) Binding of para-substituted phenyl glycosides to concanavalin A, *Carbohydr. Res.* **30,** 51–62.

Loris, R., Van Overberge, D., Dao-Thi, M. H., Poortmans, F., Maene, N., and Wyns, L. (1994) Structural analysis of two crystal forms of lentil lectin at 1.8 Å resolution, *Proteins* **20,** 330–346.

Luzatti, V. (1952) Traitement statistique des erreurs dans la determination des structures cristallines, *Acta Crystallogr.* **5,** 802–810.

Min, W., Dunn, A. J., and Jones, D. H. (1992) Non-glycosylated recombinant pro-concanavalin A is active without polypeptide cleavage, *EMBO J.* **11,** 1303–1307.

Naismith, J. H., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Wan, T. C. M., Weisgerber, S., Kalb (Gilboa), A. J., and Yariv, J. (1993) Refined structure of cadmium-substituted concanavalin A at 2.0 Å resolution, *Acta Crystallogr.* **D49,** 561–571.

Naismith, J. H., Emmerich, C., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Raftery, J., Kalb (Gilboa), A. J., and Yariv, J. (1994) Refined structure of concanavalin A complexed with methyl-α-D-mannopyranoside at 2.0 Å resolution and comparison with the saccharide-free structure, *Acta Crystallogr.* **D50,** 847–858.

Navaza, J. (1994) AMoRe: An Automated Package for Molecular Replacement, *Acta Crystallogr.* **A50,** 157–163.

Nicholls, A. (1993) GRASP: Graphical Representation and Analysis of Surface Properties, Columbia Univ., New York.

Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins: Struct. Funct. Genet.* **11,** 281–296.

Poretz, R. D., and Goldstein, I. J. (1971) Protein–carbohydrate interaction. On the mode of bonding of aromatic moieties to concanavalin A, the phytohemagglutinin of the jack bean, *Biochem. Pharmacol.* **20,** 2727–2739.

Reeke, G. N., Jr., and Becker, J. W. (1988) Carbohydrate-binding sites of plant lectins, *Curr. Topics Microbiol. Immunol.* **139,** 35–58.

Rini, J. M., Hardman, K. D., Einspahr, H., Suddath, F. L., and Carver, J. (1993) X-ray crystal structure of a pea lectin-trimannoside complex at 2.6 Å resolution, *J. Biol. Chem.* **268,** 10126–10132.

Shaanan, B., Lis, H., and Sharon, N. (1991) Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose, *Science* **254,** 862–866.

Sharon, N. (1994) When lectin meets oligosaccharide, *Nature Struct. Biol.* **1,** 843–845.

Sharon, N., and Lis, H. (1990) Legume lectins—A large family of homologous proteins, *FASEB J.* **4,** 3198–3208.

Sumner, J. B., and Howell, S. F. (1936) The identification of the haemagglutinin of Jack bean with concanavalin A, *J. Bacteriol.* **32,** 227–237.

Troganis, A., and Stassinopoulou, C. I. (1994) Modes of association of concanavalin A with α-D-glycosides, *Biochim. Biophys. Acta* **1206,** 215–224.

Vyas, N. K. (1991) Atomic features of protein-carbohydrate interactions, *Curr. Opin. Struct. Biol.* **1,** 732–740.

Vriend, G. (1990) A molecular modelling and drug design package, *J. Mol. Graph.* **8,** 52–56.

Weisgerber, S., and Helliwell, J. R. (1993) High-resolution crystallographic studies of native concanavalin A using rapid Laue data collection methods and the introduction of a monochromatic large-angle oscillation technique (LOT), *J. Chem. Soc. Faraday Trans.* **89,** 2667–2675.

Yang, J., Nelson, B. R., Weil, J. A., and Quail, J. W. (1994) The potassium salt of 9-(2,4,6-trinitroanilino)carbazole, *Acta Crystallogr.* **C50,** 1548–1550.