Acta Cryst. (1994). D**50**, 847–858

Refined Structure of Concanavalin A Complexed with Methyl α-D-Mannopyranoside at 2.0 Å Resolution and Comparison with the Saccharide-Free Structure

BY JAMES H. NAISMITH,* CHRISTIANE EMMERICH, JARJIS HABASH AND STEPHEN J. HARROP

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England

JOHN R. HELLIWELL†

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England, and SERC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, England

WILLIAM N. HUNTER AND JAMES RAFTERY

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England

A. JOSEPH KALB (GILBOA)

Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel

AND JOSEPH YARIV

Laboratoire de Cristallographie, URA 144, CNRS, Université Bordeaux I, 33405 Talence, France

(*Received* 3 *February* 1994; *accepted* 16 *May* 1994)

Abstract

The three-dimensional structure of the complex between methyl α-D-mannopyranoside and concanavalin A has been refined at 2.0 Å resolution. Diffraction data were recorded from a single crystal (space group $P2_12_12_1$, $a = 123.7$, $b = 128.6$, $c = 67.2$ Å) using synchrotron radiation at a wavelength of 1.488 Å. The final model has good geometry and an R factor of 19.9% for 58871 reflections (82% complete), within the resolution limits of 8 to 2 Å, with $F > 1.0\sigma(F)$. The asymmetric unit contains four protein subunits arranged as a dimer of dimers with approximate 222 point symmetry. Each monomer binds one saccharide molecule. Each sugar is bound to the protein by hydrogen bonds and van der Waals contacts. Although the four subunits are not crystallographically equivalent, the protein–saccharide interactions are nearly identical in each of the four binding sites. The differences that do occur between the four sites are in the structure of the water network which surrounds each saccharide; these networks are involved in crystal packing. The structure of the complex is compared with a refined saccharide-free concanavalin A structure. The saccharide-free structure is composed of crystallographically identical subunits, again assembled as a dimer of dimers, but with exact 222 symmetry. In the saccharide complex the tetramer association is different in that the monomers tend to separate resulting in fewer intersubunit interactions. The average temperature factor of the mannoside complex is considerably higher than that of the saccharide-free protein. The binding site in the saccharide-free structure is occupied by three ordered water molecules and the side chain of Asp71 from a neighbouring molecule in the crystal. These occupy positions similar to those of the four saccharide hydroxyls which are hydrogen bonded to the site. Superposition of the saccharide-binding site from each structure shows that the major changes on binding involve expulsion of these ordered solvents and the reorientation of the side chain of Tyr100. Overall the surface accessibility of the saccharide decreases from 370 to 100 Å^2 when it binds to the protein. This work builds upon the earlier studies of Derewenda *et al.* [Derewenda, Yariv, Helliwell, Kalb (Gilboa), Dodson, Papiz, Wan & Campbell (1989). *EMBO J.* **8**, 2198–2193] at 2.9 Å resolution, which was the first detailed study of lectin–saccharide interactions.

1. Introduction

Concanavalin A was first isolated and crystallized over 70 years ago (Summer, 1919). In solution the minimal molecular weight of concanavalin A is that of a dimer (Kalb & Lustig, 1968). Each of its identi-

* Present address: Research Laboratories, Howard Hughes Medical Institute, University of Texas, 5323 Harry Hines Boulevard, Dallas, TX 75235-9050, USA.

† To whom all correspondence should be addressed.

cal subunits has one saccharide-binding site [Yariv, Kalb (Gilboa) & Levitzki, 1968] and two distinct metal-binding sites: one for a transition-metal ion and another for a calcium ion (Kalb & Levitzki, 1968). When the naturally occurring metals are removed from concanavalin A in acid, the protein no longer binds saccharides. Addition of the missing metals at the correct pH restores saccharide binding (Yariv *et al.*, 1968).

After the protein was crystallized in a form suitable for X-ray crystallography (Greer, Kaufman & Kalb, 1970), its structure was determined by two independent groups (Edelman *et al.*, 1972; Hardman & Ainsworth, 1972). Initial expectations that a structure of the complex of concanavalin A and a saccharide would soon follow were not realized. Co-crystallization of concanavalin A and mannoside was reported by Hardman & Ainsworth (1976) but the crystals suffered from lattice disorders, which limited the resolution to 6 Å. The saccharide-binding site was located, however. This was consistent with observations from NMR experiments (Brewer, Brown & Koenig, 1973; Brewer, Sternlicht, Marcus & Grollman, 1973; Alter & Magnuson, 1974; Villafranca & Viola, 1974) and conclusions from crystal studies of saccharide-soaked concanavalin A crystals which had been cross-linked with glutaraldehyde (Becker, Reeke, Cunningham & Edelman, 1976). Crystallographic studies of the saccharide complex of favin (highly homologous to concanavalin A) revealed a similar binding site but were unable to position the saccharide unambiguously (Reeke & Becker, 1986). Crystallization of well diffracting cubic crystals (space group $I2_13$) of the complex of concanavalin A with methyl α-D-glucopyranoside was reported by Yariv, Kalb (Gilboa), Papiz, Helliwell, Andrews & Habash, 1987). This crystal structure has been solved at 2.0 Å resolution (Harrop *et al.*, 1993) and will be reported in detail separately.

The crystal structure of the methyl α-D-mannopyranoside concanavalin A complex was solved at 2.9 Å by Derewenda *et al.* (1989) and the details of the binding site were reported. That study showed that the methyl α-D-mannopyranoside molecule is bound in the *C*1 chair conformation, 8.7 Å from the calcium-binding site and 12.8 Å from the transition metal-binding site. There is a network of seven hydrogen bonds connecting O atoms O3, O4, O5 and O6 of the mannoside to residues Asn14, Leu99, Tyr100, Asp208 and Arg228. O2 and O1 of the mannoside extend into solvent. These results are confirmed here. We also describe the van der Waals interactions which characterize the protein–saccharide interaction and we have been able to resolve the methyl substituent on the saccharide. Moreover, we have located a water ligand to the calcium ion which is involved in stabilization of the *cis* peptide between Ala207 and Asp208.

We now report the first full description of the methyl α-D-mannopyranoside–concanavalin A tetramer complex and extend the earlier study of the saccharide-binding site from 2.9 to 2.0 Å resolution. We also provide a comparison with the saccaride-free protein also at high resolution (Emmerich *et al.*, 1994; Weisgerber & Helliwell, 1993).

2. Experimental and computational procedures

2.1. *Crystallization, data collection and processing*

Crystals were grown from stock solutions of 80–150 mg protein ml^{-1} by dialysis equilibration with a solution comprising 0.05 *M* Pipes, 0.1 *M* $NaNO_3$, 1 m*M* $MnCl_2$, 1 m*M* $CaCl_2$, 2 mg ml^{-1} NaN_3 and 0.1 *M* methyl α-D-mannopyranoside (Pfanstiehl), pH 6.8. These conditions are as previously described (Derewenda *et al.*, 1989). These crystallization conditions are as identical as possible to those for the saccharide-free (*I*222) crystals. The only differences are in the presence of the mannoside and, necessarily, the higher concentration of the protein. In the *I*222 case 24 mg ml^{-1} is used typically while for the complex, which is very soluble, the much higher concentration was used. The pH and salt concentration are about equal for both crystal forms and both are grown by dialysis.

A crystal of size 2.5 × 0.2 × 0.4 mm was used to collect all the X-ray crystallographic data on CEA Reflex photographic film at the Daresbury Synchrotron Radiation Source at Station PX7.2 Daresbury, England (Helliwell *et al.*, 1982). The storage ring was operated at 2 GeV. The circulating current was 190 mA at the start of data collection, and had fallen to 158 mA at the end of data collection, some 16 h later. The crystal capillary was aligned on the goniometer head with **a** parallel to the rotation axis and **b** initially parallel to the incident X-ray beam. The data were collected using a wavelength of 1.488 Å and a collimator of 0.3 mm diameter with the crystal-to-film distance set at 58 mm. The crystal was at room temperature (≃298 K) during data collection. An Arndt–Wonacott oscillation camera, modified for rapid oscillations at the synchrotron, was used. The oscillation interval was 0.9° from 0 to 60° and then increased to 1.5° from 60 to 90° of rotation of the crystal taking advantage of the short *c* axis, in order to collect a complete data set from one crystal in one run. The crystal was translated along the rotation axis several times during data collection to reduce the effects of radiation damage. An exposure time of 40 s per degree was used at the start and by the end of data collection this had increased to 55 s per degree in response to beam decay.

Table 1. *Data collected on the mannoside–concanavalin A complex*

Resolution (Å)	All data			Data with $F > 1.0\sigma(F)$	
	No. of reflections	Completeness	R_{merge}*	No. of reflections	Completeness
8.0–3.7	9115	85%	5.8%	9075	85%
3.7–2.8	13620	90%	7.8%	13422	89%
2.8–2.4	13777	88%	10.3%	13149	84%
2.4–2.1	17362	84%	15.0%	15983	78%
2.1–2.0	8324	84%	24.2%	7242	73%
8.0–2.0	62198	86%	7.6%	58871	82%

* $R_{merge} = \sum\sum I(h)_j - \langle I(h)\rangle / \sum\sum I(h)_j$ where $I(h)$ is the measured diffraction intensity and the summation includes all observations.

The orientation of the crystal was determined using one pair of stills measured at 0 and 90°. Stills were also recorded after each translation to monitor for crystal slippage. All films were scanned at Manchester University using a SCANDIG 3 microdensitometer with a 50 μm raster. The software for processing oscillation photographs (Arndt & Wonacott, 1977) was as follows. The pattern was generated using *OSCGEN* and the intensity measured using the profile-fitting option of *MOSFLM* and interfilm scaling was carried out using *ABSCALE*. The data were scaled, merged and structure amplitudes calculated, using *ROTAVATA/AGROVATA* (*CCP*4, SERC Daresbury Laboratory, 1979). A total of 237069 measurements were recorded which yielded 62198 unique reflections. Table 1 gives some statistics for the data set. The cell dimensions were determined as $a = 123.74$, $b = 128.62$, $c = 67.17$ Å. The space group was $P2_12_12_1$. The Matthews number (Matthews, 1968) for the crystal is 2.73 $Å^3\ Da^{-1}$ (corresponding to 55% solvent).

Initially we had hoped to merge the data from this crystal with those collected for the earlier 2.9 Å study from a crystal of essentially identical cell parameters, which had been grown under identical conditions. However, data for the two crystals were not isomorphous (mean isomorphous difference on F, approximately 40%). Retrospective comparison of the two models revealed that the positions of the centres of mass differed by 0.7 Å.

Brief details of the data collected for the saccharide-free crystal structures are given in §3.6.

2.2. *Model refinement*

Refinement was carried out using the *X-PLOR* program (Version 2.1) (Brünger, 1990) on a Silicon Graphics 4D power series workstation. The dictionary and parameter files were adjusted to accommodate the non-proline *cis* peptide and methyl α-D-mannopyranoside in each monomer. No restraints were placed on metal–ligand distances and no charge was given to the metal ions. Electron-density maps were calculated within the *CCP*4 package (SERC Daresbury Laboratory, 1979). Manual adjustment of the model relative to the $2F_o - F_c$ and $F_o - F_c$ electron-density maps was carried out using *FRODO* (Jones, 1981) adapted to run on an ESV30 (P. R. Evans, personal communication). Geometric analyses of the model were carried out with *X-PLOR* and *DSSP* (Kabsch & Sander, 1984).

The model for the tetramer from Derewenda *et al.* (1989), PDB coordinate file 4CNA, was treated as a molecular-replacement solution for the present study. The starting R factor was 45%. This model contained only the protein atoms and was refined as a rigid body against the 2.0 Å data. This was carried out for the tetramer as a whole and then each subunit treated independently. After rigid-body refinement the R factor was 37%. Restrained refinement of individual atomic positions was then carried out using a simulated-annealing protocol based on a temperature change of 3000–300 K using $F > 1.0\sigma(F)$ data in the resolution range 8–2.4 Å (35644 reflections). The model was subsequently

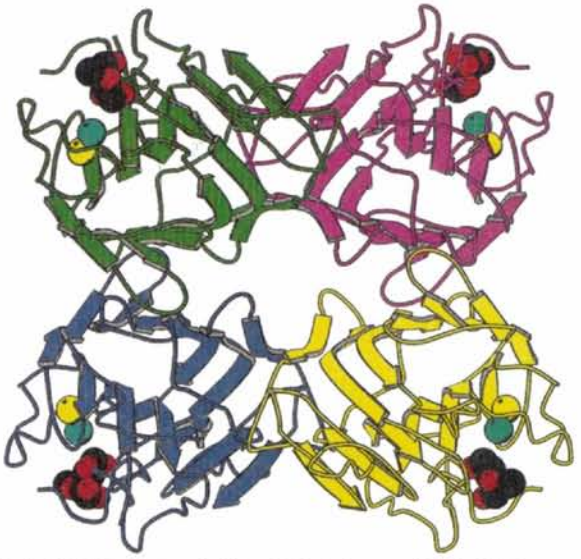


Fig. 1. Ribbon representation of the tetramer. The Mn^{2+} ions are yellow, the Ca^{2+} ions are cyan, the mannosides are shown as van der Waals spheres with red for oxygen and black for carbon. The conventional dimers are formed by the green and pink (AB) subunits and also by the blue and yellow (CD) subunits.

regularized by energy minimization (*POWELL* option in *X-PLOR*). Non-crystallographic restraints of 502 kJ mol^{-1} Å^{-2} (120 kcal mol^{-1} Å^{-2}) were employed. The data set was then expanded to include data [with $F > 1.0\sigma(F)$] in the resolution range 8–2 Å (58871 reflections). The R factor was 30%. The model was again subjected to simulated annealing (4000 to 300 K) and regularized using the *POWELL* energy-minimization option. Individual temperature factors were refined and the R factor dropped to 26%. Examination of $2F_o - F_c$ and $F_o - F_c$ electron-density maps calculated using the protein model indicated that several changes in the conformation of individual amino acids were required. Clear ($>3\sigma$) $F_o - F_c$ electron density for the sugars in the binding sites of the *A*, *B* and *D* subunits was observed. These three sugar molecules were, therefore, included in the refinement, although the methyl substituent for each was given zero occupancy at this stage. This was done in order to provide clear evidence for the position of the methyl substituent which had previously been thought to be disordered (Derewenda *et al.*, 1989). The *C* subunit was notably less well ordered than the other three subunits and no clear density for a saccharide could be identified for this subunit at this stage. This anomalous behaviour of the *C* subunit continued throughout the refinement. We can offer no explanation for the anomaly. A total of 218 water molecules were identified and included in the model. Waters were included if they were in $>3\sigma$ $F_o - F_c$ density and made plausible hydrogen bonds; they were deleted if they did not reappear in the $2F_o - F_c$ maps. The model was subjected to simulated annealing (700 to 300 K), *POWELL* energy minimization and restrained temperature-factor refinement. Non-crystallographic restraints were loosened to a value of about 125.5 kJ mol^{-1} Å^{-2} (30 kcal mol^{-1} Å^{-2}). The R factor was 21.1%. A molecule of methyl α-D-mannopyranoside could now be modelled in the *C* subunit ($>2.5\sigma$ $F_o - F_c$ electron density). The methyl substituents of the three other sugars could be seen in strong ($>3.5\sigma$) $F_o - F_c$ electron density. Two amino-acid sequence changes were also made to the model to correspond to the sequence determined by (Min, Dunn & Jones, 1992). A further 254 water molecules were added. A further cycle of simulated annealing (500 to 300 K), *POWELL* energy minimization and temperature-factor refinement was carried out, without any non-crystallographic restraints, resulting in an R factor of 20.2%. The remainder of the refinement alternated between minor manual changes in the model, such as addition/deletion of water molecules or geometry idealization followed by 20 cycles of *POWELL* energy minimization, until no further waters could be added or improvements made to the model. The final R factor was 19.9% on the 58871 reflections between 8 and 2 Å with $F > 1.0\sigma(F)$ representing 82% of the data (73% between 2.1 and 2.0 Å).

3. Results

3.1. *Quality of the final model*

The tetramer, the four saccharides and the metal ions are shown in Fig. 1. Parameters relating to the final model are given in Table 2. The final model contains 533 solvent molecules and one chloride ion. The coordinates and structure factors have been deposited with the Protein Data Bank.* The principal regions of poor density are the N-terminal residue and the surface loops at residues 120, 150, 160 and 202. Fig. 2 shows a Ramachandran plot (Ramakrishnan & Ramachandran, 1965) for each subunit. The three outlying residues are in poorly defined regions of the structure. In subunit *A* two residues Thr120 ($\varphi = 58$, $\psi = 128°$) and His121 ($\varphi = 55$, $\psi = -7°$) are significantly outside the allowed regions. In the *B* subunit only Ser204 ($\varphi = 58$, $\psi = -59°$) is significantly outside the allowed region. A Luzzati plot (Luzzati, 1952) indicates a mean coordinate error of between 0.2 and 0.4 Å (data not shown). All the atoms of the sugars are in well defined electron density and the presence of two water molecules ligated to each calcium ion has been established, in contrast to the earlier study at 2.9 Å where only one water was found (Derewenda *et al.*, 1989). Certain features of the structure were verified by *POWELL* minimization of a model structure lacking these features and examination of the resulting omit maps; for example the *cis* peptide between Ala207 and Asp208 was confirmed in this way.

3.2. *Overall structural description*

The basic topology of the concanavalin A monomer is unchanged from the saccharide-free structures (Hardman & Ainsworth, 1972; Reeke, Becker & Edelman, 1975; Naismith *et al.*, 1993) in the Protein Data Bank (Bernstein *et al.*, 1977). Almost 50% of the residues in the structure are in two large β-sheets. The tetramer contains two dimers (Reeke *et al.*, 1975), each distinguished by a 12-stranded antiparallel β-sheet on one face. The dimers are formed by subunit *A* with subunit *B* and subunit *C* with subunit *D*. The *AB* dimer interaction comprises 17 hydrogen bonds, (six of which form the extended β-sheet), seven singly bridging water molecules and

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 5CNA, R5CNASF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: LI0174).

Table 2. *Some statistics on the final model*

The *X-PLOR* Version 2.1 geometrical parameters dicationary used was 'param19.pro'.

R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles ( )	3.15
Dihedral angles ( )	26.98
Improper angles ( )	1.26
Main-chain n.c.s.* (Å)	0.4
Side-chain n.c.s.* (Å)	0.9
Bond *B* factor† (Å^2)	4.7
Angle *B* factor‡ (Å^2)	6.5

* The r.m.s. deviations from perfect non-crystallographic symmetry.

† The r.m.s. difference in temperature factor of two bonded atoms.

‡ The r.m.s. difference in temperature factor of two atoms separated by one intervening atom.

154 van der Waals contacts (defined as two atoms closer than 4 Å). The *CD* dimer interaction comprises 16 hydrogen bonds, six of which form an extended β-sheet, seven singly bridging water molecules ands 171 van der Waals contacts. *AB* dimer formation buries a total of 2300 Å^2, *CD* dimer formation buries a total of 2500 Å^2. These dimers form a tetramer through the contacts of subunits *A* with *C* and *B* with *D*. The *BD* interaction includes nine hydrogen bonds, (three of which are salt bridges), five singly bridging water molecules and 66 van der Waals contacts. The *AC* interaction comprises 14 hydrogen bonds (four of which are salt bridges), 11 singly bridging water molecules and 72 van der Waals contacts. The total surface area buried in formation of the tetramer is 9600 Å^2 comprising equal contributions from the monomer-to-monomer interactions *AB*, *AC*, *CD* and *DB* of 2400 Å^2. There are no contacts between subunits *A* and *D* or subunits *B* and *C*.

The main-chain temperature factors for each subunit are shown in Fig. 3. A similar pattern is observed for the *A*, *B* and *D* subunits with two differences. In the *A* and *B* subunits the region at Asp71 has a significantly higher temperature factor than in the *D* subunit. The carboxy terminus is, however, better ordered in the *A* and *B* subunits compared to the *D* subunit. This reflects differences in packing at these sites. As expected, the temperature factors are greatest for the loops connecting β-strands. The loops around residues 120, 150, 160,

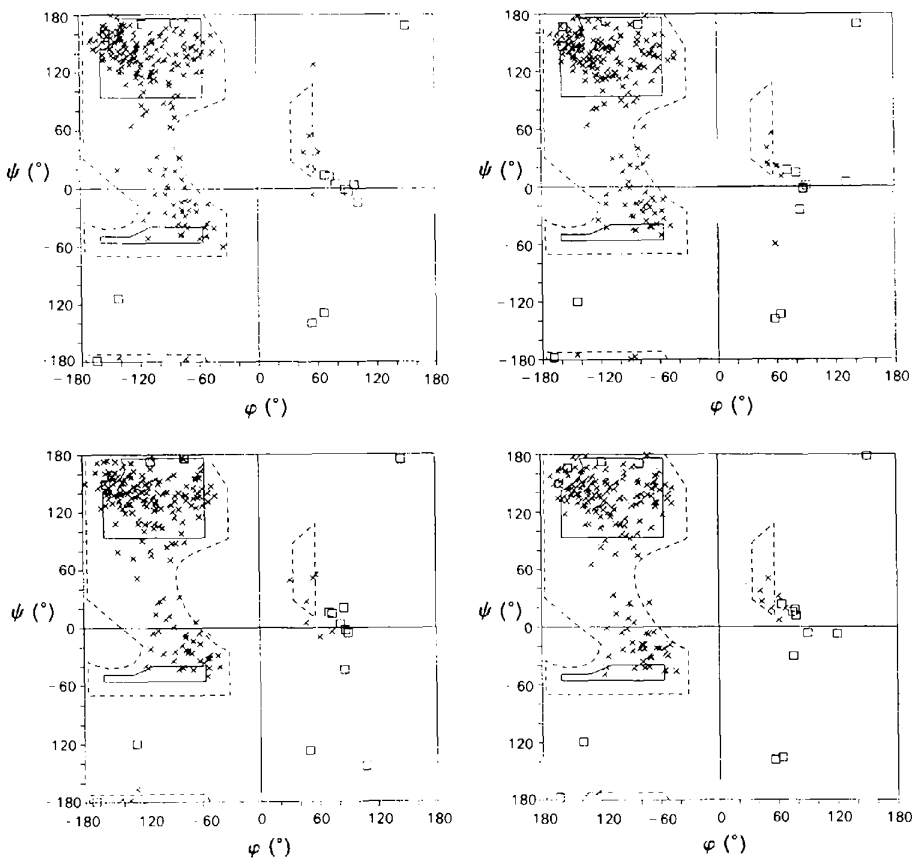


Fig. 2. Ramachandran plots for each subunit of the tetramer. The outliers for the *A* and *B* subunits are not well modelled by electron density (discussed in the text). Glycine residues are represented by a □ and non-glycine residues by a ×. The *A* subunit is top left, *B* subunit top right, *C* subunit bottom left and *D* subunit bottom right.

183 and 204 all have temperature factors above 50 Å^2. The *C* subunit is less well ordered than the other three subunits. However, the location of the highest *B* factors are in similar regions to the other subunits. The average temperature factors of the atoms in each subunit, shown in Table 3, further highlight the anomalous behaviour of the *C* subunit. The temperature factors of the water molecules range from 13 to 89 Å^2, with an average value of 49 Å^2.

Table 3. *Average temperature factors for each subunit*

Subunit	Main chain (Å)	Side chain (Å)	All atoms (Å)
A	26	30	28
B	26	31	28
C	40	44	42
D	29	34	31
Tetramer	30	35	32

3.3. *The saccharide-binding sites*

Each monomer has a saccharide-binding site in a shallow pocket near the surface of the protein. The locations can be seen from Fig. 1. The electron density for one of the bound saccharides, in a $(2F_o - F_c)$ map, is shown in Fig. 4.

Each pocket is formed by Tyr12, Asn14, Leu99, Tyr100, Asp208 and Arg228. The average distance from the centre of the saccharide to the Ca^{2+} ion is 8.2 Å. The average distance to the transition-metal ion is 12.5 Å. The average temperature factors of the saccharides are 20, 23, 53 and 30 Å^2 for subunits *A*, *B*, *C* and *D*, respectively. These follow a distribution similar to that of the subunits themselves. The average temperature factors of all the atoms in the

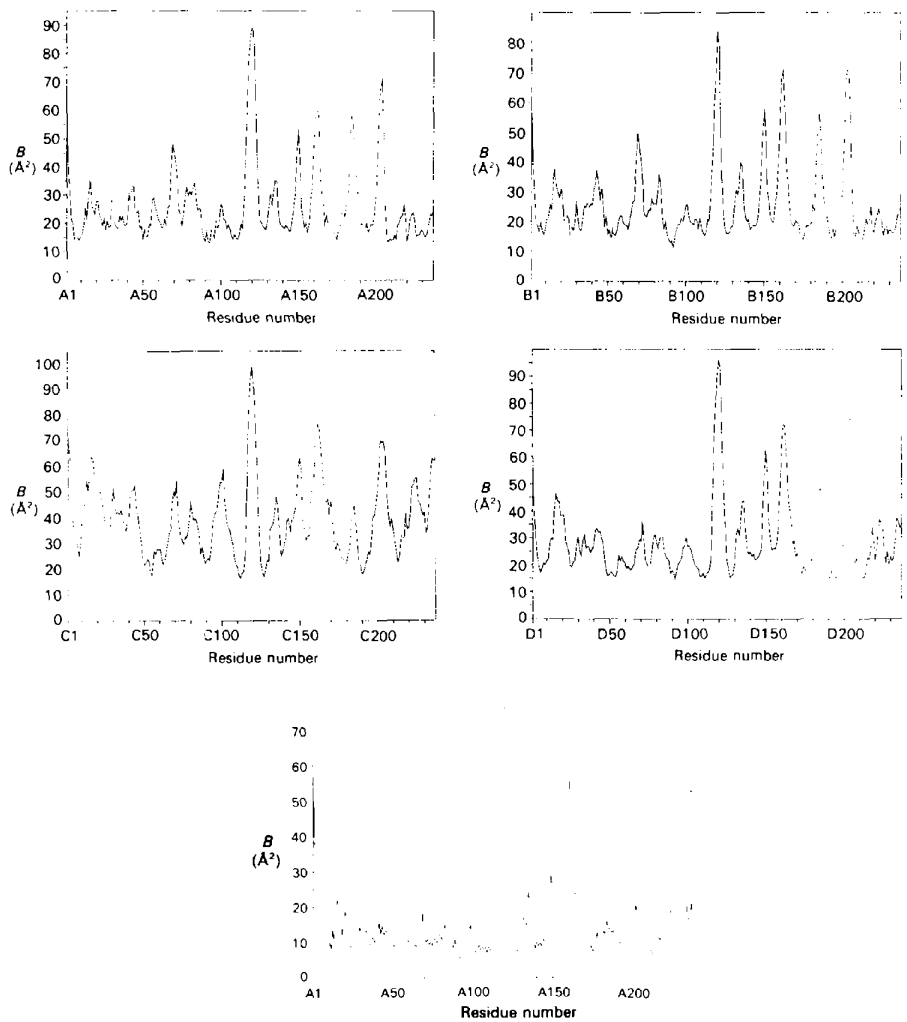


Fig. 3. Backbone temperature factor plots for each subunit of the tetramer. The peaks at residues 120, 150, and 160 occur at poorly modelled loops on the surface of the protein. The *C* subunit is anomalous in its behaviour; this is discussed in the text. The *A* subunit is top left, *B* subunit top right, *C* subunit bottom left and *D* subunit bottom right. Finally a similar plot is given for the saccharide-free structure.

saccharide-binding loops (residues 12–14, 98–100, 207–208 and 227–228) for each of the subunits are 21, 22, 53 and 29 Å^2, respectively. The direct saccharide–protein interactions are very similar for all four subunits. These interactions fall into two categories: hydrogen bonds which are listed in Table 4 and van der Waals contacts which are listed for the *B* subunit in Table 5. In addition, the sugar of subunit *A* (and *B*) is linked to Thr226 OG1 of subunit *B* (*A*) by a singly bridging water molecule (*WF* in Table 4). Fig. 5 gives the numbering scheme for the sugar and Fig. 6 shows a schematic diagram of the saccharide in the binding site. Tyr12 stacks against one face of the sugar, this type of interaction is common in protein–saccharide complexes (Vyas, 1991). The calcium ion is attached to three of the four loops which bind the saccaride. In the absence of calcium it is likely that one or more of these loops moves out of position thus abolishing saccharide binding.

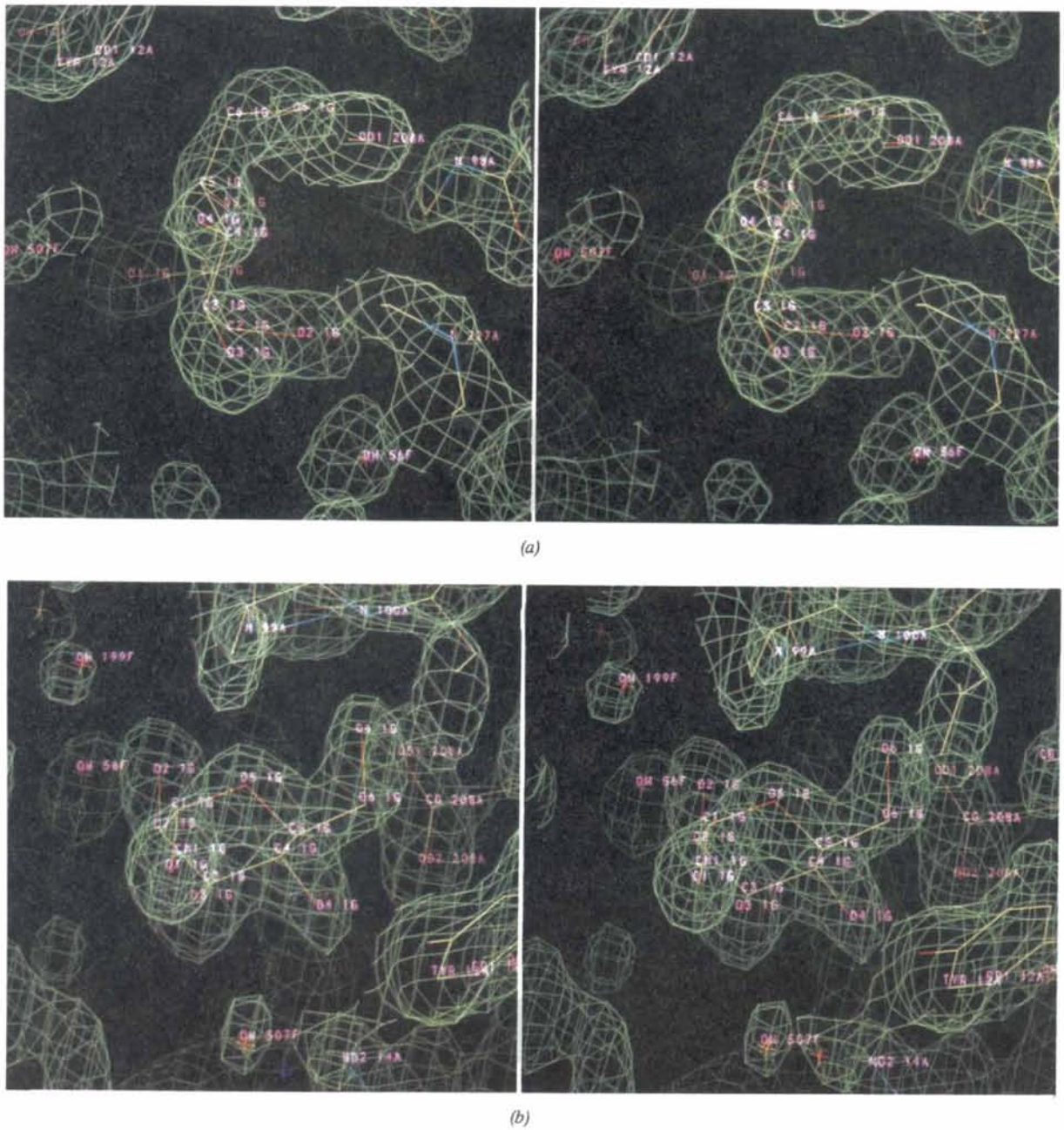


Fig. 4. Stereo diagrams (two views *a* and *b*) of a $(2F_o - F_c)$ map of the saccharide (*A* subunit) electron density, contoured at 1.5 times the r.m.s. of the electron density observed in the unit cell, calculated following several cycles of refinement of a model lacking the sugar.

Table 4. *Protein–saccharide hydrogen bonds for each sugar*

Donor	Acceptor	Hydrogen-bond distances (Å) A	B	C	D
*W*E	O1	3.5	3.3	2.7	—
O2*	*W*F	2.7	2.7	2.7	2.7
*W*G*	O2	3.0	2.8	—	3.3
Leu99 N	O2†	—	—	3.0	—
O3	*W*H	2.9	2.7	2.8	2.7
Arg228 N	O3	2.9	2.8	3.1	3.1
O4	Asp208 OD2	2.6	2.8	2.7	2.8
Asn14 ND2	O4	3.0	3.0	3.1	2.9
Arg228 N	O4†	3.2	3.3	3.2	3.3
Leu99 N	O5	3.1	3.0	2.9	3.2
O6	Asp208 OD1	2.9	2.8	2.9	2.9
Tyr100 N	O6	3.0	3.0	2.9	3.1
Leu99 N	O6†	3.1	3.0	—	3.1

* The classification as donor/acceptor here is arbitrary.
† Weak hydrogen bonds, the angle *D*—H···*A* approaches 80°.

Table 5. *van der Waals protein–saccharide contacts (less than 4.0 Å) in the B subunit*

The contacts for the other three subunits are very similar.

Sugar atom	Residues	Number of contacts
C7	Leu99	2
C1	Leu99	2
O1	—	—
C2	—	—
O2	Gly98, Leu99, Gly227	3
C3	Arg228	1
O3	Gly227, Arg228	5
C4	Asp208, Gly227, Arg228	5
O4	Tyr12, Asn14, Asp208, Arg228	7
C5	Leu99	1
O5	Gly98, Leu99	3
C6	Tyr12, Leu99, Tyr100, Ala207, Asp208	9
O6	Gly98, Leu99, Tyr100, Ala207, Asp208	11
	Total number of contacts	49

Using the classification of Baker & Hubbard (1984) O2, O3, O4 and O6 all form cooperative hydrogen bonds. O4 and O6 form a bidentate hydrogen bond with the carboxylic acid group of Asp208, and O5 and O6 form a bidentate hydrogen bond with backbone amide N atoms of Leu99 and Tyr-100.

The van der Waals contacts are particularly extensive for the region of the saccharide defined by O3, O4, C6 and O6. The surface accessbility of the saccharide decreases from 370 to 100 Å^2 when it binds to the protein. The major losses of accessibility occur at the aforementioned atoms and C7. This can be contrasted with the situation for ribose-binding protein (Mowbray & Cole, 1992) where the sugar is completely occluded from solvent binding.

The differences between the four saccharide sites occur in the water network surrounding O2 and O3 and, to a lesser extent, O1 of the sugar. These networks are involved in crystal packing. Subunits *A* and *B* share the same network whilst subunits *C* and *D*, although similar to each other, are very different from subunits *A* and *B*.

3.4. *Packing interactions of the subunits and water network*

The packing interactions for subunits *A* and *B* are very different from those for subunits *C* and *D* (Derewenda *et al.*, 1989). For the *A* and *B* subunits these occur along the *a* cell edge and are crucially dependent on the water network around the sugars. O1 and C2 of the saccharide in the *A* subunit make a total of three van der Waals contacts with Gly224 and Thr226 of the symmetry-related *B* subunit. C1, O1, C7, C2 and O2 of the sugar in the *B* subunit make a total of six van der Waals contacts with Gly224 and Thr226 of the symmetry-related *A* subunit. An extensive water structure links the saccharide of one subunit to the saccharide of the symmetry-related subunit. Water molecules form several pentagons, with shared edges, along with atoms from the protein and the sugar occupying some of the vertices. Such pentagonal arrangements of water are common in biological materials and are thought to be especially stable (Saenger, 1982). In the *D* subunit *W*G bridges between O2 of the sugar and Asn69 O of the symmetry-related *D* subunit. C1, O1 and C7 of the *D* subunit sugar make a total of five van der Waals contacts with Asn69 of the symmetry-related *D* subunit. In the *C* subunit, *W*B is bound to O2 of the sugar and makes a hydrogen bond to Asp71 OD2 from the symmetry-related *C* subunit. The saccharide-binding sites of all four subunits in this structure are involved in crystal packing.

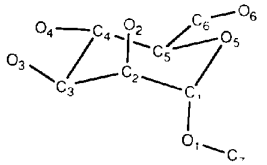


Fig. 5. Methyl α-D-mannopyranoside numbering scheme.

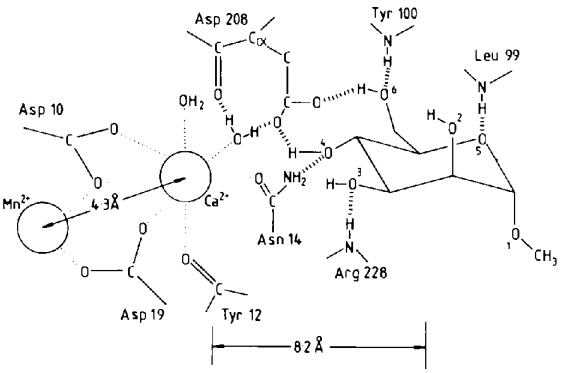


Fig. 6. Schematic diagram of the saccharide binding to concanavalin A.

Table 6. *Metal–ligand distances at* S1 (*mainly manganese*)

The average value is quoted for each ligand with the standard deviation in parentheses for the four subunits.

Ligand	*A* (Å)	*B* (Å)	*C* (Å)	*D* (Å)	Average (Å)	Saccharide-free value* (Å)	Difference (Å)
Glu8 OE2	2.30	2.21	2.38	2.16	2.26 (8)	2.19 (4)	+0.07 (9)
Asp10 OD1	1.97	1.96	2.12	2.17	2.06 (9)	2.13 (1)	−0.07 (9)
Asp19 OD1	2.15	2.31	2.27	2.46	2.30 (11)	2.21 (4)	+0.09 (12)
His24 NE2	2.51	2.53	2.45	2.32	2.45 (8)	2.19 (5)	+0.26 (10)
*W*A O	2.15	2.29	2.05	2.12	2.15 (9)	2.24 (3)	−0.08 (9)
*W*B O	2.08	2.17	2.14	2.11	2.13 (3)	2.22 (6)	−0.09 (7)
Average	2.19	2.25	2.24	2.22	2.23 (13)	2.20 (3)	

* Based on Emmerich *et al.* (1994). The averages and standard deviations here are calculated from the refined saccharide-free structures of (Co, Ca), (Ni, Ca) and native concanavalin A.

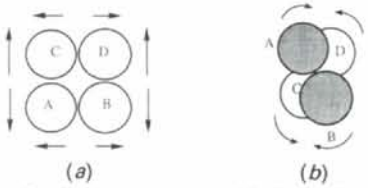
Table 7. *Ligand temperature factors at* S1 (*mainly manganese*)

Ligand	*A* (Å^2)	*B* (Å^2)	*C* (Å^2)	*D* (Å^2)	Saccharide-free value*
Glu8 OE2	20	18	40	20	8
Asp10 OD1	11	16	37	24	7
Asp19 OD1	13	21	42	34	9
His24 NE2	23	30	47	26	10
*W*A O	14	22	46	22	10
*W*B O	13	22	24	19	8
Average	16	22	39	24	9
Manganese	20	18	41	27	9

* From Emmerich *et al.* (1994) *i.e.* (Co, Ca) protein.

3.5. *Metal sites*

Each concanavalin A monomer contains two metal sites denoted S1 and S2. These are separated by 4.2–4.3 Å in the *A*, *B* and *D* subunits but by 4.6 Å in the *C* subunit. The S1 sites are occupied by Mn^{2+} and the S2 sites by Ca^{2+}. In all four subunits the coordination about the manganese ion is basically octahedral. Table 6 shows the bond distances for each ligand. The standard deviation is given for the average length of each metal–ligand bond. This is not really an error limit but a measure of consistency between the subunits. *W*A is hydrogen bonded to *W*D (a calcium ligand) at a distance of 2.8–3.2 Å, to Val32 O (2.8–3.1 Å) and to Glu8 OE1 (2.7–2.8 Å). *W*B is hydrogen bonded to Ser34 OG (2.7–2.8 Å) and to another water molecule. The temperature factors quoted in Table 7 for each ligand show that the temperature factor of the metal is close to the average temperature factor for the ligands. The *C* subunit again stands out as less well ordered than the other three. The average temperature factors of the metal ligands for the *A*, *B* and *D* subunits are below the average temperature factors of the protein, indicating that the metal site is a well ordered part of the structure.



(*a*) (*b*)

Fig. 7. Intersubunit movements on binding of saccharide. (*a*) Vertically the tetramer expands by 0.9 Å. Horizontally this expansion is less at 0.35 Å. (*b*) The overlap of the dimers increases by 3°.

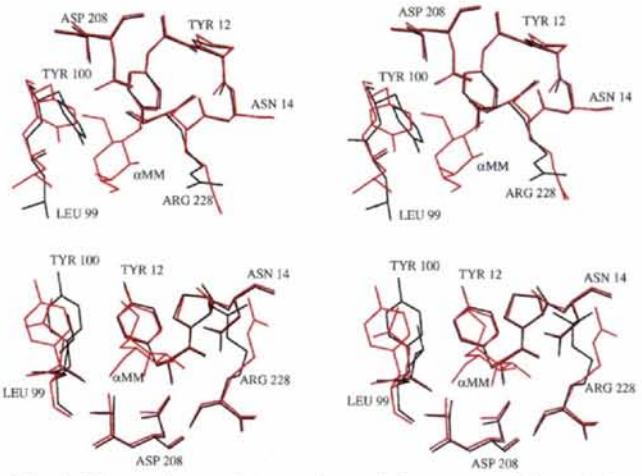


Fig. 8. Two orthogonal stereoviews of the superposition of the saccharide-binding site, with the saccharide-free structure of Emmerich *et al.* (1994) in black and the mannoside complex (*A* subunit) in red, including the sugar molecule.

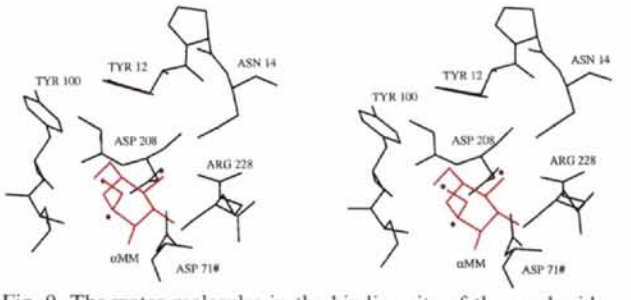


Fig. 9. The water molecules in the binding site of the saccharide-free structure (Emmerich *et al.*, 1994), which are expelled on attachment of the mannoside (latter in red). The three waters make similar hydrogen bonds to the protein as do the sugar O atoms O4, O5 and O6 to Asn14 ND2, Leu99 N and Tyr100 N, respectively (see Table 4 and Fig. 4).

Table 8. *Metal–ligand distances at* S2 *(calcium)*

The average value is quoted for each ligand with the standard deviation in parentheses for the four subunits.

Ligand	*A* (Å)	*B* (Å)	*C* (Å)	*D* (Å)	Average (Å)	Saccharide-free value* (Å)	Difference (Å)
Asp10 OD1	2.77	2.74	(2.94)	2.57	2.76 (13)	2.48 (3)	+0.28 (13)
Asp10 OD2	2.44	2.34	2.57	2.39	2.44 (9)	2.34 (5)	+0.09 (10)
Tyr12 O	2.27	2.26	2.45	2.34	2.33 (8)	2.34 (4)	0.01 (8)
Asn14 OD1	2.47	2.60	2.21	2.50	2.45 (14)	2.50 (9)	−0.05 (17)
Asp19 OD2	2.34	2.46	2.18	2.33	2.33 (10)	2.34 (4)	0.01 (11)
WC O	2.36	2.17	2.37	2.13	2.26 (11)	2.41 (3)	−0.15 (11)
WD O	2.42	2.15	2.21	2.22	2.25 (10)	2.37 (4)	−0.12 (11)
Average	2.44	2.39	2.42	2.35	2.40 (16)	2.38 (6)	

* See footnote to Table 6.

Table 9. *Ligand temperature factors at* S2 *(calcium)*

Ligand	*A* (Å²)	*B* (Å²)	*C* (Å²)	*D* (Å²)	Saccharide-free value‡
Asp10 OD1	11	16	(37)*	24	7
Asp10 OD2	9	19	39	26	9
Tyr12 O	22	24	47	30	9
Asn14 OD1	13	22	49	24	10
Asp19 OD2	14	17	37	34	10
WC O	20	15	56	18	10
WD O	14	18	31	24	7
Average	15	19	38 (37)†	26	9
Calcium	18	20	39	24	8

* The value is shown for completeness although it is outside the Ca^{2+} coordination sphere.

† The figure in parentheses includes the Asp10 OD1 temperature factor.

‡ From Emmerich *et al.* (1994) *i.e.* (Co, Ca) protein.

Using an upper limit of 2.8 Å for a calcium–oxygen bond for a bidentate carboxylic group (Einspahr & Bugg, 1981) the calcium ion is seven coordinate in the *A*, *B*, and *D* subunits but six coordinate in the *C* subunit. The individual ligand bond distances are shown in Table 8 and the temperature factors are shown in Table 9. The temperature factors show a similar distribution to that observed for the manganese ion. In the *A*, *B* and *D* subunits the calcium has pseudo octahedral geometry with Asp10 binding in a bidentate manner capping the sixth vertex of an octahedron. In the *C* subunit Ca^{2+} coordination appears to be octahedral, however, errors in bond distances mean the calcium may be seven coordinate. *W*C is hydrogen bonded to Asp208 O at a distance of 2.9–3.0 Å and to Asp208 OD2 (2.5–2.8 Å). *W*D is hydrogen bonded to Arg228 O (2.8–3.0 Å) and to *W*A (Mn^{2+} ligand) (2.9–3.0 Å). It is the *W*C interaction with Asp208 which is thought to stabilize the *cis* peptide (Hardman, Agarwal & Freiser, 1982).

3.6. *Comparison to saccharide-free concanavalin A*

Full details of the refined 2.0 Å structure of saccharide-free concanavalin A at 2.0 Å resolution have been reported firstly for (Cd, Ca) (Naismith *et al.*, 1993) and secondly, native (Weisgerber & Helliwell, 1993) concanavalin A. These refinements were performed against synchrotron film data (λ = 1.488 Å), and synchrotron Laue 'toast-rack' film data (λ_{min} = 0.5 Å and λ_{max} = 0.9 Å), respectively. The coordinates have been deposited with the Protein Data Bank (PDB entry codes 1CON and 2CTV, respectively). Refinement at 1.6 Å resolution of the (Co, Ca) saccharide-free crystal structure, against synchrotron image-plate data (λ = 1.009 Å), and refinement at 2.0 Å resolution of the (Ni, Ca) concanavalin A structure, against R-AXIS IIc Cu $K\alpha$ image-plate data, has also now been completed (Emmerich *et al.*, 1994). These structures are essentially the same, apart from differences in the case of Cd-substituted concanavalin A due to the larger ionic radius of the cadmium ion (Naismith *et al.*, 1993) and the improved precision of the 1.6 Å structure refinement. The Protein Data Bank codes for the 2.0 Å (Ni, Ca) and the 1.6 Å (Co, Ca) structures are 1SCR and 1SCS, respectively.

The saccharide-free structure crystallizes in the space group *I*222 and the asymmetric unit contains one monomer. A similar tetramer to that shown in Fig. 1 exists for the saccharide-free structure although the monomers are related by crystallographic rather than non-crystallographic twofold axes. A superposition of the tetramers indicates that the precise organization of the tetramer is subtly different in the two structures. The differences can be represented as three distinct movements which are shown schematically in Fig. 7. In the saccharide-free structure the monomers are more tightly packed together than in the complex. Hence the number of hydrogen bonds and van der Waals contacts at subunit interfaces is approximately 35% higher than in the mannoside complex. There are fewer changes for the *AB* and *CD* dimers than for the *AC* and *BD* dimers.

The fold of the concanavalin A monomer is essentially unchanged by saccharide binding and the saccharide-free monomer superimposes onto each of the monomers in the mannoside complex with an r.m.s. deviation of 0.4 Å for all backbone atoms and less than 0.2 Å for residues in β-sheets. A residue-by-

residue comparison of the mannoside complex and the saccharide-free structure shows several differences in individual amino-acid conformations remote from the saccharide-binding site.

The largest of these is centred at Trp182. In the saccharide-free structure Trp182 adopts an uncommon conformation (as defined by McGregor, Islam & Sternberg, 1987) due to crystal packing effects [these effects are discussed in detail elsewhere (Naismith *et al.*, 1993)]. In the absence of these packing constraints in the mannoside–concanavalin A complex, Trp182 is found in a more common conformation. The change at Trp182 affects the surrounding protein structure.

The comparison at the saccharide-binding site is complicated because in both this structure and the saccharide-free structure the binding sites are involved in crystal packing. A superposition of the sites is shown in Fig. 8. The main-chain positions are very similar in both structures, however the conformation of several of the side chains has altered. Tyr100 is directly affected in a major way by saccharide binding whereby it rotates out of the saccharide-binding site to avoid steric clashes with the sugar. Tyr12 does move slightly towards the saccharide site, to form van der Waals interactions with the sugar. The differences in the other side chains in the binding site, particularly Leu99 and Arg228, although bigger than for Tyr12, could well be due to the different packing interactions. In the saccharide-free structure, water molecules and the side chain of Asp71, from a neighbouring molecule in the crystal, fill the binding site and form hydrogen bonds with the same residues as does the sugar. This is shown in Fig. 9. The three waters expelled, interacted with Asn14 ND2, Leu99 N and Tyr100 N and are replaced by O4, O5 and O6 of the sugar (see also Table 4 and Fig. 4).

The metal–ligand distances at S1 and S2 for the mannoside-bound and saccharide-free structures are given in Tables 6 and 8. At S1 in the saccharide-free case the ligands sit at the vertices of an almost perfectly regular octahedron with an average distance of 2.20 (3) Å. However, in the saccharide-bound case, although the average distance is very similar, there is clearly a larger range of ligand distances [2.23 (13) Å]. The largest difference is for the distance from manganese to the ligand His24 NE2 which increases by 0.26 (10) Å. There is also a larger range of ligand distances to the Ca [2.40 (16) Å] in the saccharide-bound case compared with the saccharide-free case [2.38 (6) Å]. The largest difference is for the distance from the calcium to the ligand Asp10 (OD1) which increases by 0.28 (13) Å. Finally, the S1 to S2 distances are increased slightly in binding of saccharide to 4.2–4.3 Å in the *A*, *B* and *D* subunits (4.6 Å in *C*) compared with 4.17 (1) Å in the saccharide-free structure. These differences are, however, at the limit of the present analysis, particularly of the mannoside–concanavalin A refinement. An even higher resolution study would be of interest, therefore.

Table 10. *Temperature factors of the saccharide-binding loops in the saccharide-free and saccharide-bound structures compared with the whole subunit in each case*

Structure	Average temperature factor Loop (Å²)	Subunit (Å²)
*I*222	17	19
Subunit *A*	21	28
Subunit *B*	22	28
Subunit *C*	53	42
Subunit *D*	29	31

The temperature factors of the structures are dramatically different. The average temperature factor for the *I*222 structure is 19 Å², considerably lower than that observed for any subunit in the mannoside complex. However the average temperature factor of the saccharide-binding loops can be compared with the average for a whole subunit (Table 10). According to this analysis saccharide binding does not significantly affect the ordering of the binding loops relative to the bulk of the structure.

4. Concluding remarks

The results of the 2.9 Å study of Derewenda *et al.* (1989) with respect to the saccharide-binding site, detailing how each sugar is bound to the protein by hydrogen bonds and van der Waals contacts, are confirmed. In addition however, due to the higher resolution of this study, we have been able to resolve the methyl substituent on the saccharide. Moreover, we have been able to locate a water ligand to the calcium ion, which is involved in stabilization of the *cis* peptide bond between Ala207 and Asp208, in the saccharide-complex structure. Also, the surface accessibility of the saccharide molecule decreases from 370 to 100 Å² when it binds to the protein.

Comparison of the saccharide complex at 2.0 Å with the saccharide-free structure of concanavalin A (cobalt substituted at the transition-metal site) at 1.6 Å reveals considerable structural information. Overall, in the saccharide complex the tetramer association is different in that the monomers in the tetramer tend to separate by up to 0.9 Å resulting in only 65% of the inter-subunit interactions seen in the saccharide-free structure. Moreover, the average temperature factor of the mannoside complex is considerably higher (32 Å²) than that of the saccharide-free protein (19 Å²). Evidently there is an overall loosening of the structure on binding of the saccha-

ride molecules to each monomer. There are also small changes in the coordination distances of the metal sites on binding of the saccharide, but these are at the limits of the accuracy of these refinements. Major changes at the saccharide site on binding involve the reorientation of Tyr100 to accommodate the saccharide and the expulsion of three ordered solvent molecules.

The Carnegie Trust for the Universities of Scotland is thanked for the provision of a scholarship to JHN. The use of the CSSR and SEQNET databases (SERC, Daresbury Laboratory) is gratefully acknowledged. Synchrotron radiation facilities and support were kindly provided by SERC, Daresbury Laboratory, apart from the saccharide-free (Co, Ca) concanavalin A study which was based on data collected at EMBL, Hamburg to whom we are also very grateful. The University of Manchester, The Hasselblad Foundation and The Wellcome Trust are thanked for their support. JH and SJH were supported by the SERC under the Molecular Recognition Initiative. JY was an SERC visiting fellow during the period of this work. Gordon Leonard and Sue Bailey are thanked for helpful discussions.

References

ALTER, G. M. & MAGNUSON, J. A. (1974). *Biochemistry*, **13**, 4038–4045.

ARNDT, U. W. & WONACOTT, A. J. (1977). Editors. *The Rotation Method in Crystallography*. Amsterdam: North-Holland.

BAKER, E. N. & HUBBARD, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.

BECKER, J. W., REEKE, G. N. JR, CUNNINGHAM, B. A. & EDELMAN, G. M. (1976). *Nature (London)*, **259**, 406–409.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BREWER, C. F., BROWN, R. D. & KOENIG, S. H. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 1007–1011.

BREWER, C. F., STERNLICHT, H., MARCUS, D. M. & GROLLMAN, A. P. (1973). *Biochemistry*, **12**, 4448–4457.

BRÜNGER, A. T. (1990). *X-PLOR Version* 2.1 *Manual*. Howard Hughes Medical Institute, Yale Univ., USA.

DEREWENDA, Z., YARIV, J., HELLIWELL, J. R., KALB (GILBOA), A. J., DODSON, E. J., PAPIZ, M. Z., WAN, T. & CAMPBELL, J. (1989). *EMBO J.* **8**, 2189–2193.

EDELMAN, G. M., CUNNINGHAM, B. A., REEKE, G. N., BECKER, J. W., WAXDAL, M. J. & WANG, J. L. (1972). *Proc. Natl Acad. Sci. USA*, **69**, 2580–2584.

EINSPAHR, H. & BUGG, C. A. (1981). *Acta Cryst.* B**37**, 1044–1052.

EMMERICH, C., HELLIWELL, J. R., REDSHAW, M., NAISMITH, J. H., HARROP, S. J., RAFTERY, J., KALB (GILBOA), A. J., YARIV, J., DAUTER, Z. & WILSON, K. S. (1994). *Acta Cryst.* D**50**, 749–756.

GREER, J., KAUFMAN, H. W. & KALB, A. J. (1972). *J. Mol. Biol.* **48**, 365–366.

HARDMAN, K. D., AGARWAL, R. C. & FREISER, M. J. (1982). *J. Mol. Biol.* **157**, 69–86.

HARDMAN, K. D. & AINSWORTH, C. F. (1972). *Biochemistry*, **11**, 4910–4919.

HARDMAN, K. D. & AINSWORTH, C. F. (1976). *Biochemistry*, **15**, 1120–1128.

HARROP, S. J., NAISMITH, J. H., EMMERICH, C., HABASH, J., WEISGERBER, S., KALB (GILBOA), A. J., YARIV, J. & HELLIWELL, J. R. (1993). *Acta Cryst.* A**49**, C-94.

HELLIWELL, J. R., GREENHOUGH, T. J., CARR, P. D., RULE, S. A., MOORE, P. R., THOMPSON, A. W. & WORGAN, J. S. (1982). *J. Phys. E*, **5**, 1363–1372.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

KABSCH, W. & SANDER, C. (1984). *Biopolymers*, **22**, 2577–2637.

KALB (GILBOA), A. J. & LEVITZKI, A. (1968). *Biochem. J.* **109**, 669–672.

KALB (GILBOA), A. J. & LUSTIG, A. (1968). *Biochim. Biophys. Acta*, **168**, 366–367.

LUZZATI, P. V. (1952). *Acta Cryst.* **5**, 802–810.

MCGREGOR, M. J., ISLAM, S. & STERNBERG, M. J. E. (1987). *J. Mol. Biol.* **162**, 419–444.

MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

MIN, W., DUNN, A. J. & JONES, D. H. (1992). *EMBO J.* **11**, 1303–1307.

MOWBRAY, S. L. & COLE, L. B. (1992). *J. Mol. Biol.* **115**, 155–175.

NAISMITH, J. H., HABASH, J., HARROP, S. J., HELLIWELL, J. R., HUNTER, W. N., WAN, T. C. M., WEISGERBER, S., KALB (GILBOA), A. J. & YARIV, J. (1993). *Acta Cryst.* D**49**, 561–571.

RAMAKRISHNAN, C. & RAMACHANDRAN, G. N. (1965). *Biophys. J.* **5**, 909–933.

REEKE, G. N. JR & BECKER, J. W. (1986). *Science*, **24**, 1108–1111.

REEKE, G. N. JR, BECKER, J. W. & EDELMAN, G. M. (1975). *J. Biol. Chem.* **250**, 1525–1547.

SAENGER, W. (1982). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.

SERC Daresbury Laboratory (1979). *CCP4. A Suite of Programs for Protein Crystallography*, SERC Daresbury Laboratory, Warrington WA4 4AD, England.

SUMNER, J. B. (1919). *J. Biol. Chem.* **37**, 137–142.

VILLAFRANCA, J. J. & VIOLA, R. E. (1974). *Arch. Biochem. Biophys.* **160**, 465–468.

VYAS, N. K. (1991). *Curr. Opin. Struct. Biol.* **1**, 732–740.

WEISGERBER, S. & HELLIWELL, J. R. (1993). *J. Chem. Soc. Faraday Trans.* **89**(15), 2667–2675.

YARIV, J., KALB (GILBOA), A. J. & LEVITZKI, A. (1968). *Biochim. Biophys. Acta*, **165**, 303–305.

YARIV, J., KALB (GILBOA), A. J., PAPIZ, M. Z., HELLIWELL, J. R., ANDREWS, S. J. & HABASH, J. (1987). *J. Mol. Biol.* **195**, 759–760.