Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of concanavalin A at pH 8: bound solvent and crystal contacts

F. J. López-Jaramillo,[a]* L. A. González-Ramírez,[a] A. Albert,[b] F. Santoyo-González,[c] A. Vargas-Berenguel[d] and F. Otálora[a]

[a]Laboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra, CSIC-UGRA, Facultad de Ciencias, Campus Fuentenueva, E-18002 Granada, Spain, [b]Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química Física Rocasolano, CSIC, Serrano 119, E-28006 Madrid, Spain, [c]Instituto de Biotecnología, Facultad de Ciencias, Universidad de Granada, E-18071 Granada, Spain, and [d]Area de Química Orgánica, Universidad de Almería, 04120 Almeria, Spain

Correspondence e-mail: javier@lec.ugr.es

Received 26 May 2003
Accepted 24 March 2004

PDB Reference: Mn–Mn ConA, 1nxd, r1nxdsf.

Concanavalin A has been crystallized in the presence of the ligand (6-*S*-β-D-galactopyranosyl-6-thio)-cyclomaltoheptaose. The crystals are isomorphous to those reported for ConA complexed with peptides at low resolution (3.00–2.75 Å). The structure was solved at 1.9 Å, with free *R* and *R* values of 0.201 and 0.184, respectively. As expected, no molecules of the ligand were bound to the protein. Soaking in the cryobuffer left its fingerprint as 25 molecules of glycerol in the bound solvent, most of them at specific positions. The fact that a glycerol molecule is located in the sugar-binding pocket of each of the four subunits in the asymmetric unit and another is located in two of the peptide-binding sites suggests a recognition phenomenon rather than a displacement of water molecules by glycerol. Crystal contact analysis shows that a relation exists between the residues that form hydrogen bonds to other asymmetric units and the space group: contact Asp58–Ser62 is a universal feature of ConA crystals, while Ser66–His121, Asn69–Asn118 and Tyr100–His205 contacts are general features of the $C222_1$ crystal form.

1. Introduction

Traditionally, studies of carbohydrate–protein interactions have been carried out using lectins, proteins that are present in plants, animals and bacteria and possess the common feature of binding carbohydrates with considerable specificity. The basis of this specificity has been explained as a consequence of the variability in the size of the carbohydrate-recognition domain, post-translational modifications and the degree of oligomerization (Vijayan & Chandra, 1999). However, the relationship between structure and thermodynamics is difficult to rationalize (Bradbrook *et al.*, 1998) because lectins are a diverse group and exhibit several folds corresponding to different carbohydrate-binding motifs (Rini, 1995 and references therein; Rini & Lobsanov, 1999; Bouckaert *et al.*, 1999).

Concanavalin A (ConA) is the most studied member of the legume lectin family. It was isolated from the jack bean (*Canavalia ensiformis*) and crystallized almost nine decades ago by Sumner (1919). It exists as a dimer in solution at pH values below 6.5; above pH 6.8 it exists as a tetramer of identical subunits, each with two metal-ion sites in its native state: the S1 site, which is usually occupied by a transition-metal ion, preferentially Mn^{2+}, and the S2 site, which binds a calcium ion. The binding of these ions to metal-free ConA (the unlocked form) leads to the reorganization of the metal-binding region, the key event being the *trans*-to-*cis* isomerization of an Ala-Asp peptide bond, which yields the locked form with higher carbohydrate-binding activity (Loris *et al.*, 1998 and references therein). However, Mn^{2+} at S1 can be replaced by Ni^{2+}, Co^{2+}, Zn^{2+}, Cd^{2+} or Ca^{2+} and Ca^{2+} at S2 by

Cd^{2+} or Mn^{2+} and this phenomenon may be explained on the basis of the ionic radii of the metals: $Ni^{2+} \simeq Zn^{2+} \leq Co^{2+} < Mn^{2+} < Cd^{2+} \simeq Ca^{2+}$ (Sadhu & Magnuson, 1989). Other metal-binding sites have been described (Naismith *et al.*, 1993; Bouckaert, Loris *et al.*, 2000; Kantardjieff *et al.*, 2002), but their physiological relevance remains unclear.

Each subunit of ConA has one specific carbohydrate-binding site with a high affinity for mannose (Derewenda *et al.*, 1989) and glucose (Bradbrook *et al.*, 1998) but none for galactose (Shaanan *et al.*, 1991), although the recently solved structures of ConA complexed with peptides of different natures and lengths have demonstrated the existence of a peptide-binding region that differs from the well known sugar-binding pocket (Jain *et al.*, 2000, 2001*a*,*b*). All these peptides contain the motif Tyr-Pro-Tyr and it has been suggested that its aromatic nature corresponds to that associated with carbohydrate moieties and that its conformation may resemble a region of the complex carbohydrates that bind to ConA (Jain *et al.*, 2001*a*).

Protein–carbohydrate interactions are ubiquitous in biology and important for processes such as immune responses, cell–cell and host–pathogen interaction or inflammation. These interactions are specific and it has been predicted that they may form novel targets for new therapeutic approaches (Sharon & Lis, 1972). In addition to the design of new drugs, the specificity of the protein–carbohydrate interaction is of interest for the development of new drug carriers (Kompella & Lee, 2001), such as biodegradable nanoparticles, that would target the drug to the site of action and reduce its toxicity and systemic side effects (Soppimath *et al.*, 2001).

β-Cyclodextrins are one of these potential drug-delivery systems. They consist of seven glucose units linked by $\alpha(1–4)$ bonds in a cyclic fashion that leads to a ring-shaped structure with a hydrophilic outer surface and a hydrophobic inner cavity large enough to accommodate guest molecules. Cyclodextrins have biotechnological applications (Singh *et al.*, 2002) and in the pharmaceutical industry they are being used as complexing agents to increase the aqueous solubility of poorly water-soluble drugs and to improve their bioavailability and stability. Reviews exist describing their potential application in the delivery of opthalmic drugs (Loftssona & Järvinen, 1999), transdermal and rectal drugs (Matsuda & Arima, 1999), nasal drugs (Merkus *et al.*, 1999), topical drugs (Loftsson & Masson, 2001), peptides and proteins (Irie & Uekema, 1999) or nucleotides (Redenti *et al.*, 2001).

However, one step forward is the selective modification of the outer surface of cyclodextrins in order to control and direct their interaction with the target cells. We have focused on a model system consisting of ConA as the target and β-cyclodextrins with different chemical substituents as potential carriers. In this paper, we describe the crystallization and structure at 1.9 Å of ConA in the presence of heptakis (6-*S*-β-D-galactopyranosyl-6-thio)-cyclomaltoheptaose, a cyclodextrin with galactose as substituent which is not expected to interact with ConA, and both Mn^{2+} and Ca^{2+}. The structure contains 25 molecules of glycerol from the cryo-buffer, most of which are located at conserved positions and, as expected, no molecules of the substituted β-cyclodextrin. We analyze the crystal contacts in the orthorhombic crystals on the basis of the residues responsible for hydrogen bonding and find a relation between the residues that form hydrogen bonds and the space group.

Table 1
Summary of the main features of the data set and refinement.

Values in parentheses correspond to the highest resolution shell.

Data collection		
X-ray source	BW7B (EMBL/DESY)	Cu $K\alpha$
Wavelength (Å)	0.84	1.5418
No. crystals	1	1
Data-collection temperature (K)	100	100
Space group	$C222_1$	$C222_1$
Unit-cell parameters (Å)		
a	101.3	102.8
b	118.0	118.2
c	249.5	252.5
Data set		
Resolution range (Å)	19.00–1.90 (2.00–1.90)	14.00–3.50 (3.69–3.50)
Measured reflections	324343 (40384)	93918 (12589)
Unique reflections	111379 (15414)	18264 (2338)
Completeness $[I/\sigma(I) \geq 0]$ (%)	94.9 (93.0)	93.8 (93.8)
Completeness $[I/\sigma(I) \geq 3]$ (%)	94.0 (89.5)	93.8 (93.8)
$R_{\rm merge}$ (%)	2.0 (4.8)	14.1 (18.8)
Mean $I/\sigma(I)$	36.35 (18.15)	4.3 (3.3)
Refinement		
Resolution range (Å)	19.00–1.90 (1.97–1.90)	
Free *R* value, random 10%	0.201 (0.190)	
R value	0.184 (0.190)	
No. water molecules	764	
No. metal atoms	8 Mn	
No. glycerol molecules	25	
R.m.s.d. covalent bonds† (Å)	0.007	
R.m.s.d. bond angles† (°)	1.70	
Overall coordinate error (DPI)‡ (Å)	0.11	
Correlation coefficient§	0.934	
Optical resolution¶ (Å)	1.44	

† Deviation from the standard dictionary (Engh & Huber, 1991). ‡ Diffraction-data precision indicator (Cruickshank, 1996). § Correlation coefficient between calculated and observed structure-factor amplitude output by *SFCHECK* (Vaguine *et al.*, 1999). ¶ Expected minimum distance between two resolved peaks in the electron-density map as implemented in *SFCHECK*.

2. Materials and methods

Concanavalin A type V was purchased from Sigma (Lot 13H7045) and heptakis (6-*S*-β-D-galactopyranosyl-6-thio)-cyclomaltoheptaose (*i.e.* the substituted β-cyclodextrin) was synthesized as described by García-López *et al.* (1999). The protein was dissolved to a final concentration of 10 mg ml^{-1} in 100 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 1 m*M* $MnCl_2$, 1 m*M* $CaCl_2$, 3 m*M* NaN_3 and 4.2 m*M* heptakis (6-*S*-β-D-galacto-pyranosyl-6-thio)-cyclomaltoheptaose. The precipitating system consisted of 12%(*w*/*v*) PEG 6000 in 100 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 1 m*M* $MnCl_2$, 1 m*M* $CaCl_2$ and 3 m*M* NaN_3. Crystals were grown by the standard gel-acupuncture method (GAME; Garcia-Ruiz & Moreno, 1997). Briefly, an X-ray capillary with internal diameter of 0.5 mm was filled with about 12 µl protein solution, its upper side was sealed with nail varnish and the capillary was then punctuated into a

gel layer. The precipitating system was poured over the gel and the system was incubated at 293 K. The precipitant diffused through the gel layer and along the capillary to yield crystals, which started to appear after 10 d.

X-ray diffraction took place at 100 K using a single crystal extracted from the capillary and cryoprotected by soaking in increasing concentrations of glycerol up to 20%(*w*/*v*) in 100 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 1 m*M* $MnCl_2$, 1 m*M* $CaCl_2$, 2 m*M* NaN_3 and 6%(*w*/*v*) PEG 6000. Data collection was carried out on a MAR 345 image-plate detector (MAR Research) in dose mode at EMBL/DESY beamline BW7B (λ = 0.84 Å). Prior to starting data collection, three 0.2° oscillation images were autoindexed with *MOSFLM* 6.01 (Leslie, 1999) to predict the strategy to achieve the highest completeness. Low-resolution data to 3.35 Å were collected in 200 images with 0.25° oscillation and the detector placed at 350 mm. High-resolution data to 1.8 Å were collected with the detector at 300 mm as four non-overlapping wedges, each consisting of 75 images with 0.20° oscillation.

Data were processed with the *XDS* software package (Kabsch, 1988, 1993). Neither overloads nor incomplete reflections (*i.e.* those whose intensity estimate during profile fitting was less than 75% of the observed intensity) were included in processing. A forward-directed search strategy (Kabsch, 1993) was followed: no unit cell or space group was imposed. The program was asked to determine the reduced cell, to compare it with all 44 possible lattices and to index the data in *P*1. The diffraction parameters and unit cell were then refined in $C222_1$ and *C*222, which *XDS* gave as the space groups with the highest likelihood of being correct. The data set was finally re-indexed with the refined unit cell in $C222_1$ and *C*222 and independently scaled with *XSCALE*.

The structure was solved by molecular replacement with the program *AMoRe* (Navaza, 1994), using the atomic resolution structure at 0.94 Å reported by Deacon *et al.* (1997) (PDB code 1nls) as a search model, omitting the solvent molecules. The cross-rotation search was carried out in the resolution range 15–3 Å with a radius of the integration sphere of 43.5 Å and 101.5, 86.5 and 84.1 Å as the dimensions of the model cell. For the translation searches the resolution range was 15–2 Å. The method of Crowther & Blow (1967) was used for the first molecule and the phased translation function (Bentley, 1992) was used for the others, fixing the solutions from the previous translation-search stages. At this stage, the space group was confirmed as $C222_1$ and the possible solutions for the positions of the four molecules in the asymmetric unit were refined by fast rigid-body refinement (Castellano *et al.*, 1992) and visualized using *XtalView* (McRee, 1993).

The structure was further refined with *CNS* 1.0 (Brünger *et al.*, 1998) by torsion-angle molecular dynamics and cross-validated maximum-likelihood target function (Adams *et al.*, 1997), 10% of the data set being randomly selected for the estimation of the free *R* value (Brünger, 1992). Each cycle typically consisted of 200 steps of geometry energy minimization followed by three trials with different starting velocities of torsion molecular dynamics, starting from a temperature of 2000 K and decreasing in 25 K steps to 300 K, plus 100 steps of Cartesian molecular dynamics at a constant temperature of 300 K and 100 steps of final conjugate-gradient energy minimization followed by 30 cycles of restrained individual *B*-factor refinement. Rebuilding of the model was carried out with *XtalView* on the basis of σ_A-weighted maps and the stereochemical acceptability of the structure was verified using *PROCHECK* (Laskowski *et al.*, 1993). No σ cutoff was applied and a solvent mask was automatically determined and updated by *CNS* to use the resolution range 19.0–1.9 Å. The inclusion of solvent molecules and ligands was based on peaks of at least 3σ and 1σ in the σ_A-weighted $F_o - F_c$ and $2F_o - F_c$ maps, respectively. Evaluation of the agreement between the final model and the structure-factor data was carried out with *SFCHECK* (Vaguine *et al.*, 1999) and the crystal contacts with strong possibility of hydrogen bonding (distance < 3.3 Å) were identified using the *CCP*4-supported program *CONTACT* (Collaborative Computational Project, Number 4, 1994).

In order to identify the chemical nature of the metal centre, a 3 Å data set consisting of 500 images with 0.3° oscillation was collected using a Cu $K\alpha$ radiation on an in-house source (MAR 345 image-plate detector on an Enraf–Nonius rotating-anode generator). The data-collection strategy was carefully chosen in order to enhance the signal-to-noise ratio by increasing the crystal-to-detector distance and the data-set redundancy. Data reduction was carried out with *MOSFLM* 6.01 (Leslie, 1999) and data were scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994). All calculations to compute the anomalous difference electron-density map were performed using the *CCP*4 package (Collaborative Computational Project, Number 4, 1994).

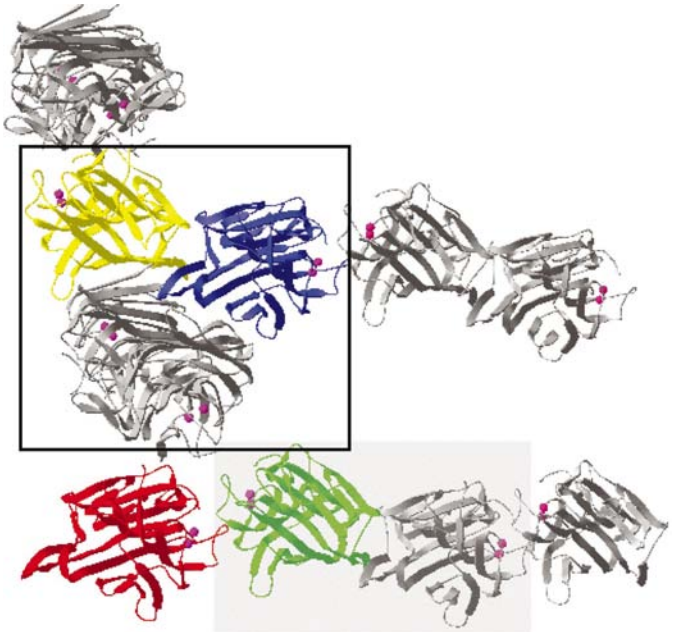

Figure 1
The four monomers contained in the asymmetric unit are shown in yellow (*A*), blue (*B*), green (*C*) and red (*D*). They are related by symmetry operations with others from different asymmetric subunits to form the tetrameric form (enclosed in the square) and the canonical dimer (shadowed). For the sake of clarity, Mn^{2+} is coloured magenta.

3. Results and discussion

This work was carried out in the frame of an effort intended to lead to a rational design of carriers to deliver a drug to its site of action. Our approach is the selective chemical modification of the outer surface of the cyclodextrin (*i.e.* the carrier) in order to control and direct its interaction with ConA (*i.e.* the model target). In order to evaluate the feasibility of our model system, it is critical to know whether the glucoses of the cyclodextrin ring play any role in the interaction with ConA. Therefore, we have crystallized ConA in the presence of heptakis (6-*S*-β-D-galactopyranosyl-6-thio)-cyclomaltoheptaose, a substituted cyclodextrin consisting of a D-glucose ring that can interact with ConA and which has D-galactose molecules as substituents on its upper side that do not bind to the ConA.

Table 2
Residues involved in the interaction with azide.

Site	Subunit	Residues forming hydrogen bonds
Az1	*A*	Ser110, Phe111, Thr112, Glu192, Thr194
Az1	*C*	Ser110, Phe111, Thr112, His127, Glu192, Wat
Az2	*B*	Lys138, Leu140, Wat
Az2	*D*	Lys138, Leu140, Wat
Az3	*B*	Thr12*B*, Pro13*B*, Wat, Wat
Az3	*D*	Thr12*D*, Pro13*B*, Wat, Wat, Wat, Wat
Az4	*A*	Thr105, Glu155, Wat, Wat
Az4	*B*	Thr103, Thr105, Glu155, Lys200, Wat

The crystallization was carried out by counter-diffusion in capillaries (Ng *et al.*, 2003 and references therein) with a molar excess of the substituted β-cyclodextrin in the presence of Ca^{2+} and Mn^{2+} at alkaline pH. The preliminary characterization using the rotating-anode generator showed that the crystals were orthorhombic, with Bravais lattice *C*, unit-cell parameters $a = 101.3$, $b = 118.0$, $c = 249.5$ Å and four molecules per asymmetric unit. This unit cell in space group $C222_1$ was described by Hardman & Ainsworth (1976) for the complex with methyl-α-D-mannopyranoside, but the structure was not solved and they concluded that the crystal lattice was disordered. More recently, Moothoo & Naismith (1999) reported it for the second crystal form of concanavalin A (ConA) co-crystallized with fructose, although the structure was again not solved. It was finally solved by Jain *et al.* (2000, 2001*a*,*b*) for ConA complexed with peptides of different lengths and compositions.

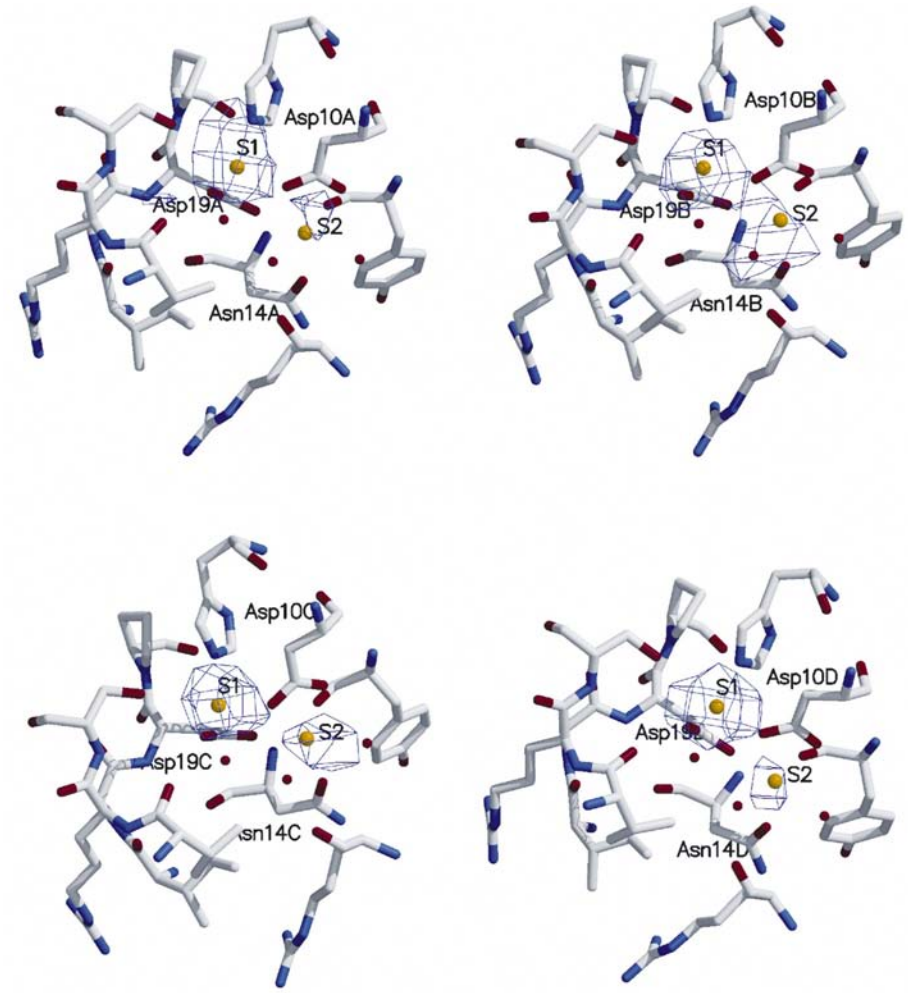

Figure 2
Difference anomalous map of the metal-binding site of each subunit contoured at 3σ. The metal ions are shown in yellow.

The facts that the estimated solvent content is 66%, which is large enough to accommodate the substituted β-cyclodextrin, that this crystal form is associated with ConA complexed with either sugars or peptides and that the best resolution for these structures was 2.75 Å encouraged us to collect a full data set from a single flash-cooled crystal at BW7B (EMBL Hamburg Outstation). It diffracted beyond 1.8 Å, but for structural purposes the data set consisted of 111 372 unique reflections to a resolution limit of 1.9 Å, with a redundancy of 2.9, an R_{sym} of 2.0% and 94.9% completeness. The structure was solved by molecular replacement using PDB entry 1nls at 0.94 Å (Deacon *et al.*, 1997) as a search model and was refined using *CNS* (Brünger *et al.*, 1998) to free *R* and *R* values of 0.201 and 0.184, respectively. The atomic coordinate error estimated by the method of Cruickshank (1996) was 0.11 Å and the density correlation coefficient between the electron density computed from calculated and observed structure-factor amplitudes as output by *SFCHECK* was 0.934 (Vaguine *et al.*, 1999). The structure has good stereochemistry, with 85.6% residues in the most favoured regions of the Ramachandran plot and the remaining 14.4% in the additional allowed regions. The

Table 3
Residues involved in the interaction with the glycerol (Gol) used as cryoprotectant.

Residues in parentheses correspond to symmetry-related molecules.

Site	Subunit	Gol	Residues forming hydrogen bonds	Remarks
G1	*A*	7	Asn14, Gly98, Leu99, Tyr100, Asp208, Wat	Asn14, Leu99, Tyr100, Asp208, Arg228 are involved in carbohydrate recognition
	B	14	Gly98, Leu99, Tyr100, Asp208, Arg228, Wat	
	C	3	Asn14, Gly98, Leu99, Tyr100, Asp208, Arg228, Wat	
	D	5	Asn14, Gly98, Leu99, Tyr100, Asp208, Arg228, Wat, Wat	
G2	*A*	12	His180, Ser113, Ala125, Wat, Wat	
	B	11	His180, Ser113, Wat, Wat	
	C	2	His180, Ser113, Ala125, (Asp139), Wat, (Wat)	
	D	13	Leu126, Pro178, His180, (Asp139), (Wat), (Wat)	
G3	*A*	22	Gly144, Ile221, Wat	
	B	18	Gly144, Ile221, Wat	
	D	19	Gln143, Gly144, Ser220, Ile221, Wat, Wat	
G4	*A*	4	(Pro13), (Thr15), Asn44, Ser201, Asp203, Wat, Wat, (Wat)	Asn44, Ser201 are key residues involved in peptide recognition
	D	10	Pro13, Thr15, Asn44, Ser201, Asp203, Wat, Wat, Wat, Azi	Pro13 and Asp203 are involved in the interaction with a 15-mer and 10-mer peptides, respectively (Jain *et al.*, 2001*a*)
G5	*A*	24	Lys138, Leu140, Wat, Wat, Wat, Wat, Wat	The cryoprotectant MPD has been found hydrogen bonded to Gln137 and Asp139 (Kantardjieff *et al.*, 2002)
	C	25	Lys138, Leu140, Wat, Wat	
G6	*A*	9	Ser72, Thr74, Wat, Wat, Wat, Wat	
	B	23	Ser72, Thr74, Wat, Wat	
	A	21	Tyr12, Pro13, Thr15,Asp16, Wat	Involved in crystal packing in *I*222
	A	6	Ser223, (Ser168), Wat, Wat, Wat, (Wat)	
	B	15	Tyr12, (Ser204), Wat, (Wat)	
	B	20	Arg228, Pro234, Ala236	
	C	8	Trp88, Ile127, Asp218, Wat	
	C	16	Tyr54, lys59, Ile181, Glu183, Wat	
	D	1	Asp235, Asn237, Wat, Wat, Wat	
	D	17	(Gly224), Wat, (Wat), (Wat), (Wat)	

r.m.s. deviations from ideality were 0.007 Å for bond lengths and 1.7° for bond angles (Table 1).

3.1. Overall structure

The asymmetric unit contains four monomers of ConA (Fig. 1), two of which are arranged in the canonical dimer (*A* and *B*), which generates the tetramer by the symmetry operation $(-x, y, 1/2 - z)$, and two of which (*C* and *D*) face each other in the region of the metal-binding loop in such a way that *D* and *C* form canonical dimers with a subunit from an asymmetric unit related by the symmetry operation $(1/2 + x, 1/2 + y, z)$ and from the next unit cell, respectively. As expected from the fact that ConA does not show affinity for galactose, no molecules of the substituted β-cyclodextrin are found in the asymmetric unit. This rules out the glucose molecules from the cyclodextrin ring as a potential target for the interaction between ConA and the substituted β-cyclodextrins and validates the feasibility of our model system.

Each monomer shows the classical 'jelly-roll' motif (Srinivas *et al.*, 2001). *A* is equivalent to *C* (r.m.s. difference 0.14 Å) and slightly different to *B* and *D* (r.m.s. differences of 0.32 and 0.30 Å, respectively) at residues Ser161–Gly163 and in loop Pro202–His205, which is involved in peptide binding (Jain *et al.*, 2001*a*). From the analysis of the packing in the crystal lattices it becomes clear that the equivalencies between *A* and *C* and between *B* and *D* are actually differences between the monomers that form the canonical dimer (*A* and *B*).

It is well known that in the presence of Mn^{2+} and Ca^{2+} ConA binds Mn^{2+} at S1 and Ca^{2+} at S2 and the molecule adopts the locked conformation (Brown *et al.*, 1977). However, although the protein was crystallized in the presence of both ions, at the final stages of the refinement with a Ca^{2+} ion placed at the S2 site the $mF_o - F_c$ maps contoured at 10σ and 15σ revealed clear electron density at the Ca^{2+}-ion site for each molecule of ConA in the asymmetric unit. Anomalous dispersion has been described as an approach for locating Mn and Ca (Einspahr *et al.*, 1985), but our difference anomalous maps were not conclusive because as the data set was collected at 0.84 Å the f' signals were very low (0.99 e^- for Mn and 0.42 e^- for Ca).

Bouckaert, Dewallef *et al.* (2000) have demonstrated the feasibility of using Cu $K\alpha$ radiation and a MAR 345 detector to confirm the chemical identity of the metal centre. Thus, in order to investigate the nature of the ion at S2, we collected an

Table 4
Comparison of different crystal forms of uncomplexed concanavalin A.

The number of residues from different asymmetric units forming hydrogen bonds of distance <3.3 Å are listed. Contacts shared by $C222_1$ as a general feature are underlined; those in bold are a universal feature of ConA crystals. The contacts shared by the $P2_12_12_1$ unlocked form and the $I222$ locked form are shown in italics and those residues with the largest r.m.s.d. when 1nls and 1nxd are compared are shown in CAPITALS.

PDB code	1gkb	ConCD	1nls	1dq2
Resolution (Å)	1.56	1.90	0.94	2.05
Space group	$C222_1$	$C222_1$	$I222$	$P2_12_12_1$
Unit-cell volume (Å^3)	1345520	2982373	497572	463830
Matthews coefficient (Å^3 Da^{-1})	3.26	3.62	2.41	2.25
Solvent content (%)	62.32	66.00	49.05	45.35
Asymmetric unit	Dimer	Tetramer	Monomer	Dimer
No. AU in cell	8	8	8	4
Contacts	Ala1–Tyr176			Ala1–Asp218
			Tyr12–SER184	
			ASP16–ASN69	
			ASP16–ASN118	
		Ser21–Lys101		
			Ser21–Lys135	
	Lys39–His205			
				Lys46–Asp235
	Thr49–His121	Thr49–HIS121		
			ASP58–Arg60	
	Asp58–Ser62	**Asp58–Ser62**	**Asp58–Ser62**	**Asp58–Ser62**
	Ser66–His121	Ser66–HIS121		
	Asn69–Asn118	ASN69–ASN118		
			ASN69–Arg228	Asn69–Thr226
				Asn69–Gly227
		Ala70–Asn118	Ala70–Asn118	
			ASP71–Arg228	
				Ser72–Ser168
				Asp80–Asp203
			Asp82–Ser204	*Asp82–Ser204*
				Asp82–His205
			Asn83–Asp203	*Asp83–Asp203*
			Asn83–Ser204	Asn83–Ser204
	Arg90–Asp218			
			Tyr100–TRP182	
	Tyr100–His205	Tyr100–His205		
	Tyr100–Asn237			
			LYS114–Glu192	
			LYS116–Glu192	
				His121–Asn131
			GLU122–Asn131	
		Thr123–Asn131	Tyr123–Asn131	
		Asn124–Asn131	Asn124–Asn131	
		ASN124–Gln132		
		Ala125–Met129	Ala125–Met129	
		Gln132–Glu183	Gln132–Glu183	
			Gln132–SER185	
				Asn136–Ser223
	Gln143–Ser220			
		Asp145–Ser223		
		Ser223–Thr226		

additional data set with Cu $K\alpha$ radiation on a MAR 345 detector. It consisted of 18 264 Friedel pairs with 93.8% completeness and a redundancy of 5.1. The data set was truncated at 3.7 Å in order to ensure a flat dependence of the mean $|F(+) - F(-)|/F^2_{\rm mean}$ on the resolution and the anomalous difference electron-density map showed two peaks at the S1 and S2 positions that range from 6.4σ to 7.4σ at site S1 and 4.2σ to 5.2σ at site S2 (Fig. 2), values that are in agreement with the peak-height ratio of 1.44 for Mn at S1 and S2 reported by Bouckaert, Dewallef *et al.* (2000) for the protein crystallized at pH 7 in the absence of Ca^{2+} (PDB code 1dq6).

These experimental results led us to suspect that Mn^{2+} may be the only ion present in the structure despite the crystallization system containing both Mn^{2+} and Ca^{2+}, although the reason for this is unclear. A plausible hypothesis may be that it is a consequence of the crystallization conditions, as the stoichiometry of Mn^{2+} binding is influenced by a variety of factors, such as the presence of sugar, the pH and the temperature; the estimated association constants for Mn^{2+} are $6.2 \times 10^5\ M^{-1}$ and $3.7 \times 10^4\ M^{-1}$ for S1 and S2, respectively, at 298 K and pH 6.5 in the absence of sugar (Sadhu & Magnuson, 1989)

A new ion-binding site has been reported by Kantardjieff *et al.* (2002) in a dimeric crystal form in space group $C222_1$, which is occupied by Mg^{2+} in a distorted octahedral coordination with Leu107 and Asn131 *via* main-chain carbonyl groups and four water molecules. Our maps show a clear electron density for a water molecule at this position. Although it is difficult to identify monovalent cations in solvent density, we have modelled it as Na^+ because (i) the crystallization system contained 200 mM NaCl, (ii) Na^+ has the same number of electrons as water and (iii) the coordination environment is characteristic of Na^+ (Nayal & Di Cera, 1996): it is surrounded exclusively by O atoms and octahedral coordination.

3.2. Solvent

The resolution of the present structure allows a clear definition of the solvent area. The final structure contains 764 water molecules, a number that is reasonable for the 66% estimated solvent content and is within the range predicted by Carugo & Bordo (1999) for a structure at 2 Å. It includes 16 peaks of electron density that have been modelled as molecules of azide, four of which are located at specific positions (Table 2).

Electron-density maps also revealed the presence of 25 glycerol molecules, the origin of which can be linked to the cryoprotectant solution. They are not randomly distributed over the asymmetric unit: 17 of them are in six positions that are conserved in at least two of the four monomers forming the asymmetric unit (Table 3). Among these six conserved positions, a glycerol molecule is found at G1 and G2 in each of

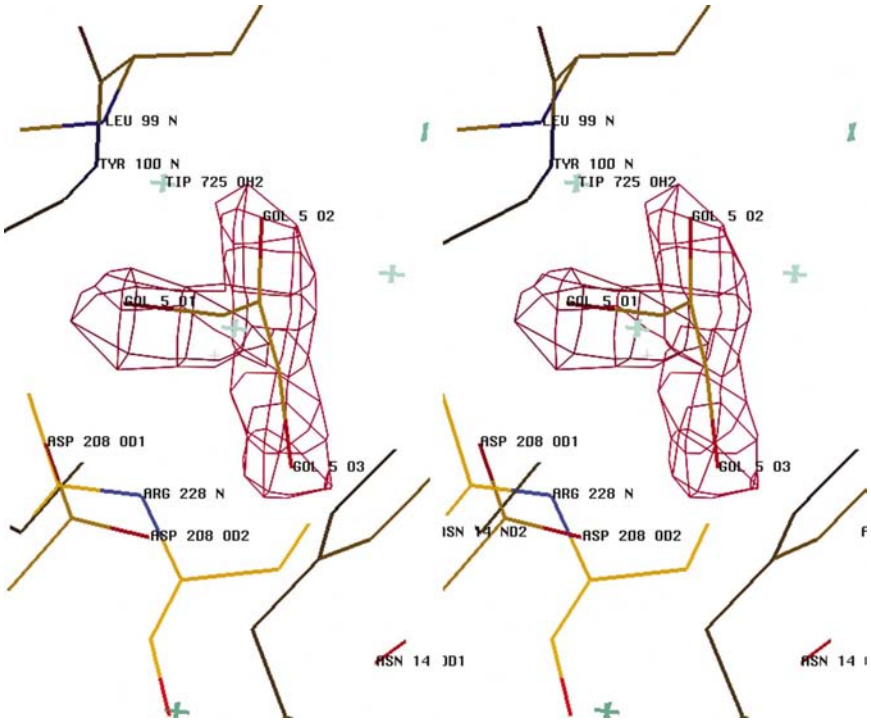

Figure 3
Stereoview of the $mF_o - F_c$ map omit map contoured at 5σ where the glycerol molecule at G1 was not included in the calculation. Residues involved in carbohydrate recognition and that form hydrogen bonds to the glycerol are labelled.

the four subunits. Each one occupies the place of two water molecules found in the structure at 0.94 Å by Deacon *et al.* (1997) (PDB code 1nls) and this fact might be experimental evidence of the replacement of water molecules by glycerol (Charron *et al.*, 2002). However, it is striking that the glycerol at G1 (Fig. 3) interacts with Asn14, Leu99, Asp208 and Arg228, residues that are involved in carbohydrate recognition (Derewenda *et al.*, 1989) and that the glycerol at G4 interacts with Asn44 and Ser201, which are key residues involved in peptide recognition, and with Pro13 and Asp203, which are involved in the interaction with 15-mer and 10-mer peptides, respectively (Jain *et al.*, 2001*a*). This result may point to a recognition phenomenon rather than a displacement of water molecules and if this is the case the fact that the G2 site is occupied by a molecule of glycerol in the four monomers of the asymmetric unit might suggest that G2 is potentially involved in the binding/recognition of ligands.

4. Crystal contacts

ConA has been crystallized in several space groups, from *P*1 (Kanellopoulos & Tucker, 1996; PDB code 1vln) to cubic $I2_13$ (Harrop *et al.*, 1996; PDB code 1gic); the most common space groups are those that belong to the orthorhombic system, with a *P* lattice for the unlocked form and an *I* lattice for the locked form. An analysis of the contacts (*i.e.* hydrogen bonding at distances of <3.3 Å) of the asymmetric units in the orthorhombic system (Table 4) shows that different space groups form different contacts, Asp58–Ser62 being the only one common to the orthorhombic system. It is also present in *P*1 and $I2_13$ crystals and hence can be considered as a universal feature of ConA crystal building regardless of the space group, unit-cell parameters or asymmetric unit. Contacts Ser66–His121, Asn69–Asn118 and Tyr100–His205 seem to be general features of the $C222_1$ crystal form and are present in the structures reported by Jain *et al.* (2000, 2001*a*,*b*) that are isomorphous to our structure (1nxd) and in the structure reported by Kantardjieff *et al.* (2002) (PDB code 1gkb) with a different unit cell and asymmetric unit.

The comparison of our structure (1nxd) and the search model used for molecular replacement (1nls) shows that they are very similar, with an average r.m.s.d. of 0.48 Å, the largest differences being at residues Asp16–Gly18, Ser56–Lys59, Pro68–Asp71, Lys117–Asn124, Ser161–Asn162, Ile181–Ala186 and Pro202–His205. A closer analysis of Table 4 reveals that all these residues are involved in contacts that if present in 1nls (space group *I*222) are not present in 1nxd (space group $C222_1$) and *vice versa*, raising the question of whether the space group is responsible for these differences or whether they were already present in the molecule during crystallization and led to the space group through the crystal contacts.

At this point, it becomes clear that both crystal contacts and local changes in residues are inter-related phenomena. Bouckaert, Dewallef *et al.* (2000) have shown that the soaking of unlocked metal-free ConA with Mn^{2+} yields locked Mn–Mn ConA crystals with a space-group transition to *I*222. This supports the idea that changes in conformation lead to different space groups. However, the removal of both Mn ions by soaking with EDTA does not reverse the space-group transition, but freezes the molecule in the locked state as the *I*222 asymmetric unit forms more contacts with its neighbour, although it does share some contacts with the $P2_12_12_1$ form. This result proves the importance of the crystal contacts in the stabilization of the molecule and shows how they can prevent a conformational change from taking place.

5. Conclusions

We have crystallized ConA in the presence of heptakis (6-*S*-β-D-galactopyranosyl-6-thio)-cyclomaltoheptaose, a complex carbohydrate consisting of a combination of a glucose ring that can interact with ConA and galactose molecules on its upper side that do not bind to ConA. The structure has been solved at 1.9 Å and experimental evidence, although inconclusive, suggest that Mn^{2+} may be present at S2 despite the crystallization system containing both $CaCl_2$ and $MnCl_2$. As expected, the structure does not contain any molecules of the

substituted β-cyclodextrin. This result rules out interaction of the glucose ring with ConA and validates the feasibility of ConA/substituted β-cyclodextrin as an attractive model for the rational design of modifications to the surface of the cyclodextrin that modulate its interaction with ConA.

The analysis of the crystal contacts on the basis of the residues responsible for the hydrogen-bonding network between asymmetric units reveals a relation between the residues forming hydrogen bonds and the space group. The contact Asp58–Ser62 is a universal feature of ConA crystals that is present regardless of the space group. Contacts Ser66–His121, Asn69–Asn118 and Tyr100–His205 are general features of the $C222_1$ crystals.

On the other hand, the cryobuffer left its fingerprint as 25 molecules of glycerol in the bound solvent, most of which are located at specific positions. The fact that one glycerol is found in the sugar-binding pocket of each monomer of the asymmetric unit and that another glycerol molecule interacts in two subunits with residues involved in the peptide binding suggests a recognition phenomenon rather than a displacement of water molecules and might point to G2 as a potential recognition/binding site.

We thank Professor García-Ruiz for his constant support and helpful discussions and suggestions, Dr Cartwright for reviewing the manuscript and Professor Kabsch for providing *XDS* free of charge. We acknowledge EMBO/DESY for beam time, Professor Martinez Ripoll for his kind support in the preliminary X-ray characterization using the rotating-anode generator at the Instituto de Química Fisica Roca Solano (CSIC) and the referees for valuable comments. Work in AA's laboratory is funded by Spanish Plan Nacional grant BMC-2002-04011-C03.

References

Adams, P. D., Pannu, N. S., Read, R. J. & Brünger, A. T. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5018–5023.

Bentley, G. A. (1992). *Proceedings of the CCP4 Study Weekend. Molecular Replacement*, edited by E. Dodson, S. Gover & W. Wolf, pp. 154–162. Warrington: Daresbury Laboratory.

Bouckaert, J., Dewallef, Y., Poortmans, F., Wyns, L. & Loris, R. (2000). *J. Biol. Chem.* **275**, 19778–19787.

Bouckaert, J., Hamelryck, T., Wyns, L. & Loris, R. (1999). *Curr. Opin. Struct. Biol.* **9**, 572–577.

Bouckaert, J., Loris, R. & Wyns, L. (2000). *Acta Cryst.* D**56**, 1569–1576.

Bradbrook, G. M., Gleichmann, T., Harrop, S. J., Habash, J., Raftery, J., Kalb, J., Yariv, J., Hillier, I. H. & Helliwell, J. R. (1998). *J. Chem. Soc. Faraday Trans.* **94**, 1603–1611.

Brown, R. D., Brewer, C. F. & Koening, S. H. (1977). *Biochemistry*, **16**, 3883–3896.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Brünger, A. T, Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Carugo, O. & Bordo, D. (1999). *Acta Cryst.* D**55**, 479–483.

Castellano, E. E., Oliva, G. & Navaza, J. (1992). *J. Appl. Cryst.* **25**, 281–284.

Charron, C., Kadri, A., Robert, M. C., Giegé, R. & Lorber, B. (2002). *Acta Cryst.* D**58**, 2060–2065.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Crowther, R. A. & Blow, D. M. (1967). *Acta Cryst.* **23**, 544–548.

Cruickshank, D. W. J. (1996). *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 11–22. Warrington: Daresbury Laboratory.

Deacon, A., Gleichmann, T., Kalb (Gilboa), A. J., Price, H., Raftery, J., Bradbrook, G., Yariv, J. & Helliwell, J. R. (1997). *J. Chem. Soc. Faraday Trans.* **93**, 4305–4312.

Derewenda, Z., Yariv, J., Helliwell, J. R., Gilboa, A. J. K., Dodson, E. J., Papiz, M. Z., Wan, T. & Campbell, J. (1989). *EMBO J.* **8**, 2189–2193.

Einspahr, H., Suguna, K. & Suddath, F. L. (1985). *Acta Cryst.* B**41**, 336–341.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

García-López, J. J., Hernández-Mateo, F., Isac-Garcia, J., Kim, J.-M, Roy, R., Santoyo-Gonzalez, F. & Vargas-Berenguel, A. (1999). *J. Org. Chem.* **64**, 522–553.

Garcia-Ruiz, J. M. & Moreno, A. (1997). *J. Cryst. Growth*, **178**, 393–401.

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

Harrop, S. J., Helliwell, J. R., Wan, T. C. M., Kalb, A. J., Tong, L. & Yariv, J. (1996). *Acta Cryst.* D**52**, 143–155.

Irie, T. & Uekema, K. (1999). *Adv. Drug Deliv. Rev.* **36**, 101–123.

Jain, D., Kaur, K. J. & Salunke, D. M. (2001*a*). *Biophys. J.* **80**, 2912–2921.

Jain, D., Kaur, K. J. & Salunke, D. M. (2001*b*). *Biochemistry*, **40**, 12059–12066.

Jain, D., Kaur, K. J., Sundaravadivel, B. & Salunke, D. M. (2000). *J. Biol. Chem.* **275**, 16098–16102.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Kanellopoulos, P. N. & Tucker, P. A. (1996). *J. Struct. Biol.* **117**, 16–23.

Kantardjieff, K. A., Höchtl, P., Segelke, B. W., Tao, F. M. & Rupp, B. (2002). *Acta Cryst.* D**58**, 735–743.

Kompella, U. B. & Lee, V. H. L. (2001). *Adv. Drug Deliv. Rev.* **46**, 211–245.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.

Loftsson, T. & Masson, M. (2001). *Int. J. Pharm.* **225**, 15–30.

Loftssona, T. & Järvinen, T. (1999). *Adv. Drug Deliv. Rev.* **36**, 59–79.

Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). *Biochim. Biophys. Acta*, **1383**, 9–36.

McRee, D. E. (1993). *Practical Protein Crystallography*. New York: Academic Press.

Matsuda, H. & Arima, H. (1999). *Adv. Drug Deliv. Rev.* **36**, 81–99.

Merkus, F. W. H. M., Verhoef, J. C., Martin, E., Romeijn, S. G., van der Kuy, P. H. M., Hermens, W. A. J. J. & Schipper, N. G. M. (1999). *Adv. Drug Deliv. Rev.* **36**, 41–57.

Moothoo, D. N. & Naismith, J. H. (1999). *Acta Cryst.* D**55**, 353–355.

Naismith, J. H., Habash, J., Harrop, S., Helliwell, J. R., Hunter, W. N., Wan, T. C. M. & Weisgerber, S. (1993). *Acta Cryst.* D**49**, 561–571.

Navaza, J. (1994). *Acta Cryst.* A**50**, 147–163.

Nayal, M. & Di Cera, E. (1996). *J. Mol. Biol.* **256**, 228–234.

Ng, J. D., Gavira, J. A. & Garcia-Ruiz, J. M. (2003). *J. Struct. Biol.* **142**, 218–231.

Redenti, E., Pietra, C., Gerloczy, A. & Szente, L. (2001). *Adv. Drug Deliv. Rev.* **53**, 235–244.

Rini, J. M. (1995). *Annu. Rev. Biophys. Biomol. Struct.* **24**, 551–577.

Rini, J. M. & Lobsanov, Y. D. (1999). *Curr. Opin. Struct. Biol.* **9**, 578–584.

Sadhu, A. & Magnuson, J. (1989). *Biochemistry*, **28**, 3197–3204.

Shaanan, B., Lis, H. & Sharon, N. (1991). *Science*, **254**, 862–866.
Sharon, N. & Lis, H. (1972). *Science*, **177**, 949–959.
Singh, M., Sharma, R. & Banerjee, U. C. (2002). *Biotechnol. Adv.* **20**, 341–359.
Soppimath, S. K., Aminabhavi, T. M., Kulkarni, A. R. & Rudzinski, W. E. (2001). *J. Control. Release*, **70**, 1–20.
Srinivas, V. R., Reddy, G. B., Chittoor, N. A., Swaminathan, C. P., Mitra, N. & Surolia, A. (2001). *Biochim. Biophys. Acta*, **1527**, 102–111.
Sumner, J. B. (1919). *J. Biol. Chem.* **37**, 137–142.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Vijayan, M. & Chandra, M. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.