Title
X-ray diffraction and atomic force microscopy analysis of twinned crystals: rhombohedral canavalin

Journal
Acta Crystallographica Section D, Structural Biology, 57(6)

ISSN
2059-7983

Authors
Ko, TP
Kuznetsov, YG
Malkin, AJ
et al.

Publication Date
2001-06-01

DOI
10.1107/s0907444901003791

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

X-ray diffraction and atomic force microscopy analysis of twinned crystals: rhombohedral canavalin

Tzu-Ping Ko,[a] Yurii G. Kuznetsov,[b] Alexander J. Malkin,[b] John Day[b] and Alexander McPherson[b]*

[a]Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan, and [b]Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697-3900, USA

Correspondence e-mail: amcphers@uci.edu

Received 18 August 2000
Accepted 26 February 2001

PDB References: rhombohedral canavalin, 1dgw; orthorhombic canavalin, 1dgr.

The structure of canavalin, the vicilin-class storage protein from jack bean, was refined to 1.7 Å resolution in a highly twinned rhombohedral crystal of space group $R3$ and unit-cell parameters $a = b = c = 83.0$ Å, $\alpha = \beta = \gamma = 111.1^\circ$. The resulting R and $R_{\rm free}$ were 0.176 and 0.245, respectively. The orthorhombic crystal structure (space group $C222_1$, unit-cell parameters $a = 136.5$, $b = 150.3$, $c = 133.4$ Å) was also refined with threefold non-crystallographic symmetry restraints. R and $R_{\rm free}$ were 0.181 and 0.226, respectively, for 2.6 Å resolution data. No significant difference in the protein structure was seen between these two crystal forms, nor between these two and the hexagonal and cubic crystal forms reported elsewhere [Ko *et al.* (1993), *Acta Cryst.* D**49**, 478–489; Ko *et al.* (1993), *Plant Physiol.* **101**, 729–744]. A phosphate ion was identified in the lumen of the C-terminal β-barrel. Lattice interactions showed that the trimeric molecule could be well accommodated in both 'top-up' and 'bottom-up' orientations in a rhombohedral unit cell of the $R3$ crystal and explained the presence of a high twin fraction. The large inter-trimer stacking interface of the $C222_1$ crystal may account for its relative stability. Atomic force microscopy (AFM) investigations of the growth of three crystal forms of canavalin indicate the rhombohedral form to be unique. Unlike the other two crystal forms, it contains at least an order of magnitude more screw dislocations and stacking faults than any other macromolecular crystal yet studied, and it alone grows principally by generation of steps from the screw dislocations. The unusually high occurrence of the screw dislocations and stacking faults is attributed to mechanical stress produced by the alternate molecular orientations in the rhombohedral crystals and their organization into discrete domains or blocks. At boundaries of alternate domains, lattice strain is relieved by the formation of the screw dislocations.

1. Introduction

Crystal twinning is problematic in the structure determination of proteins, especially when phasing is by multiple isomorphous replacement (Yeates, 1997). For merohedrally twinned crystals with low fractions of twinning, it is possible to directly recover the true diffraction intensities. When the twin fraction approaches 0.5, the observed intensities $I_{\rm obs}(\mathbf{h}_1)$ and $I_{\rm obs}(\mathbf{h}_2)$ related by hemihedral twinning will become nearly equal and additional symmetry appears in the diffraction pattern. For these nearly perfect twin crystals, direct detwinning of the data is not reliable. The situation is less acute, however, in molecular replacement. Using an accurate starting model, it is still possible to solve and refine the crystal structure. Structure determination and refinement of twinned protein crystals have become of increasing interest as seen, for

example, in the papers regarding bacteriorhodopsin (Luecke *et al.*, 1998, 1999), cephalosporin synthase (Valegard *et al.*, 1998), cytidylyltransferase (Weber *et al.*, 1999), human lactoferrin (Breyer *et al.*, 1999) and an arginase–boronic acid complex (Cox *et al.*, 1999). A more recent review on this subject can be found in the paper by Yeates & Fam (1999).

Canavalin is the vicilin-class storage protein from jack bean (*Canavalia ensiformis*) and is a trimer of 142 kDa (Smith *et al.*, 1982; McPherson, 1999). The tertiary structure of canavalin was determined by multiple isomorphous replacement and molecular-replacement methods in four different crystal forms (Ko, Ng, Day *et al.*, 1993; Ko, Ng & McPherson, 1993; McPherson, 1999). It is similar to phaseolin, another vicilin-class protein from French bean (*Phaseolus vulgaris*; Lawrence *et al.*, 1990, 1994). Both canavalin and phaseolin are trimers with exact threefold symmetry. Each subunit of the trimer, however, is composed of amino- and carboxy-terminal domains. The two domains are structurally very similar and are related to one another by pseudo-dyad axes perpendicular to the threefold axis. Thus, the canavalin molecule has a high degree of quasi-D_3 symmetry. As shown in Fig. 1, each domain is comprised of a 'core' β-barrel subdomain with an extended helical 'loop' subdomain. Within a subunit, the core subdomains are tightly associated through the larger β-sheets of A′ABIDGZ; between subunits, the smaller β-sheets of JJ′CHEF, and the loop subdomains are involved in the assembly into trimers (Ko *et al.*, 2000).

The structure of canavalin has recently assumed some additional importance, serving, along with that of phaeseolin, as the model structures for the protein superfamily known as cupins. This superfamily now includes a vast array of both plant and animal proteins, some containing metal ions, that catalyze a variety of biochemical reactions. All share the fundamental architecture of the canavalin domain (Dunwell *et al.*, 2000).

The rhombohedral canavalin crystal has special merits in that it appeared first (Sumner & Howell, 1936), grows easily and reproducibly, grew to largest size (Day & McPherson, 1992) and diffracted X-rays to highest resolution (Koszelak *et al.*, 1995). It has provided a model system for a diverse array of crystallization studies (Kuznetsov *et al.*, 1995, 1997, 1999; Land *et al.*, 1995, 1996, 1999; Malkin *et al.*, 1995, 1996, 1999) and has also been a prominent crystal employed in microgravity crystal-growth investigations (Day & McPherson, 1992; Koszelak *et al.*, 1995; McPherson, 1996). This crystal form displays a high degree of *R*32 symmetry, but the true space group was identified as *R*3 (McPherson & Spencer, 1975). It has unit-cell parameters $a = b = c = 83.0$ Å, $\alpha = \beta = \gamma = 111.1^\circ$ or $a = b = 136.9$, $c = 76.0$ Å in the equivalent hexagonal cell and contains one canavalin protomer per asymmetric unit.

In earlier lower resolution stages of study, it was disturbing to find that a crude model could be placed in different orientations, including inverting the trimer, and the model would still yield acceptable *R* values in the low 20s. Ambiguities appeared to be resolved by calculating direct cross-rotation functions between the hexagonal crystal form in which the structure had been solved by MIR, and the *R*3 crystal. This was described in Ko, Ng, Day *et al.* (1993). Refinement using higher resolution data, however, yielded only mediocre results. This was particularly perplexing because all of the other crystal forms of canavalin were refined successfully and well (Ko *et al.*, 2000). Because of the suspicious behavior of the rhombohedral crystals and their model, further analysis of structure was focused on the other crystal forms of the molecule, particularly the hexagonal form (Ko, Ng, Day *et al.*, 1993).

Here, we show that the rhombohedral crystal form is indeed hemihedrally twinned and that with the inclusion of this feature the structure can be satisfactorily refined. The orthorhombic crystal, previously refined to 2.6 Å resolution (Ko, Ng, Day *et al.*, 1993), belongs to space group $C222_1$, has unit-cell parameters $a = 136.5$, $b = 150.3$, $c = 133.4$ Å and contains the entire canavalin trimer as its asymmetric unit. This crystal has also been further refined and serves as a useful basis of comparison for the rhombohedral crystal form, particularly with regard to the atomic force microscopy investigation described below.

In parallel with our X-ray diffraction studies, we have investigated, by *in situ* AFM, the growth of three different crystal forms of canavalin: the rhombohedral, hexagonal and orthorhombic crystals (the cubic form was not examined by AFM). While both of the latter forms grow almost exclusively by the mechanism of two-dimensional nucleation, the rhombohedral form does not (Land *et al.*, 1995, 1996, 1999; Malkin *et al.*, 1995). Unique among all of the macromolecular and virus crystals we have studied (more than 20 in number), it alone grows principally owing to a vast array of screw dislocations distributed throughout its volume. This has long been a puzzling observation without any apparent explana-

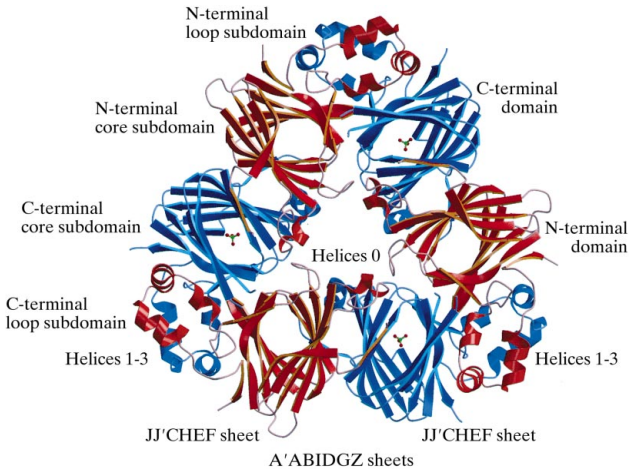


Figure 1
A ribbon diagram of the three identical subunits of the canavalin trimer shown here as viewed along the threefold axis. The N-terminal and the C-terminal domains are in red and blue, respectively. Each domain of a subunit can in turn be divided into a β-barrel subdomain and a helix–loop–helix subdomain. The former is responsible for intrasubunit domain interactions and the latter for intersubunit contacts. All other components of the molecule, including sheet strands and helices, are according to the nomenclature defined in Ko *et al.* (2000). The phosphate ions in the lumina of the C-terminal barrels are also shown, with bonds in yellow.

tion. We suggest, based on the evidence presented here, that the appearance of the screw dislocations and the many accompanying stacking faults may be correlated with the unusual twinning shown by X-ray crystallography. This would be entirely consistent with theory and experience from conventional crystal-growth research.

2. Materials and methods

Canavalin was prepared as described previously (McPherson & Spencer, 1975; Sumner & Howell, 1936) and large rhombohedral crystals were grown in microgravity experiments on the US Space Shuttle by liquid–liquid diffusion (Koszelak *et al.*, 1995). X-ray data were collected and processed using a San Diego Multiwire Systems detector with a Rigaku RU-200 rotating anode and *SDMS* software package also as previously described (Ko, Ng, Day *et al.*, 1993). These microgravity-grown crystals extended the effective diffraction resolution from 2.6 Å to about 1.7 Å. Statistics are found in Table 1. For refinement of the orthorhombic crystal structure, the data set previously used in structure determination by molecular replacement was retrieved from the Protein Data Bank (PDB code 1cax). It contains 28 836 reflections with $F > 3\sigma_F$ and extends to 2.6 Å resolution. Throughout the refinement, 8% randomly selected reflections from both crystals were set aside for cross validation with $R_{\rm free}$ (Brünger, 1992*a*). The program *X-PLOR* (Brünger, 1992*b*) was used in initial refinements, but at later stages *CNS* (Brunger *et al.*, 1998) was employed. Computation was performed on a Silicon Graphics Octane workstation. Manual rebuilding of the model and addition of solvent molecules were carried out using *O* (Jones *et al.*, 1991) and analyses of the models employed *CNS*, *CCP*4 (Collaborative Computational Project, Number 4, 1994) and *PROCHECK* (Laskowski *et al.*, 1993). Figures were produced using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merrit & Murphy, 1994).

Refinement of the rhombohedral canavalin structure was reinitiated using a new model derived from the refined cubic crystal (Ko *et al.*, 2000). After appropriate placement of the model in the rhombohedral unit cell, rigid-body refinement using *X-PLOR* yielded an *R* value of 0.32 at 3.0 Å resolution. Subsequent refinement, including addition of solvent molecules, gave marginally acceptable *R* and $R_{\rm free}$ values of only 0.28 and 0.35, respectively, at 2.0 Å resolution. The problem was manifest in the difference Fourier maps which contained unexplained densities outside the protein region. We then found that the entire model could be rotated by 180° about the 100 or 010 axis of the hexagonal unit cell and still yield comparable refinement results. Both 'top-up' and 'bottom-up' molecules were evidently present in the rhombohedral crystal, suggesting that the crystal was actually twinned with a high fraction. Further refinements were unsuccessful until *CNS* was employed.

Hemihederally twinned crystals consist of a large number of individual domains consisting alternately of molecules in one of two orientations. Thus, every measured reflection can be considered a composite of contributions from two separate crystals. Both orientations of the molecule, 'top up' and 'bottom up', contribute to the intensity of each reflection in proportion to the degree of twinning.

Table 1
Data-collection statistics of the rhombohedral canavalin crystals.

Numbers in parentheses refer to the highest resolution shell.

Space group	*R*3
Unit-cell parameters (Å,°)	$a = b = c = 82.987$; $\alpha = \beta = \gamma = 111.11$
Hexagonal system (Å,°)	$a = b = 136.876$, $c = 76.004$, $\alpha = \beta = 90$, $\gamma = 120$
No. of crystals	4
Resolution (Å)	1.69 (1.78–1.69)
No. of observations	343806 (16233)
No. of unique reflections	56649 (6947)
Completeness (%)	95.0 (81.1)
$R_{\rm merge}$ (%)	8.85 (30.4)
Average I/σ_I	18.7 (1.05)

† $R_{\rm merge} = \sum_{hkl}\sum_j |I_{\rm ave} - I_{{\rm obs},j}|/\sum_{hkl}\sum_j I_{\rm ave}$, in which $I_{\rm ave}$ is the average intensity of j equivalent reflections $I_{{\rm obs},j}$ and the sum is over all reflections that are equivalent in the unmerged data set.

For crystals where the twinning is disparate, *e.g.* 20–80%, and the twinning fraction can be accurately determined, then the contributions can be separated and refinement carried out against partial intensities. When the twinning fraction is near 50%, as it is for canavalin, detwinning the contributions becomes impossible. It is, however, possible to estimate the relative contributions of each twin to the intensity, *I*, and therefore to refine the model of the structure against *I*.

The intensities, sigmas and *R* value are calculated for all equivalent reflections by

$$I_{\rm obs}(\mathbf{h}) = [I_{\rm obs}(\mathbf{h}_1) + I_{\rm obs}(\mathbf{h}_2)]/2 = \{[F_{\rm obs}(\mathbf{h}_1)]^2 + [F_{\rm obs}(\mathbf{h}_2)]^2\}/2,$$
$$\sigma_I(\mathbf{h}) = F_{\rm obs}(\mathbf{h}_1)\sigma_F(\mathbf{h}_1) + F_{\rm obs}(\mathbf{h}_2)\sigma_F(\mathbf{h}_2),$$
$$R_{\rm average} = \sum[|I_{\rm obs}(\mathbf{h}_1) - I_{\rm obs}(\mathbf{h}_2)|]/\sum I_{\rm obs}(\mathbf{h}).$$

A routine for carrying this out has been implemented in *CNS*.

For canavalin, refinement was carried out in parallel using both the measured *R*3 intensities and a data set produced by averaging reflections according to *R*32, the apparent space group of a perfect twin, but re-expanded according to *R*3. The $R_{\rm avg}$ between hemihedrally related intensities was 0.132 for 27 473 independent reflections. Model alterations were applied only as a result of the refinement against the perfect twin data and maps were calculated using only those structure amplitudes. The original strictly *R*3 data was used for cross validation during refinement, in addition to the conventional $R_{\rm free}$ based on a sequestered data subset.

Using *CNS*, the twin fraction of the rhombohedral crystal was estimated to be 0.42–0.44 based on $\langle H\rangle$ and $\langle H^2\rangle$ (Yeates, 1997) or 0.46 based on the initial model. Starting with this model, a minimal *R* value of 0.215 for all data at 2.0 Å resolution was obtained, while the twin fraction converged to 0.426 by optimization with small intervals between 0.3 and 0.5. Initial cycles of refinement at 2.0 Å, including simulated annealing, yielded *R* and $R_{\rm free}$ values of 0.183 and 0.247.

Addition of water molecules was carried out manually using *O* in conjunction with the water-pick procedure of *CNS*. The

general criteria were a distance of less than 6 Å from appropriate protein atoms, a temperature factor of less than 60 Å^2 after refinement and a corresponding density level of more than 2σ in the Fourier maps. However, strong densities in the central cavity of the canavalin trimer, near the N- and C-termini or in the vicinity of crystal contact regions were also interpreted as ordered waters with a less stringent distance criterion. At the C-terminus of the C-terminal domain, the density initially interpreted as waters was later shown to correspond to two additional amino-acid residues Gln422 and Pro423. In the lumen of the C-terminal β-barrel, density for a cluster of three to four waters was reinterpreted as a phosphate ion. The model was repeatedly checked against omit maps for selected regions which had high temperature factors or real-space R values. Adjustments of protein side chains were made as the refinement proceeded, while the resolution was gradually increased to 1.7 Å.

Refinement of the orthorhombic crystal was carried out in parallel with that of the rhombohedral crystal. An initial R value of 0.27 for all reflections in the data set was calculated by *X-PLOR* using a trimer generated from the refined protein model of the cubic crystal (Ko *et al.*, 2000), with proper placement in the orthorhombic unit cell and rigid-body minimization of the six domains. In Fourier maps calculated from this initial model, densities for several water molecules were seen clearly, justifying inclusion of a solvent model, even at 2.6 Å. Non-crystallographic symmetry (NCS) restraints were imposed on the three protein subunits. Using *CNS*, maps were also threefold averaged with only limited improvement. In addition to the C-terminal Gln422 and Pro423 of all subunits, another extension of Met331 was identified for the *E*–*F* loop of the C-terminal domain in two subunits. A phosphate ion was also clearly seen in the C-terminal barrels of all subunits.

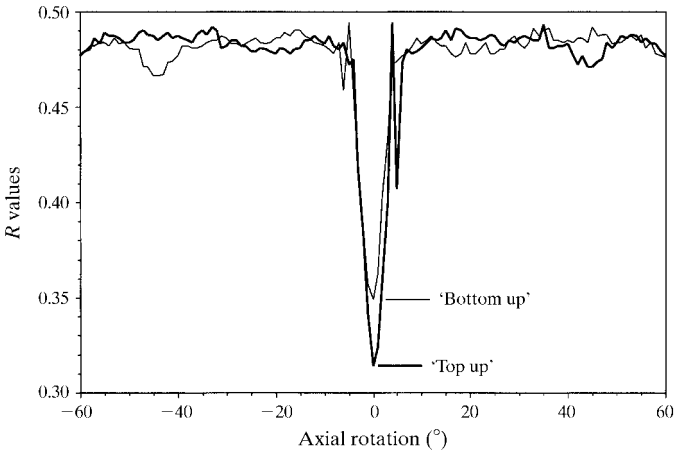


Figure 2
R-value plots of the canavalin trimer against axial rotation in the *R*3 crystal. The *R* values were calculated by *X-PLOR* using the model of the cubic crystal and 10 394 reflections having $F > 4\sigma_F$ in the resolution range 12–3.0 Å from the rhombohedral crystal. The molecule was rotated about the threefold axis with respect to the original orientation in the *R*3 crystal (thick line) or after inversion according to the hemihedral twin-symmetry operation (thin line).

For AFM studies, seed crystals of canavalin were nucleated and grown on glass substrates in 10 µl droplets by a batch method consisting of mixing 40 mg ml^{-1} of protein dissolved in water with an equal amount of 2 × Dulbecco's phosphate buffered saline (DPBS). Canavalin crystals of sizes 50–500 µm in the longest dimension were used. The glass substrate with its seed crystals was then immediately transferred to the sealed fluid cell of a Nanoscope III atomic force microscope (Digital Instruments, Santa Barbara, CA, USA) and the entire volume of the cell (approximately 50 µl) was filled with a 1:1 mixture of protein (4–20 mg ml^{-1}) in DPBS. Images were collected in contact mode using oxide-sharpened silicon nitride tips (Park Scientific Instruments, Sunnyvale, CA, USA). Cantilevers with nominal force constants of 0.01 N m^{-1} and forces of 0.3 nN or less were utilized during imaging. When scanning, care was taken to continually adjust the set-point voltage to the lowest value for which tip-to-crystal contact was maintained in order to minimize the force applied to the crystalline surface.

3. Results

3.1. Refinement

To reconfirm the molecular-replacement solution of the rhombohedral crystal, the R values of the protein model were plotted against rotation angle about the threefold axis. This was performed with either top-up or bottom-up models and the results are shown in Fig. 2. The symmetric profile was also seen in the cross-rotation function of the rhombohedral and hexagonal crystals (Fig. 5*a* of Ko, Ng, Day *et al.*, 1993). These plots indicated that only two orientations of the canavalin molecules were present in the rhombohedral crystal. To estimate the precise twin fraction, several cumulative H plots were made and are illustrated in Fig. 3. However, the linear portion of the curve was only seen within a cumulative distribution of 0–0.65; above 0.65 its slope declined rapidly. Depending on the selection of data in the calculation, the linear portion showed a range of slope for twin fraction from 0.43 (all reflections) to 0.47 (10% strongest reflections). A plausible explanation for the uncertainty of twin fraction was that data from four crystals, possibly twinned to slightly different degrees, were merged to provide the final data set. Ambiguity was removed by using an averaged data set for a perfect twin in refinement. At the same time, refinement was also carried out using the original data set with the optimized twin fraction of 0.426.

The refined model in the rhombohedral crystal contained 2783 non-H atoms in the 346 residues Asn46–Pro223, Asp246–Glu321 and Gln332–Pro423 plus a phosphate ion and 292 waters. Residues 46–223 and 246–423 comprised the two domains described previously (Ko *et al.*, 2000) in which the omitted loop between strands *E* and *F* of the C-terminal barrel (Gln322–Met331) had weak density, probably owing to flexibility. For the orthorhombic crystal, the refined model contained 8356 protein atoms, three phosphate ions and 168 waters. The polypeptide chains of the three canavalin subunits were named *A*, *B* and *C*, which all contained similar residues

Table 2
Refinement and model statistics.

Values in parentheses refer to the highest resolution shells. For the rhombohedral crystal, the averaged data set of perfect twin was used and limited with $F > 2\sigma_F$. Data for the orthorhombic crystal were limited with $F > 3\sigma_F$.

Crystal form	Rhombohedral	Orthorhombic
Resolution range (Å)	500–1.7 (1.78–1.70)	500–2.6 (2.76–2.60)
No. of reflections	49983 (4062)	28811 (2283)
Completeness (%)	85.5 (55.6)	67.8 (43.0)
R value for 92% data (working set)	0.176 (0.375)	0.181 (0.230)
$R_{\rm free}$ for 8% data (test set)	0.245 (0.386)	0.226 (0.263)
Average B (Å^2)/No. of non-H atoms	36.4/3080	34.3/8539
Protein backbone atoms†	35.5/1384	33.9/4156
Protein side-chain atoms†	36.1/1399	34.7/4200
Water molecules	41.9/292	32.9/168
Phosphate atoms	57.1/5	65.0/15
R.m.s.d. of B for backbone atoms	0.447	0.495
Side-chain atoms	0.338	0.465
R.m.s.d. from ideal bond lengths (Å)	0.0071	0.0074
Bond angles (°)	1.45	1.43
Dihedral angles (°)	25.5	25.5
Improper angles (°)	0.924	0.808
φ, ψ angles‡ (%)		
In most favored regions	84.6	90.3
In additional allowed regions	14.8	9.3

† Backbone atoms include N, C$^\alpha$, C and O; other atoms are considered to be side-chain atoms. ‡ Based on a Ramachandran plot computed with *PROCHECK* for all non-glycine and non-proline residues.

Table 3
Comparison of canavalin models.

The diagonal in bold contains numbers of amino-acid residues in the subunit models. Numbers above and below the diagonal are root-mean-square differences in Å for the backbone and side-chain atoms (defined in Table 2), respectively.

	Rhombo	OrthoA	OrthoB	OrthoC	Cubic	Hexag
Rhombo	**346**	0.468	0.490	0.457	0.425	0.415
OrthoA	1.050	**346**	0.241	0.233	0.345	0.348
OrthoB	1.073	0.367	**347**	0.236	0.347	0.382
OrthoC	1.051	0.417	0.417	**346**	0.316	0.356
Cubic	1.402	1.233	1.241	1.262	**344**	0.331
Hexag	1.434	1.286	1.288	1.315	1.197	**346**

as in the rhombohedral model, while the residue Gln321 was not observed in the *A* chain and an additional Met331 was observed for the *A* and *B* chains. Statistics of the refined models are listed in Table 2.

For the rhombohedral crystal, the R and $R_{\rm free}$ values at 1.7 Å were 0.176 and 0.245, respectively, using the averaged data set of perfect twin with $F > 2\sigma_F$ cutoff, or 0.181 and 0.248 using all data. If the original data set and a twin fraction of 0.426 were used, the R and $R_{\rm free}$ would be 0.183 and 0.250, respectively, for data with $F > 2\sigma_F$, or 0.189 and 0.252 for all data. The coordinate errors estimated by a Luzzati plot (Luzzati, 1952) using the work and test data sets were 0.24 and 0.25 Å, respectively. For the orthorhombic crystal, the R and $R_{\rm free}$ at 2.6 Å were 0.181 and 0.226, respectively, and the estimated errors were 0.27 and 0.34 Å. In *CNS*, where domains were restrained separately, the NCS-related atoms deviated by 0.09–0.10 Å for the backbone and 0.16–0.19 Å for side chains. Superposition of entire subunits using *LSQKAB* from *CCP*4 produced differences of 0.23–0.24 and 0.37–0.42 Å, respectively, as shown in Table 3. Between the rhombohedral and orthorhombic crystals the coordinate differences for backbone and side-chain atoms were 0.46–0.49 and 1.05–1.07 Å, respectively. Similar ranges of 0.32–0.43 and 1.20–1.43 Å were obtained if the cubic and hexagonal models (PDB codes 2cau and 2cav) were included in the comparison. In all models, there was only one residue per subunit, Tyr417 ($\varphi = 61 \pm 6$, $\psi = -65 \pm 10^\circ$), in the disallowed region of Ramachandran plot defined by *PROCHECK*. Another residue, Ser351 ($\varphi = 75 \pm 3^\circ$, $\psi = -8 \pm 4^\circ$), was in the generously allowed region. These residues assumed conformations in a γ-turn (Tyr417) and a type II turn (Ser351) as described previously (Richardson, 1981; Ko *et al.*, 2000).

In the lumen of the C-terminal β-barrel, the phosphate ion made four hydrogen bonds, including a salt bridge, with the side chains of His297, Asn299 and Arg376, as shown in Fig. 4. The phosphate was probably also present in the cubic and hexagonal crystals at the solvent sites of Wat515 and Wat501 (Ko *et al.*, 2000), respectively, but was now identified with certainty in the refined subunit models for the rhombohedral and orthorhombic crystals. In retrospect, this phosphate corresponded to the major mercury iodide site of the rhombohedral and hexagonal crystals used in the original structure determination (Ko, Ng, Day *et al.*, 1993). At the equivalent position in the N-terminal domain, another water molecule made hydrogen bonds with His102 and Lys161. However, the volume of density and interatomic distances precluded a phosphate ion.

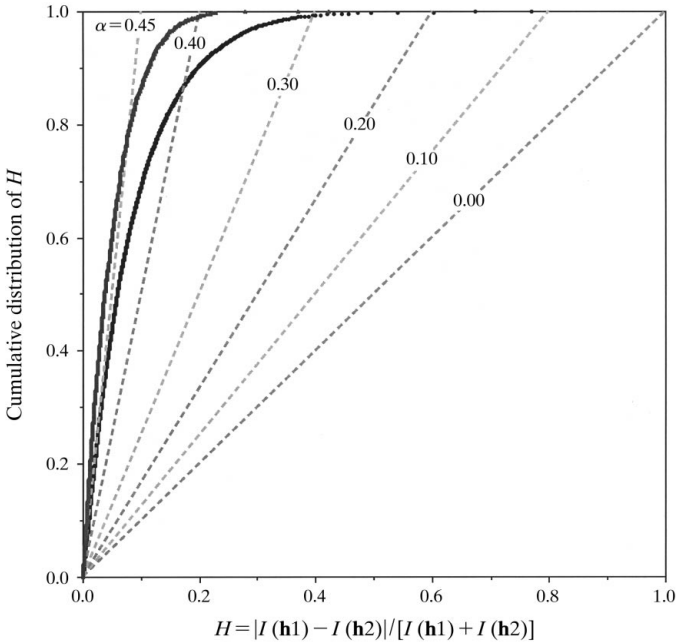


Figure 3
Cumulative H distribution plot for the twinned rhombohedral crystal of canavalin. Curves I and II were calculated using 16 000 pairs of matched reflections having $I > 2\sigma$ and the strongest 1800 reflection pairs, respectively.

3.2. Solvent and crystal packing

The Matthews coefficients, or specific volumes (V_M; Matthews, 1968), of the hexagonal, rhombohedral, orthorhombic and cubic canavalin crystals were 2.49, 2.78, 2.33 and 2.03 Å^3 Da^{-1}, respectively (Ko, Ng, Day *et al.*, 1993). Solvent models of the cubic and hexagonal crystals were analyzed in a previous paper (Ko *et al.*, 2000). In the newly refined models of the rhombohedral and orthorhombic crystals there were 292 and 168 water molecules, respectively. After superposition of the two structures based on backbone atoms, 44 pairs of water molecules per trimer matched with distances less than 1.0 Å. The average temperature factor was 23.8 Å^2. Among these, five waters in the rhombohedral crystal were observed at equivalent positions in all three subunits of the orthorhombic crystal and they were also found among the best defined waters for the cubic and hexagonal crystal structures. Specifically, waters 542, 549, 550, 759 and 762 in the rhombohedral model were equivalent to 509/556/634, 528/548/617, 523/579/616, 522/545/666 and 508/527/561 in the orthorhombic model and 506/504, 505/503, 504/502, 553/537 and 513/511 in the cubic and hexagonal models. They all bound to at least two backbone atoms of the protein. In addition, nine waters in the rhombohedral model were observed at equivalent positions in two subunits of the orthorhombic model and were also observed in either the cubic or hexagonal model or both. In the central cavity, there were 81 waters per trimer in the rhombohedral crystal but only 16 in the orthorhombic crystal.

The packing diagrams of Fig. 5 show that the disk-shaped canavalin trimer can be very well accommodated in the rhombohedral unit cell in either top-up or bottom-up orientation. Inter-trimer contacts between the rims of disks occur at the unit-cell faces. In either orientation, a trimer is in contact with six other trimers. In the top-up orientation there is only one type of interface, the top–bottom contact, which buries 310 Å^2 area on the top surface and 371 Å^2 on the bottom surface. There were at least 11 amino-acid residues involved on the top and eight on the bottom; all were polar. Interactions included a water-mediated salt bridge between Arg312 and Glu411 and a bidentated salt bridge between Asp344 and Arg392. Two other possible salt bridges of Arg52–Glu407/Glu408 and Glu314–Lys414 between the trimers were also observed at the crystal contact, but the distances were somewhat long for direct bonding. In the bottom-up orientation there were two types of interface, the top–top contact and the bottom–bottom contact, burying at least 240–250 Å^2 area on each contact surface. Precise interfaces could not be determined because the z coordinates of the model were arbitrary; the contact areas were minimized by translation of the bottom-up model along the threefold axis. There were seven and five amino-acid residues, also polar, from each subunit involved in the top–top and the bottom–bottom interfaces, respectively (see Table 4). The close contact between two pairs of twin-dyad symmetry-related side chains, Arg312 in the top–top interface and Lys414 in the bottom–bottom interface, would not be problematic since they were flexible with room for rotation away from one another. In the 'normal' top-up packing, close contacts were also observed between Lys357 and Lys414 at the top–bottom interface. More than 100 water molecules appeared in each crystal contact area.

Table 4
Specific bonds between canavalin trimers in the *R*3 crystal.

Residue 1	Atom 1	Residue 2	Atom 2	Comments
Ser53	OG	Arg392	NH1	d = 3.22 Å, hydrogen bond
Arg312	NE	Glu411	OE1	Mediated by Wat720, salt bridge
Thr339	OG1	Glu411	OE2	d = 3.49 Å, hydrogen bond
Asp344	OD1	Arg392	NH1	d = 3.00 Å, salt bridge
Asp344	OD2	Arg392	NH2	d = 2.81 Å, salt bridge

In the orthorhombic crystal, a canavalin trimer was in contact with seven neighbors through five types of interfaces. The disk-shaped molecules were stacked with their threefold axes approximately parallel to a (Ko, Ng, Day *et al.*, 1993). Fig. 6 shows a schematic diagram of the crystal packing. The most extensive interface (I) was observed between two trimers related by a twofold axis parallel to b and buried 773 Å^2 of surface area on each trimer (see Table 5). There were 22 amino-acid residues and six waters per trimer involved in interface I having dyad symmetry. Three possible salt bridges, Arg$A$254–Asp$C$128, Asp$A$255–Arg$A$388 and Lys$A$272–Asp$C$125, occurred at this interface but were not well defined, possibly owing to NCS restraints in the refinement. The second type of interface (II) was between molecules related by screw axes parallel to c and buried 444 Å^2 on one trimer and 438 Å^2 on the other. The numbers of residues involved were

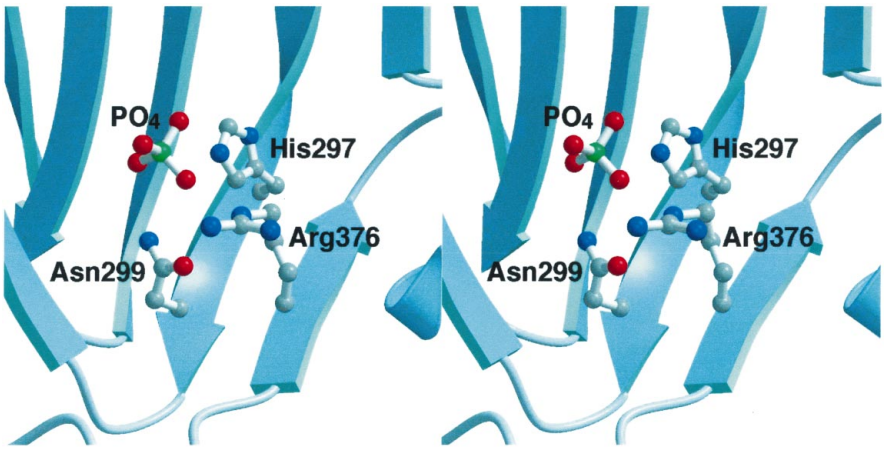


Figure 4
The bound phosphate ion in the C-terminal β-barrel. The side chains of His297, Asn299 and Arg376, which make hydrogen bonds with the phosphate, are also shown, as well as sections of the β-strands in the background. The model shown here corresponds to the A chain in the orthorhombic crystal.

Table 5
Specific bonds between canavalin trimers in the $C222_1$ crystal.

Interface I: molecules related by (1) x, y, z and (2) $1-x, y, \frac{1}{2}-z$. Residue 1 in molecule 2 also made equivalent bonds with residue 2 in molecule 1, *i.e.* the interactions were twofold redundant.

Residue 1	Atom 1	Residue 2	Atom 2	Comments
Asp*A*246	OD1	Lys*C*247	NZ	d = 3.24 Å, salt bridge
Ser*A*253	O	Glu*A*389	OE1	Mediated by Wat000, hydrogen bond
Arg*A*254	NE	Asp*C*125	OD2	d = 2.81 Å, salt bridge
Asp*A*255	OD1	Gln*A*389	NE2	d = 3.33 Å, hydrogen bond
Asp*A*255	N	Asp*C*125	O	d = 2.91 Å, hydrogen bond
Asp*A*255	O	Asp*C*125	OD1	Mediated by Wat000, hydrogen bond
Ser*A*274	OG	Lys*C*247	NZ	d = 2.87 Å, hydrogen bond
Arg*B*254	NH2	Asp*C*255	OD1	d = 3.16 Å, salt bridge

Interface II: molecules related by (1) x, y, z and (2) $\frac{1}{2}-x, \frac{1}{2}-y, \frac{1}{2}+z$.

Residue 1	Atom 1	Residue 2	Atom 2	Comments
Asn*C*54	ND2	Tyr*A*203	OH	d = 3.07 Å, hydrogen bond
Arg*C*312	NH2	Lys*A*181	O	d = 2.73 Å, hydrogen bond
Arg*C*312	NH2	Arg*A*182	O	d = 2.71 Å, hydrogen bond

Interface III: molecules related by (1) x, y, z and (2) $\frac{1}{2}-x, \frac{1}{2}+y, \frac{1}{2}-z$.

Residue 1	Atom 1	Residue 2	Atom 2	Comments
Gln*B*63	NE2	Glu*C*213	OE1	d = 3.21 Å, hydrogen bond
Lys*B*181	NZ	Ser*C*188	O	d = 3.35 Å, hydrogen bond
Lys*B*181	NA	Phe*C*190	O	d = 2.99 Å, hydrogen bond
Glu*B*213	OE1	Lys*C*192	N	Mediated by Wat000, hydrogen bond
Glu*B*213	OE2	Arg*A*335	NH2	d = 2.98 Å, salt bridge

13 and 11, respectively, plus four ordered waters. The third interface (III) buried 389 and 348 Å^2 on two trimers related by a screw axis parallel to b and involved ten and 11 amino-acid residues, respectively, plus three waters. Two other possible salt bridges, Arg*C*52–Glu*A*213 and Lys*C*55–Glu*A*207, were observed for interface II and another, Asp*B*204–Arg*A*52, for interface III. The fourth and fifth types of interface (IV and V) were between molecules related by twofold axes parallel to a [with symmetry operations $(x, -y, -z)$ and $(x, 1-y, -z)$, respectively]. Interface IV involved eight residues and buried 182 Å^2 on each trimer, while interface V involved only two residues and buried 8 Å^2.

Fig. 7 presents atomic force micrographs of the surface of a growing orthorhombic crystal of canavalin; they are also typical of those from the prismatic faces of the hexagonal crystal form. Development is characterized by the appearance of two-dimensional nuclei on the surfaces, with subsequent tangential extension by recruitment of molecules into advancing step edges. This is the mechanism that predominates, often exclusively, in the growth of most macromolecular crystals as shown by AFM (Malkin *et al.*, 1995, 1999). While stacking faults in these crystals, and also other types of imperfections, can occasionally be found, they are relatively uncommon.

Figs. 8 and 9 are AFM images of the faces obtained during growth of rhombohedral canavalin crystals. These are quite typical of many images obtained from rhombohedral canavalin crystals, including those reported by other laboratories (Land *et al.*, 1995, 1996, 1999). Their striking characteristics are the frequent appearance of screw dislocations in many forms, both double and single, left and right handed, and a high frequency of stacking faults and other related imperfections. Together, these dislocations and faults lead to a defect density that is more than an order of magnitude greater than for other macromolecular crystals and several orders of magnitude greater than for conventional crystals (Malkin *et al.*, 1996).

If one examines the step edges arising from screw dislocations at higher magnification, as in Fig. 10, the edges appear extremely irregular and rough. Such roughness is generally ascribed to a high level of impurities (Chernov, 1984; Chernov

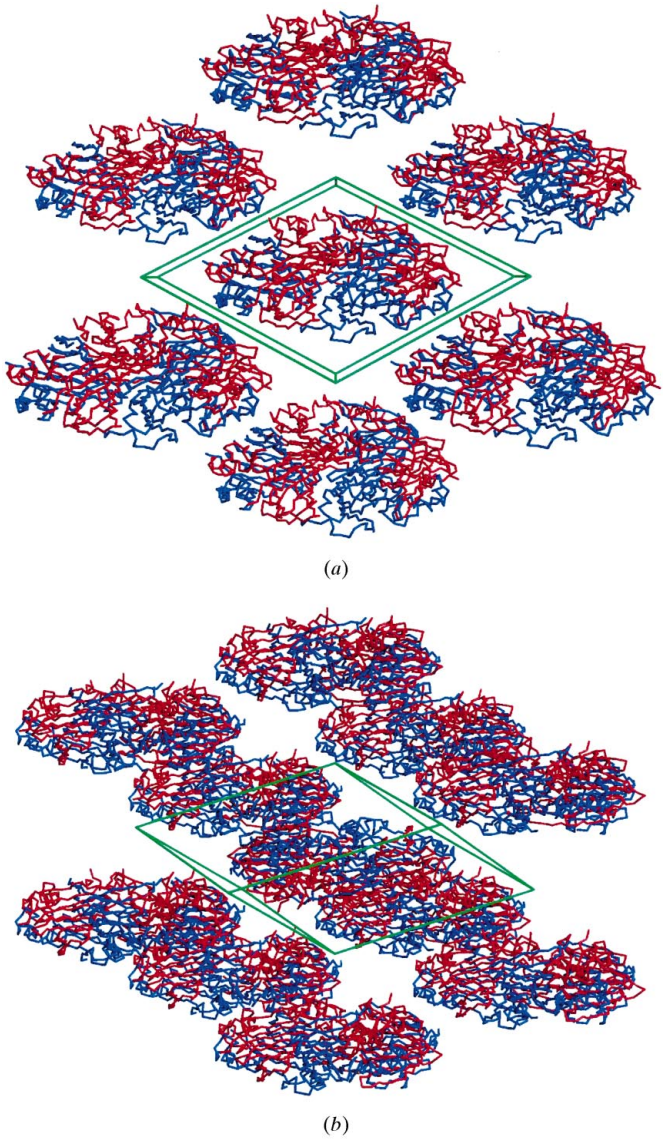


Figure 5
Packing of canavalin molecules in the twinned $R3$ crystal. (*a*) Normal arrangement with all top-up molecules. (*b*) An alternate view with a central bottom-up molecule. The rhombohedral unit cell is shown in green and the trimers in red and blue. Some molecules were omitted for clarity. In (*b*) the twofold symmetry of the top–top and bottom–bottom interfaces is apparent.

et al., 1988; Nakada *et al.*, 1999; McPherson *et al.*, 1996). Step edges arising from two-dimensional nuclei on the surfaces of hexagonal and orthorhombic crystals have a similar rough appearance, further consistent with impurities as the cause. AFM observations of amorphous material on the surfaces of growing rhombohedral and orthorhombic crystals suggests that the primary impurities may be aggregates of canavalin molecules themselves.

4. Discussion

The total interface area on a canavalin trimer was larger in the orthorhombic crystal (2 580 Å^2) than in the rhombohedral (2 043 Å^2) and the hexagonal (2 340 Å^2) crystals, but smaller than in the cubic crystal (3 282 Å^2). Excluding the cubic crystal, which had a very unique arrangement of the canavalin trimers, the other three crystal forms shared a similar arrangement of disk stacking. These crystals were seen to interconvert between one form and another, while the orthorhombic crystal was often the final and most stable form (Ko, Ng, Day *et al.*, 1993). This was consistent with its smaller specific volume and lower solvent content, as well as its larger crystal contact surface areas. In particular, interface I covered 773 Å^2 per trimer, significantly larger than all other contact interfaces of the canavalin crystals. It was also larger than the average of 280 Å^2 and was outside the range of 100–600 Å^2 surface area per contact in most crystals (Janin & Rodier, 1995). Through interface I, a pair of dyad-related trimers were bound 'bottom-to-bottom' by at least six salt bridges and ten hydrogen bonds. On the 'top' side of the trimer there were four 'normal' crystal contacts of interfaces II and III, with molecules related by screw axes (Fig. 6). The other two interfaces IV and V were smaller and occurred at the rims of disks. Perhaps the bottom-to-bottom dimer of canavalin trimers forms in solution prior to incorporation into a growing orthorhombic crystal, but there is no evidence otherwise.

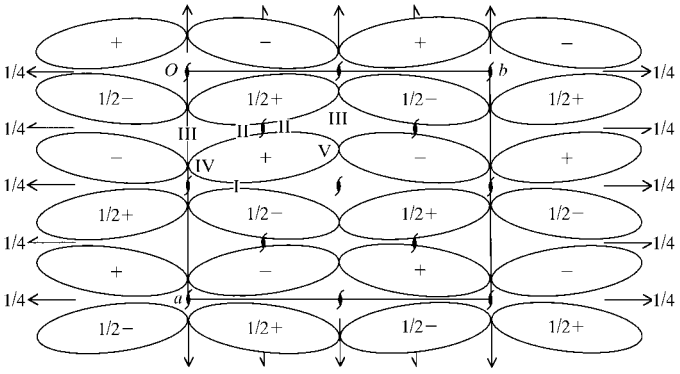
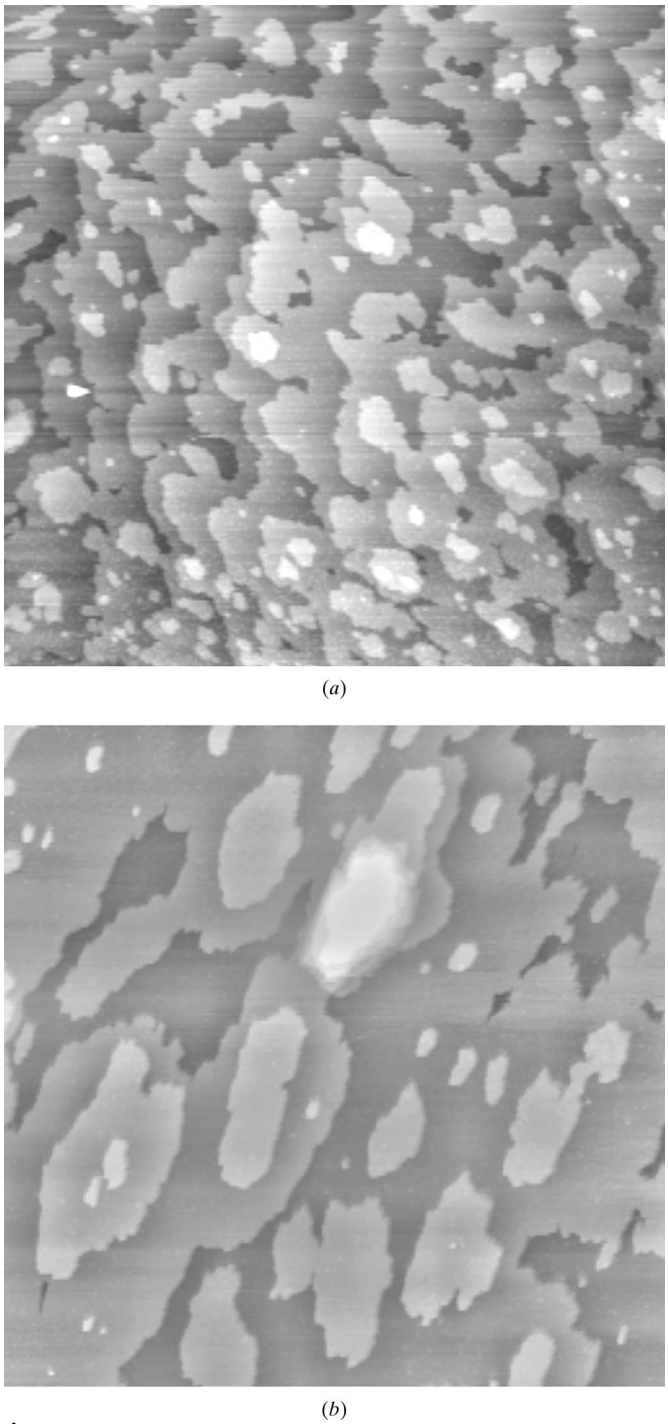


Figure 6
Schematic diagram of the $C222_1$ crystal packing, viewed along the crystallographic *c* axis. The disk-shaped trimers were stacked with the threefold axes almost parallel to the *a* axis (vertical). The reference molecule had a center of mass at fractional coordinates of approximately (0.36, 0.24, 0.18) and was in contact with seven other molecules through the labelled interfaces I–V.

Atomic force microscopy investigation of three of the four crystal forms of canavalin provides an interesting correlation. The cubic crystal form has been solved and refined by X-ray



(*a*)

(*b*)

Figure 7
In (*a*) is a 20 µm^2 area on the surface of a growing orthorhombic canavalin crystal, as visualized using atomic force microscopy, showing the extensive distribution of two-dimensional nuclei and islands. This image is typical of those from both hexagonal and orthorhombic canavalin crystals and illustrates the dominant, perhaps exclusive, growth mechanism common to these crystals. Also typical is the absence of screw dislocations and the relative scarcity of other types of dislocations or faults. In (*b*) is a higher magnification image of two-dimensional growth islands and step edges.

crystallography (Ko *et al.*, 2000), but is difficult to obtain reproducibly and has not been studied by AFM. The other three (rhombohedral, hexagonal and orthorhombic) crystal forms, however, have been examined. In addition, crystals of more than two dozen other proteins, viruses and nucleic acids have also been analyzed in our laboratory by AFM and these provide a broader context for the canavalin crystal investigations.

When the surface features and face development of the various crystals are compared, rhombohedral canavalin crystals stand out from all other crystalline macromolecules. Rhombohedral canavalin crystals contain an exceptionally large number of screw dislocations, which provide the major

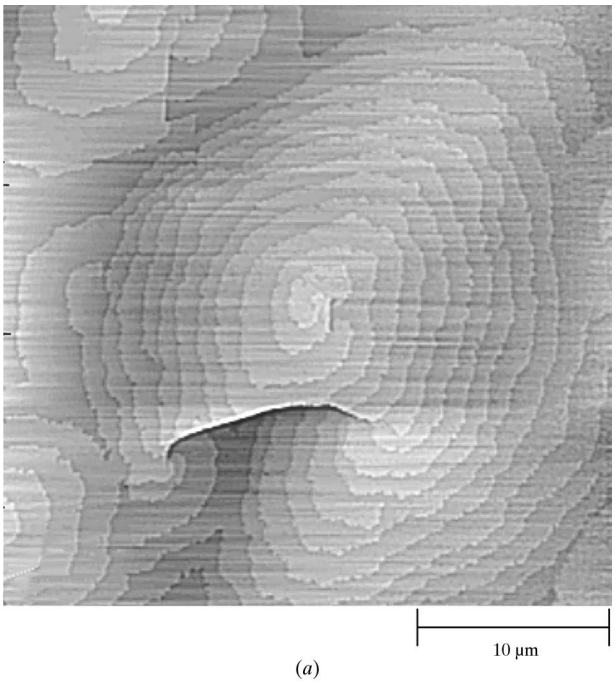


Figure 8
Representative AFM images from surfaces of growing rhombohedral canavalin crystals. In contrast to other forms of canavalin crystals, rhombohedral crystal surfaces display a vast array of screw dislocations frequently interrupted by stacking faults. Virtually no two-dimensional islands are seen, as the screw dislocations provide the sources for new step edges. Images like those shown here are unlike those recorded from all other macromolecular crystals studied using AFM.

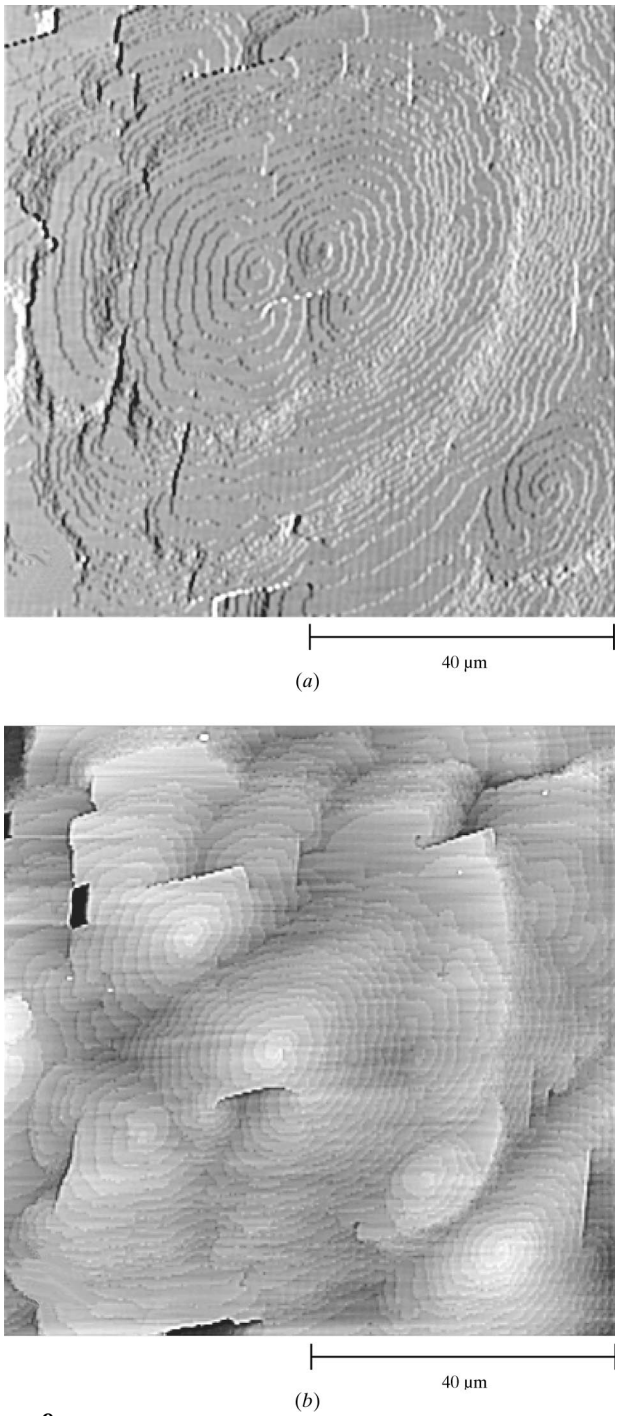


Figure 9
AFM images of screw dislocations on surfaces of rhombohedral canavalin crystals at higher resolution. The variety of types of screw dislocations and their modes of interaction are far more diverse than observed on any other protein crystals.

sources of growth steps for the crystals. The numbers of screw dislocations observed in the rhombohedral crystals, like those in Figs. 8 and 9, not only by us, but also by Land *et al.* (1995, 1996, 1999), is well over an order of magnitude greater than any other macromolecular crystal examined. They also contain an unusually high density of other defects, such as stacking faults (Malkin *et al.*, 1996). Other crystal forms of canavalin, on the other hand, exhibit virtually no screw dislocations and far fewer defects. Fig. 7, for example, is typical of the surface of a growing orthorhombic canavalin crystal and is also very similar to hexagonal prismatic surfaces. Both the hexagonal and the orthorhombic form grow almost exclusively by the deposition of two-dimensional nuclei on actively growing surfaces.

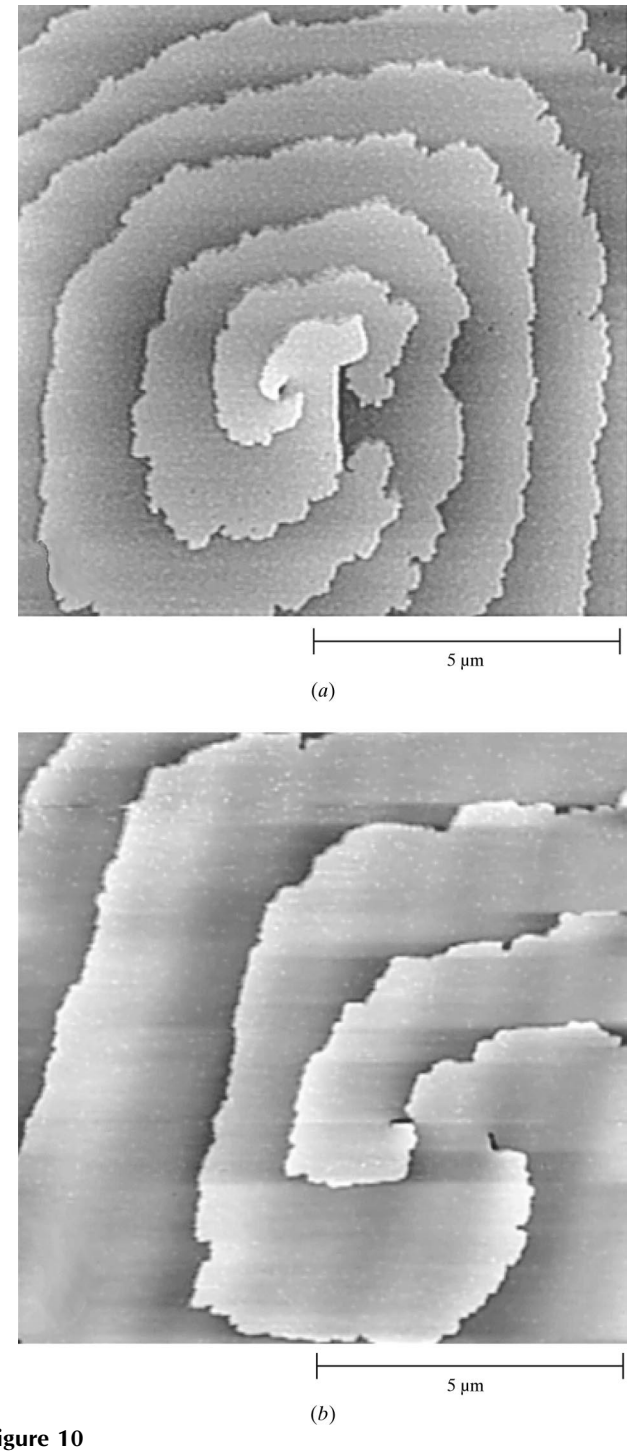


Figure 10
High-magnification AFM images of 5 µm^2 areas on surfaces of rhombohedral canavalin crystals showing the exceptional irregularity and roughness of the step edges, presumably a consequence of impurities.

Analysis for the presence of twinning according to Yeates (1997) using X-ray diffraction data show convincingly that about half of the canavalin trimers are 'top up', while the other half are 'bottom up'. The statistical analysis is consistent with a crystal composed of many domains or blocks that alternate randomly between molecules all in the top-up orientation, with domains composed of molecules all in the bottom-up orientation. Twinning, molecular misorientation and other misalignments create mechanical stress in crystal lattices. In most cases the stress is relieved, depending on the nature of the sources, by the formation of point defects or other crystal faults. For a crystal having ordered domains, half top up and half bottom up, then at domain boundaries where steps merge, stress would accumulate.

It has been well established in conventional crystal growth that stress at domain boundaries is most frequently relieved by the lattice through formation of screw dislocations or stacking faults (Burton *et al.*, 1951; Chernov, 1984; Chernov *et al.*, 1988). If canavalin trimers in the rhombohedral crystals were organized as alternative orientation domains, then the extreme occurrence of screw dislocations observed by AFM might well be anticipated.

Screw dislocations are not common to particular crystal classes or symmetries, as far as we know. They do not occur in other crystal forms of canavalin, which are not twinned as shown by X-ray analysis. Thus, it is not a consequence of some inherent property of the molecule. The only peculiarity of the rhombohedral crystals is their twinned lattice, a feature known in conventional crystals to promote formation of screw dislocations.

The canavalin molecule is unusual because of its disk shape that promotes stacking and its high degree of pseudo-symmetry, having almost 32 point-group symmetry. The high pseudo-symmetry is responsible for moderating the differences between the two kinds of domains in the rhombohedral crystals. This is probably a reflection of its physiological function, which is to store protein by packing it in seeds in a highly efficient fashion. Few simple objects pack more densely than plates into cylindrical stacks. The shape and packing propensity undoubtedly contributes to the ease of its crystallization and to the diversity of its crystal forms.

The rhombohedral crystal form has the highest water content, 56%, of the several crystal forms and makes the least intricate lattice interactions. It is, unless special precautions are taken, also somewhat unstable. Rhombohedral crystals will redissolve only a few days after growth, or dissolve and regrow into alternate forms, or their surfaces give rise directly to other crystals. The excessive dislocation and defect density probably contributes further to the instability. All of the other crystal forms of canavalin are stable once formed. The relaxed

packing of molecules and higher solvent content combined with molecular properties enable the twinning, which in turn gives rise to the defect structure.

Despite the twinning and the consequent defect density, rhombohedral canavalin crystals grow as large or larger than any other form, more rapidly and reproducibly than any other form, and from more impure preparations than any other form. In addition, when optimized they diffract to almost as high a resolution as the cubic crystals, the best diffracting form. Curiously, the twinning does not in itself appear to limit resolution.

Crystals grown in microgravity ultimately produced the highest resolution best quality diffraction data. Those crystals, however, were no less twinned than crystals grown in laboratories on earth. Similar results regarding the effect of microgravity on the twinning of alcohol dehydrogenase crystals was also reported by Esposito *et al.* (1998). Thus, gravity seems to play no role in the twinning phenomenon and suggests that crystallization of macromolecules in space will not be a therapy for this problem. On the other hand, mechanical stress produced by twinning may be less in microgravity or be better accommodated and a reduced defect density may result as a consequence.

Defect density may, we believe, have an important effect on diffraction resolution and mosaicity, and a reduction in defect density may explain the improvement in diffraction resolution for the microgravity grown crystals. We have not yet, however, had the opportunity to carry out AFM investigations of rhombohedral canavalin crystals grown in microgravity to evaluate their defect structures and densities.

The authors thank Dr Paul Adams for helpful discussion about the *CNS* system and twinning problems, Mr Aaron Greenwood for help in preparing illustrations and Dr Steven Larson for his comments regarding the manuscript. This research was supported by NASA, the NIH and DOE.

References

Breyer, W. A., Kingston, R. L., Anderson, B. F. & Baker, E. N. (1999). *Acta Cryst.* D**55**, 129–138.

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

Brünger, A. T. (1992*b*). *X-PLOR: A System for X-ray Crystallography and NMR.* New Haven, CT, USA: Yale University Press.

Brunger, 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.

Burton, W. K., Cabrera, N. & Frank, F. C. (1951). *Philos. Trans. R. Soc. London A*, **243**, 299.

Chernov, A. A. (1984). *Modern Crystallography*, Vol. III. *Crystal Growth.* Berlin: Springer–Verlag.

Chernov, A. A., Rashkovich, L. N., Smoliskii, I. L, Kuznetsov, Y. G., Mkrtchyan, A. A. & Malkin, A. J. (1988). *Growth of Crystals*, edited by E. I. Givargizov & S. A. Grinberg, Vol. 15, p. 43. New York: Consultant Bureau.

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

Cox, J. D., Kim, N. N., Traish, A. M. & Christianson, D. W. (1999). *Nature Struct. Biol.* **6** 1043–1047.

Day, J. & McPherson, A. (1992). *Protein Sci.* **1**, 1254–1268.

Dunwell, J. M., Khuri, S. & Gane, P. J. (2000). *Microbiol. Mol. Biol. Rev.* **64**, 153–179.

Esposito, L, Sica, F., Sorrentino, G., Berisio, R., Carotenuto, L., Giordano, A., Raia, C. A., Rossi, M., Lamzin, V. S., Wilson, K. S. & Zagari, A. (1998). *Acta Cryst.* D**54**, 386–390.

Janin, J. & Rodier, F. (1995). *Proteins*, **23**, 580–587.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Ko, T. P., Day, J. & McPherson, A. (2000). *Acta Cryst.* D**56,** 411–420.

Ko, T. P., Ng, J. D., Day, J., Greenwood, A. & McPherson, A. (1993). *Acta Cryst.* D**49**, 478–489.

Ko, T. P., Ng, J. D. & McPherson, A. (1993). *Plant Physiol.* **101**, 729–744.

Koszelak, S., Day, J., Leja, C., Cudney, R. & McPherson, A. (1995). *Biophys. J.* **69**, 13–19.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Kuznetsov, Yu. G., Malkin, A. J., Greenwood, A. & McPherson, A. (1995). *J. Struct. Biol.* **114**(3), 184–196.

Kuznetsov, Yu. G., Malkin, A. J., Land, T. A., DeYoreo, J. J., Barba, A. P. & McPherson, A. (1997). *Biophys. J.* **72**, 2357–2364.

Kuznetsov, Yu. G., Malkin, A. J. & McPherson, A. (1999). *J. Cryst. Growth*, **196**, 489–502.

Land, T. A., DeYoreo, J. J. & Lee, J. D. (1999). *Surf. Sci.* **384**, 136–155.

Land, T. A., Malkin, A. J., Kuznetsov, Yu. G., McPherson, A. & DeYoreo, J. J. (1995). *Phys. Rev. Lett.* **75**, 2774–2777.

Land, T. A., Malkin, A. J., Kuznetsov, Yu. G., McPherson, A. & DeYoreo, J. J. (1996). *J. Cryst. Growth*, **166**, 893–899.

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

Lawrence, M. C., Izard, T., Beuchat, M., Blagrove, R. J. & Colman, P. M. (1994). *J. Mol. Biol.* **238**, 748–776.

Lawrence, M. C., Suzuki, E., Varghese, J. N., Davis, P. C., van Donkelaar, A., Tulloch, P. A. & Colman, P. M. (1990). *EMBO J.* **9**, 9–15.

Luecke, H., Richter, H. T. & Lanyi, J. K. (1998). *Science*, **280**, 1934–1937.

Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. (1999). *J. Mol. Biol.* **291**, 899–911.

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

McPherson, A. (1996). *Crystallogr. Rev.* **6**, 157–308.

McPherson, A. (1999). *Seed Proteins*, edited by R. Casey & P. R. Shewry, pp. 241–257. London: Chapman & Hall.

McPherson, A., Malkin, A., Kuznetsov, Yu. G. & Koszelak, S. (1996). *J. Cryst. Growth*, **168**, 74–92.

McPherson, A. & Spencer, R. (1975). *Arch. Biochem. Biophys.* **169**, 650–661.

Malkin, A. J., Kuznetsov, Yu. G., Land, T. A., DeYoreo, J. J. & McPherson, A. (1995). *Nature Struct. Biol.* **2**, 956–959.

Malkin, A. J., Kuznetsov, Yu. G. & McPherson, A. (1996). *J. Struct. Biol.* **117**, 124–137.

Malkin, A. J., Kuznetsov, Yu. G. & McPherson, A. (1999). *J. Cryst. Growth*, **196**, 471–488.

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

Merrit, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Nakada, T., Sazaki, G., Miyashita, S., Durbin, S. D. & Komatsu, H. (1999). *J. Cryst. Growth*, **196**, 503–510.

Richardson, J. S. (1981). *Adv. Protein Chem.* **34**, 167–339.

Smith, S. C., Johnson, S., Andrews, J. & McPherson, A. (1982). *Plant Physiol.* **70**, 1199–1209.

Sumner, J. B. & Howell, S. F. (1936). *J. Biol. Chem.* **113**, 607–610.

Valegard, K., Terwisscha van Scheltinga, A. C., Lloyd, M. D., Hara, T, Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H. J., Baldwin, J. E., Schofield, C. J., Hajdu, J. & Andersson, I. (1998). *Nature (London)*, **394**, 805–809.

Weber, C. H., Park, Y. S., Sanker, S., Kent, C. & Ludwig, M. L. (1999). *Structure*, **7**, 1113–1124.

Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.

Yeates, T. O. & Fam, B. C. (1999). *Structure*, **7**, R25–R29.