Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of orthorhombic crystals of beef liver catalase

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

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

Correspondence e-mail: amcphers@uci.edu

Received 13 February 1999
Accepted 19 May 1999

PDB Reference: orthorhombic beef liver catalase, 4blc.

The growth mechanisms and physical properties of the orthorhombic crystal form of beef liver catalase were investigated using *in situ* atomic force microscopy (AFM). It was observed that the crystals grow in the ⟨001⟩ direction by an unusual progression of sequential two-dimensional nuclei of half unit-cell layers corresponding to the 'bottoms' and 'tops' of unit cells. These were easily discriminated by their alternating asymmetric shapes and their strong growth-rate anisotropy. This pattern has not previously been observed with other macromolecular crystals. Orthorhombic beef liver catalase crystals exhibit an extremely high defect density and incorporate great numbers of misoriented microcrystals, revealed intact by etching experiments, which may explain their marginal diffraction properties. To facilitate interpretation of AFM results in terms of intermolecular interactions, the structure of the orthorhombic crystals, having an entire tetramer of the enzyme as the asymmetric unit, was solved by molecular replacement using a model derived from a trigonal crystal form. It was subsequently refined by conventional techniques. Although the packing of molecules in the two unit cells was substantially different, with very few exceptions no significant differences in the molecular structures were observed. In addition, no statistically significant deviation from ideal 222 molecular symmetry appeared within the tetramer. The packing of molecules in the crystal revealed by X-ray analysis explained in a satisfying way the process of crystal growth revealed by AFM.

1. Introduction

Beef liver catalase (E.C. 1.11.1.6) is an iron-containing enzyme of molecular weight M_r = 240 kDa (Sumner & Garlen, 1935) which is responsible for the elimination of hydrogen peroxide without formation of free radicals. It catalyzes the reaction

$$2H_2O_2 \rightarrow 2H_2O + O_2.$$

The protein contains four heme groups (Schroeder *et al.*, 1982) and four NADPH molecules (Kirkman & Gaetani, 1984) distributed among its four identical subunits (Tanford & Lovrien, 1962), each of which consists of 506 known amino-acid residues. The properties of the enzyme have been thoroughly reviewed (Deisseroth & Dounce, 1970; Murthy *et al.*, 1981; Schonbaum & Chance, 1976). Beef liver catalase is easily crystallized in a variety of forms (Schnuchel, 1954; Rossmann & Labaw, 1967; McPherson & Rich, 1973) and was, in fact, one of the first enzymes crystallized (Sumner & Dounce, 1937).

The structure of beef liver catalase has been solved and refined at 2.5 Å resolution in a trigonal crystal form by

Rossmann and colleagues (Murthy *et al.*, 1981; Fita *et al.*, 1986). They found the tetramer to be bilobal and each subunit to be composed of four discrete domains. Based on their model, the active-site region centered on the protoporphyrin IX moiety was carefully analyzed and a mechanism for catalysis proposed (Fita & Rossmann, 1985). The structure of a catalase from the fungus *Penicillium vitale* has also been solved and refined to 2.0 Å resolution, though the amino-acid sequence of 670 residues remained somewhat ambiguous (Vainshtein *et al.*, 1986). The structure of the beef liver and the fungal proteins were compared and found to be very similar. Three other structures of catalase from *Micrococcus luteus* (Murshudov *et al.*, 1992), *Proteus mirabilis* (Gouet *et al.*, 1995) and *Escherichia coli* (Bravo *et al.*, 1995) were also determined, further demonstrating a common protein fold.

For the structure determination of both the beef liver and fungal catalases, crystals of space group $P3_121$ with strikingly similar unit-cell dimensions were utilized (Melik-Ademyan *et al.*, 1986), though each crystal was grown by quite different methods. In both cases, furthermore, the crystallographic asymmetric unit was composed of one half of the tetrameric molecule, or two protein subunits. In the case of the beef liver enzyme, however, it was demonstrated that there was virtually no statistically significant difference between non-crystallographically related subunits and that molecular 222 symmetry was maintained in the crystal (Fita *et al.*, 1986).

Beef liver catalase crystals were of particular interest to us in terms of protein crystal-growth mechanisms (Malkin *et al.*, 1995) and growth kinetics (Malkin *et al.*, 1997). This derived from the reproducibility of their nucleation and growth under a wide variety of crystallization conditions, their large size and because they had been studied extensively in the past by electron microscopy (Hall, 1950; Valentine, 1964; Labaw, 1967; Longly, 1967; Wrigley, 1968; Matricardi *et al.*, 1972), AFM (Malkin *et al.*, 1997, 1998) and X-ray diffraction (Rossmann & Labaw, 1967; Murthy *et al.*, 1981; McPherson & Rich, 1973; Fita *et al.*, 1986). The large asymmetric unit makes them ideal for many of these kinds of physical analyses and they have, in addition, provided systems for the study of protein crystallization in microgravity (McPherson, 1996; Koszelak *et al.*, 1996).

Earlier studies on catalase crystallization by AFM using PEG as the precipitant revealed that the crystals grew by two-dimensional nucleation, but that step edges were extremely rough and the surfaces were so irregular that a 'normal growth' mechanism (Malkin *et al.*, 1995) was approached. Subsequently, we found that much more uniform and ordered growth of these same crystals could be obtained under the alternative conditions used here.

To understand the molecular basis for the unusual mechanism of growth by nucleation and spread of 'half-step layers' on crystal surfaces seen by AFM, it was necessary to solve the three-dimensional structure of these crystals. Analyses of the packing of the tetrameric molecules in the crystal provides a satisfying explanation for the growth pattern observed by AFM and for the anisotropy of step-growth rates.

The crystals used in the molecular-replacement analysis and the AFM study reported here are of space group $P2_12_12_1$, with $a = 87.8$, $b = 140.6$, $c = 232.4$ Å, containing an entire molecule of four presumably identical subunits as the asymmetric unit (McPherson & Rich, 1973). These crystals have double the asymmetric unit size of the trigonal crystals. This has the consequence of placing no symmetry constraints upon the distribution or structure of subunits within the molecule. The orthorhombic crystals contain about 15% less solvent than do the trigonal crystals and on some occasions have been observed to arise by polymorphic transformation of the trigonal form.

2. Methods

2.1. Crystallization

Beef liver catalase was purchased from Sigma Chemical Co., St Louis, MO, dissolved in H_2O with a trace of thymol as preservative. Often, microcrystals of the enzyme were present in the samples either as received or upon storage at 277 K. The protein was diluted to a concentration of 40 mg ml^{-1} with water, and residual microcrystals were dissolved by the addition of a trace of NH_4OH and gentle warming under a hot-water tap. Prior to crystallization, the protein solution was centrifuged clear and filtered through 0.22 µm filters.

Crystallization of the protein for X-ray data collection was achieved using three principal sets of conditions. The crystals were isomorphous by every test that we applied, in spite of the differences in mother liquors. Vapor diffusion (McPherson, 1982) was used to grow the crystals under all three sets of conditions. In the first case, 25 µl protein droplets were equilibrated against reservoirs of 0.1 *M* sodium phosphate buffer at pH 6.2–6.8 containing 10% *n*-propanol at 290 K. The sample drops consisted of 15 µl 40 mg ml^{-1} protein plus 5 µl 1% NaCl plus 5 µl of the reservoir, having a pH of 6.2–6.8. In the second case, catalase at 40 mg ml^{-1} was combined with an equal volume of 0.05 *M* magnesium formate and equilibrated against the same formate concentration. In the third case, crystals were grown by mixing equal volumes of 40 mg ml^{-1} catalase in 0.05 *M* sodium phosphate at pH 6.4–6.8 with 12% PEG 4000 and equilibrating the protein droplets against the same concentration of PEG 4000.

2.2. X-ray diffraction

X-ray diffraction intensities were collected at 290 K from crystals mounted by conventional methods in quartz capillaries (McPherson, 1982), using a San Diego Multiwire Systems two-panel multiwire area-detector system (Hamlin *et al.*, 1981; Xuong *et al.*, 1985) with a 2θ table. Frame size was 0.12–0.14° with count times of 120 s frame^{-1}. The X-ray source was a Rigaku RU-200 rotating-anode generator fitted with a Supper graphite-crystal monochromator and operated at 45 mA and 50 kV. The radiation was Cu $K\alpha$ and a Huber three-circle goniostat was used for orientation of the crystals.

Initially, data were collected from a single crystal to 4.0 Å resolution and these data wre employed in the molecular-

replacement search. With successful completion of this analysis, data were collected to 2.8 Å, with some data extending to 2.6 Å resolution, from two additional crystals. Further data collection using three more crystals was carried out, principally over the high-resolution range 5–2.3 Å.[1] The latter data sets were merged with the first. Some statistics are presented in Table 1. The merging R factor for all data was 0.12, but for refinement only the 112000 structure amplitudes having an intensity to estimated error ratio $(F/\sigma) > 2$ were employed. This set represented 91% of the total possible reflections to 2.4 Å, but only about 50–60% of the data between 2.4 and 2.3 Å resolution.

Molecular-replacement procedures (Rossmann & Blow, 1962) used were those incorporated in the program *MERLOT* (Fitzgerald, 1988), which included the rotation function of Crowther & Blow (1967). Fourier calculations used the *FFT* of Ten Eyck (1985) and images were created and analyzed using the programs *FRODO* (Jones, 1985) and *TOM* running on an Evans and Sutherland PS390 graphics system and an SGI Indigo II workstation, respectively. All other calculations were performed on a Digital VAX cluster and an SGI system at the Institute of Molecular Biology.

Initial refinement of the catalase model was carried out using the constrained–restrained least-squares procedure in the program *CORELS* (Sussman *et al.*, 1977). Subsequent refinement of the complete model employed simulated annealing by *X-PLOR* (Brünger *et al.*, 1987; Bŕunger, 1991) and the restrained least-squares approach in *TNT* (Ten Eyck *et al.*, 1976; Tronrud *et al.*, 1987). Density modifications employed the program *DM* (Collaborative Computational Project Number 4, 1994), which included solvent flattening, histogram matching and fourfold molecular averaging. $2F_o - F_c$ and $F_o - F_c$ Fourier syntheses were used to improve the models, monitor changes and identify solvent molecules, again using *FRODO* and *TOM*.

2.3. AFM

Seed crystals of catalase for AFM experiments 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 0.05 *M* magnesium formate. Catalase crystals of sizes 50 × 500 µm were typically used in AFM experiments. For each experiment, a single thin-plate catalase seed crystal in a droplet of mother liquor was transferred and secured beneath a thin fiber (silicon carbide on tungsten, diameter 70 µm) glued to the glass substrate. The substrate with seed crystal was then immediately transferred to the sealed fluid cell of a Nanoscope E 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}) and an appropriate concentration of formate to produce the desired supersaturation. Conditions were changed by flushing the cell with a 1:1 mixture of protein and precipitant.

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–crystal contact was maintained in order to minimize the force applied to the crystalline surface. Solution supersaturation σ is defined as $\sigma = \ln(c/c_e)$, where c and c_e are the initial and equilibrium concentrations of the protein, respectively.

Table 1
Data-collection statistics.

Figures in parentheses refer to the resolution shell 2.3–2.2 Å.

Resolution (Å)	2.2
Unit-cell parameters (Å)	a = 87.8, b = 140.6, c = 232.4
Number of crystals	6
Number of observations	603135 (7434)
Unique reflections	128019 (6447)
Completeness (%)	88.5 (26.9)
$R_{\rm merge}$ (%)†	11.5 (22.6)

† $R_{\rm merge} = \sum |F_{\rm ave} - F_{\rm obs}| / \sum F_{\rm ave}$.

3. Results

3.1. X-ray

Self-rotation functions calculated directly from the observed diffraction data and expected to indicate the orientations of the molecular dyad axes contained no outstanding peaks. No convincing interpretation consistent with 222 molecular symmetry was apparent. This was also the case when data from selected resolution ranges or restricted intensity ranges were used. Because the molecule was known to possess 222 point-group symmetry in the trigonal crystal, we suspected that the axes were likely to be aligned roughly parallel with crystallographic symmetry elements.

A cross-rotation function search was carried out, again over a range of resolutions, using the structure of the catalase molecule determined from the trigonal unit cell (Fita *et al.*, 1986). The number of atoms involved exceeded the capacity of *MERLOT* if the entire tetramer were placed in a triclinic unit cell as the search probe. This was obviated by placing the catalase subunit as the asymmetric unit in an artificial orthorhombic unit cell of space group *P*222 for calculation of model structure factors, which was performed using *TNT* (Tronrud *et al.*, 1987).

For all resolution ranges employed, the maxima appeared at α = 95–105, β = 5–10, γ = 1–30° and symmetry-equivalent

[1] The structure analysis reported here was carried out using crystals grown in our laboratory at the University of California, which in general diffracted to no better than 2.7–2.5 Å resolution in the best of cases, even when of very large size. The ultimate resolution of the structure analysis was extended to 2.3 Å using crystals grown in the GN2/Dewar apparatus in microgravity aboard the Russian Space Station Mir. This approach uses a liquid–liquid diffusion method based on flash-frozen biphasic protein/precipitant samples and has been described in detail previously (Koszelak *et al.*, 1996). Crystals from microgravity experiments were, in some cases, more than 5 mm in length and 1 mm in the other dimensions. These allowed an entire set of X-ray diffraction data to be collected from a single crystal.

Table 2
Local 222 symmetry elements.

	From *TNT* rigid-body refinement (and used in NCS refinement by *X-PLOR*)	From subunit models refined by *X-PLOR* without NCS constraints
Center†	20.457, 27.469, 73.442	20.451, 27.472, 73.448
X axis	0.83791, 0.53910, −0.08529	0.83820, 0.53843, −0.08668
Y axis	−0.54566, 0.82372, −0.15407	−0.54524, 0.82410, −0.15353
Z axis	−0.01280, 0.17564, 0.98437	−0.01124, 0.17595, 0.98434

† Vectors are in orthogonal Å according to the crystallographic axial system.

orientations. Other prominent peaks occurred at $\beta = 0$ and 180° with coupled α and γ angles. Aside from these, the map was featureless. We then defined the orientation of the catalase model in the orthorhombic unit cell as $\alpha = 100$, $\beta = 8$, $\gamma = 20°$. This orientation implied rotation of the molecular dyad axes away from the crystallographic axes by only 10° in the case of *c* and ∼30° for the other two axes.

For the translation function of *MERLOT*, the extent of the coordinate set again exceeded the capacity of the program. As a consequence, the search model was taken to be a polyalanine model based on the known catalase structure with side-chain atoms removed beyond the β carbon. Nonetheless, the search model contained more than 10000 atoms, including those in the dinucleotide and the heme groups. For the three resolution ranges used, maxima were identical and fully consistent with a unique solution. These predicted the center of the tetramer to lie at (0.235, 0.195, 0.315) and symmetry-related positions. The initial *R* factor for the polyalanine model oriented and positioned according to this solution in the orthorhombic crystal was 0.45 for 4000 reflections to 6 Å resolution. The *R* factor decreased to 0.34 upon rigid-body refinement in *MERLOT* with less than a 5° change in orientation and virtually no movement of the molecular center. At 5 Å resolution, *R* was 0.37 for 7000 reflections. The entire catalase model, including side-chain atoms, was subsequently rigid-body refined using the program *TNT*. Using 16000 reflections to 4 Å resolution, an *R* factor of 0.31 was calculated with a correlation coefficient of 0.69. When each subunit of the tetramer was considered a rigid unit, *R* fell to 0.27, with a corresponding correlation coefficient of 0.76.

Although improvement in statistics was observed at 4.0 Å resolution when molecular-symmetry constraints were relaxed, in the next stages of refinement 222 molecular symmetry between the four subunits was maintained. With 222 molecular-symmetry constraints imposed, but treating the four structural domains of the catalase molecule (Murthy *et al.*, 1981) as rigid bodies, an *R* value of 0.29 was obtained at 2.8 Å resolution. The molecular center of the assembly was also allowed to shift until it demonstrated no further movement. The coordinates of the molecular center and the orientation of the molecular symmetry axes are presented in Table 2. This model became the starting point for non-crystallographic symmetry restricted refinement using *X-PLOR*.

Before further least-squares refinement, $2F_o - F_c$ difference Fourier maps were calculated for a 25 Å radius sphere about each of the four heme groups and the four NADP groups in the structure, as tests of the quality of the analysis to this point and, furthermore, to ensure that all of the groups were in fact present in the molecule. The portions of the molecules lying within the 25 Å spheres were, in each case, omitted from the calculation of the phases used for these maps. In every case, the heme or NADP groups appeared as positive difference

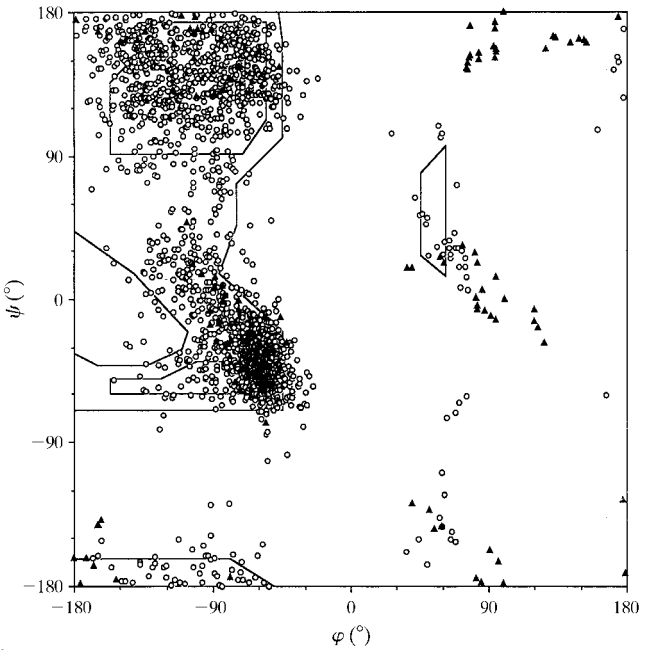

Figure 1
Ramachandran plot of the catalase model presented here. Glycine residues are shown as triangles, others as circles. Of the 1704 non-glycine and non-proline residues, 1370 were in most-favored regions and five, including four Ser216, were in the disallowed region.

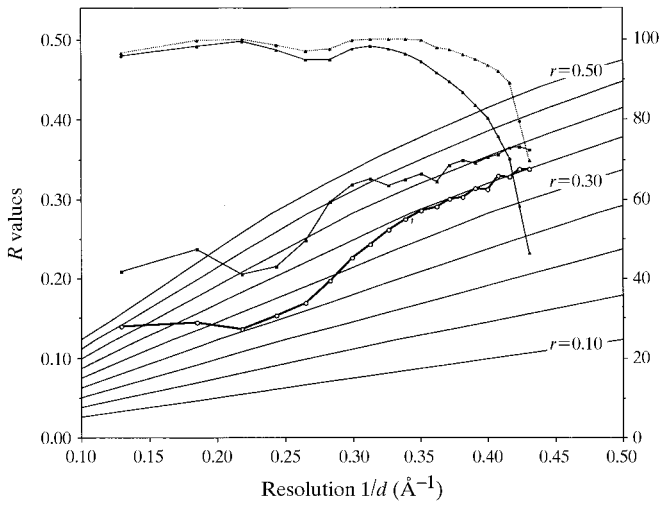

Figure 2
Luzzati plot for the refined model of catalase in the orthorhombic crystal. *R* values were calculated for 20 resolution shells of diffraction data with overall scaling factors for F_{calc}. The thick line with open circles represents the refined data set and the thin line with squares represents the test data set. The completeness of each shell is also plotted as triangles, with the dashed line representing all data and the solid line representing $F/\sigma > 2$ data.

Table 3
Crystallographic refinement.

Resolution range (Å)	20–2.3
Number of reflections ($F > 2\sigma_F$)	111957
R value based on 92% data	20.5
R_{free} for 8% test data set	27.3
Number of non-H atoms	16820
R.m.s.d. from ideal bond length (Å)	0.008
Bond angle (°)	1.344
Dihedral angle (°)	27.4
Improper angle (°)	0.960
Average B values all atoms† (Å^2)	47.8
Backbone atoms‡	46.4 (5988)
Side-chain atoms	47.8 (10084)
Heme atoms	40.6 (172)
NADP atoms	95.4 (192)
Solvent atoms	48.6 (384)

† Including N, C$^\alpha$, C. ‡ Number of atoms is given in parentheses.

density which was consistent in form and magnitude with the prosthetic groups.

Manual adjustments of the model were made with *FRODO* and *TOM* on the graphics systems where indicated by the difference Fourier maps. In every case, however, after modifications were made independently to each of the four subunits they were averaged about the dyads to restore 222 molecular symmetry. There were, in general, few significant changes from the starting model. The only noteworthy exceptions were at the amino and carboxyl termini, and the region centered on Gln21. The carboxyl terminus was extended by one residue, Glu501, which was a continuation of an α-helix. The present model consists of 499 amino-acid residues and still lacks seven apparently still disordered residues at the carboxyl terminus. Consistent with the map, the entire Gln21 was rotated 150° so that its side chain was on the opposite side of the polypeptide chain and minor adjustments were made to surrounding residues to accommodate this change.

The structure was subsequently refined using the simulated-annealing procedure in *X-PLOR*, with maintenance of strict non-crystallographic 222 symmetry. The resulting R value for the 68508 unique reflections in the resolution range 12–2.8 Å was 0.249. At this stage, no water molecules had been introduced. We believe it worth noting that maintenance of 222 molecular symmetry implied an effective asymmetric unit size of only one fourth of the tetrameric molecule. Thus, the ratio of observations to refineable parameters was roughly four times greater than would have been the case were the subunits unrestrained. The Ramachandran diagram showed the φ and ψ angles of all non-glycine residues except Ser216 to be within the allowed regions (Morris *et al.*, 1992). No temperature-factor refinement was attempted at this resolution.

Prior to refinement of the model at higher resolution, a new Fourier map was produced using the observed amplitudes and calculated phase angles from the program *DM* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The density-modification procedure included solvent flattening and fourfold molecular averaging. The effective resolution was extended to 2.5 Å, with an estimated R_{free} value of 0.268 for the data and a correlation of 0.920–0.934 between densities for the four catalase subunits. In $2F_o - F_c$ difference Fourier maps, all of the protein's backbone and side chains were clearly visible and a large number of strong densities arising from solvent molecules could also be identified. Water molecules were, therefore, added to the model in subsequent refinement cycles.

In the final phases of refinement, an additional 50000 reflections between 2.8 and 2.3 Å resolution were included, as well as low-resolution data between 20 and 12 Å. For cross-validation by R_{free}, randomly selected reflections representing 8% of the total were excluded from the working data set. The bulk-solvent approach of *X-PLOR* was employed to incorporate low-resolution data into the refinement. Non-crystallographic 222 symmetry restraints were gradually relaxed and eventually the four subunits were refined independently. Thermal parameters were also allowed to refine. In addition to the four protein subunits, the final model contained 384 waters and yielded an R value of 0.205 and an R_{free} of 0.273 for a total of 111957 reflections between 20 and 2.3 Å resolution having $F/\sigma > 2$.

Table 3 presents refinement statistics from *X-PLOR*. The r.m.s. deviations of bond lengths and angles from ideal values were 0.008 Å and 1.34°, respectively. Fig. 1 shows a plot of the φ, ψ angles (Morris *et al.*, 1992). Among the 1704 non-glycine non-

Figure 3
Right: a ribbon diagram of the beef liver catalase subunit is shown in red, with the bound heme and NADP in white. The central β arrangement surrounded by clusters of α-helices is evident. Left: four identical subunits are arranged with 222 symmetry to create the catalase tetramer of M_r = 260 kDa.

proline residues, 80.4% had φ, ψ angles in the most-favored regions. Besides the four Ser216 residues, only one residue, AlaD19, was outside the allowed regions. The plot in Fig. 2 indicates an estimated error of 0.25–0.35 Å for the atomic coordinates (Luzzati, 1952). The average temperature factor of the 16820 atoms in the model was 47.8 Å^2. The B values for four protein subunits were 39.1, 52.6, 50.5 and 46.9 Å^2. For the heme groups, they were 30.2, 39.4, 48.7 and 44.2 Å^2, for the NADP groups 91.1, 95.6, 96.0 and 98.7 Å^2, and for the 384 waters, 48.6 Å^2. Protein atoms with thermal factors greater than 80 Å^2 were confined to eight discrete regions in the polypeptide; these were the flexible loops and termini, especially the C-terminal helix and regions near B20 and C450. The structure of the 65 kDa catalase subunit with bound heme group and NADP is shown in Fig. 3(*a*). Four of these subunits, arranged in the tetramer having 222 symmetry, are shown in Fig. 3(*b*). Because the structure of the molecule in the orthorhombic crystals analyzed here almost imperceptibly varies from that described extensively in the trigonal crystals (Fita *et al.*, 1986), it serves no useful purpose to review its details. The minor differences which exist between the molecules in the two crystal forms are addressed below.

3.2. AFM

Previous studies of beef liver catalase crystals using AFM were of marginal value because the crystals, which were grown from PEG, exhibited irregular surfaces with very rough step edges. In addition, it was difficult to define supersaturation with a PEG system, and imaging with AFM was somewhat difficult because of the increased viscosity. Most of these limitations were overcome by using crystals grown by vapor diffusion from magnesium formate, which nonetheless had identical unit cells.

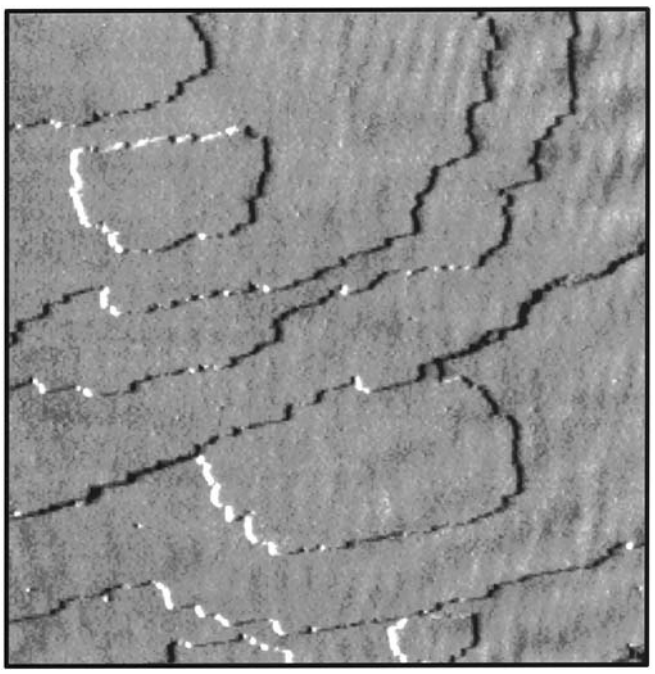

Figure 4
A 5 × 5 µm atomic force micrograph of the (001) surface of an orthorhombic beef liver catalase crystal showing step edges and two-dimensional islands. The roughness of the edges suggests the presence and incorporation of high levels of macromolecular impurities which retard normal step advancement at sites of addition.

For most macromolecular crystals that we have studied, including virus crystals, the growth-step edges into which molecules incorporate have heights of one unit-cell dimension above the surface. In addition, two-dimensional nuclei on surfaces generally exhibit the same general shape, even if the growth rates of the step edges are anisotropic. These and observations from other AFM studies further suggested that when macromolecular crystals grow they tend to complete construction of individual unit cells before initiating assembly of new unit cells. This certainly appears to be true for lysozyme and thaumatin crystals (Malkin *et al.*, 1998; Malkin, Kuznetsov & McPherson, 1996*a*). Orthorhombic catalase crystals provide a notable exception and demonstrate that other patterns of growth are indeed possible. Examination by atomic force microscopy of the (001) faces of orthorhombic catalase crystals reveals, first of all, the complete absence of screw dislocations. The face develops exclusively by two-dimensional nucleation at virtually all levels of supersaturation. Following island formation, which is the rate-limiting nucleation step, growth proceeds by tangential spread of the two-dimensional nuclei.

An unusual feature of catalase crystal growth (Malkin *et al.*, 1997) is that the heights of two-dimensional nuclei and growth steps, seen in Fig. 4, are 115 ± 2 Å, or one half of the c-axis dimension of 233.4 Å. Thus, the growth steps of the newly added layers correspond to the addition of only two of the four asymmetric units comprising the unit cell. The islands of half unit cells then spread tangentially well before their completion by subsequent formation of new two-dimensional islands atop the first. The progression of this pathway is unambiguous because the growth islands are clearly asymmetric in shape, exhibiting a 'flat side' and a 'curved side' in the ⟨100⟩ directions. Those corresponding to the 'bottom' halves of unit cells have an opposite hand, as shown in Figs. 5 and 6, to those corresponding to the 'top' halves of unit cells. Thus, the directions of the 'flat' and curved' sides alternate between successive layers, as illustrated in Fig. 7. From their visual appearance, islands can immediately be identified as being composed of 'bottom' or 'top' unit-cell halves. Observing growth *in situ* over a period a time shows that islands having 'left hands' and 'right hands' alternately add atop one another in a sequential manner.

As shown by the X-ray diffraction analysis, symmetry axes of catalase tetramers do not coincide with crystallographic directions; thus, molecules forming step edges moving in opposite crystallographic directions expose different distributions of charges, hydrogen-bond donors and acceptors, and hydrophobic clusters on the step edges, as well as differential screening by ions. This alone could explain the different rates of molecule incorporation into step edges advancing in opposite directions, though differential incorporation of impurities or differential effects of impurities cannot be ignored.

The growth of the catalase crystals by sequential addition of half unit cells is explicable when the packing of the catalase molecules in the orthorhombic unit cell is considered, as shown in Fig. 8. It is evident in Fig. 8 that molecules 1 and 2 are more extensively associated with one another than they are with 3 and 4 above. The converse is true for molecules 3 and 4. The molecules associate strongly in the (001) plane to form distinct layers, but only weakly associate with those above. It is important in this regard to emphasize that while every molecule in the interior of a crystal experiences an identical physical and chemical environment, this is not true for molecules on the surface and, therefore, for molecules as they join a growing lattice. Thus, growth rates of step edges reflect the differential strength of association in two-dimensional planes and the effects of molecules below and solvent above.

The number of contacts less than 4 Å between two layers is 22, compared with 38 contacts between catalase tetramers within each layer, a 72% difference. The surface areas on the two molecules which are buried or protected from solvent when molecules 1 and 2 associate are calculated to be 663 Å^2 on one and 637 Å^2 on the other or 1300 Å^2 in total. The solvent-inaccessible interfaces between molecules 1 and 3 or 2 and 4, however, are 379 and 401 Å^2, giving a total of 780 Å^2. Thus, the interfaces in the ⟨001⟩ plane are again about 70% greater than in the perpendicular direction. As a consequence of these different strengths of association, the nucleation and development of new unit cells in the ⟨100⟩ and ⟨010⟩ directions cause layer spread to proceed more readily than individual unit cells are completed in the ⟨001⟩ direction. This is the first case where we have observed this to be the case and it stands in marked contrast to other crystals we have studied such as lysozyme, thaumatin, lipase and xylanase.

There are vast solvent channels interrupting the crystal lattice at half unit-cell intervals along the *a* axis. These channels were also noteworthy in earlier electron-microscopy investigations of orthorhombic catalase crystals (McPherson & Rich, 1973). The existence of these pronounced solvent channels between successive growth layers may explain the forgiving nature of catalase crystal growth with respect to microcrystal incorporation (Malkin *et al.*, 1997; Malkin, Kuznetsov, Glantz *et al.*, 1996; Malkin, Kuznetsov &

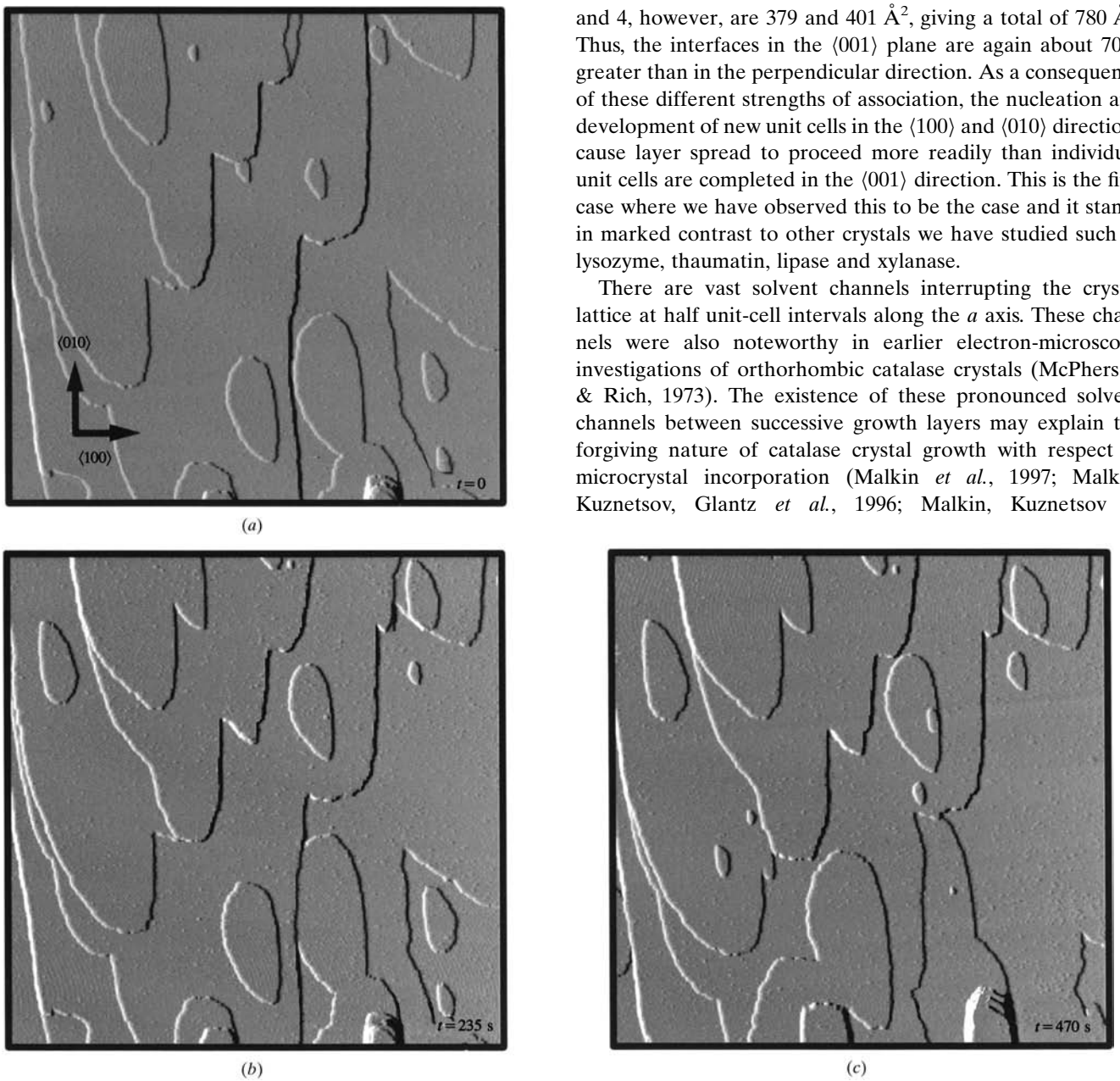

Figure 5
A sequence of AFM images of the identical surface area of 32.5 × 32.5 μm on the (001) face of an orthorhombic beef liver catalase crystal, at approximately 6 min intervals, showing the appearance of two-dimensional nuclei and their subsequent development. Note that the two-dimensional islands are asymmetric, having flat and curved sides. The two sides reverse direction depending on the asymmetry of the island on which they form, *i.e.* right-hand nucleates on the left and left-hand nucleates on the right. This can be understood if it is realised that one hand of the islands represents tops of unit cells, while the opposite hand represents layers of bottom halves of the unit cells.

McPherson, 1996*b*). As seen in Fig. 9, a remarkable number of misoriented microcrystals are readily incorporated into growing orthorhombic catalase crystals without apparent defect formation (Kuznetsov *et al.*, 1997). This is unusual when compared with impurity incorporation and defect formation in most other macromolecular crystals studied by AFM, where contaminant deposition is usually accompanied by stacking faults and spiral dislocations (Malkin, Kuznetsov, Glantz *et al.*, 1996; Malkin, Kuznetsov & McPherson, 1996*b*).

It is also noteworthy that growth by formation of half unit-cell layers appears to be independent of growth conditions and precipitant. In Fig. 10, for example, the same (001) face of an orthorhombic catalase crystal grown from PEG 4000, rather than sodium formate, is shown. The face is kinetically rough and quite different, for the most part, in its manner of growth, illustrating the dependence of growth mechanism and surface topology on the precipitant used to induce crystallization. Nonetheless, step edges are clearly visible and they too have heights of 115 ± 2 Å, corresponding to a half unit-cell dimension along *c*.

4. Discussion

The r.m.s. deviation of the model constrained by the 222 molecular symmetry from the unconstrained model is 0.703 Å. Comparison with the original model of Fita *et al.* (1986) from trigonal crystals shows that the constrained model derived here from orthorhombic crystals differs by an r.m.s. of 0.650 Å and the unconstrained model by an r.m.s. of 0.844 Å. As might have been anticipated, regions of maximum difference in coordinate positions are, in general, those displaying high temperature factors as well. Differences between the coordinates of the heme and NADP group for the PDB model and the constrained catalase model determined here are 0.397 and 1.158 Å, respectively. For the unconstrained model, they are 0.327 and 1.245 Å, respectively. For the constrained and unconstrained models of this analysis, the same values are 0.413 and 0.889, respectively. The very high temperature factors of the four NADP groups would seem to suggest some variability in binding to the protein.

The packing of catalase tetramers in the orthorhombic unit cell, shown in Fig. 8, when compared with the packing arrangement found in the trigonal crystal form shows there to be very little in common. The packing of the molecules in the two unit cells is, therefore, not consistent with the packing

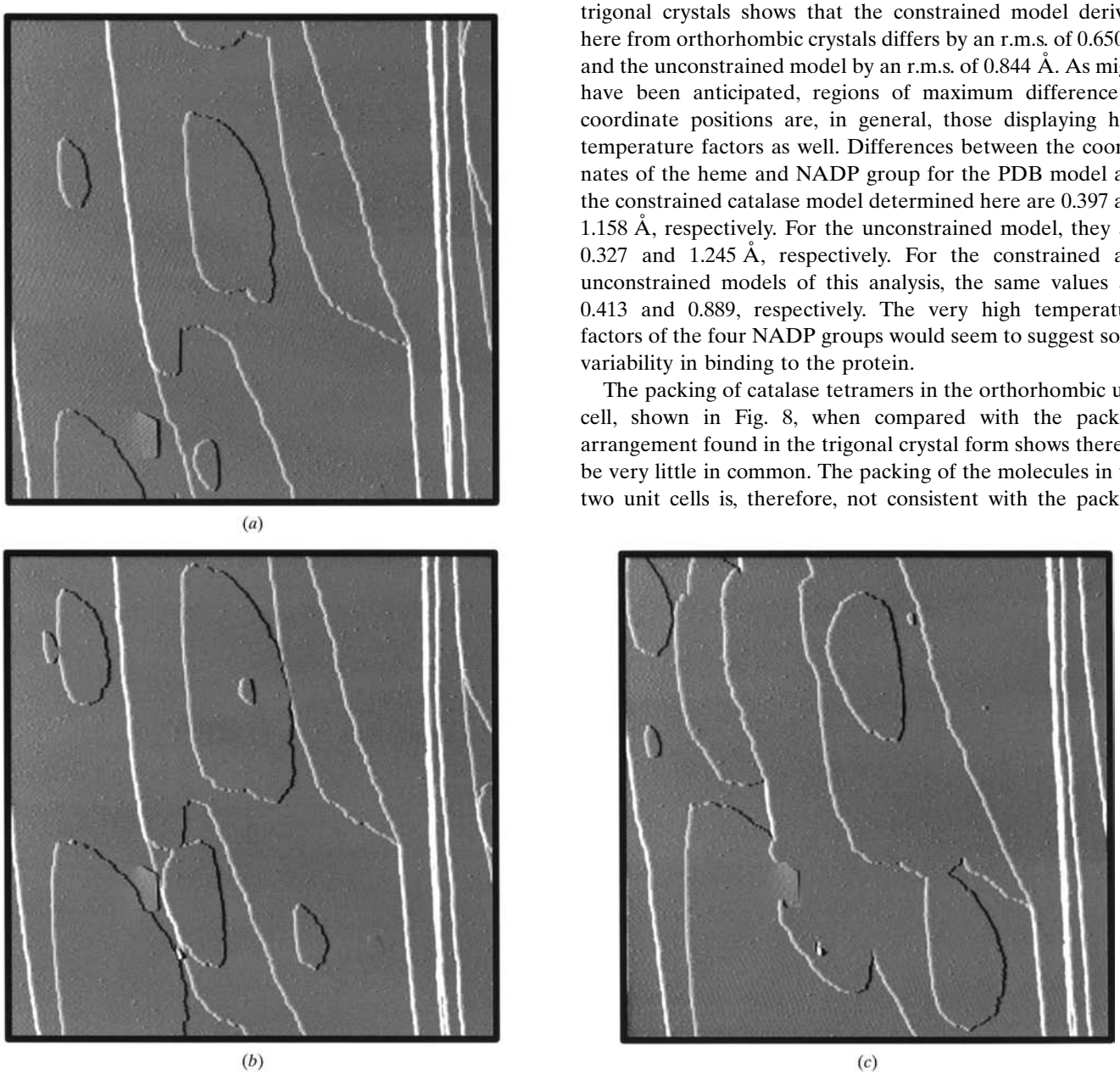

Figure 6
A second sequence of AFM images of a different area on the (001) face of a beef liver catalase crystal. The area is 32.5 × 32.5 µm and the interval between images is 12 min. Again, note the sequence of right- and left-handed islands which alternately appear on successive growth layers of the crystal.

proposal relating the two cells put forth by McPherson & Rich (1973). This suggests that the polymorphic transformation which has been observed between the two crystal forms must require dissolution of one form and *de novo* nucleation and growth of the other. This is in contrast to what is observed, for example, with yeast phenylalanine tRNA (Kim *et al.*, 1973) or the protein canavalin (McPherson & Spencer, 1975).

In the orthorhombic crystal, each catalase molecule makes contacts with four neighbors. By crystallographic symmetry relations, the intermolecular contacts comprise two different kinds of interface. The residues involved in hydrogen bonding and van der Waals contact are listed in Table 4. The first interface buries a solvent-accessible surface area of 663 Å^2 on one molecule (1) and 637 Å^2 on another (2), as calculated using the program *AREAMOL* of the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Besides the residues seen in Table 4, SerB425, AspB436 and AspB437 of the first molecule, as well as TyrC83 and GluC85 of the second molecule, are also buried in this interface. There are clearly at least two salt bridges; there are salt bridges between GluB289 and AspB427 of one molecule and ArgC105 and LysC314 of a second.

The second interface, along the *c* direction, involves a set of pseudo-twofold symmetry-related residues. The non-crystallographic dyad which relates subunits *A* and *D* of two different catalase molecules is approximately 10 Å from the crystallographic *c* axis. This interface buries 379 Å^2 of surface area on

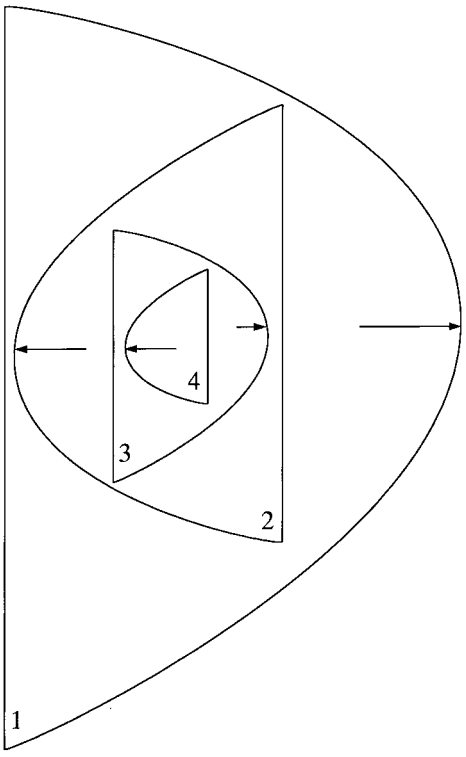

Figure 7
A diagram showing the alternation of islands of differing hands which nucleate one upon the other, alternately the bottoms and tops of crystallographic unit cells.

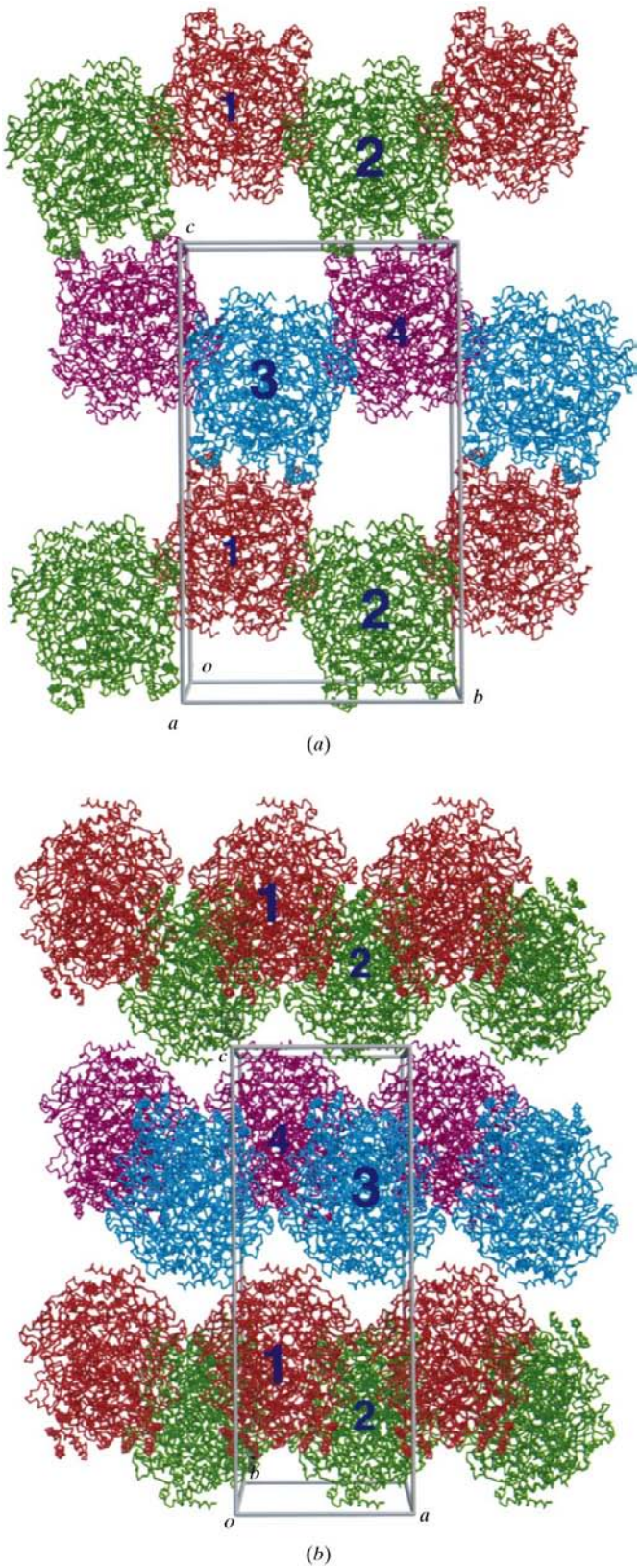

Figure 8
(*a*) and (*b*) are two orthogonal views of the packing of catalase tetramers in the $P2_12_12_1$ unit cell, from the crystallographic *a* and *b* axes, respectively. The protein molecules are colored and the unit cell is in gray. Figures were prepared using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Bacon & Anderson, 1988; Merritt & Murphy, 1994).

one molecule (1) and 401 Å^2 on another (2) and is visibly less extensive than the first interface. Two arginines, ArgA455 and ArgD455, are partially buried but not involved in direct contact. Salt bridges between LysA448 and GluD452 and between LysA452 and GluD448 are duplicated by symmetry and well defined. GluA487 also interacts with LysD448, but their symmetry-equivalent pairs are separated by more than 4.5 Å.

Among the 384 water molecules in an asymmetric unit of the orthorhombic crystal, only five reside in close interfaces between a catalase tetramer and symmetry-related neighbors. None of these, however, was found within 4 Å distance of two protein molecules and no bridging interactions could be identified. This is quite different from interface water molecules observed in many highly specific protein–protein or protein–DNA complexes (see, for example, Buckle *et al.*, 1994; Mol *et al.*, 1995). In those, waters not only form specific hydrogen bonds but also contribute significantly to the surface complementarity of opposing protein or DNA molecules. In the trigonal catalase crystal structure there are 50 waters per monomer, or 200 per tetramer. In that crystal, none are found in the vicinity of interfaces between tetramers. If entire models of both crystal forms are superimposed, C^{α} atoms agree in position with an r.m.s.d. of 0.403 Å and 99 waters from each model superimpose within 1 Å distance, if the small differences in local protein structure are taken into account. For a less stringent distance of 2 Å, the number of common waters is 135. These common solvent molecules are engaged in hydrogen bonds by plausible protein atoms and presumably play structural roles; only 10% are near the exterior protein surface. There are, nonetheless, 249 and 65 unmatched waters in the orthorhombic and trigonal models, respectively, but many of these could be a consequence, at least in part, of minor conformational differences in side chains and surface loops. These, however, are not involved in lattice interactions.

Table 4
Residues directly involved in crystal contacts.

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

Residue 1	Atom	Residue 2	Atom	d (Å)	Comments
ArgA421	NH2	TyrC279	OH	3.09	hydrogen bond
HisA423	ND1	AspC238	OD1	3.10	hydrogen bond
GluB289	OE1	ArgC105	NH2	2.95	salt bridge
GluB289	OE2	ArgC105	NE	2.74	salt bridge
IleB290	CD	ArgC379	NH1	3.61	VDW contact
HisB423	NE2	TyrC273	CD1	3.56	VDW contact
AspB427	OD2	Lys C314	NZ	2.78	salt bridge
GlnB429	OE1	HisC88	NE2	3.04	hydrogen bond
PheB431	CD1	ThrC87	O	3.80	VDW contact
AsnB432	O	IleC102	CG1	3.90	VDW contact
AsnB435	O	HisC101	NE2	3.06	hydrogen bond
ProC45	CG	ThrC87	CG2	3.51	VDW contact
ArgC46	NH1	GlyC103	O	2.83	hydrogen bond

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

Residue 1	Atom	Residue 2	Atom	d (Å)	Comments
NADP A2	N7A	GluD453	OE2	3.70	VDW contact
LysA448	NZ	GluD452	OE1	3.23	salt bridge
LysA448	O	GluD452	N	3.05	hydrogen bond
ValA449	O	AsnD451	CB	3.67	VDW contact
AsnA451	CB	ValD449	O	3.47	VDW contact
GluA452	N	LysD448	O	2.88	hydrogen bond
GluA452	OE1	LysD448	NZ	3.24	salt bridge
GluA453	OE1	NADP D2	C3′A	3.51	VDW contact
GluA487	OE2	LysD448	NZ	3.23	salt bridge

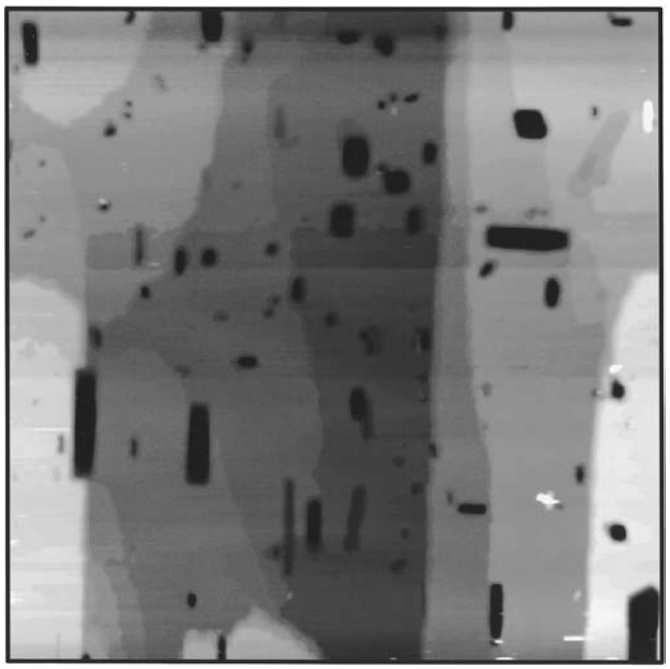

Figure 9
A 42 × 42 µm area recorded by AFM on the (001) surface of an orthorhombic catalase crystal which has been slightly etched near equilibrium conditions. This procedure reveals defects and incorporated impurities including incorporated microcrystals which are misoriented with respect to the underlying lattice. Here, it can be seen that there is a striking number of such misoriented microcrystals dispersed throughout the crystal, some as long as 10 µm.

One noteworthy feature of the packing of the molecules in the orthorhombic crystals is that combination of non-crystallographic 222 molecular symmetry with screw axes of the unit cell produces two non-intersecting pseudo-twofold axes in the unit cells. These pseudo-dyad axes, more or less parallel with the c crystallographic axis, relate adjacent tetramers. In retrospect, both of these axes were expressed in the self-rotation function and were to a large extent responsible for the difficulty of its interpretation.

In spite of the considerable differences in molecular contacts between catalase tetramers in this crystal and the trigonal form analyzed previously, there are virtually no significant structural differences in the catalase molecules. Thus, once again, intermolecular contacts in the crystal lattice are seen to have little influence on even the detailed structure of the constituent macromolecules. On the other hand, the importance of lattice interactions are seen to have a profound impact on the manner in which the crystals develop. This applies not only to the way layers spread, but to their surface nucleation as well. In the orthorhombic catalase crystals, the disparate strengths of association between molecules in the

lattice produce an unusual pattern of growth not previously seen for other macromolecular crystals. It also explains reasonably well the anisotropy in growth rates along different crystallographic directions and hence the habit development of the crystals.

It is noteworthy that orthorhombic crystals of beef liver catalase are rife with incorporated misoriented microcrystals which, nonetheless, do not serve as obstacles to continued growth. The catalase crystals, which exhibit the highest rates of microcrystal capture we have seen and suffer from more extensive sources of lattice strain, nonetheless grow to very large sizes. This suggests (as was also the case for canavalin and cubic STMV crystals) that ultimate crystal size is not strictly a function of defect accumulation, lattice stress or impurity incorporation but, perhaps, a function of those factors modulated by the extent of forgiveness or tolerance of the crystal lattice. Were this the case, then brittle rigid lattices which have many strong intermolecular interactions might nucleate more readily and grow more rapidly but reach smaller terminal sizes as a consequence of the lattices being less tolerant of stress. The high tolerance of the catalase lattice may derive from the weak bonding between the two layers making up the unit cells and from the extensive solvent channels which permeate the crystals and separate these layers.

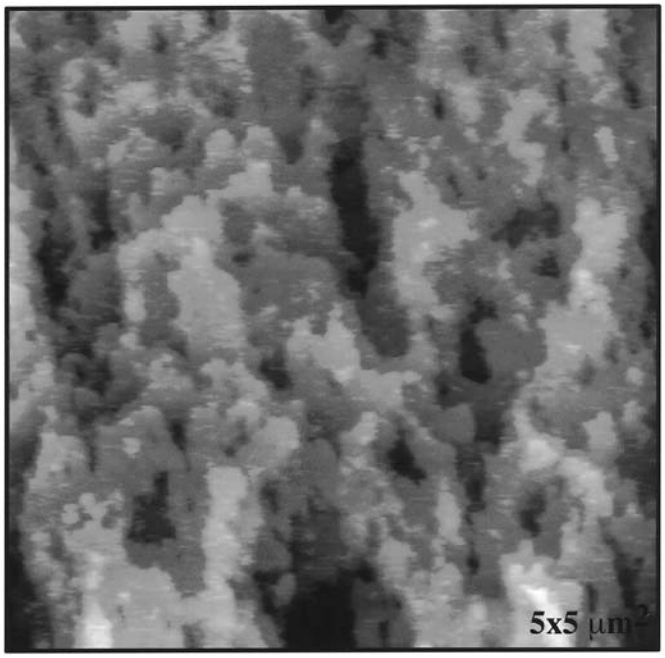

Figure 10
A 5 × 5 μm image recorded by AFM from the ⟨001⟩ face of an orthorhombic beef liver catalase crystal, but in this case grown from polyethylglycol rather than magnesium formate. Note that the surface is extremely rough, approaching the normal-growth regime of intensive random nucleation, and that two-dimensional islands are barely discernable. This image, when contrasted with other AFM images presented here, illustrates the profound influence which precipitant choice and growth conditions may have on the mechanisms by which a crystal develops.

This work was supported by a postdoctoral fellowship from Academia Sinica to T-PK and by grants from the National Aeronautics and Space Administration to AM. The authors also thank Dr Hanna S. Yuan of IMB for use of her computers.

References

Bacon, D. J. & Anderson, W. F. (1988). *J. Mol. Graph.* **6**, 219–220.
Bravo, J., Verdaguer, N., Tormo, J., Betzel, C., Switala, J., Loewen, P. C. & Fita, I. (1995). *Structure*, **3**, 491–502.
Brünger, A. T. (1991). *Annu. Rev. Phys. Chem.* **42**, 197–223.
Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.
Buckle, A. M., Schreiber, G. & Fersht, A. R. (1994). *Biochemistry*, **33**, 8878–8889.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Crowther, R. A. & Blow, D. M. (1967). *Acta Cryst.* **23**, 544–548.
Deisseroth, A. & Dounce, A. L. (1970). *Physiol. Rev.* **50**, 319–375.
Fita, I. & Rossmann, M. G. (1985). *J. Mol. Biol.* **185**, 21–37.
Fita, I., Silve, A. M., Murthy, M. R. N. & Rossmann, M. G. (1986). *Acta Cryst.* B**42**, 497–515.
Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.
Gouet, P., Jouve, H. M. & Dideberg, O. (1995). *J. Mol. Biol.* **249**, 933–954.
Hall, C. E. (1950). *J. Biol. Chem.* **185**, 45–51.
Hamlin, R., Cork, C., Howard, A., Nielson, C., Vernon, W., Matthews, D. & Xuong, N.-H. (1981). *J. Appl. Cryst.* **14**, 85–89.
Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.
Kim, S. H., Quigley, G. J., Suddath, F. L., McPherson, A., Sneden, D., Kim, J. J., Weinzierl, J. & Rich, A. (1973). *J. Mol. Biol.* **75**, 429.
Kirkman, H. N. & Gaetani, G. F. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 4343–4348.
Koszelak, S., Leja, C. & McPherson, A. (1996). *Biotechnol. Bioeng.* **52**, 449–458.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kuznetsov, Yu. G., Malkin, A. J., Land, T. A., DeYoreo, J. J., Barba, A. P. & McPherson, A. (1997). *Biophys. J.* **72**, 2357–2364.
Labaw, L. W. (1967). *J. Ultrastruct. Res.* **17**, 327–341.
Longly, W. (1967). *J. Mol. Biol.* **30**, 323–328.
Luzzati, P. V. (1952). *Acta Cryst.* **5**, 802–810.
McPherson, A. (1996). *Crystallogr. Rev.* **6**(2), 157–308.
McPherson, A. (1982). *The Preparation and Analysis of Protein Crystals.* New York: Wiley & Co.
McPherson, A. & Rich, A. (1973). *Arch. Biochem. Biophys.* **157**, 23–27.
McPherson, A. & Spencer, R. (1975). *Arch. Biochem. Biophys.* **169**, 650–661.
Malkin, A. J., Kuznetsov, Yu. G., Glantz, W. & McPherson, A. (1996). *J. Phys. Chem.* **100**, 11736–11743.
Malkin, A. J., Kuznetsov, Yu. G., Land, T. A., DeYoreo, J. J. & McPherson, A. (1995). *Nature Struct. Biol.* **2**(11), 956–959.
Malkin, A. J., Kuznetsov, Yu. G. & McPherson, A. (1996*a*). *J. Struct. Biol.* **117**, 124–137.
Malkin, A. J., Kuznetsov, Yu. G. & McPherson, A. (1996*b*). *Proteins*, **24**, 247–252.
Malkin, A. J., Kuznetsov, Yu. G. & McPherson, A. (1997). *Surf. Sci.* **393**, 95–107.
Malkin, A. J., Kuznetsov, Yu. G. & McPherson, A. (1999). *J. Cryst. Growth*, **196**, 471–488.
Matricardi, V. R., Morete, R. C. & Parsons, D. F. (1972). *Science*, **177**, 268–270.
Melik-Ademyan, W. R., Barynin, V., Vagi, A. A., Borisov, V., Vainshtein, B. K., Fita, I., Murthy, M. R. N & Rossmann, M. G. (1986). *J. Mol. Biol.* **188**, 63–72.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, S., Kavli, B., Krokan, H. E., Mosbaugh, D. W. & Tainer, J. A. (1995). *Cell*, **82**, 701–708.
Morris, A. L., McArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). *Proteins*, **12**, 345–364.
Murshudov, G. N., Melik-Adamyan, W. R., Grebenko, A. I., Barynin, V., Vagi, A. A., Vainshtein, B. K., Dauter, Z. & Wilson, K. S. (1992). *FEBS Lett.* **312**, 127–131.
Murthy, M. R. N., Reid, T. J., Sicignano, A., Tanaka, N. & Rossmann, M. G. (1981). *J. Mol. Biol.* **152**, 465–499.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Rossmann, M. G. & Labaw, L. W. (1967). *J. Mol. Biol.* **29**, 315–316.
Schnuchel, G. (1954). *Hoppe-Seyler's Z. Physiol. Chem.* **298**, 16–24.
Schonbaum, G. R. & Chance, B. (1976). *The Enzymes*, 3rd ed., Vol. 13, edited by P. D. Boyer, pp. 363–408. New York: Academic Press.
Schroeder, W. A., Shelton, J. R., Shelton, J. B., Robberson, B., Apell, G., Fang, R. S. & Bonaventura, J. (1982). *Arch. Biochem. Biophys.* **214**, 397–421.
Sumner, J. B. & Dounce, A. L. (1937). *J. Biol. Chem.* **121**, 417–424.
Sumner, J. B. & Garlen, N. (1935). *J. Biol. Chem.* **125**, 33–36.
Sussman, J. L., Holbrook, S. R., Church, G. M. & Kim, S.-H. (1977). *Acta Cryst.* A**33**, 800–804.
Tanford, C. & Lovrien, R. (1962). *J. Am. Chem. Soc.* **84**, 1892–1896.
Ten Eyck, L. F. (1985). *Methods Enzymol.* **115**, 324–337.
Ten Eyck, L. F., Weaver, L. A. & Matthews, B. W. (1976). *Acta Cryst.* A**32**, 344–350.
Tronrud, D. E., Ten Eyck, L. F & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.
Vainshtein, B. K., Melik-Adamyan, W. R., Barynin, V., Vagi, A. A., Grebenko, A. I. & Brisov, V. (1986). *J. Mol. Biol.* **188**, 49–61.
Valentine, R. C. (1964). *Nature (London)*, **204**, 1262–1264.
Wrigley, N. G. (1968). *J. Ultrastruct. Res.* **24**(5), 454–463.
Xuong, N.-H., Nielson, C., Hamlin, R. & Anderson, D. (1985). *J. Appl. Cryst.* **18**, 342–360.