CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173–178. New York: Gordon & Breach.
DEREWENDA, Z. S. & HELLIWELL, J. R. (1989). *J. Appl. Cryst.* **22**, 123–137.
FARBER, G. K. & PETSKO, G. A. (1990). *TIBS*, **15**, 228–234.
HELLIWELL, J. R., GREENHOUGH, T. J., CARR, P. D., RULE, S. A., MOORE, P. R., THOMPSON, A. W. & MORGAN, T. S. (1982). *J. Phys. E*, **15**, 1363–1372.
HENDRICKSON, W. A. (1985). *Methods Enzymol.* **115**, 252–270.
HOWARD, A. J., GILLILAND, G. L., FINZEL, B. C., POULOS, T. L., OHLENDORF, D. H. & SALEMME, F. (1987). *J. Appl. Cryst.* **20**, 383–387.
JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
MACGREGOR, A. (1988). *J. Protein Chem.* **7**, 399–415.
MCPHERSON, A. (1989). *Preparation and Analysis of Protein Crystals*, p. 96. Florida: Robert E. Krieger.
MATSUURA, Y., KUNUSOKI, M., HARADA, W. & KAKUDO, M. (1984). *J. Biochem.* **95**, 697–702.
MATTHEWS, B. W. (1966). *Acta Cryst.* **20**, 230–239.
NIXON, P. E. & NORTH, A. C. T. (1976). *Acta Cryst.* A**32**, 320–325.
PHILLIPS, D. C., STERNBERG, M. J. E., THORNTON, J. M. & WILSON, I. A. (1978). *J. Mol. Biol.* **119**, 329–351.
SERC Daresbury Laboratory (1986). *CCP4. A Suite of Programs for Protein Crystallography*. SERC Daresbury Laboratory, Warrington, England.
SIM, G. A. (1959). *Acta Cryst.* **12**, 813–815.
SWIFT, H. J., BRADY, L., DEREWENDA, Z. S., DODSON, E. J., DODSON, G. G., TURKENBURG, J. P. & WILKINSON, A. J. (1991). *Acta Cryst.* B**47**, 535–544.
WANG, B. C. (1985). *Methods Enzymol.* **115**, 90–112.

Acta Cryst. (1991). B**47**, 535–544

Structure and Molecular Model Refinement of *Aspergillus oryzae* (TAKA) α-Amylase: an Application of the Simulated-Annealing Method

BY HELEN J. SWIFT, LEO BRADY, ZYGMUNT S. DEREWENDA,* ELEANOR J. DODSON, GUY G. DODSON, JOHAN P. TURKENBURG AND ANTHONY J. WILKINSON

Department of Chemistry, University of York, Heslington, York YO1 5DD, England

(*Received* 11 *September* 1990; *accepted* 7 *February* 1991)

Abstract

Monoclinic crystals of a neutral α-amylase from *Aspergillus oryzae*, containing three molecules in the asymmetric unit, have been reported previously and studied at 3 Å resolution [Matsuura, Kunusoki, Harada & Kakudo (1984). *J. Biochem.* **95**, 697–702]. Here we report the solution of the structure of this enzyme in a different crystal form (space group $P2_12_12_1$, $a = 50{\cdot}9$, $b = 67{\cdot}2$, $c = 132{\cdot}7$ Å), with only one molecule in the asymmetric unit. The structure was solved by the molecular replacement method, using a model of acid α-amylase from a related fungus *A. niger* [Brady, Brzozowski, Derewenda, Dodson & Dodson (1991). *Acta Cryst.* B**47**, 527–535]. Conventional least-squares crystallographic refinement failed to converge in a satisfactory manner, and the technique of molecular dynamics in the form of the *XPLOR* package [Brunger (1988). *XPLOR Manual.* Yale Univ., USA] was used to overcome the problem. A large rigid-body type movement of the C-terminal domain was identified and accounted for. The final round of restrained least-squares refinement (at 2·1 Å resolution) including 3675 protein atoms and 247 water molecules resulted in a conventional crystallographic *R* factor of 0·183 and an atomic model which conforms well to standard stereochemical parameters (standard deviation of bond lengths from their expected values is 0·028 Å, while that for planar groups is 0·029 Å).

Introduction

Neutral α-amylase (α-1,4-glucan-4-glucanohydrolase, EC 3.2.1.1) from *Aspergillus oryzae* (TAKA amylase) is a glycoprotein which catalyses the hydrolysis of internal α(1–4) glycosidic linkages in various polysaccharides. The enzyme consists of a single polypeptide chain of 478 amino-acid residues for which the sequence has been determined (Toda, Konda & Narita, 1982; Tada, Imura, Gomi, Takaahaski, Hara & Yoshizawa, 1989).

The structure of TAKA α-amylase has been investigated by X-ray crystallography at 3·0 Å resolution (Matsuura, Kunusoki, Harada & Kakudo, 1984). The molecule was shown to consist of two main domains (see Fig. 1), with the N-terminal domain (residues 1 to 380) further subdivided into two smaller subdomains *A* and *B*. Domain *A* (residues 1 to 121 and 177 to 380) is folded into an $(\alpha/\beta)_8$ barrel, a pattern first noted in triose phosphate isomerase (Phillips, Sternberg, Thornton & Wilson,

* Present address to which correspondence should be sent: MRC Group in Protein Structure and Function, Department of Biochemistry, 474 Medical Sciences Building, University of Alberta, Edmonton, Canada T6G 2H7.

0108-7681/91/040535-10$03.00

1978) and later observed in 16 other proteins (Farber & Petsko, 1990). Domain *B* (residues 122 to 176) lies between the third β-strand and the third helix of domain *A* and contains a three-stranded antiparallel β-sheet and a region of less regular structure. The C-terminal residues (384 to 478) forming domain *C* are folded into an eight-stranded β-sandwich.

Two other α-amylase structures have been reported; the porcine pancreatic amylase (Buisson, Duee, Haser & Payan, 1987) at 2·9 Å resolution and *Aspergillus niger* α-amylase at 2·1 Å resolution (Boel *et al.*, 1990; Brady, Brzozowski, Derewenda, Dodson & Dodson, 1991). The latter has been refined to an *R* factor of 0·167 and constitutes the most accurate amylase structure determination to date.

The requirement of α-amylases for Ca^{2+} for maintaining their catalytic activity had long been established (Vallee, Stein, Sumerwell & Fisher, 1959). The metal-binding site was located in porcine pancreatic α-amylase (PPA) and was proposed to be essential for the structural integrity of the active site (Buisson *et al.*, 1987). The high-resolution study of *A. niger* α-amylase revealed the calcium-binding site in detail and in addition a second low-affinity inhibitory site was identified (Boel *et al.*, 1990).

TAKA amylase was originally studied in our laboratory in an effort to solve the structure of the acid α-amylase (NOVO α-amylase) from *A. niger* (Brady *et al.*, 1991). Crystals of the latter enzyme had been grown at pH 3·0, making production of heavy-atom derivatives difficult. Efforts were made to use the published 3 Å TAKA amylase coordinates (Matsuura *et al.*, 1984) as a model for molecular replacement; however, the solution was unclear. Attempts were than made to repeat the crystallization work of Matsuura *et al.* (1984) but these were also unsuccessful. Crystals of TAKA amylase were eventually grown at pH 6·0, using the protocol described below. These crystals were of space group $P2_12_12_1$, with one molecule in the asymmetric unit. They were found to be suitable for X-ray studies with a diffraction pattern extending beyond 1·8 Å (using a synchrotron source). It was therefore decided to solve and refine the TAKA amylase structure and obtain an accurate model of this enzyme, in an effort to elucidate the molecular basis of the different catalytic pH profiles of the two amylases.

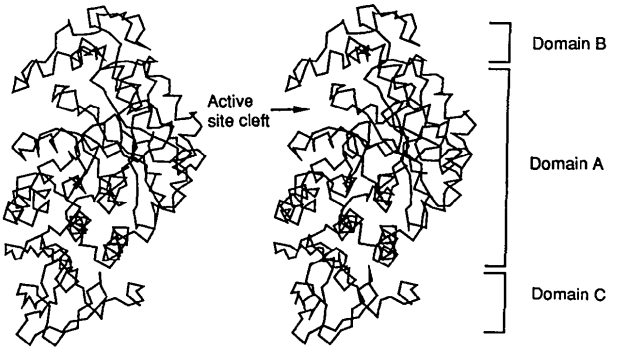

Fig. 1. A schematic stereo representation of the TAKA α-amylase molecule, showing the Cα atoms.

The crystallographic refinement of the molecular model of TAKA amylase, based on the *A. niger* structure, converged with an *R* factor of 0·28; both the value of this parameter, and the quality of the electron density map suggested that false convergence had been reached at a local minimum. The use of simulated annealing, first suggested for X-ray refinement purposes by Brunger, Kuriyan & Karplus (1987) allows one to overcome large energy barriers. Several successful applications of this procedure have been reported (*e.g.* Brunger, Karplus & Petsko, 1989; Fujinaga, Gros & van Gunsteren, 1989; Postma, Parker & Tsegnoglou, 1989; for a review see Goodfellow, Henrick & Hubbard, 1989). They illustrated that the use of molecular dynamics reduces the time usually dedicated to manual rebuilding, thereby automating the refinement process. In this paper we present the results of the application of the *XPLOR* package to refine the structure of the TAKA amylase from a molecular model of the *A. niger* enzyme. It exemplifies a successful application of crystallographic refinement with simulated annealing in a case where an atomic model was not only outside the radius of convergence of classical methods, but failed to lead to an electron density map interpretable enough to allow for manual intervention.

Crystallization

The crude TAKA amylase (Fungamyl Novo) was purified prior to crystallization studies. Low molecular weight substances were removed by ultrafiltration using tap water, followed by purification on a Q-Sepharose column equilibrated with 20 m*M* piperazine–HCl (pH 6·5). An eluent of the same buffer containing an NaCl gradient (0–0·6 *M*) was then used with the TAKA amylase being eluted at 0·2 *M* NaCl (30–33% 0·6 *M* NaCl). The amylase-containing fractions, giving a single peak on a TSK 3000 column, were pooled, concentrated by ultrafiltration and then washed with 1 m*M* calcium acetate. Finally, the material was lyophilized.

Although attempts were made to reproduce the crystallization of TAKA amylase from ammonium sulfate, as reported by Matsuura *et al.* (1979), they were unsuccessful. Instead, two crystal forms were grown using polyethylene glycol 8000 as precipitant. Monoclinic crystals were grown using the hanging-drop technique (McPherson, 1989); 50 μl volumes of protein solution (30 mg ml^{-1}) in 50 m*M* sodium acetate, pH = 6·0, 2 m*M* $CaCl_2$, 8%(*w*/*v*)

polyethylene glycol 8000, were equilibrated against a reservoir solution of 16% polyethylene glycol 8000 in the same buffer. The crystals took *ca* 14 days to grow, with one single crystal (2·0 × 0·3 × 0·3 mm) suitable for further X-ray studies (space group $P2_1$, with two molecules in the asymmetric unit, $a = 75{\cdot}0$, $b = 104{\cdot}3$, $c = 67{\cdot}4$ Å, $\beta = 104{\cdot}5°$). This form proved helpful during the molecular replacement studies of NOVO α-amylase (Brady *et al.*, 1981); further work is underway. The orthorhombic crystal form (space group $P2_12_12_1$, $a = 50{\cdot}9$, $b = 67{\cdot}2$, $c = 132{\cdot}7$ Å, one molecule per asymmetric unit) was obtained using the batch method. A protein solution of 33 mg ml^{-1} (buffered as in the previous case) and containing 17% polyethylene glycol 8000 was filtered and allowed to stand in small test tubes at room temperature for 10 days. Crystals were usually large (up to 5·0 × 0·3 × 0·2 mm). This crystal form was used for the X-ray studies described in this paper.

Data collection and processing

Crystallographic data were first collected on film using the Arndt–Wonacott oscillation camera, with synchrotron radiation at the Daresbury Laboratory, Warrington, England, on beam line 7. The crystal to film distance was 65 mm, giving a resolution of 2·1 Å at the edge of the film, with 1·488 Å radiation. The crystal was aligned in the first setting with b^* parallel to the rotation axis and a^* parallel to the beam, from which 64° of data were collected using an oscillation of 1° per exposure. A second crystal was rotated until c^* was orientated parallel to the beam. 20° of data were collected in this setting, with a 2·5° oscillation per exposure. Three films were used in each film pack. The films were scanned using a Joyce–Loebl Scandig-3 densitometer and the data integrated using the *MOSPRF* package (written by A. J. Wonacott, modified by F. Korber) incorporating profile-fitting. Oblique incidence, absorption and Lorentz–polarization corrections were applied and the data scaled. Details of this data set are given in Table 1. The overall $R_{\rm merge}(I)$ was 4·8% for 23 796 independent reflections.

Subsequent difficulties with the interpretation of molecular replacement studies have been traced to the problem of overexposure, which resulted in the saturation of the majority of low- and medium-angle intensities (see Fig. 2). It was therefore decided to complement the existing data set with medium-resolution counter data. To that end, a Xentronics (Siemens) area-detector system mounted on a conventional source was used, as described by Derewenda & Helliwell (1989) using a crystal to detector distance of 18 cm. Data were collected from a single crystal to a resolution of 4 Å and processed using the *XENGEN* suite of programs (Howard, Gilliland, Finzel, Poulos, Ohlendorf & Salemme, 1987). Further details are given in Table 2.

Table 1. *Photographic data details for native TAKA α-amylase crystals*

Total No. of packs	72
Total No. of observations input	100433
Unused partial halves	12038
Rejected observations	2878
Reflections	
Total No. used	66210
Fully recorded	48047
Partially recorded	18163
No. of independent reflections	23794
No. of reflections with negative intensities (after merging)	324
$\langle I/\sigma(I)\rangle$	9·46
Overall $R_{\rm merge}$ for all symmetry-related reflections (on intensities)	0·048
$R_{\rm merge}$ for weak intensities [$I < 5{\cdot}4\sigma(I)$]	0·043
$R_{\rm merge}$ for reflections beyond 2·2 Å	0·066
Percentage of reflections > 3σ	95·0

For the purposes of the final model refinement by least-squares methods, the two data sets were merged, using the programs *AGROVATA* and *ROTAVATA*, with the data scaled to the first batch of the film data set. An $R_{\rm merge}(I)$ for this data of 8·0% for 24 508 reflections was achieved (merged data set 1). Problems arose during scaling of these data, in the resolution range 4·0–4·2 Å, where the film data were often being included at the expense of the Xentronics data. During the final stages of refinement, the data were remerged. Each data set was scaled separately to structure factors calculated from the current model structure, using film data only in the range 2·0–4·0 Å and Xentronics data in the range 3·9–32·0 Å. The data sets, with an overlap now only in the resolution range 3·9–4·0 Å, were then merged, giving an $R_{\rm merge}(I)$ of 5·9% for 80 reflections (merged data set 2).

Molecular replacement

Two atomic models were available for molecular replacement calculations; the TAKA (entry 2TAA in

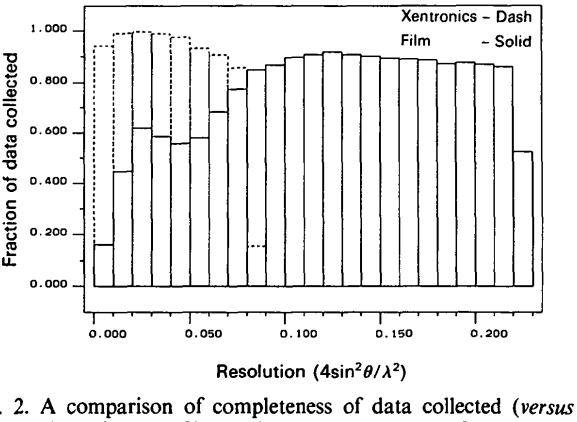

Fig. 2. A comparison of completeness of data collected (*versus* resolution) for the film and area-detector sets. The incompleteness of the film data at low resolution was found to be due almost exclusively to loss of high-intensity data caused by saturation.

Table 2. *Details of the Xentronics data collection*

Resolution limit (Å)	No. of reflections Possible	Collected	Average redundancy	No. of singly measured reflections (%)	$\langle I/\sigma(I)\rangle$	No. of reflections $< 2\sigma$ (%)	R_{merge} (%)
< 6·33	1143	1119	6·1	26 (2·3)	191·3	1100 (98·3)	3·30
< 5·02	1068	1067	3·9	54 (5·1)	97·7	1042 (97·7)	4·15
< 4·39	1041	1020	2·8	109 (10·7)	84·5	999 (97·9)	3·35
< 3·99	1030	971	2·7	100 (10·3)	71·3	951 (97·9)	3·42
< 3·70	1018	935	2·9	129 (13·8)	54·2	903 (96·6)	3·68
< 3·48	1033	794	2·3	166 (20·9)	39·7	734 (92·4)	5·74
Total	6333	5906	3·3	584 (9·9)	94·1	5729 (97·0)	3·66

the Protein Data Bank; Bernstein *et al.*, 1977) α-amylase structure and the more highly refined 2·1 Å acid NOVO amylase structure (Boel *et al.*, 1990), which has an 80% sequence homology with TAKA amylase. Both structures were tested to see which gave the better solution. All efforts to solve the rotation function using the film data set alone were unsuccessful, owing to the loss of the strong low-resolution reflections. Fig. 2 compares the completeness of the Xentronics and film data sets in resolution shells, and Fig. 3 shows the resulting rotation-function map for the β section which ought to reveal the true peak. The Xentronics/Siemens data were used for all further molecular replacement calculations.

Rotational search

Each model was placed in a *P*1 cell, with axial lengths $a = b = c = 80$ Å and $\alpha = \beta = \gamma = 90°$ and structure factors were calculated for these. The Patterson function of the search molecule was then rotated until the best match was achieved with the Patterson function of the unknown molecule, calculated from the observed structure factors (Crowther, 1972). An integration radius of 30 Å and resolution limits of 10–4 Å were used, with α and γ incremented in steps of 2·5° and β in steps of 5·0°. Clear solutions for both starting models were obtained (see Fig. 3).

Translational search

The translational parameters were determined by minimizing the *R* factor between the observed structure factors and those calculated from the properly rotated model translated on a three-dimensional grid (Nixon & North, 1976). The model molecule was orientated according to the results of the rotation function and positioned in a cell with *P*1 symmetry but proper axial lengths. A three-dimensional search is computationally intensive, therefore three two-dimensional searches were performed on 0·5 Å grids followed by a three-dimensional search on a 0·2 Å grid, around the minimum suggested by the two-dimensional searches. Data in the resolution range 10–4 Å were used throughout these calculations.

Unambiguous solutions for both models were obtained (Fig. 4), with the minimum R_{cryst} obtained for the acid amylase model with parameters $t_x = 24·2$, $t_y = 27·2$, $t_z = 20·6$ Å, $\alpha = 132·5$, $\beta = 50·0$, $\gamma = 187·5°$; and for the 3 Å TAKA amylase model using $t_x = 7·7$, $t_y = 10·6$, $t_z = 21·9$ Å, $\alpha = 125·0$, $\beta = 45·0$, $\gamma = -145·0°$.

The initial orientations of the molecules in the 3 Å TAKA amylase and the acid amylase models were different, and therefore required different solutions to the rotation and translation functions. As the *A. niger* α-amylase structure was extensively refined, it was chosen as the starting model for refinement rather than the TAKA amylase atomic model deposited in the Protein Data Bank.

Refinement of molecular replacement parameters

The molecular replacement parameters were refined using a protocol suggested by Derewenda (1989). Five cycles of KH refinement [the Konnert–Hendrickson (Hendrickson, 1985) version of geometrically restrained crystallographic least-squares refinement] were performed on the molecular replacement solution model, with the geometric terms weighted 10:8 relative to the X-ray terms. The initial TAKA model was then divided into domains *A* and C, to allow for limited movement of the domains relative to one another, with the sheets of each domain being refitted to the refined structure. This procedure gave a 1·9% reduction in the *R* factor of the model, reflecting a marginal improvement of the rigid-body fit.

Crystallographic refinement

Initial least-squares restrained refinement

Twelve cycles of KH refinement were interspersed with two sessions of model building, performed on an Evans & Sutherland PS300, using the program *FRODO* (Jones, 1978). Details of each round of refinement are given in Table 3. Electron density maps with coefficients $(2F_{obs} - F_{calc})$ and $(F_{obs} - F_{calc})$ maps were used to locate many of the side-chains not conserved between the stuctures, previously truncated back to their α-carbon. Sections of the polypeptide chain for which density was broken

or absent were removed from phase calculations, in order not to bias any subsequent density maps.

At this stage, it became clear both from the broken nature of sections of the density and the convergence of the R_{cryst} to 0·283, that there were serious errors in the model coordinates, necessitating extensive rebuilding of the atomic model. However, the electron density map was not clear enough for such purposes, and several attempts have failed. As it became clear that refinement was trapped in a local minimum, the *XPLOR* refinement package (Brunger *et al.*, 1987) incorporating simulated annealing, was used to overcome the problem.

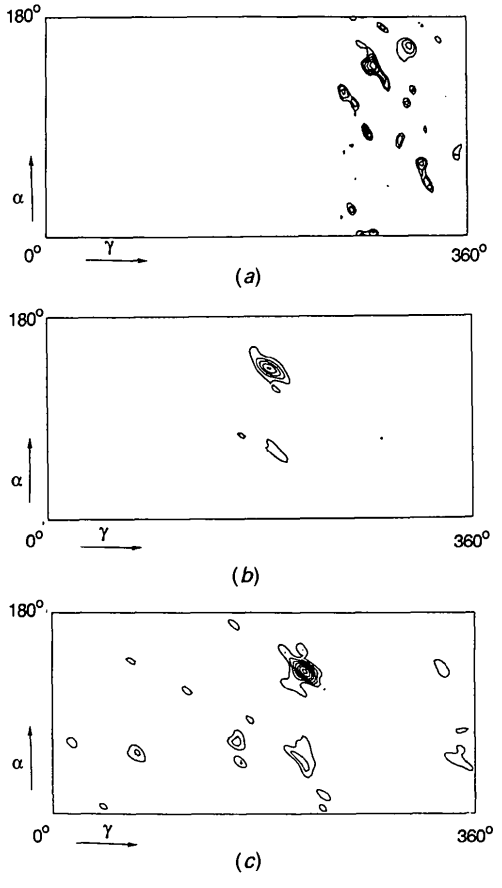

Fig. 3. Rotational searches using a step size of 2·5° along α and γ, and 5·0° along β (sections of constant β). (*a*) $\beta = 45°$ section; this calculation was carried out using film data and the *A. niger* acid α-amylase as the search model. The map has an r.m.s. deviation from the mean of 47·34 with the maximum of 1·7 times the r.m.s. deviation. The contouring is at equal intervals between 1·0 and 1·7 times the r.m.s. deviation. (*b*) $\beta = 50°$ section using the *A. niger* acid α-amylase as model structure. The map has an r.m.s. deviation from the mean of 35·0 and a maximum peak height of 307·4, contoured at equal intervals between 2·8 and 8·5 times the r.m.s. deviation. (*c*) $\beta = 45°$ section using 3 Å TAKA α-amylase as model structure. The map has an r.m.s. deviation from the mean of 53·4 and a maximum peak height of 420·8, contoured at equal intervals between 1·9 and 7·5 times the r.m.s. deviation.

Molecular dynamics

Conventional crystallographic least-squares refinement is very susceptible to local minima, and is not successful if the model atoms are outside the radius of convergence (usually less than 1 Å). Recent applications of simulated annealing (Kirkpatrick, Gellat & Vecchi, 1983) led to its introduction into crystallographic refinement by Brunger *et al.* (1987). In this method the shift which increases the minimized target function $[f(X) = f(x_1,x_2,...x_n)]$ is accepted with a probability of $\exp[-\Delta f(X)/k_bT]$, where $\Delta f(X)$ is the difference in the values of the target before and after trial, and k_b is the Boltzmann constant. The parameter T determines the amount of kinetic energy available to overcome energy barriers so that a large number of conformations can be sampled. The success of the method depends on the annealing schedule, or in effect, on the way the temperature is modified during the simulation. Molecular dynamics (Karplus & McCammon, 1983) can be successfully used to introduce a 'temperature' parameter, in order to escape from local minima. During the dynamics run, atoms fluctuate around their equilibrium positions, resulting in a non-zero gradient. From the

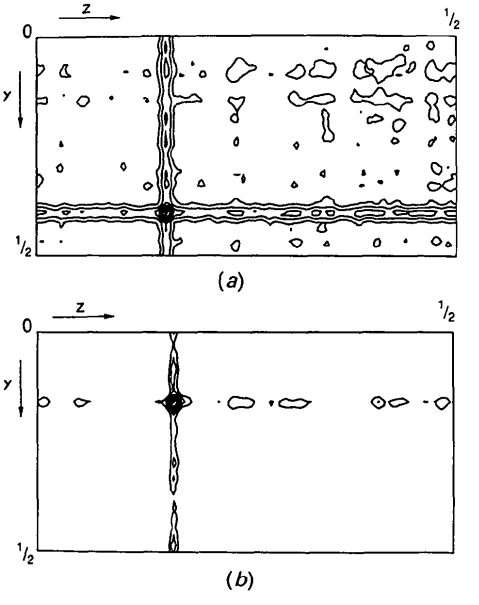

Fig. 4. Translational searches with a step size of 0·5 Å along each axis; in place of the usual R factor, these maps show the negative deviation from the mean; this means that the correct translation has the maximum value of the function. (*a*) $x =$ 24·5 Å section, using the rotated *A. niger* acid α-amylase model. The map has a maximum peak height of 0·45 and an r.m.s. deviation from the mean of 0·064 and was contoured at intervals of between 3·1 and 6·9 times the r.m.s. (*b*) $x = 8·0$ Å section, using the rotated TAKA α-amylase model. The map has a maximum peak height of 0·32 and an r.m.s. deviation from the mean of 0·015 and was contoured at equal intervals between 13·2 and 22·6 times the r.m.s.

Table 3. *Details of the course of the crystallographic refinement*

xyz denotes restrained refinement of positional parameters, *xyzB* refers to simultaneous restrained refinement of positional and isotropic parameters.

No. of cycles	*R* factor	Data set refined against	Matrix	No. of protein atoms	No. of water molecules
5 (*xyz*)	35·10	Merged data set 1	0·8	3304	0
7 (3*xyz*,4*xyzB*)	28·37	Merged data set 1	1·0	3545	0
Molecular dynamics					
5 (*xyz*,4*xyzB*)	24·90	Merged data set 1	1·0	3576	0
7 (2*xyz*,5*xyzB*)	21·60	Merged data set 1	1·0	3616	78
7 (2*xyz*,5*xyzB*)	19·70	Merged data set 1	1·0	3623	236
8 (2*xyz*,6*xyzB*)	19·71	Merged data set 2	0·5	3660	236
8 (2*xyz*,6*xyzB*)	19·33	Merged data set 2	0·5	3660	278
8 (2*xyz*,6*xyzB*)	18·25	Merged data set 2	0·5	3675	247

gradient of the energy function the weight (which relates the 'X-ray energy' term to the empirical energy term) can be calculated.

Initial minimization of the TAKA model was ill-conditioned due to several close non-bonded contacts which resulted in an unrealistically large gradient for the empirical energy. To overcome this, the van der Waals and electrostatic interactions were replaced by a 'repel' potential which is purely repulsive. Forty cycles of minimization were performed using this potential, followed by 100 cycles of full minimization in which the van der Waals and electrostatic potentials were included. A 0·1 ps dynamics simulation was then performed.

Minimization was now performed with the 'X-ray energies' included with the appropriate weighting. 40 steps of minimization were carried out in which α-carbon atoms were restrained to their initial positions, to prevent large movements of backbone atoms, caused by bad contacts, followed by 100 steps of unrestrained minimization.

For the heat stage of the protocol, velocities were assigned to atoms according to a Maxwellian distribution at 300 K. A molecular-dynamics simulation was performed for 1 ps, with rescaling of the velocities every 25 fs to maintain a constant temperature. A maximum *R* factor of 36·8% was observed. A rapid cooling protocol was then employed by assigning velocities at 300 K and equilibrating in a molecular-dynamics run for 0·5 ps with rescaling every 50 fs. Finally, the system was minimized for 25 cycles. The *R* factors at each stage of the protocol are given Fig. 5.

Final least-squares refinement

The structure output from *XPLOR* was subjected to five cycles of KH refinement, which reduced the *R* factor to 0·249 and from which a $(2F_{obs} - F_{calc})$ map of good quality was calculated. This map allowed modelling of 16 hitherto excluded side chains and the inclusion of 78 solvent molecules.

Until this time, work had proceeded using the published amino-acid sequence by Toda *et al.* (1982). Comparison of this sequence with that of the acid amylase showed a deletion in the TAKA amylase sequence, in the region connnecting the $(\alpha/\beta)_8$ barrel with the β-sandwich. When the density was examined in this region there was clearly no deletion and this was later confirmed when the nucleotide sequence became available (Tada *et al.*, 1989).

Forty-four further cycles of KH refinement were used, with five model building sessions, allowing the inclusion of 247 water molecules and the location of two Ca^{2+} ions. The water molecules were included only if chemically reasonable hydrogen bonding existed to polar protein atoms or other water molecules. On the first five cycles of refinement the X-ray terms and geometric terms were weighted in the ratio of 8:10 (matrix 0·8) with subsequent cycles weighted in the ratio of 1:1 (matrix 1·0). On the last 25 cycles

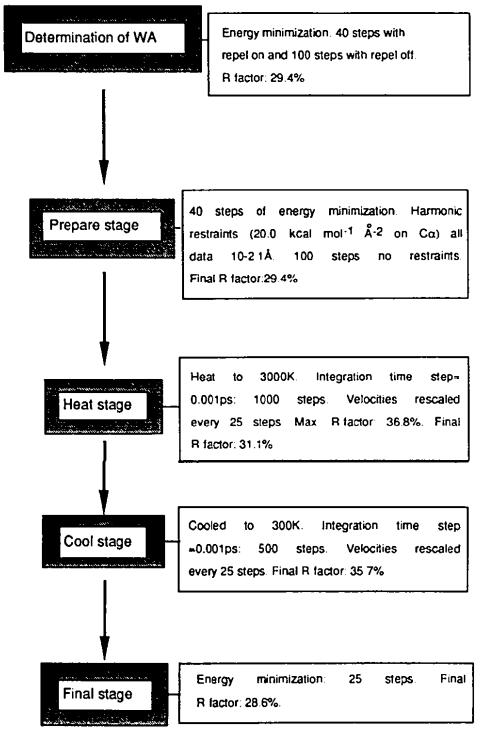

Fig. 5. The parameters used at each stage of the molecular-dynamics refinement

of refinement, a weight of 1:2 was used (matrix 0·5), to prevent unrealistic movements of the structure. A re-evaluation of the water structure was performed prior to the final nine cycles of refinement. On each cycle, 10–2·1 Å data were used for refinement, however, a 7·5 Å cutoff was used for the last nine cycles, to exclude data for which unaccounted solvent makes a significant contribution to the structure factors.

Results and discussion

Crystallographic refinement with simulated annealing was successfully used to refine the atomic model of the TAKA amylase. The technique was not used for all of the refinement calculations, but only to overcome the critical energy barrier associated with the movement of the small C-terminal domain. It was preceded, and followed by standard restrained crystallographic least-squares refinement (Hendrickson, 1985). In this section we would like to illustrate the nature and the magnitude of the observed correct shifts, as well as some of the difficulties.

Fig. 6(*a*) shows that there have been movements of up to 5 Å between the model before and after *XPLOR* with two distinct regions of the molecule where main-chain movements were greater than 1 Å. The first contains residues 279–280, which are located in a loop between helix 6 and strand 7, and is likely to be flexible. The second consists of the entire C-terminal domain (Fig. 6*b*). A comparison of the model before and after *XPLOR* shows that this is caused by a rigid-body movement of the domain with respect to domain *A*. The rigid-body parameters are a rotation of 5·4° and translations parallel to the unit-cell axes of 0·94, 0·25 and 0·42 Å. This is a large difference, and most of the atoms in this domain were well outside the radius of convergence of classical least-squares methods. Before the application of *XPLOR*, the density for the regions requiring correction was poor. Examination of the molecular model obtained with the help of *XPLOR* against the original electron density map showed that there were no gradients which could have been followed during manual rebuilding. The most dramatic illustrations of this are shown in Figs. 7 and 8.

A clear limitation of the simulated-annealing method stems from the fact that those parts of the structure unaccounted for in the atomic model will give rise to density which will cause incorrect or excessive shifts. In particular, well-ordered water molecules tend to cause problems for adjacent residues. Fig. 9 shows how a peptide plane was forced to flip, so that the carbonyl oxygen of Gly 415 was moved into previously unaccounted for solvent density. However, the quality of the electron density map calculated after the application of *XPLOR* was sufficiently clear to allow final manual intervention and subsequent least-squares refinement.

The structure has been solved to a resolution of 2·1 Å, with an *R* factor of 0·183, including all data between 10–2·1 Å. Table 4 shows the standard deviation of the stereochemical restraints obtained after the final cycle of refinement. The refined structure contains 247 solvent molecules. Of these, none are internal to the protein, 208 are in the first-layer solvent shell, forming at least one hydrogen bond to either a main- or side-chain atom with the remainder forming no hydrogen bonds to protein atoms.

In the final $2F_{obs} - F_{calc}$ electron density map the side chains of residues Ser 36, Ser 130, Val 138, Val 234 and Lys 433 are not observable. Residue 449 appears to have density indicative of a side chain; however, Gly is indicated from both reported sequences. Of the ten amino substitutions found on comparison of the sequence reported by Toda *et al.* (1982) and that by Tada *et al.* (1989), nine were detected from the calculated electron density map. Particularly clear was that of residue 80, which as a

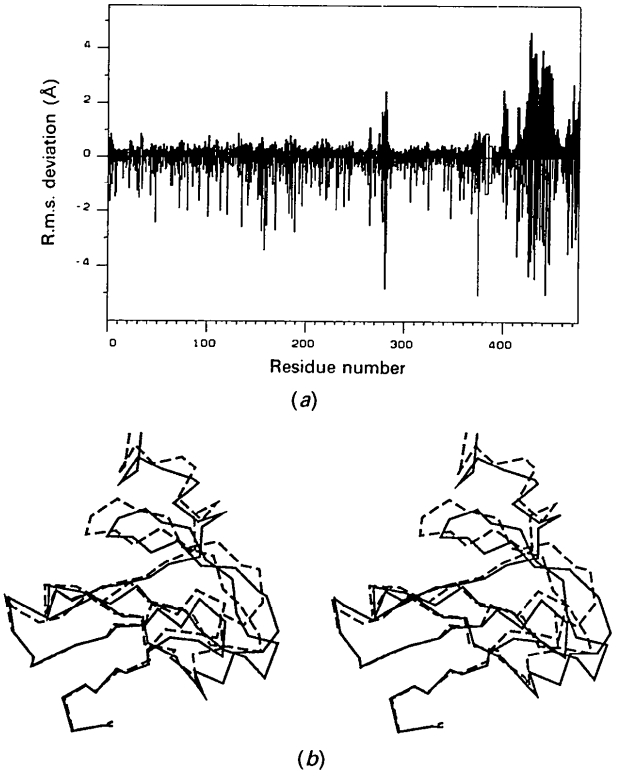

Fig. 6. (*a*) Atomic shifts in the model before and after the use of *XPLOR*. Main-chain shifts (averaged within each residue) are shown above the zero line and side-chain displacements (also averaged per residue) are shown below. (*b*) A schematic (Cα only) stereo representation of the movement of the C-terminal domain observed during the *XPLOR* refinement. The solid lines represent the model before *XPLOR*, while the dashed lines show the refined structure.

threonine residue could not be accommodated in the appropriate density. However, the histidine residue reported by the nucleotide sequencing studies (Tada *et al.*, 1989), which became available during refinement, fits this density well. The only substitution which was not observed from electron density is that of residue 174, found here to be a proline and not a leucine (see Fig. 10). The codon found by Tada *et al.* (1989) for residue 174 is CTT; as CCT codes for Pro, it is quite posible that a mistake occurred in the *c*DNA sequence. Examination of the density in the region of residue 385 also confirmed the insertion of a Trp residue at this position. The two C-terminal residues, Ser 477 and Ser 478, are not seen, indicating that they are disordered. Residues 99–105 and 181–186, both forming parts of helices, and 447–450, contained in an exposed loop between two strands of the C-terminal domain, constitute less well defined regions of the electron density map. These contain breaks in both main- and side-chain density and it is unclear how improvements can be made.

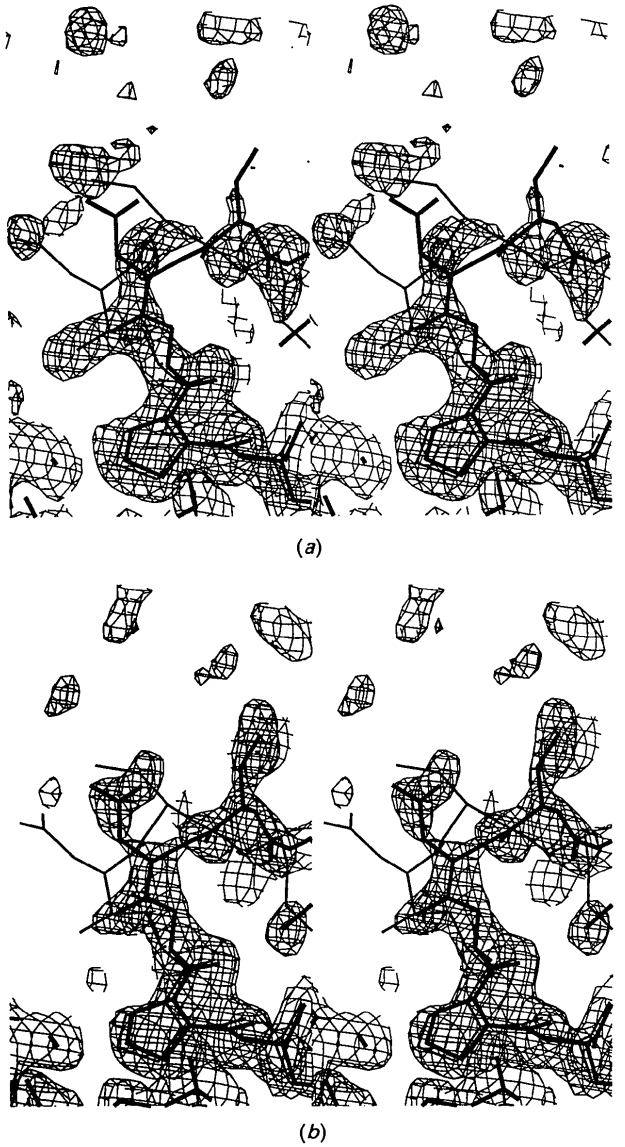

Fig. 7. The loop region (residues 279 to 281) illustrating the shifts obtained by the use of *XPLOR*; thin type denotes the input model prior to *XPLOR*, and bold type denotes the final structure. (*a*) The atomic models against an electron density map ($2F_{obs} - F_{calc}$) phased on the model prior to *XPLOR*. (*b*) The same, but seen against a electron density map ($2F_{obs} - F_{calc}$) phased on the model output by *XPLOR*.

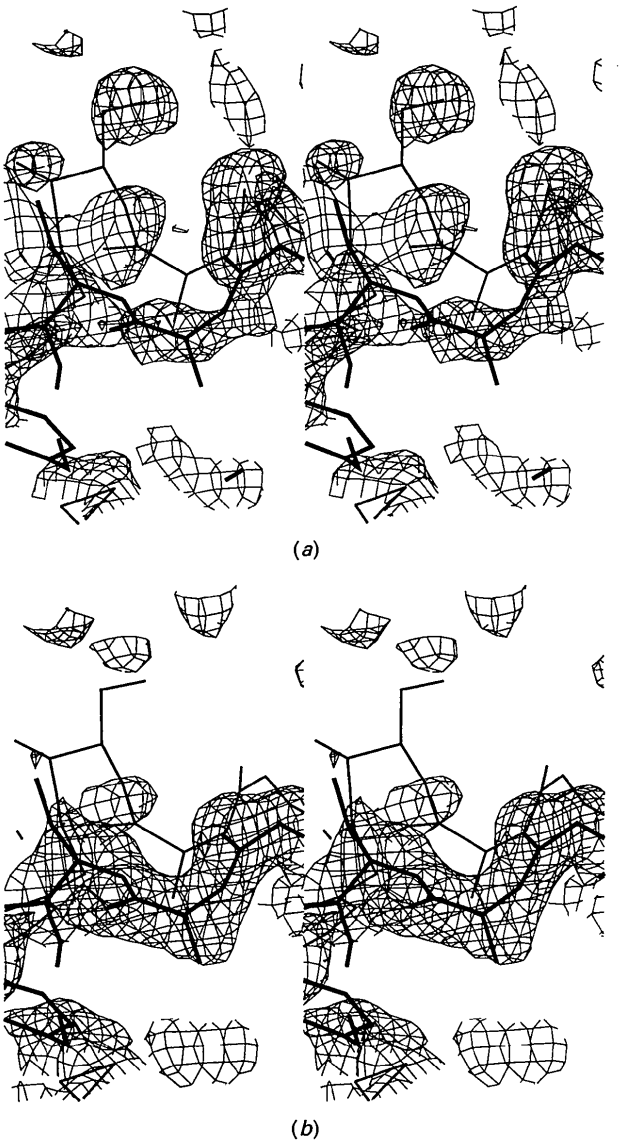

Fig. 8. Residues 424 to 427, before and after *XPLOR*. The details are the same as for Fig. 7, with (*a*) showing the map prior to *XPLOR*, and (*b*) after *XPLOR*.

The Ramachandran plot for the refined structure is shown in Fig. 11. The sterically allowed region for all amino acids, except glycine, is indicated by the solid lines. The only residues lying significantly outside these regions are glycine residues, indicated by circles, and Ser 447, which is contained in one of the less well defined regions of the map, as described above.

The details of the molecular model and the comparison of the two fungal α-amylases are in preparation and will be published elsewhere. The structure factors, molecular replacement phase angles and a full set of atomic coordinates have been deposited with the Protein Data Bank* (Bernstein *et al.*, 1987) and will be made generally available within 12 months of this publication.

Unless otherwise specified, all the programs referred to in this paper have been used in their VAX VMS versions of the *CCP*4 suite of crystallographic programs (SERC Daresbury Laboratory, 1986). *XPLOR* Version 1.5 as obtained from A. T. Brunger was used on a CRAY X-MP/416 supercomputer located in the SERC Rutherford Appleton Laboratory.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 3TAA, R3TAASF), and are available in machine-readable form from the Protein Data Bank at Brookhaven. The data have also been deposited with the British Document Supply Centre as Supplementary Publication No. SUP 37042 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. At the request of the authors, the list of structure factors will remain privileged until 1 August 1992.

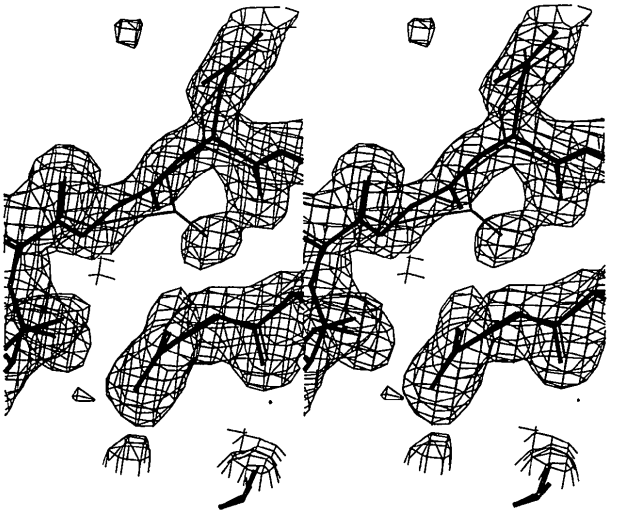

Fig. 9. Position of Gly 415 after *XPLOR* (in thin type), with the manually corrected position in bold type. The map ($2F_{obs} - F_{calc}$) is phased on the model output by *XPLOR*. Note the peak representing a water molecule, hydrogen bonded to carbonyl of 417, into which the carbonyl oxygen of Gly 416 was forced by *XPLOR*.

Table 4. *Stereochemical parameters of the refined model*

	σ	Standard deviation	No. of parameters
Bond lengths (1–2) (Å)	0·02	0·028	3772
Angle distances (1–3) (Å)	0·04	0·077	5158
Planar distances (1–4) (Å)	0·06	0·094	1441
R.m.s. deviation from planes (Å)	0·02	0·023	647
Chiral volume (Å)	0·12	0·208	563
Van der Waals contacts (Å)			
Single-torsion contact	0·5	0·234	1360
Multiple contact	0·5	0·305	1878
Possible contact	0·5	0·248	316
Torsion angles (°)			
Peptide (ω)	20·0	5·9	485
Staggered (60/120°)	20·0	23·0	563
Orthonormal (±90°)	20·0	36·0	66
Isotropic temperature factors, B (Å^2)			
Main chain (1–2)	1·0	3·36	1991
(1–3)	1·5	4·32	2532
Side chain (1–2)	1·5	5·65	1781
(1–3)	2·0	7·15	2626

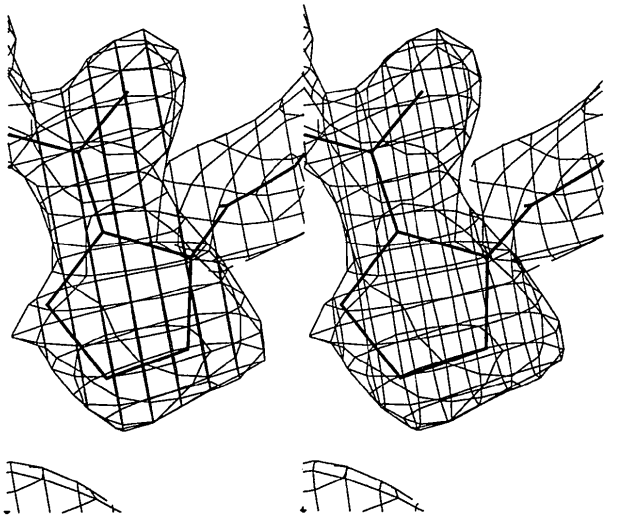

Fig. 10. Final electron density ($2F_{obs} - F_{calc}$) for Pro 174.

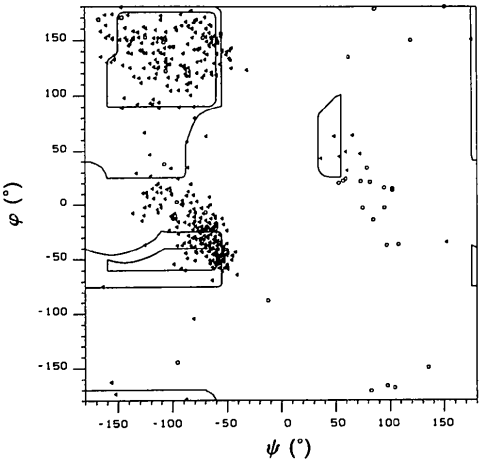

Fig. 11. Ramachandran plot (final model); open circles signify glycines (see text for further details).

The amylase project is supported by a grant from Novo–Nordisk Industri A/S, Copenhagen. The macromolecular structural studies in York are financed by a consolidated grant from the Science and Engineering Research Council, England, who are also acknowledged for the provision of CRAY time. We are grateful to the staff of the SERC Daresbury laboratory for their assistance with synchrotron data collection. We thank Dr Villy Jensen (Novo–Nordisk Research Laboratories) for supplying us with the purified protein samples.

References

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

BOEL, E., BRADY, L., BRZOZOWSKI, A. M., DEREWENDA, Z., DODSON, G. G., JENSEN, V. J., PETERSEN, S. B., SWIFT, H., THIM, L. & WOLDIKE, H. F. (1990). *Biochemistry*, **29**, 6244–6249.

BRADY, R. L., BRZOZOWSKI, A. M., DEREWENDA, Z., DODSON, E. J. & DODSON, G. G. (1991). *Acta Cryst.* B**47**, 527–535.

BRUNGER, A. T. (1988). *XPLOR Manual.* Yale Univ., USA.

BRUNGER, A. T., KARPLUS, M. & PETSKO, G. A. (1989). *Acta Cryst.* A**45**, 50–61.

BRUNGER, A. T., KURIYAN, K. & KARPLUS, M. (1987). *Science*, **235**, 458–460.

BUISSON, G., DUEE, E., HASER, R. & PAYAN, F. (1987). *EMBO J.* **6**, 3908–3916.

CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173–178. New York: Gordon & Breach.

DEREWENDA, Z. S. (1989). *Acta Cryst.* A**45**, 227–234.

DEREWENDA, Z. S. & HELLIWELL, J. R. (1989). *J. Appl. Cryst.* **22**, 123–137.

FARBER, G. K. & PETSKO, G. A. (1990). *TIBS*, **15**, 228–234.

FUJINAGA, M., GROS, P. & VAN GUNSTEREN, W. F. (1989). *J. Appl. Cryst.* **22**, 1–8.

GOODFELLOW, J., HENRICK, K. & HUBBARD, R., (1989). Editors. *Molecular Simulation and Protein Crystallography, Proceedings of the Joint CCP4/CCP5 Study Weekend.* Warrington: SERC Daresbury Laboratory.

HENDRICKSON, W. A. (1985). *Methods Enzymol.* **115**, 252–270.

HOWARD, A. J., GILLILAND, G. L., FINZEL, B. C., POULOS, T. L., OHLENDORF, D. H. & SALEMME, F. (1987). *J. Appl. Cryst.* **20**, 383–387.

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

KARPLUS, M. & MCCAMMON, J. A. (1983). *Annu. Rev. Biochem.* **52**, 263–300.

KIRKPATRICK, S., GELLAT, C. D. JR & VECCHI, M. P. (1983). *Science*, **220**, 671–680.

MCPHERSON, A. (1989). *Preparation and Analysis of Protein Crystals*, p. 96. Florida: Robert E. Krieger.

MATSUURA, Y., KUNUSOKI, M., DATE, W., HARADA, S., BANDO, S., TANAKA, N. & KAKUDO, M. (1979). *J. Biochem.* **86**, 1773–1783.

MATSUURA, Y., KUNUSOKI, M., HARADA, W. & KAKUDO, M. (1984). *J. Biochem.* **95**, 697–702.

NIXON, P. E. & NORTH, A. C. T. (1976). *Acta Cryst.* A**32**, 320–325.

PHILLIPS, D. C., STERNBERG, M. J. E., THORNTON, J. M. & WILSON, I. A. (1978). *J. Mol. Biol.* **119**, 329–351.

POSTMA, J. P. M., PARKER, M. W. & TSEGNOGLOU, D. (1989). *Acta Cryst.* A**45**, 471–477.

SERC Daresbury Laboratory (1986). *CCP4. A Suite of Programs for Protein Crystallography.* SERC Daresbury Laboratory, Warrington, England.

TADA, S., IMURA, Y., GOMI, K., TAKAAHASKI, K., HARA, S. & YOSHIZAWA, K. (1989). *Agric. Biol. Chem.* **53**, 593–599.

TODA, H., KONDA, K. & NARITA, K. (1982). *Proc. Jpn Acad. Ser. B*, **58**, 208–212.

VALLEE, B. L., STEIN, E. A., SUMERWELL, W. N. & FISHER, E. H. (1959). *J. Biol. Chem.* **34**, 2901–2929.

Acta Cryst. (1991). B**47**, 544–548

Linear Si—O—Si Fragments: Structure of 1,1,3,3-Tetramethyl-1,3-bis[(2-oxo-1-pyrrolidinyl)methyl]disiloxane Dihydrochloride and Analysis of Displacement Parameters

BY VALERY E. SHKLOVER*

Institute of Inorganic Chemistry, University of Zürich–Irchel, Winterthurerstrasse 190, *CH*-8057 *Zürich, Switzerland*

AND HANS-BEAT BÜRGI,* ANDREA RASELLI, THOMAS ARMBRUSTER AND WOLFGANG HUMMEL

Laboratory for Chemical and Mineralogical Crystallography, University of Bern, Freiestrasse 3, *CH*-3012 *Bern, Switzerland*

(*Received* 2 *September* 1990; *accepted* 26 *February* 1991)

Abstract

An X-ray diffraction study of 1,1,3,3-tetramethyl-1,3-bis[(2-oxo-1-pyrrolidinyl)methyl]disiloxane dihydrochloride, $C_{14}H_{30}N_2O_3Si_2^{2+}.2Cl^-$, $M_r = 357{\cdot}1$, has been carried out at 120 K. The crystal is monoclinic, space group $P2_1/n$, $Z = 2$; the cell dimensions at 120 K are $a = 11{\cdot}116\ (3)$, $b = 8{\cdot}176\ (3)$, $c = 12{\cdot}138\ (6)$ Å, $\beta = 114{\cdot}34\ (3)°$, $V = 1005{\cdot}1$ Å^3.

* To whom correspondence should be addressed.

0108-7681/91/040544-05$03.00