Acta Cryst. (1991). **B47**, 918–927

Structure Solution and Molecular Dynamics Refinement of the Yeast Cu,Zn Enzyme Superoxide Dismutase

BY K. DJINOVIĆ, G. GATTI AND A. CODA

Dipartimento di Genetica e Microbiologia and Centro Interuniversitario Studio Macromolecole Informazionali, Università di Pavia, Via Abbiategrasso 207, *I*-27100 *Pavia, Italy*

L. ANTOLINI

Dipartimento di Chimica, Università di Modena, Via Giuseppe Campi 183, *I*-41100 *Modena, Italy*

G. PELOSI

Istituto di Chimica Generale ed Inorganica, Università di Parma, Viale delle Scienze, I-43100 *Parma, Italy*

A. DESIDERI

Dipartimento di Chimica Organica e Biologica, Università di Messina, 98166 *Messina, Italy*

M. FALCONI

Dipartimento di Biologia, Università di Roma 'Tor Vergata', Via Orazio Raimondo, 00173 *Roma, Italy*

F. MARMOCCHI

Dipartimento di Biologia Cellulare, Università di Camerino, Via Scalziono 5, *I*-62032 *Camerino, Italy*

G. ROTILIO

Dipartimento di Biologia, Università di Roma 'Tor Vergata', Via Orazio Raimondo, 00173 *Roma, Italy*

AND M. BOLOGNESI*

Dipartimento di Genetica e Microbiologia and Centro Interuniversitario Studio Macromolecole Informazionali, Università di Pavia, Via Abbiategrasso 207, *I*-27100 *Pavia, Italy*

(*Received* 17 *January* 1991; *accepted* 23 *April* 1991)

Abstract

Cu,Zn yeast superoxide dismutase was crystallized from polyethylene glycol solutions. The crystals belong to the $P2_12_12$ space group, with cell dimensions $a = 105{\cdot}3$, $b = 143{\cdot}0$, $c = 62{\cdot}1$ Å; two dimers of $M_r = 32000$ each are contained in the asymmetric unit. Diffraction data at 2·5 Å resolution were collected with the image-plate system at the EMBL synchrotron radiation facility in Hamburg. The structure was determined by molecular replacement using as a search model the 'blue–green' dimer of the bovine Cu,Zn superoxide dismutase. The crystallographic refinement of the molecular replacement solution was performed by means of molecular dynamics techniques and resulted in an R factor of 0·268 for the data between 6·0 and 2·5 Å. The model was subsequently subjected to conventional restrained crystallographic refinement of the coordinates and temperature factors. The current R value for the data between 6·0 and 2·5 Å is 0·220. Owing to the large radius of convergence of the molecular dynamics–crystallographic refinement, the convergence of the refinement process was reached after 18·1 ps of simulation time. The geometry of the active site of the enzyme appears essentially preserved compared with the bovine superoxide dismutase. The β-barrel structure in the yeast enzyme is closed at the upper part by an efficient hydrogen-bonding scheme.

Introduction

Superoxide dismutase enzymes are metalloproteins containing either copper and zinc, a manganese or iron as natural prosthetic groups, which catalyze the disproportionation of the superoxide radical anion into molecular oxygen and hydrogen peroxide. The distribution of the three types of superoxide

* To whom correspondence should be addressed.

0108-7681/91/060918-10$03.00

dismutase (SOD) in organisms is considered to be characteristic of the evolutionary stage of the organism and of the associated cell organelles. Structural studies have shown that the iron and manganese types are closely related but bear little resemblance to the Cu,Zn SOD (Stallings, Pattridge, Strong & Ludwig, 1984). The iron SOD type has usually been isolated as a dimer of identical 20000 M_r subunits from procaryotes and anaerobic bacteria. The manganese type has been found as an oligomer of identical 20000 M_r subunits in procaryotes and mitochondria of eucaryotic cells, while the Cu,Zn enzyme is predominantly found in the cytoplasm of all higher eucaryotic cells (Stallings, Pattridge, Strong, Ludwig, Yamakura, Isobe & Steinman, 1987; Parker & Blake, 1988). The structure–function relationships in the Cu,Zn enzymes have been thoroughly investigated on the bovine Cu,Zn SOD (Bannister, Bannister & Rotilio, 1987), by high-resolution crystallographic analysis (Tainer, Getzoff, Beem, Richardson & Richardson, 1982) and by extensive spectroscopic studies (Bannister *et al.*, 1987). The isolated protein contains Cu^{2+} and Zn^{2+} ions at the active site, where the catalytically active Cu^{2+} ion is cyclically reduced and oxidized during successive encounters with the superoxide substrate. In the first step an electron from one superoxide radical is donated to the catalytic center with the formation of molecular oxygen and of a Cu^{+} ion, which in turn donates one electron to a second superoxide radical to produce, together with two protons, hydrogen peroxide (Fielden, Roberts, Bray, Lowe, Mautner, Rotilio & Calabrese, 1974). The kinetics of the process are dominated by the diffusion of the negatively charged superoxide radical anion into the active-site channel. The process is controlled by an evolutionary constant distribution of electrostatic charges around the active site (Desideri, Falconi, Parisi & Rotilio, 1989). The amino-acid sequences of Cu,Zn SOD's (Bannister *et al.*, 1987, Getzoff, Tainer, Stempien, Bell & Hallewell, 1989) show that the enzyme has been highly preserved during evolution. Among the eucaryotic systems, yeast SOD (y-SOD) has the highest sequence divergence from the bovine enzyme (55% amino-acid identities) (Bannister *et al.*, 1987). Moreover y-SOD shows lower resistance to urea (Barra, Bossa, Marmocchi, Martini, Rigo & Rotilio, 1979), and a transition temperature for reversible thermal unfolding which is 16 K lower than that reported for the irreversible thermal denaturation of the bovine enzyme (Arnold & Lepock, 1982). The functional y-SOD molecule consists of a dimer of identical subunits, 153 amino acids each (16000 M_r); one Cu^{2+},Zn^{2+} pair is associated with each subunit.

Only one three-dimensional structure of eucaryotic Cu,Zn SOD has so far been reported, for the bovine erythrocyte Cu,Zn enzyme (Tainer *et al.*, 1982). This structural investigation has shown that each subunit has as its structural core a flattened antiparallel eight-stranded β-barrel, plus three external loops. In this context we undertook a crystallographic study of a new and phylogenetically distant SOD in order to elucidate the development of the enzyme tertiary structure and its correlation with the electrostatic and functional properties during evolution.

Materials and methods

Crystallization and data collection

The crystals of y-SOD were grown by vapor diffusion techniques using the 'sitting drop' experimental setup (MacPherson, 1982). Solutions of y-SOD, at a concentration of 15 mg ml^{-1}, in 0·025 *M* citrate buffer, 0·01 *M* phosphate buffer containing 6% (*w*/*v*) polyethylene glycol (PEG, average $M_r = 4000$), at pH 6·5, were equilibrated against a reservoir solution containing 12% (*w*/*v*) PEG at the same pH, at 301 K. The crystals grew to the size of at least 0·7 mm in each dimension. The preliminary characterization of the crystal form obtained under these conditions has been reported elsewhere (Frigerio, Falconi, Gatti, Bolognesi, Desideri, Marmocchi & Rotilio, 1989). The crystals belong to the orthorhombic space group $P2_12_12$ with cell constants 105·3, 143·0, 62·1 Å; assuming that the asymmetric unit contains four SOD molecules, the solvent content of this crystalline form is 64% (*w*/*v*) (Matthews, 1968). A single crystal 0·2 × 0·2 × 0·15 mm was mounted in a thin-walled glass capillary and used for the data collection using the synchrotron radiation source ($\lambda = 1{\cdot}009$ Å) at the EMBL X11 beamline of the DORIS storage ring, DESY, Hamburg, employing a locally developed image-plate system as detector. The DORIS ring was operating in a single-bunch mode at 5·3 GeV and 40–100 mA with approximately 3 h between injections. The crystal was mounted on an Arndt–Wonacott oscillation camera (Nyborg & Wonacott, 1977) with the [110] direction parallel to the rotation axis, and was constantly kept at 277 K by a stream of dry cooled air. The crystal-to-detector distances were 370 and 250·5 mm for low- and high-resolution sets respectively, and exposure times ranged from 14 to 121 s; data were collected to 2·5 Å spacings, using 3 and 1·5° rotation angular ranges for the data recorded up to 3·5 Å resolution, and rotation ranges of 2·0 and 1·0 Å for the high-resolution data.

Refinement of orientation and integration of the intensities was performed using the *MOSFLM* suite of programs (Leslie, Brick & Wonacott, 1986) modified for processing of the image-plate data. Merging of the observed intensities into a unique

data set of reflections was carried out using the *ROTAVATA/AGROVATA* programs from the *CCP*4 suite, supplied by the SERC Daresbury Laboratory, England. The intensities were converted to structure-factor amplitudes using the program *TRUNCATE* (French & Wilson, 1978). 131852 measurements were merged to yield 32538 unique reflections (98·2% of the theoretical reflections in the 10·0 to 2·5 Å resolution range). The merging R_m factor ($R_m = \sum_i I_i - \langle I \rangle / \sum |I_i|$, where I_i is the intensity of an observation, $\langle I \rangle$ is the mean value of the reflection, and the summations are over all reflections) was 6·0%.

Molecular replacement

The starting model for the molecular replacement investigation was the 'blue–green' dimer of the bovine Cu,Zn superoxide dismutase, partially refined at 2·0 Å resolution, to an R factor of 25·5% (Tainer *et al.*, 1982), and deposited with the Protein Data Bank (Bernstein, Koetzle, Williams, Meyer, Brice, Rogers, Kennard, Shimanouchi & Tasumi, 1977) as data set 2SOD. All the side chains in y-SOD not common to the bovine SOD amino-acid sequence were modified to alanines. The search model was oriented as the 'orange–yellow' dimer of the Cu,Zu bovine SOD, and additionally rotated by 30° on the Eulerian angle β. The *MERLOT* (Fitzgerald, 1988) package of programs which includes the fast-rotation function of Crowther (1972), the rotation function of Lattman & Love (1970) and a translation function of Crowther & Blow (1967), was used throughout.

The asymmetric unit of y-SOD consists of two dimers related by a very approximate diad axis. Additionally two local twofold axes relate each pair of monomers that form the dimers. The search model superimposes twice with each of the two dimers in the asymmetric unit of the crystal structure. A rotation function investigation with a dimer as search model should therefore exhibit a maximum of four peaks related by pseudo-twofold symmetry.

For the calculation of the structure factors used to generate harmonic coefficients for the fast-rotation function the search model was placed in a triclinic reference cell with dimensions of 100 × 100 × 100 Å. The structure factors were calculated up to 3·0 Å spacings with an overall temperature factor at 0·0 Å^2 in order to sharpen the peaks. The harmonic coefficients were calculated with upper and lower limits of 4·5 and 10·0 Å respectively, the integration radius being set to 26·9 Å. The Eulerian angle increments between successive calculations of the rotation function were 2·5° for α, 1° for β and 5° for γ. Table 1 summarizes the set of solutions obtained from the cross-rotation function analysis performed.

Table 1. *Results of the fast-rotation function in Euler angles, after origin removal*

R.m.s. is the root-mean-square deviation from the mean.

Peak rank	α (°)	β (°)	γ (°)	R.m.s.	% max.
1	80·0	83·0	270·0	6·33	100·0
2	32·5	81·0	270·0	5·83	92·1
3	60·0	60·0	290·0	5·06	79·9
7	130·0	72·0	80·0	4·78	75·5

Table 2. *Results of the Lattmann–Love cross-rotation function in Euler angles and symmetry relations between the peaks in polar angles*

Peak number	α (°)	β (°)	γ (°)
1	80·0	84·0	269·0
2	32·0	80·0	273·0
3	58·0	61·0	293·0
7	133·0	71·0	78·0
Peak pairs	φ (°)	ψ (°)	κ (°)
1^{I} 3^{II}	189·0	84·0	152·0
2^{I} 7	171·0	83·0	208·0
2^{I} 3	270·0	15·0	147·0
1^{I}–7^{II}	90·0	15·0	213·0
1^{I} 2^{III}	116·0	87·0	180·0
3 7	97·0	81·0	180·0

The symmetry-equivalent Eulerian angles in $P2_12_12$ space group are coded as follows: (I) 180° + α, β, γ; (II) 180° − α, 180° − β, 180° + γ; (III) 360° − α, 180° − β, 180° + γ.

The four correct orientational peaks were selected on the basis of the rotational relationships: the pairs of peaks corresponding to one dimer are related by a pseudo-twofold axis ($\kappa = 180°$) while pairs relating the two dimers in the asymmetric unit are correlated by the rotations listed in Table 2. As the success in determining the correct translation of the oriented model is highly dependent on how accurate the rotation angles are, the rotation function of Lattman & Love (1970) was applied in order to 'refine' the determined rotation angles on a finer grid (from 11 to 13 steps of 1° around each angle). The 'refined' values of the Eulerian angles, along with the twofold symmetry relations between peaks expressed in polar angles, are reported in Table 2.

In order to position the dimers with respect to the crystallographic symmetry in only one translation search calculation, the following pairs of angles correlated by pseudo-twofold axis, were used as an input to the translation function: 1^{I}–3^{II}, 1^{I}–7^{II}, 2^{I}–7, 2^{I}–3 (see Table 2 for the explanation of symbols I, II). The correctness of the solutions was assessed on the basis of an internally consistent set of vectors from the three Harker sections in the translation function, and of packing considerations. Calculations of a translation function map on a 1·0 Å grid for the data between 8·0 and 3·5 Å for the three Harker sections ($X = \frac{1}{2}$; $Y = \frac{1}{2}$; $Z = 0$) gave a self-consistent set of vectors (Table 3).

Table 3. *Results of the Crowther–Blow translation function, after removal of self-vectors*

X, *Y* and *Z* refer to fractions of the unit cell.

Harker section	Peak pairs	*X*	*Y*	*Z*	R.m.s.
$X = \frac{1}{2}$	$1^{I}\ 3^{II}$	$\frac{1}{2}$	73	1	6·93
		$\frac{1}{2}$	18	35	8·41
	1^{I}–7^{II}	$\frac{1}{2}$	73	1	6·93
		$\frac{1}{2}$	18	35	9·23
	2^{I}–7	$\frac{1}{2}$	73	99	6·66
		$\frac{1}{2}$	18	65	9·24
	2^{I}–3	$\frac{1}{2}$	73	99	6·63
		$\frac{1}{2}$	18	65	8·28
$Y = \frac{1}{2}$	$1^{I}\ 3^{II}$	13	$\frac{1}{2}$	1	7·02
		69	$\frac{1}{2}$	35	9·72
	$1^{I}\ 7^{II}$	12	$\frac{1}{2}$	1	7·01
		69	$\frac{1}{2}$	35	9·96
	2^{I}–7	87	$\frac{1}{2}$	99	7·90
		31	$\frac{1}{2}$	65	9·90
	2^{I}–3	87	$\frac{1}{2}$	99	8·02
		31	$\frac{1}{2}$	65	9·69
$Z = 0$	1^{I}–3^{II}	63	23	0	7·41
		19	68	0	7·70
	1^{I}–7^{II}	63	23	0	7·34
		19	68	0	8·85
	2^{I}–7	37	23	0	7·24
		81	68	0	8·81
	2^{I} 3	37	23	0	7·35
		81	68	0	7·76

All four solutions were submitted to *R*-value minimization (Ward, Wishner, Lattman & Love, 1975) and the first pair of angles 1^{I}–3^{II} and corresponding translation vectors were chosen on the basis of the low *R* value and the rank of the rotation peaks used.

Rigid-body refinement

The initial *R* value for the correctly oriented and positioned model (with the side chains truncated to Ala, as described above) was 0·449, for the intensity data between 7·0 and 3·0 Å resolution. A rigid-body refinement process was started in the 10·0–6·0 Å resolution range with the overall *B* factor of 20·0 Å^2 ($R = 0{\cdot}444$), using the program *TNT* (Tronrud, TenEyck & Matthews, 1987). Three positional and three rotational parameters were allowed to vary independently for each of the four chains in the asymmetric unit. During the refinement the resolution range was increased to 10·0–4·5 Å, and convergence, as judged from the evaluation of the conventional *R* factor, was reached after six cycles at $R = 0{\cdot}406$ for the intensity data between 10·0 and 4·5 Å.

Molecular dynamics and conventional refinement

The starting model to be used for molecular dynamics refinement using the *GROMOS* suite of programs (van Gunsteren & Berendsen, 1987), was constructed by proper replacement of the truncated amino-acid side chains with those indicated by the yeast superoxide dismutase sequence (Steinman, 1980) using the molecular modelling program *FRODO* (Jones, 1978). The resulting model, however, was not manually checked nor corrected for bad contacts in any way. Only a chain interruption was left at the insertion site shown by the sequence alignment at position 31 (the numbering scheme adopted for the y-SOD molecule is that of bovine SOD). Seven residues (32, 32*A*, 32*B*, 33, 34, 35, 35*A*) were not model built, and free C and N termini were left in the polypeptide chain at positions 31 and 36 respectively, to allow the refinement for the space needed for the seven-residue insertion. The initial *R* value for this model was 0·472, for the data between 6·0 and 3·0 Å resolution.

Pseudo energy minimization was first performed (EMX – a combination of energy minimization with X-ray restraints) to remove the initial strain in the structure due to bad contacts and poor chain stereochemistry. Molecular dynamics refinement including crystallographic potential (MDX) was typically performed in a series of 0·1 to 0·6 ps per cycle, consisting of steps of $\Delta t = 2$ fs. The balance between the crystallographic *R* factor and the energy terms was controlled by applying a weight σ_x to the X-ray potential term. EMX refinement in the resolution range 6·0–3·0 Å (150 cycles) was followed by 100 cycles of MDX refinement for the data between 10·0 and 5·0 Å. The σ_x value for this stage of refinement was chosen to be twice the observed σ_f $\{\sigma_f = [\langle(|F_o| - F_c|)^2\rangle]^{1/2}\}$. Fig. 1 presents σ_f as observed during the refinement and σ_x as applied in the weighting of the X-ray energy term. The MDX refinement was initialized at $T = 300$ K and an isothermal algorithm (Berendsen, Postma, van Gunsteren, DiNola & Haak, 1984) was applied to keep the overall temperature constant, using the temperature relaxation time $\tau = 0{\cdot}1$ ps. No restraints on bond lengths were used during EMX and MDX refinement procedure, *B* factors of 20·0 Å^2 were assigned to each atom. After 0·5 ps of molecular dynamics refinement a bad contact between the Lys94 of the chains *A* and *D* and Phe48 of the symmetry-related chains *C* and *B* was manually corrected, moving the residue chains in the electron density. Subsequently the resolution was gradually increased up to 2·5 Å spacing, limiting the low-resolution data to 6·0 Å in order to avoid the contribution of solvent structure at lower resolution. After 6·7 ps MDX simulation time, followed by EMX refinement ($R = 0{\cdot}299$), an electron density map with coefficients $(2|F_o| - |F_c|)$ and calculated phases was produced. The electron density clearly showed that the alignment chosen for modelling the insertion region around residues 31–36 was not correct. Instead the y-SOD sequence alignment suggested by Getzoff *et al.* (1989) appeared more consistent with the observed electron density

features. Residues 32, 33, 34, 35, 35*A* were therefore built into the model as Ala, residues 26 to 31 were trimmed to their Cβ atoms, and a break in the polypeptide chain was introduced deleting residues 24 and 25 (in order to allow for the creation of room for residues 24, 25 and the insertions 25*A*, 25*B*). The amino-acid sequence alignment of bovine and yeast SOD's (Getzoff *et al.*, 1989) and the two alternative alignments of the enzymes in the insertion region, discussed above, are presented in Figs. 2 and 3.

The refinement proceeded by EMX minimization (6·0–3·0 Å) followed by MDX simulation, increasing the resolution and gradually lowering the σ_x weight. After 3·6 ps of simulation time ($R = 0{\cdot}280$) a new electron density map was calculated. The electron density of the trimmed side-chain residues could be clearly recognized as well as a backbone tracing for the four missing residues in three (*A*, *B*, *D*) out of the four monomers. Residues 24, 25, 25*A*, 25*B* except for those in chain *C*, were individually built into the model; for chain *C* the corresponding segment of chain *D* was adopted using the local twofold axis of the dimer. This model was then submitted to EMX and MDX refinement again. After 0·654 ps a geometry regularization of the residues 20–40 in the four chains was applied using the *FRODO* package and the refinement proceeded further with EMX/MDX minimization. As the refinement seemed to level off after the total simulation time of 11·9 ps ($R = 0{\cdot}282$ with $\sigma_x = 0{\cdot}68\sigma_f$) the temperature was increased to 600 K ($\tau = 0{\cdot}01$ ps) and after 0·6 ps of the MDX simulation, the model was cooled to 300 K ($\tau = 0{\cdot}05$–$0{\cdot}1$ ps). After MDX refinement the model was energy minimized by EMX in order to release the system's kinetic energy. This process led to an *R* factor of 0·268 for the intensity data between 6·0 and 2·5 Å.

The model was subsequently subjected to the conventional restrained crystallographic refinement of the coordinates and temperature factors using the *TNT* programs suite (Tronrud *et al.*, 1987). The refinement reached convergence at $R = 0{\cdot}220$ for the data from 6·0 to 2·5 Å.

Data processing and reduction were performed on the Ethernet-based MicroVAX cluster at EMBL, Hamburg (FRG), molecular replacement and

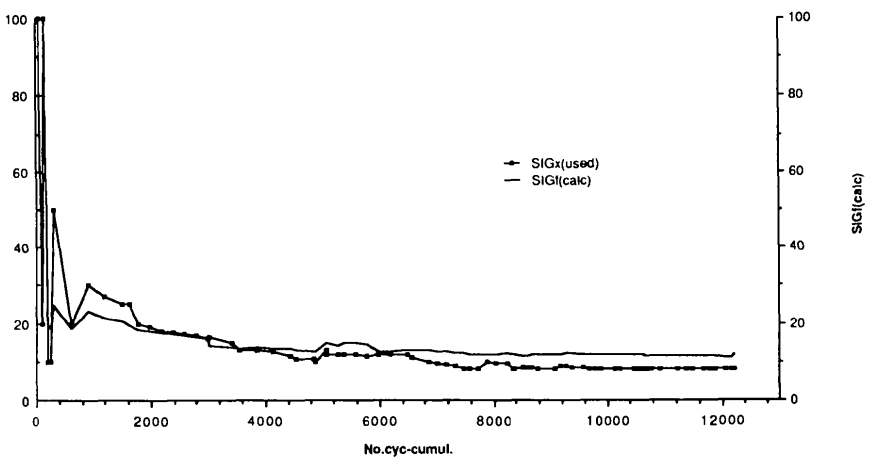

Fig. 1. σ_f as observed and σ_x as applied in the weighting of the X-ray energy term.

```
                          10                  20                  30
b-SOD   Ac-ATKAVCVLKGDGPVQGTIHFEAKGD--TVVVTGSITG-
y-SOD       VQAVAVLKGDAGVSGVVKFEQASESEPTTVSYEIAGN

            40                  50                  60                  70
b-SOD   LTEGDHGFHVHQFGDNTQGCTSAGPHFNPLSKKHGGPKD
y-SOD   SPNAERGFHIHEFGDATNGCVSAGPHFNPFKKTHGAPTD

                80                  90                  100                 110
b-SOD   EERHVGDLGNVTADKNGVAIVDIVDPLISLSGEYSIIGRT
y-SOD   EVRHVGDMGNVKTDENGVAKGSFKDSLIKLIGPTSVVGRS

                120                 130                 140                 150
b-SOD   MVVHEKPDDLGRGGNEESTKTGNAGSRLACGVIGIAK
y-SOD   VVIHAGQDDLGKGDTEESLKTGNAGPRPACGVIGLTN
```

Fig. 2. Amino-acid sequence alignment of bovine and yeast SOD (Getzoff *et al.*, 1989). Residues are numbered according to the bovine sequence and identical residues are printed in bold characters. Amino-acid residues are displayed according to their one-letter codes.

molecular dynamics calculations on a Convex C220 S computer at CILEA, Milano (Italy), while conventional refinement and molecular graphics were on a VAX8530 system and on a PS330 Evans–Sutherland graphics computer at the University of Pavia. The figures were prepared using the molecular graphics package *WHATIF* (Vriend, 1990).

Atomic coordinates and structure-factor amplitudes are deposited with the Brookhaven Protein Data Bank.*

Results and discussion

The rotation and translation searches were performed with a model whose residues non-common to those of bovine SOD primary structure were trimmed in their side chains in order to improve the signal-to-noise ratios (Schierbeek, 1988). The rotation and translation peaks, pointing to the correct solutions, had relatively high r.m.s. values (see Tables 1 and 3) even though the search was performed with 48% of the scattering matter in the asymmetric unit only. The two dimers in the asymmetric unit are related by a very approximate diad axis ($\kappa = 152°$). A similar situation has also been observed for bovine Cu,Zn SOD (Tainer *et al.*, 1982), where the two identical subunits of a SOD dimer are related by a non-crystallographic twofold axis that orients the two active sites facing away from each other. The two enzyme dimers in the bovine SOD case are related by a rotation of about 140° along a local twofold axis plus a small translation (Thomas, Rubin, Bier, Richardson & Richardson, 1974). Fig. 5 shows the y-SOD asymmetric unit in the crystal. As judged from the height of the translation function peaks, one of the two dimers in the asymmetric unit is more similar to the bovine SOD dimer than the other, suggesting that the two dimers are not strictly equivalent. The root-mean-square deviation calculated after superposition of the α-carbons of the dimers in the asymmetric unit is 0·615 Å; although the r.m.s. deviations relative to the superpositions of the two monomer chains within the dimers are not low (0·378 Å for the *A–B* superposition, and 0·459 Å for the *B–C* superposition), the pseudo-crystallographic diad axes relating the two monomers in each dimer were highly preserved ($\kappa = 177°$ in both dimers) during the refinement, in which no restraints were imposed on the non-crystallographic twofold axes.

The refinement method of molecular dynamics with crystallographic pseudo-energy terms was chosen because of its large radius of convergence owing to the presence of the kinetic energy which allows the conformational barriers to be overcome. The initial model was constructed without manual intervention, resulting in a number of close van der Waals contacts. Table 4 lists the energy terms of the initial and final molecular dynamics model, showing the decrease of different energy terms. The course of the MDX refinement monitored by *R* factor is shown in Fig. 4. The approach to the refinement was chosen in such a way as to give the MD a period of time, and a sufficient amount of freedom, for searching the conformational space (Gros, Fujinaga, Dijkstra, Kalk & Hol, 1989). The resolution was gradually increased (Fig. 4), while the initial weight of the X-ray term σ_x was kept below the σ_f value; subsequently the ratio σ_x/σ_f was gently decreased (low value of σ_x means high weight for the X-ray energy terms).

Fig. 6 superimposes the α-carbons of the initial bovine SOD search model employed in the molecular replacement analysis (MR model) and the current model after the first restrained-refinement session (TNT model). Many regions show differences between the two structures; in Table 5 the root-mean-square deviations as observed for the superposition of α-carbons during various stages of refinement are shown, the r.m.s. deviation between the initial bovine search model and the current yeast SOD being 0·965 Å. The large radius of convergence of MDX can also be appreciated from inspection of Table 6 which lists the residues that were displaced more than $2\sigma_D$ during the refinement (σ_D is the standard deviation observed upon comparison of the two coordinate sets). It can be noted that the largest

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1SDY, R1SDYSF), and are available in machine-readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Melbourne or Osaka. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37047 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

```
      20                  25                  30
b-SOD F  E  A  K  G  D  T  V  V  V  T  G  S  -
y-SOD F  E  Q  A  S  E  S  E  P  T  T  V  S  T

      33                  40
b-SOD -  I  T  G  -  L  T  E  G  D
y-SOD E  I  A  G  N  S  P  N  A  E
```

(*a*)

```
      20                  25        26        30
b-SOD F  E  A  K  G  D  -  -  T  V  V  V  T  G
y-SOD F  E  Q  A  S  E  S  E  P  T  T  V  S  T

                          36                  40
b-SOD S  I  T  G  -  L  T  E  G  D
y-SOD E  I  A  G  N  S  P  N  A  E
```

(*b*)

Fig. 3. Two alternative alignments of the bovine and yeast SOD in the region of the insertions: (*a*) corresponds to the alignment by Steinman (1980); (*b*) to the alignment by Getzoff *et al.* (1989).

Table 4. *Energy terms* (kJ mol^{-1}) *for the initial and final molecular dynamics model*

Energy term	Initial model	Final model
E(pot-tot)	$0{\cdot}27066 \times 10^{9}$	$-0{\cdot}11347 \times 10^{5}$
Bonds	$0{\cdot}6143 \times 10^{5}$	$0{\cdot}2744 \times 10^{5}$
Angles	$0{\cdot}2108 \times 10^{5}$	$0{\cdot}7355 \times 10^{5}$
Improper dihedral angle	$0{\cdot}1239 \times 10^{4}$	$0{\cdot}2028 \times 10^{5}$
Dihedral angle	$0{\cdot}2768 \times 10^{4}$	$0{\cdot}4816 \times 10^{5}$
Electrostatic energy	$-0{\cdot}1095 \times 10^{4}$	$-0{\cdot}2392 \times 10^{5}$
Lennard-Jones energy	$0{\cdot}2705 \times 10^{9}$	$-0{\cdot}1521 \times 10^{5}$
X-ray term	$0{\cdot}9585 \times 10^{5}$	$0{\cdot}2248 \times 10^{6}$
No. of atoms	4248	4432

displacements took place near the two insertion sites (residues 23, 24, 25 and residues 35, 37, 38). Another region characterized by substantial atomic shifts is at the interface between the two dimers (residues 127, 128 in the chain *D*). As the bond distances were not restrained during the MDX refinement, the r.m.s. deviations from ideality of the MDX model were relatively high, but were improved during the conventional restrained crystallographic refinement, as reported in Table 7.

An electron density map was calculated three times during the MDX refinement, and four manual interventions were performed on a computer graphics system. At stages *A* and *D* (Fig. 4) the bad contacts and poor stereochemistry at the insertion site were improved. At stage *B* the model was changed in order to fit the alignment by Tainer *et al.* (1982), moving the gap on the opposite side of the β-strand 3c, modifying 11 residues in the strand and in the following non-repetitive structure into alanines in order to avoid biasing of the refinement procedure. At stage *C* (Fig. 4) the correct and complete sequence of the model was introduced.

At the present stage of refinement the ligands as well as the symmetry of the coordination spheres around Cu^{2+} and Zn^{2+} are conserved and quite comparable with the values reported for the bovine

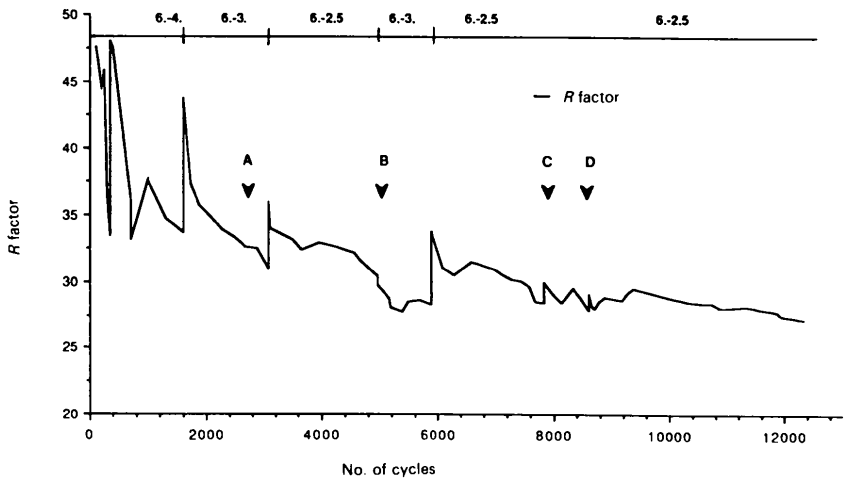

Fig. 4. The course of the refinement as indicated by the crystallographic *R* factor as a function of the number of cycles in the refinement. Labels at points along the refinement refer to models discussed in the text. The resolution limits in the refinement are shown.

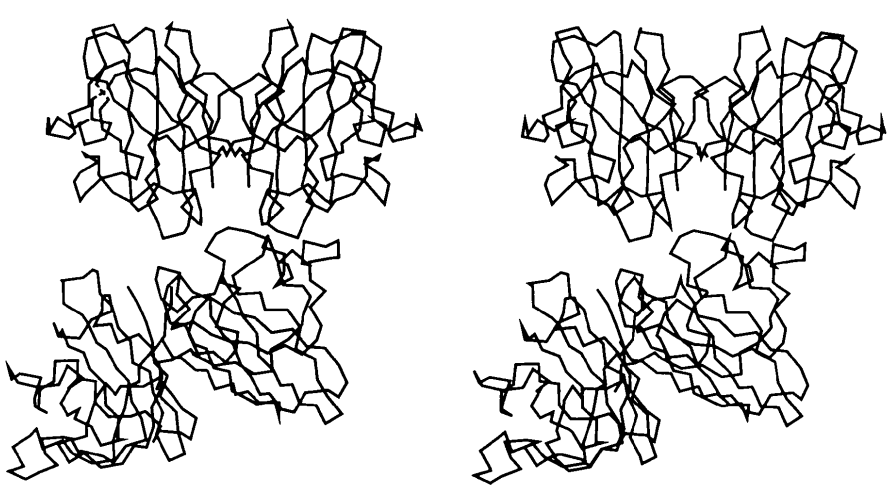

Fig. 5. Stereoview of the two y-SOD dimers present in the crystalline form examined.

Table 5. *Root-mean-square deviations* (Å) *observed upon superposition of α-carbons during various stages of the refinement*

MR – initial model (bovine SOD) used for the molecular replacement, RB – model after rigid-body refinement, MDX – model after molecular dynamics refinement, TNT – current model after the first restrained crystallographic refinement.

Stage of the refinement	MR	RB	MDX	TNT
Molecular replacement (MR)	–	0·521	0·955	0·964
Rigid-body refinement (RB)	–	–	0·819	0·820
Molecular dynamics (MDX)			–	0·092
Conventional refinement (TNT)			–	

Table 6. *Residues with the maximum root-mean-square deviations on the superposition of α-carbon atoms of the initial (MR) and the current (TNT) models*

Residue	Deviation (Å)	Residue	Deviation (Å)
Glu 25	7·57	Glu 225	6·71
Ser 24	6·25	Glu 625	5·98
Ser 624	5·98	Glu 425	5·82
Pro 37	4·65	Pro 637	4·56
Pro 437	4·48	Ser 424	4·42
Ala 623	4·34	Glu 689	4·06
Asp 688	3·88	Thr 750	3·76
Ser 436	3·73	Ser 636	3·68
Ser 36	3·61	Asn 638	3·51
Asn 38	3·49	Asn 438	3·45
Ile 302	3·05	Ala 23	2·99
Ile 305	2·96	Asp 728	2·92
Ser 224	2·89	Gly 295	2·76
Gly 727	2·76		

Table 7. *Root-mean-square deviations from ideal geometry after molecular dynamics (MDX) and after conventional refinement (TNT)*

Parameter	MDX model	TNT model
Bond length (Å)	0·022	0·013
Bond angle (°)	3·803	2·886
Torsion angle (°)	28·89	28·61
Trigonal plane	0·023	0·019
Planar groups	0·068	0·033
Non-bonded contacts (Å)	0·154	0·098
R factor	0·268	0·220

Table 8. *Comparison of the bond distances* (Å) *in the active sites of yeast and bovine superoxide dismutase, containing* Cu^{2+} *and* Zn^{2+} *ions*

		y-SOD	b-SOD			y-SOD	b-SOD
Cu^{2+}—N*D*1	His44	1·80	2·01	Zn^{2+}—N*D*1	His61	2·02	2·09
Cu^{2+}—N*E*2	His46	2·15	2·11	Zn^{2+}—N*D*1	His69	1·94	2·14
Cu^{2+}—N*E*2	His61	2·34	2·21	Zn^{2+}—N*D*1	His78	1·95	2·04
Cu^{2+}—N*E*2	His118	2·22	2·09	Zn^{2+}—O*D*1	Asp81	1·88	1·91
Cu^{2+}—N*D*1	His44	1·88	1·96	Zn^{2+}—N*D*1	His61	2·04	2·07
Cu^{2+}—N*E*2	His46	2·30	2·12	Zn^{2+}—N*D*1	His69	1·96	2·08
Cu^{2+}—N*E*2	His61	2·08	1·98	Zn^{2+}—N*D*1	His78	2·04	2·16
Cu^{2+}—N*E*2	His118	2·05	1·99	Zn^{2+}—O*D*1	Asp81	1·99	2·07
Cu^{2+}—N*D*1	His44	1·90	2·00	Zn^{2+}—N*D*1	His61	2·16	2·07
Cu^{2+}—N*E*2	His46	2·30	2·15	Zn^{2+}—N*D*1	His69	1·97	2·12
Cu^{2+}—N*E*2	His61	1·99	2·26	Zn^{2+}—N*D*1	His78	1·88	2·04
Cu^{2+}—N*E*2	His118	2·21	1·98	Zn^{2+}—O*D*1	Asp81	1·96	2·00
Cu^{2+}—N*D*1	His44	1·87	2·06	Zn^{2+}—N*D*1	His61	1·98	2·08
Cu^{2+}—N*E*2	His46	2·17	2·02	Zn^{2+}—N*D*1	His69	2·11	2·13
Cu^{2+}—N*E*2	His61	2·16	2·21	Zn^{2+}—N*D*1	His78	1·86	2·09
Cu^{2+}—N*E*2	His118	2·22	2·07	Zn^{2+}—O*D*1	Asp81	1·85	1·99

enzyme, as shown in Fig. 7 and Table 8. A more detailed analysis of the coordination spheres of the metal ions will await further crystallographic refinement. At this stage of refinement the fifth coordination site of copper, presumably occupied by a solvent molecule, cannot be unambiguously determined from electron density maps. Residue Arg141 forms part of the positively charged superoxide binding pocket above the Cu^{2+} ion. In spite of the two prolines flanking this residue in the yeast enzyme (substituting for Ser and Leu of bovine SOD), the conformation of the polypeptide backbone and side-chain orientation in all four monomers coincide with those observed in the bovine SOD. Two out of three salt bridges (Asp74···Lys126 and Arg77···Asp99) are strongly preserved in the y-SOD in comparison to the bovine enzyme. The Lys134···Glu131 salt bridge observed in bovine SOD

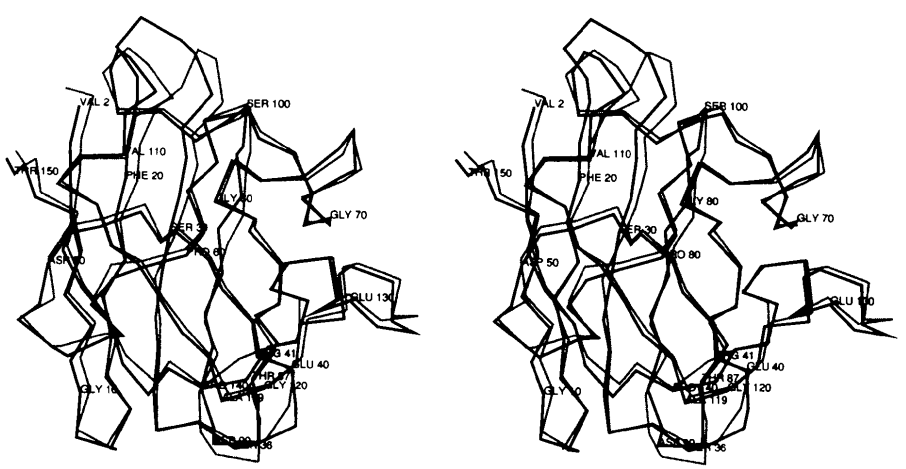

Fig. 6. *α*-Carbon superposition of the bovine search model used for the molecular replacement (dotted line) and the current refined TNT model (chain *A*).

is replaced by an ionic interaction between Lys134 and Glu130 in the yeast enzyme.

The β-barrel in the bovine enzyme is closed at opposite ends by the hydrophobic residues Leu36 and Leu104 that point towards the core of the structure. From inspection of the preliminary model it can be noticed that the Leu→Ser replacement occurring at position 36 in y-SOD, which would leave an opening in the upper part of β-barrel structure (Fig. 8), is compensated by substitutions at positions His41, Ala87, Glu119 (Arg41, Thr87, Ala119 in y-SOD), which allow efficient hydrogen bonding to Ser36 and Arg41 residues.

Concerning the intermolecular interactions which are observed in the whole asymmetric unit one should distinguish between intra-dimer contacts and the looser dimer-to-dimer interactions. The amino-acid residues which build up the monomer-to-monomer interface of bovine SOD are strongly conserved, or conservatively substituted, in the y-SOD. At position 150 the Ala→Thr substitution allows the establishment of an intra-dimer interaction with residue Asp50 (the distance $OD1$ Asp50···$OG1$ Thr350 is 2·82 Å in the *A–B* dimer, and 2·74 Å in the *B–C* dimer*), which is likely to stabilize the C terminus of y-SOD. On the other hand the main dimer–dimer interactions involve the y-SOD loop 25–26, containing one of the two inserted sequences (as compared to the bovine enzyme).

Area per la Ricerca di Trieste (Italy) is acknowledged for the financial support of Kristina Djinović. This project was supported by grants from the Italian National Research Council target oriented Project 'Biotecnologie e Biostrumentazione', special project 'Peptidi Bioattivi', and from the Ministry of University and Scientific Research.

* For identification purposes, residues of the *A* chain have been numbered 2 through 151; residues of the *B* chain 202 through 351; residues of the *C* chain 402 through 551; residues of the *D* chain 602 through 751.

References

ARNOLD, I. D. & LEPOCK, I. R. (1982). *FEBS Lett.* **146**, 302–306.
BANNISTER, J. V., BANNISTER, W. H. & ROTILIO, G. (1987). *CRC Biochem.* **22**, 111–180.
BARRA, D., BOSSA, F., MARMOCCHI, F., MARTINI, F., RIGO, A. & ROTILIO, G. (1979). *Biochem. Biophys. Res. Commun.* **86**, 1199–1205.

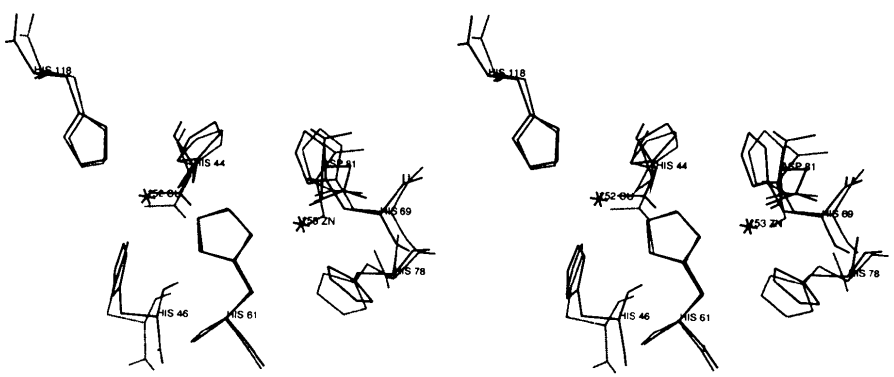

Fig. 7. Superposition of the active sites of the y-SOD monomer *A* with the 'orange' monomer of the bovine SOD (dotted line) as seen from the solvent.

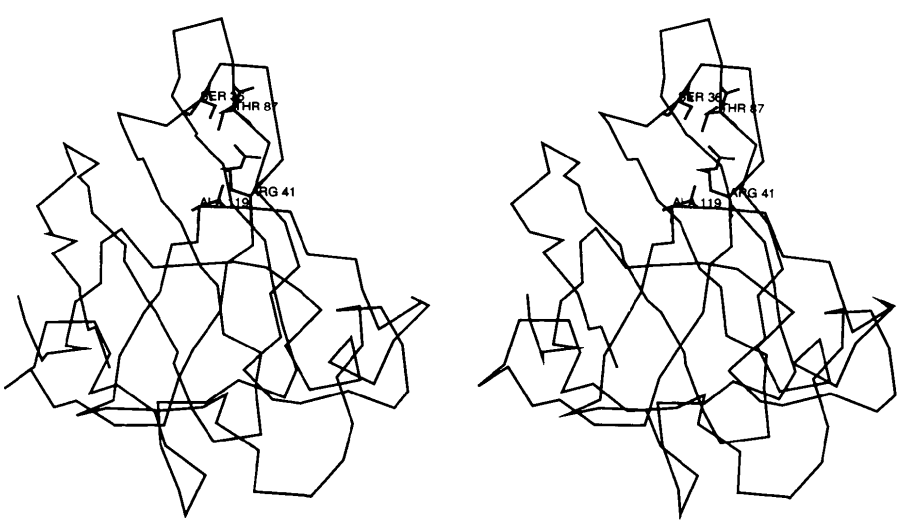

Fig. 8. α-Carbon drawing of monomer *A* showing the hydrogen-bonding network between residues Ser36, Arg41, Thr87 and Ala119 which acts as a lid to the upper part of SOD β-barrel.

BERENDSEN, H. J. C., POSTMA, J. P. M., GUNSTEREN, W. F. VAN, DINOLA, A. & HAAK, J. R. (1984). *J. Chem. Phys.* **81**(8), 3684–3690.
BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.
CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173–178. New York: Gordon & Breach.
CROWTHER, R. A. & BLOW, D. M. (1967). *Acta Cryst.* **23**, 544–548.
DESIDERI, A., FALCONI, M., PARISI, V. & ROTILIO, G. (1989). *FEBS Lett.* **250**, 45–48.
FIELDEN, E. M., ROBERTS, P. B., BRAY, R. C., LOWE, D. J., MAUTNER, G. N., ROTILIO, G. & CALABRESE, L. (1974). *Biochem. J.* **139**, 49–60.
FITZGERALD, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.
FRENCH, S. & WILSON, K. S. (1978). *Acta Cryst.* A**34**, 517–525.
FRIGERIO, F., FALCONI, M., GATTI, G., BOLOGNESI, M., DESIDERI, A., MARMOCCHI, F. & ROTILIO, G. (1989). *Biochem. Biophys. Res. Commun.* **160**(2), 677–681.
GETZOFF, E. D., TAINER, J. A., STEMPIEN, M. M., BELL, G. I. & HALLEWELL, R. A. (1989). *Proteins*, **5**, 322–336.
GROS, P., FUJINAGA, M., DIJKSTRA, B. W., KALK, K. H. & HOL, W. G. J. (1989). *Acta Cryst.* B**45**, 488–499.
GUNSTEREN, W. F. VAN & BERENDSEN, H. J. C. (1987). *BIOMOS. Biomolecular Software*. Laboratory of Physical Chemistry, Univ. of Groningen, The Netherlands.
JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
LATTMAN, E. E. & LOVE, W. E. (1970). *Acta Cryst.* B**26**, 1854–1857.
LESLIE, A. G. W., BRICK, P. & WONACOTT, A. J. (1986). *CCP4 News*, **18**, 33–39.
MACPHERSON, A. (1982). *Preparation and Analysis of Protein Crystals*, pp. 82–159. New York: Wiley.
MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
NYBORG, J. & WONACOTT, A. J. (1977). *The Rotation Method in Crystallography*, edited by U. V. ARNDT & A. J. WONACOTT, pp. 139–151. Amsterdam: North-Holland.
PARKER, M. W. & BLAKE, C. C. F. (1988). *J. Mol. Biol.* **199**, 649–661.
SCHIERBEEK, A. J. (1988). PhD Thesis, Rijskuniv. Groningen, The Netherlands.
STALLINGS, W. C., PATTRIDGE, K. A., STRONG, R. K. & LUDWIG, M. L. (1984). *J. Biol. Chem.* **259**, 10695–10699.
STALLINGS, W. C., PATTRIDGE, K. A., STRONG, R. K., LUDWIG, M. L., YAMAKURA, F., ISOBE, T. & STEINMAN, H. M. (1987). *Patterson and Pattersons*, edited by J. P. GLUSKER, B. K. PATTERSON & M. ROSSI, pp. 505–513. Oxford Univ. Press.
STEIGEMANN, W. (1974). PhD Thesis, Technische Univ. München, Germany.
STEINMAN, H. M. (1980). *J. Biol. Chem.* **225**, 6758–6765.
TAINER, J. A., GETZOFF, E. D., BEEM, K. M., RICHARDSON, J. S. & RICHARDSON, D. C. (1982). *J. Mol. Biol.* **160**, 181–217.
THOMAS, K. A., RUBIN, B. H., BIER, J. C., RICHARDSON, J. S. & RICHARDSON, D. C. (1974). *J. Biol. Chem.* **249**(17), 5677–6683.
TRONRUD, D. E., TENEYCK, L. F. & MATTHEWS, B. W. (1987). *Acta Cryst.* A**43**, 489–501.
VRIEND, G. (1990). *J. Mol. Graph.* **8**, 52–56.
WARD, K. B., WISHNER, B. C., LATTMAN, E. E. & LOVE, W. E. (1975). *J. Mol. Biol.* **98**, 161–177.

Acta Cryst. (1991). B**47**, 927–935

X-ray Studies on Crystalline Complexes Involving Amino Acids and Peptides. XXIII. Variability in Ionization State, Conformation and Molecular Aggregation in the Complexes of Succinic Acid with DL- and L-Lysine

BY G. SRIDHAR PRASAD AND M. VIJAYAN

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, *India*

(*Received* 14 *January* 1991; *accepted* 23 *April* 1991)

Abstract

Crystalline complexes of succinic acid with DL- and L-lysine have been prepared and analysed by X-ray diffraction. DL-Lysine complex: $C_6H_{15}N_2O_2^+.\frac{1}{2}C_4H_4O_4^{2-}.\frac{1}{2}C_4H_6O_4$, $M_r = 264{\cdot}2$, $P\bar{1}$, $a = 5{\cdot}506\ (4)$, $b = 8{\cdot}070\ (2)$, $c = 14{\cdot}089\ (2)$ Å, $\alpha = 92{\cdot}02\ (1)$, $\beta = 100{\cdot}69\ (3)$, $\gamma = 95{\cdot}85\ (3)^\circ$, $Z = 2$, $D_x = 1{\cdot}44$ g cm^{-3}, $R = 0{\cdot}059$ for 2546 observed reflections. Form I of the L-lysine complex: $C_6H_{15}N_2O_2^+.C_4H_5O_4^-$, $M_r = 264{\cdot}2$, $P1$, $a = 5{\cdot}125\ (2)$, $b = 8{\cdot}087\ (1)$, $c = 8{\cdot}689\ (1)$ Å, $\alpha = 112{\cdot}06\ (1)$, $\beta = 99{\cdot}08\ (2)$, $\gamma = 93{\cdot}77\ (2)^\circ$, $Z = 1$, $D_m = 1{\cdot}34\ (3)$, $D_x = 1{\cdot}34$ g cm^{-3}, $R = 0{\cdot}033$ for 1475 observed reflections. Form II of the L-lysine complex: $C_6H_{15}N_2O_2^+.\frac{1}{4}C_4H_4O_4^{2-}.\frac{1}{4}C_4H_6O_4.\frac{1}{4}(C_4H_5O_4\cdots H\cdots C_4H_4O_4)^2$, $M_r = 264{\cdot}2$, $P1$, $a = 10{\cdot}143\ (4)$, $b = 10{\cdot}256\ (2)$, $c = 12{\cdot}916\ (3)$ Å, $\alpha = 105{\cdot}00\ (2)$, $\beta = 99{\cdot}09\ (3)$, $\gamma = 92{\cdot}78\ (3)^\circ$, $Z = 4$, $D_m = 1{\cdot}37\ (4)$, $D_x = 1{\cdot}38$ g cm^{-3}, $R = 0{\cdot}067$ for 2809 observed reflections. The succinic acid molecules in the structures exhibit a variety of ionization states. Two of the lysine conformations found in the complexes have been observed for the first time in crystals containing lysine. Form II of the L-lysine complex is highly pseudosymmetric. In all the complexes, unlike molecules aggregate into separate alternating layers. The basic element of aggregation in the lysine layer in the complexes is an *S*2-type head-to-tail sequence. This element combines in different ways in the three structures. The basic element

0108-7681/91/060927-09$03.00