Structural Studies

on

Cu,Zn-Superoxide Dismutase from Spinach

A Doctoral Thesis Submitted

by

Yasuyuki Kitagawa

to

Faculty of Science

Osaka University

1987

APPROVALS

February, 1987

This thesis is approved

as to style and content

by

勝部幸輝

Member-in-chief

小林雅通

Member

中村晃

Member

田中信夫

Member

Acknowledgments

The work described in this thesis has been carried out under the guidance of Professor Yukiteru Katsube of Institute for Protein Research, Osaka University.

The author is deeply indebted to Professor Yukiteru Katsube for his incessant guidance and encouragement throughout this investigation.

The author is deeply indebted to Dr. Nobuo Tanaka for his useful discussion and advice of this work.

The author is deeply grateful to Professor Fumio Sakiyama and Dr. Susumu Tsunasawa for their guidance and advice of the primary structure determination.

The author is deeply thankful to Professor Yuhei Morita, Professor Kozi Asada and Dr. Shigeo Aibara of Research Institute for Food Science, Kyoto University for the useful guidance and advice of biochemical study.

The author is deeply grateful to Dr. Yoshinori Satow of Photon Factory, National Laboratory for High Energy Physics for his cordial support of X-ray experiment.

The author is deeply grateful to Drs. Yoshiki Matsuura, Yasuo Hata, Masami Kusunoki, Mamoru Sato and Genpei Lee for their kind guidance of X-ray structural study and to Chizuko Kitabayashi for her assistance of the primary structure determination.

The author would like to thank all the members in his laboratory for their kind assistance.

Finally, the author thanks his parents for their incessant understanding and encouragement.

北川康行

Yasuyuki Kitagawa

February, 1987

CONTENTS

Abbreviations

SOD	superoxide dismutase
Cu,Zn-SOD	copper,zinc-superoxide dismutase
Mn-SOD	manganese-superoxide dismutase
Fe-SOD	iron-superoxide dismutase
HPLC	high performance liquid chromatography
PTH	phenylthiohydantoin
SDS	sodium dodecyl sulfate
BrCN	cyanogen bromide
N-	amino
C-	carboxyl
M.I.R.	multiple isomorphous replacement
M.R.	molecular replacement
r.m.s.	root mean square
AAD	amino acid difference
ASA	accessible surface area
P.A.L.	present atmospheric level

1 Introduction

1-1 Superoxide dismutase (SOD)

Superoxide dismutase (SOD, EC 1.15.1.1) is the enzyme which catalyzes the dismutation of superoxide anion (O_2^-) to molecular oxygen (O_2) and hydrogen peroxide (H_2O_2) to protect living organisms from oxygen toxicity. The superoxide anion, which is the one electron reduction product of oxygen, is generated on reactions of oxidative metabolism or light radiation in living organisms. The anion is naturally disproportionable with a rate constant of 10^5 $M^{-1} \cdot sec^{-1}$ at physiological pH (pH 7.5), so the existence for a long time is rare. But the toxicity is so high as to induce the direct non-specific oxidation of organisms. SOD accelerates this disproportionation reaction with a rate constant of 10^9 $M^{-1} \cdot sec^{-1}$, which is 10^4 times higher than the natural disproportionation reaction, so is useful to vanish superoxide anions rapidly. Although some enzymes are requested to convert O_2^- into O_2 and H_2O as shown in Fig. 1-1, SOD working in the first step of the pathway is the most significant enzyme physiologically.

Since SOD inhibits the oxidation of the cell components or protects them from the oxygen toxicity, SOD becomes to be as pharmaceuticals. One is the medicine which eliminates O_2^- emanated by cancer radiotherapy and the other use is anti-rheumatism medicine of high security;

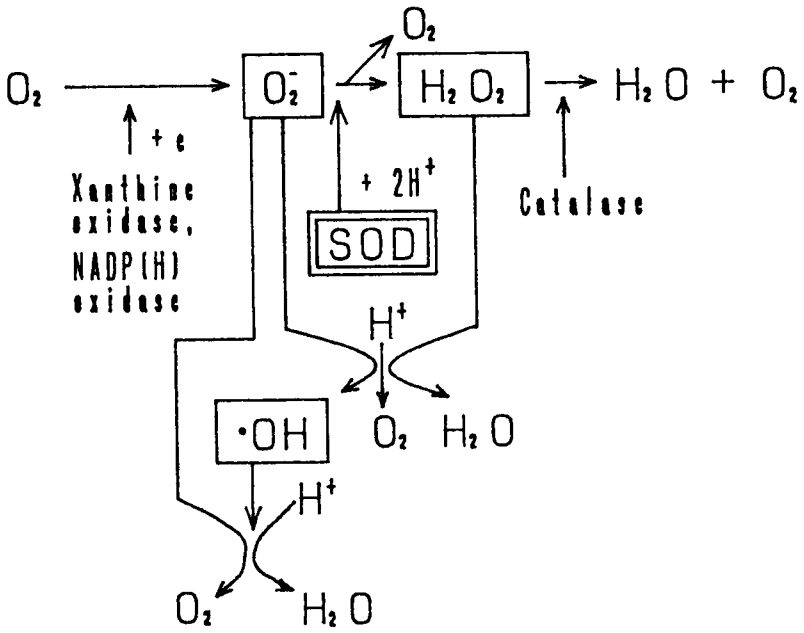

Fig. 1-1 Metabolism of active oxygen

which is now studying. These applications of SOD to medical science is very important for life science and strongly requests the analysis of the SOD structure by the X-ray crystal structure analysis which is a unique method to study the protein structure at their atomic level. In order to discuss the relation between structure and function of SOD, it is also important. For the purpose, I extracted and purified SOD from spinach leaves and determined both the primary structure and tertiary structure.

1-2 Classification of SOD

SOD was firstly characterized as a new metalloenzyme that protect living organisms from superoxide anions by Fridovich (*1*) in 1969. At the present stage, the three different types of SODs are known to exist in living organisms: SOD containing both Cu and Zn (Cu,Zn-SOD), SOD containing Mn (Mn-SOD) and SOD containing Fe (Fe-SOD) as cofactor. Although the three types are similar in functions, they are segregated into two families, Cu,Zn-SOD and Mn/Fe-SOD, of different properties. Among these three types, Mn-SOD is quite similar to Fe-SOD, not only in physico-chemical properties and N-terminal sequence, but also with regard to the three-dimensional structure (2-4). An SOD molecule with either manganese or iron is constructed with two or four identical subunits with

molecular weight of about 21,000. It is considered that these two types of SOD have evolved from an ancestral protein because of the structural similarity.

On the other hand, Cu,Zn-SOD, which was extracted and purified from a lot of organisms, is composed of two identical subunits of molecular weight 16,000 in all species. Cu,Zn-SOD is quite different from Mn/Fe-SOD in the structure (5). Most of the Cu,Zn-SODs isolated from different species can be hybridized with each other. This suggests that the tertiary structures around the subunit contact area in the dimer is similar. In an immunological cross-reactivity, however, each Cu,Zn-SOD demonstrates its individual nature (6). From these facts, the subunit contact area in Cu,Zn-SOD dimer is predicted to keep similar tertiary structure though there are some structural differences on the other parts.

The distribution of the three types of SOD among organisms is characteristic of the evolutionary levels of the organisms (Fig. 1-2). Cu,Zn-SOD and Fe-SOD are essentially only in eukaryotes and in prokaryotes, respectively. Mn-SOD is in prokaryotes and in mitochondria of eukaryotes.

1-3 Structural studies on Cu,Zn-SOD

Complete amino acid sequences of Cu,Zn-SOD have so far been established for bovine erythrocyte (7), human

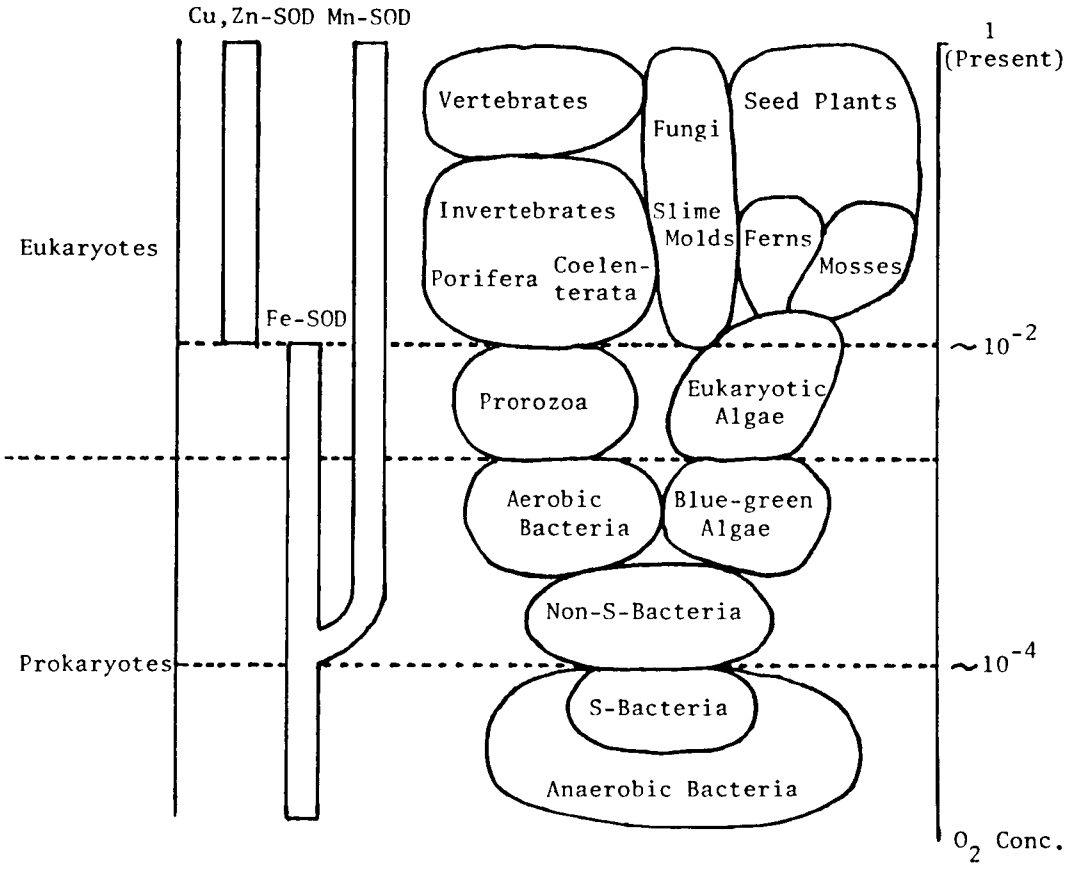

Fig. 1-2 Phylogenetic distribution of SODs

erythrocyte (*8*), horse liver (*9*), bakers' yeast (*10*), swordfish liver (*11*), fruit fly (*12*), porcine (*13-14*), *Neurospora crassa* (*15*) and a bacterium, *Photobacterium leiognathi* (*16*). The degree of their sequence conservation has been quite high, though not as high as that found among the eukaryotic cytochromes *c* (*17*). A study on the phylogenetic distribution has suggested that the Cu,Zn-SOD of *P. leiognathi*, which is the only one enzyme hitherto purified from a prokaryote, might have been acquired by gene transfer from the host fish (*18-19*). Since the evolutionary relationship of Cu,Zn-SODs has been studied only on the basis of the amino acid sequences of animals and fungi, the elucidation of the primary structure of a plant Cu,Zn-SOD is strongly requested for investigation of the molecular evolution of this enzyme. The three-dimensional structure of Cu,Zn-SOD is known only for bovine erythrocyte (*5*), which is made up of eight stranded β barrel with three loops. The structure of active site is constructed with two metals supported by four ligands respectively, one of which is held in common to these two metal ions (*5*). In plant and animal type SODs, it has been considered that they have almost the same size and shape from biochemical studies except for the immunological study (*20*). So it is worth analyzing the three-dimensional structure of the plant type Cu,Zn-SOD in order to compare the SODs of evolutionarily distinct two species (spinach and bovine).

2 Preparation

Spinach SOD was extracted and purified from its leaves by the modified method of Asada *et al.* (*21*). Spinach leaves (25 kg) were homogenized with 0.1 M potassium phosphate buffer (pH 7.8) (0.78 liter solution / 1 kg spinach leaves) using a Waring blender at medium speed for 30 seconds. The homogenate, after being forced through gauze, was centrifuged at 24,000 x *g* by an continuous centrifugation technique. The supernatant which overflowed was gathered, brought to 40% saturation with ammonium sulfate (0.243 g/ml), and left overnight at 5°C. The solution was centrifuged (10,000 x *g*) to discard the precipitates, and ammonium sulfate powder was added to the resulting supernatant up to the concentration of 80% saturation (0.285 g/ml). After being kept overnight, the precipitate was collected by centrifugation (10,000 x *g*), dissolved in 10 mM potassium phosphate buffer (pH 7.8), and dialyzed against the same buffer for three days. The subsequent purification steps were almost the same as that described previously (*21*), except that DEAE-Sepharose CL-6B and Sephacryl S-200 were used instead of DEAE-Sephadex A-50 and Sephadex G-100, respectively.

The fractions with blue to green color, separated on a Sephacryl S-200 column, were pooled and ultrafiltered to exchange the 10 mM potassium phosphate buffer (pH 7.8) for 5 mM potassium phosphate buffer (pH 6.8). The solution thus obtained was then passed through a hydroxylapatite

column equilibrated with the same buffer to remove a trace of contaminants. The enzyme in 10 mM potassium phosphate buffer (pH 7.8) was crystallized by the addition of ammonium sulfate up to about 55% saturation. The crystals collected were dissolved in water, desalted by dialysis against water, and lyophilized prior to the primary structure analysis.

Purification process of spinach Cu,Zn-SOD is shown in Fig. 2-1. Absorption spectrum in the ultraviolet and visible regions after the last stage of the purification steps is shown in Fig. 2-2. The present preparation of the spinach Cu,Zn-SOD showed a single band on sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, and the specific activity of the enzyme was 8,600 (enzyme units/ mg protein). The enzyme activity is as high as that reported previously (*21*).

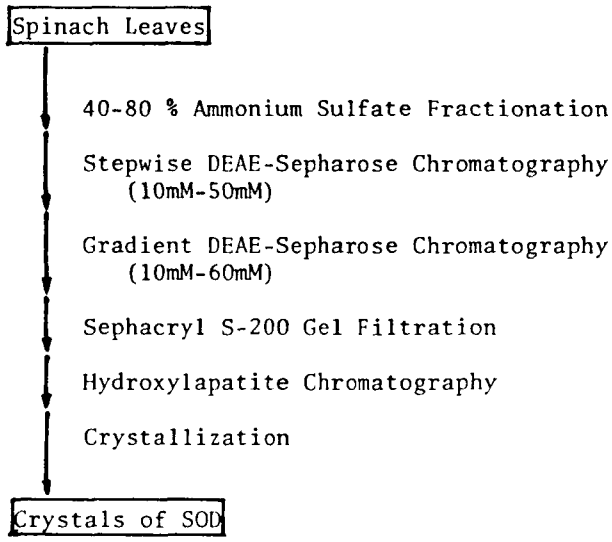

Fig. 2-1 A scheme of purification procedure of spinach superoxide dismutase

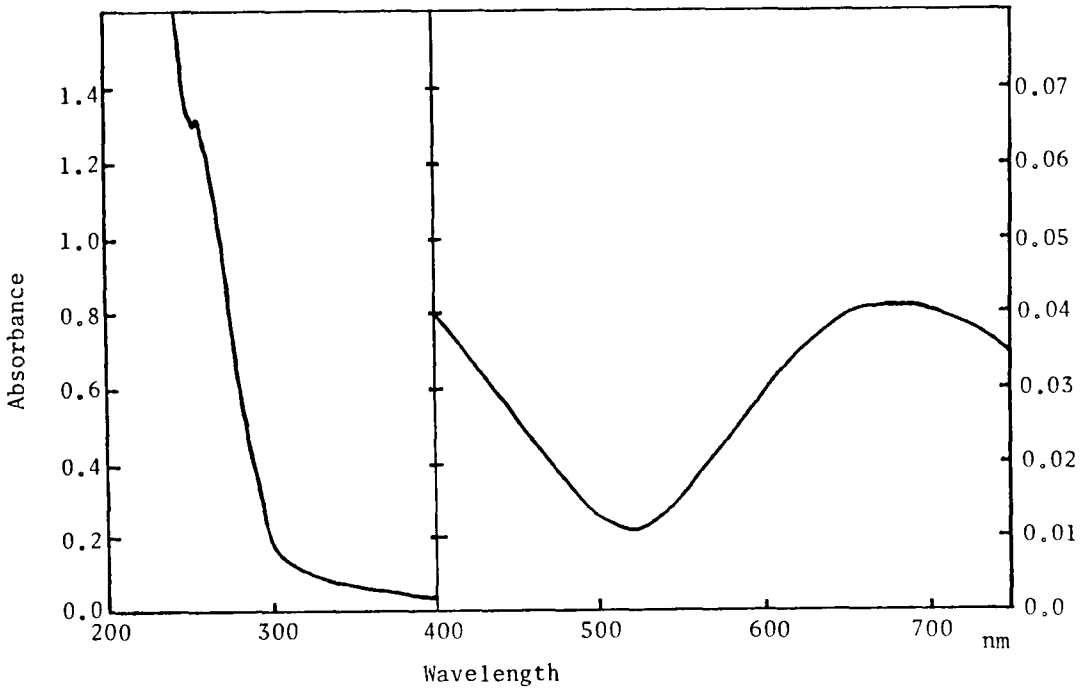

Fig. 2-2 Absorption spectrum of spinach superoxide dismutase. The concentration of the enzyme was 4.3 mg per ml. A 1-cm path length was used.

3 Amino acid sequence determination

3-1 Materials and methods

3-1-1 Materials

The following enzymes were obtained from the sources indicated in parentheses: *Staphylococcus aureus V8* protease (Miles Laboratories), thermolysin (Nakarai Chemicals) and trypsin treated with L-1-(p-tosylamino)-2-phenylethyl chloromethyl ketone (Worthington Biochemical Co.). *Achromobacter* lysyl*endo*peptidase (*Achromobacter lyticus* protease I (22)) was kindly gifted by Dr. T. Masaki of Ibaraki University. A Baker bond wide pore butyl column (0.46 x 25 cm, C4, 330 A, J. T. Baker Chemical) was used for high performance liquid chromatography (HPLC). All other chemicals were of the purest grade commercially available.

3-1-2 Amino acid analysis and sequence determination

Protein and peptides were hydrolyzed in 5.7 M HCl at 110^{o}C in evacuated, sealed tubes for 22 to 24 h. Cysteine was determined as cysteic acid after performic acid oxidation (*23*). The hydrolysates were analyzed with a Hitachi 835S amino acid analyzer. The values of amino acid analysis were not corrected for decomposition or incomplete hydrolysis.

The N-terminal sequences of the native protein, BrCN peptide, and peptides obtained by protease digestions were

determined with a 470A gas-phase automatic sequencer (Applied Biosystems, Inc.).

Phenylthiohydantoin(PTH)-amino acids were determined by reverse-phase HPLC under the same conditions as described previously (*24*). Cysteine and methionine residues were identified as PTH-cysteic acid and PTH-methionine sulfone, respectively, under the same conditions as described (*24*).

3-1-3 Acetylation and subsequent BrCN cleavage

Acetic anhydride (10 µl) was added in five portions to the native protein (11 nmol) dissolved in 100 µl of 25% trimethylamine solution over 30 min at room temperature. After the reaction mixture had been kept for another 2 h, the resulting acetylated protein was diluted with water and lyophilized. To the lyophilizate a 100-fold molar excess of BrCN over methionine in 70% formic acid (100 µl) was added and allowed to stand for 22 h at 40°C. The resulting sample was lyophilized and used for sequencing on the automatic sequencer without further purification.

3-1-4 Enzymic digestions

Prior to protease digestion, the native protein was oxidized with performic acid in order to cleave the intra-subunit disulfide bond (*23*).

Digestion of the oxidized protein with *A. lyticus* lysyl*endo*peptidase (AP) was carried out in 50 mM Tris-HCl buffer (pH 9.5) for 6 h at 35°C at an enzyme-to-substrate

ratio of 1:500 (w/w) and that with *S. aureus V8* protease (SP) was performed in 1% $(NH_4)HCO_3$ for 17 h at 35°C at an enzyme-to-substrate ratio of 1:100 (w/w). The lyophilized digests were purified by HPLC.

Fraction AP-5 was further digested with thermolysin or trypsin, each digestion was carried out in 1% $(NH_4)HCO_3$ for 16 h at 37°C at an enzyme-to-substrate ratio of 1:100 (w/w).

3-1-5 Separation of peptides by reverse-phase HPLC

Separation of peptides was performed by HPLC using a C4 column on a Hitachi 655 liquid chromatograph equipped with a 655-60 gradient processor and a Rheodyne sample injector with a 100-µl sample loop. Except for the rechromatography of fraction SP-A, the elution of peptides was carried out with a linear gradient of organic solvent (2-propanol : acetonitrile = 7 : 3, v/v) from 0 to 60% (v/v) for 1 h in 0.1% trifluoroacetic acid. For the fraction SP-A a linear gradient of acetonitrile from 0 to 40% (v/v) for 1 h in 0.1 M ammonium formate was used. The fractionated peptides were manually collected by monitoring the absorbance at 215 nm or 225 nm. The flow rate was 1.0 ml/min.

3-1-6 C-terminal sequencing

Hydrazinolysis and carboxypeptidase Y digestion were used to determine the C-terminus and the C-terminal sequence of this protein. The native protein (10 nmol) was

dissolved in anhydrous hydrazine (100 µl) and heated at 105°C for 6 h in an evacuated, sealed tube (*25*). For carboxypeptidase Y digestion, the performic acid-oxidized protein or SP-10 peptide was incubated in 300 µl of 50 mM sodium phosphate buffer (pH 7.0) at 37°C at an enzyme-to-substrate ratio of 1:100 (w/w). When 0, 1, 3, 6, and 24 h had elapsed, 50 µl of the solution was withdrawn and submitted to the amino acid analysis, respectively (Fig. 3-1).

3-1-7 Nomenclature

All peptides are named by using letters and numbers. Letters indicate the type of cleavage: AP, *A. lyticus* lysyl*endo*peptidase; SP, *S. aureus V8* protease; Th, Thermolysin; T, Trypsin. Numbers indicate the order in the final sequence of spinach SOD.

3-2 Results

3-2-1 N-terminal sequence

The amino acid composition of the enzyme is shown in Table 3-1. The values based on the amino acid analysis are in good agreement with those deduced from the amino acid sequence when low recovery of valine from four Val-Val sequences is considered.

The N-terminal sequence was determined by a gas-phase sequencer up to the 46th residue using 10 nmol of the

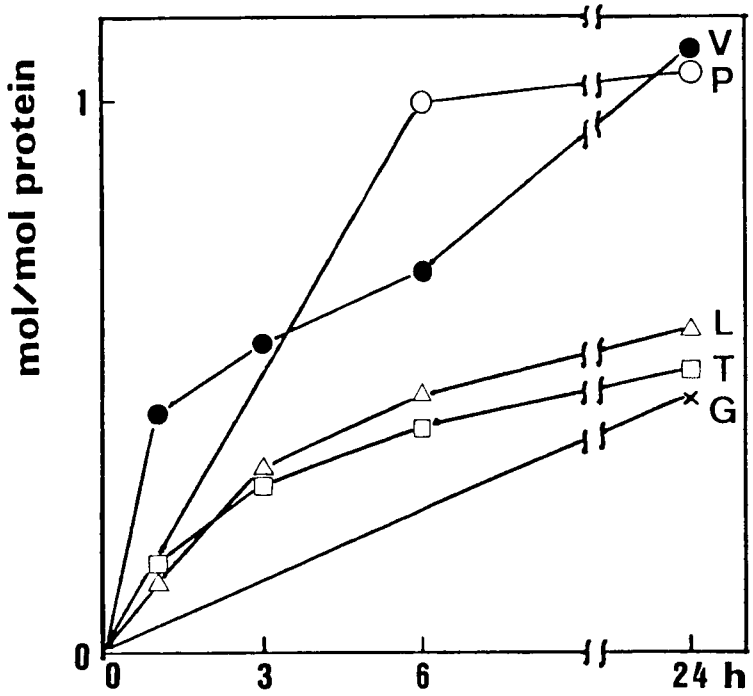

Fig. 3-1 Carboxypeptidase Y digestion of the performic acid-oxidized enzyme.

Table 3-1

Amino acid composition of spinach SOD

Amino acid	residues/mol	
Aspartic acid	18.9	(10)[a]
Asparagine		(9)
Threonine	14.9	(16)
Serine	5.3	(5)
Glutamic acid	11.2	(9)
Glutamine		(2)
Proline	9.7	(9)
Glycine	24.4	(23)
Half-cystine[b]	2.1	(2)
Alanine	12.0	(12)
Valine	16.2	(18)
Methionine	0.8	(1)
Isoleucine	3.9	(4)
Leucine	12.1	(12)
Phenylalanine	3.2	(3)
Lysine	7.0	(7)
Histidine	8.0	(8)
Arginine	4.1	(4)
Total		154

a Numbers in parentheses are based on the amino acid sequence.

b Half-cystine was determined as cysteic acid after performic acid oxidation.

intact protein. This shows that the N-terminus of spinach SOD is unblocked and dissimilar to those from animals.

3-2-2 N-terminal sequence of acetylated enzyme cleaved by BrCN

Since the spinach SOD contained only one methionine residue per subunit, it was acetylated to block the α-amino group, followed by BrCN cleavage and subjected to the automated Edman degradation without separating the cleaved products. By this method, the 30 residues following Met-58 were determined, with the exception of the residues at the 10th and 11th cycles. These two residues were finally identified as lysines by analyzing the sequences of AP-4 and SP-5. The failure of direct determination of these lysine residues is due to the acetylation of the ε-amino groups before Edman degradation.

3-2-3 Amino acid sequences of AP peptides

The elution profile of the digest of the performic acid-oxidized protein with *A. lyticus* lysyl*endo*peptidase from the C4 reverse-phase column is shown in Fig. 3-2, and the amino acid compositions of AP peptides thus obtained in Table 3-2.

The digest was resolved into six major and one minor peaks by HPLC. Peak AP-1, one of the major peaks, contained three lysines together with one each of alanine and threonine. This fraction was eventually shown to contain an equimolar mixture of a tripeptide and two

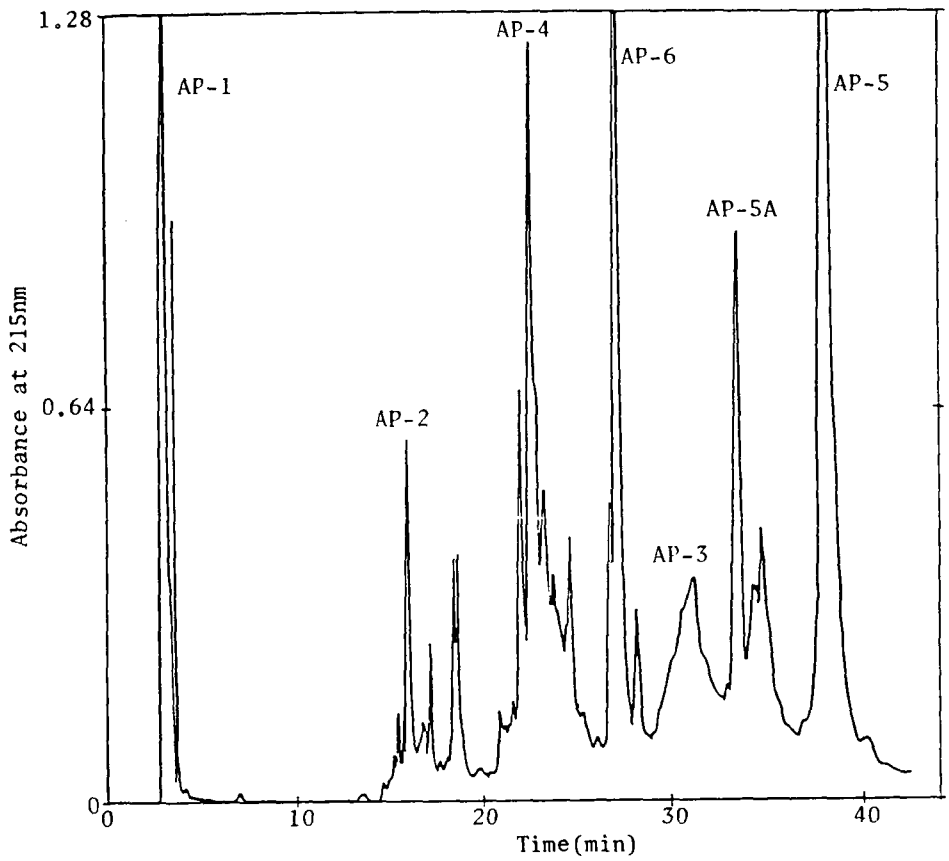

Fig. 3-2 Separation of peptides obtained by *A. lyticus* lysyl*endo*peptidase digestion of the performic acid-oxidized enzyme by HPLC. The sample was dissolved in 100 µl of 0.1% trifluoroacetic acid and chromatographed on a Baker bond wide pore butyl column (0.46 x 25 cm, C4, 330 A) at a flow rate of 1.0 ml/min. For other details, see text. Peptides in the peaks indicated were collected and lyophilized. Some other peaks in the figure were not collected because of their low yields.

Table 3-2

Amino acid compositions of AP peptides of oxidized SOD

Amino acid	AP-1	AP-2	AP-3	AP-4	AP-5[d]	AP-6	AP-5A[d]
Asp	-	-	4.1(4)	4.7[b](4)	9.2(10)	1.4(1)	7.3(8)
Thr	1.1(1)[a]	-	4.6(5)	2.8(3)	4.8(4)	2.8(3)	4.0(4)
Ser	-	-	2.1(2)	1.0(1)	1.6(1)	1.0(1)	1.4(1)
Glu	-	-	3.2(3)	1.1(1)	5.8(6)	1.4(1)	4.1(4)
Pro	-	-	2.1(2)	2.0(2)	3.4(3)	1.9(2)	2.7(3)
Gly	-	-	5.3(5)	4.4(4)	8.2(7)	6.8(7)	6.3(6)
Ala	1.0(1)	2.0(2)	1.4(1)	-	5.8(6)	2.0(2)	4.6(5)
1/2Cys[c]	-	-	-	1.1(1)	-	1.0(1)	-
Val	-	1.8(2)	4.3(5)[e]	-	6.3(8)	2.4(3)[e]	4.9(6)
Met	-	-	-	-[b](1)	-	-	-
Ile	-	-	1.0(1)	-	2.1(3)	-	2.2(3)
Leu	-	1.0(1)	2.1(2)	1.2(1)	5.1(5)	3.3(3)	2.1(2)
Phe	-	-	-	2.8(3)	-	-	-
Lys	2.9(3)	1.1(1)	0.9(1)	1.0(1)	1.7(1)	-	-
His	-	-	-	3.7(4)	3.0(3)	1.1(1)	1.8(2)
Arg	-	-	1.0(1)	-	1.9(2)	1.0(1)	1.8(2)
Positions	1-3,4,69	5-10	11-42	43-68	70-128	129-154	70-115
Yields(%)	33.0	45.9	24.9	15.8	39.9	45.1	17.3

a Numbers in parentheses are theoretical ones based on the amino acid sequence.

b Methionine sulfone was coeluted with aspartic acid.

c Determined as cysteic acid.

d This peptide is still impure.

e A Val-Val sequence is present in the peptide chain.

lysines after the completion of analysis of the whole sequence, although the analysis for the peptide in peak AP-1 was not carried out. The tripeptide was Ala-Thr-Lys in positions 1 to 3, and the two lysines were from positions 4 and 69. Peak AP-2 was not sequenced but from the amino acid composition was placed reasonably at positions 5 to 10 of the intact protein. Four other peaks, AP-3 to AP-6, were directly sequenced with the automatic sequencer. For AP-5, which was still slightly impure, the sequence over the 35th cycle was difficult to elucidate. The C-terminal consecutive residues (Pro-153 and Val-154) were not identified in AP-6. This peptide, lacking a lysine residue, was finally identified as the C-terminal peptide.

AP-5 was subfragmented with thermolysin or trypsin to establish the complete sequence and to overcome the problems arising from the impurities. The elution profiles of the digests with thermolysin and trypsin on HPLC columns and the amino acid compositions of the purified peptides are shown in Figs. 3-3 and 3-4 and Tables 3-3 and 3-4, respectively.

Th-5 and Th-6 were sequenced and identified as Val-112 - Ala-116 and Leu-117 - Lys-128, respectively, by comparison with the amino acid sequences of SP-7, SP-8, and SP-9, as described later. The other purified thermolysin peptides were assigned to the relevant positions in the amino acid sequence of AP-5 by their amino acid compositions.

As expected, three tryptic peptides, T-1 to T-3, were

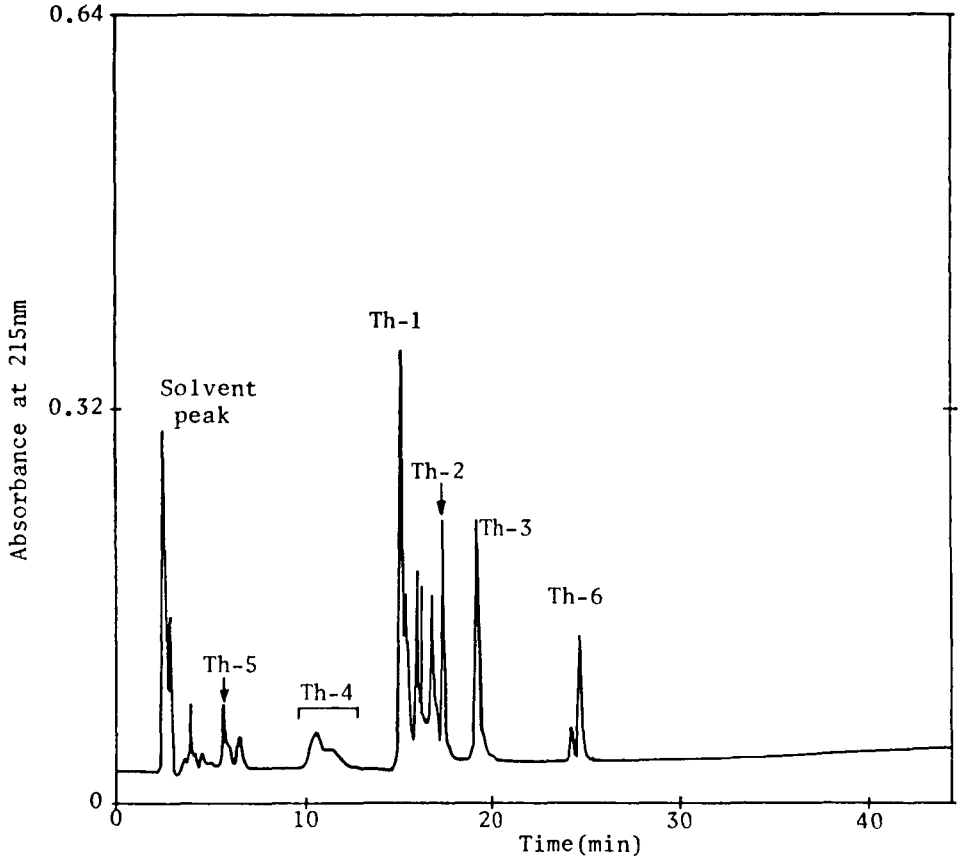

Fig. 3-3 Separation of thermolysin peptides from AP-5 by HPLC. The same conditions as described in Fig. 3-2 were used.

Table 3-3

Amino acid compositions of Th peptides from peptide AP-5

Amino acid	Th-1	Th-2	Th-3	Th-4	Th-5	Th-6
Asp	1.1(1)[a]	2.1(2)	2.3(2)	1.0(1)	-	2.0(2)
Thr	0.9(1)	1.8(2)	-	0.8(1)	-	-
Ser	-	-	-	1.1(1)	-	-
Glu	1.7(2)	1.2(1)	1.4(1)	-	-	2.2(2)
Pro	1.1(1)	-	1.1(1)	0.9(1)	-	-
Gly	1.2(1)	1.4(1)	-	1.2(1)	1.7(1)[c]	1.9(1)[c]
Ala	-	3.0(3)	-	-	1.2(1)	-
Val	-	1.5(2)	0.7(1)	-	1.1(2)[d]	1.3(2)[d]
Ile	-	-	1.5(2)	-	-	-
Leu	-	-	-	0.9(1)	-	2.7(3)
Lys	-	-	-	-	-	1.0(1)
His	1.1(1)	-	-	-	-	0.8(1)
Arg	-	-	-	-	1.0(1)	-
Positions	70-77	88-98	99-105	106-111	112-116	117-128
Yields(%)[b]	52.3	57.0	65.1	67.9	33.8	28.2

a Numbers in parentheses are theoretical ones based on the amino acid sequence.

b The yields are based on the amount of AP-5 used.

c Contamination with free glycine during manipulation is suspected.

d A Val-Val sequence is present in the peptide chain.

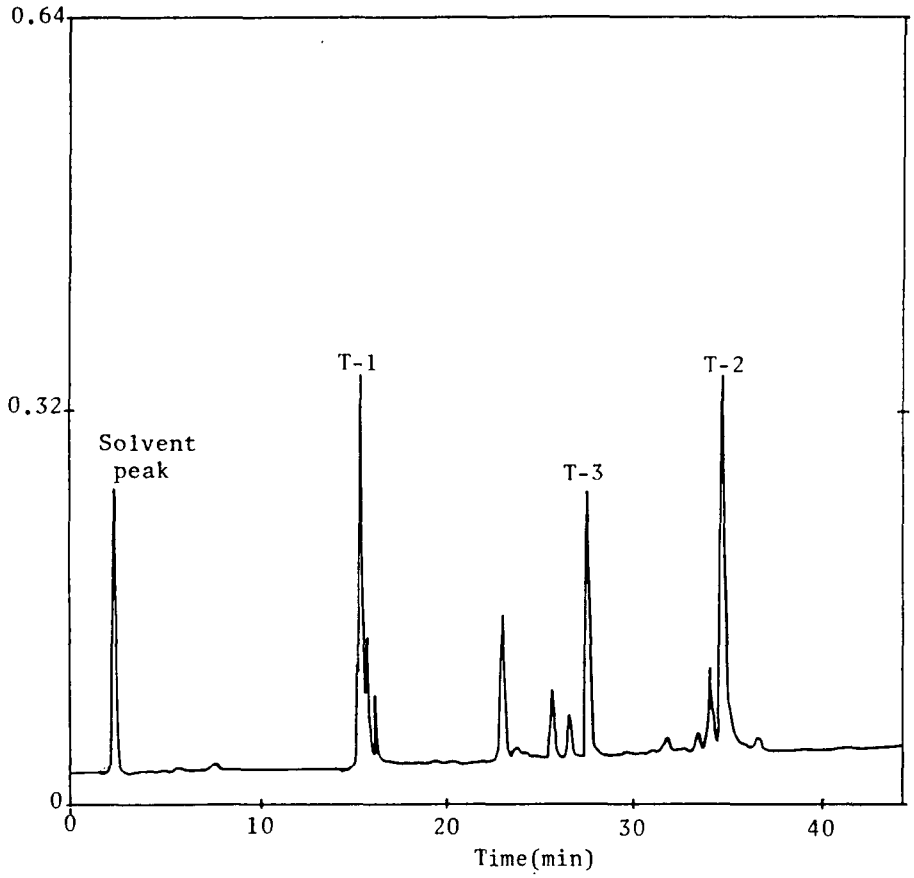

Fig. 3-4 Separation of tryptic peptides from AP-5 by HPLC. The same conditions as described in Fig. 3-2 were used.

Table 3-4

Amino acid compositions of tryptic peptides from peptide AP-5

Amino acid	T-1	T-2	T-3
Asp	1.0(1)[a]	6.5(7)	2.2(2)
Thr	0.9(1)	2.7(3)	-
Ser	-	1.5(1)	-
Glu	2.0(2)	2.1(2)	2.2(2)
Pro	1.2(1)	2.2(2)	-
Gly	1.3(1)	5.3(5)	1.4(1)
Ala	1.0(1)	3.9(4)	1.1(1)
Val	0.9(1)	3.6(5)[d]	1.3(2)[c]
Ile	-	2.1(3)[d]	-
Leu	-	1.8(2)	2.9(3)
Lys	-	-	1.0(1)
His	0.9(1)	1.0(1)	0.9(1)
Arg	0.8(1)	0.9(1)	-
Positions	70-79	80-115	116-128
Yields(%)[b]	54.1	35.9	52.8

a Numbers in parentheses are theoretical ones based on the amino acid sequence.

b The yields are based on the amount of AP-5 used.

c A Val-Val sequence is present in the peptide chain.

d A Val-Val and two Ile-Val sequences are present in the peptide chain.

isolated by HPLC. The amino acid compositions of T-1 and T-3 agreed well with those based on amino acid sequences, but a small difference was observed in the numerical values for three amino acids in T-2. The low values for valine and isoleucine are probably due to the incomplete hydrolysis of the Val-Val and two Ile-Val bonds in this peptide. T-2 was perhaps accompanied by the impurities of AP-5, which seem to be responsible for the difference of the value for serine in T-2.

The isolation of a minor peptide, AP-5A, was unexpected since this peptide is derived from unusual cleavage of the Arg-115 - Ala-116 bond in AP-5. However, the counterpart of AP-5A was not isolated.

3-2-4 Amino acid sequences of SP peptides

To obtain overlaps with AP peptides, the oxidized protein was digested with *S. aureus V8* protease. The resulting peptides were separated by HPLC on a C4 reverse-phase column. The elution profile and the amino acid compositions of the SP peptides are shown in Fig. 3-5 and in Table 3-5, respectively.

The digest was resolved into ten major peaks. From the amino acid compositions, the peptides from SP-1 to SP-3 were from residues 1 to 16, 17 to 24, and 25 to 34, respectively. The peptides from SP-4 to SP-9, except SP-5, were sequenced perfectly, the complete sequence of AP-5 being determined and the connection of AP-5 to AP-6 being established. SP-10 contained no glutamic acid and was

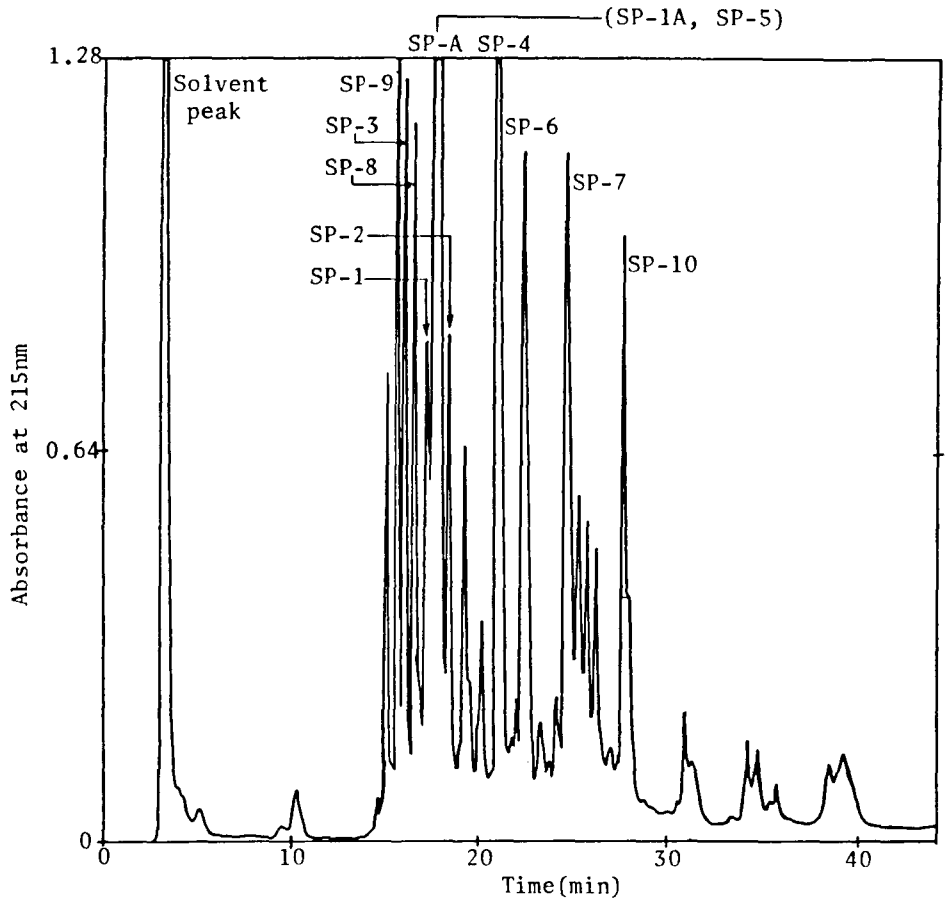

Fig. 3-5 Separation of peptides obtained by *S. aureus V8* protease digestion of the performic acid-oxidized enzyme by HPLC. The same conditions as described in Fig. 3-2 were used.

Table 3-5

Amino acid compositions of SP peptides of oxidized SOD

Amino acid	SP-1	SP-2	SP-3	SP-4	SP-5	SP-6
Asp	1.6(1)[a]	-	3.0(3)	-	6.1[b](5)	4.0(4)
Thr	1.8(2)	2.1(2)	1.9(2)	-	3.7(4)	1.1(1)
Ser	1.0(1)	-	-	1.0(1)	1.3(1)	-
Glu	1.2(1)	2.2(2)	-	1.4(1)	2.1(2)	1.2(1)
Pro	-	-	1.1(1)	1.1(1)	2.8(3)	-
Gly[f]	1.5(1)	1.4(1)	1.4(1)	3.3(3)	4.4(4)	3.4(3)
Ala	2.5(3)	-	-	1.2(1)	1.5(1)[f]	3.1(3)
1/2Cys[c]	-	-	-	-	0.7(1)	-
Val	2.8(3)	1.3(2)[e]	1.9(2)	-	-	2.5(3)
Met	-	-	-	-	-[b](1)	-
Ile	-	-	-	0.8(1)	-	0.7(1)
Leu	1.0(1)	1.0(1)	-	1.7(2)	-	1.1(1)
Phe	-	-	-	0.9(1)	1.5(2)	-
Lys	2.5(3)	-	-	0.9(1)	1.7(2)	-
His	-	-	-	2.6(3)	1.4(2)	1.1(1)
Arg	-	-	0.8(1)	-	-	0.9(1)
Positions	1-16	17-24	25-34	35-49	50-77	78-96
Yield(%)	23.2	43.9	44.1	34.0	10.0[d]	37.1

Table 3-5 (continued)

Amino acid	SP-7	SP-8	SP-9	SP-10	SP-1A
Asp	3.1(3)	-	2.0(2)	1.2(1)	1.1(1)
Thr	2.1(2)	-	-	2.9(3)	0.9(1)
Ser	1.0(1)	-	-	1.0(1)	0.9(1)
Glu	1.1(1)	1.3(1)	1.9(2)	-	1.0(1)
Pro	2.1(2)	-	-	2.1(2)	-
Gly[f]	2.4(2)	-	3.0(3)	5.2(5)	1.2(1)
Ala	1.2(1)	1.2(1)	-	2.1(2)	1.9(2)
1/2Cys[c]	-	-	-	1.1(1)	-
Val	2.2(3)[e]	1.4(2)[e]	-	2.4(3)[e]	2.2(3)
Met	-	-	-	-	-
Ile	1.6(2)	-	-	-	-
Leu	1.1(1)	1.0(1)	2.0(2)	3.0(3)	0.7(1)
Phe	-	-	-	-	-
Lys	-	-	1.1(1)	-	0.9(1)
His	-	1.1(1)	1.0(1)	-	-
Arg	1.0(1)	-	-	1.0(1)	-
Positions	97-115	116-121	122-132	133-154	5-16
Yield(%)	41.3	49.6	49.0	31.2	9.8[d]

a Numbers in parentheses are theoretical ones based on the amino acid sequence.

b Methionine sulfone was coeluted with aspartic acid.

c Determined as cysteic acid.

d Uncorrected for loss of recovery due to rechromatography.

e A Val-Val sequence is present in the peptide chain.

f The reason for over-estimation, see text.

deduced as the C-terminal peptide chain composed of residues 133 to 154. Fraction SP-A contained two peptides and was rechromatographed on the reverse-phase column. The elution profile is shown in Fig. 3-6, and the amino acid compositions of the peptides purified (SP-5, SP-1A) are in Table 3-5. SP-5 was sequenced completely and placed at positions 50 to 77. The value estimated for alanine appears to exceed the experimental error, but the established amino acid sequence is consistent with those determined for AP-4, AP-5, and Ac-BrCN. Consequently the alignment of AP-3 - AP-4 - Lys - AP-5 - AP-6 was established unequivocally.

3-2-5 C-terminal sequence

The C-terminal residue was identified as valine by hydrazinolysis. Carboxypeptidase Y digestions of the performic acid-oxidized protein and SP-10 released the same five C-terminal amino acids, Val, Pro, Thr/Leu and Gly, in this order, as shown in Fig. 3-1. By a comparison of the order of release of these amino acids with the determined sequence of the C-terminal portion of AP-6, the C-terminal sequence of spinach SOD was deduced to be Gly-Leu-Thr-Pro-Val.

The complete amino acid sequence of spinach SOD thus established is presented in Fig. 3-7, together with sequencing data of each isolated peptide.

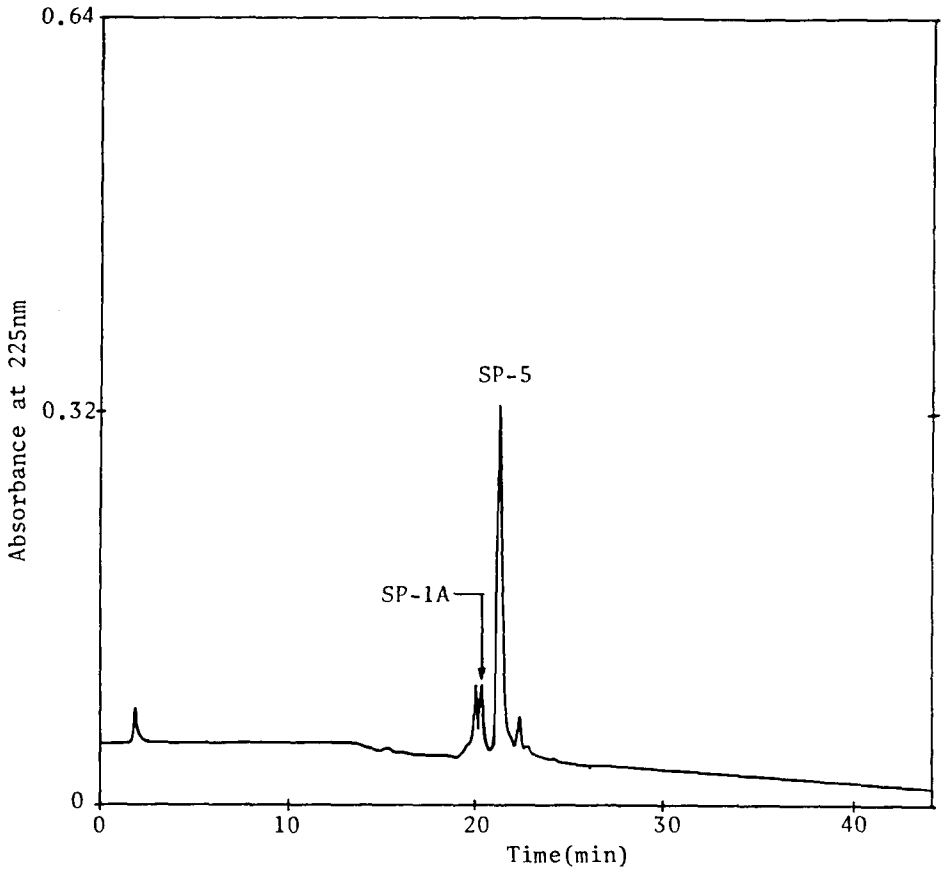

Fig. 3-6 Rechromatography of fraction SP-A by HPLC. The sample was dissolved in 100 µl of 0.1 M ammonium formate and chromatographed on a Baker bond wide pore butyl column (0.46 x 25 cm, C4, 330 A) at a flow rate of 1.0 ml/min.

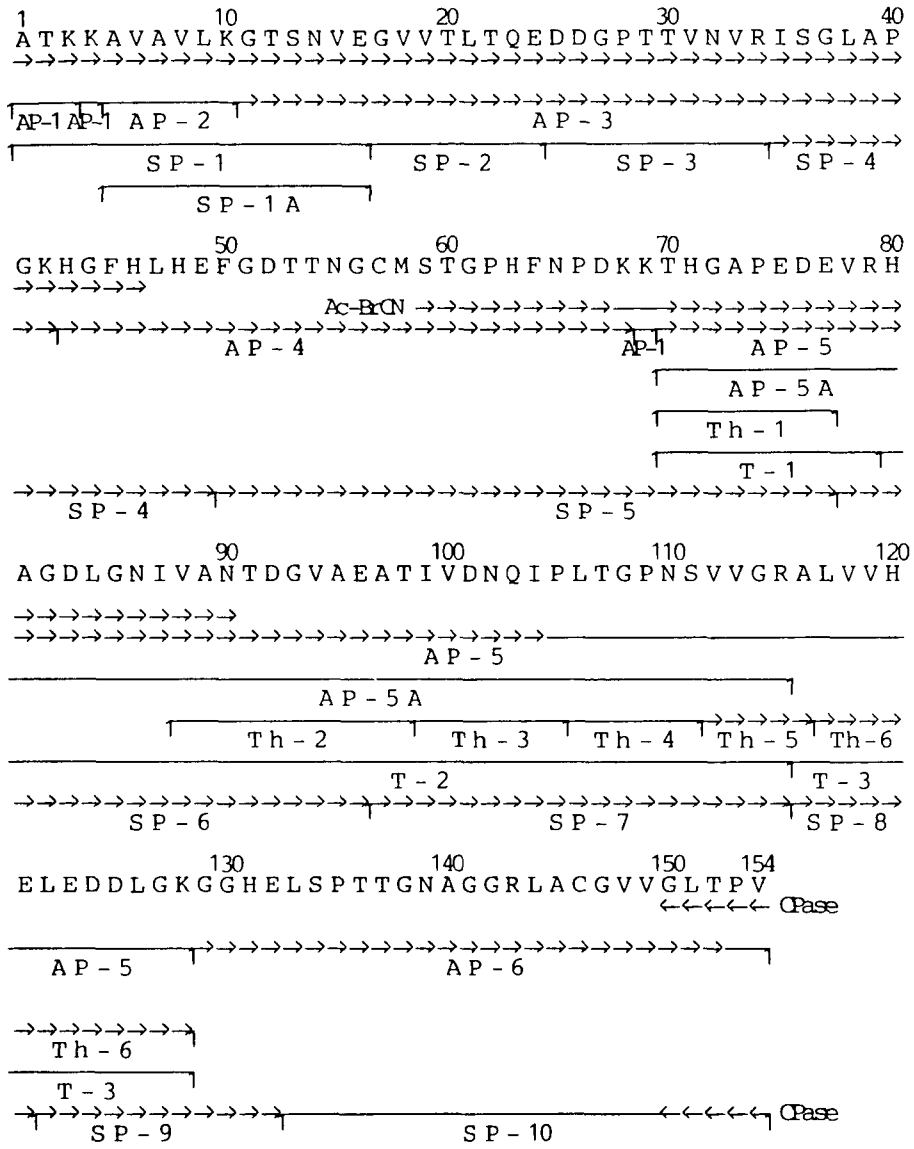

Fig. 3-7 Amino acid sequence of spinach SOD. Residues arrowed to the right and the left were identified by Edman degradation and by carboxypeptidase Y digestion, respectively. Ac-BrCN denotes the protein treated with BrCN after the acetylation of the amino groups. AP, SP, Th and T denote peptides obtained by proteolysis with *A. lyticus* lysyl*endo*peptidase, *S. aureus V8* protease, thermolysin, and trypsin, respectively. The solid lines denote the portions of the peptides whose amino acid compositions were determined but sequences were not analyzed.

3-3 Discussions

The complete amino acid sequence of the subunit of Cu,Zn-SOD from spinach leaves has been determined on the basis of peptides obtained by BrCN cleavage and by proteolysis with *A. lyticus* lysyl*endo*peptidase, *S. aureus V8* protease, trypsin, and thermolysin. The subunit consists of a total of 154 amino acid residues with an unblocked N-terminal alanine residue and a C-terminal valine residue. Therefore spinach SOD, composed of two identical subunits, has a molecular weight of 31,362 except for the two kinds of metal ions. This is the first sequence determination of a plant Cu,Zn-SOD.

The direct amino acid sequencing of the acetylated spinach SOD followed by BrCN cleavage was successful in this case. With this and the direct sequencing of the intact protein, almost half of the amino acids (76 of 154 residues) were identified.

In the digestion of performic acid-oxidized spinach SOD with *A. lyticus* lysyl*endo*peptidase, the Arg-115 - Ala-116 bond was partly hydrolyzed in addition to the cleavage of all of the seven lysyl bonds. Cleavage of the Arg-Ala bond with the same protease was also reported for salivary acidic protein-1 (SAP-1) from human salivary glands (*26*). With *S. aureus V8* protease, the quantitative cleavage of two arginyl bonds (Arg-34 - Ile-35 and Arg-115 - Ala-116) and the partial cleavage of a lysyl bond (Lys-4 - Ala-5) occurred in the performic acid-oxidized protein. This

trypsin-like activity may be caused by either impurities of the commercially available *S. aureus V8* protease or the intrinsic nature of this protease.

The discrepancy between the sequence and the amino acid composition in the values of valine and isoleucine for several peptides is probably due to the incomplete hydrolysis of Val-Val and Ile-Val bonds in their peptide chains. In amino acid analysis, glycine in Th-5 and Th-6 and alanine in SP-5 were over-estimated (0.5-0.9 mol). We suspect that a trace of these amino acids has been contaminated in the sample for analysis during manipulation and determined on analysis at picomole levels. The low recovery of histidine in SP-5 may result from partial degradation during performic acid oxidation. Eventually the sequences of Th-5, Th-6, and SP-5 are judged to be valid from the sequences unambiguously determined for any related peptides (Ac-BrCN, AP-4, AP-5, SP-7 and SP-8), which are obtained by different fragmentation.

4 Tertiary structure determination

4-1 Intensity data collection and data processing

4-1-1 Crystals

Crystallized spinach SOD was dissolved in 10 mM potassium phosphate buffer (pH 7.8) and recrystallized in order to obtain a large crystal with good quality for three-dimensional structure analysis. The recrystallization was carried out with a gradual addition of fine ammonium sulfate powder to the solution up to about 55% saturation. After several weeks, a faint green rhombic crystal was grown to the size of 1.0 x 0.3 x 0.2 mm^3 and stored in 70% saturated ammonium sulfate solution containing 10mM potassium phosphate buffer (pH 7.8). The crystal belongs to space group $C2$, with unit cell dimensions a=166.27 Å, b=45.97 Å, c=85.68 Å, β=99.38° (27-28). It contains two dimeric molecules (Mr=32,000) or four identical subunits per asymmetric unit.

4-1-2 Heavy atom derivatives

In order to analyze the crystal structure of proteins, it is necessary to obtain heavy atom derivative crystals which are useful to determine the phases of structure-factor amplitudes. Heavy atom derivatives were prepared by the conventional soaking method. Many kinds of heavy atom reagents were tested by varying the soaking conditions such as the concentration and soaking time, then two useful

heavy atom derivatives ($K_3UO_2F_5$, $K_2Pt(NO_2)_4$) were obtained. Whether the reagent bound to the protein or not, it was examined by comparing their diffraction patterns of (h 0 0), (0 k 0) and (0 0 l) zones with those of the native crystal. The soaking conditions of the two heavy atom derivatives are summarized in Table 4-1.

4-1-3 Data collection with four circle diffractometer

Three-dimensional intensity data were collected at 14°C on a Rigaku four circle diffractometer equipped with a rotating anode X-ray generator operated at 40kV-300mA, using Ni-filtered Cu-Kα radiation. A beam collimator and a beam tunnel in the receiving path were evacuated to reduce air scattering. Each crystal was mounted in a walled glass capillary tube of diameter 0.7mm to 1.0mm with a small quantity of mother liquor for stabilization so that the crystallographic a^* axis was parallel to the spindle axis. The ω-scan method was employed with the scan speed of 2 °/min to 4 °/min, and the scan width was $(0.7+0.15\tan\theta)^\circ$ to $(1.0+0.15\tan\theta)^\circ$. The background intensities were measured for 2.5 sec to 5.0 sec at each end of the scan. These measurement conditions were variable for each crystal, respectively. Three standard reflections were measured every 100 reflections to check radiation damage and crystal slipping. Twenty-five native crystals were used to get intensity data up to 2.0 Å resolution. Ten uranyl derivative crystals and six platinum derivative crystals were used to get intensity

Table 4-1

Soaking conditions

Reagent	Conc.(mM)	Buffer	Time(days)
Native	-	10mM Potassium Phosphate pH 7.8 70% Saturated Ammonium Sulfate	-
$K_3UO_2F_5$	30	Water 70% Saturated Ammonium Sulfate	3
$K_2Pt(NO_2)_4$	3	10mM Potassium Phosphate pH 7.8 70% Saturated Ammonium Sulfate	3

data including Bijvoet-pair reflections up to 2.8 Å resolution, respectively. The intensity decays of the standard reflections were about 20% for each crystal at the end of the measurement.

4-1-4 Data corrections and scalings

The original intensity data collected must be revised in the causes of absorption difference and radiation damage. The North-Phillips method (*29*) was applied to the absorption correction. Radiation damage was corrected by $I_{corr}=I_{obs}\exp(aN)$, where I_{corr} is the corrected intensity, I_{obs} is the observed intensity, a is the damage coefficient calculated from the decays of standard reflections and N is the number of order of measurement.

The intensity data from different crystals were scaled and merged with each other into a single set of data. Scaling of each data for the native, the uranyl derivative and the platinum derivative was carried out by the method of Sparks-Rollet (*30*). The three kinds of data sets were obtained, and the intensity data measurements are summarized in Table 4-2.

4-2 Structure determination

4-2-1 Multiple isomorphous replacement (M.I.R.)

Three-dimensional difference Patterson with coefficient $(F_{pH} - F_p)^2$ between the uranyl derivative and

Table 4-2

Summary of the intensity data

Crystals	R_{merg}	Resolution	N_{total}	N_{indep}	N_{cryst}
Native	0.048	2.0 Å	149090	43898	25
$K_3UO_2F_5$	0.043	2.8 Å	54249	30613[a]	10
$K_2Pt(NO_2)_4$	0.029	2.8 Å	34814	16093	6

R_{merg} $\Sigma|F_i-\langle F\rangle|/\Sigma\langle F\rangle$, for all merged reflections.

N_{total} No. of total measured reflections.

N_{indep} No. of independent processed reflections.

N_{cryst} No. of used crystals for data collection.

(a) Including Bijvoet-related reflections.

the native data at 5 Å resolution was calculated to locate the heavy atom in the crystal. Apart from the origin, the highest peak in the Harker section (Y=0) of the map (480/1000, 0/1000, 640/1000) was thought the vector between uranyl ions. The major binding site of uranyl ion in an asymmetric unit was found in this way. However, the difference Patterson map between the platinum derivative and the native data at the same resolution gave too many peaks to be interpreted uniquely in the same section. Another uranyl binding site in the U-derivative and platinum binding sites in the Pt-derivative were obtained by difference Fourier maps with phase angles calculated from the major uranyl binding site, coupled with taking care of the anomalous data. These sites were all consistent with the peaks found in the corresponding difference Patterson maps between each heavy atom derivative data and the native data.

Phase angles of each reflection were calculated by the method of Blow and Crick (*31*) considering the anomalous-dispersion effect of U-derivative. Heavy atom parameters such as atomic coordinates, occupancies and isotropic temperature factors were refined by a least-squares technique at 5 Å resolution at first. An electron density map obtained at this resolution showed the molecular boundary and the positions of the two kinds of metal ions (Cu^{2+} and Zn^{2+}) originally contained in the SOD molecule clearly. The analysis was then extended to 2.8 Å resolution in the same way and the refinement of the heavy

atom parameters was continued using about 16,000 reflections up to 2.8 Å resolution until convergence. The heavy atom parameters and the statistics for the two derivatives at 2.8 Å resolution are shown in Table 4-3, and the amplitudes *versus* their lack of closure errors are shown in Fig. 4-1. An electron density map at 2.8 Å resolution was calculated using the best phases with the mean figure of merit 0.67. The map is so clear enough to be followed easily along the polypeptide main chain. The basic structure of the spinach SOD was found to be β barrel at this stage. However, the density around the major binding site of uranyl ion was not interpretable because of the residual electron density due to the uranyl ion.

4-2-2 Molecular replacement (M.R.)

As the spinach SOD crystal had four identical subunits in an asymmetric unit, molecular replacement (M.R.) method by averaging the electron density over the four subunits was applicable to enhance the signal-to-noise ratio of the protein density (*32*). The four subunits were conveniently named Yellow (Y), Violet (V), Green (G) and Blue (B), respectively. The flow chart of this method is briefly shown in Fig. 4-2, and an example of the application of this method was previously reported by Kusunoki (*33*).

In order to refine the phase angles by this method, it is first necessary to determine an accurate molecular envelope. The initial envelope was defined from the M.I.R. map at 2.8 Å resolution, after several checks of

Table 4-3

Heavy atom parameters at 2.8 Å resolution

Heavy atoms	x [a]	y [a]	z [a]	G [b] (e)	B [c] (Å^2)
$K_3UO_2F_5$-1	0.236	0.000	0.321	43.3	11.0
$K_3UO_2F_5$-2	0.227	0.190	0.033	12.0	18.3
$K_2Pt(NO_2)_4$-1	0.104	0.040	0.401	7.9	20.0
$K_2Pt(NO_2)_4$-2	0.500	0.231	0.918	10.0	20.0
$K_2Pt(NO_2)_4$-3	0.300	0.400	0.900	8.2	15.3
$K_2Pt(NO_2)_4$-4	0.367	0.227	0.313	10.8	19.8

a Fractional coordinates along the crystallographic *a*, *b* and *c* axes, respectively.

b Site occupancy on an approximately absolute scale.

c Isotropic temperature factor.

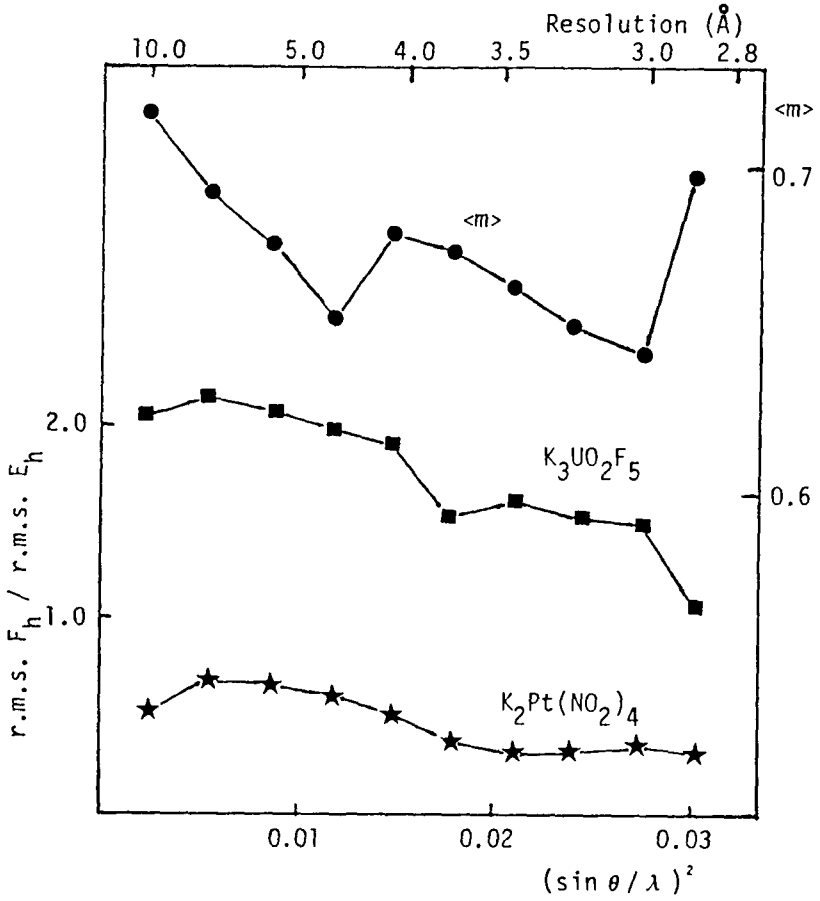

Fig. 4-1 Variation of the ratio of r.m.s. F_h to r.m.s. E_h and the mean figure of merit as a function of $(\sin\theta/\lambda)^2$.

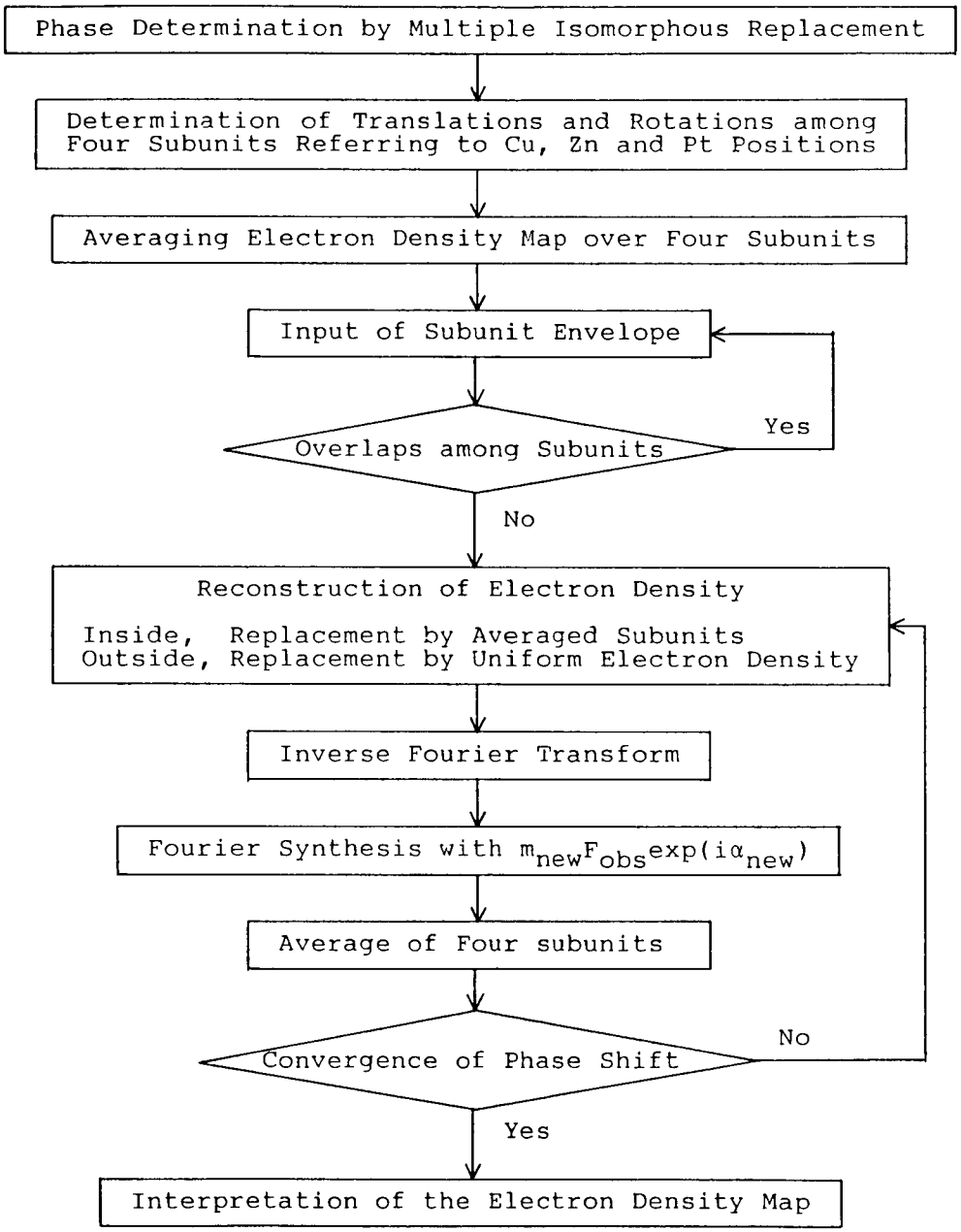

Fig. 4-2 A schematic representation of the molecular replacement method.

overlaps between adjacent subunit densities.

Generally, some identical subunits may only differ in orientation and position. It is therefore possible to define an operation involving only rotation and translation. In the present case, Cu^{2+}, Zn^{2+} and one of the heavy atom derivatives, $K_2Pt(NO_2)_4$ bound to the SOD subunit with the ratio of 1:1, respectively. The operation was obtained in such a way that these three positions in each subunit were brought into coincidence. The rotation and translation matrices in this operation were refined by least-squares method and used to search best fit positions of the electron density among the four subunits by the scans of translation and rotation. The translation scan was performed in steps of the grid spacings of Fourier map (320*80*160), that is 0.5 Å steps, and the rotation in 5^{o} steps. Rotation and translation parameters were finally refined by the least-squares technique to obtain the maximal correlation among the electron densities of the four subunits. Table 4-4 summarizes these rotation and translation parameters as the V, G and B subunits were fitted to the Y subunit. Y and V subunits are found to be related by an approximate 2-fold axis because the χ angle between them is nearly 180^{o}.

After five cycles of refinement, the molecular envelope was checked and improved. Further five cycles of refinement were carried out using the diffraction data up to 2.5 Å resolution. The density of the solvent were set to zero for every cycle of the refinement. The mean phase

Table 4-4

Rotation and translation parameters between subunits

parameters	Y-V	Y-G	Y-B
θ	62.57	105.87	99.27
ϕ	85.23	-14.54	15.46
χ	180.56	142.97	160.85
x	0.5238	0.1301	0.2705
y	0.1405	0.1111	0.3003
z	0.7430	0.9732	0.2275

Y, V, G and B denote Yellow, Violet, Green and Blue subunits, respectively. θ, ϕ, χ are rotation angles in polar coordinates with the unit of degree. x, y and z are translational shifts of fractional coordinates.

shift between the last two cycles was 2.3° and the correlation coefficient between the observed and the calculated structure factors at the last stage was 0.83. The final R-factor was 27%. Fig. 4-3 indicates the changes of each parameter during the refinement of molecular replacement. The electron density map calculated with the refined phase angles was superior to the initial M.I.R. map (Fig. 4-4).

4-2-3 Model building and atomic coordinates

The electron density map of the spinach SOD at 2.5 Å resolution after the phase refinement using M.R. method is good enough to trace a polypeptide main chain and to identify some of the residues. Four crystallographically independent subunits in an asymmetric unit, known to be chemically identical, appear to have almost the same conformations in the map. Two subunits which constructs a molecule is related by an approximate non-crystallographic 2-fold axis. On the basis of the amino acid sequence (34), a molecular model of the protein was built with a scale of 2 cm/Å, comparing with the electron density map at 2.5 Å resolution using a Richards' optical comparator (35) (Fig. 4-5). Three-dimensional atomic coordinates obtained from the Richards' box served as the starting model for a refinement by the method of Hendrickson and Konnert (36).

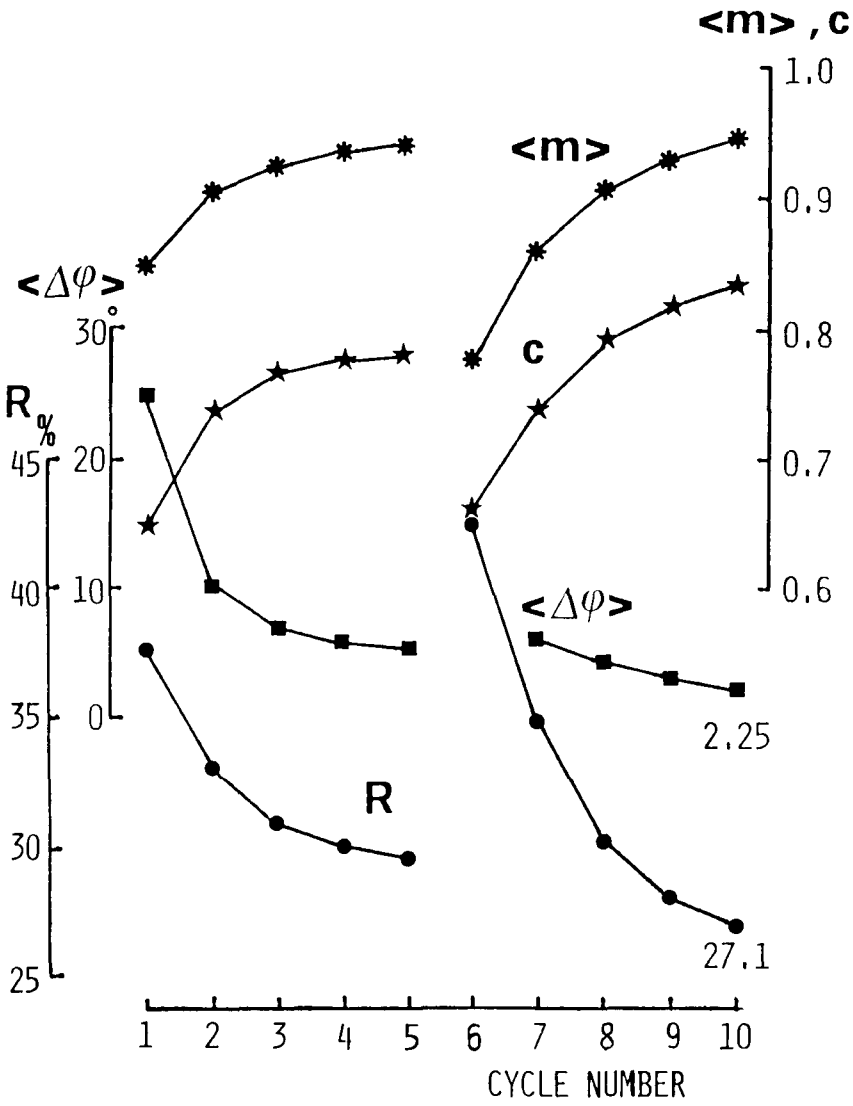

Fig. 4-3 Overall progress in molecular replacement. ⟨Δφ⟩ is the mean phase shift. c is the correlation coefficient between the observed and the calculated structure factors. R is the crystallographic residual between the observed and the calculated structure factors. ⟨m⟩ is the average of figure of merit.

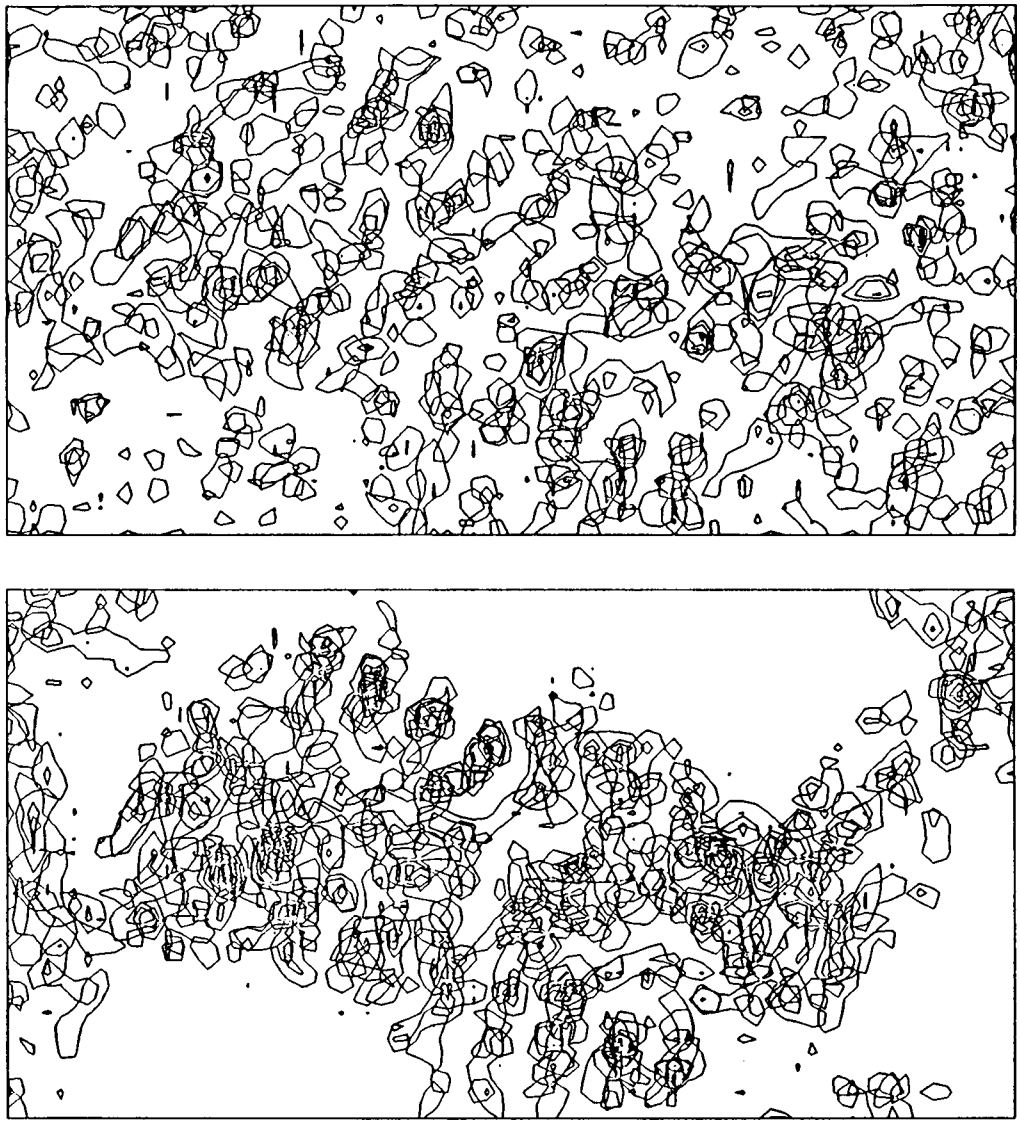

Fig. 4-4 The development of the electron density map using molecular replacement method. Upper is the original M.I.R. map at 2.8 Å resolution. Lower is the map after molecular replacement averaging at 2.5 Å resolution. Both of the map are viewed down along the 2-fold axis between the subunits.

Fig. 4-5 The Kendrew-type molecular model of the spinach SOD.

4-3 Distinction between Cu^{2+} and Zn^{2+} ions by crystallographic method

4-3-1 Theoretical background

In the 2.5 Å electron density map, eight prominent peaks in an asymmetric unit could be easily assigned to the metal ions (Cu^{2+} and Zn^{2+}). The eight peaks were grouped into four pairs, and the peaks in each pair were apart from each other by only 6.1 Å, respectively. One of the peaks in the pair was coordinated by four ligands with a distorted square planar arrangement, and the other was with a tetrahedral arrangement. From the chemical viewpoint, this difference in the coordination geometry may suggest that the former should be Cu^{2+} ion and the latter Zn^{2+} ion. In the X-ray analysis, however, it is impossible to distinguish between Cu^{2+} and Zn^{2+} ions from each other by inspecting the electron density map calculated with the usual method, because the difference in the number of electrons is only one between them, and their scattering powers must be almost the same.

Synchrotron radiation is being increasingly utilized for data collection of protein crystals owing to its brightness, parallelism and tuneability. It is worthwhile noting that anomalous-dispersion effects around specific absorption edges of bound metal ions can be optimized by finely tuning the synchrotron X-ray wavelength. In the case of proteins with more than one kind of metal cofactor, differences in both of anomalous components f' and f" can

be used to distinguish atoms or ions with similar atomic numbers. Such differences have already been utilized to locate manganese and calcium cofactors, which have a five-electron difference, in a pea lectin crystal (*37*). In the course of the structure analysis of spinach SOD, the difference in the anomalous-dispersion coefficients of Cu^{2+} (27 electrons) and Zn^{2+} (28 electrons) has been exploited in the similar way. Synchrotron radiation from the Photon Factory vertical wiggler (*38*) was used to locate the Cu^{2+} and Zn^{2+} ions unambiguously enhancing both the real coefficients f' and the imaginary coefficients f" of each ion. Wavelengths for data collection were selected close to the Cu K edge at 1.380 Å and the Zn K edge at 1.283 Å (*39*), because f' and f" values of the Cu^{2+} and Zn^{2+} ions show a pronounced change near the wavelength of their K absorption edge. The variation of f' and f" of Cu and Zn with wavelength is shown in Fig. 4-6.

In the present study, the anomalous difference Fourier maps which mainly reflect the f" effect of a specific atom or ion are compared with each other in order to locate metal ions and to distinguish between the Cu^{2+} ion and the Zn^{2+} ion in the SOD molecule. The difference Fourier maps of the different data sets, which reflect the f' effect, are also examined; the results are in agreement with those of the anomalous difference Fourier calculations.

4-3-2 Experiments

Five sets of intensity data of Bijvoet-pair

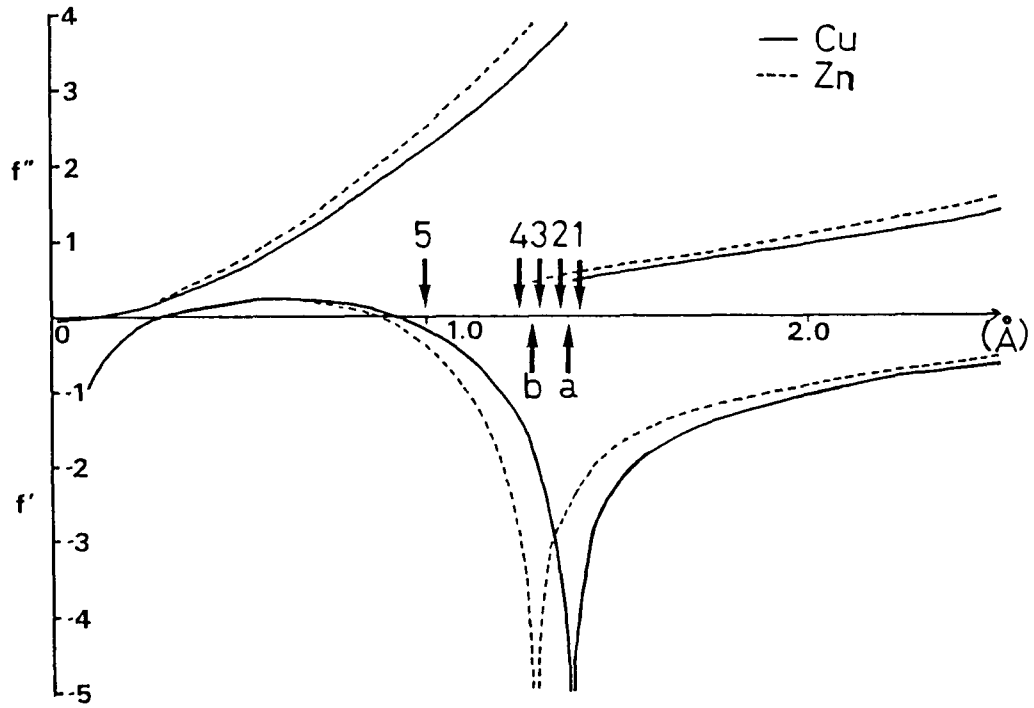

Fig. 4-6 Variation of f' and f" of Cu and Zn with wavelength. The arrows with the numbers 1 to 5 indicate the wavelengths at which data were collected, and the arrows with the characters a and b are the absorption edges of Cu and Zn, respectively.

reflections (F^+ and F^-) were measured up to 6 Å resolution at five different wavelengths where these data sets represent various anomalous-scattering effects of Cu^{2+} and Zn^{2+} ions, respectively. The five wavelengths were 1.000, 1.275, 1.299, 1.373 and 1.389 Å, and the contributions of f' and f" for each metal ion are summarized in Table 4-5 (*40*). The data set at λ = 1.000 Å is useful for the calculation of the difference Fourier maps of the other four data sets, because the f' effect at this wavelength is rather small compared with those at the other four wavelengths.

All the data were collected with a four-circle diffractometer using synchrotron radiation at the Photon Factory. The diffractometer was located on the station of the beam line 14A; it has a horizontal-type setup and utilizes the radiation from a superconducting vertical wiggler (*41*). The white X-ray beam, which has a higher degree of polarization in the vertical direction, is monochromatized by a Si(111) double-crystal monochromator and then focused by a Pt-coated fused-quartz mirror (*41*). Wavelength calibrations were made by measuring the absorption spectra of Cu and Zn foils at their K absorption edges. The intensity of the X-ray beam incident to the specimen was monitored by an ion chamber for every measurement in order to correct for intensity change (*41*). The raw data were corrected for Lorentz and polarization factors. The polarization factor was calculated, assuming that the X-ray beam incident to the monochromator was

Table 4-5

Anomalous scattering components of Cu and Zn

Data Set	f'		f"	
	Cu	Zn	Cu	Zn
1.000 Å	-0.1	-0.4	2.3	2.6
1.275 Å	-1.7	-4.5	3.4	3.8
1.299 Å	-2.0	-4.1	3.5	0.5
1.373 Å	-4.8	-2.4	3.9	0.6
1.389 Å	-4.8	-2.3	0.5	0.6

polarized 90% in the vertical direction. Damage and absorption corrections were carried out by the same method described in the section 4-1-3.

In the data set at λ = 1.275 Å, both Cu^{2+} and Zn^{2+} ions have high f" contributions (3.4, 3.8). Anomalous-dispersion effects due to both ions may be clearly observed in this wavelength.

The data sets, at λ = 1.299 and 1.373 Å, were measured to estimate anomalous-dispersion effect of only the Cu^{2+} ion, because these wavelengths are positioned between the Cu and Zn K absorption edges, and have high f" values (3.5, 3.9) for Cu^{2+} ion, and low (0.5, 0.6) for Zn^{2+} ion.

The data set, at λ = 1.389 Å, reflects low f" contributions of Cu^{2+} and Zn^{2+} ions (f" = 0.5, 0.6). These f" are comparable to those of the other light atoms such as carbon, nitrogen, or oxygen in the SOD molecule, because this wavelength is longer than those of the K absorption edges of both metal ions.

The data set, at λ = 1.000 Å, is accompanied with smaller anomalous-dispersion effects of f' and f" due to both metal ions than those of the 1.275 Å data set. This data set is useful for the calculation of difference Fourier map to the other four data sets, because the f' effect at λ = 1.000 Å is rather small compared with those at the other four wavelengths. On the basis of the difference of f' values between them, the difference Fourier map is expected to give meaningful peaks at the metal-ion positions.

4-3-3 Results

As a first approach, an anomalous difference Fourier calculation was carried out to distinguish the two kinds of metal ions from each other. The Fourier syntheses for five different data sets obtained at λ = 1.000 Å, 1.275 Å, 1.299 Å, 1.373 Å, and 1.389 Å were calculated by the method described by Kraut (42). The phases of each reflection in the calculation were derived from the M.R. method described in the section 4-2-2. The intensity differences in the structure-factor amplitudes between Bijvoet-pair reflections are primarily due to both Cu^{2+} and Zn^{2+} at λ = 1.275 Å, and mainly to Cu^{2+} at λ = 1.299 Å and 1.373 Å (Table 4-5). At λ = 1.389 Å, the difference in the amplitude should be negligibly small.

For 1.275 Å data set, the positions of the eight prominent peaks in the asymmetric unit were identical to those of the two kinds of metal ions in the original native Fourier map, though the four peaks at the possible Zn^{2+} ion sites were slightly higher than the other four at the possible Cu^{2+} ion sites, as shown in Table 4-6. This is consistent with the fact that the anomalous-dispersion component f" of a Zn^{2+} ion is larger than that of a Cu^{2+} ion by 0.4 as shown in Table 4-5. On this Fourier map, only the anomalous scatterers can be located and in fact only four Cu^{2+} and four Zn^{2+} ions in the asymmetric unit of the SOD crystal showed anomalous-scattering effects at this wavelength, because the eight prominent peaks can be completely assigned to the Cu^{2+} and Zn^{2+} ion positions of

Table 4-6

Electron density values at the Cu^{2+} and the Zn^{2+} positions in the Fourier maps

Data Set	Yellow		Violet		Green		Blue		HUP
	Cu	Zn	Cu	Zn	Cu	Zn	Cu	Zn	
N-2.8	26	37	25	29	34	29	30	40	-
N-2.5	66	86	65	93	67	94	63	92	-
1.000 Å [a,b]	30	90	62	59	90	75	32	74	123
1.275 Å [a]	103	126	73	149	83	154	108	115	70
1.299 Å [a]	101	35	63	30	109	18	96	40	73
1.373 Å [a]	113	7	134	41	139	0	124	18	80
1.389 Å [a]	20	9	25	54	27	17	53	19	77
1.275 Å [c]	14	1	8	34	11	41	24	4	86
1.299 Å [c]	12	0	0	28	17	29	8	0	83
1.373 Å [c]	56	0	32	17	68	22	61	0	81
1.389 Å [c]	30	0	21	22	66	27	49	0	75

Yellow, violet, green and blue denote four subunits in an asymmetric unit. N-2.8 and N-2.5 indicate the original native Fourier maps of 2.8 Å resolution with M.I.R. phases and 2.5 Å resolution with M.R. phases. HUP is the highest uninterpretable peak in each map. The value zero includes the density below zero.

a) Anomalous difference Fourier maps.

b) This anomalous difference Fourier map was calculated to investigate the signal-to-noise ratio.

c) Difference Fourier maps with the 1.000 Å data set.

four subunits in the asymmetric unit.

For the 1.299 Å and 1.373 Å data sets, anomalous difference Fourier maps showed almost the same electron density distributions around the eight metal-ion positions. They gave high electron density humps at four possible Cu^{2+} ion sites and almost flat density at four Zn^{2+} ion sites (Fig. 4-7, Table 4-6). This is consistent with the fact that the anomalous-dispersion component f" of a Cu^{2+} ion is about seven times larger than that of a Zn^{2+} ion at these two wavelengths (Table 4-5). This result implies that the Cu^{2+} ion has been able to distinguish from the Zn^{2+} ion in the present study and that the location of two kinds of metal ions in this analysis was identical to that expected from the chemical coordination. Therefore, the other position, which was expected as the Zn^{2+} ion site on the basis of the chemical coordination, was found to be really the Zn^{2+} ion position.

For the 1.389 Å data set at which there is little f" contribution of Cu^{2+} and Zn^{2+} ions, no significant electron density humps were found in the anomalous difference Fourier map around the previous eight positions of metal ions. The highest peak in an asymmetric unit could not be assigned to the metal-ion positions. The ratio of the peak density at the possible metal-ion position to the highest uninterpretable peak is lower than in the other three cases (Table 4-6).

As a second approach, the conventional difference Fourier calculation between the 1.000 Å data set and each

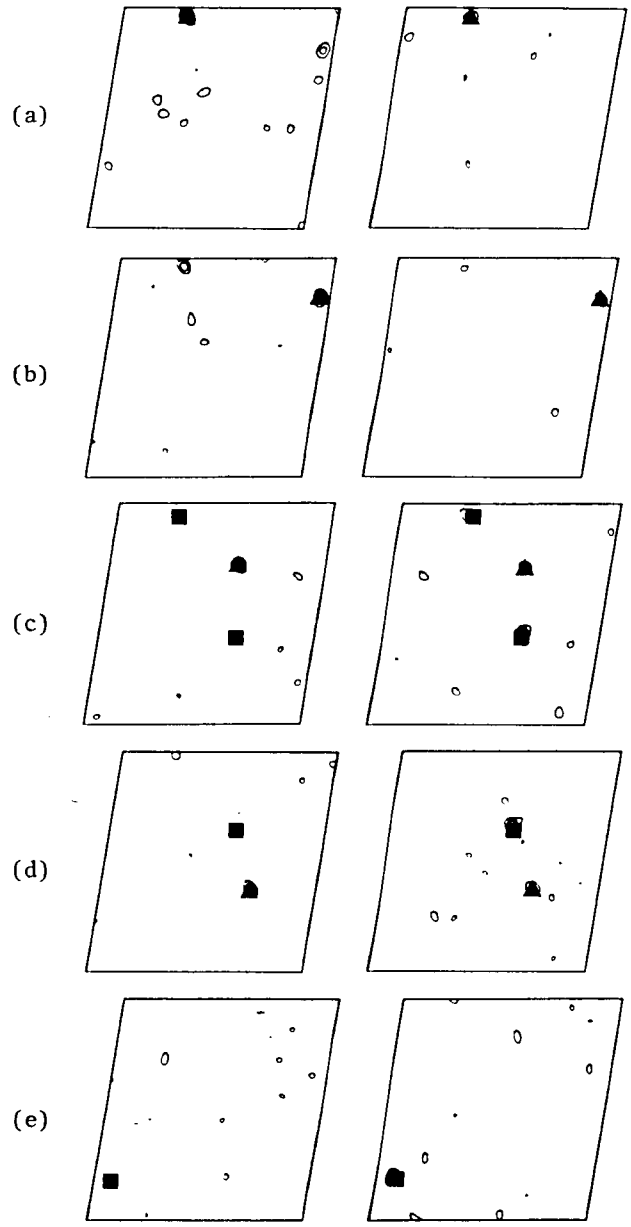

Fig. 4-7 Anomalous difference Fourier maps: left-hand side, 1.373 Å data set; right-hand side, 1.275 Å data set. The triangles are four Cu^{2+} ion positions and the squares are four Zn^{2+} ion positions in an asymmetric unit. The sections are (*a*) 4/50, (*b*) 6/50, (*c*) 13/50, (*d*) 17/50, and (*e*) 24/50 of the crystallographic *b* axis. At 4/50 and 6/50, the same peaks appear on the left-hand map. Contours are in equal intervals of positive density and are chosen to show the highest peak in each map with four contour levels. The edges of the maps correspond to half the unit cell in *a* along the horizontal direction and the full unit cell in *c* along the vertical direction.

data set of the other four wavelengths (1.275 Å, 1.299 Å, 1.373 Å, 1.389 Å) was carried out with the same phase angles used in the previous approach. The Fourier coefficients are obtained by averaging the amplitudes of Bijvoet-pair reflections (F^+ and F^-) to take off the f" contributions. The scaling factors between the different data sets were calculated by Wilson's statistics (*43*). This Fourier map reflects the differences of the amplitudes attributed to f', although anomalous difference Fourier synthesis mainly does those derived from f" contribution. The results are summarized in Table 4-6. For the 1.373 Å and 1.389 Å data sets near the Cu K absorption edge, the difference maps gave significantly higher peaks at possible Cu^{2+} ion positions than those at possible Zn^{2+} ion positions, except for one case, the violet subunit in the 1.389 Å data set (see Table 4-6). On the other hand, for the 1.275 Å and 1.299 Å data sets near the Zn K absorption edge, the difference maps did not consistenly give higher peaks at possible Zn^{2+} ion positions than those at possible Cu^{2+} ion positions. Each difference map was quite noisy compared with the anomalous difference Fourier map, and the highest peak in each map corresponds not to the position of metal ion but to an uninterpretable peak. The quality of the difference Fourier map may suffer from errors in scaling between different data sets and in the different absorption effect from different crystals.

As a third approach to confirm the accuracy of the metal-ion positions, an anomalous difference Patterson

synthesis was performed for each data set. This synthesis did not give any significant peaks which could be interpreted as Cu^{2+} or Zn^{2+} ions.

4-3-4 Discussions

Until now, distinction between two or more kinds of metal ions with similar atomic numbers has been dependent upon consideration of differences in the chemical nature of the ions. In the present study of spinach SOD, however, the distinction between Cu^{2+} and Zn^{2+} ions which have only a one-electron difference has been achieved by using anomalous difference Fourier techniques based on reliable phase angles from M.I.R. and M.R. methods, and intense tuneable synchrotron radiation. In particular, the anomalous difference Fourier maps of the 1.275 , 1.299 and 1.373 Å data sets are very useful in distinguishing the ions from each other. On the map of the 1.275 Å data set, both the Cu^{2+} and Zn^{2+} ions appeared as the higher electron density humps, while only the Cu^{2+} ions were visible on the map of the 1.299 and 1.373 Å data sets. The map of the 1.373 Å data set gave a higher signal-to-noise ratio than that of the 1.299 Å data set. This is consistent with the fact that the former has a little higher contribution of f' and f" than the latter, as shown in Table 4-5. It should be mentioned that the refined phase angles of each reflection play important roles in producing these kinds of results. As a result, the location of the Cu^{2+} and the Zn^{2+} ions which could be speculated only on the view point

of chemical coordinations has been exactly ascertained with the crystallographic methods. In contrast to the success of the Fourier syntheses, an anomalous difference Patterson synthesis failed to locate the metal ions. Patterson synthesis may therefore be less powerful than anomalous difference Fourier synthesis in the case of a protein crystallographic study, as was the case in another report (*44*).

4-4 Structure refinement

4-4-1 Method of Hendrickson and Konnert

The stereochemically restrained least-squares procedure of Hendrickson and Konnert (*36*) was used to refine the atomic coordinates of spinach SOD.

In this refinement, the least-squares solution provides the shifts in the set of model parameters that minimize the sum of the weighted squares of differences between the observed and calculated values for both structure factors and stereochemical parameters. The function which is minimized is

$$\begin{aligned}\Phi = &\sum_{i=1}^{NREF} W_i(F_{Oi} - F_{Ci})^2 + \sum_{i=1}^{NBOND} W_i(D_{Oi} - D_{Ii})^2 \\ &+ \sum_{i=1}^{NANG} W_i(D_{Oi} - D_{Ii})^2 + \sum_{i=1}^{NPLN} W_i(D_{Oi} - D_{Ii})^2 \\ &+ \sum_{i=1}^{NCHIRAL} W_i(V_{Oi} - V_{Ii})^2 + \sum_{i=1}^{NVDW} W_i(D_{Oi} - D_{Ii})^4 \\ &+ \sum_{i=1}^{NTHERM} W_i(B_{Oi} - B_{Ti})^2 + \sum_{i=1}^{NCRYP} W_i \sum_{j=1}^{4}(|\vec{r}_{i,j} - \vec{r}_i|)^2 \\ &+ \sum_{i=1}^{NCRYT} W_i \sum_{j=1}^{4}(B_{i,j} - \langle B_i \rangle)^2\end{aligned}$$

where the first term is the diffraction term and is the function usually minimized in the refinement of small molecules. The second and the third terms minimize the deviation of bond distances and angle-defining distances from their ideal values. The forth term minimizes the deviation from planarity of groups of atoms expected to be planar. The fifth term minimizes the deviations from the ideal chiral volumes (sign included) required for given chiral centers, thus ensuring the proper handedness. The sixth term prevents unreasonably short non-bonded contacts by minimizing the deviations from ideal contact distances; it is applied only to those contact distances which are deemed too short. The seventh term minimizes the unreasonable discrepancies between the thermal parameters of the neighboring atoms which are covalently bonded. This term is known as a simple model of "riding motion". The last two terms are useful for non-crystallographic symmetry restraints. The eighth term minimizes the deviations from the average structure over the four identical subunits in the asymmetric unit. The last term imposes the restraint of thermal parameters for these similar environments. All quantities subscripted with "I" are ideal values for the parameters as determined from crystal structures of small molecules. Quantities subscripted with "O" refer to the current values of the parameters in the trial model. (except for two terms, for the first term: "O" and "C" denote observed and calculated structure factors, and for the seventh term: "O" and "T"

denote origin and target temperature factors.) "W_i" denotes the weighting factor. It is obvious that the minimization of Φ with the proper weights should produce a structure that agrees well with the observed data and is consistent with almost everything known about protein structural chemistry.

4-4-2 Refinement survey

The refinement was done in four stages, increasing the resolution limit stepwise from 3.0 Å to 2.0 Å, while making appropriate changes in the restraints. Metal to ligand bonds were all restrained with the same weight as bond distances (sigma equal to 0.03 Å). Cu-His ligand bond lengths were restrained to 2.00 Å. Zn ligand bond-length restraints were 2.08 Å for the three His ligands and 2.05 Å for OD1 of Asp-83. Refinement survey were summarized in Table 4-7.

An initial R value of the model of 4420 atoms were 41.9% for 12,125 reflections beyond 2σ between 10 Å and 3 Å spacings, after the initial cycle of a stereochemical idealization of the starting model. After 18 cycles of the refinement during which the weight of non-crystallographic restraints were relaxed from 0.05 Å to 0.50 Å, the R value became 32.3%. 2Fo-Fc electron density map was produced using the calculated phases. The densities were averaged over the four subunits in an asymmetric unit, and the model was corrected manually to reduce discrepancies with the electron density map by using

Table 4-7

Refinement survey

Stage	1	2	3	4
No. of cycles	17	31	41	16
No. of parameters	13262	13262	17681	18097
No. of atoms (solvents)	4420 (0)	4420 (0)	4420 (0)	4524 (104)
Resolution (Å)	6.0-3.0	5.0-2.0	5.0-2.0	5.0-2.0
No. of F-data	10757	14295	28682	28682
$F_{obs}/\sigma(F_{obs})$[a]	2σ	7σ	3σ	3σ
$\langle F_{obs}-F_{calc} \rangle$	1.62	1.43	0.95	0.91
$\langle B \rangle$ (Å^2)	6.4	9.0	9.5	9.7
R (initial) (%)	41.9	35.9	31.4	26.1
R (final) (%)	32.3	30.0	26.0	24.9

a The F-data greater than the listed values were used for the refinement

a graphic display system. The corrected coordinates were rotated and translated to fit other three subunits in an asymmetric unit using matrices.

In the second stage, the four coordinates sets were further refined by another 31 cycles of the refinement during which the resolution limit was increased from 3 Å to 2 Å. 2Fo-Fc density map was produced as the same way and manual correction was done again. The corrections for four subunits (Y, V, G, B) were carried out individually this time. These manual corrections were applied only to the side chain and not to the polypeptide main chain.

In the third stage with 41 cycles, individual temperature factors were also refined addition to the individual coordinates of the atoms. Fo-Fc electron density map was calculated after that stage in order to find water molecules. 26 water molecules were found for a subunit using the graphic display system.

At the last stage of the refinement, these water molecules were included. A total of 105 cycles of the refinement during which the deviations of bond lengths, bond angles and the other geometrical values from ideality were monitored during the refinement was performed. After the 105 cycles of refinement, the R value was deduced to 24.9% with a 0.03 Å root mean square (r.m.s.) deviation from ideal bond distances for the 28,694 reflections beyond 3σ between 5 Å and 2 Å spacings. It is mentioned that four crystallographically independent subunits were refined independently with some weights minimizing the deviation

from the average structure.

It is important to estimate the appropriate weighting factor for each stereochemical restraint relative to that of the structure factors (45). Correctly chosen weighting factors can reduce the R value greatly and maintain reasonable stereochemistry. The best weighting scheme was determined adequately during the refinement. Table 4-8 shows the weighting scheme at the final stage, where the weight of each term was calculated by $1.0/\sigma^2$. R values for each cycle during the refinement are shown in Fig. 4-8.

4-5 Description of the Cu,Zn-SOD structure

4-5-1 Structural identity in subunits

There are two SOD dimeric molecules or four subunits in the asymmetric unit of the crystal independently. It is therefore necessary to show the structural identity among the four subunits in order to discuss the molecular structure and to compare it with the other Cu,Zn-SOD. Each subunit was superposed on other three subunits by least-squares technique as the deviation between the corresponding coordinates became minimum. Table 4-9 shows the deviations among the four subunits: right upper diagonal corresponds to the deviations only for the Cα atom distances, left down diagonal corresponds to those for all atoms in the protein. Independent four subunits have almost the same structure among one another and few

Table 4-8

Weighting parameters at the final stage

Restraints	σ[a]	r.m.s.[b] deviation	No. of parameters
Distances			
Bond length (Å)	0.030	0.027	4488
Angle-related (Å)	0.040	0.044	6120
Planar (Å)	0.050	0.045	1420
Metal coordinates and disulfide bond (Å)	0.030	0.078	36
Planar groups (Å)	0.030	0.021	796
Chiral volumes (Å^3)	0.200	0.239	724
Non-bonded contacts			
Single torsion (Å)	0.200	0.164	1427
Multiple torsion (Å)	0.200	0.203	1646
Possible hydrogen bond (Å)	0.200	0.212	278
Torsion angles			
Peptide plane (ω)	5.0	3.6	628
Staggered (60/ 120°)	15.0	23.6	688
Ortho-normal (90°)	15.0	35.7	44
Non-crystallographic symmetry			
Positional (Å)			
Yellow subunit	0.25	0.355	1131
Violet subunit	0.25	0.395	1131
Green subunit	0.25	0.443	1131
Blue subunit	0.25	0.386	1131
Thermal (Å^2)			
Yellow subunit	0.25	0.034	1131
Violet subunit	0.25	0.034	1131
Green subunit	0.25	0.033	1131
Blue subunit	0.25	0.031	1131
Isotropic thermal factors [c]			
Main chain bond (Å^2)	1.0	0.280	2604
Main chain angle (Å^2)	1.5	0.476	3320
Side chain bond (Å^2)	1.0	0.264	1956
Side chain angle (Å^2)	1.5	0.514	2836

(a) The weight for each restraint was $1/\sigma^2$

(b) R.m.s. deviation from ideality

(c) Simple model of "riding motion" was applied for the restraints of the thermal parameters of atoms related by bonds or angles.

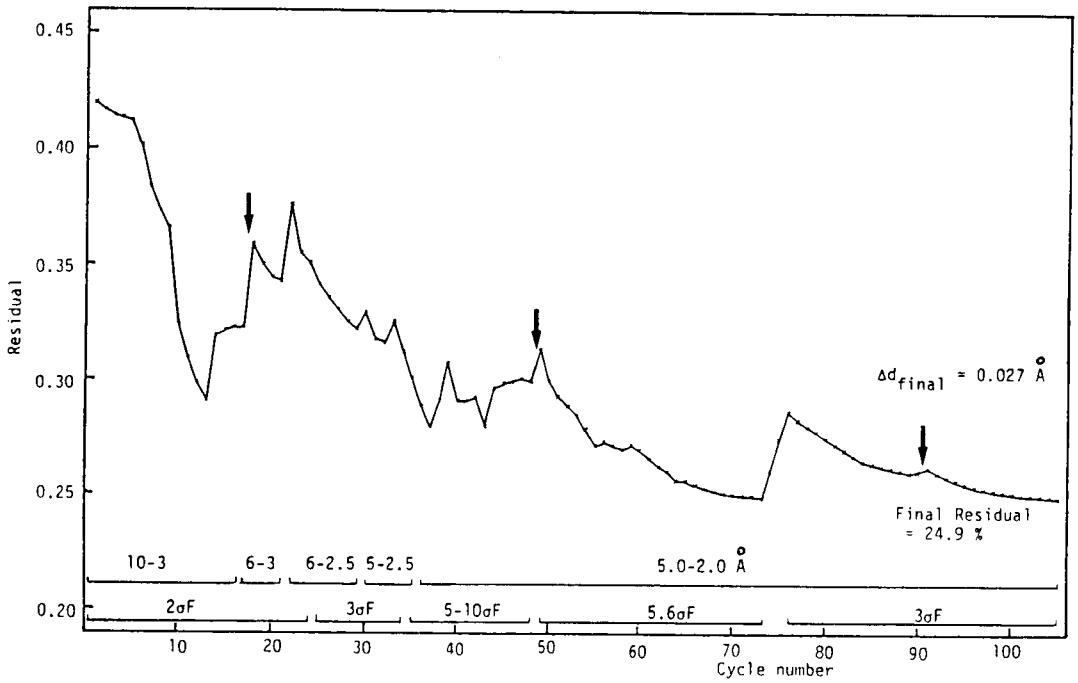

Fig. 4-8 Progress of the refinement represented by a plot of crystallographic residual error *versus* cycle number. Range of reflections used in the calculation is indicated with the resolution and the ratio to the sigma, respectively.

Table 4-9

Deviation among four subunits (Å)

(Å)	Yellow	Violet	Green	Blue
Yellow		0.33	0.33	0.31
Violet	0.61		0.34	0.35
Green	0.70	0.68		0.33
Blue	0.56	0.64	0.71	

Right upper diagonal corresponds to the deviations only for the Cα atom distances, left down diagonal corresponds to those for all atoms in the protein.

structural differences between subunits in the dimeric molecules (Yellow-Violet, Green-Blue) were observed.

4-5-2 Molecular structure in crystal

Two dimeric molecules in the asymmetric unit of the crystal are almost anti-parallel. Each of the dimers is internally related by a fairly exact non-crystallographic 2-fold axis. Viewed down the 2-fold axis, the dimeric molecule is an elongated ellipsoid about 33 Å wide, 62 Å long and 36 Å deep. The Cu^{2+} ions of each subunit are 33.4 Å apart in the molecule. In each subunit, the polypeptide main chain is folded into a cylindrical barrel comprised of eight anti-parallel β chains with three loops. The polypeptide main chain is followed as a single rope, except for two regions. One is an intra-subunit disulfide bridge and the other is a salt bridge between Arg-79 and Asp-101. The results reported in this thesis are concerned primarily with the structure of the SOD Yellow subunit.

4-5-3 Overall structure of the subunit

Each subunit of spinach SOD is composed primarily of eight anti-parallel β strands, which form a flattened cylinder, plus three external loops (Fig. 4-9 and Fig. 4-10). The active site composed of two metal ions (Cu^{2+} and Zn^{2+}) lie 6.1 Å apart at the bottom of a long channel located on the external surface of the β barrel between the two largest loops.

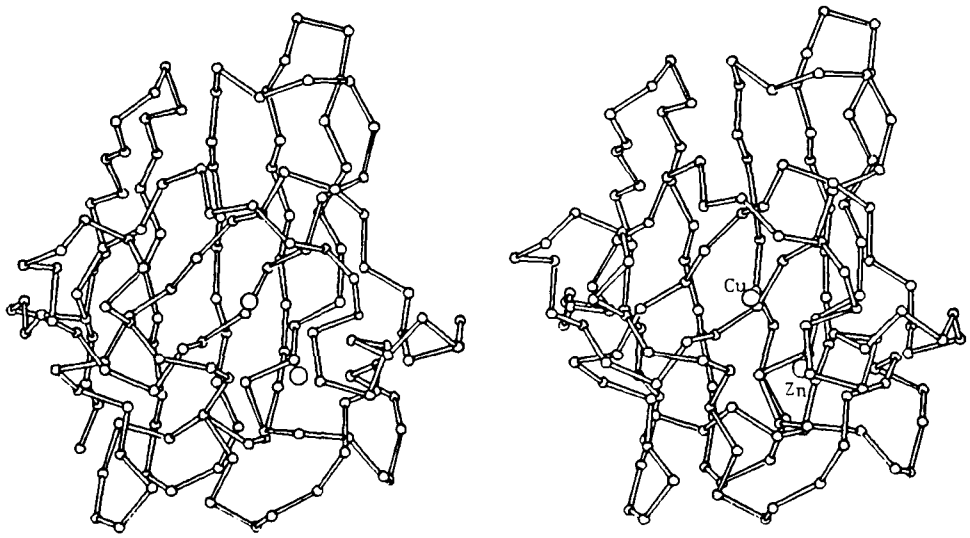

Fig. 4-9 Stereo drawing showing Cα backbone of the spinach SOD subunit viewed down the axial direction of the Cu from the solvent. Solid spheres represent the Cα atoms.

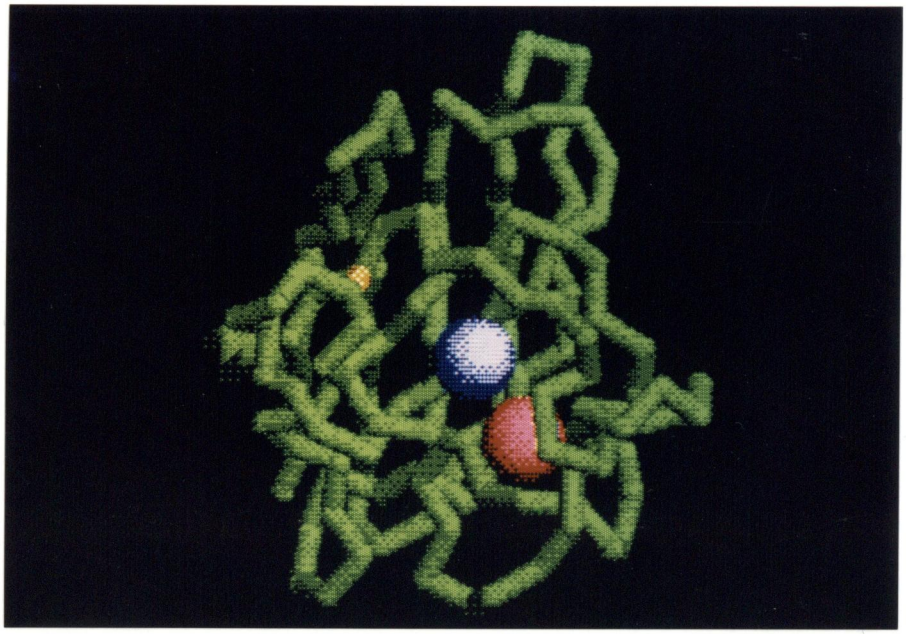

Fig. 4-10 Overall folding of the backbone in the spinach SOD with tube. A blue ball is Cu and a red is Zn.

The eight anti-parallel β strands form a Greek key β barrel with a topology of +1, +1, +3, -1, -1, +3, +1 (46) in which each β-strand connection is named for the number of strands it moves over in the sheet and the direction (clockwise or counterclockwise) that it moves around the barrel. There are no crossover connections in the SOD β barrel. Fig. 4-11 illustrates the main-chain to main-chain hydrogen-bond pattern in the β structure. Strands are identified by a number beginning at the N-terminus and proceeding clockwise around the top of the β barrel as viewed in Figs. 4-9 and 4-10, and also by a letter assigned in order of amino acid sequence. The three long loops of non-repetitive secondary structure are named for the β strands that they join (see Fig. 4-11).

The division of the structure into distinct β barrel and loop regions is an artificial, but conceptually valuable simplification. A definition of the β structure is due to both ϕ,ψ values and hydrogen-bonding pattern with neighbouring residues. Out of 154 residues, 71 residues are classified in β structure and 73 residues are in loops (36 of loop-6,5, 13 of loop-4,7 and 24 of loop-7,8). The division of the residues is shown in Table 4-10 with their torsion angles.

4-5-4 Temperature factors

Atomic temperature factors (or B values) are also obtained after the refinement. The values for all 4524 atoms including water molecules are ranging from 6 to 19

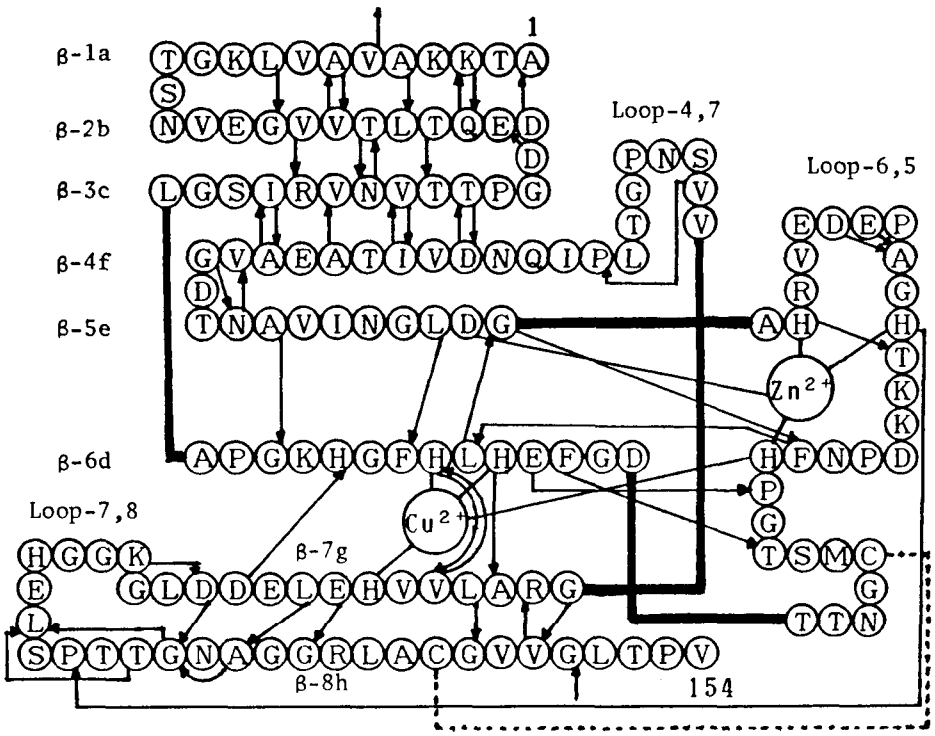

Fig. 4-11 A schematic drawing of the secondary structure and the main-chain to main-chain hydrogen bonds in spinach SOD.

Table 4-10

Main-chain dihedral angles

Residue	Group	Phi	Psi	Residue	Group	Phi	Psi
1 Ala	B	-	109	41 Gly	B	117	-163
2 Thr	B	-127	153	42 Lys	B	-92	148
3 Lys	B	-144	139	43 His	B	-117	-163
4 Lys	B	-127	159	44 Gly	B	-154	146
5 Ala	B	-143	171	45 Phe	B	-143	107
6 Val	B	-148	147	46 His	B	-136	156
7 Ala	B	-129	111	47 Leu	B	-95	96
8 Val	B	-94	107	48 His	B	-79	151
9 Leu	B	-83	141	49 Glu	L	-53	-54
10 Lys	B	-148	146	50 Phe	L	-99	165
11 Gly	B	-131	-162	51 Gly	L	-99	-26
12 Thr	T	-113	42	52 Asp	L	13	-107
13 Ser	T	-145	-166	53 Thr	L	96	-5
14 Asn	T	-103	14	54 Thr	L	-64	-39
15 Val	B	-90	118	55 Asn	L	-114	70
16 Glu	B	-94	-110	56 Gly	L	51	-136
17 Gly	B	93	162	57 Cys	L	-86	-16
18 Val	B	-163	131	58 Met	L	-74	-8
19 Val	B	-133	143	59 Ser	L	-93	3
20 Thr	B	-124	153	60 Thr	L	-95	3
21 Leu	B	-133	126	61 Gly	L	64	-171
22 Thr	B	-121	147	62 Pro	L	-77	-158
23 Gln	B	-153	133	63 His	L	-78	142
24 Glu	B	-114	148	64 Phe	L	-60	127
25 Asp	T	53	-97	65 Asn	L	-148	75
26 Asp	T	-165	43	66 Pro	L	-73	-36
27 Gly	T	-102	-154	67 Asp	L	-86	4
28 Pro	B	-72	169	68 Lys	L	48	91
29 Thr	B	-123	128	69 Lys	L	-139	-179
30 Thr	B	-93	129	70 Thr	L	-93	170
31 Val	B	-123	119	71 His	L	-75	141
32 Asn	B	-108	117	72 Gly	L	-158	-153
33 Val	B	-111	142	73 Ala	L	-108	162
34 Arg	B	-154	117	74 Pro	L	-53	-34
35 Ile	B	-126	158	75 Glu	L	-98	64
36 Ser	B	-151	162	76 Asp	L	-152	137
37 Gly	B	90	-7	77 Glu	L	-76	-42
38 Leu	B	-81	168	78 Val	L	-72	-15
39 Ala	B	-100	128	79 Arg	L	-41	152
40 Pro	T	-37	137	80 His	L	-52	-92

Table 4-10 (continued)

Residue	Group	Phi	Psi	Residue	Group	Phi	Psi
81 Ala	L	174	-49	121 Glu	L	-72	-54
82 Gly	L	-85	5	122 Leu	L	-89	172
83 Asp	L	-74	129	123 Glu	L	-79	149
84 Leu	L	-128	25	124 Asp	L	-85	140
85 Gly	B	62	-132	125 Asp	L	-69	-25
86 Asn	B	-129	142	126 Leu	L	76	11
87 Ile	B	-103	143	127 Gly	L	77	18
88 Val	B	-98	128	128 Lys	L	-79	-39
89 Ala	B	-98	129	129 Gly	L	-68	-172
90 Asn	B	-70	179	130 Gly	L	-139	32
91 Thr	T	-60	-19	131 His	L	-52	-79
92 Asp	T	-108	13	132 Glu	L	176	-38
93 Gly	T	54	64	133 Leu	L	-93	3
94 Val	B	-144	132	134 Ser	L	-56	-49
95 Ala	B	-120	137	135 Pro	L	-77	-3
96 Glu	B	-135	137	136 Thr	L	-117	-45
97 Ala	B	-149	-179	137 Thr	L	-116	-37
98 Thr	B	-149	97	138 Gly	L	93	23
99 Ile	B	-94	135	139 Asn	L	43	49
100 Val	B	-120	97	140 Ala	L	-96	-15
101 Asp	B	-108	160	141 Gly	L	-9	120
102 Asn	L	-122	12	142 Gly	L	133	97
103 Gln	L	-111	-3	143 Arg	L	-84	95
104 Ile	L	-120	65	144 Leu	L	-72	-23
105 Pro	L	-64	170	145 Ala	B	-179	142
106 Leu	L	-120	9	146 Cys	B	-143	163
107 Thr	L	-137	152	147 Gly	B	-149	151
108 Gly	L	64	-148	148 Val	B	-87	139
109 Pro	L	-72	145	149 Val	B	-76	106
110 Asn	L	84	-4	150 Gly	B	-109	147
111 Ser	L	-47	139	151 Leu	B	-66	139
112 Val	L	-104	29	152 Thr	B	-142	148
113 Val	L	-84	144	153 Pro	B	-38	-108
114 Gly	L	71	17	154 Val	B	-180	-
115 Arg	B	-110	173				
116 Ala	B	-112	150				
117 Leu	B	-112	125				
118 Val	B	-121	144				
119 Val	B	-112	126				
120 His	B	-100	162				

Phi, psi : Backbone dihedral angles measured in the range -180^{o} to $+180^{o}$.

B, L and T denote β structure, loops and turns, respectively.

$Å^2$, and the average value is 9.6 $Å^2$. They indicate the relative degree of order in the refined coordinates. The B values in the four subunits are almost the same with correlation coefficient greater than 0.99997, respectively. The highest B values are found at solvent-exposed molecular extremities in turns joining the ends of β strands. Especially in the turns of Thr-12 to Asn-14 between β strands 1a and 2b, and of Asp-25 to Gly-27 between β strands 2b and 3c, the B values are quite high. The other surface regions have high temperature factors compared with those of interior atoms in a molecule. Interior atoms in a molecule are well-ordered with lower B values. Val-113, Gly-114 and Gly-150 which are important for the dimer contact interactions have relatively low B values and are well-ordered. Around the disulfide bond, the B values are also relatively low. The B values of the Cα atoms in each residue are plotted in Fig. 4-12.

4-5-5 Torsion angles

The torsion angles (ϕ,ψ) of the polypeptide backbone are listed in Table 4-10. These angles were calculated from the coordinates after the refinement. The ϕ,ψ distribution plotted in Fig. 4-13 (47) strikingly illustrates the strongly β-strand character of this enzyme. Several glycine residues are outside the allowed region.

4-5-6 Disulfide bond and salt link

The single disulfide bridge, Cys-57 to Cys-146, forms

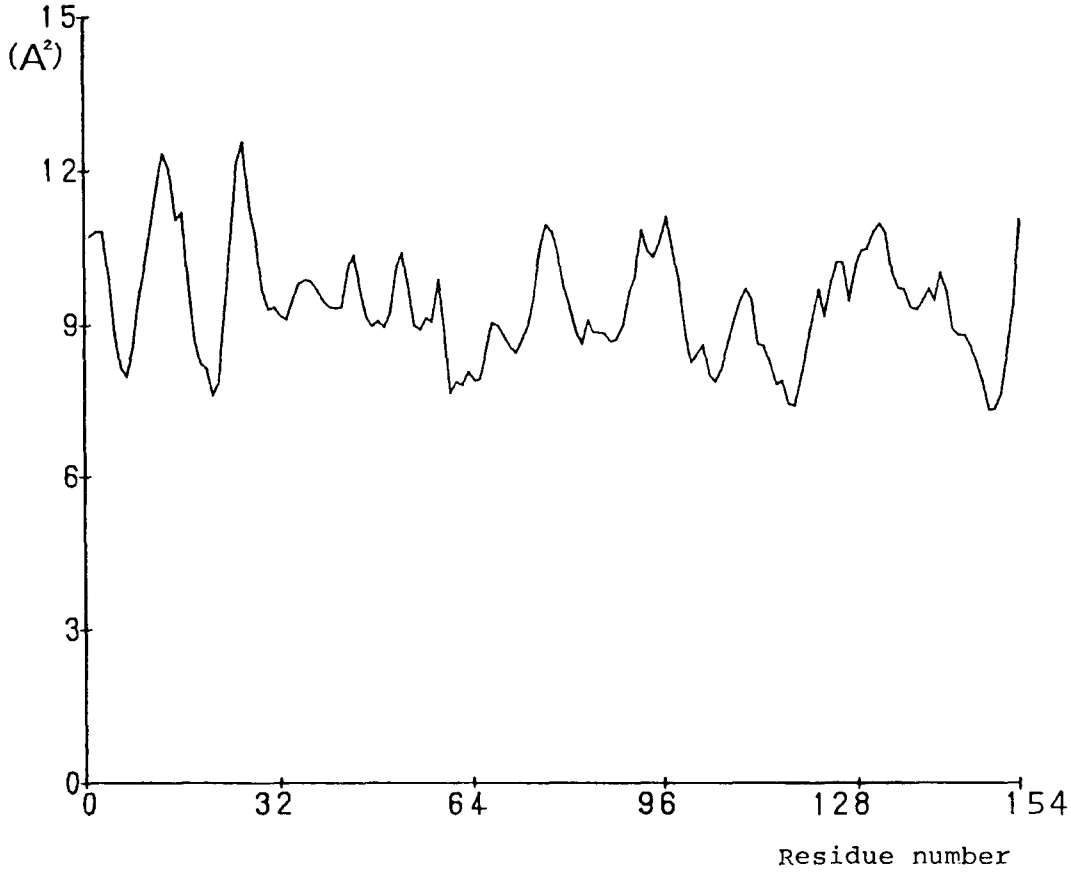

Fig. 4-12 Temperature factors of the Cα atoms.

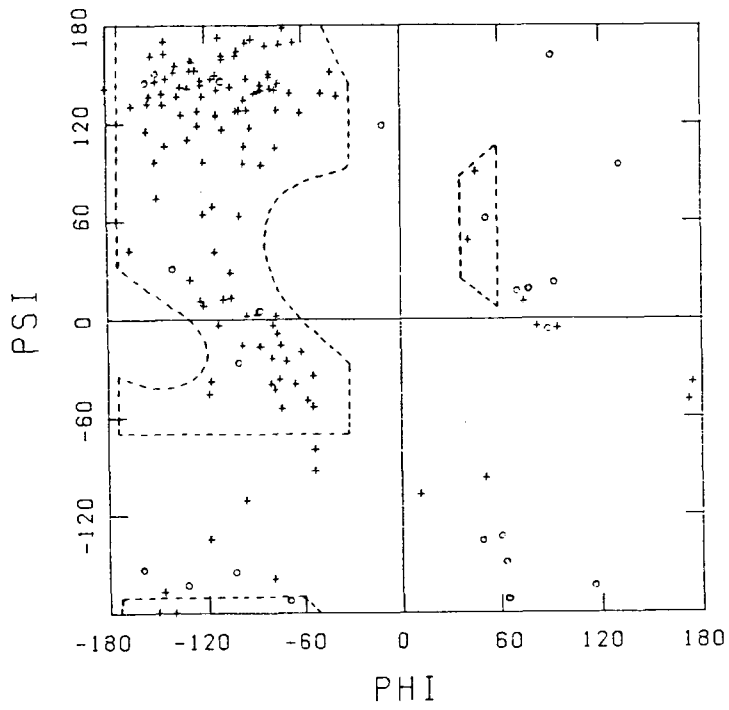

Fig. 4-13 Phi, psi plot for the coordinates of the Yellow SOD subunit. Glycine residues are indicated o; all other residues are indicated by +.

a covalent bond between loop-6,5 and the base of loop-7,8. The bond length between the sulfur atoms is 2.08 Å which is not so different from the ideal value in small molecules. Another confusing connectivity of the polypeptide main chain found in the density map is the salt link between Arg-79 and Asp-101. The longer the distance from the Cα atom, the lower the B value in these two residues. This salt link, connected between loop-6,5 and loop-4,7, may stabilize these two loops. These two types of connection which are not so popular as helix or sheet in protein structure take important parts in the stability or stiffness of the SOD structure. The distances of these connections are listed in Table 4-11.

4-5-7 Heavy atom binding sites

Two dimers in an asymmetric unit casually construct hydrophilic field among Glu-123, Asp-124 and Asp-125 in one dimer and the same residues in the other, though they are not related with a 2-fold axis. This field has a diameter of about 10 Å and corresponds to the major binding site of the uranyl ion. Concerning the platinum derivative, four binding sites were found in the asymmetric unit (Table 4-3). The platinum was therefore considered to bind quantitatively to the equivalent place in a subunit, because there were also four subunits in the asymmetric unit of the crystal. In fact, each platinum was found to bind to the sulfur atom of Met-58 which is next to the disulfide bridge, after the interpretation of the density

Table 4-11

Disulfide bond and salt link

Residue 1	Residue 2	d(Å)
57 Cys SG	146 Cys SG	2.08
79 Arg NH1	101 Asp OD2	2.65
79 Arg NH2	101 Asp OD1	2.64

map. Heavy atom binding sites are listed in Table 4-12.

4-5-8 Water molecules

26 water molecules in a subunit were identified in the Fo-Fc map after the third stage of the refinement. The positional parameters and the temperature factors have been refined in the 4th stage of the refinement. These molecules have relatively high temperature factors ranging from 11.5 to 19.1 $Å^2$ for Yellow subunit.

4-5-9 Hydrogen bonds (H-bonds)

Main-chain to main-chain, side-chain to main-chain and side-chain to side-chain H-bonds are listed in Tables 4-13, 4-14 and 4-15, respectively. Main-chain to main-chain H-bonds are also shown in Fig. 4-11 with the secondary structure demonstration. Most of the main-chain to main-chain H-bonds are between β strands and take part in the barrel-structure construction. These H-bonds were defined with the relatively narrow range of the distance between hydrogen donor and acceptor. Some of the H-bonds take important roles in the structural stability for both the metal ligand (His-71 and Asp-124) and the dimer contact (Asp-52, Thr-152, Gly-51, Gly-114 and Leu-151). The H-bonds between subunits constructing the SOD dimer are listed in Table 4-16.

4-5-10 Active site

The active site Cu^{2+} and Zn^{2+} ions lie 6.1 Å apart at

Table 4-12

Heavy atom binding sites

Heavy atoms	Subunit	Residue	d(Å)
$K_3UO_2F_5$-1	Yellow	121 Glu N	1.4
	Blue	123 Glu OE2	2.5
$K_2Pt(NO_2)_4$-1	Yellow	58 Met SD	3.4
$K_2Pt(NO_2)_4$-2	Violet	58 Met SD	2.8
$K_2Pt(NO_2)_4$-3	Green	58 Met SD	5.0
$K_2Pt(NO_2)_4$-4	Blue	58 Met SD	4.0

Table 4-13

Main-chain hydrogen bonds

Residue 1	Residue 2	d(Å)	Residue 1	Residue 2	d(Å)
1 Ala O	25 Asp N	3.03	46 His O	117 Val N	2.85
3 Lys N	23 Gln O	2.93	47 Leu O	64 Phe N	2.79
3 Lys O	23 Gln N	2.93	47 Leu N	82 Gly O	2.92
5 Ala N	21 Leu O	2.90	48 His N	116 Ala O	2.69
6 Val N	150 Gly O	2.86	49 Glu N	62 Pro O	2.91
7 Ala N	19 Val O	2.68	50 Phe N	60 Thr O	3.04
7 Ala O	19 Val N	2.98	64 Phe O	82 Gly N	2.98
9 Leu N	17 Gly O	3.04	70 Thr O	80 His N	2.83
18 Val N	34 Arg O	2.90	71 His N	135 Pro O	3.04
20 Thr N	32 Asn O	3.03	73 Ala O	75 Glu N	3.05
20 Thr O	32 Asn N	2.98	73 Ala O	76 Asp N	3.01
22 Thr N	30 Ile O	2.95	90 Asn N	94 Val O	2.92
24 Glu O	26 Asp N	2.74	90 Asn O	93 Gly N	2.85
29 Thr N	101 Asp O	2.66	105 Pro O	112 Val N	2.88
29 Thr O	101 Asp N	2.92	114 Gly N	149 Val O	2.92
31 Val N	99 Ile O	2.82	115 Arg O	149 Val N	2.59
31 Val O	99 Ile N	2.89	117 Leu N	147 Gly O	2.95
33 Val O	97 Ala N	3.02	121 Glu N	142 Gly O	2.70
35 Ile N	95 Ala O	3.01	122 Leu N	140 Ala O	3.02
35 Ile O	95 Ala N	2.89	125 Asp N	138 Gly O	2.88
41 Gly O	89 Ala N	2.64	125 Asp O	128 Arg N	2.85
43 His O	124 Asp N	3.01	133 Leu O	137 Thr N	2.70
45 Phe O	84 Leu N	2.96	133 Leu O	138 Gly N	3.02
46 His N	118 Val O	2.89	138 Gly O	140 Ala N	2.94

Table 4-14

Side-chain to main-chain hydrogen bonds

Residue 1	Residue 2	d(Å)
13 Ser OG	14 Asn N	2.82
23 Gln OE1	106 Leu N	2.89
43 His O	86 Asn ND2	2.62
48 His ND1	61 Gly O	2.80
55 Asn N	59 Ser OG	3.00
57 Cys O	60 Thr OG1	2.80
61 Gly O	143 Arg NH2	2.75
65 Asn OD1	68 Lys N	3.03
65 Asn OD1	69 Lys N	2.79
73 Ala N	76 Asp OD1	2.77
74 Pro O	79 Arg NH1	2.99
79 Arg NH2	81 Ala O	2.98
80 His N	83 Asp OD2	2.67
86 Asn ND2	124 Asp N	2.96
90 Asn ND2	94 Val O	3.04
97 Ala O	98 Thr OG1	2.93
101 Asp OD1	103 Gln N	2.77
124 Asp OD1	137 Thr O	2.62
124 Asp OD2	126 Leu N	2.75

Table 4-15

Side-chain to side-chain hydrogen bonds

Residue 1	Residue 2	d(Å)
34 Arg NE	96 Glu OE1	2.88
46 His ND1	63 His NE2	2.67
71 His ND1	83 Asp OD1	3.02
71 His NE2	124 Asp OD1	2.77
71 His NE2	124 Asp OD2	2.71
79 Arg NH1	101 Asp OD2	2.73
125 Asp Od1	134 Ser OG	2.51

Table 4-16

Hydrogen bonds in dimer contact area

Residue 1	Residue 2	d(Å)
51 Gly N	151 Leu O	2.91
114 Gly O	151 Leu N	3.05
52 Asp OD1	152 Thr OG1	2.70
151 Leu O	51 Gly N	3.00
151 Leu N	114 Gly O	2.83
152 Thr OG1	52 Asp OD1	2.69

the bottom of a long channel. Although the positions of two metal ions were clearly shown as spherical features of high density in the M.I.R. electron density map, it was difficult to distinguish between the Cu^{2+} and Zn^{2+} ions only from inspection of the density map. These were identified by using their anomalous-scattering effects measured by synchrotron radiation (section 4-3, (*48*)). As a result, the Zn^{2+} was assigned as the highest peak and the Cu^{2+} was assigned as the next highest peak in the subunit density. The identifications of these metal ions are the same result as the case of bovine SOD (*5*). The fitness between the model and the density map was quite good around the active site as shown in Fig. 4-14.

The Zn^{2+} ligands are all located in a continuous stretch of chain, which forms the second half of loop-6,5. Except for His-63, the Cu^{2+} ligands are located in the β barrel (His-46 and -48 from β strand 6d and His-120 from strand 7g). His-63 which is the ligand to both Cu^{2+} and Zn^{2+} ions bridges these ions with its imidazole ring approximately planar to both metal ions.

The geometry of the active site metals and their ligands is shown in Fig. 4-15, as viewed from the solvent. His-63, -71 and -80 attach to the Zn^{2+} with ND1. The Zn ligand geometry by these three His residues and by OD1 of Asp-83 is almost tetrahedral. The Cu^{2+} is liganded by ND1 of His-46, and by NE2 of His-48, -63 and -120 with an uneven tetrahedral distortion from square planar geometry ($\angle$63NE2-Cu-120NE2=133°; $\angle$46ND1-Cu-48NE2=134°).

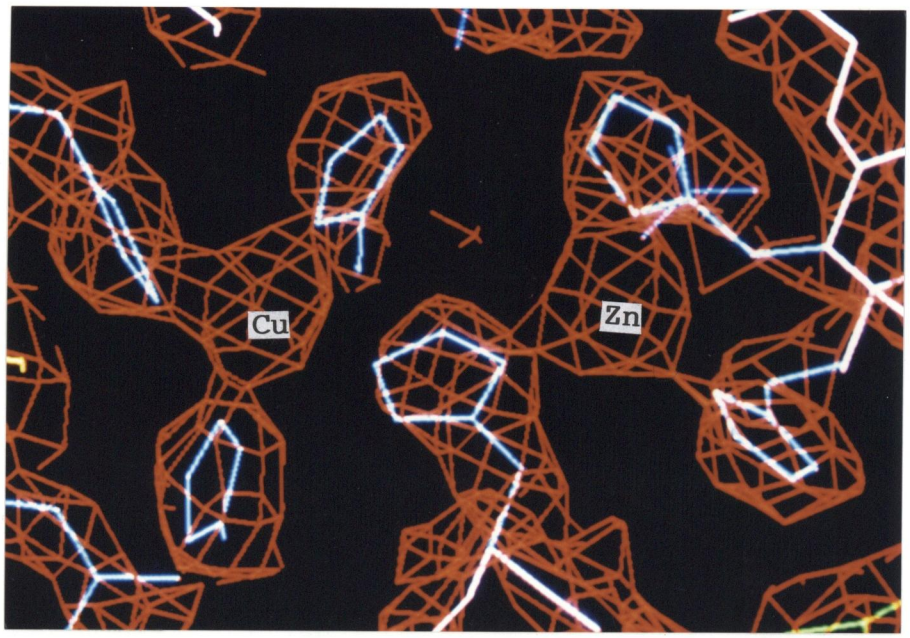

Fig. 4-14 The electron density map and model in the active site viewed down the same direction as Fig. 4-9.

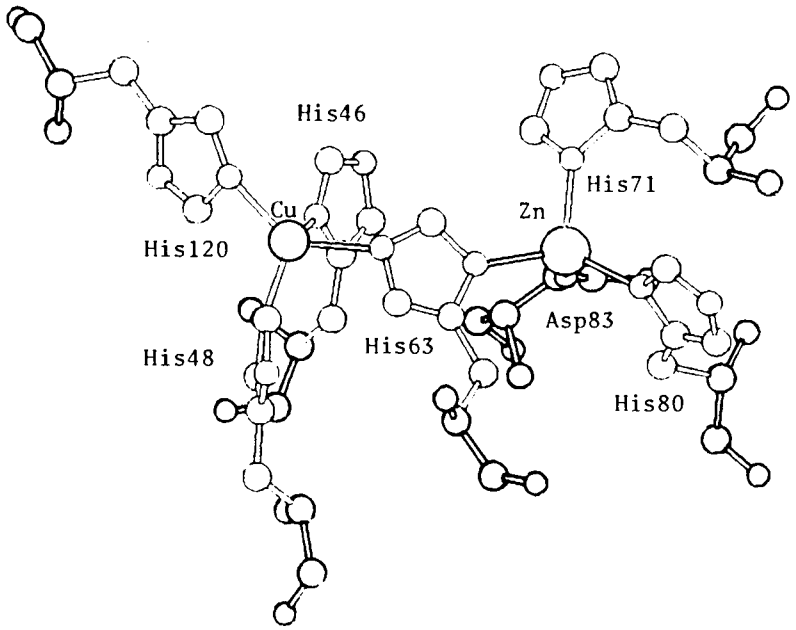

Fig. 4-15 The active site of spinach SOD viewed down the axial direction of the Cu from the solvent. The geometry of the Zn ligands is approximately tetrahedral. The Cu ligands form a distorted square plane.

The r.m.s. deviation from superposition of the ligand residues (averaged for the six possible pairs of the four SOD subunits) is 0.11 Å. If the active sites are assumed to be identical, then r.m.s. deviation from superposition provides an indication of the levels of error in the model.

The Cu^{2+}-N bond lengths of the liganding His averaged 2.07 Å for all four subunits, while Zn^{2+}-N bond lengths averaged 2.14 Å. A single Zn^{2+}-O bond length averaged 2.02 Å.

Cu^{2+} is solvent accessible and Zn^{2+} is somewhat buried on the view of the electron density map. Based on the fact, the role of the Zn^{2+} ion is considered to hold the structure and not to take part in the enzymic reaction directly. It is consistent with the catalytic process of the dismutation reaction of Cu,Zn-SOD which was demonstrated for bovine enzyme previously (49).

The active site is surrounded with two rims (see Fig. 4-16). One is the upper rim of Thr-136 to Arg-143 and the other is the lower rim of Met-58 to Asn-65. Several residues or atoms which are important for the activity are found in the upper rim around the active site. Cu^{2+} which attracts superoxide anion lies with the distance 5 to 6 Å away from both guanidinium N of Arg-143 and Oγ of Thr-137. The guanidinium N of Arg-143 has been suggested to function as a guide which let oxygen anions approach the Cu^{2+} ion (49). These two ionic groups are positioned in both the extremities of the surrounding residues in the upper rim whose sequence homology among nine Cu,Zn-SODs is

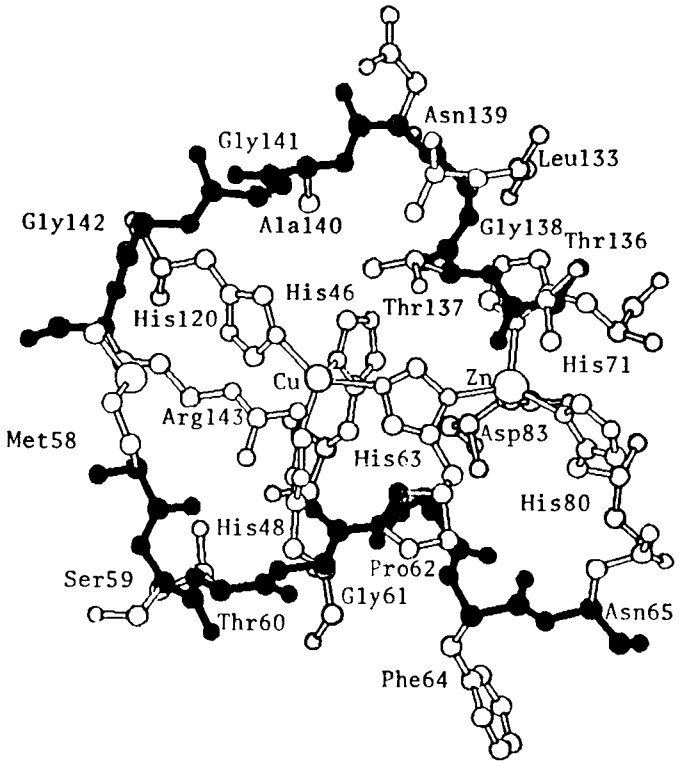

Fig. 4-16 Surroundings of the active site is shown viewed from the same direction as Fig. 4-15. The atoms consisting the two rims are represented with filled circles.

fairly high. Furthermore, Gly-141 which is conserved in all nine Cu,Zn-SODs so far sequenced (Fig. 5-7) may be considered to be important for the activity, because it has no side-chain and is quite small enough not to disturb the substrate coming. The upper rim, the end part of loop-7,8 including these residues must be important to hold the structure around the active site.

Aspartate in proteins is not solvent accessible generally and has an ionized carboxyl group. There are nine aspartate residues of spinach SOD, two of which are completely solvent inaccessible. They are Asp-83 and Asp-124. Asp-83 is one of the ligand to Zn^{2+} ion and almost buried in the molecule. Asp-124 makes three H-bonds with NE2 of His-71, carbonyl oxygen of Thr-137, and main-chain N-H of Leu-126 (see Tables 4-14 and 4-15). These H-bonds also play important roles in stabilizing the protein, especially around the active site.

4-5-11 Dimer contact

An SOD molecule is formed by two types of interactions around the 2-fold axis between the subunits. One is the hydrophobic interaction formed by Val-112, Val-113 and Leu-151. These residues in one subunit and those in the other make a hydrophobic field around the 2-fold axis. The other interaction is three H-bonds as shown in Table 4-16. These two types of interactions cause the surprising stability of Cu,Zn-SOD molecule.

5 Comparison of the structure among Cu,Zn-SODs

5-1 Superposition between spinach and bovine SODs

Tertiary structure analysis of Cu,Zn-SOD was reported only for bovine erythrocyte (5,49) so far. Therefore, the structure comparison between spinach and bovine SODs is possible. It may produce significant information about the molecular evolution of the enzyme. An amino acid sequence of bovine SOD was also reported (7) and the sequence homology between the spinach and bovine SODs is 55%. Comparing their sequences with each other, two residue insertions of the spinach SOD are found. One of them is Asp-26 in the turn region between β strands 2b and 3c, and the other is the C-terminal residue Val-154.

In order to compare the tertiary structures with each other, superposition between spinach and bovine SODs is necessary. Because a chain fold is generally well-represented by the coordinates of the Cα atoms, the tertiary structure can be compared with each other by superposing one chain on the other in such a way as to minimize the distance between corresponding Cα atoms by least-squares technique. The r.m.s. deviation can be taken as a similarity index for the comparison. The coordinates of bovine SOD was obtained from Protein Data Bank at Crystallographic Research Center, Institute for Protein Research, Osaka University.

Fig. 5-1 indicates the overall pictures of the two

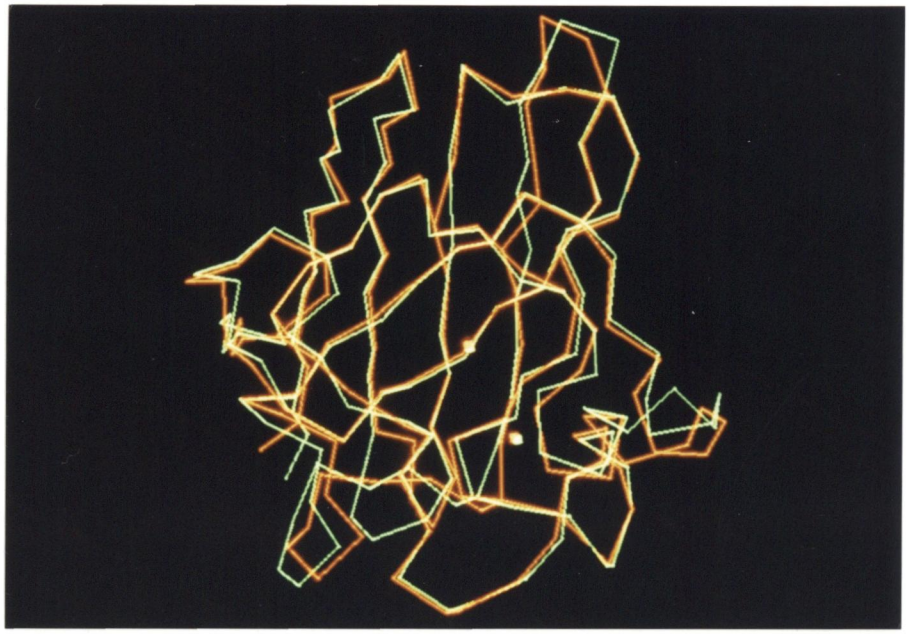

Fig. 5-1 Comparison of the Cα backbone between spinach and bovine SODs. Spinach SOD is represented with green line and bovine SOD with red.

SODs. The basic structure of β barrel with three loops is identical with each other in spite of the sequence difference of 45% when the whole structures are roughly compared. The overall r.m.s. deviation of the Cα atoms between spinach and bovine SODs is 1.52 Å except for the two inserted residues of Asp-26 and Val-154.

5-2 Structural similarities

5-2-1 General feature

The general feature of spinach SOD is quite similar to that of bovine SOD. A molecule is composed of two identical subunits related by a non-crystallographic two-fold axis in each case. The subunit structure of spinach SOD is basically eight-stranded β barrel with three loops and is topologically identical to that of bovine SOD. Fig. 5-2 shows the distances between the corresponding Cα atoms of spinach and bovine SODs. The overall r.m.s. deviation between the two SODs of 1.52 Å is considered to be quite small.

5-2-2 Active site

The active site in spinach SOD composed of two metal ions is also quite similar to that of bovine SOD. The r.m.s. deviation of the atoms in the ligand residues and metal ions (His-46, His-48, His-63, His-71, His-80, Asp-83, His-120, Cu^{2+} and Zn^{2+}) is 0.76 Å, when these are

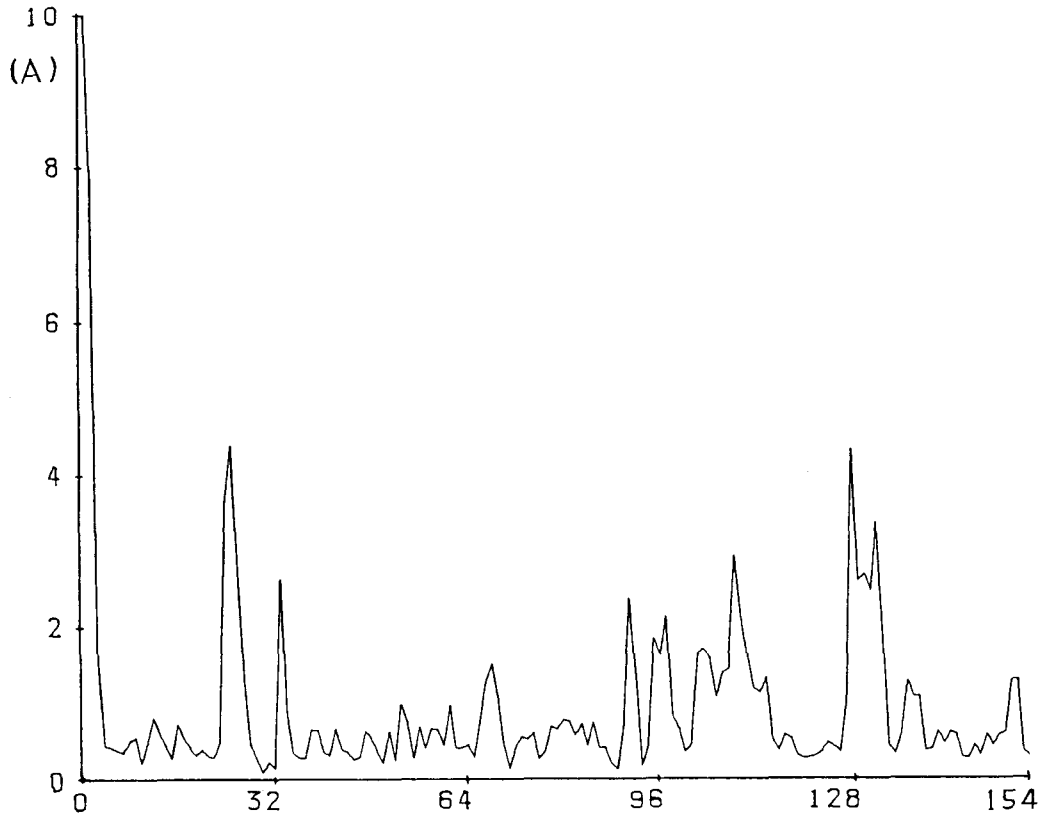

Fig. 5-2 Distances between the corresponding Cα atoms of spinach and bovine SODs.

superposed on those of bovine SOD by least-squares technique. The structure comparison at the active site is shown in Fig. 5-3.

Except for the ligand residues, functionally important residues around the active site are also conserved on the sequences in both SODs: Arg-143, Thr-137, Gly-141 and Asp-124. Arg-143 is considered to be important for the SOD activity based on the results of the modification experiment for bovine SOD (*50-51*). Asp-124 makes several hydrogen bonds shown in Tables 4-14 and 4-15 to stabilize the active site. These residues are also conserved on the sequences not only for the spinach and bovine SODs but also the other Cu,Zn-SODs whose sequences have been determined so far. It can be considered that these residues are very important for SOD activity and stability. The upper rim composed of seven residues (Thr-137, Gly-138, Asn-139, Ala-140, Gly-141, Gly-142 and Arg-143) has almost the same structure as that of bovine SOD. It is a natural result on the point of the sequence similarity between two SODs at this region except for one residue (Gly-142 is replaced with serine in bovine SOD). The only difference of the active site is the loss of a water molecule which was identified for the bovine SOD.

5-2-3 Dimer contact

A Cu,Zn-SOD molecule exists as a dimeric form with two identical subunits which are related by 2-fold axis. The interactions between the subunits in spinach SOD molecule

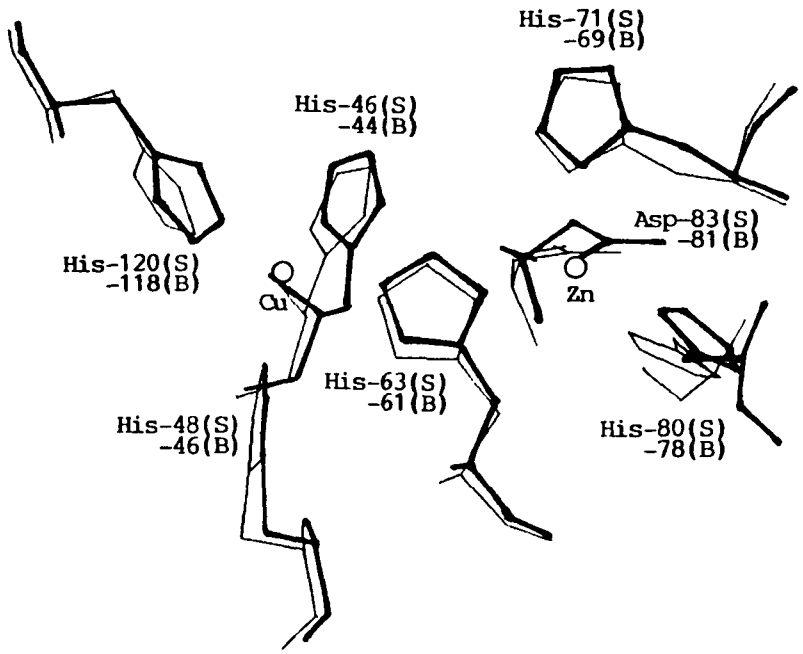

Fig. 5-3 Comparison of the active sites between spinach and bovine SODs. The thick line is spinach SOD and the thin is bovine one. (B) indicates bovine SOD and (S) indicates spinach SOD.

are hydrophobic and hydrogen bonding forces described in the section 4-5-11. Three residues (Val-112, Val-113 and Leu-151) which take part in the hydrophobic interaction for spinach SOD are replaced with isoleucine (Ile-110, Ile-111, Ile-149) for bovine SOD. The properties of the replacement pair are almost similar with each other in all three cases. Therefore, the same interactions between the subunits may be held in this case. The conformations of the side chains of the pair are certainly similar with each other. In general, replacement with the residue of the similar property is often observed in proteins. For the case of the dimer contact area in the SOD molecule, it is possible to replace the residues so as to maintain the hydrophobic interaction around the area.

There are two main-chain to main-chain and one side-chain to side-chain hydrogen bonds in the dimer contact area of spinach SOD as described in the section 4-5-11. In bovine SOD, the former two H-bonds also exist, but the latter does not exist. Thr-152 of spinach SOD is replaced with alanine for bovine SOD and cannot make the H-bond because amino acid alanine has no hydroxyl group to make H-bonds. However, the dimer contact area of spinach SOD is quite similar to that of bovine SOD as a whole (Fig. 5-4).

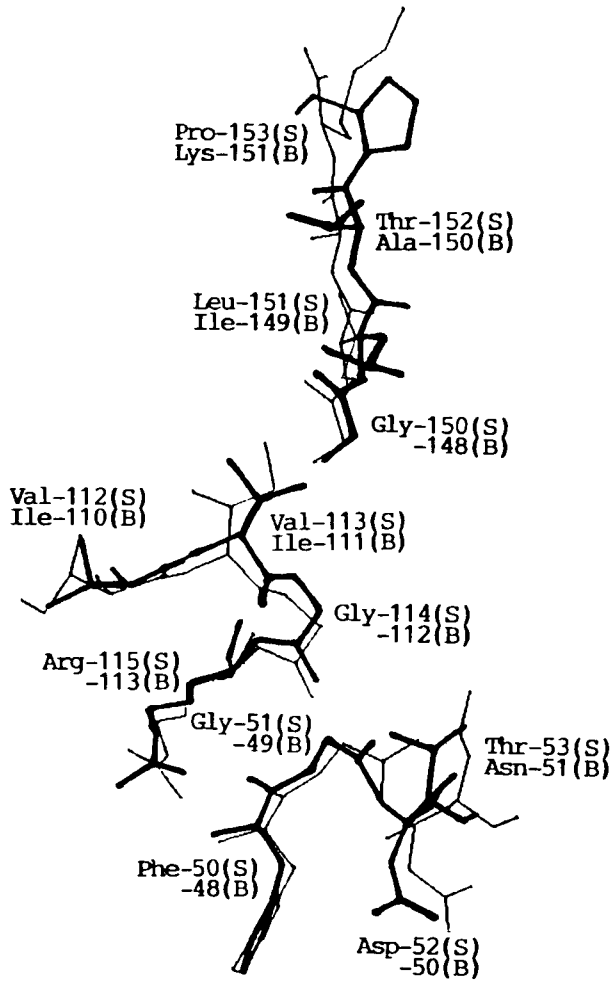

Fig. 5-4 Comparison of the dimer contact areas between spinach and bovine SODs. The thick line is spinach SOD and the thin is bovine one. (B) indicates bovine SOD and (S) indicates spinach SOD.

5-3 Structural differences

5-3-1 Definition of the structural difference

The structure of the spinach SOD was compared with that of four independent subunits of the bovine SOD (Orange, Yellow, Blue, Green), respectively. Residues whose Cα atoms differ beyond 1.52 Å among all four cases were defined as the positions with structural differences. In all the amino acid residues including Cu^{2+} and Zn^{2+} ions, eleven positions (Ala-1, Thr-2, Glu-24, Asp-25, Asp-26, Val-33, Ala-97, Gly-108, Gly-130, His-131 and Glu-132) were defined with structural differences. These residues can segregate into four groups by the structural neighbourhood. They are listed in Table 5-1 together with the average distances between the corresponding Cα atoms.

5-3-2 Group I (Hydrogen bond formation)

The first region is comprised of five residues, Ala-1, Thr-2, Glu-24, Asp-25 and Asp-26. Bovine SOD has no residue at the corresponding residue position of Asp-26 which is the inserted residue between Asp-25 and Gly-27 in spinach SOD. Therefore the turn structure including the residue is naturally different from that of bovine SOD. The turn of bovine SOD is a little shorter than that of spinach SOD because of the lack of one residue as shown in Fig. 5-5. On the other hand in spinach SOD, the hydrogen atom of N-H in Asp-25 is hydrogen bonded with the oxygen

Table 5-1

Structural differences between spinach and bovine SODs

No.	Spinach SOD	Bovine SOD	d(Å) [a]	Group
1	1 Ala	0 Acetyl	8.3	I
2	2 Thr	1 Ala	7.2	I
3	24 Glu	23 Lys	3.7	I
4	25 Asp	24 Gly	5.4	I
5	26 Asp	-	-	I
6	33 Val	31 Gly	2.6	II
7	97 Ala	95 Val	1.8	II
8	108 Gly	106 Gly	2.2	III
9	130 Gly	128 Gly	2.2	IV
10	131 His	129 Asn	3.0	IV
11	132 Glu	130 Glu	2.2	IV

a These are average distances over the four cases. See text in detail.

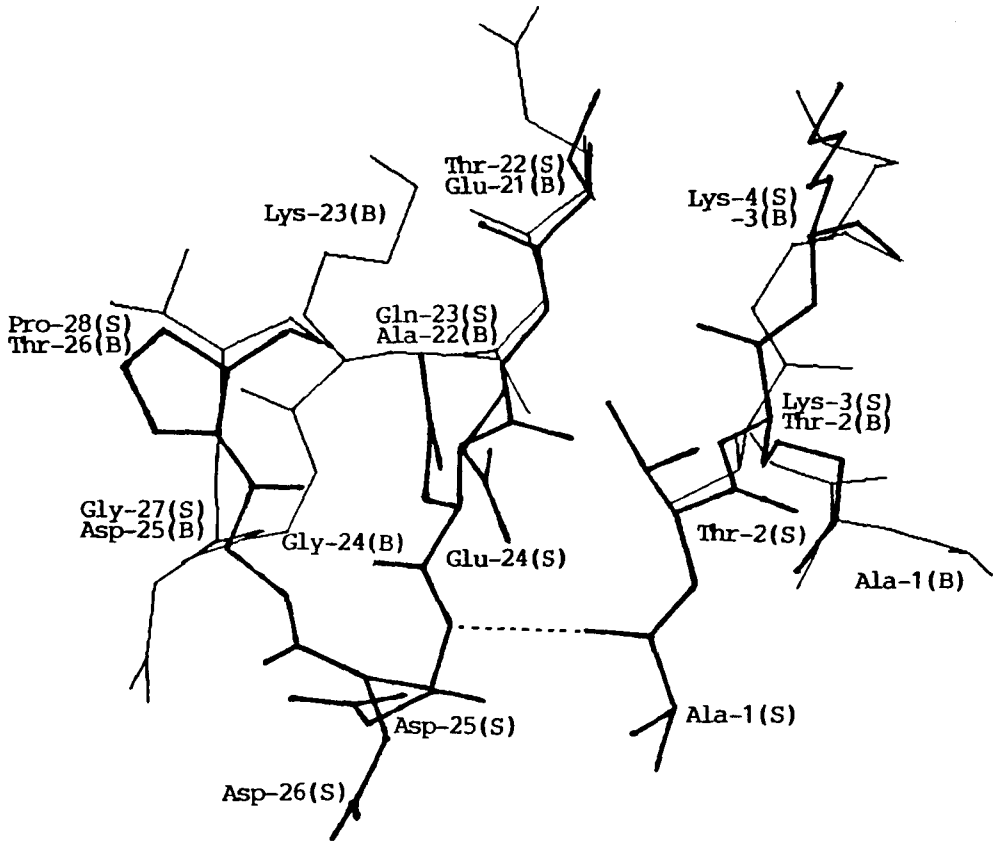

Fig. 5-5 Most different region comparing between spinach and bovine SODs (Group I in text). The thick line is spinach SOD and the thin is bovine one. (B) indicates bovine SOD and (S) indicates spinach SOD. The dotted line is the hydrogen bond between the hydrogen atom of N-H in Asp-25 and the oxygen atom of C-O in Ala-1.

atom of C-O in Ala-1. This hydrogen bond cannot be seen in bovine SOD. The N-terminus of spinach SOD must be straightly followed because of this H-bond, though that of bovine SOD bents before three residues from the N-terminus (see Fig. 5-5). This hydrogen bond of spinach SOD connects the turn region being long with the N-terminal region of spinach SOD. In general, terminal regions and turn regions in proteins are flexible. In the spinach SOD, however, one of the terminus, N-terminus and one of the turns between β-2b and β-3c are stabilized by making the H-bond. For other Cu,Zn-SODs, this turn region is expected to have a different tertiary structure compared with both bovine and spinach SODs because they have one more residue in this region than spinach SOD. Spinach SOD has one more residue at C-terminus, but no structural importance has been found about this insertion.

5-3-3 Group II (Residue replacement)

The second region is comprised of two residues, Val-33 and Ala-97, in the β strands. Val-33 in β-3c and Ala-97 in β-4f are adjacent to each other in the tertiary structure though they are apart on the sequence.

Val-33 and Ala-97 in spinach SOD is replaced with glycine (Gly-31) and valine (Val-95) in bovine SOD, respectively. It is interesting that the side chain of Val-33 in spinach SOD occupies the same space as that of Val-95 in bovine SOD: Cβ-Cβ distance is 1.5 Å, Cγ1-Cγ1 distance is 1.0 Å, Cγ2-Cγ2 distance is only 0.5 Å, if

spinach SOD were superposed on bovine SOD. The molecular surface of the side chain is calculated and indicated in Fig. 5-6 with the wire model around this region.

By comparing the sequences of all Cu,Zn-SODs hitherto sequenced only about these two positions with each other (see Fig. 5-7), an interesting result can be derived. As either site is occupied by a bulky residue, so the opposite site is occupied by a small residue, except for one case of fish SOD. This space in SOD β-barrel structure must be occupied by a bulky residue penetrated from either a strand β-3c or a strand β-4f. The residue replacement discovered in SOD structure suggests that not only the surface amino acid residues but also the residues inside the protein can be replaced with other residues so as to conserve the overall structure during the evolution.

5-3-4 Group III

The third region, Gly-108, is in loop-4,7, which is exposed to the solvent. It is not considered catalytically important because it is far from the active site. The difference may occur in the cause of amino acid replacement of Pro-109 and Asn-110 in spinach SOD with glutamate (Glu-107) and tyrosine (Tyr-108) in bovine one, respectively. In both cases, the residues of the replacement pairs have completely different properties. The torsion angles of the corresponding residues are quite different and the directions of the side chains are varied. However, the overall foldings of the loop of both SODs are

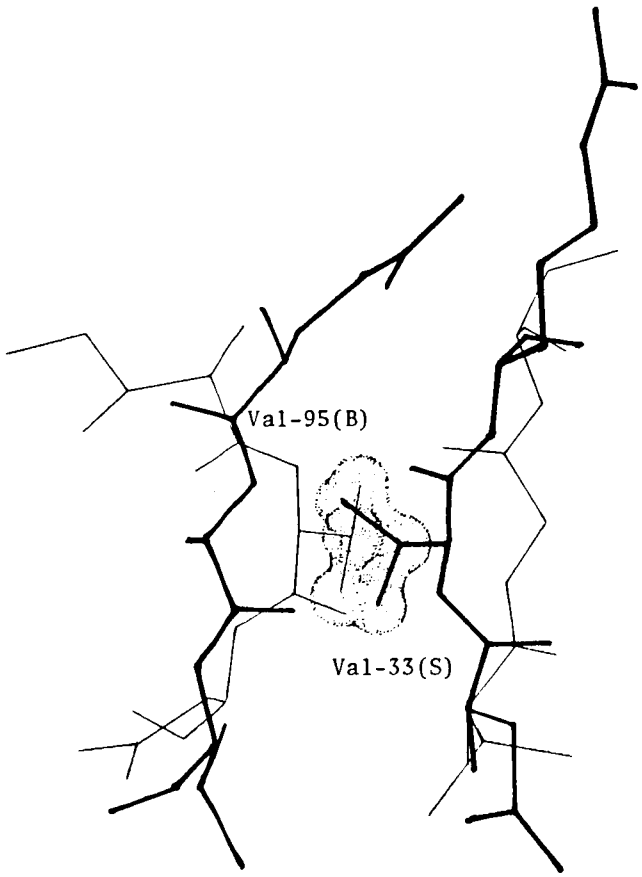

Fig. 5-6 Residue replacement (Group II in text). The thick line is spinach SOD and the thin is bovine one. Side chains of both Val-33 of spinach SOD and Val-95 of bovine SOD are shown with their molecular surfaces. (B) indicates bovine SOD and (S) indicates spinach SOD.

similar to each other and the effect of the difference is considered to be little.

5-3-5 Group IV

The fourth region, comprised of three residues, Gly-130, His-131 and Glu-132 is in loop-7,8, which is also exposed to the solvent. One amino acid replacement of histidine with aspargine at residue 131 occurs in this region, though Gly-130 and Glu-132 are conserved between these evolutionarily distinct two species.

The structural differences of Groups III and IV between two SODs may not originate from that in the species but that in the crystal fields.

5-4 Amino acid difference in nine Cu,Zn-SODs

The spinach SOD interacts the substrates in living organisms as a dimer consisting of two identical subunits. Each subunit has 154 amino acid residues with the free N-terminal residue of alanine. The complete amino acid sequence of the spinach SOD is given in Fig. 5-7, together with those of the eight species; the sequences were artificially aligned to maximize congruence. From the comparison of the primary structures, it seems likely that the blocked N-terminus may be restricted to mammalian enzymes.

Based on this alignment, the amino acid differences

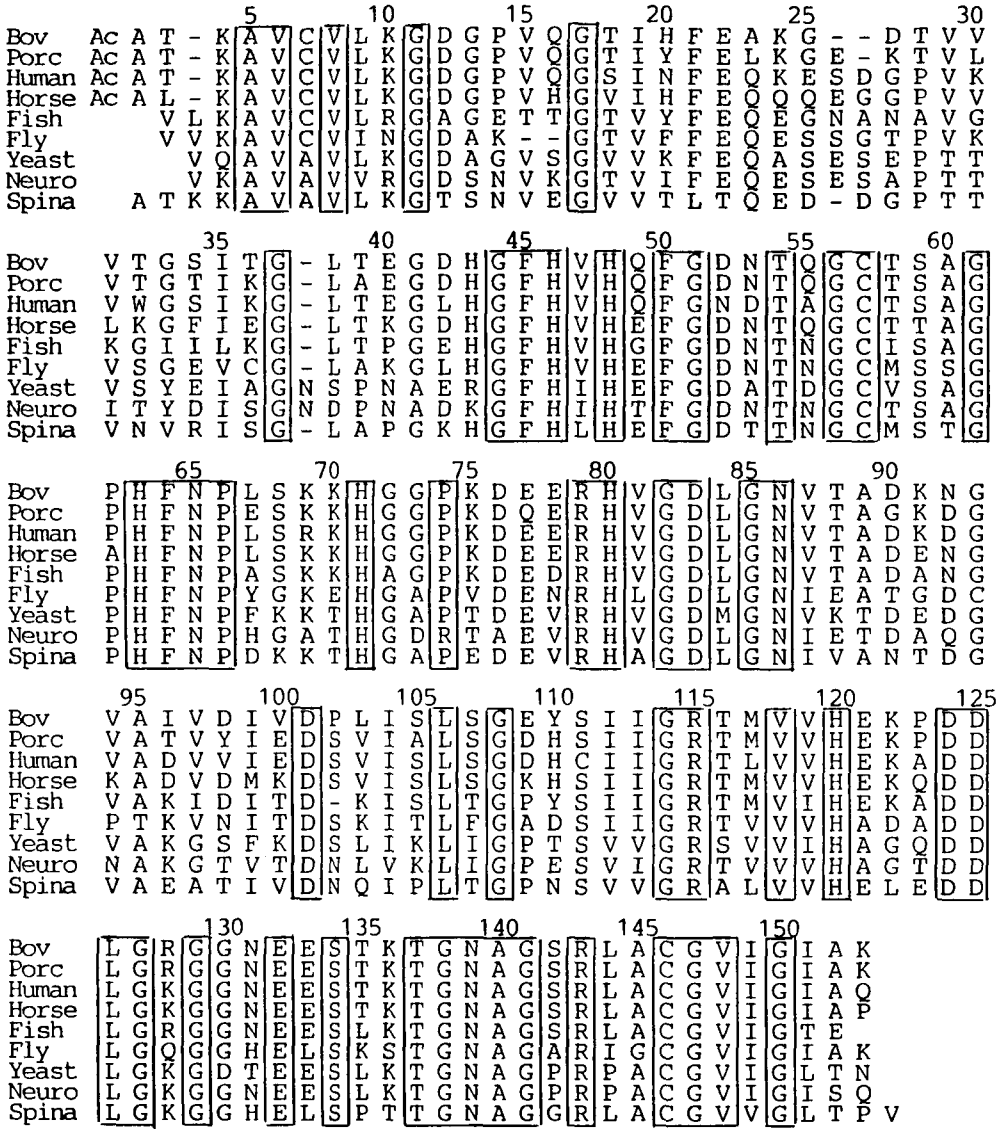

Fig. 5-7 Complete amino acid sequences of Cu,Zn-SODs from nine organisms. The numbers correspond to residue positions in the spinach sequence. The identical amino acids in all nine organisms are enclosed in boxes. Bov, Porc, Spina and Neuro show bovine, porcine, spinach and *Neurospora crassa* in the test, respectively.

(AAD) among nine Cu,Zn-SODs can be calculated for every position. If a position is occupied by the same residue for all nine Cu,Zn-SODs, AAD of that position is defined as zero. The position with a great amino acid difference is easy to be replaced. The great variability (the value of AAD is 6 or 7) occurs at nine sites of Glu-16, Thr-20, Gly-27, Asn-32, Arg-34, Ser-36, Asp-67, Thr-98 and Asn-110 in the spinach enzyme. In these sites, the hyper-variable sites are Thr-20, Gly-27 Arg-34 and Asp-67 which have the AAD value of seven.

Accessible surface area (ASA) of the dimeric molecules of spinach SOD was calculated by the method of Go *et al.* (52). The residues with the great variability are quite accessible with ASA ranging from 23 to 58% of the standard values of the residue.

Correlation coefficient among AAD, ASA, temperature factors of Cα atoms, distances from the center of the subunit, and distances between spinach and bovine Cα atoms were calculated, respectively. The results are summarized in Table 5-2. These five quantities have quite high correlation to one another. Especially, residues distributed on the molecular surface have high temperature factors and easy to be replaced during the evolution.

Table 5-2

Correlation coefficients

	1	2	3	4	5
1 Amino acid differences	1.00	0.42	0.21	0.25	0.08
2 Accessible surface areas		1.00	0.53	0.70	0.41
3 Temperature factors			1.00	0.59	0.37
4 Distances from the center of the subunit				1.00	0.48
5 Deviations between spinach and bovine SODs					1.00

5-5 Molecular evolution of Cu,Zn-SOD

Structural studies about Cu,Zn-SOD both on the primary and the tertiary ones are useful for the study of molecular evolution of this enzyme. In order to discuss the molecular evolution of Cu,Zn-SOD, the sums of amino acid differences between each pair of nine SODs were calculated (see Table 5-3).

The phylogenetic tree obtained by the matrix method is shown in Fig. 5-8. This tree shows that the divergence of spinach, fungi, fruit fly and vertebrates from one another must have been very close in time. This disagrees not only with the paleontological evidence but also with the results obtained from other proteins; according to the fossil records, the divergences among mammals, between fish and mammals, between insects and vertebrates and between plants and animals may be timed about 90, 400, 600 and 1200 million years ago, respectively. Lee *et al.* has reported that the rate of amino acid substitutions in Cu,Zn-SOD grossly departs from constancy and the apparent rate of amino acid substitutions is much greater in the evolution of mammals than at earlier times (*11*). Using the four additional data, spinach, swordfish, porcine and *N. crassa*, the same results as those of Lee *et al.* have also been obtained. The values in Table 5-3 show that the average evolutionary distance between spinach and mammals is about 2.5 times as long as that among mammals. The time elapsed since divergence between plants and animals is believed to

Table 5-3

Number of amino acid replacements between the nine Cu,Zn-SODs hitherto sequenced

	Porcine	Human	Horse	Fish	Fly	Yeast	*N.cra*	Spinach
Bovine	22	29	30	45	66	74	70	70
Porcine		28	36	50	66	73	71	70
Human			33	53	64	73	73	68
Horse				52	66	69	71	71
Swordfish					65	73	69	73
Fly						73	70	71
Yeast							48	67
N.crassa								71

The total number of residues was set to 156 for all species, and each deletion was counted as one event.

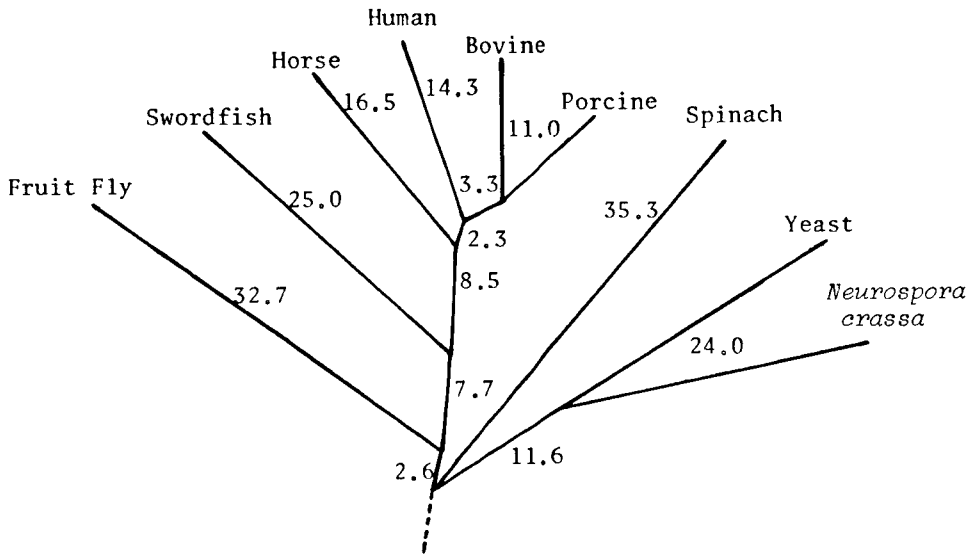

Fig. 5-8 Phylogenetic tree of Cu,Zn-SOD. The numbers along the branches show the inferred numbers of amino acid substitutions.

be more than 12 times as long as that among mammals. This discrepancy and the pattern of the phylogenetic tree may be explained by non-uniformity of the evolutionary rate of Cu,Zn-SOD.

The rate of Cu,Zn-SOD has been apparently accelerated as evolution progresses. SOD interacts superoxide anion (O_2^-) and the correlation of its evolutionary rate with the oxygen concentration cannot be ignored. As an example for the correlation, the increase of SOD contents was reported for a bacterium when a cultivation was carried out at higher oxygen concentration (*53*). Oxygen concentration in atmosphere was very low 9×10^8 years ago (0.001 present atmospheric level (P.A.L.)). The oxygen increased to 0.01 P.A.L. at 6×10^8 years ago and to 0.1 P.A.L. at 4.2×10^8 years ago. Since that time, the oxygen concentration has gradually increased with some oscillations (*54*). The increase of the oxygen concentration in atmosphere is correlative to the increase of the evolutionary rate of Cu,Zn-SOD qualitatively. It is interesting whether the rate of evolution among higher plants is as fast as that of mammals or not.

6 Conclusion and summary

The complete amino acid sequence of Cu,Zn-SOD from spinach leaves has been determined on the basis of peptides obtained by cyanogen bromide (BrCN) cleavage and by enzymic hydrolyses with *Achromobacter lyticus* lysyl*endo*peptidase, *Staphylococcus aureus V8* protease, trypsin and thermolysin. The spinach SOD consists of a total of 154 amino acid residues with alanine as the N-terminus and valine as the C-terminus. The present sequence, which has been established for the enzyme from a plant, is also highly homologous to those of the enzymes from other species. Especially, the residues essential for metal binding and enzyme activity have been extensively conserved among all of the Cu,Zn-SODs hitherto analyzed.

The spinach SOD crystallizes in space group *C*2 with two dimeric enzyme molecules per asymmetric unit. Based on the complete sequence, the electron density map has been successfully interpreted to obtain a Kendrew-type model of spinach SOD. Atomic coordinates were refined at 2.0 Å resolution using the Hendrickson and Konnert program (*36*) for stereochemically restrained refinement against structure factors, which allowed the use of non-crystallographic symmetry. The crystallographic residual error for the refined model was 24.9% with a root mean square deviation of 0.03 Å from ideal bond length and an average atomic temperature factor of 9.6 $Å^2$.

Spinach SOD is comprised of two identical subunits

which are related by a non-crystallographic 2-fold axis. Each subunit is composed primarily of eight anti-parallel β strands that form a flattened cylinder, plus three external loops. The main-chain hydrogen bonds primarily link β-strand residues. The single disulfide bridge and the salt bridge of ionic bond may stabilize the loop region of the structure. The active site Cu^{2+} and Zn^{2+} lie 6.1 Å apart at the bottom of the long channel; the Zn^{2+} is buried, while the Cu^{2+} is solvent-accessible. The imidazole side chain of His-63 forms a bridge between the Cu^{2+} and Zn^{2+}. The Cu^{2+} ligands (ND1 of His-46 and NE2 of His-48,-63 and -120) show an uneven tetrahedral distortion from a square plane. The Zn^{2+} ligands (ND1 of His-63,-71 and -80 and OD1 of Asp-83) show an almost tetrahedral geometry. Both the side chains and main chains of the metal-liganding residues are stabilized in their orientation by a complex network of hydrogen bonds.

Cu^{2+} and Zn^{2+} ions in a single crystal of spinach SOD have been exactly distinguished by enhancing the anomalous-dispersion effect of each ion with a wavelength near its K absorption edge, though these metal ions were not distinguishable in the original native Fourier map because of the small electron difference between a Cu^{2+} and a Zn^{2+}. Synchrotron radiation is useful to distinguish them because of its tuneability and strength. The X-ray diffraction data were collected on a four-circle diffractometer for protein crystallography using focused tuneable synchrotron radiation source at the Photon Factory. Anomalous

difference Fourier maps with the phases from the conventional structure analysis were calculated with the data sets at λ = 1.373 and 1.389 Å, near Cu K absorption edge, and 1.275 and 1.299 Å, near Zn K absorption edge. A comparison among these maps made it possible to distinguish between the Cu^{2+} and Zn^{2+} ions which have only one electron difference. This proves that metals or ions, whose atomic numbers are close to one another, may be located exactly by taking advantage of enhanced anomalous-dispersion data and reliable phase angles.

Two interesting differences between spinach and bovine SODs have been observed, although much of the structure is similar with each other such as active site, subunit contact area and general feature. One is the region comprised of the turn between β-2b and β-3c and N-terminus. Spinach SOD has one extra residue in the turn region and makes hydrogen bond between the region and the N-terminus to stabilize the structure. The other is the residue replacement between adjacent β strands as the space is occupied by a quite large residue from either strand. It is important that the Cu,Zn-SODs of these evolutionarily distinct two species have quite similar conformations though there are some differences in functionary unimportant places. Surprisingly loop-6,5 is quite analogous in these two species. That loop is comprised of 36 residues and the four zinc ligands are penetrated from this loop. This Zn binding loop is quite important to hold Zn^{2+} ion and to make dimeric molecule because the

subunit contact area in one molecule is between loop-6,5 in one subunit and C-terminal region in the other. In conclusion, spinach SOD has quite similar three-dimensional structure compared with bovine one though they have about 55% homology in the primary structure. Especially Cu^{2+} and Zn^{2+} ions of importance for active appearance, atmosphere around them and subunits contact area have been conserved in these two species.

Molecular evolution of Cu,Zn-SOD may be concluded in this thesis based on the nine Cu,Zn-SODs sequenced so far. Six histidine residues (His-46, -48, -63, -71, -80, and -120) and an aspartate residue (Asp-83), all of which are ligands for Cu^{2+} and Zn^{2+} in the tertiary structure of spinach SOD (*28*), are found at the same positions as those in the eight other SODs. Other residues which are functionally and structurally important in Cu,Zn-SODs are also conserved in the spinach enzyme. Those are an arginine residue (Arg-143) whose modification inactivates the enzyme (*50-51*), two cysteine residues (Cys-57 and -146) which form an intra-subunit disulfide bridge, and most of the glycine residues which can be accounted for in the β turns at the corner of the barrel in this enzyme (*28*).

Sequence homology in terms of the number of amino acid replacements among all nine Cu,Zn-SODs are summarized in Table 5-3. On average, the spinach enzyme varies from the other Cu,Zn-SODs by 70 amino acids. The fungi and fly SODs also show the same degree of replacement as the spinach one. These results suggest that Cu,Zn-SODs of

spinach, fungi, fly, and vertebrates may be evolutionarily in the same distance from each other. In conclusion, the evolutionary rate of Cu,Zn-SOD has been accelerated to evolve at the relatively high evolutionary rate as the time progresses.

References

(1) McCord, J. M. and Fridovich, I. (1969) *J. Biol. Chem.* **244**, 6049-6055

(2) Ringe, D., Petsko, G. A., Yamakura, F., Suzuki, K. and Ohmori, D. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3879-3883

(3) Stallings, W. C., Powers, T. B., Pattridge, K. A., Fee, J. A. and Ludwig, M. L. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3884-3888

(4) Stallings, W. C., Pattridge, K. A., Strong, R. K. and Ludwig, M. L. (1984) *J. Biol. Chem.* **259**, 10695-10699

(5) Tainer, J. A., Getzoff, E. D., Beem, K. M., Richardson, J. S. and Richardson, D. C. (1982) *J. Mol. Biol.* **160**, 181-217

(6) Steinman, H. (1982) in *Superoxide Dismutase* Vol.1 pp11-68

(7) Steinman, H. M., Naik, V. R., Abernethy, J. L. and Hill, R. L. (1974) *J. Biol. Chem.* **249**, 7326-7338

(8) Barra, D., Martini, F., Bannister, J. V., Schinina, M. E., Rotilio, G., Bannister, W. H. and Bossa, F. (1980) *FEBS Letters* **120**, 53-56

(9) Lerch, K. and Ammer, D. (1981) *J. Biol. Chem.* **256**, 11545-11551

(10) Steinman, H. M. (1980) *J. Biol. Chem.* **255**, 6758-6765

(11) Rocha, H. A., Bannister, W. H. and Bannister, J. V. (1984) *Eur. J. Biochem.* **145**, 477-484

(12) Lee, Y. M., Friedman, D. J. and Ayala, F. J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 824-828

(13) Hering, K., Kim, S. M. A., Michelson, A. M., Otting, F., Puget, K., Steffens, G. J. and Flohe, L. (1985) *Biol. Chem. Hoppe-Seyler.* **366**, 435-445

(14) Schinina, M. E., Barra, D., Simmaco, M., Bossa, F. and Rotilio, G. *FEBS Letters* **186**, 267-270 (1985)

(15) Lerch, K. and Schenk, E. (1985) *J. Biol. Chem.* **260**, 9559-9566

(16) Steffens, G. J., Bannister, J. V., Bannister, W. H., Flohe, L., Gunzler, W. A., Kim, S. M. A. and Otting, F. (1983) *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 675-690

(17) Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, Vol.5, suppl.3, 29-44, Natl. Biomed. Res. Found., Washington,D.C.

(18) Martin,Jr, J. P. and Fridvich, I. (1981) *J. Biol. Chem.* **256**, 6080-6089

(19) Bannister, J. V. and Parker, M. W. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 149-152

(20) Asada, K., Kanematsu, S. and Uchida, K. (1977) *Archives of Biochemistry and Biophysics*, **179**, 243-256

(21) Asada, K., Urano, M. and Takahashi, M. (1973) *Eur. J. Biochem.* **36**, 257-266

(22) Masaki, T., Fujihashi, T., Nakamura, K. and Soejima, M. (1981) *Biochim. Biophys. Acta*, **660**, 51-55

(23) Hirs, C. H. W. (1967) *Methods in Enzymology* **11**, 197-199, Academic Press, New York

(24) Tsunasawa, S., Kondo, J. and Sakiyama, F. (1985) *J. Biochem.* **97**, 701-704

(25) Narita, K., Matsuo, H. and Nakajima, T. (1975) in *Protein Sequencing Determination* (Needleman,S.B. ed.) 2nd ed., 70-79, Springen-Verlag, Berlin

(26) Isemura, S., Saitoh, E., and Sanada, K. (1984) *J. Biochem.* **96**, 489-498

(27) Morita, Y. and Asada, K. (1974) *J. Mol. Biol.* **86**, 685-686

(28) Kitagawa, Y., Tanaka, N., Katsube, Y., Kusunoki, M., Lee, G. P., Morita, Y., Asada, K., and Aihara, S. (1984) *Acta Crystallogr.* **A40**, Supplement C-43

(29) North, A. C. T, Phillips, D. C. and Mathews, F. S. (1968) *Acta Crystallogr.* **A24**, 351-359

(30) Rollet, J. S. and Sparks, R. A. (1960) *Acta Crystallogr.* **13**, 273

(31) Blow, D. M. and Crick, F. H. C. (1959) *Acta Crystallogr.* **12**, 794-802

(32) Bricogne, G. (1976) *Acta Crystallogr.* **A32**, 832-847

(33) Kusunoki, M. (1980) in *A doctoral thesis*, Faculty of Science, Osaka University, Japan

(34) Kitagawa, Y., Tsunasawa, S., Tanaka, N., Katsube, Y., Sakiyama, F. and Asada, K. (1986) *J. Biochem.* **99**, 1289-1298

(35) Richards, F. M. (1968) *J. Mol. Biol.* **37**, 225-230

(36) Hendrickson, W. A. and Konnert, J. (1981) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., ed.), vol. 1, pp. 43-47, Pergamon Press, Oxford.

(37) Einspahr, H., Suguna, K., Suddath, F. L., Ellis, G., Helliwell, J. R. and Pariz, M. Z. (1985) *Acta Crystallogr.* **B41**, 336-341

(38) Huke, K. and Yamakawa, T. (1980) *Nucl. Instr. Meth.* **177**, 253-257

(39) *International Tables for X-ray Crystallography* (1974) Vol.III pp. 171 Birmingham: Kynoch Press.

(40) Sasaki, S. (1984) in *Anomalous Scattering Factors for Synchrotron Users, Calculated Using Cromer and Liberman's Method.*, National Laboratory for High Energy Physics.

(41) Satow, Y. (1984) in *Methods and Applications in Crystallographic Computing* (Hall, S.R. and Ashida, T. eds) pp.56-64 Oxford Univ. Press, Oxford

(42) Kraut, J. (1968) *J. Mol. Biol.* **35**, 511-512

(43) Wilson, A. J. C. (1949) *Acta Crystallogr.* **2**, 318-321

(44) Cascarano, G., Giacovazzo, C., Peerdeman, A. F. and Kroon, J. (1982) *Acta Crystallogr.* **A38**, 710-717

(45) Konnert, J. H. (1976) *Acta Crystallogr.* **A32**, 614-617

(46) Richardson, J. S. *Nature*, **268**, 495-500 (1977)

(47) Ramakrishnan, C. and Ramachandran, G. N. (1965) *Biophys. J.* **5**, 909-933

(48) Kitagawa, Y., Tanaka, N., Hata, Y., Katsube, Y. and Satow, Y. (1987) *Acta Crystallogr.*, **B43**, (in press)

(49) Tainer, J. A., Getzoff, E. D., Richardson, J. S. and Richardson, D. C. (1983) *Nature*, **306**, 284-287

(50) Malinowski, D. P. and Fridovich, I. (1979) *Biochemistry*, **18**, 5909-5917

(51) Borders,Jr, C. L. Jr. and Johansen, J. T. (1980) *Biochem. Biophys. Res. Comm.* **96**, 1071-1078

(52) Go, M. and Miyazawa, S. (1980) *Int. J. Pept. Prot. Res.* **15**, 211-224

(53) Gregory, E. M. and Fridovich, I. (1973) *J. Bacteriol.* **114**, 543-548 (1973)

(54) Berkner, L. V. and Marshall, L. C. (1965) *J. Atomspheric Sciences*, **22**

List of publications

(1) The Structure Analysis of Copper,Zinc-Superoxide Dismutase from Spinach

Y.Kitagawa, N.Tanaka, Y.Katsube, M.Kusunoki, G.P.Lee, Y.Morita, K.Asada, S.Aihara

(1984) *Acta Crystallogr.*, **A40**, Supplement, C-43

(2) X-Ray Crystal Structure Analysis of Spinach Superoxide Dismutase

Y.Kitagawa

(1985) *Kessho Kaiseki Kenkyu Senta dayori* No.6, Crystallographic Research Center, Institute for Protein Research, Osaka University, Japan

(3) Structure and Molecular Evolution of Cu;Zn-Superoxide Dismutase from Spinach

Y.Kitagawa, S.Tsunasawa, N.Tanaka, Y.Hata, Y.Katsube, F.Sakiyama, and K.Asada

(1985) in *Computer Analysis for Life Science* (Chikao Kawabata and Alan R. Bishop ed.) pp 65-72 OHMSHA, Japan

(4) Amino Acid Sequence of Copper,Zinc-Superoxide Dismutase from Spinach Leaves

Y.Kitagawa, S.Tsunasawa, N.Tanaka, Y.Katsube, F.Sakiyama, and K.Asada

(1986) *J.Biochem.* **99**, 1289-1298

(5) Distiction between Cu^{2+} and Zn^{2+} Ions in a Crystal of Spinach Superoxide Dismutase by Use of Anomalous Dispersion Data and Tuneable Synchrotron Radiation
Y.Kitagawa, N.Tanaka, Y.Hata, Y.Katsube and Y.Satow
(1987) *Acta Crystallogr.* **B43**, (in press)

(6) Three-Dimensional Structure of Spinach Superoxide Dismutase at 2.0 Å Resolution
Y.Kitagawa, N.Tanaka, Y.Hata, M.Kusunoki, G.P.Lee, Y.Katsube, Y.Morita, S.Aibara, and K.Asada (in preparation)

The content of this thesis has been or will be published in the above papers and communications.