*BioMetals* 1998, **11**, 159–173

# An analysis of structural similarity in the iron and manganese superoxide dismutases based on known structures and sequences

S.M.J. Jackson & J.B. Cooper*

*Department of Crystallography, Birkbeck College, London and *Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Southampton, UK*

Received 8 October 1997; accepted for publication 3 December 1997

**There are two types of homologous enzymes catalysing the dismutation of the superoxide radical – Cu–Zn superoxide dismutases, and manganese or iron superoxide dismutases. In the latter two forms there is a high percentage of identity in the primary structures, and the tertiary structures are very similar particularly in the areas of the active site and in the residues responsible for the formation of the dimer. The quaternary structure of the dimer is also highly conserved. However, it has been found that despite this conservation there is strong metal ion specificity and many enzymes in the family will only be active if the correct metal ion is present. The purpose of this study has been to analyse solved X-ray structures for interactions common in both the manganese and iron forms and those that are specific to each, which may indicate reasons for the metal ion specificity. Initial analysis points to the probability that it is a combination of a number of residues, and not necessarily the same ones in every instance, which confer the specificity. In addition we have identified some anomalies in the currently available Fe/MnSOD structures which may require further remodelling and refinement.**

**Keywords:** homologous structures, iron and manganese superoxide dismutases, metal ion specificity

## Introduction

There are two separate homologous series of the superoxide dismutase (SOD) enzyme. The Cu–Zn SODs are found primarily in the cytosol of eukaryotes (although there are some bacterial forms) and are generally β-barrel dimers requiring the presence of both metal ions for catalytic activity (Tainer *et al.*. 1982); however, monomeric forms have been reported (Battistoni & Rotilio 1995, Battistoni *et al.* 1996). The manganese and iron SODs are also homologous but distinct from the Cu–Zn SOD family. However, the presence of only one or other metal ion is generally required, and although not structurally altered in the presence of the other metal ion, the enzymes are often significantly less active or inactive (Beyer & Fridovich 1991, Chen & Gregory 1991, Yamakura *et al.* 1991). Manganese and iron SODs are found in eubacteria, archaebacteria and organelles of eukaryotes, mostly the mitochondria, although in some plants they have been found in chloroplasts and other plant specific organelles (Almansa *et al.*. 1991). A generalization has been made that, apart from the archaebacteria and a few other exceptions, MnSODs are found in higher life forms and FeSODs in lower life forms (van Camp *et al.* 1990). The cambialistic SODs (i.e. those which retain activity whichever metal ion is present) so far identified are all bacterial and vary between those which have similar activity upon substitution of the metal ion, such as *Propionibacterium shermanii* (Meier *et al.* 1994), *Bacteroides*

Address for correspondence: J.B. Cooper, Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, UK. Tel: (+44) 1703 594338; Fax: (+44) 1703 594459

*fragilis* and *B. (Porphyromonas) gingivalis* (Amano *et al.* 1990), to those where the activity is significantly different, for example *Methylmonas J* where the Fe substituted SOD has less than 10% of the activity of the MnSOD (Matsumoto *et al.* 1991).

The reaction catalysed by both types of SOD is the dismutation of the superoxide radical into molecular oxygen and hydrogen peroxide in a two stage reaction:

$$M^{3+} + O_2^{\cdot -} \rightarrow M^{2+} + O_2$$
$$M^{2+} + O_2^{\cdot -} + 2H^+ \rightarrow M^{3+} + H_2O_2$$

The superoxide radical is a highly reactive species which, if not neutralized can cause extensive damage because of its rate of reaction. SODs have high turnover numbers so as to fulfil their function in 'mopping up' the superoxide radicals before they can inflict damage. Some prokaryotic SODs are known to be inducible, their levels of expression increasing as the oxygen concentration rises.

The manganese or iron SODs, including the cambialistic forms, have to be dimeric to function and tetramers are not uncommon. Each subunit contains around 200 residues, with a molecular weight of about 23kDa, and consists of two domains; the first is α-helical of around 90 residues or less while the second has an α+β fold (Figure 1).

The residues which ligate the metal ion are the same in all forms of this enzyme, namely three histidines and an aspartate. The first histidine is found in the first helix of the first domain (Figure 1) and the second histidine is in the last helix of the same domain. The other two ligands are four residues apart at the end of, and just after, the third β-strand of the second domain. The active site metal ion probably becomes octahedrally co-ordinated during catalytic turnover, the other two ligands being formed by a hydroxide ion (Borgstahl *et al.* 1992) and a substrate molecule (see Figure 2). It has been postulated that the metal centre moves between the five co-ordinate trigonal bipyramidal/square pyramidal form and the octahedral form depending on the presence of substrate (Lah *et al.* 1995, Tierney *et al.* 1995).

Six X-ray structures have been solved: *Bacillus stearothermophilus* (MnSOD) (Parker & Blake 1988a), *Pseudomonas ovalis* (FeSOD) (Stoddard *et al.* 1990), *Thermus thermophilus* (MnSOD) (Ludwig *et al.* 1991) , *Escherchia coli* (FeSOD) (Lah *et al.* 1995), *Mycobacterium tuberculosis* (FeSOD) (Cooper *et al.* 1995) and Human (MnSOD) (Borgstahl *et al.* 1992); of these five are publicly available from the Brookhaven database but one of the structures compares badly with the others. We have studied each of the structures

**Figure 1.** The tertiary structure of the monomer of the *M. tuberculosis* SOD showing the typical fold of Fe- and MnSOD subunits.

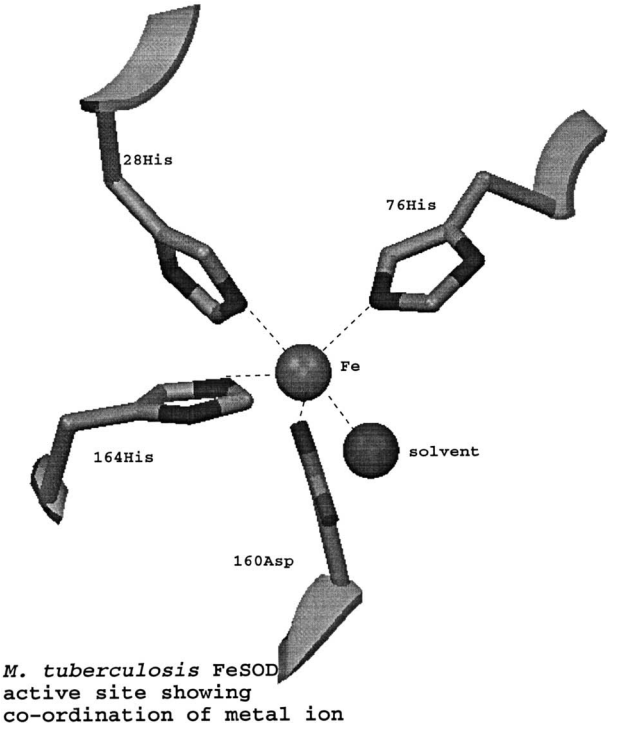


**Figure 2** The active site of *M. tuberculosis* SOD indicating the co-ordinating amino acids and solvent ligand found in Mn and FeSODs.

closely to determine common features and interactions which might help identify the basis of metal ion specificity. Discussion will principally be on the four SODs *T. thermophilus, M. tuberculosis*, *E. coli* and human. *P. ovalis* SOD will not form part of the main discussion as there are regions of secondary structure where the sequence does not align correctly with the other structures. Nevertheless there are areas where it can contribute to the overall understanding of the commonalities and differences in the Mn and FeSODs.

The information from the structures has been related to the conservation found within the primary sequences of all iron and manganese SODs. The determination of metal ion bound to each SOD has been by a number of methods. Direct detection methodologies such as electron spin resonance and atomic absorption spectroscopy are the most reliable. More indirect methods such as the action of postulated FeSOD specific inhibitors, e.g. azide, cyanide and peroxide, are not so reliable, as MnSODs also have a degree of sensitivity to these compounds. There have been several assignments of metal ion specificity by reference to the sequence alone, which our work demonstrates does not form a suitable basis from which to draw conclusions. As consideration has largely been of the monomer subunit, the interactions between the monomers, which might conceivably be significant, are not considered in any depth as they are strongly conserved in both iron and manganese SODs. Differences in the quaternary structure of the tetramer are known to occur (Wagner *et al.* 1993, Cooper *et al.* 1995) but these have not been considered in any depth here as the residues responsible for tetramer formation are never close to the active site.

## Methods

The known SOD structures were superimposed in 3D using the program MNYFIT in the COMPOSER suite of programs (Sutcliffe *et al.* 1987). The structures were aligned initially by using the Cα co-ordinates for each of the metal ligand residues in each of the SODs. The structures were studied using the program O (Jones *et al*. 1991) running on Silicon Graphics terminals. Particular attention was given to those areas where there was close alignment of 3D structure. The ϕ and ψ main chain torsion angles were analysed, as were interactions involving the backbone and the side chains. The distance cut-off between neighbours was set at 4.2 Å for van der Waals interactions and 3.2 Å for H-bond interactions. Where an interaction was found in one or more structures and not found in others, but it was reasonable to expect that the same interaction would occur, distances of up to 3.3 Å were allowed for H-bonding and up to 4.5 Å for van der Waals forces.

The program MALIGN (Johnson & Overington 1993) was used to align the primary sequence of other SODs. These were grouped on the basis of what metal ion each SOD needed for activity. The sequences were obtained from SEQNET at Daresbury Laboratory (Warrington, UK), using the program DELPHOS to search the OWL database for the consensus sequence (D)M(WEH)AY(Y), where those residues inside the brackets are fully conserved and those outside the brackets are generally conserved. Although the database is non-redundant, some of the sequences had to be eliminated as they were either mutants or there was no reliable source identifying the metal ion present in the enzyme.

## Results

After 39 cycles of least squares fitting of the X-ray structures to convergence using MNYFIT the RMS deviations (Å) in Table 1 were obtained. There were 169 residue equivalences for the four SODs for a cut-off Cα–Cα distance of 3.5Å.

The common elements of the secondary structure in the four SODs are indicated in Table 2. The results from MNYFIT show that the tertiary structure is highly conserved, with the most variable region being between the two main helices in the first domain. It is this region where *E. coli* and *T. thermophilus* SODs have more elaborate structures compared with the *M. tuberculosis* and human SODs.

Superposition of the SODs indicates an alignment of the first α-helix from 4Tyr (*M. tuberculosis*), which continues until 45Ala and re-establishes at 63Glu; the start of the second main helix. The end of the first conserved region coincides with the end of the first helix in *E. coli* and *T. thermophilus* SODs, which is about nine residues shorter than the equivalent helices in the *M. tuberculosis* and human enzymes. The *E. coli* and *T. thermophilus* SODs possess an extra discrete helical region between the

**Table 1.** RMS deviations of four Fe/Mn superoxide dismutase X-ray structures after 39 cycles of least squares fitting in MNYFIT. In all cases subunit A was used in the superpositions and subsequent comparisons

| | *E. coli* | Human | *T. thermophilus* |
|---|---|---|---|
| *M. tuberculosis* | 1.087 | 0.625 | 0.862 |
| *E. coli* | | 0.861 | 1.035 |
| Human | | | 0.735 |

**Table 2.** The common elements of the secondary structures taken as detailed in the relevant papers (Ludwig *et al.* 1991, Borgstahl *et al.* 1992, Cooper *et al.* 1995, Lah *et al.* 1995), most of which used the same algorithm (Kabsch & Sander 1983) to define the secondary structures. For this paper we have used the ϕ and ψ angles, as the segments so defined agreed more closely with the structural alignment

| | *M. tuberculosis* | *E. coli* | Human | *T. thermophilus* |
|---|---|---|---|---|
| Helix 1 | 22Gly–52Ala | 20Ala–42Ile | 20Ala–49Leu | 20Ser–45Glu |
| Helix 2 | 59Ile–82Asp | 61Gly–78Asn | 53Asp–76Ser–80Asn ($3_{10}$) | 70Gln–89Leu |
| Helix 3 | 92Gly–103Phe | 90Gly–100Phe | 91Gly–101Phe | 100Gly–108Glu |
| Helix 4 | 106Phe–118Thr | 103Phe–116Lys | 103Ser–116Val | 113Phe–125Met |
| Strand 1 | 124Gly–131Asp | 121Gly–127Lys | 122Gly–129Asn | 131Gly–137Lys |
| Strand 2 | 136Lys–143Tyr | 133Leu–139Ser | 135His–141Pro | 143Leu–149Pro |
| Strand 3 | 153Ile–160Asp | 150Thr–156Asp | 152Leu–159Asp | 160Thr–166Asp |
| Helix 5 | 174Lys–183Asn | 170Arg–178Trp | 173Arg–180Ile | 180Arg–187Ile |
| Helix 6 | 187Trp–198Thr | 183Trp–190Leu | 185Asn–197Lys | 193Trp–202Lys |

first major helix and the start of the second conserved helix, the latter running approximately antiparallel to the first. The next break in the structural alignment is in the connecting region between the two domains about four residues after the C-terminal end of the second conserved helix. The third main helix starts at the same point in each tertiary structure and continues for ten residues. The alignment is maintained up to the end of the first β-strand, where there is a break caused by the deletion of a residue in both *E. coli* and *T. thermophilus* SODs. The alignment is restored until the end of the second β-strand, where an insertion in human SOD and a deletion from the *T. thermophilus* SOD occurs. Structural variation in the loop connecting the second and third strands has been implicated in causing the different quaternary structures adopted by these SODs (Cooper *et al.* 1995). The alignment then continues to the end of the enzyme. The ϕ and ψ angles for conserved residues were very similar in the different homologues.

## Discussion

### *Description of the structures*

The close sequence identity of the four SODs means that the tertiary structures are very similar and the secondary structures, apart from a few differences, are generally the same. Presumably it is the few differences which are of the most significance for the variation in metal ion specificity and quaternary structure.

In our discussion of the similarities, residue numbering will refer to the *M. tuberculosis* SOD sequence unless otherwise stated. We will outline the apparent conservation of residues and its structural basis in a sequential manner starting at the N-terminus.

Conservation of the primary and tertiary structures starts almost at the beginning of the enzyme. The fourth residue is an aromatic which interacts with 41Asn. In all four SODs both polar atoms of the asparagine side chain are hydrogen bonded to the corresponding main chain of the aromatic residue. There is also a number of van der Waals interactions with hydrophobic residues in the area of the asparagine, and the human MnSOD has an additional H-bond between its equivalent of the aromatic residue (a histidine) and the side chain carbonyl of the asparagine. Residues 6,7,8,9 form a sequence motif L–P–X–L. Both leucines form van der Waals interactions with residues near the active site. The carbonyl on the proline is H-bonded to Nϵ2 of the histidine following the first metal ligand in *M. tuberculosis*, human and *T. thermophilus* SODs and to the hydroxyl group of the similarly placed tyrosine in *E. coli* SOD.

There is a second conserved region immediately before the first helix starting at residue 16; this has the consensus sequence L–E–P–H–I. However, it is not completely clear why they should be conserved. The proline is *cis* throughout and in all but the *E. coli* SOD the main chain carbonyl hydrogen bonds to the side chain of a hydrophilic residue just before the third helix in the second domain, 169Glu in *M. tuberculosis* SOD. The equivalent aspartate side chain in *E. coli* SOD is too short for an interaction to occur. The main chain carbonyl of the histidine in the consensus sequence H-bonds with the hydroxyl of conserved 170Tyr. There are also van der Waals interactions between the histidine and 184Val or the equivalent hydrophobic residues in the other three SODs.

The first helix starts at residue 22 and includes the first metal ion ligand and a number of other fully conserved or conservatively changed residues, some of which are involved in the reaction mechanism and the dimer formation. The motif here, starting at residue 26 is (E)–(L)–H–(H)–X–(K)–H–H, where the first H is the metal ion ligand. The bracketed histidine is equally likely to be another aromatic, particularly tyrosine or phenylalanine, and the lysine is conserved throughout apart from in the halobacterial SODs where it is a threonine. This is one of a number of residues which are different in the halobacteria, which may be a compromise because of the high electrostatic interference caused by the high salinity of their natural environment (Takao *et al.* 1990). The helix includes the first metal ion ligand 28His and the conformation of the α-helix is slightly distorted in all four structures at this point through to the lysine; some of the backbone interactions are also lost. This may be due to the liganded histidine affecting the local secondary structure in view of its strong interaction with the metal ion. There is a number of other interactions which might compensate for the loss in helix stability. The human 27His side chain is differently orientated in that, although it is in the same overall position as in the other SODs, the imidazole has turned through 180 degrees with respect to the *M. tuberculosis* and *T. thermophilus* SODs (*E. coli* has a tyrosine) and therefore appears not to have the same interactions as the rest. The conserved interactions made by the other two enzymes are H-bonds involving Nδ1 and Nϵ2. The human histidine therefore has a Cδ2 and Cϵ1 very close to polar groups. The Nϵ2 of His 27 in the *M. tuberculosis* and *T. thermophilus* SODs and the hydroxyl in the *E. coli* tyrosine at this position form H-bonds with the main chain carbonyl of 7Pro. By contrast, the human SOD has its Cϵ1 2.9 Å from the equivalent carbonyl. These effects indicate that the imidazole of 27His in the X-ray structure of human SOD is orientated 180 degrees from its correct position as both positions would give an identical fit to the electron density but only one rotamer gives rise to favourable H-bond interactions. The conserved 31Lys is thought to be involved in guiding the substrate to the active site (Takao *et al.* 1990). 32His Nϵ2 is involved in a dimer forming interaction with 167Tyr hydroxyl from the second subunit, both residues being fully conserved. 33His is also fully conserved, in all but the cambialistic SODs (*B. fragilis* and *gingivalis*) where it is a leucine, although again the human SOD's histidine side chain is orientated 180 degrees from the others and therefore would not be involved in the otherwise conserved side chain interactions. In these conserved interactions 33His Nδ1 interacts with Nϵ1 of the conserved 80Trp. 33His Nϵ2 is found to interact with the partially conserved 71His. In *E. coli.,* where the latter residue is a valine, the interaction is with the main chain carbonyl. Three residues after the above consensus sequence is the fully conserved 36Tyr, which will be discussed in more depth later as it has been implicated in the catalytic mechanism.

The C-terminal end of the first helix is the most highly variable region of the protein. However, in both the *M. tuberculosis* and the human SODs the alignment is maintained until residues 52Ala and 50Ala, respectively. The human SOD has a hairpin turn of three residues conforming to a typical helix hairpin defined by Efimov (Efimov 1991) ($\alpha\gamma\alpha_L\beta\alpha$). *M. tuberculosis* SOD has six residues between the first and second helix, the first three residues of the hairpin also conforming to the pattern of short α–α helix hairpins defined by Efimov ($\alpha\gamma\alpha_L\beta$–). The second main helix starts at 59Ile (*M. tuberculosis*) or 54Val in the human SOD. In both *T. thermophilus* and *E. coli* the first helix finishes several residues earlier than in the other SOD structures and the second main helix starts several residues later because the intervening loop contains an extra helix. It has been found that the formation of this extra helix may be indicative of a dimeric rather than a tetrameric structure (Cooper *et al.* 1995). Both *E. coli* and *P. ovalis* are dimeric and they both have extra helices here which interact with the β-sheet region of the second domain. Human and *M. tuberculosis* do not have the additional helix here, but the extended helical hairpins are found to be involved in the formation of the tetramer (Borgstahl *et al.* 1992, Cooper *et al.* 1995). However, *T. thermophilus* is tetrameric and has the additional helix, but has a very large solvent filled cavity in the centre indicating that the two dimers cannot interact extensively in this enzyme (Ludwig *et al.* 1991, Wagner *et al.* 1993).

The beginning of the last helix in the first domain is not particularly conserved in terms of primary structure, but it has several conserved interactions with the first helix. However, from residue 67 to the end of the helix there is a degree of conservation which shows up some differences between the iron and manganese SODs. In the FeSODs the motif is (F–N–N–A–A–Q)–X–(W–N)–H–X–(F–Y)–W; in MnSODs it is (Base–F–N–G–G)–G–(H)–X–(N)–H–X–(L–F)–W, where those in brackets are at least 50% conserved. The cambialistic SODs vary; both *B. fragilis* and *B. gingivalis* are more typical of the

FeSODs but are unique in replacing the fully conserved tryptophan with phenylalanine while the others are more typical of MnSODs. Two items of note are that the second ligand residue (76His) is present here, as is the conserved glutamine found only in FeSODs which may distinguish between FeSODs and MnSODs and will be discussed later. This helix runs antiparallel to the first helix, and is longest in both human and *M. tuberculosis* SODs, where it follows directly after the first main helix and, apart from in the turn region, close alignment is maintained. After the turn region, the four SODs come into structural alignment at 65Asn. The fully conserved 69Asn side chain amide nitrogen stabilizes the beginning of the helix by interacting with the main chain carbonyl of four residues earlier. Two of the X-ray SOD structures (*E. coli* and human) seem to have the side chain amide in the opposite orientation with the Oδ1 within H-bonding distance of the main chain carbonyl, indicating a possible misinterpretation of the electron density since only the amide $NH_2$ group could form a stable interaction with a carbonyl. The main chain carbonyl of 72Gly interacts with Nδ2 of 76His, the second metal ion ligand. This interaction is common to all four structures. The side chain of *M. tuberculosis* 75Asn is involved in a conserved stabilizing interaction with the main chain carbonyl of residue 145. This interaction is possible in all SODs except the FeSOD from *M. thermoautotrophicum* and the MnSODs *Halobacteria* where residue 75 is replaced by leucine. 79Trp in the *M. tuberculosis* SOD is replaced by smaller aromatic residues in the other SODs – such as tyrosine in *E. coli* and phenylalanine in the other two SODs. The *M. tuberculosis* tryptophan 79 has a larger number of contacts with other aromatic residues than do the equivalent residues in the other SODs due to their smaller size. However, both human and *T. thermophilus* SODs have more similar interactions to the *M. tuberculosis* enzyme than the *E. coli* SOD – in particular those involving the fully conserved 125Trp; they all have van der Waals contacts with 83Leu. 80Trp is conserved in all SODs (except *B. fragilis* and *B. gingivalis* where it is a phenylalanine), undergoing a number of fully conserved interactions with residues in the vicinity of the active site including the first ligand (28His), 29His and 33His. The latter interaction is an H-bond between the tryptophan Nϵ1 and the histidine Nδ1. There are also conserved van der Waals contacts with 9Leu.

There is little conservation in the connecting region between the two domains except for an almost fully conserved proline at residue 91. Some of this connecting region interacts with the loop between the fifth and sixth main helices. Specifically the main chain nitrogen of residue 84 interacts with the main chain carbonyl of residue 184 in both the *M. tuberculosis* and human SODs, and in all four enzymes the main chain nitrogen of 88 interacts with the main chain carbonyl of 185. The main chain carbonyl of 83Leu forms an H-bond with the Nϵ1 of 11Trp in the *M. tuberculosis* SOD. While this specific interaction is not replicated in the other structures as they have a tyrosine in the position of the tryptophan, there is a different interaction involving an H-bond formed between the tyrosine hydroxyl and the main chain carbonyl three residues earlier (80Trp).

The third main helix starts at the same place and continues for the same length in all four structures, although residue conservation is not marked. However, there is a number of points to note. 95Leu is only changed in *E. coli* and then to a valine, but the leucine in the human SOD is in a different position compared with the other three, which all undergo similar van der Waals contacts with the fully conserved 187Trp. This leucine also has contacts with the conservatively changed 190Val in all four SODs. The conserved 99Ile has similar contacts in all structures involving 91Pro, 103Phe, 109Phe and 187Trp, although the residues equivalent to 103 in *E. coli* SOD and 187 in the human SOD are not involved. The only SOD structure where one of these residues is different is the *T. thermophilus* SOD where the equivalent of 109Phe is a leucine. The last residue in the helix is 102Ala and again it has a series of conserved hydrophobic interactions, this time with 109Phe, 128Leu, 137Leu and 139Ile, although in the *E. coli* SOD the equivalent of 137Leu is not involved.

There are then three residues before the start of the fourth helix which form a conservatively changed motif F–G–S. The serine Oγ has a stabilizing interaction with the main chain amide three residues later in the helix. Their conformations are typical of α–α hairpins described by Efimov (1991) who observed the following conformational preferences: $\gamma\alpha_L\beta$, although in the SODs this connecting region is more accurately described as an α-corner. The conserved serine residue with the β conformation, is replaced by glycine in *T. thermophilus,* which adopts ϕψ angles found in the ϵ part of the Ramachandran plot. In the following helix there are three phenylalanines in *M. tuberculosis* and *E. coli*, two in human and one in *T. thermophilus* where the others have been replaced by leucines. Extensive contacts exist between these hydrophobic residues which are

maintained within the MnSODs (human and *T. thermophilus*) despite leucine being smaller than phenylalanine. In addition to the interactions between these three residues there are contacts with 128Leu, 139Ile, 159Leu, 182Trp and 187Trp. Although there is no specific motif, 109Phe and 113Phe are fully conserved in FeSODs but not in MnSODs. The cambialistic SOD from *P. shermanii* is unique in having methionine at position 109. The helix continues to 118Thr.

The first β-strand in the protein is immediately preceeded by two conserved residues starting at residue 122. This conserved region extends for the first five residues of the strand which have the following consensus sequence G–S–G–W–X–W–L (where G–W is fully conserved in the MnSODs and the leucine is conserved in the FeSODs). The strand, which runs centrally through the three-stranded β-sheet, starts at the same residue in each structure (124Gly); this forms a main chain H-bond with the residue immediately after the third ligand, which is at the end of the third β-strand. The previous residue 123Ser forms typical antiparallel β-sheet hydrogen bonding but the ϕ/ψ angles for β-strand are not found until 125Trp. This is a fully conserved residue forming part of the hydrophobic shield around the catalytic centre and is involved in a series of interactions which are discussed later. The residue at 127 is interesting as it is a highly conserved tryptophan, apart from in a few species such as the *M. tuberculosis* SOD where it is an alanine, or a larger hydrophobic residue, the sole exception being the cambialistic SOD from *P. shermanii* where this residue is a serine. The conserved tryptophan has extensive van der Waals contacts with the previous tryptophan (125) and with the hydrophobic part of the asparagine preceding the second histidine ligand (76). The main chain interactions between the central β-strand and the two adjacent strands are not completely identical between SODs as there is a degree of sequence variation. The first or central strand extends to 131Asp where there is a partially conserved β-hairpin preceding the second strand region. In *M. tuberculosis* and human SODs there is a four residue turn forming a 4:4 turn using the Sibanda *et al.* notation (Sibanda *et al.* 1989). The *T. thermophilus* SOD, which has a one residue deletion in the hairpin, forms the related 3:3 type β-hairpin; in the *E. coli* SOD, which also has a one residue deletion, a 3:5 turn with the expected glycine at position 4 is found. The second β-strand starts at 136Lys and continues to 143Tyr. There is little conservation in the strand, except for the second residue, which is a leucine that is conservatively changed in two SODs.

There is then a short crossover of ten residues linking the second and third strands. While there is no conservation in FeSODs, the conservatively changed motif starting at residue 144 N–Q–D is found in MnSODs. The glutamine is sometimes replaced by a histidine, but whichever residue is present here it is involved in a number of equivalent interactions which are discussed later. The loop including residue 145 is also involved in tetrameric interactions between pairs of dimers which are very extensive in the *M. tuberculosis* SOD. There is some loss of sequence alignment in this loop due to an insertion in the human enzyme. The X-ray analysis shows that this loop has very different structures in the SODs, which may arise from or give rise to their different quaternary structures.

The third strand runs antiparallel to the first or central strand and is on the opposite side to the second strand. It starts at 153Ile and continues to 160Asp, which is the third metal ion ligand. Two residues later is a fully conserved tryptophan. Its Nϵ1 forms a conserved H-bond with the main chain carbonyl of 144Asp. It also has extensive van der Waals contacts with metal ligand residue 164His, and the residue which interacts with the metal's solvent ligand, namely 145His (*M. tuberculosis*). There is a long motif running from residue 155 in the last β-strand through to the fifth main helix ($P_{155}$–L–L)–X–X–$D_{160}$–(V)–W–E–H–($A_{165}$–Y)–Y–(L–D–$Y_{170}$)–X–(N)–X–(R–$P_{175}$)–X–(Y)–X–X–$X_{180}$–(F–W). This motif contains both the third and fourth metal ion ligands (160Asp and 164His). This region, which is important for active site interactions (see later), has one invariant tyrosine and two highly conserved tyrosines, one of which is a phenylalanine in *M. tuberculosis*. A fourth tyrosine, which is also phenylalanine in *M. tuberculosis* is in the beginning of the fifth helix. The first of these residues, conserved as tyrosine (residue 166Phe in *M. tuberculosis* SOD), has hydrophobic contacts with the last conserved tyrosine (residue 177Phe in *M. tuberculosis* SOD) and there is a hydrogen bond formed between the phenolic hydroxyl group and Nϵ of a generally conserved arginine eight residues later (174Lys in *M. tuberculosis*). 167Tyr is fully conserved and in the two dimeric SODs studied, forms an H-bond with the fully conserved 32His in the other subunit. Position 177 is interesting in that when tyrosine occurs here the hydroxyl of the side chain forms an H-bond with the main chain carbonyl of the aspartate metal ligand; but there is no equivalent interaction in *M. tuberculosis* because 177 is a phenylalanine in this species. This is compensated for by a series of

van der Waals interactions between 177Phe and 79Trp (more usually a smaller aromatic residue in other SODs). Both the 177Phe and 79Trp side chains are positioned end on to main chain carbonyls enabling these to interact with the weakly ($\delta^+$ charged edge of the ring. These two carbonyls are that of the third ligand residue 160Asp and that of 118Thr, respectively. The two cambialistic *Bacteroides* SODs are the only ones that have a histidine at 177 which could interact with the phenolic group of the tyrosine they both have at 79. These aromatic residues form part of the hydrophobic shield around the catalytic centre. The hydroxyl of 170Tyr forms an H-bond with the main chain carbonyl of 19His, just before the first helix. This tyrosine is conserved in all other enzymes but the halobacteria, where it is phenylalanine, and in two of the non-metal ion specific SODs, where it is a histidine or an aspartate. The fifth helix continues through to 183Asn. In this region, the only other residue of note is 182Trp, which is generally conserved, although it can be another hydrophobic amino acid. The preceding residue is generally a phenylalanine in FeSODs but can be an isoleucine in MnSODs as it is in both human and *T. thermophilus*.

The loop before the last helix has a number of mainly hydrophobic residues which interact with much earlier parts of the structure, some of which have already been mentioned. The loop forms an α-corner and has the conserved conformation γββ (Efimov 1991). There are also interactions with the residues leading up to the first helix; in particular, 184Val and its equivalents have hydrophobic interactions with 16Leu and 19His. With the conserved 186Asn, Nδ2 hydrogen bonds with the main chain carbonyl of 81Lys in all except the *T. thermophilus* SOD, where the acceptor group is a threonine three residues downstream from the lysine position. In the *E. coli* SOD X-ray structure, the conserved asparagine side chain amide appears to have a conformation in which it is reversed with respect to the other structures with its oxygen near to a main chain carbonyl, perhaps indicating a modelling error.

Helix six starts with the fully conserved 187Trp. This is involved in extensive van der Waals and hydrophobic interactions with a number of other residues. Those common to all SODs are: 91Pro, 106Phe, 185Val and 190Val. The helix then continues until one or two residues before the C-terminal end of the enzyme, finishing at 198Thr *(M. tuberculosis)*, 190Leu (*E. coli*), 197Lys (human), and 202Lys (*T. thermophilus*). In all cases an additional H-bond was formed between the side chain of the last residue (in the case of *E. coli* the second last, 189Asn) and the main chain carbonyl four residues earlier. In the *E. coli* SOD structure, it is the carbonyl of the asparagine that is apparently within H-bond distance of the main chain carbonyl, indicating another possible modelling oversight.

### *Conservation and invariance of residues close to the catalytic centre*

In addition to the conserved ligand residues there is a number of other residues which are completely conserved. Some of these are thought to partake in the reaction mechanism or to be involved in the formation of the dimer. There is also an extensive hydrophobic shield around the active site, which extends across the interface between the two monomers. The hydrophobic shield is made up of a number of fully conserved or conservatively changed aromatic residues (Table 3). 36Tyr is also implicated in the mechanism for the reaction as it partially blocks the entrance to the active site, and is thought to be involved in a switching movement allowing access of substrate to the active site (Sines *et al.* 1990). It also interacts with bound azide (Lah *et al.* 1995) and, by inference, may interact with bound substrate during catalysis.

We have analyzed the residues forming this hydrophobic shield within 10 Å of the active site to investigate possible correlations between the occurrence of certain residues and the metal ion bound. It turns out that the patterns of conservation are not consistent throughout all members of the enzyme family and sometimes the residues in one SOD may

**Table 3.** Conserved or conservatively changed aromatic residues within 10 Å of the metal ion site

| *M. tuberculosis* (Fe) | *E. coli* (Fe) | human (Mn) | *T. thermophilus* (Mn) |
|---|---|---|---|
| 28His | 26His | 26His | 28His |
| 32His | 30His | 30His | 32His |
| 33His | 31His | 31His | 33His |
| 36Tyr | 34Tyr | 34Tyr | 36Tyr |
| 74Val | 71Trp | 72Ile | 81Leu |
| 76His | 73His | 74His | 83His |
| 78Ile | 75Phe | 76Ile | 85Leu |
| 79Trp | 76Tyr | 77Phe | 86Phe |
| 80Trp | 77Trp | 78Trp | 87Trp |
| 125Trp | 122Trp | 123Trp | 132Trp |
| 162Trp | 158Trp | 161Trp | 168Trp |
| 164His | 160His | 163His | 170His |
| 166Phe | 162Tyr | 165Tyr | 172Tyr |
| 167Tyr | 163Tyr | 166Tyr | 173Tyr |
| 177Phe | 173Tyr | 176Tyr | 183Tyr |

**Table 4.** This table is based on alignment of the primary sequences of 23 MnSODs and 17 FeSODs and shows conservation and divergence of residues in these enzymes (Notes: Yes means that Parker & Blake (1988b) identified the residue as being significant in determining metal ion usage. * indicates residues which have been identified as significant by Matsumoto *et al.* (1991) and [1] indicates residues identified by Lah *et al.* (1995).

| *M. tuberculosis* (Fe) | *E. coli* (Fe) | Human (Mn) | *T. thermophilus* (Mn) | Parker & Blake | Conservation (invariant 100%, highly conserved >80%, conserved 50–80%) |
|---|---|---|---|---|---|
| 4Tyr | 2Phe | 2His | 4Phe | No | Aromatic |
| 6Leu | 4Leu | 4Leu | 6Leu | Yes | Invariant in MnSOD, highly conserved FeSOD |
| 7Pro | 5Pro | 5Pro | 7Pro | Yes | Highly conserved |
| 9Leu | 7Leu | 7Leu | 9Leu | Yes | Invariant in MnSOD, highly conserved FeSOD |
| 11Trp | 9Tyr | 9Tyr | 11Tyr | No | Tyrhighly conserved |
| 13Tyr | 11Tyr | 11Tyr | 13Tyr | No | Highly conserved |
| 15Ala | 13Ala | 13Ala | 15Ala | Yes | Highly conserved |
| 16Leu | 14Leu | 14Leu | 16Leu | Yes | Invariant |
| 17Glu | 15Ala | 15Glu | 17Glu | No | Invariant in MnSODs, conserved FeSOD |
| 18Pro | 16Pro | 16Pro | 18Pro | Yes | Invariant |
| 19His | 17His | 17His | 19His | No | Conserved |
| 20Ile | 18Ile | 18Ile | 20Ile | No | Highly conserved MnSODs, conserved FeSODs |
| 25Asn | 23lIe | 23Met | 25Met | Yes* | Conserved MnSODs, none found in FeSODs |
| 26Glu | 24Glu | 24Gln | 26Glu | No | Polar or H-bond residue |
| 27Leu | 25Tyr | 25Leu | 27Leu | Yes* | Highly conserved: Ile or Leu MnSOD, Phe or Tyr FeSOD |
| 28His | 26His | 26His | 28His | Yes | Invariant ligand residue |
| 29His | 27Tyr | 27His | 29His | No | Aromatic except one |
| 31Lys | 29Lys | 29Lys | 31Lys | Yes | Highly conserved |
| 32His | 30His | 30His | 32His | Yes | All except one |
| 33His | 31His | 31His | 33His | Yes | Invariant |
| 36Tyr | 34Tyr | 34Tyr | 36Tyr | Yes[1] | Invariant |
| 37Val | 35Val | 35Val | 37Val | Yes[1] | All except two |
| 41Asn | 39Asn | 39Asn | 41Asn | Yes | All except two |
| 62Asn | 59Ser | 60Leu | 69Ile | Yes* | Tertiary structure not aligned |
| 64Lys | 61Gly | 62Pro | 71Thr | Yes* | No conservation |
| 65Asn | 62Gly | 63Ala | 72Thr | Yes* | No conservation |
| 67Ala | 64Phe | 65Lys | 74Arg | Yes* | MnSOD conserved Lys or Arg, FeSOD conserved Phe |
| 68Phe | 65Asn | 66Phe | 75Asn | No | MnSOD conserved Phe, FeSOD highly conserved Asn |
| 69Asn | 66Asn | 67Asn | 76Asn | No | Highly conserved |
| 71Ala | 68Ala | 69Gly | 77Gly | Yes* | MnSOD highly conserved Gly, FeSOD highly conserved Ala |
| 72Gly | 69Gln | 70Gly | 78Gly | Yes* [1] | MnSOD invariant, FeSOD conserved Gln |
| 73His | 70Val | 71His | 80His | Yes* | MnSOD highly conserved, where FeSOD 72 Gln, Alp; where Gly, His |
| 74Val | 71Trp | 72Ile | 81Leu | Yes* [1] | Where FeSOD 72 Gln, Trp; where Gly, Alp |
| 75Asn | 72Asn | 73Asn | 82Asn | Yes | Highly conserved |
| 76His | 73His | 74His | 83His | Yes | Invariant ligand residue |
| 77Thr | 74Thr | 75Ser | 84Ser | No | MnSOD conserved, FeSOD conserved |
| 78Ile | 75Phe | 76Ile | 85Leu | Yes* | FeSOD highly conserved Phe |
| 79Trp | 76Tyr | 77Phe | 86Phe | Yes* | MnSOD highly conserved, FeSOD conserved Tyr |
| 80Trp | 77Trp | 78Trp | 87Trp | Yes | Invariant |
| 82Asn | 79Cys | 80Asn | 89Leu | Yes* | Cysteines found only in FeSODs |
| 83Leu | 80Leu | 81Leu | 90Leu | Yes | MnSOD highly conserved |
| 85Pro | 82Pro | 83Pro | 92Pro | No | Highly conserved |
| 87Gly | 84Ala | 85Gly | 94Gly | No | MnSOD highly conserved, FeSOD 50% conserved |
| 91Pro | 88Pro | 89Pro | 98Pro | No | Highly conserved |
| 93Gly | 90Gly | 91Gly | 100Gly | Yes | Highly conserved |
| 95Leu | 92Val | 93Leu | 102Leu | No | MnSOD invariant, FeSOD conserved |
| 98Ala | 95Ala | 96Ala | 105Ala | Yes | Conserved |

*continued*

**Table 4** – *continued*

| *M. tuberculosis* (Fe) | *E. coli* (Fe) | Human (Mn) | *T. thermophilus* (Mn) | Parker & Blake | Conservation (invariant 100%, highly conserved >80%, conserved 50–80%) |
|---|---|---|---|---|---|
| 99Ile | 96Ile | 97Ile | 106Ile | Yes | MnSOD invariant, FeSOD highly conserved |
| 103Phe | 100Phe | 101Phe | 110Phe | Yes | Highly conserved |
| 104Gly | 101Gly | 102Gly | 111Gly | Yes | Highly conserved |
| 106Phe | 103Phe | 104Phe | 113Phe | No | Conserved |
| 109Phe | 106Phe | 107Phe | 116Leu | No | MnSOD conserved Phc, FeSOD invariant |
| 110Arg | 107Lys | 108Lys | 117Lys | No | MnSOD conserved, FeSOD conserved Lys |
| 113Phe | 110Phe | 111Leu | 120Leu | No | MnSOD conserved Phe, FeSOD invariant |
| 122Gly | 119Gly | 120Gly | 129Gly | Yes | Highly conserved |
| 123Ser | 120Ser | 121Ser | 130Ser | Yes | Highly conserved |
| 124Gly | 121Gly | 122Gly | 131Gly | No | MnSOD invariant, FeSOD highly conserved |
| 125Trp | 122Trp | 123Trp | 132Trp | Yes | MnSOD invariant, FeSOD highly conserved |
| 127Ala | 124Trp | 125Trp | 134Trp | No | MnSOD conserved, FeSOD conserved Trp |
| 128Leu | 125Leu | 126Leu | 135Leu | No | MnSOD highly conserved, FeSOD invariant |
| 137Leu | 134Leu | 135Leu | 144Leu | Yes | Highly conserved |
| 144Asp | 140Asn | 142Asn | 150Asn | No | MnSOD highly conserved, FeSOD conserved Asn |
| 145His | 141Ala | 143Gln | 151Gln | Yes* | MnSOD highly conserved (His *Halobacteria*). FeSOD Ala except where 72Gly then His or Gln |
| 146Gln | 142Gly | 144Asp | 152Asp | Yes* | MnSOD highly conserved, FeSOD no conservation, hydrophilic residues where 145 His or Gln |
| 148Asn | 144Pro | 146Leu | 154Pro | No | FeSOD highly conserved Pro |
| 150Pro | 146Thr | 148Gly | 154Met | Yes* | No conservation |
| 155Pro | 151Pro | 154Pro | 161Pro | Yes | MnSOD invariant, FeSOD highly conserved |
| 156Leu | 152Leu | 155Leu | 162Ile | No | MnSOD conserved Leu, FeSOD highly conserved |
| 157Leu | 153Leu | 156Leu | 163Val | No | MnSOD conserved Leu, FeSOD highly conserved |
| 158Leu | 154Thr | 157Gly | 164Gly | Yes* | MnSOD conserved Gly, FeSOD highly conserved Thr |
| 160Asp | 156Asp | 159Asp | 166Asp | Yes | Invariant – ligand |
| 161Met | 157Val | 160Val | 167Val | No | MnSOD highly conserved, FeSOD highly conserved Val |
| 162Trp | 158Trp | 161Trp | 168Trp | Yes | Invariant |
| 163Glu | 159Glu | 162Glu | 169Glu | Yes | Invariant |
| 164His | 160His | 163His | 170His | Yes | Invariant ligand residue |
| 165Ala | 161Ala | 164Ala | 171Ala | Yes | MnSOD conserved (Ser *Halobacteria*), FeSOD invariant |
| 166Phe | 162Tyr | 165Tyr | 172Tyr | Yes | Highly conserved Tyr |
| 167Tyr | 163Tyr | 166Tyr | 173Tyr | Yes | Invariant |
| 169Gln | 165Asp | 168Gln | 175Lys | Yes* | MnSOD conserved Gln, FeSOD highly conserved Asp |
| 170Tyr | 166Tyr | 169Tyr | 176Tyr | No | MnSOD highly conserved, FeSOD conserved |
| 171Lys | 167Arg | 170Lys | 177Gln | Yes* | MnSOD conserved Lys, FeSOD Arg conserved |
| 172Asn | 168Asn | 171Asn | 178Asn | Yes | MnSOD highly conserved (Pro *Halobacteria*), FeSOD invariant |
| 174Lys | 170Arg | 173Arg | 180Arg | No | MnSOD highly conserved, FeSOD highly conserved Arg |
| 177Phe | 173Tyr | 176Tyr | 183Tyr | Yes | MnSOD highly conserved (Phe *Halobacteria*) |
| 182Trp | 178Trp | 181Trp | 188Trp | Yes[1] | MnSOD highly conserved (Phe *Halobacteria*), FeSOD conserved |
| 184Val | 180Leu | 183Val | 190Val | Yes* | MnSOD conserved, FeSOD conserved Leu |
| 185Val | 181Val | 184Ile | 191Leu | No | FeSOD highly conserved |
| 186Asn | 182Asn | 185Asn | 192Asn | Yes | MnSOD conserved, FeSOD highly conserved |
| 187Trp | 183Trp | 186Trp | 193Trp | Yes | Invariant |
| 189Asp | 185Phe | 188Asn | 195Val | Yes* | FeSOD Phe conserved |

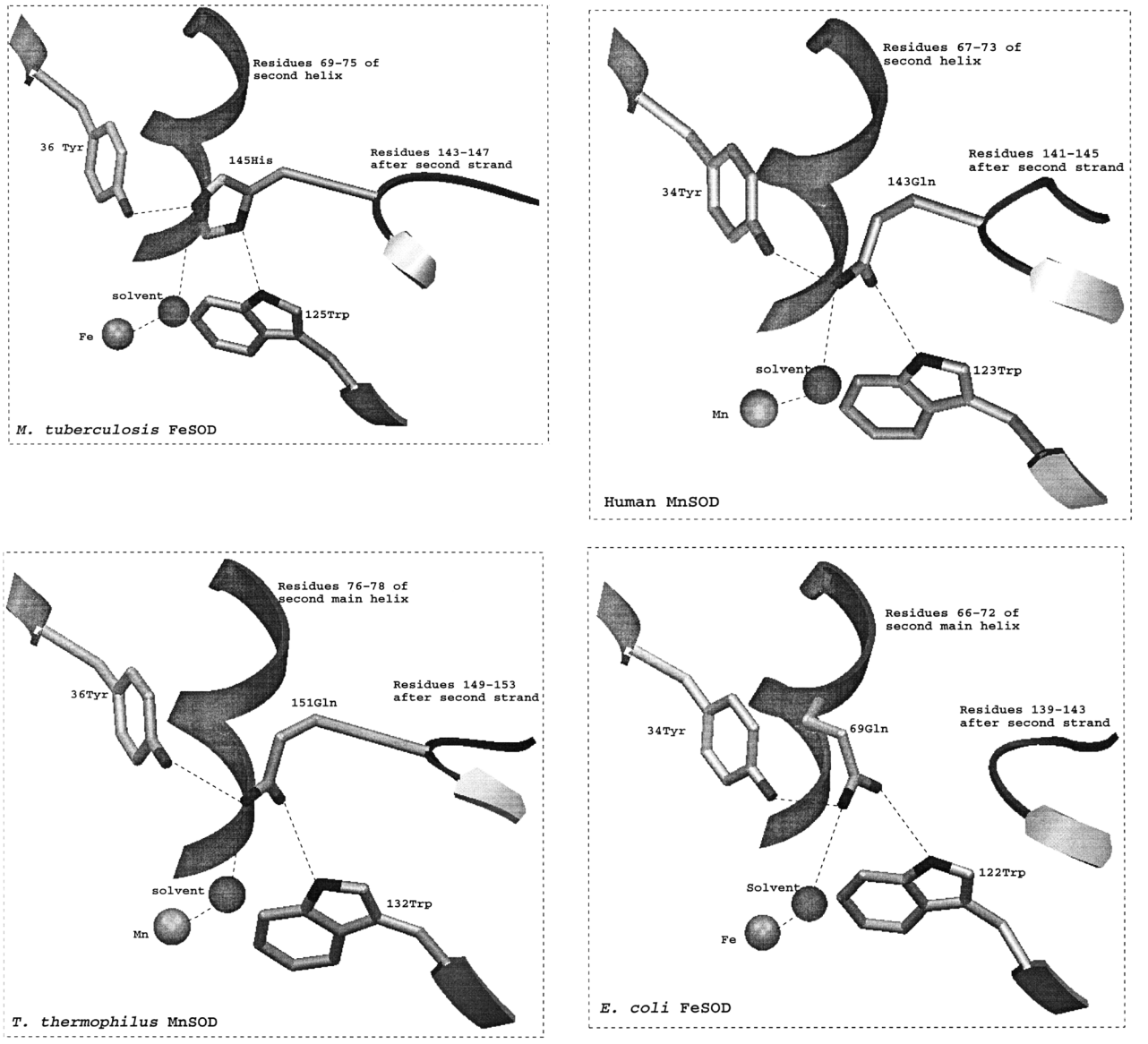


**Figure 3.** The interaction with the solvent molecule found in the four structures indicating the spatial similarity of the interactions (a) *M. tuberculosis,* (b) human, (c) *T. thermophilus*, (d) *E. coli.*

have more similarity to the pattern of residues in SODs utilizing the opposite metal ion.

There is only one absolutely consistent difference between manganese- and iron-dependent SODs, this is at residue 78. In MnSODs this is a large aliphatic residue, e.g. isoleucine or leucine, whereas in the FeSODs these residues generally do not occur and the amino acid at this position is usually aromatic. In the cambialistic SODs they are all leucine, except for *P. shermanii,* which has a valine in this position. The effects these differences would have on specificity are unclear.

Table 4 sets out those residues which are fully conserved or may assist in distinguishing between the two classes of SOD (Mn versus Fe). The previous most detailed study of metal ion determinants in the Mn or FeSOD family was undertaken by Parker & Blake (1988b) prior to the availability of highly refined SOD structures and much of the sequence data. How their results relate to this exercise is itemized in an individual column in Table 4. In comparing their data with those obtained using the solved high resolution structures it is clear that many of their conclusions remain valid.

As identified earlier in the discussion, the metal ion specificity may depend on the residue which interacts with the metal ion's solvent ligand. In the case of *M. tuberculosis* SOD this is 145His. Figure 3 a, b, c, d shows these interactions and the relationship with the metal ion for each of the four SODs under consideration. It has been postulated that the iron and manganese SODs can be distinguished by these interactions (Parker & Blake 1988b).

As can be seen the same set of interactions occurs in all four SODs with the sole difference relating to the residue which co-ordinates the solvent ligand. This is 69Gln in *E. coli,* 145His in *M. tuberculosis,* 143Gln in human and 151Gln in *T. thermophilus.* The nature and position of this residue may have an influence on the bound metal ion as it H-bonds with the metal's solvent ligand (putative $OH^-$). In MnSODs this residue is located in the connecting loop between the second and third β-strands (*T. thermophilus* 151Gln), whereas in FeSODs the functionally equivalent residue is generally in the second α-helix (*E. coli* 69Gln). While the location of this interaction remains constant in all known sequences of MnSODs, the halobacteria replace the glutamine with a histidine. In the FeSODs, however, there are four sequences where the interacting residue is in the position typical of the MnSODs. These are *M. thermoautotrophicum* and *M. tuberculosis* where the residue is a histidine (*M. tuberculosis* 145His), and *Tetrahymena pyrformis* and *Enterobacter aerogenes* where it is a glutamine. The surrounding residues in these FeSODs are also more consistent with those of an MnSOD.

A conserved aspartate which immediately follows the conserved *T. thermophilus* 151Gln (discussed above) is involved in a conserved interaction. In both *T. thermophilus* and human SODs the aspartate interacts with a base, 74Arg (*T. thermophilus*) and 65Lys (human) which are in the same positions in their respective structures. A base is conserved at this point in all but the *Halobacteria* where it is a threonine, *M. leprae* (alanine) and *P. aeruginosa* (valine). This interaction is not found in *M. tuberculosis* as the base is replaced with an alanine, and the electron density indicates that the glutamine located where the aspartate is in the *T. thermophilus* and human SODs is involved in H-bonding with the local main chain, the side chain carbonyl interacting with the main chain amide of the preceding residue. However, the interaction is not implausible for *M. thermoautotrophicum, T. pyriforma* and *E. aerogenes*, all of which have residues capable of hydrogen bonding at both points, the latter two having the aspartate and arginine to form a salt bridge. Although these polar interactions are close to the active site it is not clear how they could affect metal ion specificity.

It is clear that there may be a number of different causes of the differences in metal ion specificity. While the location and nature of the residue which ligates the putative $OH^-$ (His or Gln) could be one cause it may not be the sole or exclusive cause. What is the significance of the residue being a histidine rather than a glutamine? It is worth noting that the halobacterial SODs are the only Mn-dependent SODs with histidine at this position since all of the other MnSODs have a glutamine interacting with the metal site.

It has been observed (Takao *et al.* 1990) that the otherwise fully conserved 31Lys is changed to a threonine in halobacteria. This lysine is thought to be involved in the reaction mechanism, or the process of channelling substrate to the active site and in the halobacteria it has been postulated that the threonine substitution is a compromise between the needs of the enzyme as a catalyst and the necessity for it to be stable in a hostile environment. 67Ala is fully conserved in all mycobacterial SODs; however, in other MnSODs lysine and arginine predominate (these form the salt bridge with the aspartate which follows the residue binding the metal's solvent ligand, as described above), with the halobacteria having threonine and *P. aeruginus* valine; by contrast, whereas in FeSODs it is phenylalanine, with the exceptions of *M. thermoautotrophicum* (serine), *T. pyriforma* and *E. aerogenes* (arginine), and *Entamoeba histolyca* (tyrosine). Other than in the last species, the FeSODs which differ (i.e. do not have phenylalanine at position 67) are also those which do not conform to the typical FeSOD interaction of residue 72 (generally Gln) with the metal's solvent ligand. Both the human and *T. thermophilus* SODs have basic residues here which form salt bridges with the aspartate following the glutamine which interacts with 36Tyr (see above). This interaction is possible for all MnSODs except for those already mentioned; in the halobacterial SODs the equivalent residues might form H-bonds, but modelling indicates that this is unlikely given the distances between the residues. With the exception of *M. tuberculosis* and *E. histolyca* these H-bonds are possible in the FeSODs mentioned above, and *E. aerogenes* could form a salt bridge here. Again it is not clear how these patterns of interactions could affect metal specificity.

In the quaternary structure of *T. thermophilus* 128Phe and 75Asn from the separate subunits stack

against each other. However, in the *M. tuberculosis* SOD the equivalent interacting residues are 121Gln and 68Phe. A similar interaction can be found in all SODs except the halobacteria and the FeSOD from *P. aeruginosa*. In the tertiary structure of both Fe and MnSODs where the phenylalanine is in the second domain, the first domain residue is an asparagine. However, when the phenylalanine is in the first domain, it is glutamine in the second, except for *M. thermoautotrophicum* where it is a glutamate. However, whether the phenylalanine resides in the first or second domain does not correlate with the metal ion present.

In those FeSODs which have the interaction typical of FeSODs with the metal ion's solvent ligand (as in *E. coli*'s 69Gln), residue 74 is a tryptophan and 73 (*M. tuberculosis* numbering) is aliphatic, apart from *Coxiella burnetii* SOD where it is a histidine. However, in those few FeSODs which do not obey this rule there is a reversal. This includes *M. tuberculosis* where residue 73 is a histidine (it can be tyrosine in other FeSODs) and 74 an aliphatic. In MnSODs this is maintained, with residue 73 being a conserved histidine, except for yeast SOD where it is a phenylalanine, and 74 being aliphatic, except for yeast SOD where a threonine is found and in one of the halobacteria where tyrosine occurs at this position. Residue 72 is a fully conserved glycine in all MnSODs and those FeSODs which have the MnSOD type of interaction. However, this glycine becomes a glutamine which interacts with the metal's solvent ligand in typical FeSODs.

## Conclusion

It is clear from the above discussion that there are comparatively few areas where there are significant differences in the primary sequence of Fe and MnSODs, particularly in those areas close to the active site which may therefore have some influence on metal specificity.

There is a number of strongly conserved hydrophobic residues as set out in Table 4. Most of these have been identified as taking part in specific interactions relating to the function of the enzyme or forming the hydrophobic shield around the active site. However, there are many other hydrophobic residues which play no part in either of these roles but are strongly conserved due to their role in forming the general hydrophobic core of the protein. These are: 4, 6, 9, 11, 13, 15, 16, 20, 95, 99, 103, 106, 109, 113, 127, 128, 156, 157, 182, 187. They all form extensive van der Waals contacts with equivalent residues in each of the four structures, and therefore are probably important in the stability of the enzyme.

One of the major differences between a typical FeSOD or MnSOD relates to the residues which interact with the metal's solvent ligand. It is possible to conclude from data so far available that all MnSODs have a glutamine or sometimes a histidine at *M. tuberculosis* residue position 145. In FeSODs, a glutamine at *E. coli* residue position 69 predominates and interacts with the metal's solvent ligand instead of residue 145 which becomes alanine. However, there is a number of other FeSODs which do not conform to this, *M. tuberculosis* being typical of these, having interactions more typical of MnSODs. However, the Mn/Fe SOD *Porphyromonas gingivalis* is typical of an FeSOD, but is active whichever ion is present. The other Mn/Fe nonspecific SODs have the combination of residues more typical of an MnSOD, but as has already been established, there is a number of FeSODs which have this configuration.

It would be interesting in a typical FeSOD (such as the *E. coli* enzyme) and the typical MnSODs (such as the human and *T. thermophilus* enzymes) to do a series of site-directed mutagenesis experiments to try to determine which one or combination of changes might influence the metal ion specificity. The region of greatest variety is at the beginning of the last helix in the first domain; here not only is there the specific difference at residue 72 where a typical FeSOD would have a glutamine which interacts with the metal's solvent ion, but there is a number of other residues between 67 and 79 which are different in Mn and FeSODs. These are identified in Table 5. In the middle of this region is the fully conserved second metal ion ligand at position 76 and immediately preceding it is a highly conserved asparagine. Nevertheless, on both sides of these residues are other residues which are significantly different in the typical forms of the two types of SOD. Together with those residues at position 145 and 146, these form two areas worthy of an investigation to try and ascertain what residue profile might determine the metal ion specificity. Other specific residues worth investigation are 25 where a conserved leucine in FeSODs is a conserved methionine in MnSODs, and similarly residue 27 which changes from being aromatic to aliphatic. Similar changes are found within the proximity of ligand residue 76, where residue 74 is a tryptophan in FeSODs and aliphatic in MnSODs. A similar change happens at residue 78 from aromatic to

**Table 5.** Residues which predominate in each form of SOD (*M. tuberculosis* numbering throughout)

| Residue number | Residue type in FeSODs | Residue type in MnSODs | Possible function |
|---|---|---|---|
| 25 | Leucine | Methionine | Structural (active site) |
| 27 | Phe or Tyr | Ile or Leu | active site |
| 67 | Phenylalanine | Base residue | active site |
| 68 | Asparagine | Phenylalanine | active site |
| 70 | Alanine | Glycine | active site |
| 71 | Alanine | Glycine | active site |
| 72 | Glutamine | Glycine | active site |
| 73 | Aliphatic | Histidine | active site |
| 74 | Tryptophan | Aliphatic | active site |
| 77 | Threonine | Serine | active site |
| 78 | Phenylalanine | Leucine | active site |
| 79 | Tyrosine | Phenylalanine | active site |
| 145 | Alanine | Glutamine | active site |
| 146 | No conservation | Aspartate | active site |
| 158 | Threonine | Glycine | structural |
| 169 | Aspartate | Glutamine | active site |
| 189 | Phenylalanine | Polar | structural |

**Table 6.** List of residues in the *M. tuberculosis* SOD attributing them to the different SOD class of which they are most typical

| Residue | Typical of Mn or FeSODs |
|---|---|
| 25Asn | Neither |
| 27Leu | MnSOD |
| 67Ala | either (but found in *M. leprae*, an MnSOD) |
| 68Phe | MnSOD (but not in *M. leprae)* |
| 70Leu | Neither (but found in *M. leprae*) |
| 71Ala | FeSOD |
| 72Gly | MnSOD |
| 73His | MnSOD |
| 74Val | MnSOD |
| 77Thr | FeSOD |
| 78Ile | MnSOD |
| 79Trp | Neither (but found in *M. leprae*) |
| 127Ala | Neither (Val in *M. leprae*) |
| 130Trp | Neither |
| 144Asp | Neither (but found in *M. leprae*) |
| 145His | Neither |
| 146Gln | Neither (but found in *M. leprae*) |
| 149Phe | Neither |
| 158Leu | Neither |
| 169Gln | MnSOD |
| 189Asp | MnSOD |

aliphatic. Residue 158 may also be worth altering as it is threonine in FeSODs and glycine in MnSODs. These observations are based entirely on the known sequences and structures but the underlying reasons for the changes in metal ion specificity are not clear.

The *M. tuberculosis* and other similar FeSODs need a different consideration. Table 6 indicates which *M. tuberculosis* SOD residues are typical of the different types of SOD and which are not typical of either, e.g. 25 is an asparagine which becomes a non polar residue in all other SODs. Residues 144,145 and 146 also form a unique combination; the only MnSODs where histidine is found at 145 are the halobacterial enzymes. Residue 158 in this SOD is a leucine, i.e. distinctly different again from the threonine or glycine found in the other SODs. The presence of a tryptophan at 130 and not at 127 may also be worth investigating, even though these positions are at some distance from the active site.

Obviously further studies are necessary before the reasons for the different specificities can be established. Hopefully site-directed mutagenesis will identify the residues responsible and lead to further studies to ascertain the underlying chemical and physical reasons.

## References


Almansa MN, Palma JM, Yanez J, Del Rio LA, Sevilla F. 1991 Purification of an iron superoxide dismutase from citrus plant *Citrus limonum* R. *Free Rad Res Comm* **12–13,** 319-328.

Amano A, Shizukuishi S, Tamagawa H, Iwakura K, Tsunemitsu A. 1990 Characterisation of superoxide dismutases purified from either anaerobically maintained or aerated *Bacteroides gingivalis. J. of Bacteriology* **172,** 1457–1463.

Battistoni A, Rotilio G. 1995 Isolation of an active and heat stable monomeric form of Cu,Zn superoxide dismutase from the periplasmic space of *Escherchia coli. FEBS Lett* **374**, 199–202.

Battistoni A, Folcarelli S, Gabbianelli R, Capo C, Rotilio G. 1996 The Cu,Zn superoxide dismutase from *Escherchia coli* retains monomeric structure at high protein concentrations (evidence for altered subunit interactions in all the bacteriocupreans. *Biochem J* **320**, 713–716.

Beyer WF, Fridovich I. 1991 *In vivo* competition between iron and manganese for the occupancy of the active site region on the manganese superoxide dismutase of *Escherchia coli*. *J Biol Chem* **266** (1), 303–308.

Borgstahl GEO, Parge HE, Hickey MJ, *et al.* 1992 The structure of human mitochondrial manganese superoxide dismutase reveals a novel tetrameric interface of two 4-helix bundles. *Cell* **71,** 107–118.

Chen Y, Gregory EM. 1991 *In vivo* metal substitution in *Bacteriodes fragilis* superoxide dismutase. *Free Rad Res Comm* **12-13,** 313–318.

Cooper JB, McIntyre F, Badasso MO, Wood SP, Zhang Y, *et al.*. 1995 X-ray structure analysis of the iron dependent superoxide dismutase from *Mycobacterium tuberculosis* at 2.0 Å resolution reveals novel dimer–dimer interactions. *J Mol Biol* **246,** 531–544.

Efimov AF. 1991 Standard structures in proteins. *Prog Biophys Molec Biol* **4** (3), 245–250.

Johnson MS, Overington JP. 1993 A structural basis for sequence comparisons – an evaluation of scoring methodologies. *J Mol Biol* **233,** 716–738.

Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. 1991 Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr* **A47,** 110–119.

Kabsch W, Sander S. 1983 Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22,** 2577–2637.

Lah MS, Dixon MM, Pattridge KA, *et al.* Structure-function in *Escherichia coli* iron superoxide dismutase: comparisons with the manganese superoxide dismutase of *Thermus thermophilus. Biochemistry* **34(**5), 1646–1660.

Ludwig ML, Metzger AL, Pattridge KA, Stallings WC. 1991 Manganese superoxide dismutase from *Thermus thermophilus*: a structural model refined at 1.8Å resolution. *J Mol Biol* **219,** 335–358.

Matsumoto T, Terauchi K, Isobe T, Matsuoka K, Yamakure F. 1991 Iron- and manganese-containing superoxide dismutases from *Methylonmonas* J: identity of the protein moiety and amino acid sequence. *Biochemistry* **30**, 3210–3216.

Meier B, Sehn A, Schinina ME, Barra D. 1994 *In vivo* incorporation of copper into the iron-exchangeable and manganese-exchangeable superoxide dismutase from *Propionibacterium shermanii. Eur J Biochem* **219**, 463–468.

Parker MW, Blake CCF. 1988a Crystal structure of manganese superoxide dismutase from *Bacillus stearothermophilus* at 2.4Å resolution. *J Mol Biol* **199,** 649–661.

Parker MW, Blake CCF. 1988b Iron and manganese containing superoxide dismutases can be distinguished by analysis of their primary structures. *FEBS Lett* **229**(2), 377–382.

Sibanda BL, Blundell TL, Thornton JM. 1989 Conformation of β-hairpins in protein structures a systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J Mol Biol* **206,** 759–777.

Sines J, Allison S, Wiezbicki A, McCammon JA. 1990 Brownian dynamics simulation of the $O_2^-$ superoxide dismutase reaction: iron and manganese enzymes. *J Phys Chem* **94,** 959–961.

Stoddard BL, Howell PL, Ringe D, Petsko GA. 1990 The 2.1Å resolution structure of iron superoxide dismutase from *Pseudomonas ovalis. Biochemistry* **29,** 8885–8893.

Sutcliffe MJ, Haneef I, Carney D, Blundell TL. 1987 Knowledge-based modelling of homologous proteins I: three dimensional frameworks. *Protein Eng* **1,** 377–384.

Tainer JA, Getzoff ED, Beem KM, Richardson JS, Richardson DC. 1982 Determination and analysis of the 2Å structure of copper, zinc superoxide dismutase. *J Mol Biol* **160**, 181–217.

Takao M, Oikawa A, Yasui A. 1990 Characterisation of a superoxide dismutase gene from the archaebacterium *Methanobacterium thermoautotrophicum. Arch Biochem Biophys* **283** (1), 210–216.

Tierney DL, Fee JA, Ludwig ML, Penner-Hahn JE. 1995 X-ray absorption spectroscopy of the iron site in *Escherichia* coli Fe(III) superoxide dismutase. *Biochemistry* **34**(5), 1661–1668.

van Camp W, Bowler C, Villarroch R, *et al.* 1990 Characterisation of iron superoxide dismutase cDNAs from plants obtained by genetic complementarities in *Escherchia coli. Proc Nat Acad Sci USA* **87**(24), 9903-9907.

Wagner UG, Pattridge KA, Ludwig ML, Stallings WC, Werber MM, *et al.* 1993 Comparison of the crystal structures of genetically engineered human manganese superoxide dismutase from *Thermus thermophilus*: differences in dimer–dimer interaction. *Protein Science* **2,** 814–825.

Yamakura F, Matsumoto T, Terauchi K. 1991 Isolation of manganese superoxide dismutase and low active iron superoxide dismutase from *Methylomonas* J consisting of identical proteins. *Free Rad Res Comm* **12–13,** 329–334.