Iron Superoxide Dismutase

NUCLEOTIDE SEQUENCE OF THE GENE FROM ESCHERICHIA COLI K12 AND CORRELATIONS WITH CRYSTAL STRUCTURES*

Antoine Carlioz†, Martha L. Ludwig§, William C. Stallings‡, James A. Fee¶, Howard M. Steinman∥, and Danièle Touati†

From the Institut Jacques Monod, Tour 43, 2, Place Jussieu, 75251 Paris Cedex 05, France, the Biophysics Research Division and Department of Biological Chemistry, The University of Michigan, Ann Arbor, Michigan 48109, the Division of Isotopes and Nuclear Chemistry, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, and the Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461

The nucleotide sequence of the iron superoxide dismutase gene from *Escherichia coli* K12 has been determined. Analysis of the DNA sequence and mapping of the mRNA start reveal a unique promoter and a putative p-independent terminator, and suggest that the Fe dismutase gene constitutes a monocistronic operon. The gene encodes a polypeptide product consisting of 192 amino acid residues with a calculated M of 21,111. The published N-terminal amino acid sequence of *E. coli* B Fe dismutase (Steinman, H. M., and Hill, R. L. (1973) Proc. Natl. Acad. Sci. U. S. A. 70, 3725–3729), along with the sequences of several other peptides reported here, was located in the primary structure deduced from the K12 *E. coli* gene sequence.

A new molecular model for iron dismutase from *E. coli*, based on the DNA sequence and x-ray data for the *E. coli* B enzyme at 3.1 Å resolution, allows detailed comparison of the structure of the iron enzyme with manganese superoxide dismutase from *Thermus thermophilus* HB8. The structural similarities are more extensive than indicated by earlier studies and are particularly striking in the vicinity of the metal-ligand cluster, which is surrounded by conserved aromatic residues. The combined structural and sequence information now available for a series of Mn and Fe superoxide dismutases identifies variable regions in these otherwise very similar molecules; the principal variable site occurs in a surface region between the two long helices which dominate the N-terminal domain.

Iron and manganese superoxide dismutases from bacteria are attractive subjects for structure analysis and for genetic/functional studies (1–5). Protein or DNA sequence analyses of iron superoxide dismutases and manganese superoxide dismutases include the determination of the primary structures of four Mn superoxide dismutases (6–10), numerous N-terminal sequences (11, 12), and, most recently, the complete polypeptide sequence of the Fe superoxide dismutase from *Photobacterium leignathi* (13). In addition, x-ray structures of the Fe superoxide dismutases from *Escherichia coli* B (14) and *Pseudomonas ovalis* (15) and the Mn superoxide dismutase from *Thermus thermophilus* HB8 (16) have been reported. Taken together, these data suggest that Fe superoxide dismutases and Mn superoxide dismutases share a common polypeptide fold which is completely unlike that of Cu/Zn superoxide dismutase, and correlation of side chain shapes in the Mn superoxide dismutase electron density function with known sequences has identified the four protein ligands to the Mn(III) cofactor. The strong similarity in the polypeptide fold of Mn and Fe superoxide dismutases (17) and identities in their amino acid sequence alignments (18, 19) imply that ligands to the Fe(III) and Mn(III) cofactors would be chemically identical residues located in equivalent positions of the three-dimensional structure.

The iron superoxide dismutase gene from *E. coli* K12 has been cloned previously (20, 21). We have determined its nucleotide sequence and compared it with the reported gene sequence for *E. coli* K12 Mn superoxide dismutase (6). The protein sequence deduced from the cloned DNA is shown to be consistent with the sequence of tryptic fragments of the enzyme from the *E. coli* B bacterium. The sequence has been used to complete the three-dimensional model of the enzyme and to interpret details of the 3.1-Å electron density map, particularly in the metal-ligand cluster and its environment.

The results confirm the structural equivalence of the cofactor ligands in Fe and Mn dismutases and demonstrate that the ligand clusters are surrounded by remarkably similar chemical milieu. Comparisons of the protein sequence with other superoxide dismutases not only verify the homologies within the Fe and Mn superoxide dismutase family of enzymes but also locate variable regions in the structures.

MATERIALS AND METHODS

**Bacterial Strains and Plasmids**—Bacterial strain *E. coli* K12 71/18 (22) was used as host in the construction of the plasmid bank, and *E. coli* K12 QC774 (23) was used in complementation tests for superoxide dismutase activity. Plasmids pHS1-6 and pHS1-8, carrying the sosB gene, were used for DNA sequencing (Fig. 1). They carry bacterial DNA inserts which have been obtained by partial deletions of the insert carried by pHS1-4 (20)·pEMBL19 (24), a derivative from pUC19 (25) carrying the Fl origin of replication, was used as a vector for DNA sequencing.

**Reagents and Enzymes**—DNA restriction enzymes, T4 DNA ligase, DNA polymerase I, polynucleotide kinase, and reverse transcriptase

*This work was supported by Action Thématique Programmée 960114 from the Centre National de la Recherche Scientifique and by Grants GM16429 (to M. L. L. and F.), GM35189 (to J. A. F.), and GM24527 (to H. M. S.) from the National Institutes of Health. Institut J. Monod is a Centre National de la Recherche Scientifique/Université Paris VII institute. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EMBL Data Bank with accession number(s) J03511.

(Received for publication, May 1, 1987)
were purchased from Bethesda Research Laboratories or from Boehringer (Mannheim, Federal Republic of Germany). [35S]DATP and dCTP and [32P]dATP were from Amersham (Amersham, United Kingdom) and。www.sciencemag.org

Preparation of DNA and DNA Sequence Determination—Plasmid DNA was purified essentially as described by Birnboim and Doly (26). Single strand DNA was prepared as described (22) using the M13 derivative, M13K07 (Pharmacia, Uppsala, Sweden) as superinfecting phage instead of fi; the DNA sequence was determined by the Sanger dye-deoxy chain termination method (27). Ambiguities in a GC-rich region were eliminated by using inosine instead of guanine in sequence reactions (28).

Determination of the Starting Point of the Transcript—The mRNA start was localized by primer extension (29). RNAs were prepared as described (29) from strain AB3465/pHS1-4 (20). DNA probes were labeled using the Maxam and Gilbert method (30), and reverse transcriptase DNA synthesis was performed according to Débarbouille and Raibaud (29). Products of the reactions were analyzed on a 7% polyacrylamide gel.

Sequence Comparisons—Comparisons of primary structure were performed at the Centre Interuniversitaire de Traitement de l'Information (CITI 2, Paris) using the DIAGON program of Staden (31) which incorporates the MDM78 scoring system (32, 33).

Purification and Sequence Determination of Peptides from Superoxide Dismutase of E. coli B—Using previously published methods, the Fe superoxide dismutase was purified from E. coli B (34), then reduced and S-carboxymethylated with [14C]iodoacetic acid, and digested with tosylphenylalanyl chloromethyl ketone-treated trypsin (35). Tryptic peptides were initially fractionated on Sephadex G-50SF in 0.1 M ammonia, 5% (v/v) 1-propanol, and then individual peaks were rechromatographed on Lichrosorb RP8, using a gradient of 0-60% 1-propanol in 0.1% (v/v) trifluoroacetic acid. Purified peptides were sequenced by a micro version of the manual Edman degradation (36) at the University of Michigan Protein Sequencing Facility.

X-ray Analysis of Fe Superoxide Dismutase from E. coli B—The original Fourier map, calculated at 3.1 Å using multiple isomorphous replacement phases, was modified by averaging of the crystallographically independent subunits (14). Calculated phases derived by Fourier inversion of the solvent-leveled averaged map were combined (37) with starting phases, and the resulting map was again averaged and refined (38). The Fe superoxide dismutase model on an Evans and Sutherland graphics display with the aid of FRODO (38). The building was guided by the locations of Cα atoms which had been positioned in the earlier map.

Fitting the sequence derived from E. coli K12 DNA to the density relied on recognition of characteristic residue shapes, particularly tyrosine or phenylalanine, tryptophan, histidine, proline, and leucine. For example, the series of aromatics in the sequence 100-118 provided unambiguous markers in this region. We were able to position the entire sequence by following the published trace (14) of the polypeptide chain. At several locations the side chain density was not extensive enough to accommodate the residues expected from the sequence. With the moderate resolution of the current map, we cannot discern whether these discrepancies represent real differences between the Fe superoxide dismutase of E. coli strains B and K12. However, the x-ray results do support the presence of the seven tryptophan residues deduced from the DNA sequence.

RESULTS AND DISCUSSION

Nucleotide Sequence Determination of the Iron Superoxide Dismutase Gene (sodB) of E. coli K12—We have previously cloned the Fe superoxide dismutase gene in plasmid pH51-4 (20). In the current study, the location of the coding region within the plasmid and its direction of transcription were determined prior to sequencing. The 5' end of the structural gene was deduced to lie about 600 bp upstream of the EcoRI site (Fig. 1), given that the polypeptide chain is about 200 amino acid residues in length. This location was derived from restriction mapping of the sodB-kan fusion previously obtained by insertion of Mu transposons into the plasmid (23) and by analyzing plasmid subclones for complementation of a sodA sodB double mutant (for growth on minimal medium) and expression of Fe superoxide dismutase activity in a wild type strain, as previously described (20, 23, 35). The sequence of 970 bp, which includes the structural gene and flanking regions, was determined (Fig. 3). The sequencing strategy is summarized in Fig. 2. A single open reading frame was found beginning at the ATG at nucleotide 177. Tandem termination codons appear at nucleotides 756 and 759. A putative Shine-Dalgarno sequence (40) was identified 7 bp upstream of the ATG codon, and a putative transcriptional terminator containing an inverted repeat preceding a stretch of T residues (41) was identified at nucleotides 887-908. It is of interest to note that low Fe superoxide dismutase overproduction was observed in strains harboring pH51-6 (Fig. 1) in which this putative transcription termination signal is deleted, beginning at the ClaI site (nucleotides 818-823 in Fig. 3).

The 5' end of the mRNA was identified as A-122 by primer extension mapping (Fig. 4). The corresponding promoter sequences at -10(TACCCT) and -35 (TTGCT) agree well with known consensus promoter sequences (42, 43) and suggest a rather strong promoter (homology score, 53.8%) as predicted by the rules established by Mulligan et al. (44). This is in good agreement with the high in vivo level of the protein and with the high level of neomycin phosphotransferase in the sodB-kan fusion.2

The coding region of 576 bp predicts an amino acid sequence of 192 residues and a subunit molecular weight of 21,111 after methionine cleavage. The amino acid composition derived from this sequence is in good agreement, except for tryptophan, with that determined by hydrolysis of Fe superoxide dismutase from E. coli B (45). Furthermore, the amino-terminal 29 residues agree exactly with the amino-terminal

1 The abbreviation used is: bp, base pair(s).

2 A. Carlioz and D. Touati, unpublished observation.
Fig. 2. Sequencing strategy. The DNA sequence was determined using the Sanger dideoxy chain termination method on subfragments cloned in the plasmid pEMBL19 (see "Materials and Methods"). The Pet fragment is from phS1-8 (see Fig. 1). The arrows indicate the restriction sites used as well as the direction and extent of the restriction fragments. The duplexes are derived from nucleic acid sequencing; residues of a peptide encompassing amino acids 30-43 (beyond the starting point) were sequenced from a distinct site. The abbreviations for restriction enzymes are: N, NdeI and as in Fig. 1.

Fig. 3. Nucleotide sequence of the sodB gene region and the deduced polypeptide sequence. The Shine-Dalgarno sequence is underlined; initiation and termination codons are boxed with thick lines; relevant restriction sites are boxed with thin lines. Palindromic regions are shown by inverted horizontal arrows. P indicates the initiation site of mRNA synthesis. Promoter consensus bases are boxed with dotted lines. Small vertical arrows point to conserved metal ligands.

Fig. 4. Reverse transcriptase mapping of the starting point of transcription. In vivo synthesized RNA extracted from the strain AB2463 phS1-4 was hybridized to the 108-base Asp-718-Nle-1 restriction fragment (positions 313-320 on Fig. 3), labeled at the Asp side, which was used as a DNA probe and extended with reverse transcriptase. The products of reaction were analyzed by 7% polyacrylamide gel electrophoresis (lane 1); arrows indicate DNA probe position (down) and the point at which the extension products terminate (up). Lane 2, the sequence of the beginning of spoII G operon, used as a reference in the P. Stragier laboratory, was used as sequence ladder, and the products of sequencing reactions were electrophoresed in parallel with those of lane 1. Asp-718 is an isoschizomer of KpnI.

sequence determined from intact E. coli B Fe superoxide dismutase (46). Finally, sequences of several purified tryptic peptides of the E. coli B protein agree exactly with the sequence derived from nucleic acid sequencing: residues 44-50, 51-57, 79-91, 81-91, 92-107, 108-116, and the first 6 residues of a peptide encompassing amino acids 30-43 (beyond...
Homologies in Sequence and Structure—Pairwise comparisons of the primary structures of four Mn dismutases (10) and of the Mn sequences with the recently published sequence of the Fe enzyme from *P. leiognathi* (13) have shown substantial homologies, despite some variability in chain length and in residue composition. Using the alignment program of Staden (31), we have extended these comparisons to include the new sequence for Fe dismutase from *E. coli*. While it is evident from Fig. 5 that large sections of the sequences for Mn and Fe dismutases from *E. coli* are identical, the matching criteria used in Fig. 6 also allow definition of regions that retain sequence similarities and indicate the loci of insertions or deletions.

The x-ray structures of Mn superoxide dismutase from *T. thermophilus* (16) and Fe superoxide dismutase from *E. coli* B (14) were aligned in order to compare structural variations with differences detected solely from sequences. With the aid of the structures, insertions or deletions can be located with respect to characteristic secondary structural features of the molecule. Fig. 7 presents a superposition of the structures of *E. coli* B Fe dismutase and *T. thermophilus* Mn dismutase. At the present stage of the three-dimensional structure analyses we find no divergences in the chain tracings, nor any insertions, except at the chain termini and in the region lying between positions 42 and 56 of Fe dismutase, where the Mn superoxide dismutase chain incorporates a total of 8 "extra" residues. Six of these appear to form a single insertion which forms the loop following residue 60 of the Mn structure.

FIG. 5. Comparison of nucleotide and amino acid sequences of iron and manganese superoxide dismutase from *E. coli* K12. Identical amino acid sequences are enclosed in boxes, and matched nucleotides are indicated by double dots. The Mn superoxide dismutase sequence is from Takeda and Avila (6). No significant sequence alignment was found between the noncoding regions.

| residue | sequence |
|---------|----------|
| 35      | Glu Gly Val Phe |
| 36      | Asp Gly Val Ala |
| 37      | Glu Gly Val Ala |
| 38      | Gly Gly Val Ala |
| 39      | Asp Gly Val Ala |
| 40      | Glu Gly Val Ala |
| 41      | Gly Gly Val Ala |
| 42      | Asp Gly Val Ala |
| 43      | Gly Gly Val Ala |
| 44      | Asp Gly Val Ala |
| 45      | Gly Gly Val Ala |
| 46      | Asp Gly Val Ala |
| 47      | Gly Gly Val Ala |
| 48      | Asp Gly Val Ala |
| 49      | Gly Gly Val Ala |
| 50      | Asp Gly Val Ala |
| 51      | Gly Gly Val Ala |
| 52      | Asp Gly Val Ala |
| 53      | Gly Gly Val Ala |
| 54      | Asp Gly Val Ala |
| 55      | Gly Gly Val Ala |
| 56      | Asp Gly Val Ala |

| residue | sequence |
|---------|----------|
| 57      | Gly Gly Val Ala |
| 58      | Asp Gly Val Ala |
| 59      | Gly Gly Val Ala |
| 60      | Asp Gly Val Ala |
| 61      | Gly Gly Val Ala |
| 62      | Asp Gly Val Ala |
| 63      | Gly Gly Val Ala |
| 64      | Asp Gly Val Ala |
| 65      | Gly Gly Val Ala |
| 66      | Asp Gly Val Ala |
| 67      | Gly Gly Val Ala |
| 68      | Asp Gly Val Ala |
| 69      | Gly Gly Val Ala |
| 70      | Asp Gly Val Ala |
| 71      | Gly Gly Val Ala |
| 72      | Asp Gly Val Ala |
| 73      | Gly Gly Val Ala |
| 74      | Asp Gly Val Ala |
| 75      | Gly Gly Val Ala |
| 76      | Asp Gly Val Ala |
| 77      | Gly Gly Val Ala |
| 78      | Asp Gly Val Ala |
| 79      | Gly Gly Val Ala |
| 80      | Asp Gly Val Ala |
| 81      | Gly Gly Val Ala |
| 82      | Asp Gly Val Ala |
| 83      | Gly Gly Val Ala |
| 84      | Asp Gly Val Ala |
| 85      | Gly Gly Val Ala |
| 86      | Asp Gly Val Ala |
| 87      | Gly Gly Val Ala |
| 88      | Asp Gly Val Ala |
| 89      | Gly Gly Val Ala |
| 90      | Asp Gly Val Ala |
| 91      | Gly Gly Val Ala |
| 92      | Asp Gly Val Ala |
| 93      | Gly Gly Val Ala |
| 94      | Asp Gly Val Ala |
| 95      | Gly Gly Val Ala |
| 96      | Asp Gly Val Ala |
| 97      | Gly Gly Val Ala |
| 98      | Asp Gly Val Ala |
| 99      | Gly Gly Val Ala |
| 100     | Asp Gly Val Ala |

| residue | sequence |
|---------|----------|
| 101     | Gly Gly Val Ala |
| 102     | Asp Gly Val Ala |
| 103     | Gly Gly Val Ala |
| 104     | Asp Gly Val Ala |
| 105     | Gly Gly Val Ala |
| 106     | Asp Gly Val Ala |
| 107     | Gly Gly Val Ala |
| 108     | Asp Gly Val Ala |
| 109     | Gly Gly Val Ala |
| 110     | Asp Gly Val Ala |
| 111     | Gly Gly Val Ala |
| 112     | Asp Gly Val Ala |
| 113     | Gly Gly Val Ala |
| 114     | Asp Gly Val Ala |
| 115     | Gly Gly Val Ala |
| 116     | Asp Gly Val Ala |
| 117     | Gly Gly Val Ala |
| 118     | Asp Gly Val Ala |
| 119     | Gly Gly Val Ala |
| 120     | Asp Gly Val Ala |

| residue | sequence |
|---------|----------|
| 121     | Gly Gly Val Ala |
| 122     | Asp Gly Val Ala |
| 123     | Gly Gly Val Ala |
| 124     | Asp Gly Val Ala |
| 125     | Gly Gly Val Ala |
| 126     | Asp Gly Val Ala |
| 127     | Gly Gly Val Ala |
| 128     | Asp Gly Val Ala |
| 129     | Gly Gly Val Ala |
| 130     | Asp Gly Val Ala |
| 131     | Gly Gly Val Ala |
| 132     | Asp Gly Val Ala |
| 133     | Gly Gly Val Ala |
| 134     | Asp Gly Val Ala |
| 135     | Gly Gly Val Ala |
| 136     | Asp Gly Val Ala |
| 137     | Gly Gly Val Ala |
| 138     | Asp Gly Val Ala |
| 139     | Gly Gly Val Ala |
| 140     | Asp Gly Val Ala |

| residue | sequence |
|---------|----------|
| 141     | Gly Gly Val Ala |
| 142     | Asp Gly Val Ala |
| 143     | Gly Gly Val Ala |
| 144     | Asp Gly Val Ala |
| 145     | Gly Gly Val Ala |
| 146     | Asp Gly Val Ala |
| 147     | Gly Gly Val Ala |
| 148     | Asp Gly Val Ala |
| 149     | Gly Gly Val Ala |
| 150     | Asp Gly Val Ala |
| 151     | Gly Gly Val Ala |
| 152     | Asp Gly Val Ala |
| 153     | Gly Gly Val Ala |
| 154     | Asp Gly Val Ala |
| 155     | Gly Gly Val Ala |
| 156     | Asp Gly Val Ala |
| 157     | Gly Gly Val Ala |
| 158     | Asp Gly Val Ala |
| 159     | Gly Gly Val Ala |
| 160     | Asp Gly Val Ala |

| residue | sequence |
|---------|----------|
| 161     | Gly Gly Val Ala |
| 162     | Asp Gly Val Ala |
| 163     | Gly Gly Val Ala |
| 164     | Asp Gly Val Ala |
| 165     | Gly Gly Val Ala |
| 166     | Asp Gly Val Ala |
| 167     | Gly Gly Val Ala |
| 168     | Asp Gly Val Ala |
| 169     | Gly Gly Val Ala |
| 170     | Asp Gly Val Ala |
| 171     | Gly Gly Val Ala |
| 172     | Asp Gly Val Ala |
| 173     | Gly Gly Val Ala |
| 174     | Asp Gly Val Ala |
| 175     | Gly Gly Val Ala |
| 176     | Asp Gly Val Ala |
| 177     | Gly Gly Val Ala |
| 178     | Asp Gly Val Ala |
| 179     | Gly Gly Val Ala |
| 180     | Asp Gly Val Ala |

FIG. 5. Comparison of nucleotide and amino acid sequences of iron and manganese superoxide dismutase from *E. coli* K12. Identical amino acid sequences are enclosed in boxes, and matched nucleotides are indicated by double dots. The Mn superoxide dismutase sequence is from Takeda and Avila (6). No significant sequence alignment was found between the noncoding regions.
E. coli Fe Superoxide Dismutase: Gene Sequence and X-ray Structure

**FIG. 6. Similarities of superoxide dismutase (SOD) amino acid sequences.** The sequence of Fe superoxide dismutase from *E. coli* K12 is compared with Fe superoxide dismutase from *P. leiognathi* (13), with Mn superoxide dismutases from *B. stearothermophilus* (8), *E. coli* K12 (6), *S. cerevisiae* (9), and human liver (10), and with human Cu/Zn superoxide dismutase (58). The DIAGON matching program of Staden (31) was used. The points correspond to midpoints of 11-residue spans using a score at 1% level expectation. The N termini are at the bottom left and C termini at the top right. Numbers in parentheses indicate the number of amino acids. It is interesting that *E. coli* Fe superoxide dismutase shows more homology with the Mn protein from *B. stearothermophilus* than with its own Mn superoxide dismutase.

Thus, the available three-dimensional structures support the notion that a region with conformational variability occurs around residue 50.

Comparisons of all known Fe and Mn superoxide dismutases at the level of primary sequence (Fig. 6) suggest the generality of genetic variations in the vicinity of residues 45–65, showing divergence in both length and composition in this region. Variability in chain length near positions 90 and 150 (Fe superoxide dismutase numbering) is also evident. Examination of the x-ray structures places the latter insertion/deletion in the crossover connection between the second and third strands of an otherwise conserved β-sheet; the crossover must be longer in *E. coli* Mn superoxide dismutase than in the Mn enzyme from *T. thermophilus* (see Fig. 7B, legend). The segment near residue 90 connects the two domains and has been thought to function as a hinge in the folding and unfolding of the subunit (5, 16) even though its length and composition vary from species to species. It is worth noting that Mn and Fe superoxide dismutases from *E. coli* do not cross-react with polyclonal antibodies (39); while differences in composition and surface charge, which are distributed along the chain, may account in part for antigenic differences, it is likely that the variable regions also play roles in immunogenicity and antibody recognition.

**The Metal-binding Site**—The full three-dimensional model of the iron-binding site of *E. coli* dismutase, constructed as described under “Materials and Methods,” is represented in Fig. 8A. The new sequence information establishes the identity of the residues which serve as metal ligands; these are His-26, His-73, Asp-156, and His-160 (13, 18). As can be seen by comparing panels A and B of Fig. 8, the three-dimensional similarities of the Fe and Mn proteins extend beyond the ligands to the next shell of residues that constitutes the metal-ligand environment. Almost every residue that penetrates the metal-ligand environment is conserved in the known sequences of Fe and Mn dismutases. The only exceptions found so far are at Fe superoxide dismutase positions 76 (tyrosine or phenylalanine), 69, and 141. In Fe superoxide dismutase residue 69 is glutamine and 141 is alanine, whereas the corresponding residues in Mn superoxide dismutase are glycine and glutamine. These differences represent the exchange of a glutamine from the first domain with one from the second. In
FIG. 7. A, a stereo diagram of the polypeptide fold of E. coli Fe superoxide dismutase. The structure is composed of two domains connected near residue 87. This model has the same fold described earlier but includes 11 more residues which could be identified when the complete sequence was known; the additional residues are at or near the surface of the molecule at positions 41–42, 56–58, 80, 85, 94, 168, 175, and C terminus. B, a superposition of the folds of Mn and Fe superoxide dismutases. The polypeptide chain shown in A was aligned with the structural model of Mn superoxide dismutase from T. thermophilus using the coordinate transformation that relates the electron density maps (17). The chains coincide from the N terminus through the first helix. Just beyond residue 40 (Fe superoxide dismutase numbering) the backbones follow somewhat different courses; they correspond briefly at 52–56 and diverge again until position 61. Mn superoxide dismutase residues in this segment form subunit contacts that are responsible for assembly of the T. thermophilus enzyme into a tetrameric species. After the Mn superoxide dismutase insert which precedes the start of the second long helix the structures are very similar, with no further insertions and only minor deviations in Cα positions. However, from this figure it is apparent that insertions could be accommodated in the exterior sheet connector formed by residues 141–149 which connect the second and third strands of the antiparallel β-sheet structure in T. thermophilus Mn superoxide dismutase.

both Fe and Mn enzymes, this glutamine functions structurally as a bridge between a tyrosine residue at 34 and a tryptophan at 122 (Fig. 8).

To some extent the stereochemical equivalence of the environments of the Fe and Mn ligand clusters is surprising since a number of superoxide dismutases displays in vitro metal binding selectivity. Early reconstitution studies with several bacterial Fe and Mn superoxide dismutases (47, 48), including those from E. coli (49), indicated that the apoenzymes were able to rebind several metals but were active only with the native metal. Although no direct evidence concerning the site of ligation of the inactive metal was provided, incorporation of an “incorrect” metal inhibited binding of the native metal. These results, therefore, seemed to anticipate structural differences in the metal-binding centers of Mn superoxide dismutase and Fe superoxide dismutase. Although the interdomain relocation of an active center glutamine residue is intriguing, it is not clear how either this exchange or the Tyr→Phe substitution near the second ligand could account for the observed selectivities. More recent reports have identified superoxide dismutases, from other organisms, which are reactivated from the apoenzyme by either Mn or Fe. Determinations of the sequences of these less selective dismutases from Propionibacterium shermanii (50), Bacteroides fragilis (51), Bacteroides thetaiotaomicron (52), or Strep- tococcus mutans (53) may pinpoint the chemical origins of the observed metal-binding behavior.

In view of the near identity of the metal-binding site in the known structures, it is noteworthy that purified Fe-protein from E. coli contains virtually no Mn while the purified Mn-protein from E. coli contains virtually no iron (35, 54, 55). Moreover, the levels of the two proteins are responsive to the amounts of the respective metal ion provided in the culture medium (8, 56, 57). Knowledge of the molecular basis of this in vivo metal-binding selectivity may be important in understanding the biological behavior of these proteins.
E. coli Fe Superoxide Dismutase: Gene Sequence and X-ray Structure

FIG. 8. A, the active center of Fe superoxide dismutase from E. coli. Residue identifications are based on the new sequences reported in this paper. The metal ligands (open bonds) and several other residues are truncated at Cβ for clarity. Residues 159B and 163B (thin bonds) from the adjoining subunit penetrate into the metal-ligand environment. The extensive system of imidazole and aromatic rings which partially encloses the metal-ligand cluster is maintained by hydrogen-bonding interactions and by a herringbone network of aromatic packing interactions (69). An additional interaction which stabilizes this system of rings is a bridge in which the side chain amide moiety of Gln6 links Tyr70 with Trp73. The electron density is consistent with ligation of a water molecule (not shown) to Fe at the same site where solvent is observed in the Mn superoxide dismutase structure. B, the active center of Mn superoxide dismutase from T. thermophilus. The residue assignments are in agreement with those of the known sequences (7-10). Except for residues 86 and 151, the amino acids are identical with those found in the Fe superoxide dismutase structure, represented in Fig. 8A, and the two active sites can almost be superimposed. The bridge formed by residue 151 is functionally like that made by residue 69 in the Fe superoxide dismutase structure, although the Cα atoms are not structurally equivalent. Coordinates used to create this figure are based on the crystallographic model which has been described previously (5, 17, 19) but which now has been partially refined at 1.8-A resolution using data collected at the Multiwire Detector Facility, University of California, San Diego. Difference density maps calculated at this resolution indicated a missing residue following position 124. The present numbering reflects the insertion of this residue.

Acknowledgments—We thank Dr. M. Steinmetz for his valuable advice with various aspects of DNA sequence analysis, Dr. P. Stragier and J. Bouvier for their help in the determination of the transcriptional start, C. Mugnier for assistance in the use of CITI 2 facilities, and Dr. G. Tart of The University of Michigan Protein Sequencing Facility for executing the manual Edman degradation of peptides purified in this study.

Addendum—The amino acid sequence of Mn superoxide dismutase from T. thermophilus HB8 has recently been reported (60) and confirms all of the active center residue assignments indicated in Fig. 8B; the sequence indicates two carboxyl-terminal residues beyond those included in Fig. 7B but which could be located in difference density maps with the aid of the published sequence.

REFERENCES
1. Rotilio, G. (ed) (1986) Superoxide and Superoxide Dismutases in Chemistry, Biology and Medicine, Elsevier/North-Holland Biomedical Press, Amsterdam
2. Fridovich, I. (1986) Adv. Enzymol. Relat. Areas Mol. Biol. 58, 61-97
3. Fridovich, I. (1979) in Advances in Inorganic Biochemistry (Eichorn, G. L., and Marzilli, D. L., eds) pp. 67-90, Elsevier/North-Holland Biomedical Press, Amsterdam
4. Michelson, A. M., McCord, M. J., and Fridovich, I. (eds) (1977) Superoxide and Superoxide Dismutases, Academic Press, Orlando, FL
5. Ludwig, M. L., Pattridge, K. A., and Stallings, W. C. (1986) in Manganese in Metabolism and Enzyme Function (Schramm, V. L., and Wedler, F. C., eds) pp. 405-430, Academic Press, Orlando, FL
6. Takeda, Y., and Avila, H. (1986) Nucleic Acids Res. 14, 4577-4589
7. Steinman, H. M. (1978) J. Biol. Chem. 253, 8708-8720
8. Brock, C. J., and Walker, J. E. (1980) Biochemistry 19, 2873-2882
9. Ditlow, C., Johansen, J. T., Martin, B. M., and Svendsen, I. B. (1982) Carlsberg Res. Commun. 47, 81-91
10. Barra, D., Schinina, M. E., Simmaco, M., Banister, J. V., Banister, W. H., Rotilio, G., and Bossa, F. (1984) J. Biol. Chem. 259, 12595-12601
E. coli Fe Superoxide Dismutase: Gene Sequence and X-ray Structure

11. Walker, J. E., Auffret, A. D., Brock, C. J., and Steinman, H. M. (1980) in Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase (Bannister, J. V., and Hill, H. A. O., eds) pp. 212-222, Elsevier/North-Holland Biomedical Press, Amsterdam

12. Harris, J. I., Auffret, A. D., Northrop, F. D., and Walker, J. E. (1980) Eur. J. Biochem. 106, 297-303

13. Barra, D., Schinina, M. E., Bannister, W. H., Bannister, J. V., and Bossa, F. (1987) J. Biol. Chem. 262, 1001-1009

14. 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

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

16. Stallings, W. C., Pattridge, K. A., Strong, R. K., and Ludwig, M. L. (1985) J. Biol. Chem. 260, 16424-16432

17. Stallings, W. C., Pattridge, K. A., Strong, R. K., and Ludwig, M. L. (1984) J. Biol. Chem. 259, 10685-10699

18. Barra, D., Schinina, M. E., Bossa, F., and Bannister, J. V. (1985) FEBS Lett. 179, 329-331

19. Stallings, W. C., Pattridge, K. A., and Ludwig, M. L. (1986) in Superoxide and Superoxide Dismutase in Chemistry, Biology and Medicine (Rotilio, G., ed) pp. 195-204, Elsevier/North-Holland Biomedical Press, Amsterdam

20. Sakamoto, H., and Touati, D. (1984) J. Bacteriol. 159, 418-420

21. Nettleton, C. J., Bull, C., Baldwin, T. O., and Fee, J. A. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 4970-4973

22. Dente, L., Cesareni, G., and Cortese, R. (1985) Nucleic Acids Res. 1, 1645-1655

23. Carilo, A., and Touati, D. (1986) EMBO J. 5, 623-630

24. Dente, L., Solazzo, M., Baldari, C., Cesareni, G., and Cortese, R. (1986) in DNA Cloning, A Practical Approach (Glover, D. M., ed) Vol. 3, pp. 101-107, IRL Press, Oxford

25. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103-119

26. Birnboim, H. C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523

27. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467

28. Burd, J. F., and Wells, R. D. (1970) J. Mol. Biol. 53, 435-459

29. Débarbouillé, M., and Raibaud, O. (1983) J. Bacteriol. 153, 1221-1227

30. Maxam, A., and Gilbert, W. (1980) Methods Enzymol. 65, 499-560

31. Studen, R. (1982) Nucleic Acids Res. 10, 2951-2961

32. Dayhoff, M. O., Hunt, L. T., McLaughlin, P. J., and Barker, W. C. (1972) in Atlas of Protein Sequence and Structure (Dayhoff, M. O., ed) Vol. 5, National Biochemical Research Foundation, Washington, D. C.

33. McLachlan, A. D. (1971) J. Mol. Biol. 61, 409-424

34. Slykhouse, T. O., and Fee, J. A. (1976) J. Biol. Chem. 251, 5472-5477

35. Steinman, H. M. (1982) J. Biol. Chem. 257, 10283-10293

36. Tarr, G. E. (1986) in Methods of Protein Microcrystallization (Shively, J. E., ed) pp. 155-194, Humana Press, Clifton, NJ

37. Hendrickson, W. A., and Lattman, E. E. (1970) Acta Crystallogr. Sect. Struct. Crystallogr. Cryst. Chem. B 26, 136-143

38. Jones, T. A. (1978) J. Appl. Crystallogr. 11, 268-272

39. Touati, D. (1983) J. Bacteriol. 155, 1078-1085

40. Shin, J., and Dalcano, L. (1974) Proc. Natl. Acad. Sci. U. S. A. 71, 1342-1346

41. Pratt, T., and Bear, D. G. (1983) in Gene Function in Prokaryotes (Beckwith, J., Davies, J., and Gallant, J., ed) pp. 123-161, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

42. Rosenberg, M., and Court, D. (1979) Annu. Rev. Genet. 13, 319-353

43. Hawley, D. K., and McClure, W. R. (1983) Nucleic Acids Res. 11, 2237-2255

44. Mulligan, M. E., Hawley, D. C., Entriiken, R., and McClure, W. R. (1984) Nucleic Acids Res. 12, 789-800

45. Yost, F. J., Jr., and Fridovich, I. (1973) J. Biol. Chem. 248, 4905-4908

46. Steinman, H. M., and Hill, R. L. (1973) Proc. Natl. Acad. Sci. U. S. A. 70, 3725-3729

47. Brock, C. J., and Harris, J. I. (1977) Biochem. Soc. Trans. 5, 1537-1539

48. Yamakura, F. (1978) J. Biol. Chem. (Tokyo) 83, 849-857

49. Ose, D. E., and Fridovich, I. (1979) Arch. Biochem. Biophys. 194, 350-364

50. Meier, B., Barra, D., Bossa, F., Calabrese, L., and Rotilio, G. (1982) J. Biol. Chem. 257, 13977-13980

51. Gregory, E. M. (1985) Arch. Biochem. Biophys. 238, 83-89

52. Pennington, C. D., and Gregorio, B. P. (1986) J. Bacteriol. 166, 528-532

53. Martin, M. E., Byers, B. R., Olson, M. O., Salin, M. L., Arceneaux, J. E. L., and Tolbert, C. (1986) J. Biol. Chem. 261, 9361-9367

54. Fee, J. A., Shapiro, E. G., and Moss, T. H. (1976) J. Biol. Chem. 251, 6157-6159

55. Beyer, W. F., Jr., and Fridovich, I. (1987) Biochemistry 26, 1251-1257

56. Simons, R. S., Jocket, P. S., Carroll, M. E. W., and Lowrie, D. B. (1976) Toxicol. Appl. Pharmacol. 37, 271-280

57. Gregory, E. M., and Fridovich, I. (1973) J. Bacteriol. 114, 1193-1197

58. Halley, R. A., Masiarz, F. R., Najarian, R. C., Puma, J. P., Quiroga, M. R., Randolf, A., Sanchez Pescador, R., Scandella, C. J., Smith, B., Steiner, K. S., and Mullenbach, G. T. (1985) Nucleic Acids Res. 13, 2017-2034

59. Burley, S. K., and Petsko, G. A. (1985) Science 229, 23-28

60. Sató, S., Nakada, Y., and Nakazawa-Tomizawa, K. (1987) Biochim. Biophys. Acta 912, 178-184