JOURNAL OF BACTERIOLOGY, May 1996, p. 2564–2571 Vol. 178, No. 9
0021-9193/96/$04.00+0

Cloning and Analysis of *sodC*, Encoding the Copper-Zinc Superoxide Dismutase of *Escherichia coli*

KARIN R. C. IMLAY AND JAMES A. IMLAY*

Department of Microbiology, University of Illinois, Urbana, Illinois 61801

Received 13 November 1995/Accepted 26 February 1996

Benov and Fridovich recently reported the existence of a copper- and zinc-containing superoxide dismutase (CuZnSOD) in *Escherichia coli* (L. T. Benov and I. Fridovich, J. Biol. Chem. 269:25310–25314, 1994). We have used the N-terminal protein sequence to isolate the gene encoding this enzyme. The gene, denoted *sodC*, is located at 37.1 min on the chromosome, adjacent to *lhr* and *sodB*. A monocistronic transcript of *sodC* accumulates only in stationary phase. The presence of a conventional leader sequence is consistent with physical data indicating that the *E. coli* enzyme, like other bacterial CuZnSODs, is secreted into the periplasm. Because superoxide cannot cross membranes, this localization indicates that the enzyme has evolved to defend periplasmic biomolecules against an extracytoplasmic superoxide source. Neither the source nor the target of the superoxide is known. Although once considered an exclusively eukaryotic enzyme, CuZnSOD has now been found in species that span three subdivisions of the purple bacteria. The bacterial CuZnSODs are more homologous to one another than to the eukaryotic enzymes, but active-site residues and structural motifs are clearly shared by both families of enzymes. The use of copper and an invariant disulfide bond suggest that the ancestral gene of present-day CuZnSODs evolved in an aerobic environment, long after the evolutionary split between the eukaryotes and the eubacteria. If so, a CuZnSOD gene must have been transferred laterally between members of these domains. The eukaryotic SODs most closely resemble that of *Caulobacter crescentus*, a relatively close descendant of the mitochondrial ancestor, suggesting that *sodC* may have entered the eukaryotes during the establishment of mitochondria.

The immediate product of the oxidation of reduced electron carriers by molecular oxygen is superoxide (O_2^-), a more potent oxidant which can subsequently react with biomolecules that cannot be damaged directly by oxygen itself. Hence a universal defense against the toxicity of environmental oxygen is the synthesis of superoxide dismutase (SOD) (EC 1.15.1.1), which scavenges O_2^- at nearly diffusion-limited rates (20):

$$O_2^- + O_2^- + 2H^+ \xrightarrow{[SOD]} H_2O_2 + O_2$$

SODs have been found in virtually all of the aerobic organisms examined thus far as well as in some obligate anaerobes that can tolerate transient exposure to oxygen. The dismutation is catalyzed by the alternate reduction and oxidation of an active-site transition metal by consecutive molecules of O_2^-. Eubacteria and archaea typically contain cytosolic SODs that employ manganese or iron as the catalytic metal (denoted MnSODs and FeSODs, respectively). MnSOD is also found in the mitochondria of eukaryotes. MnSODs and FeSODs have very similar structures and presumably evolved from a common ancestral gene (24).

The cytosolic SOD isozyme in eukaryotes, however, is structurally distinct from MnSOD and FeSOD and contains copper as the redox-cycling metal, with an adjacent zinc atom playing a structural and electrostatic role (59). Plants also contain an additional, structurally homologous CuZnSOD in their chloroplasts, and mammals secrete an extracellular CuZnSOD (6). Early studies failed to find CuZnSOD in bacteria. However, in 1974 CuZnSOD was isolated from *Photobacterium leiognathi*, a symbiont of ponyfish (46). The *P. leiognathi* enzyme was sufficiently homologous to eukaryotic CuZnSODs that an evolutionary relationship was inferred, and it was initially suggested that the exceptional presence of CuZnSOD in this bacterium occurred through gene transfer from host to symbiont (7). However, the subsequent discovery of CuZnSOD in the free-living bacterium *Caulobacter crescentus* (52) suggested that the CuZnSODs might also exist in other bacteria. In fact, other bona fide CuZnSODs have since been demonstrated to exist in *Pseudomonas diminuta* and *Xanthomonas* (*Pseudomonas*) *maltophilia* (53), *Brucella abortus* (9), several *Haemophilus* species (34, 37), and *Legionella pneumophila* (2). Most recently, Kroll and colleagues used PCR technology to recover the CuZnSOD-encoding DNA sequence from *Actinobacillus pleuropneumoniae*, *Neisseria meningitidis*, and *Pasteurella multocida* (35).

Although the great majority of the bacteria that have been surveyed by either enzymic assays or DNA hybridization have appeared not to contain CuZnSOD, this spotty distribution may be more apparent than real. This possibility was underscored by the recent discovery that *Escherichia coli*, the subject of countless investigations of SOD function over the past 20 years, synthesizes a previously unrecognized CuZnSOD (12). This isozyme may have eluded other workers because of its instability in extracts and its almost complete repression during log-phase growth. The possibility of an evolutionary linkage between these bacterial CuZnSODs and their eukaryotic counterparts is intriguing.

Interestingly, sequence data and direct localization techniques have indicated that all of these CuZnSODs are secreted to the periplasms of these gram-negative bacteria. The existence of periplasmic SOD in *E. coli*, whose physiology is well understood and experimentally accessible, provides an excellent opportunity to learn its function. We wish to identify the source of the periplasmic O_2^- and the vulnerable biomolecules that CuZnSOD is designed to protect. This investigation will

* Corresponding author. Mailing address: Department of Microbiology, University of Illinois, 131 Burrill Hall, 407 S. Goodwin Ave., Urbana, IL 61801. Phone: (217) 333-5812. Fax: (217) 244-6697. Electronic mail address: Jim_Imlay@qms1.life.uiuc.edu.

require the analysis of enzyme-deficient mutants, and the isolation of the gene encoding CuZnSOD is a first step in that process.

MATERIALS AND METHODS

Strains and plasmids. The K-12 strains AB1157 (F^- *thr-1 leuB6 proA2 his-4 thi-1 argE2 lacY1 galK2 rpsL supE44 ara-14 xyl-15 mtl-1 tsx-33*) (4), JI132 (as AB1157 plus [*sodA*::Mu d PR13]*25* [*sodB-kan*]1-Δ2) (28), and MG1655 (F^- wild type) (5) were used in studies of *sodC* expression. DH5α [*supE44* Δ*lacU169 hsdR17 recA1 endA1 gyrA96 thi-1 relA1* (ϕ80 *lacZDM15*)] (13) was the host strain for plasmid constructions. Plasmid vectors used for cloning were pUC1813 (30) and pBR329 (16). The Kohara library of *E. coli* DNA was originally derived from strain W3110 [F^- IN(*rrnD-rrnE*)*1*] (33).

Media. CuZnSOD synthesis and *sodC* expression was studied using "concentrated LB," which contained 33 g of Bacto tryptone, 20 g of yeast extract, and 7.5 g of NaCl per liter of deionized water, with the pH adjusted to 7.4 with K_2HPO_4 (3). Lambda phage was propagated in 2× YT medium and LB medium (49) supplemented with 0.2% glucose. The plasmid psodC2.3 required 300 μg of ampicillin per ml to be stably maintained.

DNA manipulations. Standard techniques were used for the preparation of lambda phage DNA, plasmid preparation, subcloning, Southern blots, and hybridizations with oligonucleotide probes. mRNA was purified over a cesium chloride cushion and analyzed by the standard Northern (RNA) blotting and primer extension procedures (4). Primer extension reactions and the accompanying DNA sequence reactions were performed with a 22-base ^{32}P end-labelled oligonucleotide complementary to positions +28 to +7 of the DNA sequence, relative to the translational start site. Oligonucleotide synthesis and DNA sequencing were performed by the University of Illinois Urbana-Champaign Biotechnology Center.

Assays. In agreement with the observation of Benov and Fridovich (12), we recovered CuZnSOD in low yields when cells were lysed with a French press. Therefore, CuZnSOD was routinely extracted for assay by osmotic shock. Approximately 10^{10} cells were centrifuged, washed with 50 mM KP_i, pH 7.4, and resuspended in 5 ml of plasmolysis buffer (50 mM Tris, 2.5 mM EDTA, 20% [wt/vol] sucrose, pH 7.4) (40) at room temperature. After sitting 10 min, cells were pelleted and resuspended in 2 ml of ice-cold deionized water and held on ice for 15 min. Spheroplasts were then pelleted by centrifugation, and the supernatant was dialyzed >5 h against 2 liters of 50 mM KP_i, pH 7.4, at 4°C. SOD activity was determined by the xanthine oxidase-cytochrome *c* method (43). Cyanide (2 mM, preadjusted to pH 7.4) was included where indicated as an inhibitor of the Cu-Zn class of SODs. Total protein in the shockate fluid was measured with the Coomassie protein assay reagent from Pierce, with ovalbumin as a standard.

Time courses of *sodC* transcription and CuZnSOD activity. SOD synthesis and *sodC* transcript accumulation were measured during growth of JI132 in concentrated LB medium. An overnight culture of JI132 was subcultured and grown for more than four generations to approximately 5×10^7 CFU/ml. The log-phase culture was then diluted into concentrated LB to an optical density at 600 nm of ca. 0.01. At intervals, aliquots were removed and cells were prepared for RNA extraction or enzyme assay after osmotic shock. mRNA was prepared and blotted according to standard methods (4), with a 38-base oligonucleotide complementary to the N-terminal region of the translated *sodC* sequence used as a probe.

Sequence analysis. Predictions of the secondary structure were generated by using both the NNpredict and ProteinPredict software, available on-line at http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html and http://www.embl-heidelberg.de/predictprotein/predictprotein.html, respectively. The NNpredict program analyzes sequences individually (32), whereas the ProteinPredict software exploits protein alignments to improve the reliability of the method (47). The 16S rRNA phylogenetic subtree was extracted from the ribosomal database (http://www.rdb.life.uiuc.edu). DNA Strider 1.2 was used to scan the *sodC* sequence for DNA and RNA secondary structure. Phylogenetic analyses of bacterial CuZn SODs were performed with the assistance of Gary Olsen using PAUP (phylogenetic analysis using parsimony) (58) and PROTML (1) (maximum likelihood) programs. A cost matrix for amino acid substitution was based on Blosum 45 (25).

Nucleotide sequence accession number. The nucleotide sequence reported here has been submitted to the GenBank/EMBL data bank and assigned accession number U51242.

RESULTS

Isolation of the CuZnSOD structural gene. The CuZnSOD of *E. coli* was purified from stationary-phase cultures of AB1157 to apparent homogeneity by Benov et al., and the amino-acid sequence of the N terminus was determined to be ASEKVEMNLVTSQGVGQ (10). This sequence is not encoded by DNA present in the GenBank database and was generously provided to us to assist with the isolation of the structural gene. A four-fold degenerate oligonucleotide probe [GAAAAAGTIGA(A/G)ATGAA(C/T)CTIGTIAC, where "I" denotes inosine] was synthesized and hybridized to a filter containing an ordered set of DNA blots from the Kohara lambda library of *E. coli* DNA (Pan Vera Corp., Madison, Wis.). Phage DNA was isolated from several clones that gave positive hybridization signals at 42°C. The DNA was digested with *Eco*RI and analyzed by Southern blotting using the same oligonucleotide as a probe. Bands from two of the clones, clones 125 and 317, maintained positive signals. The *Eco*RI digests were repeated, the hybridizing bands were gel purified and digested with *Sau*3A, and the resulting fragments were cloned into *Bam*HI-restricted pUC1813 and transformed into strain DH5α. Colony-blot hybridization using the oligonucleotide probe identified corresponding clones. Plasmids containing hybridizing DNA from each of the original clones 125 and 317 were sequenced. The subclone from 317 had exact homology to the N-terminal 15 amino acids of the *E. coli* CuZnSOD but did not contain the entire gene. The insert derived from clone 125 corresponded to the *yafABCD* region and is not homologous to the CuZnSOD sequence.

Lambda phage DNA from clone 317 was then digested with several restriction enzymes, fractionated on an agarose gel, and Southern blotted with the oligonucleotide probe to determine an appropriate fragment for subcloning the whole gene. A 2.3-kb *Pvu*II fragment was subcloned into the *Pvu*II site of pBR329. The resultant plasmid, psodC2.3, was transformed into JI132, which lacks both of the cytosolic SOD isozymes. Osmotic shockates of the psodC2.3 transformant contained 23 U of SOD activity per mg of protein, while that of its JI132 parent contained 0.3 U/mg. All of the detectable SOD activity in the transformant extract was eliminated by 2 mM KCN, which uniquely inhibits the copper-zinc class of SODs. Thus, the psodC2.3 plasmid evidently included the entire structural gene for the *E. coli* CuZnSOD. This gene has been designated *sodC* in accordance with analogous genes from other bacteria. The *sodA* and *sodB* genes of *E. coli* encode its cytosolic manganese and iron SOD isozymes, respectively (48, 63).

Southern blot analysis of DNA from clone 317 indicated that *sodC* is positioned at 37.1 min on the *E. coli* chromosome. This location is deduced from (i) the known structure of this region (33), (ii) the failure of a 38-base oligonucleotide exactly complementary to *sodC* to hybridize to either of the adjacent lambda clones in the Kohara set (clones 316 and 318), and (iii) strong hybridization to a 7.0-kb *Bgl*II fragment, a 4.4-kb *Pst*I fragment, and a 2.3-kb *Pvu*II fragment. The known gene order of this region is *pdhX-sodC-rnt-lhr-sodB*. The orientation of *sodC* is not revealed by these data. Interestingly, *sodB* is a close neighbor, although these are not cotranscribed and appear to be differentially regulated (see below). The 38-base oligonucleotide did not hybridize to any other member of the Kohara library, indicating that only a single *sodC* gene is likely to exist in *E. coli.*

Features of the *sodC* gene. Both strands of the insert in psodC2.3 were sequenced (Fig. 1). The sequence encoding the amino terminus of the mature protein is preceded by a typical leader region, containing a charged N terminus (Lys-Arg in positions 2 and 3), a 16-amino-acid hydrophobic stretch, and an Ala-Ala cleavage site. This result is consistent with osmotic-shock and immunostain data (11, 12) which indicated that CuZnSOD is a periplasmic protein in *E. coli* as it is in other bacteria. The deduced amino acid sequence of the open reading frame, after removal of the leader peptide, should yield a mature protein of 15,736.6 Da. This value agrees with the molecular mass, 15,737 Da, determined by mass spectroscopy of the purified protein (9).

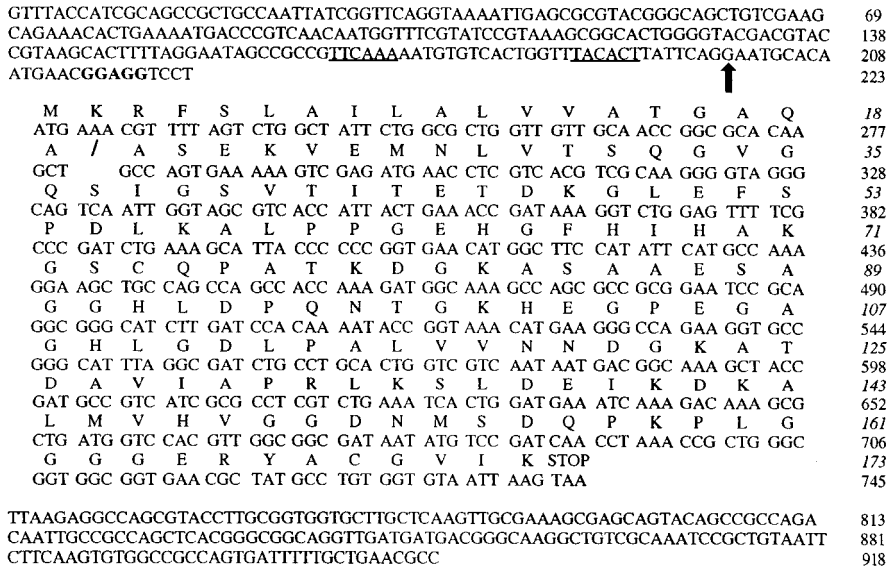

FIG. 1. Nucleotide sequence of the *sodC* region. Nucleotide numbering is arbitrary from the beginning of the region shown; amino acids are numbered from the N-terminal methionine of the *sodC* open reading frame. The cleavage site at the terminus of the leader peptide is indicated by the slash mark. The first base of the transcript is indicated by a vertical arrow. Potential −10 and −35 promoter elements are underlined, and a ribosome binding site is indicated by boldface lettering.

The ribosome binding site GGAGG lies 4 nucleotides upstream from the initiating methionine residue. Cell extracts and mRNA were prepared from cultures of JI132 at intervals as it approached and entered stationary phase, and CuZnSOD activity and *sodC* transcript accumulation were monitored. The level of CuZnSOD activity was extremely low during exponential growth and increased at least 100-fold during stationary phase (Fig. 2A). Northern blots showed that the *sodC* gene is expressed as a monocistronic 600- to 800-base transcript. This transcript was observed only in late-stationary-phase cultures, parallelling the increase in CuZnSOD activity and indicating that the growth-stage regulation is effected by transcriptional controls (Fig. 2B). Among the bacterial genes, the *sodC* of *E. coli* thus far appears to be unique in its virtually complete repression during exponential growth. Primer extension analysis determined that transcripts are initiated at a single position 24 bases upstream of the open reading frame (Fig. 2C). A good −10 promoter sequence is evident upstream of this transcriptional start site (Fig. 1). Although a moderate match to the consensus −35 site is found further upstream, the 15-base spacing between it and the −10 site is at the lower limit of the functional range. Some genes that are expressed exclusively in the stationary phase require the σ^{38} RNA polymerase factor (39) and do not have apparent −35 cassettes (61). Ongoing studies will test the involvement of σ^{38} and other global regulators in *sodC* expression.

Analysis of protein structure. Physical studies of the *E. coli* CuZnSOD have indicated that it is more unstable than other bacterial or eukaryotic CuZnSODs (12) and that, unlike those dimeric enzymes, it may be monomeric (8). Those qualities suggested to us that it might be structurally and evolutionarily distinct. Yet, homology searches in GenBank with the *E. coli* *sodC* sequence extracted the known bacterial CuZnSODs and, with a lower probability, many eukaryotic CuZnSODs. The protein sequence of the *E. coli* enzyme has been aligned in Fig. 3 with those of the five other sequenced bacterial CuZnSODs, as well as with representative eukaryotic cytosolic and chloroplast CuZnSODs. The *E. coli* enzyme is clearly homologous to the other bacterial CuZnSODs, in particular exhibiting about 50% identity with the CuZnSODs of *Haemophilus parainfluenzae*, *L. pneumophila*, and *P. leiognathi*.

While the crystal structure of bovine CuZnSOD has been defined to 2-Å resolution, none of the structures of the bacterial enzymes have yet been solved. However, close inspection of the alignments indicates that the bacterial CuZnSODs almost certainly retain the basic higher-order structure of the eukaryotic enzymes, despite having only about 30% identity. First, the residues of the bovine enzyme that are known to directly affect catalysis are invariably present in the bacterial CuZnSODs, and their relative positions in the primary sequence are unchanged. Crystallographic studies of the bovine enzyme identified His-59, His-61, His-88, and His-156 as ligands for the catalytic copper atom and His-88, His-97, His-106, and Asp-109 as ligands for zinc, while Asp-160 may modulate the metal ligand geometry in the bovine enzyme by hydrogen bonding to metal ligands His-59 and His-96 (all amino acid positions refer to the aligned sequence in Fig. 3) (59). Each of these residues is present at the corresponding site of the *E. coli* enzyme. Additionally, an arginine residue that coordinates O_2^- within the active site is present at position 179 of the aligned sequences. These ligands are widely spaced in the primary sequence, and it is economical to suppose that the bacterial and eukaryotic enzymes use similar folding arrangements to gather them into the active site.

In fact, key residues that are known to organize the tertiary folding of bovine CuZnSOD are also present in the bacterial enzymes. The bovine enzyme can be divided conceptually into a β-barrel region, comprising eight antiparallel β-strands, and three prominent loop regions (Fig. 4). An intramolecular disulfide bond between Cys-70 and Cys-182 bridges the first loop to the C-terminal β-strand, and the presence of homologous cysteine residues in the bacterial enzymes implies that this higher-order structure is retained. The β-strands are punctuated at intervals by glycine residues, which assist in protein packing, particularly at polypeptide turns near the ends of the strands. These glycine residues are universally conserved among eukaryotic CuZnSODs (6). They are also present in the bacterial enzymes, including that from *E. coli*, indicating that

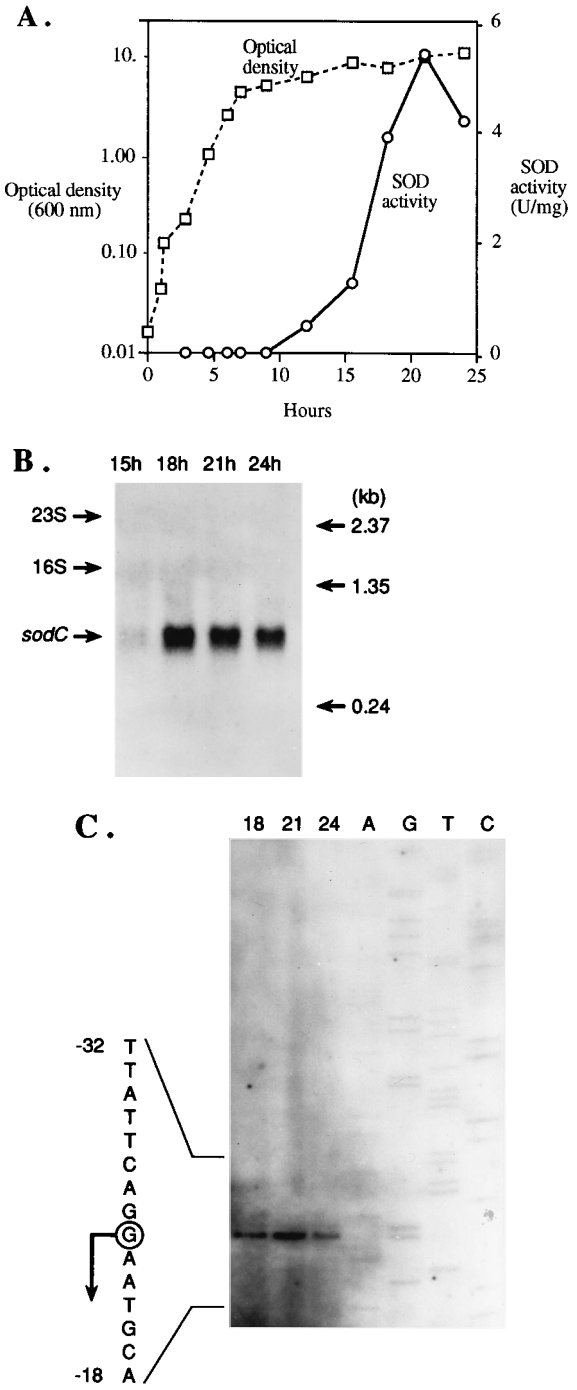

FIG. 2. Induction of CuZnSOD activity and transcript accumulation subsequent to entry into stationary phase. Bacteria were growing exponentially prior to being subcultured at time zero; see Materials and Methods for details. (A) Optical density (broken line) and periplasmic SOD activity (solid line). At early time points the SOD-specific activity of the periplasmic extract was <0.02 U/mg. (B) Northern blots of the *sodC* transcript. Total RNA (30 μg) was loaded onto each lane. Faint bands are due to minor nonspecific hybridization to 16S and 23S ribosomal RNA. No *sodC* transcript was visible on blots of log-phase cells (not shown). (C) Primer extension analysis of the *sodC* transcripts shown in panel B. The arrow marks the position of the transcript terminus.

the bacterial and bovine enzymes have similar patterns of tertiary folding.

Finally, even the many tracts of divergent sequence have been constrained to retain structural motifs. Although the comparison of the six available bacterial sequences with the eukaryotic enzyme indicates a total of 12 sites of amino acid insertions and deletions, all align within the loop and interstrand regions of the bovine enzyme, permitting the β-barrel structure to remain intact (Fig. 4). These bacterial sequences were analyzed for secondary-structure tendencies by both ProteinPredict and NNpredict programs (described in Materials and Methods). The ProteinPredict analysis indicates a likelihood of β-strand structure at precisely those regions that align with the eight β-strands of the bovine CuZnSOD (Fig. 5). A similar result was obtained with the NNpredict program (data not shown). This apparent conservation of structure is particularly striking because in these potential β-regions the degree of amino acid identity between bovine and bacterial enzymes is fairly low, ranging from 27% (*E. coli* CuZnSOD) to 37% (*C. crescentus*). Thus, despite their substantial evolutionary distances, the bacterial *sodC* genes tolerated only those mutations that would maintain the β-barrel core structure and active-site residues that are evident in the eukaryotic enzymes. These similarities strongly argue that the bacterial and eukaryotic enzymes evolved from a common ancestral CuZnSOD.

DISCUSSION

Evolutionary relationships among CuZnSODs. The known bacterial CuZnSODs are all found in purple bacteria. Their presence in members of the alpha, beta, and gamma subdivisions (Fig. 6A) indicates that the evolution of bacterial CuZn SOD at the very least predated the divergence of this division. Since surveys of SOD distribution have largely concentrated on purple bacteria, CuZnSOD could also be present in other divisions. In any case, CuZnSOD has only been found in the periplasms of gram-negative bacteria, and it seems likely that it evolved to defend component biomolecules of that compartment against damage by superoxide. This argument is based in part upon the supposition that the FeSOD-MnSOD family of cytosolic enzymes, which is universally distributed among aerotolerant microbes, is more ancient than CuZnSOD and would have rendered unnecessary the evolution of a second SOD to defend the cytosolic compartment. (Although a few bacteria, including *E. coli*, contain both MnSOD and FeSOD in the cytoplasm, this duplication may serve to ensure that the cell has active SOD even in environments where one or the other metal is lacking, a feat also achieved by "cambialistic" SODs [23, 42].) The inability of superoxide to cross membranes, however, means that SOD in one compartment has no effect upon superoxide concentrations in another. The acquisition of the periplasm as a chemically discontinuous organelle may therefore have been the event which demanded the evolution of CuZnSOD.

The structural and sequence similarity between the bacterial and eukaryotic enzymes is sufficient to conclude that these enzymes have arisen from a common ancestral gene. One possibility is that CuZnSOD evolved prior to the divergence of the protoeukaryotes and the eubacteria from their common ancestor. Although oxygen is thought not to have accumulated to substantial levels in the biosphere until long after that evolutionary split, the fact that FeSOD was apparently dispersed among bacteria and archaea before that date (51) suggests that photochemically generated oxygen may have been sufficiently abundant to provoke intracellular superoxide formation. Some models suggest that oxygen concentrations could have been as high as 0.002 atm (0.2 kPa) (64), which is sufficient to inhibit the growth of some present-day bacteria. The low-potential biochemistries that predominated in ancient microbes may

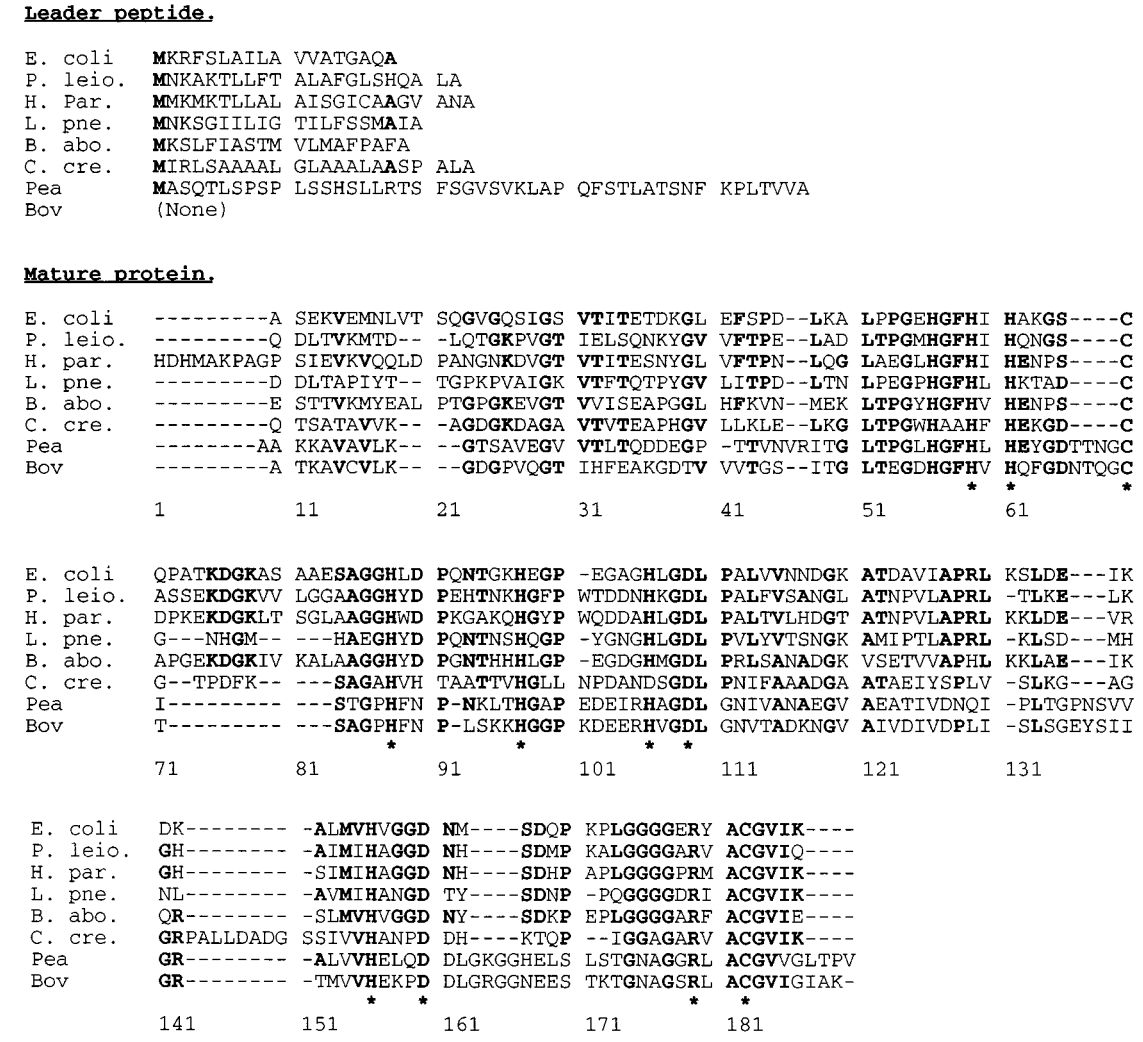

FIG. 3. Aligned amino acid sequences of the six defined bacterial CuZnSODs, a pea chloroplast CuZnSOD, and the bovine cytosolic SOD. Gaps were inserted to optimize the total homology. Amino acids conserved in at least half of the enzymes are shown in boldface letters. Amino acids that are recognized as critical for the function of bovine SOD (see text) are each marked with an asterisk. The *P. leiognathi*, *H. parainfluenzae*, *B. abortus*, *C. crescentus*, pea chloroplast, and bovine CuZnSOD sequences were obtained from the literature (9, 29, 34, 54, 56, 57). The *B. abortus* leader sequence was communicated by John Mayfield (42a), and the sequence of the *L. pneumophila* protein was retrieved from the GenBank database (2).

have particularly disposed electron carriers to form superoxide even when oxygen concentrations were slight.

However, several objections can be raised to the hypothesis that CuZnSOD arose in the common ancestor. (i) As mentioned, it would seem that the presence of FeSOD in the single compartment of that ancient microbe would have made CuZn SOD functionally superfluous. (ii) Copper was probably unavailable to organisms living in anaerobic habitats. In low-potential environments copper is reduced to its less-soluble Cu^{+} ion, and it has been estimated that the insolubility of Cu_2S restricted the copper levels of ancient seas to 10^{-17} M (44). Therefore, copper probably entered the biological repertoire only after the environment became aerobic (45)—a notion supported by the fact that other recognized copper-containing enzymes all appear to have evolved after the advent of oxygenic photosynthesis. In order to have synthesized CuZnSOD, primitive microbes would have required energy- and material-intensive scavenger and transport apparatuses that greatly exceeded in efficiency even present-day iron siderophore systems. This scenario seems implausible, particularly given the availability of dissolved iron and manganese to metallate SODs. (iii) The essential disulfide bond shared by the eukaryotic and bacterial enzymes would have been unstable in the highly reducing environments of their ancestor. This feature must have evolved in an aerobic environment. (iv) To date CuZnSOD has not been found in members of the archaea, which shared with eubacteria and eukaryotes the common ancestor but appear to contain only FeSODs. However, this failure should be regarded with some caution, since both molecular probes and enzymic assays previously failed, for example, to identify the CuZnSOD of *E. coli*.

It seems more likely that during the oxygenization of the atmosphere a gene encoding periplasmic CuZnSOD arose in gram-negative eubacteria and was then transferred laterally into the eukaryotes. Consistent with this idea, the greatest evolutionary distances between eukaryotic cytoplasmic CuZn SODs is smaller than the distances between bacterial CuZn SODs (51). A lateral transfer could have occurred when a primitive purple bacterium invaded ancient eukaryotes, eventually evolving into mitochondria. The advantage of the orig-

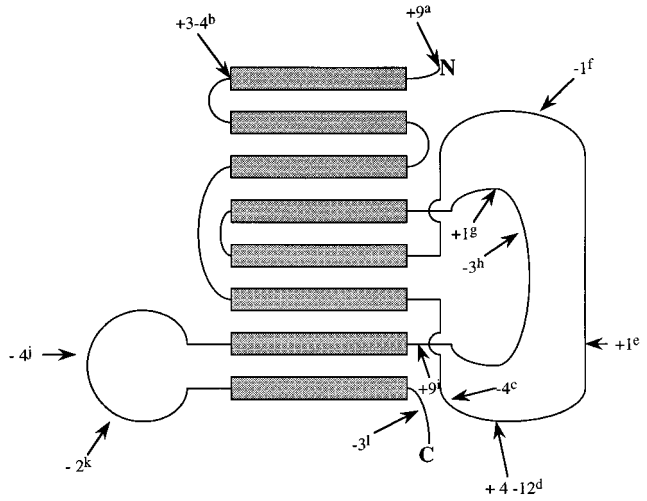

FIG. 4. Schematic diagram of bovine SOD showing the sites of the amino acid insertions (+) and deletions (−) that are indicated by the aligned bacterial sequences. The β-strands (shaded) that define the core topology of the eukaryotic enzyme are left intact. The bovine SOD structure is derived from the crystallographic studies of Tainer et al. (59, 60). Superscript letters indicate insertions or deletions as follows: a, insertion present in *H. parainfluenzae* at residues 1 to 9 of the aligned sequence; b, insertions of various lengths present in all bacterial enzymes at residues 19 to 22; c, deletion in all bacterial enzymes at residues 66 to 69; d, insertions of various lengths in all bacterial enzymes at residues 72 to 83; e, insertion in all bacterial enzymes at residue 92; f, deletion in *E. coli*, *L. pneumophila*, and *B. abortus* enzymes at residue 101; g, insertion in *E. coli*, *H. parainfluenzae*, and *B. abortus* enzymes at position 131; h, deletion in all bacterial enzymes at residues 135 to 137; i, insertion in *C. crescentus* enzyme at residue 143 to 151; j, deletion in all bacterial enzymes at residues 162 to 165; k, deletion in *C. crescentus* and *L. pneumophila* enzymes at residues 171 to 172; l, deletion in all bacterial enzymes at residues 187 to 189.

inal symbiotic relationship is thought to be that bacterial symbiosis permitted a previously obligately anaerobic eukaryote to acquire respiratory capacity and successfully compete for aerobic niches. Its entry into an aerobic habitat, however, may

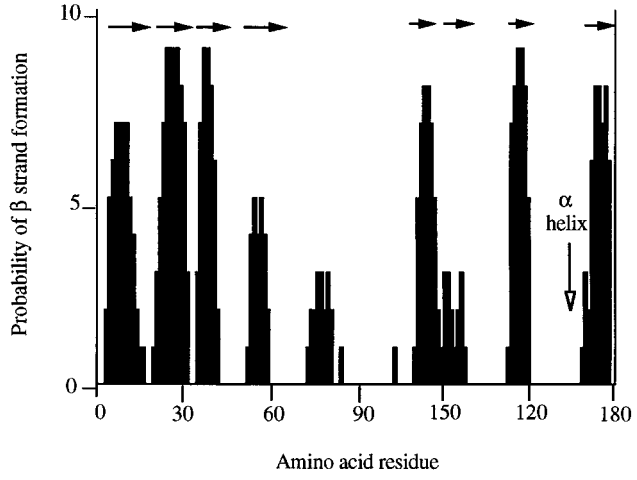

FIG. 5. Prediction of secondary structure from primary sequence of the *E. coli* CuZnSOD protein. Data were analyzed by using the ProteinPredict program (see Materials and Methods) using coaligned sequences from the other bacterial CuZnSODs. Peaks indicate likely regions of β-strand formation. The horizontal arrows demonstrate known β-strands in the bovine CuZnSOD. The vertical arrow near amino acid 135 is the sole site of predicted α-helical structure. Both the indicated α-helix and the potential β-strand at position 75 align with gaps in the eukaryotic enzymes.

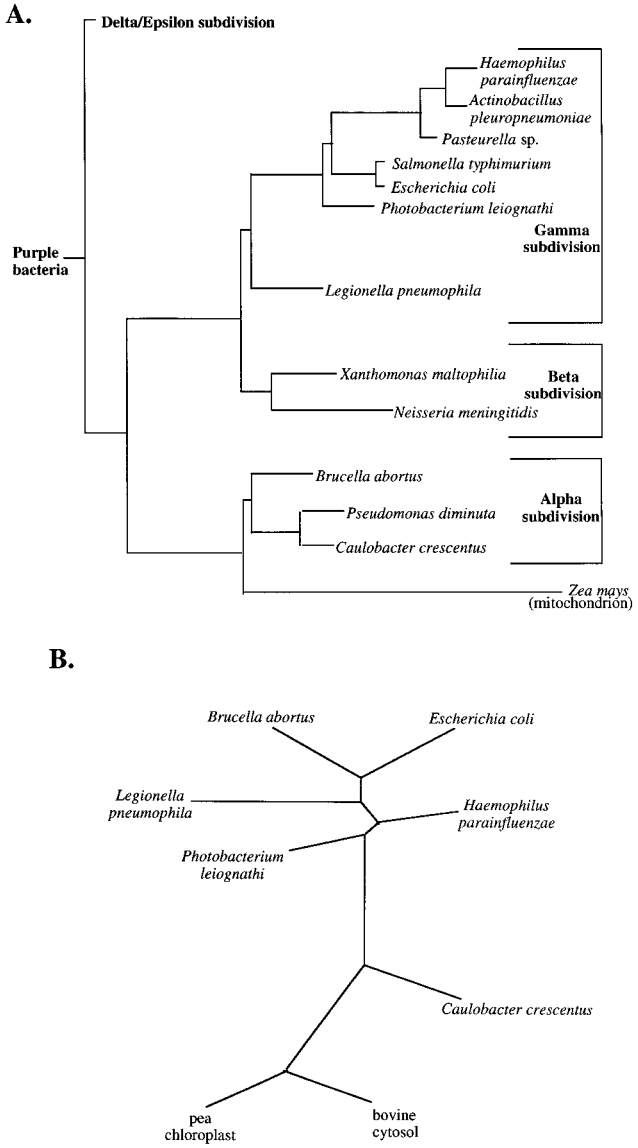

FIG. 6. (A) Suspected phylogenetic relationships of bacteria known to synthesize periplasmic CuZnSODs. The map is based on analyses of 16S ribosomal sequences. (B) Best CuZnSOD tree derived from maximum likelihood analysis of protein sequences. Branch lengths represent calculated phylogenetic distances.

have required some level of cytosolic SOD activity, since contemporary eukaryotic microbes exhibit gross defects without it (14). Superoxide freely passes through outer membrane porins, so it seems plausible that O_2^- formed in the host's cytosol was scavenged by the periplasmic CuZnSOD of the symbiotic bacterium. The presence in both SOD lineages of an invariant disulfide bond argues that the ancestral SOD was periplasmic, since disulfide bonds appear not to accumulate in cytoplasmic bacterial proteins.

In this light, it is of interest that the closest bacterial homolog to the bovine CuZnSOD appears to be the SOD of *Caulobacter* spp. a close relative of the mitochondrial ancestor (Fig. 6B). After subtraction of 21 universally conserved residues, many or all of which are essential for enzyme function, the *Caulobacter* SOD exhibits 33 amino acid identities among mutable residues with the bovine and the chloroplast SODs,

TABLE 1. Number of amino acid identities in the mature (processed) CuZnSODs[a]

CuZnSOD source (no. of residues)[b]	No. of amino acid identities							
	E. coli	*P. leiognathi*	*H. parainfluenzae*	*L. pneumophila*	*B. abortus*	*C. crescentus*	Pea chloroplast	Bovine cytosol
E. coli (154)		75	80	68	85	55	47	44
P. leiognathi (151)	75		90	67	79	57	43	47
H. parainfluenzae (164)	80	90		65	82	55	44	44
L. pneumophila (142)	68	67	65		60	44	40	41
B. abortus (154)	85	79	82	60		52	47	44
C. crescentus (151)	55	57	55	44	52		55	55
Pea chloroplast (154)	47	43	44	40	47	55		82
Bovine cytosol (151)	44	47	44	41	44	55	82	

[a] Number of identities determined on the basis of the alignment shown in Fig. 3.
[b] Number of residues in mature protein.

whereas the other bacterial enzymes average only 23 (Table 1). In fact, the *Caulobacter* SOD is as similar to the eukaryotic enzyme as it is to the *E. coli* enzyme, having 55 amino acid identities with each. Maximum likelihood analysis of the available CuZnSOD sequences derived a best tree on which the eukaryotic SODs were grouped near the *C. crescentus* enzyme (Fig. 6B), while the remaining bacterial CuZnSODs constituted a more-distant cluster. The latter cluster was poorly resolved, since changes in its internal branch order produced trees that were rated as only marginally less likely than the best and the placement of *B. abortus* CuZnSOD within this cluster contradicts its rRNA lineage. However, the *C. crescentus*-eukaryotic group was placed at a substantial distance from the other bacterial enzymes on each of the 15 best trees. This result is consistent with the notion that the eukaryotic enzyme evolved from that of a *Caulobacter* relative. Analysis of a broader distribution of bacterial CuZnSOD sequences will be required to fully test this idea.

Function of periplasmic SOD. Because O_2^- cannot cross membranes (41), the periplasmic location of the bacterial CuZnSODs suggests that they evolved to defend a cell surface target against O_2^- that is generated either extracellularly or within the periplasm itself. At present the only well-defined targets of O_2^- in any organism are a subclass of dehydratases that contain [4Fe-4S] clusters (18, 19, 21, 22, 36). Destruction of these clusters underlies the branched-chain amino acid auxotrophy, the requirement for fermentable carbon sources, and perhaps the hypermutagenesis of *E. coli* mutants that lack cytosolic SODs (15, 17, 31). However, as we are unaware of any cluster-containing enzymes in the periplasm, the need for SOD in that compartment is unclear.

The possible sources of O_2^- in the periplasm are equally mysterious. Although some O_2^- is formed by the respiratory chain within the cytoplasmic membrane, the known sites of respiratory O_2^- production are flavoenzymes, on the inner aspect of the membrane (26, 27), that would generate O_2^- in the cytoplasm rather than the periplasm. Furthermore, respiration is most active and presumably most likely to generate O_2^- during rapid growth rather than during the stationary phase, when the *E. coli sodC* is induced. The fact that periplasmic SOD is found in several invasive bacteria might suggest that it defends the cell against O_2^- that is released by phagocytes. To date, other workers have been unable to unambiguously correlate the infectivity of other pathogens with the presence or absence of the SOD (34, 38; see also reference 62). Dan Ferber in this laboratory recently determined that the stationary-phase cultures of the virulent strain *Salmonella typhimurium* ATCC 14028 contains 60 U of cyanide-inhibitable SOD activity per mg in osmotic shockates, showing that it, too, synthesizes a periplasmic CuZnSOD. The generation of a knockout mutation should clarify the role of CuZnSOD during infection.

However, even if CuZnSOD proved to be a virulence factor, this would represent a new role for an old enzyme, since the ancestral CuZnSOD-containing purple bacterium from which these species emerged had no need for phagocytic defenses. It remains unclear what extracellular O_2^- source might have threatened that bacterium. Steinman and Ely have suggested that present-day *C. crescentus*, a free-living bacterium, may use its periplasmic SOD to scavenge O_2^- that is generated by the photosynthetic apparatus of the algae with which it associates intimately (56). Consistent with that idea, *C. crescentus sodC* mutants did not exhibit any gross physiological defects when grown in pure culture, but they were killed rapidly by extracellular superoxide sources (50, 55). Studies with *E. coli* should shed more light upon the physiological function of this enzyme in bacteria.

ACKNOWLEDGMENTS

We thank Ludmil Benov, Irwin Fridovich, and Wayne Beyer for making available to us the N-terminal sequence of the *E. coli* CuZn SOD prior to publication. We are also grateful to Gary Olsen for assistance with the phylogenetic analysis of *sodC*, John Mayfield and Louisa Tabatabai for communicating the DNA sequence of *sodC* from *B. abortus*, and Howard Steinman and Dan Ferber for their helpful suggestions.

This work was supported by Public Health Service grant GM49640 from the National Institutes of Health.

REFERENCES

1. **Adachi, J., and M. Hasegawa.** 1992. MOLPHY: programs for molecular phylogenetics I—PROTML: maximum likelihood inference of protein phylogeny. Computer Monogr. **27:**1–77.
2. **Amemura, J., and H. Watanabe.** GenBank database NCBI gi:483429.
3. **Asahara, H., P. M. Wistort, J. F. Bank, R. H. Bakerian, and R. P. Cunningham.** 1989. Purification and characterization of *Escherichia coli* endonuclease III from the cloned *nth* gene. Biochemistry **28:**4444–4449.
4. **Ausubel, F. M., R. Brent, R. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1994. Current protocols in molecular biology, vol. 1. John Wiley and Sons, Inc., New York.
5. **Bachmann, B. J.** 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 1190–1219. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
6. **Bannister, J. V., W. H. Bannister, and G. Rotilio.** 1987. Aspects of the structure, function, and applications of superoxide dismutase. Crit. Rev. Biochem. **22:**111–180.
7. **Bannister, J. V., and M. W. Parker.** 1985. The presence of a copper/zinc superoxide dismutase in the bacterium *Photobacterium leiognathi*: a likely case of gene transfer from eukaryotes to prokaryotes. Proc. Natl. Acad. Sci. USA **82:**149–152.
8. **Battistoni, A., and G. Rotilio.** 1995. Isolation of an active and heat-stable

monomeric form of Cu,Zn superoxide dismutase from the periplasmic space of *Escherichia coli*. FEBS Lett. **374:**199–202.
9. **Beck, B. L., L. B. Tabatabai, and J. E. Mayfield.** 1990. A protein isolated from *Brucella abortus* is a Cu-Zn superoxide dismutase. Biochemistry **29:** 372–376.
10. **Benov, B., W. F. Beyer, and I. Fridovich (Duke University).** 1995. Personal communication.
11. **Benov, L., L. Y. Chang, B. Day, and I. Fridovich.** 1995. Copper, zinc superoxide dismutase in *Escherichia coli*: periplasmic localization. Arch. Biochem. Biophys. **319:**508–511.
12. **Benov, L. T., and I. Fridovich.** 1994. *Escherichia coli* expresses a copper- and zinc-containing superoxide dismutase. J. Biol. Chem. **269:**25310–25314.
13. **Bethesda Research Laboratories.** 1986. BRL pUC host: *E. coli* DH5α™ competent cells. Bethesda Res. Lab. Focus **8:**9.
14. **Bilinski, T., Z. Krawiec, J. Litwinska, and M. Blaszczynski.** 1988. Mechanisms of oxygen toxicity as revealed by studies of yeast mutants with changed response to oxidative stress, p. 109–123. *In* P. A. Cerutti et al. (ed.), Oxyradicals in Molecular Biology and Pathology. Alan R. Liss, Inc., New York.
15. **Carlioz, A., and D. Touati.** 1986. Isolation of superoxide dismutase mutants in *Escherichia coli*: is superoxide dismutase necessary for aerobic life? EMBO J. **5:**623–630.
16. **Covarrubias, L., and F. Bolivar.** 1982. Construction and characterization of new cloning vehicles. VI. Plasmid pBR329, a new derivative of pBR328 lacking the 482-base-pair inverted duplication. Gene **17:**79–89.
17. **Farr, S. B., R. D'Ari, and D. Touati.** 1986. Oxygen-dependent mutagenesis in *Escherichia coli* lacking superoxide dismutase. Proc. Natl. Acad. Sci. USA **83:**8268–8272.
18. **Flint, D. H., and M. H. Emptage.** 1990. Dihydroxyacid dehydratase: isolation, characterization as Fe-S proteins, and sensitivity to inactivation by oxygen radicals. *In* D. C. Z. Barak and J. V. Schloss (ed.), Biosynthesis of branched chain amino acids. Deerfield, Borch and Balaban, Philadelphia.
19. **Flint, D. H., J. F. Tuminello, and M. H. Emptage.** 1993. The inactivation of Fe-S cluster containing hydro-lyases by superoxide. J. Biol. Chem. **268:** 22369–22376.
20. **Fridovich, I.** 1989. Superoxide dismutases. An adaptation to a paramagnetic gas. J. Biol. Chem. **264:**7761–7764.
21. **Gardner, P. R., and I. Fridovich.** 1991. Superoxide sensitivity of the *Escherichia coli* 6-phosphogluconate dehydratase. J. Biol. Chem. **266:**1478–1483.
22. **Gardner, P. R., and I. Fridovich.** 1991. Superoxide sensitivity of the *Escherichia coli* aconitase. J. Biol. Chem. **266:**19328–19333.
23. **Gregory, E. M.** 1985. Characterization of the O_2-induced manganese-containing superoxide dismutase from *Bacteroides fragilis*. Arch. Biochem. Biophys. **238:**83–89.
24. **Hassan, H. M.** 1989. Microbial superoxide dismutases. Adv. Genet. **26:**65–97.
25. **Henikoff, S., and J. G. Henikoff.** 1992. Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA **89:**10915–10919.
26. **Imlay, J. A.** 1995. A metabolic enzyme that rapidly produces superoxide, fumarate reductase of *Escherichia coli*. J. Biol. Chem. **270:**19767–19777.
27. **Imlay, J. A., and I. Fridovich.** 1991. Assay of metabolic superoxide production in *Escherichia coli*. J. Biol. Chem. **266:**6957–6965.
28. **Imlay, J. A., and S. Linn.** 1987. Mutagenesis and stress responses induced in *Escherichia coli* by hydrogen peroxide. J. Bacteriol. **169:**2967–2976.
29. **Isin, S. H., J. J. Burke, and R. D. Allen.** 1990. Sequence divergence of pea Cu/Zn superoxide dismutase II cDNAs. Plant Mol. Biol. **15:**789–791.
30. **Kay, R., and J. McPherson.** 1987. Hybrid pUC vectors for addition of new restriction enzyme sites to the ends of DNA fragments. Nucleic Acids Res. **15:**2778.
31. **Keyer, K., A. S. Gort, and J. A. Imlay.** 1995. Superoxide and the production of oxidative DNA damage. J. Bacteriol. **177:**6782–6790.
32. **Kneller, D. G., F. E. Cohen, and R. Langridge.** 1990. Improvements in protein secondary structure prediction by an enhanced neural network. J. Mol. Biol. **214:**171–182.
33. **Kohara, Y.** 1990. Correlation between the physical and genetic maps of the *Escherichia coli* K-12 chromosome, p. 29–42. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
34. **Kroll, J. S., P. R. Langford, and B. M. Loynds.** 1991. Copper-zinc superoxide dismutase of *Haemophilus influenzae* and *H. parainfluenzae*. J. Bacteriol. **173:**7449–7457.
35. **Kroll, J. S., P. R. Langford, K. E. Wilks, and A. D. Keil.** 1995. Bacterial [Cu,Zn]-superoxide dismutase: phylogenetically distinct from the eukaryotic enzyme, and not so rare after all! Microbiology **141:**2271–2279.
36. **Kuo, C.-F., T. Mashino, and I. Fridovich.** 1987. α,β-Dihydroxyisovalerate dehydratase: a superoxide-sensitive enzyme. J. Biol. Chem. **262:**4724–4727.
37. **Langford, P. R., B. M. Loynds, and J. S. Kroll.** 1992. Copper-zinc superoxide dismutase in *Haemophilus* species. J. Gen. Microbiol. **138:**517–522.
38. **Latimer, E., J. Simmers, N. Sriranganathan, R. M. Roop II, G. G. Schurig, and S. M. Boyle.** 1992. *Brucella abortus* deficient in copper/zinc superoxide dismutase is virulent in BALB/c mice. Microb. Pathog. **12:**105–113.
39. **Loewen, P. C., and R. Hengge-Aronis.** 1994. The role of the sigma factor σ^{38} (KatF) in bacterial global regulation. Annu. Rev. Microbiol. **48:**53–80.
40. **Lunn, C., and V. Pigiet.** 1982. Localization of thioredoxin from *Escherichia coli* in an osmotically sensitive compartment. J. Biol. Chem. **257:**11424–11430.
41. **Lynch, R. E., and I. Fridovich.** 1978. Permeation of the erythrocyte stroma by superoxide radical. J. Biol. Chem. **253:**4697–4699.
42. **Martin, M. E., B. R. Byers, M. O. J. Olson, M. L. Salin, J. E. L. Arceneaux, and C. Tolbert.** 1986. A *Streptococcus mutans* superoxide dismutase that is active with either manganese or iron as a cofactor. J. Biol. Chem. **261:**9361–9367.

42a. **Mayfield, J.** Personal communication.

43. **McCord, J. M., and I. Fridovich.** 1969. Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). J. Biol. Chem. **244:**6049–6055.
44. **Ochiai, E.-I.** 1978. The evolution of the environment and its influence on the evolution of life. Origins Life **9:**81–91.
45. **Osterberg, R.** 1974. Origins of metal ions in biology. Nature (London) **249:**382–383.
46. **Puget, K., and A. M. Michelson.** 1974. Isolation of a new copper-containing superoxide dismutase bacteriocuprein. Biochem. Biophys. Res. Commun. **58:**830–838.
47. **Rost, B., and C. Sander.** 1994. Combining evolutionary information and neural networks to predict protein secondary structure. Proteins **19:**55–72.
48. **Sakamoto, H., and D. Touati.** 1984. Cloning of the iron superoxide dismutase gene (*sodB*) in *Escherichia coli* K-12. J. Bacteriol. **159:**418–420.
49. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
50. **Schnell, S., and H. M. Steinman.** 1995. Function and stationary-phase induction of periplasmic copper-zinc superoxide dismutase and catalase/peroxidase in *Caulobacter crescentus*. J. Bacteriol. **177:**5924–5929.
51. **Smith, M. W., and R. F. Doolittle.** 1992. A comparison of evolutionary rates of the two major kinds of superoxide dismutase. J. Mol. Evol. **34:**175–184.
52. **Steinman, H. M.** 1982. Copper-zinc superoxide dismutase from *Caulobacter crescentus* CB15. A novel bacteriocuprein form of the enzyme. J. Biol. Chem. **257:**10283–10293.
53. **Steinman, H. M.** 1985. Bacteriocuprein superoxide dismutases in pseudomonads. J. Bacteriol. **162:**1255–1260.
54. **Steinman, H. M.** 1987. Bacteriocuprein superoxide dismutase of *Photobacterium leiognathi*. J. Biol. Chem. **262:**1882–1887.
55. **Steinman, H. M.** 1993. Function of periplasmic copper-zinc superoxide dismutase in *Caulobacter crescentus*. J. Bacteriol. **175:**1198–1202.
56. **Steinman, H. M., and B. Ely.** 1990. Copper-zinc superoxide dismutase of *Caulobacter crescentus*: cloning, sequencing, and mapping of the gene and periplasmic location of the enzyme. J. Bacteriol. **172:**2901–2910.
57. **Steinman, H. M., V. R. Naik, J. L. Abernethy, and R. L. Hill.** 1974. Bovine erythrocyte superoxide dismutase. Complete amino acid sequence. J. Biol. Chem. **249:**7326–7338.
58. **Swofford, D. L.** 1990. PAUP: phylogenetic analysis using parsimony, version 3.0s. Illinois Natural History Survey, Champaign, Ill.
59. **Tainer, J. A., E. D. Getzoff, K. M. Beem, J. S. Richardson, and D. C. Richardson.** 1982. Determination and analysis of the 2 angstrom structure of copper, zinc superoxide dismutase. J. Mol. Biol. **160:**181–217.
60. **Tainer, J. A., E. D. Getzoff, J. S. Richardson, and D. C. Richardson.** 1983. Structure and mechanism of copper, zinc superoxide dismutase. Nature (London) **306:**284–286.
61. **Tanaka, K., S. Kusano, N. Fujita, A. Ishihama, and H. Takahashi.** 1995. Promoter determinants for *Escherichia coli* RNA polymerase holoenzyme containing σ^{38} (the *rpoS* gene product). Nucleic Acids Res. **23:**827–834.
62. **Tatum, F. K., P. G. Detilleux, J. M. Sacks, and S. M. Halling.** 1992. Construction of Cu-Zn superoxide dismutase deletion mutants of *Brucella abortus*: analysis of survival in vitro in epithelial and phagocytic cells and in vivo in mice. Infect. Immun. **50:**2863–2869.
63. **Touati, D.** 1983. Cloning and mapping of the manganese superoxide dismutase gene (*sodA*) of *Escherichia coli* K-12. J. Bacteriol. **155:**1078–1087.
64. **Walker, J. C. G., C. Klein, M. Schidlowski, J. W. Schopf, D. J. Stevenson, and M. R. Walter.** 1983. Environmental evolution of the archean-early proterozoic Earth. *In* J. W. Schopf (ed.), Earth's earliest biosphere. Princeton University Press, Princeton, N.J.