Volume 229, number 2, 256–260 FEB 05679 March 1988

# Isolation and characterization of complementary DNAs encoding human manganese-containing superoxide dismutase

Ye-Shih Ho and James D. Crapo

*Laboratory of Molecular Biology, Division of Allergy, Critical Care and Respiratory Medicine, Department of Medicine, Duke University Medical Center, Durham, NC 27710, USA*

Received 22 January 1988

cDNAs coding for human manganese-containing superoxide dismutase (Mn SOD) have been isolated from a human liver and a dibutyryl cyclic AMP differentiated U937 cDNA library constructed in vector λgtll. The nucleotide sequences of the insert cDNAs had an opening reading frame coding for 222 amino acid residues. The first 24 amino acids of the primarily translated polypeptide might constitute the leader peptide for transport of the precursors to the mitochondria. Differentiation of the U937 cells with dibutyryl cyclic AMP resulted in a 70% decrease in Mn SOD mRNA. The amino acid sequences of the mature Mn SODs of human, rat and mouse are highly conserved, while the sequences of the leader peptides of these species are moderately conserved.

Superoxide anion radical; cDNA cloning; Nucleotide sequence; Leader peptide; Sequence homology

## 1. INTRODUCTION

Superoxide dismutases (SODs) are thought to be the first line of cellular antioxidant defense mechanisms against the oxidative damage mediated by superoxide anion radicals that are generated intracellularly during normal metabolism [1]. These enzymes act by catalyzing the dismutation of two superoxide radicals to yield hydrogen peroxide and oxygen ($2O_2^- + 2H^+ \longrightarrow H_2O_2 + O_2$). Two distinct types of SODs possessing different molecular masses, amino acid sequences and different metal ions at the active site have been found in mammalian cells. The copper- and zinc-containing enzyme (CuZn SOD) is found principally in the cytosol [2], the manganese-containing enzyme (Mn SOD) being found exclusively in the mitochondrial matrix [3,4]. These two enzymes are encoded by two separate genes on the human chromosomes. Whereas the gene coding for CuZn SOD is located on chromosome 21, that for Mn SOD is on chromosome 6 [5,6]. While the structure and expression of the human CuZn SOD gene have been well documented [7–11], a similar analysis on the gene coding for Mn SOD is still lacking. As a first step toward understanding the gene structure and biosynthesis of Mn SOD, this report describes the isolation and characterization of cDNAs encoding human Mn SOD.

*Correspondence address:* Y.-S. Ho, Box 3177, Duke University Medical Center, Durham, NC 27710, USA

The nucleotide sequences presented here have been submitted to the EMBL/GenBank database under the accession number Y00985

## 2. MATERIALS AND METHODS

The complete amino acid sequence of Mn SOD isolated from human liver has been determined by Barra et al. [12]. In order to clone the cDNA encoding this enzyme, an artificial mini cDNA fragment corresponding to the known amino acid residues 39–64 was made from two synthetic 45-mer oligonucleotides with 15 bases complementary to each other (fig.1). The nucleotide sequences of these two oligonucleotides were derived from the preferred codon usage for humans [13]. The two oligonucleotides were reannealed to each other, followed by a Klenow fill-in reaction in the presence of $^{32}$P-labeled nucleotides and then used to screen an $N^6$-2′-*O*-dibutyryladenosine 3′:5′-cyclic monophosphate (dibutyryl



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/88/$3.50 

```
5'AATGTGACCCAGGAGAAGTACCAGGAGGCCCTGGCCAAGGGCGAT3'
                    3'GACCGGTTCCCGCTACACTGTCGGGTCTAACGGGACGTCGGACGG5'
AsnValThrGlnGluLysTyrGlnGluAlaLeuAlaLysGlyAspValThrAlaGlnIleAlaLeuGlnProAla
```

Fig.1. Sequences of oligonucleotides used to screen the cDNA library.

cyclic AMP) differentiated U937 cell cDNA library constructed in vector λgt11 [14,15]. Hybridization was carried out under moderate stringency [30% formamide, 6 × SSC (1 × SSC = 0.15 M NaCl/0.015 M sodium citrate), 5 × Denhardt's [16], 100 μg/ml single-stranded salmon sperm DNA at 42°C]. Only one positive clone was isolated from approx. 100000 recombinant bacteriophages. This clone was found to contain an approx. 0.8 kb cDNA insert partially encoding the human enzyme. The insert cDNA fragment was used to screen a human liver cDNA library constructed in vector λgt11 (Clontech Laboratories, Palo Alto, CA) under high hybridization stringency (50% formamide, 3 × SSC, 5 × Denhardt's, 100 μg/ml single-stranded salmon sperm DNA at 42°C). Eight positive clones were isolated from approx. 350000 recombinant bacteriophages. The longest cDNA insert was excised from the bacteriophage DNA with restriction endonuclease *Eco*RI and then subcloned into vector M13mp19 for determining the nucleotide sequence. DNA sequencing was performed by the dideoxy chain termination method [17,18]. Synthetic oligonucleotide primers with sequences derived from the 3'-end nucleotide sequences of previous sequencing experiments were used as the sequencing primers to perform further sequencing reactions.

DNA and RNA blot analyses were performed according to Southern [19] and Thomas [20].

## 3. RESULTS AND DISCUSSION

### 3.1. *Nucleotide sequences and deduced amino acid sequences of human Mn SOD cDNAs*

The nucleotide sequences and deduced amino

```
                                                                                    Met Leu Ser Arg
GCGG GCGGCGCAGG AGCGGCACTC GTGGCTGTGG TGGCTTCGGC AGCGGCTTCA GCAGATCGGC GGCATCAGCG GTAGCACCAG CACTAGCAGC ATG TTG AGC CGG    106

-20                                 -10                                 1                                  10
Ala Val Cys Gly Thr Ser Arg Gln Leu Ala Pro Ala Leu Gly Tyr Leu Gly Ser Arg Gln Lys His Ser Leu Pro Asp Leu Pro Tyr Asp
GCA GTG TGC GGC ACC AGC AGG CAG CTG GCT CCG GCT TTG GGG TAT CTG GGC TCC AGG CAG AAG CAC AGC CTC CCC GAC CTG CCC TAC GAC    196

                                    20                                  30                                 40
Tyr Gly Ala Leu Glu Pro His Ile Asn Ala Gln Ile Met Gln Leu His His Ser Lys His His Ala Ala Tyr Val Asn Asn Leu Asn Val
TAC GGC GCC CTG GAA CCT CAC ATC AAC GCG CAG ATC ATG CAG CTG CAC CAC AGC AAG CAC CAC GCG GCC TAC GTG AAC AAC CTG AAC GTC    286

                                    50                                  60                                 70
Thr Glu Glu Lys Tyr Gln Glu Ala Leu Ala Lys Gly Asp Val Thr Ala Gln Thr Ala Leu Gln Pro Ala Leu Lys Phe Asn Gly Gly Gly
ACC GAG GAG AAG TAC CAG GAG GCG TTG GCC AAG GGA GAT GTT ACA GCC CAG ACA GCT CTT CAG CCT GCA CTG AAG TTC AAT GGT GGT GGT    376
                                                                T
                                                               Ile
                                    80                                  90                                 100
His Ile Asn His Ser Ile Phe Trp Thr Asn Leu Ser Pro Asn Gly Gly Gly Glu Pro Lys Gly Glu Leu Leu Glu Ala Ile Lys Arg Asp
CAT ATC AAT CAT AGC ATT TTC TGG ACA AAC CTC AGC CCT AAC GGT GGT GGA GAA CCC AAA GGG GAG TTG CTG GAA GCC ATC AAA CGT GAC    466

                                    110                                 120                                130
Phe Gly Ser Phe Asp Lys Phe Lys Glu Lys Leu Thr Ala Ala Ser Val Gly Val Gln Gly Ser Gly Trp Gly Trp Leu Gly Phe Asn Lys
TTT GGT TCC TTT GAC AAG TTT AAG GAG AAG CTG ACG GCT GCA TCT GTT GGT GTC CAA GGC TCA GGT TGG GGT TGG CTT GGT TTC AAT AAG    556

                                    140                                 150                                160
Glu Arg Gly His Leu Gln Ile Ala Ala Cys Pro Asn Gln Asp Pro Leu Gln Gly Thr Thr Gly Leu Ile Pro Leu Leu Gly Ile Asp Val
GAA CGG GGA CAC TTA CAA ATT GCT GCT TGT CCA AAT CAG GAT CCA CTG CAA GGA ACA ACA GGC CTT ATT CCA CTG CTG GGG ATT GAT GTG    646

                                    170                                 180                                190
Trp Glu His Ala Tyr Tyr Leu Gln Tyr Lys Asn Val Arg Pro Asp Tyr Leu Lys Ala Ile Trp Asn Val Ile Asn Trp Glu Asn Val Thr
TGG GAG CAC GCT TAC TAC CTT CAG TAT AAA AAT GTC AGG CCT GAT TAT CTA AAA GCT ATT TGG AAT GTA ATC AAC TGG GAG AAT GTA ACT    736

Glu Arg Tyr Met Ala Cys Lys Lys .
GAA AGA TAC ATG GCT TGC AAA AAG TAA ACCACGATCG TTATGCTGAG TATGTTAAGC TCTTTATGAC TGTTTTTGTA GTGGTATAGA GTACTGCAGA ATACAG    839

TAAG CTGCTCTATT GTAGCATTTC TTGATGTTGC TTAGTCACTT ATTTCATAAA CAACTTAATG TTCTGAATAA TTTCTTACTA AACATTTTGT TATTGGGCAA GTGA    947

TTGAAA ATAGTAAATG CTTTGTGTGA TTGAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA                                          1021
```

Fig.2. Nucleotide and deduced amino acid sequences of cDNAs encoding human Mn SOD. Variations in the nucleotide sequence and deduced amino acid sequence of a truncated cDNA isolated from the differentiated U937 cell cDNA library are indicated in italics below the sequences of the full-length cDNA isolated from the human liver cDNA library.

acid sequences of the Mn SOD cDNAs isolated from the human liver and dibutyryl cyclic AMP differentiated U937 cell cDNA libraries are shown in fig.2. Comparison of the deduced amino acid sequences with the amino-terminal sequence of purified human liver Mn SOD showed that the mature human Mn SOD starts at amino acid residue 25. The preceding 24 amino acids apparently represent the signal peptide which is involved in the transportation of precursor polypeptides of Mn SOD into the mitochondria. The deduced amino acid sequences of the mature Mn SOD are essentially identical to that determined by Barra et al. [12] with a few exceptions. There are two additional amino acids (Gly-124 and Trp-125) predicted from the nucleotide sequences, and the deduced amino acids at residues 42, 88, 109 and 131 are glutamic acid instead of glutamine as determined by peptide sequencing [12].

The nucleotide sequence of the truncated cDNA insert isolated from the differentiated U937 cell cDNA library is identical to the sequence between residues 254 and 964 of the full-length cDNA insert isolated from the human liver cDNA library except for a C to T substitution at nucleotide residue 339 leading to substitution of threonine by isoleucine at amino acid residue 58. The amino acid at residue 58 was also determined to be isoleucine by peptide sequencing [12]. This sequence discrepancy may result from the polymorphisms of the gene.

### 3.2. *RNA and DNA blot analyses of human Mn SOD gene*

RNA blot analysis of poly(A)$^+$ RNAs from U937 cells and dibutyryl cyclic AMP (1 mM) differentiated U937 cells probed with the full-length cDNA fragment revealed a major hybridizable mRNA species at 1.3 kb and a minor species at approx. 6 kb (fig.3a). The quantities of these mRNAs seemed to decrease to approx. 30% upon differentiation of U937 cells from monocyte- to macrophage-like cells with dibutyryl cyclic AMP [21]. To ensure that the decrease of Mn SOD mRNA during the differentiation of U937 cells did not result from experimental variations in the amounts of mRNA loaded on the gels, the same RNA blot filter was subsequently rehybridized with a rat glyceraldehyde-3-phosphate dehydrogenase cDNA [22]. The amounts of hybridizable mRNAs were the same before and after differentiation of the U937 cells (not shown).

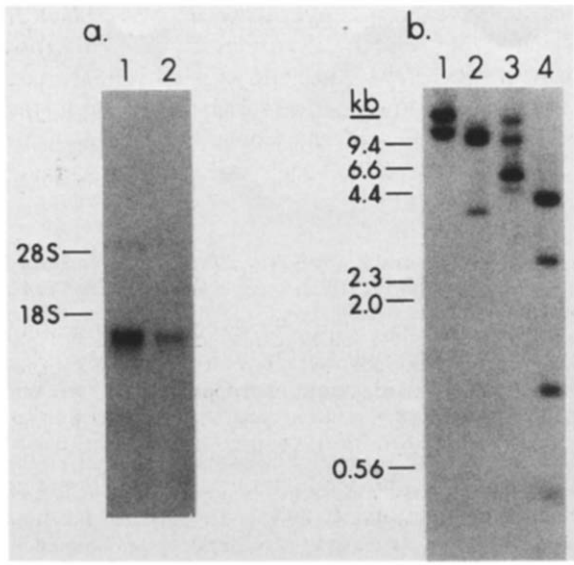


Fig.3. RNA and DNA blot analyses of human Mn SOD gene. (a) RNA blot analysis of human Mn SOD transcripts. Poly(A)$^+$ RNAs (5 $\mu$g) isolated from U937 cells (1) and U937 cells differentiated with 1 mM dibutyryl cyclic AMP for 44 h (2) were denatured, electrophoresed through agarose, transferred to nitrocellulose paper, and then hybridized with the full-length Mn SOD cDNA. The positions of the 28 S and 18 S ribosomal RNAs are shown to the left. (b) DNA blot analysis of human Mn SOD genomic DNA. U937 cell genomic DNA (10 $\mu$g) was digested with restriction endonuclease *Eco*RI (1), *Bam*HI (2), *Hind*III (3) and *Pst*I (4), electrophoresed through agarose, denatured, transferred to nitrocellulose paper, and hybridized with the full-length cDNA probe. The positions of DNA size markers are shown to the left.

In order to determine the complexity of the Mn SOD gene in human genome, DNA blot analysis of U937 cell genomic DNA digested with various restriction endonucleases was performed using the full-length cDNA fragment as a hybridization probe (fig.3b). The same DNA blot analysis of genomic DNAs isolated from three healthy donors was also performed and showed identical patterns of hybridization to those of U937 cells (not shown). The number of gene copies can be assessed from the number and intensity of restriction fragments hybridizing to the cDNA probe. Rehybridization of the DNA blot filter with a human CuZn SOD cDNA, which is derived from a single-copy gene on human chromosome 21, showed that the intensities of human genomic

restriction fragments hybridized with Mn SOD or CuZn SOD cDNA are equivalent [7,9] (not shown). Therefore, since two high molecular mass *Eco*RI genomic fragments hybridized with the probe, we estimated that there are at most two copies of Mn SOD gene per haploid human genome.

### 3.3. *Sequence homologies of Mn SOD among various species*

The amino acid sequences encoded by the long open frame of these cDNAs were aligned with known amino acid sequences of rat and mouse Mn SODs (fig.4) [23,24]. The amino acid sequences of mature Mn SODs are highly conserved among these species. The sequence homologies between the mature human Mn SOD and mature rat and mouse Mn SOD were 93 and 94%, respectively, while that between mature rat and mouse Mn SOD was even higher at 96%. The different extents of sequence homologies among these species were expected by considering that the common ancestral lineage of rat and mouse separated from the human lineage earlier than the rat and mouse species diverged during the evolution of these species [25,26]. Similar relationships of sequence homologies could also be observed from the amino acid sequences of leader peptides of Mn SODs from these species. The sequence homologies between the leader peptide of human Mn SOD and that of rat and mouse Mn SOD were the same at 50%, while the sequence homology between the leader peptides of rat and mouse Mn SOD was 83%. It is interesting to note that the rate of amino acid substitutions is much greater in the evolution of leader peptides than that of mature proteins. The leader peptides of mitochondrial proteins do not contain particular sequence homology, rather they conserve overall structural features – they are rich in positively charged and hydroxylated amino acids and usually devoid of acidic amino acids or extended hydrophobic stretches [27,28]. These features of the leader peptide allow additional 'room' for amino acid substitutions and at the same time conserve the function of this peptide in mitochondrial targeting. Furthermore, this conclusion also suggests that conservation of particular amino acid sequences in mature Mn SOD may be essential to maintain their functions.

*Acknowledgements:* We would like to thank Dr Richard Randell, Ms Mildred McAdams and Ms Kathleen Theisen of Howard Hughes Medical Institute at Duke University for synthesizing oligonucleotides and Ms Dortch Smith for excellent secretarial assistance. This research was supported by grants from National Institutes of Health (HL-39585) and RJR-Nabisco, Inc.

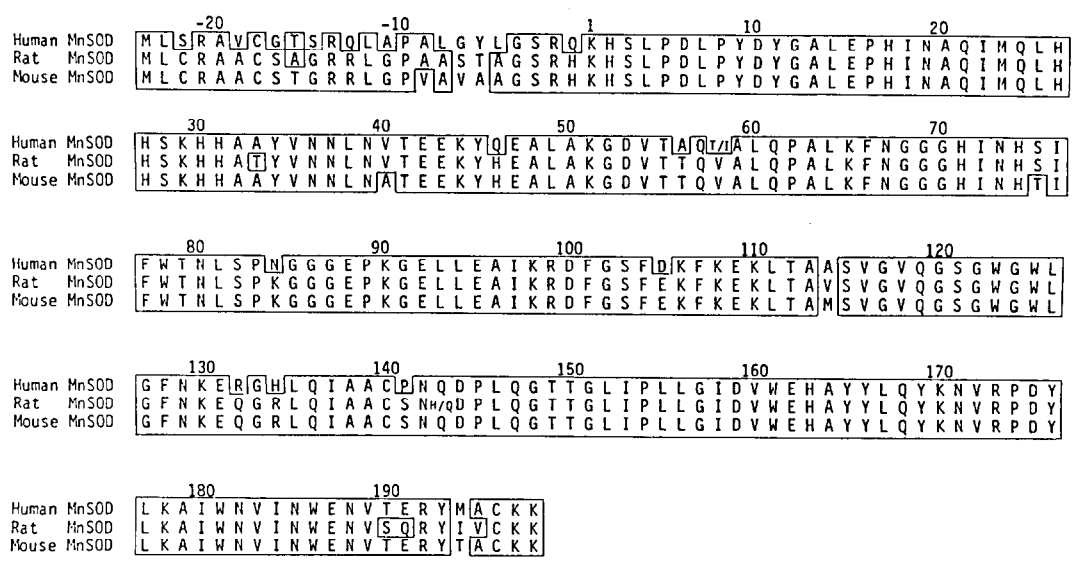


Fig.4. Comparison of the deduced amino acid sequences (designated by the single-letter code) of human, rat and mouse Mn SOD. Solid line boxes indicate positions at which residues are identical.

## REFERENCES


[1] McCord, J.M. (1979) in: Reviews of Biochemical Toxicoloty (Hodgson, E. et al. eds) vol.1, pp.109–124, Elsevier, Amsterdam, New York.

[2] McCord, J.M. and Fridovich, I. (1969) J. Biol. Chem. 244, 6049–6055.

[3] Weisiger, R.A. and Fridovich, I. (1973) J. Biol. Chem. 248, 3582–3592.

[4] Weisiger, R.A. and Fridovich, I. (1973) J. Biol. Chem. 248, 4793–4796.

[5] Tan, Y.H., Tischfield, J. and Ruddle, F.H. (1973) J. Exp. Med. 137, 317–330.

[6] Creagan, R., Tischfield, J., Ricciuti, F. and Ruddle, F.H. (1973) Humangenetik 20, 203–209.

[7] Sherman, L., Dafni, N., Lieman-Hurwitz, J. and Groner, Y. (1983) Proc. Natl. Acad. Sci. USA 80, 5465–5469.

[8] Levanon, D., Lieman-Hurwitz, J., Dafni, N., Wigderson, M., Sherman, L., Bernstein, Y., Laver-Rudich, Z., Danciger, E., Stein, O. and Groner, Y. (1985) EMBO J. 4, 77–84.

[9] Hallewell, R.A., Masiarz, F.R., Najarian, R.C., Puma, J.P., Quiroga, M.R., Randolph, A., Sanchez-Pescador, R., Scandella, C.J., Smith, B., Steimer, K.S. and Mullenbach, G.T. (1985) Nucleic Acids Res. 13, 2017–2034.

[10] Elroy-Stein, O., Bernstein, Y. and Groner, Y. (1986) EMBO J. 5, 615–622.

[11] Hartman, J.R., Geller, T., Yavin, Z., Bartfeld, D., Kanner, D., Aviv, H. and Gorecki, M. (1986) Proc. Natl. Acad. Sci. USA 83, 7142–7146.

[12] Barra, D., Schinina, M.E., Simmaco, M., Bannister, J.V., Bannister, W.H., Rotillio, G. and Bossa, F. (1984) J. Biol. Chem. 259, 12595–12601.

[13] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.

[14] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[15] Huynh, T.V., Young, R.A. and Davis, R.W. (1984) in: DNA Cloning Techniques: A Practical Approach (Glover, D. ed.) pp.49–78, IRL, Oxford.

[16] Denhardt, D. (1966) Biochem. Biophys. Res. Commun. 23, 641–652.

[17] Messing, J. (1983) Methods Enzymol. 101, 20–78.

[18] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.

[19] Southern, E. (1975) J. Mol. Biol. 98, 503–517.

[20] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.

[21] Kay, G.E., Lane, B.C. and Snyderman, R. (1983) Infect. Immun. 41, 1166–1174.

[22] Fort, P., Marty, L., Piechaczyk, M., El Sabrouty, S., Dani, C., Jeanteur, P. and Blanchard, J.M. (1985) Nucleic Acids Res. 13, 1431–1442.

[23] Ho, Y.-S. and Crapo, J.D. (1987) Nucleic Acids Res. 15, 10070.

[24] Hallewell, R.A., Mullenbach, G.T., Stempien, M.M. and Bell, G.I. (1986) Nucleic Acids Res. 14, 9539.

[25] Blomback, B. (1971) in: Molecular Evolution 2: Biochemical Evolution and the Origin of Life (Schoffeniels, E. ed.) pp.112–129, North-Holland, Amsterdam.

[26] Li, W.-H., Tanimura, M. and Sharp, P.M. (1987) J. Mol. Evol. 25, 330–342.

[27] Smeekens, S., De Groot, M., Van Binsbergen, J. and Weisbeek, P. (1985) Nature 317, 456–458.

[28] Von Heijne, G. (1986) EMBO J. 5, 1335–1342.