

Yeast Sequencing Reports

Cloning and Sequencing of a cDNA encoding a Copper-Zinc Superoxide Dismutase Enzyme from the Marine Yeast *Debaryomyces hansenii*

NORMA Y. HERNÁNDEZ-SAAVEDRA^{1*}, JEAN MARC EGLY² AND JOSÉ LUIS OCHOA¹

¹Center for Biological Research of the Northwest, Laboratory of Marine Yeast, A.P. 128, La Paz 23000, B.C.S., México

²Institut de Genetique et de Biologie Moleculaire et Cellulaire, B.P. 163, 67404 Illkirch Cedex, C.U. de Strasbourg, France

Received 4 August 1997; revised 26 October 1997

Cu-Zn superoxide dismutase (SOD-1) is a ubiquitously occurring eukaryotic enzyme with a variety of important effects on respiring organisms. A gene (*dhsod-1*) encoding a Cu-Zn superoxide dismutase of the marine yeast *Debaryomyces hansenii* was cloned using mRNA by the RT-PCR technique. The deduced amino-acid sequence shows ~70% homology with that of cytosolic superoxide dismutase from *Saccharomyces cerevisiae* and *Neurospora crassa*, as well as lower homologies (between 55 and 65%) with the corresponding enzyme of other eukaryotic organisms, including human. The gene sequence encodes a protein of 153 amino acids with a calculated molecular mass of 15.92 kDa, in agreement with the observed characteristics of the purified protein from *D. hansenii*. The *dhsod-1* sequence has been deposited in the public data library of the NCBI under Accession Number AF016383. © 1998 John Wiley & Sons, Ltd.

Yeast 14: 573–581, 1998.

KEY WORDS — marine yeast; superoxide dismutase; *Debaryomyces hansenii*; cloning

INTRODUCTION

Oxygen is essential for life on Earth. In its ground state (i.e., its normal configuration, O₂) oxygen is relatively unreactive. However, during normal metabolism, and because of various environmental perturbations and pollutants (i.e., radiation, drought, air pollutants, cigarette smoke, temperature stress, herbicides, etc.) oxygen reacts to yield various highly toxic and lethal intermediates. Toxic intermediates, referred to as reactive oxygen species (ROS) and free radicals, include the superoxide radical (O₂⁻), hydrogen peroxide (H₂O₂) and the hydroxyl radical (OH). All of these are

biologically important reactive oxygen species, cytotoxic to organisms. The effective and rapid elimination of ROS is essential to the proper functioning and survival of living organisms. For this, organisms have evolved antioxidant defence mechanisms. Such defences include non-enzymatic and enzymatic mechanisms. Enzymatic antioxidant defences include ascorbate peroxidase, glutathione reductase, catalase and peroxidases that efficiently remove hydrogen peroxide from cells, and superoxide dismutase that scavenges the superoxide anion.

Superoxide dismutase is a ubiquitous metallo-enzyme (SOD; EC 1.15.1.1) in aerobes. It catalyses the dismutation of the superoxide radical (O₂⁻)

*Correspondence to: N. Y. Hernández-Saavedra.

to molecular oxygen and hydrogen peroxide (McCord and Fridovich, 1969). Three different types of SODs have been found so far. They can be grouped into two unrelated families with respect to the bound metal ion (Fridovich, 1989). Mn-SOD and Fe-SOD form one family; they are usually found in prokaryotes and subcellular organelles. The other family comprises the Cu-Zn SODs, which occur primarily in the cytosol of eukaryotic cells and in plant chloroplasts (Amano *et al.*, 1990). An extracellular form of Cu-Zn SOD (EC SOD) has been characterized in mammalian body fluids (Marklund, 1984). Both types of Cu-Zn SODs are believed to have evolved from a common ancestor, which is different from that of Mn- and Fe-SODs (Fridovich, 1989; Matsumoto *et al.*, 1991).

As a general rule the Cu-Zn SOD is restricted to eukaryotic organisms; however, the first exception was discovered in the luminescent bacteria *Photobacterium leiognathi*. This organism contains Cu-Zn SOD, and exists in symbiotic relationship with the ponyfish, occupying a special gland that imparts a luminescent characteristic to the fish (Halliwell and Gutteridge, 1991). The first interpretation was that the bacteria obtained the gene through a gene transfer from the host fish. However, with recent sequencing studies it has been demonstrated that the bacterial enzyme is not closely related to the eukaryotic one. The same conclusion arises from the sequence analysis of the Cu-Zn SOD present in the free-living (non-symbiotic) bacterium *Caulobacter crescentus*. A Cu-Zn SOD null mutant of the free-living *C. crescentus* did not display a phenotypic deficit under ordinary growth conditions and was not hypersensitive to paraquat, but sensitive towards citrate; which is a good chelating agent able to extract metal cations from the lipopolysaccharide cell surface (Steiman, 1982). A Cu-Zn SOD null mutant of *Brucella abortus* was less able to survive in mice; thus, the enzyme probably also has a role as a pathogenicity factor (Halliwell and Gutteridge, 1991). Recently it has been discovered to have an important role in protection against oxidative damage in a double mutant strain (*soda sodb*) of *Escherichia coli* (Benov and Fridovich, 1994, 1996). To date, there are only nine sequence reports of Cu-Zn SODs for bacteria in the SwissProt data bank (update 30 May 1997).

From studies of Cu-Zn SOD in several species, the cytosolic isoenzyme is a dimeric protein of two identical, non-covalently linked and non-

glycosylated, subunits (Hong *et al.*, 1992). Each subunit has a size of 16 kDa and binds one atom of copper and one atom of zinc (Marklund, 1984; Tibell *et al.*, 1987). The extracellular form is proposed to be glycosylated through an N-linkage and contain a signal sequence of 18 amino acids (Hjalmarsson *et al.*, 1987).

The gene encoding the SOD-1 has been cloned from several species of higher eukaryotes. However, the SOD genes from moulds and yeast have not been studied extensively. To date, only four nucleotide sequences have been reported in the EMBL and GenBank data libraries. These sequences correspond to: *Saccharomyces cerevisiae*, J03279 (Birmingham-McDonogh *et al.*, 1988); *Schizosaccharomyces pombe* X66722 (unpublished data); *Neurospora crassa* M58687 (Chary *et al.*, 1990); and *Aspergillus japonicus* L32834 (Lin *et al.*, 1994).

Debaryomyces hansenii is a normal marine yeast that has provided the principal model for the study of salt tolerance in marine organisms. *D. hansenii* is strongly halotolerant and has the ability to grow in 0–24% (0–4.13 M) sodium chloride solutions (Adler, 1986; Hernández-Savedra *et al.*, 1995). When grown in high salt conditions, this organism excretes sodium and selects for potassium uptake (Norkrans and Kylin, 1969; Hobot and Jennings, 1981). Polyhydroxy alcohols are also produced and accumulated within the organism to counterbalance an increase in the osmotic potential of the environment (Adler and Gustaffson, 1980; Hernández-Saavedra *et al.*, 1995). Only sporadic molecular biology studies have been done on this species. An autonomously replicating sequence has been isolated (Gorvind and Banaszak, 1992), and the amplification and sequencing of ribosomal RNA (rRNA) fragments used as taxonomic tools has been performed (Kurtzman, 1994). This is the first work to report the cloning of a gene from the marine yeast *D. hansenii*. The aims of this study were to clone the gene *sod-1* of the marine yeast *D. hansenii* strain C-11 (Hernández-Saavedra, 1990) and compare its protein sequence with those previously reported especially from other fungi.

MATERIALS AND METHODS

Microorganisms, plasmid, chemical and enzymes

We used *D. hansenii* strain C-11 obtained from the CIBNOR Marine Yeast Collection (Hernández-Saavedra, 1990), and *S. pombe* strain

972 h⁻ and *S. cerevisiae* strain X2180 obtained from the Yeast Genetic Stock Center. The cloning vector pCRII, the expression vector pTrcHis, and *E. coli* strains INVαF' and TOP 10 were obtained from Invitrogen[®]. Enzymes were purchased from New England BioLabs and Appligene, and analytical grade chemicals from Sigma Chemical Co. (St Louis, MO, USA). Reagents to prepare culture media for yeast and bacteria were obtained from Difco (Difco Laboratories, Detroit, MI, USA).

Protein sequencing

A sample of Cu-Zn SOD protein from the marine yeast *D. hansenii* (N. Y. Hernández-Saavedra and J. L. Ochoa, submitted to *Yeast*), purified by polyacrylamide gel electrophoresis, was digested with trypsin according to Rosenfeld *et al.* (1992), to obtain internal peptide sequences. The analyses were done at the Institut de Genetique et de Biologie Moleculaire et Cellulaire (IGBMC) Peptide Service, and the sequences obtained were used to design the oligonucleotides used for reverse transcription-polymerase chain reaction (RT-PCR) amplification.

Oligodeoxynucleotides

Synthetic oligonucleotides used in these studies were obtained from the IGBMC Oligonucleotide Service as forward sequences: **NT1**, ATGAA (AG)GCIGTITG(TC)GTIATGACIGG and **NT2**, ATGGTGCAGGCCGGTGGC; and as reverse sequences: **CT1**, TC(AG)TC(TC)TC(AG)TT(TC)TC(AG)TG(GTA)AT(I)ACCAT, and **CT2**, TTA GTTGGTGAGGCCGAT. Bases between the parentheses represent randomly mixed bases at the wobble positions.

Cloning of *sod-1* gene by RT-PCR

For total RNA preparation, yeast cells (*S. cerevisiae*, *S. pombe* and *D. hansenii*) were grown until the exponential phase at room temperature in a rotatory shaker at 150 rpm in M1/2 media (glucose 20 g/l, peptone 10 g/l, yeast extract 5 g/l, and seawater with pH adjusted to 4.5 with 0.1 N-HCl). Isolation of total RNA from vegetative yeast cells was done by breaking down the cells with vigorous shaking in the presence of hot phenol (Köhler and Domdey, 1991). Poly (A)⁺ RNA was prepared using a oligo(dT)-cellulose Type 7 column (Pharmacia[®]) according to

Maniatis *et al.* (1989). RNA quantification was done at 260 nm by use of a Beckman[®] DU-7 spectrophotometer.

Single-stranded cDNAs were synthesized using the cDNA Cycle[®] Kit (Invitrogen[®]) by reverse transcription from vegetative mRNA using the PB-dT¹⁷ oligo as primer. The single-strand cDNA was then used as a template for PCR. A Perkin-Elmer thermocycler (Perkin-Elmer Corp., England) was used for PCR amplifications. Reaction mixtures (50 µl) contained 100 pmol of the specific degenerated primers NT1 and CT1, all four deoxynucleoside triphosphates at 10 µM, 5 ml of 10 × PCR buffer (Maniatis *et al.*, 1989), 10 ng of cDNA, and 2 U of *Taq* polymerase. The mixture was overlaid with 50 µl of mineral oil and subjected to 3 min at 94°C; five cycles consisting of a 1.5-min denaturation period at 94°C, a 2-min annealing period at 50°C, and a 2-min extension period at 72°C; and 25 cycles consisting of: 1 min at 94°C, 1 min at 50°C, and 1 min at 72°C. A final extension period of 5 min at 72°C was done to ensure the generation of double-stranded products. PCR products (10 µl samples) were separated on a 2% (w/v) agarose TAE 1 × gel and purified from the gel by using Gene Clean (Bio 101, La Jolla, CA, USA) when appropriate. Gels were stained with ethidium bromide and photographed under UV light. PCR products were cloned using the PCR II TA[®] Cloning Kit (Invitrogen) in the pCRII vector according to the instructions provided by the manufacturer. INVαF' cells were transformed by the heat-shock method (Maniatis *et al.*, 1989). Transformant selection was done on LB plates containing ampicillin (50 µg/ml) and X-gal (5-bromo-4-chloro-3-indolyl-β-galactopyranoside). Plasmid DNA was extracted by the alkaline method (Birnboim and Doly, 1979), whereas that used for sequencing was purified by means of Nucleobond AX100 cartridges (Macherey-Nagel[®]) according to the supplier's instructions. Total RNA from *S. cerevisiae* X2180, *S. pombe* 972 h⁻, and the *sod-1* gene (1.1 kb) from *S. cerevisiae*, provided by Dr E. Gralla (University of California, Los Angeles), were used as amplification control probes.

Nucleotide sequencing

The nucleotide sequences of the clones obtained were made on both strands by the Sanger dideoxy-mediated chain-termination method according to Maniatis *et al.* (1989) and automatically (IGBMC,

Nucleotide Sequence Service) using as primers T7 and Sp6 sequences included in the vector.

Genomic DNA PCR amplification and cloning

Genomic DNA was prepared from *D. hansenii* spheroplasts formed by zymolyase digestion according to the Cryer method (Philippsen *et al.*, 1991). Genomic DNA fragments were generated by partial digestion with *Sau3* AI and purified on a linear sucrose gradient (10–40%). Fragments between 2 and 4 kb were used as templates for PCR amplification using the specific primers NT2 and CT2. These primers were designed from the *D. hansenii* Cu-Zn SOD cDNA sequence. The reaction mixture and amplification conditions to PCR were the same as in RT-PCR, as were the cloning and sequencing methods. In addition, genomic DNA fragments (2–4 kb) were used to construct a genomic library. The fragments were ligated in the *Bam*HI site of the λ ZAP Express vector. The growth host strain was XL-1 Blue MRF' and a combinant number 5×10^4 (recombinant percentage >90%).

Sequence analysis

The FASTA program was used to identify homologies of nucleotide *D. hansenii* in cDNA sequence with those contained in the EMBL and GenBank data libraries. DNA Strider™ version 1.2 program was used to translate the nucleotide sequence to protein, and to obtain information about codon usage, full restriction map, amino acid composition, and molecular weight. The FASTA program was used to compare the deduced amino acid sequence of *D. hansenii* Cu-Zn SOD cDNA with that of the SwissProt data bank. Homology sequence analysis to construct a dendrogram was done comparing the *D. hansenii* Cu-Zn SOD sequence with others contained in the SwissProt data bank; *S. cerevisiae* (P00445), *S. pombe* (P28758), *N. crassa* (P07509), *Zea mays* (P23346), *Schistosoma mansoni* (PQ 011137), *Drosophila melanogaster* (P00444), *Xenopus laevis* (P13926), *Caretta caretta* (P80174), *Prionace glauca* (P11418), *Xiphias gladius* (P03946), *Gallus gallus* (P80566), *Mus musculus* (P08228), *Bos taurus* (P00442), *Equus caballus* (P00443), *Homo sapiens* (P00441), and *H. sapiens* (Cu-Zn SOD extracellular (P08294).

Statistical analysis

A cluster analysis was made by use of the STATISTICA software Version 4.5 StatSof. The

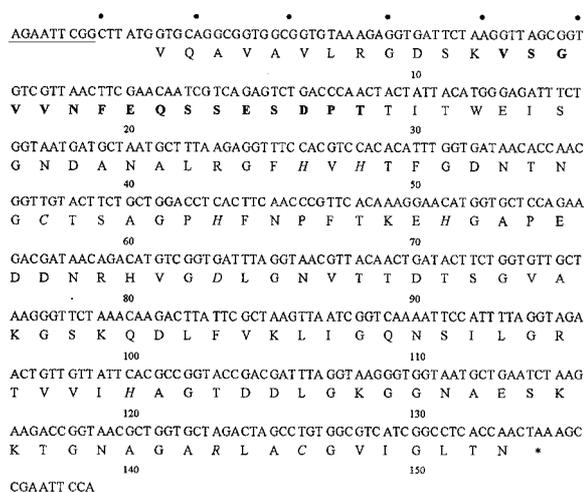


Figure 1. Nucleotide sequence of the *D. hansenii* Cu-Zn SOD cDNA (*dhsod-1*) and the deduced amino acid sequence. The translated amino acid sequence is depicted under the nucleotide sequence. Copper-zinc binding sites (italics) are formed by six His residues (H 46, 48, 63, 71, 80 and 120) and one Asp (D 83). The two Cys (C 57 and 146) form a disulfide bond. The Arg (R 143) is believed necessary to guide the superoxide radicals to the active site. The *Eco*RI cloning sites are underlined, and the asterisk denotes the stop signal. The amino acid sequence deduced from trypsin-digestion peptide sequencing is shown in bold type.

method used to measure distance was Euclidean distance and the linkage was made by unweighted pair-group average (UPGA).

RESULTS

Protein sequencing

One stretch of peptide sequence was obtained from the pure *D. hansenii* Cu-Zn SOD protein after trypsin digestion. The sequence obtained, VSGVVNFEQSSSEDPT, is located near the N-terminal end (Figure 1). This sequence showed 81.5% identity (in a 16-amino acid overlap) with the reported *S. cerevisiae* Cu-Zn SOD sequence (P00445), thus the N- and C-terminal sequences of the *S. cerevisiae* Cu-Zn SOD nucleotide sequence (J03279) were used to design the degenerate oligonucleotides (NT1 and CT1) used in the cDNA PCR amplification.

Amplification by RT-PCR and cloning

A procedure involving RT, PCR amplification, and cloning was employed to obtain clones encoding the Cu-Zn SOD enzyme. In this procedure, the

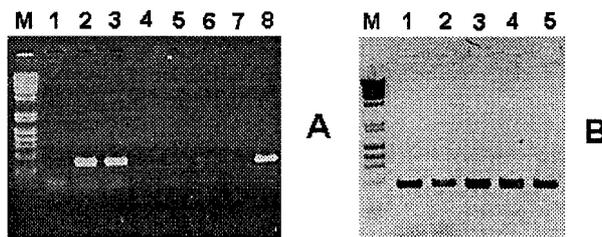


Figure 2. PCR products using: (A) degenerated (NT-1, CT-1) or (B) specific oligonucleotides (NT-2, CT-2), and as templates cDNA or DNA. Gel A: lane 1, negative control without template; lanes 2 and 3, *S. cerevisiae* cDNA, 5 and 10 ng; lanes 4, 5 and 6, *D. hansenii* cDNA, 10, 5 and 1 ng; lane 7, *D. hansenii* genomic DNA, 10 ng; lane 8, *scsod-1* gene, 10 ng. Gel B: lane 1, *scsod-1* gene, 10 ng; lane 2, *S. cerevisiae* cDNA 5 ng; lanes 3 and 4, *D. hansenii* C115 and C116 clones DNA, 10 ng each; lane 5, *D. hansenii* (2–4 kb) genomic DNA, 10 ng; lane M, markers lambda *Hind*III/*Phi* X174 *Hae*II.

primer PB-dT¹⁷ was used to reverse transcribe the *D. hansenii* vegetative mRNA into a single-stranded cDNA. The cDNA obtained was amplified by PCR using the degenerate oligonucleotides NT-1 and CT-1 (Figure 2A), and then cloned in the pCRII vector. Through this procedure, two identical clones (C115 and C116) containing fragments of 470 bp were obtained. The nucleotide sequence of these clones, named *dhsod-1*, contain a 462 bp coding region (Figure 1), which encodes a 153 amino acids protein (Figures 1 and 3) with a predicted molecular mass of 15.92 kDa. These data are in agreement with the molecular weight of the pure protein obtained from *D. hansenii* (N. Y. Hernández-Saavedra and J. L. Ochoa, submitted to *Yeast*). In addition, the same size and nucleotide sequences were obtained in several clones when the partially digested genomic DNA was used to do the PCR amplification with the specific primers NT-2 and CT-2 (Figure 2B). The positive controls (X2180 *S. cerevisiae* total RNA and DNA *sod-1* *S. cerevisiae* probe) produced fragments of 462 bp (Figure 2), which were cloned and sequenced, showing a 100% identity with the sequence of *S. cerevisiae* Cu-Zn *sod-1* gene included in the EMBL and GenBank data libraries (accession number J03279).

Sequence analysis

The sequence homology analysis (FASTA, GCG) shows clones having 55–71% identity with 15 cytosolic Cu-Zn SODs from other species, including human extracellular enzyme (39.4% identity; Figure 3). Most of the identical residues

are present at the corresponding positions in SODs from the other species, representing regions involved in metal binding and catalysis (Figure 3). The variable regions are located from amino acids 22 to 42 and 90 to 103 (Figure 3). These regions do not constitute the core structure of the protein, thus variability in these regions should not significantly affect catalytic activity and stability of the enzyme.

The nucleotide homology of *dhsod-1* with the sequences of equivalent genes from other fungi is >70% for *S. cerevisiae* and *N. crassa* and 62 and 64% when compared against *S. pombe* and *A. japonicus* respectively. As expected, in nucleotide bases there are more differences among sequences than in protein sequences. Evidently, major differences exist for codon usage among the four species considered; in particular, *N. crassa* differs from the other fungi. The amino acid balance has no strong variations: glycine content in all the enzymes is high ($\approx 15\%$), whereas tryptophan and tyrosine are scarce. The major differences in amino-acid content are found in the number of isoleucine, leucine, proline, and serine residues. Other amino acids do not show significant variations.

Cluster analysis was made considering all the protein sequences. Figure 4 shows four main branches (78% distance). One branch links higher eukaryotes (including nematodes, amphibians, reptiles, birds, and mammals), the second comprises only fungi, and the last two are formed by the arthropoda and the extracellular form of the human enzyme.

10 30 30
D.h. V Q A V A V L R G D S K V S G - V V N F E Q S S E S D P T T I T W E I S G N D
S.c. - V Q A V A V L R G D A G V S G - V V K F E Q A S E S E P T T V S Y W E I A G N S
S.p. M V R A V A V L R G D S K V S G - V V T F E Q Q V D Q N S Q V S V I V D L - G N D
N.c. - V K A V A V L R G D S N V K G T V I - F E Q E S E S A P T T I T Y D I S - G N D
Z.m. - V K A V A V L - G S S D G V K G T I F F E Q E G D - G P T A V T G S V S G L K
D.m. V V K A V C V I N G D A K - - G T V - F F E Q E S S G T P V V K V S G E V S G L K
S.m. - M K A V C V M T G T A G V K G - V V K F T T O E T D N G P V V H V H A E F V S G L K
Xl. - V K A V C V L A G S G D V K G - V V R F F E Q Q D D G D - V T V E G K I E G L -
C.c. A T K A V C V L K G E D P V K E P V K G * F E Q Q - G N G P V - L S G S I T G L -
P.g. M - K A V C V L K G T G E V T G T V L - F E Q A A D G - P V T L K G S I T G L P
X.g. V L K A V C V L R G A G E T T G T V Y - F F E Q E G N A N A V V G K G I I L K T G L -
G.g. A T L K A V C V M K G D G P V E G - V I H F F E Q Q G S - G P V K V T G K I T G L -
M.m. A M K A V C V L K G D G P V Q G - T I H F F E Q Q K A S G E P V V L S G Q I T G L -
E.c. A L K A V C V L K G D G P V H G - V I H F F E Q Q Q E G G P V V L K - G F I E G L -
B.b. A T K A V C V L K G D G P V Q G - T I H F F E - A K G D T V V T G S I T G L
H.s. A T K A V C V L K G D G P V Q G - I I N F E Q K E S N G P V K V - W G S I K G L
H.s. EC. Q V Q P S A T L D A A Q P R V T G V V L F E R Q L A P R A K L D A F F A L E G E P

40 50 60 70
D.h. A N A L R G F H V H T F G D N T N G C C T V S A G P P H F N P P F T K E H G A P E D D N V
S.c. P N A E R G F H H H E F G D N T N G C C C S A G P P H F N P P F G K T H G A D R T A A V V
S.p. A N A K R G F H H H Q F G D N T N G C C C S A G P P H F N P P F E G K T H G D R T A A V V
N.c. P N A K R G F H H H T F G D N T N G C C C S A G P P H F N P P H G T T H G D R T A A V V
Z.m. P - G L H G F H V H A L G D D T N T N G C C C M S T G P P H F N P P A S K E H G A P E D D N V
D.m. A K G L H G F H V H E F G D A T N T N G C C M S S G P P H F N P P Y G K E H G A P E D D N V
S.m. A G K - H G F H V H E F G D T T N T N G C C T S A G P P H F N P P I K Q E H G A P E D D N V
Xl. T D G N H G F H V H V F G D N T N G C C C L S A G P P H F N P P Q N K N H G S P K D A D N E
C.c. T E G K H G F H V H E F G D N T N G C C C T S A G A H F N P P P G K N H G G P K D D E E
P.g. T P G K H G F H V H A F G D N T N G C C C I S A G P P H F N P P F S K N H G G P K D D E E
X.g. I P G E H G F H V H G F G D N T N G C C T S A G P P H F N P P A S K K H A G P K D D E E
G.g. S D G D H G F H V H E F G D N T N G C C C T S A G A H F N P P E G K Q H G G P K D A D E
M.m. T E G Q H G F H V H Q Y G D N T Q G C C T T S A G P P H F N P P H S K K H G G P K D A D E
E.c. T K G D H G F H V H E F G D N T Q G C C C T A G P P H F N P P L S K K H G G P K D D E E
B.b. E G D - H G F H V H Q F G D N T Q G C C T T S A G P P H F N P P L S K K H G G P K D D E E
H.s. T E G L H G F H V H E F G D N T A G C C T T S A G P P H F N P P L S R K H G G P K D D E E
H.s. EC. N S S S R A I H V H Q F G D L S Q G C C E S T S A G P P H F N P P L A V P H - - - - P

80 90 100 110
D.h. R H V G D D L G N V T T D T S G V A K G S F K D S L F V K L I G Q N S I L G R S I V V
S.c. R H V G D D L G N V K T D D E N G V A K G S F K D S L F V K L I G P T S V V L G R S I V V
S.p. R H V G D D L G N L E S D A Q G N I K T T F S D S V I S L F G A N S I I G R T I V V
N.c. R H V G D D L G N I E T D A Q G N A K G T V T D N L V K L I G P E S V I I G R T I V V
Z.m. R H A G D D L G N V T A G A D G V A N I N V T D S Q I P L T G P N S I I G R A V V V
D.m. R H L G D D L G N I E A T A D G C P T K V N I T D S K I T L F G A D S I I G R T M V V
S.m. R H V G D D L G N V V A G A D G N A V Y N A T D K L I S L N G S H S I I G R T M V V
Xl. R H V G D D L G N V T A E - G G V A Q F K F T D P Q I S L K G E R S I I G R T A V V
C.c. R H V G D D L G N V I A N K E G V A E V C I K D S L I S - T G S Q S I I G R T M V V
P.g. R H V G D D L G N V E A N G N G V A E F E I K D R Q L H L S G E R S I I G R T L V V
X.g. R H V G D D L G N V T A D A N G V A K I D I T D K - I S L T G P Y S I I G R T M V V
G.g. R H V G D D L G N V T A - K G G V A E V E I E D S V I S L T G P H C I I G R T M V V
M.m. R H V G D D L G N V T A G K D G V A N V S I E D R V I S L S G E H S I I G R T M V V
E.c. R H V G D D L G N V T A D E N G K A D V D M K D S V I S L S G K H S I I G R T M V V
B.b. R H V G D D L G N V T A D K N G V A I V D I V D P L I S L S G E Y S I I G R T M V V
H.s. R H V G D D L G N V T A D K D G V A D V S I E D S V I S L S G D H C I I G R T L V V
H.s. EC. Q H P G D D L G N F A V R D G G L W R - Y R A G L A A S L A G P H S I V G R A V V

120 130 140 150
D.h. I H A G T D D L G K G O N A E S K K T G N A G A R R L A C G V I G L T N -
S.c. I H A G Q D D L G K G D T E E S L K T G N A G A P R P A C G V I G L T N -
S.p. I H A G E D D L G K G T S E P S L K T G N A G A P R R N A C G V I G I A V -
N.c. I H A G T D D L G K G G N B E L S L K T G N A G A P R P A C G V I G I S Q -
Z.m. V H A D P D D L G K G G H E L S K S T G N A G A G R R V A C G I I G L Q G -
D.m. V H A D A D D L G Q G G H E L S K S T G N A G A G R R I G C G V I G I A K
S.m. I H E N E D D L G R G G H E L S K V T G N A G G R R L A C G V I G I L A A E
Xl. V H E K Q D D L G K G G D D E S L K T G N A G G R R L A C G V I G I F C P -
C.c. V H E K E D D L G K G G N D E S L K T G N A G S R R L A C G V V G I A K L
P.g. V H E K E D D L G K G G D F E S L R T G N A G S R R L A C G V I G I A K D
X.g. V H E K A D D L G R G G N E E S L K - G N A G S R R L A C G V I G I - T - E
G.g. V H A K S D D L G R G G D N E S K L T G N A G P R R L A C G V I G I A K C
M.m. V H E K Q D D L G K G G N E E S T K T G N A G S R R L A C G V I G I A Q -
E.c. V H E K Q D D L G K G G N E E S T K T G N A G S R R L A C G V I G I A P -
B.b. V H E K P D D L G R G G N E E S T K T G N A G S R R L A C G V I G I A Q -
H.s. V H E K A D D L G K G G N E E S T K T G N A G S R R L A C G V I G I A Q -
H.s. EC. V H A G E D D L G R G G N Q A S V E N G N A G R R L A C C V V G V C G -

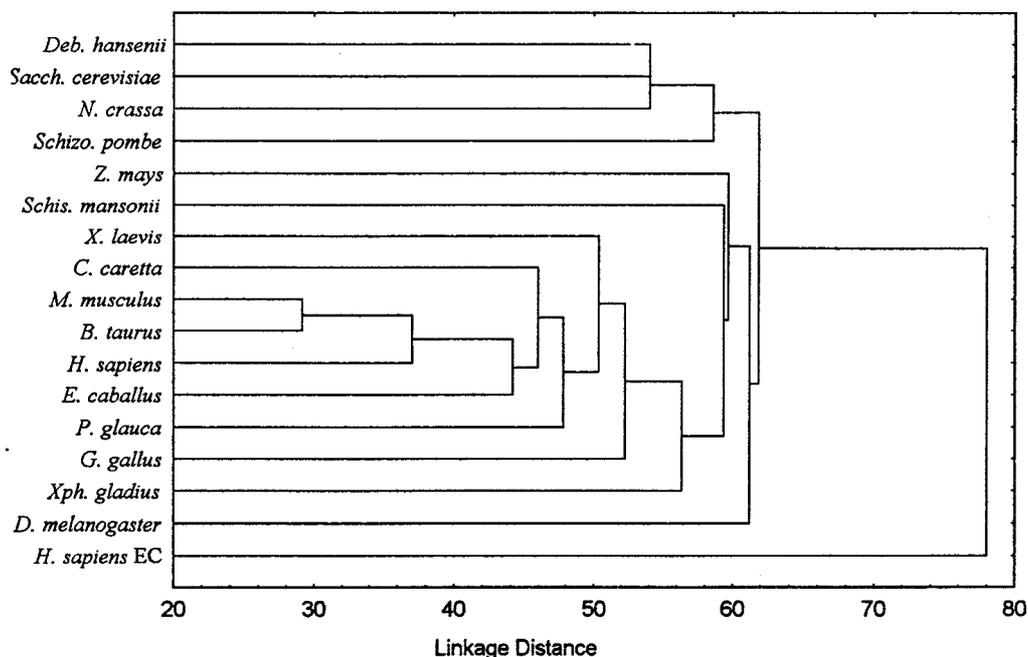


Figure 4. Tree representation of Cu-Zn SOD relationships, indicating the relative distances between 17 species of eukaryotic organisms.

DISCUSSION

Previous efforts to clone the *D. hansenii sod-1* gene (*dhsod-1*), using the *S. cerevisiae sod-1* gene as a probe, were unsuccessful (unpublished data). This was the first evidence of sequence differences between these genes, in addition to physicochemical characteristics of the pure proteins (N. Y. Hernández-Saavedra and J. L. Ochoa, submitted to *Yeast*). Because of these results, the use of the RT-PCR technique as a cloning strategy was developed. The clones obtained (C115 and C116 *dhsod-1*) contain all the necessary information to produce a functional protein (Figure 1) when compared with sequences from SwissProt data bank (Figure 3). The metal binding and disulfide amino acid residues (responsible for the correct protein folding) are present in the cloned encoding

sequence (Figure 3). Histidines 46, 48, 63, and 120 theoretically function as Cu^{2+} binding sites, and histidines 63, 71, 80 and aspartic acid 83 theoretically function as Zn^{2+} binding sites. Finally, cysteine 57 and 146 are believed to be involved in disulfide bond formation, while arginine 143 is believed to be necessary to guide the superoxide radical to the active site (Figure 3; see Parker *et al.*, 1986). Only 26 additional residues are conserved in all 16 sequences analysed (including the human extracellular Cu-Zn SOD) and, of these, 13 are glycines (Figure 3, boxes) which are connected with the three-dimensional tight packing of the molecule (Parker *et al.*, 1986). From the genomic DNA amplification using as primers NT-2 and CT-2, we obtained PCR products with the same size (Figure 2) and sequence (100% identity) as the C115 and C116 clones. We may then conclude

Figure 3. Amino acid sequence comparison of known Cu-Zn SODs from various species. The conserved histidine (H) and aspartate (D) involved in metal binding, the cysteine (C) residues involved in disulfide-bond formation, and the arginine (R) involved in directing superoxide anion to the copper atom are indicated in black boxes. Amino acid residues common to all listed SOD sequences are denoted in dark shadow, while conservative amino acid changes are denoted by light shadow. The dashes are spaces introduced for alignment of the sequences. The numbering is for the *D. hansenii* (*D.h.*) amino acid sequences. *S.c.*, *S. cerevisiae* (budding yeast); *S.p.*, *S. pombe* (fission yeast); *N.c.*, *N. crassa* (mould); *Z.m.*, *Z. mays* (corn); *D.m.*, *D. melanogaster* (fruit fly); *S.m.*, *S. mansoni* (parasite); *X.l.*, *X. laevis* (frog); *C.c.*, *C. caretta* (loggerhead); *P.g.*, *P. glauca* (blue shark); *X.g.*, *Xph. gladius* (swordfish); *G.g.*, *G. gallus* (chicken); *M.m.*, *M. musculus* (mouse); *E.c.*, *E. caballus* (horse); *B.b.*, *B. bovis* (bovine); *H.s.*, *H. sapiens* (human); *H.s.EC.*, *H. sapiens* extracellular.

that: (a) the encoding region in the *dhsod-1* gene has no introns, because no PCR products larger than 470 bp were obtained; and (b) the full-length *dhsod-1* encoding region is represented at least once in the constructed library. Our next step was to screen the library using, as a probe, the *dhsod-1* encoding sequence (clones C115 and C116) or by means of nested PCR (Rashtchian, 1995).

The codon usage among yeasts and moulds is quite different. Amino-acid differences at particular positions are mainly caused by codon substitutions in the third position. Even though these substitutions are conservative, the cumulative effect of differences of up to 30% of amino acids is a reflection of a distinct evolutive pattern. This agrees with the cluster analysis that results in four natural groups. In general, the cluster analysis produces topologies agreeing with the evolutionary tree of cytochrome *c* (Ayala, 1994), SOD (Parker *et al.*, 1986), and other trees derived from classical taxonomy and paleontology (Dobzanzky *et al.*, 1979).

ACKNOWLEDGEMENTS

We thank Drs E. B. Gralla and J. S. Valentine for providing the cloned *scsod-1* gene from which the probe for PCR amplification was prepared and used as positive control, Drs L. Adler, M. Lotz, J. Herrera and J. M. McCord for their suggestion and comments, J. C. Marinoni for technical advice, Dr Ellis Glazier for editing the English and INSERM (France) and CONACyT (México) for financial support. N.Y.H.-S. is the recipient of a CONACyT scholarship number 87142.

REFERENCES

- Adler, L. and Gustafsson, L. (1980). Polyhydric alcohol production and intracellular amino acid pool in relation to halotolerance of the yeast *Debaryomyces hansenii*. *Arch. Microbiol.* **124**, 123–130.
- Adler, L. (1986). Physiological and biochemical characteristics of the yeast *Debaryomyces hansenii* in relation to salinity. In Moss, J. (Ed.), *The Biology of Marine Fungi. 4th International Marine Mycology Symposium. Portsmouth, England*. Cambridge University Press, Cambridge, England, pp. 81–90.
- Amano, A., Shizukuishi, S., Tamagawa, H., Iwakura, K., Tsunasawa, S. and Tsunemitsu, A. (1990). Characterization of superoxide dismutases purified from either anaerobically maintained or aerated *Bacteroides gingivalis*. *J. Bacteriol.* **172**(1), 1457–1463.
- Ayala, F. J. (1994). El reloj molecular de la evolución. *Ciencia y Desarrollo* **119**, 43–45.
- Benov, L. and Fridovich, I. (1994). *Escherichia coli* express a copper- and zinc-containing superoxide dismutase. *J. Biol. Chem.* **269**(41), 25,310–25,314.
- Benov, L. and Fridovich, I. (1996). Functional significance of the Cu-Zn SOD in *Escherichia coli*. *Arch. Biochem. Biophys.* **327**(2), 249–253.
- Birmingham-McDonogh, M., Gralla, E. B. and Valentine, J. S. (1988). The copper zinc-superoxide dismutase gene of *Saccharomyces cerevisiae*: Cloning, sequencing, and biological activity. *Proc. Natl. Acad. Sci. USA* **85**, 4789–4793.
- Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**, 1513–1515.
- Chary, P., Halliwell, R. A. and Natvig, D. O. (1990). Structure, exon pattern, and chromosome mapping of the gene for cytosolic copper-zinc superoxide dismutase (sod-) from *Neurospora crassa*. *J. Biol. Chem.* **265**(31), 18,961–18,967.
- Dobzanzky, T., Ayala, F. J., Stebbins, G. L. and Valentine, J. W. (Eds) (1979). *Evolution*. H. F. Freeman and Co., San Francisco, USA.
- Fridovich, I. (1989). Superoxide dismutases: an adaptation to a paramagnetic gas. *J. Biol. Chem.* **264**, 7761–7764.
- Gorvind, N. S. and Banaszak, A. T. (1992). Isolation and characterization of an autonomously replicating sequence (ARSD) from the marine yeast *Debaryomyces hansenii*. *Mol. Mar. Biol. Biotechnol.* **1**(3), 215–218.
- Halliwell, B. and Gutteridge, J. M. C. (1991). Protection against oxidants in biological systems: the superoxide theory of oxygen toxicity. In *Free Radicals in Biology and Medicine*. Clarendon Press, Oxford, pp. 86–187.
- Hernández-Saavedra, N. Y. (1991). Bachelor Thesis. Aislamiento y Caracterización de Levaduras Marinas en la Costa Occidental de Baja California Sur, México. Universidad Nacional Autónoma de México (UNAM), Ciudad de México.
- Hernández-Saavedra, N. Y., Ochoa, J. L. and Vázquez-Duhalt, R. (1995). Osmotic adjustment in marine yeast. *J. Plankton Research* **17**(1), 59–69.
- Hjalmarsson, K., Marklund, S. L., Engström, A. and Edlund, T. (1987). Isolation and sequence of complementary DNA encoding human extracellular superoxide dismutase. *Proc. Natl. Acad. Sci. USA* **84**, 6340–6344.
- Hobot, J. A. and Jennings, D. H. (1981). Growth of *Debaryomyces hansenii* and *Saccharomyces cerevisiae* in relation to pH and salinity. *Exp. Mycol.* **5**, 217–228.
- Hong, Z., Lo Verde, P., Thakur, A., Hammarskjöld, M. L. and Rekosh, D. (1992). *Schistosoma mansoni*: cloning of a complementary DNA encoding a cytosolic Cu/Zn superoxide dismutase and high-yield expression of the enzymatically active gene product in *Escherichia coli*. *Exp. Parasitol.* **75**, 308–332.

- Köhler, K. and Domdey, H. (1991). Preparation of high molecular weight RNA. In Guthrie, Ch. and Fink, G. R. (Eds), *Methods in Enzymology, Vol. 194. Guide to Yeast Genetics and Molecular Biology*. Academic Press, New York, pp. 398–405.
- Kurtzman, C. P. (1994). Molecular taxonomy of the yeast. *Yeast* **10**, 1727–1740.
- Lin, C.-T., Lun, M.-T., Sheu, D.-C. and Duan, K.-J. (1994). Cloning and characterization of a copper/zinc-superoxide dismutase gene from *Aspergillus japonicus*. *Taiwana* **39**(1), 73–79.
- Maniatis, T., Sambrook, J. and Fritsch, E. F. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Marklund, S. L. (1984). Properties of extracellular superoxide dismutase from human lung. *Biochem. J.* **222**, 649–655.
- Matsumoto, T., Terauchi, K., Isobe, T., Matsuoka, K. and Yamakura, F. (1991). Iron and manganese containing superoxide dismutase from methylomonas: identity of the protein moiety and amino acid sequence. *Biochem.* **30**, 3210–3216.
- McCord, J. M. and Fridovich, I. (1969). Superoxide dismutase. An enzymatic function for erythrocyte (Hemocuprein). *J. Biol. Chem.* **244**, 6049–6055.
- Norkans, B. and Kylin, A. (1969). Regulation of the potassium to sodium ratio and salt tolerance in yeast. *J. Bacteriol.* 836–845.
- Parker, M. W., Bossa, F., Barra, D., Bannister, W. H. and Bannister, J. V. (1986). Structural and evolutionary relationships between the eukaryotic superoxide dismutases. In Rotilio, G. (Ed.), *Superoxide and Superoxide Dismutase in Chemistry, Biology and Medicine*. Elsevier Sci. Publ. BV, Amsterdam, pp. 237–245.
- Philippson, P., Stotz, A. and Scherf, Ch. (1991). DNA of *Saccharomyces cerevisiae*. In Guthrie, Ch. and Fink, G. R. (Eds), *Methods in Enzymology, Vol. 194. Guide to Yeast Genetics and Molecular Biology*. Academic Press, New York, pp. 169–182.
- Rashtchian, A. (1995). Novel methods for cloning and engineering genes using the polymerase chain reaction. *Curr. Op. Biotechnol.* **6**, 30–36.
- Rosenfeld, J., Capdeville, J., Guillermont, J. C. and Ferrara, P. (1992). In gel digestion of proteins for internal sequence analysis after one- or two-dimensional electrophoresis. *Anal. Biochem.* **203**, 173–179.
- Steiman, H. M. (1982). Copper-zinc superoxide dismutase from *Caulobacter crescentus* CB15. *J. Biol. Chem.* **257**, 10,283–10,289.
- Tibell, L., Hjalmarsson, K., Edlund, T., Skogman, G., Engström, A. and Marklund, S. L. (1987). Expression of human extracellular superoxide dismutase in Chinese hamster ovary cells and characterization of the product. *Proc. Natl. Acad. Sci. USA* **84**, 6634–6638.