Characterization of the manganese superoxide dismutase cDNA and gene from the human parasite *Onchocerca volvulus*

Kimberly HENKLE-DÜHRSEN*, Wilson TAWÉ, Caren WARNECKE and Rolf D. WALTER

Bernhard Nocht Institute for Tropical Medicine, Department of Biochemistry, D-20359 Hamburg, Germany

The manganese-containing superoxide dismutase (MnSOD) is a major component of the cellular defence mechanisms against the toxic effects of the superoxide radical. Within the framework of studies on anti-oxidant enzymes and their protective role in the human parasitic nematode *Onchocerca volvulus*, sequences encoding the MnSOD were isolated and examined in this study. Degenerate primers were designed based upon conserved regions of MnSOD sequences from other organisms, and were used in PCR on reverse-transcribed *O. volvulus* total RNA and genomic DNA to identify partial cDNA and genomic DNA fragments encoding the *O. volvulus* MnSOD (*OvMnSOD*). The genomic DNA PCR product was used to screen an *O. volvulus* adult worm λ unizap II cDNA library and the nucleotide sequence of the longest clone determined. The complete 5'-end of the *OvMnSOD* cDNA was obtained using the rapid amplification of cDNA ends

(RACE) procedure with *O. volvulus* total RNA and was found to possess a spliced leader sequence at the 5'-terminus. The deduced primary sequence encodes a 25 kDa protein, which has the conserved residues required for enzyme activity and metal binding. The 24 N-terminal amino acids encoded by the *OvMnSOD* cDNA comprise a putative mitochondrial transit peptide. The *OvMnSOD* gene was also isolated from an *O. volvulus* adult worm λ fix II genomic library, a restriction map was constructed and the nucleotide sequence determined. The *OvMnSOD* gene was found to possess five exons and four introns with consensus splice-site junctions. Potential regulatory elements were identified in the 5' genomic flanking sequence. Southern-blot analysis with total worm genomic DNA indicates a single-copy gene, with a restriction pattern consistent with that of the isolated gene.

INTRODUCTION

Twenty million people are infected with the parasitic nematode *Onchocerca volvulus*. The resulting disease, onchocerciasis, is the world's third leading cause of preventable blindness. These parasitic worms can survive up to 10 years in man, encapsulated in nodules in the subcutaneous tissue, where they produce and release microfilariae which are responsible for most of the pathology of the disease. During its development, the worm encounters a competent immune system and host phagocytic cells are often observed in the proximity of the parasite. Attempts to define mechanisms which aid in the long-term parasite persistence, and which explain parasite resistance to elimination by the host immune response, has led to the investigation of, among other things, the *O. volvulus* anti-oxidant enzymes.

Anti-oxidant enzymes have been postulated to play a role in protecting parasitic nematodes from host-induced oxidative damage in addition to their normal, essential cellular scavenging functions [1]. Reactive oxygen species are produced as a normal by-product of aerobic metabolism, as well as by the burst of stimulated neutrophils and macrophages [2,3]. These toxic oxygen species can damage nucleic acids, proteins and membrane lipids and therefore, can be lethal. Oxidative damage has been shown to be inversely related to life-span [4]. To protect against these potentially damaging oxygen species, aerobic organisms have developed a multi-layered anti-oxidant defence system in which the anti-oxidant enzymes superoxide dismutase (SOD), glutathione peroxidase and catalase, as well as non-enzymic anti-oxidants, play a pivotal role. A number of anti-oxidant enzymes have been characterized in parasitic nematodes [5–12].

The SODs are a family of metalloenzymes involved in the first line of cellular anti-oxidant defense by the dismutation of the

superoxide radical into hydrogen peroxide and molecular oxygen. Superoxide is the one-electron reduction product of molecular oxygen. Over-expression of SOD has been shown to increase the mean life-span of *Caenorhabditis elegans* and *Drosophila melanogaster* [13,14]. In eukaryotic cells there are three types of SODs, a cytosolic CuZnSOD, an extracellular CuZnSOD and an MnSOD. Although these enzymes catalyse the same reaction, the MnSOD is structurally distinct from the CuZnSODs and each is encoded by a different gene. The MnSOD is encoded in the nucleus, synthesized in the cytosol and imported post-translationally into the mitochondrial matrix, where 90 % of the cell's oxygen is consumed. Mitochondria are especially sensitive to oxidative damage, due to their lack of histones and inefficient DNA repair [15], and mitochondrial DNA damage is implicated in aging and in a number of degenerative conditions [16,17]. Interestingly, the MnSOD is inducible by environmental oxidative stress [18–22]. Sequences encoding the two CuZnSODs in *O. volvulus* have been described [8,10,23]. The MnSOD cDNA and gene from *O. volvulus* are described here.

MATERIALS AND METHODS

Parasites

Adult *O. volvulus* worms were isolated and prepared as previously described [24]. Briefly, onchocercal nodules were surgically removed from Liberian onchocerciasis patients and worms were isolated from the nodules by collagenase digestion. The worms were analysed microscopically for intactness and purity from contaminating host material and then maintained in liquid nitrogen until use.

Abbreviations used: SOD, superoxide dismutase; MnSOD, manganese-containing SOD; *OvMnSOD*, *O. volvulus* MnSOD; RACE, rapid amplification of cDNA ends; GST, glutathione-S-transferase.

*To whom correspondence should be addressed.

Purification of *O. volvulus* RNA and genomic DNA

Adult *O. volvulus* worms were recovered from liquid nitrogen, homogenized immediately in a guanidinium thiocyanate solution, which was layered on a CsCl step gradient, and DNA and RNA were prepared as previously described [25].

Identification of *O. volvulus* PCR fragments encoding the MnSOD

Degenerate primers (MnSOD1: 5'-GGGAATTCGTG(CTA)C-AG(A)GGC(TA)TCG(CTA)GGC(TA)TGG-3' and MnSOD2: 5'-GGGAATTCG(A)TAA(G)TAT(A)GCATGTTCCCA-3' were designed based upon two highly conserved regions of the MnSOD amino acid sequences from other organisms. A PCR reaction with the MnSOD1 and MnSOD2 primers was performed with either *O. volvulus* genomic DNA or oligo d(T)-primed *O. volvulus* cDNA. The MnSOD cDNA and genomic DNA fragments were amplified by 35 cycles at 95 °C for 2 min, 45 °C for 2 min and 72 °C for 2 min. The amplified PCR products were analysed by agarose gel electrophoresis, purified and subcloned into the pUC18 vector (Pharmacia). The nucleotide sequences were determined by standard methods [26].

Identification of λ unizap II cDNAs encoding *O. volvulus* MnSOD (*OvMnSOD*)

The isolated *OvMnSOD* genomic DNA PCR fragment was labelled with [³²P]dCTP by the random primer method (Boehringer Mannheim) and used as a hybridization probe. An *O. volvulus* λ unizap II adult worm cDNA library was screened by standard methods [26]. Four positive clones were identified and the nucleotide sequences were determined. The largest fragment of 874 bp was further characterized and the nucleotide sequence was determined completely [26].

Completion of the *OvMnSOD* cDNA using 5' RACE

Primer extension using *O. volvulus* total RNA was performed to obtain the full-length cDNA sequence. Purified *O. volvulus* RNA (1 μ g) was used as the template in 5'-rapid amplification of cDNA ends (RACE; GIBCO BRL) as described by the manufacturers. Briefly, MnSOD10 primer (5'-ATGTGTCCGCCA-GTGTG-3') was used to prime the synthesis of cDNA, followed by degradation of the template RNA and purification with GlassMax spin cartridges. The cDNA was tailed with deoxycytosine (dC). The dC-tailed cDNA was then amplified by PCR with the MnSOD4 primer (5'-CGGCTCCAAGGCTCCAT-3') and the (dG)-anchor primer. A subsequent PCR was performed with the MnSOD3 nested primer (5'-ATCGTATGGTAAAT-CAGG-3') and the anchor primer. The PCR products were analysed by 3% NuSieve (FMC) agarose gel electrophoresis. The nested PCR product was ligated directly into the pCRII plasmid vector (TA Cloning System, Invitrogen) as described by the manufacturers, and the nucleotide sequence was determined.

Identification of λ fix II genomic DNA clones encoding the *OvMnSOD*

An *O. volvulus* adult worm λ fix II (Stratagene) genomic DNA library was screened with the [³²P]-labelled MnSOD cDNA using standard procedures. Prehybridization, hybridization and washes were at 55 °C. Twenty-seven positive clones were identified, ten were purified and DNA was prepared from six independent λ fix II-Mn*OvSOD* clones. Two of the six clones were selected for further analysis following *EcoRI*/*SalI* restriction cleavage of the DNA and analysis by agarose gel electrophoresis.

Southern-blot analysis

Genomic DNA and phage DNA Southern-blot analyses were performed by standard procedures as previously described [8,26].

Nucleotide sequence accession number

The *O. volvulus* MnSOD cDNA and gene have been assigned the EMBL accession Nos. X82171 and X82172 respectively.

RESULTS

Identification of cDNA and genomic sequences encoding the *OvMnSOD*

A comparison of the amino acid sequences of the MnSOD from human [27], *C. elegans* (N. Suzuki, N. Ishii and K. Suzuki, unpublished work) and yeast [28] revealed just two regions which are highly conserved between these three evolutionarily distant species. These amino acid stretches, VQGSGW and WEHAYY (positions 143 and 186 in the *O. volvulus* sequence in Figure 2) were used in designing degenerate oligonucleotides which would be likely to identify the *O. volvulus* nucleic-acid counterpart. A 140 bp PCR fragment was identified following PCR with *O. volvulus* cDNA and a 450 bp fragment was obtained with *O. volvulus* genomic DNA. The nucleotide sequences of both fragments were determined and were found to have similarity with other MnSODs.

Characterization of the λ unizap II *OvMnSOD* cDNA

The longest cDNA fragment encoding the *O. volvulus* MnSOD that was isolated from the λ unizap II cDNA library was 874 bp, beginning at nucleotide 37 in Figure 1. The complete nucleotide sequences of both strands of this cDNA were determined as

```

5' GGTTTAATTACCCAAGTTTGAGATTATTGATGAATCTGATTATTGGCATCGCAGTTCGCT 60
   M N L I I G I A G R L 11
TATTGGTTGGCAAGAACTATTGTCTGAATACACAACTAAACATGTTCTTCTCGATT 120
   L V G K N Y C L N T Q R L K H V L P D L 31
TACCATACGATTATGGAGCCTTGGAGCCGATATTGTGCGCGCAAAATTATGCAAGTTTCATC 180
   P Y D Y G A L E P I L S A E I M Q V H 51
ACGGTAAACATCATGCTGCATATGTAATGCTCTTAATCAAGCCGAAGAGAAAGTAAAG 240
   G K H H A A Y V N A L N Q A E E K V K E 71
AAGCGTAGCAAAAGGAGATACACAGGCTGCTGTGTGCTGGCAAAATTAATGAATTTCA 300
   A L A K G D T Q A A V A T G K L M N F N 91
ACACTGGCGGACACATTAATCATACACTGTTTGGGAAGGATTAAGTCCGCTAAAGAATA 360
   T G G H I N H T L P W E G I L T A V K N S 111
GCGGAGAACCGAATTCGAATTAATGACGGCTATAAGAAGGATTTCCGGSTCTTTGGAAA 420
   G E P S E L M T A I K K D F G S L E T 131
CGATGATAGATAAGTTAAATGCTAAACAATTGCGATCCAAGGTTCCGGATGGGGATGGC 480
   M I D K L N A K T I A I Q G S G W G W L 151
TTGCTTATGACAAAGAAATGAAACGTTTTCGCACTTGTCTGTGTGCCAACCAAGATCTTC 540
   A Y D K E M K R L Q L A C C P N Q D L L 171
TTGAGCCAAACAGGTCATAATTCACATATTGTGATCGATGTCTGGGAACATGCTTATT 600
   E P T T G L I P L F C I D V W E H A Y Y 191
ATCTGCAGTATAAAATTTACGCCAGATTTCGTGAAAGCATATTGGGAAAATTCGGAATT 660
   L Q Y K N L R P D F V W A G I W K I A N S 211
GGAAAATAATCAGTGATCGCTACATTAAAGCTAGAGGATAATCTTTATGAAAGCTATTA 720
   K I I S D R Y I K A R G * 223
AATGAAATTTGCTAAAAGGTACTCGACATTCATCACTTTGCTTCTCATGAAAATTTA 780
AATATCTCTTATAATACAAATTACAATTGAATTCGGTTTCATAATCTTCCAACTTTGA 840
ATGGAAGCATTATCAATAAGTATGTATAAATAAAGCATTGTCAACAAAAA 900
AAAAAAAAA - 3' 910

```

Figure 1 *OvMnSOD* cDNA and deduced amino acid sequences

Nucleotide residues 1–22 represent the splice leader sequence (underlined). The 5'-end of the λ zap cDNA is indicated (§). The polyadenylation signal sequence is underlined and the poly(A) tail is in italics. MnSOD3, 4 and 10 (overlines) represent the oligonucleotides used in the RACE procedure as described in the text. The two *EcoRI* restriction sites are shown in bold type. The putative mitochondrial transit peptide is in bold type. The conserved metal-binding residues are bold and underlined.

<i>O. v.</i>	--MNLIIIGIAGRLLVGKNYCLNTORLKHV--LPDLPLYDYGALEPILSAE	45
<i>C. e.</i>	--LQNTVVCVSK..QPITGVAAV..S..S..AD..VI..H..	45
<i>S. c.</i>	MPAKTAANLTKKGLSLSTTAR..T..VT..KW..P..YI..GQ	47
<i>H. s.</i>	--LSRAVCVGTSLAPALGYLGS..Q..S..HIN..Q	45
<i>R.</i>	--LCEAACSAGREL..PAASTAGS..H..S..HIN..Q	45
<i>E. c.</i>	-----MNMST..S..A..D..HFDKQ	24
<i>O. v.</i>	IMQVHGHGKHAAYVNA-----LNQAEKVKKEALAKGD-T-QA-AVAGTK-L	87
<i>C. e.</i>	..L..Q..T..N-----I..LH..VS..N-VKE..I..LQPA---	88
<i>S. c.</i>	NEL..YT..QT..GFNTAVD..FQ..LSDL-----E-P--S-P--NARKM	90
<i>H. s.</i>	..L..S..T..N-----VT..YQ-----V..A..T..LQPA---	88
<i>R.</i>	..L..S..T..N-----VT..YH-----V..T..V..LQPA---	88
<i>E. c.</i>	T..EI..T..QT..NANAA..ESLP..FANLPVBELITKLDLPADKKTIV	74
<i>O. v.</i>	M--NFW---TGGHINHTLFW--EGLTAVKNSG-EPNSBLMT-AIKKDFGS	128
<i>C. e.</i>	K-----G-----SI..TN..A--DG--SA..L--S--	126
<i>S. c.</i>	IAIQQ..IKFHG..FT..C..EN..APESQG..G..PTGALAK..DEQ	139
<i>H. s.</i>	K-----G-----SI..TN..SP-NGG--KG..LE--R--	128
<i>R.</i>	K-----G-----SI..TN..SP--GG--KG..LE--R--	127
<i>E. c.</i>	RN---A-----A..S..KG--K--TTLQGD--KA..ER--	111
<i>O. v.</i>	LETMIDK-L-NAKTIA-IQGSWGWL-AYD-KEMK---RLQLACC-PNQD	169
<i>C. e.</i>	DNL-Q..Q..S..S..Q..V-----G..CP..G-----I..KV..T..A--	167
<i>S. c.</i>	DEL..T..T..L..GV-----APIV-----NLSNGGK.DVVQ-TY--	181
<i>H. s.</i>	FDKFKK..T..ASV-GV-----GPN-K..RGH---I..AC--	169
<i>R.</i>	F..KFKK..T..VSV-GV-----GPN-----QG--I..A..S--H--	168
<i>E. c.</i>	VDFKAEFEKA.ASRF-----A..VLKGD.LAVVSTAN.D-----S	151
<i>O. v.</i>	LLEPTTGLIPLF---CIDVWEHAYLYQY-KMLRPDPVKAIWKIANWKII	214
<i>C. e.</i>	P..A..V-----G-----V..Y..N--NV	212
<i>S. c.</i>	TV-TGP-V..V---A..A-----Q--KKA.YF..NVV..EA	224
<i>H. s.</i>	P..QG..L---G-----V..YL..NVI..ENV	213
<i>R.</i>	P..QG..L---G-----V..YL..NVI..ENV	213
<i>E. c.</i>	P..MGEAISGASGPFPIGL---KP-Q..R..YI..BF..NVV..DEA	200
<i>O. v.</i>	SDRYIKARG	223
<i>C. e.</i>	E..FAKAQK	221
<i>S. c.</i>	R..FDAGKI	233
<i>H. s.</i>	TE..MACKK	222
<i>R.</i>	Q..VCKK	222
<i>E. c.</i>	AA.FAAKK-	208

Figure 2 Comparison of amino acid sequences of MnSODs from *O. volvulus* (*O. v.*), *C. elegans* (*C. e.*) (N. Suzuki, N. Ishii and K. Suzuki, unpublished work), yeast (*S. c.*) [28], human (*H. s.*) [27], rat (*R.*) [32a] and *E. coli* (*E. c.*) [32b]

Dots indicate amino acid residues in common with the *O. volvulus* sequence. Amino acids are numbered in the right hand column. Gaps were introduced into the sequences (–) to maximize the homologies and compensate for the different chain lengths. The basic residues within the leader sequence are underlined. The beginning of the mature peptide is indicated with the symbol (\$). Residues involved in metal binding are indicated with the symbol (*). The residues common to all six proteins are in bold type.

shown in Figure 1. This cDNA possessed a 220 amino acid open reading frame encoding the MnSOD, which terminated at the stop codon at position 701, but which lacked sequences encoding the N-terminal portion of the protein. The cDNA possessed a complete 3' non-translated region, including a poly(A) tail. This 189 bp 3' non-translated region, excluding the poly(A) tail, contained two contiguous and overlapping polyadenylation signal sequences 5'-AATAAATAAA-3' at position 869–878 (Figure 1). The *OvMnSOD* sequence does not possess the conserved octamer 5'-TACTGAAA-3' present in the 3' non-translated regions of the *O. volvulus* and *Schistosoma mansoni* CuZnSOD cDNAs [23].

Completion of the *OvMnSOD* cDNA with 5' RACE

Total *O. volvulus* RNA was annealed with the MnSOD-specific antisense primer MnSOD10 (Figure 1, nucleotides 299–316), and the RNA was reverse-transcribed. Following tailing of the first-strand cDNA with poly(dC), a PCR was performed with a poly(dG) anchor primer and the MnSOD4 antisense primer (Figure 1, nucleotides 133–149), which produced a 170 bp fragment. To ensure specificity, a second nested PCR with the poly(dG) anchor primer and the MnSOD3 antisense primer (Figure 1, nucleotides 114–131) was performed which produced a 150 bp fragment. The nucleotide sequence of the RACE product was determined and was found to possess the missing 5' portion of the cDNA encoding the N-terminal portion of the *OvMnSOD*

Table 1 Pairwise sequence comparison of the deduced *OvMnSOD* amino acid sequence with amino acid sequences of MnSODs from other organisms

For details of abbreviations see the legend to Figure 2

	<i>O. v.</i>	<i>C. e.</i>	<i>S. c.</i>	<i>H. s.</i>	<i>R.</i>	<i>E. c.</i>
<i>O. v.</i>	**					
<i>C. e.</i>	57	**				
<i>S. c.</i>	43	45	**			
<i>H. s.</i>	56	61	43	**		
<i>R.</i>	57	62	45	93	**	
<i>E. c.</i>	40	42	39	44	42	**

enzyme. In addition, the cDNA was found to possess a 5' spliced leader sequence, 5'-GGTTTAATTACCCAAGTTTG-3' [29–31].

Analysis of the deduced amino acid sequence of the *OvMnSOD* cDNA

The complete *OvMnSOD* cDNA encodes a 223 amino acid protein with a predicted molecular mass of 25 kDa. The four residues known to be responsible for binding to manganese (His-50, His-98, Asp-184, His-188) are present in the *OvMnSOD* amino acid sequence (Figure 1). At His-98 in the *OvMnSOD* sequence, there is a single potential N-glycosylation site in this enzyme sequence. In addition, a putative mitochondrial signal is encoded by amino acids 1–24. Removal of this 2.7 kDa mitochondrial transit peptide would result in a 22.3 kDa mature enzyme. Although sequence comparison of mitochondrial transit peptides reveals no obvious homology, targeting sequences are generally rich in positively charged and hydroxylated (amphiphilic) amino acids and generally lack acidic amino acids [32,33]. The putative transit peptide of the *OvMnSOD* possesses all of these characteristics. It lacks acidic residues, has two arginine (R) and one lysine (K) residues dispersed among hydrophobic amino acids and the hydroxylated amino acids tyrosine (Y) and threonine (T). One of the arginine residues of the *OvMnSOD* lies in the penultimate position of the transit peptide, which is invariable in all eukaryotic MnSODs.

The deduced amino acid sequence of the *O. volvulus* MnSOD was aligned with other vertebrate and invertebrate MnSODs for sequence homology comparison (Figure 2). The number of absolutely conserved residues in all five sequences was found to be 17%. Pairwise comparisons of the deduced *OvMnSOD* amino acid sequence with amino acid sequences of MnSODs from other organisms are shown in Table 1. The N-terminal mitochondrial leader sequence region displayed the lowest level of primary sequence identity. The potential N-glycosylation site at His-98 is present in the other species as well, except for yeast. The N-terminal LPDL region, which is completely conserved in all eukaryotic MnSODs, is also present in the *O. volvulus* sequence.

OvMnSOD genomic structure analysis

The λ fix II-*OvMnSOD* phage DNA from two independently isolated clones (1 and 2) were analysed. A restriction enzyme cleavage map of the approximately 15 kb *O. volvulus* genomic DNA inserts was constructed. To identify *OvMnSOD* gene fragments within the inserts, the λ fix II-*OvMnSOD* phage DNA samples were digested with *EcoRI*, *EcoRI/SalI* and *SalI*, resolved

Figure 3 *OvMnSOD* genomic structure analysis

(a) Ethidium-bromide stained 0.5% agarose gel with λ fix II-MnSOD genomic phage DNA clones. Lane 1, *EcoRI*, clone 1; lane 2, *EcoRI*, clone 2; lane 3, *SalI*, clone 1; lane 4, *SalI*, clone 2; lane 5, *EcoRI/SalI*, clone 1; lane 6, *EcoRI/SalI*, clone 2; lane 7, 1 kb size marker DNA (Gibco BRL). The lengths of some of the size standards are indicated. (b) Southern blot of the λ fix II genomic phage DNA clones in (a), probed with the *OvMnSOD* cDNA. The corresponding sizes of the hybridizing fragments are indicated. (c) Southern blot of *EcoRI*-digested *O. volvulus* (lane 1) and human (lane 2) genomic DNAs, probed with the *OvMnSOD* cDNA. The corresponding length of the hybridizing fragment is indicated. (d) Organization and *EcoRI* restriction map of the *OvMnSOD* gene.

on a 0.5% agarose gel (Figure 3a) and analysed by the Southern transfer method using the radiolabelled *OvMnSOD* cDNA to probe the blot. The genomic 2.6 and 1.3 kb *EcoRI* fragments hybridized with the cDNA probe as shown in Figure 3b (lanes 1,2,5,6).

In order to determine the complexity of the *OvMnSOD* gene, *O. volvulus* genomic DNA was digested with *EcoRI*, resolved on a 0.8% agarose gel and blotted onto a nitrocellulose membrane. The blot was hybridized with the entire *OvMnSOD* cDNA (Figure 3c). Human DNA was included as a control to exclude the possibility that the MnSOD gene was derived from contaminating host material of the *O. volvulus* nodule. A 2.6 kb *O. volvulus* genomic DNA *EcoRI* fragment specifically hybridized to the *OvMnSOD* cDNA probe (Figure 3c, lane 1), which is consistent with the restriction pattern of the isolated gene. The lack of additional hybridizing fragments indicates that the *OvMnSOD* is encoded by a single-copy gene. The 1.3 kb fragment, which hybridized to the cDNA in the phage DNA Southern blot, results from the hybridization of the 100 bp 3'-end of the *OvMnSOD* cDNA (Figure 3d), which was not detected in the genomic Southern blot due to the small size of the hybridizing portion of the probe to the gene fragment and to the fragment copy number in the genomic DNA. Specific hybridization to *O. volvulus* genomic DNA and not to human genomic DNA (Figure 3c, lane 2) confirmed that the isolated MnSOD gene was of *O. volvulus* origin.

The 2.6 and 1.3 kb *OvMnSOD* gene fragments were isolated, ligated into pUC18 and the sequences were determined and orientated. Sequences of the entire MnSOD coding region, as well as the introns and 1000 bp of flanking sequence (5' and 3') were determined (Figure 4). Figure 3(d) orients the 2.6 and 1.3 kb *EcoRI* fragments with the structure of the gene. The *OvMnSOD* gene is composed of five exons with four intervening sequences. Exons 1, 2, 3, 4 and 5 encode 16, 59, 40, 60 and 48 amino acids respectively. The introns 1, 2, 3 and 4 are 197, 360, 230 and 305 bp respectively. The genomic DNA sequence analysis showed that the *EcoRI* site at position 342–347 in the cDNA is interrupted by the third intron in the genomic sequence.

An examination of the intron/exon boundaries of the *OvMnSOD* gene revealed conserved 5'-GT-AG-3' intron splice-site junctions [34]. In addition, nucleotides which may also be involved in the splicing process were observed at identical positions in the flanking regions of all four of the *OvMnSOD* gene introns (underlined in Figure 4). A thymine was observed at the sixth position downstream from the 5' splice junction and at the fifth and ninth positions upstream from the 3' splice junctions of all *OvMnSOD* gene introns.

The cDNA and genomic DNA exons were compared for nucleotide sequence differences. An adenine at position 19 in the cDNA corresponds to a guanine at position 493 in the genomic DNA. This results in the replacement of a GTC codon (Val) in the genomic DNA by an ATC codon (Ile) in the cDNA. However, this is in the leader peptide domain (amino acid number 7), and both amino acids share common characteristics (aliphatic and non-polar), hence neither enzyme activity nor mitochondrial translocation are expected to be altered.

Analysis of the 5' flanking sequences of the gene

The *OvMnSOD* gene sequence was examined for potential regulatory elements. A GC-rich region was detected between positions 40 and 180 (Figure 4), which included a GC-like box (5'-TGGGCGGGAT-3') with 90% identity to the Sp-1 binding hexanucleotide core sequence (complementary strand, position 39–48). A potential CAAT-like box was located at position 274–

5'

gtaggtggcagtcaccaggtatgcacatattgatacatcccgccacccggtcatgcat 60

cgatggcgccacagctcctgaagaagatggaaggggaaggggcaactcagatgggagga 120

gcagttggggcaagatggcgccgctgctagactcgagttggtactgagatgagaagg 180

tttgcaagtcagtgctttactcttctccaaactggtagctttgtttgttcacgcatt 240

cacgtacgagaaacgtcgaaataacatcatgattgcaatctgttatttctgttactctg 300

catatcaggccatccatcatcagctatcggttctctaaattcgcatctgtctcgata 360

cataaattatctgtgactaaaccttccagaatcgttgcataatagatcatctgttga 420

acgatagctgcagctaatattcttctcattgataagatcttttacagattattgATGAAT 480

M N

CTGATTATTGGCGTCGCGAGTTCGCTTATTTGGTGGCAAGAAgtaaaattgataagtttc 540

L I I G V A G R L L V G K N 16

aagtccccaaaataactaacaagttgtgtgtgacttttcaactcaataacttattatg 600

aaattgttggatcggaataaaagaagttattaccaaaatcttttagtgaaactttatttga 660

aagctttttaaataatagctgaaaattattttgaacctaataagattcgatattacagc 720

TATTGTCTGAATACAAACGCTAAACCATGTTCTCTGATTTACCATACGATTATGGA 780

Y C L N T Q R L K H V L P D L P Y D Y G 36

GCCTTGGAGCGGATATTGTCGGCGCAAAATATGCAAGTTTCATCAGGTAACATCATGCT 840

A L E P I L S A E I M Q V H H G K H H A 56

GCATATGTAAATGCTCTTAATCAAGCCGAAGAGAAAGTAAAGAAAGCGCTAGCAAAAGGL 900

A Y V N A L N Q A E E K V K E A L A K G 76

tttggaacaattttgtctaaattttgagaatgaaggaacaaaactcatattacttggac 960

tcacagtttagataaaatttttaagcagactcaaatcttcaataaaagaataactttcttg 1020

tatttttttaattcatcttttagacaaatcactaatcccaagattcagatgacatcgattcc 1080

ctgaagtcagaaacggttgaattgtgtaagtaaatcaattgagacatttgaagc 1140

agtagttgacgctggtttcattttttctgttgaatatttctcaatgttttagataaaaa 1200

attaagattcttcatcagttgagactcaaatcattttctagttctgtttaaattttcttg 1260

agGAGATACACAGGCTGCTGTGTGTCGCAAAAATTAATGAATTTCAACTGCGGACA 1320

D T Q A A V A G T K L M N F N T G G H 95

CATTAACTACATCTGTTCTGGGAAGGATTAATCTGCGTAAAGAAATAGCGGAGAACGAA 1380

I N H T L F W E G L T A V K N S G E P N 115

TTgtgaagttcttctcctaaataatgtttttaagaattagtttcagtgtaaaatttcat 1440

S 116

ttccatcttcatatgcaatgagatagcttcataatttcaagaaatcttttagataatgagc 1500

aaaggaaagctcttgaaactcggaatcagtcctttaaatttgaatgatggttatcaa 1560

cagatctttacaaaattttttttttgttatggcaataatttttttagatgCTGAATTAA 1620

E L M

119

TGACGGCTATAAAGAGGATTTCCGGTCTTTGGAAACGATGATAGATAAGTTAAATGCTA 1680

T A I K K D F G S L E T M I D K L N A K 139

AAACAATTCGATCCAGGTTCCGGATGGGATGCTTGTCTTATGACAAGGAATGAAAT 1740

T I A I Q G S G W G W L A Y D K E M K R 159

GTTTGAACCTTGTCTTGTGTCACCAACAGATCTTCTTGTAGCCAACACAGTgaatagt 1800

L Q L A C C P N Q D L L E P T T G 176

ataacgaaaattggataggacactccttttttataaaagtccaattataataaatcat 1860

ttctcttaatacaaatgaattgttcatatcatggtctgtctgttcttttagggagatgaaa 1920

atcatatcagtgtaataaagtaggtctgtcttggataaacattacgtgtgagatagataa 1980

gcacattatgttatgtgcaaatatgaagtgatctgagaatttttgagagaatgtata 2040

caaaacggttttagacaactgcaaatgcacattttattgtctacctatgtgttttagGTCTA 2100

L

177

ATTCCACTATTTTGCATCGATGTCGGGAACATGCTTATTATCTGCAGTATAAAATTTTA 2160

I P L F C I D V H E Y L L Q Y K N L 197

CGGCCAGATTTCTGGAAGCTATTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 2220

R P D F V K A A W K I A N W K I I S D R 217

TACATTAAGCTAGAGGATActcttctcaagaagctattaaatgaaattgtctaaaggt 2280

Y I K A R G * 223

actcgacattccatcactttgcttctctcatgaaaatttaaatttcttctataatacaa 2340

attacaattgaaattccggttccataattcttccaaatttgaatggaagcatcaataaag 2400

tatgtataataataaaagcattgtcaacatttatgaactatcatcggaatgaattaa 2460

acatttaagaatagttcattaaaaaacctagtcctcattaggaacttgtgcaatttccg 2520

gtattttgatattgttcaattgttcatgtaaaattggttagttttcgaactagaacttcgat 2580

gaactttgttcttctgcaatcgagttcgaagtcagcttttatttgaagtattgtgactaat 2640

ctcaataaaatttttctcggtattatattgtattctctcatgttaagggtttttcagatta 2700

ccaatttttaggatgatctaaatgatgaaattacaatagttcactgatcaatccttagca 2760

accatttctaccacagaacaaaagcattaaaaactgttagagtc -3' 2808

Figure 4 The complete nucleotide sequence of the *OvMnSOD* gene

Introns and flanking sequences are represented in lowercase letters and exons in uppercase letters. Underlined sequences in the 5' region represent potential regulatory sequences identified, described in the text. The conserved nucleotides at the boundaries of introns are underlined. Sequences downstream from the cDNA polyadenylation site are represented in italics. *EcoRI* sites are indicated in bold type.

282 (5'-TTGCAATCT-3') which has 66% identity to the CAAT consensus and which lies within a ten residue palindromic sequence (5'-GATTGCAATC-3'). Two copies of the 9-mer sequence 5'-GATGCGCGC-3' were found at positions 62-70 and 134-142.

DISCUSSION

MnSOD is an inducible mitochondrial enzyme which scavenges superoxide anion produced during normal metabolism or after oxidative insult. The residues known to be required for enzymic activity were conserved in the deduced *OvMnSOD* sequence. The *OvMnSOD* mature enzyme subunit molecular mass of 22.3 kDa is typical of that found in other eukaryotic species (23 kDa), which are then commonly active as homotetramers. Crystallographic studies of the *Escherichia coli* FeSOD [35] and human recombinant mitochondrial MnSODs [36] have demonstrated the metal-binding sites of these two enzymes to be located at three histidine residues and one aspartate residue corresponding to His-26, His-74, Asp-160 and His-164 of the mature *OvMnSOD* (deduced by sequence alignment). These residues are conserved in sequences of all MnSODs determined so far. In addition, amino acids adjacent to these sites are also more conserved than distant amino acids. The most highly conserved region of MnSOD sequences is the area surrounding the last two ligand sites (Asp-160 and His-164) as these two sites are only separated by four amino acids. The potential N-glycosylation site (N-X-S/T) at His-74 in the *OvMnSOD* mature protein is also conserved in all other eukaryotic species, except for yeast. The *O. volvulus* protein sequence also possesses the expected putative mitochondrial leader sequence, with characteristics similar to other mitochondrial transit peptides, including the invariable penultimate arginine residue.

The complete nucleotide sequences of two other eukaryotic MnSOD genes have been determined to date, the yeast gene [28] and the rat gene [37], as well as the 5' non-translated region of the bovine MnSOD gene [38]. The yeast gene is encoded by a single uninterrupted exon, while the rat gene has five exons and four introns. The *OvMnSOD* gene and the rat SOD gene were aligned for comparison of their genomic organization. Whereas the rat gene covers 6.3 kb of genomic DNA, the *OvMnSOD* spans only 1.9 kb of genomic DNA. However, there is a striking similarity in the overall gene organization, with both genes being arranged on five exons and four introns. The positioning of introns in both genes, as well as the lengths of the exons (particularly the third, fourth and fifth), are similar.

An *O. volvulus* genomic DNA Southern blot, using the *OvMnSOD* cDNA as the probe, resulted in a hybridization pattern consistent with that of the isolated *OvMnSOD* gene. Unlike in humans [39], rats [37] and mice [40], which have two MnSOD genes per haploid genome, these results indicate that the *OvMnSOD* gene is a single-copy gene.

Superoxide-mediated induction of SOD activity was first demonstrated by Hassan and Fridovich [41]. Since then, several factors that regulate expression of MnSOD genes in prokaryotes have been identified [19,42,43]. The expression of the MnSOD gene in *E. coli* (*sodA*) is regulated by environmental factors such as O_2^- , iron chelators, compounds capable of increasing intracellular O_2^- flux and strong oxidants capable of positively changing the redox potential of the cell [44]. In eukaryotes, MnSOD, CuZnSODs, catalase, glutathione peroxidase and glutathione-S-transferase (GST) mRNA levels have been shown to be induced in various systems by oxidants. Interestingly, in certain eukaryotic cells which contain both the CuZnSOD and MnSOD, it is the MnSOD which is increased in response to oxygen toxicity threats. It has also been suggested by another group [22] that SODs are regulated in mammalian tissue by cytokines in a co-ordinated manner, rather than as a response of individual cells to oxidants.

The 5'-flanking region of the rat MnSOD gene lacks a TATA or CAAT box, but contains two copies of the Sp-1-binding hexa-

nucleotide core sequence (G/T)GGGCGG(G/A)(G/A)CT and one copy of the consensus octanucleotide (C/G)TGACT(C/A)A for the binding of the AP-1 (*c-jun*) and *c-fos* protein complex. These elements allow activation of genes by phorbol esters. The rat MnSOD gene also possesses three copies of the SV40 core enhancer sequence GTGG(T/A)(T/A)(T/A)G and one copy of the adenovirus E1A gene enhancer (A/C)GGA-AGTG(A/C) [37]. The bovine MnSOD gene promoter also lacks a CAAT and a TATA box, is GC-rich and contains two AP-2 and eight Sp-1 consensus sequences [38]. The yeast MnSOD gene contains sequences similar to the upstream activation sites UAS1 and UAS2 [28]. Despite the appreciable homology of the *OvMnSOD* coding region to that of the rat, bovine and yeast MnSODs, the promoter sequences of the genes show a virtual lack of similarity. In addition to the promoter elements described in the rat, bovine and yeast MnSOD genes, the *OvMnSOD* 5'-flanking sequences were searched for general eukaryotic promoter elements, as well as specific promoter elements for stress-inducible genes [45]. However, only a potential CAAT box and Sp-1-binding site were identified. The significance of the Sp-1 core sequence, the palindromic 10-mer/CAAT sequence and the copies of the 9-mer sequence have not been identified, but they may play a role in *cis*-regulation of *OvMnSOD* gene transcription.

The protective role that the MnSOD plays for the parasite *O. volvulus*, along with other anti-oxidant and detoxification enzymes such as the CuZnSODs [10,46], glutathione peroxidase and catalase (K. Henkle-Dührsen, V. Eckelt, M. L. Eschbach, C. Warnecke and R. D. Walter, unpublished work) and GSTs [47–49], is currently being investigated. Parasite persistence in the host is certainly dependent upon coordinating the joint activities of these enzymes to maintain their well-being under oxidative stress.

The authors thank Norio and Kenshi Suzuki (Tokai University, Japan) for kindly providing the unpublished *Caenorhabditis elegans* MnSOD nucleotide sequence. A portion of this research was conducted by W.T. in partial fulfillment of the requirements for a Ph.D. from the University of Hamburg. This work was supported in part by the BMFT Project No. 01KA9201/12 (K.H.D.) and the DAAD-fellowship awarded to W.T.

REFERENCES

- Callahan, L. H., Crouch, R. K. and James, E. R. (1988) *Parasitol. Today* **4**, 218–225.
- Docampo, R. and Moreno, S. N. J. (1984) in *Free Radicals in Biology* (Pryor W. A., ed.), ch. 8, Academic Press, New York.
- Jesaitis, A. J. and Dratz, E. A. (1992) *The Molecular Basis of Oxidative Damage by Leukocytes*, CRC Press, Boca Raton.
- Adelman, R., Saul, R. L. and Ames, B. N. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2706–2708.
- Lied, R. W. and Suquet, C. M. (1986) *Mol. Biochem. Parasitol.* **18**, 301–311.
- Sanchez-Moreno, M., Leon, P., Garcia-Ruiz, M. A. and Monteoliva, M. (1987) *J. Helminth.* **61**, 229–232.
- Batra, S., Chatterjee, R. K. and Srivastava, V. M. L. (1990) *Biochem. Pharmacol.* **40**, 2363–2369.
- Henkle, K. J., Liebau, E., Müller, S., Bergmann, B. and Walter, R. D. (1991) *Infect. Immun.* **59**, 2063–2069.
- Cookson, E., Blaxter, M. L. and Selkirk, M. E. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5837–5841.
- James, E. R., McLean, D. C. and Perler, F. (1994) *Infect. Immun.* **62**, 713–716.
- Tang, L., Ou, X., Henkle-Dührsen, K. J. and Selkirk, M. (1994) *Infect. Immun.* **62**, 961–967.
- Sato, K., Aoki, T. and Nakano, M. (1994) *Exp. Parasitol.* **78**, 210–216.
- Larsen, P. L. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8905–8909.
- Reveillaud, I., Niedzwiecki, A., Bensch, K. G. and Fleming, J. E. (1991) *Mol. Cell. Biol.* **11**, 632–640.
- Richter, C., Park, J. and Ames, B. N. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6465–6467.
- Fraga, C. G., Shigenaga, M. K., Park, J. W., Degan, P. and Ames, B. N. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4533–4537.
- Floyd, R. A. (1991) *Science* **254**, 1597.
- White, J. A. and Scandalios, J. G. (1988) *Biochim. Biophys. Acta* **951**, 61–70.
- Touati, D. (1988) *Free Radicals Biol. Med.* **5**, 393–402.
- Shull, S., Heintz, N. H., Periasamy, M., Manohar, M., Janssen, Y. M. W., Marsh, J. P. and Mossman, B. T. (1991) *J. Biol. Chem.* **266**, 24398–24403.
- Miao, Z. and Gaynor, J. J. (1993) *Plant Mol. Biol.* **23**, 267–277.
- Stralin, P. and Marklund, S. L. (1994) *Biochem. J.* **298**, 347–352.
- Henkle, K. J., Liebau, E. and Walter, R. D. (1993) *Mol. Biochem. Parasitol.* **58**, 173–176.
- Schulz-Key, H., Albiez, E. J. and Büttner, D. W. (1977) *Trop. Med. Parasitol.* **28**, 428–430.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbour Laboratory Press, New York.
- Ho, Y. S. and Capro, J. D. (1988) *FEBS Lett.* **229**, 256–260.
- Marres, C. A. M., Van Loon A. P. G. M., Oudshoorn P., Van Steeg, H., Grivell, L. A. and Slater, E. C. (1985) *Eur. J. Biochem.* **149**, 153–161.
- Krause, M. and Hirsh, D. (1987) *Cell* **49**, 753–761.
- Zeng, W., Alarcon, C. M. and Donelson, J. E. (1990) *Mol. Cell. Biol.* **10**, 2765–2773.
- Henkle-Dührsen, K., Liebau, E. and Walter, R. D. (1994) *Trop. Med. Parasitol.* **44**, 337–339.
- Roise, D. and Schatz, G. (1988) *J. Biol. Chem.* **263**, 4509–4511.
- Ho, Y. S. and Capro, J. D. (1987) *Nucleic Acids Res.* **15**, 1007.
- Takeda, Y. and Avila, H. (1986) *Nucleic Acids Res.* **14**, 4577–4589.
- Hartl, F. U. and Neupert, W. (1990) *Science* **247**, 930–938.
- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1994) in *Molecular Biology of the Cell*, 3rd edn., pp. 365–385, Garland Publishing, New York and London.
- Carlioz, A., Ludwig, M. L., Stallings, W. C., Fee, J. A., Steinman, H. M. and Touati, D. (1988) *J. Biol. Chem.* **263**, 1555–1562.
- Borgstahl, G. E. O., Parge, H. E., Hickey, M. J., Beyer, W. F., Hallewell, R. A. and Tainer, J. A. (1992) *Cell* **71**, 107–118.
- Ho, Y. S., Howard, A. J. and Capro, J. D. (1991) *Am. J. Respir. Cell. Mol. Biol.* **4**, 278–286.
- Meyrick, B. and Magnuson, M. A. (1994) *Am. J. Respir. Cell. Mol. Biol.* **10**, 113–121.
- Church, S. L. (1990) *Biochim. Biophys. Acta* **1087**, 250–252.
- Sun, Y., Hegamyer, G. and Colburn, N. H. (1993) *Gene* **131**, 301–302.
- Hassan, H. M. and Fridovich, I. (1979) *Arch. Biochem. Biophys.* **196**, 385–395.
- Storz, G., Tartaglia, L. A. and Ames, B. N. (1990) *Science* **248**, 189–194.
- Demple, B. and Amabile-Cuevas, C. (1991) *Cell* **67**, 837–839.
- Hassan, H. M. and Sun, H. H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3217–3221.
- Rushmore, T. H., Morton, M. R. and Pickett, C. B. (1991) *J. Biol. Chem.* **266**, 11632–11639.
- Henkle-Dührsen, K., Warnecke, C., Brattig, N., Liebau, E. and Walter, R. D. (1994) *Mol. Biochem. Parasitol.* **67**, 41–47.
- Liebau, E., Walter, R. D. and Henkle-Dührsen, K. (1994) *Mol. Biochem. Parasitol.* **63**, 305–309.
- Liebau, E., Walter, R. D. and Henkle-Dührsen, K. (1994) *Exp. Parasitol.* **79**, 68–71.
- Liebau, E., Wildenburg, G., Walter, R. D. and Henkle-Dührsen, K. (1994) *Infect. Immun.* **62**, 4762–4767.