Title: Cu,Zn Superoxide Dismutases from *Tetrahymena thermophila*: Molecular Evolution and Gene Expression of the First Line of Antioxidant Defenses

Author: Diana Ferro Rigers Bakiu Cristiano De Pittà Francesco Boldrin Franco Cattalini Sandra Pucciarelli Cristina Miceli Gianfranco Santovito

PII: S1434-4610(14)00125-4
DOI: http://dx.doi.org/doi:10.1016/j.protis.2014.12.003
Reference: PROTIS 25471

To appear in:

Received date: 28-5-2014
Revised date: 5-12-2014
Accepted date: 9-12-2014

Please cite this article as: Ferro, D., Bakiu, R., De Pittà, C., Boldrin, F., Cattalini, F., Pucciarelli, S., Miceli, C., Santovito, G.,Cu,Zn Superoxide Dismutases from *Tetrahymena thermophila*: Molecular Evolution and Gene Expression of the First Line of Antioxidant Defenses, *Protist* (2014), http://dx.doi.org/10.1016/j.protis.2014.12.003

ORIGINAL PAPER

Cu,Zn Superoxide Dismutases from *Tetrahymena thermophila*: Molecular Evolution and Gene Expression of the First Line of Antioxidant Defenses

Diana Ferro[a,b], Rigers Bakiu[c], Cristiano De Pittà[a], Francesco Boldrin[a], Franco Cattalini[a], Sandra Pucciarelli[d], Cristina Miceli[d], and Gianfranco Santovito[a,1]

[a]Department of Biology, via U. Bassi 58/B, University of Padova, 35100 Padova, Italy

[b]Institute for Evolution and Biodiversity, Westfälische Wilhelms-Universität, Münster, Germany

[c]Department of Aquaculture and Fisheries, Agricultural University of Tirana, Tirana, Albania

[d]School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy

Submitted May 28, 2014; Accepted December 1, 2014

Monitoring Editor: Eric Meyer

Running title: Cu,Zn SODs from *Tetrahymena thermophila*

[1]Corresponding author; fax +390498276300

e-mail gianfranco.santovito@unipd.it (G. Santovito).

In the present study, we describe the molecular and functional characterization of two Cu,Zn superoxide dismutase (SOD) genes, named *tt-sod1a* and *tt-sod1b* from *Tetrahymena thermophila*, a free-living ciliated protozoan widely used as model organism in biological research. The cDNAs and the putative amino acid sequences were compared with Cu,Zn SODs from other Alveolata. The primary sequences of *T. thermophila* Cu,Zn SODs are unusually long if compared to orthologous proteins, but the catalytically important residues are almost fully conserved. Both phylogenetic and preliminary homology modeling analyses provide some indications about the evolutionary relationships between the Cu,Zn SODs of *Tetrahymena* and the Alveolata orthologous enzymes. Copper-dependent regulation of Cu,Zn SODs expression was investigated by measuring mRNA accumulation and enzyme activity in response to chronic exposure to non-toxic doses of the metal. Our in silico analyses of the *tt-sod1a* and *tt-sod1b* promoter regions revealed putative consensus sequences similar to half Antioxidant Responsive Elements (hARE), suggesting that the transcription of these genes directly depends on ROS formation. These data emphasize the importance of complex metal regulation of *tt-sod1a* and *tt-sod1b* activation, as components of an efficient detoxification pathway allowing the survival of *T. thermophila* in continued, elevated presence of metals in the environment.

Key words: *Tetrahymena thermophile*; ciliated protozoa; Cu,Zn superoxide dismutase; metal-induced oxidative stress; copper.

Introduction

Superoxide dismutases (SODs; EC 1.15.1.1) are metalloenzymes widely distributed in prokaryotic and eukaryotic cells (Fridovich 1995). They constitute a protein family that catalyzes the dismutation of superoxide anion ($\bullet O_2^-$) to hydrogen peroxide (H_2O_2) and molecular oxygen (McCord and Fridovich 1988). The main physiological role of these proteins is believed to be the protection of the cells by scavenging $\bullet O_2^-$, which are produced under oxidative conditions by energy-generating processes (Keele et al. 1971; McCord and Fridovich 1969).

Depending on species, there may be up to three different metal-containing SOD enzymes in the cells, acting in mitochondrion, nucleus, cytoplasm and extracellular spaces. These SODs are the products of different genes and they are classified according to the redox metal in the active site.

Copper/zinc-containing SODs (Cu,Zn SODs) use copper for the catalytic reaction and they are found in the cytoplasm of eukaryotes, in the chloroplasts of some plants and in the periplasmic space of many bacteria. Furthermore, in eukaryotes, an extracellular SOD (EC-SOD) distinct from cytosolic Cu,Zn SOD has been found (Marklund 1982).

Iron- and manganese-containing SODs (Fe SOD and Mn SOD) bind specifically Fe or Mn (Bannister et al. 1987; James 1994). Fe SODs are found in prokaryotes, plants (chloroplasts), and protozoa while Mn SODs are present both in prokaryotes and mitochondrial matrix of eukaryotes. Numerous fungi possess both cytosolic and mitochondrial Mn SODs (Fréalle et al. 2006). A few classes of bacteria are also known to possess activity with either iron or manganese incorporated in the same protein moiety (cambialistic SODs) (Santos et al. 1999; Sugio et al. 2000).

Mn SOD and Fe SOD appear as homodimers or homotetramers and exhibit a high degree of sequence and structural similarity (Jackson and Cooper 1998; Lah et al. 1995; Parker et al. 1987); it

strongly suggested that these enzymes originate from a common ancestry. In contrast, Cu,Zn and Mn/Fe SODs never shared a common ancestry, and this is shown by their distinctive amino acid sequences (Smith and Doolittle 1992) and completely different tertiary structures (Tainer et al. 1983).

The discovery of a Ni SOD in *Streptomyces* and cyanobacteria (Youn et al. 1996) established a completely new SOD group with a unique Ni-hook motif.

Copper is an essential trace element that plays a vital role as a catalytic co-factor for a variety of metalloenzymes. Beside superoxide dismutase, other enzymes are cytochrome c oxidase (mitochondrial electron transport chain), tyrosinase (pigmentation), peptidylglycine alpha-amidating monooxygenase (PAM) (neuropeptide and peptide hormone processing) and lysyl oxidase (collagen maturation) (Harris 2000; Peña et al. 1999). Moreover, copper is redox active metal that can participate in electron transfer reactions, with the consequent production of oxidants which are able to oxidize cell components (Esterbauer et al. 1992). Copper catalyzes the formation of the highly reactive hydroxyl radicals from hydrogen peroxide (H_2O_2) via the Haber–Weiss reaction and it decomposes lipid peroxides to peroxyl and alkoxyl radicals, which favor the propagation of lipid oxidation (Ziouzenkova et al. 1998).

In the present study, we report the molecular characterization and gene expression analysis of two out of three Cu,Zn SODs from the ciliated protozoon *Tetrahymena thermophila*. In order to increase the physiological response, the cells have been cultured in presence of an excess of copper, which is used as pro-oxidant. Our data suggest that *Tetrahymena tt-sod1a* and *tt-sod1b* are the components of an efficient detoxification pathway in continued, elevated presence of metals in the environment.

With the aim of evaluating differences at molecular level in relation to their evolutionary history, we also characterized Cu,Zn SOD sequences from *Oxytricha trifallax* (Swart et al. 2013) and *Euplotes focardii,* an Antarctic psychrophilic ciliate (Pucciarelli et al. 2009; Valbonesi and Luporini 1993), and compared them with those from other Alveolata species. In this context, it is particularly interesting that Cu,Zn SODs appear to have evolved rapidly in relatively recent times, whereas Mn and Fe SODs

appear to have evolved at a relatively constant rate over the entire history of eukaryotes, (Lee et al. 1985).

Results

Organization of *T. thermophila* Cu,Zn SOD Genes

In the *Tetrahymena* Genome Database (scaffolds n. 8254435, 8254811 and 8254431) we have found three sequences belonging to Copper/Zinc superoxide dismutase family. To test primer pairs for quantitative real-time PCR analyses, RT-PCR amplifications were performed using mRNA from *T. thermophila*. The obtained amplicons were cloned and sequenced; their sequences correspond to the expected partial coding sequences of only two Cu,Zn SODs. We named the respective genes *tt-sod1a* (GenBank accession number XM_001032187.2) and *tt-sod1b* (GenBank accession number XM_001007667.1). We named the unamplified gene *tt-sod1x* (GenBank accession number XM_001033543.1).

The in silico analysis of *tt-sod1a* indicates that the open reading frame (ORF) includes 609 nt and it encodes a putative protein of 166 aa, with a deduced molecular weight of 17.9 kDa (Supplementary Material Fig. S1). The comparison between the genome scaffold and the *T. thermophila* EST database (http://ciliate.org/index.php/) suggests that the 5'- and 3'-UTR regions consist of 68 nt and 40 nt, respectively. The 3′-UTR region included a consensus polyadenylation signal sequence (AATAAA) at nt 593–598. This gene is organized in 2 exons and 1 intron. The genomic portion upstream the 5'UTR region (2000 nt) did not show the expected Antioxidant Response Element (ARE), but six putative consensus sequences similar to half Antioxidant Response Elements (hAREs, TGACNNN) are present at nt –137, –483, –945, –1,134, –1,352 and –1,528 (Fig. 1).

The in silico analysis of *tt-sod1b* indicates that the open reading frame (ORF) is 591 nt long, potentially coding for a putative protein of 196 aa with a deduced molecular weight of 21.4 kDa (Supplementary Material Fig. S2). The comparison between the genome scaffold and the *T. thermophila* EST database suggests that the 5'- and 3'-UTR regions consist of 86 nt and 131 nt,

respectively. The 3′-UTR region included two putative polyadenylation signals at nt 697–702 and 725–730. The 3′-UTR also contained two ATTTA sequences, at nt 710–714 and 801–805, which signals rapid degradation in certain mammalian mRNAs (Shaw and Kamen 1986). This gene is organized in 2 exons and 1 intron. The expected ARE are missing in the 2000 nt upstream the 5'UTR but five putative consensus sequences similar to hAREs are present at nt –559, –647, –1,053, –1,473 and –1,978 (Fig. 1).

Organization of *E. focardii* Cu,ZnSOD Genes

From the *E. focardii* transcriptomic sequences (Pucciarelli et al. 2010) we identified two putative Cu,Zn SOD sequences. To obtain the entire nanochromosome sequence we performed RATE-PCR starting from *E. focardii* genomic DNA as described in Material and Methods. We named the respective genes *ef-sod1a* and *ef-sod1b* (accession Nos KF740481 and KF740482, respectively).

The ORF of *ef-sod1a* includes 700 nt and encodes a putative protein of 187 aa, with a deduced molecular weight of 20.0 kDa (Supplementary Material Fig. S3). The 5'- and 3'-UTR regions consist of 63 and 73 nt, respectively. The 3′-UTR region did not include a consensus polyadenylation signal sequence.

The ORF of *ef-sod1b* is 676 nt long, potentially coding for a putative protein of 162 aa with a deduced molecular weight of 16.8 kDa (Supplementary Material Fig. S4). The 5'- and 3'-UTR regions consist of 90 and 97 nt, respectively. The 3′-UTR region did not include a consensus polyadenylation signal sequence.

Organization of *Oxytricha* Cu,Zn SOD Genes

In the *Oxytricha trifallax* Genome Database (http://oxy.ciliate.org/index.php/home/welcome) we have identified two putative Cu,Zn SOD sequences. We named the respective genes *of-sod1a* (contig

4245.0; ENA accession number AMCR01016070.1) and *of-sod1b* (contig 15813.0, ENA accession number AMCR01005476.1).

The ORF of *of-sod1a* consist of 1191 nt, and it contains one intron that spans nucleotides 394 to 518. The ORF encodes for a putative protein of 166 aa, with a deduced molecular weight of 17.5 kDa (Supplementary Material Fig. S5). The 5'- and 3'-UTR regions are 242 nt and 70 nt, respectively. The 3′-UTR region included three putative polyadenylation signals, at nt 887–892, 968–973 and 987–992. In the 3′-UTR region there is also a putative rapid degradation signal at nt 1100–1104.

The *of-sod1b* ORF is 1212 nt long, and it contains one intron that spans nucleotides 421 to 540. It encodes a putative protein of 166 aa, with a deduced molecular weight of 17.6 kDa (Supplementary Material Fig. S6). The 5'- and 3'-UTR regions consist of 217 and 69 nt, respectively. The 3′-UTR region included three putative polyadenylation signals, at nt 908–913, 989–994 and 1008–1013. In the 3′-UTR region there is also a putative rapid degradation signal at nt 1124–1128.

Phylogenetic Analyses and Molecular Modeling

In Figure 2, the deduced amino acid sequences of the two transcribed *Tetrahymena* genes were aligned with the Cu,Zn SODs of *E. focardii*, *O. trifallax*, *Bos taurus, Xenopus leavis* and *Homo sapiens*. Amino acids in the catalytic core are highly conserved. The copper and zinc coordination environment and the active site regions are fully conserved in SOD1a, whereas in SOD1b both His^{154} and Thr^{171} are replaced by Asn. Furthermore, SOD1b lacks the two evolutionary invariant cysteine residues (at positions 58 and 147 in human SOD1), which form the unique disulphide bridge present in the enzyme. This structure is expected to contribute to the stability of the tertiary structure of SODs.

In order to identify the best-fit model to analyze Cu,Zn SOD evolution, a Bayes factor comparison was performed to test the strict clock model against the non-clock model using Cu,Zn SOD coding sequences. The marginal likelihood values are shown in Supplementary Material Table S1. The

strict-clock model (marginal likelihood value –6014.7) is almost 12 log likelihood units better than the non-clock model (marginal likelihood value –6027.27). A difference exceeding 5 log likelihood units is usually considered very strong evidence in favor of the better model (Kass and Raftery 1995).

However, in phylogenetics the unrooted model of phylogeny and the strict molecular clock model are two extremes of a continuum. Despite their dominance in phylogenetic inference, it is evident that both are biologically unrealistic and that the real evolutionary process lies between these two extremes. Local molecular clocks are another alternative to the global molecular clock, because they allow different regions of the tree to have distinct evolutionary rates, but within each region the rate must be the same. This new method conveniently allows a comparison of the strict molecular clock against a large array of alternative local molecular clock models. As it is shown in Supplementary Material Table S1, CPP model (marginal likelihood value –6004.7) is 10 log likelihood units better than the strict-clock model (–6014.7) and nearly 6 log likelihood units better than the other two local molecular clocks, TK02 (marginal likelihood value –6010.4) and IGR (marginal likelihood value –6011.2) model. Thus, we used the CPP molecular clock model in the phylogenetic tree construction.

The jModelTest 2.1.3 software determined the GTR+I+G model as being the best-fit model of Cu,Zn SOD cDNA sequence evolution with a gamma shape value (four rate categories) of 1.59 using AIC and cAIC statistical criterion (–lnL= 5923.8791). Phylogenetic relationships of Cu,Zn SOD cDNA sequences were determined using the most powerful statistical method of BI. BI and maximum likelihood (ML) methods generate phylogenies with the same topology, depicted in the cladogram of Figure 3. Ciliate SOD1s emerges together and clearly separated from the orthologs of other organisms. In particular, *T. thermophila* SOD1b and SOD1x are clearly separated from *T. thermophila* SOD1a, and are more closely related to SOD1s from *Paramecium*, *Cryptocaryon* and *Ichthyophthirius* (posterior probability 96%; bootstrap values 54%).

ProtTest3 software determined WAG+G model as being the best-fit model of Cu,Zn SOD amino

acid sequence evolution with a gamma shape value (four rate categories) of 1.0 using all statistical criterion: AIC, cAIC and BIC (-lnL= -9559.80). BI and ML methods generate phylogenies with the same topology, depicted in the cladogram of Supplementary Material Figure S7.

SwissModel identified in *Caenorabditis elegans* SOD1 (PDB code 3kbfA, 1.20 Å resolution, X-ray method) a 3D structure with a sequence that is 53.13% identical to *T. thermophila* SOD1a (Supplementary Material Fig. S8A). The resulting modeling confirmed the presence of the ligands for Zn. In contrast, *Tetrahymena* SOD1b (Supplementary Material Fig. S8B) appeared more similar to the tomato (*Solanum lycopersicum*) chloroplastic SOD1 (PDB code 3pu7A, 1.80 Å resolution, X-ray method), with a 43.87% level of sequence identity. Also in this case, the modeling confirmed the presence of the ligands for Zn. Finally, the homology modeling analysis identified in *Bombyx mori* SOD1 (PDB code 3l9yB, 1.80 Å resolution, X-ray method) a 3D structure with a sequence that is 53.21% identical to *Tetrahymena* SOD1x (Supplementary Material Fig. S8C). The resulting model confirmed the presence of the ligands for Cu and Zn.

Time Course of Cu,Zn SOD Gene Expression in *T. thermophila* Cells under Physiological Conditions

In Figure 4 are shown the time course of Cu,Zn SOD mRNA accumulation in *Tetrahymena* cells grown under physiological conditions. There is a statistically significant decrease of SOD1a mRNA levels after 1h of grow ($p<0.05$; Fig. 4A); thereafter the levels of mRNA remain more or less constant. The situation is different for SOD1b, where there is a statistically significant increase of mRNA levels with a maximum at 4h from the beginning of the experiment ($p<0.05$; Fig. 4B). Afterward the SOD1b expression decreases and it turned back to the initial level after 48h.

In control cells the SOD activity remained almost constant for the entire time of the experiment (the average is 3.5 U SOD/mg proteins; Fig. 5).

Metal Accumulation, Cu,Zn SOD Gene Expression and SOD Activity Determination in *T. thermophila* Cu-treated Cells

In order to estimate the rate of Cu uptake from *Tetrahymena* cells, first we assessed Cu concentrations in control cells following the protocol described in Material and Methods. The physiological amount of Cu in cells unexposed to the metal is only 0.01-1.21 μg/mg of total proteins. On the contrary, during Cu-exposure a high rate of uptake in *Tetrahymena* was observed (Fig. 6). In cells harvested after 0.5, 1, 2, 4, 6, 24 and 48h of exposure, the measured Cu concentrations in cell-free extracts were significantly higher than in controls ($p<0.01$ and $p<0.001$): the Cu content increased rapidly, reaching a maximum of 3.11 μg/mg of total proteins at 24h ($p<0.05$). Then, the metal concentration decreased ($p<0.05$) to 1.60 μg/mg of total proteins, after 48h of exposure.

In Figure 4 the time course of Cu,Zn SOD mRNA accumulation in *T. thermophila* cells in presence of Cu (500 μM) is also reported. As shown in Fig. 4A, *tt-sod1a* mRNA levels increased rapidly and it was statistically higher than controls after 1h of Cu-exposure ($p<0.01$). Afterwards, the mRNA returned at the control level, before rising again at 4h ($p<0.001$). At 24h from the beginning of the exposure, the mRNA accumulation was temporarily lower than controls ($p<0.01$). Similarly, *tt-sod1b* mRNA levels increased over the controls after 1h of Cu-exposure ($p<0.01$; Fig. 4B). At 2h a greater increase in mRNA levels seems to be present, but the difference with the control cells is not statistically significant, due to the high variability. Later, the mRNA returned at the control level, with a single exception registered at 24h from the beginning of the exposure, when the mRNA accumulation was lower than controls ($p<0.05$).

The SOD activity after inoculation of *Tetrahymena* cells in Cu-containing PPYG medium was also estimated, following the procedure described in Material and Methods. As shown in Figure 5, the SOD activity in Cu-treated cells was always significantly higher with respect to controls ($p<0.05$ and $p<0.01$). It increased quickly, reaching the maximum value at 1h ($p<0.05$), then it returned to the

starting level, and later it remained constant until 48h.

Discussion

In this paper we report the characterization of the Cu,Zn SOD family in *Tetrahymena thermophila*. In this organism, three genes potentially encoding Cu,Zn SOD have been identified. The presence of three genes that potentially encode cytosolic SOD is unique, in unicellular eukaryotes as well as in metazoans (Miller 2012). However, our results suggest that the functionality of one of these three genes, the *tt-sod1x,* could be limited, since its expression levels are relatively low, as also reported in the Tetrahymena Functional Genome Database (http://tfgd.ihb.ac.cn/). This characteristic is typical of gene families showing a redundancy due to gene duplication phenomena in their evolution in eukaryotic genomes, and in mammals particularly (Chen et al. 2002; Gonçalves et al. 2000). For example, in the human genome thirteen metallothionein (MT) genes are present, but only ten of them were shown to be functional (Cherian et al. 2003). Also in *T. thermophila,* duplication processes have driven the evolution of several proteins, including MTs (Boldrin et al. 2003, 2006; Diaz et al. 2007; Piccinni et al. 1999; Santovito et al. 2001, 2007; Shang et al. 2002). Regarding Cu,Zn SODs, two pseudogenes are known to be present on human chromosomes 8 and 16, in addition to the functional human SOD1 gene that is located on chromosome 21 (Mossallam 2007). Although *tt-sod1x* mRNA levels are low, it is possible that this gene is activated under particular stress conditions. For example, preliminary results indicated that in the ciliated protozoan *Euplotes rariseta* only Mn SOD and Fe SOD genes are transcribed at basal level, but in conditions favoring ROS formation, such as high environmental oxygen concentration, Cu,Zn SOD are also transcribed (personal data, unpublished). This hypothesis is also consistent with the phylogenetic relationships of *T. thermophila* Cu,Zn SODs. Our analysis indicated that *tt-sod1x* emerges together with the Cu,Zn SODs of *Paramecium tetraurelia.* The genome of this protozoan contains two Cu,Zn SOD genes, one of them showing high similarity to *T. thermophila sod1b*. The presence of two separated clusters of Cu,Zn SOD from closely related

species can be explained with an evolutionary process characterized by a duplication event that involved the ancestor gene, which probably occurred in distant time, before the speciation event that led to these two species.

From our data it is impossible to deduce if SOD1a resulted from another duplication event involving the ancestor of SOD1b or SOD1x. This unclear correlation suggests an alternative hypothesis of adaptive convergence that could lead the occurrence of structurally and functionally similar proteins from different ancestor proteins. If it is the case, this process could have happened many times during speciation. This is probably the case of Mn SOD occurrence. In fact, phylogenetic studies placed Mn SODs far from Cu,Zn SODs (Santovito et al. 2006). Preliminary results indicated that in ciliated protozoa the evolutionary relationships between SOD1s and SOD2s are similar to those of other eukaryotes. An interesting result from homology modeling analysis is the similarity between *Tetrahymena* SOD1b and *Solanum lycopersicum* chloroplastic SOD1. In plants there are two different groups of Cu,Zn SODs. The first group consists of cytoplasmic and periplasmic forms, which are homodimeric. The second group comprises the chloroplastic and extracellular Cu,Zn SODs, which are homotetrameric (Bordo et al. 1994). These enzymes are all encoded by different genes. Even animal cells export a tetrameric glycosylated Cu,Zn SOD that appears to be more closely related to fungal Cu,Zn SODs, raising the possibility that the tetrameric version is the more ancient one (Fink and Scandalios 2002). All these data suggest an independent evolution for *Tetrahymena* SOD1a and SOD2a, and the latter could have a common ancestor with the tetrameric SODs.

However, it would be interesting to also investigate further the *Tetrahymena* SOD1x, from both evolutionary and functional points of view. In fact, our analysis of homology modeling demonstrated structural similarities of SOD1x with the Cu,Zn SOD from *Bombyx mori*. Recent studies revealed that this silkworm enzyme shares 63% sequence identity with human SOD1, with some different residues at mutation sites in Cu,Zn SOD associated to the amyotrophic lateral sclerosis condition (Zhang et al.

2010).

The two expressed Cu,Zn SOD genes code for two proteins that are bigger than the typical IC-SODs (154 aa). The acquisition of proteins with long primary sequences is a characteristic of the molecular evolution in *Tetrahymena* species (Boldrin et al. 2006) and could explain the high resistance grade of these organisms to environmental stress conditions. In fact, an abnormally long sequence characterizes the Cu,Zn SODs of animals living in permanent stress conditions, such as the icefish. These organisms evolved in oceanic waters surrounding the Antarctica, exposed to high oxygen concentration as a consequence of low temperature, which increases gas solubility. Furthermore, the inability of icefish to synthesize hemoglobin implies an increased quantity of oxygen simply dissolved in their tissues. These two conditions increase the probability of ROS formation. Recent analyses demonstrated that Cu,Zn SOD of these fish is longer than orthologs from phylogenetically related fish (personal data, unpublished). This hypothesis is supported by the evidence that the SOD1a from the psychrophilic Antarctic ciliate *E. focardii* is also longer than those from other species, *Oxytricha* included.

In the case of icefish the evolutionary mechanism resulting in long proteins seems to be the mutation of stop codon (TAA → ATG). In *T. thermophila* and *E. focardii* the mechanism probably also involved the stop codons, but this is a peculiar feature of the modified genetic codes of ciliates, where TAA and TAG are not stop codons but code for Gln in *Tetrahymena*, and TGA encodes Cys in *Euplotes*.

Despite the different protein dimensions, the residues important for catalytic activity are highly conserved. One exception is His^{154} in the isoform b that is replaced by Asn through a point mutation that involved the first codon nucleotide (c → a). However, the copper coordination could be maintained by His^{165}, which could vicariate the absent amino acid. It is of note that this histidine is also present, at the same position, in *E. focardii* SOD1b (Fig. 2). The only other exception, also in SOD1b, is the

mutation of Thr^{171}, which is close to the active site channel and, together with the positively charged guanidinium group of Arg^{177} (strictly conserved), guides the superoxide anion toward the catalytic Cu^{2+} coordination site, stabilizing its bonding. In this isoform it is substituted by Asn by a point mutation that involved the second codon nucleotide (c → a). In this case, Thr^{170} could vicariate the absent amino acid. The other amino acids are fully conserved in both enzymes, such as the coordination sphere of the zinc ion that is formed by three histidine and one aspartic acid residues, assuring the enzyme functionality despite its lower level of sequence identity with other Cu,Zn SOD sequences.

An interesting result is the different pattern of mRNA accumulation observed between *tt-sod1a* and *tt-sod1b*. In particular, the latter appears to be characterized by increasing mRNA levels during the exponential phase of growth, this parameter decreasing after prolonged starvation (24 and 48 h time points). This result suggests that, although the cells are not exposed to an environmental stress factor, an endogenous stress condition could be present. During the exponential growth of the culture, the cells are in active metabolism and the rate of oxygen consumption is high. In this condition, ROS formation can increase (Acworth et al. 1997) leading to the induction of genes involved in the antioxidant defense, such as superoxide dismutases but also catalase and glutathione peroxidase, enzymes that are functionally related to SODs because they eliminate H_2O_2. Between the two Cu,Zn SODs, isoform b seems to be the one that, during evolution, assumed major importance in countering this kind of stress.

The exposure of *Tetrahymena* to a pro-oxidant, such as copper, only slightly increases Cu,Zn SOD mRNA levels. This result can be related to the absence of regulatory sequences, such as Metal Responsive Elements (MRE) and ARE, in the promoter regions, or to the low copper accumulation in the cells, or to the low ROS formation rate during the metal exposure.

It is well known that the presence of MREs and AREs in the promoter regions is a characteristic of both eukaryotic SOD as well as MT genes (Kensler et al. 2007; Yoo et al. 1999). Nevertheless, it has been shown that *Tetrahymena* MT gene transcription induced by cadmium is driven by promoter

sequences which are different from the typical MREs (Boldrin et al. 2003, 2008; Formigari et al. 2010; Piccinni et al. 1999). Therefore, it is possible that Cu-dependent induction is present, but specific metal response elements in *Tetrahymena* have not yet been identified. Instead, in the promoter regions we have found several putative h-ARE sequences, which can mediate the induction of SODs by ROS.

After Cu-treatment, the cytoplasmic concentration of this metal significantly increased until 24 h of exposure, afterwards the Cu levels slightly decreased, probably as the result of metal detoxification process, which is dependent on the biosynthesis of glutathione and MTs. Previous studies demonstrate a linear correlation between copper and MT concentrations in the cell-free extract of *Tetrahymena* experimentally exposed to this metal ion (Boldrin et al. 2002, 2008; Santovito et al. 2001). This homeostatic response can limit the presence of copper in the cells, which is not completely removed, but it remains higher than controls even after 48h. Therefore, copper continuously enters in the cells, but it is not bio-available to interact with gene regulation systems or to catalyze the production of ROS, because almost all Cu ions are bound by MTs and GSH and continuously eliminated via lysosome system. This could justify the poor induction of SOD gene activation.

Nevertheless, SOD activity in Cu-treated cells was found to be increased in comparison to controls, and this was especially evident in the first phases of exposure (0.5-1 h). The absence of a correlation between mRNA levels and proteins activity is difficult to interpret, but it is not a rare phenomenon. We can formulate two hypotheses to explain the increase of SOD activity without increase of mRNA concentrations: i) the presence of high Cu-concentrations at cytoplasmic level can be conveyed to SOD apoproteins already present in the cell; ii) copper increases the translation of Cu,Zn SOD mRNA already present in the cell. In fact mRNAs have a complex lifecycle and non-translating mRNAs, under stress conditions, can be stored in cytoplasmic granules (Olszewska et al. 2012), processing P-bodies, and stress granules for decay or future translation, respectively .(Lavut and Raveh 2012). The first hypothesis could be more plausible because copper is a fundamental component

of the Cu,Zn SOD-active site and it has been supposed that post-transcriptional regulation of Cu,Zn SOD induction is an important mechanism for oxidative stress tolerance in plants (Sunkar et al. 2006). However, we cannot exclude that the little but statistically significant increase in *tt-sod1a* and *tt-sod1b* mRNA accumulation measured at 1 h, is related also to an increase of ROS production. This hypothesis is supported by preliminary results obtained by ROS quantification experiments in *T. thermophila*, where an intense fluorescence can be observed after copper exposure, probably related to an increase of $\bullet O_2^-$ formation induced by this metal (Fig. 7).

So far, protozoa seemed to lack the Cu,Zn SOD, a unique feature that would distinguish them from plants, animals and fungi. The hypothesis was that the ancestor of protozoa could have lost these genes during evolution (Miller 2012). In this work we demonstrated that transcriptionally active genes encoding Cu,Zn SODs are also in protozoa and that the process of gene expression is active. To better explain the physiological role of Cu,Zn SODs in the antioxidant defense system of *Tetrahymena*, studies on genes coding for other antioxidant enzymes will be necessary, with the aim to reconstruct at both physiological and evolutionary levels, the complex metabolic pathways involved in the cell homeostasis of ROS.

Methods

Strains and culture conditions: *T. thermophila* wild-type strain SB210 was grown routinely in PPYG medium (0,4% proteose peptone, 1% glucose, 0.2% yeast extract) (Mori et al. 2011) axenically with shaking at 30 °C. Treated cells were grown in the same medium supplemented with $CuSO_4$ at 500 μM (the maximum concentration that does not affect the growth rate of the cells, verified by growth curves, Supplementary Material Fig. S9) for six times: 0.5, 1, 2, 4, 24 and 48h. Three independent experiments were performed for each treatment. Cultures were started by inoculating the media with exponentially growing cells (cell concentration approximately 200,000 cells/ml) to an initial density of 20,000 cells/ml.

Cell cultures of the *E. focardii* strain TN1 (Valbonesi and Luporini 1990) were used. They were isolated from sediment and sea water samples collected from the coastal waters of Terra Nova Bay (temperature, –1.8 °C; salinity, 35%; pH, 8.1-8.2) and were grown in a cold room at 4 °C, using the green alga *Dunaliella tertiolecta*, or the bacterium *Escherichia coli*, as food. *Dunaliella tertiolecta* was grown at 25 ± 1°C in sea water supplemented with Walne's medium and vitamin B12 (Gopinathan 1982). The cultures were continuously bubbled with sterile air and continuous light was provided by banks of 18W Philips Master TL-D 90 De Luxe fluorescent tubes.

Quantification of Cu and SOD activity in cell-free extracts: Cells were harvested by centrifugation at 2000×g for 15 min at the indicated times, then they were homogenized as previously described (Piccinni et al. 1990). Homogenates were centrifuged at 48,000 ×g for 60 min at 4 °C, and the resulting supernatants were used for metal and SOD activity quantification. Analyses of metal contents were performed by atomic absorption spectrophotometer (Perkin-Elmer mod. 4000). SOD activity was assayed by the method described by Beauchamp and Fridovich (1971), which is based on the inhibition

of the reduction of nitro blue tetrazolium (NBT) by $\bullet O_2^-$ produced via riboflavin photo reduction. One unit of SOD activity (U SOD) was defined as the amount of enzyme required for 50% inhibition of NBT conversion. Data refer to total protein concentrations, assayed by the Folin phenol reagent method (Lowry et al. 1951).

RNA isolation from *T. thermophila*, cDNA synthesis, RT-PCR and cDNA cloning: Total RNA was isolated from mid-log cells (3–5 × 10^6 cells/ mL) by TRIzol reagent (Invitrogen) according to the manufacturer's protocol. Total RNA was purified with LiCl 8 M in order to remove glucidic contaminants and the quantification was performed using the ND-1000 spectrophotometer (Nanodrop, Wilmington, DE); RNA integrity was assessed by capillary electrophoresis using the Agilent Bioanalyzer 2100, with the RNA 6000 Nano (Agilent Technologies, Palo Alto, CA). The first strand of cDNA was reverse-transcribed at 42 °C for 1 h from 1 μg of total RNA in a 20 μL reaction mixture containing 1 μL of ImProm II Reverse Transcriptase (Promega) and 0.5 μg oligodT Anchor primer. The coding region of SODs from *T. thermophila* was amplified with primers reported in the Table S2, which were designed on the basis of the coding sequence obtained with a BLAST search in the *Tetrahymena* genome database (http://ciliate.org/index.php/home/welcome). PCR reactions were performed with 50 ng of cDNA. The PCR program was the following: 94 °C for 2 min and 35 ×(94 °C for 30 s, Tm for 30 s, 72°C for 1 min); final elongation 72 °C for 10 min.

All the amplicons were gel-purified with the NucleoSpin Extract 2 in 1 (Macherey-Nagel), ligated into the pGEM®-T Easy Vector (Promega) and cloned in XL1-Blue *E. coli* cells (Tang et al. 1994). Positively screened clones were sequenced at the BMR Genomics (University of Padova) on an ABI PRISM 3700 DNA Analyzer (Applied Biosystems).

DNA purification from *E. focardii*, Rapid Amplification of Telomeric Ends (RATE) and subcloning: Macronuclear DNA from *E. focardii* was purified as previously described (Pucciarelli and Miceli 2002). RATE-PCR is a strategy particularly useful to amplify the macronuclear DNA molecules

(minichromosomes) of *Euplotes* (Marziale et al. 2008) which are small, mostly gene-sized and always flanked by telomeres characterized by unique sequence repeats (Hoffman et al. 1995). To obtain the entire macronuclear sequence of *E. focardii* SOD, DNA samples (0.5 μg) were first amplified through 10 cycles of two independent and subsequent PCR reactions. One was run with a forward primer represented by the oligonucleotide reported in Table S2, corresponding to the coding region of *E. focardii* SODs. The second one was run with a reverse primer represented by the oligonucleotide reported in Supplementary Material Table S2, which corresponds to the *E. focardii* SOD coding sequence. These amplifications were then completed through 20 additional cycles carried out in the presence of the oligonucleotide 5'-$(C_4A_4)_4$-3', containing four repetitions of the motif C_4A_4, that is distinctive of the *E. focardii* telomere sequence, and hence was capable of acting as reverse primer in the former reaction and as forward primer in the latter. The PCR products were then subcloned using the InsTAclone® PCR Cloning kit (Fermentas).

Quantitative real-time PCR analysis: Quantitative real-time PCR analysis was performed in a final volume of 10 μL, with 1 to 10 ng of cDNA, 1× SYBR Green PCR Master mix (Promega), 0.5 μM for each primer (Supplementary Material Table S2) or 0.5 μM for each ribosomal 17S primer (17S r-RNA sense, 17S r-RNA antisense) (Supplementary Material Table S2). To avoid the amplification of contaminant genomic DNA, total RNA samples were treated with DNAse I (Qiagen). A dissociation curve was used to confirm the specificity of the amplicon. We verified the efficiency of the primers by drawing standard curves for target genes (SOD1a, efficiency 99%, $R^2 = 0.99$; SOD1b, efficiency 98%, $R^2 = 0.99$) and 17S rRNA endogenous control (efficiency 100%, $R^2 = 0.99$). PCR reactions were performed in triplicate in a 7500 Real- Time PCR System (Applied Biosystems). Thermal cycling conditions were as follows: 2 min denaturation at 95°C followed by 38 cycles for 25 sec denaturation at 95 °C, 1 min annealing and elongation at 60 °C, and a final 3 min elongation at 72 °C. The 2-ddCt (RQ, relative quantification) method implemented in the 7500 Real Time PCR System software was

used to calculate the relative expression ratio (Livak and Schmittgen 2001). This method defines the change in expression of a nucleic acid sequence (target) in test samples relative to the same sequence in a calibrator sample, when the target gene is expressed at its lowest levels. 95% confidence intervals are associated to each time point.

Sequence alignment and phylogenetic reconstructions: The sequences used in this paper are summarized in Supplementary Material Table S3 along with their GenBank accession numbers. The 2000 nt upstream to coding sequences were analyzed with the Primer Premier 5.00 software package (Premier Biosoft International, Palo Alto, CA) to identify putative transcription factor binding sites such as Antioxidant Responsive Element (ARE) and Metal Responsive Element (MRE). Full-length of the protein sequences were aligned using the ClustalW program (Thompson et al. 1994). Later, the alignments were improved manually.

Various isoforms of SOD amino acid and coding cDNA sequences of *Tetrahymena thermophila*, *Paramecium tetraurelia*, *Cryptocaryon irritans*, *Ichthyophthirius multifiliis* and *Perkinsus marinus*, were found in GenBank (www.ncbi.nlm.nih.gov/GenBank/). *Euplotes focardii* Cu,Zn SOD sequences were obtained from the sequencing of the entire trasncriptome (Pucciarelli et al., unpubl. observ.) and *Oxytricha trifallax* Cu,Zn SOD sequences were obtained from the web site (http://oxy.ciliate.org/index.php/home/welcome). All Cu,Zn SOD sequences of were aligned using T-Coffee multiple sequence alignment software package (Notredame et al. 2000).

The jModelTest (Darriba et al. 2012) was used to carry out statistical selection of best-fit models of nucleotide substitution. Analyses were performed using 88 candidate models and three types of information criterion (Akaike Information Criterion - AIC, Corrected Akaike Information Criterion - cAIC and Bayesian Information Criterion - BIC). To select the best-fit model of analyzed protein evolution ProtTest 3 was used (Darriba et al. 2011). 122 candidate models and three types of criterion (AIC, cAIC and BIC) were used in these statistical analyses. Furthermore, in order to find the best-fit

model (clock or no clock models) to analyzed Cu,Zn SOD evolution, we used an accurate assessment of the marginal model likelihoods using the stepping-stone method. It estimates the model likelihood by sampling a series of distributions that represent different mixtures of the posterior distribution and the prior distribution (Xie et al. 2011). The stepping-stone method was applied to the Cu,Zn SOD cDNA dataset using 510,000 generations with a diagnostic frequency of 2,500 in 2 independent runs for each of the tested models.

The Cu,Zn SOD cDNA and amino acid sequences phylogenetic trees were build using the Bayesian inference (BI) method implemented in Mr. Bayes 3.2 (Ronquist et al. 2012). Four independent runs, each one with four simultaneous Markov Chain Monte Carlo (MCMC) chains, were performed for 1,000,000 generations sampled every 1000 generations. Furthermore, we also used the maximum likelihood (ML) method implemented in PhyML 3.0 (Guindon et al. 2010). Bootstrap analyses were performed on 100,000 trees using both kinds of tree topology improvement: nearest neighbor interchange (NNI) and subtree pruning and regrafting (SPR). FigTree v1.3 software was used to display the annotated phylogenetic trees.

Molecular modeling: Considering that some IC-SODs have been structurally characterized via X-ray crystallography, we employed a comparative modeling approaches to study *Tetrahymena thermophila* SOD1a, SOD1b and SOD1x folding (for which no experimental structure is currently available). In this work preliminary homology protein modeling was carried out with SwissModel bioinformatic tools (http://swissmodel.expasy.org) (Arnold et al. 2006).

Statistical analyses: All data were expressed as average of three independent experiments ± standard deviation (SD). Statistical analyses were performed with the PRIMER statistical program. One way ANOVA was followed by the Student–Newman–Keuls test to assess significant differences ($p < 0.05$).

ROS detection by fluorescence microscopy: To evaluate the ROS formation induced by copper

exposures in *T. thermophila*, we have used a common fluorescence probe, hydroethidine (dihydroethidium, HE), considered as the most adequate to detect superoxide anion ($\bullet O_2^-$) in biological systems, especially in mammalian cells (Gomes et al. 2005) but also in ciliated protozoa (Rico et al. 2009). HE can penetrate through biological membranes and be directly oxidized by intracellular $\bullet O_2^-$, originating the fluorescent product ethidium (λ excitation = 500–530 nm; λ emission = 590–620) (Benov et al. 1998). Ethidium is also retained in the nucleus due to its DNA intercalation property, which increases fluorescence emission (Walrand et al. 2003).

Acknowledgements

This work was supported by Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), University of Padova Grant PRAT2010-CPDA102125, and by the Italian National Program for Antarctic Research (PNRA).

References

Acworth IN, McCabe DR, Maher TJ (1997) The Analysis of Free Radicals, their Reaction Products, and Antioxidants. In Baskin SI, Salem H (eds) Oxidants, Antioxidants, and Free Radicals. Taylor and Francis, Washington DC, pp 23–77

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics **22**:195–201

Beauchamp C, Fridovich I (1971) Superoxide dismutase: improved assays and an assay applicable to acrylamide gels. Anal Biochem **44**:276–287

Boldrin F, Santovito G, Irato P, Piccinni E (2002) Metal interaction and regulation of *Tetrahymena pigmentosa* metallothionein genes. Protist **153**:283–291

Boldrin F, Santovito G, Negrisolo E, Piccinni E (2003) Cloning and sequencing of four new metallothionein genes from *Tetrahymena thermophila* and *T. pigmentosa*: evolutionary relationships in *Tetrahymena* MT family. Protist **154**:431–442

Boldrin F, Santovito G, Formigari A, Bisharyan Y, Cassidy-Hanley D, Clark TG, Piccinni E (2008) MTT2, a copper-inducible metallothionein gene from *Tetrahymena thermophila*. Comp Biochem Physiol C **147**:232–240

Boldrin F, Santovito G, Gaertig J, Wloga D, Cassidy-Hanley D, Clark TG, Piccinni E (2006) Metallothionein gene from *Tetrahymena thermophila* with a copper-inducible-repressible promoter. Eukaryot Cell **5**:422–425

Chen C, Gentles AJ, Jurka J, Karlin S (2002) Genes, pseudogenes and Alu sequence organization across human chromosomes 21 and 22. Proc Natl Acad Sci USA **99**:2930–2935

Cherian MG, Jayasurya A, Bay BH (2003) Metallothioneins in human tumors and potential roles in carcinogenesis. Mutat Res **533**:201–209

Darriba D., Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics **27**:1164–1165

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods **9**:772

Diaz S, Amaro F, Rico D, Campos V, Benitez L, Martin-Gonzales A, Hamilton EP, Orias E, Gutierrez JC (2007) *Tetrahymena* metallothioneins fall into two discrete subfamilies. PLoS ONE **2**:e291

Esterbauer H, Gebicki J, Puhl H, Jürgens G (1992) The role of lipid peroxidation and antioxidants in oxidative modification of LDL. Free Radic Biol Med **13**:341–390

Fink RC, Scandalios JG (2002) Molecular evolution and structure-function relationships of the superoxide dismutase gene families in angiosperms and their relationship to other eukaryotic and prokaryotic superoxide dismutases. Arch Biochem Biophys **399**:19–36

Formigari A, Boldrin F, Santovito G, Cassidy-Hanley D, Clark TG, Piccinni E (2010) Functional characterization of the 50-upstream region of MTT5 metallothionein gene from *Tetrahymena thermophila*. Protist **161**:71–77

Fréalle E, Noël C, Nolard N, Symoens F, Felipe MS, Dei-Cas E, Camus D, Viscogliosi E, Delhaes L (2006) Manganese superoxide dismutase based phylogeny of pathogenic fungi. Mol Phylogenet Evol **41**:28–39

Fridovich I (1995) Superoxide radical and superoxide dismutases. Annu Rev Biochem **64**:97–112

Gonçalves I, Duret L, Mouchiroud D (2000) Nature and structure of human genes that generate retropseudogenes. Genome Res **10**:672–678

Gomes A, Fernandes E, Lima JLFC (2005) Fluorescence probes used for detection of reactive oxygen species. J Biochem Biophys Meths **65**:45–80

Gopinathan CP (1982) Methods of culturing phytoplankton CMFRI Special publication **8**:113–118

Harris ED (2000) Cellular copper transport and metabolism. Annu Rev Nutr **20**:291–310

Jackson SM, Cooper JB (1998) An analysis of structural similarity in the iron and manganese superoxide dismutases based on known structures and sequences. Biometals **11**:159–173

James ER (1994) Superoxide dismutase. Parasitol Today **10**:481–484

Kass RE, Raftery AE (1995) Bayes factors. J Amer Stat Assoc **90**:773–795

Keele BB Jr, McCord JM, Fridovich I (1971) Further characterization of bovine superoxide dismutase and its isolation from bovine heart. J Biol Chem **246**:2875–2880

Kensler TW, Wakabayashi N, Biswal S (2007) Cell survival responses to environmental stresses via the Keap1–Nrf2–ARE pathway. Annu Rev Pharmacol Toxicol **47**:89–116

Lah MS, Dixon MM, Pattridge KA, Stallings WC, Fee JA, Ludwig ML (1995) Structure-function in *Escherichia coli* iron superoxide dismutase: comparisons with the manganese enzyme from *Thermus thermophilus*. Biochemistry **34**:1646–1660

Lavut A, Raveh D (2012) Sequestration of highly expressed mRNAs in cytoplasmic granules, P-bodies, and stress granules enhances cell viability. PLoS Genet **8**:e1002527

Lee YM, Friedman DJ, Ayala FJ (1985) Superoxide dismutase: an evolutionary puzzle. Proc Natl Acad Sci USA **82**:824–828

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods **25**:402–408

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the folin phenol reagent. J Biol Chem **193**:265–275

Marklund SL (1982) Human copper-containing superoxide dismutase of high molecular weight. Proc Natl Acad Sci USA **79**:7634–7638

Marziale F, Pucciarelli S, Ballarini P, Melki R, Uzun A, Ilyin VA, Detrich HW 3rd, Miceli C (2008) Different roles of two gamma-tubulin isotypes in the cytoskeleton of the Antarctic ciliate

Euplotes focardii: remodelling of interaction surfaces may enhance microtubule nucleation at low temperature. FEBS J **275**:5367–5382

McCord JM, Fridovich I (1969) Superoxide dismutase: an enzymic function for erythrocuprein (Hemocuprein). J Biol Chem **244**:6049–6055

McCord JM, Fridovich I (1988) Superoxide dismutase: the first twenty years (1968-1988). Free Radic Biol Med **5**:363–369

Miller AF (2012) Superoxide dismutases: ancient enzymes and new insights. FEBS Lett **586**:585–595

Mori K, Kashiwagi A, Yomo T J (2011) Single-cell isolation and cloning of *Tetrahymena thermophila* cells with a fluorescence-activated cell sorter. J Eukaryot Microbiol **58**:37–42

Mossallam AAA (2007) Identification and assignment of human superoxide dismutase-1 retropseudogenes using comparative mapping and bioinformatic analysis. Biotechnology **6**:393–407

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol **302**:205–217

Olszewska M, Bujarski JJ, Kurpisz M (2012) P-bodies and their functions during mRNA cell cycle: mini-review. Cell Biochem Funct **30**:177–182

Parker MW, Blake CC, Barra D, Bossa F, Schinina ME, Bannister WH, Bannister JV (1987) Structural identity between the iron- and manganese-containing superoxide dismutases. Protein Eng **1**:393–397

Peña MM, Lee J, Thiele DJ (1999) A delicate balance: homeostatic control of copper uptake and distribution. J Nutr **129**:1251–1260

Piccinni E, Irato P, Guidolin L (1990) Cadmium-thionein in *Tetrahymena thermophila* and *Tetrahymena pyriformis*. Eur J Protistol **26**:176–181

Piccinni E, Bertaggia D, Santovito G, Miceli C, Kraev A (1999) Cadmium metallothionein gene of

Tetrahymena pyriformis. Gene **234**:51–59

Pucciarelli S, Miceli C (2002) Characterization of the cold-adapted alpha-tubulin from the psychrophilic ciliate *Euplotes focardii*. Extremophiles **6**:85–89

Pucciarelli S, Devaraj RR, Barchetta S, Yu T, La Terza A, Miceli C (2010) Characterization of new biomarkers from the Antarctic ciliate *Euplotes focardii* by a transcriptomic approach. Comp Biochem Physiol A **157**:S40–S41

Pucciarelli S, La Terza A, Ballarini P, Barchetta S, Yu T, Marziale F, Passini V, Methé B, Detrich HW 3rd, Miceli C (2009) Molecular cold-adaptation of protein function and gene regulation: the case for comparative genomic analyses in marine ciliated protozoa. Mar Genomics **2**:57–66

Rico D, Martín-González A, Díaz S, de Lucas P, Gutiérrez J-C (2009) Heavy metals generate reactive oxygen species in terrestrial and aquatic ciliated protozoa. Comp Biochem Physiol C **149**:90–96

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol **61**:539–542.

Santos R, Bocquet S, Puppo A, Touati D (1999) Characterization of an atypical superoxide dismutase from *Sinorhizobium meliloti*. J Bacteriol **181**:4509–4516

Santovito G, Cassini A, Piccinni E (2006) Cu,Zn superoxide dismutase from *Trematomus bernacchii*: functional conservation and erratic molecular evolution in Antarctic teleosts. Comp Biochem Physiol C **143**:444–454

Santovito G, Formigari A, Boldrin F, Piccinni E (2007) Molecular and functional evolution of *Tetrahymena* metallothioneins: new insights into the gene family of *Tetrahymena thermophila*. Comp Biochem Physiol C **144**:391–397

Santovito G, Irato P, Palermo S, Boldrin F, Sack R, Hunziker P, Piccinni E (2001) Identification, cloning and characterisation of a novel copper- metallothionein in *Tetrahymena pigmentosa*. Sequencing of cDNA and expression. Protist **152**:203–213

Shang Y, Song X, Bowen J, Corstanje R, Gao Y, Gaertig J, Gorovsky MA (2002) A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in *Tetrahymena thermophila*. Proc Natl Acad Sci USA **99**:3734–3739

Shaw G, Kamen R (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46**:659–667

Smith MW, Doolittle RF (1992) A comparison of evolutionary rates of the two major kinds of superoxide dismutase. J Mol Evol **34**:175–184

Sugio S, Hiraoka BY, Yamakura F (2000) Crystal structure of cambialistic superoxide dismutase from *Porphyromonas gingivalis*. Eur J Biochem **267**:3487–3495

Sunkar R, Kapoor A, Zhu JK (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. Plant Cell **18**:2051–2065

Swart EC, Bracht JR, Magrini V, Minx P, Chen X, Zhou Y, Khurana JS, Goldman AD, Nowacki M, Schotanus K, Jung S, Fulton RS, Ly A, McGrath S, Haub K, Wiggins JL, Storton D, Matese JC, Parsons L, Chang WJ, Bowen MS, Stover NA, Jones TA, Eddy SR, Herrick GA, Doak TG, Wilson RK, Mardis ER, Landweber LF (2013) The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. PLoS Biol **11**:e1001473.

Tainer JA, Getzoff ED, Richardson JS, Richardson DC (1983) Structure and mechanism of copper, zinc superoxide dismutase. Nature **306**:284–287

Tang X, Nakata Y, Li HO, Zhang M, Gao H, Fujita A, Sakatsume O, Ohta T, Yokoyama K

(1994) The optimization of preparations of competent cells for transformation of *E. coli*. Nucleic Acids Res **22**:2857–2858

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting; position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673–4680

Valbonesi A, Luporini P (1990) Description of two new species of *Euplotes* and *Eupiotes rariseta* from Antarctica. Polar Biol **11**:47–53

Valbonesi A, Luporini P (1993) Biology of *Euplotes focardii* an Antarctic ciliate. Polar Biol **13**:489–493

Walrand S, Valeix S, Rodriguez C, Ligot P, Chassagne J, Vasson M-P (2003) Flow cytometry study of polymorphonuclear neutrophil oxidative burst: a comparison of three fluorescent probes. Clin Chim Acta **331**:103–110.

Xie W, Lewis PO, Fan Y, Kuo L, Chen MH (2011) Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Sys Biol **60**:150–160

Yoo HY, Chang MS, Rho HM (1999) Heavy metal-mediated activation of the rat Cu/Zn superoxide dismutase gene via a metal-responsive element. Mol Gen Genet **262**:310–313

Youn HD, Kim EJ, Roe JH, Hah YC, Kang SO (1996) A novel nickel-containing superoxide dismutase from *Streptomyces* spp. Biochem J **318**:889–896

Zhang NN, He YX, Li WF, Teng YB, Yu J, Chen Y, Zhou CZ (2010) Crystal structures of holo and Cu-deficient Cu/Zn-SOD from the silkworm *Bombyx mori* and the implications in amyotrophic lateral sclerosis. Proteins **78**:1999–2004

Ziouzenkova O, Sevanian A, Abuja PM, Ramos P, Esterbauer H (1998) Copper can promote oxidation of LDL by markedly different mechanisms. Free Radic Biol Med **24**:607–623

Figure Legends

Figure 1. Arrangement of putative hAREs in upstream regions of *tt-sod1a* and *tt-sod1b* genes. hAREs carrying consensus core sequences and TATA box are indicated by filled arrows and an open box, respectively.

Figure 2. Multiple alignment of the amino acid sequences of Cu,Zn SODs from *T. thermophila* species as well as from *E. focardii*, *O. trifallax*, *Xenopus laevis, Bos taurus* and *Homo sapiens*. The shaded letters refer to the β-strand conformation of the bovine structure. Unshaded letters between β-strand regions refer to the 7 loops. The boxed letters refer to the amino acids that play an important role in the Cu^{2+}, Zn^{2+} coordination environments and active site region. Underlined and double-underlined letters refer to the amino acids that play an important role in the formation of the lower and upper rims, respectively. The symbols at the bottom of the SOD sequences correspond to the definitions of the CLUSTALW program: (*) fully conserved; (:) highly conserved; (.) conserved substitution.

Figure 3. Phylogenetic relationships among Cu,Zn SODs of various organisms reconstructed on the basis of cDNA sequences and using both BI (arithmetic mean = –10397.65; harmonic mean = –10419.72) and ML (arithmetic mean = –10336.9) methods. Posterior probability (first number) and bootstrap values (second number, if present) higher than 50% are indicated on each node. The scale for branch length (0.07 substitution/site) is shown below the tree. *T. thermophila* Cu,Zn SODs are boxed.

Figure 4. Expression levels of SOD1a (A) and SOD1b (B) in control and Cu-treatment conditions.

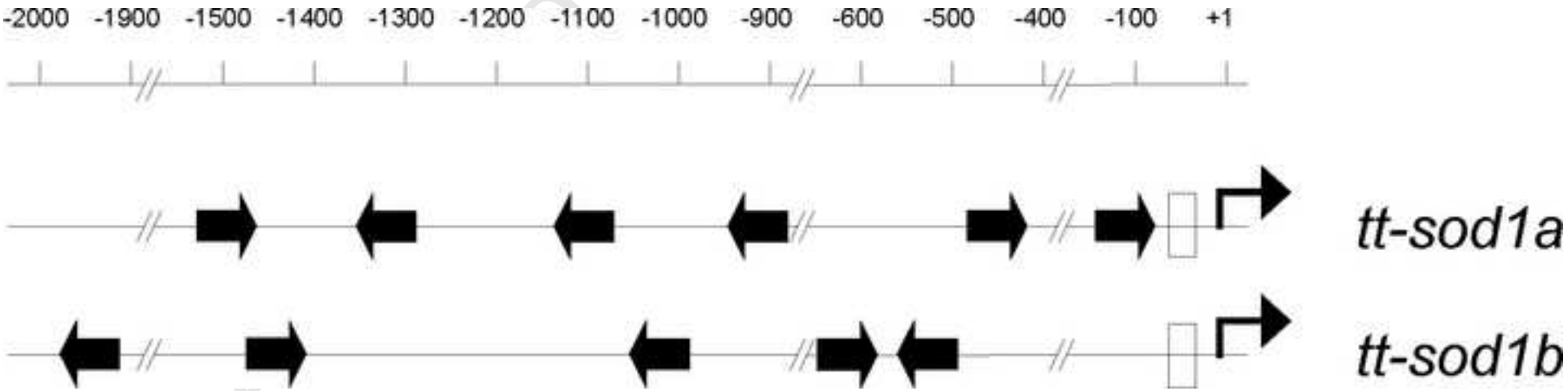
Tetrahymena cells harvested at 0.5, 1, 2, 4, 24, and 48h. The results are reported as mean of 3 independent experiments ± SD. Statistical differences between the two experimental groups are reported: * $p < 0.05$.

Figure 5. Metal accumulation in control and Cu-treatment conditions. *Tetrahymena* cells harvested at 0.5, 1, 2, 4, 24, and 48h. The results are reported as mean of 3 independent experiments ± SD. Statistical differences between the two experimental groups are reported: ** $p < 0.01$; *** $p < 0.001$.

Figure 6. SOD activity in control and Cu-treatment conditions. *Tetrahymena* cells harvested at 0.5, 1, 2, 4, 24, and 48h. The results are reported as mean of 3 independent experiments ± SD. Statistical differences between the two experimental groups are reported: * $p < 0.05$; ** $p < 0.01$.

Figure 7. Selected examples of fluorescent images obtained after exposure of *T. thermophila* living cells to dihydroethidium, in order to detect superoxide anions induced by heavy metals. (**A**) Control and (**B**) Cu (500 μM for 0.5h) treated cells.

Figure 1



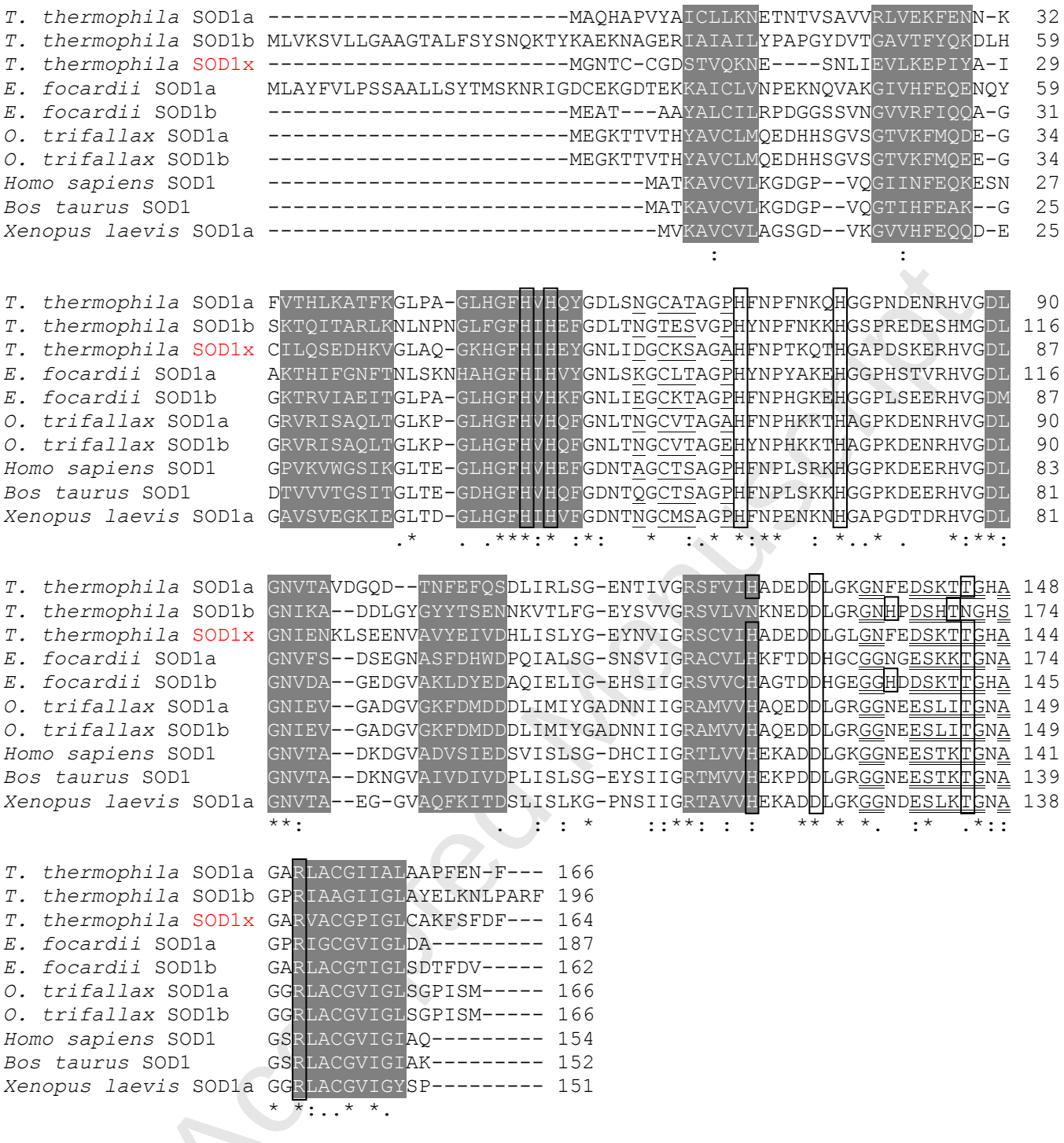
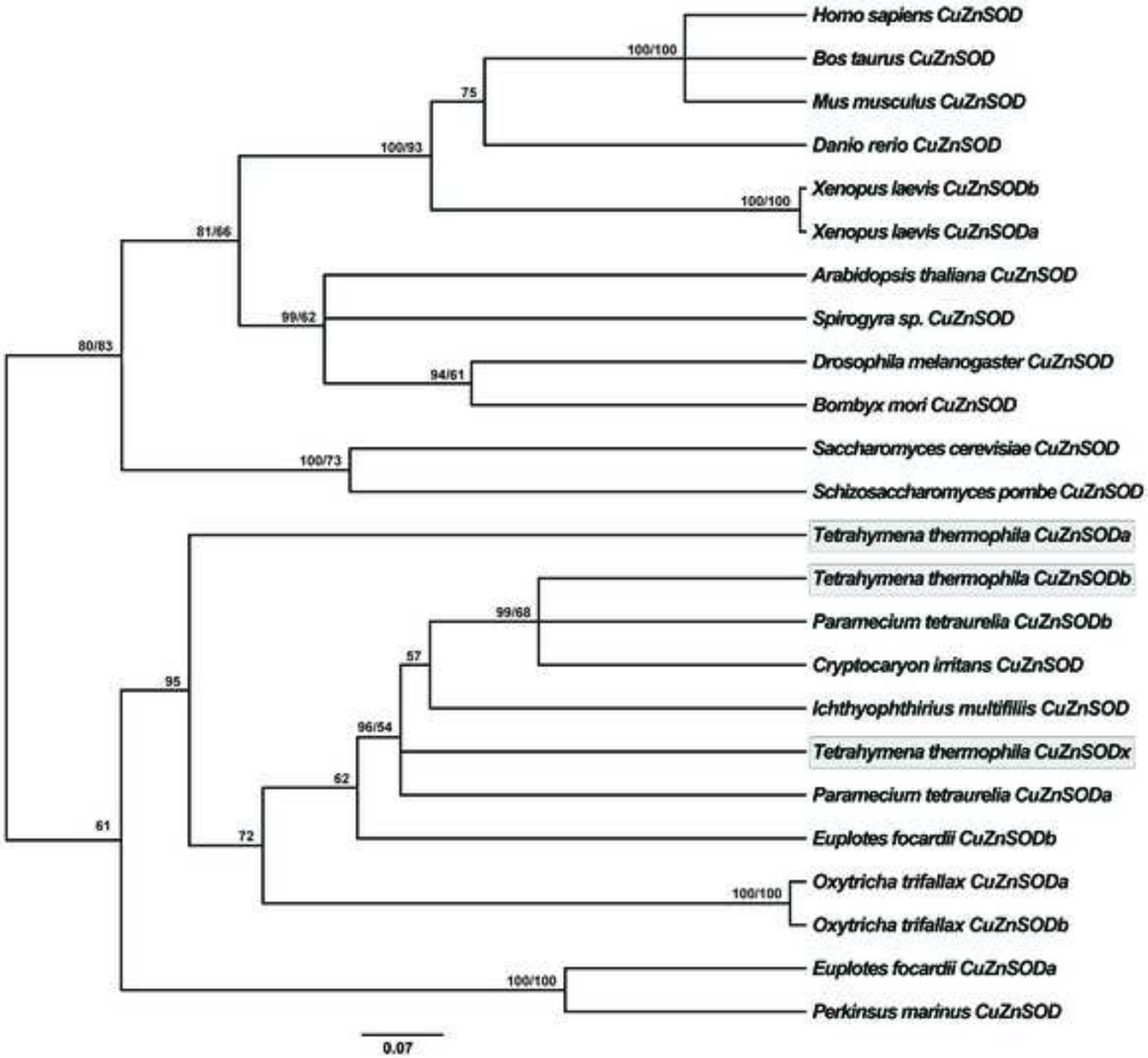
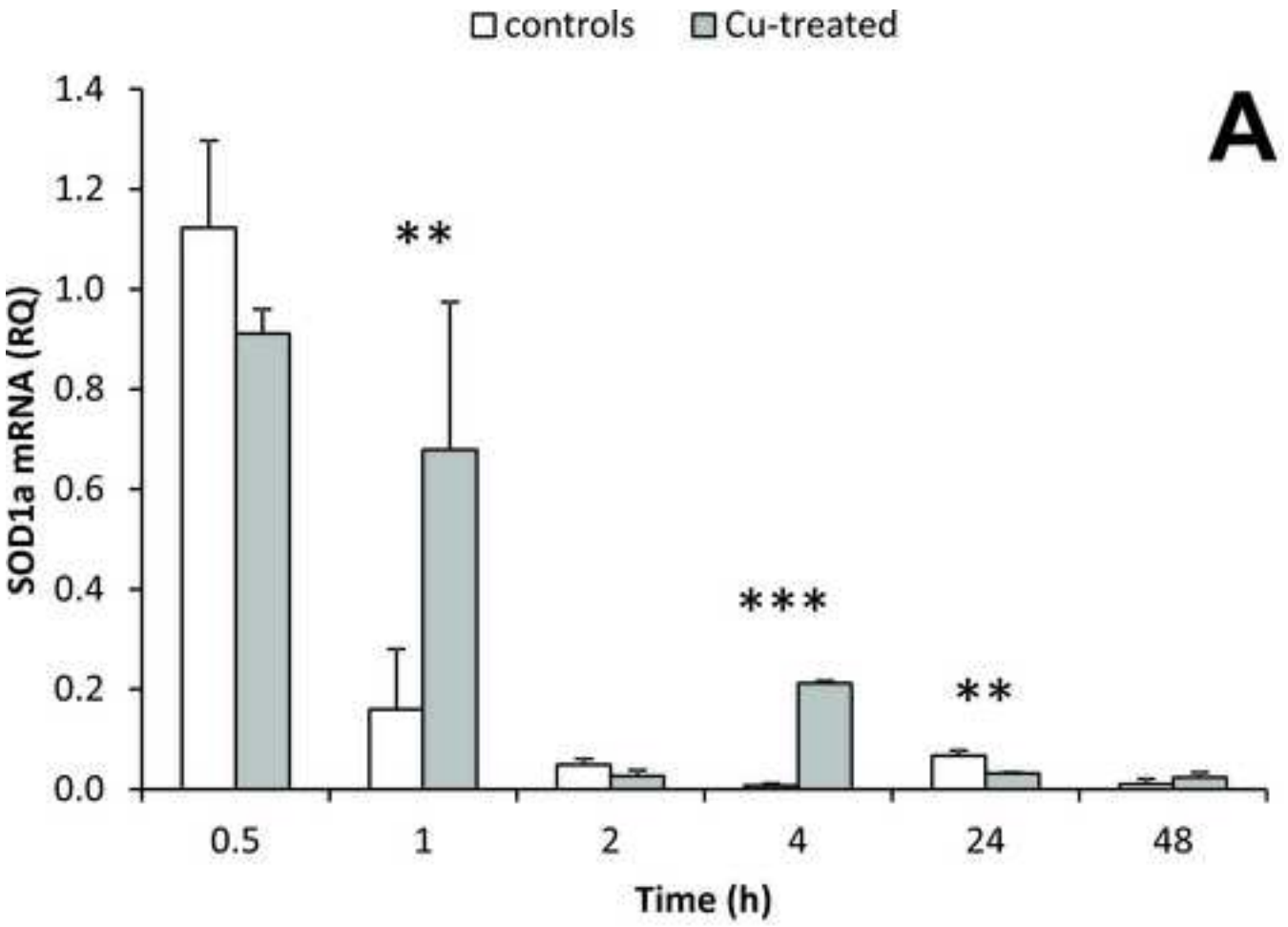


Figure 3


Homo sapiens CuZnSOD
Bos taurus CuZnSOD
Mus musculus CuZnSOD
Danio rerio CuZnSOD
Xenopus laevis CuZnSODb
Xenopus laevis CuZnSODa
Arabidopsis thaliana CuZnSOD
Spirogyra sp. CuZnSOD
Drosophila melanogaster CuZnSOD
Bombyx mori CuZnSOD
Saccharomyces cerevisiae CuZnSOD
Schizosaccharomyces pombe CuZnSOD
Tetrahymena thermophila CuZnSODa
Tetrahymena thermophila CuZnSODb
Paramecium tetraurelia CuZnSODb
Cryptocaryon irritans CuZnSOD
Ichthyophthirius multifiliis CuZnSOD
Tetrahymena thermophila CuZnSODx
Paramecium tetraurelia CuZnSODa
Euplotes focardii CuZnSODb
Oxytricha trifallax CuZnSODa
Oxytricha trifallax CuZnSODb
Euplotes focardii CuZnSODa
Perkinsus marinus CuZnSOD
100/100
75
100/93
100/100
81/66
99/62
80/83
94/61
100/73
99/68
57
95
96/54
62
61
72
100/100
100/100
0.07



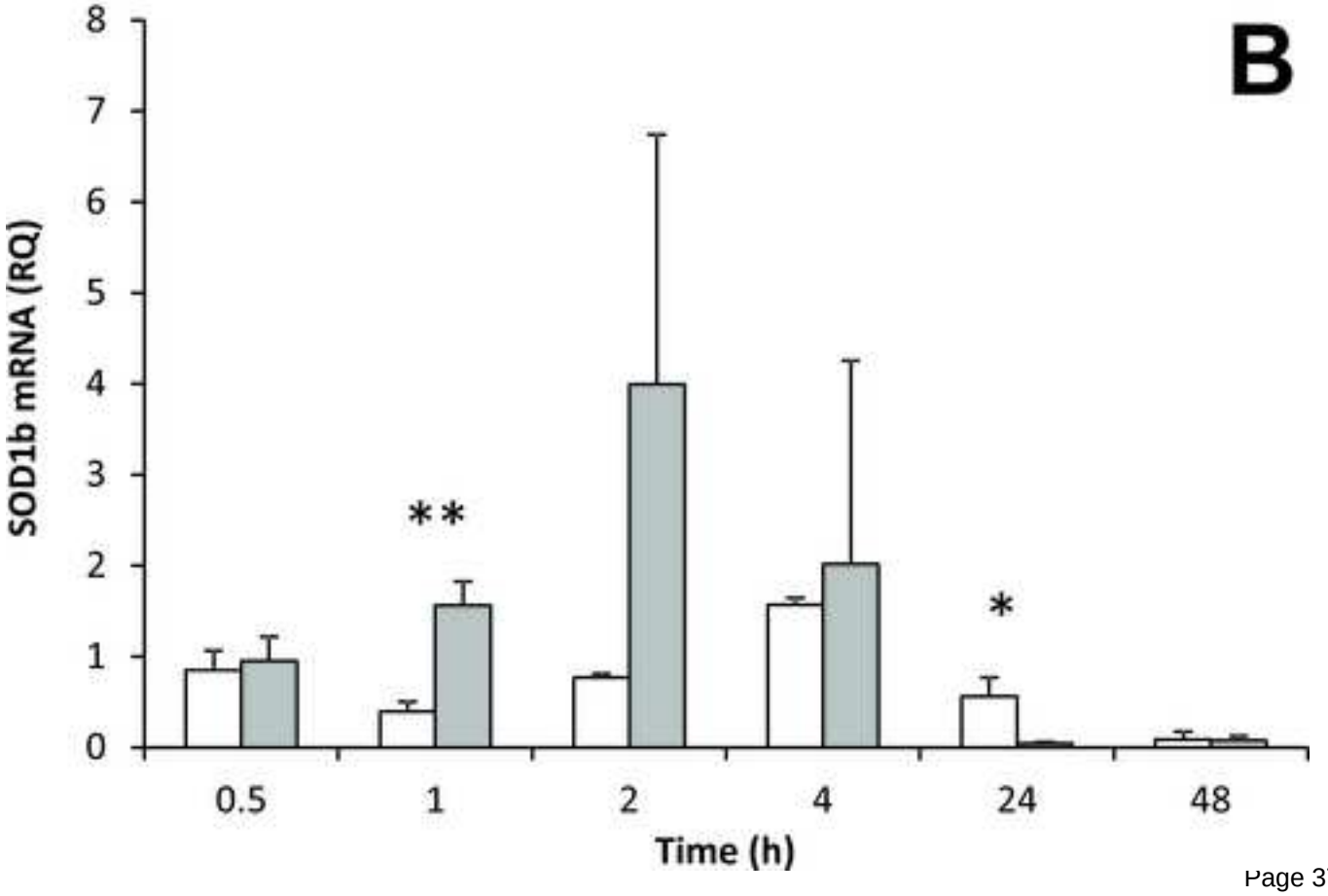


Figure 5

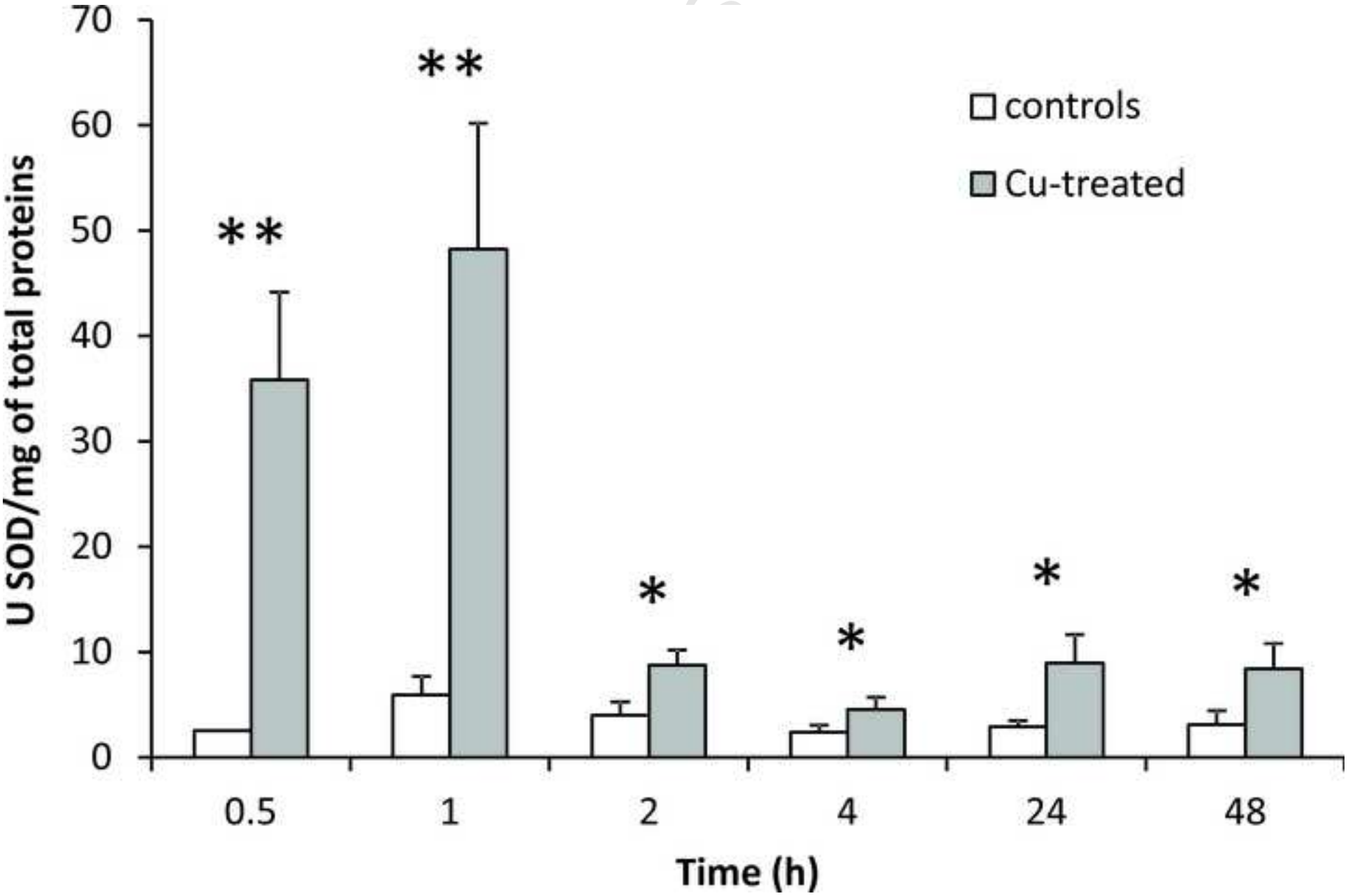

70
60
50
40
30
20
10
0
U SOD/mg of total proteins
**
**
*
*
*
*
controls
Cu-treated
0.5
1
2
4
24
48
Time (h)

Figure 6

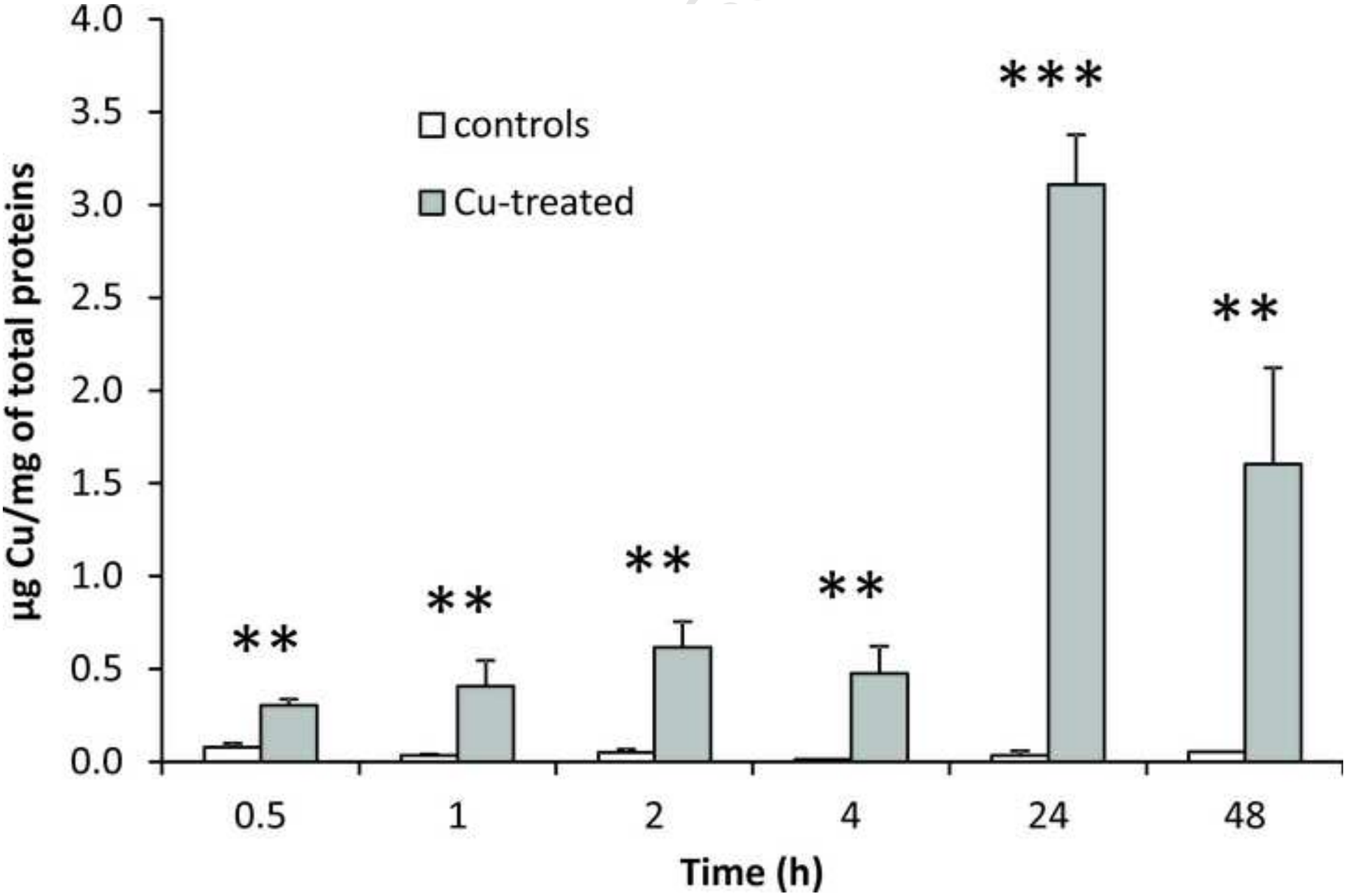

4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
µg Cu/mg of total proteins
controls
Cu-treated
**
**
**
**
***
**
0.5
1
2
4
24
48
Time (h)

Figure 7