COMPARATIVE SEQUENCE AND STRUCTURAL ANALYSIS OF SUPEROXIDE DISMUTASE (SOD), AN AGING RELATED PROTEIN IN DIFFERENT VERTEBRATES

Rituparna Sarma[1] and D. K. Sharma*[2]
[1]Bioinformatics Centre, Gauhati University, Guwahati=781014
[2]Department of Zoology, School of Biological Sciences, USTM, Baridua-793101, India
*Author for correspondence: dksgu@yahoo.co.uk

ABSTRACT
Superoxide dismutase gene SOD1 and SOD2 have been established as responsible for ageing process in animals. The sequence data of the said genes were taken from the National Centre for Biotechnology Information (NCBI) database and subjected to various processes of bioinformatics tools. The expressed proteins of these sequences taken from various species of the short-lived Drosophila melanogaster to the long lived turtles including Homo sapiens were modeled and attempted to compare their various properties. Veracity of the modeled proteins was attempted using Rampage scores. The findings may elucidate some logic in understanding the ageing process; however, it demands further study both at wet and dry lab condition.
Key Words : Structure, Superoxide dismutase, aging, vertebrates

INTRODUCTION

Aging in the living domain is directly regulated by many of the antioxidants. Metabolism by products of oxidants seems to cause extensive damages to DNA (Floyd, 1990), protein and lipids (Sies, 1993; Stadtman, 1992) which directly influence the aging process (Ames et al., 1993). Reactive oxygen species (ROS) include the Hydroxyl radical (OH), superoxide anion radical (O_2^-) ad hydrogen peroxide (H_2O_2). Unbalanced concentration of ROS often contributes to diseases like cancer, diabetes, premature aging, inflammation and hypertension (Zelko *et al*., 2002). SODs are the most prominent proteins in organisms which regulate the Reactive Oxygen Species (ROS). Cu, Zn SOD is mainly found in the cytosol of cells and outer mitochondrial space whereas Mn SOD, in the inner mitochondria space (Landis and Tower, 2005). Cu, Zn SOD plays vital role in Immune responses and removal of free radicals from cytoplasm. Mn SOD protects the cells from injuries and suppresses apoptosis (Scott *et al*. 1989). So far, SOD have been extensively studied in invertebrates and sequenced in many organisms. In case of longest living turtles, SOD1 and SOD2 has been sequenced completely only in two freshwater turtles (Rhodin *et al*. 2010), *Pelodiscus sinensis* Wiegmann, 1835 and *Mauremys reevesii* Grey, 1831. Whereas, SODs have been sequenced in shortest living organisms like Mayfly, Drosophila, and Anopheles etc. but without such structural information. Therefore, an attempt has been made in this study to analyze the SOD sequences and structural conformation from different classes of organisms, shortest living to longest living organisms.

Methods and Materials

Sequence retrieval

SOD protein sequences of flies (*Drosophila melanogaster* and *Anopheles aquasalis, Anopheles darlingi*), fish (*Larimichthys crocea, Latescal carifer, Danio rario, Callorhinchusmilii, Hemibarbus mylodon, Hypophthalmichthys nobilis, Channa striata*), amphibian (*Xenopus (Silurana) tropicalis, Xenopus laevis*), Reptiles (*Pelodiscus sinensis, Mauremys reevesii, Chelonia mydas, Thamnophi selegans*) and *Homo sapiens* were downloaded from NCBI

Database (Appendix-1). Sequences were analyzed in CLC workbench package (CLC bio) (Chang *et al*., 2010) to analyze sequence properties, alignment of the sequences.

Homology Modeling of Protein and Evaluation

The 3D structures of SOD1 and SODs were generated using comparative method SWISS MODEL (Marco *et al.,* 2014). Validation of the models was done by and RAMPAGE (Lovell *et al*., 2003). The minimization of energy and refinement of protein structures were carried out by Discovery Studio package (Accelrys) (Zhang *et al.,* 2012).

Solvents Accessible Surface Area (SASA) and active site predictions

SASA and active site predictions were carried out by using the best refined modeled structures of the proteins. Discovery Studio Client 4.0 was used to find out the SASA percentage.

RESULTS AND DISCUSSION

Sequence analysis:

SOD Sequence analysis in CLC work bench suggested that the in SOD1, SOD2 and SOD3 protein length, weight, Isolectric Point, Aliphatic index scores varies from organisms to organisms. Sequence to structure analysis of SOD1 and SOD2 of the two freshwater turtles *P. sinensis* and *M. revesii* has been done recently (Sarma and Sharma, 2015) and results were compared with other SOD structures in this study. Length of Sequences varies from 154 to 167 in SOD1, 140 to 264 in SOD2 and 240 to 274 in CCS respectively (Table 1, 2, 3). Molecular weights were around the range of 15.755 to 29.971 kDa in SOD1 and 26.317 to 29.533 KDa in CCS respectively. Isoelectric points indicated that the pH of all the SOD1 and CCS were acidic when not carrying any net electrical charge. While, in case of SOD2 weight varied from 19.444 to 28.792 KDa and isoelectric points were suggesting Basic pH scores. Aliphatic indexes were high in all SOD1, SOD2 and CCSs indicating that they have stability in different temperature ranges. CCS (gene) Copper chaperone for superoxide dismutase is a metalloproteinase that is responsible for the delivery of Cu to superoxide dismutase (SOD1). CCS is a 54kDa protein that is present in mammals and most eukaryotes including yeast. The structure of CCS is composed of three distinct domains that are necessary for its function. Although CCS is important for many organisms, there are CCS independent pathways for SOD1, and many species lack CCS all together, such as *C. elegans*. In humans the protein is encoded by the CCS gene. Alignment of SOD1s, SOD2s and CCSs separately suggested that there were variations in the respective protein groups which might be due to mutations or genetic drifts in evolution time span.

Table 1: SOD1 Sequence analysis and Comparison

UniProt Accession No.	NP_00126 1700.1	AET14 835.1	ETN65 426.1	NP_5713 69.1	ACR5 6339.1	NP_0010 16252.1	NP_000 445.1	AFX959 18.1	AGC247 78.1
Length	167aa	164aa	288aa	154aa	154aa	151aa	154aa	155aa	155aa
Weight	17.377 kDa	17.404 kDa	29.971 kDa	15.953 kDa	15.755 kDa	15.697 kDa	15.936 kDa	16.022 kDa	15.819 kDa
Isoelectric point	6.35	5.94	5.81	6.35	6.54	6.03	5.94	6.18	6.47
Aliphatic index	78.743	81.463	77.188	75.26	75.26	79.338	78.442	82.968	76.71

Table 2: SOD2 sequence analysis and comparison

UniProt Accession No.	NP_476925.1	NP_956270.1	NP_001290282.1	ACR23312.1	AIL29305.1	ADM26563.1	NP_001279581.1	AAQ63483.1	AAH16934.1	AFS63895.1	NP_001005694.1	AFX95919.1
Length	217aa	210aa	264aa	224aa	225aa	224aa	223aa	224aa	140aa	175aa	210aa	226aa
Weight	24.659 kD	23.319 kDa	28.792 kDa	25.113 kDa	25.141 kDa	24.962 kDa	25.01 kDa	25.165 kDa	15.745 kDa	19.444 kDa	23.445 kDa	25.049 kDa
Isoelectric point	8.43	8.37	8.8	8.88	8.72	8.88	9.07	8.59	8.94	9.05	8.35	9.04
Aliphatic index	85.945	85.952	80.909	81.83	83.733	84.509	83.139	82.321	79.429	80.857	84.571	83.761

Table 3: CCS sequence analysis and Comparison

UniProt accession No.	NP_001011020.1	NP_001086811.2	AFS63896.1
Length	274aa	273aa	240aa
Weight	29.518 kDa	29.522 kDa	26.317 kDa
Isoelectric point	5.82	6.12	9.52
Aliphatic index	82.555	81.062	92.292

Homology Modeling and analysis:

3D structures of all the SODs were predicted using SWISS Model Program and were evaluated using RAMPAGE. SOD structures scored more than 90.0% suggestive of their reliability for further studies. (Table- 4). Analysis of 3D structures indicated that, most SOD1s had 2 chains and were alpha helix dominant, Whereas SOD2s contained 4 chains and more or less similar number of alpha helix and beta sheet numbers. SOD1s had the similar frequency of hydrophobic residues, but there were variations in frequencies of hydrophilic residue with both the negative and positive charges. Where as in case of SOD2s, charged residue frequencies were of same but there were variations in frequencies of hydrophobic and hydrophilic residues. Glycine (G) residue distribution was found to be dominant in both the SOD1, while the distribution frequency of Leucine (L) was dominant in SOD2s. Due to dominance of Glycine, helix forming probabilities were low in SOD1, while such helix forming probabilities were evident in the SOD2 with leucine dominance. Moreover, all the four SOD sequences might be highly conserved since the Glycine and Leucine have low mutability and are more frequent in conserved sequence elements.

Table 4. Rampage scores of all the SOD and CCS structures

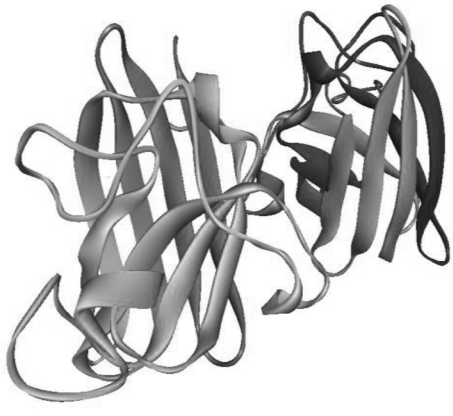
Drosophila SOD1	97.6
Drosophila SOD2	97.2
Copper zinc superoxide dismutase 3B [*Anopheles aquasalis*]	96.6
Copper zinc superoxide dismutase 3B [*Anopheles sinensis*]	97.7
superoxide dismutase [*Anopheles darlingi*]	97.6
Copper chaperone for superoxide dismutase [*Larimichthyscrocea*]	90.9
Copper zinc superoxide dismutase [*Danio rerio* SOD1]	96.7
extracellular Cu-Zn superoxide dismutase *[Larimichthys crocea*]	97.7
Extracellular superoxide dismutase [Cu-Zn] [*Larimichthys crocea*]	96.7
Intracellular Cu-Zn superoxide dismutase [*Larimichthys crocea*]	94.2
Superoxide dismutase 2, mitochondrial [*Callorhinchus milii*]	96.7
Channa striata SOD1	96.7
Hemibarbus mylodon SOD2	97.2
Larimichthys crocea SOD2	97.2
Manganese superoxide dismutase [*Hypophthalmichthys nobilis*]	91.5
Chaperone for superoxide dismutase [*Xenopus (Silurana) tropicalis*]	97.3
Copper chaperone for superoxide dismutase [*Xenopus laevis*]	98.6
Manganese superoxide dismutase [*Xenopus laevis*]	91.9
Superoxide dismutase [Cu-Zn] [*Xenopus (Silurana) tropicalis*]	97.4
Superoxide dismutase [Cu-Zn] B [*Xenopus laevis*]	97.2
Superoxide dismutase [Mn], mitochondrial [*Xenopus (Silurana) tropicalis*]	97.2
Superoxide dismutase 2, mitochondrial [*Xenopus (Silurana) tropicalis*]	95.1
Superoxide dismutase 2, mitochondrial [*Xenopus laevis*]	98.6
SOD2 isoform 1 [*Thamnophis elegans*]	97.2
SOD3 [*Thamnophis elegans*]	97.4
SOD1 [*Pelodiscus sinensis*]	98.0
SOD2 [*Pelodiscus sinensis*]	98.7
SOD1 [*M. reevesii*]	97.4
SOD2 [*M. reevesii*]	98.6
SOD1 [*Homo sapiens*]	98.0
SOD2 [*Homo Sapiens*]	93.8

Appendix-1: SOD sequence retrieval

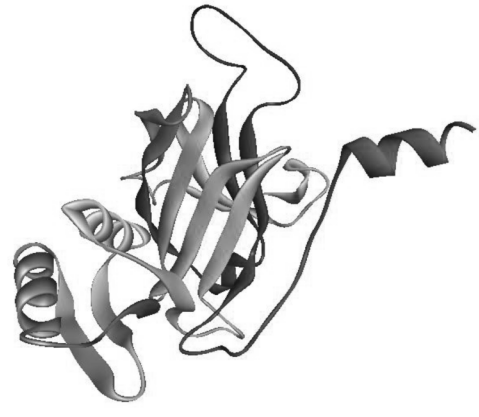
Superoxide dismutase, isoform D [*Drosophila melanogaster*]	167aa	NP_001261700.1
Superoxide dismutase 2 (Mn), isoform A [*Drosophila melanogaster*]	217 aa	NP_476925.1
Cu/Zn superoxide dismutase 3B [*Anopheles aquasalis*]	164aa	AET14835.1
Superoxide dismutase [*Anopheles darlingi*]	288aa	ETN65426.1
Superoxide dismutase [Cu-Zn] [*Danio rerio*]	154aa	NP_571369.1
Superoxide dismutase [Mn], mitochondrial [*Danio rerio*]	224aa	NP_956270.1
Extracellular superoxide dismutase [Cu-Zn] [*Larimichthys crocea*]	264aa	AIL29306.1
Hemibarbus mylodon SOD1	154aa	ACR56339.1
Hemibarbus mylodon SOD2	224aa	ACR23312.1
Intracellular Cu-Zn superoxide dismutase [*Larimichthys crocea*]	154aa	AIL29307.1
Mn superoxide dismutase [*Larimichthys crocea*]	225aa	AIL29305.1
Mn superoxide dismutase [*Hypophthalmichthys nobilis*]	224 aa	ADM26563.1
Superoxide dismutase 2, mitochondrial [*Callorhinchus milii*]	223aa	NP_001279581.1
Cu chaperone for superoxide dismutase [*Xenopus (Silurana) tropicalis*]	274aa	AAH82734.1
Cu chaperone for superoxide dismutase [*Xenopus laevis*]	274aa	NP_001086811.2

Mn superoxide dismutase [*Xenopus laevis*]	224aa	AAQ63483.1
Superoxide dismutase [Cu-Zn] [*Xenopus tropicalis*]	151aa	NP_001016252.1
Superoxide dismutase [Cu-Zn] B [*Xenopus laevis*]	151aa	NP_001080933.1
Superoxide dismutase [Mn], mitochondrial [*Xenopus tropicalis*]	224aa	CAJ81602.1
SOD2 isoform 1 [*Thamnophis elegans*]	225aa	AFS63894.1
SOD3 [*hamnTophis elegans*]	240aa	AFS63896.1
SOD1 [*Homo sapiens*]	154aa	CAG46542.1
SOD2 [*Homo sapiens*]	140aa	AAH16934.1
Mitochondrial manganese superoxide dismutase [*Mauremys reevesii*]	155aa	AFX95919.1
Cu/Zn superoxide dismutase [*Mauremys reevesii*]	210aa	AFX95918.1
Superoxide dismutase 2, mitochondrial [*Pelodiscussinensis*]	155aa	NP_001303978.1
Superoxide dismutase 1 [*Pelodiscussinensis*]	216aa	AGC24778.1

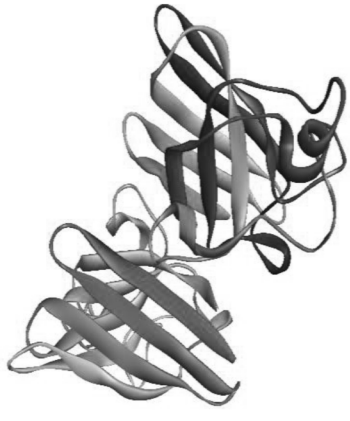
Figure A :3D structures of all the SODs



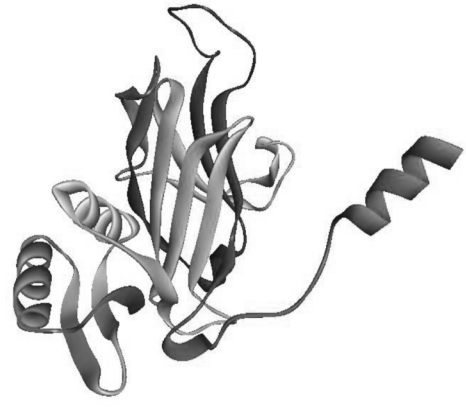
1. *Channa striata* SOD1



2. Chaperone for superoxide dismutase [*Xenopus (Silurana) tropicalis*]



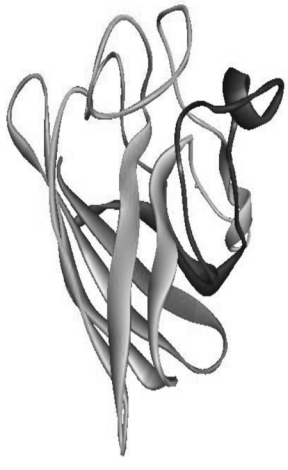
3. Copper chaperone for superoxide dismutase [*Larimichthys crocea*]



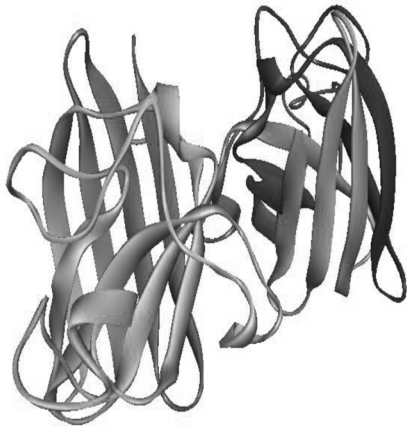
4. Copper chaperone for superoxide dismutase [*Xenopus laevis*]

5. Copper zinc superoxide dismutase [*Lates calcarifer*]

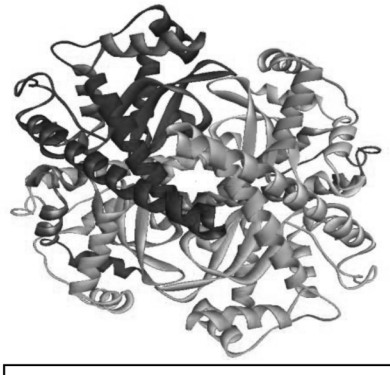
6. copper zinc superoxide dismutase 3B [*Anopheles aquasalis*]



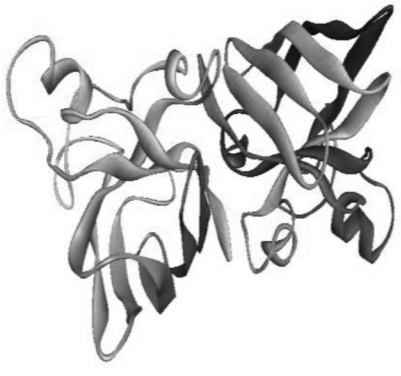
7. copper zinc superoxide dismutase 3B [*Anopheles sinensis*]



8. copper chaperone for superoxide dismutase *Chelonia mydas*



9. Copper Zinc for superoxide dismutase *Chelonia mydas*



10. *Danio rerio* SOD1

CONCLUSION

CCS (gene) Copper chaperone for superoxide dismutase is a metalloprotein that is responsible for the delivery of Cu to superoxide dismutase (SOD1). CCS is a 54kDa protein that is present in mammals and most eukaryotes including yeast. The structure of CCS is composed of three distinct domains that are necessary for its function. Although CCS is important for many organisms, there are CCS independent pathways for SOD1, and many species lack CCS all together, such as *C. elegans*. In humans the protein is encoded by the CCS gene. It may well be argued thatf the activity of various antioxidants including superoxide dismutases may be diminished or the structural identity may enhance/extend life span. In fruit fly *Drosophila melanogaster*, the absence of CuZn superoxide dismutase diminishes lifespan by ~80% and increases sensitivity to oxidative stress, whereas flies with one-half the normal amount of this enzyme have a normal lifespan. Flies lacking Mn superoxide dismutase do not develop to adulthood, but those with only 50% reduction of this enzyme are slightly shorter-lived than control (Duttaroy *et al.,*(2003). Though the present study has been attemtepted to predict the structural domain of these two proteins may induce certain light in the aging process yet demand through inputs through various means.

ACKNOWLEDGEMENT

The authors acknowledge the DBT Govt. of India for extending the computational facility in the Bioinformatics Centre, Gauhati University

REFERENCE

Ames BN, Shigenaga MK, Hagen TM (1993): Oxidants, antioxidants, and the degenerative diseases of aging. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(17): 7915–7922.

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Schwede T (2014): SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research 42*(W1), 252–258.

Chang G, hui Lin L, Luo Y, jun Cai L, jun Wu, yan X, mei H, Zhu Q. yu. (2010): Sequence analysis of six enterovirus 71 strains with different virulences in humans. *Virus Research* 151(1): 66–73.

Duttaroy A, Paul A, Kundu M, Belton AA (2003): Sod2 Null Mutation Confers Severely Reduced Adult Life Span in Drosophila. *Genetics* 165: 2295–2299.

Floyd RA (1990): Tine role off 8-hydroxygMSuniiiinie to caranogeeesns. *Commentary* (9): 1447 1450.

Fraczkiewicz R and Braun W (1998): Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *Journal of Computational Chemistry 19*(3): 319–333. .

Landis GN, and Tower J (2005): Superoxide dismutase evolution and life span regulation. *Mechanisms of Ageing and Development* 126(3): 365–379

Lovell SC, Davis IW, Adrendall WB, de Bakker PIW, Word JM, Prisant MJ, Richardson JS, Richardson DC (2003). Structure validation by C α geometry: Φ Ψ and C β deviation. *Proteins-Structure Function and Genetics* 50: 437–450.

Rhodin AG, van Dijk PP, Iverson JB, Shaffer HB, (2010). In Turtles of the world, 2010 update: annotated checklist of taxonomy, synonymy, distribution, and conservation status. *Chelonian Research Monographs* 5(3):1-85.

Sarma R and Sharma DK (2015): Sequence to Structure Analysis of SOD1 and SOD2 from Fresh Water Turtles. *J Bioinfo Proteomics Rev.* 1(2): 1–6

Scott J, Wallis SC, Davies MS, Wynne JK, Powell LM, Driscoll DM (1989): RNA editing: a novel mechanism for regulating lipid transport from the intestine. *Gut* 30: 35–43.

Sies H (1993): Strategies of antioxidant defense. *European Journal of Biochemistry / FEBS* 215(2): 213–219

Stadtman ER (2001): Protein oxidation in aging and age-related diseases. *Annals of the New York Academy of Sciences* 928: 22–38

Turtle Taxonomy Working Group – (van Dijk P, Iverson JB, Shaffer HB, Bour R, Rhodin, AGJ (2012). Turtles of the World, 2011 Update: Annotated Checklist of Taxonomy, Synonymy, Distribution, and Conservation Status. *Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group. Chelonian Research Monographs* (5):243-.328.

Wolfe-Simon F, Starovoytov V, Reinfelder JR, Schofield O, Falkowski PG (2006): Localization and role of manganese superoxide dismutase in a marine diatom. *Plant Physiol*. 142(4): 1701–1709.

Zelko IN, Mariani TJ, Folz RJ (2002): Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. *Free Radical Biology and Medicine* 33(3): 337-349.

Zhang YR, Zhao YQ, Huang JF (2012): Retinoid-binding proteins: Similar protein architectures bind similar ligands via completely different ways. *PLoS ONE* 7(5): 1–8