Molecular Modeling and Phylogeny of the Manganese Superoxide Dismutase from the Camel, *Camelus dromedarius*

Dalia Fouad^{1,2}

¹Department of Zoology, College of Science, King Saud University. P.O. Box 22452, Riyadh 11459, Saudi Arabia ²Department of Zoology and Entomology, Faculty of Science, Helwan University, Ein Helwan, Cairo, Egypt

Abstract.- The manganese superoxide dismutase (MnSOD or SOD2) is an important antioxidant enzyme in mammals as it eliminates the reactive oxygen species (ROS) produced by the aerobic reactions. The one-humped camel (*Camelus dromedarius*) is adapted to live in the widely varying arid climate and in many intrinsic and extrinsic ROS producing agents. Studying the MnSOD in *C. dromedarius* could help understand the impact of exposure to such factors on the health status of camel. The coding sequence of *MnSOD* of *C. dromedarius* (*cSOD2*) was amplified by reverse transcription PCR from the liver. The cDNA sequencing revealed an open reading frame of 666 nucleotides encoding a protein of 222 amino acids which is comparable to the *SOD2* genes from many eukaryotic organisms. The calculated molecular weight and isoelectric point of cSOD2 were 24.6 kDa and 8.15, respectively. The amino acid sequence analysis revealed the presence of the mitochondria targeting sequence at the N-terminus, the conservation of the characteristic MnSOD motif signature and the four manganese binding sites in cSOD2. The alignment and phylogenetic analysis of the cSOD2 with sequences from 18 organisms indicated that cSOD2 groups with mammals which took late evolutionary line different from SOD2 from birds, amphibians, fish and invertebrates.

Keywords: One-humped camel, superoxide dismutase, manganese superoxide dismutase, reactive oxygen species, anti-oxidant enzyme.

INTRODUCTION

During the normal cell activity, the molecular oxygen (O2) is reduced to superoxide anion (O_2) when absorbs excited electron released from the metabolic reactions. O_2^- causes severe damage to the cellular components as it denatures or inactivates enzymes (Fouad, 2015; Gardner et al., 1995), oxidizes lipids and damages the DNA (Alscher et al., 2002). Superoxide dismutase enzymes (SODs) represent the first line of defense against oxidative stress through the catalysis of O₂⁻ radical conversion into either O₂ or to less harmful hydrogen peroxide (H₂O₂) by adding or removing an electron from the O₂ (Turrens, 2003; Fridovich, 1997) in a reaction requiring a metal ion in the enzyme's active site capable of changing its oxidation state in oxidation-reduction reaction.

There are three types of mammalian SODs classified according to their cellular localization, protein folding and the catalytically important bound metal cofactor. The cytosolic and extracellular SOD (SOD1 and SOD3, respectively) contain copper and zinc, while the mitochondrial (SOD2) contains manganese (Zelko *et al.*, 2002).

The general reaction outline can be summarized as follows:

$$M^{(n+1)+}-SOD + O_2^- \rightarrow M^{n+}-SOD + O_2$$
$$M^{n+}-SOD + O_2^- + 2H^+ \rightarrow M^{(n+1)+}-SOD + H_2O_2.$$

where M = Cu (n=1); Mn, or Ni (n=2)

Mammalian SOD2 is mitochondrial enzyme consists of a tetramer harboring one manganese in each monomer. Its presence in mitochondria protects the cell against the large amount of superoxide produced during the oxidative reactions in Kreps cycle. In case of low level of SOD2, the superoxide inactivates the aconitase enzyme of the TCA cycle, ceases the energy metabolism, and releases the potentially toxic iron from aconitase (Gardner *et al.*, 1995).

^{*} Correspondence author: <u>dibrahim@ksu.edu.sa</u> 0030-9923/2015/0004-1015 \$ 8.00/0 Copyright 2015 Zoological Society of Pakistan

Abbreviations: MnSOD, manganese superoxide dismutase; TCA cycle, Tri Carboxylic acid cycle; ROS, reactive oxygen species.

Lacking SOD2 causes mice death several days after birth (Li *et al.*, 1995). Mutation and/or polymorphisms in SOD2 gene is associated with DNA damage (Van Remmen *et al.*, 2003), infertility (Mruk *et al.*, 2002), premature aging (Muller *et al.*, 2007), lipid peroxidation (Strassburger *et al.*, 2005), carotid atherosclerosis (Kakko *et al.*, 2003), and many other diseases. A new trend is emerging for using of SOD2 expression level as a biomarker for breast cancer prediction and prognosis but it is still under investigation (Becuwe *et al.*, 2014).

The one humped camel, C. dromedarius lives in harsh arid climatic conditions in the Arabian Gulf region and is continuously exposed to both endogenous and exogenous factors that necessitate powerful enzymatic and non-enzymatic antioxidant defense mechanisms capable of minimizing the impact of ROS. Camel is well adapted to such harsh desert life. So, it is proposed that it could have robust mechanisms for eliminating ROS. To the best of our knowledge, no study has been done on the molecular structure of the camel's SOD2 gene. In this study, a molecular and modelling approach has been used to identify and predict the phylogeny and structure of camel's SOD2. Comparing camel's SOD2 sequence with well-studied enzymes from many organisms provides an important perspective on SOD2 diversity and potential in the cellular defense against high ROS levels in desert living animals.

MATERIALS AND METHODS

Tissue samples and materials

Liver tissues was collected from three male camels (2 years old), immediately after killing the animal in Western Riyadh slaughterhouse by skilled veterinarian. Tissues were submerged in RNAlater® solution (Qiagen, Ambion, Courtabeuf, France) to avoid RNA degradation and stored until used at -20° C. *E. coli* strains were used and grown in LB medium supplemented with 100 µg/mL ampicillin.

Oligonucleotide design

Primers for PCR were designed from the highly conserved regions of known *SOD2* genes available in the GenBank. The primers were Fwd: 5'-CGGGCGGCGTCCACCAT-3' and Rev:5'-

GCTACATTCAGAAACGCTACAACA-3'.

Total RNA preparation and cDNA synthesis

Fifty mg of each tissue in RNAlater were homogenized in RTL lysis buffer (Qiagen) containing 1% 2-mercaptoethanol. Total RNA was extracted using AllPrep DNA/RNA Mini kit (Qiagen, Cat# 80204), following the manufacturers manual and the RNA was eluted with 50 μ L nuclease free water. Quantification of the extracted RNA was determined using NanoDrop-8000 and its integrity was assessed by formaldehyde agarose gel (1%) electrophoresis. Two μ g of the total RNAs were retrotranscribed into single stranded cDNA using ImProm-II Reverse Transcription System (Promega, Cat # A3800,) according to the manufacturer manual.

PCR amplification SOD2 gene and cloning

Gradient annealing temperature PCR was carried out from 50 to 60°C in a final volume of 50 μ L as follow: 25 μ L of GoTaq® Green Master Mix (Promega, Cat # M712c), 5 μ L of cDNA, 3 μ L of each forward and reverse primers (30 pmole) then the final volume was adjusted to 50 μ L with nuclease free water. Initial denaturation was carried out at 95°C for 1 min followed by 40 repeated cycles of 94°C for 30 seconds, 50–60°C for 45 seconds and 68°C for 60 seconds. A final extension was carried out at 72°C for 5 min and cooling to 5°C. The amplified product was analyzed by electrophoresis using 1.2% agarose gel in TAE buffer.

The band of amplified DNA of the expected size was cut from the gel after electrophoresis and purified using Wizard SV and PCR Clean-up kit (Promega, Cat # A9282), then cloned into the pGEM[®]-T Easy vector (Promega, Cat # A1360). Ligation was performed by mixing the gel-purified PCR products (2 μ L) with 1 μ L pGEM-T Easy vector (50 ng) and 5 μ L of 2× rapid ligation buffer. Finally, the reaction was initiated by the addition of 3 units of T4 DNA ligase enzyme. The final volume of the ligation reaction was adjusted to 10 μ L by the addition of nuclease free water and the ligation mixture was incubated at 15°C for 16 h. Transformation of *E. coli* JM 109 competent cells was carried out according to Sambrook *et al.* (1989)

and the positive clones were screened in selective LB/IPTG/X-gal/Ampicillin/agar plates. Moreover, colony PCR was conducted to ensure the presence of the recombinant plasmid using the universal T7/SP6 multiple cloning site primers. A small part of each bacterial colony was transferred to a clean sterile Eppendorff tube containing 5µL nuclease free water, mixed well and the rest of the PCR reaction components was added as described earlier. The colony-PCR condition was as follows, 1 cycle at 95°C for 5 min followed by 30 cycles at 94°C for 1 min, 50°C for 1 min and 72°C for 2 min. The PCR products were analyzed by 1.2% agarose gel electrophoresis.

Sequencing of the PCR product and prediction of amino acid sequence

Sequencing of SOD2 gene cloned onto pGEM-T Easy vector was carried out using DNA Analyzer (Applied Biosystems) using chain termination sequencing reaction (Sanger *et al.*, 1977) and the T7 or SP6 primers. The emerged sequences from both directions were analyzed using the Seqman of the DNAstar PROGRAM (2003). The amino acid sequence was predicted by translating the open reading frame of the sequenced DNA and compared with sequences from NCBI Protein Database using the BLASTP algorithm (http://blast.ncbi.nlm.nih.gov/Blast).

Phylogenetic analysis and multiple sequence alignment of SOD2

The deduced amino acid sequence of camel SOD2 was used to identify similar sequences of other living organisms using the PSI-BLAST facility of the NCBI website (www.ncbi.nlm.nih.gov). Homologous sequences from 18 different organisms were aligned with camel SOD2 sequence by ClustalW and the phylogenetic tree for these sequences was built using BLOSUM62 (MAFT program 2011, Jalview program 2011) from MAFFT Multiple Sequence Alignment.

Prediction of the secondary structure of cSOD2, antigenicity and hydrophilicity

The secondary structure of camel SOD2 was predicted using PSIPRED program (2008). The

subcellular localization was determined bv ProtComp 9.0 program (http://linux1.softberry.com/berry.phtml?topic=prot compan&group=programs&subgroup=proloc). The antigenic determinant in the cSOD2 were predicted according to the methods of Kolaskar and Tongaonkar method (1990) and the antigenicity score of more than 1.0 for at least six amino acid residues was considered. The hydrophilicity of cSOD2 was calculated according to the method of Parker et al. (1986).

RESULTS AND DISCUSSION

cSOD2 gene

The full length *cSOD2* obtained by PCRbased technique showed a cDNA fragment of 789 bp (Fig. 1). Figure 2 shows the sequence of SOD2 gene. It covers the full coding region preceded by 15 bases from the 5'side and 105 bases after the stop codon. cSOD2 sequence was given accession number KR023951 in the GenBank database. The open reading frame consists of 666 bases which gives deduced amino acid sequence of 222 amino acid residues (Fig. 2).



Fig. 1. Electrophoretic separation of the PCR product of camel's *SOD2* (Lane 2). Amplification was performed at 58°C annealing temperature. Lane 1 contains 1 kb ladder.

The high percentage of sequence identity between cSOD2 and other compared species

D. FOUAD

													Mito	chond	ria ta	rgetin	g sea	uence													
1	cgq	geg	qcq	tcc	acc	ATG	TTG	TGC	CGG	GCG	GCG	TGC	AGC	GCG	AGC	AGG	AAG	CTG	GTG	CCG	GCT	TTG	GGG	TCG	CTG	GGT	TCC	AGG	CAG	AAG	90
1						М	L	С	R	A	A	С	s	A	S	R	K	L	v	Ρ	A	L	G	S	L	G	s	R	Q	ĸ	25
																										•					
91	CAC	AGC	CTC	CCC	GAC	CTG	ccc	TAC	GAC	TAT	GGC	GCC	\mathbf{CTG}	GAG	CCT	CAT	ATC	AAC	GCC	GAG	ATC	ATG	CAG	CTG	CAC	CAC	AGC	AAG	CAC	CAC	180
26	н	s	L	р	D	L	₽	Y	D	У	G	Α	L	E	р	H	1	N	A	E	I	м	Q	L	H	H	s	ĸ	н	н	55
1.01		~~~	mac	080			000		ama	000		~~~			6 36	~~~	000	080	010		COR		3 000		000	aaa			0.00	~~~	070
101	666	GCC	IAC	GIG	AAT	AAL	CIG	AAT	GIU	GCC	GAG	GAG	AAG	TAT	LAG	GAG	979	CIG	GAG	AAG	GGT	GAC	AII	ACA	GCI	CAG	GIA	GUT	C1G	CAG	270
26	A	A	x	v	14	IN	11	N	¥	A	£.	E	r	I	Q	24	A	11	E	r.	6	D	1	T.	A	Ŷ	v	A	ц	ž	85
271	CCG	GCA	АТА	AAG	TTC	AAC	GGT	GGA	GGC	CAT	GTC	ААТ	CAT	TCC	ATT	TTC	TGG	ACA	AAC	CTG	AGT	CCT	ААТ	GGT	GGA	GGA	GAA	CCC	ААА	GGG	360
86	P	A	I	ĸ	F	N	G	G	G	Н	v	N	H	s	I	F	W	T	N	L	s	P	N	G	G	G	E	P	K	G	115
													=																		
361	GAA	TTA	CTG	GAA	GCC	ATC	AAC	CGT	GAC	TTT	GGT	TCC	TTC	AGC	AAA	TTT	AAG	GAG	AAG	TTG	ACC	GCT	GTA	TCC	GTT	GGC	GTC	CAA	GGC	TCG	450
116	Е	\mathbf{L}	\mathbf{L}	Е	A	I	N	R	D	F	G	s	F	s	ĸ	F	ĸ	E	K	L	т	А	v	s	v	G	v	Q	G	s	145
451	GGT	TGG	GGT	TGG	CTT	GGT	TTC	AAT	AAG	GAA	CAG	GGA	CGC	TTA	CAG	ATT	GCT	GCT	TGT	TCT	AAC	CAG	GAT	ccc	TTG	CAA	GGA	ACA	ACA	GGT	540
146	G	W	G	W	L	G	F	N	K	E	Q	G	R	L	Q	I	A	A	С	s	N	Q	D	P	L	Q	G	т	T	G	175
									N	1n - SC	DD m	otif sig	gnatur	e																	
541	CTT	ATT	CCA	TTG	CTG	GGA	ATT	GAT	GTG	TGG	GAG	CAC	GCT	TAC	TAC	CTT	CAG	TAT	AAA	AAT	GTT	AGA	CCT	GAT	TAC	CTG	AAA	GCT	ATT	TGG	630
176	L	I	Ð	\mathbf{r}	\mathbf{r}	G	I	D	V	W	E	H	Α	Y	Y	L	Q	Y	ĸ	N	v	R	P	D	Y	L	K	А	I	W	205
694						~ ~		0.57.3	1.07	~ ~ ~	101				maa																700
631	AAT	GIA	ATC	AGC	TGG	GAG	AAT	GTA	ACT	CAG	AGA	TAC	CTG	GCG	TGC	AAA	AAG	TAG	agc	gtc	age	OTT	acc	etg	agt	aca	cgg	age	tdd	ττα	120
206	N	v	1	s	W	E	N	v	T	Q	R	¥	Ц	A	С	ĸ	К	*													
721	tga	cta	tag	tag	tgo	aga	gto	oog	cgg	tat	acc	agt	aag	atg	oto	tgt	tgt	age	gtt	tet	gaa	tgt	age	78	39						

Fig. 2. The nucleotide and the deduced amino acids sequences of the cSOD2. The sequences were submitted to NCBI GenBank (accession number KR023951). The mitochondrial targeting sequence and the MnSOD motif signature were highlighted. The metal binding residues and their corresponding codons are labelled in blue colour.

indicate that the gene encoding this protein originates from common ancestor that maintains high conservancy during evolution (Figs. 3, 4). The Mn-SOD motif signature (DVWEHAYY starting from 186 to 193 on the graph) and the four manganese binding sites (H54, H102, D187 and H191 on the graph) are conserved in cSOD2 and in all compared organisms including mammals, amphibians, birds, fish and arthropods. The identification of both the metal-binding residues and the signature sequence support that our cMnSOD belongs to the MnSOD family.

SOD2 is a nuclear gene that is translated in the cytoplasm and exported to the mitochondria where it confers protection from ROS induced oxidative damage. The analysis of the cSOD2 aminoacid sequence using MitProt website (http://ihg.gsf.de/ihg/mitoprot.html) revealed the presence of putative N-terminal region of 20 aminoacids (MLCRAACSASRKLVPALGSL) that strongly supports a mitochondrial targeting with high probability (0.9907) (Claros and Vincens, 1996). Most mitochondrial MnSODs contain signal peptides which would be essential for translocation into the mitochondria (Fukuhara et al., 2002). The multiple sequence alignment indicated that this sequence is poorly conserved in the compared sequences (Fig. 3). The first 20 amino acids of cSOD2 contains the major characteristics of the mitochondrial targeting sequence like the abundance of the positively charged (lysine and arginine) and the hydroxylated residues (mostly serine). These residues together with some hydrophobic residues form an amphipathic α -helices which is thought to be important for the translocation through the mitochondrial outer and inner membranes (von Heijne 1986; Roise and Schatz, 1988).

Amino acid composition and protein secondary structure

The bioinformatics analysis of the 222-amino acid sequence of cSOD2 using the program PROTEAN (2003) showed that it has a calculated molecular weight of 24.6 KDa and isoelectric point (pI) of 8.15. The predicted protein contains 64 charged amino acids (28.8%), 79 hydrophobic (35.6%), 20 acidic (9%), 22 basic (9.9%) and 59 polar amino acids (26.6%). The detailed amino acid analysis of the predicted protein is given in Table I.

Multiple sequence alignment

The PSI-protein blast comparison between the predicted amino acid sequence of cSOD2 and sequences from the GenBank of the bestcharacterized representatives of SOD2 from different organisms was carried out. The cSOD2

		Mitoc	hondri	a target	ing sequer	ice										
			10		20		30		40		50	V	60		70	
Camelus dromedarius	1	ML CR	AACS-	- ASRI	KLVPAL	SLGS	ROKHS	PDL	PYDYGAI	EPHI		HHSK	HHAAY	VNNLNV	AEEKY	QEA72
Sus scrofs	1	ML C R/	A A C S -	- SSRI	NL V <mark>P</mark> AL (GVL G S	R <mark>Q K</mark> HTI	LPDL	PY DY GAI	LEPHII	N A <mark>Q</mark> I M <mark>Q</mark> I	L HHS <mark>K</mark>	HHAAY	V <mark>NNL N</mark> V	VEEKY	Q <mark>E A</mark> 72
Ovis aries	1	· · ML SR/	A A C S -	- TSR	KL V <mark>P</mark> AL (GVL G <mark>S</mark>	R <mark>QK</mark> HSI	L <mark>PDL</mark> I	PY DY G A I	L <mark>E P H I</mark> I	N A <mark>Q</mark> I M <mark>Q</mark> I	HHSK	HHAAY	VNNL NV	AEE KY	R <mark>E A</mark> 72
Bos taurus	1	ML S <mark>R</mark> /	AACS-	- T SRE	RL V <mark>P</mark> AL S	SVL G S	RQKHSI		PY DY GAI	LEPHII	N A Q I MQI	HHSK	HHAAY	VNNL NV	AEEKY	R <mark>E A</mark> 72
Equus caballus	1	ML C R/	AACS-	- TSR	KL VPAL (GSLG <mark>S</mark>	RQKHSI		Q Y D Y G A I			HHSK	HHAAY	VNNL NV	TEEKY	Q EA 72
Homo sapieus	1	ML SR/	AVCG-	- TSRC		GYLGS	ROKHSI		PYDYGAI	EPHII		HHSK	HHAAY		NEEKY	Q EA 72
Pongo abelli	1	ML SRO	SVCG-	- TSRC	DLAPALO	YLGS	ROKHS	PDL	PYDYGAI	EPHI		HHSK	HHAAY	VNNLNV	IEEKY	QEA72
ALUS BRUSCHUS Collas gollas	1		AACS-	- I GRI			RHKHS		PTDTGAI	EPHI		HHSN	HHAAY		TEENT	HEATZ
Canas ganas Vanonus teopicalis	1			DCDI					DYDYCAL		SAEIMO				TEEN	AEA74
Danio reria	1			RCAAT		AVTS	ROKHA	PDI	TYDYGAL	EPHI	CAFIMO	HHSK	HHATY	VNNI NV	TEEKY	
Salmo salar	1		GOVR	RCAAT		TVAA	RWKHS	PDI	TYDYGAI	EPHI	NAFIMO	HHSK	HHATY	VNNI NV	TEEKY	OFA76
Cancer pagurus	1	MLLA	RAFFP	ARS	SLVAAG	SWC	ROKHT	PDL	PYDYGAI	EPTI	SAEIMO	HHSK	нноту	VNNLNI	GEEKL	AEA70
Argopecten irradians	1	ML SAT	TTVL	KSIP	SVGAFO	ALAS	RLKHT	PDL	PYDYNAI	LEPYI	SADIMQ	HHSK	HHAAY	VNNLNI	AEEKL	AEA74
Schistosoma japonicum	1	MVFF	R <mark>SS</mark> MS	LVHIL	L R <mark>S S V L</mark> F	RNYGV	RFKHT	PPL	PYDP SAI	LEPVI	SKEIMQI	HHSK	HHAAY	VNNLNI	AEEQF	ADA73
Echinococcus granulosus	1	MLWL	<mark>g l h</mark> c r	t HL (GSVFRC1	RVL S	A K <mark>K H T</mark> I	LPDL	PYD F G A I	L EPVI	SAEIMK	снуок	HHAAY	VNNLNI	VEEKM	T <mark>E A</mark> 72
Apis cerana cerana	1		M F A	VRRII	FSNTV	(DIFV	R T KHT	L PDL I	P <mark>y dy</mark> kai	LEPIIS	S A E I MQI	HHSK	HHATY	VNNL NV	AEE KM	K <mark>E A</mark> 68
Hydra valgaris	1	MF SF	GIHH-	L	SVFRKIS	SRIAF.	ANKHTI	LPEL	GYEY NAI		S GQ I ME	I HHR <mark>K</mark>	HHQAY	VNNLNT	AEEQL	A <mark>E A</mark> 70
Bombyx mori	1		ML M	ISQRI	G <mark>s</mark> lirv <i>i</i>	G - AS	RQKHT	LPELI	PYEYNAI	LEPVI	SR <mark>eim</mark> si	HHSK	HHATY	INNLNV	AEEKL	A <mark>Q A</mark> 67
		80		90		100		110		120		130		140		150
Cametus dromedarius	73					UVNU				CELLE		CELS		TAVEV	CVOCS	TAIC 149
Sus scrola	73				KENGGO	HINH				GELLE	AIKPDE	GSEE	KEKEK	TAVSV	GVOGS	SWG 148
Ovis aries	73	EKGDV			KENGGO	HINH			GGGEP	GELLE	AIKRDE	GSEA	KEKEK	TAVSV	GVOGS	WG 148
Bos taurus	73	LEKGDV			KENGGO	HINH	FWT	IL SPI	GGGEPC	GELLE	AIKRDE	GSEA	KEKEK	TAVSV	GVOGS	GWG 148
Equus caballus	73	LAKGDV	AQIA		KFNGGG	HINH	I FWT		GGGEP	GKLL	AIKRDE	GSFD	KEKEK	TAVSA	GVQGS	GWG 148
Homo sapieus	73	LA <mark>kgdv</mark> i			K F NGGG	HINH	SIFWT	IL SPI	GGGEP	GELLE	AIKRDE	GSFD	K F <mark>K E K</mark> I	LTAASV	GVQGS	GWG 148
Pongo abelii	73	L A <mark>k</mark> gdv1	Γ Α Ο Ι Α		L	HINH	S I FWT I	NL SPI	GGGEP	< <mark>ge</mark> ll <mark>e</mark>	A I KRD F	GSFD	K F <mark>K E K</mark> I	LTAASV	GVQGS	G <mark>WG</mark> 148
Mus musculus	73	L A <mark>kgdv</mark> i	T <mark>Q V</mark> A	L <mark>Q P</mark> A L	L <mark>K</mark> F <mark>N</mark> G G G	HINH	E I FWTI	VL SPI	K G G G E P H	(<mark>GE</mark> LL <mark>E</mark>	A I KRD F	GSFE	K F <mark>K E K</mark> I	L <mark>T A V S</mark> V	GVQGS	G <mark>WG</mark> 148
Gallus gallus	75	L A <mark>kgd</mark> v1	raqvs		KFNGGG	HINH	I FWT	VL SP	SGGGEPH	GELME	AIKRDE	GSFA	NFKEKI	LTAVSV	GVQGS	GWG 150
Actiopus tropicaiis	75		TOVS	LQAAL	KFNGGG	HINH	I FWT	IL SPI	GGGEPO	GELLI	DAIKRDE	GSFE	KEKEKI		GVQGS	GWG 150
Salmo rerio	75		TOVS		KENGGG	HINH	FWI		GGGEPO	GELLE	AIKRDE	GSFQ	KMKEK	SAATV	AVQGS	SWG 150
Cancer Daguras	71				KENGGO					GELL		CEVE		SAATV	AVUGS GVOGS	SWG 152
Argopecten irradians	75	TETKNI			KENGGO					CDLL		CELE		SEAST	AVOCS	CMC 150
Schistosoma japonicum	74	MSKSDVI	KMIS		RENGGO	HINH	FWH		GGGGVPI	GSLA		GSED	NEKSR	SATTI	ALOGS	SWG 149
Echinococcus granulosus	73	LHKGDA	TIIS	LOPAR	FKFNGGG	HINH	SI FWC		GGGEPS	GPLA	AIKRDE	GSFE	AFKEK	TNATY	SVEGS	GWG 148
Apis cerana cerana	69	VAKGDVI	TOVA	LSPAI	KFNGGG	HLNH	SI FWC	IL SPI	GG - KPL	AALL		GSLE	ЕМККО	SENTV	AIQGS	GWG 143
Hydra valgaris	71	QHKGDTS	KIIS		KFNGGG	HINH	SIFWT	NL SPI	GGGKPI	GELLE	AIIKDE	GSFE	AMKTR	SSSAV	AVQGS	GWG 146
Bombyx mori	68	Q <mark>a k g d i</mark> i	DTIIN	IL A <mark>p</mark> al	L <mark>K</mark> F <mark>N</mark> G G G	HINH	S I FWH	VL SPI	GG-KPS	SDVLT	AVE KD F	G SWD	NIKNQI	STASV	AVQGS	G <mark>WG</mark> 142
							3.5	60D -								
							NIR-	ອບມະ	nom signa	ture 1						
			160	-	170	_	180	١V			200	_	210	22	20	
Camelus dromedarius	149	WL G F NK	EQGRL		CSNQDPL	QGTT	GL I PL I	GID	WEHAY	LOYKI		KAIW		ENVTOR	YLACK	222 222 222
Sus scrola	149	WLGFNK	QGRL		CSNQDPL	QGII	GLVPLI	GID	WEHAY	LOYK	VRPDYL	KAIW		ENVTER	YAACK	222
Ovis arres	149	WL GFNK	COCRU	QI AA	CSNQUPL	QGIT		GID	VWEHAY	LQYN	VRPDYL	AIW			TACS	222
Dos thurns Fanns caballus	149	AL CENK		O I VA		OCTT		GID	VVVE HAT			A IVV			TACS	222
Homo seniens	149	M GENK	RGHI	OI AA		OGTT		GID	WEHAY	LOYK		KAIW	NVINW		YMACK	222
Pongo shchi	149	MGENK	RGHI	OLAA	CPNODPI	OGTT	GLIPII	GID	WEHAY	I OY K	VRPDYI	KAIW	NVINW	ENVIER	YMACK	222
Mus musculus	149	WLGFNK		QIAA		OGIT	GLIPLI	GID	WEHAY	LOYK	VRPDYL	KAIW	NVINW	ENVIER	YTACK	4 - 222
Gallus gallus	151	WL GYNK	EQGRL	QIAA	CANODPL	QGTT	GLIPLI		WEHAY	LQYK	VRPDYL	KAIW		ENVSOR	YESCR	4 224
Xenopus tropicalis	151	WL GYNK	ESNRL	QLAA	CANQDPL	QGTT	GLIPLI		WEHAY	LQYK	VRPDYN	KAIW	NVINW	ENVAER	YRASKI	224
Danio rerio	151	WL G F E <mark>K</mark>	ESGRL	RIAA	CANQDPL	QGTT	GLIPLI	L <mark>G I D</mark> '	WEHAY	LOYK	VRPDYN	/KAIW	NVVNW	ENVSER	FQAAK	- 224
Salmo salar	153	ML G F D K	ESGKL	RITA	CPNQDPL	QGTT	GL VPL I		WEHAY	LOYK	VRPDYN	KAIW		ENVSER		226
Cancer pagurus	146	WL GYNK	RGTL	QIAT	CPNQDPL	EATT	GLVPL	GID	WEHAY	LOYK	VRPDY	KAIW	NIANW	KDITAR	FTAAK	218
Argopecten Hradians	151	WL GF NP	VSKRL	RIAA	CANODPL	COPT 1		GID	WEHAY	LUYK	A B B B B Y	NAIW	NVVNW	DCV SQK		NDC 226
Senistosoma japonicum	150	M GLED	SSCR	RIAT	CANODDI	EGTT		GID	AMEHAY		ARPUY	A IW		KDVAKK	EARAG	224
r.cumococcus granulosus	14.9	MGYCO	KSKEL	RIAT	CANODDI	QATT	GLIDI	EGID	WEHAY	(LOV P	ARPDY	KALE	DVVNM	NDVNSD	KKAL	S . 219
Apis cerana cerana Nedeo entronio	147	MGYDS	VTKR	ALTA	PNODP	QATT	GLIPI	GID	WEHAY	(LOYK	VRIDY		NIIDW	KNVSAR	EVAAK	219
Bombyx mori	143	WL GYNK	MKKL	QIAT	CQNQDPL	QATT		FGID	WEHAY	LOYK	VRADY	KAIF	DVANW	ND I SQR	YEKAL	216

Fig. 3. Amino acid sequence alignment of cSOD2 with 18 different animal proteins using the MAFFT Multiple Sequence Alignment program. Residues are colored according to their conservancy. The mitochondrial targeting sequence and the MnSOD motif signature were marked and the Mn binding residues are labeled by $(\mathbf{\nabla})$.

sequence was aligned with 18 different organisms belonging to different families in the animal kingdom using ClustalW (*MAFFT* program, version 6.864, 2008 and *Jalview*, version 2.3, 2011) (Fig. 3). The cSOD2 shared high similarity with SOD2 from many mammalian species. The highest identity was found with pig *S. scrofa*, cattle *B. taurus*, horse *E. caballus*, human *H. sapiens*, the monkey Sumatran orangutan *P. abelii*, and house mouse *M. musculus* (93-89%) (Table II, Fig. 3). Such high identity



Fig. 4. The phylogeny of the cSOD2 and potentially related proteins. The amino acid sequence of cSOD1 was compared with 18 different sequences of many animal families from the GenBank database. The phylogenetic tree was generated with the BLOSUM62 from MAFFT Multiple Sequence Alignment.

Amino Acid	Number	% by	% by
	count	weight	frequency
Ala (A)	20	5.78	9.01
Cys (C)	4	1.68	1.80
Asp (D)	7	3.28	3.15
Glu (E)	13	6.82	5.86
Phe (F)	6	3.59	2.70
Gly (G)	21	4.87	9.46
His (H)	9	5.02	4.05
Ile (I)	11	5.06	4.95
Lys (K)	15	7.82	6.76
Leu (L)	24	11.04	10.81
Met (M)	2	1.07	0.90
Asn (N)	14	6.49	6.31
Pro (P)	10	3.95	4.50
Gln (Q)	12	6.25	5.41
Arg (R)	7	4.44	3.15
Ser (S)	14	4.96	6.31
Thr (T)	6	2.47	2.70
Val (V)	12	4.84	5.41
Trp (W)	6	4.54	2.70
Tyr (Y)	9	5.97	4.05
Charged amino acids	64	35.03	28.83
(RKHYCDE)			
Acidic (DE)	20	10.10	9.01
Basic (KR)	22	12.26	9.91
Polar (NCQSTY)	59	27.82	26.58
Hydrophobic	79	34.85	35.59
(AILFWV)			

 Table I. Predicted chemical composition of the camel's SOD2 using Protean program.

proposed a close evolutionary relationship which best figured by drawing the phylogenetic tree of the examined proteins. Figure 4 indicates that cSOD2 groups with *S. scrofa* and other mammals which took late evolutionary line different from birds, amphibians, fish and invertebrates.

The secondary structure of cSOD2 was predicted using PSIPRED program (2008) (Fig. 5) which indicated that the cSOD2 is composed of 10 major alpha helices and only 2 small β -sheets.

It was also predicted that most of the sequence of cSOD2 is hydrophilic and, as a result, it is highly antigenic as the majority of the protein surface is exposed to the aqueous medium. There are at least 8 potential antigenic peptides (Fig. 6) according to the method of Kolaskar and Tongaonkar (1990) <u>http://emboss.bioinformatics.nl/cgi-bin/emboss/</u> antigenic. The predicted antigenic sequences and their positions are presented in Table III.

CONCLUSIONS

Camel's MnSOD share the same structure characteristics of mammalian SOD2. Its predicted amino acid sequence and the alignment with sequences from many candidates from the animal kingdom helped in building and determine the phylogenecity and antigenicity of the cSOD2.



Fig. 5. The predicted secondary structure of the cSOD2 sequence using PSIPRED program. Pink cylinders indicated alpha helix and yellow arrows indicated β -sheets.



Fig. 6. Predicted hydrophilicity and antigenic properties of cSOD2 using the method of Parker *et al.*, (1986) and Kolaskar and Tongaonkar method (1990), respectively.

Table II.- Comparison of cSOD2 and other SOD2 enzymes from different organisms including mammals, birds, amphibians, fish, and invertebrates. The comparison included number of amino acid sequence, percent identity and E-value.

Organism	(NCBI Ref.)	a.a. Residues	Total score	Identity (%)	Positive (%)	Gap (residue)	E-Value
Camelus dromedarius	KR023951	222		100	100	0	
Sus scrofa	NP 999292	222	440	93	97	Ő	9e-155
Bos taurus	NP 963285	222	435	92	96	Ő	7e-153
Equus caballus	NP 001075986	222	429	90	95	0	2e-150
Homo sapiens	CAA42066	222	424	90	94	0	1e-148
Pongo abelii	NP_001127035	222	424	89	93	0	2e-148
Mus musculus	NP_038699	222	422	89	92	0	8e-148
Gallus gallus	NP_989542	224	398	83	91	2	3e-138
Xenopus tropicalis	NP_001005694	224	382	79	88	2	8e-132
Danio rerio	NP_956270	226	380	78	88	6	5e-131
Salmo salar	ACN10263	226	370	76	86	6	3e-127
Cancer pagurus	CAR82596	218	325	70	81	3	2e-109
Schistosoma japonicum	AAW26480	223	307	62	78	1	2e-102
Apis cerana cerana	AET7406	218	299	70	82	1	3e-99
Bombyx mori	NP_001037299	216	291	67	80	1	5e-96
Hydra vulgaris	XP_002160626	219	292	67	81	0	2e-96

N	Commence	Posi	tion	Case	Man accession	T	
INO.	Sequence	Start	End	Score	Max. score pos.		
1	KLTAVSVGVQ	134	143	1.159	140	10	
2	RAACSASRKLVPALGSL	4	20	1.155	5	17	
3	TAQVALQPAI	79	88	1.142	10	85	
4	GLIPLLGIDVWEHAYYLQYK NVRPDYLKAIWNV	175	207	1.142	33	179	
5	RLQIAACSNQ	158	167	1.138	10	162	
6	MQLHHSKHHAAYVNNL	47	62	1.116	16	56	
7	KHSLPDLPYDYGALEPH	25	41	1.095	17	30	
8	GELLEA	115	120	1.042	6	118	

Table III.- The list of the predicted antigenic amino acid sequences and their positions.

ACKNOWLEDGMENTS

This research project was supported by a grant from the "Research Center of the Female Scientific and Medical Colleges", Deanship of Scientific Research, King Saud University.

REFERENCES

- ALSCHER, R.G., ERTURK, N. AND HEATH, L.S., 2002. Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. J. exp. Bot., 53: 1331–1341.
- ATAYA, F. S., AL-JAFARI, A. A., DAOUD, M.S., AL-HAZZANI, A.A., SHEHATA, A.I., SAEED, H.M. AND FOUAD, D., 2014. Genomics, phylogeny and in silico analysis of mitochondrial glutathione Stransferase-kappa from the camel Camelus dromedarius. Res. Vet .Sci., 97: 46-54.
- ATAYA, F.S., FOUAD, D., AL-OLAYAN, E. AND MALIK, A., 2012. Molecular cloning, characterization and predicted structure of a putative copper-zinc SOD from the camel, *Camelus dromedarius*. Int. J. mol. Sci., 13: 879-900.
- BECUWE, P., ENNEN, M., KLOTZ, R., BARBIEUX, C. AND GRANDEMANGE, S., 2014. Manganese superoxide dismutase in breast cancer: from molecular mechanisms of gene regulation to biological and clinical significance. *Free Radic. Biol. Med.*, **77**: 139-151
- CLAROS, M.G. AND VINCENS, P., 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.*, **241**: 770-786.
- FOUAD, D., 2015. Antioxidant and modulatory effect of melatonin on hepatotoxicity and oxidative stress induced by orange yellow s in male rats. *Pakistan J. Zool.*, **47**: 383-391.
- FRIDOVICH, I., 1997. Superoxide anion radical (O₂), superoxide dismutases, and related matters. J. biol.

Chem., 272: 18515-18517.

- FUKUHARA, R., TEZUKA, T. AND KAGEYAMA, T., 2002. Structure, molecular evolution, and expression of primate superoxide dismutases. *Gene*, **296**: 99e-109.
- GARDNER, P.R., RAINERI, I., EPSTEIN, L.B. AND WHITE, C.W., 1995. Superoxide radical and iron modulate aconitase activity in mammalian cells. *J. biol. Chem.*, 270: 13399–13405.
- JALVIEW, 2011. version 2.3; University of Dundee: Scotland, UK.
- KAKKO, S., PAIVANSALO, M., KOISTINEN, P., KESANIEMI, Y.A., KINNULA, V.L. AND SAVOLAINEN, M.J., 2003. The signal sequence polymorphism of the MnSOD gene is associated with the degree of carotid atherosclerosis. *Atherosclerosis*, 168: 147–152.
- KOLASKAR, A.S. AND TONGAONKAR, P.C., 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.*, 276: 172– 174.
- LI, Y., HUANG, T.T., CARLSON, E.J., MELOV, S., URSELL, P.C., OLSON, J.L., NOBLE, L.J., YOSHIMURA, M.P., BERGER, C., CHAN, P.H., WALLACE, D.C. AND EPSTEIN, C.J., 1995. Dilated cardiomyopathy and neonatal lethality in mutant mice lacking manganese superoxide dismutase. *Nat. Genet.*, **11:** 376–381.
- MAFFT, 2011. version 6.864; Computational Biology Research Center (CBRC): Tokyo, Japan.
- MRUK, D.D., SILVESTRINI, B., MO, M. AND CHENGA, C.Y., 2002. Antioxidant superoxide dismutase review: its function, regulation in the testis, and role in male fertility. *Contraception*, **65**: 305e-311.
- MULLER, F.L., LUSTGARTEN, M.S., JANG, Y., RICHARDSON, A. AND VAN REMMEN, H., 2007. Trends in oxidative aging theories. *Free Radic. Biol. Med.*, 43: 477–503.
- PARKER, J.M., GUO, D. AND HODGES, R.S., 1986. New hydrophilicity scale derived from high-performance

liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*, **25**: 5425–5432.

- PROTEAN, 2003. version 5.07; DNASTAR, Inc.: Madison, WI, USA.
- PSIPRED, 2008. version 3.0; University College London: London, UK.
- ROISE, D. AND SCHATZ, G., 1988. Mitochondrial presequences. J. biol. Chem., 263: 4509-4511.
- SAEED, H.M., ALANAZI, M.S., SHALABY, M.A., ALSHAHRANI, O., ATAYA, F.S., PATHAN, A.A. AND ABDULJALEEL, Z.A., 2014. Molecular cloning and cDNA characterization of Camelus dromedarius putative cytochrome P450s 1A, 2C, and 3A. *Genet. Mol. Res.*, 13: 2886-2905.
- SAMBROOK, J., FRITSCH, E. AND MANAIATIS, T., 1989. Molecular cloning: a laboratory manual. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- SANGER, F., NICKLEN, S. AND COULSON, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. natl. Acad. Sci. USA*, 74: 5463–5467.
- SEQMAN, 2003. version 5.07; DNASTAR, Inc. Madison, WI, USA.
- STRASSBURGER, M., BLOCH, W., SULYOK, S., SCHULLER, J., KEIST, A.F., SCHMIDT, A., WENK,

J., PETERS, T., WLASCHEK, M., KRIEG, T., HAFNER, M., KUMIN, A., WERNER, S, MULLER, W. AND SCHARFFETTER-KOCHANEK, K., 2005. Heterozygous deficiency of manganese superoxide dismutase results in severe lipid peroxidation and spontaneous apoptosis in murine myocardium in vivo, *Free Radic. Biol. Med.*, **38**: 1458–1470.

- TURRENS, J.F., 2003. Mitochondrial formation of reactive oxygen species. J. Physiol., 15: 335–344.
- VAN REMMEN, H., IKENO, Y., HAMILTON, M., PAHLAVANI, M., WOLF, N., THORPE, S.R., ALDERSON, N.L., BAYNES, J.W., EPSTEIN, C.J., HUANG, T.T., NELSON, J., STRONG, R. AND RICHARDSON, A., 2003. Life-long reduction in MnSOD activity results in increased DNA damage and higher incidence of cancer but does not accelerate aging. *Physiol. Genomics*, **16**: 29–37.
- VON HEIJNE, G., 1986. Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J.*, **5**: 1335-1342.
- ZELKO, I.N., MARIANI, T.J. AND FOLZ, R.J., 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 Radic. Biol. Med.*, 33: 337-349.

(Received 7 April 2015, revised 11 May 2015)