First publ. in: Flavins and flavoproteins 1999 : proceedings of the Thirteenth International Symposium, Konstanz, Germany, August 29 - September 4, 1999, pp. 591-594

Comparison of the amino acid sequence of *Calloselasma rhodostoma* L-amino acid oxidase to other FAD-dependent oxidases

Oliver Seth[1,4*], Lo-Chun Au[2], Sandro Ghisla[3*], Claus Bollschweiler[4], Manfred Kurfürst[5*] and Peter Macheroux[1*]

1 Eidgenössische Technische Hochschule Zürich, Institute of Plant Sciences, CH-8092 Zürich, Switzerland; 2 Veterans General Hospital-Taipei, Taipei, Taiwan; 3 Universität Konstanz, Fakultät für Biologie, D-78567 Konstanz, Germany; 4 Knoll AG, D-67008 Ludwigshafen, Germany and 5 Knoll AG, D-25436 Uetersen, Germany

Introduction

L-amino acid oxidase (LAAO) occurs in many different organisms such as bacteria (*Corynebacterium*, *Proteus rettgeri*), cyanobacteria (*Synechococcus*), fungi (*Neurospora crassa*), green algae (*Chlamydomonas rheinhardtii*) and most prominently, venomonous snakes (such as *Crotalidae* and *Viperidae*). Ophidian LAAOs contribute to the toxicity of snake venom while in all other cases the enzyme appears to be involved in utilisation of ammonia as a nitrogen source. There are also major differences in molecular weight, substrate specificity, glycosylation and the physiological regulation of LAAO. In view of this diversity it can be envisaged that these proteins evolved from different ancestral proteins. In order to aid investigations aimed at the elucidation of the three-dimensional structure and the analysis of the oligosaccharide moiety of LAAO from the Malayan pit viper (*Calloselasma rhodostoma*) we have cloned and sequenced the gene from a cDNA library. Here we report the phylogenetic relationship of *C. rhodostoma* LAAO to other FAD-dependent oxidases and describe a highly conserved FAD binding fold.

Results and Discussion

The mature LAAO from *C. rhodostoma* is 498 amino acids in length and has a molecular mass of 56178 Da. The amino acid sequence is 83% identical to the closely related LAAOs from *Crotalus adamanteus* (1) and *Crotalus atrox* (accession number AF093248) (see Figure 1). The ophidian LAAOs are more distantly related to the interleukin 4-induced Fig-1 protein from *Mus musculus* (1), a putative LAAO from *Bacillus subtilis*, and a protein of unknown function from *Bacillus cereus* (denoted #6, #1, and #2 in Figure 1 and 2). However, there is no apparent relationship to the LAAOs from *Neurospora crassa* (#6), *Chlamydomonas rheinhardtii* (#7) and *Synechococcus* (#8) suggesting that LAAOs have evolved from different ancestral proteins.

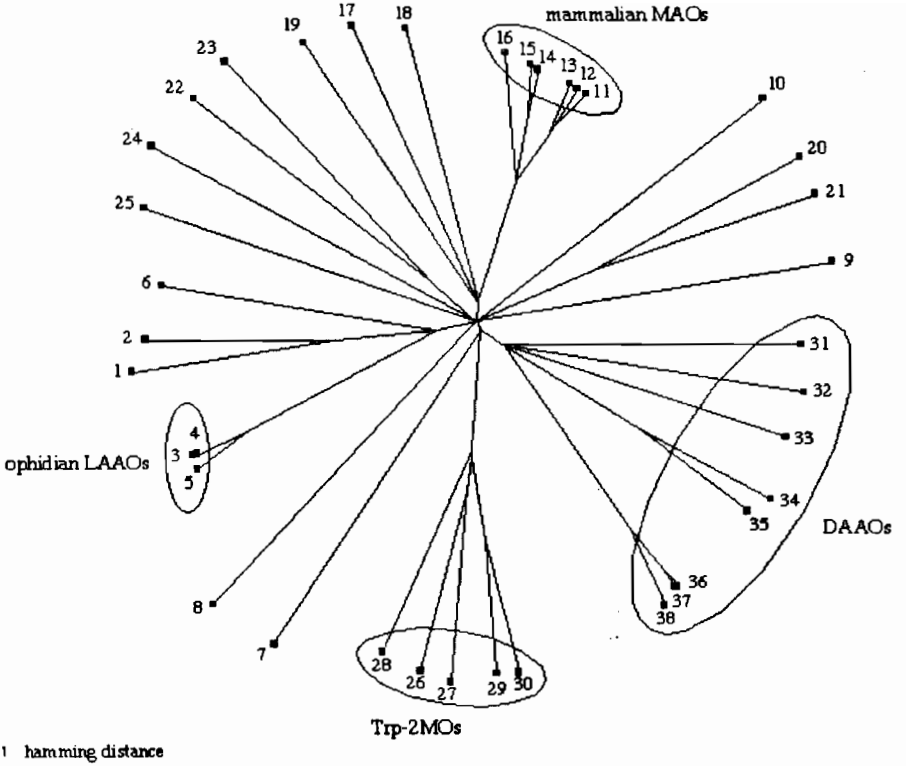

Figure 1. Phylogenetic reconstruction was carried out with the programm SplitsTree2.4 by David Huson based on the split decomposition method described recently (5). The complete amino acid sequences of the proteins were used for analysis. The circles indicate the suggested families of proteins. The protein numbering is as follows: 1, putative LAAO (*Bacillus subtilis*); 2 LAAO (?) (*Bacillus cereus*); 3, LAAO (*Crotalus adamanteus*); 4, LAAO (*Crotalus atrox*); 5, LAAO (*Calloselasma rhodostoma*); 6, interleukin 4-induced Fig-1 protein (*Mus musculus*); 7, LAAO (*Neurospora crassa*); 8, LAAO (*Chlamydomonas reinhardtii*); 9, LAAO (*Synechococcus*); 10 MAO (*Aspergillus niger*); 11, MAO (*Bos taurus*); 12, MAO A (*Homo sapiens*); 13, MAO A (*Rattus norvegicus*); 14, MAO B (*Homo sapiens*); 15, MAO B (*Rattus norvegicus*); 16, MAO (*Oncorhynchus mykiss*); 17, putrescine oxidase (*Micrococcus rubens*); 18, hypothetical MAO (*Mycobacterium tuberculosis*); 19, protein of unknown function (*Synechocystis*); 20, achacin (*Achatina fulica*); 21, aplysianin-A (*Aplysia kurodai*); 22, corticosteroid-binding protein (*Candida albicans*); 23, steroid-binding protein FMS1 (*Saccharomyces cerevisiae*); 24, protein of unknown function (*Caenorhabditis elegans*); 25, polyamine oxidase (*Zea mays*); 26, Trp-2MO (*Agrobacterium rhizogenes*); 27, Trp-2MO (*Agrobacterium tumefaciens*); 28, Trp-2MO (*Agrobacterium vitis*); 29, Trp-2MO (*Erwinia herbicola*); 30, Trp-2MO (*Pseudomonas syringae*); 31, DAAO (*Trigonopsis variabilis*); 32, DAAO (*Fusarium solani*); 33, DAAO (*Rhodotorula gracilis*); 34, DAAO (*Mycobacterium tuberculosis*); 35, DAAO (*Mycobacterium leprae*); 36, DAAO (*Rattus norvegieus*); 37, DAAO (*Mus musculus*); 38, DAAO (*Homo sapiens*).

#	SPECIES	AMINO ACID SEQUENCE	PROTEIN FUNCTION
1	*B. subtilis*	-------MSGLVSASLLKNAG------HRVTILEASGRAGGRVCT	L-AMINO ACID OXIDASE
2	*B. cereus*	IIVVGAGMAGLISASLLKAAGLLKAAGHEVKIFEANNRVGGRIET	L-AMINO ACID OXIDASE ?
3	*C. adamanteus*	VVIVGAGMAGLSAAYVLAGAG------HQVTVLEASERVGGRVRT	L-AMINO ACID OXIDASE
4	*C. atrox*	VVIVGAGMAGLSAAYVLAGAG------HQVTVLEASERVGGRVRT	L-AMINO ACID OXIDASE
5	*C. rhodostoma*	VVIVGAGMAGLSAAYVLAGAG------HQVTVLEASERPGGRVRT	L-AMINO ACID OXIDASE
6	*M. musculus*	VVVVGAGVAGLVAAKMLSDAG--------VTILEADNRIGGRIFT	FIG-1 PROTEIN
7	*N. crassa*	AIAVGAGMSGLMTYLCLTQAGLTQAGMTNVSIIEGGNRLGGRVHT	L-AMINO ACID OXIDASE
8	*C. reinhardtii*	VVVVGGGCGGIYSAYRLL-SGTTLK--PSVCTFEATNRVGGRIFS	L-AMINO ACID OXIDASE
9	*Synechococcus*	VFTLGAGMAGLTAALSLLRRG------HQVTVIEYQNRIGGRLLS	L-AMINO ACID OXIDASE
10	*A. niger*	VIVIGGGYCGLTATRDLTVAG------FKTLLLEARDRIGGRSWS	MONOAMINE OXIDASE
11	*B. taurus*	VVVIGGGISGLSAAKLLAEHE------VNVLVLEARERVGGRTYT	MONOAMINE OXIDASE
12	*H. sapiens*	VVVIGGGISGLSAAKLLTEYG------VSVLVLEARDRVGGRTYT	MONOAMINE OXIDASE A
13	*R. norvegicus*	VGLIGGGISGLAAAKLLSEYK------INVLVLEARDRVGGRTYT	MONOAMINE OXIDASE A
14	*H. sapiens*	VVVVGGGISGMAAAKLLHDSG------LNVVVLEARDRVGGRTYT	MONOAMINE OXIDASE B
15	*R. norvegicus*	VIVVGGGISGMAAAKLLHDCG------LSVVVLEARDCVGGRTYT	MONOAMINE OXIDASE B
16	*O. mykiss*	VIVIGGGISGLSAAKLLKEKG------LSPVVLEARDRVGGRTFT	MONOAMINE OXIDASE
17	*M. rubens*	VVVVGAGPAGLMAARTLVAAG------RTVAVLEARDRVGGRTWS	PUTRESCINE OXIDASE
18	*M. tuberculosis*	VVVVGAGFAGLAAARELTRQG------HEVLVFEGRDRVGGRFTS	MONOAMINE OXIDASE ? (AL021646)
19	*Synechocystis*	V---GSGLSGLIAARNLSRVN------YSVLVIEAQERLGGRMYG	UNKNOWN (D64000)
20	*A. fulica*	VAVVGAGPSGTYSAYKLRNKG------QTVELFEYSNRIGGRLFT	ACHACIN
21	*A. kurodai*	IAIVGAGPSGAYSAYKMRHSG------KDVGLFEYCNRVGGRLYT	APLYSIANIN-A
22	*C. albicans*	VLIIGAGVSGLKAAETILSKSFLTG--DDVLVVEAQNRIGGRLKT	CORTICOSTEROID BIND. PROT.
23	*S. cerevisiae*	VIIIGAGIAGLKAASTLHQNGI-----QDCLVLEARDRVGGRLQT	STEROID-BINDING PROT. FMS1
24	*C. elegans*	IAIVGAGISGLSTARRLIELGI-----DDFDIYEGLDRIGGRIHA	UNKNOWN (AF067936)
25	*Z. mays*	VIVVGAGMSGISAAKRLSEAGI-----TDLLILEATDHIGGRMHK	POLYAMINE OXIDASE
26	*A. rhizogenes*	VAIIGAGFSGLVAASELLHAGV-----DDVTVYEASDRLGGKLWS	TRYPTOPHAN 2-MONOOXYGENASE
27	*A. tumefaciens*	VAVIGAGISGLVVASELLHAGV-----DDVTIYEAGDRVGGKLWS	TRYPTOPHAN 2-MONOOXYGENASE
28	*A. vitis*	VAVIGAGISGLVSATLLLRNGI-----DDVTIFEAKNVVGGRAHT	TRYPTOPHAN 2-MONOOXYGENASE
29	*E. herbicola*	VAIIGAGISGLIAATELLRAGV-----RDITLFEARDRLGGRAWS	TRYPTOPHAN 2-MONOOXYGENASE
30	*P. syringae*	VAIVGAGISGLVAATELLRAGV-----KDVVLYESRDRIGGRVWS	TRYPTOPHAN 2-MONOOXYGENASE
	Consensus	VhVVGAG-SGL-AA--L---G-------V-VhEA-DRVGGRh-T	

Figure 2. Sequence alignment of an extended dinucleotide binding fold in L-amino acid oxidases, monoamine oxidases and tryptophan 2-monooxygenases. The species names are shown on the left and the function of the protein on the right. The alignment was performed using ClustalX (6). The consensus sequence is shown below (s=small aa; h=hydrophobic aa; p=polar aa). Amino acids found in more than 80% of the sequences are underlined.

Moreover, inspection of the phylogenetic tree (Figure 1) reveals that ophidian LAAOs (#3-5) share no similarity with the families of mammalian monoamine oxidases (MAO, #11-16), D-amino acid oxidases (DAAO, #31-38) and tryptophan 2-monooxygenases (Trp-2MO, #26-30). Despite this dissimilarity on a whole sequence level, the FAD binding fold shows a very high degree of conservation as shown in Figure 2. The main feature of this fold is a highly conserved R-h-G-G-R-h-x-(T or S) motif extending beyond the dinucleotide binding fold described by Wierenga (2).This conserved region is found in all LAAOs, MAOs, Trp-2MOs, putrescine and polyamine oxidase, as well as a number of other less well characterized proteins (see Figure 1 and 2). These proteins appear to form a subfamily in the FAD-containing superfamily described recently (3). Therefore we like to suggest that the presence of this motif is indicative of a FAD-binding site. It is, however, not required for FAD binding as, for example, DAAOs do not have the conserved region. In this context it is also noteworthy that the second conserved arginine and the following, highly conserved, threonine (see Figure 2) were found to be essential for covalent incorporation of FAD in human MAO B (4).

Sequence alignment of *C. adamanteus*, *C. atrox* and *C. rhodostoma* LAAO revealed one highly conserved possible N-glycosylation site at the asparagine in position 361. In addition to this conserved site, *C. rhodostoma* LAAO has a second possible glycosylation site at asparagine 172. Studies on the glycosylation of *C. rhodostoma* LAAO suggest glycosylation at only one site which is tentatively assigned to the conserved asparagine at position 361.

References

1. Raibekas, A. A. and Massey, V. (1998): Biochem. Biophys. Res. Commun. **248,** 476-478.

2. Wierenga, R. K., Terpstra, P. and Hol, W. G. J. (1986): J. Mol. Biol. **187**, 101-107.

3. Dailey, T. A. and Dailey, H. A. (1998): J. Biol. Chem. **273**, 13658-13662.

4. Kirksey, T. J., Kwan, S.-W., and Abell, C. W. (1998): Biochemistry **37**, 12360-12366.

5. Bandelt, H.-J. and Dress, A. W. M. (1992): Mol. Phylogenetics Evo. **1**, 242-252.

6. Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994): Nucleic Acid Res. **22**, 4673-4680.