Molecular Phylogenetics and Evolution 139 (2019) 106571

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Evolution and functional classification of mammalian copper amine oxidases

Leonor Lopes de Carvalho[a], Eva Bligt-Lindén[a], Arunachalam Ramaiah[a,b,1], Mark S. Johnson[a], Tiina A. Salminen[a,*]

[a] Structural Bioinformatics Laboratory, Biochemistry, Faculty of Science and Engineering, Åbo Akademi University, Turku, Finland
[b] Sri Paramakalyani Centre for Environmental Sciences, Manonmaniam Sundaranar University, Alwarkurichi, Tamil Nadu 627412, India

ARTICLE INFO

Keywords:
Phylogenetics
Copper amine oxidase
Active site motif
Functional classification
Three-dimensional structure

ABSTRACT

Mammalian copper-containing amine oxidases (CAOs), encoded by four genes (*AOC1-4*) and catalyzing the oxidation of primary amines to aldehydes, regulate many biological processes and are linked to various diseases including inflammatory conditions and histamine intolerance. Despite the known differences in their substrate preferences, CAOs are currently classified based on their preference for either primary monoamines (EC 1.4.3.21) or diamines (EC 1.4.3.22). Here, we present the first extensive phylogenetic study of CAOs that, combined with structural analyses of the CAO active sites, provides in-depth knowledge of their relationships and guidelines for classification of mammalian CAOs into AOC1-4 sub-families. The phylogenetic results show that CAOs can be classified based on two residues, X1 and X2, from the active site motif: T/S-X1-X2-N-Y-D. Residue X2 discriminates among the AOC1 (Tyr), AOC2 (Gly), and AOC3/AOC4 (Leu) proteins, while residue X1 further classifies the AOC3 (Leu) and AOC4 (Met) proteins that so far have been poorly identified and annotated. Residues X1 and X2 conserved within each sub-family and located in the catalytic site seem to be the key determinants for the unique substrate preference of each CAO sub-family. Furthermore, one residue located at 10 Å distance from the catalytic site is different between the sub-families but highly conserved within each sub-family (Asp in AOC1, His in AOC2, Thr in AOC3 and Asn in AOC4) and likely contributes to substrate selectivity. Altogether, our results will benefit the design of new sub-family specific inhibitors and the design of *in vitro* tests to detect individual CAO levels for diagnostic purposes.

1. Introduction

Copper-containing amine oxidases (CAOs) are a large class of enzymes that catalyze the oxidation of primary amines to the corresponding aldehyde associated with the reduction of molecular oxygen to hydrogen peroxide (Reviewed in Klema and Wilmot, 2012). Because of their wide distribution in mammalian tissues, the CAO family members are involved in a range of complex processes, and changes in their expression and catalytic activity have been connected with several diseases e.g. histamine intolerance, rheumatoid arthritis, psoriasis, atopic eczema, multiple sclerosis, diabetes, respiratory diseases and cancer (Boomsma et al., 2003; Maintz and Novak, 2007; Dunkel et al., 2008; Jalkanen and Salmi, 2017).

Mammals typically have either three (*AOC1-3*) or four (*AOC1-4*) genes encoding CAOs (Schwelberger, 2006). Schwelberger (2010) identified *AOC1-4* genes in the genomes of cows, horses, dogs, chimpanzees and macaques. The complete *AOC4* gene was also found in the human genome but it is disrupted by a stop codon and does not form a functional protein. The genomes of rat and mouse only have a partial *AOC4* gene (Schwelberger, 2006). The *AOC1* gene has five exons and it is not linked with the other AOC genes, whereas the *AOC2*, *AOC3* and *AOC4* genes are tandemly arranged on the same chromosome and each gene has four exons (Schwelberger, 2006).

The *AOC1* gene encodes diamine oxidase (DAO; hereafter called AOC1), which is expressed e.g. in placenta, kidney, gut and lung (Elmore et al., 2002). Decreased levels of human AOC1 (hAOC1) are correlated with histamine intolerance (Maintz and Novak, 2007). In accordance with the substrate preference for the aromatic diamines histamine and 1-methylhistamine (Elmore et al., 2002), hAOC1 has been suggested to have *in vivo* functions in cell proliferation, inflammation, allergic response and ischemia (McGrath et al., 2009). There is also evidence that some pharmaceuticals might interfere with the normal function of hAOC1 and, thus, cause harmful side effects (McGrath et al., 2009).

The *AOC2* gene encodes retina-specific amine oxidase (RAO; hereafter called AOC2), which was originally identified in the ganglion cell

* Corresponding author.
E-mail address: tiina.salminen@abo.fi (T.A. Salminen).
[1] Present address: Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, United States.

https://doi.org/10.1016/j.ympev.2019.106571
Received 30 January 2019; Received in revised form 5 July 2019; Accepted 23 July 2019
Available online 24 July 2019

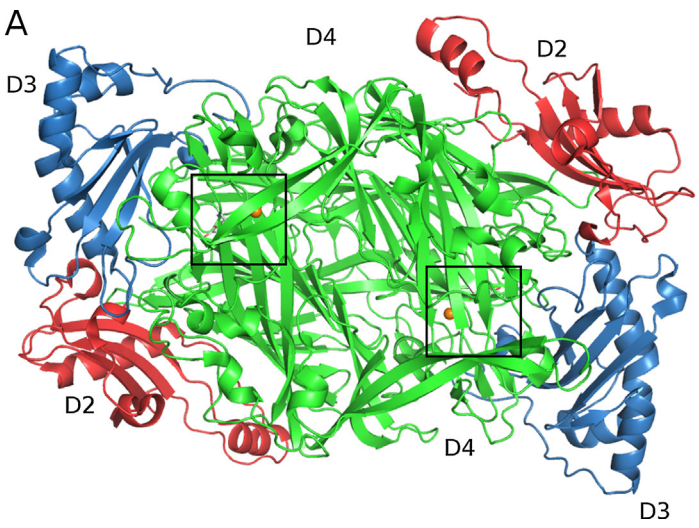

Fig. 1. A, Overall view of the CAO homodimer (PDB ID 4BTX). The D2 (red) and D3 (blue) domains surround the central catalytic D4 (green) domains, which form extensive interactions in the dimer. The catalytic sites in each monomer are boxed. B, Close up view of the catalytic site showing TPQ (off-copper), the catalytic aspartate (D) and the three histidines (H) coordinating the copper ion (orange sphere). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

layer of the retina and has an N-terminal transmembrane segment. Its physiological role is uncertain (Imamura et al., 1997; Kaitaniemi et al., 2009) but human AOC2 (hAOC2) has been suggested to have a role in hereditary retinal diseases (Imamura et al., 1997, 1998). hAOC2 oxidizes aromatic monoamines like *p*-tyramine, tryptamine and 2-phenylethylamine (Kaitaniemi et al., 2009).

The *AOC3* gene encodes vascular adhesion protein-1 (VAP-1; hereafter called AOC3), which is primarily expressed on the endothelial cell surface (Salmi et al., 1993; Smith et al., 1998), but also in smooth muscle cells and adipocytes (Salmi et al., 1993; Zorzano et al., 2003). In addition to its amine oxidase activity, human AOC3 (hAOC3) functions as an adhesin (Smith et al., 1998; Salmi and Jalkanen, 2001). Upon inflammation, the expression of hAOC3 is upregulated (Salmi et al., 1993; Jaakkola et al., 2000) and it has an essential role in extravasation of leukocytes from the blood into tissues (Tohka et al., 2001; Lalor et al., 2002; Bonder et al., 2005; Merinen et al., 2005; Jalkanen et al., 2007). The physiological substrates of hAOC3 include methylamine and aminoacetone, it exhibits high activity towards the exogenous aromatic amine, benzylamine, and it is weakly active with 2-phenylethylamine (Smith et al., 1998; Kaitaniemi et al., 2009). Additionally, the Siglec-9 and Siglec-10 proteins on leukocyte surfaces have been identified as physiological ligands for the adhesive function of hAOC3 (Kivi et al., 2009; Aalto et al., 2011; Salmi and Jalkanen, 2014). The level of hAOC3 activity is low in healthy humans but it is elevated upon inflammation due to hAOC3 translocation to endothelial cell surface. Thus, a Siglec-9 derived peptide binding to hAOC3 can be used as a tracer in positron emission tomography (PET) to detect sites of inflammation *in vivo* (Aalto et al., 2011).

The *AOC4* gene encodes a soluble plasma amine oxidase (SAO; hereafter called AOC4) (Schwelberger, 2006). Bovine AOC4 (bAOC4) prefers the physiological polyamines spermidine and spermine (Di Paolo et al., 2003; Bonaiuto et al., 2010). It is primarily expressed in the liver from where it is secreted into the bloodstream, constituting the major part of the serum amine oxidase activity in some mammalian species (Schwelberger, 2006). Recently, delivery of bAOC4 to cancer cells has been reported as a method for *in situ* killing of cancer cells (Reviewed in Agostinelli et al., 2010). Cancer cells have elevated levels of endogeneous polyamines, which are catalyzed to cytotoxic reaction products that selectively reduce tumor growth upon the targeted delivery of bAOC4.

Some mammals e.g. humans, mice and rats, do not have an *AOC4* gene product. Therefore, the detected CAO activity in the blood of these three species is derived from AOC3, which is proteolytically released from the endothelial cell surface (Kurkijärvi et al., 1998; Abella et al., 2004; Stolen et al., 2004; Jalkanen and Salmi, 2008). In humans, elevated levels of plasma AOC3 activity are associated with diabetes mellitus, congestive heart failure and liver cirrhosis (Boomsma et al., 2005). In adipocytes, AOC3 is involved in glucose metabolism (Boomsma et al., 2005) and its activity in these cells has been shown to be down-regulated by hypoxia (Repessé et al., 2015). Besides having a role in various diseases and being suitable as drug-targets and therapeutics, CAOs are also effective biomarkers for several diseases including histamine intolerance, cancers and inflammatory diseases (Agostinelli et al., 2010; Mušič et al., 2013; Pannecoeck et al., 2015).

The mammalian CAOs have a highly conserved T/S-X-X-N-Y-D motif in the active site (Mu et al., 1994; Salminen et al., 1998) where the tyrosine is post-translationally modified to 2,4,5-trihydroxyphenylalanine quinone (topaquinone; TPQ) (Janes et al., 1990). To date, the crystal structures of hAOC1 (McGrath et al., 2009), hAOC3 (Airenne et al., 2005; Jakobsson et al., 2005; Elovaara et al., 2011; Bligt-Lindén et al., 2013), and bAOC4 (Lunelli et al., 2005; Holt et al., 2008) are known but no X-ray structure for AOC2 has been solved. Based on the 68% sequence identity between hAOC2 and hAOC3, the overall structure of hAOC2 is predicted to be very similar to that of hAOC3 (Kaitaniemi et al., 2009). The CAOs as a family are all tightly bound homodimers with a similar three-dimensional (3D) fold consisting of the D2 and D3 domains, and the catalytic D4 domain (Fig. 1). The buried active site is highly conserved and has several common features: the TPQ cofactor and the conserved catalytic aspartic acid residue are both involved in the catalytic reaction, and the three conserved histidines coordinate the copper ion involved in TPQ biogenesis. The catalytic reaction mechanism of the CAOs is divided into two half-reactions and has been extensively studied in bacteria and yeast (Reviewed in Dawkes, 2001; Klema and Wilmot, 2012; Shepard and Dooley, 2015). In the reductive half-reaction, the amine substrate binds to TPQ of an oxidized enzyme forming a substrate Schiff base, and the α-carbon proton of the substrate abstracted by the catalytic base to form the product Schiff base, which is hydrolyzed to release the aldehyde product. In the oxidative half-reaction, the reduced TPQ is re-oxidized by molecular oxygen with the release of ammonia and hydrogen peroxide.

CAOs are currently sub-classified by the Enzyme Commission (EC) based on their substrate preference for either primary monoamines (EC 1.4.3.21) or diamine (EC 1.4.3.22). Due to the lack of an appropriate classification to distinguish between AOC3 and AOC4, the majority of AOC4 proteins are currently annotated as “membrane primary amine oxidase-like” proteins in the sequence databases. In the earlier literature, bAOC4 and hAOC3 were often regarded as the same protein, although they are products of different genes and present clear differences e.g. in the substrate preference and enzyme kinetics (Holt et al.,

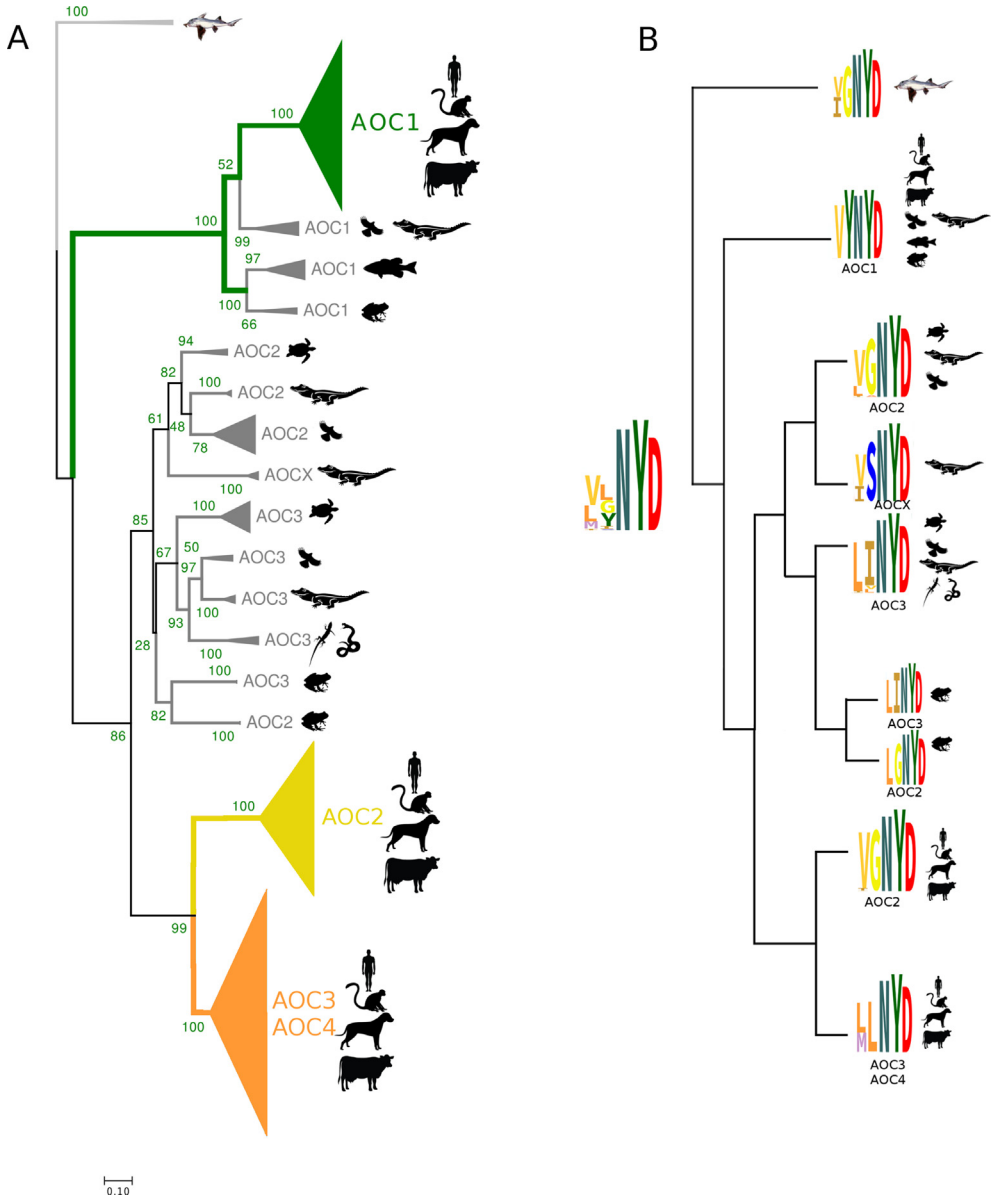

Fig. 2. (A) ML phylogenetic tree of vertebrate CAOs. The branches are scaled to show evolutionary distance and the bootstrap values are shown for the main branches (green). The cartilaginous fishes form an out-group in the tree. (B) The HMM motif logos for the main branches of the phylogenetic tree. Pictures of the representative organisms within each branch are shown in black. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2008; Bonaiuto et al., 2010). Similarly, hAOC2 and hAOC3 have clear differences in the substrate preference and respectively prefer aromatic and aliphatic amines (Kaitaniemi et al., 2009). Despite the reported differences of the EC 1.4.3.21 monoamine oxidases, there are currently no guidelines to further classify them into AOC2-4 sub-families.

In this study, we have combined a phylogenetic study of vertebrate CAOs with analyses of the variable X1 and X2 residues from the T/S-X1-X2-N-Y-D active site motif and linked the results to known substrate preferences of hAOC1-3 and bAOC4 in order to derive general rules that aid in the classification of the CAOs to distinct sub-families of AOC1, AOC2, AOC3 and AOC4. Most notably, our comprehensive analysis allowed us to distinguish the poorly annotated AOC4 proteins from the highly similar AOC3s. By associating similarities and differences in the active sites of CAOs with the unique substrate preference of each CAO sub-family, we were able to identify residues that contribute to the substrate selectivity. The results from our work will benefit the design of new AOC3 sub-family specific inhibitors, which would not bind to the other CAOs, and the design of *in vitro* tests to detect specific CAO levels for diagnostic purposes.

2. Materials and methods

2.1. Sequence data

The sequences for human AOC1 (hAOC1; UniProt ID: P19801), AOC2 (hAOC2; UniProt ID: O75106), AOC3 (hAOC3; UniProt ID: Q16853) and bovine AOC4 (bAOC4; UniProt ID: Q29437) proteins (ranging from 756 to 763 residues) were obtained from UniProt Knowledgebase (Jain et al., 2009). To identify vertebrate CAO sequences, the hAOC1, hAOC2, hAOC3 and bAOC4 sequences were used as a query to search vertebrate genomes with protein BLAST (Altschul et al., 1990) at the National Center for Biotechnology Information (NCBI) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Only the full-length sequences and the sequences, which had the conserved tyrosine that is

post-transcriptionally modified to TPQ, the catalytic aspartic acid and the three histidines that coordinate the copper ion, were used in the analysis. Gene information for the selected protein sequences was collected from NCBI gene database (https://www.ncbi.nlm.nih.gov/gene/).

As the degree of conservation is higher at the 3D structural level than at the sequence level (Illergård et al., 2009), we built a structure-based alignment of hAOC1 (PDB: 3HI7) (McGrath et al., 2009), hAOC3 (PDB: 4BTX) (Bligt-Lindén et al., 2013), bAOC4 (PDB: 1TU5) (Lunelli et al., 2005) 3D structures in BODIL (Lehtonen et al., 2004). This alignment was then used as a starting point for aligning the sequences from the BLAST search to ensure the correct positioning of gaps and the conservation of structural features. All sequence alignments were done using the MALIGN (Johnson et al., 1996) in BODIL (Lehtonen et al., 2004) by first creating a separate multiple sequence alignments (MSAs) for the sequences retrieved by each BLAST search. These alignments were then combined in BODIL to generate a MSA with all studied sequences. The MSA was analyzed with special focus on the residue near the active site and on other functionally important residues cited in the literature to confirm that they are conserved in the final MSA. The Hidden Markov Model (HMM) motifs were generated with Skylign (Wheeler et al., 2014).

Signal peptide predictions were performed in the SignalIP 4.1 server (Petersen et al., 2011) and transmembrane domain predictions were performed in the TMHMM 2.0 server (http://www.cbs.dtu.dk/services/TMHMM-2.0/) for all AOC3 and AOC4 sequences.

2.2. Phylogenetic reconstruction

Before performing the phylogenetic analysis, isoforms with the same active site motif were removed from the alignments previously generated in BODIL to avoid redundancies in the trees and to include only one representative for each sequence of each species in the study. The model that best fits the data was chosen and the phylogenetic reconstructions were performed in MEGA-CC 7 (Kumar et al., 2012, 2016). Phylogenetic tree of vertebrate CAOs were estimated using Maximum Likelihood (ML) and Neighbour-joining (NJ) (Saitou and Nei, 1987) methods. The NJ method is based on distance matrix calculated from the MSA, whereas the ML method uses a random tree as a starting point, which is then modified until it reaches the highest likelihood of producing the alignment that was used. The NL trees were generated with the Jones-Taylor-Thornton (JTT) (Jones et al., 1992) distance matrix.

To generate the ML trees, the best fitting model for each alignment was assessed using MEGA-CC 7 (Kumar et al., 2012, 2016) and chosen based on the Bayesian Information Criteria. The Poisson model (Zuckerkandl and Pauling, 1965) with gamma distributions with invariant sites (i.e. P + G + I) was the best model for the vertebrate CAO alignment. Since the ML and NJ phylogenetic trees were congruent, we used the ML tree for further analysis (Fig. 2A; NJ in Supplementary Fig. S1). JTT with gamma distributions (JTT + G) was chosen as the best model for the mammalian AOC1 and AOC2 alignments while the JTT model with gamma distributions and invariant sites (JTT + G + I) fitted best for the alignment of mammalian AOC3 and AOC4 sequences. The areas with gaps were excluded from the phylogenetic analysis. Statistical support for the topology of the tree was inferred from using the bootstrap method (Felsenstein, 1985) with 500 replicates. In general, bootstrap values of 70–100% indicate significant support for a branch (Soltis and Soltis, 2003b). To generate the 3D multivariate plots, the distance matrices from mammalian MSAs were generated in MEGA-CC 7 (Kumar et al., 2012, 2016) using JTT substitution matrix. Thereafter, the JTT distance data was supplied to a C-program to perform a multivariate analysis (PCA, MS Johnson, unpublished). The program displays coordinates for each sequence and their locations such that the variance among the data is a maximum, and projections for various numbers of dimensions are possible. The program returns the three most informative dimensions as a pseudo-PDB coordinate file, which can be visualized in BODIL (Lehtonen et al., 2004) as multivariate 3D plots.

We first generated a phylogenetic tree containing all vertebrate CAO sequences to aid a rough classification of the mammalian CAOs into the AOC1-4 sub-families. In the resulting tree, we named the proteins as AOC1 or AOC2 based on the similarity of their active site motif to hAOC1 and hAOC2, respectively. The proteins with an active site motif similar to either hAOC3 or bAOC4 were named as AOC3/4. Thereafter, individual phylogenetic trees were constructed for the *AOC1* and *AOC2* gene products and a joint tree of the *AOC3/4* gene products that grouped on the same clade within the phylogenetic tree of vertebrate CAOs.

With an aim to produce phylogenetic trees with improved bootstrap values, we also constructed the phylogenetic trees for the conserved D4 domain using the nucleotide sequences corresponding to the proteins listed in the Supplementary Tables S3 and S4. The nucleotide sequences that fulfilled the same functional criteria as we applied for protein sequences were aligned with ClustalW in MEGA-CC 7 (Kumar et al., 2012, 2016). The ML trees were calculated similarly as previously (the best fitting model for each alignment listed in the Supplementary Table S5).

2.3. Homology modeling

The original structure-based alignment was used as a starting point to generate a model for hAOC2 and the crystal structure of hAOC3 (PDB: 4BTX) (Bligt-Lindén et al., 2013) was used as a template. Ten models were generated using MODELLER (Sali and Blundell, 1993). From the generated models, the one with the lowest value of the MODELLER objective function was chosen for analysis. All structures were compared by superimposition using VERTAA in BODIL (Lehtonen et al., 2004) and quality of the model was evaluated using PROCHECK (Laskowski et al., 1993) and QMEAN (Benkert et al., 2009). Visual analysis of the model and crystal structures was performed in Pymol (The PyMOL Molecular Graphics System, Version 1.23, Schrödinger, LLC.).

3. Results and discussion

3.1. Evolution of vertebrate CAOs

In this study, we used a large number of sequences from a variety of vertebrate species, which gives us the benefit of obtaining a more in-depth knowledge of the phylogenetic relationship of the mammalian and non-mammalian CAO sequences. No extensive phylogenetics analysis of the CAO family has been performed earlier, since Lai and co-workers analyzed only ten mammalian CAO sequences (Lai et al., 2012) and Schwelberger (2006, 2010) solely focused on the porcine genome. To collect sequences homologous to CAOs for the phylogenetic tree of vertebrate CAOs, we performed BLAST searches using hAOC1, hAOC2, hAOC3 and bAOC4 as queries against vertebrate genomes. The sequences without (1) the tyrosine post-transcriptionally modified to TPQ, (2) the catalytic aspartate or (3) the three histidines coordinating the copper ion (Fig. 1) were excluded to ensure that the selected sequences are catalytically active. Isoforms with identical active site motifs were not taken into account as highly similar sequences induce redundancy in the trees. The final dataset for inferring the phylogenetic relationships of the collected vertebrate CAOs consists of 369 sequences originating from 143 vertebrate species and includes 283 sequences from 90 mammalian species. The collected sequences share sequence identities ranging from 29.9% to 99.9% (Supplementary Tables S1 and S2). The AOC1, AOC2 and AOC3 proteins were found in Amphibians and Amniotes but the AOC4 proteins were exclusively found in placental mammals. Additionally, bony fishes had AOC1s and Cartilaginous fishes (out-group) had only AOC2s.

We used the CAO sequences from cartilaginous fishes *Rhincodon*

typus (whale shark) and *Callorhinchus milii* (elephant shark) as an outgroup in the phylogenetic analysis. *C. milii* diverged from the bony vertebrates ~ 450 million years ago and has the slowest evolving genome of all known vertebrates, which makes it a good model for inferring the state of ancestral genomes (Venkatesh et al., 2014). Of the ancestral CAO proteins, CAO from *C. milii* annotated as "retina-specific amine oxidase-like" also shares the highest sequence identity with hAOC2 (43.8%) and has almost the same identity to hAOC3 (43.4%) whereas its identity to hAOC1 (40.9%) is somewhat lower. Similarly, *R. typys* CAO has the highest identities to hAOC2 and hAOC3 (42.5% and 41.8%, respectively) and somewhat lower identity to hAOC1 (39.2%). Since the phylogenetic trees inferred using MEGA-CC 7 (Kumar et al., 2012, 2016) with ML and NJ methods were congruent, we used the ML tree for further analysis (Fig. 2A; NJ in Supplementary Fig. S1).

The phylogenetic tree of vertebrate CAO is divided into two major branches of the AOC1 and AOC2-4 proteins (Fig. 2A). The distances of the three branches, which are scaled to the relative evolutionary distance, show that the AOC2 and AOC3/AOC4 proteins are closely related but AOC1s have a longer evolutionary distance to them. The analysis of the gene information from the NCBI gene database (http://www.ncbi.nlm.nih.gov/gene/) revealed that the *AOC2-4* genes in the studied organisms are located in the same chromosome, whereas the AOC1 gene is located in another chromosome, which is in accordance with previous studies (Schwelberger, 2006, 2010).

In the phylogenetic tree of vertebrate CAOs, the mammalian and non-mammalian AOC1 proteins share the same clade with a strong bootstrap support of 100% (Fig. 2A). The phylogenetic tree suggests a similar substrate preference for AOC1 from fish, amphibians, reptiles, birds and mammals. Accordingly, Almeida and Beaven (1981) have reported AOC1 activity towards histamine in the brain of lower vertebrates.

Unlike for AOC1s, the non-mammalian AOC2-4s segregate from the mammalian proteins with a bootstrap value of 86%. The non-mammalian AOC2 and AOC3 proteins form their own sub-branches with a bootstrap support of 85% for the node, whereas mammalian AOC2-4s are further divided into the AOC2 and AOC3/AOC4 sub-branches with high bootstrap values (Fig. 2A). Within the non-mammalian AOC2 clade, the crocodilian CAO-like proteins (labeled AOCX) are found in a separate clade (48% bootstrap value). The non-mammalian AOC3 branch is further divided with a low bootstrap support of 28% into the major AOC3 branch and the Amphibian branch including both AOC2s and AOC3s. The phylogenetic tree of vertebrate CAOs was also constructed using the nucleotide sequences and shows similar branching topology as the tree based on protein sequences (Supplementary Fig. S6).

3.2. Classification of mammalian CAOs to distinct AOC1-4 sub-families

Our earlier structural and sequence analysis of CAOs indicated that variable residues X1 and X2 of the T/S-X1-X2-N-Y-D active site motif might follow the substrate preference of mammalian CAOs (Salminen et al., 1998; Airenne et al., 2005; Kaitaniemi et al., 2009). This led us to propose that the residues at position X1 and X2 might be useful in distinguishing the sub-families of mammalian CAOs. In the vertebrate CAO tree, the three clades of the mammalian CAOs (Fig. 2A) agree with their known substrate preference: (1) diamines for the AOC1 clade, (2) aromatic monoamines for the AOC2 clade, and (3) aliphatic amines for the AOC3/AOC4 clade. Therefore, we next focused our study at positions X1 and X2 and analyzed the collected vertebrate CAO sequences with a special focus on the mammalian CAOs to find out if these two residues are suitable for categorizing the mammalian CAO sub-families.

When we visualized the vertebrate CAO alignment using the Hidden Markov Model (HMM) motifs (Wheeler et al., 2014) and analyzed the X1-X2-N-Y-D motif in the mammalian CAOs, the results revealed that residue X2 is always a Tyr in AOC1s, a Gly in AOC2s and a Leu in the AOC3/AOC4 proteins (Fig. 2B). Thus, residue X2 coincides with both the substrate preference and the phylogenetic classification of the three CAO sub-families in Mammalia (Fig. 2A). Consistently, all vertebrate AOC1s have a Tyr at position X2 and most of the AOC2s a Gly. The additional CAO-like proteins in crocodilians (labeled AOCX) have a Ser at position X2 indicating that they have a different substrate preference than the other non-mammalian AOC2s. In non-mammalian AOC3s, residue X2 is less conserved than the totally conserved Leu in mammalian AOC3/AOC4 proteins. It is always a hydrophobic residue but more frequently an Ile than a Leu and even a Val in some of the non-mammalian AOC3s.

As mammalian AOC3s and AOC4s both have a Leu at position X2 (Fig. 2B), and they are in the same AOC3/AOC4 clade of the phylogenetic tree (Fig. 2A), their classification needed further inspection. Therefore, we limited the vertebrate CAO data set to mammalian sequences including 85 AOC1, 77 AOC2, 74 AOC3 and 47 AOC4 sequences from 90 different mammalian species (Supplementary Table S1) and created a multiple sequence alignment (MSA). The average sequence identity of AOC1s to AOC2 is 62% and to AOC3/AOC4s 60.2%, whereas the AOC2s and AOC3/AOC4 proteins share a higher average sequence identity of 73.2%. Furthermore, the intra-species sequence identities of AOC3/AOC4 proteins range from a high identity of 84% to an extremely high value of 98%. As an alternative to the phylogenetic reconstruction methods, principal components analysis (PCA) was used to provide a multivariate analysis of the pairwise sequence distance data. The resulted 3D multivariate plot with three clusters representing sequences similar to hAOC1, hAOC2 and hAOC3/bAOC4 (Fig. 3A) is in accordance with the phylogenetic tree of vertebrate CAOs (Fig. 2A). Since position X2 did not differentiate AOC3 and AOC4, we next analyzed residues at position X1. Comparison of hAOC3 and bAOC4 sequences revealed that residue X1 was a Leu in hAOC3 and a Met in bAOC4. Next we analyzed the origin of the AOC3 and AOC4 proteins using the information in the gene and sequence databases and noticed, with few exceptions (see below), that the proteins having a Leu at position X1 originated from a different gene than those with a Met. Furthermore, the fact that a Leu is predominantly found at position X1 in the non-mammalian AOC3s gave further support for our hypothesis that X1 might distinguish mammalian AOC3s from AOC4s. To test our hypothesis, we named the mammalian CAOs with a L-L-N-Y-D motif as AOC3 and those with a M-L-N-Y-D motif as AOC4 and generated a 3D multivariate plot for them (Fig. 3B). In the resulting plot, the AOC3 cluster separates from the AOC4 cluster suggesting that the CAOs with a Leu at position X1 should be classified as AOC3 and those with a Met as AOC4.

Thereafter, we generated the Hidden Markov Model (HMM) motifs for the AOC1-4 subfamilies of mammalian CAOs (Wheeler et al., 2014). The HMM motifs (Fig. 3C) further confirmed that residue X2 of the X1-X2-N-Y-D motif reveals the substrate preference of mammalian CAOs while residue X1 classifies the mammalian AOC3 and AOC4 proteins into distinct sub-families. As a summary, a Tyr at position X2 is consistent with a preference for aromatic diamines (AOC1), a Gly for aromatic monoamines (AOC2) and a Leu for aliphatic amines (AOC3 and AOC4). A Leu at position X1 allows preference for small aliphatic monoamines (AOC3) and a Met for aliphatic polyamines (AOC4), whereas AOC1 and AOC2 that prefer aromatic substrates both have a Val at position X1.

3.3. Phylogenetic analysis of mammalian CAO sub-families

Our results concluded that mammalian CAOs can be classified into sub-families according to the X1 and X2 residues in the active site motif, which is (1) T-V-Y-N-Y-D in AOC1, (2) T/S-V-G-N-Y-D in AOC2s, (3) T-L-L-N-Y-D in AOC3s and (4) T-M-L-N-Y-D in AOC4s (Fig. 3C). To analyze the evolution of the CAO sub-families, we now generated separate trees for the AOC1, AOC2 and AOC3/4 subfamilies characterized by a Tyr, a Gly and a Leu at position X2, respectively. In general, the phylogenetic trees of CAO sub-families are supported by low bootstrap

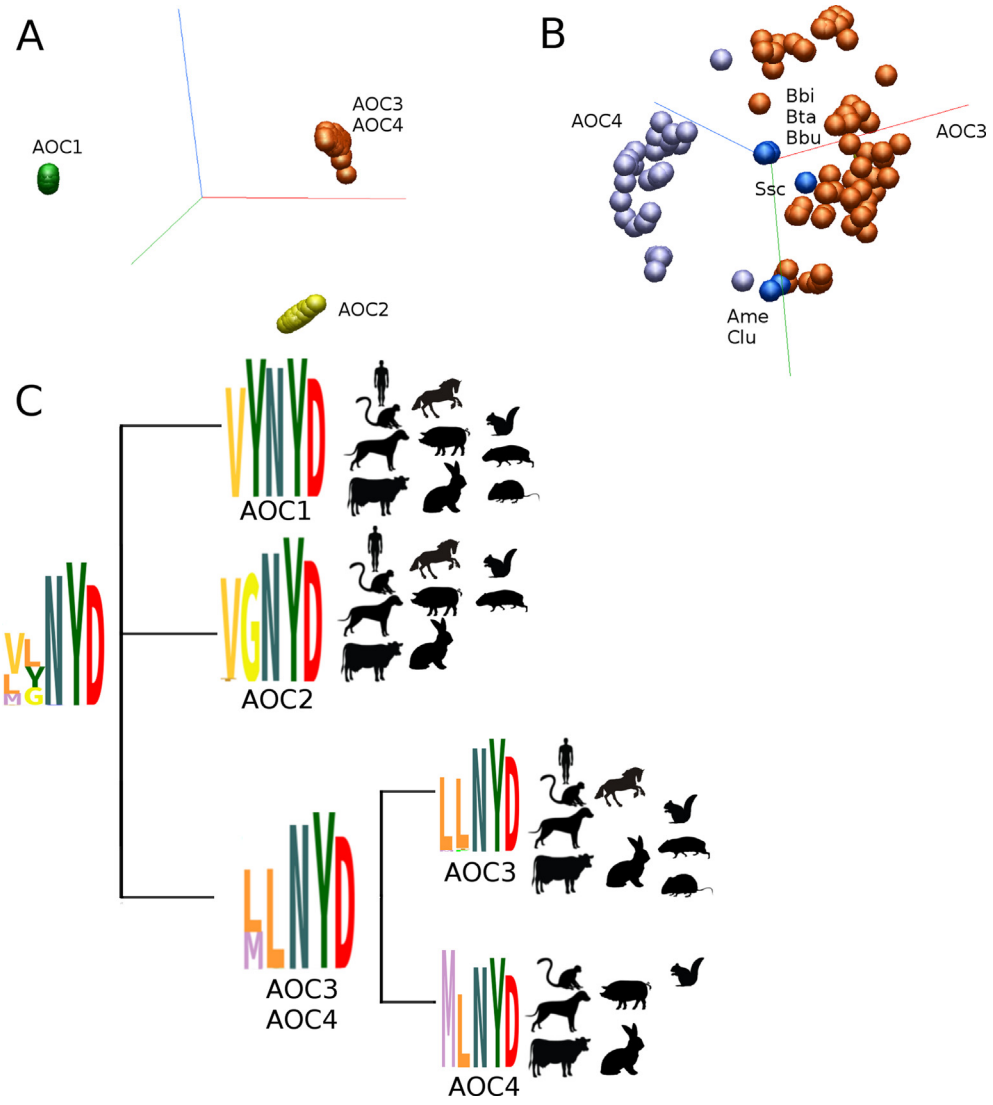

Fig. 3. (A) 3D multivariate plot of mammalian CAO sequences. The plot shows the relationship between the sequences for the three most informative dimensions and supports the grouping of mammalian CAOs in the phylogenetic tree (Fig. 2A). The plot shows three clusters: AOC1s as green spheres, AOC2s as yellow spheres and AOC3/4 as orange spheres. (B) 3D multivariate of mammalian AOC3/4 sequences. The AOC3 proteins are represented as orange spheres and the AOC4 proteins as violet spheres. The plot shows that the AOC3 and AOC4 proteins cluster based on residue X1 with the exception of some outliers (blue), which are labelled. These outliers (blue) are all AOC4 proteins that group closer to the AOC3 cluster than the AOC4 cluster. Bbi – *Bison bison bison*; Bta- *Bos taurus*; Bbu- *Bubalus bubalis;* Ssc – *Sus scrofa*; Ame- *Ailuropoda melanoleuca*; Clu – *Canis lupus familiaris*. (C) HMM motif logos for mammalian CAOs. The HMM motif logos illustrate the classification of mammalian CAOs based on their active site motif. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

values (Figs. 4–6), which reflect the short evolutionary distances of CAOs in each sub-family. When the evolutionary time frame is short, the number of substitutions supporting each node and the dataset as a whole remain small and it is particularly sensitive to random fluctuations in observed mutations and small differences, which leads to low bootstrap values (Soltis and Soltis, 2003a). Indeed, CAOs were one of the few protein families accepted into a study that explicitly searched for proteins related by short evolutionary distances (Lai et al. (2012).

Among AOC1s, the inter-species sequence identity varies between 68.3 and 99.1% (the average identity is 80.0%). The AOC1 tree is divided into two major branches (Fig. 4). One of the branches has the majority of AOC1s and the other one consists of AOC1s from Myomorpha (mouse-like rodents; Fig. 4, green). Thus, the rodent proteins are divided into two distinct branches, despite AOC1s from the same taxonomic order generally group together. The other rodent branch with AOC1s from the Hystricomorpha, Sciuromorpha and Castorimorpha suborders of rodents (Fig. 4, orange) is in the same sub-branch as primate AOC1s. This rodent branch also includes AOC1s from Lagomorpha, Scandentia and Dermoptera. In the phylogenetic tree constructed using the nucleotide sequences of the D4 domain of AOC1s (Supplementary Fig. S7), the rodent proteins are distributed among several branches and *Pteropus alecto*, *Pteropus vampyrus* and *Rousettus aegyptiacus* did not group together with the other Chiroptera but the topology of the tree was otherwise similar.

Interestingly, the tricyclic antidepressant amitriptyline had an opposite effect on *Rattus norvegicus* (rat; Myomorpha) and *Cavia porcellus* (guinea pig; Hystricomorpha) AOC1s (Rajtar and Irman-Florjanc, 2007), which also belong to different branches of the AOC1 tree (Fig. 4). Amitriptyline was reported to act as an inhibitor of *R. norvegicus* AOC1, whereas it increased the activity of *C. porcellus* AOC1 leading to faster histamine inactivation. Thus, the phylogenetic relationship of the two proteins is also reflected at the molecular level in their response to ligand binding.

The tree of the AOC2 proteins has two main branches (Fig. 5), which in turn have diverged into seven sub-branches. The branch length in the seven sub-branches is short and the bootstrap values supporting each branch range from 17% to 98% and the bootstrap values of the tree of AOC2 nucleotide sequences are not significantly higher for the main sub-branches (Supplementary Fig. S8). All the AOC2 proteins are predicted to have a N-terminal transmembrane helix. The prediction could not however be done for *Orincus orca* and *Ictiodomys tridecemlinearus* AOC2s, since they lack part of the N-terminal sequence. The inter-species sequence identity of AOC2s (67.1–100%, average 87.0%) is higher than that of AOC1s and, similar to AOC1, the sequences group in the phylogenetic tree according to the taxonomic rank of species (Fig. 4). Unlike AOC1s, all rodent AOC2s group together in the same sub-tree with *O. cuniculus* AOC2 (Lagomorpha). The AOC2 proteins from rodents and *O. cuniculus* are closely related but differ from Primates; therefore studies conducted on the physiological function of AOC2 from rodents and *O. cuniculus* should not be extended to Primates

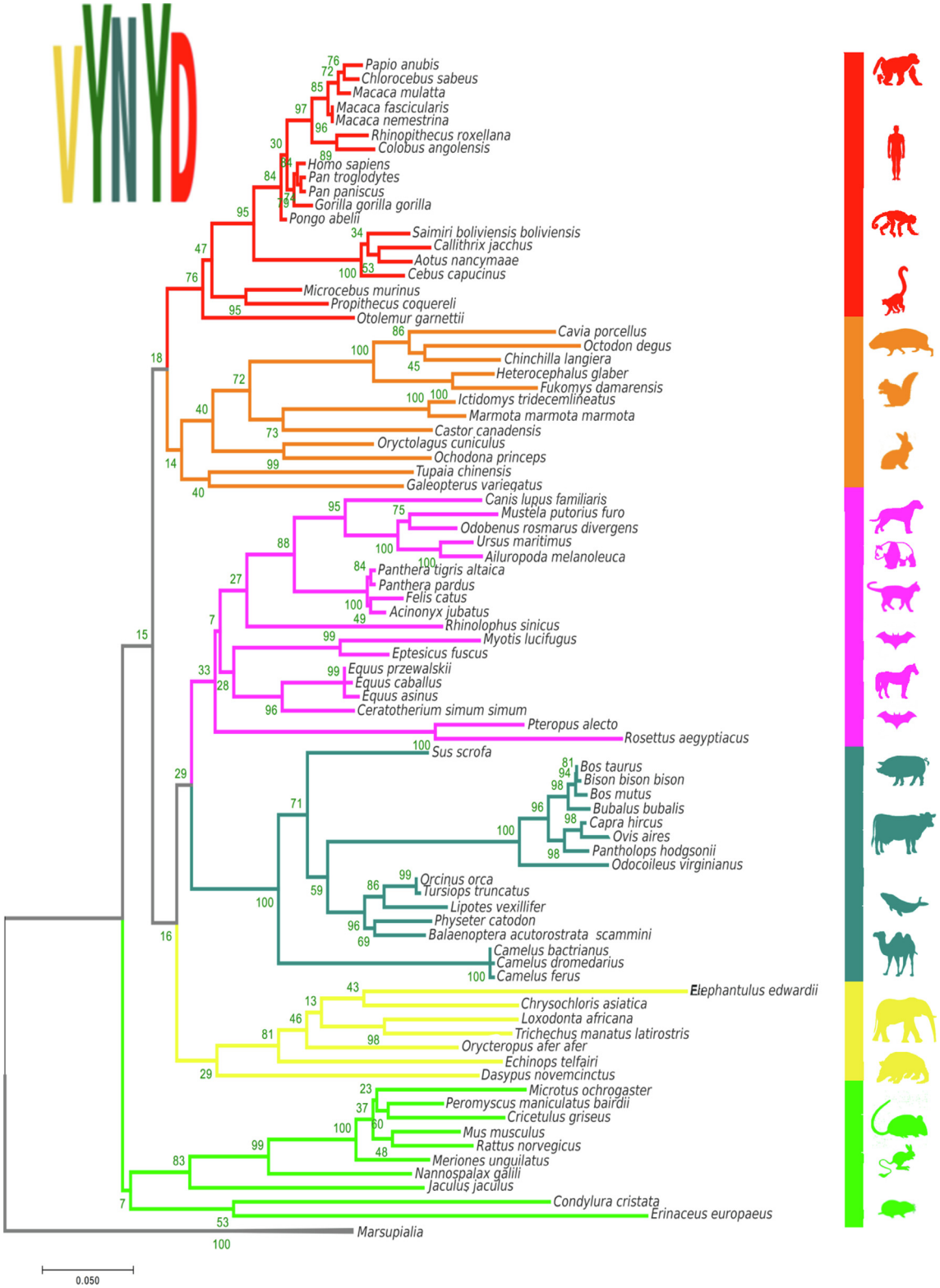

Fig. 4. ML phylogenetic tree of mammalian AOC1s. The HMM logo for AOC1 is shown in the upper left corner and pictures of the representative organisms on the right side of the tree. The branches are scaled to show evolutionary distance and the bootstrap values shown for the main branches (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

without careful comparison of the proteins at sequence and structural level. Of the rodents, *R. norvegicus* lacks AOC2 due to a stop codon that terminates the peptide at 127 amino acids (Zhang et al., 2003). The function of AOC2 is still unknown and further characterization is needed to gain insights into its contribution to the CAO activity in mammals. As rats only have AOC1 and AOC3, detailed and comparable studies of CAO activities in rats and mouse (AOC1-3 proteins) could give hints for the function of AOC2.

The average sequence identity of 81.1% shows that the AOC3 and AOC4 proteins are closely related. The identity of all AOC3 and AOC4

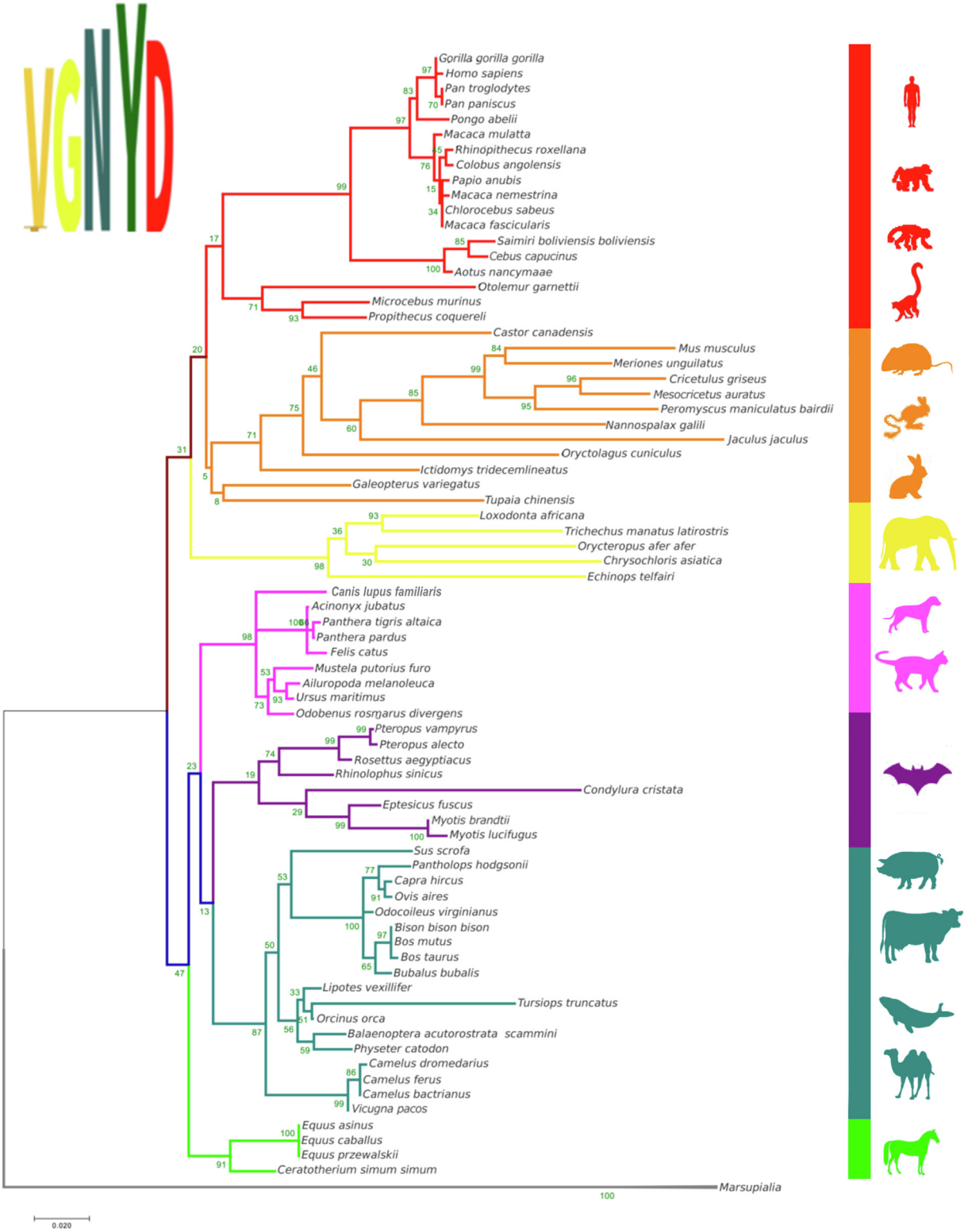

Fig. 5. ML phylogenetic tree of mammalian AOC2s. The HMM logo for AOC2 is shown in the upper left corner and pictures of the representative organisms on the right side of the tree. The branches are scaled to show evolutionary distance and the bootstrap values shown for the main branches (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

proteins varies between 36.7% and 99.9% and is much higher for the catalytic D4 domain (59.5–100%). Consistent with the AOC1 and AOC2 trees (Figs. 4 and 5), the main branches of the AOC3/AOC4 tree (Fig. 6) have diverged over a relatively short evolutionary time frame. Although the main branches have low bootstrap values, the bootstrap values are on average higher within the sub-branches. For example, the bootstrap of the sub-branch containing Artiodactyla and Cetacea proteins is low (9%) but the bootstrap support for the node of the AOC3 proteins from Cetacea is high (99%) and, similarly, the bootstrap support for grouping the AOC3 and AOC4 proteins from Artiodactyla into different clades is

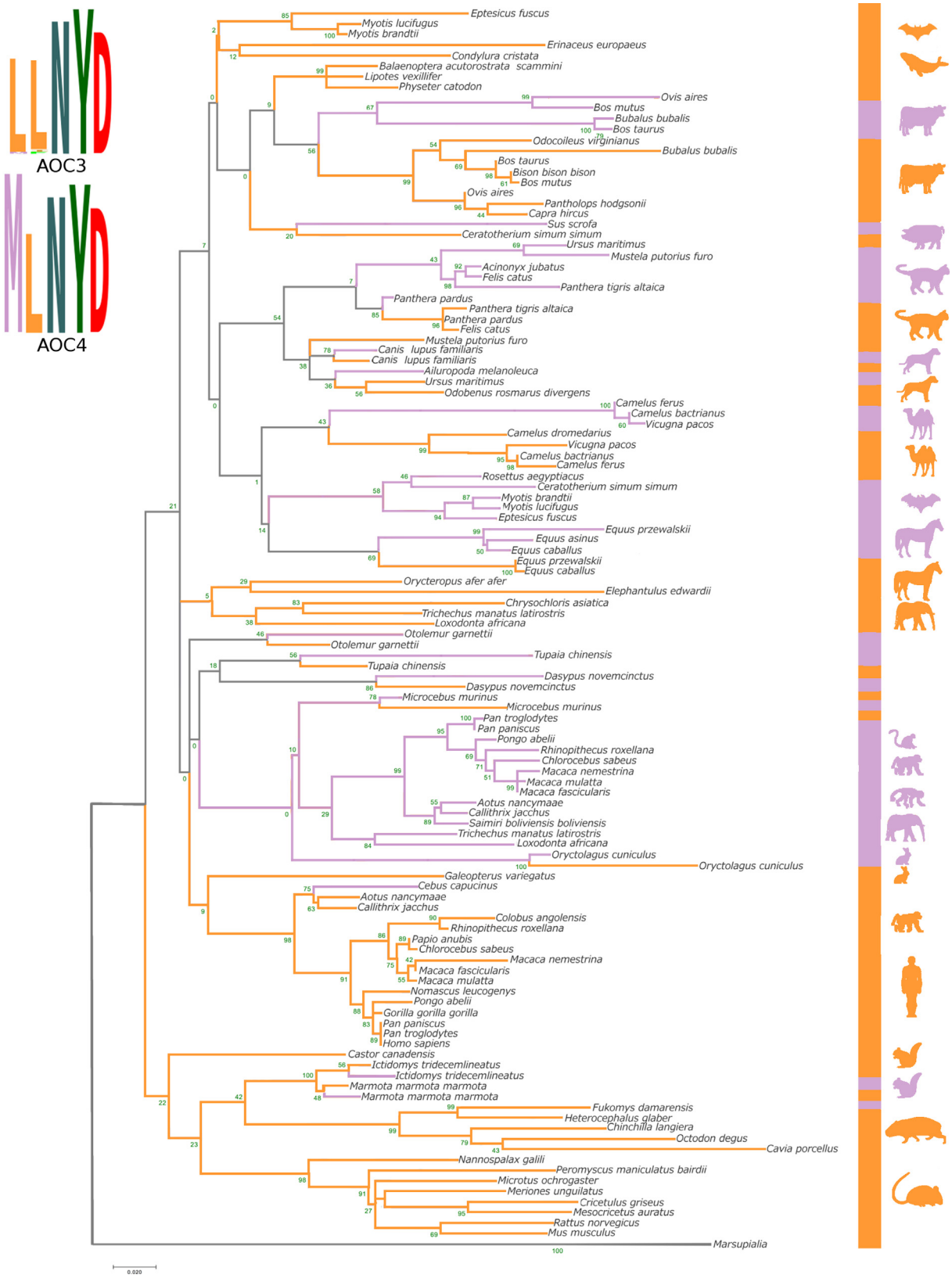

Fig. 6. ML phylogenetic tree of mammalian AOC3s (orange) and AOC4s (violet). The HMM logos for AOC3s and AOC4s are shown in the upper left corner and pictures of the representative organisms on the right side of the tree. The branches are scaled to show evolutionary distance and the bootstrap values shown for the main branches (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

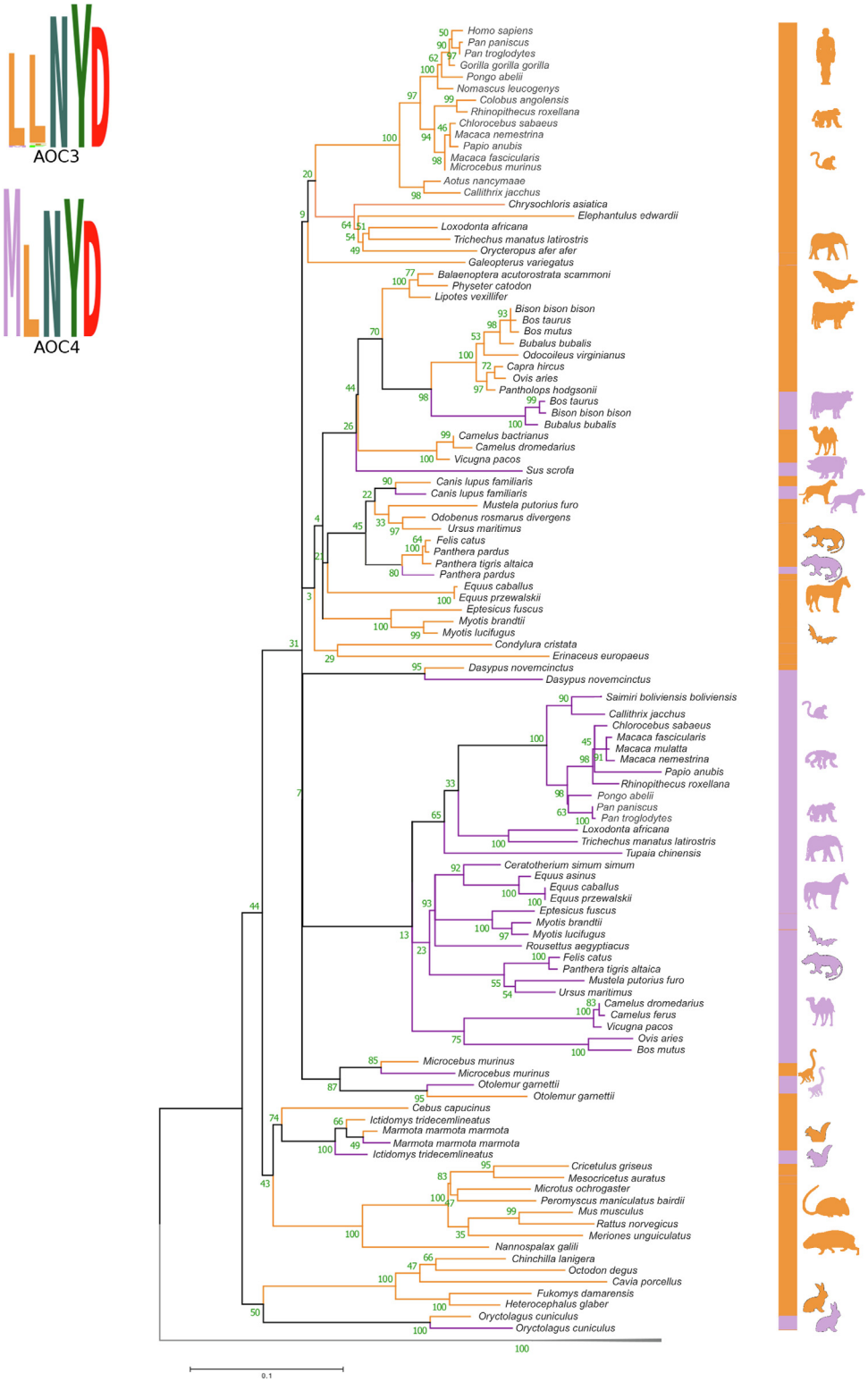

Fig. 7. ML phylogenetic tree of mammalian AOC3s (orange) and AOC4s (violet) constructed using the nucleotide sequences. The HMM logos for AOC3s and AOC4s are shown in the upper left corner and pictures of the representative organisms on the right side of the tree. The branches are scaled to show evolutionary distance and the bootstrap values shown for the main branches (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

99% for the AOC3 clade and 67% for the AOC4 clade. In the case of AOC3/AOC4, the same taxonomic order group together much better in the tree of the nucleotide sequences (Fig. 7) than in the tree of the protein sequences (Fig. 6).

Based on the phylogenetic trees (Figs. 6 and 7), the divergence between AOC3 and AOC4 proteins has occurred in a species-specific manner, probably due to different evolutionary pressure as no consensus among all species was found. In some species, a recent divergence has occurred, whereas in other species the divergence has happened earlier. For example, the AOC3 and AOC4 proteins from monkeys

do not share the same most recent last common ancestor, whereas bovine AOC3 and AOC4 proteins do.

Previously, it has been suggested that those species that have an *AOC4* gene also have a high level of serum CAO (Schwelberger, 2006). Much higher levels of CAO activity ha been measured in the serum of *Ovis aires* (sheep), *Bos taurus* (cow), *O. cuniculus, Capra hircus* (goat), *Canis lupus familiaris* (dog) and *Sus scrofa* (pig) than in the serum of human, *R. norvegicus* (rat), *M. musculus* (mouse) and *C. porcellus* (Guinea pig) (Boomsma et al., 2003; Schwelberger, 2006). The suggested relationship is otherwise in accordance with our results but we could identify only one *AOC3/4* gene in the *C. lupus familiaris* and *S. scrofa* genomes. The two AOC3/4 proteins from *C. lupus familiaris* are special since they turned out to be isoforms from the same *AOC3* gene. They share a sequence identity of 97.8% but one of them has Leu (like AOC3) and the other one Met (like AOC4) at position X1. In the multivariate plot, the AOC4-like protein from *C. lupus familiaris* appears as an outlier clustering together with the AOC3 proteins (Clu in Fig. 3B). Analysis of the gene information of those species that have both AOC3 and AOC4 proteins revealed eight other species in which the AOC3/4 proteins are isoforms from the same gene (Supplementary Table S6). Unlike *C. lupus familiaris* proteins, these other proteins are not found in the same branch of the phylogenetic trees (Figs. 6 and 7). In the *S. scrofa* genome, we could identify only one protein in the AOC3/4 clade and it is also an outlier in the multivariate analysis. The *S. scrofa* protein is currently annotated as an *AOC3* gene product and it consistently clusters with AOC3s (Ssc in Fig. 3B) despite it having Met at position X1 (like AOC4). In accordance with the multivariate plot, *S. scrofa* AOC3/AOC4 groups into the main clade of the AOC3 proteins in the phylogenetic tree (Fig. 7).

In addition to *C. lupus familiaris* and *S. scrofa* proteins, there are some other similar outliers in the multivariate plot (Fig. 3B) but there is no literature on CAO serum activity levels in these organisms. Of these outliers, *Ailuropoda melanoleuca* (panda) protein (Ame in Fig. 3B) should be classified as an AOC4-like protein, since it has Met at position X1 despite it is encoded by *AOC3* and has a predicted transmembrane helix. The additional outliers in the multivariate plot are the AOC4 proteins from *Bos taurus* (cattle), *Bison bison bison* (plains bison) and *Bubalus bubalis* (water buffalo), which group closer to AOC3s than AOC4s (Bta, Bbi and Bbu in Fig. 3B). All these species also have an AOC3 protein, which clusters together with the other AOC3 proteins from Artiodactyla (Figs. 6 and 7). Thus, their *AOC4* gene product might still be differentiating from AOC3.

In accordance with previous research stating that the *AOC3* gene duplication is the origin of the *AOC4* gene (Schwelberger, 2010), we did not find any *AOC4* encoded protein with an AOC3 motif (a Leu at position X1), whereas we found the above mentioned *AOC3* encoded proteins with an AOC4 motif (a Met at position X1). The predominant codon for a Leu is TTG, which turns to a Met coding ATG by a single nucleotide substitution. The need for a new protein with a unique CAO activity in mammals probably results from environmental pressure. For example, the AOC4 proteins from *O. cuniculus* and *S. scrofa* plasma have relatively high activity towards mescaline, which is a hallucinogenic plant-derived amine, suggesting a dietary influence on the selection of the metabolized substrates (Lyles, 1996). Moreover, the fact that extracellular proteins, like AOC3 and AOC4, have a constant need to adapt to new functions and niches and, therefore, evolve fast (Sojo et al., 2016), supports the inferred phylogenetic tree and the relationships between AOC3 and AOC4.

It is interesting that there is variation in the number of *AOC* genes even between closely related species e.g. we found a protein-coding *AOC4* gene in chimpanzees and macaques, whereas human and gorilla have a pseudo gene. In this study, we noticed that all species from Cetacea (Supplementary Table S1) and the majority of rodents have only the AOC3 protein and the *AOC4* gene is missing or non-functional. In fact, *Ictidomys tridecemlineatus* (thirteen-lined ground squirrel) and *Marmota marmota marmota* are the only rodents that have both AOC3 and AOC4.

Furthermore, our results show that it is important to analyze the genomic origin of the encoded protein. We suggest that a CAO with a predicted AOC4 activity (X1 = Met) but originating from *AOC3* should not be named as AOC3 but rather as an AOC4-like protein. The proposed classification system could be used to correctly annotate new mammalian CAO sequences and revise wrongly annotated sequences.

3.4. Active site differences and key residues for substrate preference in mammalian CAOs

Next we extended our study from the phylogenetic analysis to the level of the 3D structure with a particular focus on the catalytic region near TPQ (Fig. 1). We used the following X-ray structures as the representative structures: hAOC1 (Protein Data Bank (PDB) Identification code (ID) 3MPH; McGrath et al., 2010), hAOC3 (PDB ID 4BTX; Bligt-Lindén et al., 2013), and bAOC4 (PDB ID 1TU5; Lunelli et al., 2005), which are homodimers consisting of the D2, D3 and D4 domains (Fig. 1). Since no crystal structure is available for hAOC2 we built a homology model based on the hAOC3 structure (PDB: 4BTX) (Bligt-Lindén et al., 2013). The high sequence identity of 68% between hAOC2 and the hAOC3 template is much higher than 30%, which is generally regarded as the sequence identity criteria for an accurate homology model (John and Sali, 2003). The quality of the hAOC2 model is good, since 90% of the residues in the Ramachandran plot are in the favored regions and the QMEAN quality score of 0.704 (range 0–1; where 1 is good) is within the range of scores for reference structures of the same size. The folds of the hAOC2 model and the X-ray structure of hAOC3 are very similar with a root mean square deviation (RMSD) of 0.67 Å for the C-alpha atoms of the superimposed structures.

In the studied CAO structures, half of the residues near TPQ are fully conserved (white in Fig. 8). Most of the conserved residues are involved in TPQ biogenesis and/or in the catalytic mechanism (Airenne et al., 2005; Ernberg et al., 2010; Klema and Wilmot, 2012). Of these residues, the catalytic Asp (D373 in hAOC1) and a nearby Tyr (Y371 in hAOC1) form the conserved Y-X-D motif. The corresponding Tyr has been identified as a “gate” residue in all mammalian CAOs (Salminen et al., 1998; Lunelli et al., 2005; Kaitaniemi et al., 2009; McGrath et al., 2009). In the hAOC3 − 2-hydrazino-pyridine (2HP) complex, the corresponding Y384 stacks with the pyridine ring of 2HP (Jakobsson et al., 2005) and docking studies also suggested that this residue interacts with the aromatic ring of benzylamine (Kaitaniemi et al., 2009). Therefore, it is likely that the conserved “gate” Tyr interacts with the amine substrates of all mammalian CAOs. Two conserved Tyrs and an Asn (Y359, Y463 and N460 in hAOC1) orient TPQ (TPQ461 in hAOC1; (Airenne et al., 2005; Ernberg et al., 2010), whereas the conserved Ser (S456 in hAOC1) forms hydrogen bonds that stabilize the active site architecture (Fig. 8). The rest of the strictly conserved residues originate from the T/S-X1-X2-N-Y-D active site motif, where the first residue is a Ser in most of the AOC2 proteins, but conservatively replaced by a Thr in all the other mammalian CAOs. The Ser/Thr residue (T457 in hAOC1) and the Asp (D462 in hAOC1) form a hydrogen bond that stabilizes the tip of the β-hairpin formed by the active site motif. Both of them are also involved in an extensive network of inter-monomeric hydrogen bonds together with the fully conserved Arg (R429 in hAOC1) and His (H430 in hAOC1) from the other monomer. A similar hydrogen bonding network has been identified e.g. in the crystal structure of *Escherichia coli* amine oxidase (Parsons et al., 1995) but its functional significance is currently unknown.

Since the unique substrate preferences of each CAO sub-family is likely determined by the active site residues, which are conserved within a sub-family but variable between the sub-families, we next investigated these residues in each CAO sub-family (Fig. 8; Table 1). hAOC1 has a unique preference for aromatic diamines as substrates (e.g. histamine; Elmore et al., 2002) and, accordingly, the AOC1 sub-family does not group together with the other CAO sub-families (Fig. 2).

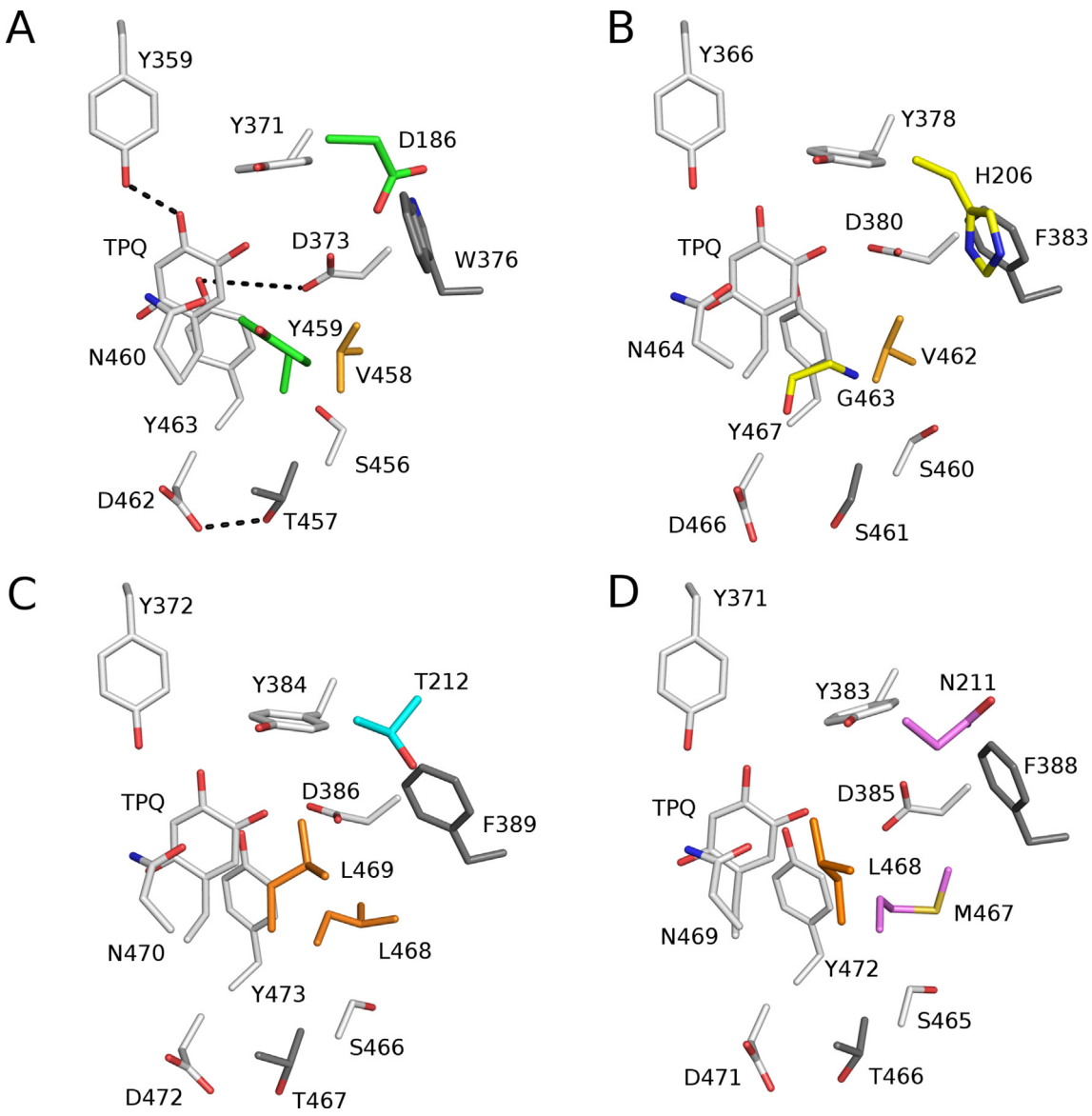

Fig. 8. Active sites residues in CAOs within 6 Å distance from TPQ. The fully conserved residues represented as white sticks and residues with conserved features as grey sticks. (A) hAOC1. The conserved hydrogen bonding network in the active site of CAOs is shown with black dashes. The residues putatively important for the diamine substrate preference of AOC1s are shown as green sticks. (B) hAOC2. The residues that putatively provide AOC2s the preference for aromatic substrates are shown as yellow sticks. (C) hAOC3. The residues shown as orange and cyan sticks likely contribute to the preference of AOC3s for small aliphatic substrates. (D) bAOC4. The residues that differ between AOC3 and AOC4 are shown as violet sticks. These residues allow AOC4 to use aliphatic polyamines as substrates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In addition to the residues involved in the catalytic mechanism and orienting TPQ, three other residues are strictly conserved only within the AOC1 sub-family (V458, Y459 and D186 in hAOC1) (Fig. 8A; Table 1). Of these residues, the totally conserved Val from the active site motif (V458 in hAOC1; X1) is also conserved in the AOC2 sub-family and, based on our results, it contributes to the preference for aromatic substrates. The Tyr (Y459 in hAOC1; X2) and the other aromatic residue in the active site (W376 in hAOC1; Trp/Tyr in all AOC1s) are suitably positioned to form stacking interactions with the aromatic ring of the substrate (Fig. 8A). In addition, D186 in hAOC1 has been predicted to be involved in histamine binding, since its position is suitable for binding the second amine group of the diamine substrates (McGrath et al., 2009). Our results support this prediction, as the corresponding residue is an Asp in all the studied AOC1s (Fig. 8A; Table 1). In the crystal structure of the hAOC1-pentamidine complex (McGrath et al., 2009), S380 forms a hydrogen bond with the amidinium group of pentamidine. The longer diamine substrates of hAOC1, spermidine and spermine, have nitrogen in the linker region between the primary and second amines that could interact with D186, whereas the position of S380 is perfect for binding the second amine of these longer diamines (McGrath et al., 2009). The role of S380 in binding the second amine is supported by the fact that the corresponding residue always has a hydroxyl group (Ser/Thr) in the studied mammalian AOC1s (Table 1).

AOC2 prefers the aromatic amine 2-phenylethylamine and also binds the polar aromatic amines tryptamine and *p*-tyramine (Kaitaniemi et al., 2009). The small variable Gly and Val residues in the active site motif, make AOC2 optimal for aromatic amines by providing the space required for binding the aromatic ring in the substrates (G463 and V462 in hAOC2; Fig. 8B; Table 1). Of these, the Gly has been shown to be important for the *p*-tyramine activity since the replacement of the corresponding Leu in hAOC3 by a Gly (L469G) acquired enzymatic activity towards the aromatic *p*-tyramine, which cannot be oxidized by wild-type hAOC3 (Elovaara et al., 2011). In the position corresponding to the conserved Asp in AOC1s (D186 in hAOC1), AOC2s have a His

Table 1
Comparison of the active site residues in CAOs within 6 Å distance from TPQ. Conserved residues are marked in bold and the residues in the active site motif with *. The residues that form the basis for the classification into the CAO sub-families are shown with coloured background: Val with light orange (aromatic preference), Tyr with green (diamine preference), Leu with orange (monoamine preference) and Met with violet (polyamine preference).

hAOC1	hAOC2	hAOC3	bAOC4	Role
Y371	**Y378**	**Y384**	**Y383**	The gate residue involved in substrate binding
D373	**D380**	**D386**	**D385**	The catalytic aspartate
T457	S461	**T467**	**T466**	Involved in inter-monomeric interactions*
V458	**V462**	**L468**	**M467**	X1, discriminates AOC3 and AOC4*
Y459	**G463**	**L469**	**L468**	X2, discriminates AOC1-2-3/4*
N460	**N464**	**N470**	**N469**	Positions TPQ*
D462	**D466**	**D472**	**D471**	Involved in inter-monomeric interactions*
S456	**S460**	**S466**	**S465**	Forms hydrogen bonds that stabilize the active site
Y359	**Y366**	**Y372**	**Y371**	Positions TPQ
Y463	**Y467**	**Y473**	**Y472**	Positions TPQ
D186	**H206**	T212	N211	Key residue suggested to have a role in substrate binding
W376	F383	**F389**	**F388**	Aromatic residue that putatively positions substrates

(H206 in hAOC2; Fig. 8B), which could interact with polar and/or aromatic substrates. This His can be considered as an AOC2-specific residue as it is exclusively conserved in AOC2s (Table 1), with the exception of the bat proteins where a Gln conservatively substitutes for His. It is worth noticing that rats do not have a functional AOC2 and, thus, mice are preferable model organisms for AOC2 studies.

The phylogenetic and sequence analysis of the ancestral CAOs from *C. milii* and *R. typys* suggest that they could use both diamines and monoamines as substrates. A Gly at position X2 classifies both *C. milii* and *R. typys* CAOs as AOC2 proteins, which is in accordance with their highest sequence identities to human AOC2 and their predicted N-terminal transmembrane helices. Val at position X1 makes *R. typys* CAO even more similar to hAOC2 than *C. milii* CAO, which has an Ile. Further comparison of *C. milii* and *R. typys* CAOs with hAOC1 (Supplementary Fig. S2) reveals that they might also have diamine oxidase activity since they have a conserved Asp at the position of D186 in hAOC1 instead of a highly conserved His (H206 in hAOC2).

AOC3 and AOC4 proteins both prefer aliphatic amines; AOC3 prefers small aliphatic amines and AOC4 long polyamines (Kaitaniemi et al., 2009; Bonaiuto et al., 2010; Elovaara et al., 2011). Based on the active site comparison of hAOC3 and bAOC4 (Fig. 8C, D), both proteins have a Phe (F389 in hAOC3) in the active site channel 6 Å from TPQ towards the surface a Leu (Leu469 in hAOC3; X2), which are totally conserved in all the studied proteins (Table 1). The Leu has been associated in blocking/allowing access to TPQ (Airenne et al., 2005; Jakobsson et al., 2005; Kaitaniemi et al., 2009; McGrath et al., 2009), and mutational studies have proven that it is involved in the substrate preference of hAOC3 (Kaitaniemi et al., 2009; Elovaara et al., 2011). Unlike hAOC1 and hAOC2, both AOC3 and AOC4 prefer aliphatic amines, for which Leu and Phe provide a hydrophobic environment that can well accommodate the carbon tails of the aliphatic substrates (Fig. 8C, D). Residue X1 from the active site motif is the only active site residue that is not conserved between the AOC3 (Leu) and AOC4 (Met) sub-families but almost exclusively conserved in each of them. The Met in AOC4s is slightly more polar and much more flexible than the Leu in AOC3s. Similar amino acid difference was found in *Arabidopsis thaliana* and *Euphorbia lagascae* sterol carrier protein-2 in which the Met/Leu difference was respectively crucial for sterol sensitivity and insensitivity (Viitanen et al., 2006). However, the exact role of the X1 residue in CAOs has not been studied experimentally. Furthermore, it has been indicated that the K393/Y392 residue difference of hAOC3/bAOC4 contributes to the substrate preference as the positively charged Lys in hAOC3 could hinder binding of long polyamines (Bonaiuto et al., 2010). The high conservation of the corresponding Lys in AOC3s observed in our study lends support to this suggested role in selecting preferred substrates.

Interestingly, the residue corresponding to D186 in AOC1 (Fig. 8A) is a highly conserved Thr in AOC3s (T212 in hAOC3; Fig. 8C) and a highly conserved Asn in most of the studied mammalian AOC4s (N211 in bAOC4; Fig. 8D). Mutational studies on hAOC3 (T212A) have shown that T212 slightly contributes to substrate preference but merely has a role as an additional gate residue (Elovaara et al., 2011). Intriguingly, both AOC3 and AOC4 from *I. tridecemlineatus* (rodent) have an Asn in this position, whereas all primate AOC3 and AOC4 proteins have a Thr. It should also be noted that only human and gorilla lack AOC4 but most of the studied Primates have it, whereas *I. tridecemlineatus* and *M. marmota marmota* are the only rodent species with AOC4. Therefore, based on our analysis, rats and mice are better model organisms for studying AOC3 than those Primates that have both AOC3 and AOC4.

The substrate preference of CAOs has mainly been studied for human and rodent proteins (Lyles, 1996; Elmore et al., 2002; Zhang et al., 2003; Shen et al., 2012; Finney et al., 2014). Rodents have been used as model organisms in human AOC3-related diseases, drug design and *in vivo* imaging applications, whereas the research on AOC4 has focused on bAOC4. Comparison of rodent and primate CAO proteins suggests that for binding studies of small substrates and inhibitors binding covalently to TPQ, rodents and other primate species can serve as good model organisms for human CAOs. In drug design projects, problematic species-specific differences in ligand binding between human and mouse AOC3 have been demonstrated for a number of AOC3 inhibitors (Foot et al., 2012; Bligt-Lindén et al., 2013; Inoue et al., 2013a; Inoue et al., 2013b). These problems are mainly caused by residue differences in the 15 Å long active site channel of CAOs, which leads from TPQ to the surface of the protein. Therefore, our analysis suggests that the active site channel in CAOs from model organisms should be studied in detail when larger substrates and inhibitors are designed. In addition to the species-specific binding properties, the other CAO proteins expressed in the used model organism should also

be taken into consideration in the CAO-related research.

4. Conclusion

In this work, we combined the phylogenetic analysis of vertebrate CAOs with the multivariate plots of mammalian CAOs to suggest a novel way to classify mammalian CAOs based on their active site motif. We conclude that residue X2 from the X1-X2-N-Y-D motif separates the mammalian CAOs in the AOC1, AOC2 and AOC3/AOC4 sub-families and residue X1 can be used to discriminate between the AOC3 and AOC4 proteins. Moreover, we aimed to understand what makes each mammalian CAO unique and, by relating phylogenetic analyses with sequence and structural data, we gained new insights into the substrate preference of mammalian CAOs. Residues X1 and X2, which are conserved within each CAO sub-family and located in the vicinity of the catalytic site, can be considered as the key determinants for selective substrate binding. Both AOC1 and AOC2 that uniquely prefer aromatic substrates have a Val at position X1. In AOC1, residue X2, which is a conserved Tyr, might be the most important residue for its preference for aromatic diamines. Additionally, two other key positions are conserved in the AOC1 sub-family but variable in the other sub-families. The totally conserved Asp (D186 in hAOC1) and the conservatively replaced Ser/Thr (S380 in hAOC1) could bind the second amine group in shorter and longer diamine substrates, respectively. The position corresponding to the totally conserved Asp is also highly conserved in the other sub-families (a His in AOC2, a Thr in AOC3 and an Asn in AOC4) and, thus, it seems to have an effect on the substrate selectivity in the other subfamilies as well. In AOC2, the highly conserved His (H206 in hAOC2) and the totally conserved Gly (X2), which creates space for large aromatic substrates, form the foundation for substrate preference. The respective Thr and Asn residues in AOC3 (T212) and AOC4 (N211) likely contribute to their substrate selectivity together with the X1 residue (Leu in AOC3 and Met in AOC4) whereas the X2 residue (Leu) makes them distinctively prefer aliphatic rather than aromatic substrates. Due to the species-specific expression profile of CAOs, our results also highlight that it is important to be aware of the presence of all CAO family members expressed in the studied organism, even thought the research would focus on one of the sub-families.

Acknowledgement and funding

We thank the bioinformatics (J.V. Lehtonen), translational activities and structural biology infrastructure support from Biocenter Finland and Instruct-FI, and CSC IT Center for Science for computational infrastructure support at the Structural Bioinformatics Laboratory, Åbo Akademi University. We thank Käthe M. Dahlström and Mia Åstrand for critical reading of the article and Julia Fredman for assistance with the revision documents. This work was supported by the National Doctoral Programme in Informational and Structural Biology (E.B.-L. and L.L.C.) and Sigrid Juselius Foundation (T.A.S. and M.S.J.), the Magnus Ehrnrooth foundation (T.A.S), the Tor, Joe, and Pentti Borgs Foundation (E.B.-L., L.L.C., T.A.S. and M.S.J), Medicinska Understödsföreningen Liv och hälsa (T.A.S., L.L.C. and E.B.-L.), the Finnish Cultural Foundation (L.L.C.), the Centre for International Mobility (CIMO; A.R.), the University Grants Commission, India under Indo-Finnish Cultural Exchange Programme (UGC; A.R.) and the Academy of Finland (No. 132998 and No. 288534 T.A.S.).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106571.

References

Aalto, K., Autio, A., Kiss, E.A., Elima, K., Nymalm, Y., Veres, T.Z., Marttila-Ichihara, F., Elovaara, H., Saanijoki, T., Crocker, P.R., Maksimow, M., Bligt, E., Salminen, T.A., Salmi, M., Roivainen, A., Jalkanen, S., 2011. Siglec-9 is a novel leukocyte ligand for vascular adhesion protein-1 and can be used in PET imaging of inflammation and cancer. Blood 118, 3725–3733.

Abella, A., García-Vicente, S., Viguerie, N., Ros-Baró, A., Camps, M., Palacín, M., Zorzano, A., Marti, L., 2004. Adipocytes release a soluble form of VAP-1/SSAO by a metalloprotease- dependent process and in a regulated manner. Diabetologia 47, 429–438.

Agostinelli, E., Tempera, G., Viceconte, N., Saccoccio, S., Battaglia, V., Grancara, S., Toninello, A., Stevanato, R., 2010. Potential anticancer application of polyamine oxidation products formed by amine oxidase: A new therapeutic approach. Amino Acids 38, 353–368.

Airenne, T.T., Nymalm, Y., Kidron, H., Smith, D.J., Pihlavisto, M., Salmi, M., Jalkanen, S., Johnson, M.S., Salminen, T.A., 2005. Crystal structure of the human vascular adhesion protein 1: Unique structural features with functional implications. Protein Sci. 14, 1964–1974.

Almeida, A.P., Beaven, M.A., 1981. Phylogeny of histamine in vertebrate brain. Brain Res. 208, 244–250.

Altschul F., S., Miller, W., Myers W., E., Lipman J., D., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Benkert, P., Kunzli, M., Schwede, T., 2009. QMEAN server for protein model quality estimation. Nucleic Acids Res. 37, W510–W514.

Bligt-Lindén, E., Pihlavisto, M., Szatmári, I., Otwinowski, Z., Smith, D.J., Lázár, L., Fülöp, F., Salminen, T.A., 2013. Novel pyridazinone inhibitors for vascular adhesion protein-1 (VAP-1): Old target-new inhibition mode. J. Med. Chem. 56, 9837–9848.

Bonaiuto, E., Lunelli, M., Scarpa, M., Vettor, R., Milan, G., Di Paolo, M.L., 2010. A structure-activity study to identify novel and efficient substrates of the human semicarbazide-sensitive amine oxidase/VAP-1 enzyme. Biochimie 92, 858–868.

Bonder, C.S., Norman, M.U., Swain, M.G., Zbytnuik, L.D., Yamanouchi, J., Santamaria, P., Ajuebor, M., Salmi, M., Jalkanen, S., Kubes, P., 2005. Rules of recruitment for Th1 and Th2 lymphocytes in inflamed liver: A role for alpha-4 integrin and vascular adhesion protein-1. Immunity 23, 153–163.

Boomsma, F., Bhaggoe, U.M., Van Der Houwen, A.M.B., Van Den Meiracker, A.H., 2003. Plasma semicarbazide-sensitive amine oxidase in human (patho)physiology. Biochim. Biophys. Acta - Proteins Proteomics 1647, 48–54.

Boomsma, F., Hut, H., Bhaggoe, U., Van Der, Houwen A., Van Den, Meiracker A., 2005. From cell to circulation. Heart Fail. 11, 122–126.

Dawkes, H., 2001. Copper amine oxidase: cunning cofactor and controversial copper. Curr. Opin. Struct. Biol. 11, 666–673.

Dunkel, P., Gelain, A., Barlocco, D., Haider, N., Gyires, K., Sperlágh, B., Magyar, K., Maccioni, E., Fadda, A., Mátyus, P., 2008. Semicarbazide-sensitive amine oxidase/vascular adhesion protein 1: recent developments concerning substrates and inhibitors of a promising therapeutic target. Curr. Med. Chem. 15, 1827–1839.

Elmore, B.O., Bollinger, J.A., Dooley, D.M., 2002. Human kidney diamine oxidase: heterologous expression, purification, and characterization. J. Biol. Inorg. Chem. 7, 565–579.

Elovaara, H., Kidron, H., Parkash, V., Nymalm, Y., Bligt, E., Ollikka, P., Smith, D.J., Pihlavisto, M., Salmi, M., Jalkanen, S., Salminen, T.A., 2011. Identification of two imidazole binding sites and key residues for substrate specificity in human primary amine oxidase AOC3. Biochemistry 50, 5507–5520.

Ernberg, K., McGrath, A.P., Peat, T.S., Adams, T.E., Xiao, X., Pham, T., Newman, J., McDonald, I.A., Collyer, C.A., Guss, J.M., 2010. A new crystal form of human vascular adhesion protein 1. Acta Crystallogr Sect. F Struct. Biol. Cryst. Commun. 66, 1572–1578.

Felsenstein, J., 1985. Confidence-Limits on phylogenies - an approach using the bootstrap. Evolution (N. Y) 39, 783–791.

Finney, J., Moon, H.J., Ronnebaum, T., Lantz, M., Mure, M., 2014. Human copper-dependent amine oxidases. Arch. Biochem. Biophys. 546, 19–32.

Foot, J.S., Deodhar, M., Turner, C.I., Yin, P., van Dam, E.M., Silva, D.G., Olivieri, A., Holt, A., McDonald, I.A., 2012. The discovery and development of selective 3-fluoro-4-aryloxyallylamine inhibitors of the amine oxidase activity of semicarbazide-sensitive amine oxidase/vascular adhesion protein-1 (SSAO/VAP-1). Bioorg. Med. Chem. Lett. 22, 3935–3940.

Holt, A., Smith, D.J., Cendron, L., Zanotti, G., Rigo, A., Di Paolo, M.L., 2008. Multiple binding sites for substrates and modulators of semicarbazide-sensitive amine oxidases: kinetic consequences. Mol. Pharmacol. 73, 525–538.

Illergård, K., Ardell, D.H., Elofsson, A., 2009. Structure is three to ten times more conserved than sequence–a study of structural response in protein cores. Proteins 77, 499–508.

Imamura, Y., Kubota, R., Wang, Y., Asakawa, S., Kudoh, J., Mashima, Y., Oguchi, Y., Shimizu, N., 1997. Human retina-specific amine oxidase (RAO): cDNA cloning, tissue expression, and chromosomal mapping. Genomics 40, 277–283.

Imamura, Y., Noda, S., Mashima, Y., Kudoh, J., Oguchi, Y., Shimizu, N., 1998. Human retina-specific amine oxidase: genomic structure of the gene (AOC2), alternatively spliced variant, and mRNA expression in retina. Genomics 51, 293–298.

Inoue, T., Morita, M., Tojo, T., Nagashima, A., Moritomo, A., Imai, K., Miyake, H., 2013a. Synthesis and SAR study of new thiazole derivatives as vascular adhesion protein-1 (VAP-1) inhibitors for the treatment of diabetic macular edema: part 2. Bioorg. Med. Chem. 21, 2478–2494.

Inoue, T., Morita, M., Tojo, T., Nagashima, A., Moritomo, A., Miyake, H., 2013b. Novel 1H-imidazol-2-amine derivatives as potent and orally active vascular adhesion protein-1 (VAP-1) inhibitors for diabetic macular edema treatment. Bioorg. Med. Chem. 21, 3873–3881.

Jaakkola, K., Nikula, T., Holopainen, R., Vahasilta, T., Matikainen, M.T., Laukkanen, M.L., Huupponen, R., Halkola, L., Nieminen, L., Hiltunen, J., Parviainen, S., Clark, M.R., Knuuti, J., Savunen, T., Kääpä, P., Voipio-Pulkki, L.M., Jalkanen, S., 2000. In vivo detection of vascular adhesion protein-1 in experimental inflammation. Am. J.

Pathol. 157, 463–471.
Jain, E., Bairoch, A., Duvaud, S., Phan, I., Redaschi, N., Suzek, B.E., Martin, M.J., McGarvey, P., Gasteiger, E., 2009. Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinf. 10, 136.
Jakobsson, E., Nilsson, J., Ogg, D., Kleywegt, G.J., 2005. Structure of human semicarbazide-sensitive amine oxidase/vascular adhesion protein-1. Acta Crystallogr. Sect. D Biol. Crystallogr. 61, 1550–1562.
Jalkanen, S., Karikoski, M., Mercier, N., Koskinen, K., Henttinen, T., Elima, K., Salmivirta, K., Salmi, M., 2007. The oxidase activity of vascular adhesion protein-1 (VAP-1) induces endothelial E- and P-selectins and leukocyte binding. Blood 110, 1864–1870.
Jalkanen, S., Salmi, M., 2008. VAP-1 and CD73, endothelial cell surface enzymes in leukocyte extravasation. Arterioscler. Thromb. Vasc. Biol. 28, 18–26.
Jalkanen, S., Salmi, M., 2017. Vascular adhesion protein-1: a cell-surface amine oxidase in translation. Antioxid. Redox Signal 00:ars.2017.7418.
Janes, S.M., Mu, D., Wemmer, D., Smith, A.J., Kaur, S., Maltby, D., Burlingame, A.L., Klinman, J.P., 1990. A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site of bovine serum amine oxidase. Science 248, 981–987.
John, B., Sali, A., 2003. Comparative protein structure modeling by iterative alignment, model building and model assessment. Nucleic Acids Res. 31, 3982–3992.
Johnson, M.S., May, A.C.W., Rodinov, M.A., Overington, J.P., 1996. Discrimination of common protein folds: Application of protein structure to sequence/structure comparisons. Methods Enzymol. 266, 575–598.
Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.
Kaitaniemi, S., Elovaara, H., Grön, K., Kidron, H., Liukkonen, J., Salminen, T., Salmi, M., Jalkanen, S., Elima, K., 2009. The unique substrate specificity of human AOC2, a semicarbazide-sensitive amine oxidase. Cell. Mol. Life Sci. 66, 2743–2757.
Kivi, E., Elima, K., Aalto, K., Nymalm, Y., Auvinen, K., Koivunen, E., Otto, D.M., Crocker, P.R., Salminen, T.A., Salmi, M., Jalkanen, S., 2009. Human Siglec-10 can bind to vascular adhesion protein-1 and serves as its substrate. Blood 114, 5385–5392.
Klema, V.J., Wilmot, C.M., 2012. The role of protein crystallography in defining the mechanisms of biogenesis and catalysis in copper amine oxidase. Int. J. Mol. Sci. 13, 5375–5405.
Kumar, S., Stecher, G., Peterson, D., Tamura, K., 2012. MEGA-CC: Computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. Bioinformatics 28, 2685–2686.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.
Kurkijärvi, R., Adams, D.H., Leino, R., Möttönen, T., Jalkanen, S., Salmi, M., 1998. Circulating form of human vascular adhesion protein-1 (VAP-1): increased serum levels in inflammatory liver diseases. J. Immunol. 161, 1549–1557.
Lai, J., Jin, J., Kubelka, J., Liberles, D.A., 2012. A phylogenetic analysis of normal modes evolution in enzymes and its relationship to enzyme function. J. Mol. Biol. 422, 442–459.
Lalor, P.F., Shields, P., Grant, A., Adams, D.H., 2002. Recruitment of lymphocytes to the human liver. Immunol. Cell Biol. 80, 52–64.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.
Lehtonen, J.V., Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., Pentikäinen, O., Nyrönen, T., Salminen, T., Gyllenberg, M., Johnson, M.S., et al., 2004. BODIL: a molecular modeling environment for structure-function analysis and drug design. J. Comput. Aided. Mol. Des. 18, 401–419.
Lunelli, M., Di Paolo, M.L., Biadene, M., Calderone, V., Battistutta, R., Scarpa, M., Rigo, A., Zanotti, G., 2005. Crystal structure of amine oxidase from bovine serum. J. Mol. Biol. 346, 991–1004.
Lyles, G.A., 1996. Mammalian plasma and tissue-bound settsitive amine oxidases: biochemical, pharmacological and toxicological aspects. Science (80-.) 28.
Maintz, L., Novak, N., 2007. Histamine and histamine intolerance. Am. J. Clin. Nutr. 85, 1185–1196.
McGrath, A.P., Hilmer, K.M., Collyer, C.A., Shepard, E.M., Elmore, B.O., Brown, D.E., Dooley, D.M., Guss, J.M., 2009. Structure and inhibition of human diamine oxidase. Biochemistry 48, 9810–9822.
McGrath, A.P., Hilmer, K.M., Collyer, C.A., Dooley, D.M., Guss, J.M., 2010. A new crystal form of human diamine oxidase. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66, 137–142.
Merinen, M., Irjala, H., Salmi, M., Jaakkola, I., Hänninen, A., Jalkanen, S., 2005. Vascular adhesion protein-1 is involved in both acute and chronic inflammation in the mouse. Am. J. Pathol. 166, 793–800.
Mu, D., Medzihradszky, K.F., Adams, G.W., Mayer, P., Hines, W.M., Burlingame, A.L., Smith, A.J., Cai, D., Klinman, J.P., 1994. Primary structures for a mammalian cellular and serum copper amine oxidase. J. Biol. Chem. 269, 9926–9932.
Mušič, E., Korošec, P., Šilar, M., Adamič, K., Košnik, M., Rijavec, M., 2013. Serum diamine oxidase activity as a diagnostic test for histamine intolerance. Wien. Klin. Wochenschr. 125, 239–243.
Pannecoeck, R., Serruysa, D., Benmeridjaa, L., Delangheb, J.R., Geelc, N., Speeckaertc, R., Speeckaerta, M.M., 2015. Vascular adhesion protein-1: Role in human pathology and application as a biomarker. Crit. Rev. Clin. Laboraty Sci. 8363, 1–17.
Di Paolo, M.L., Stevanato, R., Corazza, A., Vianello, F., Lunelli, L., Scarpa, M., Rigo, A., 2003. Electrostatic compared with hydrophobic interactions between bovine serum amine oxidase and its substrates. Biochem. J. 371, 549–556.
Parsons, M.R., Convery, M.A., Wilmot, C.M., Yadav, K.D.S., Blakeley, V., Corner, A.S., Phillips, S., McPherson, M.J., Knowles, P.F., 1995. Crystal structure of a quinoenzyme: copper amine oxidase of Escherichia coli at 2 Å resolution. Structure 3, 1171–1184.
Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.
Rajtar, S., Irman-Florjanc, T., 2007. Amitriptyline affects histamine-N-methyltransferase and diamine oxidase activity in rats and guinea pigs. Eur. J. Pharmacol. 574, 201–208.
Repessé, X., Moldes, M., Muscat, A., Vatier, C., Chetrite, G., Gille, T., Planes, C., Filip, A., Mercier, N., Duranteau, J., Fève, B., 2015. Hypoxia inhibits semicarbazide-sensitive amine oxidase activity in adipocytes. Mol. Cell. Endocrinol. 411, 58–66.
Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evo. 4, 406–425.
Sali, A., Blundell, T., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.
Salmi, M., Jalkanen, S., 2001. VAP-1: an adhesin and an enzyme. Trends Immunol. 22, 211–216.
Salmi, M., Jalkanen, S., 2014. Ectoenzymes in leukocyte migration and their therapeutic potential. Semin. Immunopathol. 36, 163–176.
Salmi, M., Kalimo, K., Jalkanen, S., 1993. Induction and function of vascular adhesion protein-1 at sites of inflammation. J. Exp. Med. 178, 2255–2260.
Salminen, T.A., Smith, D.J., Jalkanen, S., Johnson, M.S., 1998. Structural model of the catalytic domain of an enzyme with cell adhesion activity: human vascular adhesion protein-1 (HVAP-1) D4 domain is an amine oxidase. Protein Eng. Des. Sel. 11, 1195–1204.
Schwelberger, H.G., 2006. Origins of plasma amine oxidases in different mammalian species. Inflamm. Res. 55 (Suppl 1), S57–S58.
Schwelberger, H.G., 2010. Structural organization of mammalian copper-containing amine oxidase genes. Inflamm. Res. 59 (Suppl 2) S223-5. A.
Shen, S.H., Wertz, D.L., Klinman, J.P., 2012. Implication for functions of the ectopic adipocyte copper amine oxidase (AOC3) from purified enzyme and cell-based kinetic studies. PLoS ONE 7, 1–11.
Shepard, E.M., Dooley, D.M., 2015. Inhibition and oxygen activation in copper amine oxidases. Acc. Chem. Res. 48, 1218–1226.
Smith, D.J., Salmi, M., Bono, P., Hellman, J., Leu, T., Jalkanen, S., 1998. Cloning of vascular adhesion protein 1 reveals a novel multifunctional adhesion molecule. J. Exp. Med. 188, 17–27.
Sojo, V., Dessimoz, C., Pomiankowski, A., Lane, N., 2016. Membrane proteins are dramatically less conserved than water-soluble proteins across the tree of life. Mol. Biol. Evol. 33, 2874–2884.
Soltis, D.E., Soltis, P.S., 2003a. Update on molecular systematics the role of phylogenetics in comparative genetics. Plant Physyology 132, 1790–1800.
Soltis, D.E., Soltis, P.S., 2003b. Applying the bootstrap in phylogeny reconstruction. Stat. Sci. 18, 256–267.
Stolen, C.M., Yegutkin, G.G., Kurkijärvi, R., Bono, P., Alitalo, K., Jalkanen, S., 2004. Origins of serum semicarbazide-sensitive amine oxidase. Circ. Res. 95, 50–57.
Tohka, S., Laukkanen, M., Jalkanen, S., Salmi, M., 2001. Vascular adhesion protein 1 (VAP-1) functions as a molecular brake during granulocyte rolling and mediates recruitment in vivo. FASEB J. 15, 373–382.
Venkatesh, B., Lee, A.P., Ravi, V., Maurya, A.K., Lian, M.M., Swann, J.B., Ohta, Y., Flajnik, M.F., Sutoh, Y., Kasahara, M., Hoon, S., Gangu, V., Roy, S.W., Irimia, M., Korzh, V., Kondrychyn, I., Lim, Z.W., Tay, B.H., Tohari, S., Kong, K.W., Ho, S., Lorente-Galdos, B., Quilez, J., Marques-Bonet, T., Raney, B.J., Ingham, P.W., Tay, A., Hillier, L.W., Minx, P., Boehm, T., Wilson, R.K., Brenner, S., Warren, W.C., 2014. Elephant shark genome provides unique insights into gnathostome evolution. Nature 505, 174–179.
Viitanen, L., Nylund, M., Eklund, D.M., Alm, C., Eriksson, A., 2006. Characterization of SCP-2 from Euphorbia lagascae reveals that a single Leu / Met exchange enhances sterol transfer activity. FASEB J. 273, 5641–5655.
Wheeler, T.J., Clements, J., Finn, R.D., 2014. Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. BMC Bioinf. 15, 7.
Zhang, Q., Mashima, Y., Noda, S., Imamura, Y., Kudoh, J., Shimizu, N., Nishiyama, T., Umeda, S., Oguchi, Y., Tanaka, Y., Iwata, T., 2003. Characterization of AOC2 gene encoding a copper-binding amine oxidase expressed specifically in retina. Gene 318, 45–53.
Zorzano, A., Abella, A., Marti, L., Carpéné, C., Palacín, M., Testar, X., 2003. Semicarbazide-sensitive amine oxidase activity exerts insulin-like effects on glucose metabolism and insulin-signaling pathways in adipose cells. Biochim. Biophys. Acta - Proteins Proteomics 1647, 3–9.
Zuckerkandl, E., Pauling, L., 1965. Evolutionary divergence and convergence in proteins. Evol. Genes Proteins 97.