Computational and Structural Biotechnology Journal xxx (2017) xxx–xxx

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

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

Amplification and bioinformatics analysis of conserved FAD-binding region of L-amino acid oxidase (*LAAO*) genes in gastropods compared to other organisms

Wipawadee Suwannapan [a], Pramote Chumnanpuen [b,c], Teerasak E-kobon [a,c,*]

[a] Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
[b] Department of Zoology, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
[c] Computational Biomodelling Laboratory for Agricultural Science and Technology, Kasetsart University, Bangkok 10900, Thailand

ARTICLE INFO

Article history:
Received 23 October 2017
Received in revised form 25 February 2018
Accepted 25 February 2018
Available online xxxx

Keywords:
L-amino acid oxidase (LAAO)
FAD-binding region
Gastropods
Pattern analysis
Clustering

ABSTRACT

This study aimed to investigate the conserved FAD-binding region of the L-amino acid oxidase (*LAAO*) genes in twelve gastropod genera commonly found in Thailand compared to those in other organisms using molecular cloning, nucleotide sequencing and bioinformatics analysis. Genomic DNA of gastropods and other invertebrates was extracted and screened using primers specific to the conserved FAD-binding region of *LAAO*. The amplified 143-bp fragments were cloned and sequenced. The obtained nucleotide sequences of 21 samples were aligned and phylogenetically compared to the *LAAO*-conserved FAD-binding regions of 210 other organisms from the NCBI database. Translated amino acid sequences of these samples were used in phylogenetics and pattern analyses. The phylogenetic trees showed clear separation of the conserved regions in fungi, invertebrates, and vertebrates. Alignment of the conserved 47-amino-acid FAD-binding region of the LAAOs showed 150 unique sequences among the 231 samples and these patterns were different from those of other flavoproteins in the amine oxidase family. An amino acid pattern analysis of five sub-regions (bFAD, FAD, FAD-GG, GG, and aGG) within the FAD-binding sequence showed high variation at the FAD-GG sub-region. Pattern analysis of secondary structures indicated the aGG sub-region as having the highest structural variation. Cluster analysis of these patterns revealed two major clusters representing the mollusc clade and the vertebrate clade. Thus, molecular phylogenetics and pattern analyses of sequence and structural variations could reflect evolutionary relatedness and possible structural conservation to maintain specific function within the FAD-binding region of the LAAOs in gastropods compared to other organisms.

1. Introduction

Gastropods are the most diverse taxonomic group of molluscs living in a wide range of habitats, particularly in tropical and subtropical areas. The gastropods are generally classified into operculum-bearing prosobranchs, air-breathing pulmonates, and opisthobranchs (or sea slugs and sea hares). As Thailand is one of the biodiversity hotspots of the world and gastropods can live in diverse habitats, many gastropods are commonly found in Thailand, including pond snails (*Physa* sp., *Lymnaea* sp., *Biomphalaria* sp., *Bithynia* sp., *Viviparus* sp., *Pomacea canaliculata* and *Melanoides* sp.), garden snails (*Achatina fulica* and *Cryptozona siamensis)*, forest snails (*Cyclophorus volvulus)*, land slugs (*Semperula siamensis)*, sea snails (*Babylonia areolata*) and sea slugs (*Jorunna funebris*). Some of these are nutritious human foods (*Viviparus* sp., *B. areolata* and *C. volvulus)*, while other gastropods are harmful agricultural pests (*P. canaliculata*, *C. siamensis*, and *A. fulica*) and recognized as intermediate hosts of many nematode parasites in humans. Gastropods secrete mucus to facilitate their locomotion and hydrate their body surfaces. The mucus allows the snails to trail home, adhere to substrates and protect themselves from skin damage and microbial infection [1]. Gastropod mucus contains mostly water, electrolytes, minerals, glycoconjugates, proteoglycans, small peptides and glycoproteins including L-amino acid oxidases (LAAOs), a member of the amine oxidase family (PF01593) [2–8].

LAAO is a flavoenzyme that requires flavin adenine dinucleotide (FAD) or a quinone as a cofactor [9,10]. This enzyme acts as an innate immune defence by catalysing the oxidative deamination of an L-amino acid substrate to produce an alpha-keto acid, ammonia and hydrogen peroxide [11]. LAAO is also recognized as an important bioactive protein in defence against bacterial infections and cancers

* Corresponding author at: Department of Genetics, Faculty of Sciences, Kasetsart University, Bangkok 10900, Thailand.
E-mail address: fscitse@ku.ac.th (T. E-kobon).

https://doi.org/10.1016/j.csbj.2018.02.008

Please cite this article as: Suwannapan W, et al, Amplification and bioinformatics analysis of conserved FAD-binding region of L-amino acid oxidase (*LAAO*) genes in gastropods compared to other organisms, Comput Struct Biotechnol J (2017), https://doi.org/10.1016/j.csbj.2018.02.008

[12–16]. LAAOs are widely distributed in several organisms including bacteria, fungi, fish, snakes, and mammals as well as gastropods [11]. The protein structures of these enzymes and their related amine oxidases have been elucidated in some bacteria (*Streptomyces* sp. and *Streptococcus oligofermentans)* and reptiles (*Calloselasma rhodostoma*), showing three common domains: an FAD-binding domain, a substrate-binding domain, and a helical domain [17,18]. The FAD-binding domain consists of two conserved motifs, the FAD-binding motif (GxGxxG…hxhxE) and the GG motif (RxGGRxxS/T), where x indicates any amino acid residue and h indicates a hydrophobic residue [9]. The first and second glycine residues of the FAD-binding motif are important to the formation of the conserved structure of this motif, while the second glycine assists in making close contact with the phosphate group of the FAD molecule [19]. The first glycine of the GG motif is close to the FAD-binding motif, and the second residue interacts with the ribose molecule of the FAD cofactor [9].These conserved motifs form an initial beta-alpha-beta fold which is the most conserved part of the Rossmann fold structure of the LAAO proteins [20]. Variations within these two motifs have been previously reported, such as distinct conserved GG motifs in the LAAOs of *Streptococcus* spp. (RxGKK) and *Pseudoalteromonas* spp. (RxGGH) [21]. Variations within the non-conserved residues (x) of the FAD-binding motif were shown when comparing LAAOs from fish (*Sebastes schlegelii* and *Scomber japonicus*), a snake (*Pseudechis australis*) and gastropods (*Aplysia kurodai*, *Aplysia californica*, and *Achatina fulica*) [22], but a detailed analysis of these "x" residues has not been reported.

A previous phylogenetic analysis of 54 partial amino acid sequences inclusive of the FAD-binding region of LAAOs from bacteria and animals by Hughes [11] proposed the classification of LAAOs into two clusters. These two clusters were the mollusc-related subfamily and the vertebrate-related subfamily. The first subfamily included all LAAO sequences from a few gastropod species, three bacterial phyla (*Bacteroidetes*, *Firmicutes*, and *Proteobacteria*), and *Marseillevirus*. The second cluster included all sequences from vertebrates and three bacterial phyla (*Actinobacteria*, *Chloroflexi*, and *Firmicutes*). The occurrence of these two clusters suggested that the separation of the two subfamilies may have occurred when the two major clades of bacteria diverged, approximately three billion years ago [23]. Similarly, Campillo-Brocal et al. [10] reported that the phylogenetic relationships of LAAO enzymes from animals were separated into vertebrate-related and gastropod-related groups. These findings suggest the possible separate evolution of LAAO enzymes in the innate immune systems of both animal groups [11,24].

However, little information is known about the variation of LAAO enzymes in gastropods or in other invertebrates of the second cluster in the previous research due to limited number of the gastropod LAAO samples. Therefore, this study began investigating the FAD-binding region, which consists of the most conserved Rossmann fold of the *LAAO* genes in 12 gastropod species (representing prosobranch, pulmonate, and opisthobranch) together with nine other invertebrates commonly found in Thailand by gene amplification, molecular cloning, and nucleotide sequencing. The molecular relationships of this conserved region in gastropods with those in other organisms were analysed at the levels of nucleotide and amino acid sequences, and protein structures were compared by using phylogenetics and pattern analyses. The information of this conserved region would assist further investigation of sequences and structures of the gastropod LAAOs in more details.

2. Materials and methods

2.1. Selection and collection of gastropod and other invertebrate samples

Triplicate samples used in this study included 12 gastropod genera covering the three major groups (six pulmonates, six prosobranchs, and one opisthobranch), other classes of mollusks (five bivalves and one cephalopod considered as sister taxa) and other invertebrates in different phyla (two arthropods and one annelid) (see Supplementary Table 1). These commonly-found gastropod samples were collected from different habitats (lawns and gardens, freshwater ponds, forest, river and sea) in Thailand and were maintained in rearing chambers or preserved in 95% ethanol before the DNA extraction process. While the samples of commonly-found bivalves, cephalopods, and other invertebrates were selected and purchased from local fresh markets and preserved at −80 °C until use.

2.2. Amplification and sequencing of the conserved FAD-binding region in LAAOs

Genomic DNA was extracted using the GF-1 Nucleic Acid Extraction Kit (Vivantis Technologies, Oceanside, CA, USA). The FAD-binding and GG conserved motifs of the FAD-binding domain were amplified from the genomic DNAs using primers specific to the most conserved FAD-binding region (143-bp) of an *LAAO* gene, namely achacin in *A. fulica*: *LAAO2*-F (5′-TAGACGTTGCTGTGGTCGG-3′) and *LAAO2*-R (5′-GGGGAC GTTAGGCAAGTG-3′). The achacin was the first full-length *LAAO* gene that was sequenced and characterized in gastropod [25]. PCR was conducted using 50 ng of the genomic DNA, 2 mM of the forward and reverse primers, 2 mM of dNTPs**, 10× Buffer A**, 50 mM $MgCl_2$, and 5 U of Taq polymerase (Vivantis Technologies). Thermal cycling was performed with an initial denaturation step at 94 °C for 3 min, followed by 34 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 1 min, and a final extension step of 72 °C for 5 min using an Eppendorf MasterCycler EP Gradient Thermal Cycler (Eppendorf, Hauppauge, NY, USA). PCR products were separated by 2% agarose gel electrophoresis and stained with 0.8% ethidium bromide before visualization under UV light using a Gel Doc™ XR+ system (Bio-Rad, Hercules, CA, USA). The PCR products were then cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA). Ligation was conducted using 50 ng of the PCR products, 2× Rapid ligation buffer, 3 Weiss units/μl of ligase (Promega) and 50 ng of the vector. The ligation reaction was incubated overnight at 4 °C before mixing with 50 μl of *E. coli* DH5α competent cells and incubating on ice for 30 min. Transformation was conducted by heat-shocking at 42 °C for 1 min before placing on ice again for 2 min. Next, 400 μl of LB medium was added and the mixture was incubated in a shaker (Lab Companion, Shelburne, VT, USA) for 1 h at 37 °C and 180 rpm. Then, 300 μl of the transformation reaction was spread onto LA (Luria agar) plates containing 100 mg/mL of ampicillin (Vivantis Technologies), 0.1 M IPTG (Vivantis Technologies) and 20 mg/mL X-gal (Vivantis Technologies), and the plates were incubated overnight at 37 °C. White colonies were selected and checked by PCR amplification using the T7 (5′-TAATACGACTCACTATAGGG-3′) and SP6 (5′-TATT TAGGTGACACTATAG-3′) primers specific to the plasmid and the *LAAO2*-F and *LAAO2*-R primers. The transformed colonies harbouring the 143-bp products were picked and cultured in 5 ml of LB medium containing 100 mg/mL of ampicillin and incubated overnight on a shaker at 37 °C and 180 rpm. The recombinant plasmids were isolated from the cultured bacteria using BioFact™ Plasmid Mini Prep Kit (BioFact, Yuseong Gu, Daejeon, Korea). The nucleotide sequences of the conserved motifs were amplified and sequenced (Macrogen, Geumcheon-gu, Seoul, Korea).

2.3. Phylogenetic analyses of nucleotide and amino acid sequences of the conserved FAD-binding region of LAAOs in gastropods and other organisms

The newly obtained nucleotide sequences of the conserved FAD-binding sequence of *LAAO*s from gastropods and other invertebrates were aligned, edited and compared with those of other organisms downloaded from the National Center for Biotechnology Information (NCBI) database using MAFFT version 7 [26]. Two-hundred and six *LAAO* sequences obtained from the NCBI database were from eukaryotes, while four sequences were from the *LAAO* and its distantly-related genes

Please cite this article as: Suwannapan W, et al, Amplification and bioinformatics analysis of conserved FAD-binding region of L-amino acid oxidase (*LAAO*) genes in gastropods compared to other organisms, Comput Struct Biotechnol J (2017), https://doi.org/10.1016/j.csbj.2018.02.008