SHORT COMMUNICATION

Isolation and cloning of an aryl-aldehyde dehydrogenase gene from the white-rot fungus Pycnoporus cinnabarinus strain MUCL 39533

Khai Lun Ong1, Siew Ling Liew1*, Sahilah Abdul Mutalib1, Abdul Munir Abdul Murad2, Farah Diba Abu Bakar2*

1School of Chemical Sciences and Food Technology, 2School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor D.E., Malaysia.

Email: fabylf@ukm.edu.my; siew@ukm.edu.my

Received 31 December 2014; Received in revised form 23 March 2015; Accepted 4 April 2015

ABSTRACT

Aims: The white rot fungus Pycnoporus cinnabarinus MUCL 39533 is able to reduce vanillic acid to vanillin. Reduction of vanillic acid to vanillin catalysed by the key enzyme aryl-aldehyde dehydrogenase has been reported. Here we report the isolation and cloning of aryl-aldehyde dehydrogenase from P. cinnabarinus strain MUCL 39533.

Methodology and results: An aryl-aldehyde dehydrogenase gene (PcALDH) was isolated from P. cinnabarinus by producing a partial cDNA sequence fragment of an aryl-aldehyde dehydrogenase gene through PCR. Degenerate PCR primers were designed based on codons corresponding to conserved amino acid regions of aryl-aldehyde dehydrogenases of several fungi and bacteria. The full-length PcALDH cDNA was obtained through Reverse-Transcription-Polymerase Chain Reaction (RT-PCR) and Rapid Amplification cDNA Ends (RACE) PCR. PcALDH cDNA comprises an open reading frame of 1,506 bp that encodes a protein of 501 amino acids. The PcALDH predicted protein showed the highest amino acid sequence identity (84%) to ALDH from Trametes versicolor. In silico analysis of PcALDH indicated that it belongs to the ALDH super-family and Class 3 ALDH.

Conclusion, significance and impact study: PcALDH cDNA was successfully isolated and characterized. Important motifs identified from the highly conserved PcALDH protein indicated that it belongs to the aldehyde dehydrogenase super-family. The cDNA clone will be used in expression studies to confirm the catalytic function of the enzyme.

Keywords: Aldehyde dehydrogenase super-family, vanillin, vanillic acid

INTRODUCTION

Vanillin (4-hydroxy-3-methoxybenzaldehyde) is the main component of vanilla, which is a widely used flavouring in the food industry (Priefert et al., 2001). Vanillin is produced from the beans of Vanilla planifolia Andrews, a member of the orchid family (Orchidaceae) (Berger, 2007). Vanillin from vanilla beans has a very tenacious, creamy, vanilla-like odour, is intensely sweet and forms white, needle-like crystals (Priefert et al., 2001). More than 12,000 tons of vanillin are produced per year, but natural vanillin makes up only 1% of the total (Walton et al., 1996). According to Falconnier et al. (1994), the reduction of vanillic acid to vanillin can be catalysed by the extracellular aryl aldehyde dehydrogenase (ALDH) a result which has been confirmed by others (Frost, 2002). Similarly, the white rot fungus P. cinnabarinus is able to reduce vanillic acid to vanillin (Lesage-Meessen et al., 1996). Aryl aldehyde dehydrogenases are widespread amongst living organisms where they detoxify aromatic aldehydes to their corresponding carboxylic acids (Nakamura et al., 2010).

In filamentous fungi (Aspergillus niger and Pycnoporus cinnabarinus), vanillin is formed in a two-stage process in which ferulic acid is converted to vanillic acid then reduced to vanillin (Lesage-Meessen et al., 1996). As far as we are aware, ALDH of P. cinnabarinus has yet to be elucidated or isolated/cloned. Our aim here is to studies on the synthesis of vanillin in microbes (Mitsui et al., 2010).
isolate the gene and use it to produce bio-vanillin from vanillic acid.

MATERIALS AND METHODS

Strains

Pycnoporus cinnabarinus MUCL 39533 was obtained from the Mycothèque de l’Université Catholique de Louvain, Louvain, Belgium. The strain was stored at –20 °C in 10% (v/v) glycerol. Escherichia coli JM109 was obtained from Promega (Fitchburg, WI, USA) and stored at –80 °C.

Growth conditions

Pycnoporus cinnabarinus was grown on malt-extract agar (MEA) for 5 days after which the spores were transferred to and grown in basal medium containing 20 g/L maltose, 1.82 g/L di-ammonium tartarate, 0.5 g/L yeast extract, 0.2 g/L KH₂PO₄, 1.32 mg/L CaCl₂·2H₂O, 0.5 g/L MgSO₄·7H₂O and 2.5 mg/L thiamine hydrochloride (Gross et al., 1990) for 10 days. Mycelia were collected, mixed and fragmented using a Waring® blender. All steps were performed aseptically. Five millilitres of mycelial suspension was inoculated into 100 mL fresh basal medium in 250 mL Erlenmeyer flasks, incubated at 30 °C and shaken at 120 rpm. Three days later, 0.3 g/L vanillic acid was added and after 7 days of growth, the mycelia were harvested, frozen in liquid nitrogen and then kept at –80 °C (Falconnier et al., 1994; Ong et al., 2014).

Isolation and cloning of a partial aryl aldehyde dehydrogenase cDNA

Frozen mycelia were ground into powder and total ribonucleic acid (RNA) was isolated using TRIzol® reagent (Invitrogen/Life Technologies, Grand Island, NY 14072 USA). Reverse transcription (RT-PCR) was carried out using SuperScript™ III First-Strand Synthesis System (Invitrogen/Life Technologies, USA). The 20 µL reverse transcription mixture contained 1.5 µg of total RNA, 50 µM oligo(dT)₅, annealing buffer, 2× first-strand reaction mix and 200 U SuperScript™ III reverse transcriptase. Incubation was at 65 °C for 5 min, followed by 50 min at 50 °C for 50 min. Reactions were terminated by heating to 85 °C for 5 min. Partial PcALDH was amplified from cDNA fragments using Go Taq® Green Master Mix (Promega, USA). Twenty-five microlitres of the PCR reaction mixture contained 1× Go Taq Green Master Mix, 1 µL of template cDNA and 100 µM each of pALDHF (5'-CCNTGGAAAYTYYCCNNT-3') and pALDHR (5'-GNNCCRADAAGTYCCTC-3') degenerate primers. After an initial denaturation at 95 °C for 2 min, PCR conditions comprised 25 cycles of 95 °C for 1 min, 56 °C for 1 min and 72 °C for 1 min with a final extension at 72 °C for 5 min. The partial PcALDH PCR products were cloned into a pGEM-T Easy vector (Promega, USA) and the ligation products transformed into competent cells of Escherichia coli JM109. Plasmids were sequenced using the universal primer M13.

Isolation and cloning of full length aryl-aldehyde dehydrogenase cDNA

Total RNA was isolated using TRizol® reagent as described above, de-phosphorylated and de-capped. Ligation of RNA oligonucleotides (GeneRacer® RNA oligo) to the de-capped mRNA then followed. Subsequently, reverse transcription was carried out using SuperScript™ III RT (Invitrogen, USA) to generate the GeneRacer® cDNA. Amplification of 5’ and 3’ cDNA ends (RACE) was carried out using KOD Hot Start DNA polymerase (Merck KGaA, Darmstadt, Germany). Primers (10 µM) used for 5’ RACE were GeneRacer® 5’ (Life Technologies, USA) (5’-CGACTGGAAGAGGACGACT-3’ and 10 µM RPRACE (5’-TGAAGTGAAGGAGCAGCAGCCTAC GTCAGC-3’) and 10 µM FPRACE (5’-GTGCTCCAGGTTCCAACAACTACG3’-3’)) primers whereas 3’ RACE employed 10 µM GeneRacer® 3’ (5’-GCTGTCACGATCGCTA GATAAGC-3’) and 10 µM FPRACE (5’-GTGCTCCAGGTTCCAACAACTACG3’-3’) primers. In a 50 µL PCR reaction contained 1 U KOD Hot Start DNA polymerase (Merck KGaA, Germany), 10 µM of each primer, 10× KOD Hot Start DNA polymerase buffer, 25 mM MgSO₄, 2 mM dNTPs and 1 µL template cDNA. After initial denaturation at 95 °C for 2 min, 30 cycles of denaturation at 95 °C for 20 sec, annealing at 65 °C for 10 sec, extension at 70 °C for 20 sec with final extension at 70 °C for 5 min were performed. Finally, the full-length aryl-aldehyde dehydrogenase cDNA (PcALDH) fragment was amplified using 1 U KOD Hot Start DNA polymerase, 10 µM of F-ALDH (5’-AGCTGCCAGAAGGATCC-3’) and R-ALDH (5’-TGCTGTCACGATCGCTA GATAAGC-3’) primers, 10× KOD Hot Start DNA polymerase buffer, 25 mM MgSO₄, 2 mM dNTPs and 1 µL template cDNA. The PCR conditions were described as above. The full-length PcALDH PCR products were cloned into pGEM-T Easy vector (Promega, USA). Subsequently, the ligation products were transformed into competent cells of E. coli JM109. Plasmids were sequenced using the universal primer M13.

Bioinformatics analyses

Comparative sequence analyses of PcALDH were performed using NCBI (National Center for Biotechnology Information) blastx (Basic Local Alignment Search Tool) programme (http://blast.ncbi.nlm.nih.gov/) (Altschul et al., 1997). The theoretical iso-electric point and mass of PcALDH were computed using Compute pI/Mw (http://expasy.org/compute_pi/) (Bjellqvist et al., 1993, 1994; Gasteiger et al., 2005). The signal peptide of PcALDH was predicted using Phobius (http://phobius.sbc.su.se/) (Kall et al., 2007). Disulphide bonds of PcALDH were predicted using DIANNA (http://clavius.bc.edu/~clotelab/DIANNA/) (Ferré and Clote, 2005, 2006). Multiple sequence alignments of amino acid sequences of ALDHs were conducted using Clustal W2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/)
(Larkin et al., 2007) and BoxShade server version 3.21
(http://www.ch.embnet.org/software/BOX_form.html).

![Sequence alignment]

**Figure 1**: The full-length 1,506 bp sequence of the cDNA of aryl-aldehyde dehydrogenase of *Pycnoporus cinnabarinus* MUCL 39533 (GenBank Accession KP342121). Capital letters represent nucleotides while capital letters marked in bold show the translated amino acid sequence. Boxed ‘ATG’ and ‘TGA’ represent translational start and stop codons respectively.
Figure 2: Amino acid sequence Alignments ALDH of: *Pycnoporus cinnabarinus* (this study, GenBank Accession KP342121), *Trametes versicolor* (GenBank accession no. EIWI65328.1), *Coniophora puteana* (GenBank accession no. EIW76574.1), *Dichomitus squalens* (GenBank accession no. EJF56132.1), *Punctularia strigosozonata* (GenBank accession no. EIN03470.1) and *Fomitiporia mediterranea* NAD-dependent ALDH (GenBank accession no. EJDO4639.1). Asterisks and full stops indicate residues with identity and weak similarity, respectively. Invariant residues are putative active sites common to the ALDH super family and are indicated by arrows. The 16 residues conserved in the ALDH super family are indicated by capital letters. The positions of the ten motifs in the ALDH super family are shown at the top. The glycine motif involved in NAD(P)⁺ binding, G-x-G-x-G, is also shown (x positions without a clear residue consensus).
Figure 3: Phylogenetic tree of aryl-aldehyde dehydrogenases (ALDHs). The tree was constructed by neighbour-joining. The scale indicates an evolutionary distance of 0.1%. Bootstrap values (%) are indicated at the nodes. The GenBank accession numbers of the ALDH sequences are as followed: Pycnoporus cinnabarinus (this study, KP342121), Trametes versicolor (EJW65328.1), Dichomitus squalens (EJF56132.1), Punctularia strigosozonata (EIN03470.1), Coniophora puteana (EIW75974.1), Fomitiporia mediterranea (EJD04639.1), Laccaria bicolor (XP_001883222.1), Coprinopsis cinerea (XP_001834689.1), Auricularia delicata (EJD50206.1), Dacryopinax sp. (EJT99149.1), Stereum hirsutum (EIM91726.1), Puccinia graminis (XP_003333586.1), Trichosporonah sableh (EJT46540.1), Glomerella graminicola (EFO35179.1), Pseudozyma antarctica (GAC71960.1), Neosartorya fischeri (XP_001258260.1), Aspergillus fumigatus (XP_751026.1), Aciculosporium take (BAF57023.1), Trichoderma reesei (EGR48076.1), Verticillium dahlia (EGY19997.1), Aspergillus clavatus (XP_001268240.1), Metarhizium anisopliae (EFZ02176.1), Coccioidoides posadasii (EFW13700.1), Aspergillus kawachii (GAA87440.1), Cordyceps militaris (EGX90970.1), Aspergillus niger (XP_001392844.1), Gaeumannomyces graminis (EJT79328.1) and Homo sapiens (NP_000680.2).
RESULTS AND DISCUSSION
Cloning and sequence analysis

RNA was extracted from cultures 4 d after induction with vanillic acid. The first cDNA strand was produced from RNA via reverse transcription. A partial PcALDH fragment (710 bp) was amplified from the cDNA strands using the degenerate primers pALDHF and pALDHR. The resulting cDNA fragment was then translated and compared with protein databases of NCBI using blastx. The fragment had 64% identity to Aspergillus fumigatus AT293 (GenBank Accession XP_751026.1) ALDH mRNA at the amino acid level suggesting that the sequence encoded part of the ALDH gene of *P. cinnabarinus*.

Two gene-specific primers based on the partial cDNA sequence were designed (FPRACE and RPRACE) and used in RLM-RACE (RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends) to clone the full length cDNA of PcALDH. RACE-PCR amplification yielded 5' and 3' flanking regions of ~720 bp and ~630 bp respectively. Sequencing the entire construct showed a full length PcALDH cDNA of 1.7 kbp consisting of a 5' untranslated region, a 3' poly (A) tail and an open reading frame of 1,506 bp encoding ALDH (GenBank Accession KP342121) (Figure 1). The deduced amino acid sequence contained 501 residues with a calculated molecular mass of 55.04 kDa and a pI of 6.20. PcALDH contained three disulphide bonds and the absence of a signal peptide indicated that it was an intracellular protein.

The deduced amino acid sequence was used to perform BLAST searches which clearly showed that it belonged to the ALDH super family. The best matches were to: (a) hypothetical protein BN946_scf184803.g29 of *Trametes cinnabaria* (GenBank Accession CDO69831.1) (98% identity); (b) ALDH of *Trametes versicolor* (GenBank Accession EIW65328.1) (84% identity) and (c) ALDH of *Dichomitus squalens* (GenBank Accession EJF56132.1) (78% identity). Other matches of more than 65% identity were to ALDH of *Punctularia strigosozonata* (GenBank Accession EIN03470.1), NAD-dependent ALDH of *Fomitiporia mediterranea* (GenBank Accession EJD46391.1) and ALDH of *Coniophora puteana* (GeneBank Accession EIW73974.1). Sixteen amino acid residues were fully conserved in more than 95% of the sequences (Perzich et al., 1999; Lo and Chen, 2010; Yao et al., 2012) (Figure 2). PcALDH has two active-site residues Glu268 and Cys296 that are responsible for catalytic activity (Lo and Chen, 2010). A glycine motif, presumably involved in NAD(P)⁺ binding, GXGXXG (GYGHVG) was found between motifs 3 and 4. A phylogenetic tree based on 1,000 replicates (Figure 3) was generated using the neighbor-joining method and ALDH of *Homo sapiens* as the outlier. High bootstrap values supported the classification of each monophyletic group. As expected, Ascomycota and Basidiomycota clustered together into a group separate from other fungal ALDHs. Also, the PcALDH of *P. cinnabarinus* (Polyporaceae, Basidiomycota) was placed exactly where the taxonomic data suggested it should be.

ALDH catalysis involves acylation and deacylation. During acylation, a cysteine nucleophile interacts with the carbonyl carbon of aldehyde forming a thiohemiacetal intermediate, followed by hydride transfer from a tetrahedral thiohemiacetal intermediate to the pyridine ring of NAD(P)⁺. Then, deacylation occurs involving hydrolysis of the resulting thioester intermediate (Marchal et al., 2001; Lo and Chen, 2010; Munoz-Clares et al., 2011; Tsybovsky et al., 2013). Glu residues play important roles in activation of water molecules (Marchal et al., 2001). Activation of the catalytic Cys upon binding of the cofactor to the enzyme is via a local conformational rearrangement involving the Glu and cofactor (Wang and Weiner, 1995; Marchal et al., 2001). Glu268 and Cys296 of PcALDH are potential active site residues since they correspond to Glu268 and Cys302 of human liver ALDH and are conserved in the aligned enzymes (Figure 2) (Abriola et al., 1990; Wang and Weiner, 1995). Based on this evidence, we conclude that PcALDH is a typical member of the ALDH superfamily enzymes and that PcALDH may be a key enzyme in the conversion of vanillic acid to vanillin in *P. cinnabarinus*.

CONCLUSION

PcALDH cDNA was isolated and characterised. Motifs identified from the highly conserved PcALDH protein indicated that it belongs to the aldehyde dehydrogenase super-family. The cDNA clone will be used in expression studies to confirm the catalytic function of the enzyme.

ACKNOWLEDGEMENT

This research was supported by the Universiti Kebangsaan Malaysia Arus Perdana Research Grant, AP-2013-013. The authors wish to thank Professor William Broughton for his very useful comments and suggestions on this manuscript.

REFERENCES


types defined in a pH scale where isoelectric points correlate with polypeptide compositions. Electrophoresis 15, 529-539.


