Genome Mining and Expression Characterization of a Novel R-Mandelate Dehydrogenase: Post-print

Authors: Tang Cunduo, Shi Hongling, Ma Yue, DING Pengju, Xu Jianhe, Kan Yunchao, Yao Lunguang

Date: 2017-12-22T00:00:00+00:00

Abstract

Dehydrogenases are of great significance. To obtain an ideal R-mandelate dehydrogenase, this study employed genome mining technology to acquire a novel R-mandelate dehydrogenase, LhDMDH, from the Lactobacillus harbinensis strain. The specific activity of recombinant LhDMDH reached 1264.3 U/mg, approximately 4-fold that of the probe, positioning it at the leading edge among reported R-mandelate dehydrogenases. Concurrently, the primary enzymatic properties of four recombinant enzymes were characterized, with optimal reaction temperatures of 25~30°C and optimal pH values of 9.0~9.5. Kinetic parameter analysis revealed that the Kcat value of LhDMDH for the substrate was 30.28 s^{-1}, significantly higher than other recombinant enzymes. Moreover, substrate spectrum analysis demonstrated that LhDMDH exhibits superior advantages in the chiral resolution of racemic mandelic acid and the biosynthesis of phenylglyoxylic acid. This study achieved promising results in the gene mining of R-mandelate dehydrogenase, establishing a solid foundation for further modification and application, while also providing valuable experience for mining other enzymes.

Full Text

Gene Mining, Expression and Characterization of Novel R-Mandelate Dehydrogenases

TANG Cun-duo[1],[2]§, SHI Hong-ling[1]§, MA Yue[1], DING Peng-ju[1], XU Jian-he[2], KAN Yun-chao[1]**, YAO Lun-guang[1]****

(1 Henan Provincial Engineering Laboratory of Insect Bio-reactor and Henan Key Laboratory of Ecological Security for Water Source Region of Mid-line of South-to-North, Nanyang Normal University, Nanyang 473061, China)

(2 State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: R-mandelate dehydrogenase plays a key role in the biosynthesis of phenylglyoxylic acid, and mining novel R-mandelate dehydrogenases with high catalytic activity and stability holds significant importance. To obtain an ideal R-mandelate dehydrogenase, this study employed genome mining technology to isolate a novel R-mandelate dehydrogenase, designated LhDMDH, from *Lactobacillus harbinensis*. The specific activity of recombinant LhDMDH reached 1264.3 U/mg, approximately four times that of the probe enzyme, placing it at the leading level among reported R-mandelate dehydrogenases. Meanwhile, the key enzymatic properties of four recombinant enzymes were investigated, revealing optimal reaction temperatures of 25–30°C and optimal pH values of 9.0–9.5. Kinetic parameter analysis demonstrated that the Kcat value of LhDMDH toward its substrate was 30.28 S ^{1}, significantly higher than other recombinant enzymes. Furthermore, substrate spectrum analysis indicated that LhDMDH offers distinct advantages in the chiral resolution of racemic mandelic acid and the biosynthesis of phenylglyoxylic acid. This study achieved promising results in the genome mining of R-mandelate dehydrogenases, establishing a solid foundation for further modification and application while providing valuable experience for mining other enzymes.

Keywords: mandelate dehydrogenase, genome mining, expression, enzymatic characterization, biocatalysis

Phenylglyoxylic acid (PGA), also known as benzoylformic acid, belongs to the -keto acid family and serves as an important synthetic building block for various pharmaceutical intermediates [1, 2]. Due to its susceptibility to oxidation, decarboxylation, and decarbonylation, phenylglyoxylic acid is difficult to synthesize. Reported synthetic methods primarily include benzoyl cyanide hydrolysis, styrene oxidation, mandelic acid oxidation, Friedel-Crafts acylation, and biocatalytic synthesis [3], among which mandelate dehydrogenase plays a crucial role in the biosynthesis of phenylglyoxylic acid [4]. Mandelate dehydrogenases are widely distributed in microorganisms capable of metabolizing mandelic acid, with the enzymes from *Pseudomonas putida* and *Rhodotorula graminis* being the most extensively studied [5]. The former acts on S-mandelic acid, while the latter acts on R-mandelic acid. NAD -dependent mandelate dehydrogenases are mainly derived from yeast, *Lactobacillus*, and *Enterococcus* genera.

Yamazaki et al. isolated an R-mandelate dehydrogenase from *Streptococcus faecalis* IFO 12964, which existed as a homodimer with a specific activity of 8.85 U/mg, and conducted preliminary application studies [6]. Hummel obtained an R-mandelate dehydrogenase from *Lactobacillus curvatus* DSM 20019 with a specific activity of approximately 3 U/mg for the purified enzyme [7]. Fan et al. cloned an R-mandelate dehydrogenase from *Lactobacillus brevis*, achieving a catalytic activity of 330 U/mg for the recombinant enzyme toward R-mandelic acid [5, 8], and successfully applied it to the cascade synthesis of phenylglycine [8]. Despite significant improvements in the catalytic activity of R-mandelate

dehydrogenases through persistent efforts by researchers, these enzymes still fall short of meeting industrial production demands. Moreover, current reports on R-mandelate dehydrogenases lack investigations into temperature stability and organic solvent tolerance.

Nature harbors abundant unexploited novel enzyme resources, and how to discover these unknown enzymes and modify them appropriately to meet industrial demands has attracted increasing attention from researchers [9, 10]. To date, over 2000 microbial genome sequences have been determined and made publicly available, with this data continuing to grow [11], providing rich resources for novel enzyme discovery [12]. Genome mining refers to the exploitation of genomic information to identify new processes, targets, and products [13]. This technology bridges the gap from genomic databases to real enzyme databases, further enriching enzyme resources available for utilization or modification [10, 14]. Numerous successful cases of novel enzyme discovery through genome mining have been reported in recent years [15, 16]. This study aimed to mine novel R-mandelate dehydrogenases from reported genomic information using genome mining technology, followed by expression and characterization, to establish a solid foundation for subsequent rational modification, high-level expression, and application in biosynthesis.

1.1.1 Strains and Plasmids

Escherichia coli BL21(DE3) and *E. coli* DH5 /pET28a were preserved in our laboratory; *Lactococcus lactis* CICC 20209 and *Leuconostoc citreum* CICC 23234 were purchased from the China Center of Industrial Culture Collection; *Klebsiella oxytoca* was kindly provided by Professor Wu Jianping from Zhejiang University; *Staphylococcus aureus* was kindly provided by Dr. Shi Hongfei from Nanyang Normal University.

1.1.2 Reagents and Kits

Restriction endonucleases BamH I and Xho I were purchased from NEB; PrimeSTAR® HS (Premix) and DNA Marker were purchased from Takara; the One-Stop His-Tagged Protein Miniprep Kit from Beijing Tian Enze Gene Technology was used for rapid affinity purification of recombinant enzymes; R-mandelic acid, S-mandelic acid, R-o-chloromandelic acid, S-o-chloromandelic acid, L-lactic acid, and phenylglyoxylic acid were purchased from Shanghai Aladdin Biochemical Technology; NAD , NADP , NADH, and NADPH were purchased from Bangtai Bioengineering (Shenzhen); other reagents were of analytical grade from domestic or international sources.

1.1.3 Media

Skim milk medium: 12% skim milk powder, natural pH, sterilized at 113°C for 20 min, used for *Lactococcus lactis* cultivation; MRS medium: 1% casein peptone, 1% beef extract, 0.5% yeast powder, 0.5% glucose, 0.5% sodium acetate,

0.2% diammonium citrate, 0.1% Tween 80, 0.2% K HPO , 0.02% MgSO ·7H O, 0.005% MnSO ·H O, 2% CaCO , pH adjusted to 6.8, sterilized at 121°C for 20 min, used for *Leuconostoc citreum* cultivation; LB medium: 1% peptone, 1% NaCl, and 0.5% yeast extract, natural pH, sterilized at 121°C for 20 min, used for *E. coli* cultivation.

1.1.4 Software

DNAMAN 6.0 for simple sequence alignment and analysis; BioEdit for sequencing chromatogram analysis; Oligo 7 for primer design; Geneious 4.7.5 for gene and protein sequence storage and management; Clustal X2 for multiple sequence alignment; MEGA 6.0 for phylogenetic tree construction.

1.2.1 Genome Mining of Novel Mandelate Dehydrogenases

Using the well-characterized and highly active R-mandelate dehydrogenase from *Lactobacillus brevis* (LbDMDH) [8] as a probe, BLAST analysis was performed to identify a series of putative 2-dehydropantoate 2-reductases from genomic databases that had not been expressed or characterized. Phylogenetic trees were constructed and analyzed for these sequences, and 1–2 representative gene sequences from each branch were selected for cloning, expression, and characterization.

1.2.2 Primer Design and Gene Synthesis

Based on the above analysis, specific forward and reverse primers were designed for enzymes whose original strains were readily available, with BamH I and Xho I restriction sites added to the 5' ends. Primer sequences are listed in Table 1 and were synthesized by Suzhou Hongxun Biotechnology. For mandelate dehydrogenase genes from *Lactobacillus pasteurii*, *Staphylococcus succinus*, *Clostridium botulinum*, and *Lactobacillus harbinensis* (LpDMDH [WP_009559388.1], SsDMDH [WP_073505393.1], CbDMDH [WP_035784277.1], and LhDMDH [WP_027828400.1]) whose original strains were not easily accessible, the genes were codon-optimized and synthesized in full by Suzhou Hongxun Biotechnology.

1.2.3 Gene Cloning, Expression, and Purification of R-Mandelate Dehydrogenases

Genomic DNA was extracted from *S. enterica*, *L. lactis*, *S. aureus*, *K. oxytoca*, and *L. citreum*, and target genes were amplified by PCR using corresponding primers. The obtained genes and other synthesized genes were digested with BamH I and Xho I and ligated into similarly digested pET28a plasmid. The constructs were transformed into BL21 competent cells, followed by antibiotic selection, colony PCR verification, and sequencing confirmation. Expression and purification were performed according to the method of Ma et al. [17] with

minor modifications: induction was carried out with 0.1 mM IPTG at 16°C for 20 h, while other conditions remained unchanged.

1.2.4 Enzyme Activity and Protein Analysis

Mandelate dehydrogenase activity was determined using the method described by Fan et al. [8]. The reaction mixture (1 mL) contained 1 mM NAD (or NADP), 6 mM substrate (R-mandelic acid), and 100 mM glycine-NaOH buffer (pH 9.5). After incubation at 30°C for 2 min, 100 L of appropriately diluted enzyme solution was added, and the change in absorbance at 340 nm was measured over 1 min. One unit (U) was defined as the amount of enzyme required to produce 1 mol NADH (or NADPH) per minute under these conditions. Activity was calculated as: $U = EW \cdot V \cdot 1000/6220 \cdot L = EW/6.22$, where EW is the change in A per minute, V is the reaction volume (mL), 6220 is the molar extinction coefficient ($L \cdot mol^{-1} \cdot cm^{-1}$), and L is the path length (cm). Protein concentration was determined by the Bradford method using bovine serum albumin as standard. SDS-PAGE was performed using 12.5% separating gel, and apparent molecular weights were calculated using Quantity One software [18].

1.2.5 Temperature Characteristics of Recombinant Enzymes

Using the method described in Section 1.2.4, catalytic activities were measured at temperatures ranging from 20–60°C (at 5°C intervals) to determine optimal reaction temperatures. For thermal stability assessment, enzymes were incubated at 20–50°C (at 5°C intervals) for 1 h, after which residual activities were measured at their respective optimal temperatures, with activity of ice-bath-treated samples defined as 100%.

1.2.6 pH Characteristics of Recombinant Enzymes

Using the method described in Section 1.2.4, catalytic activities were measured at pH values from 7.0–11.0 (at 0.5-unit intervals) at their optimal temperatures to determine optimal pH values. For pH stability assessment, enzymes were incubated in buffers of pH 7.0–11.0 (at 0.5-unit intervals) on ice for 1 h, after which residual activities were measured at their respective optimal temperatures and pH values, with the highest residual activity defined as 100%.

1.2.7 Kinetic Parameters of Recombinant Enzymes

With coenzyme concentration fixed at 1 mM, substrate concentration was gradually increased to measure reaction rates at different substrate concentrations. Km, Kcat, and Vmax values were obtained by non-linear fitting using Origin 9.0 software. Similarly, with substrate concentration fixed at 10 mM, coenzyme concentration was gradually increased to determine kinetic parameters for the coenzyme.

1.2.8 Substrate Spectrum Analysis

Using the method described in Section 1.2.4, oxidation activities toward S-mandelic acid, R-o-chloromandelic acid, S-o-chloromandelic acid, and L-lactic acid were measured to provide theoretical basis for phenylglyoxylic acid biosynthesis and coenzyme recycling.

2.1 Gene Cloning of R-Mandelate Dehydrogenases

Using the method described in Section 1.2.3, four putative mandelate dehydrogenase genes (LlDMDH1, LlDMDH2, KoDMDH, and LcDMDH) were amplified from respective genomic DNA templates by PCR. As shown in Figure 1 [Figure 1: see original paper], the PCR products were relatively pure and were recovered using a PCR purification kit. Despite multiple attempts, genes SeDMDH and SaDMDH could not be successfully amplified. Additionally, four other putative mandelate dehydrogenase genes (LpDMDH, SsDMDH, CbDMDH, and LhDMDH) were obtained through full-gene synthesis. The eight obtained genes were digested and ligated into pET28a, transformed into BL21 competent cells, and selected on kanamycin plates. Colony PCR verification (Figure 2 [Figure 2: see original paper]) identified eight positive clones, which were sequenced by Suzhou Hongxun Biotechnology. The results confirmed that all eight gene sequences were identical to those deposited in NCBI.

2.2 Expression and Purification of Recombinant R-Mandelate Dehydrogenases

The eight recombinant strains harboring putative mandelate dehydrogenase genes were induced at low temperature with low IPTG concentration, using BL21/pET28a as negative control and BL21/pET28a-LbDMDH as positive control. Cells from 100 mL cultures were harvested by centrifugation, resuspended, and lysed by sonication. After high-speed centrifugation for 30 min, the supernatant was filtered through a 0.45 m membrane and purified using the One-Stop His-Tagged Protein Miniprep Kit. Imidazole was removed and samples were concentrated using 10 kDa centrifugal filter units, with final volumes adjusted to 1.5 mL. SDS-PAGE analysis of crude lysates and purified products revealed soluble expression of all eight putative mandelate dehydrogenases in BL21, with apparent molecular weights of approximately 38 kDa calculated using Quantity One software. Representative SDS-PAGE profiles of four active enzymes are shown in Figure 3 [Figure 3: see original paper].

Activity assays of crude enzymes showed expression levels of 28.3, 233.1, 1.4, and 1.1 U/mL culture for LbDMDH, LhDMDH, LcDMDH, and LlDMDH-2, respectively, under identical conditions. The remaining five genes were expressed but showed no R-mandelate dehydrogenase activity. Notably, LhDMDH expression level was nearly 10-fold higher than the probe LbDMDH. Specific activities of purified enzymes were 340.0, 1264.3, 40.6, and 11.7 U/mg for LbDMDH, LhDMDH, LcDMDH, and LlDMDH-2, respectively, with LhDMDH showing nearly

four-fold higher specific activity than the probe.

2.3 Temperature Characteristics of Recombinant Enzymes

Optimal reaction temperatures and thermal stability of the four recombinant enzymes were determined as described in Section 1.2.5, with results shown in Figure 4 [Figure 4: see original paper] and Figure 5 [Figure 5: see original paper], respectively. LhDMDH shared the same optimal temperature as the probe (30°C), while LcDMDH and LlDMDH-2 exhibited optimal activity at 25°C. All four enzymes retained over 80% activity after incubation at 20°C for 1 h. However, LhDMDH and LlDMDH-2 retained less than 50% residual activity after incubation at 30°C for 1 h, which may limit their application in the biosynthesis of phenylglyoxylic acid and its derivatives and necessitates future engineering for improved thermal stability.

2.4 pH Characteristics of Recombinant Enzymes

Optimal pH values and pH stability were determined as described in Section 1.2.6, with results presented in Figure 6 [Figure 6: see original paper] and Figure 7 [Figure 7: see original paper]. All four recombinant enzymes were alkalophilic, with LlDMDH-2 showing an optimal pH of 9.0 and the other three enzymes exhibiting optimal pH at 9.5. LbDMDH, LhDMDH, and LcDMDH remained stable across pH 9.5–10.5, retaining over 80% residual activity after 1 h incubation on ice, whereas LlDMDH-2 displayed a narrower pH stability range, being stable only at pH 9.0 and rapidly losing activity below pH 8.5 or above pH 9.5.

2.5 Kinetic Parameters of Recombinant Enzymes

Kinetic parameters for R-mandelic acid and NAD were determined as described in Section 1.2.8, with results summarized in Table 2 and Table 3 . LhDMDH exhibited a Kcat of 30.82 S^{1} and a Kcat/Km ratio of 28.80 $S^{1} \cdot mM^{1}$ for R-mandelic acid, significantly outperforming the other three recombinant enzymes. Notably, high concentrations of R-mandelic acid caused pronounced substrate inhibition in all four enzymes, with Ki values of 24.37, 26.40, 24.48, and 16.25 mM for LbDMDH, LhDMDH, LcDMDH, and LlDMDH-2, respectively. LhDMDH showed a Kcat of 29.05 S^{1} for NAD , similar to its value for the substrate, while the Km values for NAD were approximately one-tenth of those for the substrate across all four enzymes.

2.6 Substrate Spectrum of Recombinant Enzymes

Substrate specificity was investigated as described in Section 1.2.8 by measuring oxidation activities toward S-mandelic acid, R-o-chloromandelic acid, S-o-chloromandelic acid, and L-lactic acid (Table 4). All four enzymes showed

high activity toward R-mandelic acid but extremely low activity toward R-o-chloromandelic acid, indicating that substrate specificity engineering would be required for synthesizing o-chlorophenylglyoxylic acid. Importantly, none of the four recombinant enzymes exhibited activity toward S-mandelic acid or L-lactic acid, suggesting their suitability for chiral resolution of racemic mandelic acid and for establishing coenzyme recycling systems with L-lactate dehydrogenase.

Phenylglyoxylic acid is an important building block in chemical synthesis, and developing cost-effective, high-quality green processes for its biosynthesis holds significant research value and scientific importance. Mandelate dehydrogenase plays a crucial role in the biosynthesis of phenylglyoxylic acid [3], and mining highly active and stable mandelate dehydrogenases can substantially reduce production costs. In this study, genome mining was employed to screen ten putative mandelate dehydrogenase genes from vast genomic databases. Eight gene fragments were obtained through PCR amplification or full-gene synthesis and successfully expressed as soluble proteins in BL21. Among these, LhDMDH exhibited a remarkable specific activity of 1264.3 U/mg, nearly four times that of the probe enzyme and ranking at the forefront among reported natural R-mandelate dehydrogenases. Additionally, LhDMDH demonstrated superior potential in expression level and catalytic constant compared to the probe and other candidates, although its thermal stability requires improvement. Fortunately, with the emergence of various bioinformatics tools, particularly the maturation of molecular dynamics simulation technology [19], rational engineering for thermal stability has become increasingly feasible, which will be the focus of our future work. Moreover, substrate spectrum analysis revealed that LhDMDH lacks catalytic activity toward S-mandelic acid and L-lactic acid, a desirable property for chiral resolution of racemic mandelic acid and phenylglyoxylic acid biosynthesis. Theoretically, effective NAD recycling during chiral resolution of racemic mandelic acid could be achieved using sodium pyruvate as a co-substrate in coordination with L-lactate dehydrogenase, which warrants further investigation.

In summary, this study successfully obtained a novel R-mandelate dehydrogenase, LhDMDH, with excellent catalytic properties through genome mining, and systematically characterized its enzymatic properties, laying a solid foundation for high-level expression, rational modification, and application in phenylglyoxylic acid biosynthesis. Furthermore, this work enriches the successful application cases of genome mining technology and provides valuable experience for mining other enzymes.

Acknowledgments: We thank Professor Wu Minchen from Jiangnan University for technical support. We also acknowledge funding from the Youth Project of Nanyang Normal University (17359).

References:

[1] Zhang JL, Lei JH. Synthesis and Application of Benzoylformic acid. *Zhejiang Chemical Industry*, 2008(12):13-15.

[2] Dai ZH. Synthesis of benzoylformic acid and its methyl ester. Master degree thesis, Hefei University of Technology, 2012.

[3] Zhen Z. Cloning and expression of mandelate racemase gene and S-mandelate dehydrogenase. Master degree thesis, Nanjing University of Science and Technology, 2009.

[4] Wang J, Feng J, Li W et al. Characterization of a novel (R)-mandelate dehydrogenase from *Pseudomonas putida* NUST506 [J]. *J Mol Catal B: Enzym*, 2015, 120:23-27.

[5] Fan CW. Discovery and application of novel D-mandelate dehydrogenase multienzymatic synthesis of L-phenylglycine. Master degree thesis, East of China University of Science and Technology, 2014.

[6] Yamazaki Y, Maeda H. Enzymatic Synthesis of Optically Pure (R)-Mandelic Acid and Other 2-Hydroxycarboxylic Acids: Screening for the Enzyme, and Its Purification, Characterization and Use [J]. *Agricultural and Biological Chemistry*, 1986, 50(10):2621-2631.

[7] Hummel W, Schütte H, Kula M-R. Mandelic acid dehydrogenase from *Lactobacillus curvatus* [J]. *Appl Microbiol Biotechnol*, 1988, 28(4):433-439.

[8] Fan CW, Xu GC, Ma BD et al. A novel D-mandelate dehydrogenase used in three-enzyme cascade reaction for highly efficient synthesis of non-natural chiral amino acids [J]. *J Biotechnol*, 2015, 195:67-71.

[9] Tang CD, Shi HL, Yu XH. Gene Mining, Expression and Characterization of Novel GH5 Family -Mannanases. *Food Science*, 2016(11):90-96.

[10] Tang CD, Shi HL, Tang QH. Rencent trends in discovery and protein engineering of biocatalysts. *China Biotechnology*, 2014(9):113-121.

[11] Furuya T, Kino K. Genome mining approach for the discovery of novel cytochrome P450 biocatalysts [J]. *Appl Microbiol Biotechnol*, 2010, 86(4):991-1002.

[12] Gong J-S, Lu Z-M, Li H et al. Metagenomic technology and genome mining: emerging areas exploring novel nitrilases [J]. *Appl Microbiol Biotechnol*, 2013, 97(15):6603-6611.

[13] Tang CD, Shi HL, Tang QH et al. Genome mining and motif truncation of glycoside hydrolase family 5 endo-beta-1,4-mannanase encoded by *Aspergillus oryzae* RIB40 for potential konjac flour hydrolysis or feed additive [J]. *Enzyme Microb Technol*, 2016, 93-94:99-104.

[14] Xiao Z, Wu M, Grosse S et al. Genome mining for new alpha-amylase and glucoamylase encoding sequences and high level expression of a glucoamylase from *Talaromyces stipitatus* for potential raw starch hydrolysis [J]. *Appl Biochem Biotechnol*, 2014, 172(1):73-86.

[15] Zhu D, Mukherjee C, Biehl ER et al. Discovery of a mandelonitrile hydrolase from *Bradyrhizobium japonicum* USDA110 by rational genome mining [J]. *J Biotechnol*, 2007, 129(4):645-650.

[16] Barriuso J, Prieto A, Martínez MJ. Fungal genomes mining to discover novel sterol esterases and lipases as catalysts [J]. *BMC Genomics* 2013, 14:712.

[17] Ma CL, Tang CD, Shi HL et al. Semi rational modification of cephalosporin C acylase and biosynthesis of 7-ACA. *China Biotechnology*, 2015(12):65-71.

[18] Hu D, Tang C, Li C et al. Stereoselective Hydrolysis of Epoxides by reVrEH3, a Novel *Vigna radiata* Epoxide Hydrolase with High Enantioselectivity or High Complementary Regioselectivity [J]. *J Agr Food Chem*, 2017, 65(45):9861-9870.

[19] Tang CD, Li JF, Wei XH et al. Fusing a Carbohydrate-Binding Module into the *Aspergillus usamii* -Mannanase to Improve Its Thermostability and Cellulose-Binding Capacity by In Silico Design [J]. *Plos One*, 2013, 8(5):e64766.

Note: Figure translations are in progress. See original paper for figures.