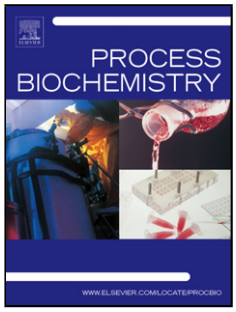


Stereoselective synthesis of the key intermediate of ticagrelor and its diverse analogs using a new alcohol dehydrogenase from *Rhodococcus kyotonensis*

Junmei Hu, Gudong Li, Chaoqun Liang, Saira Shams, Shaozhou Zhu, Guojun Zheng

PII: S1359-5113(19)31316-9

DOI: https://doi.org/10.1016/j.procbio.2020.01.015

Reference: PRBI 11900

To appear in: *Process Biochemistry*

Received Date: 30 August 2019

Revised Date: 17 December 2019

Accepted Date: 16 January 2020

Please cite this article as: Hu J, Li G, Liang C, Shams S, Zhu S, Zheng G, Stereoselective synthesis of the key intermediate of ticagrelor and its diverse analogs using a new alcohol dehydrogenase from *Rhodococcus kyotonensis*, *Process Biochemistry* (2020), doi: https://doi.org/10.1016/j.procbio.2020.01.015

Stereoselective synthesis of the key intermediate of ticagrelor and its diverse analogs using a new alcohol dehydrogenase from *Rhodococcus kyotonensis*

Junmei Hu[1,‡], Gudong Li[1,‡],Chaoqun Liang[1], Saira Shams[1], Shaozhou Zhu[1,*] and Guojun Zheng[1,*]

[1]State Key Laboratory of Chemical Resources Engineering, Beijing University of Chemical Technology, Beijing, People's Republic of China

[‡]These authors contributed equally.

*Correspondence

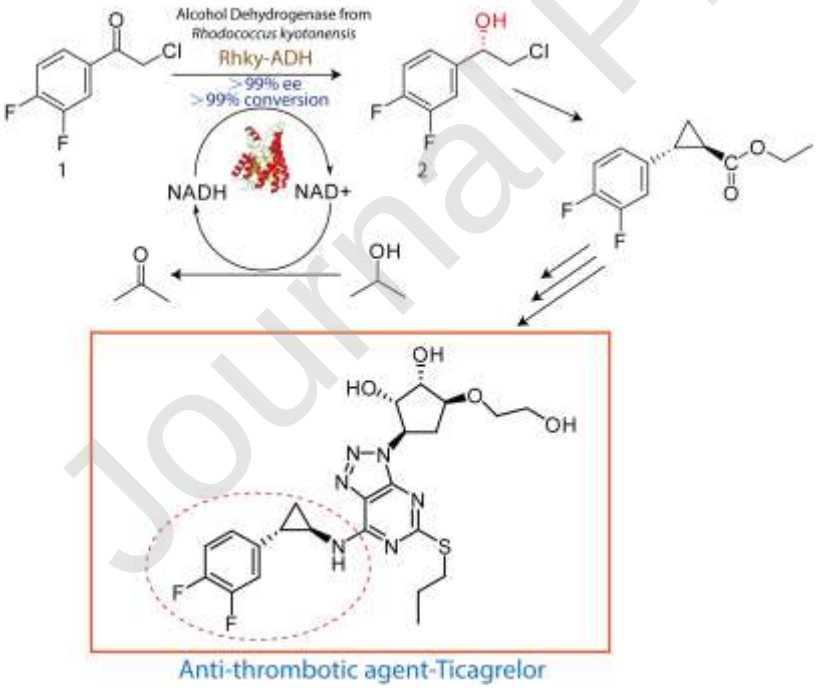
Shaozhou Zhu, E-mail: zhusz@mail.buct.edu.cn

Guojun Zheng, E-mail: zhenggj@mail.buct.edu.cn

State Key Laboratory of Chemical Resources Engineering, Beijing University of Chemical Technology, Beijing, People's Republic of China

Tel: +86 10 64437507.

Graphical abstract



Highlights

- •Discover a novel alcohol dehydrogenase (Rhky-ADH) from *R. kyotonensis.*
- •First enzyme from *Rhodococcus* for synthesis of key ticagrelor intermediates.
- •Rhky-ADH exhibited notable activity towards several aryl ketones.

Abstract

Bioreduction catalyzed by alcohol dehydrogenase/reductase is one of the most valuable biotransformation processes widely used in industry. The (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol is a key chiral synthon for synthesizing the antithrombotic agent ticagrelor. Herein, a new alcohol dehydrogenase (named Rhky-ADH) identified from *Rhodococcus kyotonensis* by an enzyme promiscuity-based genome mining method was successfully cloned and functionally expressed in *Escherichia coli*. The whole cell biocatalyst harboring Rhky-ADH was biochemically characterized and was shown to be able to convert 2-Chloro-1-(3, 4-difluorophenyl) ethanone to (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol with more than 99% enantiomeric excess (ee) and 99% conversion. Our data showed that the optimum temperature and pH for Rhky-ADH were 25 °C and pH 8.0, respectively. The addition of NADH and an appropriate concentration of isopropanol enhanced the activity of Rhky-ADH, and 1 mM Mn^{2+} increased the enzyme activity by about 8%. Substrate specificity experiments showed that Rhky-ADH had notable enzyme promiscuity and could reduce several ketones with high stereoselectivity. Our investigation on this novel enzyme adds another rare biocatalyst to the toolbox for producing chiral alcohols, which are widely used in the pharmaceutical industry.

Keywords: Alcohol dehydrogenase, (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol, Chiral alcohols, Whole-cell biocatalyst, Ticagrelor

1. Introduction

Chiral alcohols belong to a versatile class of chiral synthons and play a pivotal role in synthesizing modern pharmaceuticals and agricultural chemicals, since they can be incorporated into the active pharmaceutical ingredient structures directly as esters or ethers [1, 2]. For example, aromatic chiral alcohols are frequently used as building blocks to prepare diverse pharmaceuticals such as the antanacathartic drugs fosaprepitant, rolapitant, and aprepitant (Fig. 1), which function as NK-1 receptor antagonists and are very effective for treating nausea and vomiting caused by chemotherapy [3]; the selective cholesterol absorption inhibitor ezetimibe, which is used to treat high blood cholesterol by reducing plasma LDL cholesterol levels and increasing plasma HDL levels [4]; the anti-cancer drug crizotinib, which

acts as an anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) inhibitor and was diffusely used for treating non-small cell lung carcinoma [5]; and the platelet aggregation inhibitor ticagrelor, which is widely used for preventing thrombotic events [6].

The ever-growing demand for these chiral alcohols has increased in recent years, and therefore, new strategies to effectively prepare these compounds at a large scale have attracted more attention [7]. Current chemical approaches for preparing these chiral synthons require multiple reaction steps and involve the intensive use of metal catalysts and chiral auxiliaries [8]. For example, the chemical methods for preparation of chiral alcohol (*S*)-2-chloro-1-(3, 4-difluorophenyl) ethanol ((*S*)-CFPL) for ticagrelor synthesis are accomplished by the asymmetric reduction of the ketone precursor. These processes often involve the used of plenty of organic solvents and reducing agent [9]. Therefore, these processes normally could result in a large amount of waste and low environmental sustainability [10].Lately, the pharmaceutical industry has sought to address these challenges by embracing the principles of green chemistry [11]. Enzymes are largely employed in the chemical process, offering the possibility of highly enantio-, regio-, and chemoselective transformations with high catalytic turnover rates [12].

For the chemo-enzymatic synthesis of ticagrelor, several biocatalytic approaches have been developed [13-16]. The first approach requires the preparation of the cyclopropyl intermediate 3 of ticagrelor. This can be accomplished either by kinetic resolutions of the corresponding racemic amide by amidase or the racemic esters by lipase [15]. So far, only the whole cell biocatalyst *Rhodococcus rhodochrous* has been shown to have amidase activity and has been used to resolve cyclopropyl-containing carboxamides [15]. On the other hand, only the lipase from *Thermomyces lanuginosus* (TlL) showed notable activity toward the corresponding racemic ethyl ester [15]. Recently, Arnold et al. reported a novel biocatalytic approach for synthesizing the key cyclopropane intermediate to ticagrelor, employing a truncated globin from *Bacillus subtilis* [17]. The engineered enzyme could mediate the cyclopropanation of 3,4-difluorostyrene with ethyl diazoacetate with notable yield, high diastereoselectivity (> 99% dr) and enantioselectivity (98% ee), providing another green process for preparing ticagrelor [17]. Meanwhile, Fasan et al reported an enzymatic method to access chiral cyclopropane-containing intermediates via an engineered myoglobin catalysts. The reaction could be performed in gram-scale [16].

Among all the biocatalytic processes developed so far, the fourth strategy employing enzymatic asymmetric ketone reduction to synthesize chiral alcohol 2 represents another valuable strategy for ticagrelor biosynthesis since the chiral alcohol 2 can be easily transformed into the cyclopropyl amine (Fig. 2A) [7]. Compared with racemic resolution methods using

hydrolases (maximum yield: 49. 9%), there is no substrate loss with this strategy (maximum yield: 99.9%). Additionally, this biocatalytic process is more efficient compared with the racemic resolution methods whereas the activities of lipase or amidase used were very low, and thus more attractive for pharmaceutical manufacturing [1]. This strategy is increasingly being applied to synthesize ticagrelor intermediates. However, like other strategies, only a few biocatalysts have been reported for this process. In 2016, Turner et al. screened several ketoreductases from the Codexis Codex® KRED Screening Kit plus the RasADH from *Ralstonia* sp. and identified three promising biocatalysts with different enantioselectivity for ketone reduction of compound 1 [15]. Meanwhile, Wu et al. found that the ketoreductases from *Chryseobacterium* sp. CA49 could catalyze the reduction of ketone precursor 1 with excellent stereoselectivity (> 99 % ee) [18]. Through engineering, several mutants with significantly increased activity were discovered [18]. Very recently, Chen et al. also screened a set of ketoreductases, and several promising reductases were applied for ketone reduction of compound 1 [13].

Alcohol dehydrogenases belong to the 'classical' family within the short-chain alcohol dehydrogenase/reductase (SDR) superfamily, and are known to catalyze the reduction of a broad variety of ketones [19]. In this study, a rational genome mining strategy based on the natural function of a 1-(4-hydroxyphenyl)-ethanol dehydrogenase from *Aromatoleum aromaticum* EbN1 was designed and a new alcohol dehydrogenase capable of asymmetric reduction of ketone 2 from *R. kyotonensis* was discovered. The Rhky-ADH gene was synthesized and functionally overexpressed in *E. coli* BL21. Whole *E. coli* cells harboring Rhky-ADH were shown to be an ideal catalyst for the stereoselective reduction of several ketones. Remarkably, Rhky-ADH was used for the enantioselective reduction of 2-Chloro-1-(3, 4-difluorophenyl) ethanone to (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol with more than 99% enantiomeric excess (ee) and 99% conversion. Our study offers another efficient process for preparing the key intermediate of ticagrelor.

2. Material and methods

2.1 Materials

E. coli DH5*α* and BL21 (DE3) strains were purchased from TransGen Biotech (China) and used for gene cloning and expression respectively. Q5 High-Fidelity DNA Polymerase, restriction enzymes and T_4 DNA ligase were acquired from NEB (USA). Plasmid pET-28(+) was from Invitrogen. Purified Oligonucleotides were obtained from Sangon Biotech (Beijing, China). Antibiotics were from Sigma. All the kits used for DNA, plasmid and gel extraction

and genomic DNA purification were procured from Tiangen (Beijing, China). Other chemicals of analytical grade were purchased from commercial sources.

2-Chloro-1-(3, 4-difluorophenyl) ethanone (compound 1) were synthesized according to a previous established method [20]. Non-stereoselective reduction of 1 was accomplished by the same approach [20]. Rhky-ADH gene from *R. kyotonensis* was synthesized by Synbio Technologies (Suzhou, China). DNA dideoxy sequencing, accomplished by Ruibio Biotech (Beijing, China) was used to confirm the identity of all plasmids. Other ketones were obtained from Beijing Ouhe Technology CO., LTD (China).

2.2 Discovery of Rhky-ADH via an enzyme promiscuity-based genome mining method

To discover a useful biocatalyst for the stereoselective synthesis of the intermediate of Ticagrelor, we then examined the Genebank for a known enzyme with a similar substrate specificity. It was noticed that *A. aromaticum* EbN1 contains a pathway for the degradation of *p*-ethylphenol [21]. The Hped alcohol dehydrogenase (CAI06282.1) was then selected as a suitable inquiry to perform a PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool) in NCBI (National Center for Biotechnology Information). The searching database was limited to Non-redundant protein sequences (nr) and the organism was limited to *Rhodococcus* (taxid: 1827). The algorithm parameters (Max target sequences:500, Expect threshold:10 and Matrix: BLOSUM62) was taken. 500 promising dehydrogenases were identified by adopting this searching. The gene (WP_089252776.1) from *R. kyotonensis* named as Rhky-ADH was cloned into pUC57 and employed for further studies. Rhky-ADH was then amplified by polymerase chain reaction using pUC57-Rhky-ADH with primers as mentioned (Table 1). The PCR products were then purified. On the other hand pET28a were also amplified with primers listed in Table 1 and purified. These two PCR products were then ligated with Gibson Assembly Master Mix according to the manufacture's protocols. Thus, Rhky-ADH was cloned into pET28a with the *Nde*I and *Xho*I restriction sites, followed by addition of an N-terminal His6 which is origined from the pET28a vector. This construction leads to the production of an expression plasmid containing Rhky-ADH gene (pET28- Rhky-ADH). The recombinant plasmids were retransformed into *E. coli* DH5α cells and prepared from individual transformants. The right plasmids were confirmed by DNA sequencing.

2.3 Heterologous expression of Rhky-ADH in *E. coli* BL21

After the construction of the expression vector for Rhky-ADH, *E. coli* BL21 (DE3) cells was used to produce recombinant proteins. pET28-Rhky-ADH was retransformed into *E. coli* BL21 (DE3) cells by heat shock method [22]. Single colony harboring kanamycin resistance plasmid was inoculated in LB media (50 mL) and incubated overnight at 37 °C. 2 mL of the

overnight culture was then transformed into 200 mL (10 times) fresh LB medium supplemented with kanamycin (50 μg/L). The fermentation was performed at 37 °C until the optical density of culture was reached to 0.8, following cooling down of culture by keeping in ice for 15 min. Furthermore, the culture was then induced with 0.1 mM isopropyl-β-D-dthiogalactopyranoside (IPTG) with constant shaking for further 14 h at 23 °C. After fermentation, 1 mL of the culture was centrifuged (10000 rpm, 5 min) and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). 200 mL of the fermentation cultures (OD≈2.1), harvested by centrifugation (4000 rpm, 25 min) were resuspended with phosphate buffer (30 mL, pH, 7.0). 1 mL of the cells were then repacked to 2 ml Eppendorf tube and centrifugated. The pellets were stored at -80 °C for further use.

2.4 Synthesis of 2-chloro-1-(3, 4-difluorophenyl) acetone

8.4 g of Aluminum trichloride (31.50 mM, 0.9 eq) was added to 6.92 mL of 1, 2-difluorobenzene (36.06 mM, 1.0 eq) at room temperature. The mixture was heated at 50 °C and then 5.58 mL of chloroacetyl chloride (35.06 mmol, 1.0 eq) was added within 20 minutes [20]. Furthermore, the mixture was stirred for 60 minutes until the reaction was completed. After the completion of reaction, 30 mL of hydrochloric acid (1N) and a small amount of water were slowly added. The temperature was kept below 60 °C during the whole process. The reaction mixture was extracted with dichloromethane used in triplicate and the collected oil layer was washed with saturated NaCl solution. The organic layer was further dried with Na_2SO_4 and evaporated in vacuum. The crude product was obtained and further purified by silica gel column chromatography. Purified 2-chloro-1-(3, 4-difluorophenyl) acetone (5-10 mg) in $CDCl_3$ were then analyzed by ^1H NMR [20].

2.4 Enzymatic assays

Enzymatic reduction of 2-chloro-1-(3,4-difluorophenyl) acetone was achieved by the whole cells expressed with Rhky-ADH. In a typical assay, 100 mg of wet cells in phosphate buffer (1 mL , pH 7.0, 0.2 M) with 2-chloro-1-(3,4-difluorophenyl) acetone (6 mg/mL), NADH (0.5 mg/mL) and 40 μL isopropanol. Reaction with normal *E. coli* without induction was selected as control. All the reactions were performed at 25 °C for 24 h and extracted with 1 mL ethyl acetate. After centrifugation, the organic layer was analyzed by chiral HPLC. (Shimadzu system, controler; Shimadzu CMB-20A, diode array detector; Shimadzu SPD-20A, pumps; Shimadzu LC-20A) equipped with a Chiralcel IF column. UV absorption was monitored at 260 nm by employing the gradient with 90% acetonitrile and 10% isopropanol. Optically pure 2-chloro-1-(3,4-difluorophenyl) ethanol was used to determine the configuration of the products. The retention time for (*S*)-2-chloro-1-(3,4-difluorophenyl)

ethanol and (*R*)-2-chloro-1-(3,4-difluorophenyl) ethanol was 5.72 min and 6.706 min respectively.

2.5 Bioinformatics and structure analysis of Rhky-ADH

The sequences of putative alcohol dehydrogenases were download from NCBI (https://www.ncbi.nlm.nih.gov/). Phylogenetic trees of the alcohol dehydrogenases were constructed using the minimum evolution and neighbor joining method. Phylogeny was tested by bootstrapping with 1000 replications and the Poisson-correction amino acid model was used. Homology model of Rhky-ADH was constructed by the protein threading and fold recognition server I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/)[23]. 3D structures were compared by PyMOL. Multiple sequence alignments were performed by using the program ClustalW and visualized with ESPript 3.0 (http://endscript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) to identify strictly conserved amino acid residues. To validate the catalytic tetrad, an alanine scan was performed by site-directed ligation-independent mutagenesis (SLIM) [24].

2.6 Optimization of reaction conditions

The reaction conditions for Rhky-ADH were further optimized to develop a proper process for preparation of optically pure 2-chloro-1-(3,4-difluorophenyl) ethanol. The standard assay and factors affecting the enzymatic activity were investigated. In biocatalytic reactions, the concentration of substrate and catalysts could significantly affect the yield and conversion. Therefore, different concentrations of the substrates (10, 20, 30, 40, 50, 60, 80, 100 mg/mL) were first test using 100 mg/mL wet cells as catalysts. The assays were performed in phosphate buffer (1 mL, pH=7.0, 0.2 M) containing isopropanol (40 μL, 0.04 equiv) and 0.5 mg/mL NADH. The reaction mixtures were kept at 25 °C for 24 h with continuous shaking (220 rpm) and extracted with ethyl acetate (1 mL) . Moreover, the organic layer obtained by centrifugation was then analyzed by chiral HPLC. 2 mg/mL benzoylamide were added as a reference to determine the conversion. The experiments were performed in triplicate.

To probe the optimal concentration of biocatalysts used in the reaction, different concentration of wet cells (20, 30, 50, 100, 150, 200 mg) were suspended in 1 mL phosphate buffer containing 80 mg/mL 2-chloro-1-(3,4-difluorophenyl) acetone, 0.5 mg/mL NADH and 150 μL isopropanol (0.15 equiv). All the reactions were shaken at 25 °C for 24 h (220 rpm) and extracted with 1 mL ethyl acetate. Conversions were monitored by chiral HPLC. The experiments were performed in triplicate.

The effect of temperature on enzyme activity was also studied. A commonly used buffer (50 mM Tris, 50 mM boric acid, 33 mM citric acid and 50 mM Na_2PO_4, pH 8.0) was employed. 1 mL of this buffer containing 80 mg/mL substrate, 0.5 g/LNADH, 150 μL isopropanol (0.15 equiv) were assayed with 100 mg/mL wet cells under different temperature (15, 20, 25, 30, 35, 40 °C). Each reaction was performed at 24 h. Thermostability of Rhky-ADH was test under 25 °C, 30 °C or 35 °C. Similarly, the wet cells were incubated for different time intervals (1, 2, 3, 4, 8 and 12 h) and assayed with the substrate as described previous. Each reaction was performed at 220 rpm for 24 h. All the experiments were performed in triplicate.

Moreover, the effect of pH on enzyme activity was investigated using the same buffer. The pH was adjusted by NaOH or HCl. The reaction system was same as described above except the pH was from 6 to 11. The substrates are organic, oily compounds and are almost insoluble in water. Therefore, the effects of organic cosolvent were investigated using seven different kinds of organic solvents including dimethyl sulfoxide (DMSO), methanol (MeOH), ethanol (EtOH), isopropanol (IPA), N,N-dimethylformamide (DMF), tetrahydrofuran (THF) and acetonitrile (ACN). 100 μL of each cosolvent were added and assayed. The system lacking no cosolvent was taken as a control. All the experiments were performed in triplicate.

To investigate the optimal amount of isopropanol on enzyme activity, different volume of isopropanol (50-300 μL) were added in 1 mL assay buffer which contains 100 mg/mL wet cells, 80 mg/mL substrate and 0.5 g/LNADH. All the reactions were performed at 25 °C for 24 h (220 rpm) and analyzed as described before. All the experiments were performed in triplicate.

Cofactor is another important factor that could significantly affect the activity of Rhky-ADH. To study the effect of coenzyme amount on the reaction, four different concentration of NADH (0.5 mg/mL, 0.1 mg/mL, 0.05 mg/mL NADH and no coenzyme) were set up in the experiments. Time course of the reaction (0.5 h, 1 h, 2 h, 4 h, 12 h, 18 h, 24 h and 48 h) under these four conditions were also monitored. All the reactions were carried out at 25 °C for 24 h (220 rpm). The effects of metal ions (K^+, Ca^{2+}, Na^+, Mg^{2+}, Zn^{2+}, Fe^{2+}, Cu^{2+}, Mn^{2+}, Ni^{2+}, Co^{2+} and Li^+) on activity of Rhky-ADH were studied at a final concentration of 1 mM. System treated with EDTA was selected as a control. Other conditions are the same as described before. All the experiments were performed in triplicate.

2.7 Substrate specificity

To determine the substrate specificity of Rhky-ADH, seven different ketones including Acetophenone, 2-Chloro-1-(4-fluorophenyl) ethanone, 2-Chloro-1-(4-fluorophenyl) ethanone, 3,4-Difluoroacetophenone, 2-Fluoroacetophenone, 3-Chloroacetophenone, 3,5-bis(trifluoromethyl) acetophenone, 5-(4-Fluorophenyl)-5-oxopentanoic acid were selected and

assayed. Reactions were performed in 1 ml buffer (pH=7.0) containing 150 μL isopropanol, 100 mmol ketone, 0.5 mg/mL NADH and 100 mg wet cells and kept at 25 °C for 4 h with constant shaking at speed 220 rpm. After reaction the mixture was extracted with ethyl acetate and analyzed by chiral HPLC. For the analysis of substrates Acetophenone, 3-Chloroacetophenone and 5-(4-Fluorophenyl)-5-oxopentanoic acid, a Chiral OJ-H column was used. For the analysis of substrates 2-Chloro-1-(4-fluorophenyl) ketone, 3,4-Difluoroacetophenone and 2-Fluoroacetophenone, a Chiral IF column was used. The analysis 3, 5-bis(trifluoromethyl) acetophenone was accomplished by a Chiral OD-H column.

3. Results and discussion

3.1 Genome mining and identification of Rhky-ADH from *R. rhodochrous*.

Ticagrelor is one of the most popular medicines for preventing atherothrombotic events in adults with acute coronary syndromes [25]. Sales of this drug reach hundreds of millions of US dollars each year. However, only a few biocatalytic processes have been reported for preparing its key intermediates (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol (Fig. 2A) [13-15]. To develop an efficient bioprocess for preparing (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol, we focused on biocatalysts that could perform a similar reaction. We noticed that various (facultative) anaerobic bacteria have evolved a few strategies for biodegrading aromatic compounds. For example, *A. aromaticum* EbN1 has evolved a biosynthetic gene cluster that is responsible for the anaerobic degradation of *p*-ethylphenol (Fig. 2B). Past studies have shown that HpeD, which encodes a 1-(4-hydroxyphenyl)-ethanol dehydrogenase, can catalyze the reduction of 4-hydroxyacetophenone to 1-(4-hydroxyphenyl)-ethanol with (*R*)-specificity (Fig. 2C) [21]. These enzymes belong to the 'classical' family within the SDR superfamily and have great potential to be developed as biocatalysts for reducing 2-Chloro-1-(3, 4-difluorophenyl) ethanone and its analogs. Therefore, a genome mining strategy based on HpeD was developed and applied to search for new enzymes (Fig. 2D). The protein sequence of HpeD (accession number: CAI06282.1) was downloaded from the NCBI and used as a query for a Blast search. A past study showed that *Rhodococcus* spp. is a rich source of alcohol dehydrogenases/reductases [26]. Therefore, the position-specific iterated model was applied and the organism was limited to *Rhodococcus* (taxid: 1827). Using these settings, 500 protein sequences producing significant alignments with E-values above the threshold were discovered. Notably, 12 hypothetical reductases from *Rhodococcus* with more than 47% identities were selected as candidates for further investigation. First, a phylogenetic tree derived from the protein sequences was constructed by the neighbor joining method to illustrate the relative evolutionary history of the identified proteins (Fig. 3). The bioinformatics analysis

showed that all protein sequences are closely related, though four clear clades could be observed. To further confirm our hypothesis, the hypothetical reductase (Rhky-ADH, GenBank accession: WP_089252776.1) with low identity (48%) to HpeD from *R. kyotonensis* was chosen for further studies. *R. kyotonensis* is a gram-positive, non-spore-forming actinomycete that has the potential to degrade nitroaromatic compounds [27]. Rhky-ADH is a SDR family oxidoreductase that showed similarity to 3-ketoacyl-(acyl-carrier-protein) reductase, enoyl-(Acyl carrier protein) reductase, and short-chain alcohol dehydrogenase. Unlike HpeD, which is involved in the *p*-ethylphenol degradation pathway, Rhky-ADH does not belong to an obvious gene cluster. It is possible that this new reductase might be a novel family of alcohol dehydrogenases/reductases, and may be suitable for industrial application.

3.2 Heterologous expression of Rhky-ADH in *E. coli*.

To show that Rhky-ADH is indeed a new aromatic alcohol dehydrogenase, the Rhky-ADH gene from *R. kyotonensis* was directly synthesized by Synbio Technologies. The corresponding gene was then amplified by PCR and assembled into pET28a (+) by Gibson Assembly in the *Nde*I and *Xho*I sites. An N-terminal His_6-tag was thereby introduced into the recombinant protein. The resulting pET28a-Rhky-ADH was then transformed into *E. coli* BL21(DE3) and induced overnight with the addition of IPTG. One liter of the overnight culture was centrifuged and the pellets were resuspended in 30 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole; pH 8.0) and lysed by sonication on ice. The soluble supernatants were separated from the insoluble pellets by centrifugation and both were analyzed by SDS-PAGE to examine the expression of Rhky-ADH. As shown in Fig. 4A, Rhky-ADH was successfully expressed in *E. coli* with the predicted molecular mass of 27 KDa.

With the assistance of the N-terminal His_6-tag, Rhky-ADH could be readily purified by Ni-NTA affinity chromatography since most of the recombinant proteins were expressed in the soluble form. SDS-PAGE analysis showed that the purified Rhky-ADH was more than 95% pure. The purified protein was buffer exchanged using an Amicon Ultra filter and used for *in vitro* studies. Our previous studies on an alcohol dehydrogenase used for preparation of t-butyl 6-chloro-(3R,5S)-dihydroxyhexanoate showed that these dehydrogenases can easily become inactive after purification [28]. As expected, Rhky-ADH has the same properties and no activity was observed even though different parameters including cofactors were optimized. It is possible that the nicotinamide cofactors are not properly binding to the active site or the purification step is detrimental to the folding of Rhky-ADH. Nevertheless, using whole cells as the biocatalyst for alcohol dehydrogenase is a frequently used strategy since the expensive

nicotinamide cofactors can be easily regenerated and the biocatalysts are more stable inside cells [13]. Therefore, we directly used the whole cells expressing Rhky-ADH as biocatalysts and tested the bioactivity. The initial activity of Rhky-ADH was monitored by HPLC using 2-Chloro-1-(3, 4-difluorophenyl) ethanone as the substrate. As shown in Fig. 4B, the whole cells showed high enzymatic activity. Most of the 2-Chloro-1-(3, 4-difluorophenyl) ethanone was successfully converted into (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol as indicated by comparison to the reference compound. Notably, the enantiomeric excess was more than 99%. Thus, we successfully discovered a new alcohol dehydrogenase from *R. kyotonensis* that has potential for industrial application. The genome mining strategy based on enzyme promiscuity was shown to be feasible and could be used as a general method for discovering new enzymes.

3.3 Homology modeling and active sites of Rhky-ADH

After confirming that Rhky-ADH is indeed a 2-Chloro-1-(3, 4-difluorophenyl) ethanone reductase, bioinformatic studies combined with homology modeling were used to investigate the catalytic mechanism and enantioselectivity of this enzyme. A previous genome mining strategy identified 12 new alcohol dehydrogenases from *Rhodococcus* spp. Past studies showed that members of the SDR superfamily have a common $(\alpha/\beta)_8$-barrel fold and a similar catalytic tetrad [29]. Notably, the structure of HpeD, which is involved in the *p*-ethylphenol degradation pathway, was solved in 2015 and possesses a similar fold [21]. Sequence alignments of the identified 12 reductases and HpeD were performed. The alignments clearly showed that the active sites of these alcohol dehydrogenases are highly conserved (Fig. 5A). The catalytic tetrad of HpeD consists of the four amino acids Asn114, Ser142, Tyr155, and Lys159, which are also present in the other 12 reductases from *Rhodococcus* spp. and shared a similar catalytic mechanism. In Rhky-ADH, these four amino acids are Asn116, Ser144, Tyr 157, and Lys161. To better understand the mechanism of Rhky-ADH, a homology modeling structure was built by the protein threading and fold recognition server I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/). A structural model of Rhky-ADH was generated (Fig. 5B). The TM score for the Rhky-ADH model is 0.88 ± 0.07, indicating a correctly predicted overall topology [23]. Moreover, the root-mean-square deviation (rmsd) of the Cα atoms between the generated model and template was only 3.4 ± 2.4 Å, which further confirms the propriety of the structure. Generally, the modeling studies showed that the structure of Rhky-ADH is similar to that of similar proteins and should share a common catalytic mechanism [21]. Tyr 157 functions as the catalytic base with the help of Ser 144 to orient the substrate for hydride transfer. Lys 161 may lower the pKa of the OH in Tyr 155 by interacting with the nicotinamide ribose. Asn 116 is likely part of the proton relay system. To

confirm our hypothesis and validate the catalytic tetrad, an alanine scan was performed for those residues. The mutants were transformed and expressed in *E. coli* BL21 (DE3) as previously described and assayed again with 2-Chloro-1-(3, 4-difluorophenyl) ethanone to determine the reductase activity. The results were in agreement with our hypothesis: reductase activity was completely lost in all mutants. Our data confirmed that Asn116, Ser144, Tyr157, and Lys161 are indeed involved in the catalytic process.

To explain the enantioselectivity of Rhky-ADH, the modeling structure was compared with the alcohol dehydrogenase (Ped) with an opposite enantioselectivities from *A. aromaticum* EbN1 (PDB: 2EW8) [30]. As shown in Fig. 5C, those two alcohol dehydrogenases also shared a very similar fold. However, the Phe189 in the active site of Rhky-ADH (Leu186 in Ped) could confine the substrate-binding pocket to assure the (*R*)-stereospecificity of the hydrogenation reaction. On the other hand, the Tyr93 in Ped (Gly95 in Rhky-ADH), with opposite enantioselectivity, could completely change the binding pocket and force Ped to have a different enantioselectivity [30].

3.4 Effect of substrate concentration and biocatalyst loading

Rhky-ADH was shown to be capable of catalyzing the reduction of 2-Chloro-1-(3, 4-difluorophenyl) ethanone into (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol with an enantiomeric excess of more than 99%. The products could directly be used to synthesize the cyclopropyl amine intermediate 3 of ticagrelor. The high stereoselectivity make Rhky-ADH another rare biocatalyst suitable for biosynthesis of ticagrelor in industry. To develop an effective enzymatic process, we further optimized the reaction conditions for this biocatalytic system. When using cofactor-dependent enzymes, such as alcohol dehydrogenase, for biocatalysis, whole cells are frequently used as the catalysts. The main reasons for this are the low stability of such enzymes when they are outside cells, and the high cost of using cofactors in the process. Therefore, all reaction conditions were optimized using the whole cells as the catalysts. First, the effect of the substrate concentration on conversion with 100 mg/mL wet cells was investigated. As shown in Figure 6A, the conversion was above 90% when the substrate concentration was lower than 40 mg/mL. On the other hand, the conversion decreased to 65.75%±0.61 when the substrate concentration increased to 80 mg/mL. Therefore, 80 mg/mL substrate concentration was selected to study the influence of other different conditions on enzyme activity.

Next, the optimal concentration of biocatalysts was studied to maximize their utilization and reduce costs. As shown in Fig. 6B, when 100 mg/mL wet cells were used, the conversion

was more than 90% and increased slowly. Almost all the substrate was transformed when 200 mg/mL wet cells were used in the system. In order to use biocatalysts effectively, 100 mg/mL wet cells was selected as a suitable concentration for further experiments.

3.5 Effects of temperature and pH

We next determined the optimal temperature and pH for Rhky-ADH. Past studies have shown that *R. kyotonensis* can grow at 10–37 °C, with optimal growth at 27–30 °C [27]. The optimal pH for growth is around pH 7.0 [27]. The enzyme from this strain might have similar properties. As shown in Fig. 7A, the conversion rate increased slowly with rising temperature, and reached a peak at 25 °C. The conversion rate dropped quickly with further increases in temperature, and the conversion was only 40% at 40 °C, which is consistent with the fact that *R. kyotonensis* cannot grow at 40 °C. Rhky-ADH could theoretically perform the reaction at room temperature with high enantiomeric excess, which is significantly superior to traditional chemical methods.

Rhky-ADH showed high activity over a pH range about 7.0-10.0 (Fig. 7B). The conversion was about 55% at pH 6.0 and 30% at pH 11.0. The optimal pH for Rhky-ADH was found to be around pH 8.0, which is slightly different from the optimal pH for *R. kyotonensis* (pH 7.0). Since the whole cells are used to determine the optimal pH, this difference might due to that the intracellular pH is not in equilibrium with the external pH. Unexpectedly, we found that higher pH values led to decreased of ee values of the product. This may have occurred because some products racemize in over-alkaline conditions, or the selectivity of the enzyme may change depending on the pH value. Therefore, pH 8.0 was selected as a suitable condition for further studies.

To evaluate the thermostability of Rhky-ADH, three different temperatures were selected. The whole cells were incubated at 25 °C, 30 °C or 35 °C for various times in the assay buffer without any substrate. The biocatalysts were then assayed with the substrate and compared with the untreated Rhky-ADH. Our data showed that Rhky-ADH could maintain most of its activity under heat treatment (Fig. 8A). Rhky-ADH lost 40% of its activity at 30 °C after 12 h, indicating that it is not a very thermostable enzyme. Future works to improve its thermostability is necessary for further industrial applications.

3.6 Effects of an NADH cofactor and metal ions

2-Chloro-1-(3, 4-difluorophenyl) ethanone is only partially soluble in aqueous solutions. To improve the solubility of the substrates and enhance the activity of Rhky-ADH, seven

cosolvents were tested. The solvents DMSO, IPA, and THF had positive effects on enzyme activity while MeOH, EtOH, DMF, and ACN could inhibit the conversion. Isopropanol was the best cosolvent for Rhky-ADH. This might be due to the fact that isopropanol is also a hydrogen donor involved in the reduction process [13]. We next determined the optimal amount of IPA for use in the system. As shown in Fig. 8B, conversion was remarkably improved with increasing concentrations of IPA. The highest conversion was observed at 150 μL/mL. At higher concentrations of IPA, the activity of Rhky-ADH gradually decreased, indicating that high solvent concentrations could inactivate the enzyme.

NADH is another important cofactor for alcohol dehydrogenase [19]. Though whole cells contain certain some NADH, adding more NADH is an option to improve the enzyme activity. As expected, additional NADH remarkably strengthened the performance of Rhky-ADH (Fig. 9A). Besides, NADH is superior to NADPH (data not shown). When 0.5 mg/mL NADH was used, the conversion reached 88.37%±0.67 after 4 h. On the other hand, whole cells without additional NADH only exhibited 45.80%±1.23 conversion. Considering the cost of NADH, 0.5 mg/mL was selected as an appropriate concentration for future applications.

The effects of 11 different kinds of metal ions were evaluated at a concentration of 1 mM (Fig. 9B). Compared with the control, Mg^{2+}, Zn^{2+}, Cu^{2+}, and Mn^{2+} had a slight inhibitory effect on enzyme activity. K^{+} and Na^{+} had no effect, while others had positive effect on the activity of Rhky-ADH. Mn^{2+} increased enzyme activity by 7.59%±0.1. Based on these results, Rhky-ADH behaves more like a metal-independent enzyme. Therefore, no additional metal ions were added in subsequent assays.

3.7 Substrate spectra of Rhky-ADH

Rhky-ADH was discovered based on the promiscuity of the 1-(4-hydroxyphenyl)-ethanol dehydrogenase from *A. aromaticum* EbN1 [21]. Generally, enzymes belonging to the SDR superfamily are known to have a wide substrate specificity [19]. This example further showed that promiscuity-based genome mining may have a wide application in the future for searching novel dehydrogenases. We suspected that Rhky-ADH could also be used in other biocatalytic reactions. As shown in Fig. 1, diverse analogs of 2-Chloro-1-(3, 4-difluorophenyl) ethanol are also versatile intermediates for pharmaceutical industry. To study the substrate spectra of Rhky-ADH, 7 different ketones were examined (Table 2). Our data showed that the *E. coli* host cells have no activity towards these substrates. On the other hand, *E. coli* with Rhky-ADH

could catalyze most of the analogs with substitutions at various positions on the aromatic ring. The ee value and conversation were remarkably high. Only 5-(4-Fluorophenyl)-5-oxopentanoic acid showed no activity due to the long chain carboxylic acid. In summary, Rhky-ADH has the potential to be developed as a general catalyst for aromatic chiral alcohols, which are the most valuable intermediates used in the pharmaceutical industry [12].

4. Conclusion

With the quick development of synthetic biology and the advances in DNA synthesis, now it is much quicker and cheaper to synthesize new genes, and hence obtain new enzymes [31]. Enzymes are key components in the development of sustainable manufacturing processes for active pharmaceutical ingredients, such as the direct enzymatic asymmetric synthesis of chiral alcohols [7]. The inherent features of enzymes that catalyze complex reactions in a highly efficient, highly selective manner are significantly superior to the conventional heterogeneous and homogenous catalysts. Currently, the key focus of biocatalysis research is to find novel enzymes that enable new chemical transformations to be performed, thereby expanding the biocatalytic toolbox [32]. So far, a few strategies have been used to discover new biocatalysts with known or unknown substrate scopes, including bioinformatics methods, metagenomics methods, directed evolution, high-throughput screening methods, and de novo design [33]. Among all these strategies, genome mining represents one of the most efficient methods that has been widely used to discover new enzymes and biosynthetic pathways over recent decades. Several different genome mining methods have been developed, including homology-based genome mining, mass spectrometry-guided genome mining, transcription-factor centric genome mining, and precursor-centric genome mining strategies [33-36]. All these methods have led to a rapid increase in the number of novel enzymes that are available for biocatalysis and biosynthesis.

We began this study to look for suitable ketoreductases to design a process for synthesizing the important intermediates of ticagrelor. Generally, ketoreductases are used to stereoselectively reduce prochiral ketones to obtain optically pure alcohols in industry. Past studies have shown that conserved amino acid residues are essential for this family of enzymes and could be used for genome mining studies [1, 28]. To discover a new enzyme for stereoselective reduction of the corresponding ketones, we developed a modified homology-based genome mining method. Based on the natural function of a 1-(4-hydroxyphenyl)-ethanol dehydrogenase, HpeD, and the enzyme promiscuity, a potential alcohol dehydrogenase, Rhky-ADH, was identified [21]. The whole *E. coli* cells expressing Rhky-ADH were shown to be

useful for reducing 2-Chloro-1-(3, 4-difluorophenyl) ethanone. This clearly demonstrated that enzyme promiscuity-based genome mining methods could be developed as a general method for discovering novel enzymes for a given process. This is especially useful when biocatalysts for the process of interest are not currently known or represented. After the successful discovery of the biocatalysts, further substrate specificity studies further provide information on enzyme promiscuity and could be used to guide the search for new enzymes.

Rhky-ADH was functionally expressed in *E. coli* BL21 in a soluble form. However, the endogenous NADH regeneration system in *E. coli* makes the whole cells a more suitable catalyst for the process. When assayed with the whole cells, 2-Chloro-1-(3, 4-difluorophenyl) ethanone was reduced, giving (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol as the sole observable product. Further optimization showed that Rhky-ADH has an optimum temperature and pH of 25 °C and 8.0, respectively. Adding NADH, along with an appropriate concentration of IPA and metal ions including Mn^{2+}, was shown to enhance the activity of Rhky-ADH. Compared with other reported reductases, Rhky-ADH was shown to be a superior biocatalyst yielding the (*S*)-enantiomer of 2-Chloro-1-(3, 4-difluorophenyl) ethanol with an ee of 99.9% and almost 100% conversion after 4 h. By far, only a few reductases have such high activity and stereoselectivity such as the KR-01 from *Leifsonia* sp. and KR-12 from *Lactobacillus brevis*. Besides 2-Chloro-1-(3, 4-difluorophenyl) ethanone, Rhky-ADH was also very effective and highly enantioselective toward other similar ketones. 2-Chloro-1-(4-fluorophenyl) ethanone was also reduced to the corresponding (*S*)-alcohols, whereas Acetophenone, 3, 4-Difluoroacetophenone, 2-Fluoroacetophenone, 3-Chloroacetophenone, and 3, 5-bis(trifluoromethyl) acetophenone were converted into the corresponding alcohol with different stereobiases, giving (*R*)-alcohols with high enantioselectivity. Rhky-ADH cannot reduce 5-(4-Fluorophenyl)-5-oxopentanoic acid, indicating that the aliphatic chain in the ketones must fit to the active pocket of Rhky-ADH.F

In conclusion, a new alcohol dehydrogenase for enzymatic preparation of the key intermediate in ticagrelor and its diverse analogues with high enantioselectivity by reducing the corresponding ketones has been identified and expressed in *E. coli.* Our data suggests that whole cells have potential as useful catalysts for industry applications. Further customization and scaling-up could provide another useful process for generating diverse chiral alcohols. Moreover, the enzyme promiscuity-based genome mining method represents a promising strategy for identifying novel enzymes to meet the challenges of pharmaceutical development.

Author Statement

SZ and GZ designed the experiments. JH and GL performed the experiments. SS and CL provided experimental support. GZ and SZ analyzed the data and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was supported bythe National Megaproject for Innovative DrugsGreat Science and Technology Projects (20198ZX09721001-007-002), the Fundamental Research Funds for the Central Universities (XK1802-8 and XK1803-06), and National Natural Science Foundation of China (NSFC, Grant No. 21706005).

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

[1] Y. Ni, J.H. Xu, Biocatalytic ketone reduction: a green and efficient access to enantiopure alcohols, Biotechnol. Adv 30(6) (2012) 1279-88.
[2] G.W. Huisman, J. Liang, A. Krebber, Practical chiral alcohol manufacture using ketoreductases, Curr. Opin. Chem. Biol 14(2) (2010) 122-9.
[3] M. Aapro, A. Carides, B.L. Rapoport, H.J. Schmoll, L. Zhang, D. Warr, Aprepitant and fosaprepitant: a 10-year review of efficacy and safety, Oncologist 20(4) (2015) 450-8.
[4] A. Pandor, R.M. Ara, I. Tumur, A.J. Wilkinson, S. Paisley, A. Duenas, P.N. Durrington, J. Chilcott, Ezetimibe monotherapy for cholesterol lowering in 2,722 people: systematic review and meta-analysis of randomized controlled trials, J. Intern. Med 265(5) (2009) 568-80.
[5] J.E. Frampton, Crizotinib: a review of its use in the treatment of anaplastic lymphoma kinase-positive, advanced non-small cell lung cancer, Drugs 73(18) (2013) 2031-51.
[6] G. Lemesle, G. Schurtz, C. Bauters, M. Hamon, High on-treatment platelet reactivity with ticagrelor versus prasugrel: a systematic review and meta-analysis, J. Thromb. Haemost 13(6) (2015) 931-42.
[7] M. Kataoka, K. Kita, M. Wada, Y. Yasohara, J. Hasegawa, S. Shimizu, Novel bioreduction system for the production of chiral alcohols, Appl. Microbiol. Biotechnol 62(5-6) (2003) 437-45.
[8] T.J. Ward, K.D. Ward, Chiral separations: a review of current topics and trends, Anal. Chem 84(2) (2012) 626-35.
[9] J.-P. Dejonghe, K. Peeters, M. Renard, Chemical process for preparation of aromatic cyclopropane esters and amides, Google Patents, 2008.

[10] T.J. Ward, K.D. Ward, Chiral separations: fundamental review 2010, Anal. Chem 82(12) (2010) 4712-22.
[11] M. Poliakoff, J.M. Fitzpatrick, T.R. Farren, P.T. Anastas, Green chemistry: science and politics of change, Science 297(5582) (2002) 807-10.
[12] S.M. Thomas, R. DiCosimo, V. Nagarajan, Biocatalysis: applications and potentials for the chemical industry, Trends. Biotechnol 20(6) (2002) 238-42.
[13] X. Guo, J.-W. Tang, J.-T. Yang, G.-W. Ni, F.-L. Zhang, S.-X. Chen, Development of a Practical Enzymatic Process for Preparation of (S)-2-Chloro-1-(3,4-difluorophenyl)ethanol, Org. Process. Res. Dev 21(10) (2017) 1595-1601.
[14] K.G. Hugentobler, M. Rasparini, L.A. Thompson, K.E. Jolley, A.J. Blacker, N.J. Turner, Comparison of a Batch and Flow Approach for the Lipase-Catalyzed Resolution of a Cyclopropanecarboxylate Ester, A Key Building Block for the Synthesis of Ticagrelor, Org. Process. Res. Dev 21(2) (2017) 195-199.
[15] K.G. Hugentobler, H. Sharif, M. Rasparini, R.S. Heath, N.J. Turner, Biocatalytic approaches to a key building block for the anti-thrombotic agent ticagrelor, Org. Process. Res. Dev 14(34) (2016) 8064-8067.
[16] P. Bajaj, G. Sreenilayam, V. Tyagi, R. Fasan, Gram-Scale Synthesis of Chiral Cyclopropane-Containing Drugs and Drug Precursors with Engineered Myoglobin Catalysts Featuring Complementary Stereoselectivity, Angew. Chem. Int. Ed. Engl 55(52) (2016) 16110-16114.
[17] K.E. Hernandez, H. Renata, R.D. Lewis, S.B. Jennifer Kan, C. Zhang, J. Forte, D. Rozzell, J.A. McIntosh, F.H. Arnold, Highly Stereoselective Biocatalytic Synthesis of Key Cyclopropane Intermediate to Ticagrelor, ACS Catalysis 6(11) (2016) 7810-7813.
[18] F.J. Zhao, Y. Liu, X.Q. Pei, C. Guo, Z.L. Wu, Single mutations of ketoreductase ChKRED20 enhance the bioreductive production of (1S)-2-chloro-1-(3, 4-difluorophenyl) ethanol, Appl. Microbiol. Biotechnol 101(5) (2017) 1945-1952.
[19] Y.G. Zheng, H.H. Yin, D.F. Yu, X. Chen, X.L. Tang, X.J. Zhang, Y.P. Xue, Y.J. Wang, Z.Q. Liu, Recent advances in biotechnological applications of alcohol dehydrogenases, Appl. Microbiol. Biotechnol 101(3) (2017) 987-1001.
[20] H. Zhang, J. Liu, L. Zhang, L. Kong, H. Yao, H. Sun, Synthesis and biological evaluation of ticagrelor derivatives as novel antiplatelet agents, Bioorg. Med. Chem. Lett 22(11) (2012) 3598-602.
[21] I. Busing, H.W. Hoffken, M. Breuer, L. Wohlbrand, B. Hauer, R. Rabus, Molecular Genetic and Crystal Structural Analysis of 1-(4-Hydroxyphenyl)-Ethanol Dehydrogenase from '*Aromatoleum aromaticum*' EbN1, J. Mol. Microbiol. Biotechnol 25(5) (2015) 327-39.
[22] H. Inoue, H. Nojima, H. Okayama, High efficiency transformation of *Escherichia coli* with plasmids, Gene 96(1) (1990) 23-28.
[23] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, Nat. Methods 12(1) (2015) 7-8.
[24] J. Chiu, P.E. March, R. Lee, D. Tillett, Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h, Nucleic. Acids. Res 32(21) (2004) e174-e174.
[25] J. Li, W. Tang, R.F. Storey, S. Husted, R. Teng, Population pharmacokinetics of ticagrelor in patients with acute coronary syndromes, Int. J. Clin. Pharmacol. Ther 54(9) (2016) 666-74.
[26] E. Martinez-Rojas, T. Kurt, U. Schmidt, V. Meyer, L.A. Garbe, A bifunctional enzyme from *Rhodococcus erythropolis* exhibiting secondary alcohol dehydrogenase-catalase activities, Appl. Microbiol. Biotechnol 98(22) (2014) 9249-58.
[27] B. Li, K. Furihata, L.X. Ding, A. Yokota, *Rhodococcus kyotonensis* sp. nov., a novel actinomycete isolated from soil, Int. J. Syst. Evol. Microbiol 57(Pt 9) (2007) 1956-9.

[28] T. Xu, C. Wang, S. Zhu, G. Zheng, Enzymatic preparation of optically pure t-butyl 6-chloro-(3R,5S)-dihydroxyhexanoate by a novel alcohol dehydrogenase discovered from *Klebsiella oxytoca*, Process. Biochcem 57 (2017) 72-79.
[29] U. Oppermann, C. Filling, M. Hult, N. Shafqat, X. Wu, M. Lindh, J. Shafqat, E. Nordling, Y. Kallberg, B. Persson, H. Jornvall, Short-chain dehydrogenases/reductases (SDR): the 2002 update, Chem-Biol. Interact 143-144 (2003) 247-53.
[30] H.W. Hoffken, M. Duong, T. Friedrich, M. Breuer, B. Hauer, R. Reinhardt, R. Rabus, J. Heider, Crystal structure and enzyme kinetics of the (S)-specific 1-phenylethanol dehydrogenase of the denitrifying bacterium strain EbN1, Biochemistry 45(1) (2006) 82-93.
[31] P.N. Devine, R.M. Howard, R. Kumar, M.P. Thompson, M.D. Truppo, N.J. Turner, Extending the application of biocatalysis to meet the challenges of drug development, Nat. Rev. Chem 2(12) (2018) 409-421.
[32] M. Ferrer, F. Martinez-Abarca, P.N. Golyshin, Mining genomes and 'metagenomes' for novel catalysts, Curr. Opin. Biotechnol 16(6) (2005) 588-93.
[33] T. Davids, M. Schmidt, D. Bottcher, U.T. Bornscheuer, Strategies for the discovery and engineering of enzymes for biocatalysis, Curr. Opin. Chem. Biol 17(2) (2013) 215-20.
[34] J.S. Gong, Z.M. Lu, H. Li, Z.M. Zhou, J.S. Shi, Z.H. Xu, Metagenomic technology and genome mining: emerging areas for exploring novel nitrilases, Appl. Microbiol. Biotechnol 97(15) (2013) 6603-11.
[35] S. Zhu, G. Zheng, Transcription-factor centric genome mining strategy for discovery of diverse unprecedented RiPP gene clusters, bioRxiv (2018) 467670.
[36] H. Li, Y. Qiu, C. Guo, M. Han, Y. Zhou, Y. Feng, S. Luo, Y. Tong, G. Zheng, S. Zhu, Pyrroloindoline cyclization in tryptophan-containing cyclodipeptides mediated by an unprecedented indole C3 methyltransferase from *Streptomyces* sp. HPH0547, Chem. Commun 55(58) (2019) 8390-8393.

Figure legends

Fig. 1 Chiral alcohol building blocks used for diverse pharmaceuticals.

Fig. 2 A Schematic diagram of the chemo-enzymatic synthesis of Ticagrelor using ketoreductase. B, Gene clusters that for the responsible for the anaerobic degradation of *p*-ethylphenol. C. Biosynthetic pathways for the degradation of *p*-ethylphenol. D, Using HpeD as inquiry for genome mining of new alcohol dehydrogenase for reduction of 2-Chloro-1-(3, 4-difluorophenyl) ethanone.

Fig. 3 Phylogenetic tree indicating evolution relationships between HpeD and newly identified alcohol dehydrogenases. HpeD (CAI06282.1) was marked in red and the newly identified Rhky-ADH (WP_089252776.1) was marked in blue.

Fig. 4 A, SDS-PAGE (12%) of Rhky-ADH expressed by *E. coli* BL21 (DE3). Lane 1, Protein marker; Lane 2, *E. coli* BL21/pET28-Rhky-ADH before induction; Lane 3, After induction; B.

HPLC analysis of the bioreduction of 2-Chloro-1-(3, 4-difluorophenyl) ethanone to (*S*)-2-Chloro-1-(3, 4-difluorophenyl) ethanol. The reference was used to monitor the retention time.

Fig. 5 A. Alignments of alcohol dehydrogenase from this study and other alcohol dehydrogenase from bacteria; B. Homology model of Rhky-ADH (sky blue) overlaid with the studied structure (5H5X) of the carbonyl reductase from *Streptomyces coelicolor,* the active pocket was shown in the box. The catalytic tetrad and the binding NADH was shown on the right; C. Comparison of the structure model of Rhky-ADH (sky blue) with the alcohol dehydrogenase (Ped), residues that might be essential for the stereoselectivity was shown in detail.

Fig. 6 A. Effect of substrate concentration on the bioreduction of 2-Chloro-1-(3, 4-difluorophenyl) ethanone. Reaction assays: different concentrations of ketone 1, 100 mg/mL wet cells, 0.5 mg/mL NADH, 40 μL IPA and phosphate buffer (1 mL, pH 7.0) for 24 h at 25 °C and 220 rpm; B. Effect of biocatalyst loading on the influence of conversion. Reaction assays: different amount of wet cells, 80 mg/mL of ketone 1, 0.5 mg/mL NADH, 150 μL IPA and phosphate buffer (1 mL, pH 7.0) for 24 h at 25 °C and 220 rpm

Fig. 7 A. Effect of temperature on the activity of Rhky-ADH. Reaction assays: 80 mg/mL of ketone 1, 100 mg/mL wet cells, 0.5 mg/mL NADH, 150 μL IPA and common buffer (1 mL, pH 8.0) for 24 h at different temperature and 220 rpm； B. Effect of pH on the activity of Rhky-ADH. Reaction assays: 80 mg/mL of ketone 1, 100 mg/mL wet cells, 0.5 mg/mL NADH, 150 μL IPA and common buffer with different pH for 24 h at 25 °C and 220 rpm.

Fig. 8 A, Thermolstability of Rhky-ADH under the thermal treatment at Activities of KR-01 to thermal treatment at 25, 30 and 35 °C. Reaction assays: 100 mg/mL wet cells were incubated for different time intervals (1, 2, 3, 4, 8 and 12 h) at 25 °C, 30 °C or 35 °C and assayed with 80 mg/mL of ketone 1, 0.5 mg/mL NADH, 150 μL IPA and common buffer (1 mL, pH 8.0) for 24 h at 25 °C and 220 rpm; B, Effect of concentration of IPA on the activity of Rhky-ADH. Different volume of isopropanol (50-300 μL) were added in 1 mL assay buffer and assayed with 100 mg/mL wet cells, 80 mg/mL substrate and 0.5 mg/mLNADH for 24 h at 25 °C and 220 rpm.

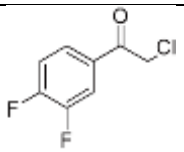
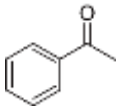
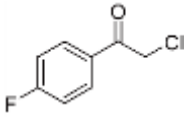
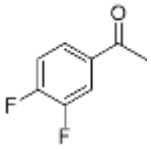
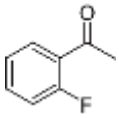
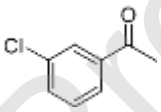
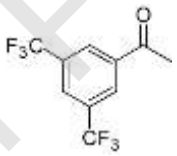
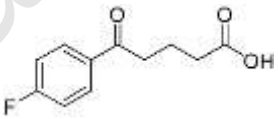
Fig. 9 A, Effect of NADH loading on the activity of Rhky-ADH. Reaction assays: Different concentration of NADH were added in 1 mL assay buffer and assayed with 100 mg/mL wet cells, 80 mg/mL substrate for 24 h at 25 °C and 220 rpm; B, Effect of metal ions on the conversion. Reaction assays: 1 mM of different metal ions were added in 1 mL assay buffer

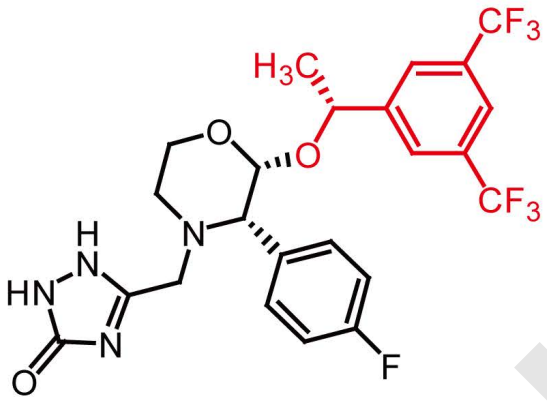
and assayed with 100 mg/mL wet cells, 80 mg/mL substrate and 0.5 mg/mL NADH for 24 h at 25 °C and 220 rpm.

Table. 1 Primers used in this study

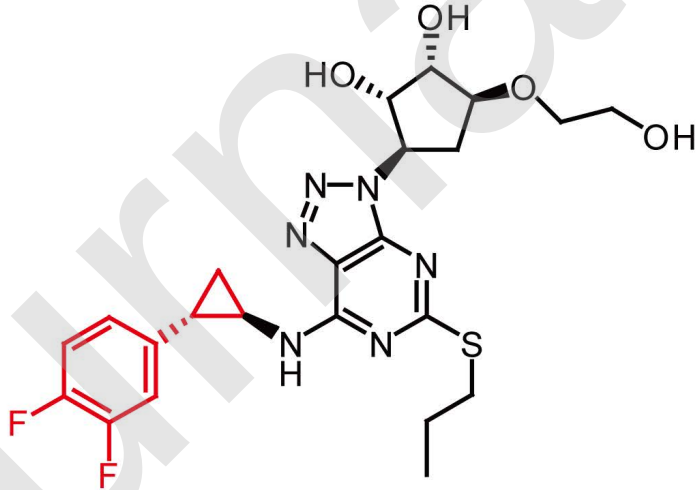
Name	DNA sequence (5'-3')	Description
Rhky-ADH-FP	ATGGCCACGTACGACGTCGCGGATCGTTCG	Rhky-ADH amplification
Rhky-ADH-RP	CTACTGCGCTGCGTACCCGCCGTCGACGAG	Rhky-ADH amplification
pET28 FP	GCGGGTACGCAGCGCAGTAGCTCGAGCACCA CCACCACCACCACTGAGATCC	pET28a amplification
pET28 RP	GCGACGTCGTACGTGGCCATATGGCTGCCGCG CGGCACCAGGCC	pET28a amplification
N116A P1	GAC CTT GCG CCA GCT GTC GAC G	N116A mutation
N116A P4	ATC GAC GTC GCA CTC AAC GCC GTG TTC	N116A mutation
N116A P2	GTG TTC TAC GGA ATG CAG AAG CAG ATC CC	N116A mutation
N116A P3	GGCGTTGAGTGCGAC GTC GAT GAC CTT GCG CCA GC	N116A mutation
S144A P1	GAC GAT CGA TCC ACC GCC CGC T	S144A mutation
S144A P4	AAC ATG GCG GCA GTT CTG GGC AGC GT	S144A mutation
S144A P2	AGC GTC GGC ATC GCG AAC TCC	S144A mutation
S144A P3	AAC ATG GCG GCA GTT CTG GGC AGC GT	S144A mutation
Y157AP1	CGC GGA GTT CGC GAT GCC GAC	Y157A mutation
Y157A P4	AGCGCAGTG GCT GCG AAG CAC GGA CTC CT	Y157A mutation
Y157AP2	GGA CTC CTC GGC CTC ACC AAA ACC GCC GCA CTC G	Y157A mutation
Y157A P3	GTGCTTCGCAGCCACTGCGCT CGC GGA GTT	Y157A mutation
K161A P1	CGC GGA GTT CGC GAT GCC GAC GCT G	K161A mutation
K161A P4	AGCTACGTGGCTGCGGCACACGGA CTC CTC	K161A mutation
K161A P2	GGA CTC CTC GGC CTC ACC AAA ACC G	K161A mutation
K161A P3	GTGTGCCGC AGC CAC GTA GCT CGC G	K161A mutation

Table 2. Substrate specificity of Rhky-ADH. All the experiments were performed in triplicate.

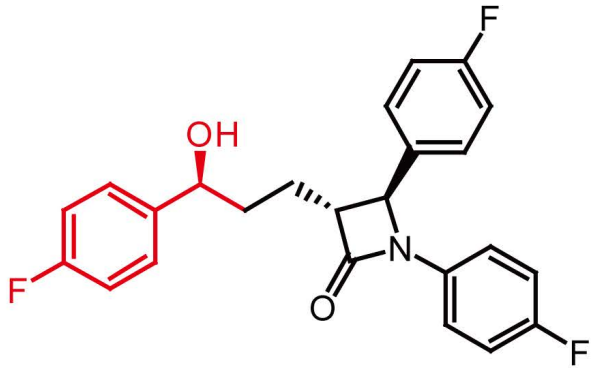
No	Substrate	Structure	Relative activity (%)	e.e or d.e (%)
1	2-Chloro-1-(3, 4-difluorophenyl) ketone		100	>99.9 (S)
2	Acetophenone		76.3±1.3	98.4 (R)
3	2-Chloro-1-(4-fluorophenyl) ketone		98.5±0.2	>99.9 (S)
4	3,4-Difluoroacetophenone		89.5±3.5	>99.9(R)
5	2-Fluoroacetophenone		40.8±1.8	92.4 (R)
6	3-Chloroacetophenone		77.0±0.4	99.5(R)
7	3, 5-bis(trifluoromethyl) acetophenone		46.4±1.1	>99.9(R)
8	5-(4-Fluorophenyl)-5-oxopentanoic acid		n.d.	n.d.



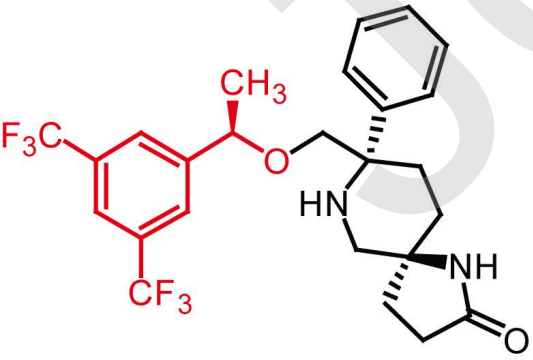
Aprepitant



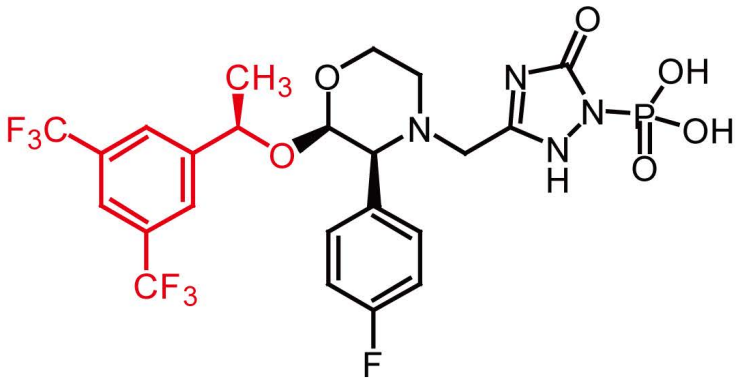
Ticagrelor



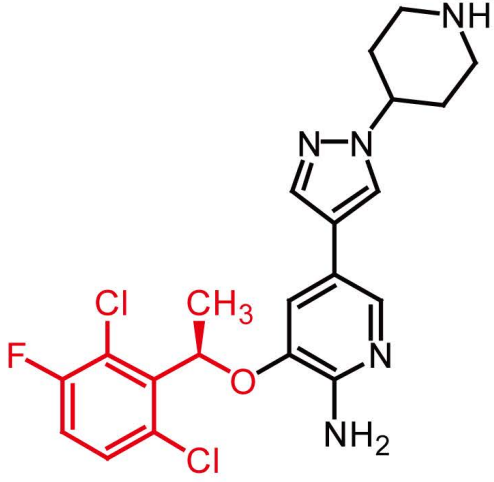
Ezetimibe



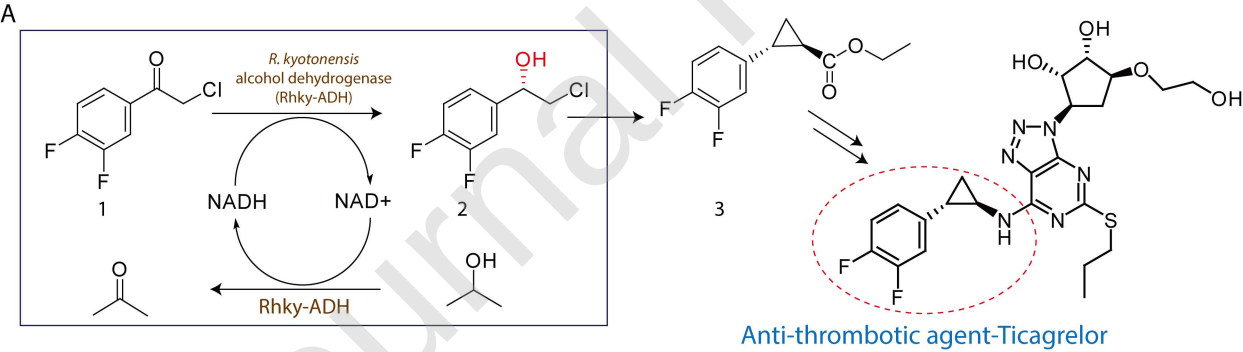
Rolapitant

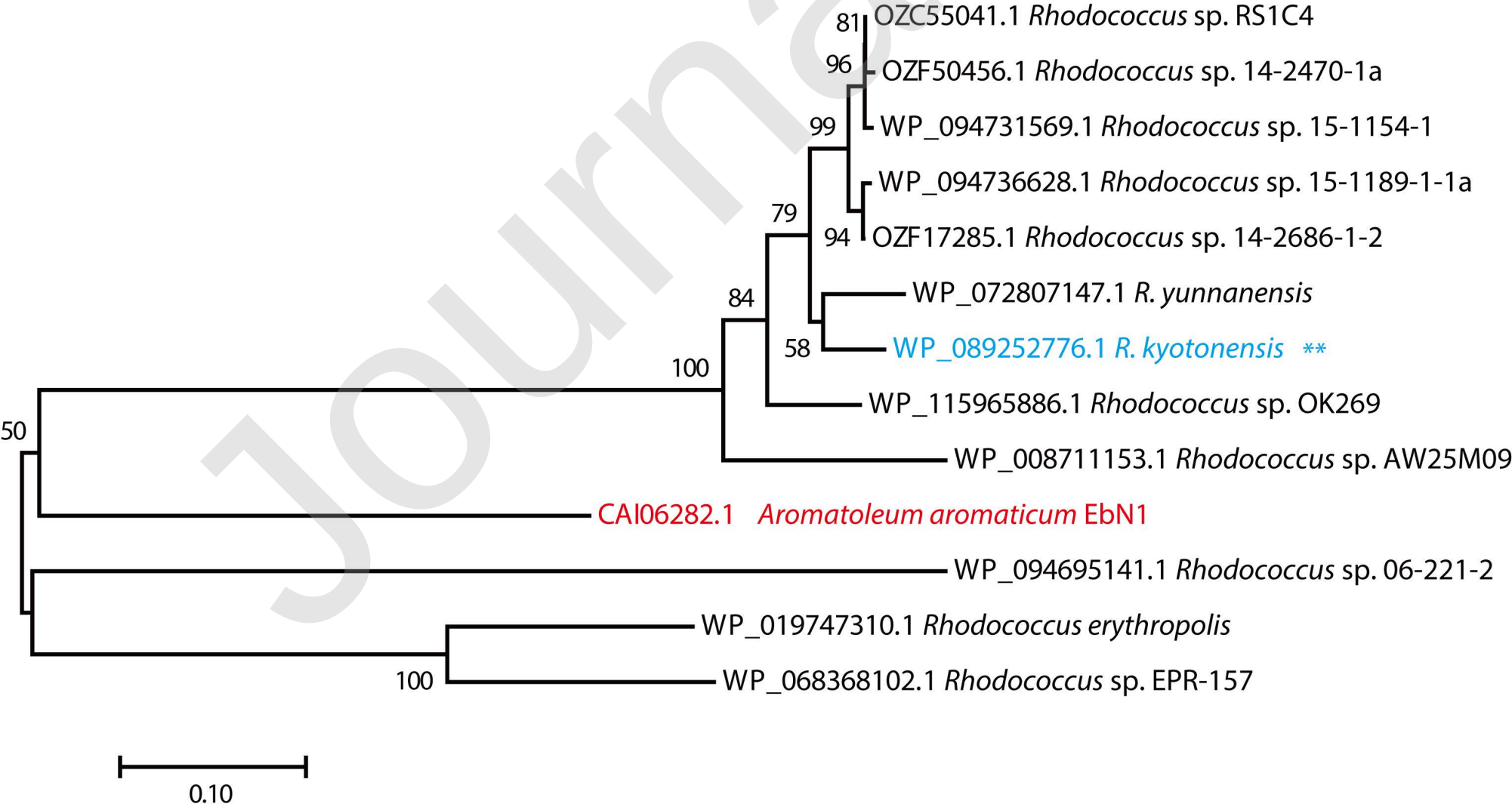


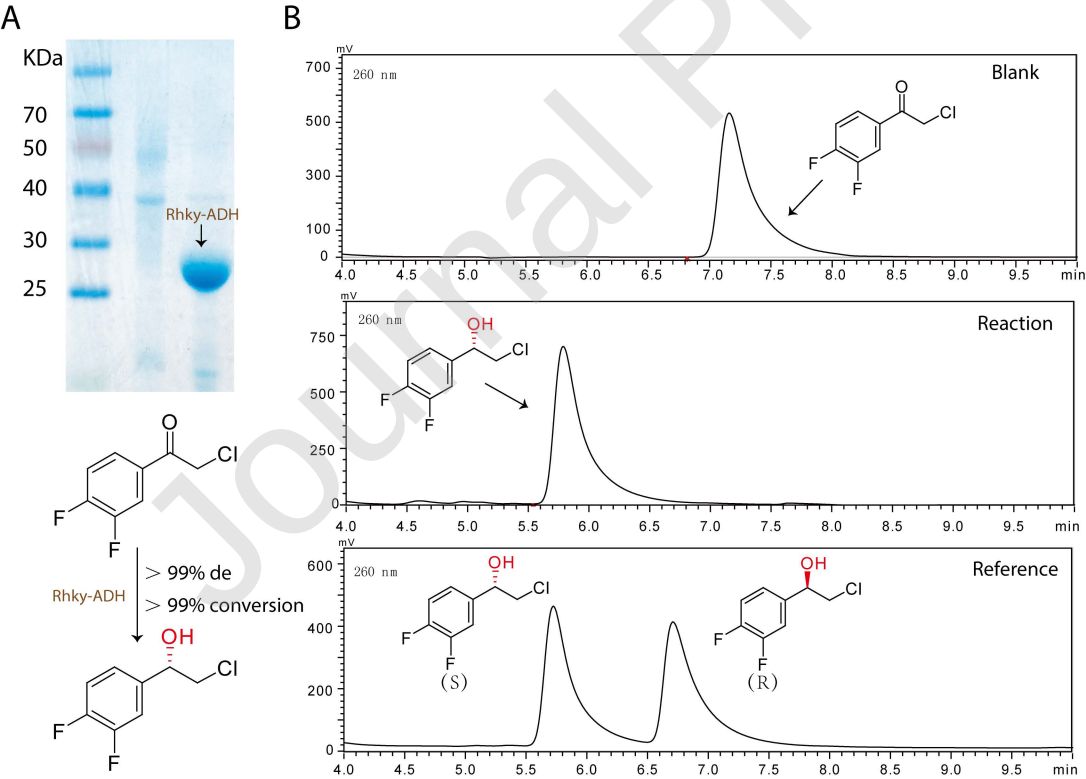
Fosaprepitant

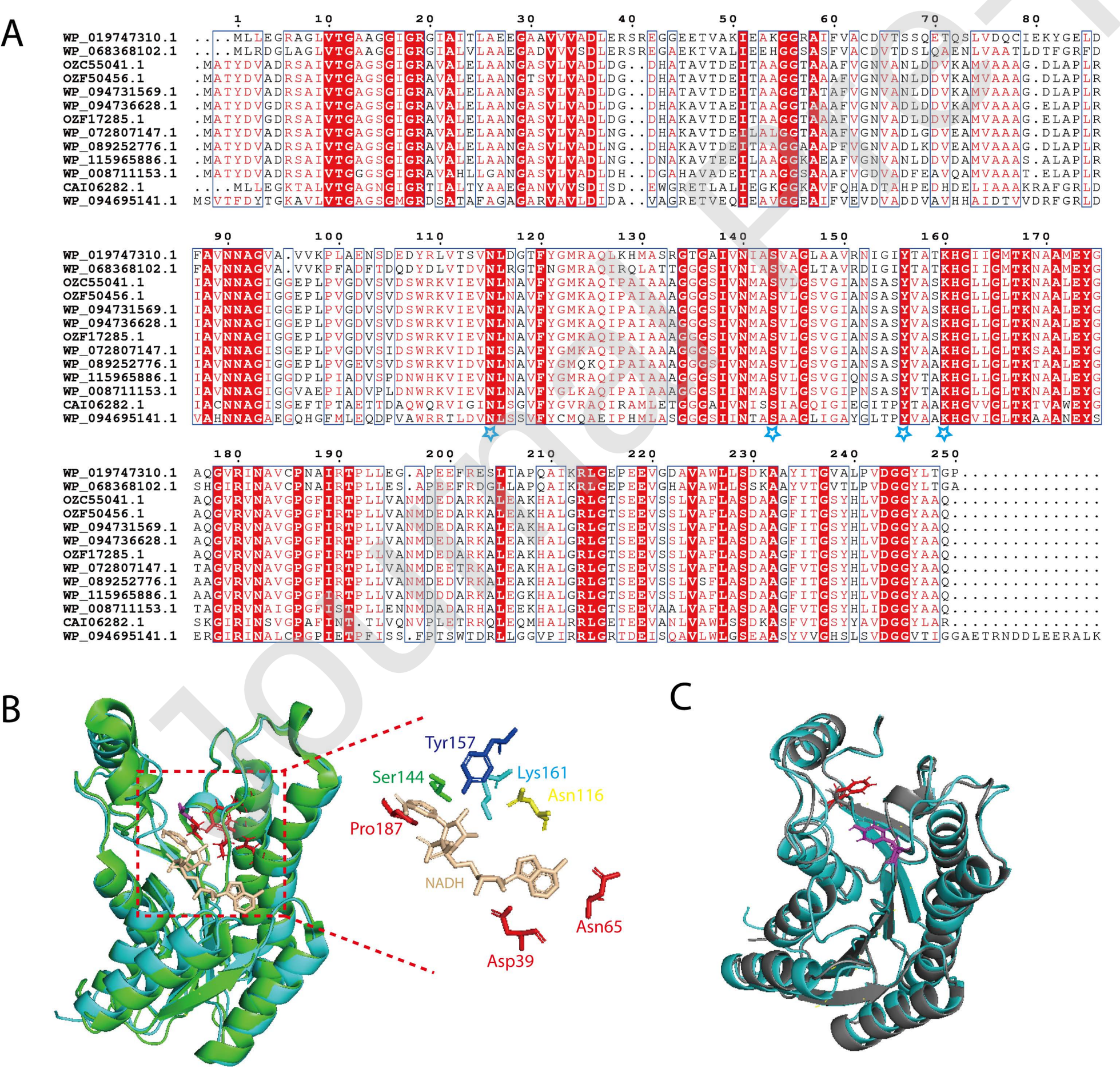


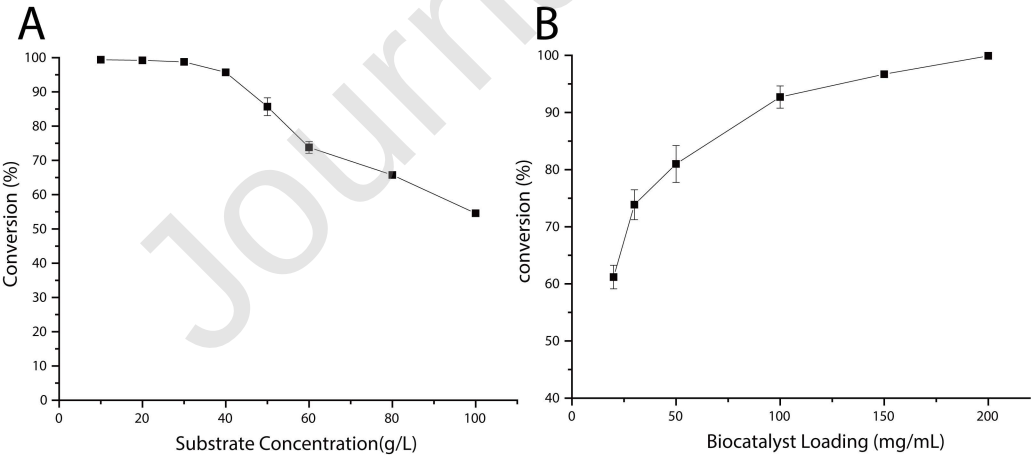
Crizotinib

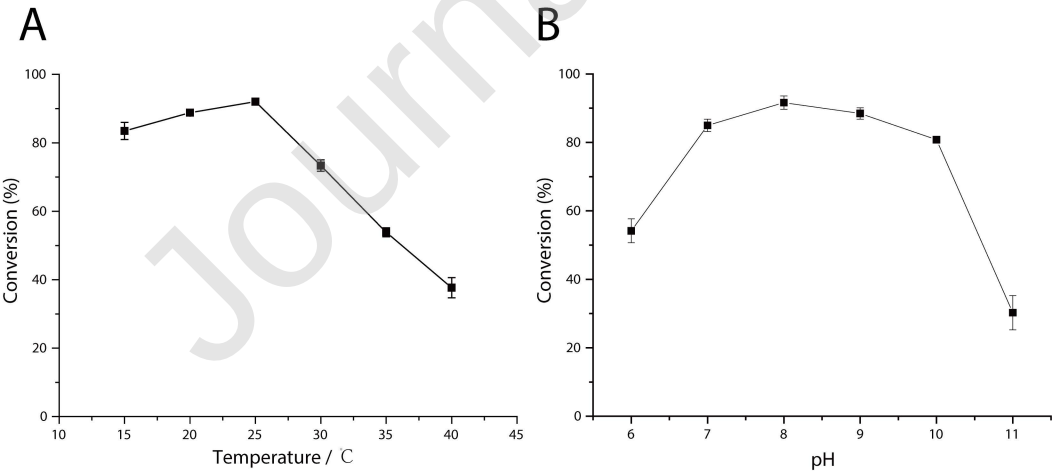

A
R. kyotonensis
alcohol dehydrogenase
(Rhky-ADH)
NADH
NAD+
Rhky-ADH
1
2
3
Anti-thrombotic agent-Ticagrelor
B
Azoarcus aromaticum EbN1
Hped
tioL
ebA305
ebA303
ebA302
pchF
pchC
C
H2O
2[H]
pchCF
NAD+
NADH
HpeD
p-Ethylphenol
1-(4-hydroxyphenyl)-ethanol
4-hydroxyacetophenone
D
HpeD
Genome mining
Rhky-ADH
promiscuousity
New dehydrogenase

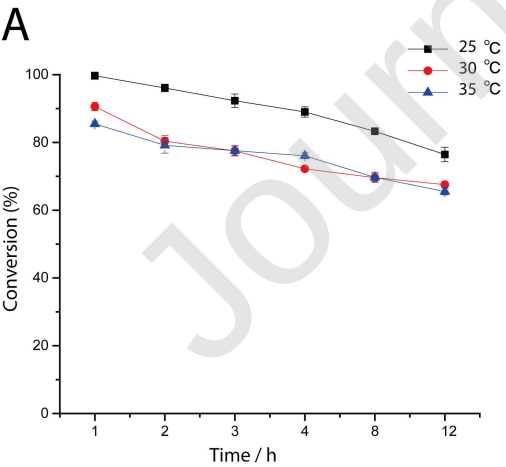

OZC55041.1 Rhodococcus sp. RS1C4
OZF50456.1 Rhodococcus sp. 14-2470-1a
WP_094731569.1 Rhodococcus sp. 15-1154-1
WP_094736628.1 Rhodococcus sp. 15-1189-1-1a
OZF17285.1 Rhodococcus sp. 14-2686-1-2
WP_072807147.1 R. yunnanensis
WP_089252776.1 R. kyotonensis **
WP_115965886.1 Rhodococcus sp. OK269
WP_008711153.1 Rhodococcus sp. AW25M09
CAI06282.1 Aromatoleum aromaticum EbN1
WP_094695141.1 Rhodococcus sp. 06-221-2
WP_019747310.1 Rhodococcus erythropolis
WP_068368102.1 Rhodococcus sp. EPR-157
81
96
99
79
94
84
58
100
50
100
0.10

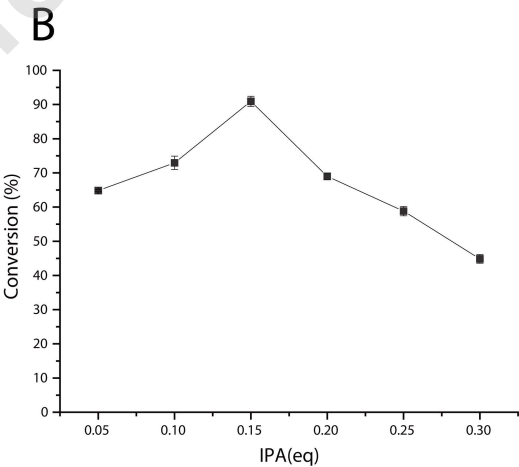

A
KDa
70
50
40
30
25
Rhky-ADH
B
mV
260 nm
Blank
Reaction
Reference
OH
Cl
F
(S)
(R)
min
Rhky-ADH
> 99% de
> 99% conversion

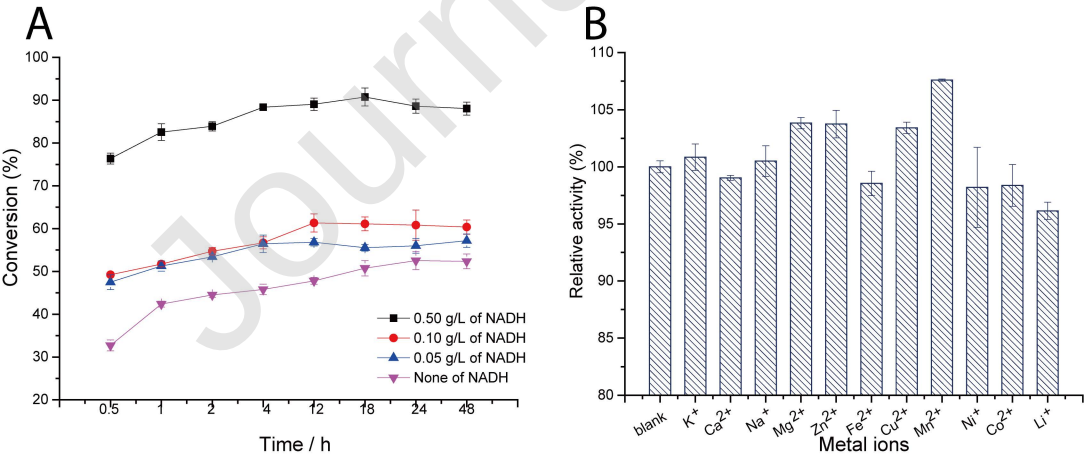

A
WP_019747310.1
WP_068368102.1
OZC55041.1
OZF50456.1
WP_094731569.1
WP_094736628.1
OZF17285.1
WP_072807147.1
WP_089252776.1
WP_115965886.1
WP_008711153.1
CAI06282.1
WP_094695141.1
B
Tyr157
Ser144
Lys161
Asn116
Pro187
NADH
Asn65
Asp39
C


A
Conversion (%)
Substrate Concentration(g/L)
B
conversion (%)
Biocatalyst Loading (mg/mL)


A
B
Conversion (%)
Temperature / ℃
pH
0
20
40
60
80
100
10
15
20
25
30
35
40
45
6
7
8
9
10
11


A
25 ℃
30 ℃
35 ℃
Conversion (%)
Time / h
0
20
40
60
80
100
1
2
3
4
8
12


B
Conversion (%)
IPA(eq)
0
10
20
30
40
50
60
70
80
90
100
0.05
0.10
0.15
0.20
0.25
0.30


A
Conversion (%)
Time / h
0.50 g/L of NADH
0.10 g/L of NADH
0.05 g/L of NADH
None of NADH
B
Relative activity (%)
Metal ions
blank
K+
Ca2+
Na+
Mg2+
Zn2+
Fe2+
Cu2+
Mn2+
Ni+
Co2+
Li+