



## Co-evolution of activity and thermostability of an aldo-keto reductase *KmAKR* for asymmetric synthesis of statin precursor dichiral diols

Shuai Qiu<sup>a,b,c</sup>, Feng Cheng<sup>a,b,c</sup>, Ling-Jun Jin<sup>a,b,c</sup>, Yi Chen<sup>a,b,c</sup>, Shu-Fang Li<sup>a,b,c</sup>, Ya-Jun Wang<sup>a,b,c,\*</sup>, Yu-Guo Zheng<sup>a,b,c</sup>

<sup>a</sup> Key Laboratory of Bioorganic Synthesis of Zhejiang Province, College of Biotechnology and Bioengineering, Zhejiang University of Technology, Hangzhou 310014, PR China

<sup>b</sup> Engineering Research Center of Bioconversion and Biopurification of the Ministry of Education, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, PR China

<sup>c</sup> The National and Local Joint Engineering Research Center for Biomanufacturing of Chiral Chemicals, Zhejiang University of Technology, Hangzhou 310014, PR China

### ARTICLE INFO

#### Keywords:

Aldo-keto reductase  
Co-evolution  
Asymmetric reduction  
Dichiral diol  
Biocatalysis

### ABSTRACT

Aldo-keto reductase *KmAKR*-catalyzed asymmetric reduction offers a green approach to produce dichiral diol *tert*-butyl 6-substituted-(3*R*,5*R*/*S*)-dihydroxyhexanoates, which are important building blocks of statins. In our previous work, we cloned a novel gene of NADPH-specific aldo-keto reductase *KmAKR* (WT) from a thermo-tolerant yeast *Kluyveromyces marxianus* ZJB14056 and a mutant *KmAKR*-W297H/Y296W/K29H (Variant III) has been constructed and displayed strict diastereoselectivity towards *tert*-butyl 6-cyano-(5*R*)-hydroxy-3-oxohexanoate ((5*R*)-1) but moderate activity and stability. Herein, to further co-evolve its activity and thermostability, we performed semi-rational engineering of Variant III by using a combinational screening strategy, consisting of tertiary structure analysis, loop engineering, and alanine scanning. As results, the “best” variant *KmAKR*-W297H/Y296W/K29H/Y28A/T63M (Variant VI) was acquired, whose  $K_m$ ,  $k_{cat}/K_m$  towards (5*R*)-1 was 0.66 mM and 210.77 s<sup>-1</sup> mM<sup>-1</sup>, respectively, with improved thermostability (half-life of 14.13 h at 40 °C). Combined with 1.5 g dry cell weight (DCW) L<sup>-1</sup> *Exiguobacterium sibiricum* glucose dehydrogenase (*EsGDH*) for NADPH regeneration, 4.5 g DCW L<sup>-1</sup> Variant VI completely reduced (5*R*)-1 of up to 450 g L<sup>-1</sup> within 7.0 h at 40 °C, yielding the corresponding optically pure *tert*-butyl 6-cyano-(3*R*,5*R*)-dihydroxyhexanoate ((3*R*,5*R*)-3, > 99.5% *d.e.p.*) with a space-time yield (STY) of 1.24 kg L<sup>-1</sup> day<sup>-1</sup>, and this was the highest level documented in literatures so far on substrate loading and STY of producing (3*R*,5*R*)-3. Besides (5*R*)-1, Variant VI displayed strong activity on *tert*-butyl 6-chloro-(5*S*)-hydroxy-3-oxohexanoate ((5*S*)-2). 4.5 g DCW L<sup>-1</sup> Variant VI completely reduced 400 g L<sup>-1</sup> (5*S*)-2, within 5.0 h at 40 °C, yielding optically pure *tert*-butyl 6-chloro-(3*R*,5*S*)-dihydroxyhexanoate ((3*R*,5*S*)-4, > 99.5% *d.e.p.*) with a STY of 1.34 kg L<sup>-1</sup> day<sup>-1</sup>. In summary, Variant VI displayed industrial application potential in statins biomanufacturing.

### 1. Introduction

Asymmetric hydrogenation of prochiral ketones or aldehydes to chiral alcohols is an important transformation in organic synthesis [1,2]. Over the past several decades, a variety of chemical approaches have been developed [3–5], most of which are based on boride-catalyzed asymmetric hydrogenation [6,7], which requires harsh conditions, especially cryogenic treatment ( $T < -60$  °C) [8,9], bringing about huge energy consumption and waste disposal problems [10], along with unsatisfactory product optical purity [11,12]. Biocatalysis generally proceeds under moderate conditions and is highly selective [13], therefore being extensively applied in the commercial chemical

production, especially pivotal chiral compounds [13–18]. Unfortunately, natural enzymes commonly display poor performance on unnatural substrates [19], most of which fail to withstand the harsh industrial conditions [20,21]. Catalysis at high temperatures is advantageous to raise substrate solubility, diffusion and reaction rate, and to reduce viscosity of reaction mixture [22]. However, co-evolution of enzyme activity and thermostability is particularly challenging due to the well-known stability-activity trade-off [18,22]. In such cases, one needs to take into account enzymes' activity and thermostability simultaneously towards unnatural substrates with high concentration [23,24].

Aldo-keto reductases (AKRs) have been used in prochiral ketones

\* Corresponding author at: College of Biotechnology and Bioengineering, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, PR China.

E-mail address: [wangyj@zjut.edu.cn](mailto:wangyj@zjut.edu.cn) (Y.-J. Wang).

<https://doi.org/10.1016/j.bioorg.2020.104228>

Received 28 May 2020; Received in revised form 28 July 2020; Accepted 11 August 2020

Available online 26 August 2020

0045-2068/ © 2020 Elsevier Inc. All rights reserved.

and aldehydes asymmetric synthesis [18,25,26]. Dichiral diol *tert*-butyl 6-substituted-(3*R*,5*R*/*S*)-dihydroxyhexanoates are important building blocks of statins. Asymmetric reduction of *tert*-butyl 6-cyano-(5*R*)-hydroxy-3-oxohexanoate ((5*R*)-1) [27] and *tert*-butyl 6-chloro-(5*S*)-hydroxy-3-oxohexanoate ((5*S*)-2) gave optically pure *tert*-butyl 6-cyano-(3*R*,5*R*)-dihydroxyhexanoate ((3*R*,5*R*)-3) and *tert*-butyl 6-chloro-(3*R*,5*S*)-dihydroxyhexanoate ((3*R*,5*S*)-4) [28], which are pivotal chiral intermediates of atorvastatin (Lipitor®) [29,30,31], rosuvastatin (Crestor®) and pitavastatin (Livalo®) [32,33], respectively. These statins possess excellent lipid-lowering efficacy, safety and long-term clinical benefits in reducing cardiovascular and cerebrovascular disease incidence and mortality [27,34,35].

In our previous work, wild type (WT) *KmAKR* has been cloned from a thermotolerant yeast *Kluyveromyces marxianus* ZJB14056 [25], and its variant *KmAKR*-W297H/Y296W/K29H (Variant III) was constructed via one round site-directed saturation mutagenesis (SSM) and two rounds iterative saturation mutagenesis (ISM), which retained good thermostability and had 16.69 folds improved catalytic efficiency in comparison with those of WT [26]. Variant III took only 3.5 h to completely reduce 100.0 g L<sup>-1</sup> (5*R*)-1, yielding 237.4 mM (3*R*,5*R*)-3 in *d.e.p* value above 99.5%, and the corresponding space–time yield (STY) was 372.8 g L<sup>-1</sup> d<sup>-1</sup>. Still, engineering of *KmAKR* to develop a robust biocatalyst for green biomanufacturing of statin intermediates is in huge demand.

In recent years, owing to the fast development of computer simulation technology [36–39], semi-rational engineering [40] of targeted residues based on sequence alignment [41] and structure analysis [14,41,42] has been developed. A general strategy is substituting the unstable residues or segments with their stable counterparts, thus rigidifying the overall protein structure and enhancing the thermostability and activity [15–17]. In such protocols, choosing suitable hot spots for mutation is critical and challenging, which depends on the molecular understanding of enzymes' structure–function relationship [43].

In the present study, we dug out the critical hot spots by a combinatorial screening strategy that consists of tertiary structure analysis, loop engineering, and alanine scanning, and then performed ISM to co-evolve the activity and thermostability of *KmAKR*. For the “best” variant, its enzymatic properties towards (5*R*)-1 and catalytic performances towards both (5*R*)-1 and (5*S*)-2 were studied in details. Moreover, the underlying molecular mechanism was investigated by docking analysis and molecular dynamics (MD) simulations.

## 2. Materials and methods

### 2.1. Materials

*Escherichia coli* BL21(DE3) used for cloning and recombinant protein expression was purchased from Novagen (Shanghai, China). The *E. coli* BL21(DE3)/pET28b(+)-*esgdh* and *E. coli* BL21(DE3)/pET28a(+)-*kmakr*-Y296W/W297H/K29H were constructed in our previous work [26], and preserved in our laboratory. Primers synthesis and DNA sequencing were conducted by Tsingke Zixi Biotechnology Co., Ltd (Hangzhou, China). (5*R*)-1 was generously presented by Zhejiang Lepu Pharmaceutical Co., Ltd (Taizhou, China), and (5*S*)-2 presented by Zhejiang Jingxin Pharmaceutical Co., Ltd (Xinchang, China). (3*R*,5*R*)-3 was purchased from Toronto Research Chemicals Co., Ltd. (Toronto, Canada), (3*R*,5*S*)-4 from J&K Scientific (Beijing, China). NADPH and NADP<sup>+</sup> were purchased from Roche Co., Ltd. (Basel, Switzerland), Phanta Super-Fidelity DNA polymerase from Vazyme Biotech Co., Ltd (Nanjing, China). The plasmid miniprep kit, DNA gel extraction kit and PCR cleanup kit were supplied by AxyPrep (Suzhou, China). Kanamycin was acquired from Solarbio Science & Scientific Co., Ltd. (Beijing, China), isopropyl β-D-1-thiogalactopyranoside (IPTG) from Heli Biotechnology Co., Ltd (Shanghai, China). Unless otherwise specified, other chemicals were of analytical grade and obtained from the standard commercial sources.

### 2.2. Methods

#### 2.2.1. Mutagenesis

Alanine scanning technology was used for primary screening of 6 hot spots based on Variant III. A one-round of SSM was conducted to assess the functions of Tyr28 based on Variant III and one-round of ISM was conducted to assess the functions of Thr63 in the *KmAKR*-W297H/Y296W/K29H/Y28A (Variant V), in which the degenerate codon NNS was substituted for the codon of the target amino acids. The primers used in this study were listed in Table S1 (Supporting Information). The template was digested by *DpnI* at 37 °C for 0.5 h (refer to Table S3 for details) and PCR amplified under PCR reaction conditions, which were supplemented in Table S2 and Figure S1. The resultant PCR products were transformed into *E. coli* BL21 (DE3) competent cells for enzyme expression

#### 2.2.2. Enzyme expression and purification

The Variant III and its variants were cultivated at 37 °C, 200 rpm, in 100 mL of Luria-Bertani (LB) medium containing 50 μg mL<sup>-1</sup> kanamycin sulfate. When optical density at 600 nm (OD<sub>600</sub>) of the culture reached 0.6–0.8, protein expression was induced with 0.15 mM IPTG for 12 h at 28 °C, 200 rpm. The cells were collected by centrifugation (8000 rpm) at 4 °C for 10 min, and washed twice with 0.9% (w/v) saline. The harvested wet cells were re-suspended in 100 mM sodium phosphate buffer (pH 7.0) at 50 g L<sup>-1</sup>, and the mixture was sonicated for 10 min (power 400 W, work for 1 s, break for 1 s). After centrifugation at 8000 rpm for 10 min to remove the precipitate, the as-prepared cell-free extract was loaded onto a Ni<sup>2+</sup>-nitrilotriacetic acid column (Bio-Rad, USA), which was pre-equilibrated with buffer A (pH 8.0, 20 mM sodium phosphate buffer, supplemented with 300 mM NaCl and 20 mM imidazole). The target protein was eluted with buffer B (pH 8.0, 20 mM sodium phosphate buffer supplemented with 300 mM NaCl and 500 mM imidazole), and desalted against 20 mM phosphate buffer (pH 7.0) overnight for 12 h using the cellulose dialysis bag (molecular weight cut-off: 14000 Dal). Enzyme purity was assessed by SDS-PAGE, and enzyme concentration was determined using the Bradford method [44] at 562 nm using bovine serum albumin as the standard. The purified enzymes were stored at 4 °C until further use.

#### 2.2.3. Enzyme activity assay and kinetics characterization

The activities of Variant III and its mutants were assayed at 35 °C, 500 rpm, in a total of 1.0 mL reaction mixture, composed of pH 7.0, 100 mM sodium phosphate buffer supplemented with 2 mM (5*R*)-1, and 0.25 mM NADPH together with appropriate amounts of the purified *KmAKRs*. An aliquot of 500 μL was withdrawn after reaction for 5 min, which was quenched by adding 2 μL, 6.0 mol L<sup>-1</sup> HCl. The resultant mixture was centrifuged at 12,000 rpm for 3 min. The supernatant was combined, and subjected to microfiltration with 0.22 μm PTFE organic membranes. The resulting filtrate was subsequently analyzed by High Performance Liquid Chromatography (HPLC).

One unit (U) of enzyme activity was defined as the amount of *KmAKR* required for 1 μmol of (3*R*,5*R*)-3 formed per minute at 35 °C, pH 7.0. Specific activity was calculated as the units of enzyme activity per milligram of proteins, U mg<sup>-1</sup> protein.

The Michaelis-Menten kinetics of *KmAKRs* were determined through measuring the initial rate at varying concentrations of (5*R*)-1 in a range of from 0.1 to 10 mM at a fixed NADPH concentration (0.2 mM). The kinetic parameters were obtained by nonlinear regressing the experimental data to Michaelis-Menten equation [18]. All data were handled by the software Origin 8.5, in the Michaelis-Menten model. 10 μL of each reaction mixture was sampled to assay (5*R*)-1, and (3*R*,5*R*)-3 by HPLC equipped with a C18 column (4.6 × 250 mm, 5 μm particle size, J&K Scientific). Separation was achieved with mobile phase consisting of acetonitrile and deionized water (75 : 25, v/v) at 40 °C, running at a flow rate of 1.0 mL min<sup>-1</sup>. Wavelength of the ultraviolet detector was set at 210 nm.

### 2.2.4. Substrate spectrum

The specific activity of purified *KmAKR-WT* and its mutants towards various ketones or ketoesters was determined spectrophotometrically by monitoring the change in NADPH absorbance at 340 nm ( $\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$ ) in phosphate buffer (pH 7.0, 100 mM) at 35 °C. The assay mixture (0.5 mL) contained 10 mM substrates (Sub1-Sub7), 0.25 mM NADPH, and an appropriate amount of **WT**, **Variant III**, **Variant V** and **Variant VI**. The test substrates included model substrates (5*R*)-1 (Sub1), (5*S*)-2 (Sub2), ethyl 4-chloro-3-oxobutanoate (Sub3), ethyl 3-oxobutanoate (Sub4), acetophenone (Sub5), ethyl 4,4,4-trifluoro-3-oxobutanoate (Sub6), *tert*-butyl 3-oxobutanoate (Sub7).

### 2.2.5. Thermostability assay

Determination of half-life  $t_{1/2}$

The half-lives ( $t_{1/2}$ ) of purified *KmAKRs* were determined by incubating each purified protein ( $1.0 \text{ mg mL}^{-1}$ ) at 40 °C, 50 °C, respectively, for appropriate time, followed by measurement of the residual activity. The half-lives ( $t_{1/2}$ ) were calculated according to the first order deactivation function 1 and Equation 2:

$$\ln(A/A_0) = -k_d t \quad (1)$$

$$t_{1/2} = \ln 2/k_d \quad (2)$$

where  $A_0$  is the initial activity,  $A$  the residual activity at time  $t$  during the thermal deactivation,  $k_d$  the deactivation rate constant ( $\text{h}^{-1}$ ).

Determination of  $T_{50}$ <sup>15</sup>

$T_{50}$ <sup>15</sup> is the temperature at which enzymes lose 50% activity after incubated for 15 min. 50  $\mu\text{L}$  of purified *KmAKRs* ( $1.0 \text{ mg mL}^{-1}$ ) were added into Eppendorf tubes, which were incubated at temperatures ranging from 20 °C to 50 °C for 15 min. After cooled down on ice bath, the residual activities of *KmAKRs* were assayed.

Determination of melting temperature ( $T_m$ )

$T_m$  values were measured by circular dichroism (CD) spectrometer (Applied Photophysics Ltd., United Kingdom). The melting curve data of *KmAKRs* were collected from 180 nm to 260 nm under temperatures of from 10 °C to 90 °C continuously with a gradient of 2 °C  $\text{min}^{-1}$ . Data measured at wavelength of 222 nm were converted to fraction folded  $\alpha$  according to the following Equation (3). The enzyme samples were diluted in potassium phosphate buffer (20 mM, pH 7.0) at the final concentration of 0.1  $\text{mg mL}^{-1}$  purified protein, and then loaded to a 10 mm quartz cuvette for CD assay [45].

$$\alpha = \frac{\theta_t - \theta_U}{\theta_F - \theta_U} \quad (3)$$

where  $\theta_t$  is the observed ellipticity at any temperature,  $\theta_F$  is the ellipticity of the fully folded form and  $\theta_U$  is the ellipticity of the unfolded form. The melting temperature ( $T_m$ ) values were calculated from the CD using Origin 8.5 software.

### 2.2.6. Modelling, docking and MD simulation

The three-dimensional (3-D) model of *KmAKR-WT* was created via homology modeling by using the SWISSMODEL web server (<http://www.swissmodel.expasy.org/>). Modeler models structure (PDB ID: 3WG6) was used as the starting structure of **WT** for model-constructing. The structures of **Variant III**, **Variant V**, **Variant VII** and **Variant VI** were constructed individually by foldx software with its inherent script.

Docking was carried out using YASARA 16.3.5 software under default processes using (5*R*)-1 as the ligand and the *KmAKR-Variant III*/NADPH complex, *KmAKR-Variant V*/NADPH complex, *KmAKR-Variant VI*/NADPH complex and *KmAKR-Variant VII*/NADPH complex as the receptor molecule. The center of the grid box was located at the catalytic tetrad region and each dimension of the grid box was set at 10 Å. Once the docking was finished, the optimal conformation was selected from the 10 results according to the binding energies and the attacking direction of NADPH to the carbonyl group.

Then, the **Variant III** and mutants were subjected to MD simulation at 298 K for 20 ns with the software YASARA [46,47]. A cubic box was

constructed to perform MD calculation. In the MD simulation, the Amber 03 force field was applied and the structures were put in a cubic box in each XYZ direction. The transferable intermolecular potential 3 points (TIP3P) model was used for water molecules and 150 mM NaCl was added. The model systems were relaxed by a series of minimizations and short dynamic simulations. The time step was 2 fs, and trajectories were saved every 100 ps. For *KmAKR-WT* and its variants, three independent MD simulations were performed [46,47].

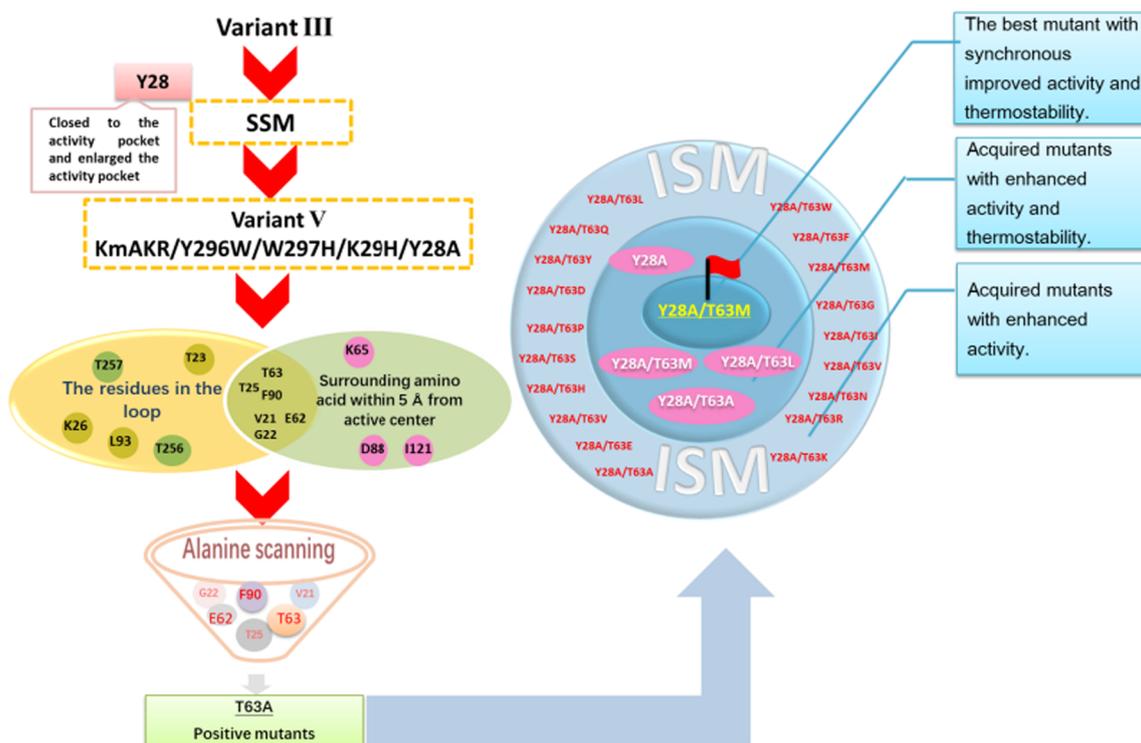
### 2.2.7. Biotransformation of (5*R*)-1 and (5*S*)-2

As for (5*R*)-1, a 100 mL scale reaction mixture consisting of substrate (45.0 g), glucose (1.0 equivalent), and 6.0 g DCW  $\text{L}^{-1}$  containing **Variant VI** mixed with *EsGDH*, in potassium phosphate buffer (100 mM, pH 7.0) was incubated at 40 °C and 800 rpm. The pH of reaction mixture was maintained at 7.0 by adding of  $\text{Na}_2\text{CO}_3$  solution (2.0 M) with a pH-stat (Mettler Toledo, Switzerland). The stereoselectivity and yield of (3*R*,5*R*)-3 were detected as previously described [26]. As for (5*S*)-2, the reaction conditions were the same as (5*R*)-1 except that the substrate addition amount was 40.0 g. The stereoselectivity and yield of (3*R*,5*S*)-4 were detected as previously described [28].

## 3. Results and discussion

### 3.1. Design of mutants based on **Variant III**

In our previous study, three *KmAKR* mutants harboring three beneficial substitutions, Y296W, W297H and K29H (variant *KmAKR-W297H* was designated as **Variant I**, variant *KmAKR-W297H/Y296W* designated as **Variant II**, *KmAKR-W297H/Y296W/K29H* designated as **Variant III**) were characterized [25,26]. These variants exhibit strictly selectivity and improved activity compared with **WT**, demonstrating that protein engineering is a powerful tool to modulate diastereoselectivity and activity. The homology model of *KmAKR* was constructed based on the crystal structure of a reductase C1 from *Candida parapsilosis* complexed with NADPH CPR-C1 (GenBank accession No.: BAD01652.1, PDB ID: 3WG6) since its crystal structure shares 49% sequence identity with *KmAKR* [26]. A previous study [26] has reported side chain's size of the residues spatially adjacent to the active center affects catalytic efficiency. In this study, molecular docking of (5*R*)-1 into **Variant III** were firstly performed, and we tried to replace the larger amino acids with smaller amino acids to enlarge the volume of the substrate-binding pocket. Initially, Tyr28 was chosen as the mutagenesis target (Figure S2). **Variant III** possesses a substrate binding cavity with a calculated volume of 124.2 Å<sup>3</sup>, which was enlarged to 140.7 Å<sup>3</sup> by replacing Tyr28 with much smaller alanine (computed by CAVER Analyst 2.0 BETA) based on its homology model (Figure S3). To further verify the effect of Tyr28, SSM was performed on Tyr28 based on **Variant III**, and *KmAKR-W297H/Y296W/K29H/Y28A* (**Variant V**) had 55.4% improved relative activity compared with **Variant III** (see 3.2 for detailed data). Subsequently, a screening strategy based on structure analysis was employed to dig out potential hot spots (Scheme 1). We mainly focused on the amino acid residues within 5 Å from the catalytic residues for the consideration that the maximum distance in which the hydride of C4 from the nicotinamide ring of NADPH attacked the carbon atom of the C3 carbonyl group of (5*R*)-1 was 4.5 Å [40,48,49], and on the surface loop within 30 Å from the catalytic residues, as residues within catalytic center played important roles in the activity and stability of enzyme [43]. Finally, we chose six sites Val21, Gly22, Thr25, Gln62, Thr63, Phe90 (Scheme 1) according to the intersection, by taking the amino acid residues within 5 Å from the catalytic residues and the surface loop within 30 Å from the catalytic residues into account. Attempt was made to identify the potential hot spots using SSM. However, screening saturation mutant libraries of 6 residues is an enormous task. Numerous reports have demonstrated that "low-energy screening" strategies such as alanine



Scheme 1. Strategy for constructing highly focused *KmAKR* library.

scanning, which are efficient in identifying potential amino acids that may affect catalytic activity [50,51]. Therefore, alanine scanning in combination with ISM was adopted in this study.

### 3.2. SSM and ISM for co-evolving *KmAKR*'s thermostability and activity

As shown in Table 1, only Y28A had 55.4% improved relative activity compared with Variant III in the 1st SSM, and *KmAKR*-W297H/Y296W/K29H/Y28A (Variant V) was acquired. The mutation sites of Val21, Gly22, Thr25, Gln62, Thr63 and Phe90 were targeted and the potential positive mutants were obtained via alanine scanning based on Variant V. Compared with Variant V, only T63A showed 1.12-fold improved relative activity, which was 0.8 times higher in  $k_{cat}/K_m$  than Variant III (Table 2). The second round SSM was set T63 as the target to conduct ISM based on Variant V. As shown in Table 1, a "best" mutant Variant VI (*KmAKR*-W297H/Y296W/K29H/Y28A/T63M) was acquired with 4.1-fold activity improvement, compared with Variant V. Variant VI exhibited 7.8-fold and 3.8-fold improvement in  $k_{cat}/K_m$  in comparison with Variant III and Variant V, respectively (Table 2).

Among all the saturated mutations, 10 mutants were positive, most of which were caused by replacements with hydrophobic amino acids. We speculated that the hydrophobic amino acid substitutions at position 63 was beneficial to promote the formation of the hydrogen bond and the protonation of (5*R*)-1. In addition to Variant VI, *KmAKR*-W297H/Y296W/K29H/Y28A/T63L (Variant VII) was 3.0 times higher in relative activity than Variant V.

Variant V, Variant VI, Variant VII were selected for thermostability verification. The half-lives of Variant V, Variant VI, and Variant VII at 40 °C were 8.77 h, 14.13 h, 16.96 h, respectively. Although Variant V acquired enhanced activity, unfortunately, a decreased thermostability was observed (Table 3). In addition,  $T_m$  and  $T_{50}^{15}$ , representing the thermodynamic stability and kinetic stability, respectively, were also characterized. Both the thermodynamic stability and kinetic stability of Variant VI and Variant VII were significantly improved compared with Variant III (Fig. 1). The  $T_{50}^{15}$  values of Variant VI and Variant VII were 49.8 and 50.4 °C, respectively, higher

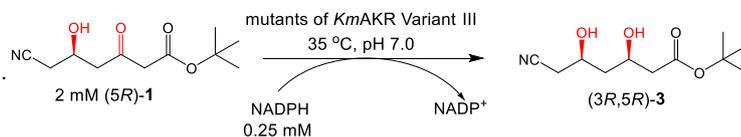
than that of Variant III (48.1 °C) (Fig. 1A). CD spectroscopy analysis showed that the  $T_m$  values of Variant VI (56.0 °C) and Variant VII (56.8 °C) were enhanced by 3.8 and 4.6 °C, respectively, compared with Variant III (52.2 °C) (Fig. 1B). Herein, a desired robust mutant Variant VI was obtained, whose activity and thermostability were improved synchronously.

### 3.3. Substrate specificity analysis of *KmAKRs*

The specific activities of the purified *KmAKRs* towards 7 different ketones or ketoesters were investigated. Results were summarized in Table 4. As documented in our previous study [25], *KmAKR*-WT showed low specific activity towards the test ketones, ketoesters. Variant III, Variant V, Variant VI showed much higher activity towards all the test compounds except Sub5. And two trends emerged from the analysis of *KmAKRs* catalytic activity towards artificial ketones or ketoesters. 1) Variant VI's activity decreased with skeleton length of the test substrates, and it exhibited the highest specific activity towards Sub7, which shares *tert*-butyl moiety with (5*R*)-1 and (5*S*)-2 and has the least number of carbon atoms in the skeleton. 2) Variant III, Variant V, Variant VI showed much reduced activity towards bulky aromatic ketone Sub5. Since tailoring *KmAKR* was targeted at (5*R*)-1, these two trends illustrated that the catalytic activity of variants were more adaptable for the aliphatic short-chain ketoesters, structural analogs of (5*R*)-1. Furthermore, Variant III, Variant V, Variant VI displayed excellent activity towards Sub3, which was the precursor of both (5*R*)-1 and (5*S*)-2.

### 3.4. The mechanism clarification of activity enhancement in *KmAKR*

The reaction mechanism of NAD(P)H-dependent dehydrogenase has been well studied. The prereaction states reflect early substrate binding with enzymes and co-factors via computational simulations. Based on the reaction mechanism of AKRs, Asp59 stabilizes the substrate; Tyr64 functions as the catalytic base, and the pKa of the hydroxyl group of Tyr64 is lowered by Lys89 [52,53]. The reaction is initiated by proton

**Table 1**Effect of mutation on activity and stereoselectivity of *KmAKR* towards (5*R*)-1.

Variant	Relative activity (%)	<i>d.e.</i> <sub>p</sub> (%)	Variant	Relative activity (%)	<i>d.e.</i> <sub>p</sub> (%)
<b>Variant III</b> <sup>a</sup>	100.0 ± 0.66	> 99.5	3rd round ISM <sup>d</sup>		
1st round SSM <sup>b</sup>			Y28A/T63P	15.9 ± 2.13	> 99.5
Y28F	71.8 ± 0.85	> 99.5	Y28A/T63F	263.2 ± 2.11	> 99.5
Y28G	18.7 ± 1.24	> 99.5	Y28A/T63M	793.3 ± 1.19	> 99.5
Y28H	55.9 ± 1.51	> 99.5	Y28A/T63I	358.1 ± 1.47	> 99.5
Y28I	9.9 ± 0.77	> 99.5	Y28A/T63R	308.2 ± 2.55	> 99.5
Y28L	23.4 ± 2.13	> 99.5	Y28A/T63L	654.1 ± 2.28	> 99.5
Y28P	25.4 ± 1.20	> 99.5	Y28A/T63Q	416.6 ± 3.23	> 99.5
Y28R	63.1 ± 2.24	> 99.5	Y28A/T63K	157.4 ± 3.87	> 99.5
Y28S	37.9 ± 1.13	> 99.5	Y28A/T63S	140.3 ± 2.22	> 99.5
Y28T	25.2 ± 1.78	> 99.5	Y28A/T63H	93.9 ± 2.98	> 99.5
Y28V	54.3 ± 1.82	> 99.5	Y28A/T63E	8.1 ± 0.11	> 99.5
Y28W	23.4 ± 0.69	> 99.5	Y28A/T63G	117.3 ± 0.87	> 99.5
Y28A (Variant V)	155.4 ± 3.81	> 99.5	Y28A/T63N	112.1 ± 2.33	> 99.5
2nd Alanine Scanning <sup>c</sup>			Y28A/T63C	263.4 ± 3.94	> 99.5
V21A	95.1 ± 3.22	> 99.5	Y28A/T63Y	290.1 ± 2.71	> 99.5
G22A	34.1 ± 1.33	> 99.5	Y28A/T63W	153.4 ± 1.99	> 99.5
T25A	53.2 ± 2.31	> 99.5	Y28A/T63D	8.4 ± 0.49	> 99.5
E62A	104.4 ± 4.94	> 99.5	Y28A/T63V	374.2 ± 5.89	> 99.5
T63A	329.8 ± 1.23	> 99.5			
F90A	13.6 ± 2.25	> 99.5			

<sup>a</sup> Variant III is *KmAKR*-W297H/Y296W/K29H, whose activity is designated as 100%. Activity was measured under the standard assay conditions. The values are the means of three individual experiments.

<sup>b</sup> 12 variants at Y28 based on *KmAKR*-W297H/Y296W/K29H were shown in 1st round SSM. The remaining 7 variants at Y28 were not shown in Table 1 because they lost activity completely after mutation.

<sup>c</sup> Alanine scanning was performed based on *KmAKR*-W297H/Y296W/K29H/Y28A.

<sup>d</sup> ISM was performed based on *KmAKR*-W297H/Y296W/K29H/Y28A.

**Table 2**Kinetic parameters of *KmAKR*-Y296W/W297H/K29H and other variants towards (5*R*)-1<sup>a</sup>

Variants	<i>v</i> <sub>max</sub> (U mg <sup>-1</sup> )	<i>k</i> <sub>cat</sub> (s <sup>-1</sup> )	<i>K</i> <sub>m</sub> for (5 <i>R</i> )-1 (mM)	<i>k</i> <sub>cat</sub> / <i>K</i> <sub>m</sub> (s <sup>-1</sup> mM <sup>-1</sup> )
<i>KmAKR</i> <sup>b</sup> -WT	4.48	2.72	2.01	1.35
Variant I <sup>b</sup>	17.09	10.36	1.58	6.56
Variant II <sup>b</sup>	49.50	29.84	1.42	21.01
Variant III	56.56	33.67	1.41	23.88
Variant V	75.76	45.37	1.04	43.62
Variant VI	141.89	139.69	0.66	210.77
Variant VII	128.62	160.17	0.81	198.09

<sup>a</sup>Kinetics data were measured by using purified proteins. All enzymatic assays were performed in triplicate and average values were adopted.

<sup>b</sup>The kinetics data of *KmAKR*-WT, Variant I and Variant II were referred to our previous work [25,26].

**Table 3**

The half-lives of Variant III and its variants.

Variants	40 °C		50 °C	
	<i>k</i> <sub>d</sub> (h <sup>-1</sup> ) <sup>a</sup>	<i>t</i> <sub>1/2</sub> (h) <sup>b</sup>	<i>k</i> <sub>d</sub> (h <sup>-1</sup> ) <sup>a</sup>	<i>t</i> <sub>1/2</sub> (h) <sup>b</sup>
Variant III	0.065	10.66	0.319	2.17
Variant V	0.079	8.77	0.669	1.04
Variant VI	0.049	14.13	0.242	2.86
Variant VII	0.041	16.96	0.221	3.15

<sup>a</sup> *k*<sub>d</sub>: inactivation rate constant.

<sup>b</sup> *t*<sub>1/2</sub>: half-life, the time taken for the residual activity of enzyme to fall to half its original activity at a specified temperature.

transfer from Tyr64-OH to the carbonyl oxygen atom of (5*R*)-1. Then, the hydride of NADPH is transferred to the carbonyl carbon, converting the carbonyl group to a hydroxyl group [54–56]. Hence, there are two

distances that can be used to evaluate whether the enzyme-substrate complex is developed to the prereaction state. One is defined by the distance between the carbonyl oxygen (C3) of (5*R*)-1 and the H atom of Tyr64-OH, designated as *d*<sub>1</sub>, indicating the formation of the hydrogen bond and the protonation of (5*R*)-1.

The other is defined by the distance between the carbonyl carbon (C3) of (5*R*)-1 and the hydrogen atom (H4) at C4 of NADPH nicotinamide, designated as *d*<sub>2</sub>, indicating the process of nucleophilic attack (Figure S4). *d*<sub>1</sub> ≤ 3.4 Å and *d*<sub>2</sub> ≤ 4.5 Å are prerequisites for hydrogen bond formation and hydride transfer, respectively [40,48,49]. Therefore, it is a reasonable mechanism explanation to clarify the improvement in activity of positive mutants through the changes of two distances in enzyme-substrate complex indirectly. According to the results of docking studies, the *d*<sub>1</sub> and *d*<sub>2</sub> of enzyme-substrate complex on three positive mutants were notably shortened compared with Variant III (Fig. 2), which explained why the formation of hydrogen bond, the process of hydride transfer and the process of nucleophilic attack became easier.

Interaction studies of substrate and neighbouring amino acids using CDOCKER (Discovery Studio 4.0) were performed with Variant III, Variant V, Variant VI, Variant VII and (5*R*)-1 as the substrate. The structural characteristics of prereaction states of Variant III and different mutants with (5*R*)-1 were further studied. Major interactions of both systems were calculated, as shown in Figure S6, respectively. For Variant III, the interactions of (5*R*)-1 and amino acids were mainly concentrated on the *tert*-butyl group, and except van der Waals force, few amino acid directly interacts with C3 carbonyl oxygen of (5*R*)-1. In comparison, in addition to the interactions of the surrounding amino residue with the *tert*-butyl group, the interactions of (5*R*)-1 with neighboring amino acids of Variant V, Variant VI and Variant VII were mainly maintained through hydrogen bonds with amino acid residues including Thr25 and Tyr64. It was demonstrated that these mutations changed the interactions between the amino acid residues in active center and (5*R*)-1, contributing to the increased interactions,

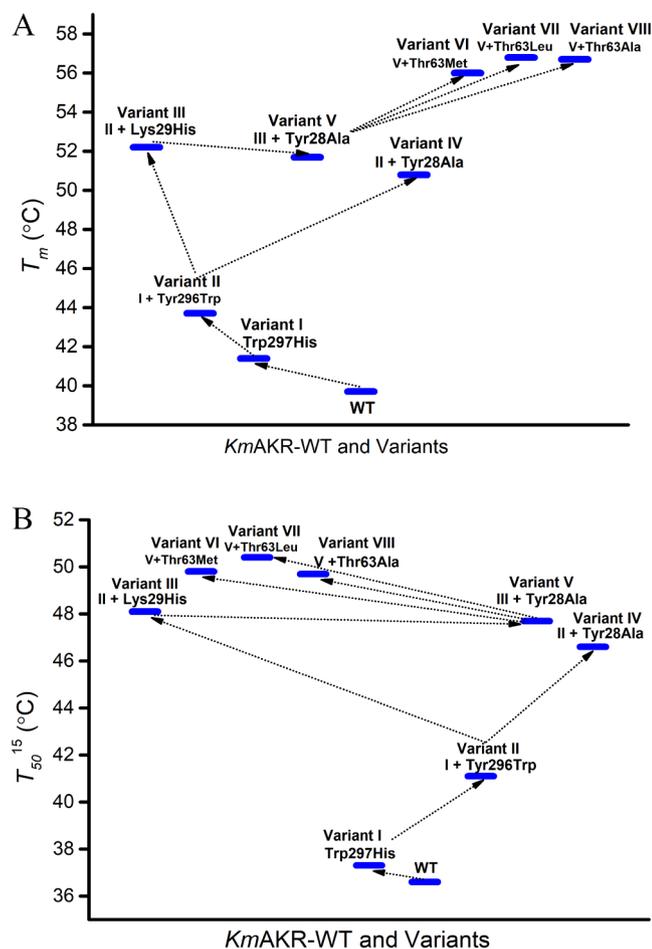


Fig. 1. The trend in thermostability of *KmAKRs* by mutation,  $T_m$  (A),  $T_{50}^{15}$  (B).

Table 4

Specific activity of *KmAKR*-WT and its variants towards various ketones and ketoesters.

Entry	Specific activity ( $\text{U mg}^{-1}$ ) <sup>a</sup>			
	WT	Variant III	Variant V	Variant VI
Sub1 <sup>b</sup>	4.44 ± 0.81	55.32 ± 3.33	72.69 ± 3.28	135.44 ± 5.45
Sub2 <sup>b</sup>	3.21 ± 0.32	41.42 ± 2.10	68.12 ± 1.23	151.42 ± 3.98
Sub3	19.51 ± 1.11	61.13 ± 3.98	152.34 ± 3.76	209.77 ± 2.21
Sub4	5.87 ± 2.21	80.11 ± 4.87	187.01 ± 4.49	250.57 ± 7.89
Sub5	1.57 ± 0.11	< 1.00	N.A. <sup>c</sup>	N.A. <sup>c</sup>
Sub6	14.24 ± 2.23	66.71 ± 5.92	181.77 ± 7.21	154.43 ± 5.62
Sub7	7.07 ± 0.03	89.94 ± 4.22	205.03 ± 5.98	291.55 ± 8.87

<sup>a</sup> Specific activity was determined using the purified enzymes by monitoring the change in NADPH absorbance at 340 nm ( $\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$ ). All enzymatic assays were performed in triplicate, and the averaged values were provided.

<sup>b</sup> Herein, the specific activities towards Sub1 and Sub2 obtained are a little bit (varied between 5 and 8%) lower than those determined by HPLC, owing to the methodological errors.

<sup>c</sup> N.A. stands for no measurable activity.

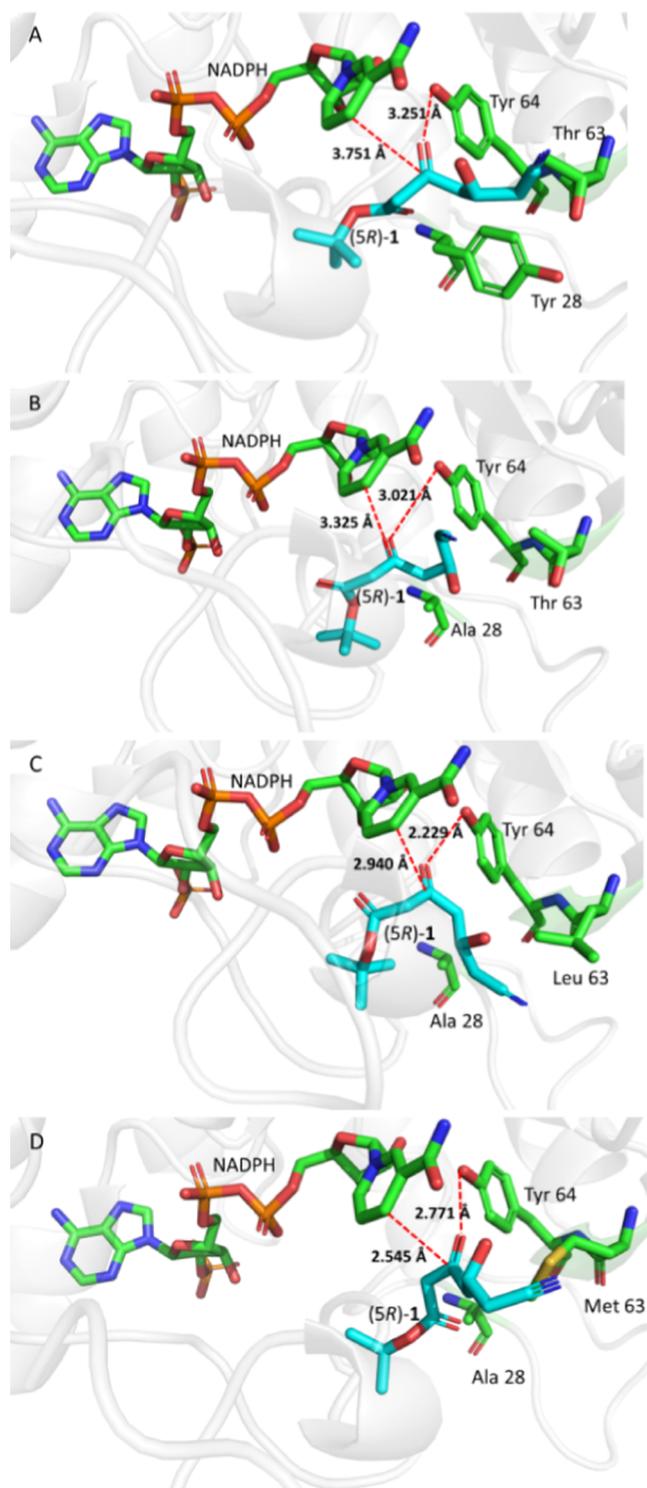
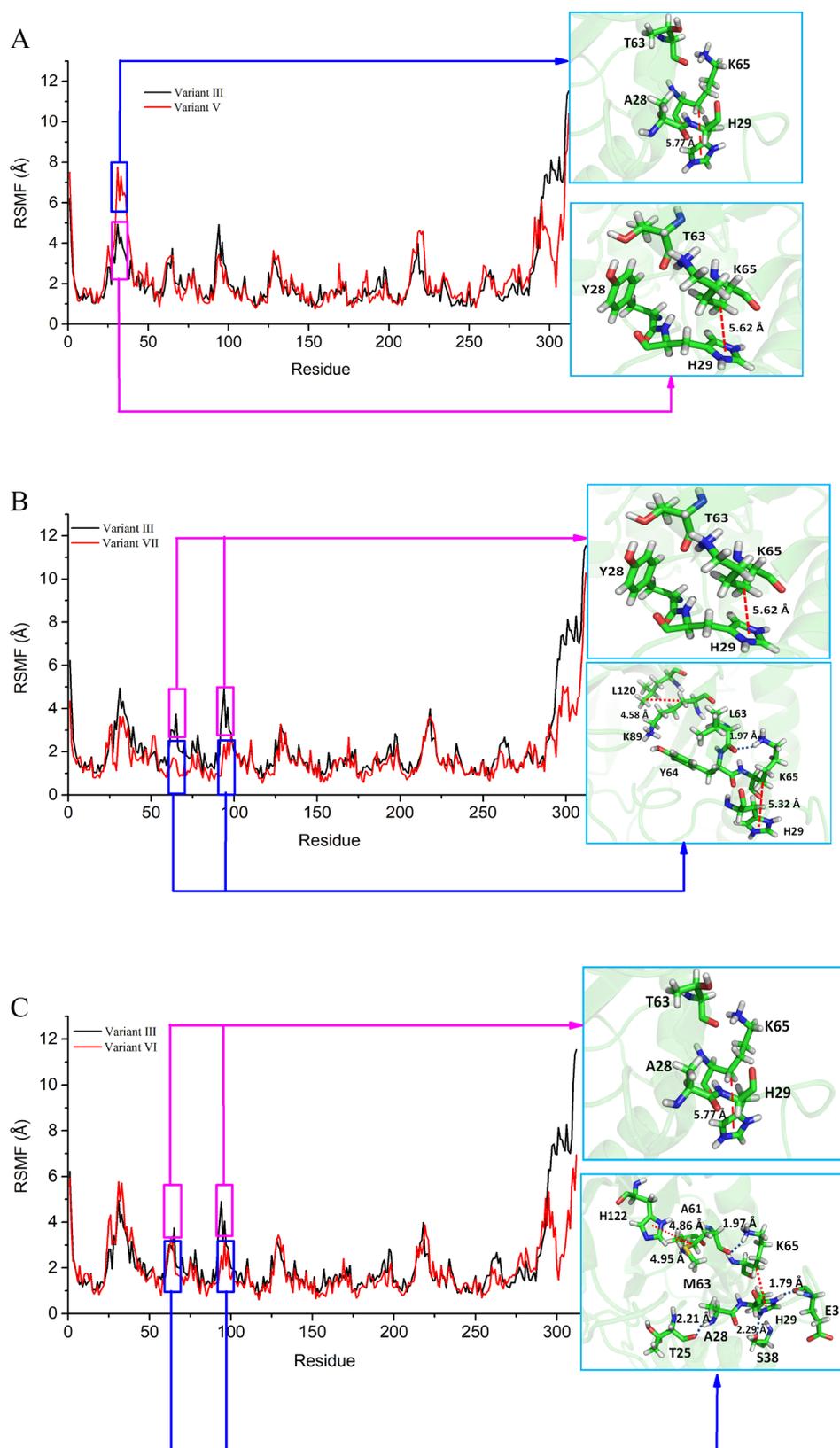


Fig. 2. Docking binding orientations of the substrate (5R)-1 and NADPH in the active site of Variant III (A), Variant V (B), Variant VII (C), Variant VI (D). Docking were performed with YASARA.

therefore being more conducive to the catalysis of the enzyme and the substrate. This interaction analysis is consistent with the enhanced  $k_{\text{cat}}/K_m$ . Compared with the volume size of substrate binding pocket of Variant III ( $124.2 \text{ \AA}^3$ ), all of Variant V, Variant VI and Variant VII showed an expanding trend in the volume size (computed by CAVER



**Fig. 3.** MD analysis and intramolecular interactions analysis in flexible region of **Variant III** and its mutants at 313 K using YASARA during a 20-ns simulation, and comparison between **Variant III** and **Variant V** (A), **Variant III** and **Variant VII** (B), **Variant III** and **Variant VI** (C). Hydrophobic interaction is shown in red, hydrogen bond shown in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

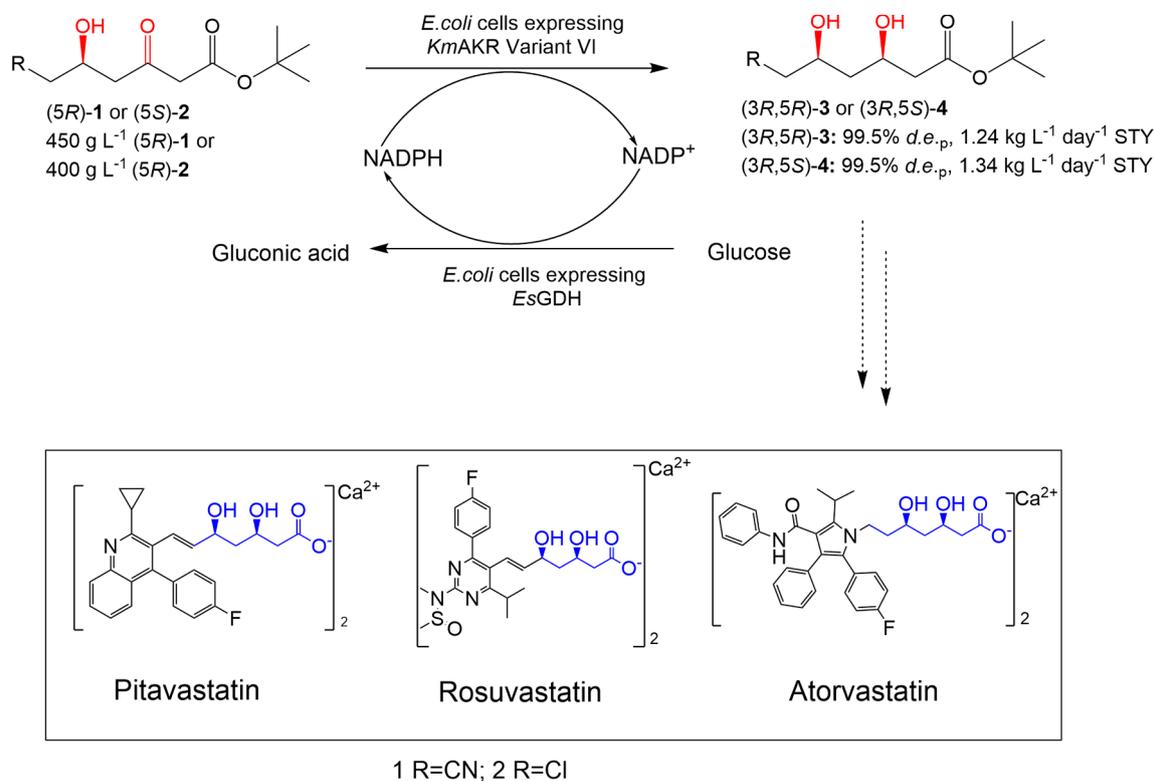


Fig. 4. Asymmetric reduction of (5R)-1 and (5S)-2 by *E. coli* cells harboring *KmAKR* Variant VI and *EsGDH* to synthesize (3R,5R)-3, (3R,5S)-4 on 100 mL scale.

Analyst 2.0 BETA), which were 140.7, 294.1 and 245.7 Å<sup>3</sup>, respectively (Figure S7). A larger substrate binding pocket can accommodate more (5R)-1 molecules, raising the binding probability of substrate and active site, and hence accelerating the reaction rate.

### 3.5. Clarifying the mechanism underpinning thermostability enhancement

Protein interaction calculation and MD simulation were used to analyze the difference in interactions and structures between various mutants. Protein interaction calculator online services (<http://pic.mbu.iisc.ernet.in/>) was applied to analyze the protein intramolecular interactions including hydrophobic interactions, hydrogen bonds, ionic interactions, aromatic-aromatic interactions, aromatic-sulphur interactions, cation- $\pi$  interactions and disulfide bridges. As shown in Table S4, positive mutants didn't possess increased overall protein interactions, compared to Variant III. As for Variant V, mutation lost 1 hydrophobic interaction, 1 hydrogen bond, 1 aromatic-aromatic interaction. For Variant VI and Variant VII, mutation lost 2, 3 hydrogen bonds and increased 3, 1 hydrophobic interactions, respectively. Besides, mutation increased 1 aromatic-sulphur interaction in Variant VI.

As shown in the inserts of Fig. 3, introduced Ala28, Met63, Leu63 were all hydrophobic amino acids, however, compared with Variant III, no additional hydrophobic interaction was introduced in the corresponding flexible region of Variant V and Variant VII. The hydrophobic distance between His29 and Lys65 of Variant V and Variant VII was 5.77 and 5.32 Å, respectively. However, the distance between His29 and Lys65 of Variant III was 5.62 Å, much shorter than that in Variant V, which is an explanation for decreased thermostability of Variant V. Instead, 2 additional hydrophobic interactions were introduced in the flexible region of Variant VI, and the hydrophobic distance between His29 and Lys65 was 5.37 Å, between His122, Ala61 and Met63 was 4.86 and 4.95 Å, respectively.

Among all the 3 mutants, hydrophobic distance between His29 and Lys65 in Variant VII was the shortest (5.32 Å), contributing to excellent thermostability. Based on model analysis, in the case of mutation of K29H located at flexible loop region, the hydrophobic distance

between His29 and Lys65 might play a significant role in thermostability. Pace *et al.* reported the hydrophobic interaction contributed 60 ± 4% to protein stability and the more hydrophobic groups, and the larger the volume, the stronger the hydrophobic interaction. Hydrogen bonds contribute 40 ± 4% to protein stability.[57]

MD simulations of Variant III and mutants with improved thermostability Variant VI and Variant VII was performed to obtain the profound insight into the molecular stabilization mechanism [43]. Root mean square fluctuation (RMSF) reflects protein stability, and high RMSF values on individual residues indicate high flexibility for a given residue. Loops that comprise turns and random coils are the most flexible structure elements in enzymes. The active-site-decorating loops involved in the enzyme substrate interaction influence enzyme functions, such as activity and stability [43]. As shown in Fig. 3, in addition to the flexible C-terminal RMSF value reduction, the RMSF values of loop 91–99 and loop 49–56 decreased. We hypothesized that (5R)-1 bound to the active center, forming stable interactions with the surrounding amino acids, which not only reduced the volatility of the amino acids near the catalytic center, but also reduced the volatility of loop 91–99 and loop 49–56.

Compared with Variant III, Variant VI had an additional hydrogen bond between Ala28 and Thr25 on loop17-38, and His29 formed a hydrogen bond with the Glu32 and Ser38 on loop 17–38, weakening the flexibility of loop 17–38 (Fig. 3C). Met63 forms a hydrogen bond with Lys65, along with hydrophobic interactions with Ala61 and His122, which play important role in reducing flexibility of loop 49–56 and loop 91–99 loop. When Thr63 was substituted by Leu63, additional hydrophobic interactions were formed between Lys89 and Leu120, between His29 and Tyr64, decreasing the flexibility of loop 91–99 and loop 49–56 (Fig. 3B). Hence, we concluded that introducing new hydrophobic interactions and hydrogen bonds in flexible regions raises *KmAKR*' thermostability.

### 3.6. Biocatalysis of (5R)-1 and (5S)-2

The engineered *KmAKR* mutants were capable of catalyzing the asymmetric reduction of prochiral statin intermediate (5R)-1 and (5S)-2

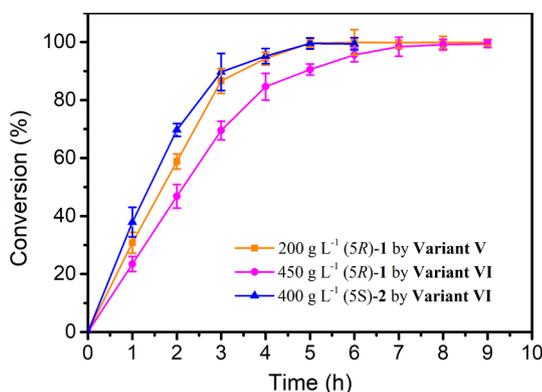


Fig. 5. Conversion profile of Variant V and Variant VI catalyzed reduction of (5R)-1 and (5S)-2 to (3R,5R)-3 and (3R,5S)-4.

to pharmacologically active (3R,5R)-3, (3R,5S)-4, accompanied by the consumption of NADPH. Therefore, a glucose dehydrogenase from *Exiguobacterium sibiricum* was introduced to regenerate the desired cofactor NADPH from NADP<sup>+</sup> using glucose as the co-substrate (Fig. 4).

The enzymatic asymmetric reduction of (5R)-1 at 200 g L<sup>-1</sup> was conducted using the Variant V, the reaction was accomplished completely within 4.5 h using 15.0 g DCW L<sup>-1</sup> of *E. coli* cells expressing Variant V and 5.0 g DCW L<sup>-1</sup> of *E. coli* cells expressing EsGDH on 100 mL scale. Under the same scale, 450 g L<sup>-1</sup> (5R)-1, 400 g L<sup>-1</sup> (5S)-2 were completely converted by 4.5 g DCW L<sup>-1</sup> Variant VI. Due to the outstanding thermal stability of the Variant VI, we tried to raise the reaction temperature from 35 °C to 40 °C to accelerate reaction rate. As shown in Fig. 5, the reaction catalyzed 200 g L<sup>-1</sup> (5R)-1 by Variant V achieved > 99% conversion in 4.5 h, and the corresponding STY reached 670.5 g L<sup>-1</sup> d<sup>-1</sup> with *d.e.p.* > 99.5%. The “best” mutant Variant VI catalyzed 450 g L<sup>-1</sup> (5R)-1 for 7.0 h, giving 99% conversion and STY of 1.24 kg L<sup>-1</sup> d<sup>-1</sup>. 400 g L<sup>-1</sup> (5S)-2 was converted into (3R,5S)-4 within 5 h, STY reaching 1.34 kg L<sup>-1</sup> d<sup>-1</sup>. The “best” mutant Variant VI significantly decreased the catalyst load, resulting in a dramatic increase of substrate/catalyst ratio of from 10.0 g g<sup>-1</sup> to 75.0 g g<sup>-1</sup> (Table S5).

Reviewing the entire evolution campaign from WT to Variant VI, five mutation sites were acquired to enhance the stereoselectivity, activity and thermostability. As shown in Figure S13, the location of five mutation site can be divided into 3 portions, core shell, middle shell and surface shell. Core shell is the active center, NADPH binding region and substrate binding region, and site 63 is located in the core shell. Surface shell is the enzyme surface region, where site 296 and 297 are located. Between the core shell and surface shell is the middle shell, where site 28 and 29 are located. The sites surrounding active center is the common hot spot for activity enhancement via rational and semi-rational engineering. However, the sites surrounding active center represents only a small portion of the whole enzyme. Identifying more hot spots from surface shell and middle shell via other methods, such as error prone PCR, to discover further reinforcing mutations other than only core shell for enzyme activity and thermostability engineering is helpful to develop ideal industrial biocatalysts with better catalytic performance.

To date, only limited alcohol dehydrogenases/reductases have been reported for chiral diol (3R,5R)-3 and (3R,5S)-4 biocatalysis. In our previous work, an aldo-keto reductase KlAKR-WT was cloned from the *Kluyveromyces lactis* (Table 5), which had excellent *R*-stereoselectivity towards (5R)-1, a sequence identity of 89% and the same catalytic tetrad with the KmAKR [58]. Our previous work on KlAKR was mainly focused on activity improvement through modification on amino acid residues that were situated in the neighborhood of tunnels of substrate-entrance and product-release, (5R)-1 and NADPH binding pockets, and within 4 Å distance from the bound (5R)-1, yielding a positive variant KlAKR-Y295W-W296L-I125V-S30P-Q212R-I63W, whose catalytic

efficiency towards (5R)-1 was 36.31 s<sup>-1</sup>·mM<sup>-1</sup>. Under the optimized conditions, it completely reduced (5R)-1 at load of up to 80 g·L<sup>-1</sup> in 1.5 h, giving (3R,5R)-3 in *d.e.p.* > 99.5% and STY of 660.0 g·L<sup>-1</sup>·d<sup>-1</sup>. In present work, to engineer KmAKR, we used the combinational screening strategy to screen hot spots around the intersection of the neighborhoods of the active center and surface loop. A synchronously improvement in both activity and thermostability was achieved. The Variant VI is much more active and thermo-stable than KlAKR-Y295W-W296L-I125V-S30P-Q212R-I63W [31].

### 3.7. Environmental factor analysis and economic analysis

Asymmetric bioreduction of (5R)-1 and (5S)-2 to (3R,5R)-3, (3R,5S)-4 by Variant VI is greener and more economic than the chemical reduction. As for the conditions of reaction, a variety of organic reagents consumption is zero for Variant VI. In contrast to the chemical asymmetric reduction using metal catalysts and NaBH<sub>4</sub>, this enzymatic process used glucose as sustainable co-substrates, affording water as a clean byproduct, making this method a promising green chemistry approach to manufacture atorvastatin intermediate (3R,5R)-3 and “Super statins” chiral intermediate (3R,5S)-4. Biocatalytic process employing Variant VI avoids the use of potentially hazardous hydrogen and heavy metal catalysts throughout the process, thus obviating concerns for their removal from waste streams and/or contamination of the product. More importantly, the enzyme catalysts and the co-substrate glucose are derived from renewable raw materials and are completely biodegradable. The by-products of the reaction are gluconate, a tiny amount of residual glucose, enzyme, cell debris and minerals, and the waste water is directly suitable for biotreatment. In term of economy, in this bioprocess, no exogenous NADPH or NADP<sup>+</sup> was added and the atom efficiency is 56.3%, attributing to that the use of glucose as the reductant for cofactor regeneration is cost effective but not particularly atom efficient. [30] Fortunately, glucose is a renewable resource and the co-product gluconate is fully biodegradable.

## 4. Conclusion

In summary, KmAKR mutants Variant VI and VII that exhibit simultaneously improved thermostability and activity compared with the parent enzyme Variant III were constructed using the combinational screening strategy and ISM. It was worth noting that the trade-off between activity and thermostability was avoided. Analysis of Docking, protein interaction calculator and MD stimulation, mutations of Y28A and T63M or T63L shortened the distance between the H atom of catalytic site Tyr64-OH and the carbonyl oxygen atom of (5R)-1, accelerating the transfer rate of protons. As for thermostability, flexible loops around the active site were rigidified through introducing additional hydrogen bonds and hydrophobic interactions in mutants Variant VI and Variant VII, leading to improvements in *t*<sub>1/2</sub>, *T*<sub>m</sub>, *T*<sub>50</sub><sup>15</sup>. Most importantly, Variant VI allowed for asymmetric bioreduction of (5R)-1 and (5S)-2 to (3R,5R)-3 and (3R,5S)-4 at elevated temperatures (e.g., 40 °C), dramatically shortening the reaction times and enhancing the productivity significantly. Finally, a STY of (3R,5R)-3 up to 1.24 kg·L<sup>-1</sup>·d<sup>-1</sup> was achieved with a 75.0 g·g<sup>-1</sup> S/C mass ratio, and a STY of (3R,5S)-4 up to 1.34 kg·L<sup>-1</sup>·d<sup>-1</sup> was achieved with a 66.6 g·g<sup>-1</sup> S/C mass ratio. Hence, Variant VI is a very promising biocatalyst for large scale production of statins chiral intermediates (3R,5R)-3 and (3R,5S)-4.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Table 5**  
Comparison of (3R,5R)-3 synthesis by Variant VI with other reductases.

Enzyme	Microorganism source	Substrate load (g L <sup>-1</sup> )	Reaction time (h)	NADPH or NADP <sup>+</sup> (mM)	S/C ratio <sup>c</sup> (g g <sup>-1</sup> )	P/C ratio <sup>d</sup> (g g <sup>-1</sup> )	STY (g L <sup>-1</sup> h <sup>-1</sup> )	Reference
Variant VI	<i>K. marxianus</i>	450	7.0	0	75.0	60.0	1224.0	This study
M3	<i>K. marxianus</i>	100	3.5	0	5.0	2.9	372.8	[26]
KRED	The engineered keroreductase SEQ ID NO.316	300	22	0.1	85.0	18.9	281.0	[59]
LbCR <sub>M8</sub> <sup>a</sup>	<i>Lactobacillus brevis</i>	300	5.0	0.1	27.0	21.6	1152.0	[18]
LbCR <sub>M8</sub> <sup>b</sup>	<i>L. brevis</i>	300	6.0	0.1	300.0	262.5	1050.0	[18]
KIAKR <sub>M8</sub>	<i>K. lactis</i>	80	1.5	0	4.0	2.1	660.0	[31]

<sup>a</sup> LbCR<sub>M8</sub> is the lyophilized cells expressing LbCR<sub>M8</sub>.

<sup>b</sup> LbCR<sub>M8</sub> is the lyophilized *E. coli* cells coexpressing LbCR<sub>M8</sub> and BmGDH.

<sup>c</sup> Mass ratio of substrate load to catalyst load.

<sup>d</sup> Mass ratio of product formed to catalyst load.

## Acknowledgments

This research was financially supported by National Natural Science Foundation of China (21878274).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bioorg.2020.104228>.

## References

- X.P. Zeng, Z.Y. Cao, Y.H. Wang, F. Zhou, J. Zhou, Catalytic enantioselective desymmetrization reactions to all-carbon quaternary stereocenters, *Chem. Rev.* 116 (2016) 7330–7396.
- J. Merad, M. Candy, J.M. Pons, C.J.S. Bressy, Catalytic enantioselective desymmetrization of Meso compounds in total synthesis of natural products: Towards an economy of chiral reagents, *Synthesis* 49 (2017) 1938–1954.
- M. Muller, Chemoenzymatic synthesis of building blocks for statin side chains, *Angew. Chem. Int. Ed.* 44 (2005) 362–365.
- Y. Yamashita, T. Yasukawa, W.J. Yoo, T. Kitano, S. Kobayashi, Catalytic enantioselective aldol reactions, *Chem. Soc. Rev.* 47 (2018) 4388–4480.
- R.A. Sheldon, Fundamentals of green chemistry: efficiency in reaction design, *Chem. Soc. Rev.* 41 (2012) 1437–1451.
- S. Qiu, Y.J. Wang, H. Yu, F. Cheng, Y.G. Zheng, *t*-Butyl 6-cyano-(3R,5R)-dihydroxyhexanoate synthesis via asymmetric reduction by immobilized cells of carbonyl reductase and glucose dehydrogenase co-expression *E. coli*, *Process Biochem.* 80 (2019) 43–51.
- X. Luo, Y.J. Wang, W. Shen, Y.G. Zheng, Activity improvement of a *Kluyveromyces lactis* aldo-keto reductase KIAKR via rational design, *J. Biotechnol.* 224 (2016) 20–26.
- M.E. Maier, Design and synthesis of analogues of natural products, *Org. Biomol. Chem.* 13 (2015) 5302–5343.
- Y.J. Wang, X.P. Chen, W. Shen, Z.Q. Liu, Y.G. Zheng, Chiral diol *t*-butyl 6-cyano-(3R,5R)-dihydroxyhexanoate synthesis catalyzed by immobilized cells of carbonyl reductase and glucose dehydrogenase co-expression *E. coli*, *Biochem. Eng. J.* 128 (2017) 54–62.
- N. Ran, L. Zhao, Z. Chen, J. Tao, Recent applications of biocatalysis in developing green chemistry for chemical synthesis at the industrial scale, *Green Chem.* 10 (2008) 361–372.
- J.M. Woodley, New opportunities for biocatalysis: making pharmaceutical processes greener, *Trends Biotechnol.* 26 (2008) 321–327.
- P.L. Brower, D.E. Butler, C.F. Deering, T.V. Le, A. Millar, T.N. Nanninga, B.D. Roth, The synthesis of (4R-cis)-1,1-dimethylethyl 6-cyanomethyl-2,2-dimethyl-1,3-dioxane-4-acetate, a key intermediate for the preparation of CI-981, a highly potent, tissue selective inhibitor of HMG-CoA reductase, *Tetrahedron Lett.* 33 (1992) 2279–2282.
- U.T. Bornscheuer, G.W. Huisman, R.J. Kazlauskas, S. Lutz, J.C. Moore, K. Robins, Engineering the third wave of biocatalysis, *Nature* 485 (2012) 185–194.
- Y. Cen, D. Li, J. Xu, Q. Wu, X. Lin, Highly focused library-based engineering of *Candida antarctica* Lipase B with (S)-selectivity towards sec-alcohols, *Adv. Synth. Catal.* 361 (2019) 126–134.
- F.F. Chen, G.W. Zheng, L. Liu, H. Li, Q. Chen, F.L. Li, C.X. Li, J.H. Xu, Reshaping the active pocket of amine dehydrogenases for asymmetric synthesis of bulky aliphatic amines, *ACS Catal.* 8 (2018) 2622–2628.
- X. Chen, H. Zhang, J. Feng, Q. Wu, D. Zhu, Molecular basis for the high activity and enantioselectivity of the carbonyl reductase from *Sporobolomyces salmonicolor* toward  $\alpha$ -haloacetophenones, *ACS Catal.* 8 (2018) 3525–3531.
- X. Chen, H. Zhang, M.A. Maria-Solano, W. Liu, J. Li, J. Feng, X. Liu, S. Osuna, R.T. Guo, Q. Wu, D. Zhu, Y. Ma, Efficient reductive desymmetrization of bulky 1,3-cycloketones enabled by structure-guided directed evolution of a carbonyl reductase, *Nature Catal.* 2 (2019) 931–941.
- X.M. Gong, Z. Qin, F.L. Li, B.B. Zeng, G.W. Zheng, J.H. Xu, Development of an engineered ketoreductase with simultaneously improved thermostability and activity for making a bulky atorvastatin precursor, *ACS Catal.* 9 (2019) 147–153.
- W. Song, X. Chen, J. Wu, J. Xu, W. Zhang, J. Liu, J. Chen, L. Liu, Biocatalytic derivatization of proteinogenic amino acids for fine chemicals, *Biotechnol. Adv.* 40 (2020) 107496.
- B.M. Bonk, J.W. Weis, B. Tidor, Machine learning identifies chemical characteristics that promote enzyme catalysis, *J. Am. Chem. Soc.* 141 (2019) 4108–4118.
- L.D. Unsworth, J. van der Oost, S. Koutsopoulos, Hyperthermophilic enzymes—stability, activity and implementation strategies for high temperature applications, *FEBS J.* 274 (2007) 4044–4056.
- K.S. Siddiqui, Defying the activity–stability trade-off in enzymes: taking advantage of entropy to enhance activity and thermostability, *Crit. Rev. Biotechnol.* 37 (2017) 309–322.
- F.H. Arnold, Directed evolution: Bringing new chemistry to life, *Angew. Chem. Int. Ed.* 57 (2018) 4143–4148.
- Z. Wu, S.B.J. Kan, R.D. Lewis, B.J. Wittmann, F.H. Arnold, Machine learning-assisted directed protein evolution with combinatorial libraries, *Proc. Natl. Acad. Sci. USA* 116 (2019) 8852–8858.
- Y.J. Wang, B.B. Ying, W. Shen, R.C. Zheng, Y.G. Zheng, Rational design of *Kluyveromyces marxianus* ZJB14056 aldo–keto reductase KmAKR to enhance diastereoselectivity and activity, *Enzyme Microb. Technol.* 107 (2017) 32–40.
- H. Yu, S. Qiu, F. Cheng, Y.N. Cheng, Y.J. Wang, Y.G. Zheng, Improving the catalytic efficiency of aldo-keto reductase KmAKR towards *t*-butyl 6-cyano-(3R,5R)-dihydroxyhexanoate via semi-rational design, *Bioorg. Chem.* 90 (2019) 103018.
- P. Hoyos, V. Pace, A. Alcántara, Biocatalyzed synthesis of statins: A sustainable strategy for the preparation of valuable drugs, *Catalysts* 9 (2019) 260.
- Z.Q. Liu, H.H. Yin, X.J. Zhang, R. Zhou, Y.M. Wang, Y.G. Zheng, Improvement of carbonyl reductase activity for the bioproduction of *tert*-butyl (3R,5S)-6-chloro-3,5-dihydroxyhexanoate, *Bioorg. Chem.* 80 (2018) 733–740.
- J.M. Patel, Biocatalytic synthesis of atorvastatin intermediates, *J. Mol. Catal. B-Enzym.* 61 (2009) 123–128.
- S.K. Ma, J. Gruber, C. Davis, L. Newman, D. Gray, A. Wang, J. Grate, G.W. Huisman, R.A. Sheldon, A green-by-design biocatalytic process for atorvastatin intermediate, *Green Chem.* 12 (2010) 81–86.
- W. Shen, Y. Chen, S. Qiu, D.N. Wang, Y.J. Wang, Y.G. Zheng, Semi-rational engineering of a *Kluyveromyces lactis* aldo-keto reductase KIAKR for improved catalytic efficiency towards *t*-butyl 6-cyano-(3R,5R)-dihydroxyhexanoate, *Enzyme Microb. Technol.* 132 (2020) 109413.
- S. Tartagaglia, S. Fogal, R. Motterle, C. Ferrari, M. Pontini, R. Aureli, O. De Lucchi, Chemoenzymatic synthesis of  $\delta$ -keto  $\beta$ -hydroxy esters as useful intermediates for preparing statins, *Eur. J. Org. Chem.* 2016 (2016) 3162–3165.
- P.A. Spreider, B. Breit, Palladium-catalyzed stereoselective cyclization of in situ formed allenyl hemiacetals: synthesis of rosuvastatin and pitavastatin, *Org. Lett.* 20 (2018) 3286–3290.
- K.F. Croom, G.L.J.D. Plosker, Atorvastatin, *Drugs* 65 (2007) 137–152.
- G.A. Roth, C. Johnson, A. Abajobir, F. Abd-Allah, S.F. Abera, G. Abyu, M. Ahmed, B. Aksut, T. Alam, K. Alam, F. Alla, N. Alvis-Guzman, S. Amrock, H. Ansari, B. Arnlöv, H. Asayesh, T.M. Atey, L. Avila-Burgos, A. Awasthi, A. Banerjee, A. Barac, T. Barnighausen, L. Barregard, N. Bedi, E. Belay Ketema, D. Bennett, G. Berhe, Z. Bhutta, S. Bitew, J. Carapetis, J.J. Carrero, D.C. Malta, C.A. Castaneda-Orjuela, J. Castillo-Rivas, F. Catala-Lopez, J.Y. Choi, H. Christensen, M. Cirillo, L. Cooper Jr., M. Criqui, D. Cundiff, A. Damasceno, L. Dandona, R. Dandona, K. Davletov, S. Dharmaratne, P. Dorairaj, M. Dubey, R. Ehrenkranz, M. El Sayed Zaki, E.J.A. Faraon, A. Esteghamati, T. Farid, M. Farvid, V. Feigin, E.L. Ding, F. Fowkes, T. Gebrehiwot, R. Gillum, A. Gold, P. Gona, R. Gupta, T.D. Habtewold, N. Hafezi-Nejad, T. Hailu, G.B. Hailu, G. Hankey, H.Y. Hassen, K.H. Abate, R. Havmoeller, S.I. Hay, M. Horino, P.J. Hotez, K. Jacobsen, S. James, M. Javanbakht, P. Jeemon, D. John, J. Jonas, Y. Kalkonde, C. Karimkhani, A. Kasaeian, Y. Khader, A. Khan, Y.H. Khang, S. Khera, A.T. Khoja, J. Khubchandani, D. Kim, D. Kolte, S. Kosen, K.J. Krohn, G.A. Kumar, G.F. Kwan, D.K. Lal, A. Larsson, S. Linn, A. Lopez, P.A. Lotufo, H.M.A. El Razeq, R. Malekzadeh, M. Mazidi, T. Meier, K.G. Meles, G. Mensah, A. Meretoja, H. Mezgebe, T. Miller, E. Mirzakhimov, S. Mohammed, A.E. Moran, K.I. Musa, J. Narula, B. Neal, F. Ngallesoni, G. Nguyen, C.M. Obermeyer,

- M. Owolabi, G. Patton, J. Pedro, D. Qato, M. Qorbani, K. Rahimi, R.K. Rai, S. Rawaf, A. Ribeiro, S. Safiri, J.A. Salomon, I. Santos, M. Santric Milicevic, B. Sartorius, A. Schutte, S. Sepanlou, M.A. Shaikh, M.J. Shin, M. Shishebor, H. Shore, D.A.S. Silva, E. Sobngwi, S. Stranges, S. Swaminathan, R. Tabares-Seisdedos, N. Tadele Atnafu, F. Tesfay, J.S. Thakur, A. Thrift, R. Topor-Madry, T. Truelsen, S. Tyrovolas, K.N. Ukwaja, O. Uthman, T. Vasankari, V. Vlassov, S.E. Vollset, T. Wakayo, D. Watkins, R. Weintraub, A. Werdecker, R. Westerman, C.S. Wiysonge, C. Wolfe, A. Workicho, G. Xu, Y. Yano, P. Yip, N. Yonemoto, M. Younis, C. Yu, T. Vos, M. Naghavi, C. Murray, Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015, *J. Am. Coll. Cardiol.* 70 (2017) 1–25.
- [36] M. Mohammadi, A. Sakhteman, S. Ahrari, K. Hassanpour, S.E. Hashemi, G. Farnooosh, Disulfide bridge formation to increase thermostability of DFPase enzyme: A computational study, *Comput. Biol. Chem.* 77 (2018) 272–278.
- [37] J. Wang, J. Chen, X. Tang, Y. Li, R. Zhang, L. Zhu, Y. Sun, Q. Zhang, W. Wang, Catalytic mechanism for 2,3-dihydroxybiphenyl ring cleavage by nonheme extradiol dioxygenases BphC: Insights from QM/MM analysis, *J. Phys. Chem. B* 123 (2019) 2244–2253.
- [38] I. Zoi, J. Suarez, D. Antoniou, S.A. Cameron, V.L. Schramm, S.D. Schwartz, Modulating enzyme catalysis through mutations designed to alter rapid protein dynamics, *J. Am. Chem. Soc.* 138 (2016) 3403–3409.
- [39] M.A. Galmés, E. García-Junceda, K. Świderek, V. Moliner, Exploring the origin of amidase substrate promiscuity in CALB by a computational approach, *ACS Catal.* 10 (2019) 1938–1946.
- [40] Z. Wang, S. Zhou, S. Zhang, S. Zhang, F. Zhu, X. Jin, Z. Chen, X. Xu, Semi-rational engineering of a thermostable aldo-keto reductase from *Thermotoga maritima* for synthesis of enantiopure ethyl-2-hydroxy-4-phenylbutyrate (EHPB), *Sci. Rep.* 7 (2017) 4007.
- [41] W. Zhang, M. Jia, S. Yu, T. Zhang, L. Zhou, B. Jiang, W. Mu, Improving the thermostability and catalytic efficiency of the d-psicose 3-epimerase from *Clostridium bolteae* ATCC BAA-613 using site-directed mutagenesis, *J. Agric. Food Chem.* 64 (2016) 3386–3393.
- [42] S.H. Baik, F. Michel, N. Aghajari, R. Haser, S. Harayama, Cooperative effect of two surface amino acid mutations (Q252L and E170K) in glucose dehydrogenase from *Bacillus megaterium* IWG3 on stabilization of its oligomeric state, *Appl. Environ. Microbiol.* 71 (2005) 3285–3293.
- [43] X. Feng, H. Tang, B. Han, B. Lv, C. Li, Enhancing the thermostability of  $\beta$ -glucuronidase by rationally redesigning the catalytic domain based on sequence alignment strategy, *Ind. Eng. Chem. Res.* 55 (2016) 5474–5483.
- [44] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [45] N. Greenfield, Analysis of circular dichroism data, *Methods enzymol.* 383 (2004) 282–317.
- [46] F. Cheng, J. Yang, M. Bocola, U. Schwaneberg, L. Zhu, Loop engineering reveals the importance of active-site-decorating loops and gating residue in substrate affinity modulation of arginine deiminase (an anti-tumor enzyme), *Biochem. Biophys. Res. Commun.* 499 (2018) 233–238.
- [47] E. Krieger, G. Koraimann, G. Vriend, Increasing the precision of comparative models with YASARA NOVA-A self-parameterizing force field, *Proteins* 47 (2002) 393–402.
- [48] H. Jornvall, B. Persson, M. Krook, S. Atrian, R. Gonzalez-Duarte, J. Jeffery, D. Ghosh, Short-chain dehydrogenases/reductases (SDR), *Biochemistry* 34 (1995) 6003–6013.
- [49] V. Carbone, M. Giglio, R. Chung, T. Huyton, J. Adams, R. Maccari, R. Ottana, A. Hara, O. El-Kabbani, Structure of aldehyde reductase in ternary complex with a 5-arylidene-2,4-thiazolidinedione aldose reductase inhibitor, *Eur. J. Med. Chem.* 45 (2010) 1140–1145.
- [50] K.L. Griffith, R.E. Wolf, A comprehensive alanine scanning mutagenesis of the *Escherichia coli* transcriptional activator SoxS: identifying amino acids important for DNA binding and transcription activation, *J. Mol. Biology* 322 (2002) 237–257.
- [51] A.M. Gazzali, C. Boura, T. Peterlini, L. Colombeau, S. Acherar, C. Frochet, R. Vanderesse, Modification of KDKPPR peptide through alanine-scanning technique to investigate the effect on its binding on neuropilin-1 receptor for photodynamic therapy application, *Photodiagnosis Photodyn. Ther.* 17 (2017) A27.
- [52] T.M. Penning, The aldo-keto reductases (AKRs): Overview, *Chem. Biol. Interact.* 234 (2015) 236–246.
- [53] G. Sanli, J.I. Dudley, M.J.C.B. Blaber, Biophysics, structural biology of the aldo-keto reductase family of enzymes: Catalysis and cofactor binding, *Cell Biochem. Biophys.* 38 (2003) 79–101.
- [54] A.C. Reyes, T.L. Amyes, J.P. Richard, Enzyme architecture: A startling role for Asn270 in glycerol 3-phosphate dehydrogenase-catalyzed hydride transfer, *Biochemistry* 55 (2016) 1429–1432.
- [55] S.H. Hong, H.K. Nam, K.R. Kim, S.W. Kim, D.K. Oh, Molecular characterization of an aldo-keto reductase from *Marivirga tractuosa* that converts retinal to retinol, *J. Biotechnol.* 169 (2014) 23–33.
- [56] T. Shi, L. Liu, W. Tao, S. Luo, S. Fan, X.L. Wang, L. Bai, Y.L. Zhao, Theoretical studies on the catalytic mechanism and substrate diversity for macrocyclization of pikromycin thioesterase, *ACS Catal.* 8 (2018) 4323–4332.
- [57] C.N. Pace, H. Fu, K.L. Fryar, J. Landua, S.R. Trevino, B.A. Shirley, M.M. Hendricks, S. Iimura, K. Gajiwala, J.M. Scholtz, G.R. Grimsley, Contribution of hydrophobic interactions to protein stability, *J. Mol. Biol.* 408 (2011) 514–522.
- [58] X. Luo, Y.J. Wang, Y.G. Zheng, Cloning and characterization of a NADH-dependent aldo-keto reductase from a newly isolated *Kluyveromyces lactis* XP1461, *Enzyme Microb. Technol.* 77 (2015) 68–77.
- [59] L.J. Giver, L.M. Newman, E. Mundorff, G.W. Huisman, S.J. Jenne, J. Zhu, B. Behrouzian, J.H. Grate, J. Lalonde, Ketoreductase enzymes and uses thereof, U.S. Patent 7879585B2 (2011).