ACS Catalysis

Research Article

pubs.acs.org/acscatalysis

Efficient Synthesis of Non-Natural L-2-Aryl-Amino Acids by a Chemoenzymatic Route

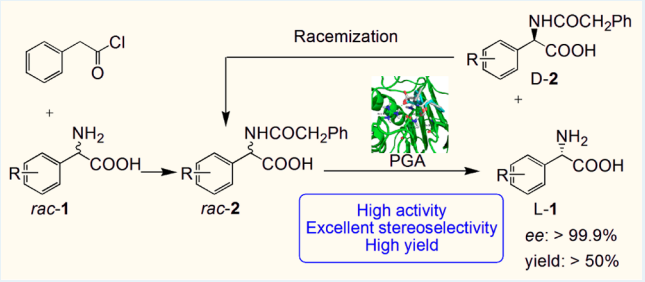
Ya-Ping Xue,[†,‡] Yu-Guo Zheng,*[,†,‡] Zhi-Qiang Liu,[†,‡] Xue Liu,[†,‡] Jian-Feng Huang,[†,‡] and Yin-Chu Shen[†,‡]

[†]Institute of Bioengineering, Zhejiang University of Technology, Hangzhou 310014, China

[‡]Engineering Research Center of Bioconversion and Biopurification of Ministry of Education, Zhejiang University of Technology, Hangzhou 310014, China

Supporting Information

ABSTRACT: Enantiopure non-natural 2-aryl-amino acids are important intermediates for synthesizing pharmaceuticals. To develop an efficient chemoenzymatic process to produce non-natural 2-aryl-amino acids, a penicillin G acylase (PGA) gene from *Bacillus megaterium* was cloned and expressed in *Bacillus subtilis* WB800. The recombinant PGA exhibited a high hydrolytic activity and excellent enantioselectivity (E > 200) toward *N*-phenylacetyl derivatives of non-natural 2-aryl-amino acids. The L-2-aryl-amino acids were obtained in >99.9% enantiomeric excess (ee) and >49% conversion within 3 h. The position and type of the substituent in the substrate influence the recombinant PGA activity but do not affect the PGA enantioselectivity. The kinetic parameters of the recombinant PGA for different substrates were determined and compared. The mechanisms of enantioselectivity of PGA with respect to different substrates were elucidated. The chiral discrimination of PGA with respect to *rac*-**2a**−**e** was mainly because the D-substrates used in this study cannot interact with the active residues and bind to the active pocket as stably as the L-substrates. The unreacted *N*-phenylacetyl-D-2-aryl-amino acids can be completely racemized at 170 °C and then used as the substrate. A gram-scale production of L-2-aryl-amino acids was successfully achieved with approximately theoretical conversion, indicating that the chemoenzymatic approach appears to be promising for industrial applications.



KEYWORDS: *amino acids, chiral resolution, deracemization, enzyme catalysis, hydrolysis, penicillin G acylase*

INTRODUCTION

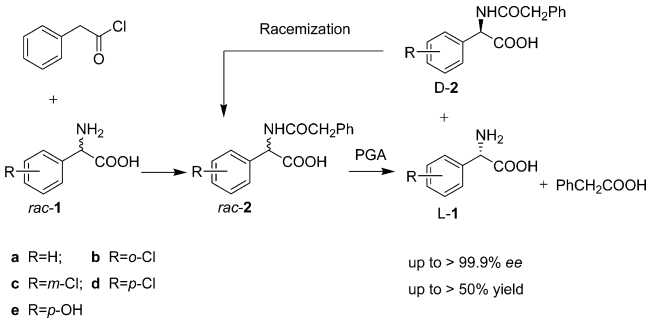
Enantiopure non-natural 2-aryl-amino acids have been utilized as important building blocks or starting materials for the production of pharmaceuticals.[1] For instance, L-phenylglycine and its derivatives are important intermediates for synthesizing proctolin analogues[2] and HIV protease inhibitor or paclitaxel.[3] Because of their importance, tremendous efforts have been made in exploring efficient methods to produce single-isomer compounds.[4] The traditional chemical route for preparing enantiopure non-natural aryl-amino acids was realized by using the classical Strecker reaction combined with a resolution step via the formation of diastereomeric salts using camphor sulfonic acid or tartaric acid, followed by fractional crystallization.[5] Although this method is one of the most important ways for preparing chiral non-natural 2-aryl-amino acids, it does not always work because of the low enantioselectivity or yield (<50%), which is unfavorable for commercial production. As an alternative to the chemical approach, fermentation-based routes appear to be economical for the preparation of enantiopure L-hydroxyphenylglycine[6] or D-phenylglycine[7] by use of metabolic engineering methods. However, enantiopure 2-aryl-amino acids are not easily accessible with short biosynthesis pathways in organisms but need complicated multistep engineering. An enzymatic route using hydantoinase-carbamoylase as the biocatalyst has been proven to be a commercially successful application for the preparation of a large number of chiral non-natural 2-aryl-amino acids.[4,8] Unfortunately, the hydantoinase-mediated process is still limited to the production of the D-enantiomer. The search for economically industrial processes for the preparation of non-natural L-2-aryl-amino acids with high yield and enantiomeric excess (ee) is a very important goal in the chemical industry.[1c,4]

Chemoenzymatic synthesis, combining the flexibility of chemical synthesis and the high selectivity of enzymatic synthesis, is a powerful method for the synthesis of chiral compounds[9] such as enantiopure amines,[10] alcohols,[11] and amino acids.[12] Although a few chemoenzymatic processes have also been reported for synthesizing the enantiopure non-natural 2-aryl-amino acids,[12] the large-scale application of these methods was limited because of the low activity of the

Received: December 29, 2013
Revised: July 27, 2014
Published: July 30, 2014

biocatalyst, low space−time yield, or narrow substrate scope.[4,12a] Therefore, an efficient process for the industrial production of chiral non-natural L-2-aryl-amino acids remains a challenge. In this work, we provided a simple methodology using the chemoenzymatic route for the efficient synthesis of non-natural L-2-aryl-amino acids in high yield and ee, which is based on enantioselective hydrolysis of *N*-phenylacetyl derivatives of non-natural 2-aryl-amino acids by a highly active and stereoselective recombinant penicillin G acylase (PGA)[12c] (Scheme 1). Various substituted non-natural L-2-aryl-amino

Scheme 1. Production of Enantiopure Non-Natural L-2-Aryl-Amino Acids by a Chemoenzymatic Route Using PGA[a]



[a]The residual D-**2** can be completely racemized and then used as the enzyme substrate.

acids can be prepared with near-perfect stereopurity and approximately theoretical conversion. The mechanism of enantioselectivity of PGA toward those substrates was elucidated using molecular modeling and docking. The unreacted D-enantiomer can be easily racemized by heating for recycling. Enantioselective deacylation combined with chemical racemization of the unreacted D-enantiomer has allowed dynamic resolutions giving a >50% yield. This chemoenzymatic approach was explored to improve the atom economy and develop an efficient "green" process for synthesizing L-2-aryl-amino acids.

RESULTS AND DISCUSSION

Recombinant PGA. To achieve economic feasibility and competitiveness for industrially chemoenzymatic synthesis, finding an ideal PGA that shows high stereoselectivity and activity toward *rac*-**2** is desirable. A genome mining approach[13] was used to search for a robust PGA. Sequence searching in the GenBank with "penicillin G acylase" as keywords was conducted, and 12 representative precursor sequences that have been experimentally confirmed from different bacteria were selected to perform the multiple-sequence alignment using DNAMAN (Supporting Information, Figure S1). The precursor of PGA consisted of a signal peptide, an α subunit, a linker peptide, and a β subunit. There were several conserved regions. The probe searching region should be located at the β subunit because it is the catalytic subunit. A conserved motif of NWNNKP was detected, in which one of the N's is crucial for the hydrolysis mediated by PGA,[14] and K exists at a higher frequency. This conserved motif was selected as a probe, and a protein BLAST was conducted to search homologous protein sequences in the GenBank. Eleven hits named PGA were obtained from microbial genomes. Via deletion of the highly homologous hits that are from the same kind of bacteria with >90% homology and the published sequences,[15] the PGA genes from *Arthrobacter viscosus* (AAA22077.1), *Bacillus badius* (DQ115799.1), *Bacillus metaterium* (AF161313.1), and *Kluyvera cryocrescens* (M15418.1) were selected. After being synthesized *in vitro* and cloned into pGEM-T, the four PGA genes were then inserted into expression vector pPZW103 and expressed in *Bacillus subtilis* WB800. *B. subtilis* WB800 was used for protein expression because it has outstanding secretion ability, displays fast growth, and is a nonpathogenic bacterium free of endotoxins.[16] Moreover, secreted proteins usually remain in biologically active forms, and the downstream purification step is greatly simplified.[17] However, only the gene with GenBank accession number AF161313.1 (*B. megaterium* CA4098) was actively expressed in *B. subtilis* WB800. Fortunately, the expressed PGA from *B. megaterium* CA4098 showed satisfactory activity and perfect enantioselectivity toward *rac*-**2b**, a representative substrate. The cultivation conditions for preparing the recombinant PGA with *B. subtilis* WB800/pPZW103-PGA were then optimized by single-factor experiments. The optimal fermentation medium contained 10.0 g/L soluble starch, 12.0 g/L peptone, 3.0 g/L yeast extract, and 10.0 g/L sodium chloride at an initial pH of 7.5. The optimal temperature, air flow rate, and agitation speed for cell growth and PGA expression were 37 °C, 0.5 vvm (air volume per culture volume per minute), and 150 rpm, respectively. Under these conditions, the cultivation process in a 5 L fermentor (working volume of 3 L) yielded a total PGA activity of 15.6 units/mL and a biomass of 6.3 g (dry weight)/L after fermentation for 24 h (Supporting Information, Figure S2). The overexpressed recombinant PGA from the fermentation broth of *B. subtilis* WB800/pPZW103-PGA was then isolated and purified. Through cell-free extraction using ammonium sulfate at 80% saturation, DEAE-IEX chromatography, and butyl HIC chromatography, approximately 4.6-fold purification with an overall yield of 55.5% was achieved. The final purified recombinant PGA gave two bands on sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) (Figure 1), which showed the enzyme consisted of an α subunit and a β subunit with molecular masses of approximately 26 and 61 kDa, respectively.

PGA-catalyzed hydrolytic kinetic resolution of *rac*-**2b** was performed as a model reaction to study the recombinant PGA properties. Operation parameters, including temperature, pH, substrate concentration, and product concentration, of the hydrolysis mediated by the recombinant PGA were inves-

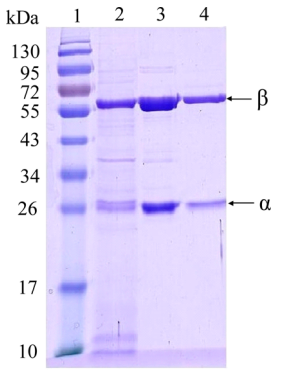


Figure 1. SDS−PAGE of the purified recombinant PGA: lane 1, molecular weight standards; lane 2, crude enzyme; lane 3, enzyme sample from the DEAE-IEX column; lane 4, enzyme sample from the final purification step.

dx.doi.org/10.1021/cs500535d | *ACS Catal.* 2014, 4, 3051−3058

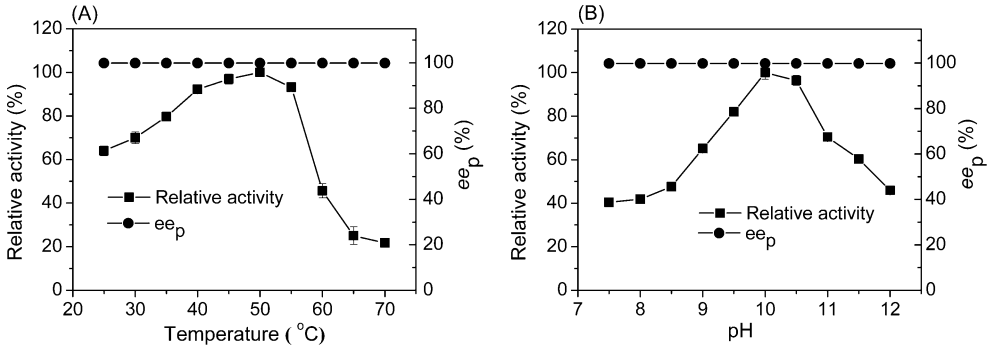


Figure 2. Effects of (A) temperature and (B) pH on the kinetic resolution of *rac*-**2b** catalyzed by the recombinant PGA.

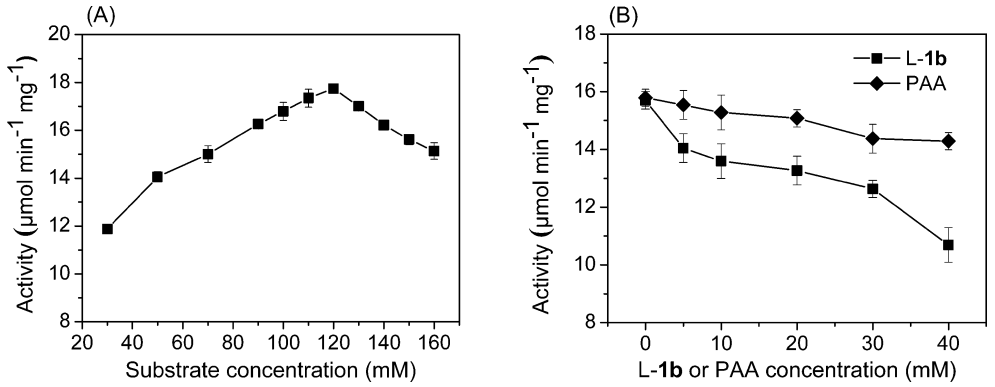


Figure 3. Effect of (A) substrate concentration and (B) product concentration on the kinetic resolution of *rac*-**2b** catalyzed by the recombinant PGA.

tigated. The effect of temperature on the PGA-catalyzed hydrolysis was determined between 25 and 70 °C. The activity and ee$_p$ (ee of product) at different temperatures are shown in Figure 2A. The recombinant PGA showed maximal activity at 50 °C. The temperature change showed no effect on the ee$_p$ (>99.9%), which suggested the enantioselectivity of the recombinant PGA-catalyzed resolution is not sensitive to the temperature. This differs from the PGA extracted from *Escherichia coli*, for which the higher temperature led to a decrease in enantioselectivity.[18] The thermostability was then investigated by assaying enzyme residual activity after preincubation at different temperatures (40, 50, and 60 °C) for the required time intervals. The half-lives of the recombinant PGA were calculated to be 75.3, 24.9, and 2.0 h at 40, 50, and 60 °C, respectively. The thermostability of this recombinant PGA is higher than that of acylase from *Pseudomonas* sp. SE83,[19] *Alcaligenes faecalis*, and *E. coli*.[20] The half-life of the *A. faecalis* PGA at 55 °C was 15 min. More than 50% of the *E. coli* PGA become inactive after 10 min at 50 °C.[20] The effect of pH on the PGA-catalyzed resolution was determined at pH 7.5−12.0. The highest activity was achieved at pH 10.0 (Figure 2B). When the pH was >11.0 or <9.0, the activity decreased sharply. The ee$_p$ showed little variation and remained >99% at pH 7.5−12.0.

Figure 3A shows the effect of substrate concentration on the kinetic resolution in the range of 30−160 mM. The PGA activity was increased from 11.8 to 17.7 μmol min^{-1} mg^{-1} when the *rac*-**2b** concentration was increased from 30 to 120 mM, indicating that a high concentration of substrate up to 120 mM did not result in any decrease in enzyme activity. Phenylacetic acid (PAA) and L-**1b** are two PGA-catalyzed reaction products. Figure 3B shows the enzyme activities after the addition of different L-**1b** or PAA concentrations to the reaction mixture. Both products (L-**1b** and PAA) exhibited an inhibitory effect on recombinant PGA. L-**2b** showed more significant inhibition than PAA did when both were present at the same concentration.

PGA-Catalyzed Kinetic Resolution of *rac*-2a−e. The kinetic resolution of *rac*-**2a−e** with recombinant PGA was performed and compared. The conversion of substrate, ee$_p$, and enantioselectivity were determined. As shown in Table 1, the recombinant PGA exhibited an excellent enantioselectivity (E > 200) and a remarkably high conversion (up to 49.6% conversion within 180 min). The position and type of the substituent in substrates *rac*-**2a−e** influence the recombinant PGA activity but do not affect the PGA enantioselectivity. This recombinant PGA may serve as a versatile catalyst for the synthesis of enantiopure non-natural L-2-aryl-amino acids. It is

Table 1. Resolution of *rac*-2a−e with the Recombinant PGA[a]

substrate	reaction time (min)	conversion (%)	ee$_p$[b] (%)	E[c]
rac-**2a**	140	49.8	≥99.9	>200
rac-**2b**	180	49.9	≥99.9	>200
rac-**2c**	110	49.9	≥99.9	>200
rac-**2d**	50	49.9	≥99.9	>200
rac-**2e**	160	49.6	≥99.9	>200

[a]Reaction conditions: 60 mM racemic compounds **2a−e**, 126 μg/mL PGA, 200 mL reaction volume, pH 10.0, 40 °C, 180 rpm. [b]ee value of products L-**1a−e**. ee$_p$ = ([L-**1**] − [D-**1**])/([L-**1**] + [D-**1**]) × 100%, where [L-**1**] and [D-**1**] are the concentrations of L-**1** and D-**1**, respectively. [c]Enantioselectivity (E) was calculated according to the method of Chen et al.:[21] $E = \ln[(1 - c)(1 + ee_p)]/\ln[(1 - c)(1 - ee_p)]$.

dx.doi.org/10.1021/cs500535d | *ACS Catal.* 2014, 4, 3051−3058

well-known that enantioselectivity is the intrinsic property of an enzyme with respect to a specific substrate.[21] A high enantioselectivity (E > 200) of the PGA will be more favorable for industrial application.

In general, hydrolases show different enantioselectivities toward ortho-, meta-, and para-substituted substrates because of steric hindrance.[22] It is necessary to find different enzymes or substrate structure to meet the requirement of enantioselectivity. The recombinant PGA used in this paper, surprisingly, exhibited excellent enantioselectivity for all the *N*-phenylacetyl derivatives of the 2-aryl-amino acids tested. Among these substrates, this enzyme showed the highest hydrolytic activity toward *rac*-**2d**. The conversion reached 49.9% after reaction for only 50 min. Compared with that of *rac*-**2d**, a relative longer reaction time was needed for *rac*-**2a**−**c** and *rac*-**2e** to achieve this conversion of ∼50% (Table 1).

The kinetic behavior of recombinant PGA toward *rac*-**2a**−**e** was described well by the Michaelis−Menten model at low substrate concentrations. The Michaelis−Menten constant (K_m), maximal activity (V_{max}), and turnover number (k_{cat}) for the hydrolysis of different substrates by recombinant PGA were on the same order of magnitude (Table 2). The PGA

Table 2. Kinetic Parameters of the Purified PGA toward *rac*-2a−e

substrate	V_{max} (μmol min^{-1} mg^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
rac-**2a**	25.1	21.0	35.8	1.70×10^3
rac-**2b**	19.8	19.5	28.3	1.45×10^3
rac-**2c**	27.3	20.8	39.0	1.88×10^3
rac-**2d**	27.4	18.8	39.1	2.08×10^3
rac-**2e**	22.0	21.1	31.4	1.49×10^3

maintained the high catalytic efficiency and excellent enantioselectivity (E > 200) when the chlorine substituent in the phenyl ring was shifted from the para position to the ortho position. The highest specificity constant was observed for *rac*-**2d**. The specificity constant of *rac*-**2b** is lower than those of the other substrates.

Homology Modeling and Molecular Docking. PGA is one member of the N-terminal nucleophile (Ntn) hydrolase protein superfamily. The Ntn functions as the catalytic residue.[23] The acyl carbon in the substrate is first attacked by the N-terminal serine B1 Oγ atom in the B chain. The formed tetrahedral intermediate is stabilized by interactions with the main-chain amides of Ala B69 and the Nδ atom of Asn B245, resulting in an oxyanion hole as seen in the serine proteinases.[14] This tetrahedral intermediate collapses to form a seryl acyl enzyme and releases the free 2-aryl-amino acid. The acyl enzyme is attacked by water to form a second tetrahedral intermediate, stabilized by the same interactions as the first, which can in turn collapse to release the free phenylacetic acid (Supporting Information, Figure S3).

To elucidate the observed enzymatic activities and enantioselectivities of recombinant PGA toward substrates *rac*-**2a**−**e**, homology modeling with the PGA from *Providencia rettgeri* [Protein Data Bank (PDB) entry 1CP9, resolution of 2.5 Å] as the template and molecular docking was conducted. Figure 4 shows the binding modes of the substrates in the active site. The results indicate that L-**2a**−**e** were successfully docked into the PGA active site. All the L-substrates formed hydrogen bonds with Asn B245 by N-phenylacetylated amino acids and oxygen. The distances between the key residues in the active site and the L-substrates are listed in Table 3. The favored substrate L-**2d** has a relatively shorter distance between the catalytic residues. The nucleophilic attack from the N-terminal serine B1 Oγ atom in the B chain is enhanced because of an interaction with a shorter distance. On the other hand, the shorter distance of interaction with the main-chain amide of Ala B69 makes the tetrahedral intermediate more stable for nucleophilic attack. The recombinant PGA showing the higher hydrolytic activity on L-**2d** could also be attributed to the electron-withdrawing substituent of the chloro moiety.[24] The recombinant PGA has the lowest activity toward substrate L-**2b**, which was due to the distance between the carbonyl of substrate L-**2b** and the catalytic residues being the longest. A steric factor associated with the ortho position substituent of the benzene ring in L-**2b** might also have a negative impact on PGA activity. For substrate L-**2e**, the electron-donating group in the phenyl ring might inhibit the PGA activity by increasing the bond energy of the amide group. In the case of D-substrates, the distances between the Ser B1 oxygen and the carbonyl carbon of amide bond are greater than 3.5 Å, which is too remote for catalytic interaction (Table 3). The attack by the N-terminal serine B1 Oγ atom in the B chain becomes very weak. In addition, the hydrogen bonds between the Asn B245 nitrogen and the carbonyl oxygen of N-phenylacetylated amino acids were not formed normally, which may break the stabilization of the tetrahedral intermediate (Table 4). These factors may be the most critical for the enantioselectivity of the PGA. Arg B266 is also involved in substrate binding and helps to position Asn B245.[23b] The interaction between the side chain of Arg B266 and the carboxyl of substrates that stabilizes the binding conformation also disappeared (Figure 4). Therefore, the perfect chiral discrimination of PGA with respect to *rac*-**2a**−**e** was mainly because the D-substrates used in this study cannot interact with the active residues and bind to the active pocket as stably as the L-substrates.

Scale-Up Preparation of L-2-Aryl-Amino Acids. The bioprocess was scaled up to confirm the feasibility for the production of non-natural L-2-aryl-amino acids in practical applications. A 500 mL-scale reaction was performed for the production of L-**1b** with a 1.0 L stirred-tank bioreactor at 40 °C and pH 10.0. Total amounts of 63 mg of enzyme powder and 12.14 g of substrate *rac*-**2b** were added to the reaction mixture. The enzymatic reaction was monitored by chiral high-performance liquid chromatography (HPLC) analysis. The result is shown in Figure 5. The substrate conversion reached a value of ∼50% after reaction for 4 h; 39.88 mM L-**1b** was formed, and substrate L-**2b** was almost exhausted. An excellent optical purity of product (ee_p > 99.9%) and enantioselectivity of PGA (E > 200) were observed. The catalyst productivity (grams of product per gram of catalyst) was calculated to be as high as 58.7 g/g. The supernatant solution was acidified with 6.0 M hydrochloric acid to pH 2.0 after being cooled on ice. Then the unreacted D-**2b** and PPA in the solution were precipitated. The aqueous solution was harvested and the pH adjusted to 7.0. L-**1b** was obtained via the evaporation of an aqueous solution with a rotavapor. The product was then redissolved in hot 2-propanol. The 2-propanol layer was evaporated to obtain L-**1b** (3.50 g, isolated yield of 94.59%, ee_p > 99.9%): $[\alpha]_D^{25}$ +90 (c 1, 1 M HCl) [lit.[12c] +89 (c 1, 1 M HCl)]; ^{1}H NMR (500 MHz, D_2O) δ 7.52 (d, J = 8.0 Hz, 1H), 7.50−7.34 (m, 3H), 5.15 (s, 1H); MS (ESI) m/z 186 $(M)^+$.

dx.doi.org/10.1021/cs500535d | ACS Catal. 2014, 4, 3051−3058

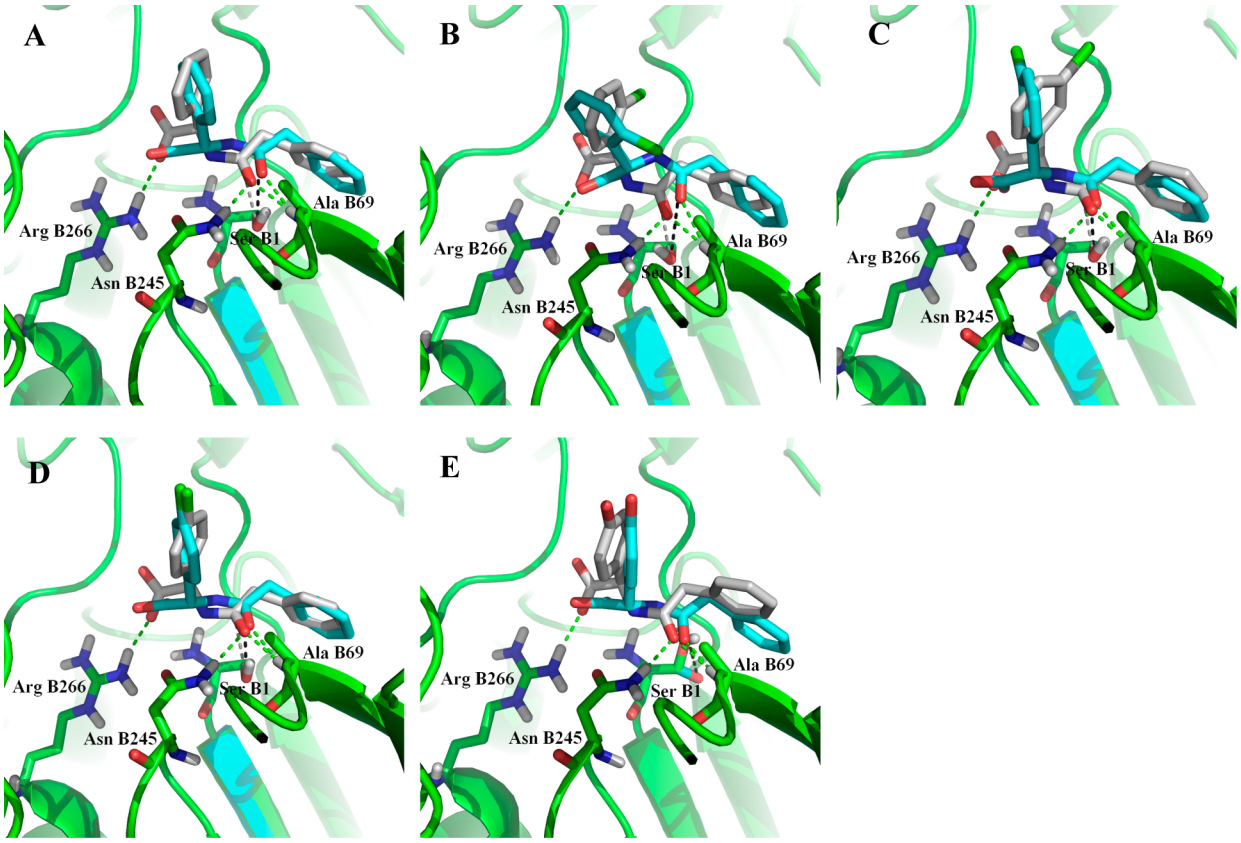


Figure 4. Docking of substrates D,L-**2a** (A), D,L-**2b** (B), D,L-**2c** (C), D,L-**2d** (D), and D,L-**2e** (E) to the active pocket of the recombinant PGA. The recombinant PGA protein is shown as a green cartoon. The carbons of L-**2a**−**e** and D-**2a**−**e** are colored white and blue, respectively. The oxygen and nitrogen in the active pocket residues and substrates are colored red and dark blue, respectively. The hydrogen bonds are represented by green dashed lines. The distances between the serine B1 Oγ atom and the carbonyl carbon are represented by gray dashed lines for L-**2a**−**e** and black dashed lines for D-**2a**−**e**.

Table 3. Results of Docking Experiments with D,L-2a−e

	distance (Å)[a]		
substrate	Ser B1 O−C═O	Ala B69 NH−O═C	Asn B245 NH−O═C
L-**2a**	3.304	2.232	1.906
L-**2b**	3.437	2.198	1.924
L-**2c**	3.013	2.125	1.925
L-**2d**	3.007	2.000	2.053
L-**2e**	3.168	2.231	1.918
D-**2a**	3.556	2.156	ND[b]
D-**2b**	4.168	2.433	ND[b]
D-**2c**	3.536	2.246	ND[b]
D-**2d**	3.588	2.105	ND[b]
D-**2e**	4.411	1.889	ND[b]

[a]Ser B1 O−C═O is the distance between the Ser B1 oxygen and the carbonyl carbon of the amide bond. Ala B69 NH−O═C is the distance between the Ala B69 nitrogen and the carbonyl oxygen of the amide bond. Asn B245 NH−O═C is the distance between the Asn B245 nitrogen and carbonyl oxygen of the amide bond. [b]No hydrogen bond detected.

Table 4. Results of Racemization of D-2a−e

substrate	racemization rate (%)	*rac*-**2** (g)	extraction yield (%)
D-**2a**	100	1.46	90.5
D-**2b**	100	1.72	94.7
D-**2c**	100	1.69	92.8
D-**2d**	100	1.53	84.0
D-**2e**	100	1.54	90.0

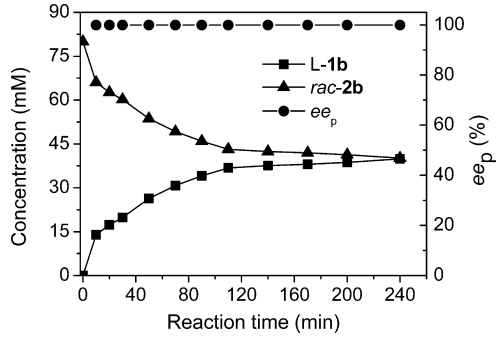


Figure 5. PGA-catalyzed kinetic resolution of *rac*-**2b** for synthesizing L-**1b** in a 500 mL reaction mixture. Reaction conditions: 80 mM substrate, 126 μg/mL PGA, pH 10.0, 40 °C, 180 rpm.

The large-scale result confirmed that the PGA-catalyzed hydrolytic kinetic resolution of racemic *N*-phenylacetyl derivatives of 2-aryl-amino acids to L-2-aryl-amino acids has the potential for practical application. Therefore, the PGA reported in this work should be considered as a much more promising catalyst in the industrial synthesis of chiral L-2-aryl-amino acids in the future.

Chemical Racemization of Unreacted D-2a−e. Using 200 mL of an enzymatic reaction mixture at a substrate concentration of 60 mM, the reaction solution containing 30

dx.doi.org/10.1021/cs500535d | *ACS Catal.* 2014, 4, 3051−3058

mM D-**2a−e** determined by chiral HPLC was acidified with hydrochloric acid (6.0 M) to pH 2.0 after being cooled in ice. The unreacted D-**2a−e** and PAA in the solution were then precipitated. The precipitate was filtered and dried to obtain the mixture of D-**2a−e** and PAA. The mixture of D-**2a−e** and PAA recovered after the enzymatic reaction was heated at 170 °C in an oil bath for 15 min and then cooled. The racemization was confirmed by chiral HPLC analysis. The results indicated that D-**2a−e** were completely racemized to *rac*-**2a−e** (Table 4). The mixture was extracted with cyclohexane, and the residue gave *rac*-**2a−e** for recycling. Racemization thus makes the process overcome the limitation of the maximal theoretical yield of 50% encountered in the classical resolution of racemic non-natural L-2-aryl-amino acids.

CONCLUSION

In summary, we have described a practical procedure for the synthesis of L-2-aryl-amino acids using recombinant PGA as a biocatalyst. The PGA gene from *B. megaterium* was cloned and expressed in *B. subtilis* WB800. The recombinant PGA exhibited a high hydrolytic activity and excellent enantioselectivity (E > 200) toward *N*-phenylacetyl derivatives of non-natural 2-aryl-amino acids to prepare non-natural L-2-aryl-amino acids. Molecular modeling and docking with substrates were then performed to elucidate the PGA's substrate specificities and enantioselectivities. The preparative reaction for the production of L-**1b** was successfully performed. Enantiopure L-**1b** (>99.9% ee) was obtained in high yield. The results indicated that this recombinant PGA represents a potentially promising biocatalyst for the preparation of L-2-aryl-amino acids in practical applications. The unreacted *N*-phenylacetyl D-2-aryl-amino acids can be completely racemized at 170 °C and then used as the substrate. The PGA-mediated chemoenzymatic route offers significant advantages over conventional kinetic resolution approaches because of its high activity and excellent enantioselectivity. In addition, this method provides ∼100% conversion. The recombinant PGA was used to catalyze resolution of racemic mixtures of 2-aryl-amino acids, which might boost the frequency of application of PGA in the synthesis of enantiopure amino acids and their derivatives. Further studies, including protein engineering to improve the product tolerance and immobilization to recycle the PGA based on this work, are in progress.

EXPERIMENTAL SECTION

Synthesis and Cloning of the PGA Gene. A protein BLAST was conducted to search homologous protein sequences in the GenBank with the conserved motif of NWNNKP as a probe. The PGA genes from *Ar. viscosus* (AAA22077.1), *B. badius* (DQ115799.1), *B. metaterium* (AF161313.1), and *K. cryocrescens* (M15418.1) were selected. Nucleotide sequences of PGA from these strains were synthesized using the polymerase chain reaction assembly method.[25] The synthesized genes were cloned into pGEM-T and inserted into expression vector pPZW103 between *Sma*I and *Bam*HI restriction endonuclease sites. Ligated plasmid pPZW103 was transformed into *B. subtilis* WB800. For the selection of *B. subtilis* WB800 transformants, 50 μg/mL kanamycin was added to the LB medium. DNA manipulation, plasmid isolation, and agarose gel electrophoresis were performed according to the standard protocol[26] unless stated otherwise.

Enzyme Expression and Purification. *Cultivation of Recombinant B. subtilis WB800/pPZW103-PGA. B. subtilis* WB800/pPZW103-PGA was cultivated on LB agar medium containing 50 μg/mL kanamycin at 37 °C for 20 h. The colony was picked and inoculated in 60 mL of LB medium containing 50 μg/mL kanamycin in a 500 mL flask. After incubation at 37 °C and 150 rpm for 12 h, 2% (v/v) of the seed cultures was inoculated with 3 L of fermentation medium [10.0 g/L soluble starch, 12.0 g/L peptone, 3.0 g/L yeast extract, and 10.0 g/L sodium chloride (pH 7.5)] in a 5 L fermentor. Fermentation was conducted at 37 °C, 0.5 vvm, and 150 rpm. After being cultivated for 24 h, the supernatant was harvested by centrifugation (12000*g* and 4 °C) for 10 min.

Purification of the Recombinant PGA. Ammonium sulfate was slowly added to the supernatant to 80% saturation while the mixture was being stirred at 0 °C. The precipitate was harvested and dissolved in sodium phosphate buffer (0.05 M, pH 7.8). The solution was dialyzed overnight with the same buffer. The dialyzed solution was loaded on an anion ion-exchange column (DEAE-IEX, 5 mL, Bio-Rad Laboratories) using a mobile phase (0.05 M, pH 7.8 phosphate buffer) at a flow rate of 1.0 mL/min. The anion ion-exchange column was eluted with a linear gradient of 0 to 1.0 M sodium chloride buffer. Active PGA fractions were supplemented with ammonium sulfate to a final concentration of 2.0 M by being stirred. The resulting solution was loaded on a hydrophobic interaction chromatography column (Butyl HIC, 10 mL, Bio-Rad Laboratories), which was equilibrated with 20 mM sodium phosphate buffer [2.0 M ammonium sulfate in the buffer (pH 7.2)]. The PGA was then eluted by a linear gradient from 2.0 to 0.6 M ammonium sulfate in 20 mM sodium phosphate buffer (pH 7.2) at a flow rate of 1.0 mL/min. The active PGA fractions were collected and dialyzed. The protein purity was determined by SDS−PAGE. The protein concentration was measured by the Bradford method using the protein assay kit (Bio-Rad Laboratories) with bovine serum albumin as a standard.

Synthesis of *rac*-2a−e. *rac*-**1a−e** (0.1 mol) were dissolved in 4.0 M sodium hydroxide (100 mL) and the mixtures stirred at 0 °C (water/ice bath). Phenylacetyl chloride (1.1 equiv) was added while the mixtures were vigorously stirred. In the case of *rac*-**2d**, the reaction mixture was stirred overnight and a white solid was obtained. The white solid was filtered and dried to yield *rac*-**2d**. For *rac*-**2a−c** and *rac*-**2e**, the reaction mixture was stirred overnight and extracted using dichloromethane. Subsequently, the resulting aqueous layer was acidified to pH 2.0 with hydrochloric acid (6.0 M) after being cooled on ice. The precipitate was filtered, washed with cold water, and dried to yield the four compounds.

*rac-**2a**:* white solid; yield 25.3 g (94%); ^{1}H NMR (500 MHz, DMSO) δ 12.89 (s, 1H), 8.87 (s, 1H), 7.46−7.37 (m, 4H), 7.34 (s, 1H), 7.29 (d, J = 4.4 Hz, 4H), 7.22 (d, J = 4.2 Hz, 1H), 5.35 (s, 1H), 3.57 (s, 2H).

*rac-**2b**:* white solid; yield 28.35 g (93.4%); ^{1}H NMR (500 MHz, DMSO) δ 13.10 (s, 1H), 8.95 (s, 1H), 7.48 (s, 1H), 7.37 (s, 2H), 7.28 (s, 5H), 7.22 (s, 1H), 5.79 (s, 1H), 3.56 (s, 2H).

*rac-**2c**:* light green solid; yield 20.28 g (67%); ^{1}H NMR (500 MHz, DMSO) δ 13.09 (s, 1H), 8.95 (s, 1H), 7.49 (s, 1H), 7.40 (d, J = 1.3 Hz, 3H), 7.29 (d, J = 4.6 Hz, 4H), 7.22 (d, J = 4.3 Hz, 1H), 5.41 (s, 1H), 3.58 (s, 2H).

*rac-**2d**:* white solid; yield 28.80 g (95%); ^{1}H NMR (500 MHz, DMSO) δ 8.06 (s, 1H), 7.27 (qd, J = 14.0, 7.3 Hz, 9H), 4.84 (s, 1H), 3.43 (s, 2H).

dx.doi.org/10.1021/cs500535d | *ACS Catal.* 2014, 4, 3051−3058

rac-**2e**: gray solid; yield 15.5 g (54%); ^{1}H NMR (500 MHz, DMSO) δ 12.68 (s, 1H), 8.72 (s, 1H), 7.44−6.97 (m, 7H), 6.76 (d, J = 8.6 Hz, 2H), 5.17 (s, 1H), 3.53 (s, 2H).

Effect of Reaction Conditions on the Activity and Enantioselectivity of Recombinant PGA. The optimal temperature and pH of recombinant PGA were investigated by incubating 126 μg/mL recombinant PGA with 20 mM *rac*-**2b** at a certain temperature in the range of 25−70 °C and a certain pH in the range of 7.5−12.0. The thermal stability of recombinant PGA was determined by preincubating the enzyme at a certain temperature (40, 50, or 60 °C) for required time intervals. The PGA residual activity was measured at 40 °C and pH 10.0 with 20 mM *rac*-**2b** and 126 μg/mL recombinant PGA. The effect of substrate concentration on the recombinant PGA activity was examined at different *rac*-**2b** concentrations ranging from 30 to 160 mM. The effect of product concentration on recombinant PGA activity was investigated by adding different concentrations of L-**1b** or PAA to the reaction system. The experiments were performed at 40 °C and pH 10.0 with 20 mM *rac*-**2b** and 126 μg/mL recombinant PGA. For the determination of PGA activity, the reactions were performed for 6.0 min. During all these reactions, the pHs were kept constant with pH-stat by adding a 5% ammonia solution. One unit of enzyme activity was defined as the amount of enzyme that produced 1.0 μmol of L-**1b**/min. For the determination of ee_p, the reactions were performed for 24 h. ee_p = ([L-**1b**] − [D-**1b**])/([L-**1b**] + [D-**1b**]) × 100%, where [L-**1b**] and [D-**1b**] are the concentrations of L-**1b** and D-**1b**, respectively.

PGA-Catalyzed Resolution of *rac*-2a−e. *rac*-**2a**−**c** were dissolved by adding a 5% ammonia solution. The pHs were adjusted to 10.0 with 2.0 M hydrochloric acid. *rac*-**2d** and *rac*-**2e** were dissolved in water and the pHs adjusted to 10.0 with a 5% ammonia solution. During the reactions, the pHs were kept constant with pH-stat by adding a 5% ammonia solution. The reactions were conducted at 180 rpm and 40 °C. The resolution process was analyzed by chiral HPLC.

Determination of Kinetic Parameters. The kinetic parameters were determined by performing the experiments at 40 °C and pH 10.0 in the batch reactors at substrate concentrations ranging from 1 to 50 mM. The initial rates were measured using the method described previously.[27] The curve slope during the first 6 min was designated as the initial rate. The maximal activity (V_{max}) and Michaelis−Menten constant (K_m) of the recombinant PGA were obtained by fitting the initial rate versus initial substrate concentration to the Michaelis−Menten equation employing Orgin 6.0 (Microcal Software, Inc.). The turnover number (k_{cat}) was calculated as V_{max}/[PGA].

Analytical Methods. The enantiomers of *rac*-**1** and *rac*-**2** were determined by chiral HPLC as described previously.[12b] The D- and L-enantiomers of the various 2-aryl-amino acids and N-phenylacetylated 2-aryl-amino acids were efficiently separated (Supporting Information, Table S1).

Homology Modeling and Molecular Docking. The Build Homology Models (MODELER) module in Discovery Studio (DS) version 2.1 (Accelrys, Inc.) was used for predicting the three-dimensional (3D) structure of PGA. The crystal structures of PGA from *Providencia rettgeri* (PDB entry 1CP9) was used as the template. 3D models were generated with Modeler 9v8 by Procheck (http://nihserver.mbi.ucla.edu/SAVES/) with the loop refined sufficiently. Several models were achieved and validated by Procheck. The best quality model was shown in the Ramachandran plot, where 91.5% of residues were located in the most favored regions, 6.8% of residues in the additional allowed regions, 1.3% of residues in the generously allowed regions, and two residues (0.4%) in the disallowed regions. The root-mean-square deviation between the model and β subunit of the 1CP9 crystal structure was 0.69 Å. Then molecular docking was performed using Autodock version 4.0.1. The protein structure diagram was prepared with PyMOL.[28]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Time course of PGA production from *B. subtilis* WB800/pPZW103-PGA, proposed catalytic mechanism of PGA, characterization data of new compounds, and HPLC data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: zhengyg@zjut.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge the financial support from the 863 Program (2011AA02A210) and the 973 Project (2011CB710806). We thank Cory Sarks from the Department of Chemical Engineering and Materials Science of Michigan State University (East Lansing, MI) for his assistance.

■ REFERENCES

(1) (a) Bommarius, A. S.; Schwarm, M.; Drauz, K. *J. Mol. Catal. B: Enzym.* **1998**, *5*, 1−11. (b) Patel, R. N. *Coord. Chem. Rev.* **2008**, *252*, 659−701. (c) Resch, V.; Fabian, W. M. F.; Kroutil, W. *Adv. Synth. Catal.* **2010**, *352*, 993−997.

(2) (a) Szeszel-Fedorowicz, W.; Lisowski, M.; Rosinski, G.; Issberner, J.; Osborne, R.; Konopinska, D. *Polym. J. Chem.* **2001**, *75*, 411−417. (b) Bhaskar, G.; Rao, B. V. *Tetrahedron Lett.* **2003**, *44*, 915−917. (c) Pawelczak, K.; Makowski, M.; Kempny, M.; Dzik, J. M.; Golos, B.; Rode, W.; Rzeszotarska, B. *Acta Biochim. Pol.* **2002**, *49*, 407−420.

(3) Wang, Z. M.; Kolb, H. C.; Sharpless, K. B. *J. Org. Chem.* **1994**, *59*, 5104−5105.

(4) Servi, S.; Tessaro, D.; Pedrocchi-Fantoni, G. *Coord. Chem. Rev.* **2008**, *252*, 715−726.

(5) (a) Boesten, W. H. J.; Seerden, J. P. G.; de Lange, B.; Dielemans, H. J. A.; Eisenberg, H. L. M.; Kaptein, B.; Moody, H. M.; Kellogg, R. M.; Broxterman, Q. B. *Org. Lett.* **2001**, *3*, 1121−1124. (b) Lin, S. S.-S.; Chen, C.-C. U.S. Patent 20040176637A1, 2004. (c) Maheshwari, K. K.; Sarma, R. K.; Joshi, S. V.; Barde, A. R.; Sutar, R. P.; Ranade, P. V. U.S. Patent 20040073057A1, 2004. (d) Battula, S. R. WO Patent 2006003671A1, 2006.

(6) (a) Choroba, O. W.; Williams, D. H.; Spencer, J. B. *J. Am. Chem. Soc.* **2000**, *122*, 5389−5390. (b) Hubbard, B. K.; Thomas, M. G.; Walsh, C. T. *Chem. Biol.* **2000**, *7*, 931−942.

(7) Muller, U.; van Assema, F.; Gunsior, M.; Orf, S.; Kremer, S.; Schipper, D.; Wagemans, A.; Townsend, C. A.; Sonke, T.; Bovenberg, R.; Wubbolts, M. *Metab. Eng.* **2006**, *8*, 196−208.

(8) Altenbuchner, J.; Siemann-Herzberg, M.; Syldatk, C. *Curr. Opin. Biotechnol.* **2001**, *12*, 559−563.

(9) (a) Porcar, R.; Sans, V.; Rios-Lombardia, N.; Gotor-Fernandez, V.; Gotor, V.; Burguete, M. I.; Garcia-Verdugo, E.; Luis, S. V. *ACS Catal.* **2012**, *2*, 1976−1983. (b) Yu, H.; Chokhawala, H. A.; Huang, S.; Chen, X. *Nat. Protoc.* **2006**, *1*, 2485−2492. (c) Tao, J.; Zhao, L.; Ran, N. *Org. Process Res. Dev.* **2007**, *11*, 259−267.

(10) Paetzold, J.; Bäckvall, J. E. *J. Am. Chem. Soc.* **2005**, *127*, 17620−17621.
(11) Lee, J. H.; Han, K.; Kim, M.-J.; Park, J. *Eur. J. Org. Chem.* **2010**, *2010*, 999−1015.
(12) (a) Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Kesseler, M.; Stürmer, R.; Zelinski, T. *Angew. Chem., Int. Ed.* **2004**, *43*, 788−824. (b) Xue, Y. P.; Jiang, T.; Liu, X.; Zheng, Y. G. *Biochem. Eng. J.* **2013**, *74*, 88−94. (c) Fadnavis, N. W.; Vedamayee Devi, A.; Swarnalatha Jasti, L. *Tetrahedron: Asymmetry* **2008**, *19*, 2363−2366.
(13) (a) Xu, G.-C.; Yu, H.-L.; Zhang, X.-Y.; Xu, J.-H. *ACS Catal.* **2012**, *2*, 2566−2571. (b) Vergne-Vaxelaire, C.; Bordier, F.; Fossey, A.; Besnard-Gonnet, M.; Debard, A.; Mariage, A.; Pellouin, V.; Perret, A.; Petit, J.-L.; Stam, M.; Salanoubat, M.; Weissenbach, J.; De Berardinis, V.; Zaparucha, A. *Adv. Synth. Catal.* **2013**, *355*, 1763−1779.
(14) Duggleby, H. J.; Tolley, S. P.; Hill, C. P.; Dodson, E. J.; Dodson, G.; Moody, P. C. *Nature* **1995**, *373*, 264−268.
(15) Cheng, T. F.; Chen, M. L.; Zheng, H. B.; Wang, J. G.; Yang, S.; Jiang, W. H. *Protein Expression Purif.* **2006**, *46*, 107−113.
(16) Lin, H.-H.; Yin, L.-J.; Jiang, S.-T. *J. Agric. Food Chem.* **2009**, *57*, 7779−7784.
(17) Zhang, X.-Z.; Cui, Z.-L.; Hong, Q.; Li, S.-P. *Appl. Environ. Microbiol.* **2005**, *71*, 4101−4103.
(18) Abian, O.; Mateo, C.; Palomo, J. M.; Fernández-Lorente, G.; Guisán, J. M.; Fernández-Lafuente, R. *Biotechnol. Prog.* **2004**, *20*, 984−988.
(19) Zhu, X.; Luo, H.; Chang, Y.; Su, H.; Li, Q.; Yu, H.; Shen, Z. *World J. Microbiol. Biotechnol.* **2011**, *27*, 823−829.
(20) Verhaert, R.; Riemens, A. M.; Van der Laan, J.; Van Duin, J.; Quax, W. J. *Appl. Environ. Microbiol.* **1997**, *63*, 3412−3418.
(21) Chen, C. S.; Fujimoto, Y.; Girdaukas, G. *J. Am. Chem. Soc.* **1982**, *104*, 294−1299.
(22) Zheng, G.-W.; Yu, H.-L.; Zhang, J.-D.; Xu, J.-H. *Adv. Synth. Catal.* **2009**, *351*, 405−414.
(23) (a) Brannigan, J. A.; Dodson, G.; Duggleby, H. J.; Moody, P. C.; Smith, J. L.; Tomchick, D. R.; Murzin, A. G. *Nature* **1995**, *378*, 416−419. (b) McVey, C. E.; Walsh, M. A.; Dodson, G. G.; Wilson, K. S.; Brannigan, J. A. *J. Mol. Biol.* **2001**, *313*, 139−150.
(24) (a) Zhu, D.; Rios, B. E.; Rozzell, J. D.; Hua, L. *Tetrahedron: Asymmetry* **2005**, *16*, 1541−1546. (b) Smekal, O.; Reid, G. A.; Chapman, S. K. *Biochem. J.* **1994**, *297*, 647−652.
(25) Rydzanicz, R.; Zhao, X. S.; Johnson, P. E. *Nucleic Acids Res.* **2005**, *33*, W521−W525.
(26) Chong, L. *Science* **2001**, *292*, 446−446.
(27) Xue, Y. P.; Xu, M.; Chen, H. S.; Liu, Z. Q.; Wang, Y. J.; Zheng, Y. G. *Org. Process Res. Dev.* **2013**, *17*, 213−220.
(28) Hallaway, D.; Feiner, S.; Hollerer, T. *Applications of Artificial Intelligence* **2004**, *18*, 477−500.