

Cite this: *Catal. Sci. Technol.*, 2019, 9, 398

# Engineering of lysine cyclodeaminase conformational dynamics for relieving substrate and product inhibitions in the biosynthesis of L-pipecolic acid†

Hanxiao Ying,<sup>a</sup> Jing Wang,<sup>a</sup> Ting Shi,<sup>b</sup> Yilei Zhao,<sup>b</sup>  
Pingkai Ouyang<sup>a</sup> and Kequan Chen \*<sup>a</sup>

Substrate and product inhibitions in enzyme-catalyzed reactions are major limitations in the preparative biosynthesis of valuable chemicals. In the present study, we described a modulation of the conformational dynamics of lysine cyclodeaminase from *Streptomyces pristinaespiralis* (*SpLCD*) for synchronously reducing substrate and product inhibitions. LCD is the key enzyme in the biosynthesis of piperidine derivatives, but incurs both severe substrate and product inhibitions due to steric hindrance by the narrow delivery tunnels of the substrate and product. Conformational dynamics studies *via* molecular simulations, which revealed the detailed atomic structures of both substrate and product delivery processes of *SpLCD*, indicated two separate intrinsic motions affected by the position of NAD<sup>+</sup>. Two key residues, Ile61 and Ile94, were observed to play a key role in regulating the shape of the substrate and product delivery processes. *Via* saturation mutagenesis studies, the Val61-Val94-*SpLCD* variant, which improved the  $K_M/k_{cat}$ ,  $K_i$ -lys and  $K_i$ -LPA by 3.6, 19.4 and 9.2 times, respectively, was obtained. The structure analysis showed that the superior catalytic performance of the Val61-Val94-*SpLCD* variant is mainly due to the expanded substrate and product delivery tunnels when compared with the wild-type enzyme. By using recombinant *Escherichia coli* containing Val61-Val94-*SpLCD* as the whole-cell biocatalyst, a 2.5-fold higher substrate loading concentration with a total space-time yield of 0.83 g L<sup>-1</sup> h<sup>-1</sup> was achieved. Moreover, the L-pipecolic acid titer increased to 73.4 g L<sup>-1</sup> without a decrease in yield, which was 4.2-fold higher than that achieved with the original recombinant whole-cell biocatalyst. The results of this study provide insights into the application of conformational dynamics of protein substrate and product delivery processes for simultaneously reducing both substrate and product inhibitions.

Received 7th November 2018,  
Accepted 30th November 2018

DOI: 10.1039/c8cy02301h

rsc.li/catalysis

## Introduction

Natural enzymes can rarely meet the needs of different applications in industrial catalysis due to highly regulated enzymatic reactions.<sup>1</sup> In typical enzymatic processes, substrate binding and product dissociation are usually considered to determine the overall catalytic rate.<sup>2</sup> Thus, from a practical standpoint, substrate and product inhibitions of enzymes are

considered to be the dominant limitations in biosynthesis of valuable chemicals.<sup>3</sup> Traditional protein engineering methods targeting shared binding sites of substrates and products based on static crystal structure evidence have provided a way to design specific enzyme mutations that can reduce individual substrate or product inhibition.<sup>4</sup> However, due to the distinct tunnel of the substrate entrance and product exit, some enzymes are inhibited by both the substrate and product simultaneously.<sup>5</sup> As static structural or functional studies are inadequate to search for all the delivery modes of substrates and products, the identification, analysis, and re-design procedures for relieving substrate and product inhibitions by separate substrate and product delivery processes are more difficult.<sup>6</sup>

Protein conformational dynamics affects the essential steps in enzyme catalysis, including substrate recognition and binding, allosteric regulation, transition state formation, and product release.<sup>7</sup> Recent reports on protein conformational

<sup>a</sup> College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, 30 South Puzhu Road, Nanjing, Jiangsu, China. E-mail: kqchen@njtech.edu.cn

<sup>b</sup> School of Life Sciences and Biotechnology, Shanghai Jiaotong University, 800 Dongchuan Road, Shanghai, China

† Electronic supplementary information (ESI) available: Experimental procedures including information on molecular chaperone plasmids, RMSFs and RMSDs of MD simulation systems, detailed substrate delivery motions and product release motions, and visualization of substrate tunnel 2 of variant Val61-*SpLCD*. See DOI: 10.1039/c8cy02301h

dynamics have indicated its significant advantages in studying integrated substrate binding and product dissociation processes by revealing the detailed atomic structures of rare states.<sup>8</sup> For example, by using NMR and stopped-flow spectroscopy, the area of product release in dihydrofolate reductase was investigated. The conformational dynamics of ligand dissociation processes showed another efficient product exit pathway which was not observed in the crystal structures, and is likely to occur with allosteric effects.<sup>6b</sup> Despite many advances in understanding the fundamental aspect of protein conformational dynamics, practical implementation of such knowledge for designing superior enzymes by targeting separate substrate and product delivery processes for synchronously reducing both substrate and product inhibitions remains a challenge.<sup>9,10</sup>

L-Pipecolic acid (LPA) is a key chiral component for the synthesis of many important groups of drugs including local anesthetics ropivacaine, bupivacaine and chlorprocaine as well as predominant immunosuppressants rapamycin, FK-506 and FK-520.<sup>11</sup> Similar to the preparation of other enantiomeric organics, pure LPA is mainly produced by chemical enantioselective synthesis and stereoselective transformation.<sup>12</sup> However, the chemical methods for commercial-scale preparation of LPA require harmful substrates, expensive metal catalysts and complicated product purification processes.<sup>13</sup> With the development of genome mining, lysine cyclodeaminase (LCD) catalyzes the heterocyclic bioconversion of linear L-lysine to circular LPA in one step by combining cyclization and deamination processes, which was discovered from *Streptomyces* (Fig. 1).<sup>14</sup> Although a few studies have focused on the bio-production of LPA and its derivatives since the discovery of LCD in 2006, owing to the poor soluble protein expression level and severe substrate and product inhibitions, efficient accumulation of LPA could not be achieved.<sup>15</sup> In our previous studies, the *pipA* gene from *Streptomyces pristinaespiralis* ATCC 25486 encoding lysine cyclodeaminase (*Sp*LCD) was cloned and introduced into recombinant *E. coli*. Using *E. coli*-*Sp*LCD as the original recombinant whole-cell biocatalyst, by optimizing the reaction conditions,

the highest LPA yield of 17.4 g L<sup>-1</sup> was obtained. However, the long reaction time (average of 48–72 h), severe substrate and product inhibitions, and high biocatalyst consumption (cell density of the whole-cell biocatalyst controlled at OD<sub>600</sub> = 200) made it impractical to perform preparative and sustainable production of LPA.<sup>15a</sup>

In the present study, a rational protein engineering strategy was developed by modifying the protein conformational dynamics of the substrate and product delivery processes for simultaneously relieving substrate and product inhibitions. Molecular dynamics (MD) simulations revealed two substrate delivery tunnels and two distinct product exits for LCD, which were affected by the location of NAD<sup>+</sup>. The first substrate delivery tunnel was observed between Asp236 and Glu264 with the corresponding product exit in the vicinity of Ile61 and NAD<sup>+</sup>. The other substrate delivery tunnel was observed between Ile61 and Asp236 with the corresponding product exit near Ile94 and NAD<sup>+</sup>. *Via* analyzing the distance between the C4 position of the NAD<sup>+</sup> ribose ring and α-NH<sub>2</sub> of L-lysine (substrate), both the substrate and product delivery processes were confirmed to be functional. Subsequently, the key role of two non-polar residues, Ile61 (involved in substrate delivery process 2 and product release process 1) and Ile94 (involved in product release process 2), was confirmed through CAVER tunnel computation and mutagenesis. By screening saturation mutations on these two positions, the variant Val61-Val94-*Sp*LCD was obtained and exhibited maximum improvement in the biosynthesis ability of LCD, improving the  $K_M/k_{cat}$ ,  $K_i$ -lys and  $K_i$ -LPA by 3.6, 19.4 and 9.2 times, respectively, and achieving the highest concentration of LPA reported so far. With the rapid development of new enzyme structures and molecular modeling dynamics software, more and more valuable enzymes affected by both substrate and product inhibitions could be improved by modifying their distinct substrate and product delivery processes using a similar strategy, making them attractive biocatalysts.

## Results and discussion

### Construction and catalytic performance of the *Sp*LCD biocatalyst

The *pipA* gene from *S. pristinaespiralis* ATCC 25486 encoding *Sp*LCD was amplified and cloned into the vector pET-28a to generate the recombinant plasmid pET-28a-*pipA*, followed by transformation into *E. coli* BL21(DE3) cells and overexpression. However, SDS-PAGE revealed the formation of a large amount of inclusion bodies. To obtain higher *Sp*LCD expression levels, molecular chaperones were used to assist in the folding of the *Sp*LCD protein. Four recombinant *E. coli* strains were constructed by introducing different molecular chaperones including pG-KJE8, pGro7, pKJE7 and pTf16 (information on plasmids is summarized in ESI† Table S1), into the *E. coli*-*Sp*LCD biocatalyst. Positive effects were observed with the co-expression of molecular chaperones pG-KJE8, pGro7 and pKJE7 (Fig. S1†). However, pTf16 did not appear to be suitable for co-expression with *Sp*LCD, because the enzyme was barely

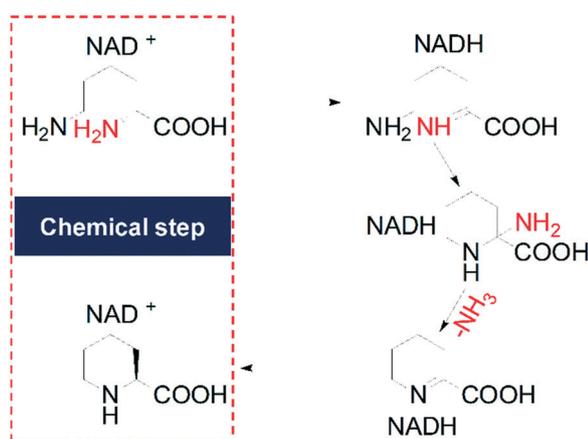


Fig. 1 Reaction process of lysine cyclodeaminase.

generated. In contrast, SDS-PAGE revealed that the amount of intracellular soluble *Sp*LCD increased by approximately three fold when using pGro7, indicating that pGro7 could significantly promote correct folding of *Sp*LCD. Using the purified enzyme, the enzymatic kinetics of *Sp*LCD was determined and is summarized in Table 1. As expected, both substrate LYS and product LPA inhibit *Sp*LCD severely with a relatively low  $K_i$  value of 1.0 mM and 1.7 mM, respectively.

## Two ligand delivery processes detected by MD simulations of *Sp*LCD

The active site of many enzymes is buried inside the protein core and is connected to the bulk solvent through transport tunnels.<sup>16</sup> The anatomy, physiochemical properties and dynamics of these tunnels have a large influence on enzymatic catalysis by determining the exchange rate of substrates or products between the active sites and solvent environment.<sup>17</sup> The importance of transport processes for enzymatic catalysis has prompted many engineering studies to focus on protein tunnels and overcome many natural limitations.<sup>18</sup> In our previous studies of *Sp*LCD crystal structures,<sup>19</sup> a buried active site and a potential substrate delivery tunnel were identified, indicating that the substrate delivery or product release steps might be a rate-limiting bottleneck.

To overcome substrate and product inhibitions, as well as to improve the efficiency of *Sp*LCD, the conformational dynamics of the substrate delivery and product release processes was investigated *via* theoretical computational calculations based on *Sp*LCD crystal structures. Firstly, MD simulations of the LCD-NAD<sup>+</sup>-lysine (*Sp*LCD-LYS) system were conducted for 100 ns to explore the substrate delivery mechanism (Fig. S2<sup>†</sup>). MD simulations of *Sp*LCD-LYS showed the substrate dynamic movement from the initial binding state to a temporary steady state. Although substrate lysine was not actually located in the appropriate reaction position at the end of the simulation, our calculations revealed the tendency of two substrate delivery states. Based on the MD trajectory, these two substrate delivery tunnels were visualized using the CAVER program to obtain an overall perspective and bottlenecks.<sup>20</sup> First, consistent with our previous analysis of the *Sp*LCD crystal structures, a relatively short and uniform substrate delivery tunnel 1 ('state 1') was revealed in the MD simulations. As shown in Fig. 2a and c, the entrance of substrate delivery tunnel 1 existed in the vicinity of Glu264 and Asp236 (Fig. 2c), and no obvious bottleneck was observed as the radius of the tunnel seemed to be similar (radius of 2.0–2.5 Å). The MD simulations indicated that Asp236 forms

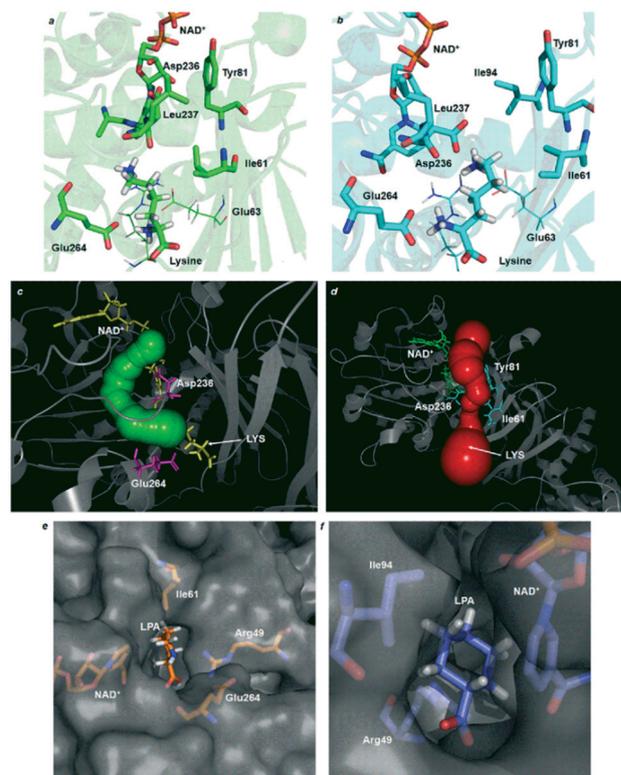


Fig. 2 Distinct substrate delivery tunnels and product release exits of *Sp*LCD analysed by MD simulation. a. Representative conformations of substrate delivery state 1. b. Representative conformations of substrate delivery state 2. The residues, cofactor and ligand are presented in green in substrate delivery state 1 and presented in cyan in substrate delivery state 2. c. Visualization of substrate tunnel 1 *via* CAVER. The tunnel is presented in green, residues are presented in pink and NAD<sup>+</sup> and LYS are presented in yellow. d. Visualization of substrate tunnel 2 *via* CAVER. The tunnel is presented in red, residues are presented in cyan and NAD<sup>+</sup> and LYS are presented in green. e. Representative conformations of product release exit 1. f. Representative conformation of product release exit 2.

a stable H-bond with cofactor NAD<sup>+</sup>, whereas the H-bond network of the substrate could only be formed between Glu63 and Glu264 (Fig. S3<sup>†</sup>). As a result, lysine was pulled down by Glu264 and is dragged along the gap between Glu264 and Leu237 to the active site. On the other hand, owing to the possible alternative location of the pyridine ring of NAD<sup>+</sup>, and featuring the rotation of Asp236, another relatively long and irregular substrate delivery tunnel 2 was revealed (Fig. 2b and d). The entrance of substrate delivery tunnel 2 existed at the opposite site of Glu264 and a bottleneck could be observed between Asp236 and Ile61 (Fig. 2d) (radius of 1.6

Table 1 Kinetic parameters of WT-*Sp*LCD and its variants

Enzymes	$K_m$ (mM)	$k_{cat}$ (min <sup>-1</sup> )	$k_{cat}/K_m$ (min <sup>-1</sup> mM <sup>-1</sup> )	$K_i$ -LYS (mM)	$K_i$ -LPA (mM)
WT- <i>Sp</i> LCD	1.7 ± 0.4	0.61 ± 0.04	0.36 ± 0.03	1.0 ± 0.3	1.7 ± 0.2
V61- <i>Sp</i> LCD	4.9 ± 0.5	6.79 ± 0.28	1.38 ± 0.11	14.2 ± 2.1	4.9 ± 0.8
V94- <i>Sp</i> LCD	3.5 ± 0.4	4.74 ± 0.65	1.35 ± 0.42	8.1 ± 1.4	9.5 ± 1.2
V61-V94- <i>Sp</i> LCD	9.6 ± 0.2	12.5 ± 2.11	1.30 ± 0.31	19.4 ± 3.7	15.7 ± 2.6

Å). The MD simulations of *Sp*LCD-LYS suggested that the carboxylate moiety of Asp236 in substrate delivery state 2 rotated and formed a stable H-bond with lysine, thereby helping to anchor the substrate among Asp236, Glu63, and Glu264. Once lysine was captured, delivery occurred through the synergistic action of these three residues (Fig. S4†).

In addition, the real-time distance between the C4 position of the NAD<sup>+</sup> ribose ring and  $\alpha$ -NH<sub>2</sub> of substrate lysine ( $d_{C-N}$ ) was measured during the MD simulation. Although the substrate lysine is not actually located in the reaction position at the end of the simulation but in a temporary stable situation,  $d_{C-N}$  could be interpreted as the probability and tendency of the occurrence of substrate delivery processes from the protein surface to the active site (near the NAD<sup>+</sup>).<sup>21</sup> As shown in Fig. 3, in substrate delivery state 1, the distance sharply shortened from 10 ns and subsequently stabilized at about 13 Å, indicating that the temporary steady state came earlier and directly. In contrast, in substrate delivery state 2, the distance started to shorten after 40 ns with subsequent fluctuation for several times until finally reaching stability at about 12.5 Å, indicating that the temporary steady state came with more difficulty and tortuously. However, the distances in both of the substrate delivery processes decreased and remained at a similar level after 80 ns, revealed that both substrate delivery states were stable and lysine has the tendency to be delivered to the active site through either way.

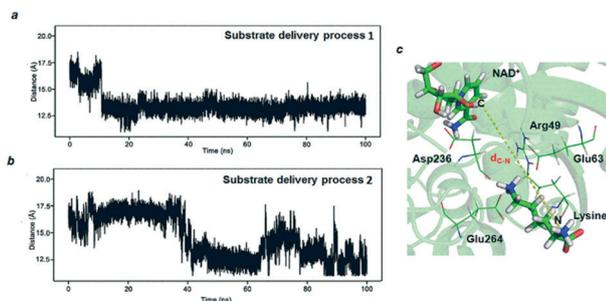
Furthermore, MD simulations of the LCD-NAD<sup>+</sup>-LPA (*Sp*LCD-LPA) system were also conducted for 100 ns to explore the product release exits (Fig. S2†). As a result, two corresponding product exits were observed which are functioned in pairs with the corresponding substrate delivery tunnels. In substrate delivery state 1, owing to the different locations of Glu264, a cavity (exit 1) on the LCD surface occurred near the original position of Glu264 (Fig. 2e) and the product may be released directly from the cavity among Glu264, Ile61, and NAD<sup>+</sup> following completion of the enzymatic reaction. The second exit was on the other side of *Sp*LCD (Fig. 2f). In substrate delivery state 2, owing to the specific location of NAD<sup>+</sup>, a cavity (exit 2) of considerable size occurred around Ile94 and the pyridine ring of NAD<sup>+</sup>, and the product L-pipecolic acid might be released from exit 2 directly. Two non-polar

amino acid residues, Ile61 in product release exit 1 and Ile94 in product release exit 2, appeared to influence the corresponding product release process *via* their side chains. Although the MD simulations suggested that two exits may exist, it appeared that only one exit remained active during the enzymatic process at any particular time. In other words, when exit 1 was open in substrate delivery state 1, exit 2 was blocked by the nicotinamide nucleoside moiety of NAD<sup>+</sup> (Fig. S5a†). Similarly, when exit 2 was open in substrate delivery state 2, the location of Glu264 blocked the product release in exit 1 (Fig. S5b†).

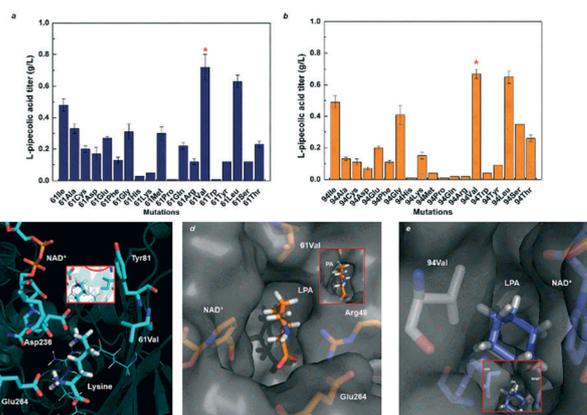
### Mutagenesis of Ile61 and Ile94 of *Sp*LCD

As demonstrated *via* MD simulations, several key residues, including Asp236, Glu264, Glu63, Ile61, and Ile94, were involved in the substrate delivery and product release sites. The highly conserved residues in the LCDs, Asp236, Glu264, and Glu63 formed significant H-bonds with lysine in the substrate delivery tunnels or with the intermediate in the active site, and mutations on these positions may result in a serious loss of catalytic activity, due to the detrimental effects on the structural integrity around the substrate of the product-binding site.<sup>19</sup> Therefore, the other two non-polar residues, Ile61 and Ile94, which both appeared to be the bottlenecks, were saturatedly mutated to further explore their key roles in the substrate and product delivery processes.

As shown in Fig. 4a and b, among all the constructed variants, the LPA titer of Val61-*Sp*LCD and Val94-*Sp*LCD was 1.5 and 1.4 times higher than that of WT-*Sp*LCD, respectively. Moreover, the Val61-Val94-*Sp*LCD combination variant exhibited a 1.8-fold higher LPA titer when compared with WT-*Sp*LCD, reaching almost 100% yield (data not shown). In addition, the kinetic data of Val61-*Sp*LCD, Val94-*Sp*LCD, and Val61-Val94-*Sp*LCD are summarized in Table 1, which showed



**Fig. 3** Distance between the C4 position of the NAD<sup>+</sup> ribose ring and  $\alpha$ -NH<sub>2</sub> of substrate lysine during the MD simulation. a. Distance change in substrate delivery state 1. b. Distance change in substrate delivery state 2. c. Schematic diagram of the measured distance.



**Fig. 4** Saturation mutation studies of 61 and 94 positions and partial structures of Val61-*Sp*LCD and Val94-*Sp*LCD. a. Saturation mutation study of Ile61. b. Saturation mutation study of Ile94. c. Structure of Val61 in substrate delivery process 2. d. Structure of Val61 in product release exit 1. e. Structure of Val94 in product release exit 2. The bottlenecks of the substrate delivery tunnel structures of wild-type *Sp*LCD are presented in red panes in c–e.

that the enhanced LPA production ability was primarily due to the higher  $k_{\text{cat}}$  values (11-, 7-, and 20-fold higher when compared with WT-*SpLCD*, respectively) and  $K_i$  values (14/3-, 8/5-, 19/9-fold higher when compared with WT-*SpLCD*). These kinetics results revealed that the substrate and product inhibitions of these valine variants were significantly reduced.

Structure analysis of Val61-*SpLCD* and Val94-*SpLCD* indicated that the reduction of the substrate and product inhibitions was probably due to the significant decrease of steric hindrance of the flaps of product release exit 1, exit 2 and particularly substrate delivery tunnel 2, thus facilitating the substrate and product delivery processes (Fig. 4c–e). Additional simulation and visualization of substrate delivery process 2 of the Val61-*SpLCD* variant using CAVERdock confirmed this concept. As shown in Fig. S6,<sup>†</sup> the bottleneck (Asp236-Ile61) of substrate delivery tunnel 2 of the Val61-*SpLCD* variant broadened critically in comparison with that of wild-type *SpLCD* owing to the substitution of Val in the 61 position, which might enhance the catalytic performance significantly (radius from 1.6 Å to 2.3 Å).

According to our previous studies, the whole-cell LCD biocatalyst (200 OD cell density) will be fully inactivated with more than 25 g L<sup>-1</sup> substrate L-lysine or 15 g L<sup>-1</sup> product LPA.<sup>15a</sup> To further ascertain the improvement in the catalytic activity of Val61-Val94-*SpLCD*, the effect of high substrate concentration and product concentration on the recombinant whole-cell biocatalyst *E. coli* Val61-Val94-*SpLCD* was investigated (Fig. 5a and b). The results showed that a consistent LPA molar yield of about 95% was achieved by *E. coli*-Val61-Val94-*SpLCD* with a lysine concentration ranging from 1 to 20 g L<sup>-1</sup>. In addition, the bioconversion of L-lysine was not significantly affected by the initial exogenous addition of LPA ranging from 5 to 20 g L<sup>-1</sup>, and only a slight decrease in the

LPA yield was observed in *E. coli* Val61-Val94-*SpLCD* supplemented with more than 20 g L<sup>-1</sup> exogenous LPA. These results revealed that the reduction of substrate/product inhibition using the LCD biocatalyst might be an alternative promising approach in future LPA production.

### LPA production using Val61-Val94-*SpLCD*

LPA production was then performed with Val61-Val94-*SpLCD* as the whole-cell biocatalyst. First, the substrate concentration was optimized. The production of LPA was conducted in 1 L bioreactors containing a 500 mL bioconversion system comprising 10–70 g L<sup>-1</sup> lysine, the whole-cell biocatalyst (with a final cell density of OD<sub>600</sub> = 50), 0.1% w/v Triton X-100, and 2.5 mM FeSO<sub>4</sub>. As shown in Fig. 5c, with an increasing initial substrate concentration from 10 to 50 g L<sup>-1</sup>, the LPA titer increased with a molar yield of more than 95%. A maximum LPA titer of 43.2 g L<sup>-1</sup> and a space-time yield of 1.74 g L<sup>-1</sup> h<sup>-1</sup> were achieved with an initial substrate concentration of 50 g L<sup>-1</sup>. With a further increase in the initial substrate concentration to 60 and 70 g L<sup>-1</sup>, the molar yield of the product considerably decreased to 84.9% and 79.7%, respectively. Thus, 50 g L<sup>-1</sup> initial substrate concentration was selected for further investigation.

During the batch process based on the optimized substrate concentration, when the initial amount of lysine was totally consumed, additional 50 g L<sup>-1</sup> lysine was supplemented to the reaction system two times (Fig. 5d). In the first cycle, the catalytic efficiency was the highest, the space-time yield at 0–11 h reached 2.9 g L<sup>-1</sup> h<sup>-1</sup>, and a LPA titer of 41.0 g L<sup>-1</sup> with a molar yield of 93.1% was achieved. In the second cycle, the catalytic efficiency decreased gradually, and a LPA titer of 22.6 g L<sup>-1</sup> with a molar yield of 51.2% was obtained. In the final reaction cycle, the substrate consumption rate and LPA yield significantly decreased, partially due to cell lysis. Nevertheless, 73.4 g L<sup>-1</sup> LPA was accumulated, with a total space-time yield of 0.83 g L<sup>-1</sup> h<sup>-1</sup>. Thus, the method developed in the present study achieved significant improvement in both the LPA titer and catalytic efficiency, which are the highest reported so far, making it a promising alternative for potential sustainable and environment-friendly industrial applications.

## Experimental

### General

Previously-established protocols were employed for the cultivation of recombinant *E. coli*, preparation of the whole-cell biocatalyst, and construction, expression, purification, kinetic and enzymatic assays of *SpLCD*.<sup>14,15a,19</sup> To be specific, the *pipA* gene encoding lysine cyclodeaminase from *Streptomyces pristinaespiralis* ATCC 25486 is artificially synthesized by Genescript. co, Nanjing Jiangsu. *Escherichia coli* strain BL21(DE3) cells harboring the pET28b-*pipA* plasmid were grown at 37 °C. N-Terminally His6-tagged LCD was over-expressed following induction with 100 μM isopropyl-β-D-1-thiogalactopyranoside overnight at 25 °C. After cell lysis, LCD was purified *via*

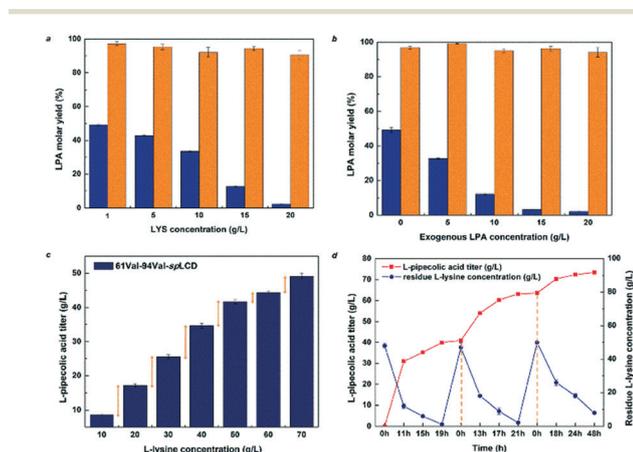


Fig. 5 Bio-production of L-pipecolic acid using the Val61-Val94-*SpLCD* whole-cell biocatalyst. a. The effect of high substrate LYS concentration on the recombinant whole-cell biocatalysts *E. coli SpLCD* and *E. coli* Val61-Val94-*SpLCD*. b. The effect of high exogenous LPA concentration on the recombinant whole-cell biocatalysts *E. coli SpLCD* and *E. coli* Val61-Val94-*SpLCD*. The blue bar represents *E. coli SpLCD* and the orange bar represents *E. coli* Val61-Val94-*SpLCD*. c. The assay of the optimal substrate loading concentration. d. The batch bio-production process of L-pipecolic acid.

immobilized Ni<sup>2+</sup> affinity chromatography (HisTrap HP; GE Healthcare), followed by both anion exchange chromatography (Resource Q; GE Healthcare) and size-exclusion chromatography (Superdex 200 10/300 GL; GE Healthcare). Purified LCD (10 mM HEPES-NaOH buffer, pH 7.0) was concentrated to 10 mg ml<sup>-1</sup> using an Amicon Ultra-430 K filter (Millipore) and stored frozen at -80 °C prior to use. All chromatography protocols were performed using an ÄKTA-explorer apparatus (GE Healthcare). Protein concentrations were determined using a Thermo Nanodrop 2000c spectrophotometer, and the purity was monitored by SDS-PAGE.

A representative activity assay of WT and variants was performed using the above-mentioned purified LCD. To be specific, the reaction system (1 ml) consists of 10 mM HEPES-NaOH buffer (pH 7.0), 1 mg ml<sup>-1</sup> purified enzymes and 2 g L<sup>-1</sup> substrate; after incubation for 30 min, the reactions were stopped by liquid nitrogen, and the reaction system was frozen and stored at -80 °C for HPLC determination. The reactions were run in triplicate and initiated by the addition of the substrate.

The kinetics assays of WT and variants were performed using the above-mentioned purified LCD. To be specific, the purified enzymes were diluted to 0.1 mg ml<sup>-1</sup> and incubated in a final volume of 100 µl with 50 µM NAD<sup>+</sup> (pre-incubated for 5 min) and a substrate concentration ranging from 10–1000 µM. The reactions were run in triplicate and initiated by the addition of the substrate. At time points of 15, 30, 45, and 60 min, the corresponding reactions were stopped by liquid nitrogen, and the reaction system was frozen and stored at -80 °C for HPLC determination. The empirical model equation  $v = V_{\max}/(1 + K_m/S + S/K_i)$  was used to process the obtained data.

### Biotransformation

A standard whole-cell biocatalyst property assay was conducted in 2 mL tubes (total reaction volume of 1 mL) containing recombinant *E. coli* cells as the whole-cell biocatalyst, L-lysine as the substrate, and HEPES-NaOH buffer (pH = 7.0) as the solvent. The cell density of the reaction system was controlled at OD<sub>600</sub> = 10 ( $g_{dcw} = 5 \text{ g L}^{-1}$  or concentration given in the Results and discussion section; dcw, dry cell weight), and L-lysine was added at a concentration of 1 g L<sup>-1</sup>.

Biotransformation of L-lysine for efficient LPA production was conducted in 1 L bioreactors (total reaction volume of 500 mL) containing recombinant *E. coli* cells as the whole-cell biocatalyst, L-lysine as the substrate, HEPES-NaOH buffer as the solvent, 0.1% (w/v) Triton X-100 as the surfactant for better transport of the substrate, and 2.5 mM FeSO<sub>4</sub> as the activator for SpLCD. The cell density of the reaction system was controlled at OD<sub>600</sub> = 50, and the initial L-lysine concentration ranges are given in the Results and discussion section.

During the reaction, the product concentration was monitored by using high-performance liquid chromatography (HPLC). All the experiments were conducted in triplicate.

### Analytical methods

The concentrations of L-lysine were analyzed by using an SBA-40C biosensor analyzer (Shandong Province Academy of Sciences, China).<sup>22</sup>

The exact LPA and L-lysine concentrations were determined by using HPLC. LPA and L-lysine were first labelled with phenyl isothiocyanate (PITC), and then analyzed by using a high-performance liquid chromatograph (1290, Agilent Technologies, USA) equipped with a C18 column (5 µm, 250 mm × 4.6 mm, Grace, USA) at room temperature. Mobile phase A consisted of 7% v/v acetonitrile in 0.1 M sodium acetate aqueous solution, while mobile phase B consisted of 80% v/v acetonitrile in water. The PITC derivatives of LPA and L-lysine were separated with a gradient of 97:3–30:70 v/v of A:B for 50 min and detected using absorbance at 254 nm.

### MD simulations

Two complex systems, including SpLCD-NAD<sup>+</sup>-lysine (SpLCD-LYS) and SpLCD-NAD<sup>+</sup>-pipercolic acid (SpLCD-LPA), were constructed based on the protein complex structures of SpLCD-LYS and SpLCD-LPA (PDB code: 5GZL and 5GZM, respectively).<sup>19</sup> To be specific, in each complex system, Repeat 1 and Repeat 2 were performed based on the different locations of LYS or LPA in different monomers of the corresponding crystal structures. Parameter preparation for the substrate (L-lysine) and product (LPA) was performed by using the Antechamber package.<sup>23</sup> First, optimization was accomplished using the Gaussian 09 program at the level of HF/6-31G(d)<sup>24</sup> followed by calculation of the electrostatic surface potential (ESP) charge. Subsequently, restrained electrostatic potential (RESP)<sup>25</sup> charge fitting was applied. Finally, using the AMBER program suite (with ff03.r1 force field),<sup>26</sup> MD simulations were conducted with the prepared structures of the two systems. The protonation state of each residue was determined at pH 7.0 using the PROPKA program on the PDB2PQR website.<sup>27</sup> The proteins were solvated in a cubic box of TIP3P water molecules, with a water thickness extending at least 10 Å away from the protein surface, and sodium ions were added to the system as counter ions to create a neutral simulation system.

In an attempt to preclude instability that might occur during MD simulations, the solvated system was subjected to 10 000 steps of minimization, and was changed from the steepest descent algorithm to a conjugate gradient algorithm after 1000 cycles. The system was then gradually heated from 0 K to 300 K in 50 ps steps controlled by Langevin dynamics with a collision frequency of 2 ps<sup>-1</sup>. Subsequently, the system was switched to constant pressure and temperature (NPT) and equilibrated for 50 ps to facilitate adjustment of the system to the appropriate density. Finally, production simulations were performed without any restraint under NPT conditions, followed by a 100 ns MD simulation. The Particle Mesh Ewald (PME) method<sup>28</sup> was used to calculate long-range electrostatic interactions. The lengths of the bonds involving

hydrogen atoms were fixed with the SHAKE algorithm.<sup>29</sup> During the simulations, an integration time step of 2 fs was adopted, and structural snapshots were flushed every 500 steps (1 ps). The non-bonded cut-off was set to 10.0 Å. This protocol was applied to all the simulation systems, and all the MD simulations were performed using the parallel processing version of PMEMD.cuda in the AMBER12 suite.<sup>30</sup> The structure figures were prepared using Pymol.<sup>31</sup> The tunnel visualization was computed and prepared using CAVER Analyst 1.0 and CAVERdock 1.0.<sup>32</sup>

### Mutagenesis

Point mutants and saturation mutagenesis of SpLCD were generated using the whole-plasmid PCR and *DpnI* digestion methods. The identities of the successfully-generated constructs were verified by DNA sequencing. The recombinant *E. coli*-SpLCD whole-cell biocatalyst variants were prepared using the above-mentioned procedure.

### Conclusions

This study showed a practical implementation for reducing substrate and product inhibitions simultaneously *via* the identification of the multiple substrate and product delivery tunnels, confirmation of key residues and redesigning of rate-limiting bottlenecks in the enzymatic processes. Using whole-cell biocatalyst *E. coli* co-expressing the molecular chaperones and Val61-Val94-SpLCD variant, efficient bio-production of the chiral drug intermediate LPA was performed. The highest LPA accumulation (73.4 g L<sup>-1</sup>) reported so far was achieved. The structure–function relationship, structure-based engineering strategy of lysine cyclo-deaminase, and preparative biosynthesis of L-pipecolic acid reported here allow a potential industrial application in the future. The enzyme catalytic efficiency modification, the cell and protein stability, and the cofactor binding affinity are engineering targets to improve the L-pipecolic acid production yield in long-time reaction cycles.

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 21390200), the National Key Research and Development Program (2016YFA0204300), and the Synergetic Innovation Center for Advanced Materials. We sincerely thank Prof. Yilei Zhao and Dr. Ting Shi for the kindly help in MD simulation experiments.

### Notes and references

- (a) F. Cedrone, A. Menez and E. Quemeneur, *Curr. Opin. Struct. Biol.*, 2000, **10**, 405–410; (b) D. N. Bolon, C. A. Voigt and S. L. Mayo, *Curr. Opin. Chem. Biol.*, 2002, **6**, 125–129; (c) R. A. Chica, N. Doucet and J. N. Pelletier, *Curr. Opin. Biotechnol.*, 2005, **16**, 378–384.
- S. S. Han, H. H. Kyeong, J. M. Choi, Y. K. Sohn, J. H. Lee and H. S. Kim, *ACS Catal.*, 2016, **6**, 8440–8445.
- R. Alves and M. A. Savageau, *Biophys. J.*, 2000, **79**, 2290–2304.
- (a) X. Hu, S. Robin, S. O'Connell, G. Walsh and J. G. Wall, *Appl. Microbiol. Biotechnol.*, 2010, **87**, 1773–1782; (b) Y. Kai, T. Kashiwagi, K. Ishikawa, M. K. Ziyatdinov, E. I. Redkina, M. Y. Kiriukhin, M. M. Gusyatiner, S. Kobayashi, H. Takagi and E. Suzuki, *Protein Eng., Des. Sel.*, 2006, **19**, 163–167; (c) R. L. Silveira and M. S. Skaf, *J. Phys. Chem. B*, 2015, **119**, 9295–9303; (d) M. E. Atreya, K. L. Strobel and D. S. Clark, *Biotechnol. Bioeng.*, 2016, **113**, 330–338; (e) R. L. Silveira and M. S. Skaf, *J. Phys. Chem. B*, 2015, **119**, 9295–9303; (f) M. E. Atreya, K. L. Strobel and D. S. Clark, *Biotechnol. Bioeng.*, 2016, **113**, 330–338.
- (a) M. Klvana, M. Pavlova, T. Koudelakova, R. Chaloupkova, P. Dvorak, Z. Prokop, A. Stsiapanava, M. Kutý, I. K. Smatanova, J. Dohnalek, P. Kulhanek, R. C. Wade and J. Damborsky, *J. Mol. Biol.*, 2009, **392**, 1339–1356; (b) Z. Wen, J. Baudry, M. R. Berenbaum and M. A. Schuler, *Protein Eng., Des. Sel.*, 2005, **18**, 191–199; (c) R. Sprangers, A. Gribun, P. A. Hwang, W. A. Houry and L. E. Kay, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 16678–16683.
- (a) J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern and T. Alber, *Nature*, 2009, **462**, 669–673; (b) D. Oyen, R. B. Fenwick, R. L. Stanfield, H. J. Dyson and P. E. Wright, *J. Am. Chem. Soc.*, 2015, **137**, 9459–9468; (c) X. D. Kong, S. G. Yuan, L. Li, S. Chen, J. H. Xu and J. H. Zhou, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15717–15722.
- (a) A. G. Palmer, *J. Magn. Reson.*, 2014, **241**, 3–17; (b) G. Jimenez-Oses, S. Osuna, X. Gao, M. R. Sawaya, L. Gilson, S. J. Collier, G. W. Huisman, T. O. Yeates, Y. Tang and K. N. Houk, *Nat. Chem. Biol.*, 2014, **10**, 431–436; (c) D. D. Boehr, H. J. Dyson and P. E. Wright, *Chem. Rev.*, 2006, **106**, 3055–3079; (d) D. D. Boehr, R. Nussinov and P. E. Wright, *Nat. Chem. Biol.*, 2009, **5**, 789–796; (e) N. M. Goodey and S. J. Benkovic, *Nat. Chem. Biol.*, 2008, **4**, 474–482.
- (a) E. Kim, S. Lee, A. Jeon, J. M. Choi, H. S. Lee, S. Hohng and H. S. Kim, *Nat. Chem. Biol.*, 2013, **9**, 313; (b) M. H. Seo, J. Park, E. Kim, S. Hohng and H. S. Kim, *Nat. Commun.*, 2014, **5**, 3724; (c) N. Basse, J. L. Kaar, G. Settanni, A. C. Joerger, T. J. Rutherford and A. R. Fersht, *Chem. Biol.*, 2010, **17**, 46–56; (d) D. Oyen, R. B. Fenwick, R. L. Stanfield, H. J. Dyson and P. E. Wright, *J. Am. Chem. Soc.*, 2015, **137**, 9459–9468; (e) Y. Hu and H. Liu, *J. Phys. Chem. A*, 2014, **118**, 9272–9279.
- (a) M. K. Tiwari, R. Singh, R. K. Singh, I. W. Kim and J. K. Lee, *Comput. Struct. Biotechnol. J.*, 2012, **2**, e201209002; (b) S. E. Nichols, R. V. Swift and R. E. Amaro, *Curr. Top. Med. Chem.*, 2012, **12**, 2002–2012.
- (a) G. Parigi, N. Rezaei-Ghaleh, A. Giachetti, S. Becker, C. Fernandez, M. Blackledge, C. Griesinger, M. Zweckstetter and C. Luchinat, *J. Am. Chem. Soc.*, 2014, **136**, 16201–16209; (b) J. G. Chandraseelan, S. M. Oliveira, A. Hakkinen, H.

- Tran, I. Potapov, A. Sala, M. Kandhavelu and A. S. Ribeiro, *Mol. BioSyst.*, 2013, 9, 3117–3123.
- 11 (a) S. N. Sehgal, H. Baker and C. Vezina, *J. Antibiot.*, 1975, 28, 727–732; (b) T. Kino, H. Hatanaka, M. Hashimoto, M. Nishiyama, T. Goto, M. Okuhara, M. Kohsaka, H. Aoki and H. Imanaka, *J. Antibiot.*, 1987, 40, 1249–1255.
- 12 B. Adger, U. Dyer and G. Hutton, *Tetrahedron Lett.*, 1996, 37, 6399–6402.
- 13 F. Couty, *Amino Acids*, 1999, 16, 297–320.
- 14 G. J. Gatto, M. T. Boyne, N. L. Kelleher and C. T. Walsh, *J. Am. Chem. Soc.*, 2006, 128, 3838–3847.
- 15 (a) H. X. Ying, J. Wang, Z. Wang, J. Feng, K. Q. Chen, Y. Li and P. K. Ouyang, *J. Mol. Catal. B: Enzym.*, 2015, 117, 75–80; (b) H. X. Ying, J. Wang, X. Wang and K. Q. Chen, *Microb. Cell Fact.*, 2017, 16, 52.
- 16 (a) Z. Prokop, A. Gora, J. Brezovsky, R. Chaloupkova, V. Stepankova and J. Damborsky, ed. S. Lutz and U. T. Bornscheuer, *Protein engineering hand book*, Wiley-VCH, Weinheim, Germany, 2012, vol. 3, pp. 421–464; (b) L. J. Kingsley and M. A. Lill, *Proteins: Struct., Funct., Genet.*, 2015, 83, 599–611.
- 17 (a) M. Pavlova, M. Klvana, Z. Prokop, R. Chaloupkova, P. Banas, M. Otyepka, R. C. Wade, M. Tsuda, Y. Nagata and J. Damborsky, *Nat. Chem. Biol.*, 2009, 5, 727–733; (b) L. Biedermannova, Z. Prokop, A. Gora, E. Chovancova, M. Kovacs, J. Damborsky and R. C. Wade, *J. Biol. Chem.*, 2012, 287, 29062–29074.
- 18 (a) M. C. C. J. C. Ebert, S. L. Durr, A. A. Houle, G. Lamoureux and J. N. Pelletier, *ACS Catal.*, 2016, 6, 7426–7437; (b) J. Brezovsky, P. Babkova, O. Degtjarik, A. Fortova, A. Gora, I. Iermak, P. Rezacova, P. Dvorak, I. Smatanova, Z. Prokop, R. Chaloupkova and J. Damborsky, *ACS Catal.*, 2016, 6, 7597–7610.
- 19 H. X. Ying, J. Wang, T. Shi, Y. L. Zhao, X. Wang, P. K. Ouyang and K. Q. Chen, *Biochem. Biophys. Res. Commun.*, 2018, 495, 306–311.
- 20 B. Kozlíková, E. Šebestová, V. Šustr, J. Brezovský, O. Strnad, L. Daniel, D. Bednář, A. Pavelka, M. Maňák, M. Bezděka, P. Beneš, M. Kotry, A. W. Gora, J. Damborský and J. Sochor, *J. Bioinform.*, 2014, 30, 18.
- 21 T. Shi, L. X. Liu, W. T. Tao, S. G. Luo, S. B. Fan, X. L. Wang, L. Q. Bai and Y. L. Zhao, *ACS Catal.*, 2018, 8, 4323–4332.
- 22 H. X. Ying, X. He, Z. Wang, P. P. Yuan, K. Q. Chen and P. K. Ouyang, *Appl. Biochem. Biotechnol.*, 2014, 172, 3835–3843.
- 23 J. M. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, 25, 247–260.
- 24 (a) R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, 54, 724–728; (b) W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, 56, 2257–2261; (c) H. Stoll, G. Wagenblast and H. Preuss, *J. Am. Chem. Soc.*, 1978, 100, 3787–3794.
- 25 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gaudl, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1996, 118, 2309–2309.
- 26 (a) A. L. De Fonseca, R. R. Nunes, V. M. L. Braga, M. Comar Jr, R. J. Alves, F. R. Varotti and A. G. Tatanto, *J. Mol. Graphics Modell.*, 2016, 66, 174–186; (b) Y. Jiang, H. Y. Zhang, W. Feng and T. W. Tan, *J. Chem. Inf. Model.*, 2015, 55, 2575–2586; (c) G. A. Pantelopulos, S. Mukherjee and V. A. Voelz, *Proteins*, 2015, 83, 1665–1676.
- 27 (a) H. Li, A. D. Robertson and J. H. Jensen, *Proteins*, 2005, 61, 704–721; (b) D. C. Bas, D. M. Rogers and J. H. Jensen, *Proteins*, 2008, 73, 765–783.
- 28 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, 98, 10089–10092.
- 29 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, 23, 327–341.
- 30 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, F. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, 117, 5179–5197.
- 31 *The Pymol Molecular Graphics System*, Version 2.0, Schrödinger, LLC.
- 32 B. Kozlíková, E. Šebestová, V. Šustr, J. Brezovský, O. Strnad, L. Daniel, D. Bednář, A. Pavelka, M. Maňák, M. Bezděka, P. Beneš, M. Kotry, A. W. Gora, J. Damborský and J. Sochor, *Bioinformatics*, 2014, 30, 18.