

SCIENTIFIC REPORTS

OPEN

Engineering catechol 1, 2-dioxygenase by design for improving the performance of the *cis*, *cis*-muconic acid synthetic pathway in *Escherichia coli*

Received: 23 February 2015

Accepted: 27 July 2015

Published: 26 August 2015

Li Han^{1,*}, Pi Liu^{2,*}, Jixue Sun², Yuanqing Wu¹, Yuanyuan Zhang¹, Wujiu Chen¹, Jianping Lin^{1,2}, Qinhong Wang¹ & Yanhe Ma¹

Regulating and ameliorating enzyme expression and activity greatly affects the performance of a given synthetic pathway. In this study, a new synthetic pathway for *cis*, *cis*-muconic acid (*ccMA*) production was reconstructed without exogenous induction by regulating the constitutive expression of the important enzyme catechol 1,2-dioxygenase (*CatA*). Next, new *CatA*s with significantly improved activities were developed to enhance *ccMA* production using structure-assisted protein design. Nine mutations were designed, simulated and constructed based on the analysis of the *CatA* crystal structure. These results showed that mutations at Gly72, Leu73 and/or Pro76 in *CatA* could improve enzyme activity, and the activity of the most effective mutant was 10-fold greater than that of the wild-type *CatA* from *Acinetobacter* sp. ADP1. The most productive synthetic pathway with a mutated *CatA* increased the titer of *ccMA* by more than 25%. Molecular dynamic simulation results showed that enlarging the entrance of the substrate-binding pocket in the mutants contributed to their increased enzyme activities and thus improved the performance of the synthetic pathway.

The microbial production of platform chemicals, which can be either directly used or further processed for the production of industrial-scale and high value-added products in the chemical industry, has recently gained tremendous interest and attention, especially with recent rapid advances in synthetic biology and metabolic engineering^{1–3}. Adipic Acid, six-carbon dicarboxylic acid, is linked almost 90% to nylon production, and can be further processed into fibers for the production of carpeting, automobile tire cords and clothing⁴. Most commercial adipic acid is produced from cyclohexane through a two-step oxidation process. This inferior process exhibits a poor yield, and the chemicals used are environmentally harmful⁵. Alternatively, adipic acid can be produced from *cis*, *cis*-muconic acid (*ccMA*) by chemical hydrogenation. The production of *ccMA* from renewable resources by fermentation, followed by hydrogenation to adipic acid, would be desirable because such a route to adipic acid would be more environmentally friendly than the traditional petrochemical route⁶. Moreover, *ccMA* possesses reactive carboxylate groups and a conjugated double bond configuration that can be useful as a raw material for new functional resins, pharmaceuticals and agrochemicals⁷.

*Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China. ²State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin 300071, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.L. (email: jianpinglin@nankai.edu.cn) or Q.W. (email: wang_qh@tib.cas.cn)

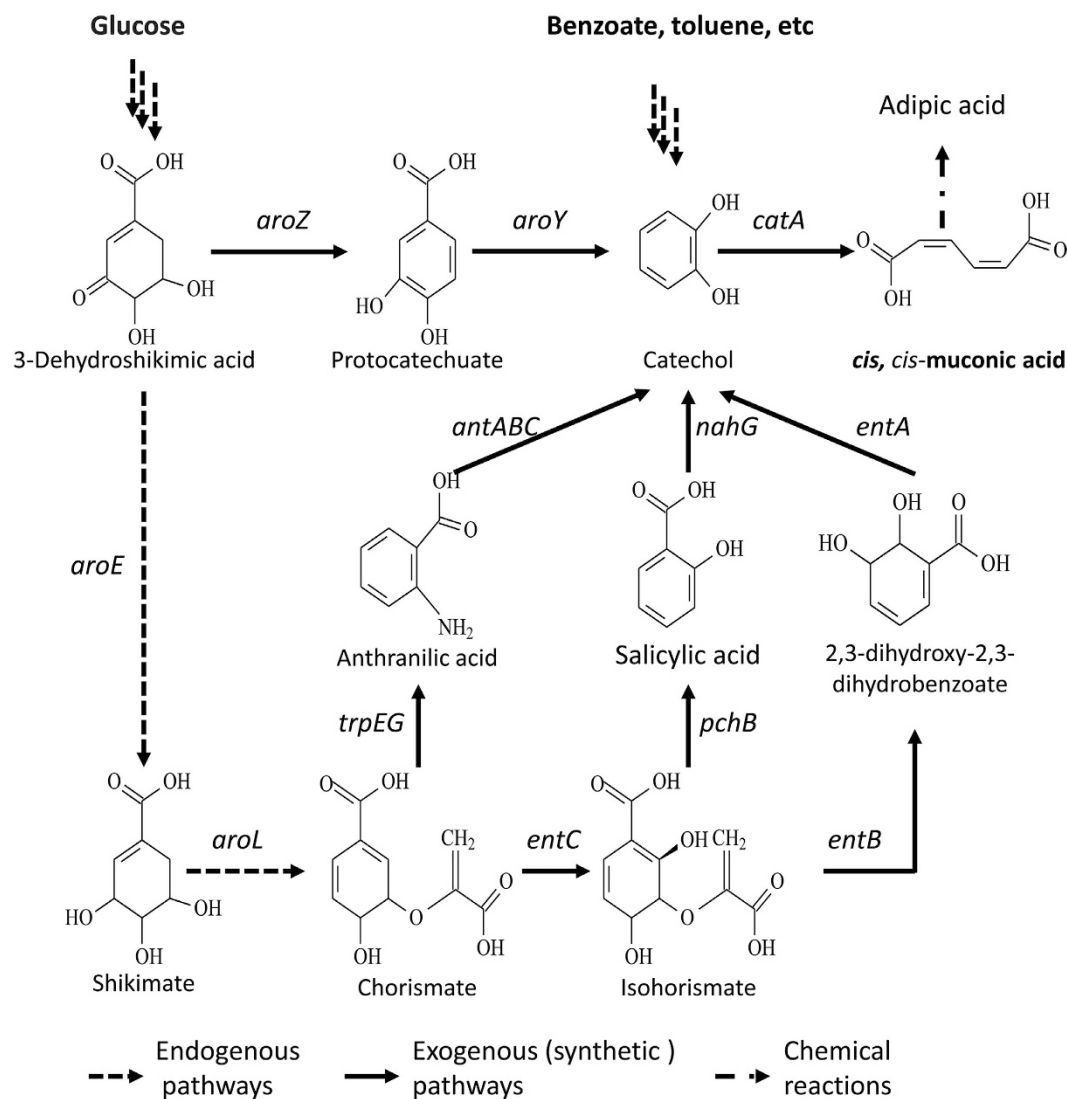


Figure 1. Summary of synthetic pathways for the production of *cis, cis*-muconic acid from glucose.

aroZ encodes 3-dehydroshikimate dehydratase; *aroY* encodes protocatechuic acid decarboxylase; *catA* encodes catechol 1,2-dioxygenase; *trpEG* encodes anthranilate synthase; *antABC* encodes anthranilate 1,2-dioxygenase; *entC* encodes isochorismate synthase; *pchB* encodes isochorismate pyruvate lyase; *nahG* encodes salicylate 1-monoxygenase; *entB* encodes isochorismatase; *entA* encodes 2,3-dihydro-2,3-DHBA dehydrogenase.

ccMA is not endogenously produced from renewable sugars by any known organisms, but it can be produced via the catabolism of aromatic compounds, such as phenol, toluene, benzyl alcohol and benzoic acid, by many organisms, including yeast, such as *Candida* sp., and bacteria, such as *Acinetobacter* sp., *Pseudomonas* sp., *Rhodococcus* sp., and *Sphingobacterium* sp.^{7–9}. However, *ccMA* can also be produced from renewable glucose via the introduction of heterologous synthetic pathways that draw from naturally occurring intermediates in the shikimate pathway (Fig. 1)^{5,10–14}. In all of these *ccMA* synthetic pathways, as well as the catabolic pathways of aromatic compounds, catechol 1,2-dioxygenase (CatA) is the last and most important enzyme for converting catechol to produce *ccMA* (Fig. 1). To improve the production of *ccMA*, it is necessary to obtain a CatA with higher activity.

The comprehensive substrate diversity and catalytic properties of many CatAs from *Acinetobacter*, *Alcaligenes*, *Arthrobacter*, *Geobacillus*, *Pseudomonas*, *Rhodococcus*, *Sphingomonas*, *Stenotrophomonas* and *Trichosporon* have been reported and described^{15,16}. Additionally, global microbial genome and environmental metagenome sequencing efforts are also contributing ever-increasing genetic information on *catA*. These studies facilitate screening efforts to find CatAs with higher activities, which are needed for the efficient production of *ccMA*^{15,17}. Moreover, many CatAs have been structurally elucidated using crystal structures and modeling studies^{18–21}. The crystal structure of catechol 1,2-dioxygenase is formed from an $\alpha\alpha$, $\alpha\beta$ and $\beta\beta$ dimer by two identical or non-identical subunits. A non-heme Fe(III) ion is

| Strains | ccMA (g/L) | Catechol (g/L) | Yield** (%) |
|-----------------------|-------------|----------------|-------------|
| <i>E. coli</i> WZK | 0.34 ± 0.05 | 0.92 ± 0.08 | 3.9 ± 0.6 |
| <i>E. coli</i> WZC | 0.43 ± 0.02 | 0.85 ± 0.07 | 5.3 ± 0.3 |
| <i>E. coli</i> WZT | 0.71 ± 0.03 | 0.45 ± 0.03 | 9.3 ± 0.3 |
| <i>E. coli</i> WZPL25 | 1.21 ± 0.07 | 0.07 ± 0.02 | 22.2 ± 0.5 |
| <i>E. coli</i> WZI* | 1.15 ± 0.12 | 0.09 ± 0.02 | 20.1 ± 0.3 |

Table 1. ccMA production by engineered strains with different constitutive promoters or an inducible promoter*. This strain was a control. During fermentation, 1 mM IPTG was added to the fermentation medium to induce the expression of catechol 1,2-dioxygenase. Errors (standard deviations) were calculated based on at least three independent experiments. **Yield: produced ccMA against consumed glucose.

incorporated into the catalytic center and ligated by two histidine residues, two tyrosine residues, and a hydroxyl ion in a trigonal-bipyramidal geometry²¹. The interactions of the enzyme with the substrate catechol to form a binding pocket involve Leu73, Pro76, Ile105, Pro108, Leu109, Arg221, Phe253 and Ala254 in *Acinetobacter* sp. ADP1, in which Ile105, Pro108, Phe253, and Arg221 are conserved in the catechol 1,2-dioxygenase family^{18–22}. This structural information lays a solid foundation for the rational redesign of CatA to improve its catalytic activities and other related functions. Recently, the catalytic properties of CatA from *A. radioresistens* LMG S-13 were fine-tuned via the tailoring of a pocket shape to improving catalysis at the active site²². The mutation of residue Leu69 changed the substrate specificity, and some variants of Ala72 enhanced the k_{cat} toward chlorinated substrates. Although an increasing number of proteins have been engineered for altered substrate specificity/selectivity by rational design, the engineering of highly efficient enzymatic pathways for industrial-scale fuel and chemical production by increasing the catalytic activity of the key enzymes remains an overwhelming challenge and requires expanded efforts in metabolic engineering and synthetic biology²³.

In this study, a ccMA synthetic pathway in *Escherichia coli* was reengineered to produce ccMA without exogenous induction by regulating the constitutive expression of CatA. Next, new mutants derived from wild-type *Acinetobacter* sp. ADP1 CatA (WT) were designed and constructed to enhance enzyme activity by reshaping the substrate-binding pocket. Successful mutants increased the titer of ccMA in *E. coli*. The biochemical studies and computational modeling of the mutated CatAs further elucidated the underlying molecular mechanism of the improved enzymatic activities. Our results show that rational redesign can enable the improved performance of key enzymes and synthetic pathways and can be applied to many other metabolic engineering and synthetic biology studies.

Results

Reconstructing the ccMA synthetic pathway by constitutive promoter engineering. At present, the expression of the catechol 1,2-dioxygenase gene (*catA*) for ccMA production in *E. coli* is typically under the control of the lac promoter and induced by exogenous isopropyl- α -D-thiogalactopyranoside (IPTG)^{5,10,11,14}. Thus, the utility of an inducible promoter-based ccMA synthetic pathway is limited in future industrial applications. Therefore, we investigated the possibility of replacing the lac promoter with constitutive promoters. Three constitutive promoters from the antibiotic cassettes of pACYC177 or pACYC184, P_{kan} , P_{cm} , and P_{tc} , as well as PL25 from our synthetic promoter library²⁴, were used to replace the lac promoter of the catechol 1,2-dioxygenase gene (Fig. S1A).

The constitutive promoter-*catA* expression cassettes were constructed by either fusion PCR or chemical synthesis and then used to replace the inducible-*catA* expression cassette in pKD8.292⁵ to obtain four recombinant plasmids, pKD8.292K, pKD8.292C, pKD8.292T and pKD8.292PL25, with constitutive promoters P_{kan} , P_{cm} , P_{tc} and PL25, respectively (Fig. S1B). Then, pKD8.292K, pKD8.292C, pKD8.292T and pKD8.292PL25 were each co-transformed with the plasmid pKD8.243⁵ containing 3-dehydroshikimate dehydratase (*aroZ*) and protocatechuate decarboxylase (*aroY*) into *E. coli* AB2834 to obtain new engineered strains (*E. coli* WZK, WZC, WZT, and WZPL25, respectively) with ccMA heterologous synthetic pathways. All strains with constitutive promoters produced different amounts of ccMA without IPTG induction (Table 1), and the strain with the synthetic promoter PL25 (*E. coli* WZPL25) exhibited a similar titer as that of the strain with the inducible lac promoter (*E. coli* WZI). These results indicate that it is possible to replace the inducible promoter with a suitable constitutive promoter. However, these reconstructed ccMA synthetic pathways still accumulated the metabolic intermediate catechol (Table 1). This result implied that it would be necessary to increase the enzymatic activity of CatA.

Prediction and analysis of the structure-based CatA redesign to increase enzyme activity.

Based on the crystal structure of CatA from *Acinetobacter* sp. ADP1²¹, as well as data from the Protein Data Bank (PDB ID: 1DLT), the substrate-binding pocket model of wild-type CatA was reconstructed using PyMOL (Fig. 2). As shown in this substrate-binding pocket model, residues 105–109 (purple loop) above the substrate catechol, residues 199–203 and 218–221 below the substrate (red parallel loop), and

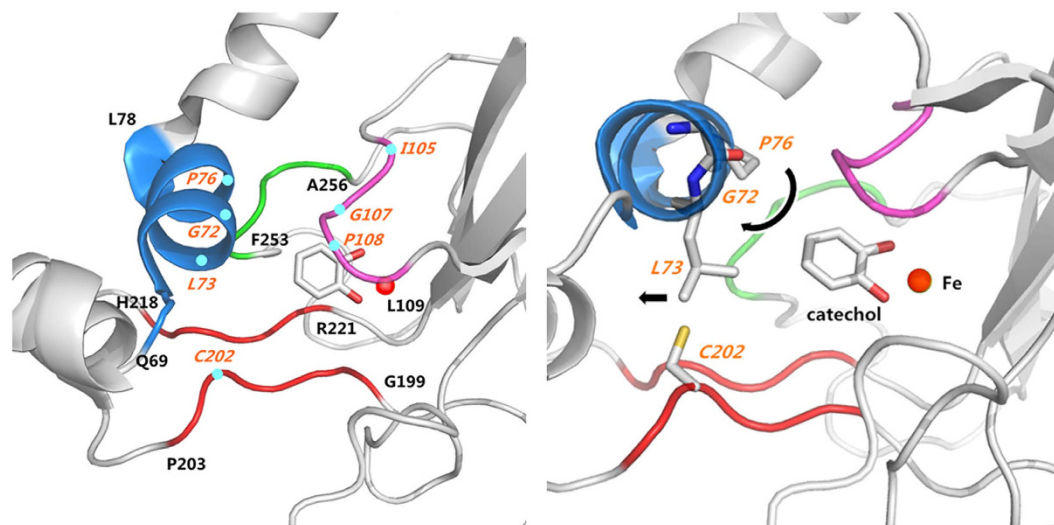


Figure 2. Substrate-binding pocket model of CatA *Acinetobacter* sp. ADP1 with different viewpoints. The blue helix, purple loop, green loop and red parallel loop are the key structures that determine pocket volume. G72 interacts with P108, and P76 interacts with I105. Both of these residues have interactions with the substrate (Left panel). The mutation of P76 and G72 might change the dihedral angles ϕ and ψ (Black arrows), thus twisting the blue helix and moving L73 to yield a wider entrance (Right panel).

residues 253–256 beside the substrate (green loop) are critical to maintain the shape of the binding pocket, and we hypothesized that introducing any mutation would likely seriously lower or destroy the enzymatic activity of CatA. However, residues 69–78 (blue helix) to the upper left of the substrate assist in regulating the size of the substrate-binding pocket. At this location, these residues act as a cap on the edge of the binding pocket, making these residues a suitable choice for mutation to improve enzymatic activity. Thus, two design strategies were considered for changing enzyme activity. The first strategy was to increase the substrate-enzyme binding energy, and the second strategy was to enlarge the binding cavity to increase substrate access.

As shown in Fig. 2, Gly72 is located immediately above the substrate and is one of the key residues interacting with the substrate catechol. We posited that if Gly72 is mutated to Ala or Val to increase hydrophobicity, this mutation might influence interactions with catechol, as well as the location of Pro108, to change the size of the pocket. Additionally, the interaction of Pro108 with hydrophobic residues at position 72 might twist the blue helix slightly. Thus, Leu73 might yield more space for a substrate to enter the binding pocket. Moreover, the larger branched chain of Val72 may produce a bigger cavity. Leu73 directly controls the size of the pocket like a switch. If it is mutated to Phe or Met, these hydrophobic amino acid residues might push away Cys202, thus enlarging the binding-pocket entrance while maintaining hydrophobicity. However, if the extended side chains of Phe and Met point toward the substrate after mutation, they might also block entry into the binding pocket.

Similar to Gly72, we considered the mutation of Pro76 to Ala, Gly or Val. Based on our structural predictions, the mutation of Pro76 might lead to two outcomes, either altering the hydrophobic interactions with the substrate catechol or giving structural flexibility to the blue helix. If Pro76 is mutated into Ala76 or Gly76, these changes may decrease hydrophobic interactions with the substrate catechol and Ile105 and result in enlargement of the binding pocket. If Pro76 is mutated to Val76, this mutation may increase hydrophobic interactions with the substrate but may also push away Ile105 farther. Thus, this mutation might also enlarge the binding pocket. Furthermore, the mutation of Pro76 to Ala76 can bring flexibility to the blue helix, thus bringing flexibility to the direction of Leu73. This change will enlarge the entrance of the binding pocket. In addition, double mutations at residues Leu73 and Pro76 may have synergistic effects on the size of the binding pocket. Based on the predictions and analyses above, eight single mutations and one double mutation were designed for experimental validation (Table 2).

Experimental validation of the structure-based redesign of CatA for increasing enzyme activity and improving the production of ccMA. To verify our structure-based functional predictions and efficiently screen the mutated CatAs, site-specific mutagenesis was used to generate the mutants G72P, G72A, G72V, L73F, L73M, P76A, P76G, P76V and L73F/P76A with the recombinant plasmid pKD8.29PL25. pKD8.29PL25 derivatives with different mutations were co-transformed with pKD8.243 containing the genes *aroZ* and *aroY* into *E. coli* AB2834 to obtain engineered *E. coli* strains for the analysis of CatA enzymatic activity, as well as for the improved production of ccMA. In *E. coli* AB2834, *aroE* was mutated to redistribute the flux and thus improve the production of ccMA⁶.

| Strains [*] | Specific Activity (U/mg) ^{**} |
|----------------------|--|
| WT | 2.489 ± 0.112 |
| G72P | 0.006 ± 0.002 |
| G72A | 15.828 ± 0.491 |
| G72V | 0.004 ± 0.001 |
| L73F | 6.425 ± 0.301 |
| L73M | 0.0652 ± 0.003 |
| P76A | 25.2629 ± 0.626 |
| P76G | 0.043 ± 0.002 |
| P76V | 0.005 ± 0.002 |
| L73F/P76A | 14.019 ± 0.549 |

Table 2. Specific activities of catechol 1,2-dioxygenase in crude cell extracts from different engineered strains with wild-type and mutated *catA*. ^{*}WT indicates *E. coli* AB2834 containing pKD2.834A and pKD2.292 with wild-type *catA*. Other strains contain different mutated *catAs*. ^{**}1U = 1 μmol *ccMA* produced per minute at 24 °C. Errors (standard deviations) were calculated based on at least three independent experiments.

Crude cell extracts from the engineered *E. coli* strains with wild-type or mutated CatAs exhibited different specific activities of catechol 1,2-dioxygenase (Table 2). Four variants, G72A, L73F, P76A, and L73F/P76A, showed higher specific activities, and the mutant P76A had the highest specific activity with more than 10-fold that of the WT toward the substrate catechol. Based on our previous modeling and mutation designs, the behavior of nearly half of the mutations were consistent with our prediction analysis. The successful rate for predictions was not comparatively high, but the proposed hypothesis for improving enzymatic activity via enlarging the binding pocket was still helpful. In addition, engineering the substrate specificity of CatA was reported in the previous report by tailoring the catalysis at the active sites²², but improving the enzymatic activity of CatA significantly by structural-based redesign was first reported here.

To test the influence of the mutated CatAs on the performance of the *ccMA* synthetic pathway in *E. coli*, the production of *ccMA* and certain related metabolic intermediates were measured during growth in a minimal medium. As shown in Fig. 3A, the strains harboring mutated CatAs with higher enzymatic activities showed increased *ccMA* production. Although the G72A strain did not show a significant difference versus WT, the other three mutants, L73F, P76A and L73F/P76A, led to significantly increased production of *ccMA*. The strain with mutation P76A produced the highest amount of *ccMA*, 1.53 g/L, an approximately 26.4% improvement compared with the strain with WT CatA, whereas L73F and L73F/P76A produced approximately 1.42 g/L (a 17.4% improvement) and 1.48 g/L *ccMA* (a 22.3% improvement), respectively.

Catechol accumulated in the medium when the enzymatic activity of CatA decreased (Fig. 3B). As shown in Fig. 3C, increased CatA activity also contributed to a decrease in the intermediate protocatechuic acid. Our results indicate that the performance of the *ccMA* synthetic pathway may be improved by redesigning key enzymes. However, some protocatechuic acid still accumulates in our engineered strains. This finding demonstrates that the activity of AroY must be further increased (Fig. 1). Although we could largely increase the enzymatic activity of CatA, the titer of *ccMA* from the engineered *E. coli* with mutated CatAs was not increased dramatically. The supply of precursors in the *ccMA* synthetic pathway was likely not adequate, thus limiting the increase of the *ccMA* titer. This hypothesis was confirmed by our ongoing engineering studies for improving the production of 3-dehydroshikimic acid (Fig. 1), and subsequently increasing *ccMA* production (unpublished data).

Biochemical studies and computational simulation to elucidate the mechanisms of the effective mutants.

To investigate how the mutation of CatA affected enzymatic activity and the production of *ccMA*, biochemical studies of wild-type and mutated CatAs were first conducted to elucidate the underlying functions. Based on previous results, three variants, L73F, P76A and L73F/P76A, as well as WT were expressed and purified from *E. coli* for biochemical characterization (Fig. 4). K_m and k_{cat} values of the recombinant proteins were measured and calculated based on the Michaelis-Menten equation by nonlinear fitting (Table 3). The K_m values of L73F and P76A were nearly 20- and 9-fold higher than that of WT, respectively. This result indicates that the residues at positions 73 and 76 greatly influence substrate binding. However, the k_{cat} of P76A and L73F/P76A increased by more than 5-fold versus WT, emphasizing their increased catalytic activities and resulting in increased *ccMA* production. The L73F mutation exhibited a clear effect on K_m , but its k_{cat} value was similar to that of WT. Hence, the production of *ccMA* by the L73F-containing strain was less than that of P76A or L73F/P76A.

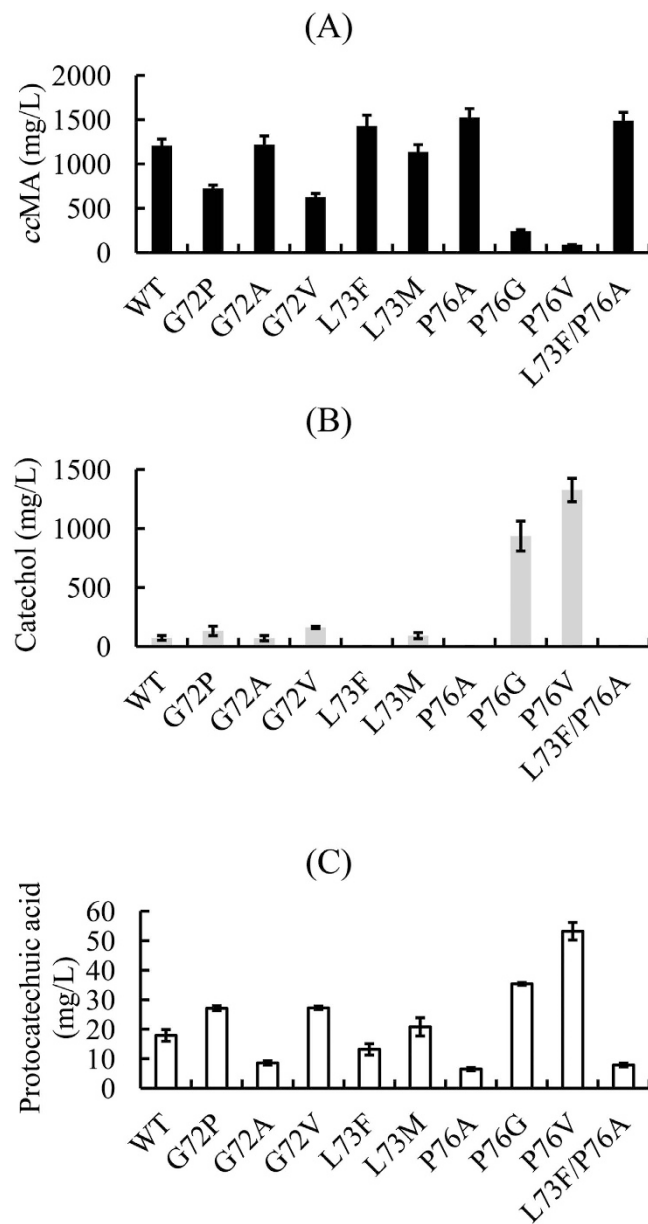


Figure 3. ccMA, catechol and protocatechuic acid production in engineered *E. coli* strains with wild-type or mutated CatAs.

To further investigate how the L73F, P76A and L73FP76A mutations affect the catalytic activity of CatA and ccMA production, molecular dynamic (MD) simulations were used to elucidate the underlying mechanisms. Based on the previous structural analysis²⁰, a non-heme Fe(III) ion is a cofactor in the binding pocket of CatA, and this non-heme Fe(III) ion plays a key role in maintaining enzyme activity. When the substrate catechol enters, the metal should be ligated by five ligands: catechol, two histidine residues, one tyrosine residue and a water molecule, leading to the formation of a hexacoordinated octahedron-shaped structure¹⁹. To better analyze the molecular mechanism of CatA, an iron force field (Fig. S2) was constructed using Pang's method²⁵. With this iron force field, wild-type CatA and the three mutants L73F, P76A and L73F/P76A were subjected to 2-ns MD simulations to verify the experimental results.

Because the distances between residues 73 and Cys202 or Leu109 and between residues 76 and Ile105 or Gly107 will decide the volume of the binding pocket, these distances were measured in WT and the three mutants based on the simulation results (Table 4, Figs S3 and S4). As shown in Table 4, the mutation of Leu73 to Phe decreased the distance between residues 73 and Leu109 or Cys202. However, the mutation of Pro76 to Ala increased the distance between residues 76 and Ile105 or Gly107. The double mutation L73F/P76A exhibited a similar distance to Leu109 or Cys202 and an increased distance to Ile105 or Gly107 compared with those in WT. The distance changes should affect the volume of the

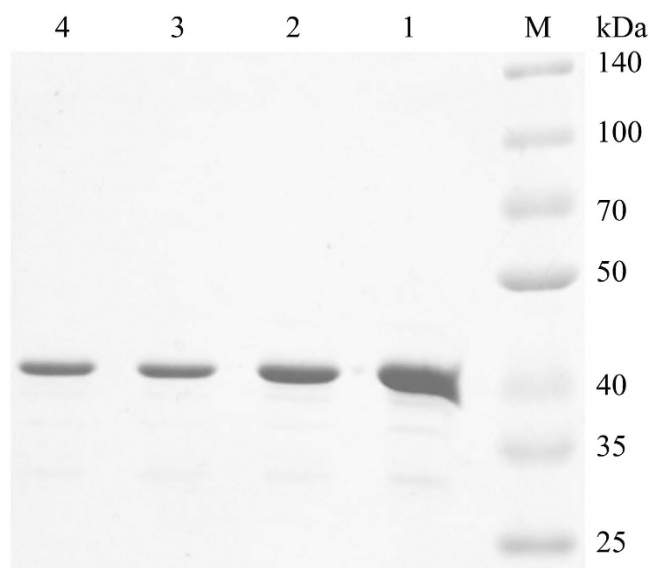


Figure 4. SDS-PAGE analysis of purified wild-type and mutant CatA proteins. Lane M, protein marker; lane 1, WT; lane 2, L73F; lane 3, P76A; lane 4, L73F/P76A.

| Enzymes | K_m (μM) | k_{cat} (S^{-1}) |
|-----------|-------------------------|-------------------------------|
| WT | 7.88 ± 1.239 | 7.86 ± 0.248 |
| L73F | 151.09 ± 10.299 | 5.65 ± 0.163 |
| P76A | 65.91 ± 7.753 | 46.56 ± 0.201 |
| L73F P76A | 45.96 ± 7.691 | 39.50 ± 0.221 |

Table 3. Enzymatic kinetic parameters of wild-type and mutated CatAs after purification against the substrate catechol*. *WT indicates wild-type CatA. Others are mutated proteins. Errors (standard deviations) were calculated based on at least three independent experiments.

| Characteristics | L73F | P76A | L73F/P76F | WT |
|--|---------------|---------------|---------------|---------------|
| Leu73 or Phe73 to Leu109 | 6.7 ± 0.4 | 7.9 ± 0.4 | 7.6 ± 0.4 | 7.6 ± 0.4 |
| Leu73 or Phe73 to Cys202 | 3.7 ± 0.4 | 4.3 ± 0.4 | 3.9 ± 0.4 | 3.9 ± 0.3 |
| Pro76 or Ala76 to Ile105 | 4.1 ± 0.3 | 4.2 ± 0.3 | 4.7 ± 0.4 | 3.7 ± 0.3 |
| Pro76 or Ala76 to Gly105 | 4.5 ± 0.3 | 6.7 ± 0.2 | 6.7 ± 0.2 | 5.1 ± 0.3 |
| Size of active pocket (\AA^3) | 123 ± 10 | 178 ± 13 | 152 ± 11 | 132 ± 12 |

Table 4. Distances (\AA) between certain key residues and the size of the active pocket for wild-type and mutated CatAs*. *50 snapshots of last 250 ps trajectories (5 ps per snapshot) during molecular dynamic simulation were used to calculate the distance and the volume. The results represent the mean \pm standard deviation.

binding pocket. Therefore, the pocket volumes of the WT and mutants were simulated and measured by POVME²⁶. As shown in Table 4 and Fig. 5, the pocket volumes of WT, L73F, P76A and L73F/P76A were 132, 123, 178 and 152\AA^3 , respectively. The shortened distance between Phe73 and Leu109 narrowed the width of the pocket gate, but the cavity volume in L73F did not change considerably. This result explains why the k_{cat} value of L73F is lower than that of WT. For P76A, Gly107 pushed Ala76 away when Pro76 was mutated to Ala, and this resulted in an expanded cavity and pocket volume. With the combinatorial effect of L73F and P76A, the pocket volume of the double mutant L73F/P76A was smaller than that of P76A but larger than that of L73F. The volumes of the mutants corresponded with their k_{cat} values. Our results showed that the computational modeling was in good agreement with the experimental outcome, and increasing the volume of the binding pocket will facilitate substrate entry or product release, thus

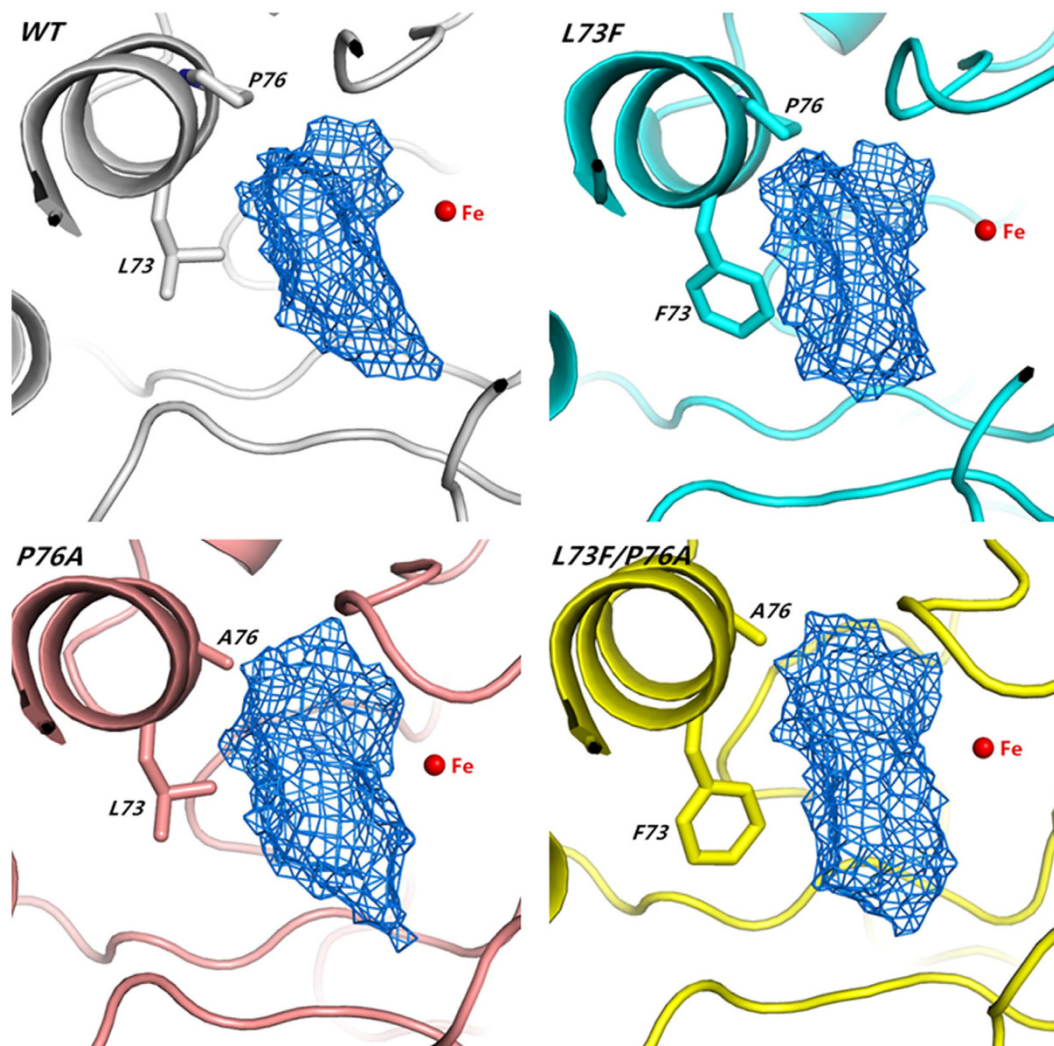


Figure 5. Pocket volume of wild-type and mutated CatA. The blue lattice displays the volume of the binding-pocket cavity. The pocket volumes of WT, L73F, P76A and L73F/P76A are 132, 123, 178 and 152 Å³, respectively.

enhancing the catalytic rate of the enzyme. This finding implies that the binding pocket is a possible target for improving the catalytic efficiency of any enzyme.

Discussion

As a high value-added bioproduct with reactive dicarboxylic groups and conjugated double bonds, *cc*MA has gained increasing interest due to its potential applications in the manufacture of commercially important bulk chemicals, such as adipic acid, terephthalic acid and trimellitic acid, new functional resins, bioplastics, food additives, agrochemicals, and pharmaceuticals²⁷. Although many microorganisms can convert some aromatic compounds, such as catechol, benzoate and toluene, to *cc*MA, it is almost impossible to develop an economically feasible and environmentally friendly process based on these bioconversions. Hence, recent efforts have focused on a promising approach for more sustainable production of *cc*MA based on the design of a *de novo* biosynthetic pathway (Fig. 1). As a result, this compound was produced by engineered *E. coli* or yeast with biomass-derived glucose as the substrate, although the performances of the majority of these synthetic pathways must be further improved for future industrial applications.

As shown in Fig. 1, all synthetic pathways for *cc*MA production utilize the same last biochemical reaction catalyzed by CatA. Therefore, the selection of a CatA with suitable protein expression and activity is key for improving the performance of the synthetic pathways^{15,17}. Unfortunately, the expression of catechol 1,2-dioxygenase in many reported *cc*MA synthetic pathways used the *lac*-derived expression system^{5,10,11,14}. This protein expression system requires an inducer in the medium, the un-metabolizable compound IPTG. IPTG is expensive and cannot be used on an industrial scale. It is highly toxic, and it yields a highly intense response. Hence, in this study, we engineered the expression of CatA by a

constitutive promoter instead. According to our results (Table 1), the selected synthetic promoter PL25 performed similarly to, even better than the inducible lac promoter. This genetic modification will contribute to future industrial applications. Moreover, this selected synthetic promoter PL25 was from our designed promoter library with different strengths, which would be very useful to regulating other synthetic pathways constitutively²⁴.

Enhancing the activity or catalytic efficiency of key enzymes by protein redesign is another important strategy for improving the performance of synthetic pathways^{23,28}. Many proteins can be engineered for altered substrate specificity/selectivity, increased catalytic activity, reduced mass-transfer limitations through specific protein colocalization, and reduced substrate/product inhibition. There are many reports of engineered enzymes that increase the overall performance of a pathway. Previous studies have demonstrated that substrate specificity/selectivity can be improved by reshaping or enlarging the binding pocket of an enzyme^{22,29}; however, in this study, we demonstrated that enlarging the binding pocket could also enhance enzymatic activity (Table 4 and Fig. 5), thus improving the performance of our ccMA synthetic pathway.

Based on computational modeling, we speculated that the larger binding pocket facilitates substrate entry or product release, resulting in increased catalytic activity. Enlarging the substrate-binding pocket of an enzyme would make it have ability to catalyze the bigger substrate. However, for the smaller substrate or original substrate, enlarging the binding pocket would weak the interaction energy. As a consequence, the K_m of the enzyme would increase. On the other hand, the bigger binding pocket would give substrate more chance to enter the pocket. Therefore, it would also increase k_{cat} . In this study, we try to enlarge the entrance of binding pocket to increase k_{cat} while maintaining the volume of pocket bottom to prevent K_m from increasing dramatically (Table 3). This strategy should also be applied to other protein redesigns for increasing enzymatic activity and improving the performance of metabolic pathways.

In addition, in this study, we did dramatically improve the enzymatic activity of CatA by structure-based rational design, but the improvement of ccMA production for the strain with effective mutated CatA was not so significant. This confirmed the fact that decreasing one specific metabolic bottleneck will automatically make another one to rise. Hence, further engineering other enzymes for new limited steps with similar redesign strategy presented in this study is very necessary to further improve the production of ccMA.

Methods

Bacterial strains and media. *E. coli* trans5a from Beijing Transgen Biotech (Beijing, China) was used for plasmid construction and molecular cloning. *E. coli* AB2834 with a 3-dehydroshikimate dehydrogenase (AroE) deficiency (from The Coli Genetic Stock Center at Yale University) was used for ccMA production. *E. coli* BL21(DE3) was used for protein expression to obtain the purified recombinant proteins for kinetic analysis. All strains used in this study are shown in Table S1 (Supplementary data). LB medium containing 10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl was used for cell growth. Modified M9 minimal medium containing 40 mg/L shikimate, 10 g/L glucose, 6 g/L Na₂HPO₄, 0.5 g/L NaCl, 3 g/L KH₂PO₄, 2 g/L NH₄Cl, 246.5 mg/L mg/L MgSO₄, and 14.7 mg/L CaCl₂ was used for ccMA production. If necessary, chloramphenicol (Cm^R), spectinomycin (Spe^R), kanamycin (Kan^R), or IPTG was added to the medium at 34 μg/mL, 50 μg/mL, 50 μg/mL, or 1 mM, respectively.

Plasmid construction. The plasmid pKD8.243 was obtained from *E. coli* ATCC 69875⁵. The primers used for plasmid construction in this study are listed in Table S2, and the restriction enzyme sites are underlined in the primers. The 220-bp constitutive promoters P_{kan}, P_{cm}, and P_{tc} from the Kan^R, Cm^R, and Tc^R cassettes in pACYC177 or pACYC184 were amplified by polymerase chain reaction (PCR) with the primers Pkan-EcoRI-5/Pkan-3, Pcm-EcoRI-5/Pcm-3, and Ptc-EcoRI-5/Ptc-3, respectively. The 1080-bp *catA* gene for each constitutive promoter cassette was amplified from pKD8.292⁵ by PCR with the primers CatA-kan-5/CatA-EcoRI-3, CatA-cm-5/CatA-EcoRI-3, or CatA-tc-5/CatA-EcoRI-3, respectively. Then, the PCR products of each constitutive promoter and the *catA* gene were ligated by fusion PCR using the overlapping sequences that were introduced by each primer pair. After *EcoRI* digestion, the 1.3-kb fusion PCR products were each ligated with the 4.2-kb *EcoRI*- and calf intestinal alkaline phosphatase-digested pKD8.292 product to obtain the new recombinant plasmids (Fig. S1B). In pKD8.292K, pKD8.292C and pKD8.292T, the lac promoters are replaced by the constitutive promoters from the Kan^R, Cm^R, or Tc^R cassettes (Fig. S1A), respectively. Wild-type *catA* from *Acinetobacter* sp. ADP1⁵ with the constitutive synthetic promoter PL25 (TCTAGAATATGTTATCTCTGGCGGTGTTGACAACGAGCTGGACAACCTGGTATAATGCCACATGAGCGGATAACAATTTCAAGGAGGACAGCTCCATGG)²⁴ was chemically synthesized by GENEWIZ, Inc. in Suzhou, China (Fig. S1C). This synthetic DNA fragment was cloned into pKD8.292 to obtain the recombinant plasmid pKD8.29PL25 by replacing the original *catA* with the inducible lac promoter via *EcoRI* restriction digestion and ligation. The recombinant plasmids with different mutated *catAs* were constructed by inverse PCR of pKD8.29PL25. All recombinant plasmids in this study were confirmed by DNA sequencing and are shown in Table S1.

Microbial production of ccMA and metabolite analysis. Different plasmids with wild-type or mutated *catA*, as well as pKD8.243, were transformed into *E. coli* AB2834 to evaluate the production

of *ccMA*. The engineered strains were inoculated in 3 mL of LB medium. Overnight cultures of these strains were diluted 1:10 into modified M9 minimal medium containing 10 g/L glucose, appropriate antibiotics and/or IPTG (for the lac promoter) and cultivated at 37 °C for 36 h with shaking at 220 rpm. From each culture, 1 mL was collected and centrifuged at 10,000 g to remove cells for metabolite analysis. Metabolites in the culture supernatants, including glucose, *ccMA* and other metabolites, were quantified by HPLC using a reverse-phase ZORBAX C18 (Agilent Technologies) or HPX-87H column (Bio-Rad Laboratories). All experiments were performed minimally in triplicate to ensure reproducibility, and the results represent the mean \pm standard deviation.

Expression, purification and enzyme assay of catechol 1,2-dioxygenase. Wild-type and mutant *catA* were cloned into the expression plasmid pET30a(+) between *NdeI* and *XhoI* to obtain recombinant plasmids (pET-CatAWT, pET-CatAL73F, pET-CatAP76A and pET-CatAL73F/P76A, Table S1) for protein overexpression to obtain the purified recombinant protein for kinetic assays. *E. coli* BL21(DE3) with the different recombinant plasmids (*E. coli* CatAWT, *E. coli* CatAL73F, *E. coli* CatAP76A and *E. coli* CatAL73F/P76A) were induced using 1 mM IPTG for 7 h at 20 °C in LB medium for protein production. For purification, the cells were harvested by centrifugation and disrupted by ultrasonication in 20 mM phosphate buffer (pH 7.4) containing 20 mM imidazole and 0.25 M NaCl. The supernatant with the recombinant protein was obtained by centrifugation at 10,000 g for 30 min, filtered through a 0.22- μ m membrane and then purified by Ni²⁺-affinity chromatography from GE Healthcare Life. Imidazole in the purified protein was removed by tube ultrafiltration. The resulting enzyme solution was used for activity assays. The reaction was performed in a 1 mL system containing 0.2 mM catechol, 100 μ M potassium phosphate buffer (pH 7.4), and 100 μ L of enzyme solution at 25 °C for 20 min. Enzyme activity was measured and calculated by the production of *ccMA* via HPLC analysis. The protein concentration was determined with the Thermo Scientific Pierce BCA quantification assay kit.

Kinetic parameter determination. Enzyme catalytic parameters (Michaelis-Menten constant K_m and k_{cat}) were calculated by measuring the initial linear rates of the enzymatic reaction after the addition of different concentrations of catechol ranging from 0 to 200 μ M in 100 mM potassium phosphate buffer (pH 7.4). After the addition of the purified enzyme, the reactions were performed, and the OD₂₆₀ was monitored at 25 °C using with Spectramax M^{2e} apparatus (Molecular Devices, USA). The extinction coefficient of the catechol oxidation product was determined as $\epsilon_{260nm} = 16,800 \text{ M}^{-1} \text{ cm}^{-1}$. Three independent measurements were performed. K_m and k_{cat} were calculated based on the Michaelis-Menten equation. One unit was defined as the amount of enzyme producing 1 μ mol *ccMA* per minute.

Computational modeling and simulations. The wild-type model of CatA from *Acinetobacter* sp. ADP1 was constructed using PyMOL (Version 1.5.0, Schrödinger LLC) based on the crystal structure of CatA (PDB: 1DLT)²¹. The mutant structural models were constructed based on the 1DLT PDB file by renaming the target residue, removing the line of the different atom name between the original residue and mutated residue, and using the leap program from Amber Tools 12 and Amber12³⁰ to add the missing atoms. Finally, the entire system was optimized using 5,000 steps of minimization. The iron force field was constructed using Pang's method²⁵. MD simulation for the three mutants L73F, P76A, and L73F/P76A and WT was conducted using the Leap module of Amber12^{30,31} with the standard Amber99SB force field and manually creating an iron force field based on the crystal structure from the Protein Data Bank (PDB ID: 1DLT). All simulations were run with a 0.5-fs time step and gradually reduced constraints on protein and langevin dynamics³² for temperature control. Then, the final phase of equilibration of every system was used to run a 2 ns MD simulation. POVME was used to calculate the volume of the binding pocket²⁶. Some detailed information on computational modeling and simulation are provided in the supplementary methods.

References

- Chen, Y. & Nielsen, J. Advances in metabolic pathway and strain engineering paving the way for sustainable production of chemical building blocks. *Curr. Opin. Biotechnol.* **24**, 965–972 (2013).
- Chen, Z. & Zeng, A. P. Protein design in systems metabolic engineering for industrial strain development. *Biotechnol. J.* **8**, 523–533 (2013).
- Yadav, V. G., De Mey, M., Lim, C. G., Ajikumar, P. K. & Stephanopoulos, G. The future of metabolic engineering and synthetic biology: towards a systematic practice. *Metab. Eng.* **14**, 233–241 (2012).
- Polen, T., Spelberg, M. & Bott, M. Toward biotechnological production of adipic acid and precursors from biorenewables. *J. Biotechnol.* **167**, 75–84 (2013).
- Draths, K. M. & Frost, J. W. Environmentally compatible synthesis of adipic acid from D-glucose. *J. Am. Chem. Soc.* **116**, 399–400 (1994).
- Niu, W., Draths, K. M. & Frost, J. W. Benzene-free synthesis of adipic acid. *Biotechnol. Prog.* **18**, 201–11 (2002).
- Yoshikawa, N., Mizuno, S., Ohta, K. & Suzuki, M. Microbial production of *cis, cis*-muconic acid. *J. Biotechnol.* **14**, 203–210 (1990).
- van Duuren, J. B. *et al.* pH-stat fed-batch process to enhance the production of *cis, cis*-muconate from benzoate by *Pseudomonas putida* KT2440-JD1. *Biotechnol. Prog.* **28**, 85–92 (2012).
- Wu, C. M. *et al.* Microbial synthesis of *cis, cis*-muconic acid from benzoate by *Sphingobacterium* sp. mutants. *Biochem. Eng. J.* **29**, 35–40 (2006).

10. Sun, X., Lin, Y., Yuan, Q. & Yan, Y. Biological production of muconic acid via a prokaryotic 2, 3-dihydroxybenzoic acid decarboxylase. *ChemSusChem* **7**, 2478–2481.
11. Lin, Y., Sun, X., Yuan, Q. & Yan, Y. Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in *Escherichia coli*. *Metab. Eng.* **23**, 62–69 (2014).
12. Curran, K. A., Leavitt, J. M., Karim, A. S. & Alper, H. S. Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. *Metab. Eng.* **15**, 55–66 (2013).
13. Weber, C. *et al.* Biosynthesis of *cis, cis*-muconic acid and its aromatic precursors, catechol and protocatechuic acid, from renewable feedstocks by *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **78**, 8421–30 (2012).
14. Sun, X., Lin, Y., Huang, Q., Yuan, Q. & Yan, Y. A novel muconic acid biosynthesis approach by shunting tryptophan biosynthesis via anthranilate. *Appl. Environ. Microbiol.* **79**, 4024–4030 (2013).
15. Guzik, U., Hupert-Kocurek, K., Sitnik, M. & Wojcieszynska, D. High activity catechol 1, 2-dioxygenase from *Stenotrophomonas maltophilia* strain KB2 as a useful tool in *cis, cis*-muconic acid production. *Antonie Van Leeuwenhoek* **103**, 1297–1307 (2013).
16. Strachan, P. D., Freer, A. A. & Fewson, C. A. Purification and characterization of catechol 1, 2-dioxygenase from *Rhodococcus rhodochrous* NCIMB 13259 and cloning and sequencing of its *catA* gene. *Biochem. J.* **333**, 741–747 (1998).
17. Kim, B. J., Choi, W. J., Lee, E. Y. & Choi, C. Y. Enhancement of *cis, cis*-muconate productivity by overexpression of catechol 1, 2-dioxygenase in *Pseudomonas putida* BCM114. *Biotechnol. Bioprocess Eng.* **3**, 112–114 (1998).
18. Micallella, C. *et al.* X-ray crystallography, mass spectrometry and single crystal microspectrophotometry: a multidisciplinary characterization of catechol 1, 2 dioxygenase. *Biochim. Biophys. Acta.* **1814**, 817–23 (2011).
19. Matera, I. *et al.* Catechol 1, 2-dioxygenase from the Gram-positive *Rhodococcus opacus* 1CP: quantitative structure/activity relationship and the crystal structures of native enzyme and catechols adducts. *J. Struct. Biol.* **170**, 548–564 (2010).
20. Di Nardo, G. *et al.* Structural roles of the active site iron(III) ions in catechol 1,2-dioxygenases and differential secondary structure changes in isoenzymes A and B from *Acinetobacter radioresistens* S13. *Arch. Biochem. Biophys.* **431**, 79–87 (2004).
21. Vetting, M. W. & Ohlendorf, D. H. The 1.8 Å crystal structure of catechol 1, 2-dioxygenase reveals a novel hydrophobic helical zipper as a subunit linker. *Structure* **8**, 429–440 (2000).
22. Caglio, R. *et al.* Fine-tuning of catalytic properties of catechol 1, 2-dioxygenase by active site tailoring. *ChemBiochem* **10**, 1015–1024 (2009).
23. Eriksen, D. T., Lian, J. & Zhao, H. Protein design for pathway engineering. *J. Struct. Biol.* **185**, 234–242 (2014).
24. Wu, Y. Q., Zhang, Y. Y., Tu, R., Liu, H. & Wang, Q. Construction of the synthetic promoter for *Escherichia coli* and application in the biosynthesis of *cis, cis*-muconic acid. *Chin. J. Biotech.* **29**, 760–771 (2013).
25. Pang, Y. P. Novel zinc protein molecular dynamics simulations: Steps toward antiangiogenesis for cancer treatment. *J. Mol. Model.* **5**, 196–202 (1999).
26. Durrant, J. D., Votapka, L., Sørensen, J. & Amaro, R. E. POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics. *J. Chem. Theory Comput.* **10**, 5047–5056 (2014).
27. Xie, N. Z., Liang, H., Huang, R. B. & Xu, P. Biotechnological production of muconic acid: current status and future prospects. *Biotechnol. Adv.* **32**, 615–622 (2014).
28. Davids, T., Schmidt, M., Böttcher, D. & Bornscheuer, U. T. Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr. Opin. Chem. Biol.* **17**, 215–220 (2013).
29. Zhang, K., Sawaya, M. R., Eisenberg, D. S. & Liao, J. C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc. Natl. Acad. Sci. USA.* **105**, 20653–20658 (2008).
30. Nurisso, A., Daina, A. & Walker, R. C. A practical introduction to molecular dynamics simulations: applications to homology modeling. *Methods Mol. Biol.* **857**, 137–173 (2012).
31. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
32. Pastor, R. W., Brooks, B. R. & Szabo, A. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **65**, 1409–1419 (1988).

Acknowledgments

This work was supported by the National Basic Research Program (973 Program, 2011CBA00800), the National Science Foundation of China (No. 31070640), and the Research Equipment Program of the Chinese Academy of Sciences (No. YZ201153).

Author Contributions

Q.W., L.H., P.L., J.L. and Y.M. designed the experiments; L.H., P.L., J.S., Y.W., Y.Z. and W.C. performed the experiments; L.H., P.L., J.L. and Q.W. analyzed the data; L.H., P.L., J.L. and Q.W. wrote the manuscript; all authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Han, L. *et al.* Engineering catechol 1, 2-dioxygenase by design for improving the performance of the *cis, cis*-muconic acid synthetic pathway in *Escherichia coli*. *Sci. Rep.* **5**, 13435; doi: 10.1038/srep13435 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>