Article

Improving Theaflavin-3,3′-digallate Production Efficiency Optimization by Transition State Conformation of Polyphenol Oxidase

Ying Huang [1], Changzheng Gao [2], Wei Song [3], Wanqing Wei [4], Xiulai Chen [4], Cong Gao [4], Jia Liu [4], Jing Wu [3] and Liming Liu [1,4,*]

[1] School of Food Engineering, Anhui Science and Technology University, Chuzhou 233100, China
[2] Department of Cardiology, Affiliated Hospital of Jiangnan University, Wuxi 214122, China
[3] School of Life Sciences and Health Engineering, Jiangnan University, Wuxi 214122, China
[4] State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, China
[*] Correspondence: mingll@jiangnan.edu.cn; Tel./Fax: +86-510-85197875

Abstract: Theaflavins (TFs) are good for health because of their bioactivities. Enzymatic synthesis of TFs has garnered much attention; however, the source and activity of the enzymes needed limit their wide application. In this study, a microbial polyphenol oxidase from *Bacillus megaterium* was screened for the synthesis of theaflavin-3,3′-digallate (TFDG). Based on structural and mechanistic analyses of the enzyme, the O-O bond dissociation was identified as the rate-determining step. To address this issue, a transition state (TS) conformation optimization strategy was adopted to stabilize the spatial conformation of the O-O bond dissociation, which improved the catalytic efficiency of tyrosinase. Under the optimum transformation conditions of pH 4.0, temperature 25 °C, (−)-epigallocatechin gallate/epicatechin gallate molar ratio of 2:1, and time of 30 min, Mu_4 (BmTyr$^{V218A/R209S}$) produced 960.36 mg/L TFDG with a 44.22% conversion rate, which was 6.35-fold higher than that of the wild type. Thus, the method established has great potential in the synthesis of TFDG and other TFs.

Keywords: polyphenol oxidase; protein engineering; transition state; theaflavin-3,3′-digallate

1. Introduction

Theaflavin (TFs) is a flavanol with a benzotoketone ring system produced by the condensation of gallocatechin and catechin [1]. TFs have various health benefits, including antioxidant, antiobesity, anti-inflammatory, and anticancer properties, and they can also prevent fatty liver and cardiovascular diseases [2–6]. TFs are naturally derived from the enzymatic oxidation of catechin monomers in fresh green tea leaves during the fermentation process. In addition to the four main compounds, namely theaflavin (TF), theaflavin-3,3′-digallate (TFDG), theaflavin-3′-gallate (TF-3′-G), and theaflavin-3-gallate (TF-3-G), more than 30 species of theaflavins have been found [7]. Among these, TFDG is generally regarded as the most effective bioactive component [8].

The production of TFDG involves direct extraction, chemical oxidation, or enzymatic methods [9]. However, the direct extraction is not suitable for industrial-scale production because of high costs, wastage of resources, and a low extraction rate. The chemical oxidation synthesis method has disadvantages such as a low extraction rate of TFs and a complicated operational process. Furthermore, according to the European and United States legislation, chemically synthesized additives cannot be used in cosmetics, foods, or beverages [10]. Consequently, the enzymatic synthesis of TFDG has garnered increasing attention, which is a natural synthesis method and can greatly decrease the production cost [11].

Polyphenol oxidase (PPO) and peroxidases (PODs) are two groups of enzymes that are important for the synthesis of TFs. Zhang et al. [12] integrated transcriptome and

metabolite analysis of tea leaves to identify PPO and POD genes that may be involved in TF production during tea processing. The results showed that PPOs mainly catalyze the oxidation of catechins to TFs, whereas PODs catalyze the oxidation of catechins to TFs and further catalyze the rapid polymerization of TFs to thearubigins. Therefore, PPO are the key enzymes in TF synthesis. To synthesize TFDG, purified PPOs from tea leaves were used as catalysts, which obtained a final yield of 11 mg/L at pH 5.5 and 33 °C [13]. In a different study, PPOs taken from pear fruits were purified, immobilized, and applied to the synthesis of TFDG, using (−)-epigallocatechin gallate (EGCG) and epicatechin gallate (ECG) as substrates, and resulted in a maximum yield of 42.23% [11]. These previous studies have made remarkable progress in the development and utilization of plant PPO. However, the process of isolating and purifying PPO directly from plant tissues is complex and inefficient. Therefore, identifying new and efficient PPO for TF production is important.

Computational strategies have emerged in recent years that become an effective and cost-effective way to obtain target mutants by virtual screening. Rational design is based on structural analysis and applies various computational strategies to design novel enzymes or to optimize protein properties using a limited number of mutations [14,15]. This approach has achieved remarkable success in industrial biology and synthetic biology. The transition state (TS) is a metastable state during a chemical reaction that bridges reactants and products. Enzymatic catalysis and TS theory indicate that the catalytic effect of enzymes is owed to the TS, which binds to the enzyme much more tightly than the substrate [16]. Based on this TS theory, much progress has been made in de novo enzyme design and protein modification. Therefore, the use of simple and efficient computational methods is highly preferable for enzyme engineering. Combining the advantages of directed evolution and rational design, we used a method of enzyme redesign based on the Rosetta enzyme design for the computer simulation of stable TS in order to improve oxidation efficiency.

In this study, a microbial PPO named *Bm*Tyr was screened. Based on the analysis of its structure and mechanism, the O-O bond dissociation (TS2) was found to be the rate-determining step. Therefore, we proposed a strategy to optimize the conformation of TS2, based on the transition state and structure. Compared with the traditional approach of random mutation, which required a lot of screening work, this strategy effectively reduced the workload. The best variant was *Bm*Tyr$^{V218A/R209S}$, which increased the specific activity by 4.67-fold and V_{max}/K_m by 6.97-fold toward EGCG. Finally, the optimization of the transformation conditions obtained a titer of 960.36 mg/L TFDG.

2. Results

2.1. Screening Enzymes for TFDG Synthesis

PPO from three categories including tyrosinases, catechol oxidases, and laccases, which could produce quinone, were selected as probes on the basis of the specified product query tool in the BRENDA Enzyme database. Using these probes, seven functionally known microbial PPOs from the NCBI database were selected through data mining. The selected PPOs were subcloned into the pET28a (+) vector and were heterologously expressed in *E. coli* BL21 (DE3). SDS–PAGE analyses were performed to observe soluble PPO proteins (Figure 1a, which were then purified to produce TFDG). Among them, the enzyme *Bm*Tyr from *Bacillus megaterium* showed the highest yield (127.62 mg/L TFDG) (Figure 1c). The product was identified by HPLC (Figure S1), and mass spectroscopy showed the [M−H]$^{-}$ ion at m/z 867, which was consistent with the standard sample (Figure 1b). Therefore, *Bm*Tyr was selected as a catalyst for TFDG production. However, the conversion rate was only 7.35%. Hence, improved *Bm*Tyr catalytic efficiency is required. To guide the engineering-improved *Bm*Tyr, the structure and mechanism of *Bm*Tyr were determined.
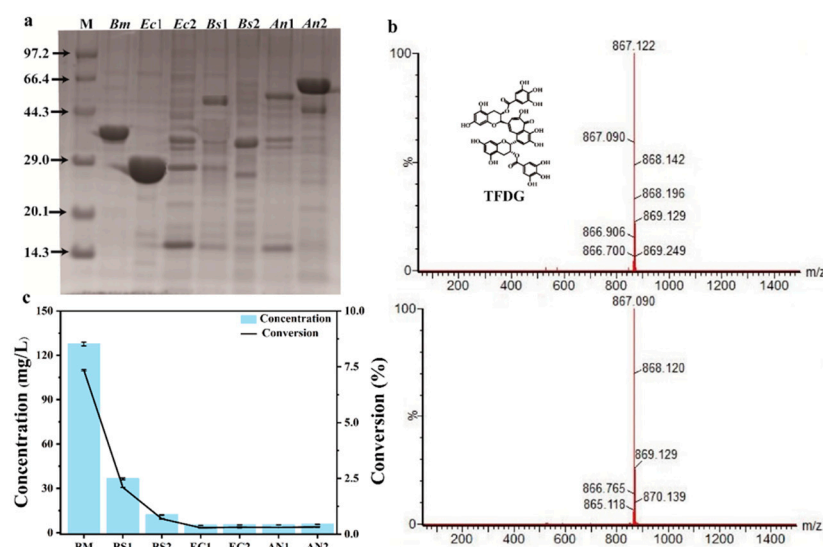
Figure 1. (**a**) SDS-PAGE analysis of the purified protein (M, Maker; *Bm*, *Bacillus megaterium* ACC86108.1, *Ec*1, *Escherichia coli 8739* AYB96806.1, *Ec*2, *Escherichia coli* AAC75642.1, *Bs*1, *Bacillus subtilis* MG571271.1, *Bs*2, *Bacillus subtilis strain N2* KY849863.1, *An*1, *Aspergillus niger ATCC 1015* EHA19536.1, *An*2, *Aspergillus niger ATCC 13496*) CAK40046.1. (**b**) Characterization of TFDG with HPLC-MS. (**c**) Effect of different activities of seven recombinant PPOs (5 mM substrate in 200 mM Mcilvaine Buffer (pH 5.0), 30 °C, 0.1 mM Cu^{2+}) on TFDG production. The data represent mean $\pm$ s.d., as determined from three independent experiments.

2.2. Structure and Catalytic Mechanism of BmTyr

*Bm*Tyr structure was obtained from the PDB database (PDB:3nm8) (Figure 2a), which was similar to that of the type 3 copper protein family containing two Cu (II) ions. These two Cu(II) ions were coordinated by six conserved histidine residues, which serve as the major cofactors. We aimed to identify the complex structure of *Bm*Tyr with Cu (II) ions and a ligand (EGCG); however, only the *apo-Bm*Tyr structure was obtained with 2.56 Å resolution (PDB ID:8HPI) (Figure S2). Therefore, based on 3nm8, a reactive binding mode M1 was created (Figure 2b). M1 contained two copper ions that were 3.5 Å apart and two bridging solvent molecules that were designated as oxygens. Furthermore, an MD analysis of M1 was performed, and the conformation with the highest evaluation rate was obtained through cluster analysis (Figure 2c). This conformation was named model 2 (M2), which was used for subsequent docking experiments with EGCG. The docking experiments produced a theoretically complex structure of M2–EGCG, which was named M3 (Figures 2d and S3).

The docking results are shown in Figure 3a. The flexibility of R209 facilitated the stabilization of EGCG in the *Bm*Tyr binding pocket. F197, a bulky residue, showed a weak π–π interaction with the aromatic ring of EGCG, forming a specific gate that prevented a substrate with incorrect orientation from entering the pocket (Figure 3b). The conserved water molecules, which were stabilized by E195 and N205, acted as proton donors and receptors during the process of hydroxylation. The other residues, including M184, P201, G216, V217, and V218, interacted to fix EGCG to facilitate the reaction (Figure 3b).

Based on the docking results and the reported mechanism underlying PPO-mediated reactions, a hypothetical catalytic mechanism of *Bm*Tyr was inferred (Figure 3c) [17–20]. This reaction contained the following steps: (i) phenolic -O11H bond formation with the bridging oxygen atom (TS1), (ii) O-O bond dissociation of the bridging oxygen species (TS2), (iii) another phenolic -O9H bond formation, and (iv) proton transfer from the catechol to the bridging oxygen group accompanied by electron transfer to afford the quinone product. A detailed analysis of the energy profile showed that the rate determination step is the O-O bond dissociation, regardless of the active copper center state [21]. Similar to the catechol oxidase diphenolase cycle, three protons are required to enter the stage where

the O-O bond is cleaved [22]. The O-O bond cleavage was a prohibitively large barrier for this reaction. MD simulations showed that the distance between the -O11H of the EGCG and the O1 of the bridging oxygen atom of the active center was 4.66 ± 0.57 Å (Figure S8), which made the hydride transfer and proton transfer difficult. Thus, optimizing this critical step required the optimum conformation of TS2.
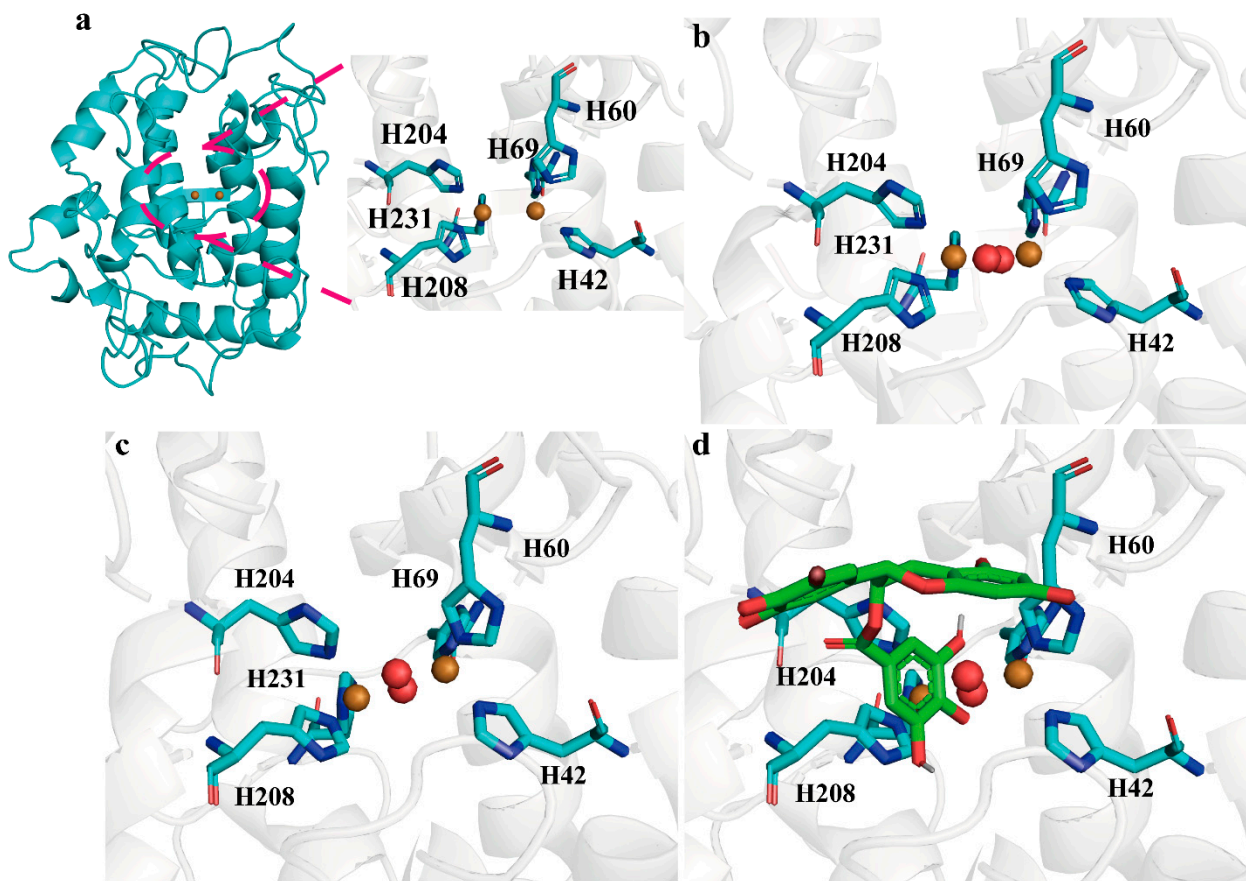

Figure 2. The construction process of the theoretical complex structure. (**a**) The whole structure and detailed active site view of 3nm8. (**b**) The detailed active site view of M_1. (**c**) The detailed active site view of M_2. (**d**) The detailed active site view of M_3. The carbon of histidines is shown in cyan stick and Cu^{2+} in brown sphere.

2.3. Improving BmTyr Catalytic Efficiency via TS2 Conformation Optimization

To determine the optimum conformation of TS2, restrained MD simulations were conducted. First, the MD state model was constructed by defining geometric criteria on the basis of the existing M3 and previous studies on the mechanism [20]. During the restrained MD calculations, the coordinates of the active site atoms were fixed. Subsequently, stable conformations were selected by cluster analysis to construct the active site cluster model, namely the TS2 state model (Figure S5). A total of 15 residues (M61, E158, M184, E195, F197, P201, N205, R209, Q214, M215, G216, V217, V218, A221, and F227) within 5 Å of the TS2 were selected for the subsequent Rosetta enzyme design (Figure 4a).

During Rosetta enzyme design, to accommodate the large backbone changes introduced at each step, we used PSSMs to limit amino acid choices to identities that were common in the multiple-sequence alignment (Figure S6) [23]. Thus, 119 mutants were designed in this study. Computational mutation scanning was performed to predict the effects of each mutation ($\Delta\Delta G$). Mutants with scores lower than that of the WT at the same residue position were selected as hits, which created 22 variants for experimental valida-

tion (Figure S7). Using PSSMs and $\Delta\Delta G \leq 0$ R.e.u. as the cutoff, severely destabilizing mutations were eliminated.
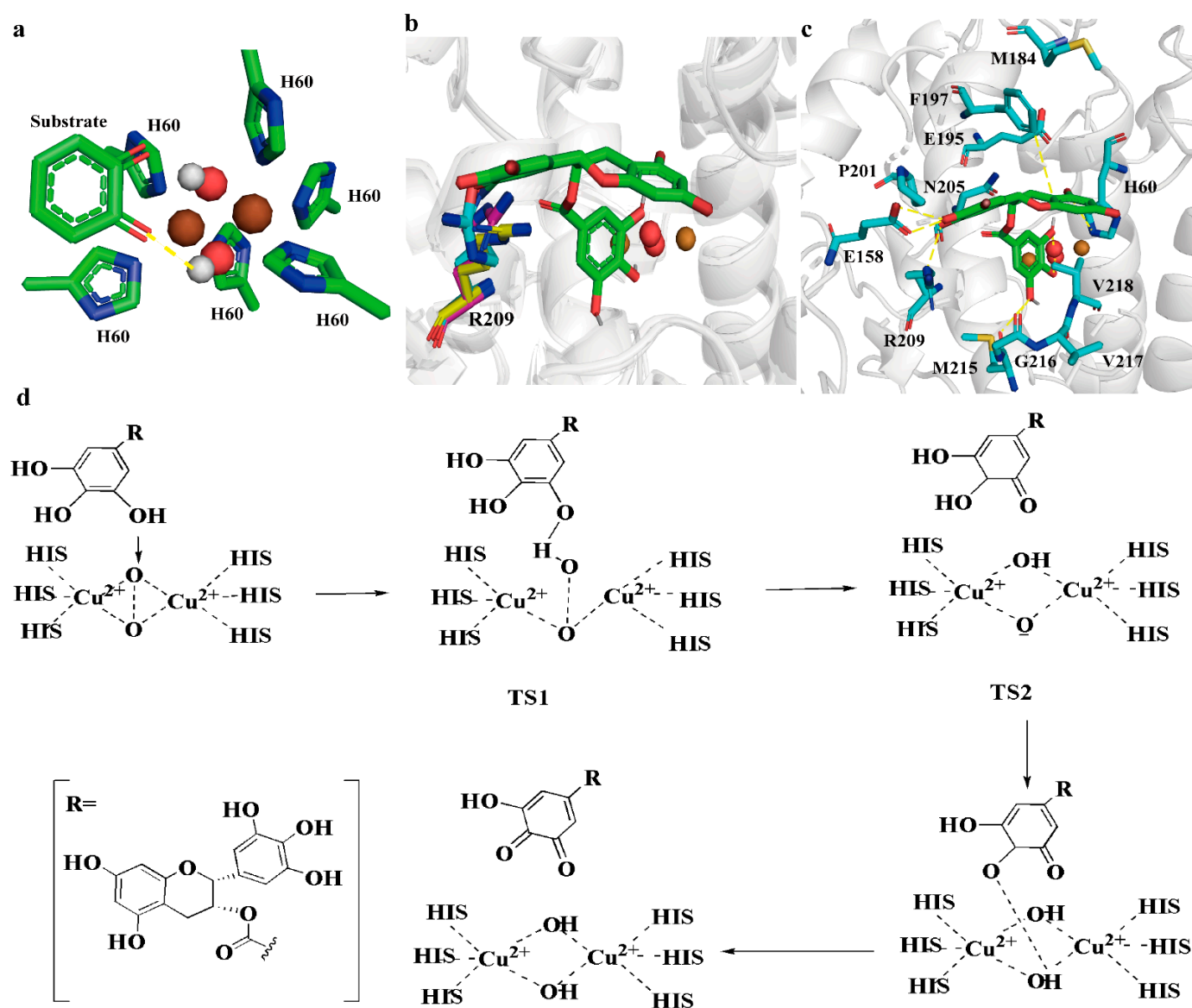

Figure 3. EGCG binding and catalytic mechanism of *Bm*Tyr. (**a**) Possible 3D view structure for TS3. (**b**) The three conformations of R209 identified in the molecular dynamics are presented. (**c**) The binding of EGCG in the active site of *Bm*Tyr. The carbon of EGCG is shown in green stick and the remaining residues are shown in cyan. (**d**) Catalytic mechanism of *Bm*Tyr.

The 22 variants were expressed in *E. coli* BL21 (DE3), which were purified and tested for TFDG production (Figure S7). Compared with the TFDG titers of WT, 19 of the 22 variants showed improved TFDG titers, which confirmed the effectiveness of this method. Among the 19 improved variants, Mu$_1$ (*Bm*Tyr^{R209S}), Mu$_2$ (*Bm*Tyr^{V217L}), and Mu$_3$ (*Bm*Tyr^{V218A}) showed an almost two-fold increase in TFDG titer; Mu$_3$, the most efficient mutant, showed a 2.43-fold increase in TFDG titer (438.19 mg/L) compared with that of WT. To further increase this activity, R209S and V217L were combined with V218A, resulting in the variants Mu$_4$ (*Bm*Tyr$^{V218A/R209S}$) and Mu$_5$ (*Bm*Tyr$^{V218A/V217L}$). Mu$_4$ showed the highest TFDG titer (675.30 mg/L), corresponding to a 4.29-fold increase compared with that of the WT (Figure 4b). The kinetic parameters of WT and its variants are shown in Table 1. Compared with WT, Mu$_4$ showed increased activity toward ECG and EGCG. The specific activities of

Mu$_4$ toward ECG and EGCG, which were 51.14 and 86.58 U/mg, respectively, were 3.68- and 4.67-fold higher than that of the WT. The V$_{max}$/K_m of Mu$_4$ toward ECG and EGCG were 14.59 U and 27.17 U min^{-1} mg^{-1} mM^{-1}, respectively, which were 6.05- and 6.97-fold higher than that of the WT.
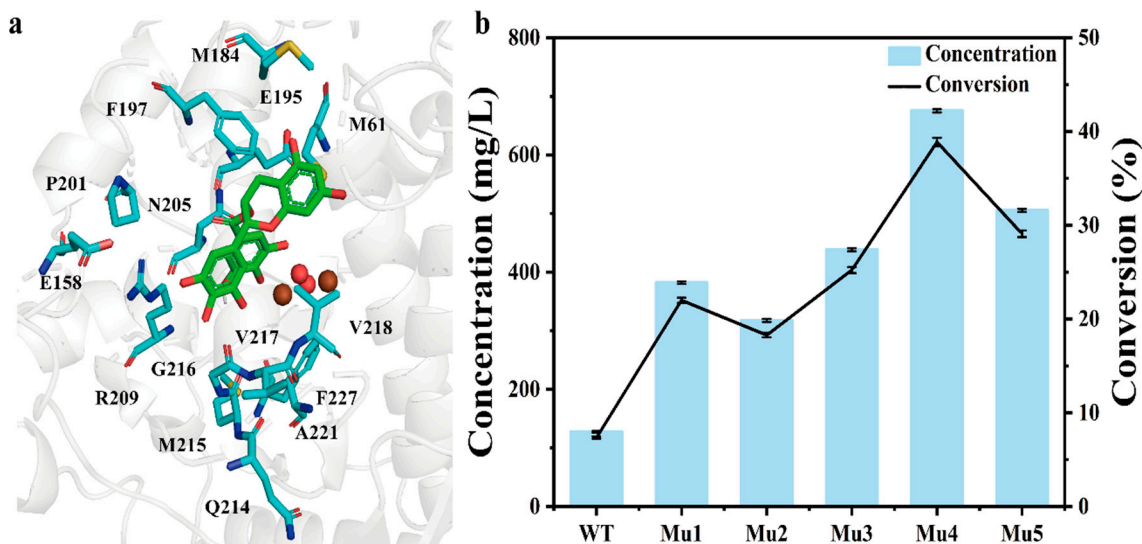

Figure 4. Redesign of *Bm*Tyr. (**a**) Residues (cyan sticks) within 5 Å from the substrate (green sticks) in the *Bm*Tyr. (**b**) The activity of the designed enzymes on TFDG production (5 mM substrate in 200 mM Mcilvaine Buffer pH 5.0, 30 °C, 0.1 mM Cu^{2+}). The data represent mean ± s.d., as determined from three independent experiments.

Table 1. Kinetic constants of purified *Bm*Tyr and its variants.

Enzyme	Substrate	Specific Activity (U mg^{-1})	V$_{max}$ (U min^{-1} mg^{-1})	K_m (mM)	V$_{max}$/K_m (U min^{-1} mg^{-1} mM^{-1})	E$_0$ (J/mol k)
WT	ECG	10.93 ± 0.82	123.24 ± 1.19	59.62 ± 0.91	2.07 ± 0.05	1.28×10^{-5}
	EGCG	15.26 ± 0.78	35.26 ± 1.25	10.35 ± 0.36	3.41 ± 0.03	1.25×10^{-5}
Mu$_1$ (R209S)	ECG	16.57 ± 0.96	146.26 ± 1.53	43.87 ± 0.12	3.33 ± 0.83	1.29×10^{-5}
	EGCG	28.63 ± 0.96	116.59 ± 0.95	18.62 ± 0.48	6.26 ± 0.65	1.29×10^{-5}
Mu$_2$ (V217L)	ECG	31.23 ± 0.89	174.53 ± 1.69	28.51 ± 0.48	6.12 ± 0.51	1.29×10^{-5}
	EGCG	26.04 ± 0.55	145.89 ± 1.04	26.58 ± 0.37	5.48 ± 0.46	1.29×10^{-5}
Mu$_3$ (V218A)	ECG	27.66 ± 0.78	154.69 ± 1.68	36.75 ± 1.27	4.20 ± 0.41	1.29×10^{-5}
	EGCG	33.78 ± 0.78	165.62 ± 0.98	28.51 ± 0.19	6.51 ± 0.63	1.29×10^{-5}
Mu$_4$ (V218A/R209S)	ECG	51.14 ± 1.05	474.53 ± 10.56	32.51 ± 4.54	14.59 ± 0.64	1.31×10^{-5}
	EGCG	86.58 ± 0.95	1265.62 ± 13.32	46.58 ± 3.15	27.17 ± 0.88	1.34×10^{-5}
Mu$_5$ (V218A/V217L)	ECG	57.58 ± 1.31	524.69 ± 9.32	33.75 ± 2.56	15.54 ± 0.93	1.32×10^{-5}
	EGCG	73.16 ± 1.45	865.62 ± 13.46	45.41 ± 4.23	19.06 ± 0.67	1.33×10^{-5}

The data represent mean ± s.d., as determined from three independent experiments.

To elucidate the mechanism underlying the increased enzymatic activity, MD simulations were performed on the Mu$_4$–EGCG complex. Compared with *Bm*Tyr–EGCG, the hydride transfer distance of the Mu$_4$–EGCG complex was shortened from 4.66 ± 0.57 to 3.02 ± 0.31 Å (Figures 5a,b and S8). This suggests that the distance between the -O11H of the EGCG and the O1 of the bridging oxygen atom of the active center was shortened in Mu$_4$. Furthermore, V218A and R209S eliminated steric hindrance (Figure 5c,d). The side chain residues of valine and arginine in Mu$_4$ were larger than that of the WT, which increased the steric hindrance of the transition state. Providing more space for the substrate, where the catalytic residues are present, is conducive to the stability of the transition state, which may enable the release of the products.
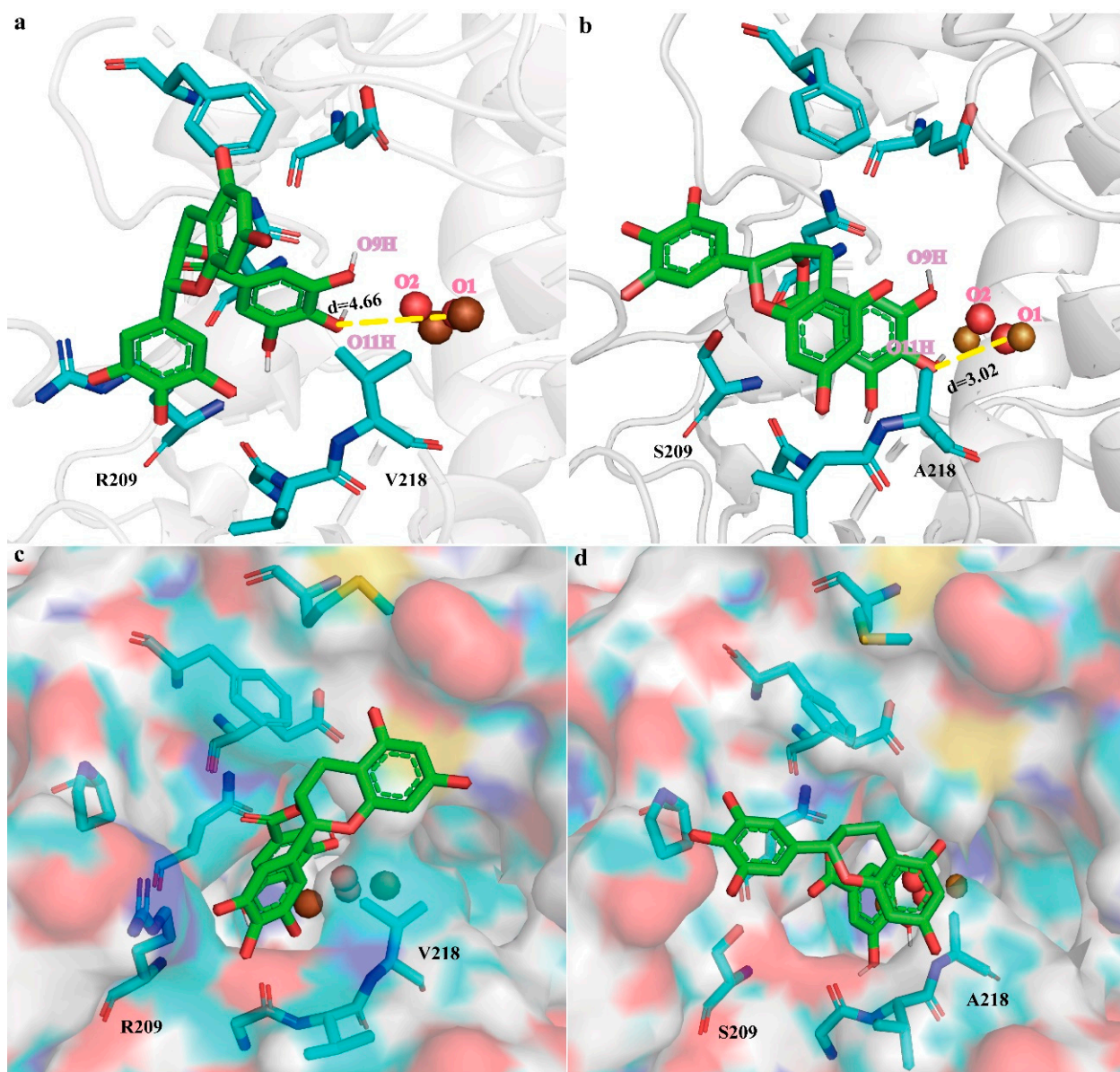
Figure 5. MD simulations on the *Bm*Tyr-EGCG and Mu$_4$-EGCG complexes. Representative MD snapshots of the *Bm*Tyr-EGCG (**a**) and Mu$_4$-EGCG (**b**) complexes. Binding site of WT with bound EGCG (**c**) and Mu$_4$ with EGCG (**d**). Residues are shown in the cartoon and surface, respectively. The carbon of EGCG is shown as a green stick, and the remaining residues are shown in cyan.

2.4. Optimize the Transformation System to Produce TFDG

Purified Mu$_4$ was used as the catalyst to produce TFDG. To increase the TFDG titer by Mu$_4$, the reaction parameters that affect TFDG synthesis were optimized (Figure 6a,d), including pH (3.0–7.0), EGCG/ECG molar ratio (1:1, 1:2, 1:3, 2:1, and 3:1), reaction temperature (15–35 °C), and reaction time (10–60 min). The optimum conditions for the efficient synthesis of TFDG by Mu$_4$ were as follows: Mcilvaine Buffer pH 4.0, 2:1 molar ratio of EGCG/ECG, temperature at 25 °C, and a reaction time of 30 min. The highest TFDG titer was 960.36 mg/L in 30 min, which was 6.35-fold higher than that of WT (127.62 mg/L).
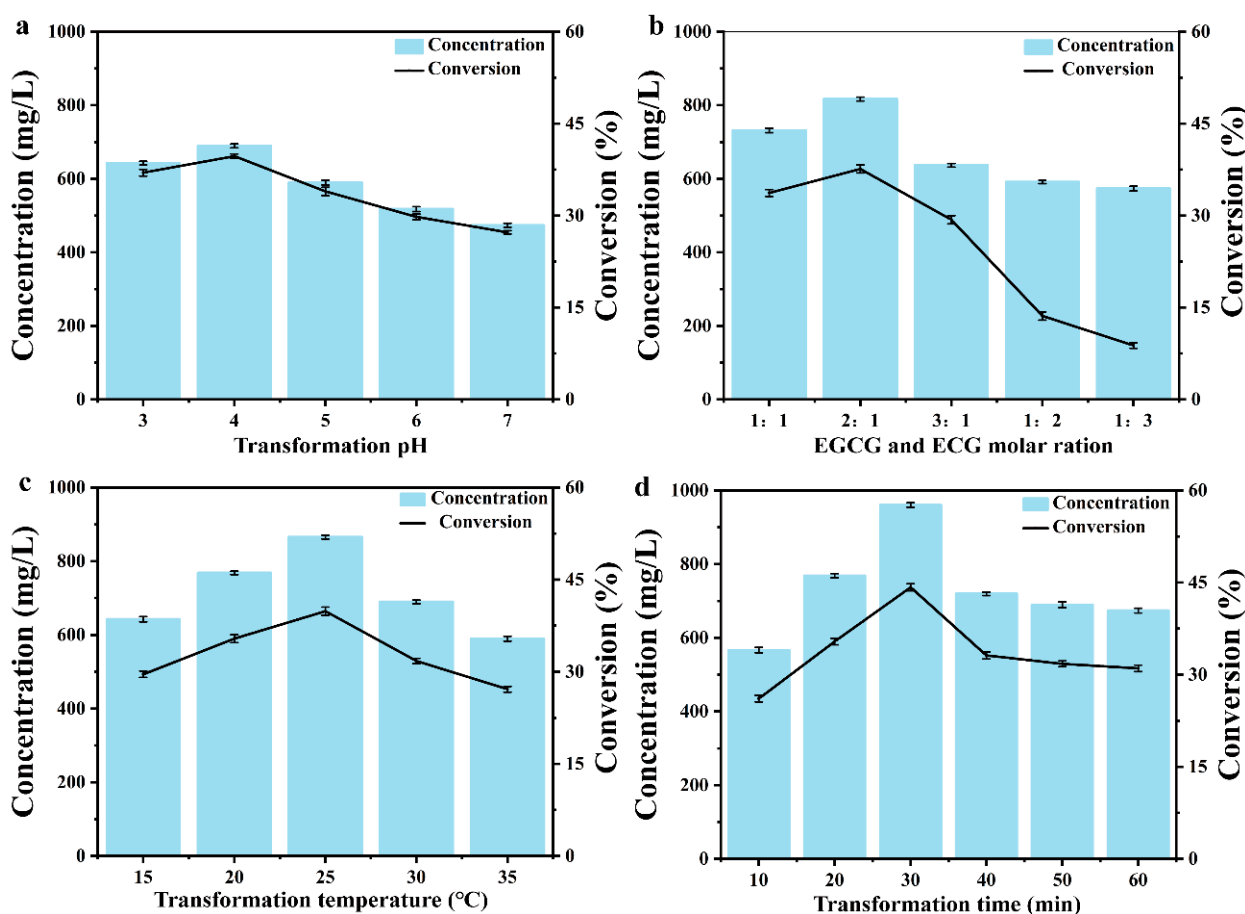
Figure 6. Optimization of TFDG production conditions. (**a**) Effect of buffer pH on TFDG production. (**b**) Effect of EGCG and ECG ration on TFDG production. (**c**) Effect of temperature on TFDG production. (**d**) Effect of time on TFDG production. This system contains 200 mM MacIlvaine buffer, 0.1 mM Cu^{2+} and 0.06 mM pure enzymes with 200 rpm. The data represent mean $\pm$ s.d., as determined from three independent experiments.

3. Discussion and Conclusions

In this study, the efficacy of PPO from different sources in TFDG synthesis was compared, and *Bm*Tyr, with a high conversion rate and wide application value, was screened. To improve the conversion of TFDG, we performed mechanism analysis and computational design to optimize the TS2 conformation. TFDG was effectively synthesized using ECG and EGCG as substrates and the optimum mutants as biocatalysts, increasing the conversion of TFDG by 6.35-fold. Therefore, the established method has practical value for TFDG synthesis and great potential for the synthesis of other TFs.

A stable enzyme source-supporting method, using microbiological fermentation, was selected. To date, most enzyme sources related to in vitro catalytic synthesis of TFDG were derived from plant extracts, which are mainly concentrated in natural PPOs that are directly isolated from tea leaves, mushrooms, or pear fruits [13,24]. Plant PPOs are not conducive to the industrialization of tea pigment production because it requires a large number of tea leaves to meet the application requirements. Furthermore, there were large differences in the enzyme composition of plants in different seasons, and their effects on the synthesis of TFs were inconsistent. The large-scale production of TFs requires a stable enzyme source. We selected seven functionally known microbial PPOs from the NCBI database via data mining. The recombinant strain *Bm*Tyr was selected to improve the biological preparation of TFDG because it has the advantages of a simple culture method, abundant raw material sources, high enzyme yield, and good enzyme production stability.

A transition-state conformation optimization strategy was adopted, which provided an effective tool for obtaining beneficial mutations in *Bm*Tyr to improve TF synthesis. Many strategies have been adopted to engineer proteins and to improve their properties. For example, Zhou et al. [25] obtained activating mutations from *Bm*Tyr after screening almost 20,000 colonies for random mutations. This approach required enormous screening efforts. In this study, *Bm*Tyr was efficiently engineered on the basis of the Rosetta enzyme design for the computer simulation of stable TS to improve oxidation efficiency. A total of 19 beneficial mutants were obtained from the designed library (22 mutants) that produced TFDG. Compared with the traditional random mutation, this strategy effectively reduces the workload. In a different study, theaflavins were synthesized by extracting polyphenol oxidase from eggplants. The results showed that the total content of theaflavins could reach 40.7 mg under optimum reaction conditions [26]. Ngure, F. M. et al. [27] used PPO from tea to oxidize catechins, and the final theaflavin content reached 26.99 μmol/g. PPOs from snow lotus fruits have also been used to oxidize catechins to form theaflavins. The final estimate of theaflavins in the reaction system was 65 μM dry weight/h [28]. Under the catalysis of the optimum mutant *Bm*Tyr$^{V218A/R209S}$, the titer of TFDG reached 960.36 mg/L with 44.22% conversion, which was 6.35-fold higher than that of the WT. These results showed the effectiveness of our strategy and may be useful for improving the properties of other enzymes.

Despite the abovementioned results, several factors restrict the production of TFs. Although the TFDG generation rate is high, the industrial application of free enzymes is impeded because they are either inactivated or cannot be recovered for reuse. Immobilized enzymes can effectively overcome the limitations of free enzymes. A suitable carrier should be found to immobilize PPO, which can make the enzyme reaction continuous and automatic, reduce production costs, and improve efficiency. Our ongoing study will address these issues and improve the utilization of *Bm*Tyr.

4. Materials and Methods

4.1. Strains, Plasmids, and Chemicals

All plasmids and strains are listed in Table S1, and their genetic constructions are shown in the supporting information. Commercial solvents, standards, and reagents were purchased from Meryer Chemicals (Shanghai, China), Sigma Aldrich (Shanghai, China), J&K Chemicals (Beijing, China), TCI Chemicals (Shanghai, China), and Aladdin Reagents (Shanghai, China) and used without further purification.

4.2. Access Codes

Biosynthetic protein accession codes: *Bm*, ACC86108.1, *Ec*1, AYB96806.1, *Ec*2, AAC75642.1, *Bs*1, MG571271.1, *Bs*2, KY849863.1, *An*1, EHA19536.1, *An*2, CAK40046.1.

4.3. Culture and Purification

All desired purified proteins and mutants were obtained under the following operating conditions. Strains were first grown in LB medium with suitable antibiotics at 37 °C and 200 rpm. When the OD$_{600}$ of the culture broth reached 0.6–0.8, IPTG was added to make a final concentration of 0.1 mg/mL. The cells were inducted at 16 °C for 16 h, collected by centrifugation (5000 rpm, 10 min), suspended in lysis buffer (20 mM Tris-HCl (pH 8.0), 20 mM imidazole, 500 mM NaCl) and then disrupted using a high-pressure homogenizer. The resulting crude lysate was centrifuged at 10,000 rpm at 4 °C for 30 min, and the protein containing a His-tag was trapped on Ni-NTA Superflow resin for 30 min. Subsequently, the elution buffer (20 mM Tris-HCl (pH 8.0), 0.3 M NaCl, 0.5 M imidazole) was used to release the protein. The protein concentration was determined using a NanoDrop 2000c spectrophotometer (Thermo Scientific). The purity of the proteins was determined by SDS-PAGE.

4.4. Construction and Screening

The seven samples of PPOs obtained by genome mining were overexpressed in *E. coli* via pET28a and purified to obtain pure enzymes that were used to catalyze the conversion of ECG and EGCG to TFDG. This 2 mL screening system contained 0.2 M MacIlvaine buffer, 5 mM substrate, 0.1 mM Cu^{2+}, and 0.06 mM pure enzymes at pH 5.0, 30 °C, and 200 rpm.

4.5. HPLC Analysis

The titer of TFDG was determined using high-performance liquid chromatography (HPLC). The HPLC analytical conditions were as follows: XBridge®-C18 column (5 μm, 4.6×250 mm), column temperature of 30 °C, detection wavelength of 280 nm, injection quantity of 10 μL, and the flow rate was 1.0 mL/min. Chromatography was performed with an elution of 30% A for 5 min, 34% A for 3 min, and finally, 30% A for 2 min, where mobile phase A was 100% acetonitrile and mobile phase B was 2% acetic acid aqueous solution.

The conversion was calculated as follows:

$$\text{Conversion of TFDG (\%)} = (a/b) \times 100$$

where a is the amount of TFDG synthesized, and b is the amount of ECG added to the system. The conversion values of the reaction catalyzed for TFDG synthesis were calculated.

4.6. Protein Crystallization and Structure Determination

Crystals of apo-*Bm*Tyr were obtained after 3 days at 20 °C in 96-well plates, in which 0.8 μL protein solution (25 mg/mL) and an equal volume of reservoir solution (0.1 M zinc acetate dehydrate, 0.1 M sodium acetate trihydrate; pH 4.6; 12% *w/v* PEG4000) were mixed, as per the optimized conditions for crystallization. All crystals were flash frozen with 25% glycerol as an antifreezing solution in liquid nitrogen. X-ray diffraction data were collected using a Bruker D8 QUEST diffractometer, and the data sets were indexed, integrated, and merged using XDS [29]. The structure of *Bm*Tyr was solved using the molecular replacement, with the protein structure of 3nm8 as the search model. Iterative cycles of model building and refinement were performed in Coot [30] and Refmac [31] programs. The data collection and refinement statistics of the apo-*Bm*Tyr crystal structure are listed in Table S4. Structural figures were prepared using PyMOL v2.3.3 (Schrödinger, LLC, New York, NY, USA).

4.7. Initial Structural Preparation

The initial structure of *Bm*Tyr was obtained using *Bacillus megaterium* as a template (PDB ID:3nm8 [32]). Based on 3nm8, a reactive binding mode, namely *Bm*Tyr (M1), was created. The protonation states of the charged residues were determined at a constant pH of 5.0 on the basis of the pKa calculations using H^{++} [33] while taking into account the local hydrogen bonding network. The His residues were assigned as HID. Asp and Glu residues were deprotonated, whereas Lys and Arg were protonated. The prepared protein was neutralized by adding Na^+ ions and solvated into a truncated octahedral TIP3P [34] water box with a 10 Å buffer distance on each side. The M1 system consisted of 27,473 atoms. The system was equilibrated through a series of minimization steps that was interspersed by short molecular dynamics (MD) simulations, during which the restraints on the protein backbone heavy atoms were released gradually (with a force constant of 10, 2, 0.1, and 0 kcal/mol·$Å^2$). The system was gradually heated up to 303 K in 50 ps in which harmonic potentials were used to positionally restrain the protein backbone heavy atoms (with a force constant of 10 kcal/mol·$Å^2$). Finally, the standard unrestrained MD simulation with periodic boundary condition at 303 K and 1 atm was carried out for up to 100 ns. Based on the dynamically stable *Bm*Tyr (M2) binary complex structure determined by MD simulations, EGCG was docked into the active center to obtain the possible binding mode of EGCG (*Bm*Tyr–EGCG, M3). A representative pre-equilibrated structure of *Bm*Ty–EGCG was used as a template to create the starting coordinates of Mu_4–EGCG by changing the

two mutated residues (V218A/R209S). To check for statistical convergence, each simulation was carried out in triplicate with different initial atomic velocities that were randomly generated according to the Maxwell–Boltzmann distribution.

4.8. Molecular Docking

In the equilibrium trajectories of the M2 binary complex, 2000 snapshots were selected and divided into ten groups through a hierarchical agglomerative (bottom-up) approach [35]. EGCG was fully optimized at the B3LYP/6-31G(d) level using the Gaussian 16 [36] package, and docked into the active site of each group representative snapshot. Molecular docking was performed by the Lamarckian genetic algorithm local search method in AutoDock 4.2 and AutoDockTools-1.5.6 [37]. The docking approach was employed on a rigid receptor conformation. One-hundred independent docking runs were undertaken. Reasonable conformations were then selected as the binding conformations for MD simulations.

4.9. MD Simulations

The pre-equilibrated M3 structures and possible catalytically reactive binding modes of EGCG were used as the starting conformations for the MD simulations of the protein–ligand complexes. The partial charges of EGCG were fitted with HF/6-31G(d) calculations, and the restrained electrostatic potential (RESP) [38,39] protocol was implemented by the Antechamber module in the Amber 16 package [40]. Each system was initially neutralized with Na^+ counter ions and solvated with TIP3P [34] water in a truncated octahedron box at a 10 Å buffer distance. The resulting systems M2, M3, and Mu_4–EGCG contained 27,473, 27,188, and 27,230 atoms, respectively. Next, each system was equilibrated with a series of minimization that was interspersed by short MD simulations, during which the restraints on the protein backbone heavy atoms were released gradually (with force constants of 10, 2, 0.1, and 0 kcal/mol·$Å^2$). The system was then gradually heated from 0 to 303 K for 50 ps, in which a 10 kcal/(mol·$Å^2$) restraint was applied to heavy atoms on the protein backbone. Finally, an extensive MD simulation of 50 ns was performed at constant temperature and pressure. The pressure was maintained at 1 atm and coupled with isotropic position scaling. The temperature was controlled at 303 K using the Berendsen thermostat method [41]. Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) method [42], and a 12 Å cutoff was applied to both PME and van der Waals (vdW) interactions. A time step of 2 fs was employed along with the SHAKE algorithm [43] for hydrogen, and periodic boundary conditions were used. The atomic positions were stored every 2 ps for further analysis. Each system was checked for stability (structure, energy, and temperature fluctuations) and convergence (root mean square deviations (RMSD) of the structures).

4.10. Rosetta Design

Active-site positions were selected for design, and at each position, position-specific scoring matrices (PSSMs) were used to constrain amino acid choices to identities. We generated a sequence alignment, and the less favorable amino acids were removed by natural selection. We then used Rosetta computational mutation scanning. Each "allowed" mutation was modeled in the context of the WT structure, and the energy difference ($\Delta\Delta G$) between the WT and the mutant was calculated. We defined the space of "conformation optimization" mutations as mutations with $\Delta\Delta G \leq 0$ Rosetta energy units (R.e.u.), generating many designs for experimental testing.

4.11. Directed Evolution Experiments

Directed evolution experiments were performed by whole-plasmid polymerase chain reaction (PCR) using KOD-Plus-Neo (TOYOBO). The primers used are listed in Table S2. The resulting PCR products were digested with D*pn*I to remove the template plasmid. Furthermore, 10 μL of digested products was transformed into *E. coli* BL21 (DE3) cells

for subsequent screening or DNA sequencing (GENEWIZ, China). These mutants were cultured into LB medium and shaken at 37 °C for 12 h. Mutants were then diluted 1:10 into 150 mL fresh Terrific Broth (TB) medium. After shaking at 37 °C for 3 h (for cell growth), the temperature was decreased to 25 °C for 16 h (for protein expression). Cells were then collected by centrifugation.

4.12. Activity Assay

The activities of *Bm*Tyr and the mutants were spectrophotometrically measured using ECG and EGCG as substrates [44,45]. The activity was determined by measuring the initial rate of quinone formation by an enzyme-labeled instrument under the following conditions: 5 mM substrate, 0.6 mM enzymatic solution containing 0.1 mM Cu^{2+}, 200 mM Mcilvaine Buffer (pH 5.0) in a final volume of 200 μL, at 200 rpm and 30 °C for 5 min. Finally, the reaction was terminated by adding 50 μL of 2 M trichloroacetic acid and centrifuging the solution at $12,000\times g$ for 5 min. The absorbance of the samples was analyzed at 405 nm, and a reaction mixture without the enzymatic solution was used as a negative control. One unit of enzymatic activity was defined as the amount of enzyme that changes the absorbance by 0.01/min.

4.13. Kinetic Assay

The kinetic parameters of the enzymes were determined by measuring the initial rates of product formation at different concentrations of the substrate (EGCG:0–16 g/L, and ECG:0–8 g/L) for 5 min. Other assay conditions were the same as those described for the corresponding activity assays. The samples were withdrawn, extracted, and analyzed using HPLC. The K_m and V_{max} values were calculated using nonlinear regression according to the Michaelis–Menten equation in the Origin software.

4.14. Optimum Reaction Conditions for TFDG Synthesis

The selection of optimal conditions is crucial for the conversion of TFDG. Therefore, the effects of different pH values (pH 3.0, 4.0, 5.0, 6.0, and 7.0), EGCG/ECG molar ratios (1:1, 1:2, 1:3, 2:1, and 3:1, based on a 2.5 mM substrate ECG), reaction temperatures (15, 20, 25, 30, 35 °C), and reaction times (10, 20, 30, 40, 50, and 60 min) on TFDG synthesis were investigated. Optimum parameters were screened by analyzing the effect of each factor on TF production.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/molecules28093831/s1, Table S1. Strains and plasmids used in this study. Table S2. The primers for gene amplification. Table S3. The primers for the saturated mutation. Table S4. Data collection and refinement statistics. Figure S1. The HPLC chromatograms of TFDG reaction products. Figure S2. The whole structure of *Bm*Tyr (PDB ID:8HPI). Figure S3. The construction process of the whole theoretical complex structure. Figure S4. RMSD and RMSF of *Bm*Tyr^{WT} and *Bm*Tyr$^{V218A/R209S}$. Figure S5. The whole TS2 state model. Figure S6. Multiple sequence alignment of *Bm*Tyr with homologous tyrosinase sequences from superfamily. Figure S7. Redesign of *Bm*Tyr. Figure S8. The distance between the -O11H of the hydroxy oxygen atom of EGCG and the O1 of the bridging oxygen atom of the active center.

Author Contributions: All authors contributed to the study conception and design. Material preparation and methodology, L.L., W.W., Y.H. and W.S.; validation, Y.H., W.W. and W.S.; investigation, Y.H.; data curation, Y.H. and W.S.; writing—original draft preparation, Y.H.; writing—review and editing, L.L., Y.H. and W.S.; supervision, L.L.; funding acquisition, L.L. and J.W; Conceptualization, C.G. (Changzheng Gao), X.C., C.G. (Cong Gao) and J.L. All authors have read and agreed to the published version of the manuscript.

References

1. Shi, J.; Yang, G.; You, Q.; Sun, S.; Chen, R.; Lin, Z.; Simal-Gandara, J.; Lv, H. Updates on the chemistry, processing characteristics, and utilization of tea flavonoids in last two decades (2001–2021). *Crit. Rev. Food Sci. Nutr.* **2021**, 1–28. [CrossRef] [PubMed]
2. Gao, Y.; Rankin, G.O.; Tu, Y.; Chen, Y.C. Inhibitory effects of the four main theaflavin derivatives found in black tea on ovarian cancer cells. *Anticancer Res.* **2016**, *36*, 643–651. [PubMed]
3. Shi, M.; Lu, Y.; Wu, J.; Zheng, Z.; Lv, C.; Ye, J.; Qin, S.; Zeng, C. Beneficial effects of theaflavins on metabolic syndrome: From molecular evidence to gut microbiome. *Int. J. Mol. Sci.* **2022**, *23*, 7595. [CrossRef] [PubMed]
4. Hu, X.; Ping, Z.; Gan, M.; Tao, Y.; Wang, L.; Shi, J.; Wu, X.; Zhang, W.; Yang, H.; Xu, Y. Theaflavin-3,3′-digallate represses osteoclastogenesis and prevents wear debris-induced osteolysis via suppression of ERK pathway. *Acta Biomater.* **2017**, *48*, 479–488. [CrossRef] [PubMed]
5. Cui, X.; Zhou, R.; Huang, C.; Zhang, R.; Wang, J.; Zhang, Y.; Ding, J.; Li, X.; Zhou, J.; Cen, S. Identification of theaflavin-3,3′-digallate as a novel Zika virus protease inhibitor. *Front. Pharmacol.* **2020**, *11*, 514313. [CrossRef] [PubMed]
6. Jang, M.; Park, Y.I.; Cha, Y.E.; Park, R.; Namkoong, S.; Lee, J.I.; Park, J. Tea polyphenols EGCG and theaflavin inhibit the activity of SARS-CoV-2 3CL-protease in vitro. *Evid. Based Complement. Altern. Med.* **2020**, *2020*, 5630838. [CrossRef] [PubMed]
7. Novak, A.J.; Trauner, D. Complex natural products derived from pyrogallols. In *Progress in the Chemistry of Organic Natural Products*; Springer: Cham, Switzerland, 2022; pp. 1–46.
8. Leung, L.K.; Su, Y.; Chen, R.; Zhang, Z.; Huang, Y.; Chen, Z.-Y. Theaflavins in black tea and catechins in green tea are equally effective antioxidants. *J. Nutr.* **2001**, *131*, 2248–2251. [CrossRef] [PubMed]
9. Hua, J.; Wang, H.; Jiang, Y.; Li, J.; Wang, J.; Yuan, H. Influence of enzyme source and catechins on theaflavins formation during in vitro liquid-state fermentation. *LWT Food Sci. Technol.* **2021**, *139*, 110291. [CrossRef]
10. Xu, P.; Hua, D.; Ma, C. Microbial transformation of propenylbenzenes for natural flavour production. *Trends Biotechnol.* **2007**, *25*, 571–576. [CrossRef]
11. Lei, S.; Xie, M.; Hu, B.; Zhou, L.; Sun, Y.; Saeeduddin, M.; Zhang, H.; Zeng, X. Effective synthesis of theaflavin-3,3′-digallate with epigallocatechin-3-O-gallate and epicatechin gallate as substrates by using immobilized pear polyphenol oxidase. *Int. J. Biol. Macromol.* **2017**, *94*, 709–718. [CrossRef]
12. Zhang, G.; Yang, J.; Cui, D.; Zhao, D.; Li, Y.; Wan, X.; Zhao, J. Transcriptome and Metabolic Profiling Unveiled Roles of Peroxidases in Theaflavin Production in Black Tea Processing and Determination of Tea Processing Suitability. *J. Agric. Food Chem.* **2020**, *68*, 3528–3538. [CrossRef] [PubMed]
13. Teng, J.; Gong, Z.; Deng, Y.; Chen, L.; Li, Q.; Shao, Y.; Lin, L.; Xiao, W. Purification, characterization and enzymatic synthesis of theaflavins of polyphenol oxidase isozymes from tea leaf (Camellia sinensis). *LWT Food Sci. Technol.* **2017**, *84*, 263–270. [CrossRef]
14. Sharma, A.; Gupta, G.; Ahmad, T.; Mansoor, S.; Kaur, B. Enzyme engineering: Current trends and future perspectives. *Food Rev. Int.* **2021**, *37*, 121–154. [CrossRef]
15. Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **2018**, *361*, 360–365. [CrossRef]
16. Jiang, L.; Althoff, E.A.; Clemente, F.R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J.L.; Betker, J.L.; Tanaka, F.; Barbas III, C. De novo computational design of retro-aldol enzymes. *Science* **2008**, *319*, 1387–1391. [CrossRef]
17. Polatoğlu, İ.; Karataş, D. Modeling of molecular interaction between catechol and tyrosinase by DFT. *J. Mol. Struct.* **2020**, *1202*, 127192. [CrossRef]
18. Rolff, M.; Schottenheim, J.; Decker, H.; Tuczek, F. Copper–O$_2$ reactivity of tyrosinase models towards external monophenolic substrates: Molecular mechanism and comparison with the enzyme. *Chem. Soc. Rev.* **2011**, *40*, 4077–4098. [CrossRef]
19. Hamann, J.N.; Herzigkeit, B.; Jurgeleit, R.; Tuczek, F. Small-molecule models of tyrosinase: From ligand hydroxylation to catalytic monooxygenation of external substrates. *Coord. Chem. Rev.* **2017**, *334*, 54–66. [CrossRef]
20. Jiang, H.; Lai, W. Monophenolase and catecholase activity of *Aspergillus oryzae* catechol oxidase: Insights from hybrid QM/MM calculations. *Org. Biomol. Chem.* **2020**, *18*, 5192–5202. [CrossRef]

21. Inoue, T.; Shiota, Y.; Yoshizawa, K. Quantum chemical approach to the mechanism for the biological conversion of tyrosine to dopaquinone. *J. Am. Chem. Soc.* **2008**, *130*, 16890–16897. [CrossRef]

22. Koval, I.A.; Gamez, P.; Belle, C.; Selmeczi, K.; Reedijk, J. Synthetic models of the active site of catechol oxidase: Mechanistic studies. *Chem. Soc. Rev.* **2006**, *35*, 814–840. [CrossRef] [PubMed]

23. Altschul, S.F.; Gertz, E.M.; Agarwala, R.; Schäffer, A.A.; Yu, Y.-K. PSI-BLAST pseudocounts and the minimum description length principle. *Nucleic Acids Res.* **2009**, *37*, 815–824. [CrossRef] [PubMed]

24. Mauracher, S.G.; Molitor, C.; Michael, C.; Kragl, M.; Rizzi, A.; Rompel, A. High level protein-purification allows the unambiguous polypeptide determination of latent isoform PPO4 of mushroom tyrosinase. *Phytochemistry* **2014**, *99*, 14–25. [CrossRef] [PubMed]

25. Zhou, J.; Liu, C.; Zhao, S.; Liu, Y.; Zhang, S.; Zhao, Q.; Wang, F.; Xu, G.; Huang, J.; Liu, Z. Improved yield of theaflavin-3,3′-digallate from Bacillus megaterium tyrosinase via directed evolution. *Food Chem.* **2022**, *375*, 131848. [CrossRef] [PubMed]

26. Fang, W.P.; Wang, L.P.; Yu, J.; Yue, P.X.; Jiang, X.; Feng, W.Y.; Chenzhou, Y.Q.; Li, X.H. In Studies on optimum conditions of synthesizing theaflavins by using bio-enzyme method. *Appl. Mech. Mater.* **2012**, *138*, 929–932. [CrossRef]

27. Ngure, F.M.; Wanyoko, J.K.; Mahungu, S.M.; Shitandi, A.A. Catechins depletion patterns in relation to theaflavin and thearubigins formation. *Food Chem.* **2009**, *115*, 8–14. [CrossRef]

28. Van der Westhuizen, M.; Steenkamp, L.; Steenkamp, P.; Apostolides, Z. Alternative pathway implicated as an influencing factor in the synthesis of theaflavin. *Biocatal. Biotransform* **2015**, *33*, 298–309. [CrossRef]

29. Kabsch, W. Integration, scaling, space-group assignment and post refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 133–144. [CrossRef]

30. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [CrossRef]

31. Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 355–367. [CrossRef]

32. Sendovski, M.; Kanteev, M.; Ben-Yosef, V.S.; Adir, N.; Fishman, A. First structures of an active bacterial tyrosinase reveal copper plasticity. *J. Mol. Biol.* **2011**, *405*, 227–237. [CrossRef] [PubMed]

33. Gordon, J.C.; Myers, J.B.; Folta, T.; Shoja, V.; Heath, L.S.; Onufriev, A. H++: A server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, *33* (Suppl. S2), W368–W371. [CrossRef] [PubMed]

34. 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.* **1983**, *79*, 926–935. [CrossRef]

35. Ward, J.H., Jr. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [CrossRef]

36. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16 Rev. A.03*; Gaussian Inc.: Wallingford, CT, USA, 2016.

37. Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662. [CrossRef]

38. Jakalian, A.; Bush, B.L.; Jack, D.B.; Bayly, C.I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132–146. [CrossRef]

39. Jakalian, A.; Jack, D.B.; Bayly, C.I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641. [CrossRef]

40. Case, D.A.; Betz, R.M.; Cerutti, D.S.; Cheatham, T.E., III; Darden, T.A.; Duke, R.E.; Giese, T.J.; Gohlke, H.; Götz, A.W.; Homeyer, N.; et al. *Amber 16*; University of California: San Francisco, CA, USA, 2016.

41. Berendsen, H.J.; Postma, J.V.; Van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690. [CrossRef]

42. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092. [CrossRef]

43. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [CrossRef]

44. Garcia-Molina, F.; Muñoz, J.L.; Varón, R.; Rodriguez-Lopez, J.N.; García-Cánovas, F.; Tudela, J. A review on spectrophotometric methods for measuring the monophenolase and diphenolase activities of tyrosinase. *J. Sci. Food Agric.* **2007**, *55*, 9739–9749. [CrossRef] [PubMed]

45. Narai-Kanayama, A.; Kawashima, A.; Uchida, Y.; Kawamura, M.; Nakayama, T. Specificity of tyrosinase-catalyzed synthesis of theaflavins. *J. Mol. Catal. B Enzym.* **2016**, *133*, S452–S458. [CrossRef]