November 2023 Vol.66 No.11: 3232–3241
https://doi.org/10.1007/s11426-023-1755-4

Xinying Ge[1†], Yan Long[1†], Jun Wang[2†], Bo Gu[1], Zixuan Yang[1], Yinyin Feng[1], Shuo Zheng[1], Yingying Li[1], Wupeng Yan[2*] & Heng Song[1*]

[1]*College of Chemistry and Molecular Sciences, Wuhan University, Wuhan 430072, China;*
[2]*School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China*

Received June 14, 2023; accepted August 14, 2023; published online September 19, 2023

Indole alkaloids have attracted considerable attention from synthetic chemists and biochemists for their structural diversity and important biological activities. Compared with traditional organic synthesis methods, the strategy of using cytochrome P450s' extraordinary abilities to selectively activate carbon-hydrogen bonds to assist in the synthesis of various indole alkaloids has the characteristics of short synthetic route, mild conditions and high atomic economy. Here, we utilized P450 monooxygenases HinD and TleB to synthesize a novel 6/5/8 tricyclic product from (*S*)-*N*-((*S*)-1-(4-fluoro-1*H*-indol-3-yl)-3-hydroxypropan-2-yl)-2-mercapto-3-methylbutanamide through the substrate structure-directed strategy. TleB was more effective in catalyzing C–S coupling, and was used to synthesize a series of 6/5/8 tricyclic indololactam derivatives to provide drug candidates. Interestingly, the S–S coupling product was observed in HinD catalysis, which was a minor product in the wild-type TleB catalysis. With the help of protein engineering, we accurately regulated the catalytic flow and reversed the selectivity of TleB to obtain the S–S coupling product. At the same time, the reaction mechanism was reasonably speculated by means of site blocking and protein-substrate complex analysis.

Ge X, Long Y, Wang J, Gu B, Yang Z, Feng Y, Zheng S, Li Y, Yan W, Song H. Regulation of P450 TleB catalytic flow for the synthesis of sulfur-containing indole alkaloids by substrate structure-directed strategy and protein engineering. *Sci China Chem*, 2023, 66: 3232–3241, https://doi.org/10.1007/s11426-023-1755-4

Natural products are the treasure house of drug discoveries. Mining new natural products and analogues and developing new synthetic routes of natural products are of great significance in the field of new drug research [1–4]. However, since the discovery of natural products with new structures is laborious, and the synthetic routes of natural products with complex structures are difficult to exploit, the development of new drugs has slowed down. In synthetic chemistry, site-selective activation of the C–H bond is one of the most ideal synthesis methods. Nevertheless, it is not easy to selectively activate a C–H bond of a complex substrate in traditional organic synthesis [5–7]. Fortunately, in biocatalysis, cytochrome P450s (P450s) are highly selective C–H activation catalysts, which are involved in the biosynthetic pathways of many natural products [8–10]. Therefore, incorporating P450 enzyme-catalyzed reactions into the synthetic routes of natural products can greatly reduce the synthesis steps and improve the selectivity. Although enzyme specificity makes the major product structurally specific, it is worth mention-

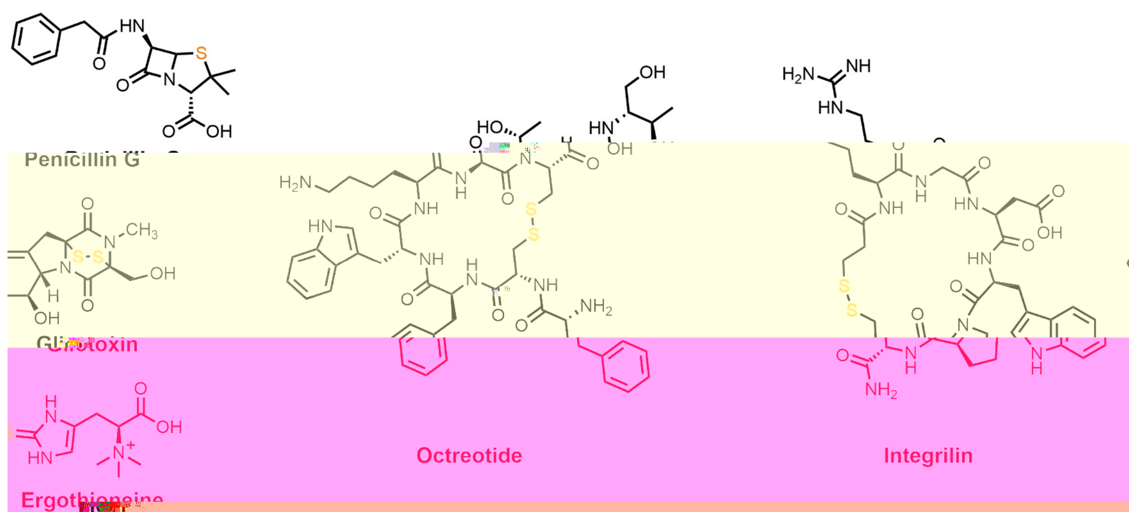
†These authors contributed equally to this work.
*Corresponding authors (email: yanwupeng@sjtu.edu.cn; hengsong@whu.edu.cn)

ing that in enzyme-catalyzed reactions, there are sometimes other minor by-products with novel structures that exhibit good biological activities [11–13]. Exploring these structurally new compounds is of great significance for the development of new drugs. Regulating enzyme catalytic selectivity through reasonable methods to make minor products into major products can improve the potential of enzyme catalysis in synthetic chemistry [14,15].

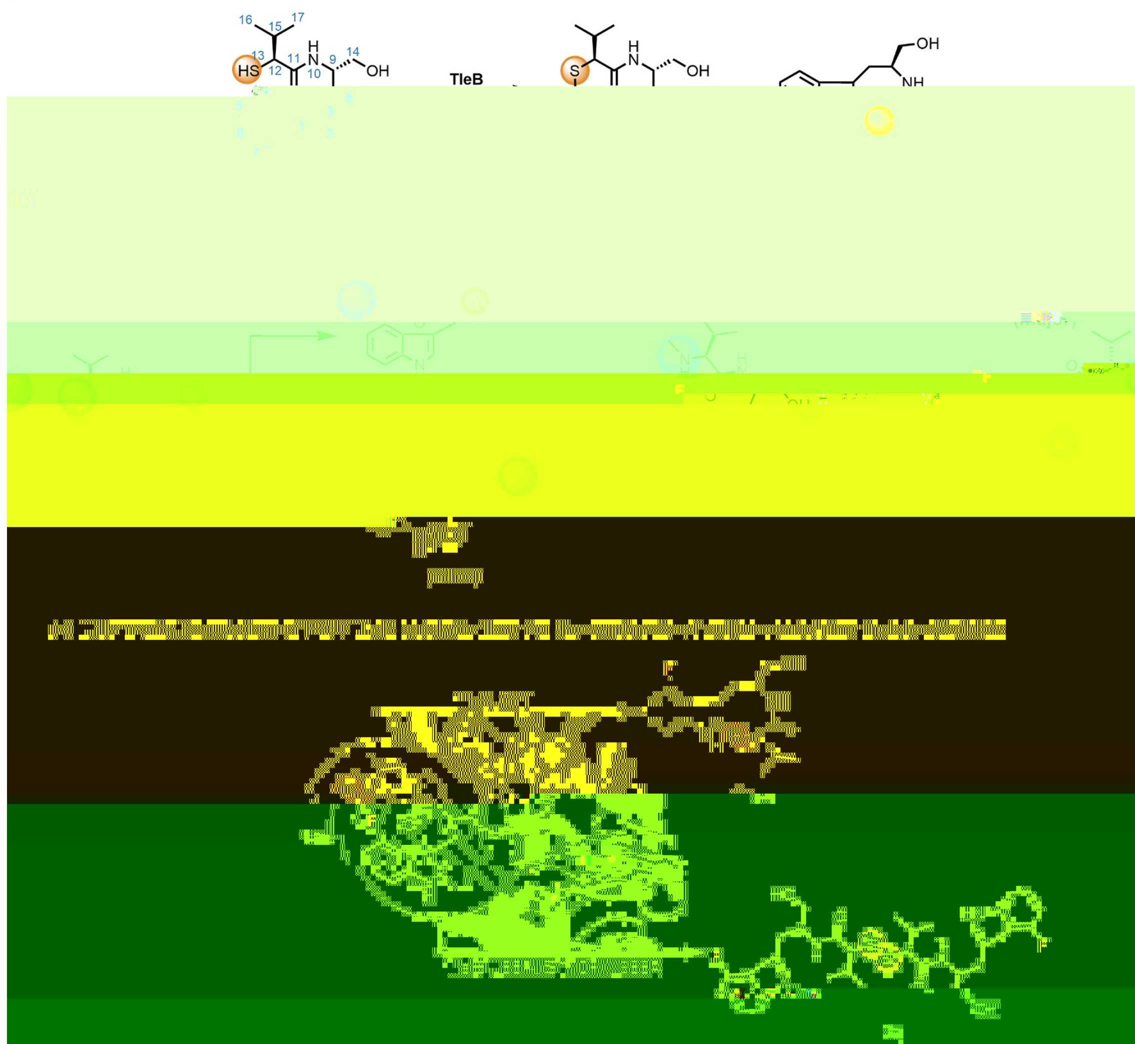
Many natural products containing sulfur element have significant biological activities, such as various penicillins with antibacterial effects [16], gliotoxin [17], and ergothioneine with antioxidant effects [18] (Figure 1). Reactions such as those forming C–S and S–S bonds exist in the biosynthetic pathways of many natural products and have received extensive attention. At present, there are relatively few reports on mechanisms of C–S or S–S formation that are associated with P450 catalysis in biosynthesis. However, P450s have the ability to activate and initiate the formation of S–X bonds, which has important potential for the development of biocatalysts to generate sulfur-containing compounds. For example, cytochrome P450 monooxygenase CxnD catalyzes thioheterocyclization in the biosynthesis pathway of chuangxinmycin [19]. TleB, another P450 monooxygenase, has been reported to catalyze the intramolecular C–S coupling reaction, resulting in a sulfur-containing indolactam derivative and an unusual indole-fused 6/5/8 tricyclic product [20] (Scheme 1a). This tricyclic indololactam derivative has a novel structure and is similar to a class of protein kinase D (PKD) inhibitors [21], so it may have potential anti-tumor activity. It is a challenging task to obtain such a 6/5/8 tricyclic product with two chiral centers by chemical synthesis. Enzymatic catalysis would be a good choice to produce this product. However, when this product is generated by TleB, a sulfur-containing indolactam derivative is also generated. In order to improve the efficiency and the atomic economy of the target product 6/5/8 tricyclic product, we can specifically generate 6/5/8 tricyclic product by regulating the biocatalytic reaction.

At present, the regulation of biocatalytic reactions mainly focuses on the engineering of enzyme-active cavities. Protein engineering can regulate the compatibility between enzymes and target substrate analogues at the molecular scale, thereby improving the efficiency of enzymatic reactions [22–25]. In traditional organic synthesis, a novel reaction is usually achieved by modifying the reactants to change the selectivity [26–28]. Combining the advantages of the above two methods will provide a good opportunity to expand the type of enzymatic reaction. In this strategy, the chemical selectivity of the reaction is regulated by rationally designing the substrate structure, and protein engineering is used to increase the compatibility of the enzyme and the substrate. In our previous work, we designed a series of F-substituted *N*-methyl-L-valyl-L-tryptophanol (NMVT) analogues, which successfully reversed the catalytic flow of P450 enzyme HinD and efficiently synthesized indole-fused 6/5/6 tricyclic products [15]. In this reaction, HinD catalyzed the C–O coupling of NMVT and its analogues to form 6/5/6 tricyclic products through hydroxyl nucleophilic attack (Scheme 1b). Therefore, in our work, we continue to use cytochrome P450 to catalyze sulfur-containing NMVT substrate analogues as template reactions and hope to use the strategy of regulating the reaction catalytic flow to introduce the F atom into the indole C4 position of the 13-SH substituted substrate analogue, so that the reaction catalytic flow is carried out in the direction of generating indole-fused 6/5/8 indolactam derivatives.

In this work, the nucleophilic nitrogen methyl in 4-F-NMVT was replaced by a more nucleophilic thiol group to obtain the substrate (*S*)-*N*-((*S*)-1-(4-fluoro-1*H*-indol-3-yl)-3-hydroxypropan-2-yl)-2-mercapto-3-methylbutanamide ( ).



Sulfur-containing bioactive molecules (color online).

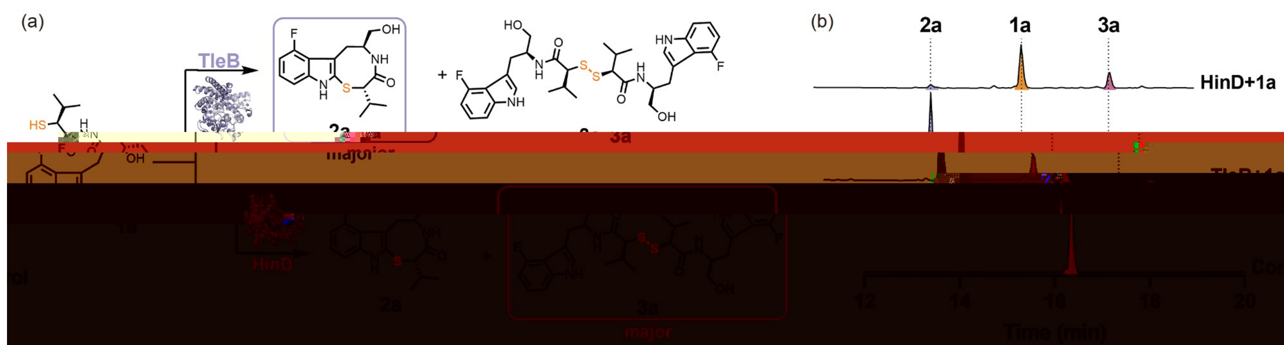


(a) Two C–S coupling products catalyzed by P450 TleB. (b) Regulation of P450 HinD catalytic flow. The original minor product, indole-fused 6/5/6 tricyclic compound, was manipulated to be the major product in HinD catalysis by an interception and inversion strategy. (c) Regulation of P450 TleB catalytic flow for the synthesis of sulfur-containing indole alkaloids (color online).

This further regulated the selectivity of HinD catalysis to obtain a 6/5/8 tricyclic C–S coupling product ( ) (Scheme 1c). At the same time, TleB, the homologous protein of HinD was found to be more effective in catalyzing C–S coupling, and could efficiently convert and its analogues into 6/5/8 tricyclic products. During the study, the catalytic activity of the P450 enzyme for S–S coupling was unexpectedly observed. In view of the important role of peptide drugs containing disulfide bonds [29,30] (Figure 1), TleB was selected as the research object, and site-directed mutagenesis was carried out to efficiently catalyze the formation of the thiol dimerization product ( ) (Scheme 1c). In addition, we obtained the crystal of TleB- complex and solved its complex structure to analyze substrate interactions with protein active sites. Finally, the formation mechanism of C–S coupling products and S–S coupling products was reasonably speculated. This work demonstrates the feasibility of combining substrate structure design with protein engineering. This strategy enriches the existing means of regulating enzyme catalytic selectivity, provides the possibility to further expand the type of enzyme catalytic reaction, and promotes the wider application of enzyme catalytic reaction in the field of synthetic chemistry.

All chemical reagents were purchased from commercial suppliers (Energy Chemical, Heowns, Innochem, Sigma Aldrich, Amatek Scientific, *etc.*) if not otherwise stated and were used without further purification. HinD, TleB and CamAB plasmids were purchased from GENEWIZ, Inc.

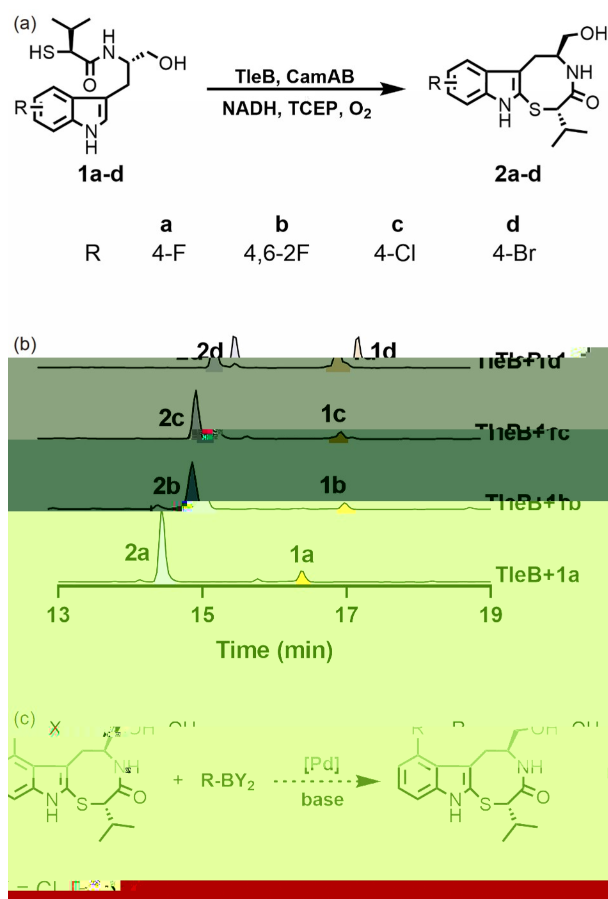
E. coli BL21(DE3) cells transformed with the expression plasmid for CamAB protein were inoculated into 10 mL



Enzymatic reaction of HinD and TleB with . (a) Reaction schemes for the conversions of with HinD and TleB. (b) HPLC profiles of the enzymatic reaction. The chromatograms were monitored at 290 nm (color online).

of the fusion protein CamAB as the P450 redox partner, 10 mM NADH, 0.5 mM TCEP hydrochloride, and 125 µg/mL catalase. The reactions were monitored by HPLC at 290 nm. Compared with the control, was converted into two new products with different proportions (Figure 2b), indicating a significant difference in the catalytic activity of homologous proteins HinD and TleB for . It was found that TleB could catalyze the C–S coupling of to form a new thioindole lactam derivative (m/z $[M+H]^+$ = 323.1222). To confirm the structure of product , it was isolated from the reaction mixture and characterized using proton nuclear magnetic resonance (^{1}H NMR), ^{13}C NMR and high-resolution mass spectrometry (HRMS) (Figures S2–S4). Interestingly, was a minor product in HinD catalysis while the main product catalyzed by HinD was speculated to be the S–S coupling product (m/z $[M+H]^+$ = 647.2500) according to HRMS (Figure 2 and Figure S5), which was unexpected. This result indicates that both TleB and HinD can accept as the substrate and the introduced 4-F group indeed blocks the original nine-membered sulfur-containing indolelactam ring formation leading to the conversion for . However, TleB exhibited significant catalytic efficiency for the formation of new thioindole lactam derivative while HinD catalyzed the provision of the thiol dimerization product ( ) as the leading product. This result indicated that the subtle differences in homologous proteins led to different product selectivity.

The product from TleB catalysis, , possessed a structure similar to the skeleton of a class of PKD inhibitors [21] and exhibited good biological activity in inhibiting the growth of breast cancer cell line MCF-7 and gastric cancer cell line SGC-7901 (Figure S6). Therefore, we attempted to extend the substrate range to other halogen-substituted analogues – in order to provide more structural analogues (Figure 3a), laying a foundation for the diversity synthesis of thioindole lactam derivatives for further biological activity research. In the field of drug development, the introduction of halogen atoms, especially F and Cl, can improve drug properties [38,39], and the introduction of Cl and Br atoms can reserve reaction sites for post-modification [40]. Products – were purified from a scaled-up reaction *via* semi-preparative HPLC, and their structural assignments



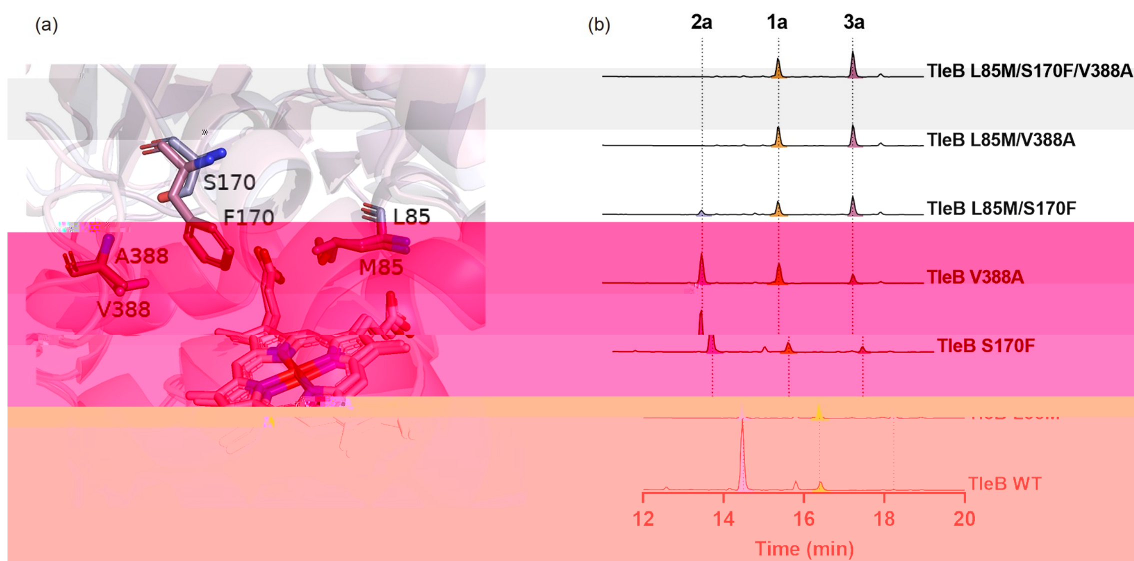
Exploring the substrate scope of wild-type TleB. (a) Reaction schemes for the conversions of – with TleB. (b) HPLC chromatograms of the enzyme reactions of TleB with – . The chromatograms were monitored at 290 nm. (c) Subsequent derivatization of / by Suzuki reaction (color online).

were confirmed by NMR and HRMS (Figures S7–S15). Among the four substrates, the kinetic parameter of TleB for was determined as k_{cat}/K_m = 44.20 min^{-1} mM^{-1}, reflecting its highest catalytic efficiency. The kinetic parameter of TleB for substrate with disubstituted F atoms was determined to be k_{cat} = 3.90 ± 0.15 min^{-1}, indicating that the reaction rate was the fastest (Table 1). By introducing Cl or Br at the C4 position of the indole ring, the obtained compounds were also accepted by TleB (Figure 3b). The kinetic parameters were determined to be k_{cat}/K_m = 23.31 min^{-1} mM^{-1} for and k_{cat}/K_m = 22.17 min^{-1} mM^{-1} for (Table 1). The catalytic efficiency of TleB for and was lower than that of , possibly due to steric hindrance. Notably, the Cl/Br-substituted products / could undergo coupling reactions under the catalysis of transition metals, such as the Suzuki reaction, which provided a good reaction site for subsequent derivatization (Figure 3c).

Kinetic parameters of wild-type TleB in reactions of –

Substrate	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (min^{-1} mM^{-1})
	1.80 ± 0.04	40.72 ± 4.93	44.20
	3.90 ± 0.15	94.37 ± 13.61	41.33
	1.16 ± 0.02	49.76 ± 4.39	23.31
	1.27 ± 0.03	57.29 ± 5.04	22.17

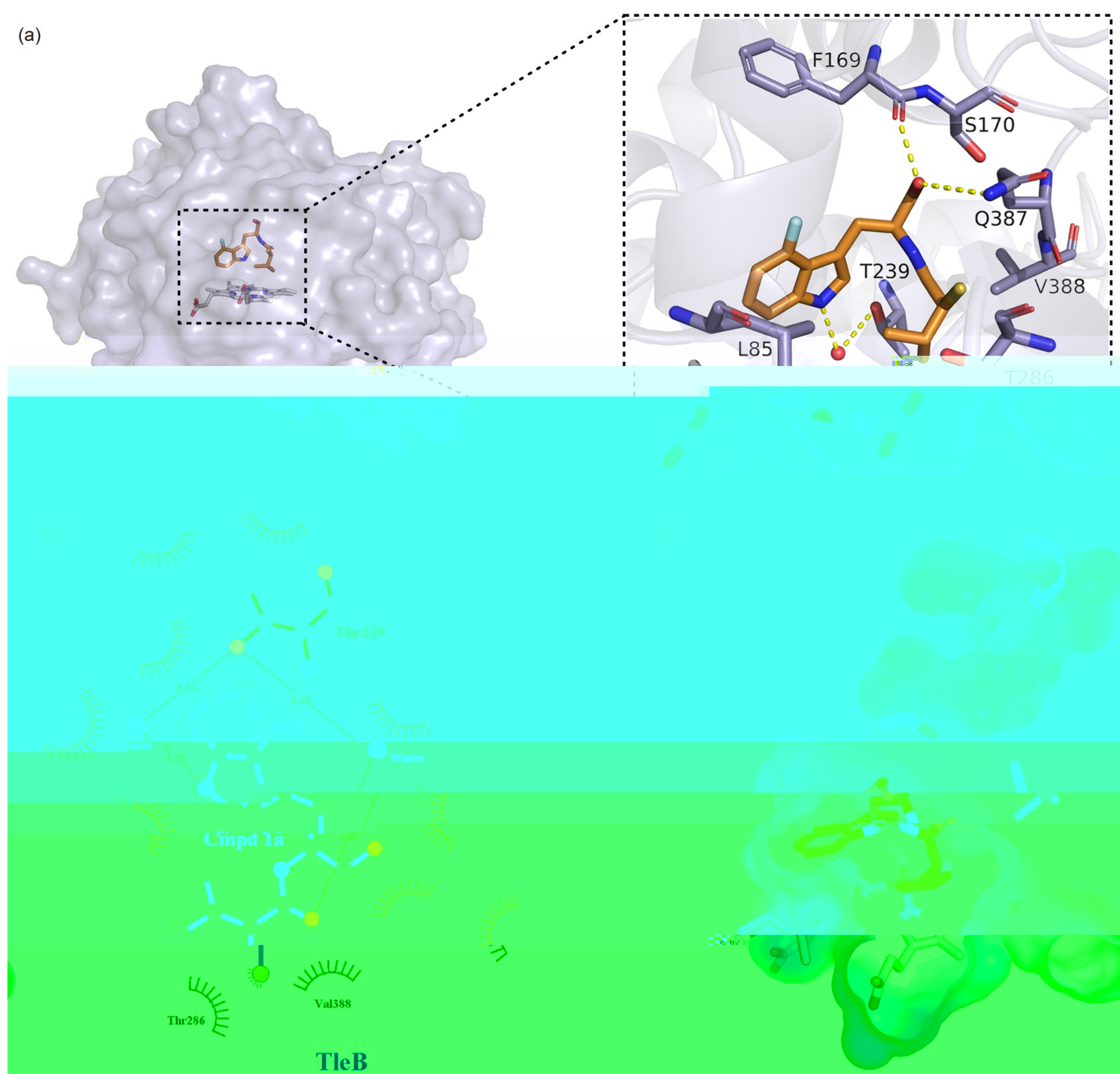
Disulfide-rich peptides (DRPs) are widely distributed in nature, and many bioactive natural products and drug molecules also contain disulfide bonds [41–43]. At present, the synthesis of disulfide bonds is mainly based on chemical methods [44–47]. There are relatively few reports on the formation of disulfide bonds by enzyme catalysis [48,49]. As mentioned above, the catalytic effects of HinD and TleB for were significantly different. HinD tended to catalyze to form a thiol-coupled product, which was a potential biocatalyst for the formation of compounds with disulfide bonds. However, the low expression of HinD in *E. coli* BL21 (DE3) made us worry about its application potential. Size exclusion chromatography results suggested as well it seemed like more HinD proteins were not folded correctly even if they were in solution. Therefore, TleB with higher expression was selected as a potential biocatalyst for further exploration. We planned to regulate the catalytic function of TleB through protein engineering so that it can effectively catalyze the formation of disulfide bonds. The mimic strategy was adopted to engineer the TleB protein to mimic the properties of HinD catalyzing the disulfide bond formation. In order to achieve the regulation of TleB function, it is necessary to first identify the key residues that lead to differences in the activity of TleB and HinD, laying the foundation for subsequent engineering studies. Therefore, the crystal structures of HinD (PDB ID: 6J85) and TleB (PDB ID: 6J82) were compared, and it was found that the active centers of the two proteins were relatively conservative, with only three residues different, which were 85, 170 and 388 residues, respectively (Figure 4a). Then, site-directed mutagenesis of TleB with HinD as a reference was performed to investigate the roles of these residues. Indeed, when Leu85, Ser170 and Val388 in TleB were uniquely replaced with Met85, Phe170 and Ala388 in HinD, respectively, it was found that all three



Engineering of TleB with HinD as reference for improving its catalytic activity for S–S coupling. (a) Comparison of the active centers of TleB and HinD. TleB is shown in purple. HinD is shown in pink. The heme molecules are depicted by white stick models. (b) Enzyme reactions of TleB WT and mutants in Tris-HCl (pH 8.0) (color online).

mutants increased the amount of dimerization product . In particular, the amount of 6/5/8 tricyclic product significantly decreased and that of increased after the L85 of TleB was mutated to methionine. Therefore, TleB L85M mutant was selected and further mutated at 170 and 388 residues to improve the formation of dimerization product . Combining L85M with two other mutations, S–S coupling product became the main product in the reaction catalyzed by TleB L85M/S170F and was the only product observed in the reaction catalyzed by TleB L85M/V388A and TleB L85M/S170F/V388A (Figure 4b). The kinetic parameters of TleB L85M/S170F/V388A were determined as k_{cat} = 8.90 ± 1.13 min^{-1} and K_m = 5.11 ± 1.20 mM for (Table S2). In addition, protein engineering not only changed the catalytic function of TleB for (Figure 4b), but also resulted in a similar ratio change between 6/5/8 tricyclic products (*i.e.*, –) and dimerized products (*i.e.*, –) when – were catalyzed (Figure S25). Therefore, protein engineering can regulate the product selectivity of TleB to achieve efficient synthesis of disulfides.

To understand the structural details of TleB involved in the catalysis of the intra-molecular C–S bond formation, the complex structure of TleB with was solved by crystallography (Figure 5a, b). The additional electron density at the active site clearly indicated the existence of the substrate (Figure S29). The binding conformation of was similar to that of NMVT in TleB (Figure S30). Inspired by the mechanism of TleB-catalyzed intra-molecular C–N coupling of

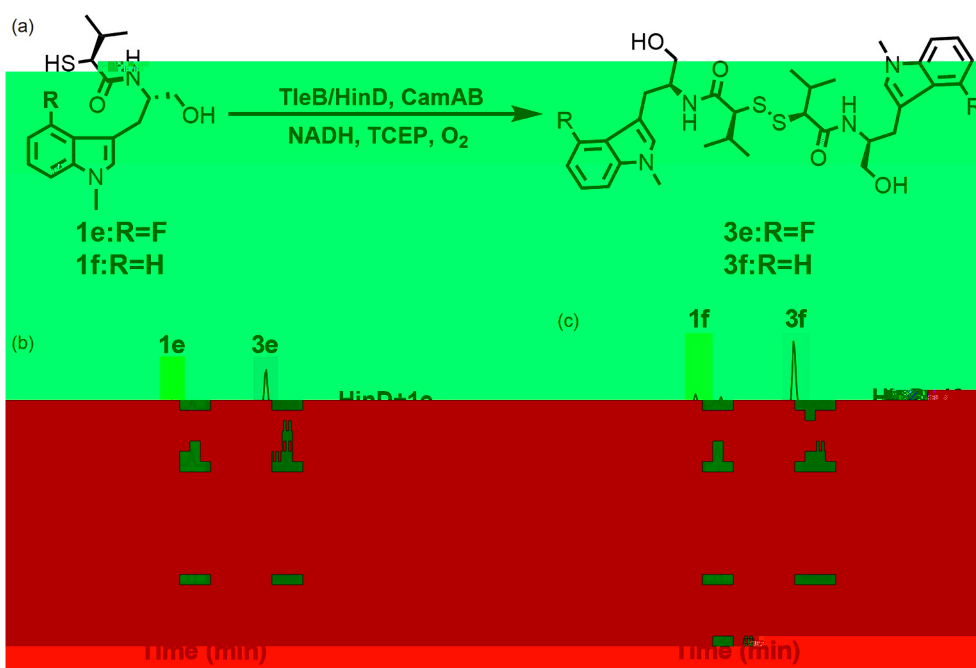


(a) The binding mode of in the active site of TleB. Heme, and residues are shown as white, orange and purple sticks, respectively. Dashed yellow lines represent hydrogen bonds. Water molecule is shown as red nb_sphere. (b) The diagram of interactions between the substrate and surrounding residues of TleB. The figure was prepared by Ligplot. (c) The substrate binding pocket of TleB. Here, the substrate binding pocket is shown as the electrostatic surface with the solvent entrance highlighted. Compound is shown in an orange stick and the heme group in a white stick. The three residues of TleB (L85, S170 and V388) are shown in blue sticks (color online).

NMVT [31], we speculated that the formation of intramolecular C–S bond begins with the abstraction of the N1 hydrogen of the indole ring by a ferryl-oxo species compound I. The distance between N1 of and an iron-coordinated water molecule is 2.5 A (Figure 5b). Thus, the N1 hydrogen is located close enough to be abstracted by compound I, to generate a radical to initiate the reaction. Further analysis of the binding mode of in TleB revealed that tryptophanol 14-OH forms hydrogen bonds with the side chain of Gln387 and the main chain of Phe169 (Figure 5a). In addition, the complex structure of TleB with revealed that N1 hydrogen interacts with Thr239 *via* a water molecule, and there is no significant hydrogen bonding between 13-SH and TleB. As the distance between C2 and S13 for C–S bond formation in TleB (5.6 A, Figure S31) is out of proximity for a covalent bond formation, movement of the radical intermediate might be needed, presumably driven by conformational changes of the enzyme. Similar movement and conformational changes have been suggested for TleB-catalyzed C–N bond formation [50]. The residue L85 is located on the surface of the substrate binding pocket and acts as a ridge to separate the indole group and the isopropyl tail of . Thus, mutation on L85 would likely change the relative arrangement of these two groups and induce a different orientation of N1 as well as the sulfur on to alter the initial step of attack. The residues S170 and V388, on the other hand, are located on the bottleneck of the solvent channel of the active site and thus govern the substrate entrance. Mutations on these two residues would change the substrate entrance and probably make it more favorable of the active site to tolerate a second molecule of to catalyze the thiol dimerization product. In summary, the TleB L85M, S170F and V388A mutants decreased the intramolecular C–S bond formation activity for (Figure 4b and 5c).

In order to further verify our conjecture on the reaction mechanism, N1 hydrogen was substituted with methyl, and two substrates and were synthesized to test the catalytic effects of HinD and TleB (Figure 6a). HPLC analysis showed that neither HinD nor TleB could catalyze the formation of C–S coupling products when the N1 hydrogen of the indole ring was substituted by methyl (Figure 6b, c), indicating that the generation of nitrogen radical might be necessary for the formation of C2–S13 bond. Since most was converted into dimer product (Figure 6c), as confirmed by NMR and HRMS (Figures S26–S28), it was speculated that the nitrogen radical could not be initiated after the N1 hydrogen of the indole ring was substituted by methyl, while 13-SH possessed high activity and S13 hydrogen might be abstracted to form radical. The formation of the disulfide bond may be due to the coupling of sulfur radicals.

Based on the above experimental results, a possible reaction mechanism was proposed. For the formation of 6/5/8 tricyclic product, ferryl-oxo species compound I first abstract the N1 hydrogen of the indole ring to form a radical, which resonates to the C3 position. Subsequently, the hydroxyl group of the C3 radical is replenished by Fe-OH (IV) to form a hydroxyl intermediate containing a quaternary carbon center. The sulfhydryl group of the substrate carries out a nucleophilic attack on C2, and finally removes a molecule of

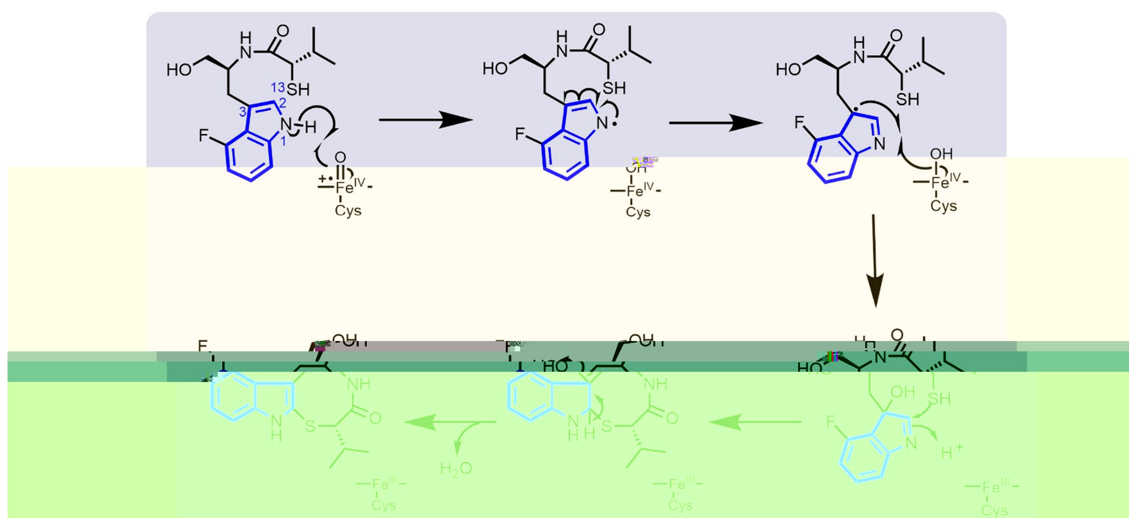


Enzymatic reaction of HinD and TleB with – in MOPS-NaOH (pH 6.5). (a) Reaction schemes for the conversions of – with HinD and TleB. (b) HPLC chromatograms of the enzyme reactions of TleB and HinD with . (c) HPLC chromatograms of the enzyme reactions of TleB and HinD with . The chromatograms were monitored at 290 nm (color online).

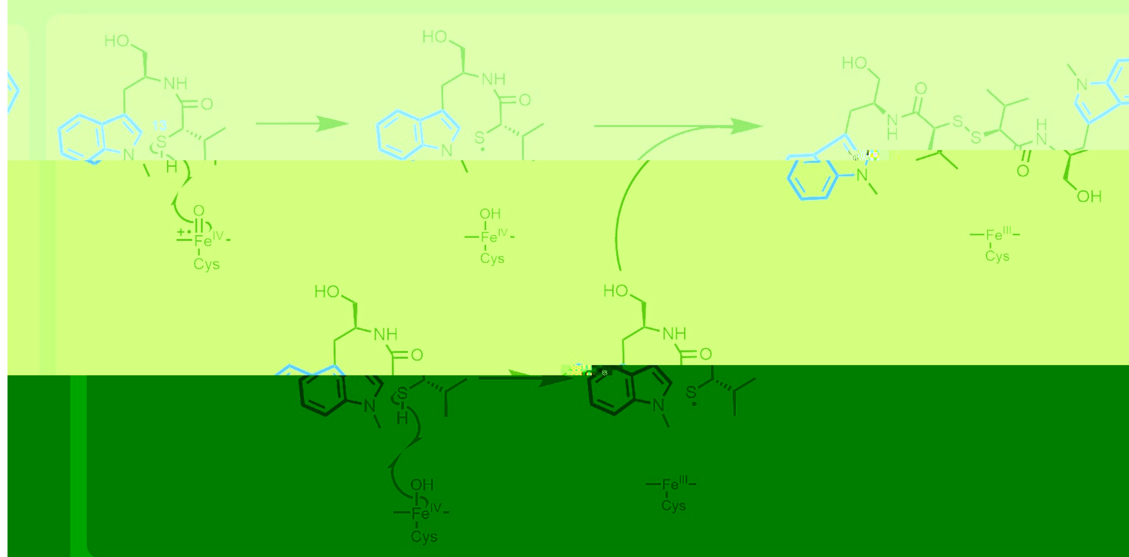
water to form the C–S coupling product (Figure 7a). For the formation of the thiol dimerization product, the S13 hydrogen might be abstracted by the highly active intermediate compound I to form a radical, and then Fe-OH (IV) continues to abstract the S13 hydrogen of the second substrate molecule, resulting in the polymerization of the two radicals to form a disulfide bond (Figure 7b). Other possible mechanisms are also under further investigation.

Several 4-halogenated (*S*)-*N*-((*S*)-1-hydroxy-3-(1*H*-indol-3-yl)propan-2-yl)-2-mercapto-3-methylbutanamide analogues were designed and synthesized based on the substrate structure-directed strategy. Wild-type TleB was used to catalyze these substrates to synthesize new drug candidates. In addition, the formation of a thiol dimerization product was observed. The selectivity of TleB was changed by site-directed mutagenesis to achieve efficient synthesis of the thiol dimerization product. Through site blocking and binding mode analysis of TleB with substrate, it was speculated that the formation of the products was initiated by the abstraction of active hydrogen by compound I in the active center of TleB to generate radicals. This study demonstrates the feasibility of combining substrate design and protein engineering to regulate biocatalytic selectivity. This strategy is helpful for achieving more reaction types by limited enzyme types. In the future, with a further understanding of enzyme reaction mechanisms and more mature protein engineering

Proposed mechanism for 6/5/8 tricyclic product formation



Proposed mechanism for thiol dimerization product formation



Proposed mechanism for the TleB and HinD enzymatic reactions (color online).

methods, this strategy could be applied to discover new catalytic flows of natural enzymes and obtain new bioactive molecules.

This work was supported by the National Key Research and Development Program of China (2018YFA0903200 to H.S.), Wuhan University, Undergraduate Training Programs for Innovation and Entrepreneurship of Wuhan University to Y.L., the Science and Technology Commission of Shanghai, and the National Natural Science Foundation of China (31900033, 21ZR1433700). The authors thank staff from BL19U1 beamline at Shanghai Synchrotron Radiation Facility (SSRF), China, for their assistance in X-ray crystal data collection.

The authors declare no conflict of interest.

The supporting information is available online at chem.scichina.com and link.springer.com/journal/11426. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

1 Newman DJ. *Natl Sci Rev*, 2022, 9: nwac206
2 Newman DJ, Cragg GM. *J Nat Prod*, 2020, 83: 770–803
3 Rodrigues T, Reker D, Schneider P, Schneider G. *Nat Chem*, 2016, 8: 531–541
4 Zhou B, Yue JM. *Natl Sci Rev*, 2022, 9: nwac112
5 Godula K, Sames D. *Science*, 2006, 312: 67–72
6 Labinger JA, Bercaw JE. *Nature*, 2002, 417: 507–514
7 Yamaguchi J, Yamaguchi AD, Itami K. *Angew Chem Int Ed*, 2012, 51: 8960–9009
8 Barry SM, Kers JA, Johnson EG, Song L, Aston PR, Patel B, Krasnoff SB, Crane BR, Gibson DM, Loria R, Challis GL. *Nat Chem Biol*, 2012, 8: 814–816
9 Devine AJ, Parnell AE, Back CR, Lees NR, Johns ST, Zulkepli AZ, Barringer R, Zorn K, Stach JEM, Crump MP, Hayes MA, van der Kamp MW, Race PR, Willis CL. *Angew Chem Int Ed*, 2023, 62: e202213053
10 Grandner JM, Cacho RA, Tang Y, Houk KN. *ACS Catal*, 2016, 6: 4506–4511
11 Li S, Tietz DR, Rutaganira FU, Kells PM, Anzai Y, Kato F, Po- [illegible]