P450-Catalyzed Biaryl Macrocyclization of Leaderless Ribosomal Peptides

Runze Liu[1#], Hongyu Ren[2#], Gengfan Wu[1], Bei-Bei He[1], Cuiping Pang[2], Zhuanglin Shen[2], Jing Liu[1], Zhi-Man Song[1], Jiahai Zhou[2, 3*] and Yong-Xin Li[1*]

[1] Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China

[2] Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[3] School of Food Science and Pharmaceutical Engineering, Nanjing Normal University, Nanjing 210023, China

[#] These authors contributed equally to this work.

Email: yxpli@hku.hk, jiahai@nnu.edu.cn

Abstract

Macrocyclic peptides, an intriguing class of molecules, embody promising frameworks for the development of chemical probes and prospective therapeutic agents. However, the lack of an established wide-substrate scope synthesis methodology has made it challenging to achieve both efficiency and selectivity simultaneously. P450-catalyzed ribosomal synthesized and post-translational modification peptides (RiPPs) are a new family of macrocyclic peptides with typical biaryl crosslinks. In this study, we report the discovery of a RiPP-modified P450 macrocyclase, GpeC, which natively catalyzes the biaryl C(sp^2)-C(sp^2) crosslink of tryptophan-tyrosine in a tetrapeptide module. Unlike most previously documented P450-RiPPs that possess a substantial leader peptide region, hindering biocatalytic applications, our P450 GpeC exhibits catalytic potential with leaderless precursor peptides and a wide substrate tolerance, highlighting its adaptability and potential for creating diverse macrocyclic structures. The crystal structure of GpeC, combined with in silico calculations and site-directed mutagenesis, provides strong evidence for the extreme flexibility of intra-cyclic amino acids within the biaryl-cyclized tetrapeptide. This flexibility enables the substitution of any of the other 19 amino acids, indicating that GpeC is a robust and adaptable catalyst for biaryl cyclization, providing a valuable resource for the biocatalytic synthesis of complex macrocyclic compounds.

Introduction

Macrocyclic peptides, an intriguing class of molecules, embody promising frameworks for the development of chemical probes and prospective therapeutic agents, owing to the fine-tuned rigid three-dimensional structures and unexpected biophysical features[1]. Those with biaryl linkages are particularly noteworthy[2,3], exemplified by prominent antibiotics vancomycin[4], arylomycin[5], and darobactin[6] (Figure 1A). Through macrocyclization, these molecules demonstrate resistance to proteolysis, enhancing the cell permeability and the ability to adopt the active conformation, which facilitates effective binding to their respective targets[7]. The significance of peptide biaryl crosslink has prompted the development of numerous catalytic methods for their synthesis. Transitional metals (TM) play a crucial role in these reactions as they connect two aryl groups by pre-functionalization or oxidation reactions with relatively well-effectiveness(Figure 1B)[8,9-12], while the implementation of multistep reactions involving high-cost metal catalysts, harmful reagents, and atom-inefficient protective groups poses substantial difficulties, particularly at the scale requisite for peptide therapeutics[13]. Despite their proven therapeutic potential, the structural complexity of biaryl-cyclized macrocyclic peptides, along with the inefficient cyclization protocols, presents significant hurdles to their synthetic preparation[14].

Biocatalysts are becoming increasingly essential in synthetic chemistry, which offers a superior alternative by modifying scaffolds using substrate-tolerant enzymes[15-18]. By their benign operating conditions, selectivity, and evolvability, enzymes could provide more efficient and sustainable routes, making these medically valuable agents more available[19-21]. Bacterial cytochrome P450 enzymes, particularly, have been extensively studied as versatile biocatalysts[22,23]. The biaryl-bridged macrocyclization through P450-mediated oxidative couplings greatly expands the chemical space and complexity of macrocyclic peptides[24]. These highly modified structures exhibit favorable drug-like properties, positioning them as promising candidates for drug discovery[25-28]. P450-catalyzed macrocyclization in non-ribosomal peptides (NRPs) such as vancomycin[29], however, requires a peptidyl carrier protein-tethered precursor for substrate recognition, which complicates the utilization of these P450s and limits their biocatalytic potential[30-32]. In contrast, the unexplored P450s that modified ribosomal synthesized and post-translational modification peptides (RiPPs), due to their concise and versatile biosynthetic pathway and promiscuity, present a more direct biocatalytic pathway for the production of biaryl-crosslinked cyclic peptides[33].

The biosynthesis of P450-catalyzed RiPPs involves a relatively simple sequential process as the RiPPs biosynthetic logic, ribosomes synthesize precursor peptides, followed by post-translational modifications (PTMs) facilitated by P450 enzymes[3,34]. These enzymes catalyze oxidative cross-coupling reactions between aromatic rings through the homolysis of phenolic O-H or tryptophan indole N-H bonds[35]. The rapidly expanding P450 families are known to catalyze several macrocyclic crosslinking types, including Trp-Tyr, Trp-Trp, Trp-His linkages in atropitides[36-41], Tyr-Tyr linkages in cittilins[39,41,42], and Tyr-His, His-His linkages in biarylitides[38,43] (Figure 1A). Leveraging the straightforward biosynthetic logic and the potential of tolerating various substrates, the RiPP-P450 enzymes represent a promising and underexplored class of biocatalysts for peptide biaryl cyclization[44]. A recent study on $P450_{blt}$-catalyzed tricyclic biaryl module formation involving Tyr-His and Tyr-Tyr cyclization has demonstrated the enzyme's ability to carry out biaryl-bridged macrocyclic reactions[45,46], while the substrate scope of $P450_{blt}$ limited its catalytic ability on the stereo-hindered amino acid residues inside the tricyclic module. Apart from $P450_{blt}$, the landscape of RiPP-P450s, particularly catalytic potential in vitro, remains largely unexplored[36-43], limiting their application as biocatalysts.

In our recent research, we employed the SPECO genome mining strategy to uncover the untouched potential of P450s in RiPP modification. This approach led to the identification and in vivo characterization of four versatile P450s[38]. However, their potential for in vitro enzymatic activity remains unexplored, thus hindering their use in peptide bioengineering. By implementing the SPECO genome mining strategy on a significantly larger dataset, in this study, we further expanded the scope of RiPP-P450s and unveiled GpeC, a previously uninvestigated RiPP-P450 enzyme that demonstrates a diverse catalytic ability in biaryl cyclization in vitro. Unlike known P450s, where the long leader region significantly hinders in vitro biocatalytic potential, GpeC maintains its catalytic ability by accepting the 9-mer core peptide region in a leader-independent manner. Further investigation revealed GpeC's efficient biaryl crosslink formation between Trp-Tyr, Tyr-Trp, and Tyr-Tyr (Figure 1C). By integrating the GpeC crystal structure, molecular dynamics calculations, and site-directed mutagenesis, we concluded that the 9-mer substrate binding phase accommodates intra-cyclic amino acid variations. GpeC's catalytic capabilities position it as a promising template for bioengineering in biaryl macrocyclic peptide synthesis, advancing the biocatalysis of biaryl macrocyclic peptides.

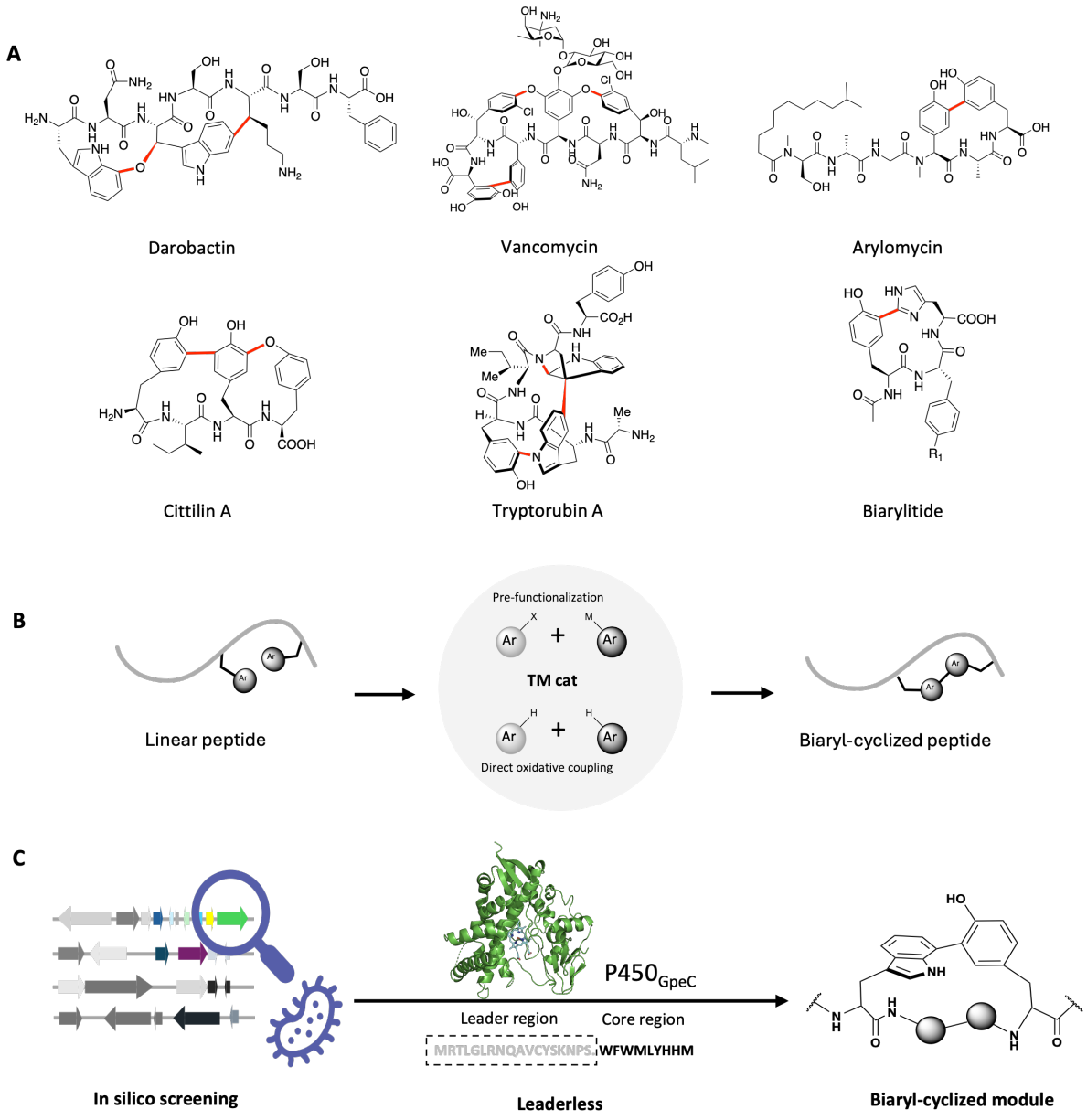

Figure 1. Representative biaryl crosslinked macrocyclic peptides and their synthetic strategies. A) Natural-derived biaryl crosslinked macrocyclic peptides. B) Chemical synthesis of biaryl crosslink containing macrocyclic peptides. C) Genomics-guided discovery of the P450 GpeC enzyme, which facilitates the formation of the biaryl crosslink under mild conditions, exhibiting both leader-independent properties and specificity.

Results

Genomics-guided discovery of a new P450 enzyme GpeC

We have recently developed a genome mining workflow called 'Short Peptide Enzyme Co-Occurrence' (SPECO) analysis, which effectively identifies RiPP BGCs without relying on sequence similarity to known RiPPs[47-49]. Intrigued by the fascinating chemistry of P450-catalyzed biaryl crosslinking[38,40,45], we previously harnessed the SPECO strategy to investigate the P450-catalyzed RiPPs landscape within actinobacteria and validate P450 enzymatic potential in vivo. However, the task of isolating pure and soluble RiPP-P450s for in vitro studies presents a significant challenge, despite its crucial role in comprehending their catalytic efficiency and biocatalysis potential[44]. In this study, we first evaluated the cytochrome P450 enzymes, previously reported in our research, specifically KstB, MciB, SgrB, and ScnB, but didn't find enzymatic activity in vitro. Thus, through ongoing efforts using SPECO, we expanded our investigation of the small peptide-P450s landscape from actinobacteria to all bacterial phyla, expanding the small peptide-P450s sequence pairs to 10, 228 (SI). Notably, the *gpe* family from Gram-negative bacteria *Gallaecimonas pentaromativorans* DSM 21945, previously undiscovered, garnered our attention due to their diverse precursor peptides.

We first characterized the function of unraveled P450 GpeC primarily via co-expression of the precursor hexahistidine-SUMO-GpeA$_2$ and tailoring enzyme GpeC in *Escherichia coli*, digested by trypsin which led to the identification of a new P450 GpeC-catalyzed Try-Tyr crosslinked product through comparative ultra-performance liquid chromatography-high resolution mass spectrometry (UPLC-HRMS) analysis (Extended data). Comparative UPLC-HRMS analysis showed hexahistidine-SUMO-GpeA$_2$-GpeC presented one new peak with 2 Da minor than the GpeA$_2$ (Extended data). Tandem mass analysis of fragments further supported that the modification was installed on the WMLY motif (Extended data). Based on the biosynthetic logic of P450-catalyzed RiPPs, we hypothesized that the crosslink was formed between W21 and Y24.

To verify the crosslinking pattern catalyzed by GpeC, we conducted 18 liters fermentation of *Escherichia coli* coexpressing both GpeA$_2$ and GpeC before incorporating chymotrypsin to digest the catalyzed peptide into a 7-mer core region WMLYHHM (**1**), obtaining 3 mg of compound **1** for NMR analysis (SI). The intensive NMR analysis confirmed the biaryl crosslink between Try1-C7 and Tyr4-C3, as evidenced by the singlet signal of Tyr4-H2 and the HMBC

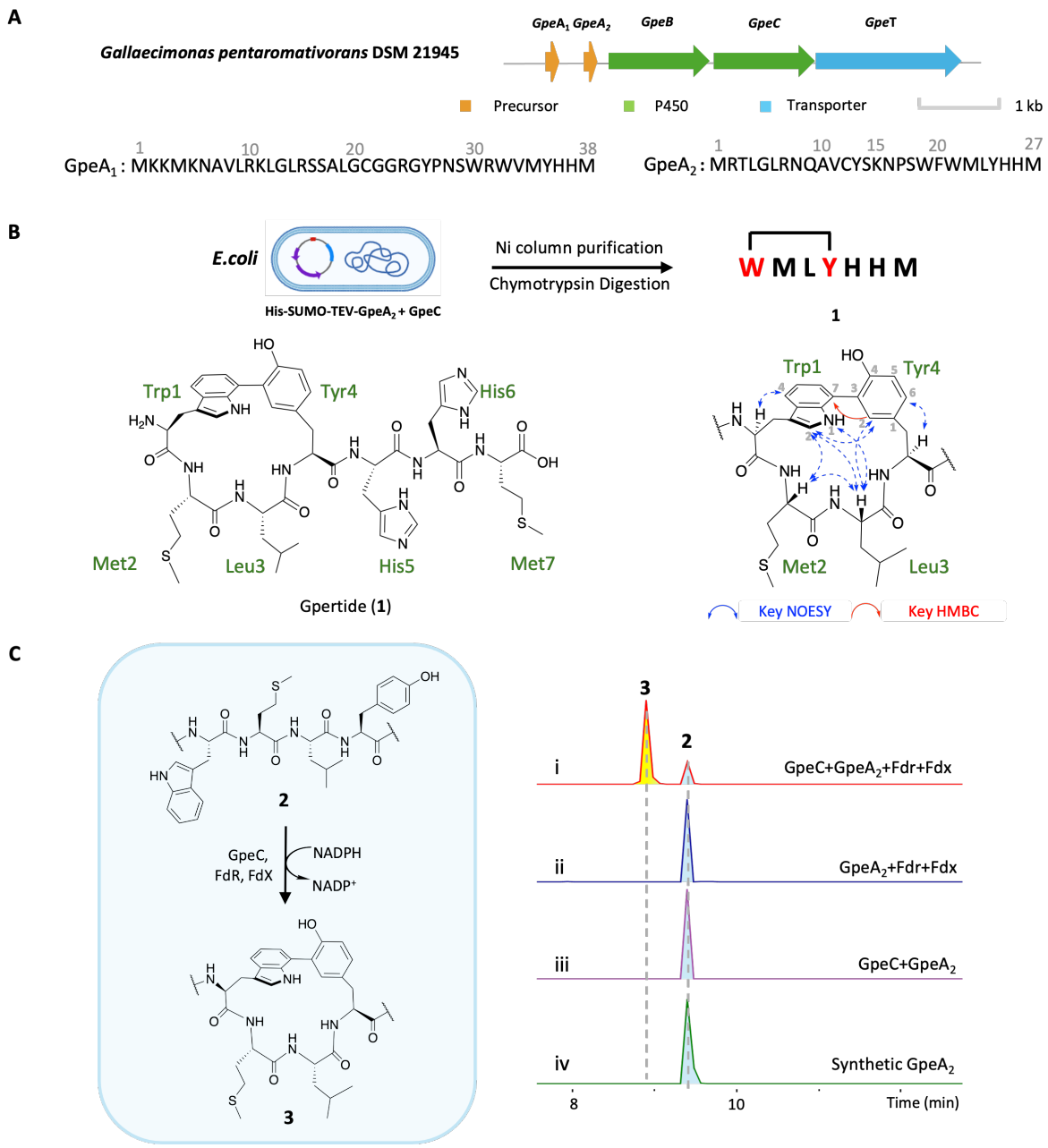

Figure 2. Structure elucidation of Gpertide and in vitro assay. A) BGC architecture of the *gpe* family. B) In vivo characterization of GpeC and structural elucidation of GpeC's enzymatic product Gpertide (**1**). C) In vitro assay of GpeC catalyzed biaryl crosslink in WMLY region with the extracted mass profile.

signal from Tyr4-H2 to Try1-C7. The NOESY spectra further verified the biaryl system, showing interactions between Tyr4's phenolic OH-group hydrogen and Try1-H6, and between Try1-H2 and Tyr4-H2. Advanced Marfey's analysis of compound **1** revealed L-configured residues of methionine, leucine, and histidine (SI). The biaryl system's axial chirality was assumed to be *Sa* based on the NOESY spectra, and this was further confirmed by comparing the experimental Electronic Circular Dichroism (ECD) spectrum with calculated isomers (SI).

To characterize the P450 enzyme in vitro, we integrated a chaperone and 5-aminolevulinic acid (5-ALA) into the *E.coli* expression system and successfully obtained a functional in vitro GpeC (Figure 2C). The functionality of GpeC was confirmed through a Carbon Monoxide Reduction Assay[50], which demonstrated a typical peak shift in the UV spectra from 426nm to 450nm (SI). The in vitro assay of GpeC with $GpeA_2$ (**2**), involving NADPH and spinach-oriented ferredoxin-ferredoxin reductase pairs, produced a catalytic product after 6 hours at 28°C (Figure 2C). The UPLC-HRMS together with the tandem mass profile of the enzymatic product (**3**) corresponded with that of the in vivo modified peptide, indicating that GpeC can catalyze the biaryl crosslink formation of the WMLY motif in vitro (Extended Data).

Enzymatic characterization of GpeC

Typically, the biocatalytic process of RiPPs is contingent on the leader peptide's binding to the PTMs, which often restricts these enzymes' catalytic scope. To explore GpeC's enzymatic potential, we performed a leader removal assay on the native precursor $GpeA_2$, removing over 18 residues from $GpeA_2$'s N-terminus with a truncation step size of three amino acids. We chemically synthesized the truncated constructs and assayed them with GpeC in vitro (Figure 3A). To determine the smallest catalytic precursor, we performed a comparative UPLC-HRMS analysis between synthesized precursor and enzymatic product. By testing 9-mer substrate WFWMLYHHM (**4,** $[M+3H]^{3+}$=450.8685), the profile exhibited that a -2 Da loss peak (**5**, $[M+3H]^{3+}$=450.1964) was observed in the enzymatic assayed product (Figure 3A). Tandem mass analysis of **5** further supported that the modification was installed on the WMLY motif (Figures 3A and SI), consistent with the reaction of native precursor $GpeA_2$ (**3**) with a conversion ratio of 36%. However, further reduction into the 7-mer peptides WMLYHHM or WFWMLY did not yield any cyclized product, suggesting that the amino acids on the positions including N-terminal Trp1, Phe2, and C-terminal tail His7, His8, Met9 residues are crucial for the catalytic process.

Taking into account the distinctive ability of GpeC to catalyze the leaderless precursor $GpeA_2$, we delved deeper into examining the catalytic capabilities of GpeC. The *Gpe* biosynthetic gene cluster from *G. pentaromativorans* includes two precursors, $GpeA_1$ and $GpeA_2$ (Figure 2A). We carried out an in vitro biocatalytic reaction of GpeC with the 9-mer substrate $SgpeA_1$, which resulted in a Trp-Tyr cyclized $SGpeA_1$ product with a conversion ratio of 78% (Figure 3B) (SI). To investigate the flexibility of GpeC, we altered the ring-forming residues to other aromatic residues. We discovered that GpeC facilitated the crosslinking between Trp3-Tyr6 ($SgpeA_1$, $SgpeA_2$), Tyr3-Tyr6 (**7**, **8**), and Tyr3-Trp6 (**9**, **10**) (Figure 3B). The enzymatic process was

demonstrated to be highly efficient, particularly on the linear precursor WRYVMYHHM (**6**), with a total conversion ratio to the biaryl-cyclized product (**7**) of 92%. In the case of the HHM region, we found that H8S, H8N, M9I, and M9L are also compatible with the catalytic cyclization process facilitated by GpeC (SI).

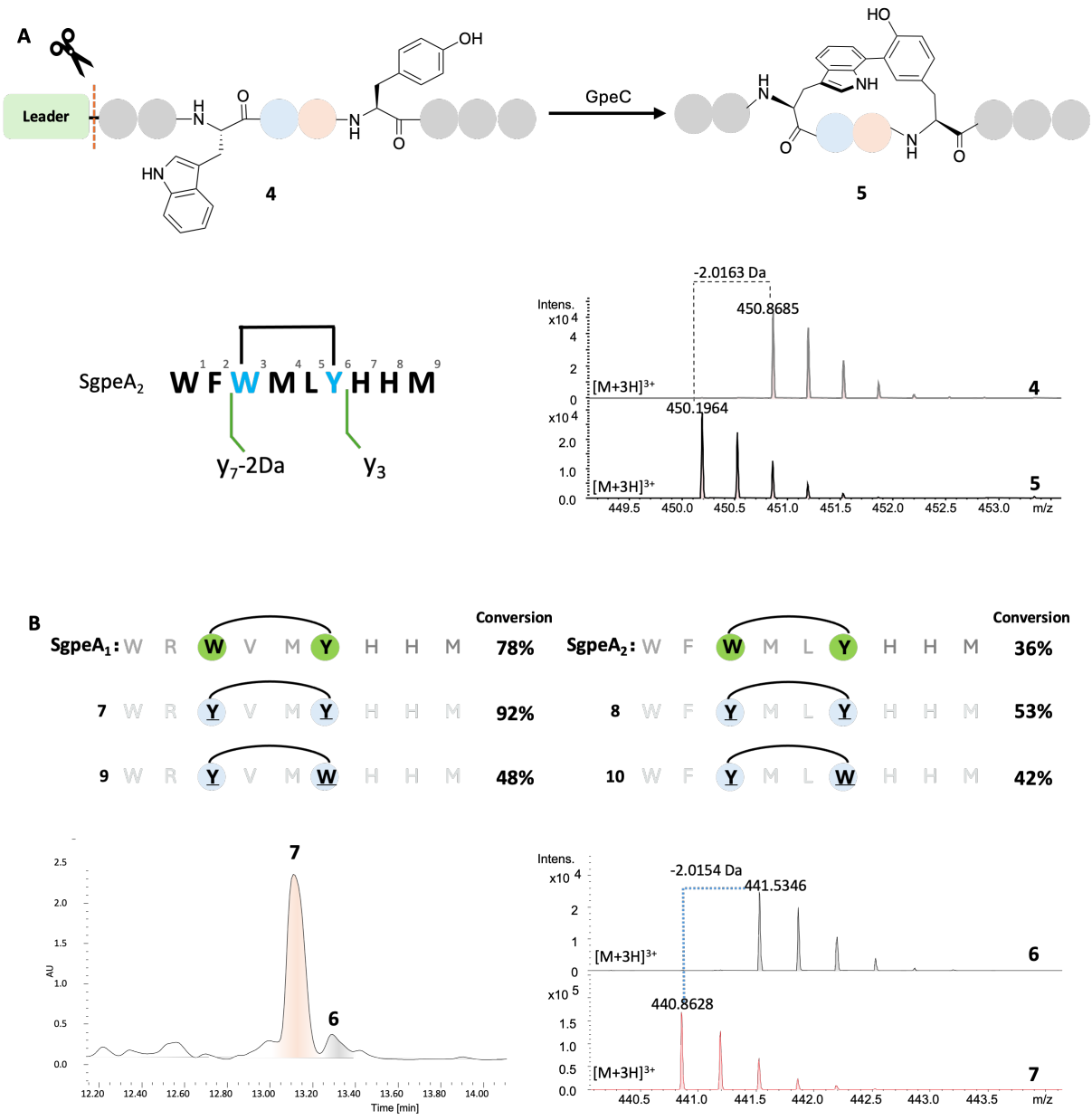

Figure 3. In vitro characterization of GpeC with truncated precursors A) Leader-removal with a truncation step size of 3 amino acids assay of Gpe_{A2} and assays with precursor Gpe_{A1}. B) Primary substrate promiscuity of GpeC and generating a new-to-nature YVMY module.

Crystal structure and mutagenesis study of GpeC

To understand the catalytic reaction mechanism, we determined the X-ray crystal structure of GpeC. To obtain high-resolution diffraction data, we have applied several optimization strategies to protein sequences, such as sumo fusion expression, C-terminal truncation, and site-directed mutagenesis (SI). The 2.86 Å crystal structure of GpeC-F62G was determined using the molecular replacement method, with $P450_{GpeC}$ coordinates predicted by Alphafold3[51] serving as the molecular replacement model. The crystal belongs to the space-group $P4_32_12$ and contains two molecules (chain A and chain B, Figure 4A) in each asymmetric unit. GpeC-F62G displays the typical P450 architecture, mainly consisting of 12 α-helical secondary structures, multiple β-sheets, and a centrally positioned heme group coordinated by a thiolate (Figure 4B). The heme iron in the active center of GpeC-F62G is axially coordinated by a conserved proximal residue Cys325, forming a cysteine-iron bond. The two propionate groups of the heme are coordinated through electrostatic interactions with two highly conserved histidine residues (H57 and H323) and one arginine residue (R261). In the P450 catalytic cycle, the conserved acid/alcohol residue, typically located in the I-helix above the heme (Ser-219), plays a central role in oxygen activation (SI). Structural alignments were performed by using the Darli Sever[52], which confirmed that GpeC-F62G exhibits high structural similarity to $P450_{BM3}$ (PDB: 2NNB) and $P450_{Blt}$ (PDB: 8U2M)[53], with Z-scores of 30.8 and 25.6, respectively. Alignment GpeC with $P450_{Blt}$ revealed highly conserved I-helix and L-helix structures, while the F-helix, G-helix, and BC-loop show significant secondary structure differences (Figure SI). This disparity may be attributed to the larger substrate catalyzed by $P450_{GpeC}$ compared to the substrate (MRYLH) of $P450_{Blt}$, necessitating the outward movement of the F-helix and G-helix to create a larger substrate-binding pocket. The volume of the active cavity of $P450_{GpeC}$ (3558 $Å^3$) was found to be 38% larger than that of $P450_{Blt}$ (2574 $Å^3$) through CavitOmiX calculation[54] (SI).

Generally, the substrate-binding site of P450 is located in close proximity to the heme and involves the B-C and F-G loops, as well as helices F, G, and I. The B-C loop and F-G region constitute the entrance to the active site and are responsible for substrate recognition (Figure 4B). To evaluate how the substrate binds to GpeC, we docked the 9-mer substrate WFWMLYHHM to the active pocket of GpeC and conducted molecular dynamics simulations (Figure 4C). The results indicated that the active pocket is a large cavity where the polypeptide assumes a V-shaped conformation (Figure 4C). The carbon atoms of the tyrosine and tryptophan involved in the coupling reaction of the 9-mer substrate are positioned 2.60 Å away

from the oxygen of the Fe-oxygen intermediate, respectively, suggesting that the reaction can proceed as anticipated (Figure 4D). To verify the impact of substrate pocket amino acids on catalytic efficiency, site-directed mutagenesis was performed on the relevant amino acids. GpeC lost its catalytic activity after mutations on N-terminal tryptophan interacting residues W257G (Figure 4E). For residues binding N-terminal Trp1-Phe2-Trp3, H57, and D59 are necessary for binding Trp3 of the substrate, while mutations on H57G and D59G significantly reduce the catalytic ability (Figure 4E). The F62G mutation increased GpeC activity in catalyzing the 9-mer substrate (Figure 4E). Reverting the 62G in the mutant back to phenylalanine and conducting molecular dynamics simulations based on the crystal structure showed no significant differences between the two proteins, with a root mean square deviation (RMSD) of 1.206 Å for the backbone atoms across all chains (Figure SI). Compared to the wild type, the efficiency of substrate binding is approximately 80% for the F62G mutant (Figure SI), indicating a higher catalytic activity of F62G (Figure 4E). The mutagenesis study elucidated the functions of amino acids related to the substrate pocket of GpeC, and we hypothesized that a range of tolerance may exist in the inner-cyclic residues within the cyclized tetrapeptide region.

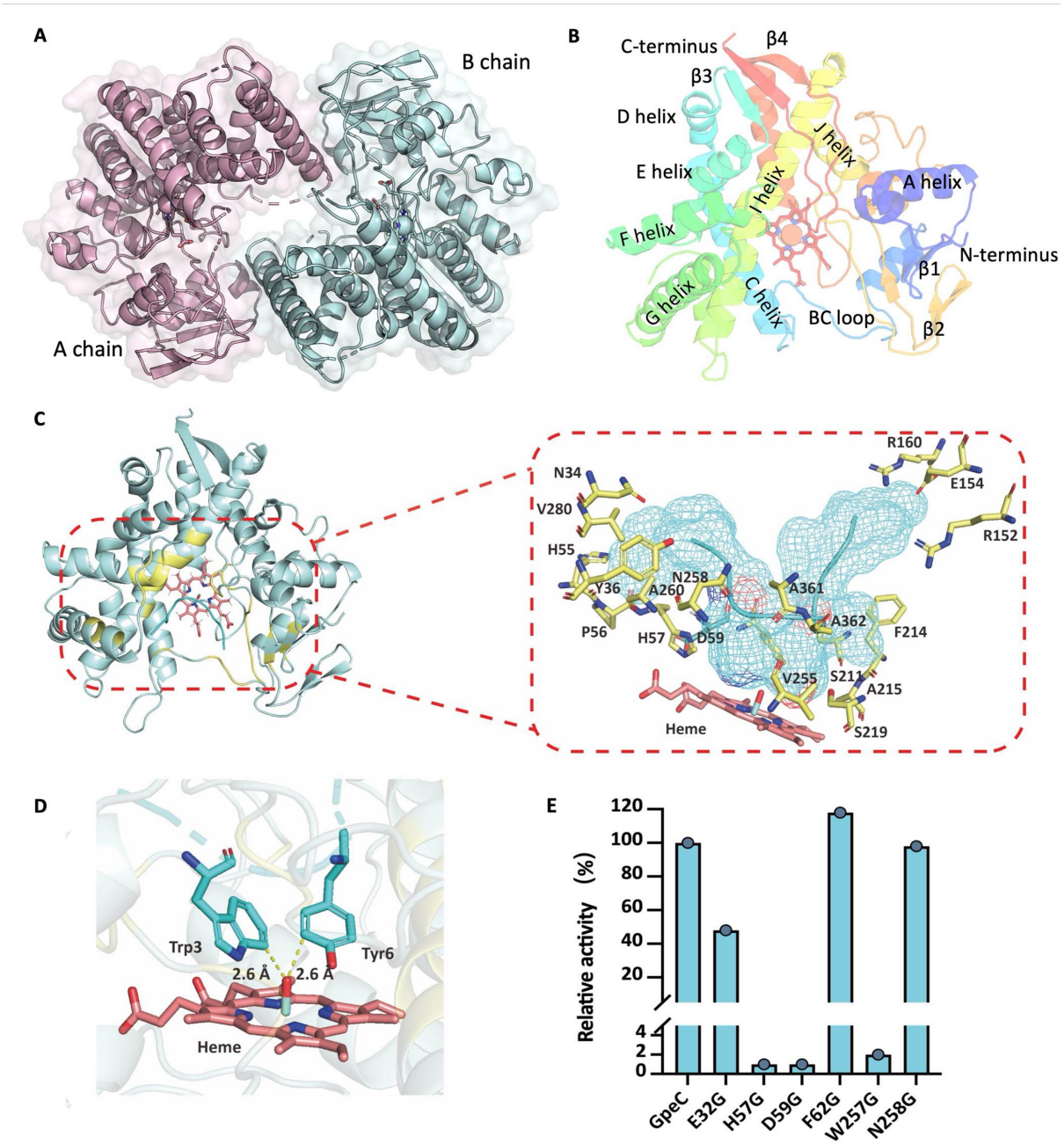

Figure 4. Analysis of the structural characteristics and site-directed mutagenesis of GpeC A) Asymmetric unit of the P450GpeC-F62G. Pink represents A chain. Palecyan represents B chain. B) Secondary and tertiary structure of each monomer. C) Docking and MD results of 9-mer substrate complex. Cyans represents the substrate WFWMLYHHM. Paleyellow represents amino acid residues within a 4 Å range around the substrate. Red represents heme. D) The binding pose and distance of Trp3 and Tyr6 of the substrate WFWMLYHHM with Fe-oxygen in heme. E) Functional validation of amino acid residues related to the substrate pocket of GpeC.

Substrate scope of GpeC

Based on the assumption of GpeC's adaptability regarding the intra-cyclic amino acids in the biaryl-cyclized module, we examined its substrate scope using an in vitro assay for a series of precursor mutants. We chemically synthesized a range of 9-mer peptide precursors based on the motifs of WRWX_1MYHHM and WRWVX_2YHHM, where X_1 and X_2 could represent any of the 20 native amino acids, except for selenocysteine. To simplify the chemical synthesis process, the 20 amino acids were classified into five groups according to the properties of their side chains. Enzymatic reactions were conducted in the presence of redox partners for 6 hours at 28°C and halted by methanol. The enzymatic product was then subjected to UPLC-HRMS analysis (Figure 5A). GpeC showed excellent substrate promiscuity that it could accept a wide substrate scope for the intra-cyclic amino acids for an average conversion ratio of 19 reacted amino acids on both X_1 (45.6%) and X_2 (66.7%). For the amino acid residues with electrically charged side chains including R, H, K, D, E (**11-15)**. Position X_1 exhibited an average conversion ratio of 53% and exhibited greater catalytic ability on H (78.5%) and R (61%), while position X_2 exhibited an average conversion ratio of 66.9%. GpeC also demonstrated its catalytic ability on amino acids with polar uncharged side chains, including S, T, N, and Q (**16-19**). The enzymatic activity for catalyzing amino acids N and Q on the X_1 position has proved to be efficient with an average conversion ratio of 69.5%. However, the substitution of S and T on the X_1 position significantly decreased the conversion ratio of the enzymatic reaction to 15%, suggesting the incompatibility of aliphatic amino acids with terminal hydroxyl groups. Conversely, at the X_2 position, the enzymes demonstrated a robust average conversion ratio of 83%. For amino acids with hydrophobic side chains or without side chains, including G, A, L, I, V, W, M, Y, and F (**21-29**), GpeC was found to exhibit lower catalytic activity for amino acids without side chain or with small side chains (G, A) on the X_1 position (13%). However, when it included L, I, V, and those with sterically hindered aromatic side chains, such as W and F, Y, GpeC's catalytic efficiency was higher (50.1% for X_1 and 66.5% for X_2). Yet, the introduction of proline (**30**) at either X_1 or X_2 positions did not yield a biaryl-cyclized product. In conclusion, for both X_1 and X_2 positions, 19 of the native amino acids could be successfully catalyzed, achieving a conversion ratio of up to 94% (Figure 5B).

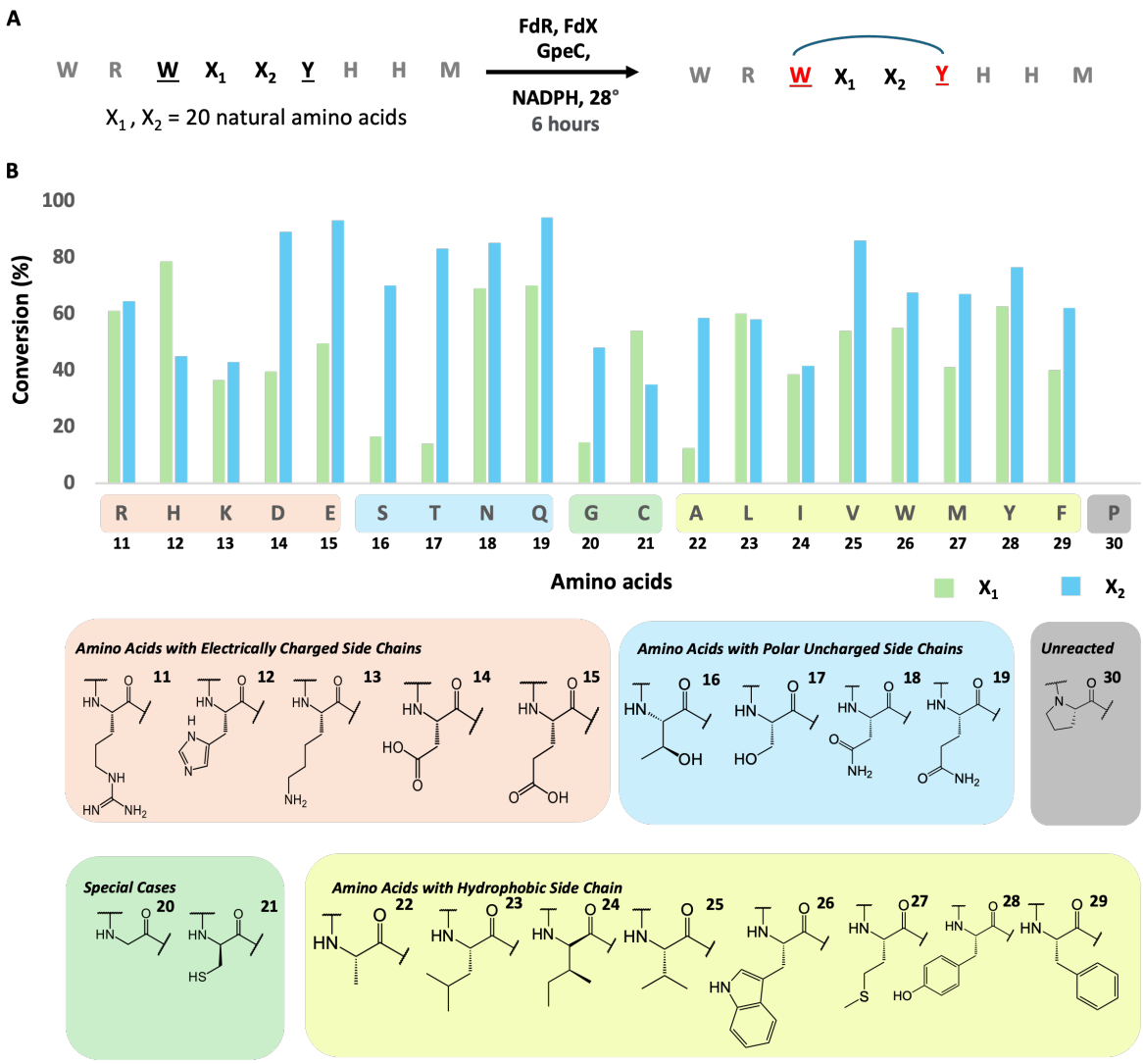

Figure 5. Substrate scope of GpeC A) General in vitro reaction process. B) GpeC can accept 19 kinds of amino acids with diverse side chain properties for the inner-cyclic biaryl-cyclized tetrapeptide motif.

Discussion

Macrocyclic peptides, as potent therapeutic leads, are highly attractive due to their metabolic stability and cell permeability. Recent advancements in peptide cyclization have predominantly focused on methods utilizing amino acids with reactive side chains. While synthesizing cyclic peptides with non-reactive amino acids is feasible, expanding these methods to create novel cyclic modules and apply them on a broad scope remains challenging, particularly in forming biaryl bonds. Encouragingly, the biosynthesis of biaryl-cyclized peptides by P450 offers a potential solution to these challenges. P450 enzymes have demonstrated their efficacy in the biaryl cyclization of non-ribosomal peptides. However, their utilization is often hindered by

substrate binding to the peptidyl carrier protein (PCP), which obstructs direct catalytic activity. The recent discovery of P450 enzymes involved in the cyclization of RiPPs has shed light on their potential to bypass these limitations, paving the way for more efficient biaryl cyclization. The identification of $P450_{blt}$, in particular, reveals that the RiPP-P450 enzyme is capable of catalyzing the formation of a biaryl bond within a tetrapeptide module, underscoring its usefulness in complex cyclic peptide synthesis[45,46,53]. Given their potential in catalyzing biaryl cyclization reactions, further discovery and characterization of novel RiPP-P450s is of paramount importance. The widely-recognized SPECO (Short Peptide-Enzyme Co-Occurrence) method is currently being employed to identify and characterize new RiPPs-P450 enzymes. By marrying the principles of RiPP biosynthesis with bioinformatic tools, SPECO overcomes restrictions in traditional genome mining methods, especially those related to the absence of hallmark genes, expanding the untapped landscape of RiPPs-P450s.

Based on SPECO analysis, we redefined the landscape of RiPPs-P450s, identifying 10,228 peptide-P450 pairs and characterizing the functional P450 enzyme, GpeC, which catalyzes the formation of a biaryl C-C bond between the C-7 of tryptophan and the C-4 of tyrosine in a tetrapeptide module. Using an optimized *E. coli* expression system, we obtained functional P450 GpeC in vitro. Remarkably, GpeC not only retained its catalytic activity with leaderless substrates but also demonstrated its ability to synthesize a novel tetracyclic biaryl cyclized module, YVMY. We determined GpeC's crystal structure together with docking and molecular dynamics simulation and inferred that the binding of the 9-mer substrate predominantly relies on the C-terminal tryptophan and the N-terminal histidine residues. However, it is presumed that there is a degree of tolerance for variations in the amino acids within the intracyclic region. Based on the assay with synthetic substrates varying in the intra-cyclic peptide within the biaryl module, we are excited to find that GpeC could accept 19 amino acids in the intracyclic tetrapeptide region.

Our investigation into GpeC has resulted in the development of a biocatalytic approach for synthesizing biaryl-cyclized macrocyclic peptides. The enzymatic activity of GpeC demonstrated remarkable efficiency and broad substrate tolerance, which is advantageous for the creation of a cyclic peptide library aimed at discovering novel bioactivities. We believe that this enzyme-catalyzed reaction will also inspire organic chemists to devise chemo-enzymatic strategies for constructing complex cyclic peptides featuring biaryl-cyclized peptide motifs. We aim to enhance the enzymatic activity and expand the substrate scope through further

engineering of GpeC. Ultimately, we foresee that the discovery of GpeC will evolve into a sustainable and effective tool for the synthesis of biaryl-cyclized peptides.

Method

General chemical reagents

All chemicals and culture medium, including kanamycin, Chloramphenicol, Spectinomycin and isopropyl β-D-1-thiogalactopyranoside (IPTG) are purchased from either Sangon Biotech (Shanghai) or Sigma (USA). Marfey's reagent Nα-(5-Fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA), Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide (D-FDLA), and Tioglycolic acid (TGA) were purchased from TCI and MACKLIN. Oligonucleotide primers were synthesized from Ruibotech (Guangzhou). DNA polymerase Phanta Max Master Mix was purchased from Vazyme. TEV protease, endonucleases BamHI, Hind III, XhoI, NcoI, NdeI, XbaI, and Gibson assembly kit NEBuilder® HiFi DNA Assembly were purchased from New England Biolabs. Trypsin and PD-10 MidiTrap desalting columns were purchased from Sigma. HISPUR NI−NTA resin was purchased from Thermo Fisher. NMR solvent Dimethyl sulfoxide-d6 was purchased from Cambridge Isotope Laboratories.

Bioinformatic analysis

SPECO pipeline was used to calculate co-occurred sORF-P450 pairs. Sequence clustering analysis of sORF and P450 enzyme sequences was conducted using MMseqs2[55] with the following parameters: mmseqs --easy-cluster --min-seq-id [0.6 for P450, 0.35 for small peptide] --seq-id-mode 1 --cluster-mode 1 --similarity-type 2 --single-step-clustering --dbtype 1. Multilayer sequence similarity network (MSSN) analysis of sORF-P450 pairs were performed by employing Python package NetworkX and visualized by using Matplotlib.

Cloning and plasmid construction

Redox partner gene synthesis and construction**.** For spinach ferredoxin and ferredoxin reductase, we synthesized the same nucleotide sequences (sequence id: *TRX-Spinach ferredoxin* and *Spinach ferredoxin reductase*) as reported by Prof. Xudong Qu, et al.[56]. The *E. coli* thioredoxin gene was fused to the N-terminal of ferredoxin to increase the solubility as reported by Prof. Xudong Qu, et al[56]. The maltose binding protein (MBP) was fused to the N-terminal of ferredoxin reductase to increase the solubility. These two fusion proteins were respectively cloned into multiple cloning sites 1 of pCDF-Duet-1 and pRSF-Duet-1 vectors, generating pRSF-fdr and pCDF-fdx.

Precursor and P450 enzyme cloning and plasmid construction. For precursor cloning, $GpeA_1$, $GpeA_2$ were individually cloned by using each genome as a template and inserted into the multiple cloning site 1 of pRSF-Duet-1 with N-terminal His^6-SUMO-TEV tag, yielding pRSF-SUMO-$GpeA_1$ and pRSF-SUMO-$GpeA_2$ plasmids. For precursor and P450 coexpression, corresponding P450 enzyme GpeC was cloned and inserted into the multiple cloning site 2 of pRSF-SUMO-$GpeA_1$ and pRSF-SUMO-$GpeA_2$, generating pRSF-SUMO-$GpeA_1$C and pRSF-SUMO-$GpeA_2$C.

P450 enzyme cloning, GpeC was individually cloned by using each genome as a template and inserted into the multiple cloning site 1 of pRSF-Duet-1 with N-terminal His^6-tag, yielding pRSF-His-GpeC.

Transformation, protein overexpression, and purification

For precursor expression, each redox partner, including pRSF-fdr and pCDF-fdx., was individually transformed with *E. coli* Rosetta (DE3) to generate strain E. coli-redox. Plasmids pRSF-SUMO-$GpeA_2$ and pRSF-SUMO- $GpeA_1$ were respectively transformed with *E. coli* Rosetta (DE3). For co-expression of precursor and P450, pRSF-SUMO-$GpeA_2$-*GpeC* were respectively transformed with *E. coli*-redox or *E. coli* Rosetta (DE3).

After transformation, a single colony of each recombinant strain was inoculated into 5 mL fresh LB broth with 34 μg/mL chloramphenicol (native resistance for *E. coli* Rosetta (DE3)) and 50 μg/mL kanamycin (for pRSF-Duet-1 backbone) and shaken at 37 °C overnight. Then, 1 mL of the culture was incubated into 100 mL fresh TB medium (24 g/L yeast extract, 12 g/L casein hydrolysate, 9.4 g/L K_2HPO_4 and 2.2 g/L KH_2PO_4 and 8 ml/L glycerol) with antibiotics. After an additional 2-4 hours of growth, the culture broth was cooled to 25 °C when OD_{600} reached 0.8. A final concentration of 0.4 mM IPTG was added to induce protein expression followed by a 20-hour induction at 16 °C, 150 rotations per minute (rpm.). The cells were harvested at 4 °C and centrifuged at 10,000 rpm for five minutes.

Each His^6-SUMO-tagged peptide was purified using a denaturing method. The following buffers were used for purification: denature buffer A (50 mM Tris, 500 mM NaCl, 6 M guanidine hydrochloride, adjusted pH to 8.0 using HCl), denature buffer B (30 mM imidazole in denature buffer A) and buffer C (50 mM Tris, 500 mM NaCl, 500 mM imidazole, adjusted pH to 8.0 using HCl). The cells were first resuspended using denaturing buffer A and then lysed using Sonics VCX-750 Vibra-Cell Ultrasonic Liquid Processor with a 13 mm probe, 3 seconds on and 3 seconds off for 10 minutes. The cell lysis was then centrifuged at 4 °C, 10,000 rpm

for 1 hour, and the supernatant was loaded into pre-equilibrated Ni−NTA resin (using denaturing buffer A) for three times to maximize target peptide yield. 5 mL denaturing buffer B was used to wash away non-specifically bound proteins. The target peptide was then eluted by using 1 mL buffer C for 4 times, i.e., 4 mL in total. The purified peptide was desalted using PD-10 desalting column. TEV protease and trypsin digestion were conducted under 30 °C for 1 hour and 37 °C for 12 hours, respectively. After digestion, the reaction mixture was centrifuged at 10000 rpm for 10 minutes, and the supernatant was subjected to UPLC-HRMS analysis.

For purifying the P450 enzymes, the chaperone plasmid was transformed into *E. coli* BL21 first, generating a chaperone-based *E. coli* BL21-c. Then, plasmid GpeB and GpeC were individually transformed into E.coli BL21-c. After transformation, a single colony of each recombinant strain was inoculated into 3 mL fresh LB broth with 34 μg/mL chloramphenicol (native resistance for chaperone plasmid) and 50 μg/mL kanamycin (for pRSF-Duet-1 backbone) and shaken at 37 °C overnight. Then, 1 mL of the culture was incubated into 500 mL fresh LB medium (10g/L tryptone, 5g/L yeast extract, 10g/L Nacl) with antibiotics. After an additional 3 hours of growth, when OD_{600} reached 0.5, the culture broth was cooled to 30 °C, and a final concentrate of 6g/L arabinose, 1.5uM 5-aminolevulinic acid were incubated into the culture medium. Keep shaking until OD_{600} reaches 0.8. A final concentration of 0.4 mM IPTG was added to induce protein expression followed by a 20-hour induction at 24 °C, 150 rotations per minute (rpm.). The cells were harvested at 4 °C and centrifuged at 4900 rpm for an hour.

Each His-tagged P450 enzyme was purified using a naturing method. The following buffers were used for purification: buffer A (50 mM Tris, 500 mM NaCl, 5% glycerine, adjusted pH to 8.0 using HCl), buffer B (30 mM imidazole in buffer A) and buffer C (50 mM Tris, 500 mM NaCl, 500 mM imidazole, adjusted pH to 8.0 using HCl). The cells were first resuspended using buffer A and then lysed using Sonics VCX-750 Vibra-Cell Ultrasonic Liquid Processor with 13 mm probe, 3 seconds on and 3 seconds off for 15 minutes. The cell lysis was then centrifuged at 4 °C, 10,000 rpm for 1 hour, and the supernatant was loaded into pre-equilibrated Ni−NTA resin (using buffer A) for three times to maximize target protein yield. 5 mL buffer B was used to wash away non-specifically bound proteins. The target peptide was then eluted by using 1 mL buffer C for 4 times, i.e., 4 mL in total. The purified protein was desalted using PD-10 desalting column.

GpeC biaryl-cyclization assay

The P450 enzyme GpeC and substrate mixtures (50 μL) containing 0.2mM substrate, 1uM P450, 1.25uM Ferredoxin reductase, 2.5 uM Ferredoxin and 1mM NADPH in 50mM Tris-HCL buffer with 5% glycerol (pH 7.5) were incubated at 28 °C for 6h. The reactions were quenched by the addition of 100 μL methanol and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis.

Protein crystallization and structure determination

The GpeC protein was crystallized at a concentration of 10-14 mg/mL in SEC buffer with reservoir solution (10% w/v PEG 4000, 20% v/v glycerol, 0.03 M of magnesium chloride, 0.03M calcium chloride, 0.1 M MES/imidazole pH 7.25) by sitting drop vapor diffusion technique. After 7 days of growth at 16°C, crystals were observed in 0.8 μL sitting drops composed of a 1:1 mixture of protein solution and reservoir solution. The crystals were then cryoprotected by briefly soaking in the corresponding reservoir solution containing an additional 5% (v/v) ethylene glycol before being transferred to liquid nitrogen. Data collection was carried out on the BL18U1 beamline at Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, China). The dataset was indexed and integrated using XDS[57] and scaled by AIMLESS from the CCP4i suite[58]. The GpeC crystals achieved a diffraction resolution of 2.86 Å and belonged to the space group $P4_32_12$, with unit cell parameters a=b=78.40 Å, c=327.02 Å, α=β=γ=90.0°. There were two molecules in the asymmetric unit. The structure was phased by PHASER2[59] using predicted GpeC coordinates by Alphafold3[51] as the molecular replacement model. The models were completed through manual model building using COOT4[60], interspersed with crystallographic refinement using REFMAC5[59] with TLS restraints. Phenix.table_one[59] was utilized to create the crystallographic statistics table (excluding the Rwork/Rfree values), which indicated that 97.18% of the residues were in the most favored regions of the Ramachandran plot. Data collection, phasing, and refinement statistics are detailed in Table 1. Structure figures were generated using PyMOL2.3.4.

System Setup

The structure of protein P450GpeC was used to build the model system by crystallize structure. The native protein was called WT protein. And we used PyMol[61] Mutagenesis modular to mutate one residue of the original protein, which is called WT or F62G protein. The peptide ligands were docked into the active site of P450 protein via AutoDock Vina[62]. The associated atomic interactions between residues were described by the classical force field Amber ff14sb.

The molecular structures of peptides were optimized at the MP2/6-31G* level of theory by Orca[63]. The HEM parameters are from Cpd I[64].

The classical molecular dynamics simulation

The solvents were explicitly represented by the TIP3P model[65] under periodic boundary conditions (PBC). Na+ and Cl- were added to make neutralize the system. The initial atomic coordinates of WT/F62G systems were with implicit solvent. In an isothermal–isobaric (NPT) ensemble, the system temperature and pressure were maintained at 300 K and 1 bar using the respectively Langevin thermostat[66] and Monte Carlo barostat[67]. The electrostatic interactions were calculated via the Particle mesh Ewald (PME) method[68]. The cutoff of electrostatic and vdW interactions are both 1.2 nm. The simulation box is cubic, and its size is adjusted so that its boundary is more than 1.0 nm away from the protein. After energy minimization and NPT equilibrium, 200 ns of MD simulations were performed. The simulation of each system is performed three times in parallel. The time step was set as 2 fs. All classical MD simulations were performed with OpenMM[69]. Each system was simulated in parallel with both WT and F62G proteins. Some details of simulation systems are listed in Supplementary Table S1. And the binding free energy is calculated by MMPBSA.py[70] in AmberTools[71].

Acknowledgments

This work is partially funded by the Research Grants Council of Hong Kong (17115322, 17102123, and 17101324) to Y.-X.L. The authors would like to acknowledge and thank Yi-Man Eva Fung, Jo Yip, and Bonnie Yan for their help in MS and NMR analysis. We thank Yuqi Shi for her help with the primary bioassay.

Author contributions

R-Z.L. and B-B.H and Y-X.L. designed the research. H-Y.R. and C-P.P characterized the crystal structure. G-F.W. conducted the bioinformatic analysis. R-Z. L. characterized the BGC, isolated the compounds, and characterized the structures. S-Z.L performed molecular dynamics, R-Z. L. and J.L. contributed to in vitro enzyme characterization. R-Z. L and Z-M.S. conducted Advanced Marfey's experiment. All authors analyzed data and discussed results. All authors participated in preparing the manuscript. J-H. Z. and Y.X. L. supervised the project.

Competing Interests

The authors declare no competing interests.

\References

1 White, C. J. & Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nat Chem* **3**, 509-524 (2011). https://doi.org:10.1038/nchem.1062

2 Sussmuth, R. D. & Mainz, A. Nonribosomal Peptide Synthesis-Principles and Prospects. *Angew Chem Int Ed Engl* **56**, 3770-3821 (2017). https://doi.org:10.1002/anie.201609079

3 Montalban-Lopez, M. *et al*. New developments in RiPP discovery, enzymology and engineering. *Nat Prod Rep* **38**, 130-239 (2021). https://doi.org:10.1039/d0np00027b

4 Levine, D. P. Vancomycin: A History. *Clinical Infectious Diseases* **42:S5–12** (2005).

5 Smith, P. A. *et al*. Optimized arylomycins are a new class of Gram-negative antibiotics. *Nature* **561**, 189-194 (2018). https://doi.org:10.1038/s41586-018-0483-6

6 Imai, Y. *et al*. A new antibiotic selectively kills Gram-negative pathogens. *Nature* **576**, 459-464 (2019). https://doi.org:10.1038/s41586-019-1791-1

7 Ji, X., Nielsen, A. L. & Heinis, C. Cyclic Peptides for Drug Development. *Angew Chem Int Ed Engl* **63**, e202308251 (2024). https://doi.org:10.1002/anie.202308251

8 Yang, Y., Lan, J. & You, J. Oxidative C-H/C-H Coupling Reactions between Two (Hetero)arenes. *Chem Rev* **117**, 8787-8863 (2017). https://doi.org:10.1021/acs.chemrev.6b00567

9 Kozlowski, M. C. Oxidative Coupling in Complexity Building Transforms. *Acc Chem Res* **50**, 638-643 (2017). https://doi.org:10.1021/acs.accounts.6b00637

10 Yu, L., Nagata, Y. & Nakamura, H. Atroposelective Total Synthesis of Cihunamide B. *J Am Chem Soc* **146**, 2549-2555 (2024). https://doi.org:10.1021/jacs.3c11016

11 Zhang, J., Yu, L., Ogawa, H., Nagata, Y. & Nakamura, H. Modular, Scalable Total Synthesis of Lapparbin with a Noncanonical Biaryl Linkage. *Angew Chem Int Ed Engl*, e202409987 (2024). https://doi.org:10.1002/anie.202409987

12 Ben-Lulu, M., Gaster, E., Libman, A. & Pappo, D. Synthesis of Biaryl-Bridged Cyclic Peptides via Catalytic Oxidative Cross-Coupling Reactions. *Angew Chem Int Ed Engl* **59**, 4835-4839 (2020). https://doi.org:10.1002/anie.201913305

13 Sun, H., Zhang, H., Ang, E. L. & Zhao, H. Biocatalysis for the synthesis of pharmaceuticals and pharmaceutical intermediates. *Bioorg Med Chem* **26**, 1275-1284 (2018). https://doi.org:10.1016/j.bmc.2017.06.043

14 Rivera, D. G., Ojeda-Carralero, G. M., Reguera, L. & Van der Eycken, E. V. Peptide macrocyclization by transition metal catalysis. *Chem Soc Rev* **49**, 2039-2059 (2020). https://doi.org:10.1039/c9cs00366e

15 Buller, R. *et al*. From nature to industry: Harnessing enzymes for biocatalysis. *Science* **382**, eadh8615 (2023). https://doi.org:10.1126/science.adh8615

16 Nicholas S. Sarai1†‡, T. J. F., Ryen L. O'Meara1†, Kadina E. Johnston2§, Sabine Brinkmann-Chen1, Ryan R. Maar3, Ron E. Tecklenburg3, John M. Roberts3, Jordan C. T. Reddel, Dimitris E. Katsoulis, Frances H. Arnold. Directed evolution of enzymatic silicon-carbon bond cleavage in siloxanes. *Science* (2024).

17 Longwitz, L., Leveson-Gower, R. B., Rozeboom, H. J., Thunnissen, A. W. H. & Roelfes, G. Boron catalysis in a designer enzyme. *Nature* **629**, 824-829 (2024). https://doi.org:10.1038/s41586-024-07391-3

18 Kissman, E. N. *et al*. Expanding chemistry through in vitro and in vivo biocatalysis. *Nature* **631**, 37-48 (2024). https://doi.org:10.1038/s41586-024-07506-w

19 Xu, G. *et al*. Cryptic enzymatic assembly of peptides armed with beta-lactone warheads. *Nat Chem Biol* (2024). https://doi.org:10.1038/s41589-024-01657-7

20 Budimir, Z. L. *et al*. Biocatalytic cyclization of small macrolactams by a penicillin-binding protein-type thioesterase. *Nat Chem Biol* **20**, 120-128 (2024). https://doi.org:10.1038/s41589-023-01495-z

21 Maolin Li, Y. Y., Wesley Harrison, Zhengyi Zhang, Huimin Zhao. Asymmetric photoenzymatic incorporation of fluorinated motifs into olefins. *Science* **385, 416–421** (2024).

22 Podust, L. M. & Sherman, D. H. Diversity of P450 enzymes in the biosynthesis of natural products. *Nat Prod Rep* **29**, 1251-1266 (2012). https://doi.org:10.1039/c2np20020a

23 Zetzsche, L. E. *et al*. Biocatalytic oxidative cross-coupling reactions for biaryl bond formation. *Nature* **603**, 79-85 (2022). https://doi.org:10.1038/s41586-021-04365-7

24 Richter, D. & Piel, J. Novel types of RiPP-modifying enzymes. *Curr Opin Chem Biol* **80**, 102463 (2024). https://doi.org:10.1016/j.cbpa.2024.102463

25 Ben-Lulu, M., Gaster, E., Libman, A. & Pappo, D. Synthesis of Biaryl-Bridged Cyclic Peptides via Catalytic Oxidative Cross-Coupling Reactions. *Angewandte Chemie* **132**, 4865-4869 (2020).

26 Hansen, M. H., Stegmann, E. & Cryle, M. J. Beyond vancomycin: recent advances in the modification, reengineering, production and discovery of improved glycopeptide antibiotics to tackle multidrug-resistant bacteria. *Curr Opin Biotech* **77**, 102767 (2022).

27 Molinaro, C. *et al*. Engineered Cytochrome P450-Catalyzed Oxidative Biaryl Coupling Reaction Provides a Scalable Entry into Arylomycin Antibiotics. *Journal of the American Chemical Society* **144**, 14838-14845 (2022).

28 Nam, H. *et al*. Exploring the Diverse Landscape of Biaryl-Containing Peptides Generated by Cytochrome P450 Macrocyclases. *Journal of the American Chemical Society* (2023).

29 Dan Kahne, C. L., Wei Lu, Christopher Walsh. Glycopeptide and Lipoglycopeptide Antibiotics

. *Chemical Reviews* **105**, 425−448 (2005).

30 Forneris, C. C. & Seyedsayamdost, M. R. In Vitro Reconstitution of OxyC Activity Enables Total Chemoenzymatic Syntheses of Vancomycin Aglycone Variants. *Angew Chem Int Ed Engl* **57**, 8048-8052 (2018). https://doi.org:10.1002/anie.201802856

31 Woithe., K. *et al*. Oxidative Phenol Coupling Reactions Catalyzed by OxyB: ACytochrome P450 from the Vancomycin Producing Organism.Implications for Vancomycin Biosynthesis. *J Am Chem Soc* **129**, 6887-6995 (2007).

32 Yim, G., Thaker, M. N., Koteva, K. & Wright, G. Glycopeptide antibiotic biosynthesis. *J Antibiot (Tokyo)* **67**, 31-41 (2014). https://doi.org:10.1038/ja.2013.117

33 Mordhorst, S., Ruijne, F., Vagstad, A. L., Kuipers, O. P. & Piel, J. Emulating nonribosomal peptides with ribosomal biosynthetic strategies. *RSC Chem Biol* **4**, 7-36 (2023). https://doi.org:10.1039/d2cb00169a

34 Hetrick, K. J. & van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr Opin Chem Biol* **38**, 36-44 (2017). https://doi.org:10.1016/j.cbpa.2017.02.005

35 Tang, M. C., Zou, Y., Watanabe, K., Walsh, C. T. & Tang, Y. Oxidative Cyclization in Natural Product Biosynthesis. *Chem Rev* **117**, 5226-5333 (2017). https://doi.org:10.1021/acs.chemrev.6b00478

36 Wyche, T. P., Ruzzini, A. C., Schwab, L., Currie, C. R. & Clardy, J. Tryptorubin A: A Polycyclic Peptide from a Fungus-Derived Streptomycete. *J Am Chem Soc* **139**, 12899-12902 (2017). https://doi.org:10.1021/jacs.7b06176

37 An, J. S. *et al*. Discovery and Biosynthesis of Cihunamides, Macrocyclic Antibacterial RiPPs with a Unique C-N Linkage Formed by CYP450 Catalysis. *Angew Chem Int Ed Engl* **62**, e202300998 (2023). https://doi.org:10.1002/anie.202300998

38 He, B. B. *et al*. Bacterial Cytochrome P450 Catalyzed Post-translational Macrocyclization of Ribosomal Peptides. *Angew Chem Int Ed Engl* **62**, e202311533 (2023). https://doi.org:10.1002/anie.202311533

39 Hu, Y. L. *et al*. P450-Modified Ribosomally Synthesized Peptides with Aromatic Cross-Links. *J Am Chem Soc* **145**, 27325-27335 (2023). https://doi.org:10.1021/jacs.3c07416

40 Nam, H. *et al*. Exploring the Diverse Landscape of Biaryl-Containing Peptides Generated by Cytochrome P450 Macrocyclases. *J Am Chem Soc* (2023). https://doi.org:10.1021/jacs.3c07140

41 Liu, C. L. *et al*. P450-Modified Multicyclic Cyclophane-Containing Ribosomally Synthesized and Post-Translationally Modified Peptides. *Angew Chem Int Ed Engl* **63**, e202314046 (2024). https://doi.org:10.1002/anie.202314046

42 Hug, J. J. *et al*. Biosynthesis of Cittilins, Unusual Ribosomally Synthesized and Post-translationally Modified Peptides from Myxococcus xanthus. *ACS Chem Biol* **15**, 2221-2231 (2020). https://doi.org:10.1021/acschembio.0c00430

43 Zdouc, M. M. *et al*. A biaryl-linked tripeptide from Planomonospora reveals a widespread class of minimal RiPP gene clusters. *Cell Chem Biol* **28**, 733-739 e734 (2021). https://doi.org:10.1016/j.chembiol.2020.11.009

44 Pattanayakanahalli Henjarappa, K., Das, S. & Mahanta, N. A new class of macrocyclic peptides. *Nat Chem Biol* (2024). https://doi.org:10.1038/s41589-023-01507-y

45 Zhao, Y. *et al*. Cytochrome P450(Blt) Enables Versatile Peptide Cyclisation to Generate Histidine- and Tyrosine-Containing Crosslinked Tripeptide Building Blocks. *Angew Chem Int Ed Engl* **61**, e202204957 (2022). https://doi.org:10.1002/anie.202204957

46 Treisman, M. *et al*. An Engineered Biarylitide Cross-Linking P450 from RiPP Biosynthesis Generates Alternative Cyclic Peptides. *Org Lett* **26**, 1828-1833 (2024). https://doi.org:10.1021/acs.orglett.3c04366

47 He, B. B. *et al.* Expanded Sequence Space of Radical S-Adenosylmethionine-Dependent Enzymes Involved in Post-translational Macrocyclization. *Angew Chem Int Ed Engl* **61**, e202212447 (2022). https://doi.org:10.1002/anie.202212447
48 Cheng, Z. *et al.* Rule-based omics mining reveals antimicrobial macrocyclic peptides against drug-resistant clinical isolates. *Nat Commun* **15**, 4901 (2024). https://doi.org:10.1038/s41467-024-49215-y
49 Cai, P. *et al.* Uncovering Untapped Carboxylic Acid Reductases (CARs) for One-Step Biosynthesis and Diversification of Bioactive Nitrogen-Containing Heterocycles. *ACS Catalysis* **13**, 15404-15416 (2023). https://doi.org:10.1021/acscatal.3c03670
50 Phillips, I. R. & PaulR.OrtizdeMontellano, E. A. S. (Methods in Molecular Biology).
51 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024). https://doi.org:10.1038/s41586-024-07487-w
52 Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Research* **50**, W210-W215 (2022). https://doi.org:10.1093/nar/gkac387
53 Hansen, M. H. *et al.* Structural Insights into a Side Chain Cross-Linking Biarylitide P450 from RiPP Biosynthesis. *ACS Catalysis* **14**, 812-826 (2024). https://doi.org:10.1021/acscatal.3c05417
54 Hetmann, M. *et al.* Identification and validation of fusidic acid and flufenamic acid as inhibitors of SARS-CoV-2 replication using DrugSolver CavitomiX. *Scientific Reports* **13**, 11783 (2023). https://doi.org:10.1038/s41598-023-39071-z
55 Steinegger, M. & Soding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* **35**, 1026-1028 (2017). https://doi.org:10.1038/nbt.3988
56 Tian, W. *et al.* Efficient biosynthesis of heterodimeric C(3)-aryl pyrroloindoline alkaloids. *Nat Commun* **9**, 4428 (2018). https://doi.org:10.1038/s41467-018-06528-z
57 Brehm, W., Trivino, J., Krahn, J. M., Uson, I. & Diederichs, K. XDSGUI: a graphical user interface for XDS, SHELX and ARCIMBOLDO. *Journal of Applied Crystallography* **56**, 1585-1594 (2023). https://doi.org:doi:10.1107/S1600576723007057
58 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011). https://doi.org:10.1107/s0907444910045749
59 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). https://doi.org:10.1107/s0907444909052925
60 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004). https://doi.org:10.1107/s0907444904019158
61 Maya, C. Using PyMOL to Understand Why COVID-19 Vaccines Save Lives. *J Chem Educ* **100**, 1351-1356 (2023). https://doi.org:10.1021/acs.jchemed.2c00779
62 Forli, S. *et al.* Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc* **11**, 905-919 (2016). https://doi.org:10.1038/nprot.2016.051

63 Neese, F., Wennmohs, F., Becker, U. & Riplinger, C. The ORCA quantum chemistry program package. *J Chem Phys* **152**, 224108 (2020). https://doi.org:10.1063/5.0004608

64 Shahrokh, K., Orendt, A., Yost, G. S. & Cheatham, T. E., 3rd. Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle. *J Comput Chem* **33**, 119-133 (2012). https://doi.org:10.1002/jcc.21922

65 Mark, P. & Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A* **105**, 9954-9960 (2001).

66 Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry* **18**, 1463-1472 (1997).

67 Gomez, Y. K. *et al*. Taking the Monte Carlo gamble: how not to buckle under the pressure! *Biophysical Journal* **121**, 35a (2022).

68 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics* **98**, 10089-10092 (1993).

69 Eastman, P. *et al*. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology* **13**, e1005659 (2017).

70 Miller III, B. R. *et al*. MMPBSA. py: an efficient program for end-state free energy calculations. *Journal of chemical theory and computation* **8**, 3314-3321 (2012).

71 Case, D. A. *et al*. The Amber biomolecular simulation programs. *Journal of computational chemistry* **26**, 1668-1688 (2005).