LimF is a versatile prenyltransferase catalyzing histidine-*C*-geranylation on diverse nonnatural substrates

Yuchen Zhang[1,†], Keisuke Hamada[2,†], Dinh Thanh Nguyen[1], Sumika Inoue[1], Masayuki Satake[1], Shunsuke Kobayashi[2], Chikako Okada[2], Kazuhiro Ogata[2], Masahiro Okada[3], Toru Sengoku[2,*], Yuki Goto[1,*], and Hiroaki Suga[1,*]

Affiliation

[1]Department of Chemistry, Graduate School of Science, The University of Tokyo, Bunkyo, Tokyo 113-0033, Japan

[2]Department of Biochemistry, Graduate School of Medicine, Yokohama City University, Yokohama 236-0004, Japan

[3]Department of Material and Life Chemistry, Kanagawa University, Rokkakubashi, Kanagawa 221-8686, Japan

[†] These authors contributed equally.

[*] Corresponding authors: Suga, Hiroaki (hsuga@chem.s.u-tokyo.ac.jp), Goto, Yuki (y-goto@chem.s.u-tokyo.ac.jp), and Sengoku, Toru (tsengoku@yokohama-cu.ac.jp)

Abstract

Prenylation plays an important role in diversifying structure and function of secondary metabolites. Although several cyanobactin prenyltransferases have been characterized, their modes of action are mainly limited to the modification of electron-rich hetero atoms. Here we report a unique prenyltransferase originating from *Limnothrix sp.* CACIAM 69d, referred to as LimF, which catalyzes an unprecedented His-*C*-geranylation. Interestingly, LimF executes the geranylation on not only its native peptide substrate but also a wide range of exotic peptides, including thioether-closed macrocycles. We have also serendipitously uncovered an ability of Tyr-*O*-geranylation as the secondary function of LimF, indicating it is an unusual bifunctional prenyltransferase. Crystallographic analysis of LimF complexed with a pentapeptide substrate and a prenyl donor analog provides structural basis for its unique His recognition and its bifunctionality. Lastly, we show the LimF's prenylation ability on various bioactive molecules containing an imidazole group, highlighting its potential as a versatile biocatalyst for site-specific geranylation.

Introduction

Prenylation is a ubiquitous modification distributed in primary metabolism as well as biosynthetic pathways of diverse natural products, including terpenoids,[1-3] alkaloids,[4,5] polyketides,[6,7] flavonoids,[8,9] nonribosomal peptides,[10,11] and ribosomally synthesized and post-translationally modified peptides (RiPPs)[12-14]. Prenyl groups can contribute an enhancement of lipophilicity to the molecules, aiding their interaction with lipid membranes; and consequently they grant better membrane permeability and bioavailability.[15,16] Since unmodified peptides are generally deficient in such properties,[17,18] installation of prenyl group(s) to peptides potentially enhances not only bioactivities but also pharmacokinetics[19,20]. Thus, a set of prenylation biocatalysts can provide a unique toolbox to generate *de novo* peptides and even pseudo-natural products[21] with desired therapeutical properties.

Prenyltransferases (PTases) involved in the biosynthesis of cyanobactins,[22-24] a class of RiPPs, are enzymes that prenylate cognate peptidic substrates. Such enzymes belong to ABBA-type PTases[25,26] and commonly adopt a truncated α/β PTase barrel fold with a unique solvent-exposed cavity where binding of peptide substrate to form a hydrophobic active site is necessary for the activity.[27] This unique architecture leads to strict selectivity for a particular prenylation mode (*i.e.*, specific recognition of a prenyl donor and a modifying residue), but yet relaxed preference for global substrate peptide sequences, making them promising biocatalysts for peptide alkylation. Despite their similar tertiary structures, these PTases catalyze diverse modes of prenylation, including Tyr-*O*-prenylation,[27-30] Trp-*C*- and *N*-prenylation,[31-33] Arg-*N*-prenylation,[34] Ser/Thr-*O*-prenylation,[35-39] and terminal-*N*- and *O*-prenylation[40,41]. While many cyanobactin PTases execute prenylation on electron-rich heteroatoms, only a single cyanobactin *C*-PTase, KgpF catalyzing Trp-*C*-dimethylallylation in kawaguchipeptin biosynthesis, is currently characterized.[31] To develop the *C*-PTase toolbox, we have been interested in discovering a new cyanobactin PTase with a novel *C*-prenylation mode.

Here, we report the discovery of a PTase from *Limnothrix sp*. CACIAM 69d, referred to as LimF, that catalyzes unprecedented His-*C*-geranylation. We have demonstrated that LimF is the first PTase capable of catalyzing two distinct modes of prenylations, His-*C*-geranylation and Tyr-*O*-geranylation. Our X-ray crystallographic analysis of LimF in complex with a substrate and a prenyl donor analog has revealed the structural basis for its imidazole ring recognition and its bifunctionality of prenylation towards two distinct residues. Most importantly, LimF catalyzes *C*-geranylation on His residue(s) in an array of acyclic and cyclic peptides as well as imidazole-containing small molecules, thus providing a versatile His-*C*-geranylation

catalyst.

Results

LimF is a histidine-*C*-geranyltransferase

To mine potential peptide PTases with *C*-prenylation activity, proteins homologous to the sole characterized cyanobactin *C*-PTase, KgpF, were searched by means of the basic local alignment search tool (BLAST)[42]. To explore the sequence-function relationships of the potential peptide PTases, a sequence similarity network (SSN)[43] was generated for the KgpF-queried BLAST hits. The resulting SSN revealed several clusters with putative functions including Tyr-*O*-prenylation, terminal-*N*-/*O*-prenylation, Arg-*N*-prenylation, Trp-*N*-/*C*-prenylation and uncharacterized ones (**Fig. S1**). Among the uncharacterized proteins, an outlier protein from *Limnothrix* sp. CACIAM 69d drew our attention. A phylogenetic analysis of this protein with other characterized cyanobactin PTases revealed that it forms a distinct clade rooted between Tyr-*O*-PTases and Trp-*C*-/*N*-PTases (**Fig. 1a**). Although this protein was annotated as a homolog of PirF (Tyr-*O*-PTase) in a previous report[44], both the SSN and phylogenetic analyses implied that it would exhibit unique activity. Therefore, we decided to name this protein LimF and investigate it further.

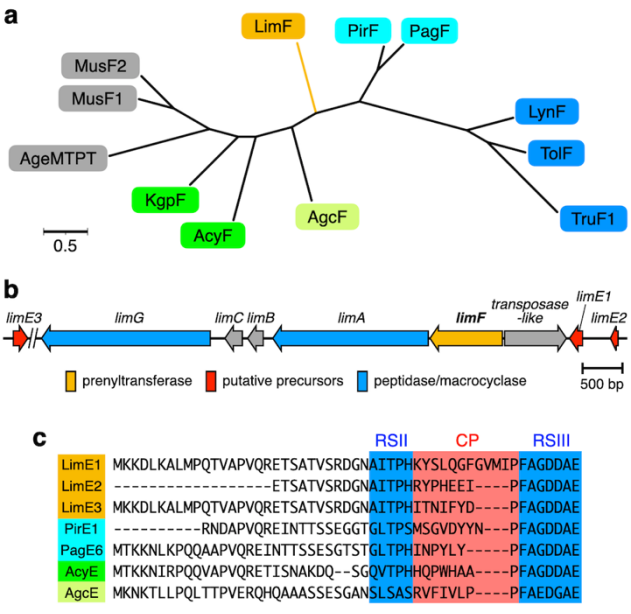


Fig 1. A putative cyanobactin BGC from *Limnothrix* sp. CACIAM 69d including a unique PTase LimF. a, Phylogenetic analysis of LimF with characterized cyanobactin PTases. For AgeMTPT, the PTase domain region (275-559) was used for alignment. LimF, Ser/Thr-*O*-PTases, Tyr-*O*-PTases, Arg-*N*-PTase, Trp-*N*/*C*-PTases, and terminal-*N*/*O*-PTases are highlighted with orange, blue, light blue, light green, green, and gray, respectively. **b**, Cyanobactin-related genes in the cyanobacterium *Limnothrix* sp. CACIAM 69d. The *limE3* gene is encoded in remote genetic loci.

c, Sequence alignment of the putative precursor peptides in the *lim* BGC (LimE1–3) with some characterized cyanobactin precursors. Recognition sequences for A-family proteases (RSII) and G-family macrocyclases (RSIII) are highlighted with blue. The core peptide (CP) region is highlighted with red. For LimE2 and PirE1, a part of the N-terminal leader peptide sequence is missing probably because these genes are located at the end of sequencing contig.

The *limF* gene is included in a typical cyanobactin BGC, which also contains cyanobactin protease (*limA*) and macrocyclase (*limG*) as well as putative precursor peptides (*limE1* and *limE2*) (**Fig. 1b** and **Table S1**). Manual curation pointed to an additional putative precursor encoded in remote genetic loci (*limE3*). The LimE1–3 peptides shared notable similarity with previously characterized cyanobactin precursors[45], and their sequence alignment analysis highlighted signature recognition sequences (RSII and RSIII) for LimA and G (**Fig. 1c**). In the general biosynthesis pathways of cyanobactins, the A-family protease clips off the N-terminal RSII, and the G-family macrocyclase subsequently cleaves at the C-terminal RSIII to yield backbone-cyclized core peptide (CP) region.[22,46,47] The resulting macrocyclic CPs are often modified further by other tailoring enzymes including PTases to give the final cyanobactin product.[48,49] On the basis of analogy with other cyanobactin BGCs and the conserved RS motifs, we identified the CP regions of LimE1–3 peptides and postulated that their backbone-cyclized forms (**Fig. S2**; bcLimE1, bcLimE2, and bcLimE3) could be potential substrates for LimF.

To examine the enzymatic activity of LimF, it was heterologously expressed in *Escherichia coli* (**Fig. S3**) and incubated with chemically synthesized bcLimE1, E2, and E3 in the presence of $MgCl_2$ and various prenyl donors, such as GPP and others (**Fig. S4**). LC-MS analysis of the products revealed that geranylation of bcLimE2 quantitatively occurred in the presence of GPP, while other prenyl donors could not promote efficient prenylation (**Fig. 2a**). On the other hand, neither bcLimE1 nor E3 were active substrates for LimF regardless of the donors (**Fig. S4**). MS/MS analysis of the geranylated bcLimE2 suggested that a His residue present in bcLimE2 was the geranylation site (**Fig. S5**). NMR analysis of the geranylated bcLimE2 gave heteronuclear multiple bond correlations (HMBC) including the key interrelationship between geranyl-C1-methylene and imidazole-C2 (**Fig. 2b** and **S6**). These observations allowed us to conclude that the geranylation occurred at the C2 position of the His sidechain in the forward direction, *i.e.,* prenylation at the primary carbocation. Collectively, we identified the structure of a putative *lim* BGC product, limnothamide (**Fig. 2b**), and revealed that LimF is a unique PTase catalyzing the His-*C2*-forward geranylation.

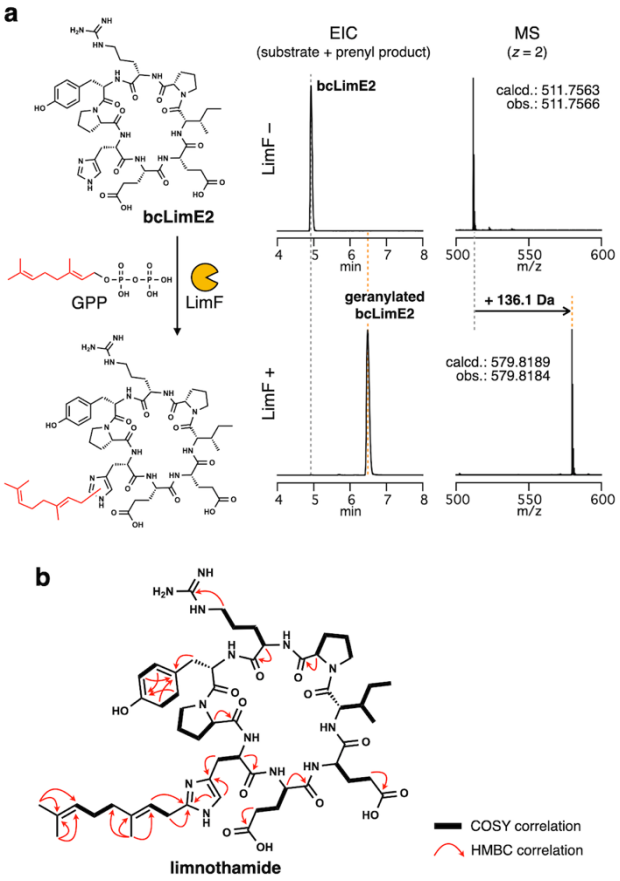


Fig 2. *In vitro* LimF treatment of bcLimE2 confirms the His-*C*2-forward geranylation in limnothamide. a, LC-MS analysis of bcLimE2 before and after 12 h LimF treatment. The displayed extracted ion chromatograms (EICs) include all possible charge states of the substrate bcLimE2 and singly geranylated product. Mass spectra integrated over the peak observed in the corresponding EICs are also shown. **b**, Structure of limnothamide characterized by NMR analyses. The key COSY and HMBC correlations, identifying the forward geranylation on His-C2, are shown. The NMR spectra are shown in **Fig. S6**.

To further characterize LimF, we first optimized the bcLimE2 geranylation reaction conditions (**Supplementary Text 1** and **Fig. S7**), being 1 mM GPP, 40 mM $MgCl_2$, 1 mM DTT, pH 7.2, and incubation at 25°C. We then performed the steady-state kinetic study, determining its parameters to be k_{cat} = 13.8 ± 0.6 min^{-1}, K_M = 0.31 ± 0.04 mM (**Table 1 entry 1**, **Fig. S7g**). The parameters revealed that LimF is one of the most efficient cyanobactin PTases characterized so far (**Table S2**). We also found that LimF was able to geranylate not only the native bcLimE2 substrate but also other His-containing molecules, *e.g.* a dipeptide (Ac-Ala-His-NH_2) and simple His monomers (H-His-OH and Fmoc-His-OH in **Fig. S8**). The NMR analyses of the geranylated dipeptide confirmed the geranylated His structure consistent with the native limnothamide (**Fig. S9**). Collectively, these results suggested that LimF possesses a catalytic function of selective His-*C2*-

forward geranylation but also substrate tolerance. This encouraged us to gain more insights into the substrate scope of LimF.

FIT-LimF system enables *in vitro* biosynthesis of diverse geranylated teMPs

The thioether-closing macrocycle has proven to offer a remarkable scaffold for devising *de novo* peptide ligands against target proteins of interest.[21,50] We wondered if LimF could modify a His residue in such thioether-macrocyclic peptides (teMPs). To test this possibility, a teMP (teLimE2) that mimics the sequence of bcLimE2 was designed and expressed by means of genetic code reprogramming powered by a custom-made cell-free translation, so-called flexible in vitro translation (FIT) system (**Fig. 3a**).[51] The *in vitro* expressed teLimE2 was then incubated with LimF and analyzed by LC-MS, indicating that teLimE2 underwent quantitative geranylation to form teLimE2-H^{Ger}. We refer to the LimF-coupled *in vitro* translation system as the FIT-LimF system, enabling us to readily conduct preparation of diverse substrate variants and following LimF-catalyzed geranylation in a one-pot manner. We envisaged that it would facilitate in-depth studies on the catalytic scope of LimF geranylating various teMPs.

As RiPP enzymes are often affected by the local sequence environment nearby the modification site,[52-56] we commenced with mutagenesis of the residues adjacent to the geranylation site. We expressed thirty-six teLimE2 mutants, in which the His–1 and His+1 positions are replaced with 19 proteinogenic amino acids (expect Cys), in the FIT-LimF system, and the reaction outcomes were analyzed by LC-MS (**Fig. 3b** and **S10**). Clearly, LimF has a preference to certain amino acids at the –1 position; small or hydrophobic residues (Thr, Ala, Met, Pro, Ser and Val) at this position are favored whereas bulky or charged residues are not well tolerated. On the other hand, the +1 position is highly variable with a single exception for Pro that abolishes the geranylation (**Fig. 3c** and **S11**).

In the next series of experiments, we examined whether LimF could geranylate teMPs with various ring sizes. We designed 13 teMP variants bearing the favored local sequence motif (Ser-His-Ala/Cys) in different ring sizes (from 4 to 18 residues). One-pot expression and modification of these teMPs in the FIT-LimF system showed that all of them underwent near-quantitative geranylation (**Fig. 3d** and **S12a**), demonstrating that LimF is capable of geranylating teMPs regardless of the ring size.

Encouraged by the observed relaxed substrate preference of LimF, we next tested whether LimF could modify teMPs with diverse global sequence compositions. Five distinct teMPs composed of arbitrarily chosen sequences (teMP1–5) were assessed by the FIT-LimF system (**Fig. 3e** and **S12b**). As anticipated,

LimF was able to geranylate teMP1(–1P) and teMP2(+1W), which contain a tolerable residue at –1 and +1 positions (**Fig. 3e**). When this residue was intentionally mutated to a disfavored residue, teMP1(–1K) and teMP2(+1P), according to the data shown in **Fig. 3b** and **3c**, their modification was drastically suppressed. On the other hand, teMP3–5, which have an unfavorable residue at either position, gave poor or no geranylation expectedly. However, replacement of the respective residues to preferred residues in teMP3(–1S), teMP4(–1A), and teMP5(+1A) completely rescued the LimF modification. These results suggest that the rule determined in **Fig. 3b** and **3c** is reliable to design the "LimF-modifiable" peptide sequences regardless of their length, global sequence composition, and ring size.

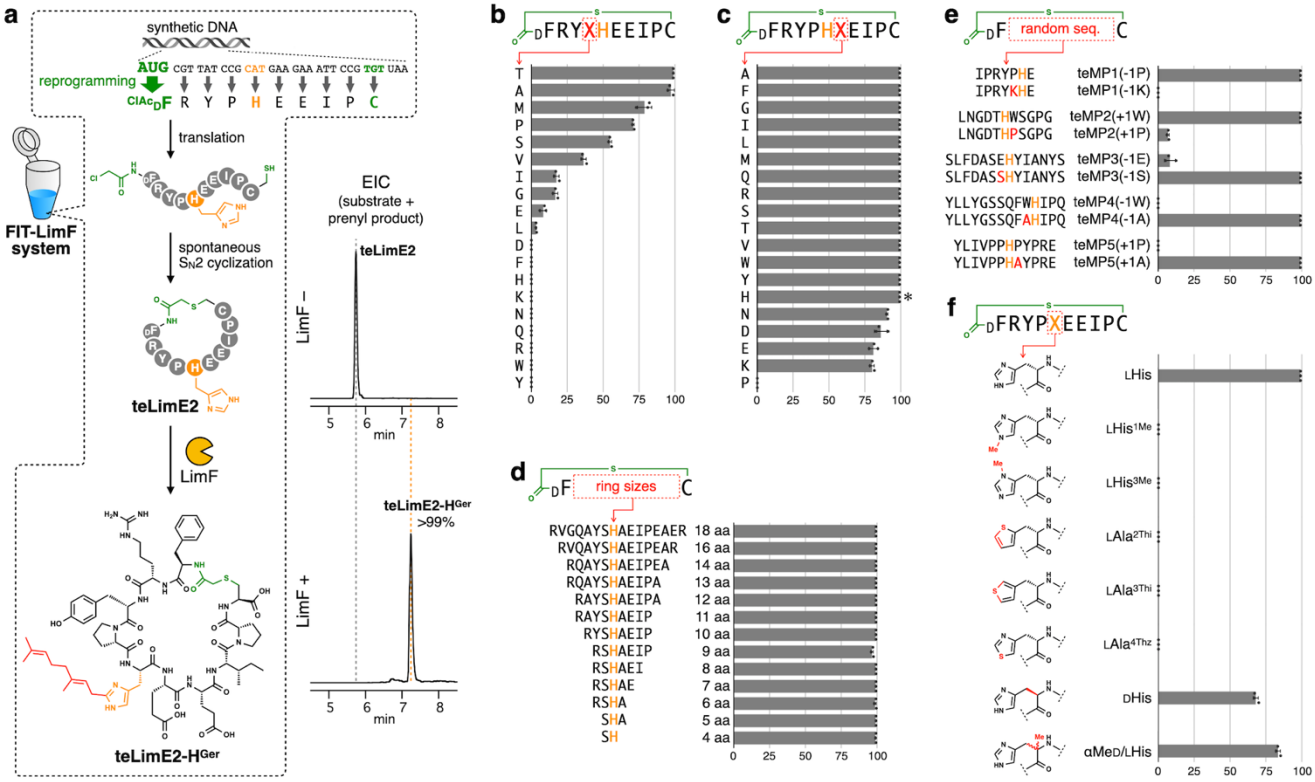


Fig 3. One-pot in vitro biosynthesis of geranylated teMPs in the FIT-LimF system validates the broad substrate tolerance of LimF. a, One-pot expression and modification of teLimE2 in the FIT-LimF system. In this experiment, the AUG initiation codon was reprogrammed with *N*-chloroacetyl-d-phenylalanine (^{ClAc}DF) to express a liner precursor bearing an N-terminal ClAc group, which underwent spontaneous ring-closing S_N2 reaction with a downstream Cys thiol to form a thioether linkage. The displayed EICs include all possible charge states of the substrate teLimE2 and its corresponding geranylated product. The conversion efficiency calculated based on EIC peak areas is shown in the chromatogram. **b**, The –1 position mutagenesis of teLimE2. The bar graph indicates conversion efficiencies of the mutants in LimF-catalyzed geranylation. Data points in triplicate experiments are shown as black dots, and the error bars indicate standard deviations. **c**, The +1 position mutagenesis of teLimE2. The asterisk indicates that His at the –1 position was modified in this mutant. **d**, LimF modification of teMPs with different ring sizes. **e**, LimF modification

of teMPs with arbitrary global sequence compositions (teMP1–5) and their –1 or +1 position mutants. **f**, Modification of non-canonical His analogs incorporated via genetic code reprogramming. The LHis data indicates a control experiment, in which teLimE2 was expressed by the genetic code reprogramming with LHis. The observed comparable geranylation efficiency to the teLimE2 translated by the canonical genetic code confirmed that the reprogrammed FIT-LimF system works well. Structural alternation on each analog compared with the canonical His residue is highlighted in red. The raw LC-MS data are shown in **Fig. S10–S13**.

LimF modifies exotic His analogs with altered mainchain structures

Intrigued by the broad substrate tolerance of LimF, we wondered if the His residue in teLimE2 could be replaceable to its exotic analogs, yielding non-canonical geranylated residues. We expressed teLimE2 derivatives in which the His residue is substituted with various analogs by the genetic code reprogramming facilitated by the FIT-LimF system (**Fig. 3f** and **S13**). N_τ- and N_π-methylated His analogs were not geranylated by LimF, suggesting that the imidazole NH could be involved in the recognition by LimF (**Fig. 3f**, LHis^{1Me} and LHis^{3Me}). Likewise, LimF could not geranylate noncanonical residues bearing thiophene (LAla^{2Thi} and LAla^{3Thi}) or thiazole (LAla^{4Thz}). In contrast, LimF was clearly able to geranylate DHis and α-methyl-D/LHis (αMeD/LHis). Thus, LimF critically recognizes the imidazole sidechain but not the mainchain structure. This intriguing property of LimF in turn indicates a high utility potential of the FIT-LimF system for the production of teMPs bearing exotic His analogs with altered mainchain structures.

LimF is an unusual bifunctional prenyltransferase

While testing various teMPs with arbitrary sequences, we serendipitously discovered a teMP (teMP6) that gave two geranylated products after LimF treatment (**Fig. 4a**). Even in a mutant in which His is mutated with Ala (teMP6-H5A), one geranylated product was produced (**Fig. S14**), indicating that LimF would install geranyl group on a non-His residue. Notably, a loss of the geranyl moiety ($C_{10}H_{17}$) during ESI ionization was observed in such peptides (**Fig. 4b**), hinting us to speculate geranylation on a heteroatom.[28,45] The MS/MS analysis of the geranylated teMP6-H5A identified that LimF geranylated Tyr at the 7th position (**Fig. S15**). To facilitate the NMR characterization, a model tripeptide (Ac-Ala-Thr-Tyr-NH_2) was synthesized and modified by LimF, giving the corresponding Tyr-geranylated product (**Fig. S16**). The key COSY and HMBC correlations of the product determined the geranylation site to be the phenolic OH group (**Fig. 4c** and **Fig. S17**). Collectively, we have demonstrated that LimF is capable of catalyzing two distinct modes of prenylation, His-*C*-geranylation and Tyr-*O*-geranylation.

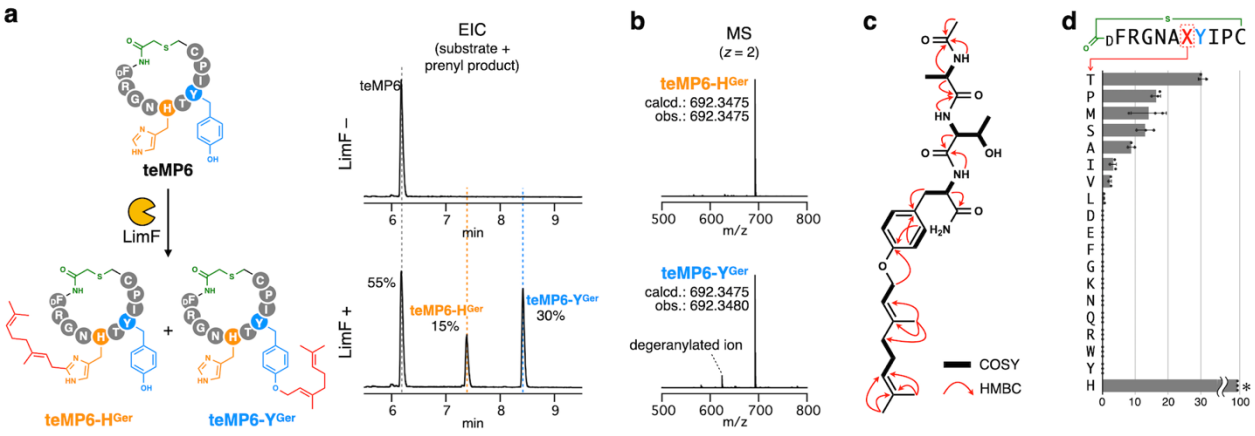


Fig 4. Tyr-*O*-geranylation occurs as the secondary prenylation mode of LimF. a, LimF modification of teMP6 forming two different geranylated products. The displayed EICs include all possible charge states of the substrate teMP6 and its singly geranylated product. The peaks corresponding to His-geranylated and Tyr-geranylated products (teMP6-H^{Ger} and teMP6-Y^{Ger}) are highlighted with orange and light blue dotted lines. Conversion efficiency calculated based on EIC peak areas are shown in the figure. **b.** The mass spectra of teMP6-H^{Ger} and teMP6-Y^{Ger}. The peak derived from a loss of the geranyl moiety during ionization is annotated. **c**, Structure of a geranylated tripeptide characterized by NMR analyses. The key 2D-NMR correlations identifying Tyr-*O*-geranylation are shown. The NMR spectra are shown in **Fig. S17**. **d,** The –1 position mutagenesis of teMP6-H5A. The bar graph indicates conversion efficiencies of the mutants in LimF-catalyzed Tyr-*O*-geranylation. The asterisk indicates that His at the –1 position was modified in this mutant. The raw LC-MS data are shown in **Fig. S18**.

To study the local sequence preference of the Tyr-*O*-geranylation by LimF, we tested a series of Tyr–1 mutants of teMP6-H5A (**Fig. 4d** and **Fig. S18**). This illustrated the similar –1 position preference to His-*C*-geranylation while the overall conversion efficiency of Tyr-*O*-geranylation was lower. To further compare the two distinct modes of geranylation by LimF, we performed kinetic analyses by using chemically synthesized two teMPs (teMP6-H5A/Y7H and teMP6-H5A), which were designed for efficient His-*C*- and Tyr-*O*-geranylation, respectively. The obtained K_M and k_{cat} values for His-*C*-geranylation of teMP6-H5A/Y7H were comparable to those of bcLimE2, suggesting that the ring-closing scaffold does not significantly affect the catalytic efficiency (**Table 1**, entries 1 and 2, and **Fig. S19**). LimF modification of teMP6-H5A with different GPP/Mg^{2+} concentrations revealed that their concentrations required for the half-maximal Tyr-*O*-geranylation activity was similar to His-*C*-geranylation (**Fig. S20a-e**), suggesting that two modes of geranylation may share similar active site. However, the k_{cat} and K_M values for Tyr-*O*-geranylation were significantly worse than His-*C*-geranylation (**Table 1**, entries 2 and 3, and **Fig. S20f**), implying that the

His-*C*-geranylation would be the primary function of LimF. It should be, however, noted that the turnover rate of the Tyr-*O*-geranylation by LimF (k_{cat} = 1.0 min^{-1}) is comparable to some previously characterized cyanobactin PTases and aromatic *O*-/*N*-PTases (**Table S2**). Therefore, LimF has the Tyr-*O*-geranylation activity as the secondary function but yet is quite active as an *O*-PTase.

Table 1. The steady-state kinetic parameters for LimF (wild type) and LimF-H172L mutant in different geranylation modes.

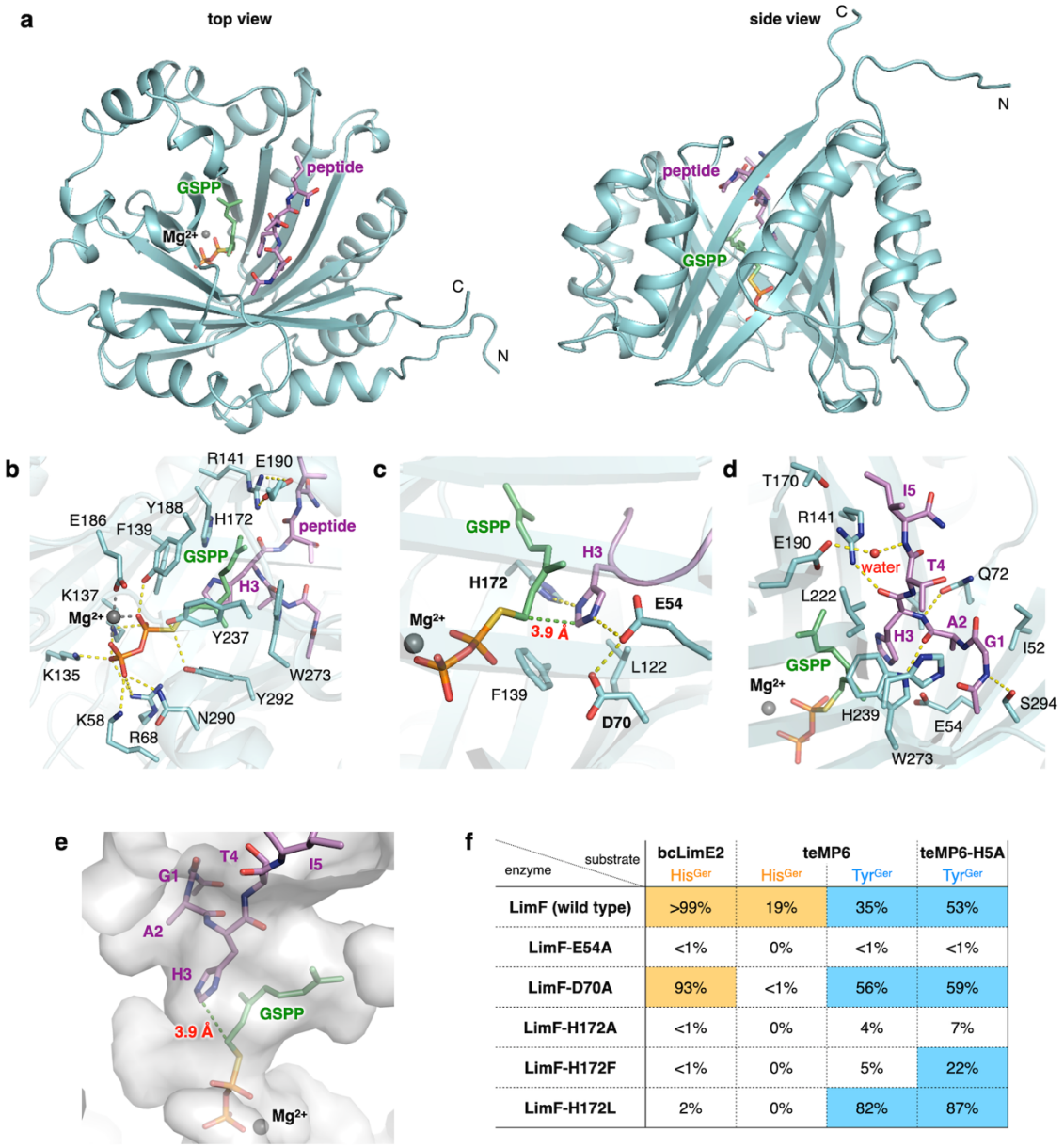
entry	enzyme	substrate	geranylating residue	K_M (mM)	k_{cat} (min^{-1})	k_{cat}/K_M ($min^{-1}·mM^{-1}$)
1	LimF	bcLimE2	His	0.31 ± 0.04	13.8 ± 0.6	44.1
2	LimF	teMP6-H5A/Y7H	His	0.58 ± 0.07	11.3 ± 0.6	19.4
3	LimF	teMP6-H5A	Tyr	7.3 ± 1.8	1.0 ± 0.1	0.14
4	LimF-H172L	teMP6-H5A	Tyr	1.7 ± 0.4	1.2 ± 0.1	0.68

Crystallographic characterization of LimF

We solved the co-crystal structures of the binary complex of LimF with a non-hydrolyzable GPP analog, geranyl S-thiolodiphosphate (GSPP), as well as the ternary complex with GSPP and a 5-mer substrate peptide (Ac-Gly-Ala-His-Thr-Ile-NH_2) at 1.77 Å and 1.93 Å, respectively (**Fig. 5a**, **S21**, and **Table S3**). LimF adopts the typical ABBA-fold[25,26], in which GSPP, Mg^{2+}, and the peptide are bound to the active site inside the β-barrel. In the complex, GSPP is precisely fitted into a hydrophobic pocket via Mg^{2+}-mediated coordination and intensive hydrophobic interactions (**Fig. 5b** and **S22**; see **Supplementary Text 2.1** for detailed description), supporting the LimF's selectivity towards the C10 prenyl donor. Moreover, the GSPP-C1 carbon is arranged at 3.9 Å distance from the imidazole-C2 of the substrate's His3 (**Fig. 5c**), which is consistent with the His-*C2*-forward geranylating activity of LimF.

The acceptor His3 sidechain is intensively recognized by LimF as the histidine N1 and N3 nitrogen are located within hydrogen-bonding distances from His172 sidechain and Glu54/Asp70 sidechains, respectively, supporting the requirement of the *N*-free imidazole moiety in the substrates predicted by the His analog study (**Fig. 3f, 5c,** and **Supplementary Text 2.2** for detailed description). On the other hand, LimF does not significantly recognize the Gly1, Ala2, Thr4, and Ile5 sidechains of the substrate peptide and mainly interacts with their main chain nitrogens and oxygens, substantiating the broad substrate scope of LimF (**Fig. 5d** and **Supplementary Text 2.3** for detailed description). The Ala2 is located in a narrow pocket composed

of Ile52, Glu54, Gln72, Leu122, and Ser294 of LimF (**Fig. 5e**), validating the observed His-1 position preference toward relatively small residues. The orientation of Thr4 sidechain, facing the bulk solvent, proved the tolerance of various residues at the His+1 position. When the complex structure, in which Thr4 is replaced with a Pro residue, is modelled, its Cδ atom would sterically clash with the main chain carbon and nitrogen atoms of Ala2 (**Fig.S23**), supporting the exceptional interdiction of Pro at the His+1 position in LimF substrates. Overall, the co-crystal structure has elucidated the structural basis for the selective His-*C2*-forward geranylation and broad substrate tolerance of LimF.



f

enzyme \ substrate	bcLimE2 HisGer	teMP6 HisGer	teMP6 TyrGer	teMP6-H5A TyrGer
LimF (wild type)	>99%	19%	35%	53%
LimF-E54A	<1%	0%	<1%	<1%
LimF-D70A	93%	<1%	56%	59%
LimF-H172A	<1%	0%	4%	7%
LimF-H172F	<1%	0%	5%	22%
LimF-H172L	2%	0%	82%	87%

Fig 5. Crystallographic characterization of the LimF-GSPP-peptide complex reveals the structural basis for the His-*C2*-forward geranylation and broad substrate tolerance. **a**, Overall structure. (Left) top view of the antiparallel β-barrel from the peptide binding site. (Right)

side view of the left figure rotated about 90° on the horizontal axis. LimF is shown in cyan cartoon; GSPP and peptide are shown in sticks; the geranyl group is colored in light green, a sulfur atom in yellow, phosphorus atoms in orange, and carbon atoms of peptide in violet; Mg^{2+} is shown in a gray sphere. **b**, GSPP recognition. Yellow and gray dotted line indicates hydrogen bonds and coordination bonds with Mg^{2+}, respectively. **c**, Recognition of the His3 imidazole. The LimF residues interacting with the His3 sidechain are shown in stick model. Green dotted line indicates the distance between the GSPP-C1 atom and the His3-C2 atom. **d**, Substrate peptide recognition. A water molecule that mediates the interaction between LimF and the peptide is shown as a red sphere. **e**, The surface of the active site cavity, where the peptide and GSPP bind. **f**, Geranylation efficiency of LimF and its mutants towards different reaction modes. The mean values of end-point conversion ratios in three independent reactions are shown. The experiments showing significant His- and Tyr-geranylation are highlighted with orange and light blue, respectively. The raw LC-MS data are shown in **Fig. S25**.

Site-directed mutagenesis of LimF reveals the critical residues

To validate the importance of the residues locating nearby the substrate imidazole ring (**Fig. 5c**; Glu54, Asp70, and His172), we performed site-directed mutagenesis experiments. Sequence alignment analysis of cyanobactin PTases indicated that Glu54 is highly conserved among cyanobactin PTases (**Fig. S24**). On the other hand, Asp70 and His172 are unique in LimF while other characterized PTases modifying aromatic sidechains have a conserved aliphatic residue (Leu, Met, or Val) at the His172 position. We prepared a Glu54 mutant (LimF-E54A), a Asp70 mutant (LimF-D70A), and three His172 mutants (LimF-H172A, LimF-H172F, and LimF-H172L) and reacted with the canonical substrate bcLimE2 to check their His-*C*-geranylation abilities. The LimF mutants except for LimF-D70A gave no or little geranylated product (**Fig. 5f**), clearly demonstrating that Glu54 and His172 play an essential role in His-geranylation by LimF. We also conducted modification of teMP6 and teMP6-H5A substrates by these LimF mutants to evaluate their Tyr-*O*-geranylation abilities. The Glu54 mutant resulted in a drastic loss of Tyr-*O*-geranylation, which is reasonable given the strong conservation of this residue in cyanobactin PTases. In contrast, LimF-H172L, in which the unique His in LimF is replaced with Leu conserved in Tyr-*O*-PTases, gave better end-point yields in both Tyr-containing substrates than the wild type LimF. Further steady-state kinetic assay of LimF-H172L validated its enhanced Tyr-*O*-geranylation activity with a 5-fold increase in k_{cat}/K_M than the wild type (**Table 1**, entries 3 and 4, and **Fig. S26**). Overall, in contrast to His-*C*-geranylation where both Glu54 and His172 are essential, His172 does not play an important role in Tyr-*O*-geranylation, and rather its mutation to Leu enhances the Tyr-*O*-geranylating activity.

LimF as a biocatalyst for facile alkylation of imidazole-containing bioactive molecules

As a final demonstration, we have examined whether the LimF-catalyzed geranylation is applicable for peptide late-stage alkylation. Due to electron deficiency of the imidazole ring, selective His modification of unprotected peptides remains challenging with chemical approaches.[57-59] Two FDA-approved His-containing peptidic drugs, gonadotropin-releasing hormone analogue leuprorelin and amylin analogue pramlintide, were treated by LimF, resulting in nearly quantitative geranylation (**Fig. 6a** and **b**). The MS/MS analysis of the products confirmed His-selective geranylation in the presence of various amino acid sidechains (Fig. **S27a** and **b**). To our surprise, LimF was also able to geranylate nonpeptidic, small-molecule drug cimetidine bearing 4,5-dialkylated imidazole moiety efficiently (Fig. **6c** and **S27c**). We conclude that LimF is a unique and remarkable biocatalyst capable of chemo- and regio-selective geranylation of His in various-sized bioactive peptides as well as imidazole derivatives in nonpeptidic molecules.

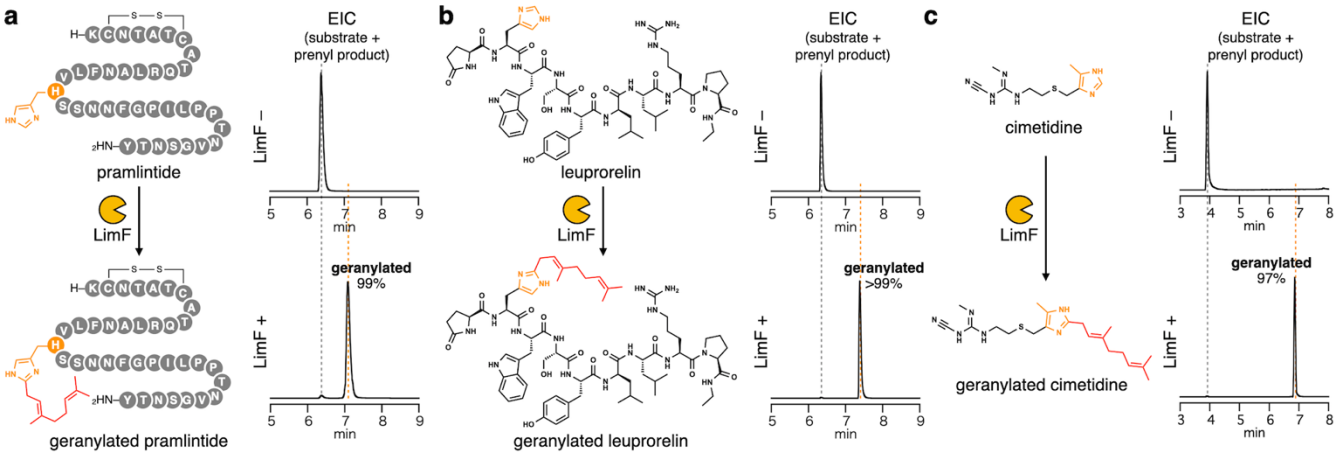


Fig 6. Selective *C*-geranylation of imidazole-containing bioactive molecules. Combined EICs including all possible charge states of the substrate [(a) pramlintide, (b) leuprorelin, or (c) cimetidine] and the corresponding geranylated product are shown. Conversion efficiency of each reaction determined by EIC peak areas are shown in the figure. The peaks corresponding to the substrates and geranylated products are highlighted with gray and orange dotted lines, respectively.

Discussion

Histidine is one of the four proteinogenic amino acids bearing aromatic sidechain and plays often critical roles in protein/peptide functions including metal binding, hydrogen bonding, proton shuttling, and nucleophile.[60] However, the *C*-modification on His residue by post-translational modification enzymes is a rare event probably due to the electron-deficient characteristic compared with other aromatic amino acids. Only a few small-molecular natural products with *C*-prenylated imidazole structures have been characterized

thus far,[61,62] and only one of them has a *C2*-substituted imidazole moiety.[63] As yet, enzymes responsible for His-*C*-alkylation remained unexplored due to the rareness of such structures in nature. To the best of our knowledge, LimF is the first functionally and structurally characterized PTase to catalyze *C*-alkylation of His.

The co-crystal structure of LimF with a peptide substrate has provided the structural basis of its unprecedented function. Further mutagenesis study indicated that the unique hydrophilic His172 and the conserved Glu54 cooperatively recognize the imidazole ring. These residues would also contribute to deprotonation of imidazole and/or stabilization of the intermediary σ-complex, accelerating the *C*-geranylation. In the co-crystal structure, the prenyl donor analog GPPS was bound to a relatively large binding pocket involving Gly224. The fact that PirF, the only cyanobactin geranyltransferase characterized prior to LimF,[64] also has Gly at the corresponding position suggests that a small residue at this position provides an sterically less-hindered active site for the accommodation of GPP, which is likely a general determinant of cyanobactin geranyltransferases.[64]

Although LimF is the only characterized His-PTase, further SSN analysis with LimF-queried BLAST hits revealed a putative PTase co-clustered with LimF, which possibly possesses His-*C*-prenylation activity (**Fig. S28**). Besides that, the recently reported Arg-*N*-PTase AgcF[34] and its homologs with potential Arg-*N*-prenylation activities shared a His residue at the position corresponding to His172 in LimF (**Fig. S24**). We assume that the equivalent His in AgcF likely plays a role in the recognition of NH in Arg sidechain. This unique His residue could serve as a probe in mining novel cyanobactin PTases that modify basic/hydrophilic sidechains.

Beside the dominant His-*C*-geranylation activity, LimF exhibits the alternative Tyr-*O*-geranylation function. This is in stark contrast to other previously characterized cyanobactin PTases, which selectively catalyzes a sole mode of prenylation. We cannot rule out the possibility that the Tyr-*O*-geranylation mode may not take place in its original host, but it can be regarded as an intrinsic activity of LimF, considering that the k_{cat} of LimF's Tyr-geranylation is comparable to or significantly larger than some other PTases (**Table S2**). A model structure of LimF bound with a Tyr-peptide by replacing His3 in the co-crystal structure with Tyr shows that the OH group of Tyr is located close to the C1 atom of GSPP (4.0 Å), supporting the Tyr-*O*-geranylation activity of LimF (**Fig. S29**). The LimF-H172L mutant conducted Tyr-geranylation with a comparable k_{cat} but an approximately 5-fold lower K_M than the wild type (**Table 1**), indicating that the corresponding aliphatic residues present in Tyr-*O*-PTases enhanced the activity majorly by increasing substrate binding affinity.

The LimF's bifunctionality to geranylate His and Tyr residues clearly implies its plasticity, and such a property could be evolutionarily advantageous for a PTase to evolve into another PTase modifying different residues. The phylogenetic analysis of LimF with other cyanobactin PTases (**Fig. 1a**) showed that LimF is located between the clades of Tyr-*O*-PTases and Trp-*C*-/*N*-PTases. This implies that LimF might have evolved from such Tyr- or Trp-PTases and *vice versa*, and the Tyr-*O*-geranylation activity of LimF could be regarded as a relic of the Tyr-*O*-PTase through the evolution. Collectively, the two modes of geranylation in LimF could provide evolutional evidence that supports the diverse repertoire of prenylated peptides produced by cyanobactin PTases.

Non-traditional peptide-based molecules are emerging as a therapeutic modality due to their high potency and specificity to various protein targets, even those previously considered as "undruggable".[65,66] As a platform to provide therapeutically relevant peptide ligands, we have devised the RaPID (random nonstandard peptides integrated discovery) system, which enables the construction and screening of massive libraries of teMPs.[21,67] Given that LimF is proved to accept the exotic teMP scaffold, it allows us to expand of the structural diversity of teMPs accessible by the RaPID system and construct geranylated teMP libraries with an improved lipophilicity, providing a new opportunity to identify membrane permeable peptide ligands. Moreover, LimF also offers geranylation of various imidazole-containing bioactive molecules. The currently available peptide modification strategies, despite significant development, are largely relied on nucleophilic residues such as Cys and Lys.[59] Although there are some chemical His-*C*-functionalization methodologies,[68-72] the direct modification on His is still in its infancy due to the intrinsic low-reactivity of imidazole and the presence of various interfering functional groups on peptides. Highlighted by remarkable chemo- and regio-selectivity, and relaxed substrate tolerance in LimF-catalyzed geranylation, we envision that this enzyme serves as a powerful biocatalyst for the late-stage modification of various imidazole-containing molecules.

In conclusion, our results have demonstrated a novel post-translational modification and provided structural basis for the LimF-catalyzed His-*C*-geranylation. The discovery of LimF also indicates that cyanobactin biosynthetic pathways remain a rich source for mining novel natural products and biosynthetic machinery which would further expand the chemical space of peptide modification. Furthermore, these results have expanded the toolbox of prenylation biocatalysts applicable for production of alkylated peptides with enhanced pharmacokinetics.

Acknowledgements

We thank Prof. Hiroshi Nagai (Tokyo University of Marine Science and Technology) and Dr. Toshio Nagashima (RIKEN Center for Biosystems Dynamic Research) for technical assistance with the NMR analysis. We thank all beamline staffs at BL32XU (SPring-8) and BL1A (Photon Factory) for their technical support. This work was supported by KAKENHI (JP16H06444 to H.S. and Y.G.; JP17H04762, JP19H01014, JP19K22243, JP20H02866 to Y.G.; JP20H05618 to H.S.; JP21K06051 to K.H.; JP19H02842 and JP21K19056 to M.O.) from the Japan Society for the Promotion of Science, and by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from Japan Agency for Medical Research and Development (AMED) under Grant Number JP19am0101070 (support number 1698).

Author contributions

Y.G. and H.S. conceived and supervised the study. Y.Z., K.H., D.T.N., S.I., M.S., M.O., T.S., Y.G., and H.S. designed experiments. Y.Z., K.H., D.T.N., S.I., S.K., and C.O. prepared recombinant LimF and its mutants. Y.Z., S.I., M.O., and Y.G. performed bioinformatic analyses. Y.Z., D.T.N., S.I., and Y.G. performed in vitro LimF reactions and substrate tolerance study. Y.Z., S.I., and M.S. performed NMR experiments. K.H., S.K., C.O., T.S., and Y.G. performed crystallographic study. All authors analyzed the experimental results. Y.Z., K.H., K.O., M.O., T.S., Y.G., and H.S. wrote the manuscript with input from all authors.

Competing interests

The following authors (H.S., Y.G., Y.Z., and M.O.) are co-inventors on patent application related to preparation of site-directed geranylated chemical entities with LimF.

Data availability

The coordinates and structure factors of LimF-GSPP-peptide complex and LimF-GSPP complex were deposited in the Protein Data Bank (PDB codes 7VMW and 7VMY, respectively).

Reference

1 Vickery, C. R., La Clair, J. J., Burkart, M. D. & Noel, J. P. Harvesting the biosynthetic machineries that

cultivate a variety of indispensable plant natural products. *Curr. Opin. Chem. Biol.* **31**, 66-73, (2016).
2 Zhao, H. *et al.* Dimericbiscognienyne A: A meroterpenoid dimer from Biscogniauxia sp. with new skeleton and its activity. *Org. Lett.* **19**, 38-41, (2017).
3 He, H. *et al.* Discovery of the cryptic function of terpene cyclases as aromatic prenyltransferases. *Nat. Commun.* **11**, 1-13, (2020).
4 Li, S.-M. Prenylated indole derivatives from fungi: structure diversity, biological activities, biosynthesis and chemoenzymatic synthesis. *Nat. Prod. Rep.* **27**, 57-78, (2010).
5 Awakawa, T. & Abe, I. Molecular basis for the plasticity of aromatic prenyltransferases in hapalindole biosynthesis. *Beilstein J. Org. Chem.* **15**, 1545-1551, (2019).
6 Chooi, Y.-H. *et al.* Genome Mining of a Prenylated and Immunosuppressive Polyketide from Pathogenic Fungi. *Org. Lett.* **15**, 780-783, (2013).
7 Matsuda, Y. & Abe, I. Biosynthesis of fungal meroterpenoids. *Nat. Prod. Rep.* **33**, 26-53, (2016).
8 Shi, S., Li, J., Zhao, X., Liu, Q. & Song, S.-J. A comprehensive review: Biological activity, modification and synthetic methodologies of prenylated flavonoids. *Phytochemistry* **191**, 112895, (2021).
9 Kuzuyama, T., Noel, J. P. & Richard, S. B. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435**, 983-987, (2005).
10 Sattely, E. S., Fischbach, M. A. & Walsh, C. T. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Nat. Prod. Rep.* **25**, 757-793, (2008).
11 Edwards, D. J. & Gerwick, W. H. Lyngbyatoxin biosynthesis: sequence of biosynthetic gene cluster and identification of a novel aromatic prenyltransferase. *J. Am. Chem. Soc.* **126**, 11432-11433, (2004).
12 Arnison, P. G. *et al.* Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30**, 108-160, (2013).
13 Montalbán-López, M. *et al.* New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **38**, 130-239, (2021).
14 Okada, M. *et al.* Structure of the Bacillus subtilis quorum-sensing peptide pheromone ComX. *Nat. Chem. Biol.* **1**, 23-24, (2005).
15 Jeong, A., Suazo, K. F., Wood, W. G., Distefano, M. D. & Li, L. Isoprenoids and protein prenylation: implications in the pathogenesis and therapeutic intervention of Alzheimer's disease. *Crit. Rev. Biochem. Mol. Biol.* **53**, 279-310, (2018).
16 Hatano, T. *et al.* Phenolic Constituents of Licorice. VIII. Structures of Glicophenone and Glicoisoflavanone, and Effects of Licorice Phenolics on Methicillin-Resistant Staphylococcus aureus. *Chem. Pharm. Bull. (Tokyo)* **48**, 1286-1292, (2000).
17 Fosgerau, K. & Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discov. Today* **20**, 122-128, (2015).
18 Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J. Am. Chem. Soc.* **141**, 4167-4181, (2019).
19 Zhang, L. & Bulaj, G. Converting Peptides into Drug Leads by Lipidation. *Curr. Med. Chem.* **19**, 1602-1618, (2012).
20 Offerman, S. C. *et al.* N-tert-Prenylation of the indole ring improves the cytotoxicity of a short antagonist G analogue against small cell lung cancer. *MedChemComm* **8**, 551-558, (2017).
21 Goto, Y. & Suga, H. The RaPID Platform for the Discovery of Pseudo-Natural Macrocyclic Peptides. *Acc. Chem. Res.* **54**, 3604-3617, (2021).
22 Czekster, C. M., Ge, Y. & Naismith, J. Mechanisms of cyanobactin biosynthesis. *Curr. Opin. Chem. Biol.* **35**, 80-88, (2016).
23 Gu, W., Dong, S.-H., Sarkar, S., Nair, S. K. & Schmidt, E. W. The biochemistry and structural biology of

cyanobactin pathways: enabling combinatorial biosynthesis. *Methods Enzymol.* **604**, 113-163, (2018).

24 Leikoski, N. *et al.* Genome Mining Expands the Chemical Diversity of the Cyanobactin Family to Include Highly Modified Linear Peptides. *Chem. Biol.* **20**, 1033-1043, (2013).

25 Tello, M., Kuzuyama, T., Heide, L., Noel, J. P. & Richard, S. B. The ABBA family of aromatic prenyltransferases: broadening natural product diversity. *Cell. Mol. Life Sci.* **65**, 1459-1463, (2008).

26 Saleh, O., Haagen, Y., Seeger, K. & Heide, L. Prenyl transfer to aromatic substrates in the biosynthesis of aminocoumarins, meroterpenoids and phenazines: The ABBA prenyltransferase family. *Phytochemistry* **70**, 1728-1738, (2009).

27 Hao, Y. *et al.* Molecular basis for the broad substrate selectivity of a peptide prenyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 14037-14042, (2016).

28 McIntosh, J. A., Donia, M. S., Nair, S. K. & Schmidt, E. W. Enzymatic basis of ribosomal peptide prenylation in cyanobacteria. *J. Am. Chem. Soc.* **133**, 13698-13705, (2011).

29 Morita, M. *et al.* Post-Translational Tyrosine Geranylation in Cyanobactin Biosynthesis. *J. Am. Chem. Soc.* **140**, 6044-6048, (2018).

30 Martins, J. *et al.* Sphaerocyclamide, a prenylated cyanobactin from the cyanobacterium Sphaerospermopsis sp. LEGE 00249. *Sci. Rep.* **8**, 1-9, (2018).

31 Parajuli, A. *et al.* A Unique Tryptophan C-Prenyltransferase from the Kawaguchipeptin Biosynthetic Pathway. *Angew. Chem., Int. Ed.* **55**, 3596-3599, (2016).

32 Okada, M. *et al.* Stereospecific prenylation of tryptophan by a cyanobacterial post-translational modification enzyme. *Org. Biomol. Chem.* **14**, 9639-9644, (2016).

33 Dalponte, L. *et al.* N-prenylation of tryptophan by an aromatic prenyltransferase from the cyanobactin biosynthetic pathway. *Biochemistry* **57**, 6860-6867, (2018).

34 Phan, C.-S. *et al.* Argicyclamides A–C Unveil Enzymatic Basis for Guanidine Bis-prenylation. *J. Am. Chem. Soc.* **143**, 10083-10087, (2021).

35 Donia, M. S., Ravel, J. & Schmidt, E. W. A global assembly line for cyanobactins. *Nat. Chem. Biol.* **4**, 341-343, (2008).

36 Tianero, M. D. B., Donia, M. S., Young, T. S., Schultz, P. G. & Schmidt, E. W. Ribosomal route to small-molecule diversity. *J. Am. Chem. Soc.* **134**, 418-425, (2012).

37 Ruffner, D. E., Schmidt, E. W. & Heemstra, J. R. Assessing the combinatorial potential of the RiPP cyanobactin tru pathway. *ACS Synth. Biol.* **4**, 482-492, (2015).

38 Tianero, M. D. *et al.* Metabolic model for diversity-generating biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1772-1777, (2016).

39 Purushothaman, M. *et al.* Genome mining based discovery of the cyclic peptide tolypamide and TolF, a Ser/Thr forward O-prenyltransferase. *Angew. Chem., Int. Ed.* **60**, 8460-8465, (2021).

40 Sardar, D. *et al.* Enzymatic N- and C-Protection in Cyanobactin RiPP Natural Products. *J. Am. Chem. Soc.* **139**, 2884-2887, (2017).

41 Mattila, A. *et al.* Biosynthesis of the bis-prenylated alkaloids muscoride A and B. *ACS Chem. Biol.*, (2019).

42 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410, (1990).

43 Atkinson, H. J., Morris, J. H., Ferrin, T. E. & Babbitt, P. C. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One* **4**, e4345, (2009).

44 Lima, A. R. J. *et al.* Insights Into Limnothrix sp. Metabolism Based on Comparative Genomics. *Front Microbiol.* **9**, (2018).

45 Donia, M. S. & Schmidt, E. W. Linking chemistry and genetics in the growing cyanobactin natural products family. *Chem. Biol.* **18**, 508-519, (2011).

46 Koehnke, J. *et al.* The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nat. Struct. Mol. Biol.* **19**, 767, (2012).
47 Sarkar, S., Gu, W. & Schmidt, E. W. Expanding the chemical space of synthetic cyclic peptides using a promiscuous macrocyclase from prenylagaramide biosynthesis. *ACS Catalysis* **10**, 7146-7153, (2020).
48 Gu, W., Sardar, D., Pierce, E. & Schmidt, E. W. Roads to Rome: Role of multiple cassettes in cyanobactin RiPP biosynthesis. *J. Am. Chem. Soc.* **140**, 16213-16221, (2018).
49 Ge, Y. *et al.* Insights into the mechanism of the cyanobactin heterocyclase enzyme. *Biochemistry* **58**, 2125-2132, (2019).
50 Goto, Y. *et al.* Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **3**, 120-129, (2008).
51 Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **6**, 779, (2011).
52 Burkhart, B. J., Schwalen, C. J., Mann, G., Naismith, J. H. & Mitchell, D. A. YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function. *Chem. Rev.* **117**, 5389-5456, (2017).
53 Sinha Roy, R., Belshaw, P. J. & Walsh, C. T. Mutational Analysis of Posttranslational Heterocycle Biosynthesis in the Gyrase Inhibitor Microcin B17: Distance Dependence from Propeptide and Tolerance for Substitution in a GSCG Cyclizable Sequence. *Biochemistry* **37**, 4125-4136, (1998).
54 Vinogradov, A. A. *et al.* Promiscuous Enzymes Cooperate at the Substrate Level En Route to Lactazole A. *J. Am. Chem. Soc.* **142**, 13886-13897, (2020).
55 Wiebach, V. *et al.* An Amphipathic Alpha-Helix Guides Maturation of the Ribosomally-Synthesized Lipolanthines. *Angew. Chem., Int. Ed.* **59**, 16777-16785, (2020).
56 Song, I. *et al.* Molecular mechanism underlying substrate recognition of the peptide macrocyclase PsnB. *Nat. Chem. Biol.* **17**, 1123-1131, (2021).
57 Narayanan, S., Vangapandu, S. & Jain, R. Regiospecific Synthesis of 2,3-Disubstituted-l-Histidines and Histamines. *Bioorg. Med. Chem. Lett.* **11**, 1133-1136, (2001).
58 Koniev, O. & Wagner, A. Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.* **44**, 5495-5551, (2015).
59 deGruyter, J. N., Malins, L. R. & Baran, P. S. Residue-specific peptide modification: a chemist's guide. *Biochemistry* **56**, 3863-3873, (2017).
60 Liao, S.-M., Du, Q.-S., Meng, J.-Z., Pang, Z.-W. & Huang, R.-B. The multiple roles of histidine in protein interactions. *Chem. Cent. J.* **7**, 1-12, (2013).
61 Larsen, T. O., Frisvad, J. C. & Jensen, S. R. Aurantiamine, a diketopiperazine from two varieties of Penicillium aurantiogriseum. *Phytochemistry* **31**, 1613-1615, (1992).
62 Kanoh, K. *et al.* (−)-Phenylahistin: a new mammalian cell cycle inhibitor produced by Aspergillus ustus. *Bioorg. Med. Chem. Lett.* **7**, 2847-2852, (1997).
63 Jan, C., Dippenaar, A. & Holzapfel, C. W. Crystal structure of the metal complexes of viridamine. *S. Afr. J. Chem.* **30**, 161-168, (1977).
64 Estrada, P., Morita, M., Hao, Y., Schmidt, E. W. & Nair, S. K. A Single Amino Acid Switch Alters the Isoprene Donor Specificity in Ribosomally Synthesized and Post-Translationally Modified Peptide Prenyltransferases. *J. Am. Chem. Soc.* **140**, 8124-8127, (2018).
65 Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. The future of peptide-based drugs. *Chem. Biol. Drug Des.* **81**, 136-147, (2013).
66 Henninot, A., Collins, J. C. & Nuss, J. M. The current state of peptide drug discovery: back to the future? *J. Med. Chem.* **61**, 1382-1414, (2018).
67 Yamagishi, Y. *et al.* Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin

ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol.* **18**, 1562-1570, (2011).

68 Liu, M., Zhang, Z., Cheetham, J., Ren, D. & Zhou, Z. S. Discovery and Characterization of a Photo-Oxidative Histidine-Histidine Cross-Link in IgG1 Antibody Utilizing 18O-Labeling and Mass Spectrometry. *Anal. Chem.* **86**, 4940-4948, (2014).

69 Xu, C.-F. *et al.* Discovery and Characterization of Histidine Oxidation Initiated Cross-links in an IgG1 Monoclonal Antibody. *Anal. Chem.* **89**, 7915-7923, (2017).

70 Noisier, A. F. *et al.* Late-Stage Functionalization of Histidine in Unprotected Peptides. *Angew. Chem., Int. Ed.* **58**, 19096-19102, (2019).

71 Chen, X. *et al.* Histidine-Specific Peptide Modification via Visible-Light-Promoted C–H Alkylation. *J. Am. Chem. Soc.* **141**, 18230-18237, (2019).

72 Nakane, K. *et al.* Proximity Histidine Labeling by Umpolung Strategy Using Singlet Oxygen. *J. Am. Chem. Soc.* **143**, 7726-7731, (2021).