

# Biosynthesis and engineering of the nonribosomal peptides with a C-terminal putrescine

Received: 5 January 2023

Accepted: 9 October 2023

Published online: 19 October 2023

 Check for updates

Hanna Chen<sup>1,3,4</sup>, Lin Zhong<sup>2,4</sup>, Haibo Zhou<sup>1,4</sup>, Xianping Bai<sup>1</sup>, Tao Sun<sup>1</sup>, Xingyan Wang<sup>1</sup>, Yiming Zhao<sup>1</sup>, Xiaoqi Ji<sup>1</sup>, Qiang Tu<sup>2</sup>, Youming Zhang<sup>1,2</sup> & Xiaoying Bian<sup>1</sup>✉

The broad bioactivities of nonribosomal peptides rely on increasing structural diversity. Genome mining of the *Burkholderiales* strain *Schlegelella brevitalea* DSM 7029 leads to the identification of a class of dodecapeptides, glidonins, that feature diverse N-terminal modifications and a uniform putrescine moiety at the C-terminus. The N-terminal diversity originates from the wide substrate selectivity of the initiation module. The C-terminal putrescine moiety is introduced by the unusual termination module 13, the condensation domain directly catalyzes the assembly of putrescine into the peptidyl backbone, and other domains are essential for stabilizing the protein structure. Swapping of this module to another two nonribosomal peptide synthetases leads to the addition of a putrescine to the C-terminus of related nonribosomal peptides, improving their hydrophilicity and bioactivity. This study elucidates the mechanism for putrescine addition and provides further insights to generate diverse and improved nonribosomal peptides by introducing a C-terminal putrescine.

Nonribosomal peptide synthetases (NRPSs) are large multidomain enzymes that operate in an assembly line fashion to produce a broad variety of nonribosomal peptides (NRPs) with diverse bioactivities and functions<sup>1–3</sup>. Each module of NRPS consists of an adenylation (A) domain, a thiolation (T) domain (also as a peptidyl carrier protein (PCP) domain), and a condensation (C) domain. The A domain recognizes and activates the target amino acids at the expense of ATP. The adenylation intermediate is transferred to the phosphopantetheinyl arm tethered to a T domain. The C domain catalyzes peptide bond formation between sequential amino acids. Last, the additional thioesterase (TE) domain releases the mature peptide product via hydrolysis or cyclization<sup>3,4</sup>.

Nonribosomal lipopeptides have fatty acyl chains at the N-terminus, which are introduced by a starter condensation (Cs) domain<sup>5–7</sup> or a

fatty acyl ligase<sup>8,9</sup> frequently located at the initiation module of NRPS, and the fatty acyl chains of lipopeptides show significant effects on their biological activities and toxicities<sup>10,11</sup>. Interestingly, nonacylated products were found in some lipopeptide NRPS gene clusters<sup>12–15</sup>, mainly from the deacylation of lipopeptides mediated by acylases or peptidases<sup>16</sup>. The typical acylase PvdQ from pyoverdine biosynthesis belonging to the Ntn-hydrolase family was characterized, which was responsible for the cleavage of the myristate or myristoleate moiety from the N-terminal amino acid<sup>12</sup>. To date, the reported acylases PvdQ play vital roles in siderophore maturation and quorum sensing, as well as detoxification, by cleaving acyl chains in different bacterial taxa<sup>12,14,15</sup>.

In addition to the well-known diverse acyl units at the N-terminus of nonribosomal lipopeptides, they also possess unusual moieties

<sup>1</sup>Helmholtz International Lab for Anti-infectives, Shandong University–Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, Shandong University, 266237 Qingdao, Shandong, China. <sup>2</sup>CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, 518055 Shenzhen, China. <sup>3</sup>Present address: School of Medicine, Linyi University, Shuangling Road, 276000 Linyi, China. <sup>4</sup>These authors contributed equally: Hanna Chen, Lin Zhong, Haibo Zhou. ✉e-mail: [bianxiaoying@sdu.edu.cn](mailto:bianxiaoying@sdu.edu.cn)

assembled at the C-terminus beyond the prediction of biosynthetic assembly lines. The additional C-terminal moieties include the amino acid residues that might be mediated by the TE domain, such as bacillothiazols<sup>17</sup> and a threonine-tagged lipopeptide family<sup>18–22</sup>, and diverse terminal amines, such as diamine putrescine (Put), spermidine (Spe), agmatine (Agm), phenylethylamine (Pea), tryptamine (Tra), and tyramine (Tya)<sup>14,23–31</sup>. Putrescine is ubiquitously and abundantly distributed in living organisms<sup>32</sup> and is used as a precursor of other polyamines, surfactants, siderophores, and agrochemicals<sup>33,34</sup>. Putrescine is fully protonated at physiological pH and is usually detected in the unmodified form, although Put is also modified for purposes of metabolic regulation and catabolism, such as N-acetylation and N-glutamylolation, or for incorporation into specialized metabolites, such as N-methylation and N-hydroxylation<sup>35</sup>. Furthermore, the direct incorporation of unmodified Put into the C-terminus of several natural products was prevalent in some NRPs, which was widespread in *Burkholderiales*<sup>14,23–25</sup>. However, the biosynthetic mechanism of the Put moiety participating in the C-terminus of these NRPs remains elusive and controversial. At present, two speculations have been proposed based on the features of their NRPS biosynthetic gene clusters (BGCs): one was that the C-terminal C domain might catalyze nascent peptide chain condensation with Put, such as malleobactins<sup>24</sup> and bicornutin A<sup>28</sup>, and the other was that a VibH-like protein elsewhere in the chromosome involved in the condensation of putrescine, such as crochelin<sup>14</sup>.

Genome sequencing reveals that the microbial genome is a massive and untapped trove of new metabolites<sup>36,37</sup>. Efficient strategies for the genome-driven discovery of novel natural products by awakening silent BGCs have been developed, such as functional promoter insertion in native producers and heterologous expression<sup>38,39</sup>. In our previous study, we successfully activated several silent/ cryptic NRP BGCs in *Burkholderiales* strains by in-situ promoter insertion using the efficient recombineering system Red $\alpha$ 7029 from *Schlegelella brevitalea* DSM 7029<sup>19,40–42</sup>, a potential chassis for gram-negative bacterial natural products<sup>39,43</sup>. Herein, we identified a class of linear dodecapeptides, glidonins A–L (**1–12**), by triggering an NRPS BGCII in the DSM 7029 strain<sup>40</sup>. The glidonins feature diverse modifications at the N-terminus and a putrescine moiety at the C-terminus. The biosynthetic pathway of glidonins was elucidated via in vivo genetic manipulation and in vitro biochemical assays. Moreover, the putrescine moiety could be added into potential peptides to generate novel and improved non-natural NRPs by bioengineering the termination module of NRPSs.

## Results

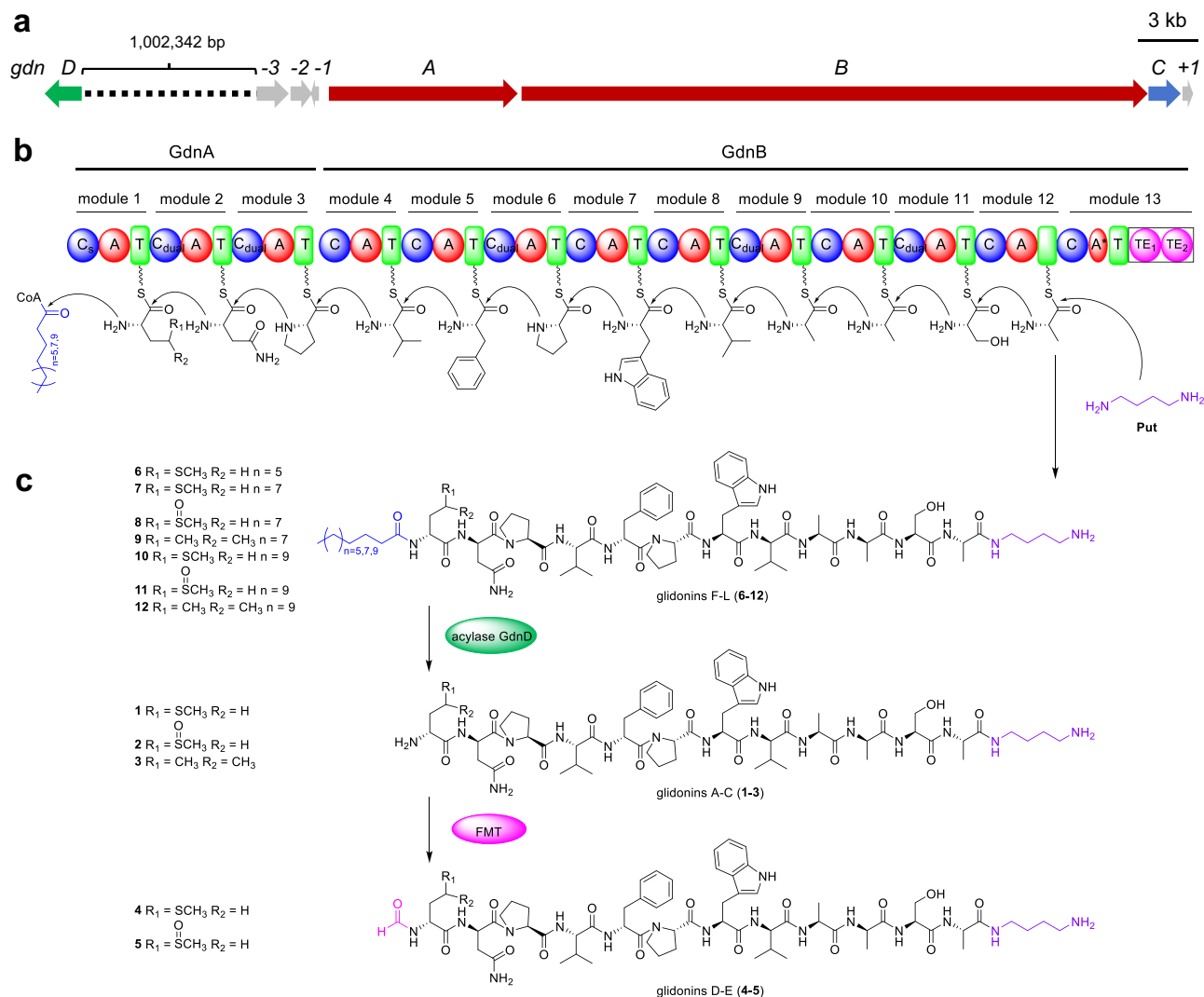
### Characterization of the glidonin biosynthetic gene cluster

Our previous study showed that the silent NRPS BGCII from *S. brevitalea* DSM 7029 was activated successfully by in-situ constructive promoter P<sub>Apra</sub> insertion using Red $\alpha$ 7029 recombineering<sup>40</sup> and produced a series of products, designated glidonins here, in the activated mutant *S. brevitalea* DSM 7029 P<sub>Apra</sub>-BGCII, as shown by comparative metabolic profile analysis<sup>40</sup> (Supplementary Fig. 1). Bioinformatic analysis showed that the glidonin (*gdn*) BGC included two essential core NRPS genes (*gdnA* and *gdnB*) and several additional genes (Fig. 1a and Supplementary Table 1). To determine the additional genes involved in the biosynthesis of glidonins, we inactivated each gene through Red $\alpha$ 7029 recombineering<sup>40</sup>. The LC–MS profiles showed that deletion of the gene *gdnC* (AKJ30060.1), which encodes an ABC transporter ATP-binding permease<sup>44</sup>, abolished all products, indicating its critical role in the efficient transportation of glidonins, while the other mutants failed to impact the production of glidonin (Supplementary Fig. 2). Thus, the core genes of the *gdn* gene cluster are *gdnA*, *gdnB*, and *gdnC*, spanning approximately 44 kb contiguous DNA region.

The first NRPS GdnA contains an initiation Cs domain condensing a fatty acid chain to the peptidyl backbone to form a lipopeptide<sup>5</sup>. Its three modules (modules 1–3) are predicted to activate Phe, Asn, and Pro by A domain specificity analysis<sup>45</sup> (Supplementary Table 9). The NRPS GdnB includes nine canonical modules (modules 4–12) activating Val, Phe, Pro, X (defined as an unknown amino acid by AntiSMASH), Val, Ala, Ala, Ser, and Ala, respectively<sup>45</sup>, as well as an incomplete termination module 13 containing a C domain, a partial A domain (A\*), a T domain, and a noncanonical TE domain with two putative TE domains (TE<sub>1</sub> and TE<sub>2</sub>) identified by the presence of two conserved active motifs (GX SXG)<sup>46</sup> according to bioinformatics analysis<sup>47</sup> (Supplementary Figs. 19 and 20 and Supplementary Table 9). The absence of the N-terminal subdomain (A<sub>core</sub>) and the Stachelhaus codes in the A\* domain leads to the A\* domain retaining only a C-terminal subdomain (A<sub>sub</sub>), which means it fails to function in amino acid activation (Supplementary Fig. 19b, c). The A\* domain shows similarity (Coverage: 190/409; Identities: 41%–45%) to other A domains of GdnB, but it is not possible to predict the specificity of the A\* domain due to the absence of the Stachelhaus codes (Supplementary Table 15 and Supplementary Fig. 19d). This kind of A\* is rare in nature besides the non-functional A\* domain of the kutzneride assembly line, which is critical for the stand-alone A domain KtzN-mediated transfer of Glu<sup>48</sup>. Further investigation is needed to determine the function of unusual termination module 13 of the *gdn* BGC, and structural elucidation of glidonins could provide valuable insights.

### Structural elucidation of glidonins

Eight target products were purified from the fermentation of *S. brevitalea* DSM 7029 P<sub>Apra</sub>-BGCII. Glidonin A (**1**, Fig. 1c, Supplementary Note 1) was obtained as a colorless oil with the molecular formula C<sub>65</sub>H<sub>98</sub>N<sub>16</sub>O<sub>14</sub>S (HR-ESI-MS, *m/z* 680.3699 [*M* + 2H]<sup>2+</sup>, calcd 1359.7242). Interpretation of its <sup>1</sup>H, <sup>13</sup>C, and DEPT NMR spectra (Supplementary Table 2, in MeOD-*d*<sub>4</sub>) revealed that **1** was a dodecapeptide containing the following amino acid moieties: Met, Asn, two Pro, two Val, Phe, Trp, three Ala, Ser, and one Put. The sequence of the amino acids and the location of Put were established by detailed interpretation of the HMBC correlations from amide protons and adjacent carbonyl groups (Supplementary Fig. 3, in DMSO-*d*<sub>6</sub>). The NMR data of glidonin B (**2**) and glidonin D (**4**), whose molecular formulas were shown to be C<sub>65</sub>H<sub>98</sub>N<sub>16</sub>O<sub>15</sub>S and C<sub>66</sub>H<sub>98</sub>N<sub>16</sub>O<sub>15</sub>S by HR-ESI-MS, were also quite similar to the NMR data of **1**, except for the obvious chemical shift variation of the methionine sulfoxide [Met(O)] residue in **2** and an additional aldehyde group ( $\delta_{\text{H}}/\delta_{\text{C}}$  8.15/164.0) in **4** (Supplementary Tables 3 and 4). The Met(O) in **2** is isomer and, in fact, was observed as a mixture of two epimeric compounds (major: **2a**, minor: **2b**). The molecular formulas of glidonins F–J (**6–10**) were found by HR-ESI-MS to be C<sub>75</sub>H<sub>116</sub>N<sub>16</sub>O<sub>15</sub>S, C<sub>77</sub>H<sub>120</sub>N<sub>16</sub>O<sub>15</sub>S, C<sub>77</sub>H<sub>120</sub>N<sub>16</sub>O<sub>16</sub>S, C<sub>78</sub>H<sub>122</sub>N<sub>16</sub>O<sub>15</sub>, and C<sub>79</sub>H<sub>124</sub>N<sub>16</sub>O<sub>15</sub>S. A comparison of the NMR data of **6**, **7**, and **10** (Supplementary Tables 4, 5, and 6) with **1** and analysis of their 2D NMR spectra (Supplementary Fig. 3) revealed that a decanoyl in **6**, a lauroyl in **7**, and a myristoyl in **10** are connected to the amino group of the Met moiety. Similarly, an additional lauroyl in **8** is connected to the amino group of Met(O) compared with **2**, which was supported by 2D NMR correlations (Supplementary Fig. 3). Compound **8** also existed as a mixture of isomers, and the NMR data for the major rotamer of **8** are summarized in Supplementary Table 5. The 1D NMR data (Supplementary Table 6) of **9** are also quite similar to those of **7** except for the replacement of Met in **7** by Leu in **9**. The structural compositions of the above eight compounds were confirmed by HR-ESI-MS/MS fragmentation (Supplementary Figs. 9, 10c, 10d, 11c, 11d, 12, and 13). Due to low yields, the structures of glidonin C (**3**), glidonin E (**5**), glidonin K (**11**), and glidonin L (**12**) were characterized by detailed comparative analysis of HR-ESI-MS/MS fragmentation (Supplementary Figs. 10a, 10b, 11a, 11b, and 14).



**Fig. 1 | The biosynthetic pathway of glidonin. a** Genetic organization of the *gdn* gene cluster. **b** Organization of modular enzymes involved in the biosynthesis of glidonin. **c** Structures and formation of glidonins A-L (1–12). FMT methionyl-tRNA

formyltransferase, Cs starter condensation, C<sub>dual</sub> dual condensation (C/E), C condensation, A adenylation, T thiolation, TE thioesterase, A\* partial adenylation.

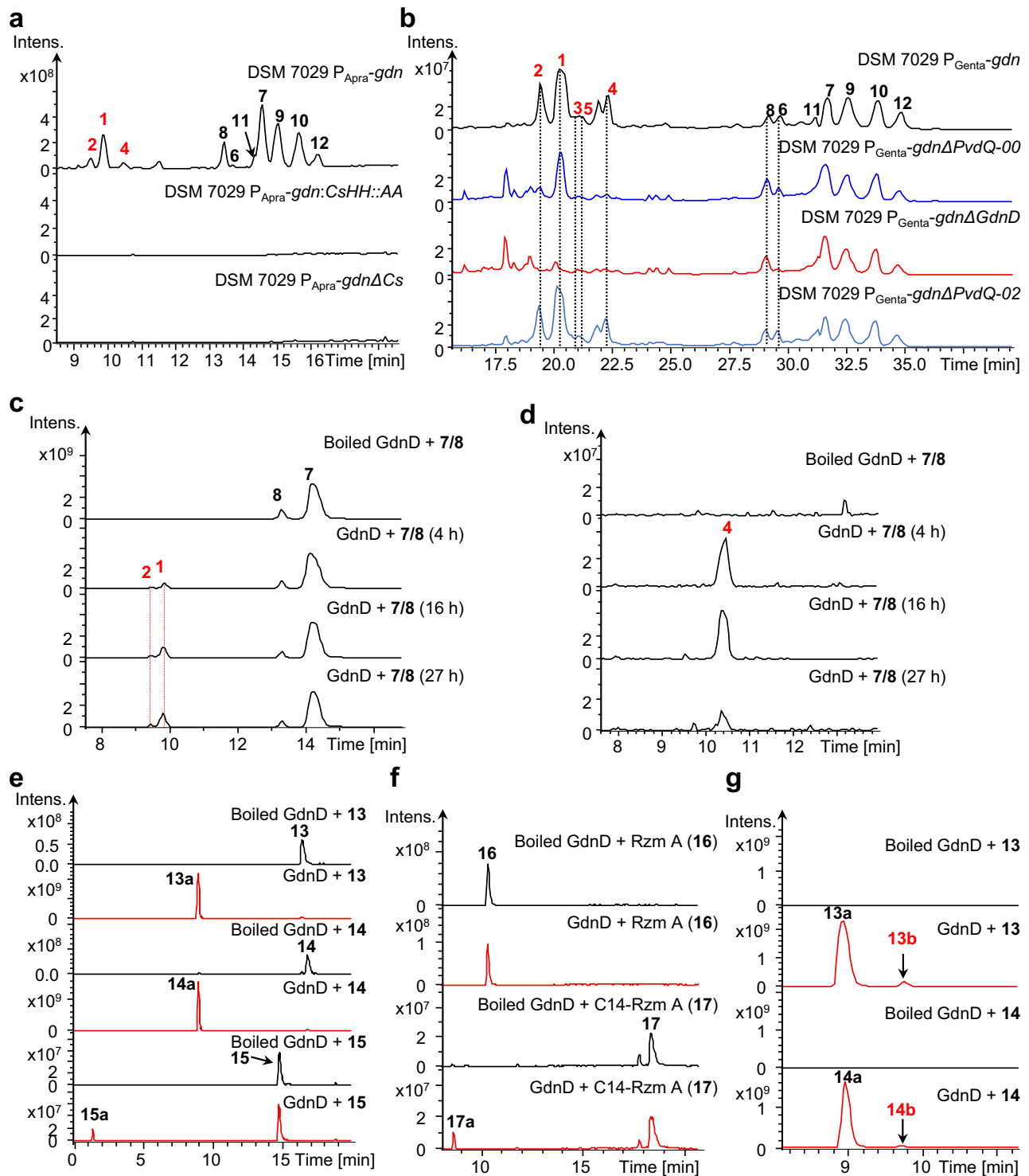
The absolute configuration of the amino acid monomers in these compounds was determined by Marfey's analysis<sup>49</sup> combined with bioinformatic analysis of five dual C domains with epimerization and condensation activities<sup>50</sup> in the Gdn NRPS assembly line (Fig. 1b, Supplementary Fig. 23, and Supplementary Table 10). The results demonstrated that the amino acid constitutions of these compounds are D-Met<sup>1</sup>(or D-Met(O)<sup>1</sup>)/D-Leu<sup>2</sup>-D-Asn<sup>2</sup>-L-Pro<sup>3</sup>-L-Val<sup>4</sup>-D-Phe<sup>5</sup>-L-Pro<sup>6</sup>-L-Trp<sup>7</sup>-D-Val<sup>8</sup>-L-Ala<sup>9</sup>-D-Ala<sup>10</sup>-L-Ser<sup>11</sup>-L-Ala<sup>12</sup> (Fig. 1c).

### The formation of N-terminal diversity of glidonins

The final products of the Cs domain-embedded *gdn* gene cluster should be lipopeptides, but the nonacylated products (**1**, **2** and **3**) were identified here. To test whether **1-3** were derived from the incomplete function of the Cs domain, we inactivated the Cs domain in-situ by two patterns, which failed to produce any glidonins, indicating that the Cs domain was required for the biosynthesis of all glidonins (Fig. 2a). Therefore, we hypothesized that a deacylase is responsible for the deacylation of lipopeptide glidonins as a post-assembly modification. The genome of *S. brevitalea* DSM 7029 contains three PvdQ-like acylase enzymes (AKJ26698.1, AKJ30942.1, AKJ29580.1). The deletion of AKJ30942.1 (named GdnD), which showed 68% similarity to PvdQ and was far away from the *gdn* BGC (1,002,342 bp), abolished the

production of **1-5**, as shown by HPLC–MS analysis (Fig. 2b and Supplementary Fig. 5d). Subsequently, we used the cell lysate containing acylase GdnD (85 kDa) to test its deacylation function because we failed to obtain the soluble protein (Supplementary Fig. 5a). The nonacylated products **1** and **2** were detected in the reaction mixture containing GdnD incubated with **7** and **8** for 4 h, 16 h, and 27 h, respectively, but not in the boiled cell lysate (Fig. 2c). Therefore, GdnD is responsible for the removal of the fatty acid chain as a post-assembly modification in the biosynthesis of **1** and **2**. In addition, GdnD was also functional for other lipopeptides with mid- or long acyl chains and lipopeptides with the L-/D- configuration of the first amino acids and different lengths of peptide chains, exhibiting some substrate diversity (Fig. 2e, f and Supplementary Note 2). Notably, we detected formylated **1** (**4**), **13a** (**13b**) and **14a** (**14b**) in the reaction mixture of the GdnD assay (Fig. 2d, g and Supplementary Fig. 22), indicating that the formylation could be catalyzed by a formyltransferase from *Escherichia coli*, and **4** should be derived from tailoring formylation of **1**. The in vivo gene deletion and in vitro assays allowed us to determine that **1** was formylated by methionyl-tRNA formyltransferase to form **4** (Supplementary Note 3).

The glidonins contain different fatty acyls and the first amino acid residues, which is generally due to the substrate selectivity of the Cs



**Fig. 2 | Identification of Cs domain and acylase PvdQ enzymes.** **a** LC-MS analysis of crude extracts at EIC 650–800 of activated mutant DSM 7029  $P_{Apra-gdn}$  and two inactivated mutants of Cs domain, DSM 7029  $P_{Apra-gdn:C5HH::AA}$  and DSM 7029  $P_{Apra-gdn\Delta Cs}$ . **b** HPLC-MS profile of crude extracts of activation mutant DSM 7029  $P_{Genta-gdn}$ , and three deletion mutants of acylases PvdQ. PvdQ-00: AKJ26698.1, GdnD: AKJ30942.1, PvdQ-02: AKJ29580.1. **c–g** Biochemical characterization of the acylase GdnD-catalyzed reactions by LC-MS analysis. **13**: lauroyl-D-Met-D-Asn-L-Pro-L-Val-D-Phe (EIC 789 [ $M+H$ ]<sup>+</sup>), **13a**: D-Met-D-Asn-L-Pro-L-Val-D-

Phe (EIC 607 [ $M+H$ ]<sup>+</sup>), **14**: lauroyl-L-Met-D-Asn-L-Pro-L-Val-D-Phe (EIC 789 [ $M+H$ ]<sup>+</sup>), **14a**: L-Met-D-Asn-L-Pro-L-Val-D-Phe (EIC 607 [ $M+H$ ]<sup>+</sup>), **15**: lauroyl-D-Met-D-Asn (EIC 446 [ $M+H$ ]<sup>+</sup>), **15a**: D-Met-D-Asn (EIC 264 [ $M+H$ ]<sup>+</sup>), **16**: Rzm A (Rhizomide A) (EIC 732 [ $M+H$ ]<sup>+</sup>), **17**: C14-Rzm A (C14-Rhizomide A) (EIC 900 [ $M+H$ ]<sup>+</sup>), **17a**: deacylation of C14-Rzm A (EIC 690 [ $M+H$ ]<sup>+</sup>), **4**: glidonin D (EIC 693.2 [ $M+2H$ ]<sup>2+</sup>), **13b**: formyl-D-Met-D-Asn-L-Pro-L-Val-D-Phe (EIC 633 [ $M+H$ ]<sup>+</sup>), **14b**: formyl-L-Met-D-Asn-L-Pro-L-Val-D-Phe (EIC 633 [ $M+H$ ]<sup>+</sup>).

domain and A<sub>1</sub> domain in initiation module 1. The formation of an unusual Met(O) of **2** might be a Met oxidized derivative during the isolation process from facile oxidation of the sulfide in Met of **1**. However, **1** and **2** were detected in the fermentation broth, and purified **1** was dissolved in MeOH for 3 days without changing to **2**, suggesting that the Met(O) residue was not generated by a tailoring reaction that took place after peptide formation. We assumed that the A<sub>1</sub> domain might select the nonproteinogenic amino acid Met(O). To test this hypothesis, we constructed and obtained the soluble protein CsA<sub>1</sub>T<sub>1</sub> from *E. coli* BAP1<sup>51</sup> (Supplementary Fig. 7a). The in vitro adenylation activity toward the A<sub>1</sub> domain, as determined by an ATP/PPI exchange assay<sup>52</sup>, demonstrated that the A<sub>1</sub> domain could indeed select L-Met(O), but the preferred substrate was L-Met (Supplementary Fig. 6a). The successful loading of substrates L-Met(O), L-Met, and L-Leu onto the T<sub>1</sub> domain was confirmed by using A<sub>1</sub>T<sub>1</sub> and CsA<sub>1</sub>T<sub>1</sub> proteins (Supplementary Fig. 24). We also tested the selectivity of Cs using lauroyl-CoA (C12-CoA) and myristoyl-CoA (C14-CoA) as donor substrates and L-Leu, L-Met, and L-Met(O) as acceptor substrates. Lauroyl products **18–20** and myristoyl products **21–23** were detected in vitro (Supplementary Fig. 6b, c), demonstrating that Cs could incorporate lauroyl and myristoyl moieties into the peptide assembly, consistent with the in vivo results. Therefore, the diverse substrates of initiation module 1 might be employed as an exchange element for the combinatorial biosynthesis of NRPS to generate structural diversity in pharmacologically relevant compounds.

### Assembly of the putrescine moiety

Genetic and biochemical analyses of the assembly of the Put moiety at the C-terminus of nonribosomal peptides remain elusive. Comparing BGCs encoding nonribosomal peptides with a C-terminal Put moiety, we found that the C-terminal multi-domains of the last module in the different NRPS assembly lines were different, such as C-T di-domains, C-A-T-TE domains, and C domain (Fig. 3a and Supplementary Fig. 8). The function of termination module 13 (C<sub>13</sub>A\*T<sub>13</sub>TE) of glidonin was investigated here to elucidate the mechanism of Put incorporation into the peptide chain. We first performed site-directed mutagenesis to inactivate the conserved active sites of the C<sub>13</sub>, T<sub>13</sub>, and TE domains in vivo (Fig. 3b). The metabolic profiles showed that the inactivation of the C<sub>13</sub> domain or T<sub>13</sub> domain abolished glidonins, and the mutants ΔTE<sub>2</sub>(S10624A) and ΔTE<sub>1/2</sub>(S10391A/S10624A) reduced the yields of glidonins, but the mutant ΔTE<sub>1</sub>(S10391A) had no obvious influence on the yields of glidonins compared with that of the wild type (Fig. 3c, d). Therefore, the C<sub>13</sub> domain and T<sub>13</sub> domain on module 13 were essential for glidonin biosynthesis in vivo, and the TE domain seemingly was not involved in the release of final products from the NRPS assembly line due to the presence of the final products in the mutants of the TE domain.

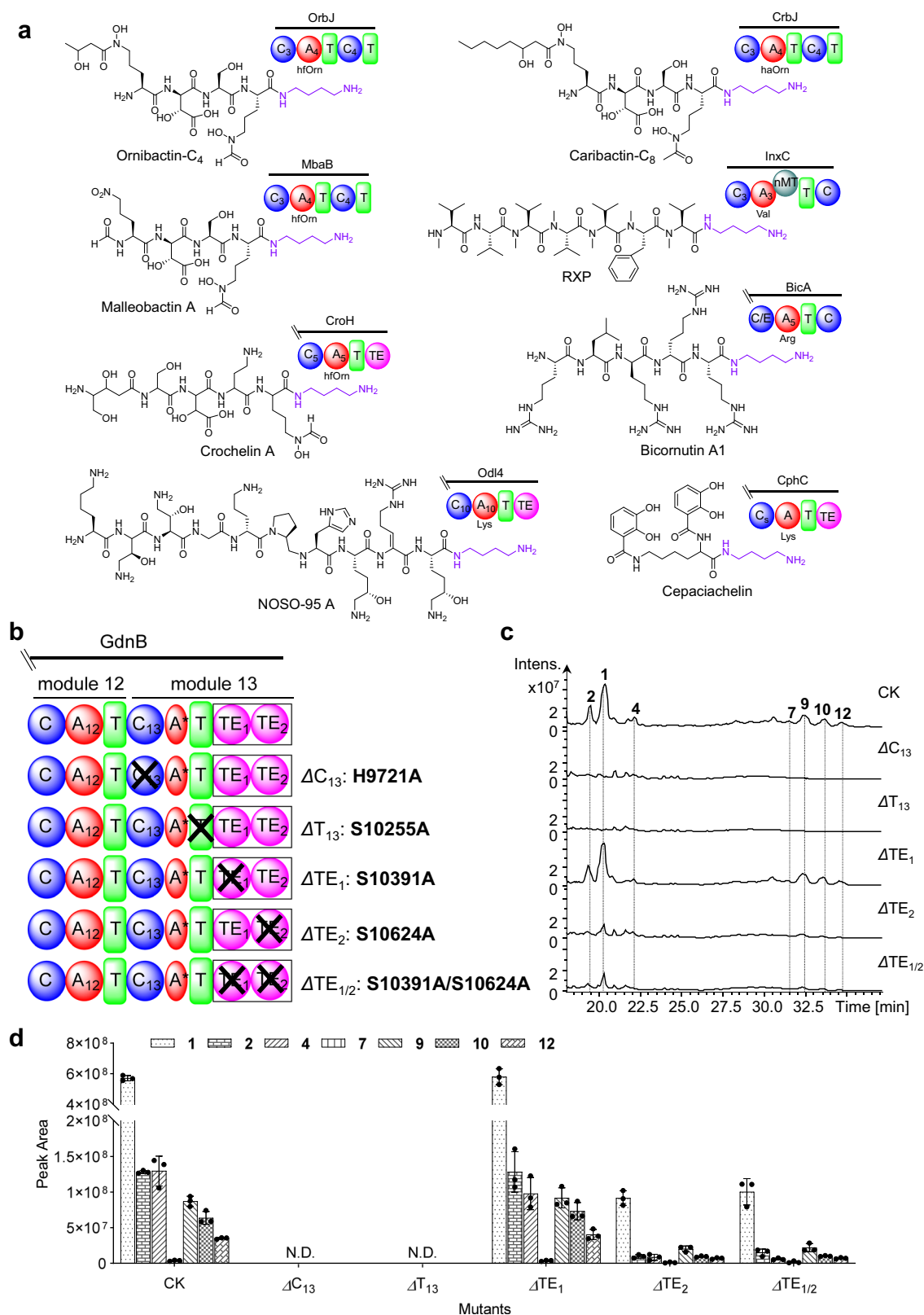
To gain more insight into the assembly of the Put moiety in glidonin biosynthesis by biochemical assays, we cloned and expressed eight NRPS fragments of GdnB, A<sub>12</sub>T<sub>12</sub>, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1</sub>(S10391A), A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut2</sub>(S10624A), and A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1/2</sub>(S10391A/S10624A) (Supplementary Fig. 7b, c). All soluble proteins were obtained except A<sub>12</sub>T<sub>12</sub>C<sub>13</sub> and A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1/2</sub>, hinting that A\* and the two conserved Ser residues of TE may be involved in the correct folding of the protein A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE. Moreover, the importance of the T<sub>13</sub> domain in maintaining protein structural stability was demonstrated, as the purified A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\* protein was relatively unstable and prone to precipitation after thawing. An ATP/PPI exchange assay<sup>52</sup> was conducted to evaluate the activity of the six mutants. A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub> and A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE maintained the same bioactivity toward substrate L-Ala as the control protein A<sub>12</sub>T<sub>12</sub>, but the others showed decreased activity of A<sub>12</sub> (Fig. 4a), which might be due to the instability or incorrect folding of A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1</sub>(S10391A), and A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut2</sub>(S10624A).

Based on the results of the ATP/PPI exchange assay, we divided these proteins into two groups: group 1 included A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>, A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE and the negative control A<sub>12</sub>T<sub>12</sub>, and group 2 included A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1</sub>(S10391A) and A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut2</sub>(S10624A), in the subsequent condensation assay to remove the impact of L-Ala<sub>12</sub> loading on the yield comparison of T<sub>12</sub>-linked Ala condensing with Put.

Afterward, we incubated substrates L-Ala and Put with the six proteins for the condensation assay (Fig. 4b). The target product **24a** was observed in all reactions, including the reaction of protein A<sub>12</sub>T<sub>12</sub> with relatively low yield, which indicated that the free amino of Put could attack the carbonyl of T<sub>12</sub>-linked Ala to form Ala-Put (**24**). In group 1, the highest yield of **24a** was obtained in the reaction containing protein A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>. The yield in A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\* was higher than that in A<sub>12</sub>T<sub>12</sub>, although the reduced activity of A<sub>12</sub> to L-Ala showed that the C<sub>13</sub>A\*T<sub>13</sub> domains and C<sub>13</sub>A\* domains both greatly promoted the intermolecular condensation of T<sub>12</sub>-linked Ala with Put in vitro (Fig. 4b). Notably, the protein A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE showed a greatly decreased yield of **24a** compared with A<sub>12</sub>T<sub>12</sub>, revealing that the TE domain preserved the hydrolysis function and resulted in a reduced yield of intermediate T<sub>12</sub>-linked Ala in vitro. In group 2, despite reduced activity toward the substrate L-Ala compared with A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE and A<sub>12</sub>T<sub>12</sub> in group 1, as mentioned above, the mutants A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1</sub>(S10391A) and A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut2</sub>(S10624A) exhibited higher yields of the target compound (**24a**) than A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE and A<sub>12</sub>T<sub>12</sub> in the condensation test. These results further support the notion that the TE domain has hydrolysis activity, and this function is abolished when the typical Ser active site is mutated. In addition, the yield of **24a** in the reaction containing A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut2</sub> was higher than that in the reaction containing A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub>TE<sub>Mut1</sub> under similar activity for substrate L-Ala (Fig. 4a, b), which implied that mutation of the Ser site of the TE<sub>1</sub> and TE<sub>2</sub> domains both led to reduce the hydrolysis function of TE in vitro, and the effect of the Ser site mutation was more pronounced in the TE<sub>2</sub> domain than in the TE<sub>1</sub> domain, which was consistent with the results observed in vivo (Fig. 3c, d). Based on the aforementioned results, we can infer that the C<sub>13</sub> domain is directly involved in the intermolecular condensation of T<sub>12</sub>-linked Ala and free Put. Moreover, the A\* and T<sub>13</sub> domains are essential for protein folding and stability in vitro.

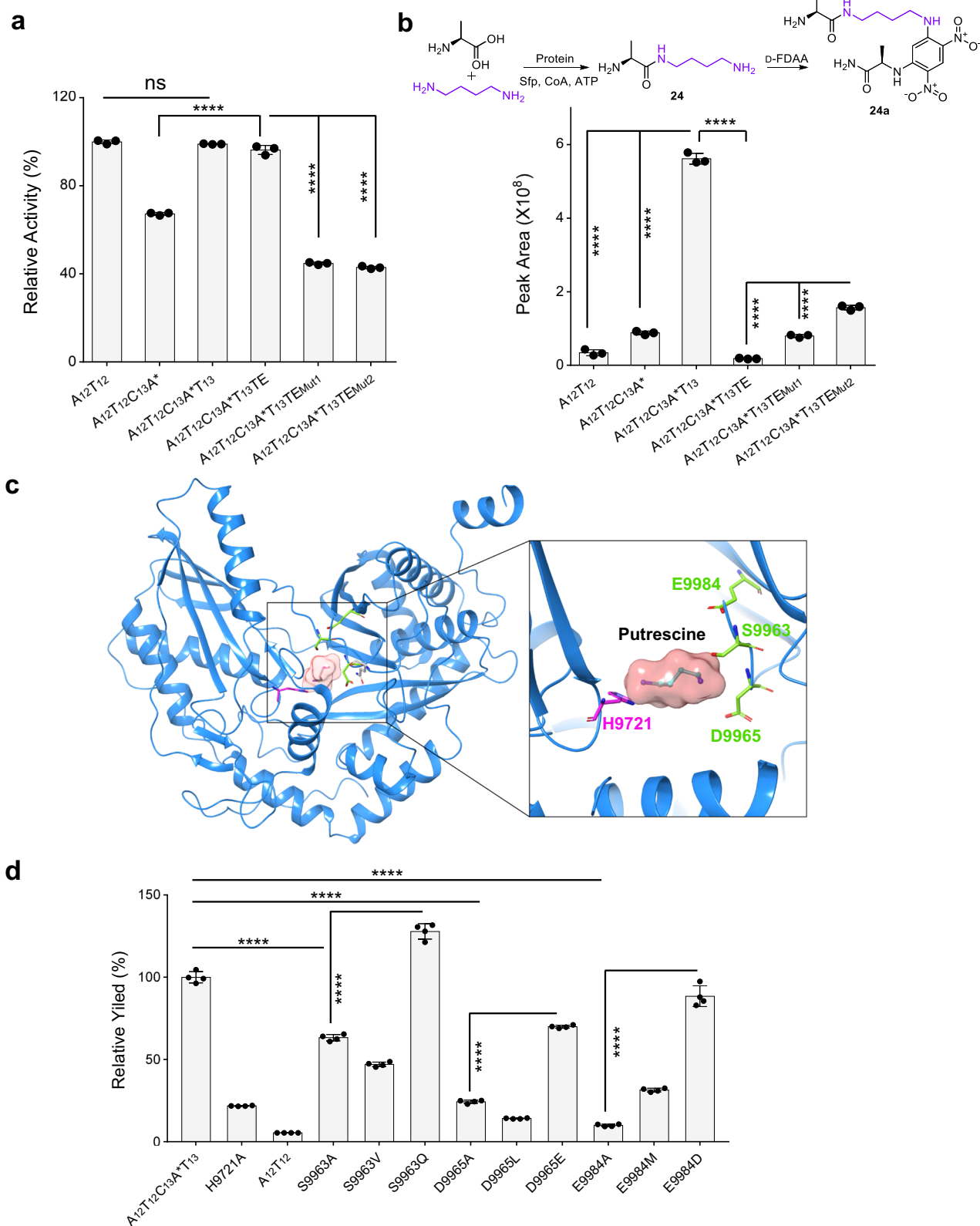
Combining the in vivo and in vitro results, it can be concluded that the C<sub>13</sub> domain selects Put as an intermolecular nucleophile for elongated peptide chain release, resulting in diamine incorporation at the C-terminus, and the A\*T<sub>13</sub> di-domain is essential for stabilizing the NRPS structure. As we presented, the TE domain (both for TE<sub>1</sub> and TE<sub>2</sub>) retains its hydrolysis function in vitro (Fig. 4b), which was further supported by the TE deletion test in vivo (Supplementary Fig. 32). The final products can still be obtained for the deletion of TE<sub>2</sub>, and the product without Put can be produced for P<sub>13</sub>-TE (deletion of C<sub>13</sub>A\*T<sub>13</sub> and insertion of the P<sub>13</sub> promoter for TE). Despite having a hydrolysis function, the TE domain does not directly participate in the release of final products. Instead, it can impact the folding or stability of the NRPS, as shown both in vivo and in vitro (Figs. 3d and 4a). This is supported by the observation that deletion of the TE domain and TE<sub>2</sub> significantly decreased the yield of the final products compared with the wild type (Supplementary Fig. 32). Moreover, the TE domain may also assist in the efficient operation of the full-length NRPS pipeline in vivo by hydrolyzing any stagnant by-products resulting from the malfunction of the Put release step, as we previously discussed regarding glidomide<sup>41</sup>. This was also supported by the observation of an intermediate (**9a**) without the Put moiety found in the crude extract of both wild-type and mutant strain with a constitutive promoter P<sub>13</sub><sup>53</sup> inserted in front of the TE domain (Supplementary Fig. 32).

The C<sub>13</sub> domain could condense free putrescine with T<sub>12</sub>-linked Ala, suggesting that the acceptor binding pocket of the C<sub>13</sub> domain might be appropriate for Put. The docking between Put and the pocket of the C<sub>13</sub> domain protein predicted by AlphaFold2<sup>54</sup> showed that three



**Fig. 3 | Identification of the domains in charge of Put introduction in vivo.** **a** The nonribosomal peptides with a C-terminal Put and their termination modules of NRPSs. hfOrn: N<sup>5</sup>-formyl-N<sup>5</sup>-hydroxyornithine, haOrn: N<sup>5</sup>-acetyl-N<sup>5</sup>-hydroxyornithine, RXP: Rhabdopeptide/xenortide-like peptide. **b** Schematic diagram of the different mutations of module 13. **c** HPLC-MS chromatogram obtained from the crude extract of the culture of the different mutants. **1** ( $m/z$  1359.7 [ $M+H$ ]<sup>+</sup>), **2**

( $m/z$  1375.7 [ $M+H$ ]<sup>+</sup>), **4** ( $m/z$  1385.7 [ $M+H$ ]<sup>+</sup>), **7** ( $m/z$  771.4 [ $M+2H$ ]<sup>2+</sup>), **9** ( $m/z$  762.4 [ $M+2H$ ]<sup>2+</sup>), **10** ( $m/z$  785.4 [ $M+2H$ ]<sup>2+</sup>), **12** ( $m/z$  776.4 [ $M+2H$ ]<sup>2+</sup>). **d** Yield comparison of the compounds from the culture of the different mutants and wild type. CK indicates wild type, N.D. indicates no products. Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent samples. Source data are provided as a Source Data file.



residues (S9963, D9965, E9984) of the C<sub>13</sub> domain, which were not conserved in other C domains, might be relevant to the substrate Put (Fig. 4c and Supplementary Fig. 17). Then, we constructed nine mutants of the A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub> protein to investigate the impact of these residues on the function (Supplementary Fig. 7d, e, f). The in vitro assays showed that the yield of **24a** from three mutants (S9963A, D9965A and E9984A) was greatly decreased, almost to that of mutant

A<sub>12</sub>T<sub>12</sub>C<sub>13</sub>A\*T<sub>13</sub> (H9721A) and A<sub>12</sub>T<sub>12</sub>, at the premise of their similar activity toward the substrate L-Ala (Fig. 4d and Supplementary Fig. 16a, c). Moreover, the three mutants altered the substrate specificity of the C<sub>13</sub> domain for the substrates putrescine and 1,5-diaminopentane, resulting in a reduced catalytic activity toward putrescine and increased catalytic activity toward 1,5-diaminopentane (Supplementary Fig. 16b), indicating that these three crucial residues are located

**Fig. 4 | Identification of the Put assembly by in vitro biochemical experiments.** **a** The relative activity of  $A_{12}T_{12}C_{13}A^*T_{13}$  domain from different proteins for the substrate L-Ala. The activity of  $A_{12}T_{12}C_{13}A^*T_{13}$  for activation of L-Ala was quantified as a reference (100%). Two mutants are  $A_{12}T_{12}C_{13}A^*T_{13}TE_{Mut1}(S10391A)$  and  $A_{12}T_{12}C_{13}A^*T_{13}TE_{Mut2}(S10624A)$ . TE contains TE<sub>1</sub> and TE<sub>2</sub>. Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent samples. **b** In vitro bioactivity assays of L-Ala condensing with Put mediated by the different proteins through comparing the yield of **24a** (**24** derivatized with D-FDAA). Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent

samples. **c** The docking of  $C_{13}$  domain and substrate Put, and zoom-in view of the conserved active site residue H9721 and key site residues S9963, D9965, E9984. **d** Yield comparison of the target product **24a** obtained from  $A_{12}T_{12}C_{13}A^*T_{13}$ , its nine mutants as well as negative control  $A_{12}T_{12}$  and  $A_{12}T_{12}C_{13}A^*T_{13}(H9721A)$ . The yield of **24a** generated from the wild type  $A_{12}T_{12}C_{13}A^*T_{13}$  was quantified as a reference (100%). Data are presented as mean values  $\pm$  SD,  $n = 4$  biologically independent samples. All  $P$  values were determined by two-tailed unpaired  $t$  test. \*\*\*\* $p < 0.0001$ . Source data and the exact  $p$  values are provided as a Source Data file.

within the binding pocket. The recovery of catalytic activity in the mutants with the original amino acid properties (S9963A vs. S9963Q, D9965A vs. D9965E and E9984A vs. E9984D) suggested the importance of the properties of these three residues (Fig. 4d). This indicated that these three residues might be key sites for substrate binding, especially E9984, which is located on the  $\beta$ -sheet of the  $C_{13}$  domain. The impact of the length of the amino acid side chain on the product yield (S9963A vs. S9963V, D9965A vs. D9965L and E9984A vs. E9984M) suggested that it also created steric hindrance in the substrate binding pocket (Fig. 4d). In summary, the three residues were essential for interacting with substrate Put, and our mutation results indicated that engineering the substrate specificity of this releasing module is a viable strategy.

### Introduction of putrescine into NRPs by multi-domain swapping

Putrescine is also a precursor of some specialized compounds, such as industrial polyamide nylon 4,6<sup>33</sup> and siderophore alcaligin<sup>55</sup>. To clarify the function of the Put moiety of glidonins, we synthesized lipopeptide **10a** without the Put group. Compound **10** was easily soluble in water, methanol and acetonitrile, while compound **10a** was soluble only in dimethyl sulfoxide (DMSO), which indicated that the Put moiety could enhance the solubility of this lipopeptide. Moreover, the antitumor activities of **10a** against the tumor cell lines A549 and HepG2 were lower than those of **10**, demonstrating that the Put group might improve the bioactivities of lipopeptides (Supplementary Table 11). These results indicated that introducing the Put moiety into lipopeptides might have the potential to improve the chemical properties and bioactivities of lipopeptides.

To confirm this speculation, we tried to introduce the Put moiety into the assembly backbone of other lipopeptides by engineering the C-terminal module of NRPs. Two distinct types of lipopeptides, cyclic C8-rhizomide A<sup>56</sup> (C8-Rzm A, **28**) and linear holrhizin A<sup>42,57</sup> (**30**), from *Mycetohabitans rhizoxinica* HKI 454, were used as the target products. To obtain the desired products, the different fusion regions using interdomain linkers from *gdn* BGC to *rzmA-RI48A* (*rzmA*<sup>\*</sup>) and *holA* BGCs were assessed. Considering that *rzmA*<sup>\*</sup> encodes the cyclic lipopeptide, the strategy of substitution was used for engineering *rzmA*<sup>\*</sup> (Fig. 5a, b), and we detected the expected mass of product **29** ( $m/z$  904.5882  $[M+H]^+$ ) in both mutants. Additionally, the yield of **29** in  $C_{13}A^*T_{13}TE$  substitution of RzmA<sup>\*</sup>-M2 was higher than that in recombinant RzmA<sup>\*</sup>-M1 (Fig. 5c). We then purified and elucidated target derivative **29**, whose molecular formula was C<sub>45</sub>H<sub>78</sub>N<sub>9</sub>O<sub>10</sub>, by HR-ESI-MS (Supplementary Fig. 15a). NMR data and analysis of its 2D NMR spectra confirmed that **29** indeed included a Put moiety connected to the carboxyl group of Val (Supplementary Table 7).

For the *holA* BGC, we first explored four different fusion regions to replace C-terminal module 6 of HolA (Fig. 5d, e). Metabolite profiles showed that the expected product **31** ( $m/z$  774.5122  $[M+H]^+$ ) was detected in different recombinants, and the optimal yield of **31** was from  $T_{12}C_{13}A^*T_{13}TE$  substitution of HolA-M1, which was better than the yield from  $C_{13}A^*T_{13}TE$  substitution of HolA-M2 (Fig. 5f). Moreover, the yield of **31** from the  $T_{12}C_{13}A^*$  substitution of HolA-M3 was higher than that from the  $C_{13}A^*$  substitution of HolA-M4 (Fig. 5f), which indicated that the preservation of the  $T_{12}$  domain in multi-domain substitution contributed to the increased production of derivative **31**.

Subsequently, the structure of derivative **31** was elucidated by HR-ESI-MS and NMR, which showed a Put moiety connected to the carboxyl group of Ala (Supplementary Table 8 and Supplementary Fig. 15b). To verify the effect of a Put moiety on the bioactivities of these compounds, two Put-lacking lipopeptides **29a** and **31a** were synthesized. Bioactivity assays showed that **29** and **31** indeed exhibited higher anti-inflammatory and antitumor activities than **29a** and **31a** (Supplementary Table 12 and Supplementary Fig. 29).

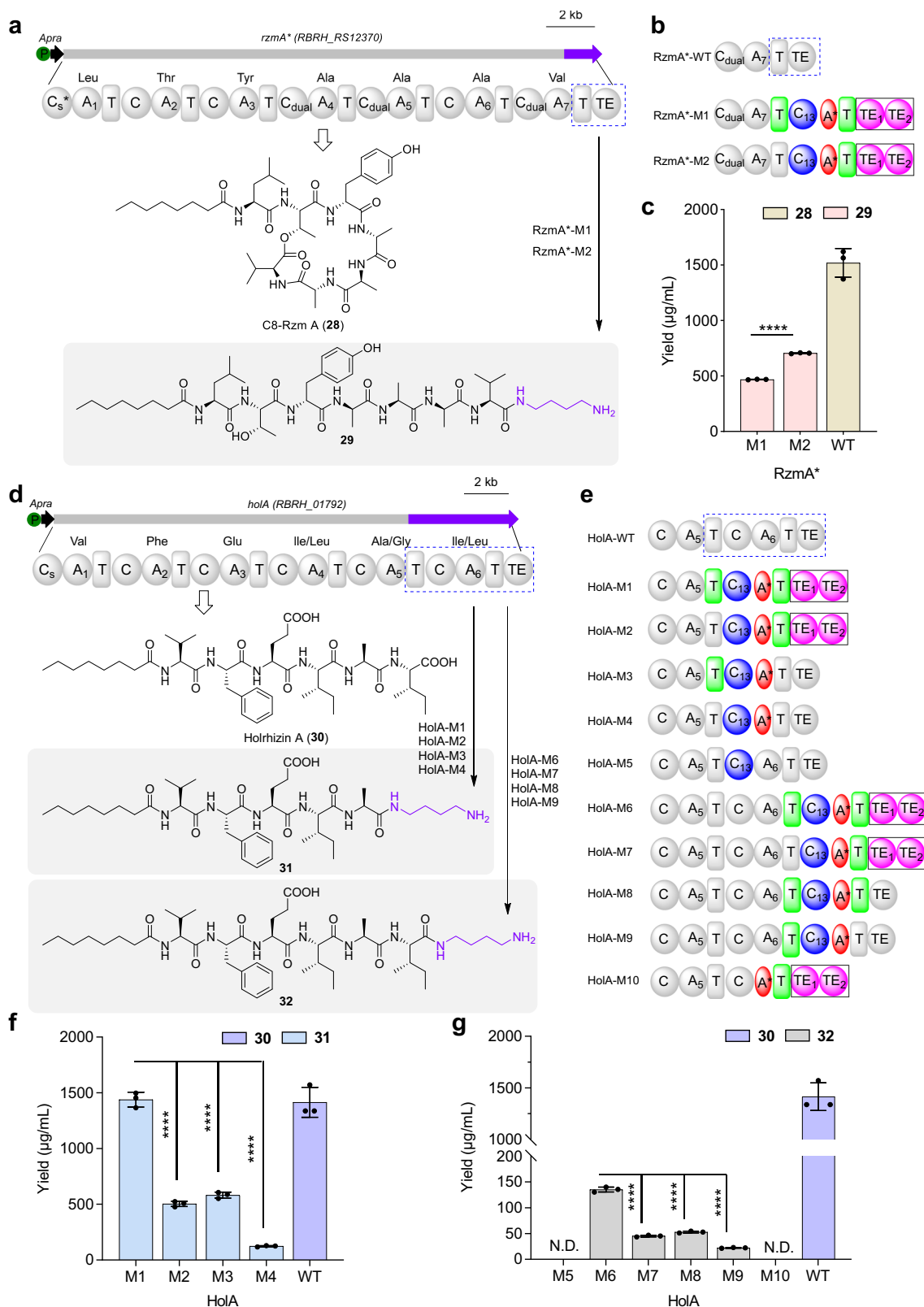
To further explore the minimal exchange units in charge of Put moiety assembly, six different exchange units were used to engineer the *holA* BGC (Fig. 5e). We found that four engineered strains produced another derivative (**32**) of holrhizin A, but HolA-M5 with  $C_{13}$  domain substitution and HolA-M10 with  $A^*T_{13}TE$  domain substitution did not (Fig. 5g), demonstrating that the minimal domain  $C_{13}A^*$  was required for Put assembly in vivo, which was consistent with the aforementioned in vitro biochemical assays indicating that the  $C_{13}$  domain was in charge of Put incorporation and that  $A^*$  might be very important for the structure of NRPs. The substitution sites of *rzmA*<sup>\*</sup> and *holA*, along with their corresponding amino acid sequences and protein structure characteristics, are summarized in Supplementary Fig. 25. During the substitution domain process, we also detected some short intermediates (**28a**, **30a**, **30b**) that were hydrolyzed in advance from the assembly line in different recombinants, as shown in Supplementary Fig. 30. In summary, the efficient substitution domains for the Put moiety engineered into NRPs by comprehensive analysis are the  $C_{13}A^*T_{13}TE$  multi-domains, defined as the “Put unit”, which has application prospects for improving the activity and enhancing the water solubility of target compounds in the bioengineering field.

### Introduction of other diamines into NRPs

We subsequently investigated the substrate tolerance of  $C_{13}$  using other potential diamines and polyamides, including 1,7-diaminoheptane, 1,6-diaminohexane, 1,5-diaminopentane, 1,2-diaminocyclohexane, spermine, and spermidine, in addition to the native substrate Put in vitro. Only four corresponding compounds (**24a**, **25a**, **26a** and **27a**) were observed in the reaction mixtures of Put, 1,5-diaminopentane, 1,6-diaminohexane, and 1,7-diaminoheptane by LC-MS analysis. Moreover, the yield of **25a** was higher than that of **26a** and **27a** but lower than that of **24a** (Supplementary Fig. 31a), revealing that the available substrates of the  $C_{13}$  domain were linear diamines, analogs of Put, and the longer lengths of diamines had a negative effect on the final product yield.

In addition, to extend the application of the “Put unit” in bioengineering, three usable diamines (1,7-diaminoheptane, 1,6-diaminohexane, and 1,5-diaminopentane) were added into the culture of the engineered strain *E. coli* GB05-MtaA:RzmA<sup>\*</sup>-M2. We found three corresponding desired derivatives **33** ( $m/z$  918.6032  $[M+H]^+$ , calc. 918.6023), **34** ( $m/z$  932.6167  $[M+H]^+$ , calc. 932.6179), and **35** ( $m/z$  473.8216  $[M+2H]^{2+}$ , calc. 946.6336) in the crude extracts in addition to the original compound **29**, as confirmed by HR-ESI-MS/MS fragmentation (Supplementary Fig. 31b, c, d). Bioengineering the “Put unit” can serve as a decent approach to enhance the diversity of NRPs with potential biological activities by feeding various diamines, and it can complement the Put binding pocket mutation method mentioned above.





### The bioactivity and function of glidonins

The bioactivity of glidonins was screened by several assays (Supplementary Note 4). The cytotoxicity assay showed that lauroyl products **7** and **9** and myristoyl product **10** exhibited not only broad-spectrum cytotoxicity but also strong activities against several tumor lines (Supplementary Table 11). However, nonacylated compounds **2** and **3**

showed no activities, and **1** had moderate activities against the Kasumi and A549 cell lines.

To clarify the function of the deacylation step in glidonin biosynthesis for the host *S. brevitalea* DSM 7029, we compared the growth profiles and biofilm formation between the wild type (WT) and two mutants, DSM 7029  $P_{Genta}$ -*gdn* and DSM 7029  $P_{Genta}$ -*gdnΔgdnD*. The

**Fig. 5 | Introduction of putrescine moiety on the two NRPs by multi-domain swapping.** **a** The flow chart of engineering *rzmA*\* BGC and corresponding products. **b** A diagram of the different multi-domain substitutions of the C-terminus of RzmA-R148A (RzmA\*). **c** Yield comparison of C8-Rzm A (**28**) and its derivative **29** between wild type (WT) and mutants. The absolute yield of related compounds was determined using the standard curve of **28** and **29**, respectively. Data are presented as mean values  $\pm$ SD,  $n = 3$  biologically independent samples. **d** The flow chart of engineering *holA* BGC and corresponding products. **e** A diagram of the different

multi-domain substitutions of the C-terminus of holrhizin A. **f, g** Yield comparison of holrhizin A (**30**) as well as their derivatives **31** and **32** between WT and mutants. The absolute yield of related compounds was determined using the standard curve of **30a** and **31**, respectively. N.D. indicates no products. Data are presented as mean values  $\pm$ SD,  $n = 3$  biologically independent samples. All  $P$  values were determined by two-tailed unpaired  $t$  test. \*\*\*\* $p < 0.0001$ . Source data and the exact  $p$  values are provided as a Source Data file.

growth profiles of the two mutants exhibited unusual growth curves because the mutants produced unknown pigments (black line and red line, respectively) (Fig. 6a and Supplementary Fig. 18a, b). Once pigment absorption was deduced (black and red dotted lines, respectively), the WT showed better cell growth than the two mutants, while the growth of the activation mutant DSM 7029 P<sub>Genta</sub>-*gdn* was superior to that of the  $\Delta$ *gdnD* mutant (Fig. 6a), which indicated that glidonins could inhibit the growth of DSM 7029 and that acylated glidonins might play a decisive role in this process. To directly evaluate the effect of acylated and nonacylated glidonins on wild-type DSM 7029, nonacylated compounds **1** and **2** and acylated compounds **7** and **10** were incubated with DSM 7029. The results revealed that the acylated compounds could obviously inhibit the growth of DSM 7029, but the nonacylated compounds had no effect (Fig. 6c). These results demonstrated that the deacylation of *gdn* BGC could be a detoxification mechanism for the native host DSM 7029.

The biofilm formation assays showed that DSM 7029 P<sub>Genta</sub>-*gdn* and  $\Delta$ *gdnD* increased biofilm formation compared to the WT, and the biofilm-forming capability of the activated mutant DSM 7029 P<sub>Genta</sub>-*gdn* was higher than that of the  $\Delta$ *gdnD* mutant (Fig. 6b). The colonies of the two mutants appeared smaller than those of the WT in the swarming motility assay (Supplementary Fig. 18d), which was consistent with the biofilm assays because swarming motility is often oppositely regulated and antagonistic to biofilm formation<sup>58,59</sup>. In addition, the colony morphology of the  $\Delta$ *gdnD* mutant was dry and wrinkled compared to that of the WT and the activated mutant, and the cell morphology of the two mutants was sticky relative to that of the wild-type strain, as observed by scanning electron microscopy (SEM), due to the enhanced biofilm formation of the two mutants (Fig. 6e and Supplementary Fig. 18c). Additionally, acylated compounds **7** and **10** indeed increased biofilm formation of DSM 7029 wild type, but nonacylated compounds **1** and **2** had less effect on biofilm formation (Fig. 6d). Thus, acylated glidonins could improve biofilm formation by inhibiting the swarming motility of DSM 7029, leading to alterations in the colony morphology of DSM 7029. Based on the above results, acylated glidonins could affect the growth state of DSM 7029, including growth speed, biofilm formation, and pigment secretions, but nonacylated products could not.

## Discussion

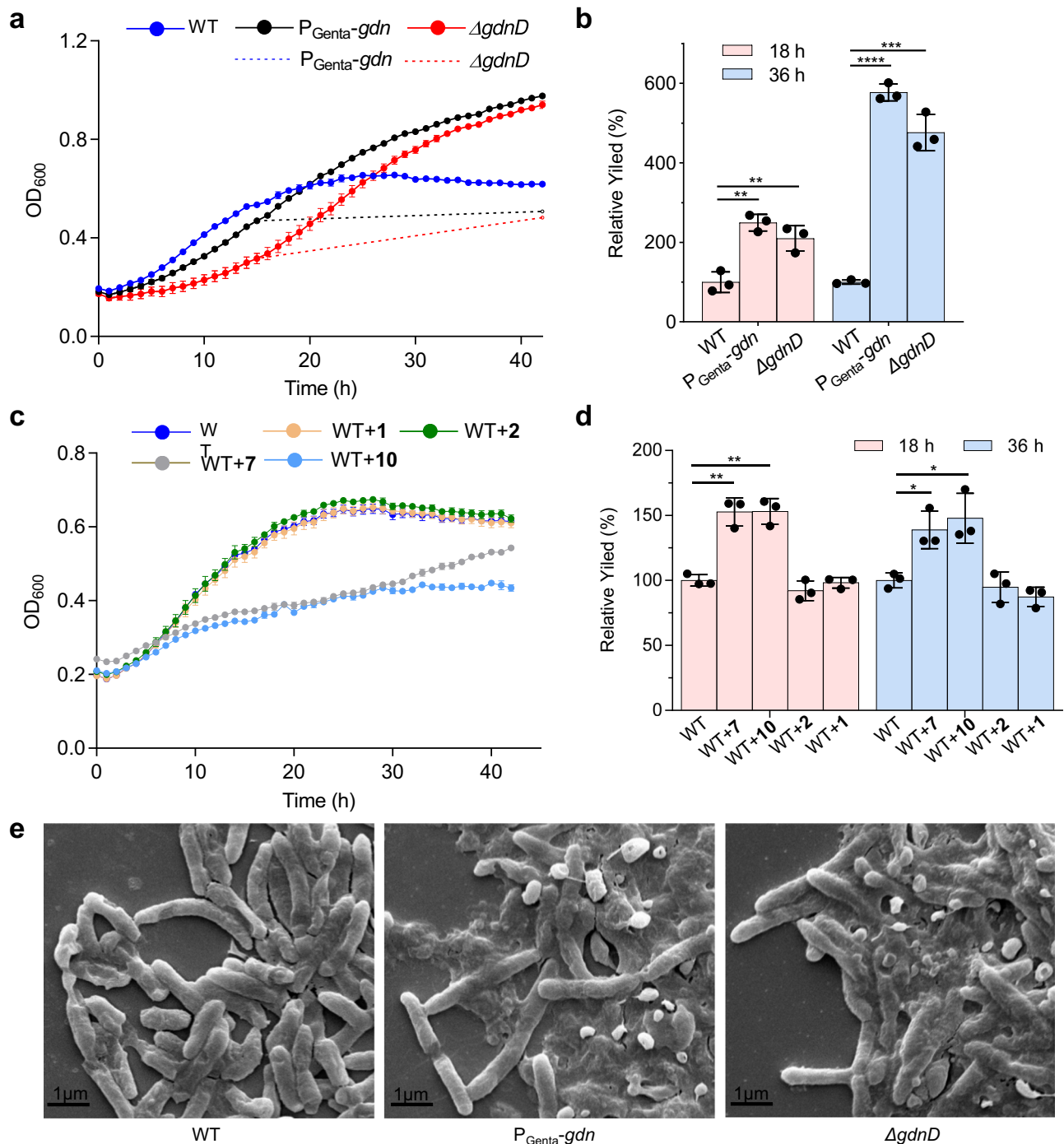
The biosynthesis of Put, as the precursor for some natural products, has been intensively studied in various bacteria. However, the current issue of the assembly of free Put into natural products remained unclear. Only two similar studies about siderophores vibriobactins<sup>60</sup> and serratiochelins<sup>61</sup>, whose free-standing amide synthases, VibH and SchH, incorporate free norspermidine/1,3-diaminopropane into the NRPS assembly line, were reported. In this work, we verified that the C-terminal C<sub>13</sub>A\*T<sub>13</sub> domains embedded into the NRPS assembly line were essential for condensing Put with thioesterified intermediate products through in vivo and in vitro experiments. Therein, A\*T<sub>13</sub> domains may play a vital role in maintaining the correct folding of NRPS protein, and the C<sub>13</sub> domain is more similar to the individual amide synthases in condensation function. The second histidine in the "HHxxxDC" motif also functions similarly to the typical C domain by playing a substrate positioning role (Fig. 4d, H9721A mutant vs. A<sub>12</sub>T<sub>12</sub>

mutant)<sup>56</sup>. In addition, the other representative NRPS assemble lines (Fig. 3a) possessing a C domain fixed into the C-terminus of termination module allowed us to deduce that the function of these C domains might be similar to that of amide synthases for Put introduction.

The addition of polyamines into natural products could enhance bioactivity and solubility, such as the anticancer F14512 linking a spermine chain to the epipodophyllotoxin core, showing priority efficiency in the pre- and clinical trials<sup>62,63</sup>. Here we also demonstrated the remarkable feature of the Put moiety to improve the solubilities and bioactivities of glidonins, and further supported by bioengineering the lipopeptides rhizomide and holrhizin to generate new-to-nature lipopeptides with increased activities. This finding provides further insights to biologists and chemists on the modification of potential lead compounds. The optimized "Put unit" could accomplish putrescine assembly, as well as the assembly of other diamines, which provides a decent strategy for expanding the diversity of lipopeptides by bioengineering in terms of combinatorial biosynthesis. In addition, the A<sub>1</sub> domain of *gdn* BGC showed good selectivity for the unusual nonproteinogenic amino acid Met(O) in vivo and in vitro, in addition to the proteinogenic amino acids Met and Leu, and can thus guide the introduction of unusual amino acids into NRPs through swapping of A domain or module to obtain diverse nonribosomal peptides.

The family of linear lipopeptides glidonins with diverse fatty acid chains (C10-C14) incorporates three different first amino acids. Interestingly, a high yield of nonacylated product from lipopeptide glidonins was found in the crude extract, which was processed post-modification by the characterized acylase GdnD. The acylase GdnD is clearly encoded outside the *gdn* BGC and deacylates a wide variety of lipopeptides. It is possible that the deacylation step might be a global detoxification or self-resistance mechanism for the native strain DSM 7029, similar to the role of the acylase Tem25 located inside the telomycin BGC<sup>15</sup>. Currently, the mechanisms of acylases in siderophore maturation and quorum sensing have been elucidated<sup>12,14</sup>. However, the natural function of several acylases, such as GdnD and Tem25, has yet to be resolved. Lipopeptides consist of a lipid moiety linked to a linear or cyclic oligopeptide, resulting in amphiphilic properties that confer enormous functional versatility. Their natural functions, some of which may be unique to the biology of the producing microorganisms, are involved in diverse ecological roles as biosurfactants and mediators of competitive advantage in interactions with coexisting organisms, as well as in biofilm formation<sup>64,65</sup>. The lipopeptides, acylated glidonins, not only increased the biofilm formation of DSM 7029 to improve its competitive advantage but also had negative effects on its growth. To balance this contradiction, DSM 7029 chooses to silence the *gdn* BGC under laboratory conditions and evolve the acylase GdnD. In addition, GdnD preferentially deacylates lipopeptides with long fatty chain lengths, which also have better bioactivities. This finding expands the knowledge of the self-protection for strains to adapt to the natural competitive environment.

In conclusion, the complex biosynthetic pathway of glidonin was elucidated, which starts with Cs domain to introduce fatty acids to the first amino acid, followed by condensation of the C-terminal Put by the C<sub>13</sub> domain of module 13, forming the final lipopeptide. The elucidation of the putrescine addition mechanism and the "Put unit" swapping strategy provided further insights into the engineering of diverse NRPs



**Fig. 6 | The growth and biofilm formation between the wild type (WT) and mutants.** **a** The growth profiles of DSM 7029 WT and two mutants were measured by culturing in CYMG medium. Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent samples. The dotted line means the authentic growth curves which the pigment absorption was deducted. **b** The comparison of the biofilm forming capabilities between WT and two mutants. The yield of the biofilm formation of the wild type strain was quantified as a reference (100%). Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent samples. **c** The growth profiles of DSM 7029 WT and wild-type strain dealt with different compounds were

measured by culturing in CYMG medium. Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent samples. **d** The comparison of the biofilm forming capabilities between WT and wild-type strain dealt with different compounds. The yield of the biofilm formation of the wild type strain treated with H<sub>2</sub>O was quantified as a reference (100%). Data are presented as mean values  $\pm$  SD,  $n = 3$  biologically independent samples. All  $P$  values were determined by two-tailed unpaired  $t$  test. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . **e** Cell morphology determination by using SEM of WT and two mutants. Source data and the exact  $p$  values are provided as a Source Data file.

with improved bioactivities by introducing a C-terminal Put tail. The structural characterization, biosynthetic machinery and bioengineering of lipopeptides with a C-terminal Put offer a versatile platform for the combinatorial biosynthesis of lipopeptides.

## Methods

### Bacterial strains and reagents

All strains, plasmids, and primers used in this study are supplied in Supplementary Data 1. *S. brevitalea* DSM 7029 and mutant strains were

cultured on CYMG (Casein peptone 8 g/L, Yeast extract 4 g/L, MgCl<sub>2</sub> 6H<sub>2</sub>O 8.66 g/L, Glycerol 5 mL/L) broth or agar plates with kanamycin (km, 30 µg/mL), gentamicin (genta, 20 µg/mL) and apramycin (apra, 30 µg/mL) at 30 °C. *E. coli* cultures were grown in Luria-Bertani (LB) broth or on LB-agar plates (1.2% agar) with ampicillin (amp, 50 µg/mL or 100 µg/mL), km (10 µg/mL or 15 µg/mL), or chloramphenicol (cm, 10 µg/mL or 15 µg/mL). Regents were purchased from Sigma-Aldrich, New England Biolabs, Invitrogen, and Thermo Fisher Scientific. DNA primers were purchased from Sangon Biotech, Tsingke Biological Technology, and RuiBiotech. The synthesized peptides (from Sangon Biotech) were shown in Supplementary Table 14, and their structures were confirmed by HPLC-ESI-MS, <sup>1</sup>H and <sup>13</sup>C spectra, which were shown in supplementary figures.

### Gene replacement in the chromosome of the activation mutant *S. brevitalea* DSM 7029 P<sub>Genta</sub>-*gdn*

Our previous study demonstrated the *gdn* gene cluster was silent, and only three antibiotics (kanamycin/apramycin/gentamicin) were used in the *S. brevitalea* DSM 7029, as well two antibiotics (apramycin and gentamicin) are overlapping<sup>40,43</sup>. Thus, we first constructed the activated mutant *S. brevitalea* DSM 7029 P<sub>Genta</sub>-*gdn*. The polymerase chain reaction (PCR) product of the constructive P<sub>Genta</sub> promoter with 50 bp homology arm was amplified. The template for gentamicin resistance gene is pFox-genta-P<sub>tet</sub>-LTX. The PCR product was transformed into *S. brevitalea* DSM 7029 containing recombineering system Redy-Redαβ7029. Correct recombinants were verified by colony PCR and DNA sequencing.

The single genes (*gdnC*, +1, -1, -2, -3, three PvdQ-like acylase genes, two formyltransferase PurH and PurN genes) were replaced by apramycin resistance gene (*apra*<sup>R</sup>) using the recombineering system Redy-Redαβ7029 as well, respectively. The apramycin resistance genes flanked with homology arms (50 bp) were generated by PCR amplification using 2 × ApexHF HS DNA Polymerase FS Master Mix, and the template for *apra*<sup>R</sup> is plasmid RK2-*apra*-cm. Purified PCR products of apramycin resistance genes were transformed into *S. brevitalea* DSM 7029 P<sub>Genta</sub>-*gdn*/pBBR1-Rha-Redy-Redαβ7029-km, respectively. Correct recombinants were verified by colony PCR.

### Site-directed mutagenesis and knockout of starter condensation (Cs) domain and module 13 on the *gdn* BGC

Due the site mutation of the *gdn* BGC was unachieved in DSM 7029, the *gdn* BGC was firstly cloned from the genome via direct cloning mediated by ExoCET<sup>66</sup> and verified by digestion by the restriction enzyme *Apa*LI. Then, the constructive P<sub>Apra</sub> promoter and *amp-attP* fragment were inserted into the plasmid p15A-cm-*gdn* by the linear plus circle homology recombination (LCHR) mediated by the recombineering system Redy-Redαβ, respectively, to drive the whole *gdn* BGC and complement *gdn* gene cluster into DSM 7029. The correct plasmid p15A-amp-attP-P<sub>Apra</sub>-*gdn* was confirmed by restriction enzyme and DNA sequencing.

We introduced several modifications into the plasmid p15A-amp-attP-P<sub>Apra</sub>-*gdn* using the RedEx strategy<sup>67</sup>, including knockout of the Cs domain and module 13, deletion of TE<sub>2</sub> domain (271 amino acids), TE domain (502 amino acids), and C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>TE (1237 amino acids), specific mutations of Cs and module 13 at various sites (including H9721A in C domain, S10255A in T domain, and ΔTE<sub>1</sub>-S10391A, ΔTE<sub>2</sub>-S10624A, and ΔTE<sub>1/2</sub>-S10391A/S10624A in TE domain), as well as the insertion of promoter P<sub>13</sub> in the front of TE domain (Supplementary Fig. 4a). The *BstZ171*-cm-*ccdB*-*BstZ171* cassette containing mutation site with 40 bp homology arm by PCR amplification can be stitched with the insertion sequence using overlap extension PCR. Two *BstZ171* sites were flanked by 30-bp direct repeats. The *BstZ171*-cm-*ccdB*-*BstZ171* cassette was inserted into the target site by LCHR in the *E. coli* strain GBRed-gyr462. Correct recombinant plasmids were selected by restriction digest analysis. The *cm-ccdB* cassette was removed by *BstZ171* digestion and

the terminal 30-bp overlaps were exposed. The linear fragments were circled by T4pol reaction in vitro and then the mixture was transformed into *E. coli* GB05. The resulting recombinant constructs were analyzed by restriction analysis and confirmed by DNA sequencing.

The fragments containing mutations and knockout of Cs domain were obtained by PCR using primer *gdnCs*-150HA-primer-S/*gdnCs*-150HA-primer-A, respectively. Then, the fragments were transferred into the mutant *S. brevitalea* DSM 7029Δ*Cs-genta* containing recombineering system Redy-Redαβ7029, which was constructed by the *genta* fragment replacing the Cs domain, generating two mutants of Cs domain. The region (-49.5 kb) containing *gdn* gene cluster was replaced in-situ by the *attB-genta* fragment in the strain *S. brevitalea* DSM 7029 containing the recombineering system Redy-Redαβ7029, and we obtained a recombinant *S. brevitalea* DSM 7029:*attB-genta* (Supplementary Fig. 4c, d). In addition, the above nine correct mutant plasmids containing different deletions and mutations of module 13, the promoter P<sub>13</sub> insertion<sup>53</sup>, and wild type plasmid p15A-amp-attP-P<sub>Apra</sub>-*gdn* were transformed into the recombinant *S. brevitalea* DSM 7029:*attB-genta* generating different recombinants. The recombinants were verified by PCR and DNA sequencing. All plasmids generated in this study and oligonucleotides are listed in Supplementary Data 1.

### Multi-domain swapping of C-terminus of *gdn* BGC

The different kind of multi-domain swapping was conducted by LCHR and linear plus linear homologous recombination (LLHR). For engineering *rzmA*\* gene cluster, a linear fragment *cm-ccdB* cassette flanked with homology arms was amplified from plasmid R6K-cm-*ccdB* by PCR. Then, the *cm-ccdB* cassette and plasmid p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\* (p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*CsR148A) were transformed into *E. coli* GBRed-gyr462 and recombined to form p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\*-cm-*ccdB*. Another linear fragment T<sub>12</sub>C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>TE amplified from *gdn* gene cluster recombined with a linear p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\*-cm-*ccdB* digested by *Pme*I yielding target plasmid p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\*-MI containing T<sub>12</sub>C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>TE domains of *gdn*. For another construct of *rzmA*\* gene cluster, another fragment *cm-ccdB* cassette with homology arms recombined with p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\*-MI to obtain p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\*-MI-cm-*ccdB*, which was linearized by digestion of *Pme*I in the next. The linear T domain of *rzmA* by PCR recombined with the above linear *rzmA*\*-MI-cm-*ccdB* to form plasmid p15A-*apra*-phiC31-P<sub>km</sub>-*rzmA*\*-M2 containing C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>TE domains of *gdn* BGC.

For constructing the recombinants of the *holA* BGC, the procedure was the similar to the mutants of *rzmA*\*. The linear fragment *cm-ccdB* cassettes flanked with homology arms were amplified from R6K-cm-*ccdB* by PCR using different primers. The acquired *cm-ccdB* fragments were recombined to p15A-*apra*-phiC31-P<sub>km</sub>-*holA* in *E. coli* GBRed-gyr462 by LCHR, respectively, which yielded different recombinant plasmids containing *cm-ccdB*. Then, fragments T<sub>12</sub>C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>TE, T<sub>12</sub>C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>, C<sub>13</sub>A\*<sup>T</sup><sub>13</sub>TE, T<sub>12</sub>C<sub>13</sub>A\*, A\*<sup>T</sup><sub>13</sub>TE, C<sub>13</sub>A\*, and C<sub>13</sub> were amplified from *gdn* gene cluster, respectively. The different plasmids containing *cm-ccdB* were linearized by digestion of *Pme*I. The PCR products and corresponding linear plasmids recombined to form target recombinant plasmids p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-MI, p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M6, p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M3, p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M9, p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M4, p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M5, and p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M8. In addition, two fragments T<sub>5</sub> and T<sub>6</sub> were amplified from *holA* BGC by PCR, which were fused with *cm* cassette, respectively, obtaining two linear products T<sub>5</sub>-*cm* and T<sub>6</sub>-*cm*. The fragment T<sub>5</sub>-*cm* recombined with p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-MI in *E. coli* GB08Red. Likewise, T<sub>6</sub>-*cm* recombined with p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M6 in *E. coli* GB08Red. Subsequently, the final plasmids p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M2 and p15A-*apra*-phiC31-P<sub>km</sub>-*holA*-M7 were obtained from the two recombinant plasmids by RedEx<sup>67</sup>. The melting sites were shown in Supplementary Fig. 25.

## Metabolic analysis

The fermentation and extraction of metabolites from the activation mutant and other mutants were the same as mentioned in our previous study<sup>40</sup>. The crude extracts were analyzed by UHPLC–HR–MS and LC–MS, respectively. The UHPLC system was performed using an ODS column (Luna RP-C<sub>18</sub>, 4.6 × 250 mm, 5 μm, 0.75 mL/min) with gradient elution. Mass spectra was acquired in centroid mode ranging from 100 to 1500 *m/z* with positive-mode electrospray ionization and auto MS<sup>2</sup> fragmentation. HPLC parameters were as follows: solvent A, H<sub>2</sub>O with 0.1% formic acid (FA); solvent B, 0.1% FA in acetonitrile (ACN); gradient at a constant flow rate of 0.75 mL/min, 0–5 min, 5% B; 5–55 min, 5–95% B; 55.1 min, 95% B; 55.1–60 min, 95% B; detection by UV spectroscopy at 200–400 nm. The HPLC–MS or LC–MS system was performed using Acclaim<sup>TM</sup> RSLC 120 C<sub>18</sub> column (2.1 × 100 mm, 2.2 μm) and its condition was: 0–3 min, 5% ACN; 3–18 min, 5–95% ACN; 18.1 min, 95% ACN; 18.1–22 min, 95% ACN; 22.1 min, 5% ACN; 22.1–25 min, 5% ACN, ACN and H<sub>2</sub>O contained 0.1% FA; flow rate by 0.3 mL/min; detection by UV spectroscopy at 200–400 nm.

## Purification of glidonins and lipopeptide derivatives

For compounds produced by *gdn* BGC of *S. brevitalea* DSM 7029 P<sub>Apra</sub><sup>-</sup> *gdn*, gradient elution (CH<sub>2</sub>Cl<sub>2</sub>: MeOH) resulted in 40 fractions. Fr.34–Fr.36 (CH<sub>2</sub>Cl<sub>2</sub>: MeOH 1:1) were purified by preparative MPLC (ODS; 5 μm, 250 × 20 mm, gradient elution 0–5 min, 20% MeOH; 5–55 min, 20–70% MeOH; 55.1 min, 95% MeOH; 55.1 min–60 min, 95% MeOH; H<sub>2</sub>O with 0.1% trifluoroacetic acid (TFA), 8 mL/min) to afford to Fr.1–Fr.4. Fr.2 was further purified by semipreparative HPLC (ODS; 5 μm, 250 × 10 mm, gradient elution 0–5 min, 30% ACN; 5–35 min, 30–60% ACN; 35.1 min, 95% ACN; 35.1–40 min, 95% ACN; H<sub>2</sub>O with 0.2% TFA, 3 mL/min) to yield compound **1** (100 mg). Fr.28–Fr.29 (CH<sub>2</sub>Cl<sub>2</sub>: MeOH 1:1) were purified by preparative MPLC (ODS; 5 μm, 250 × 20 mm, gradient elution 0–5 min, 10% MeOH; 5–55 min, 10–70% MeOH; 55.1 min, 95% MeOH; 55.1 min–60 min, 95% MeOH; H<sub>2</sub>O with 0.1% TFA, 8 mL/min) to afford to Fr.1–Fr.7. Fr.5 was further purified by semipreparative HPLC (ODS; 5 μm, 250 × 10 mm, isogradient elution 30% ACN, H<sub>2</sub>O with 0.2% TFA, 3 mL/min) to yield compound **2** (30 mg), Fr.7 was further purified by semipreparative HPLC (ODS; 5 μm, 250 × 10 mm, gradient elution 0–5 min, 34% ACN; 5–25 min, 34–48% ACN; 25.1 min, 95% ACN; 25.1 min–30 min, 95% ACN; H<sub>2</sub>O with 0.2% TFA, 3 mL/min) to yield compound **4** (28 mg). Fr.20–Fr.24 (CH<sub>2</sub>Cl<sub>2</sub>: MeOH 3:1) were isolated by negative column HPLC (MeOH: H<sub>2</sub>O) and resulted in 10 fractions. Fr.8 (MeOH: H<sub>2</sub>O, 60%) was further purified by semipreparative HPLC (ODS; 5 μm, 250 × 10 mm, gradient elution 0–5 min, 55% ACN; 5–38 min, 55–85% ACN; 38.1 min, 95% ACN; 38.1–43 min, 95% ACN; H<sub>2</sub>O with 0.2% TFA, 3 mL/min) to yield compound **7** (14 mg), compound **9** (20 mg), and compound **10** (15 mg). Compound **6** (7 mg) and **8** (5 mg) were purified from Fr.7 by semipreparative HPLC with the condition: 0–5 min, 55% ACN, 5–23 min, 55–65% ACN, 23.1 min, 95% ACN, 23.1–28 min, 95% ACN; H<sub>2</sub>O with 0.2% TFA, 3 mL/min.

Glidonin A (**1**): colorless oil; [α]<sup>20</sup>D-25 (c 0.1, MeOH); UV (MeOH) λ<sub>max</sub> 192 nm; IR (KBr) ν<sub>max</sub> 3308, 2969, 1868, 1681, 1540, 1455, 1207, 1138, 841, 802, 724 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 2; HR-ESI-MS: *m/z* 1359.7254 [M+H]<sup>+</sup> (calculated for C<sub>66</sub>H<sub>99</sub>N<sub>16</sub>O<sub>14</sub>S, 1359.7242). Glidonin B (**2**): colorless oil; [α]<sup>20</sup>D-27 (c 0.17, MeOH); UV (MeOH) λ<sub>max</sub> 191 nm; IR (KBr) ν<sub>max</sub> 3311, 2969, 1868, 1681, 1541, 1454, 1204, 1137, 838, 802, 723 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 3; HR-ESI-MS: *m/z* 1375.7206 [M+H]<sup>+</sup> (calculated for C<sub>66</sub>H<sub>99</sub>N<sub>16</sub>O<sub>15</sub>S, 1375.7191). Glidonin D (**4**): colorless oil; [α]<sup>20</sup>D-20 (c 0.12, MeOH); UV (MeOH) λ<sub>max</sub> 191 nm; IR (KBr) ν<sub>max</sub> 3309, 2969, 1868, 1653, 1540, 1455, 1205, 1138, 839, 802, 723 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 4; HR-ESI-MS: *m/z* 1385.7381 [M+H]<sup>+</sup> (calculated for C<sub>66</sub>H<sub>99</sub>N<sub>16</sub>O<sub>15</sub>S, 1385.7046). Glidonin F (**6**): colorless oil; [α]<sup>20</sup>D-22 (c 0.1, MeOH); UV (MeOH) λ<sub>max</sub> 221 nm; IR (KBr) ν<sub>max</sub> 3305, 2928, 1795, 1652, 1540, 1455, 1204, 1136, 836, 801, 723 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 4;

HR-ESI-MS: *m/z* 757.4335 [M+2H]<sup>2+</sup> (calculated for C<sub>75</sub>H<sub>117</sub>N<sub>16</sub>O<sub>15</sub>S, 1513.8600). Glidonin G (**7**): colorless oil; [α]<sup>20</sup>D-23 (c 0.12, MeOH); UV (MeOH) λ<sub>max</sub> 221 nm; IR (KBr) ν<sub>max</sub> 3306, 2929, 1868, 1651, 1540, 1455, 1204, 1137, 837, 802, 722 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 5; HR-ESI-MS: *m/z* 771.4479 [M+2H]<sup>2+</sup> (calculated for C<sub>77</sub>H<sub>121</sub>N<sub>16</sub>O<sub>15</sub>S, 1541.8913). Glidonin H (**8**): colorless oil; [α]<sup>20</sup>D-28 (c 0.1, MeOH); UV (MeOH) λ<sub>max</sub> 221 nm; IR (KBr) ν<sub>max</sub> 3310, 2969, 1868, 1681, 1541, 1454, 1204, 1137, 838, 802, 723 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 5; HR-ESI-MS: *m/z* 779.4772 [M+2H]<sup>2+</sup> (calculated for C<sub>77</sub>H<sub>121</sub>N<sub>16</sub>O<sub>16</sub>S, 1557.8862). Glidonin I (**9**): colorless oil, [α]<sup>20</sup>D-25 (c 0.15, MeOH); UV (MeOH) λ<sub>max</sub> 221 nm; IR (KBr) ν<sub>max</sub> 3310, 2928, 1795, 1652, 1540, 1455, 1204, 1137, 837, 802, 723 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 6; HR-ESI-MS: *m/z* 762.4718 [M+2H]<sup>2+</sup> (calculated for C<sub>78</sub>H<sub>123</sub>N<sub>16</sub>O<sub>15</sub>, 1523.9348). Glidonin J (**10**): colorless oil, [α]<sup>20</sup>D-21 (c 0.18, MeOH); UV (MeOH) λ<sub>max</sub> 222 nm; IR (KBr) ν<sub>max</sub> 3305, 2927, 1868, 1669, 1540, 1454, 1204, 1136, 836, 802, 722 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR, Supplementary Table 6; HR-ESI-MS: *m/z* 785.4645 [M+2H]<sup>2+</sup> (calculated for C<sub>79</sub>H<sub>123</sub>N<sub>16</sub>O<sub>15</sub>S, 1569.9226).

To obtain the two derivatives **29** and **31**, a 16 L scale fermentation of the mutant *E. coli* GB05-MtaA:RzmA\*-M2 and a 12 L scale fermentation of the mutant *E. coli* GB05-MtaA:HolA-M1 were prepared. The crude extracts of bacterial fermentation were purified by normal phase column and semipreparative HPLC. The condition of the gradient elution for generating **29** (3.6 mg) was 0–5 min, 40% ACN; 5–30 min, 40–65% ACN; 30.1 min, 95% ACN; 30.1–35 min, 95% ACN; H<sub>2</sub>O with 0.1% TFA, 3 mL/min. And the condition of the gradient elution for generating **31** (3.6 mg) was 0–5 min, 46% ACN; 5–24 min, 46–58% ACN; 24.1 min, 95% ACN; 24.1–29 min, 95% ACN; H<sub>2</sub>O with 0.1% TFA, 3 mL/min. In addition, the derivatives and intermediates (or original products) produced by the engineering *rzmA\** and *holA* BGCs were quantified using the standard curve of compounds **28** (for quantification of compound **28**), **29a** (for quantification of compound **28a**), **29** (for quantification of compound **29**), **30a** (for quantification of compounds **30**, **30a**, **30b**), and **31** (for quantification of compound **31**) using HPLC–MS measurement as described above (5%–95% ACN in 15 min). Triplicates of all experiments were measured.

The mutant strain *E. coli* GB05-MtaA:RzmA\*-M2 was cultured with three different diamines 1,7-diaminoheptane, 1,6-diaminohexane, and 1,5-diaminopentane with the concentration of 1.6 μg/mL, respectively. After three days, the XAD-16 resin was added into the fermentation, and the mixture was cultured continually for one day. Then, the fermentations were extracted using MeOH as previous study<sup>40</sup>. The crude extracts were analyzed by UHPLC–MS (5%–95% ACN in 50 min) as the above procedure.

## Marfey's analysis of the constituents of glidonins

Approximate 0.5–1 mg of compounds was hydrolyzed with 6 M HCl (0.5 mL) at 90 °C overnight. The solutions were dried by nitrogen, and the mixture was dissolved in 200 μL of H<sub>2</sub>O. The standard amino acids were also dissolved in 200 μL of H<sub>2</sub>O. Then, the above samples were added 1 M NaHCO<sub>3</sub> (25 μL) and 1% 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide (L-FDAA, 200 μL)<sup>49</sup>, and the mixture was vortexed and incubated at 40 °C for 60 min. Then, 100 μL of 2 M HCl was added to quench the reaction and the filtered solution was analyzed by LC–MS using Acclaim<sup>TM</sup> RSLC 120 C<sub>18</sub> column (2.1 × 100 mm, 2.2 μm) with a linear gradient of ACN with elution conditions (5%–50% ACN in 35 min) at a flow rate of 0.3 mL/min and UV detection at 330 nm. The LC–MS condition: 0–3 min, 5% ACN; 3–38 min, 5–50% ACN; 38.1 min, 95% ACN; 38.1–42 min, 95% ACN; 42.1 min, 5% ACN; and 42.1–45 min, 5% ACN, ACN, and H<sub>2</sub>O contained 0.1% FA.

## Purification and characteristics of enzymes

Primers used for amplification of acylase *GdnD*, methionyl-tRNA formyltransferase (*Fmt*) from the genomic DNA of *S. brevitalea* DSM 7029, primers used for amplification of *CsA1T1*, *A1T1*, *A12T12*, *A12T12C13*,

$A_{12}T_{12}C_{13}A^*$ ,  $A_{12}T_{12}C_{13}A^*T_{13}$ ,  $A_{12}T_{12}C_{13}A^*T_{13}TE$  from the plasmid p15A-cm-gdn are listed in Supplementary Data 1. The mutated fragments including  $A_{12}T_{12}C_{13}A^*T_{13}TE_{Mut1}$  (S10391A),  $A_{12}T_{12}C_{13}A^*T_{13}TE_{Mut2}$  (S10624A), and  $A_{12}T_{12}C_{13}A^*T_{13}TE_{Mut1/2}$  (S10391A/S10624A), and others were obtained by fused PCR. The resulting DNA fragments were cloned into the commercial pET-28b(+) vector via LLHR were transformed into *E. coli* GB05Dir. Then, the resulting constructs were verified by restriction digestion and DNA sequencing. The correct plasmids were transformed into *E. coli* BL21 or *E. coli* BAP1 for expressing the His-tagged protein. Each recombinant *E. coli* BL21 or *E. coli* BAP1 was incubated in LB medium supplemented with km (10  $\mu$ g/mL) for overnight at 37 °C. Overnight 30 mL culture was transformed into 2 L fresh LB medium supplemented with km at 37 °C, 180 rpm for 2 h, and the growth was continued to an OD<sub>600</sub> reached to 0.6 at 16 °C, 160 rpm. Then, the addition of 0.2 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) to the cultures that were further incubated at 16 °C for additional 20 h. For protein purification according to our previous study<sup>41</sup>, the cells were harvested via centrifugation (8,228  $\times$  g, 4 °C, 10 min). The pellet was resuspended in buffer A (50 mM Tris, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, 40 mM imidazole, pH 7.6), and lysed by cell disruption. Cell debris was removed by centrifugation (60,325  $\times$  g, 4 °C, 30 min), and the supernatant was bound to Ni-NTA resin. The protein was eluted with buffer B (50 mM Tris, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, 400 mM imidazole, pH 7.6). The buffer of the elution fraction was exchanged with PD10 column using buffer C (50 mM Tris, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.6). Then, proteins were concentrated with centrifugal filters with a molecular weight cut off (30 kDa). Finally, proteins were portioned into 100  $\mu$ L aliquots, flash frozen on liquid N<sub>2</sub>, and stored at -80 °C.

For pyrophosphate production assays<sup>68</sup>, the reaction system contains 10 mM amino acid substrates, 2.5 mM ATP (Solarbio), 0.2 mM MESG (Medchemexpress), 1 U/mL purine nucleoside phosphorylase (Shanghai yuanye Bio-Technology Co., Ltd), 0.4 U/mL inorganic pyrophosphatase (Sigma), 150 mM hydroxylamine (Sigma) and 1  $\mu$ M proteins in 100  $\mu$ L buffer (50 mM Tris, 300 mM NaCl and 5 mM MgCl<sub>2</sub>, pH 7.6). The value was read at A360 in a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, California) in 96 well, clear-bottomed plates (Corning). A360 values were converted to pyrophosphate release, by comparing with a standard curve (KH<sub>2</sub>PO<sub>4</sub> from 2  $\mu$ M to 128  $\mu$ M) for known quantities of pyrophosphate.

In vitro assay for CsA<sub>1</sub>T<sub>1</sub> domain was performed. Reaction buffer (50 mM Tris, 300 mM NaCl and 5 mM MgCl<sub>2</sub>, pH 7.6) contained 80  $\mu$ M protein, 3  $\mu$ M *Bacillus subtilis* phosphopantetheinyl transferase (Sfp), 200  $\mu$ M Coenzyme A (CoA), 2.5 mM substrates of amino acids (i.e. L-Met, L-Met(O) and L-Leu respectively), and 5 mM ATP was incubated at 30 °C for 30 min. Then, 5 mM acyl-CoA (C12-CoA and C14-CoA respectively) was added to make final volume of 200  $\mu$ L. The reaction was further incubated at 30 °C for 4 h. Products was hydrolyzed as reported<sup>7</sup>. In brief, triple amount of cold methanol was added to quench the reaction, and the precipitate was collected at 10,000  $\times$  g for 8 min. The pellet dissolved in 125  $\mu$ L of 0.1 M KOH and was heated at 70 °C for 10 min after washed twice and dried. Then the protein was neutralized with HCl after cooling to room temperature. One milliliter of the methanol was added to precipitate the protein overnight at -20 °C. The hydrolysate was centrifuged at 10,000  $\times$  g for 15 min to remove precipitated protein, and the supernatant was applied LC-MS analysis (5%–95% ACN in 15 min).

Protein-MS of T<sub>1</sub> domain linked substrates was conducted as the previous study<sup>69</sup>. Each sample (100  $\mu$ L) was completely mixed with 400  $\mu$ L methanol, 100  $\mu$ L chloroform, and 300  $\mu$ L ddH<sub>2</sub>O by gentle vortexing. Samples were centrifuged for 15 min at 18,407  $\times$  g, 4 °C. The top supernatant of each sample was aspirated carefully to not disturb thin layer of the precipitated protein located within the organic/aqueous interface. An additional 400  $\mu$ L of methanol was added to each sample again followed by gentle vortexing, and the mixture centrifuged for 15 min. Solvent was completely removed and the pellet was

dried briefly. To ensure complete removal of interfering substances, an additional 400  $\mu$ L of methanol was added to each pellet. The resulting pellets were chilled at -20 °C for 15 min and incubated in 25  $\mu$ L of 80% FA at -20 °C for 2 min. The samples were then mixed and incubated at -20 °C for an additional 15 min. 75  $\mu$ L of cold ddH<sub>2</sub>O was then added to each sample, followed by a final 1:50 dilution in 49.95% ACN, 49.95% ddH<sub>2</sub>O, and 0.1% FA (v/v). In all, 4  $\mu$ L sample was detected using a UHPLC system coupled to an impact HD QTOF (Bruker) with ACQUITY UPLC Protein BEH C<sub>4</sub> column (300 Å, 1.7  $\mu$ m, 2.1 mm  $\times$  100 mm). For the LC, Solvent A: 95% H<sub>2</sub>O, 5% ACN, 0.2% FA and Solvent B: 5% H<sub>2</sub>O, 95% ACN, 0.2% FA were used with the following gradient: 0–5 min, 5% B; 5–42 min, 5–60% B; 42–44 min, 60–95% B; 44–46 min, 95% B; 46–47 min, 95–5% B; 47–60 min, 5% B, at a flow rate of 0.2 mL/min with the mass range at 600–2000 m/z.

In vitro assays for the different enzymes of the C-terminal multi-domains were performed. Reaction buffer contained 80  $\mu$ M protein, 3  $\mu$ M Sfp from *B. subtilis*, 200  $\mu$ M CoA, 2.5 mM L-Ala and 5 mM ATP was incubated at 30 °C for 30 min. Then, 2.5 mM diamines or polyamines (i.e. putrescine, 1,7-diaminoheptane, 1,6-diaminohexane, 1,5-diaminopentane, 1,2-diaminocyclohexane, spermine, and spermidine, respectively) was added to make final volume of 200  $\mu$ L. The reaction was further incubated at 30 °C for 1 h or 4 h. Triple amount of cold methanol was added to quench the reaction, and the supernatant was dried by SpeedVac after removed the precipitate at 10,000  $\times$  g for 8 min. The reaction mixture was dissolved with 200  $\mu$ L ddH<sub>2</sub>O, and then the mixture was derived with D-FDAA according to the procedure of the above Marfey's analysis. The final reaction mixture was suspended using 100  $\mu$ L 2 M HCl, which subsequently was dried by SpeedVac. Finally, the mixture dissolved with 100  $\mu$ L MeOH to detect the target products (**24a**, **25a**, **26a**, and **27a**) using LC-MS analysis (5%–95% ACN in 15 min).

The hydrolysis function of PvdQ-like acylase GdnD was tested. GdnD is predicted to be a periplasmic protein by PSORTb<sup>13</sup>. The collected pellet from 50 mL LB dissolved with 8 mL buffer C (50 mM Tris, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, pH 7.6), and lysed by cell disruption. Then, one hundred microliter of the cell lysate was incubated with 5  $\mu$ L 10 mM substrates **7** and **8** for 4 h, 16 h, and 27 h at 30 °C. The negative control was boiled cell lysate. The other substrates (**13**, **14**, **15**, **16**, and **17**) incubated with the cell lysate dealt for 16 h at 30 °C. Then, the reaction was stopped by the addition of the 100  $\mu$ L MeOH, and the final mixture was checked by LC-MS analysis (5%–95% ACN in 15 min) according to the above condition.

The in vitro assay of formyltransferase was tested<sup>70</sup>. In brief, the formyl-donor co-substrate intermediate N<sup>5</sup>, N<sup>10</sup>-methenylH<sub>4</sub>F was first synthesized as follow: 14 mg of N<sup>5</sup>-fH<sub>4</sub>F was dissolved in 3 mL of water and was converted to N<sup>5</sup>, N<sup>10</sup>-methenylH<sup>4</sup>F through addition of 0.1 M HCl, until pH 1.9 was reached. The solution was then brought to a final volume of 4.4 mL with water and continually incubated at room temperature for 4 h. Then, the transformylation assay was setup in a 50  $\mu$ L volume, in the presence of 50 mM HEPES buffer pH 7.5, 1 mM substrates (compounds **1**, **1a**, **13a**), 1.5 mM N<sup>5</sup>, N<sup>10</sup>-methenylH<sub>4</sub>F, and 25  $\mu$ M Fmt (35 kDa). Before the addition of the compounds and the enzyme, the intermediate N<sup>5</sup>, N<sup>10</sup>-methenylH<sub>4</sub>F was preincubated in the reaction buffer at 30 °C for 30 min, to permit the final pH-dependent conversion to N<sup>10</sup>-fH<sub>4</sub>F. After the addition of the compounds and the enzyme, the reaction was allowed to proceed for 4 h at 30 °C, then ceased by the addition of 2  $\mu$ L FA, and analyzed via HPLC-MS. The condition for **1**: 0–3 min, 20% ACN; 3–18 min, 20–60% ACN; 18.1 min, 95% ACN; 18.1–22 min, 95% ACN; 22.1 min, 20% ACN; 22.1–25 min, 20% ACN, ACN and H<sub>2</sub>O contained 0.1% FA; detection by UV spectroscopy at 200–400 nm. The condition for **1a** and **13a**: 0–3 min, 5% ACN; 3–18 min, 5–95% ACN; 18.1 min, 95% ACN; 18.1–22 min, 95% ACN; 22.1 min, 5% ACN; 22.1–25 min, 5% ACN, ACN and H<sub>2</sub>O contained 0.1% FA; detection by UV spectroscopy at 200–400 nm.

## Bioactivity assays screening of glidonins and derivatives

The two human acute lymphoblastic leukemia cells K562 and Kasumi, two human breast adenocarcinoma cells MCF7 and MDA-MB-231, human hepatoma cell HepG-2, human lung adenocarcinoma cell A549, human gastric cancer cell SGC-7901, human colon cancer cell HCT-116, human ovarian cancer cell line SKOV3, or human lung epithelial normal cell 2B were obtained from Beyotime Biotechnology and utilized to test the cytotoxicity of compounds by MTT assays as previous study<sup>71</sup>. The anti-inflammatory effect of the compounds was detected by the influence on the increase in the level of nitric oxide caused by lipopolysaccharide (LPS) based on the classic Griess method. RAW 264.7 cells (Shanghai Institutes for Biological Science) were seeded in 96-well plates and incubated for 24 h. Then, cells were treated with compounds **29**, **29a**, **31**, and **31a** (0–80  $\mu$ M) for 30 min, followed by co-treatment with LPS (10 mg/mL) for another 24 h. The Nitric Oxide assay kit with Griess reagents (Beyotime, Lot: S0021, Shanghai, China) was used to examine cellular supernatant nitrite accumulation, which represents cellular NO levels. The OD value of 540 nm absorbance was detected with a microplate reader. Their authentication in the laboratory was performed regularly based on morphology and gene/protein expression. The tested micro-organisms including Gram-negative bacteria *E. coli* ATCC 35218 and *Pseudomonas aeruginosa* ATCC 27853, and Gram-positive bacteria *Staphylococcus aureus* ATCC 29213 and *B. subtilis* ATCC 6633, and fungi *Candida albicans*, *Rhizoctonia solani* and *Aspergillus fumigatus* were treated with different concentrations of compounds (1  $\mu$ M, 10  $\mu$ M, 100  $\mu$ M) using Bauer-Kirby disk diffusion method to evaluate the antimicrobial activities<sup>72</sup>.

Comparing the concentration of compounds (**7**, **10**, **10a**) in extracellular and intracellular after the HepG2 cell line dealt with compounds. HepG2 cells were plated in six-well plates at a density of  $1.5 \times 10^5$  cells/well and grew overnight. After cell treated with compounds (5  $\mu$ g/mL) for 3 h, the compounds were extracted with MeOH for cells culture medium and cells separately as followed the reported method<sup>73</sup>. The compounds were quantified by LC–MS (5%–95% ACN in 15 min). The cell numbers of HepG2 with the compounds treatments was determined.

## Physiology and biochemistry test

The growth curve was tested by Bioscreen C (Labsystems, Finland) as follow. Overnight culture was inoculated into 400  $\mu$ L fresh CYMG medium in the ratio of 1:100 (the concentration of compounds was 0.1  $\mu$ g/ $\mu$ L) and shook in normal speed. The OD<sub>600</sub> value was gained in every hour for 42 h in total. Biofilm formation was measured at 18 h and 36 h respectively, as reported previously<sup>74</sup>. In brief, the free-floating cells were removed, and the microtitre plate was washed with PBS (pH 7.4) after incubation for 18 h and 36 h. The biofilm was stained by adding 200  $\mu$ L of crystal violet (0.1%) solution. Crystal violet solution was removed after 15 min and 200  $\mu$ L of ethanol (95%) was added. The absorbance at OD<sub>600</sub> was read using a microplate reader (Infinite M200, Tecan, Männedorf, Switzerland). Meanwhile, free-floating cells were used to test the secretion of pigment. The bacteria were removed and the supernatant was measured at OD<sub>600</sub>.

## Rigid receptor docking

The structure of GdnD was predicted by AlphaFold2<sup>54</sup>. The structure of the PvdQ in complex with the pyoverdine precursor PVDIq (PDB number: 5UBK) was retrieved from RCSB Protein Data Bank (PDB). GdnD was docked with substrate C14-D-Met(O) in the PVDIq binding pocket of PvdQ using AutoDock Vina. To cover the entire binding site, a 2 nm side cubic grid box was positioned at the geometric center of residues S486 (PvdQ S268), N497 (PvdQ D269), R516 (PvdQ R297), and H683 (PvdQ H457) as reported<sup>75</sup>. A total of 20 poses for one substrate were generated through the docking procedure, and the complex that closely matched the binding pose of PVDIq in PvdQ was selected independently for further analysis.

The docking procedure for the C<sub>13</sub> domain with Put followed a similar approach. RzmA mutant H140A/R148A in complex with C8-CoA and Leu-SNAC (PDB number: 7CIS) was used as the model complex and 2 nm side cubic grid box was positioned at the geometric center of residues H9721 (RzmA H140), T9879 (RzmA T289), V9877 (RzmA S287), and L9860 (RzmA P271) as reported<sup>56</sup>. The complex that closely matched the binding pose of Leu-SNAC in RzmA and 4'-phosphopantetheine in the condensation domain of holo-AB3403 (PDB number: 4ZXI<sup>76</sup>) was selected for further analysis.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. The following public databases were utilized for the analysis of the protein structures: 5UBK, 7CIS, 4ZXI, 3L9I, IJMK. Source data are provided with this paper.

## References

- Schwarzer, D., Finking, R. & Marahiel, M. A. Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* **20**, 275–287 (2003).
- Winn, M., Fyans, J. K., Zhuo, Y. & Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* **33**, 317–347 (2016).
- Süssmuth, R. D. & Mainz, A. Nonribosomal peptide synthesis—principles and prospects. *Angew. Chem. Int. Ed. Engl.* **56**, 3770–3821 (2017).
- Miller, B. R. & Gulick, A. M. Structural biology of nonribosomal peptide synthetases. *Methods Mol. Biol.* **1401**, 3–29 (2016).
- Wittmann, M., Linne, U., Pohlmann, V. & Marahiel, M. A. Role of DptE and DptF in the lipidation reaction of daptomycin. *FEBS J.* **275**, 5343–5354 (2008).
- Kraas, F. I., Helmetag, V., Wittmann, M., Strieker, M. & Marahiel, M. A. Functional dissection of surfactin synthetase initiation module reveals insights into the mechanism of lipoinitiation. *Chem. Biol.* **17**, 872–880 (2010).
- Imker, H. J., Krahn, D., Clerc, J., Kaiser, M. & Walsh, C. T. N-acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF. *Chem. Biol.* **17**, 1077–1083 (2010).
- Hansen, D. B., Bumpus, S. B., Aron, Z. D., Kelleher, N. L. & Walsh, C. T. The loading module of mycosubtilin: an adenylation domain with fatty acid selectivity. *J. Am. Chem. Soc.* **129**, 6366–6367 (2007).
- Auerbach, D., Yan, F., Zhang, Y. M. & Müller, R. Characterization of an unusual glycerate esterification process in vioprolide biosynthesis. *ACS Chem. Biol.* **13**, 3123–3130 (2018).
- Götze, S. & Stallforth, P. Structure, properties, and biological functions of nonribosomal lipopeptides from pseudomonads. *Nat. Prod. Rep.* **37**, 29–54 (2020).
- Baltz, R. H., Miao, V. & Wrigley, S. K. Natural products to drugs: daptomycin and related lipopeptide antibiotics. *Nat. Prod. Rep.* **22**, 717–741 (2005).
- Drake, E. J. & Gulick, A. M. Structural characterization and high-throughput screening of inhibitors of PvdQ, an NTN hydrolase involved in pyoverdine synthesis. *ACS Chem. Biol.* **6**, 1277–1286 (2011).
- Gauglitz, J. M., Iinishi, A., Ito, Y. & Butler, A. Microbial tailoring of acyl peptidic siderophores. *Biochemistry* **53**, 2624–2631 (2014).
- Baars, O. et al. Crochelins: siderophores with an unprecedented iron-chelating moiety from the nitrogen-fixing bacterium *Azotobacter chroococcum*. *Angew. Chem. Int. Ed. Engl.* **57**, 536–541 (2018).

15. Fu, C. Z. et al. Biosynthetic studies of telomycin reveal new lipopeptides with enhanced activity. *J. Am. Chem. Soc.* **137**, 7692–7705 (2015).
16. Reimer, D., Pos, K. M., Thines, M., Grün, P. & Bode, H. B. A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat. Chem. Biol.* **7**, 888–890 (2011).
17. Shen, Q. Y. et al. Characterization of a cryptic NRPS gene cluster in *Bacillus velezensis* FZB42 reveals a discrete oxidase involved in multithiazole biosynthesis. *ACS Catal.* **12**, 3371–3381 (2022).
18. Thongkongkaew, T. et al. Two types of threonine-tagged lipopeptides synergize in host colonization by pathogenic *Burkholderia* species. *ACS Chem. Biol.* **13**, 1370–1379 (2018).
19. Zheng, W. T. et al. Establishment of recombineering genome editing system in *Paraburkholderia megapolitana* empowers activation of silent biosynthetic gene clusters. *Microb. Biotechnol.* **13**, 397–405 (2020).
20. Chen, H. N. et al. Genomics-driven activation of silent biosynthetic gene clusters in *Burkholderia gladioli* by screening recombineering system. *Molecules* **26**, 700 (2021).
21. Li, R. J. et al. Development and application of an efficient recombineering system for *Burkholderia glumae* and *Burkholderia plantarii*. *Microb. Biotechnol.* **14**, 1809–1826 (2021).
22. Yoshimura, A. et al. Unlocking cryptic metabolites with mass spectrometry-guided transposon mutant selection. *ACS Chem. Biol.* **15**, 2766–2774 (2020).
23. Franke, J., Ishida, K. & Hertweck, C. Evolution of siderophore pathways in human pathogenic bacteria. *J. Am. Chem. Soc.* **136**, 5599–5602 (2014).
24. Franke, J., Ishida, K., Ishida-Ito, M. & Hertweck, C. Nitro versus hydroxamate in siderophores of pathogenic bacteria: effect of missing hydroxylamine protection in malleobactin biosynthesis. *Angew. Chem. Int. Ed. Engl.* **52**, 8271–52013 (2013).
25. Wang, X. Y. et al. Recombineering enables genome mining of novel siderophores in a non-model *Burkholderiales* strain. *Eng. Microbiol.* **3**, 100106 (2023).
26. Lanois-Nouri, A. et al. The odororhadin antibiotic biosynthetic cluster and acetyltransferase self-resistance locus are niche and species specific. *mBio* **13**, e0282621 (2022).
27. Perlatti, B., Nichols, C. B., Alspaugh, J. A., Gloer, J. B. & Bills, G. F. Sphaerostilbellins, new antimicrobial aminolipopeptide peptaibiotics from *Sphaerostilbella toxica*. *Biomolecules* **10**, 1371 (2020).
28. Fuchs, S. W. et al. Neutral loss fragmentation pattern based screening for arginine-rich natural products in *Xenorhabdus* and *Photorhabdus*. *Anal. Chem.* **84**, 6948–6955 (2012).
29. Zhao, L., Kaiser, M. & Bode, H. B. Rhabdopeptide/Xenortide-like peptides from *Xenorhabdus innexi* with terminal amines showing potent antiprotozoal activity. *Org. Lett.* **20**, 5116–5120 (2018).
30. Cai, X. F. et al. Entomopathogenic bacteria use multiple mechanisms for bioactive peptide library design. *Nat. Chem.* **9**, 379–386 (2017).
31. Müller, S. et al. Paenilamicin: structure and biosynthesis of a hybrid nonribosomal peptide/polyketide antibiotic from the bee pathogen *Paenibacillus larvae*. *Angew. Chem. Int. Ed. Engl.* **53**, 10821–10825 (2014).
32. Michael, A. M. Polyamines in eukaryotes, bacteria, and archaea. *J. Biol. Chem.* **291**, 14896–14903 (2016).
33. Schneider, J. & Wendisch, V. F. Biotechnological production of polyamines by bacteria: recent achievements and future perspectives. *Appl. Microbiol. Biotechnol.* **91**, 17–30 (2011).
34. Wendisch, V. F., Mindt, M. & Pérez-García, F. Biotechnological production of mono- and diamines using bacteria: recent progress, applications, and perspectives. *Appl. Microbiol. Biotechnol.* **102**, 3583–3594 (2018).
35. Li, B. et al. Alternative pathways utilize or circumvent putrescine for biosynthesis of putrescine-containing rhizoferrin. *J. Biol. Chem.* **296**, 100146 (2021).
36. Scherlach, K. & Hertweck, C. Mining and unearthing hidden biosynthetic potential. *Nat. Commun.* **12**, 3864 (2021).
37. Li, Z. Y., Zhu, D. Y. & Shen, Y. M. Discovery of novel bioactive natural products driven by genome mining. *Drug Discov. Ther.* **12**, 318–328 (2018).
38. Baral, B., Akhgari, A. & Metsä-Ketelä, M. Activation of microbial secondary metabolic pathways: avenues and challenges. *Synth. Syst. Biotechnol.* **3**, 163–178 (2018).
39. Huo, L. J. et al. Heterologous expression of bacterial natural product biosynthetic pathways. *Nat. Prod. Rep.* **36**, 1412–1436 (2019).
40. Wang, X. et al. Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in *Burkholderiales* species. *Proc. Natl. Acad. Sci. USA* **115**, E4255–E4263 (2018).
41. Chen, H. N. et al. Biosynthesis of glidomides and elucidation of different mechanisms for formation of  $\beta$ -OH amino acid building blocks. *Angew. Chem. Int. Ed. Engl.* **61**, e202203591 (2022).
42. Chen, H. N. et al. Identification of holrhizins E-Q reveals the diversity of nonribosomal lipopeptides in *Paraburkholderia rhizoxinica*. *J. Nat. Prod.* **83**, 537–541 (2020).
43. Liu, J. Q. et al. Rational construction of genome-reduced *Burkholderiales* chassis facilitates efficient heterologous production of natural products from proteobacteria. *Nat. Commun.* **12**, 4347 (2021).
44. Li, X. Z. & Nikaido, H. Efflux-mediated drug resistance in bacteria: an update. *Drugs* **69**, 1555–1623 (2009).
45. Stachelhaus, T., Mootz, H. D. & Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493–505 (1999).
46. Horsman, M. E., Hari, T. P. A. & Boddy, C. N. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat. Prod. Rep.* **33**, 183–202 (2016).
47. Blin, K. et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **49**, W29–W35 (2021).
48. Strieker, M., Nolan, E. M., Walsh, C. T. & Marahiel, M. A. Stereo-specific synthesis of *threo*- and *erythro*- $\beta$ -hydroxyglutamic acid during kutzneride biosynthesis. *J. Am. Chem. Soc.* **131**, 13523–13530 (2009).
49. Fujii, K. et al. A Nonempirical method using LC/MS for determination of the absolute configuration of constituent amino acids in a peptide: elucidation of limitations of Marfey's method and of its separation mechanism. *Anal. Chem.* **69**, 5146–5151 (1997).
50. Bloudoff, K. & Schmeing, T. Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. *Biochim. Biophys. Acta Proteins Proteom.* **1865**, 1587–1604 (2017).
51. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E. & Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**, 1790–1792 (2001).
52. Schoenafinger, G., Schracke, N., Linne, U. & Marahiel, M. A. Formylation domain: an essential modifying enzyme for the non-ribosomal biosynthesis of linear gramicidin. *J. Am. Chem. Soc.* **128**, 7406–7407 (2006).
53. Ouyang, Q. et al. Promoter screening facilitates heterologous production of complex secondary metabolites in *Burkholderiales* strains. *ACS Synth. Biol.* **9**, 457–460 (2020).
54. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
55. Brickman, T. J. & Armstrong, S. K. The ornithine decarboxylase gene *odc* is required for alcaligin siderophore biosynthesis in *Bordetella* spp.: putrescine is a precursor of alcaligin. *J. Bacteriol.* **178**, 54–60 (1996).



56. Zhong, L. et al. Engineering and elucidation of the lipoinitiation process in nonribosomal peptide biosynthesis. *Nat. Commun.* **12**, 296 (2021).
57. Niehs, S. P., Scherlach, K. & Hertweck, C. Genomics-driven discovery of a linear lipopeptide promoting host colonization by endofungal bacteria. *Org. Biomol. Chem.* **16**, 8345–8352 (2018).
58. Verstraeten, N. et al. Living on a surface: swarming and biofilm formation. *Trends Microbiol.* **16**, 496–506 (2008).
59. Kearns, D. B. A field guide to bacterial swarming motility. *Nat. Rev. Microbiol.* **8**, 634–644 (2010).
60. Keating, T. A., Marshall, C. G. & Walsh, C. T. Vibriobactin biosynthesis in *Vibrio cholerae*: VibH is an amide synthase homologous to nonribosomal peptide synthetase condensation domains. *Biochemistry* **39**, 15513–15521 (2000).
61. Seyedsayamdost, M. R. et al. Mixing and matching siderophore clusters: structure and biosynthesis of serratiochelins from *Serratia* sp. V4. *J. Am. Chem. Soc.* **134**, 13550–13553 (2012).
62. Barret, J.-M. et al. F14512, a potent antitumor agent targeting topoisomerase II vectored into cancer cells via the polyamine transport system. *Cancer Res.* **68**, 9845–9853 (2008).
63. Kruczyński, A. et al. F14512, a polyamine-vectorized anti-cancer drug, currently in clinical trials exhibits a marked preclinical anti-leukemic activity. *Leukemia* **27**, 2139–2148 (2013).
64. Gutiérrez-Chávez, C., Benaud, N. & Ferrari, B. C. The ecological roles of microbial lipopeptides: where are we going? *Comput. Struct. Biotechnol. J.* **19**, 1400–1413 (2021).
65. Raaijmakers, J. M., Bruijn, I. D., Nybroe, O. & Ongena, M. Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol. Rev.* **34**, 1037–1062 (2010).
66. Wang, H. L. et al. ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. *Nucleic Acids Res.* **46**, e28 (2018).
67. Song, C. Y. et al. RedEx: a method for seamless DNA insertion and deletion in large multimodular polyketide synthase gene clusters. *Nucleic Acids Res.* **48**, e130 (2020).
68. Wilson, D. J. & Aldrich, C. C. A continuous kinetic assay for adenylation enzyme activity and inhibition. *Anal. Biochem.* **404**, 56–63 (2010).
69. Donnelly, D. P. et al. Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat. Methods* **16**, 587–594 (2019).
70. Bosello, M., Mielcarek, A., Giessen, T. W. & Marahiel, M. A. An enzymatic pathway for the biosynthesis of the formylhydroxyornithine required for rhodochelin iron coordination. *Biochemistry* **51**, 3059–3066 (2012).
71. Mosmann, T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods.* **65**, 55–63 (1983).
72. Barry, A. L., Coyle, M. B., Thornsberry, C., Gerlach, E. H. & Hawkins, R. W. Methods of measuring zones of inhibition with the Bauer-Kirby disk susceptibility test. *J. Clin. Microbiol.* **10**, 885–889 (1979).
73. Baati, T., Schembri, T., Villard, C., Correard, F., Braguer, D. & Estève, M. A. An ultrasensitive LC-MS/MS method with liquid phase extraction to determine paclitaxel in both cell culture medium and lysate promising quantification of drug nanocarriers release in vitro. *J. Pharm. Biomed. Anal.* **115**, 300–306 (2015).
74. Zhong, L. et al. Attenuation of *Pseudomonas aeruginosa* quorum sensing by natural products: virtual screening, evaluation and bio-molecular interactions. *Int. J. Mol. Sci.* **21**, 2190 (2020).
75. Clevenger, K. D. et al. Substrate trapping in the siderophore tailoring enzyme PvdQ. *ACS Chem. Biol.* **12**, 643–647 (2017).
76. Drake, E. J. et al. Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235–238 (2016).

## Acknowledgements

We thank Haiyan Sui, Xiangmei Ren, Zhifeng Li, Jingyao Qu, Jing Zhu, and Guannan Lin from State Key Laboratory of Microbial Technology of Shandong University for help and guidance in NMR, LC and HRMS. This work was supported by National Key R&D Program of China (2019YFA0905700), National Natural Science Foundation of China (32070060, 32000040, 32100052, 32371488, 32161133013, 32201195), Shandong Provincial Natural Science Foundation (ZR2020QH345), the 111 project (B16030), and Youth Interdisciplinary Innovative Research Group (2020QNKQT009) and Fundamental Research Funds (2023QNTD001) of Shandong University.

## Author contributions

X.Bi. designed the research; X.Bi. and Yo.Z. supervised the experiments; H.C., L.Z., T.S., and X.Ba. performed gene cloning, site mutations, protein purification, in vivo deletion; H.C. and L.Z. performed bioinformatic analyses, in vitro experiments; H.C., X.W., and X.J. purified compounds and analyzed data; H.Z. performed structural elucidation; L.Z. performed structural prediction of proteins; Yi.Z. performed bioactivities of compounds; H.C., L.Z., H.Z., Q.T., and X.Bi. wrote and revised the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-42387-z>.

**Correspondence** and requests for materials should be addressed to Xiaoying Bian.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023