

Total In Vitro Biosynthesis of the Thioamide Thioholgamide and Investigation of the Pathway

Asfandyar Sikandar,[†] Maria Lopatniuk,[‡] Andriy Luzhetskyy,[‡] Rolf Müller^{†,‡,*} and Jesko Koehnke^{§, &, *}

[†] Department of Microbial Natural Products (MINS), Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Center for Infection Research (HZI) and Department of Pharmacy at Saarland University (UdS), Campus E8.1 66123, Saarbrücken, Germany.

[‡] Department of Pharmacy, Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Germany.

[#] German Centre for Infection Research (DZIF), Partner Site Hannover-Braunschweig Germany

[§] Workgroup Structural Biology of Biosynthetic Enzymes, HIPS, HZI, UdS, 66123 Saarbrücken, Germany

ABSTRACT: Thioholgamides are ribosomally synthesized and post-translationally modified peptides (RiPPs) with potent activity against cancerous cell lines and an unprecedented structure. Despite being one of the most structurally and chemically complex RiPPs, very few biosynthetic steps have been elucidated. Here, we report the complete *in vitro* reconstitution of the biosynthetic pathway. We demonstrate that thioamidation is the first step and acts as a gatekeeper for downstream processing. Thr dehydration follows thioamidation, and our studies reveal that both these modifications require the formation of protein complexes – ThoH/I and ThoC/D. Harnessing the power of AlphaFold we deduce that ThoD acts as a lyase and also propose putative catalytic residues. ThoF catalyzes the oxidative decarboxylation of the terminal Cys and the subsequent macrocyclization is facilitated by ThoE. This is followed by Ser dehydration, which is also carried out by ThoC/D. ThoG is responsible for histidine bis-*N*-methylation, which is a prerequisite for His β -hydroxylation – a modification carried out by ThoJ. The last step of the pathway is the removal of the leader peptide by ThoK to afford mature thioholgamide.

Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) have emerged as a major family of natural products, with broad chemical diversity and a wide range of bioactivities.¹⁻³ Thioamitides are a group of RiPPs that contain thioamides in place of amides in the peptide backbone as their class-defining feature.³ A subset of thioamitides display potent cytotoxic activity by inhibiting mitochondrial ATP synthase, which leads to mitochondrial dysfunction and in turn triggers apoptosis.^{4,6} The founding member of this group, thioviridamide, was isolated from *Streptomyces olivoviridis* NA005001,⁷ and the identification of its biosynthetic gene cluster (BGC)⁸ initiated the genomics-led discovery of other thioamitides including TVA-YJ-2, thiostreptamide S4, thiosparsoamide, and thioholgamide (Figure 1a).⁹⁻¹² Thioholgamide contains extensive post-translational modifications (PTMs) that are common to most thioamitides including four thioamide bonds, a β -hydroxy-N1, N3-dimethylhistidinium (hdmHis), and a C-terminal 2-aminovinyl-3-methyl-cysteine (AviMeCys) that forms part of the macrocycle.

Enzymes responsible for installing some of the modifications have either been predicted by their homology with proteins from other RiPP pathways, or functionally characterized by co-expression studies.^{11, 13, 14} Recently, the order of modifications has been proposed based on gene deletions and untargeted metabolomics studies (Figure S1a; we will

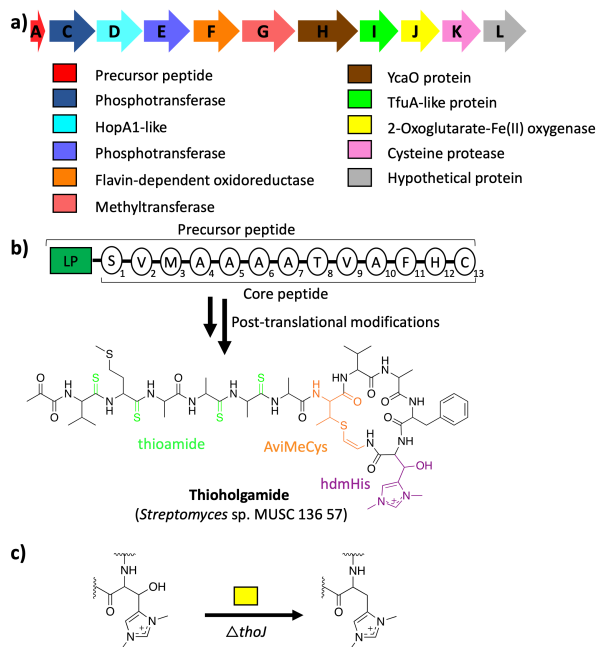


Figure 1. a) Organization of the thioholgamide BGC. b) The thioholgamide precursor peptide, composed of leader peptide (LP) and core peptide, is modified extensively. The order of PTMs is not fully understood. c) Based on gene deletion studies, hydroxylation of hdmHis is carried out by *thoJ*.¹⁶

use *Streptomyces malaysiense* MUSC 136 57 nomenclature).^{10, 15} The first step was reported to be the iterative addition of four thioamides to the precursor peptide, ThoA, likely by the concerted action of ThoH and ThoI. Subsequently, the dehydration of Thr8 and Ser1 is catalyzed by ThoC, ThoD and possibly by ThoE.

The intermediate is then acted upon by ThoF that carries out the oxidative decarboxylation of the terminal cysteine – a pre-requisite for AviMeCys formation. Due to the absence of an obvious cyclase in the BGC, it has been proposed that macrocyclization may be non-enzymatic; however, cyclization is facilitated by ThoE.¹⁵ Surprisingly, co-expression studies using proteins from the thiosparsoamide and thioviridamide BGCs are not in agreement with the metabolomic studies. ThoF homologues from the thioviridamide (TvaF) and thiosparsoamide (TsaF) pathways have been shown to act on unmodified precursor peptide in *E. coli* BL21 (DE3) leading the authors to propose decarboxylation to be the first modification installed during biosynthesis (Figure S1b and S2a).^{11, 13} ThoC and ThoD homologues present in the thiosparsoamide BGC (SpaC and SpaD) were also reported to be inactive, and instead dehydration was shown to be carried out by a lanthipeptide synthetase (SpaK/C) encoded outside the BGC.¹¹ This synthetase can also form a complex with the decarboxylase to assist in AviMeCys formation.¹¹ Somewhat similarly for Tva-YJ-2, the formation of a four component complex involving the homologues of ThoC, ThoD, ThoE and ThoF has been proposed to carry out the dehydration followed by AviCys formation (Figure S2b).¹⁴ The dehydrated and cyclized precursor peptide is then acted upon by a protease responsible for the removal of the leader peptide – it is not clear if the putative protease, ThoK, is responsible for this step. The last steps of the biosynthesis involve histidine bis-*N*-methylation followed by β -hydroxylation, modifications catalyzed by ThoJ and ThoG (Figure 1a), respectively.^{15, 16} In short, a basic understanding of thioholgamides and related molecules biosynthesis exists, but our understanding lacks key details, such as the order of the biosynthetic steps, and mechanistic insights for most of the enzymes involved. To address these issues, we set out to reconstitute the entire pathway *in vitro*, which enabled us to functionally characterize every post-translational modification in thioholgamide biosynthesis.

Results and Discussion

Understanding ThoF catalyzed oxidative decarboxylation. To investigate the function of the decarboxylase in thioholgamide biosynthesis, we attempted to express ThoF, but it could not be expressed in a soluble form. We instead expressed and purified its close homologue MusF from *Streptomyces sp.* MUSC 14 (96.5% sequence identity to ThoF with 100% coverage). This protein originates from a cluster that is homologous and highly similar to those reported to produce thioholgamide-like molecules (Figures S1, S2 and S4). The protein purified as a yellow-colored, soluble protein, indicating the presence of flavin as a co-factor. MusF was heat-denatured, and the flavin was identified to be flavin mononucleotide (FMN) (Figure S5). Unexpectedly, upon incubation of MusF with the precursor peptide, ThoA (1), we did not observe decarboxylation of Cys13 (Figures 2a and S6a). Our results are at odds with thioviridamide

and thiosparsoamide *in vivo* studies.^{11, 13, 14} Although we thought it highly unlikely that the sequence of biosynthetic steps was strain specific, we decided to co-express the precursor peptide, ThoA^{co-express} (2) and ThoF in *E. coli* BL21 (DE3). High-performance liquid chromatography electrospray ionization mass spectrometry (HPLC-ESI-MS) analysis led to the identification of at least three modified derivatives of 2 that displayed similar mass losses reported earlier for TvaF/TvaA co-expression studies.¹³ Compared to 2, product 3a displayed a mass loss of 43.9380 Da, corresponding to oxidative decarboxylation of the terminal cysteine followed by reduction. The generated ene-thiol is unstable and underwent spontaneous hydrolysis to yield 3b (- 59.9252 Da) and 3c (- 104.0022 Da), which were further characterized by MS² (Figures S6 - 9). Degradation of the thio-enol product from peptide decarboxylation is not unprecedented and has been reported for other AviCys-containing RiPPs.^{17, 18} However, unlike previous thioamitides co-expression studies we observed a significant proportion of unmodified precursor peptide. These findings, coupled with observed *in vitro* inactivity implied that oxidative decarboxylation is not the first tailoring step.

***In vitro* reconstitution of thioamidation and Thr dehydration.** To resolve this discrepancy and determine the initial biosynthetic step, we decided to test 1 systematically with all proteins. To successfully implement this strategy, we had to overcome one major bottleneck: soluble expression of proteins. Unlike other proteins of the pathway attempts to produce ThoH, ThoC and ThoD as individual proteins through heterologous expression in *E. coli* yielded inclusion bodies. However, on close inspection we observed that ThoH/ThoI and ThoC/ThoD had overlapping stop and start codons, hinting that these protein pairs are translationally coupled. Therefore, we decided to co-purify these pairs, which yielded soluble proteins (Figure S3). This result has implications for thioamide biosynthesis: both, ThoH/ThoI and ThoC/ThoD, form stable, active complexes (from now referred to as ThoH/I and ThoC/D).

Prior studies have demonstrated that YcaO and TfuA domain proteins are required for thioamidation in archaea and bacteria.¹⁹⁻²⁴ We therefore predicted that ThoH (YcaO protein) and ThoI (TfuA protein) would carry out the thioamidation step. Only when 1 was treated with ThoH/I in the presence of ATP and sodium sulfide, a new peak with a mass of + 63.9720 Da was observed, which was consistent with the replacement of four oxygens by sulfur (4; Figures 2 and S10a). MS² analysis localized the thioamidations at the expected sites: Val2, Met3, Ala5 and Ala6 (Figure S10b). An in-frame deletion of *thoH* also resulted in complete abolishment of thioholgamide production in our heterologous host (Figure S11). These findings are also in agreement with metabolomic studies,¹⁵ thereby leading us to conclude that thioamidation is the first biosynthetic step *in vitro* and acts as gatekeeper for downstream PTMs. Recently, it has been shown that the TfuA found in methanogens carries out hydrolysis of thiocarboxylated ThiS to provide a sulfur donor for its cognate YcaO protein.²² Since ThiS-like proteins are not present within the thioholgamide BGC, or the 1kb flanking regions used for heterologous expression, we searched for ThiS homologues in the genome of *S. lividans* delY8A. Only one homologue (WP_003976798.1) was found, which

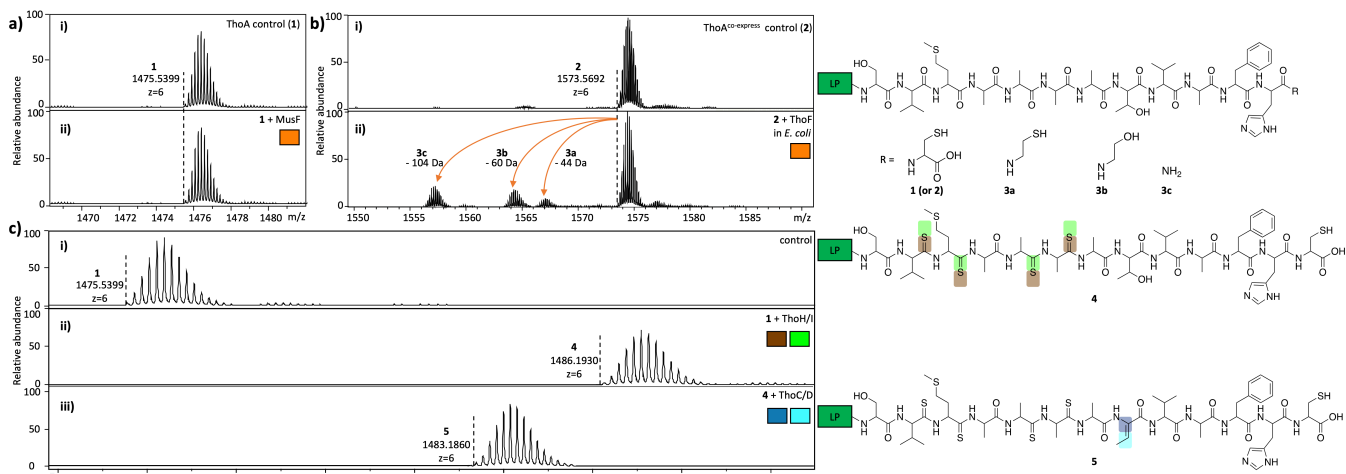


Figure 2. Enzymatic modification of **1**, **2** and **4** by MusF, ThoF, ThoH/I and ThoC/D. a) HPLC-ESI-MS analysis of **1** after incubation with MusF. b) Analysis of **2** co-expressed with ThoF in *E. coli*. c) Analysis of **1** and **4** after incubation with ThoH/I and ThoC/D. MS² analyses of **3b**, **3c**, **4** and **5** can be found in Figures S8, S9, S10 and S17.

is likely to fulfil the role of sulfur donor in a manner similar to the proposed thioamidation in methanobacteria.

Interestingly, thioamidation occurs at specific positions of the core peptide in all known thioamitides (Figures S1 and S2). In an effort to shed some light on the factors that determine this specificity we tested a small panel of altered precursor peptides (Figure S12). Truncation of the core peptide completely abolished ThoH/I activity (Figures S12 - S14). In contrast, we observed up to four thioamidations for ThoA^{A(6-11)} (Figures S12 and S15). These data suggest that the length, not the sequence, of the core peptide dictates the position of thioamidation. Several leader peptide-dependent reactions in RiPP biosynthesis have been shown to display directionality: heterocyclization in microcin B17 and dehydration/cyclization reactions in nisin take place from the N-terminus of the core peptide to the C-terminus.^{25,26} In contrast, for plantazolicin and cyanobactins, predominantly C-to-N terminal processing has been reported.^{27,28} To see which processing mode thioamidation displays, we monitored the reaction over a time course. MS² analysis revealed that thioamidation occurred predominantly in a C-to N-terminal fashion (Figure S16).

When fully thioamidated ThoA (**4**) was incubated with ThoC/D and ATP/Mg²⁺, we observed a loss of 17.9862 Da, consistent with dehydration (Figures 2c and S17a). The presence of 2, 3-dehydrobutyryne (Dhb) at position 8 was confirmed by MS² (Figure S17b). This places dehydration of Thr8 as the second modification of the pathway. Intriguingly, our results contradict the thiosparsoamide co-expression studies: ThoC/D homologues (SpaC/D vs. ThoC/D sequence identity is 76.8% and 84.1%, respectively, with 100% coverage; Figure S18) were reported to be inactive as co-expression with the precursor peptide yielded no detectable modification.¹¹ We think that the observed inactivity is most likely due to poor conversion of non-thioamidated precursor peptide by SpaC/D. Considering that co-expression studies with the ThoC/D homologues from TVA-YJ-2 (TvaC/D) exhibited dehydratase activity on unmodified precursor peptide, the involvement of additional factors cannot be excluded for the observed SpaC/D inactivity.⁴ In short, our findings confirm that thioamitide clusters are fully capable of carrying out dehydration reactions, without the need for enzymes located outside

the BGC as postulated for thiosparsoamide. This places SpaK/C as a dehydratase with relaxed substrate specificity.¹¹

ThoC contains an aminoglycoside phosphotransferase (APH)-like domain (pfam01636). In addition, a Brenner motif (HXDX₄N)²⁹, found in many enzymes catalyzing phosphoryl transfer reactions, is located in ThoC and its homologues (Figure S19). ThoD belongs to the HopA1 effector protein family (pfam17914). To gain further insights into the possible role of ThoD, we generated a 3D model using AlphaFold (Figure S20).³⁰ This model was subsequently used to find structural homologues via the DALI server.³¹ The ThoD model shares weak structural homology to effector proteins: HopA1 (pdb: 4rsx), OspF (pdb: 3iou) and VirA (pdb: 3bo6). These proteins act as phospholyases that catalyze the elimination of phosphate from either phosphothreonine or phosphoserine (though less efficiently) to generate 2,3 dehydrobutyryne (Dhb) and 2,3 dehydralanine (Dha), respectively.³² The catalytic Lys residue that acts as a general base to abstract hydrogen was found to be spatially conserved in ThoD (Ly208; Figure S20a).^{32,33} If lyase activity proceeds via an E1cB mechanism, as proposed for aforementioned effector proteins,³⁴ Tyr225 and Arg176 (located nearby) may form an oxyanion hole to stabilize the carbanion intermediate (Figure S20a). Alternatively, Mg²⁺ may also perform this task. Given that we cannot decouple the activity of ThoC from ThoD, both reaction mechanisms are possible, but considering the overall similarity of ThoD to the effector proteins we propose the former. Mutagenesis of ThoD Lys208 (K208A) resulted in a ThoC/D complex that still phosphorylated Thr8, but did not display dehydration activity (Figures S3, S22b and S22c). We thus propose that ThoC is responsible for phosphorylation, while ThoD acts as lyase to catalyse the elimination of the ThoC-installed phosphate group to yield the dehydrated product. However, a potential involvement of ThoD in facilitating precursor peptide binding and/or modulation of its cognate partner protein, ThoC, cannot be ruled out. Very recently, three new Class V lanthipeptides that feature lanthionine and AviMeCys motifs – cacaoidin,³⁵ lexapeptide³⁶ and pristinin³⁷ – were reported. The respective BGCs encode homologues

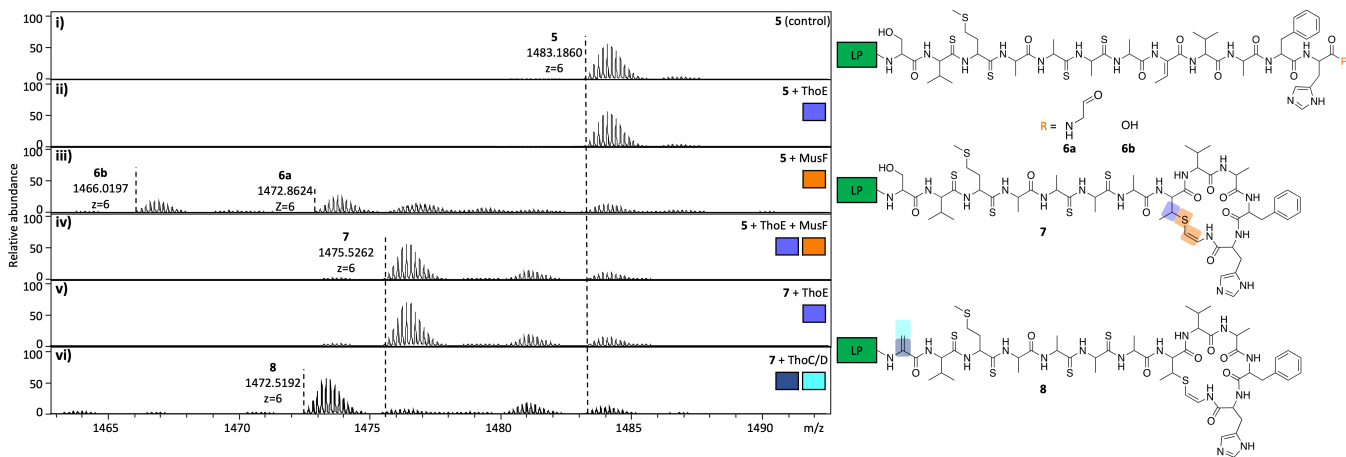


Figure 3. Enzymatic reconstitution of AviMeCys and Ser1 dehydration. (i) HPLC-ESI-MS analysis of **5**. Mass spectrum of **5** after incubation with ii) ThoE; (iii) MusF and (iv) ThoE and MusF (without ATP). Mass spectrum of **7** after incubation with (v) ThoE and ATP and (vi) ThoC/D and ATP. The predicted structures are shown on the right, and the corresponding MS² analyses of **6a**, **6b**, **7** and **8** can be found in Figures S21, S22 and S24.

of ThoC and ThoD, and we propose that these proteins may also form complexes to carry out the dehydration step.

In vitro reconstitution of AviMeCys formation and Ser dehydration. ThoE shares weak homology with the APH-like phosphotransferase domain (pfam01636); therefore, it may perform a function similar to ThoC and carry out the dehydration of Ser1 to Dha. The incubation of **5** with ThoE and ATP/Mg²⁺ did not lead to a detectable mass shift (Figure 3), but we thought that the absence of the lyase, ThoD, may be the cause. Subsequent incubation of **5** in the presence of both, ThoE and ThoC/D, in the presence of ATP/Mg²⁺ also did not yield doubly dehydrated product (Figure 3), which suggested that either ThoE may play a different role in the biosynthesis of thioholgamide or that dehydration of Ser1 followed Cys13 decarboxylation. To rule out the latter, we tried to capture the decarboxylated intermediate by adding MusF to **5**. As observed for co-expression studies with ThoF, the addition of MusF to **5** resulted in a series of ThoA-related peptides that differed at the C-terminus (Figures 3 and S21a). The two main species (**6a** and **6b**) were further characterized by MS² (Figure 21a and S21b). The addition of ThoE to the reaction significantly simplified the product profile, leaving **7** as the only major product (Figures 3 and S22). The identity of **7** was confirmed by MS² (Figure S22b). Next, we incubated **7** with ThoE in the presence of ATP/Mg²⁺ to see if Ser dehydration is macrocycle-dependent. However, even after extended incubation times we did not observe any change (Figures 3). Therefore, it appears that ThoE is an inactive phosphorylase with an overall fold of APH-like phosphotransferase that assists in macrocyclization, possibly by interacting with MusF and/or the precursor peptide. As a result, the unstable thioenol intermediate is protected from hydrolysis long enough for Michael-type cyclization to take place, which yields AviMeCys. It is not clear if ThoE acts as a cyclase, or if cyclization is non-enzymatic. The proposed role of ThoE in assisting AviMeCys formation is further supported by binding studies showing weak interaction of MusF with ThoE - measured by microscale thermophoresis (K_d : 32 ± 22 μM; Figure S23). The results of recently reported *in vivo* and co-expression studies on ThoE homologues (TsaE and TvaE) support the role we propose. For thiostreptamide, gene deletion of *tsaE*

abolished thiostreptamide production in the heterologous host and no macrocyclized metabolite was produced.¹⁵ Gene deletion of *tvaE* resulted in a significant decrease in TVA-YJ-2 production, which could largely be compensated by enhancing transcriptional level using a strong constitutive promoter.¹⁴ In this particular case, AviMeCys formation may proceed rapidly enough to allow its formation in the absence of TvaE, albeit at much lower yields. Co-expression studies and substrate stabilization studies clearly point towards the involvement of TvaE in assisting AviMeCys formation.¹⁴ The absence of reported multiple diastereoisomers of thioamitides^{7, 9, 10, 12} also suggest stereochemical control during AviMeCys formation. As both, MusF and ThoE, play a role in AviMeCys formation, it is likely that one of these proteins also controls the stereochemical outcome during cyclization. Structural studies are underway to help explain how ThoE or ThoF exert their function(s).

Our data left us to wonder which enzyme catalyzed the dehydration of Ser1: was a native Ser dehydratase from the producer strain responsible, or is ThoC/D remarkably promiscuous and responsible for this dehydration? To test the possible role of ThoC/D in the formation of Dha1, we incubated it with **7** (Figures 3 and S24a). To our pleasant surprise we observed a loss of 18.033 Da, confirming that ThoC/D is also responsible for the formation of Dha1 (**8**; Figure S24). This finding also explains why previous co-expression studies of precursor peptide with ThoC/D homologues (TvaC/D) resulted in the formation of products containing either Dhb or Dha/Dhb, with double-dehydrated species being a minor product¹⁴ - Ser dehydration preferably occurs after the formation of the macrocycle. The observed interplay led us to envision a scenario that involves the interaction of ThoF with ThoC/D and ThoE for efficient and coordinated dehydration and AviMeCys formation - reminiscent of Class II, III and IV lanthipeptide biosynthesis that rely on multidomain proteins to install lanthionine (Lan) or methyl-lanthionine ((Me)-Lan) motifs.^{3, 38} Even though ThoC/D is fully functional without the other two proteins, we wondered if it can also interact with MusF. Remarkably, MST analysis revealed ThoC/D to interact with MusF tightly (K_d : 180 ± 56 nM; Figure S25). These findings, in conjunction with the observed MusF/ThoF interaction

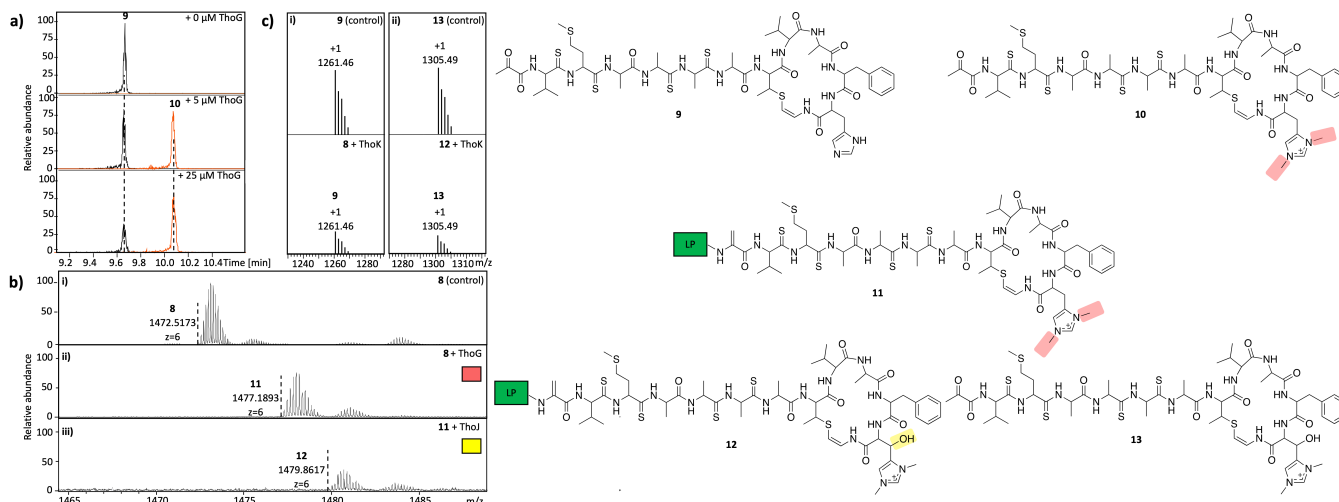


Figure 4. *In vitro* reconstitution of ThoG, ThoJ and ThoK activities. a) Incubation of **9** with ThoG in the presence of SAM yielded **10**. Increasing the concentration of ThoG beyond 25 μM did not enhance the yield of **10**. Extracted ion chromatograms (EICs) of **9** and **10** are shown. b) HPLC-ESI-MS analysis of **8** (i); mass spectrum of **8** after incubation with ThoG and SAM (ii); mass spectrum of **10** after reaction with ThoJ in the presence of Fe^{2+} and α -ketoglutarate (iii). c) Processing of **8** and **12** by ThoK leads to cleavage of the leader peptide to yield **9** and **13**. MS² analyses of **9**, **10**, **11**, **12** and **13** can be found in Figures S26, S28, S31 and S32.

suggest presence of a supramolecular enzyme complex for dehydration (x2) and AviMeCys formation with the decarboxylase acting as an anchoring protein in thioholgamide biosynthesis. Similarly, co-expression studies for lexapeptide biosynthesis also suggest a requirement of supramolecular assembly (LxmK/X/Y/D; analogous to ThoC/D/E/F) for Dha, Dhb, Lan and AviMeCys formation.³⁶ Nevertheless our attempts to co-purify ThoC/D/E/F were unsuccessful, likely due to the formation of a low affinity complex (data not shown).

***In vitro* reconstitution of histidine bis-*N*-methylation, β -hydroxylation and leader peptide removal.** With the biosynthesis of the core scaffold established, we focused on the last three enzymes of the pathway, namely ThoG^S, ThoJ and ThoK. Previous *in vivo* work in our group identified ThoJ as the enzyme responsible for histidine hydroxylation.¹⁶ Subsequently, the deletion of ThoG led to the production of a thioholgamide derivative (**9**; Figure S26), which lacks both histidine modifications but is otherwise fully mature. These results suggested that methylation precedes hydroxylation. To ascertain if ThoG^S required the presence of leader peptide for activity, we incubated it with **8** and **9**. Surprisingly, histidine bis-*N*-methylation was observed for neither substrate. On closer inspection, we found multiple in-frame start codons located upstream of ThoG^S, suggesting that the annotation of *thoG* may be incomplete (Figure S27). When we expressed and purified the longest version, ThoG (Figures S3 and S27), its incubation with **8** or **9** led to the bis-*N*-methylation of both substrates; however, only the reaction with **8** went to completion, which suggest that the leader peptide is important for efficient methylation, but not essential (**10** and **11**; Figures 4a, 4b and S28). Genetic complementation experiments also confirm the requirement of these additional residues for ThoG activity (Figure S29). Based on sequence and 3D model analysis, the glycine rich motif found in ThoG (DXGXGXG)³⁹, a hallmark for SAM binding,

is not a major factor determining substrate specificity. This would also explain why both, **9** and **13**, were observed upon treatment with ThoK, i.e. methylation and hydroxylation are not a prerequisite for leader peptide removal. In addition, these studies also support that

along with the putative active site residues were found to be present in a deep pocket with roughly the first 80 amino acid residues forming a lid on top of the putative active site pocket (Figures S27 and S30). Hence, it is likely that the lid is involved in substrate binding and/or occluding the active site from bulk solvent to aid catalysis.

To ascertain if ThoJ, a non-heme monooxygenase, was responsible for β -hydroxylation we incubated it with **11**. As expected, a mass shift of + 16.0344 Da was observed (**12**; Figures 4b and S31), consistent with hydroxylation. Previously, we had reported a lack of ThoJ activity on *S.lividan*:pTho-delThoJ derivative **10**.¹⁶ This implies that the activity of ThoJ is both, leader- and substrate-specific.

This left removal of the leader peptide as the last unresolved step in thioholgamide biosynthesis. ThoK was predicted to belong to the C1A cysteine protease family, and is conserved amongst thioholgamide-like BGCs. Although studies on thiostreptamide biosynthesis indicate that *tsaK* (*thoK* homologue) is not essential for heterologous production of fully mature compound.¹⁵ This suggests that either *thoK* is inactive or a native protease is responsible for the removal of the leader peptide. To rule out the former we incubated **8** and **12** with ThoK. After extended incubation, we detected a small amount of mass corresponding to thioholgamide (**9** and **13**; Figures 4c, S26 and S32). The observed poor conversion could be due to inefficient proteolytic conditions and/or need for zymogen-activation – a phenomenon well studied for cathepsins that also belong to the family of cysteine proteases.⁴⁰ The role of C1A cysteine proteases in RiPP maturation is not unprecedented and they have been shown to catalyse the removal of the leader peptide in polytheonamide biosynthesis.⁴¹

Typically, the substrate specificity of C1A cysteine proteases is based on the interaction of the protease with seven residues surrounding the cleavage site of the substrate (P4 - P3')^{42,43}, so it is likely that the thioholgamide macrocycle

the pyruvyl group originates from a dehydrated Ser spontaneously after removal of the leader peptide instead of a pyruvyl transferase. This is supported by *in vivo* studies of thiostreptamide S4.¹²

The only uncharacterized protein left in the thioholgamide BGC is ThoL, which contains no clear catalytic domain when analyzed with Pfam. Although, *thoL*-like genes are conserved in almost all thioviridamide-like BGCs (Figure S1 and S2), our data suggests no apparent biosynthetic role for this protein. Similarly, a gene knockout of the *thoL* homologue *tsaL* was reported to yield fully processed thio-streptamide S4.¹⁵ Based on protein sequence analysis using HMMER⁴⁴ (v. 3.3.2) and TMHMM⁴⁵ (v. 2.0) ThoL is predicted to contain a transmembrane domain, so it may be involved in transport. However, further studies are needed to confirm this role.

Conclusions

The data presented here provide the first complete *in vitro* reconstitution of a thioholgamide(-like) pathway and resolve the apparent disparities between different *in vivo* studies. In addition, the dissection of each PTM lead us to propose a revised biosynthetic scheme (Figure 5): Thioamidation, carried out by the ThoH/I complex, is the first PTM installed in ThoA, thereby acting as the first gate-keeper of the pathway. The activity of ThoH and ThoI homologues from methanogen and thiopeptides have already been reconstituted,^{19,22,24} but this is the first reported case of a successful *in vitro* reconstitution of iterative YcaO/TfuA catalyzed thioamidation. The second PTM is the dehydration of Thr8 to Dhb, catalyzed by the ThoC/D complex. We propose that ThoC is responsible for phosphorylation of the Thr side-chain, followed by the elimination of the ThoC-installed phosphate group by ThoD to yield the dehydroamino acid. This is followed by the formation of AviMeCys, which first involves terminal Cys decarboxylation to generate a reactive thioenolate. The unspecific hydrolysis products of the thioenolate are avoided by addition of ThoE, hinting at the deep-seated cooperation between ThoF and ThoE for efficient macrocyclization. After cyclization, the transformation of Ser1 to Dha is also carried out by ThoC/D. Given the requirement of ThoE for efficient AviMeCys formation and the observed interaction of ThoC/D with MusF, it is likely that dehydration (x2) and macrocyclization are carried out by a ThoC/D/E/F supra-complex. Such a complex would ensure coordination, thereby preventing multiple enzymes from competing for the same substrate, which would result in unwanted side products that could ultimately lead to shunt products.

The next step is histidine bis-*N*-methylation, which is carried out by ThoG. To the best of our knowledge this modification is not present in any other natural product and provides a positive charge that may be important for biological activity. Gene deletion and *in vitro* studies indicate that methylation is a prerequisite for β -hydroxylation, a modification that has been reported to be crucial for the activity of two other RiPPs, duramycin and cinnamycin.^{46, 47} However, it remains to be seen if the hydroxyl group has any effect of

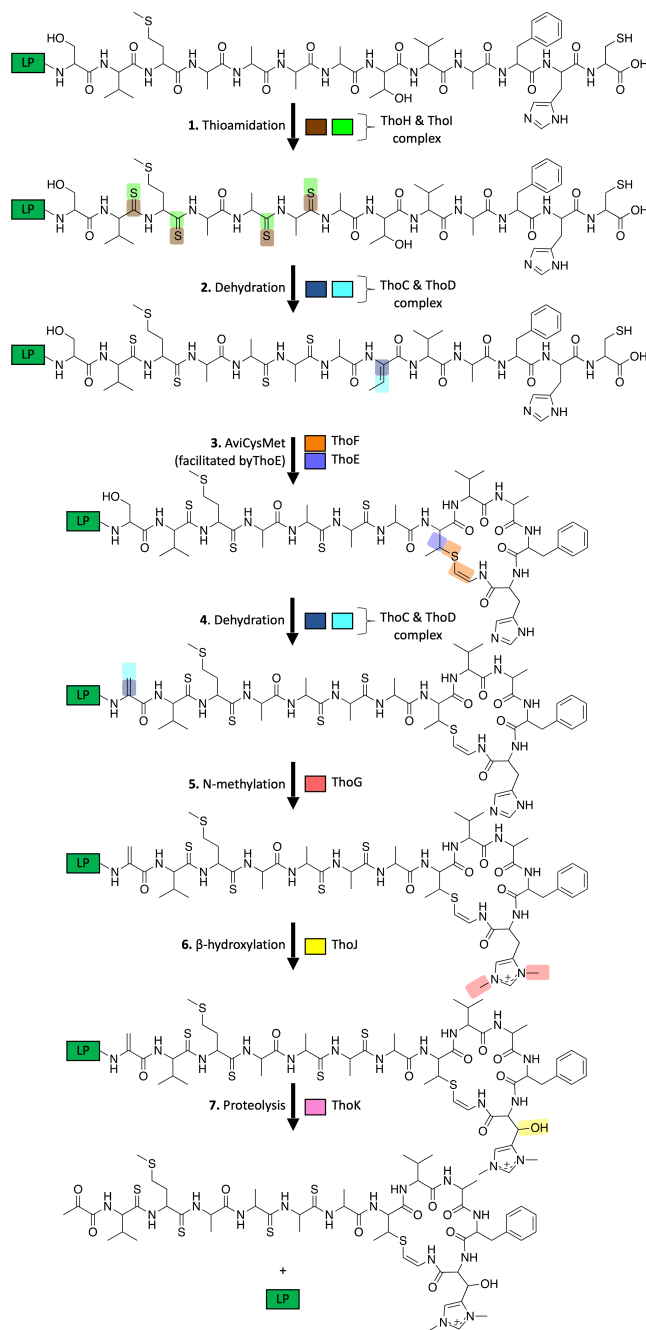


Figure 5. Proposed pathway to thioholgamides.

thioholgamide activity. The final step involves the removal of the leader peptide by ThoK to yield fully mature thioholgamide. Further functional and structural studies are underway to help explain: (1) the positional specificity of thioamidation and (2) why certain ThoH and ThoI homologues can install different numbers of thioamide moieties (e.g. thioviridamide⁷ and thioalbamide⁹) (3) the exact role of ThoE in macrocyclization. Additional remaining questions include delineating the substrate specificity of each enzyme of the pathway to help future bioengineering applications. This includes the generation of combinatorial libraries to guide the development of thioholgamides' antiproliferation activity towards cancer cells.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental methods and supporting figures (file type, i.e., PDF)

AUTHOR INFORMATION

Corresponding Authors

rolf.mueller@helmholtz-hips.de
jesko.koehnke@glasgow.ac.uk

Present Addresses

& School of Chemistry, University of Glasgow, Glasgow G12 8QQ, U.K.

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript

Funding Sources

J.K. thanks the BBSRC for support (BB/V016059/1). R.M. would like to acknowledge DFG (Leibniz Award: MU 1254/32-1).

ACKNOWLEDGMENT

We appreciate the assistance provided by Alexander Voltz and Jake Haeckl for LC-MS, and Nestor Zaburanyi for bioinformatic support. We would also like to acknowledge Prof. Olga Kalinina for help with AlphaFold.

REFERENCES

1. Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* **2013**, *30* (1), 108-60.
2. Hudson, G. A.; Mitchell, D. A., RiPP antibiotics: biosynthesis and engineering potential. *Curr Opin Microbiol* **2018**, *45*, 61-69.
3. Montalban-Lopez, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martinez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H. G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Sussmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J. K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A., New developments in RiPP discovery, enzymology and engineering. *Nat Prod Rep* **2021**, *38* (1), 130-239.
4. Dahlem, C.; Siow, W. X.; Lopatniuk, M.; Tse, W. K. F.; Kessler, S. M.; Kirsch, S. H.; Hopstadter, J.; Vollmar, A. M.; Muller, R.; Luzhetskyy, A.; Bartel, K.; Kiemer, A. K., Thioholgamide A, a New Anti-Proliferative Anti-Tumor Agent, Modulates Macrophage Polarization and Metabolism. *Cancers (Basel)* **2020**, *12* (5).
5. Takase, S.; Kurokawa, R.; Kondoh, Y.; Honda, K.; Suzuki, T.; Kawahara, T.; Ikeda, H.; Dohmae, N.; Osada, H.; Shin-Ya, K.; Kushiro,

T.; Yoshida, M.; Matsumoto, K., Mechanism of Action of Prethioviridamide, an Anticancer Ribosomally Synthesized and Post-Translationally Modified Peptide with a Polythioamide Structure. *ACS Chem Biol* **2019**, *14* (8), 1819-1828.

6. Hayakawa, Y.; Sasaki, K.; Adachi, H.; Furihata, K.; Nagai, K.; Shin-ya, K., Thioviridamide, a novel apoptosis inducer in transformed cells from *Streptomyces olivoviridis*. *J Antibiot (Tokyo)* **2006**, *59* (1), 1-5.

7. Hayakawa, Y.; Sasaki, K.; Nagai, K.; Shin-ya, K.; Furihata, K., Structure of thioviridamide, a novel apoptosis inducer from *Streptomyces olivoviridis*. *J Antibiot (Tokyo)* **2006**, *59* (1), 6-10.

8. Izawa, M.; Kawasaki, T.; Hayakawa, Y., Cloning and heterologous expression of the thioviridamide biosynthesis gene cluster from *Streptomyces olivoviridis*. *Appl Environ Microbiol* **2013**, *79* (22), 7110-3.

9. Frattaruolo, L.; Lacret, R.; Cappello, A. R.; Truman, A. W., A Genomics-Based Approach Identifies a Thioviridamide-Like Compound with Selective Anticancer Activity. *ACS Chem Biol* **2017**, *12* (11), 2815-2822.

10. Kjaerulff, L.; Sikandar, A.; Zaburanyi, N.; Adam, S.; Herrmann, J.; Koehnke, J.; Muller, R., Thioholgamides: Thioamide-Containing Cytotoxic RiPP Natural Products. *ACS Chem Biol* **2017**, *12* (11), 2837-2841.

11. Lu, J.; Wu, Y.; Li, Y.; Wang, H., The Utilization of Lanthipeptide Synthetases Is a General Strategy for the Biosynthesis of 2-Aminovinyl-Cysteine Motifs in Thioamitides*. *Angew Chem Int Ed Engl* **2021**, *60* (4), 1951-1958.

12. Li, Y.; Liu, J.; Tang, H.; Qiu, Y.; Chen, D.; Liu, W., Discovery of New Thioviridamide-Like Compounds with Antitumor Activities. *Chinese Journal of Chemistry* **2019**, *37* (10), 1015-1020.

13. Lu, J.; Li, J.; Wu, Y.; Fang, X.; Zhu, J.; Wang, H., Characterization of the FMN-Dependent Cysteine Decarboxylase from Thioviridamide Biosynthesis. *Org Lett* **2019**, *21* (12), 4676-4679.

14. Qiu, Y.; Liu, J.; Li, Y.; Xue, Y.; Liu, W., Formation of an aminovinyl-cysteine residue in thioviridamides occurs through a path independent of known lanthionine synthetase activity. *Cell Chem Biol* **2021**, *28* (5), 675-685 e5.

15. Eyles, T. H.; Vior, N. M.; Lacret, R.; Truman, A. W., Understanding thioamidite biosynthesis using pathway engineering and untargeted metabolomics. *Chem Sci* **2021**, *12* (20), 7138-7150.

16. Sikandar, A.; Lopatniuk, M.; Luzhetskyy, A.; Koehnke, J., Non-Heme Monooxygenase Thoj Catalyzes Thioholgamide beta-Hydroxylation. *ACS Chem Biol* **2020**, *15* (10), 2815-2819.

17. Wiebach, V.; Mainz, A.; Siegert, M. J.; Jungmann, N. A.; Lesquame, G.; Tirat, S.; Dreux-Zigha, A.; Aszodi, J.; Le Beller, D.; Sussmuth, R. D., The anti-staphylococcal lipolanthines are ribosomally synthesized lipopeptides. *Nat Chem Biol* **2018**, *14* (7), 652-654.

18. Kupke, T.; Kempter, C.; Jung, G.; Gotz, F., Oxidative decarboxylation of peptides catalyzed by flavoprotein EpiD. Determination of substrate specificity using peptide libraries and neutral loss mass spectrometry. *J Biol Chem* **1995**, *270* (19), 11282-9.

19. Nayak, D. D.; Mahanta, N.; Mitchell, D. A.; Metcalf, W. W., Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea. *Elife* **2017**, *6*.

20. Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A., Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proc Natl Acad Sci U S A* **2018**, *115* (12), 3030-3035.

21. Liu, J.; Lin, Z.; Li, Y.; Zheng, Q.; Chen, D.; Liu, W., Insights into the thioamidation of thiopeptides to enhance the understanding of the biosynthetic logic of thioamide-containing thiopeptides. *Org Biomol Chem* **2019**, *17* (15), 3727-3731.

22. Liu, A.; Si, Y.; Dong, S. H.; Mahanta, N.; Penkala, H. N.; Nair, S. K.; Mitchell, D. A., Functional elucidation of TfuA in peptide backbone thioamidation. *Nat Chem Biol* **2021**, *17* (5), 585-592.

23. Mahanta, N.; Szantai-Kis, D. M.; Petersson, E. J.; Mitchell, D. A., Biosynthesis and Chemical Applications of Thioamides. *ACS Chem Biol* **2019**, *14* (2), 142-163.

24. Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A., Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. *J Am Chem Soc* **2018**, *140* (30), 9494-9501.

25. Kelleher, N. L.; Hendrickson, C. L.; Walsh, C. T., Posttranslational heterocyclization of cysteine and serine residues in the antibiotic microcin B17: distributivity and directionality. *Biochemistry* **1999**, *38* (47), 15623-30.
26. Lubelski, J.; Khusainov, R.; Kuipers, O. P., Directionality and coordination of dehydration and ring formation during biosynthesis of the lanthibiotic nisin. *J Biol Chem* **2009**, *284* (38), 25962-72.
27. Melby, J. O.; Dunbar, K. L.; Trinh, N. Q.; Mitchell, D. A., Selectivity, directionality, and promiscuity in peptide processing from a *Bacillus* sp. Al Hakam cyclodehydratase. *J Am Chem Soc* **2012**, *134* (11), 5309-16.
28. Koehnke, J.; Bent, A. F.; Zollman, D.; Smith, K.; Houssen, W. E.; Zhu, X.; Mann, G.; Lebl, T.; Scharff, R.; Shirran, S.; Botting, C. H.; Jaspars, M.; Schwarz-Linek, U.; Naismith, J. H., The cyanobactin heterocyclase enzyme: a processive adenylase that operates with a defined order of reaction. *Angew Chem Int Ed Engl* **2013**, *52* (52), 13991-6.
29. Brenner, S., Phosphotransferase sequence homology. *Nature* **329** (21).
30. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D., Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-589.
31. Holm, L., Using Dali for Protein Structure Comparison. *Methods Mol Biol* **2020**, *2112*, 29-42.
32. Chambers, K. A.; Abularrage, N. S.; Scheck, R. A., Selectivity within a Family of Bacterial Phosphothreonine Lyases. *Biochemistry* **2018**, *57* (26), 3790-3796.
33. Chen, L.; Wang, H.; Zhang, J.; Gu, L.; Huang, N.; Zhou, J. M.; Chai, J., Structural basis for the catalytic mechanism of phosphothreonine lyase. *Nat Struct Mol Biol* **2008**, *15* (1), 101-2.
34. Ke, Z.; Smith, G. K.; Zhang, Y.; Guo, H., Molecular mechanism for eliminylation, a newly discovered post-translational modification. *J Am Chem Soc* **2011**, *133* (29), 11103-5.
35. Ortiz-Lopez, F. J.; Carretero-Molina, D.; Sanchez-Hidalgo, M.; Martin, J.; Gonzalez, I.; Roman-Hurtado, F.; de la Cruz, M.; Garcia-Fernandez, S.; Reyes, F.; Deisinger, J. P.; Muller, A.; Schneider, T.; Genilloud, O.; Cacaoidin, First Member of the New Lanthidin RiPP Family. *Angew Chem Int Ed Engl* **2020**, *59* (31), 12654-12658.
36. Xu, M.; Zhang, F.; Cheng, Z.; Bashiri, G.; Wang, J.; Hong, J.; Wang, Y.; Xu, L.; Chen, X.; Huang, S. X.; Lin, S.; Deng, Z.; Tao, M., Functional Genome Mining Reveals a Class V Lanthipeptide Containing a d-Amino Acid Introduced by an F420 H₂ -Dependent Reductase. *Angew Chem Int Ed Engl* **2020**, *59* (41), 18029-18035.
37. Kloosterman, A. M.; Cimermancic, P.; Elsayed, S. S.; Du, C.; Hadjithomas, M.; Donia, M. S.; Fischbach, M. A.; van Wezel, G. P.; Medema, M. H., Expansion of RiPP biosynthetic space through integration of pan-genomics and machine learning uncovers a novel class of lanthipeptides. *PLoS Biol* **2020**, *18* (12), e3001026.
38. Hegemann, J. D.; Sussmuth, R. D., Matters of class: coming of age of class III and IV lanthipeptides. *RSC Chem Biol* **2020**, *1* (3), 110-127.
39. Kozbial, P. Z.; Mushegian, A. R., Natural history of S-adenosylmethionine-binding proteins. *BMC Struct Biol* **2005**, *5*, 19.
40. Verma, S.; Dixit, R.; Pandey, K. C., Cysteine Proteases: Modes of Activation and Future Prospects as Pharmacological Targets. *Front Pharmacol* **2016**, *7*, 107.
41. Helf, M. J.; Freeman, M. F.; Piel, J., Investigations into PoyH, a promiscuous protease from polytheonamide biosynthesis. *J Ind Microbiol Biotechnol* **2019**, *46* (3-4), 551-563.
42. Berger, A.; Schechter, I., Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philos Trans R Soc Lond B Biol Sci* **1970**, *257* (813), 249-64.
43. Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogyo, M.; Bromme, D.; Ellman, J. A.; Craik, C. S., Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *J Biol Chem* **2006**, *281* (18), 12824-32.
44. Eddy, S. R., Accelerated Profile HMM Searches. *PLoS Comput Biol* **2011**, *7* (10), e1002195.
45. Sonnhammer, E. L.; von Heijne, G.; Krogh, A., A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **1998**, *6*, 175-82.
46. Huo, L.; Okesli, A.; Zhao, M.; van der Donk, W. A., Insights into the Biosynthesis of Duramycin. *Appl Environ Microbiol* **2017**, *83* (3).
47. Okesli, A.; Cooper, L. E.; Fogle, E. J.; van der Donk, W. A., Nine post-translational modifications during the biosynthesis of cinnamycin. *J Am Chem Soc* **2011**, *133* (34), 13753-60.

LP S V M A A A A T V A F H C

