bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Formation of an Aminovinyl-Cysteine Residue in Thioviridamide Non-Lanthipeptides Occurs through a Path Independent of Known Lanthionine Synthetase Activity in *Streptomyces sp.* NRRL S-87

Yanping Qiu[1,#], Jingyu Liu[1,#], Yuqing Li[1,2], Yanqing Xue[1], Huan Wang[2], and Wen Liu[1,3,*]

[1] State Key Laboratory of Bioorganic and Natural Products Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China;

[2] State Key Laboratory of Coordination Chemistry, Chemistry and Biomedicine Innovation Center of Nanjing University, Jiangsu Key Laboratory of Advanced Organic Materials, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210093, China;

[3] Huzhou Center of Bio-Synthetic Innovation, 1366 Hongfeng Road, Huzhou 313000, China.

* To whom correspondence should be addressed: Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Rd., Shanghai 200032, China. Wen Liu, Email: wliu@sioc.ac.cn, Tel: 86-21-54925111, Fax: 86-21-64166128.

[#] These authors contributed to this work equally.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ABSTRACT:

2-Aminovinyl-cysteine (AviCys) is an unusual thioether amino acid shared by a variety of ribosomally synthesized and posttranslationally modified peptides (RiPPs), as part of a macrocyclic ring system that contains the *C*-terminal 4 or 6 residues of a precursor peptide. This amino acid is nonproteinogenic and arises from processing the *C*-terminal Cys residue and an internal Ser/Thr residue to form an unsaturated thioether linkage. Enzyme activities for forming lanthionine (Lan), a distinct saturated thioether residue characteristic of lanthipeptide-related RiPPs, has long been speculated to be necessary for AviCys formation. Based on investigations into the biosynthesis of thioviridamide non-lanthipeptdes in *Streptomyces sp.* NRRL S-87, we here report an alternative path for AviCys formation that is independent of known Lan synthetase activity. This path relies on four dedicated enzymes for posttranslational modifications of the precursor peptide, in which $TvaE_{S-87}$, a phosphotransferase homolog, plays a critical role. It works with LanD-like flavoprotein $TvaF_{S-87}$ to form a minimum AviCys synthetase complex that follows the combined activity of $TvaCD_{S-87}$ for Thr dehydration and catalyzes Cys oxidative decarboxylation and subsequent Michael addition of the resulting enethiol nucleophile onto the newly formed dehydrobutyrine residue for cyclization. With $TvaE_{S-87}$, $TvaF_{S-87}$ activity for Cys processing can be coordinated with $TvaCD_{S-87}$ activity for minimizing competitive or unexpected spontaneous reactions and forming AviCys effectively.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

INTRODUCTION

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a large class of natural products that arise from post-translational modifications (PTMs) for precursor peptide maturation and the achievement of their biological activities.[1] Using 20 proteinogenic amino acids as building blocks, nature creates stunningly diverse RiPPs largely because of unique enzymes for PTMs. The biosynthesis of thioviridamides (TVAs, **Figure 1A**),[2] a growing group of structurally related thioamitide RiPPs with potent antitumor activity, is such a case. TVAs were initially isolated from *Streptomyces olivoviridis* and then produced in *S. lividans* where the TVA biosynthetic gene cluster (*tva*) was expressed heterologously.[3] Recent survey of the published bacterial genome sequences revealed many *tva*-like gene clusters, and subsequent product mining in related microorganisms led to a rapid enrichment of the TVA family, in which several new members appear to be promising for cancer treatment.[4,5] In addition to a characteristic thioamide peptide backbone, TVAs contain an unusual thioether residue, 2-aminovinyl-(3-methyl)-cysteine (AviCys),[6] which is shared by a variety of RiPPs including linaridins (e.g., cypemycin[7] and cacaoidin[8]) and certain lanthipeptide-related RiPPs (e.g., epidermin[9] and microvionin lipolanthines[10]), as part of a macrocyclic ring system that contains the *C*-terminal 4 or 6 residues of a precursor peptide (**Figure 1A**).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The process through which an AviCys residue is formed remains poorly understood. Available insights have gained largely from the biosynthesis of lanthipeptide-related RiPPs, which, in fact, are characterized by a distinct thioether amino acid, (3-methyl-)lanthionine (Lan) (**Figure 1B**). The formation of a Lan residue has been well established over the past two decades.[11] It results from dehydration of a Ser or Thr residue and subsequent cyclization by nucleophilic addition of a Cys residue onto the newly formed dehydroamino acid, 2,3-didehydroalanine (Dha) or 2,3-didehydrobutyrine (Dhb), to furnish a thioether linkage in the precursor peptide (**Figure 1B**). The two reactions, i.e., dehydration and cyclization that occur in tandem, can be attributed to the dedicated activities of a LanB dehydratase and a LanC cyclase for Class I lanthipeptides or to the combined activity of a multifunctional protein, i.e., LanM, LanKC or LanL for Class II, III or IV lanthipeptides, respectively (here collectively referred to as Lan synthetase).[11] In comparison, the formation of an AviCys residue, which was proposed to share a similar biosynthetic route, relies on specifically processing the *C*-terminal Cys residue and the 2 or 4-aa-upstream, internal Ser/Thr residue of a precursor peptide (**Figure 1B**). It likely involves **1)** the dehydration of internal Ser or Thr and **2)** the oxidative decarboxylation of *C*-terminal Cys, followed by **3)** Michael addition of the resulting enethiol nucleophile onto the preceding Dha or Dhb residue to provide a distinct, unsaturated thioether linkage. In this route, processing *C*-terminal Cys to an enethiol, which is unique for AviCys formation, has been characterized to be dependent of the activity of LanD, a flavoprotein belonging to the family of homo-oligomeric, flavin-containing cysteine decarboxylases

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(HFCDs).[6]

Given mechanistic similarity in the formation of Lan and AviCys thioether residues, Lan synthetase activity for Ser/Thr dehydration and thiol addition/cyclization has long been speculated to be involved in AviCys formation.[6,12] Recently, the combination of a LanKC-type, Class III multifunctional Lan synthetase with a LanD flavoprotein proved to be necessary for the formation of an AviCys-containing avionin motif in the biosynthesis of lipolanthines,[10] supporting the relevance between these two distinct thioether residues in PTM enzymes. With great interest in how residue AviCys is formed and whether Lan synthetase activity participates in the biosynthesis of AviCys-containing non-lanthipeptides, we here investigated the biosynthetic pathway of TVA RiPPs in *S. sp.* NRRL S-87. Based on the characterization of a LanD-like, HFCD-fold flavoprotein and three previously functionally unknown, PTM-related enzymes, we demonstrated that in this *Streptomyces* strain, the formation of an AviCys residue does not require previously known Lan synthetase activity and instead occurs through a path biochemically similar to but phylogenetically different from that in the biosynthesis of lanthipeptide-related RiPPs.

RESULTS

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TVA biosynthesis in *S. sp.* NRRL S-87 does not depend on Known Lan synthetase activity. We and others have recently validated the production of TVAs in *S. sp.* NRRL S-87, where a gene cluster (i.e., $tva_{S\text{-}87}$) sharing nearly head-to-tail homology with *tva*, the prototype for TVA biosynthesis in *S. olivoviridis*, was identified (**Figures 2A** and **S2**).[2,4d] Notably, the $tva_{S\text{-}87}$ cluster does not encode any homologs of Lan synthetases, either monofunctional (i.e., LanB and LanC) or multifunctional (LanM, LanKC or LanL), consistent with the absence of residue Lan in the structure of TVAs. Whether TVA biosynthesis involves genes outside the $tva_{S\text{-}87}$ cluster for Lan formation needed to be determined. In the genome of *S. sp.* NRRL S-87, we identified two genes potentially related to Lan formation, i.e., *lanKC1* and *lanKC2*, which are $tva_{S\text{-}87}$-unclustered and encode two homologs of LanKC-type, Class III multifunctional Lan synthetases (**Table S2**). To exclude the participation of these two genes in TVA biosynthesis, we cloned and heterologously expressed the $tva_{S\text{-}87}$ cluster in *S. laurentii*, a strain producing the thiopeptide antibiotic thiostrepton (**Figure 3A**).[13] As expected, TVA-related compounds were observed, including TVA-YJ-1, the newly characterized member that shares an AviCys residue but differs from TVA-YJ-2 (previously identified in *S. sp.* NRRL S-87[4d]) by possessing a reduced lactyl unit instead of a pyruvyl unit at the *N*-terminus (**Figures 1A**, **S3** and **S25** and **Table S1**). Careful analysis of the *S. laurentii* genome revealed $lanL_{SL}$, a gene coding for a homolog of LanL-type, Class IV multifunctional Lan synthetases (**Table S2**). We then inactivated $lanL_{SL}$ in *S. laurentii* and introduced the $tva_{S\text{-}87}$ cluster into the resulting mutant strain. As a result, the inactivation of $lanL_{SL}$ had little effect on the production

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of TVAs in *S. laurentii* (**Figure 3A**), demonstrating that the $tva_{S\text{-}87}$ cluster is sufficient for TVA biosynthesis and therefore contains all necessary genes for AviCys formation.

Identification of genes coding for AviCys formation in the $tva_{S\text{-}87}$ cluster. In addition to $tvaA_{S\text{-}87}$, the gene encoding the precursor peptide $TvaA_{S\text{-}87}$ composed of a 74-aa leader peptide (LP) and the 13-aa core peptide (CP), **S_1VMAAAAT$_8$VAFHC$_{13}$** (numbering relates to residue position, as below), the $tva_{S\text{-}87}$ cluster contains 10 functional genes, i.e., $tvaCDEFGHIKL_{S\text{-}87}$, being organized into the same operon for co-expression to facilitate the coordination of their protein products in TVA biosynthesis (**Figure 2**). $TvaL_{S\text{-}87}$ is a membrane protein potentially participating in TVA transportation, while the other 9 proteins were believed to be responsible for the transformation of the CP sequence of $TvaA_{S\text{-}87}$ into mature molecules. During this PTM process, the LP sequence of $TvaA_{S\text{-}87}$ was supposed to be necessary for the engagement of related enzyme catalysts and can be removed in the late biosynthetic stage as usual, assumedly by protease $TvaK_{S\text{-}87}$.[4b] During the maturation of TVAs, the specific PTMs include **1)** thioamidation of the peptidyl backbone, for which $TvaH_{S\text{-}87}$ and $TvaI_{S\text{-}87}$ might function together to form a complex composed of a discrete YcaO protein and a TfuA-fold protein;[14] and **2)** functionalization within the AviMeCys-containing, *C*-terminal ring system, for which $TvaG_{S\text{-}87}$ and $TvaJ_{S\text{-}87}$ could work on the same residue, His12, for *N*,*N*-dimethylation and β-hydroxylation (**Figure 2B**).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Based on the above analysis, the PTM proteins for AviCys formation were narrowed to TvaF$_{S-87}$, a LanD-like, HFCD-fold flavoprotein, along with remaining TvaC$_{S-87}$, TvaD$_{S-87}$ and TvaE$_{S-87}$, all of which were functionally unassigned. The counterparts of TvaF$_{S-87}$ were known in the biosynthesis of various AviCys-containing RiPPs. Either as a LanD protein (for lanthipeptides[10-12,15]) or as a LanD-like protein (for non-lanthipeptides[6,8,16]), these flavoproteins share the oxidative decarboxylation activity necessary for AviCys formation and can process the *C*-terminal Cys residue of a precursor peptide to an enethiol before Michael addition to the upstream Dha/Dhb residue. The homologs of TvaC$_{S-87}$ and TvaE$_{S-87}$ were previously annotated as hypothetic proteins, with the exception in ref. 4b, where they were proposed to be phosphotransferases by Truman *et al*. Here, careful sequence analysis allowed the identification of a protein kinase domain in both TvaC$_{S-87}$ and TvaE$_{S-87}$, which share moderate sequence identity to each other (27.2 % identity). Different from TvaC$_{S-87}$, which is likely an active phosphotransferase possessing the necessary residues for Ser/Thr activation and phosphorylation with adenosine triphosphate (ATP), TvaE$_{S-87}$ appears to be an inactive homolog of phosphotransferases given the absence of these conserved residue in its sequence (**Figure S4**). TvaD$_{S-87}$ is a hypothetic protein sharing no sequence homology with any proteins of known functions. Secondary structure prediction revealed that it resembles phosphothreonine lyases, the bacterial effector proteins capable of inactivating MAP kinase activity via phosphate elimination of a phosphorylated Thr residue, which leads to Dhb production, and

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

thus attenuating the immune response of infected hosts (**Figure S5**).[17] Consequently, while TvaE$_{S-87}$ activity remained unclear based on sequence analysis alone, putative lyase TvaD$_{S-87}$ likely associates with phosphotransferase TvaC$_{S-87}$ for the dehydration of the TvaA$_{S-87}$ precursor peptide, specifically by targeting residues Ser1 and Thr8 (**Figure 2B**). Processing these two residues is a critical step in the proposed biosynthetic pathway of TVAs. Thr8 dehydration leads to a Dhb8-containing precursor to set the structural stage for AviMeCys formation, and Ser1 dehydration results in a Dha1-containing intermediate, where, following the hydrolysis of the LP sequence, the deamination of Dha1 can occur to form the *N*-terminal pyruvyl unit of TVA-YJ-2 (which might then undergo reduction to produce TVA-YJ-1).

Necessity of *tvaCDEF*$_{S-87}$ for the biosynthesis of TVAs. To determine the functions of *tvaC*$_{S-87}$, *tvaD*$_{S-87}$, *tvaE*$_{S-87}$ and *tvaF*$_{S-87}$ in TVA biosynthesis, we first inactivated these genes individually in *S. sp.* NRRL S-87 (**Figure 3B**). The precursor peptide-encoding gene *tvaA*$_{S-87}$ was also inactivated, yielding the *ΔtvaA*$_{S-87}$ mutant as a TVA non-producing, control strain. Compared to the wild-type strain, the *ΔtvaC*$_{S-87}$, *ΔtvaD*$_{S-87}$ and *ΔtvaF*$_{S-87}$ mutants share a phenotype with the *ΔtvaA*$_{S-87}$ mutant and failed to produce TVAs (**Figure 3B**). Phylogenetically, TvaC$_{S-87}$ and TvaD$_{S-87}$ share no sequence homology to any known LanM, LanKC or LanL-type, multifunctional Lan synthetases, suggesting that they evolve from different ancestors for dehydration activity development. Intriguingly, TVA-YJ-2 was identified in

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the *ΔtvaE*$_{S\text{-}87}$ mutant, albeit with a significant decrease of yield. To ascertain TVA production, the *ΔtvaE*$_{S\text{-}87}$ mutant was further engineered by introduction of a constitutive promoter, *PermE**, to enhance the transcription and subsequent expression of the mutant *tva* cluster in which *tvaE*$_{S\text{-}87}$ was deleted. Both TVA-YJ-2 and TVA-YJ-1 were observed, thereby indicating that *tvaE*$_{S\text{-}87}$ is closely related to but not indispensable for the production of TVAs in *S. sp.* NRRL S-87.

TvaC$_{S\text{-}87}$ and TvaD$_{S\text{-}87}$ activities in the dehydration of TvaA$_{S\text{-}87}$. We utilized a heterologous expression strategy to examine dehydrated intermediates in *Escherichia coli*, where the precursor peptide-encoding gene *tvaA*$_{S\text{-}87}$ was co-expressed with related PTM genes (**Figure 4A**). Expressing *tvaA*$_{S\text{-}87}$ with *tvaCD*$_{S\text{-}87}$ resulted in two dehydrated variants of TvaA$_{S\text{-}87}$, i.e., **1** (− 18 Da, − 1 × H_2O) and **2** (− 36 Da, − 2 × H_2O), with a ratio of ~3:1. Both variants were characterized by high-performance liquid chromatography with mass spectrometric (HPLC-MS) and high resolution (HR)-MS/MS detection (**Figures S6**, **S7** and **S8**). Compared to the major product **1**, which bears a Dhb residue derived from Thr8, the minor product **2** is di-dehydrated and has an additional Dha derived from the first Ser1 residue in the CP sequence of TvaA$_{S\text{-}87}$. In contrast, expressing *tvaA*$_{S\text{-}87}$ alone or co-expressing it with *tvaDE*$_{S\text{-}87}$ in *E. coli* only produced unmodified TvaA$_{S\text{-}87}$. To access the dehydration mechanism, we took advantage of the above *E. coli* heterologous system that produces **1** and **2** and conducted site-specific mutagenesis to trace possible

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

intermediate(s) (**Figures 4A** and **S9**). This approach can minimize unexpected effects resulting from the variety of constructions and thus facilitate product profile comparison. TvaD$_{S\text{-}87}$ was subjected to the mutation of residues His24 and Arg26 to Ala at the proposed active site.[17] This double mutation must have inactivated TvaD$_{S\text{-}87}$, because the production of **1** and **2** was completely abolished. Instead, a new variant of TvaA$_{S\text{-}87}$ (**3**), which likely arises from the TvaC$_{S\text{-}87}$-catalyzed phosphorylation of residue Thr8, was accumulated in *E. coli*, consistent with the lack of lyase activity. We did not observe any putative TvaA$_{S\text{-}87}$ variants in which Ser1 is phosphorylated. Most likely, Ser1 processing relies on completing the TvaD$_{S\text{-}87}$-catalyzed conversion of **3** into **1** via phosphate elimination during **2** production.

Using the similar site-specific mutagenesis method, we examined the necessity of residues N241 and D258 in TvaC$_{S\text{-}87}$catalysis (**Figures 4A** and **S4**), where these residues are involved in the Mg^{2+}-chelating of importance to kinase activity.[18] Variants **1-3** were not produced by expressing *tvaA$_{S\text{-}87}$* with mutant *tvaC$_{S\text{-}87}$-N241A/D258A* and wild-type *tvaD$_{S\text{-}87}$* in *E. coli*. Further, replacing wild-type *tvaC$_{S\text{-}87}$* with mutant *tvaC$_{S\text{-}87}$-N241A/D258A* in the *tva* cluster abolished the production of TVAs in *S. sp.* NRRL S-87 (**Figure 3B**). These results supported the phosphotransferase activity of TvaC$_{S\text{-}87}$ and the lyase activity of TvaD$_{S\text{-}87}$. Considering expressing *tvaA$_{S\text{-}87}$* with *tvaC$_{S\text{-}87}$* alone failed to produce **3**, the two PTM enzymes might function together to form a two-component dehydratase complex, in which

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TvaD$_{S\text{-}87}$ activity immediately follows TvaC$_{S\text{-}87}$ activity to install Dhb8 first and Dha1 next in TvaA$_{S\text{-}87}$.

TvaF$_{S\text{-}87}$ activity for the oxidative decarboxylation of TvaA$_{S\text{-}87}$. We next expressed *tvaA$_{S\text{-}87}$* with *tvaF$_{S\text{-}87}$* in *E. coli*, to examine the LanD-like activity of TvaF$_{S\text{-}87}$, i.e., for Cys13 oxidative decarboxylation in the treatment of the precursor peptide (**Figure 2B**). Analyzing the resulting product profile revealed a series of TvaA$_{S\text{-}87}$-related peptides that differ at the *C*-terminus (**Figure 4B**). Among them, the -46 Da ($-$ [2H + CO_2]) product was characterized to be variant **4** (**Figure S10**), which was believed to be the direct product of TvaF$_{S\text{-}87}$ and arises from oxidative decarboxylation of the precursor peptide at the *C*-terminus. This enethiol variant is unstable and tends to undergo spontaneous conversions to various derivatives in *E. coli*, including thiol **4a** by hydrogenation, aldehyde **4b** by dethiolation and alcohol **4c** by further reduction (**Figures 4B** and **S10-S13**). Similar results were observed when assaying the activities of other LanD and LanD-like proteins involved in the biosynthesis of AviCys-containing RiPPs.[12,19] Most likely, related enethiol intermediates are immediately transformed during AviCys formation, otherwise unexpected spontaneous conversions occur because of their high reactivity.

TvaE$_{S\text{-}87}$ activity is necessary for Effective AviMeCys formation. We then tested whether residue AviCys can be formed by combining TvaCD$_{S\text{-}87}$ and TvaF$_{S\text{-}87}$ activities together during the conversion of TvaA$_{S\text{-}87}$ (**Figure 4B**).

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Compared with the combination of *tvaA*$_{S\text{-}87}$ with *tvaF*$_{S\text{-}87}$ alone in *E. coli*, expressing *tvaA*$_{S\text{-}87}$ with *tvaCD*$_{S\text{-}87}$ and *tvaF*$_{S\text{-}87}$ further complicated the product profile, in which the individual products of both TvaCD$_{S\text{-}87}$ (e.g., **1**) and TvaF$_{S\text{-}87}$ (eg., **4** and its derivatives **4a-c**) were observed, indicating the competition of TvaCD$_{S\text{-}87}$ and TvaF$_{S\text{-}87}$ activities in the utilization of substrate TvaA$_{S\text{-}87}$. Intriguingly, we observed a small amount of **5**, a -64 Da (i.e., $-[H_2O + 2H + CO_2]$), relatively stable variant of the precursor peptide (**Figures 4B** and **S14**). Thus, dehydration and oxidative decarboxylation occurred. HR-MS/MS analysis narrowed the difference of this variant from TvaA$_{S\text{-}87}$ to the 6-aa *C*-terminal sequence, **T$_8$VAFHC$_{13}$**. Unlike TvaA$_{S\text{-}87}$ and the above linear variants, no fragment ions were observed within this sequence when analyzing **5**. It was resistant to thiol derivatization with iodoacetamide (IAA); in contrast, TvaA$_{S\text{-}87}$ and its thiol-containing variants were sensitive to this treatment, yielding the derivatives conjugated by one IAA molecules (**Figures S15** and **S16**). Together, **5** was deduced to be the cyclized variant of TvaA$_{S\text{-}87}$ that contains AviCys. This variant likely results from intramolecular Michael addition that follows TvaCD$_{S\text{-}87}$-catalyzed Thr8 dehydration and TvaF$_{S\text{-}87}$-catalyzed, Cys13 oxidative decarboxylation for cyclizing the resulting linear intermediate (**6**) (**Figure 4B**). Similar to **4**, enethiol intermediate **6** appears to be highly reactive and thus was not observed in the *E. coli* system where *tvaACDF*$_{S\text{-}87}$ were overexpressed. The presence of this linear intermediate was supported by the observation of **6c** (**Figures 4B** and **S17**), the Dhb8-containing alcohol derivative of **6**. Although residue AviCys can be formed, the accumulation of the above linear TvaA$_{S\text{-}87}$ variants/derivatives indicates the lack

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of functional cooperation between TvaCD$_{S-87}$ and TvaF$_{S-87}$ activities and particularly the inefficiency of the subsequent cyclization step.

We further added *tvaE$_{S-87}$* into the above *E. coli* system in which *tvaA$_{S-87}$* was expressed with *tvaCD$_{S-87}$* and *tvaF$_{S-87}$*, to examine whether the efficiency in AviCys formation can be improved. Remarkably, the product profile was significantly simplified, leaving **5** and **7**, the two AviCys-containing cyclized variants of TvaA$_{S-87}$ with relatively high yields (**Figures 4B** and **S18**). Compared to **5**, **7** contains an additional dehydrated residue, Dha1, which is derived from Ser1 by TvaCD$_{S-87}$ activity. Clearly, the incorporation of TvaE$_{S-87}$ activity greatly facilitates AviCys formation. Substrate TvaA$_{S-87}$ and its linear variants/derivatives arising from TvaCD$_{S-87}$ activity (e.g., **1** and **2**), TvaF$_{S-87}$ activity (e.g., **4** and **4a-c**) or both of them (e.g., **6** and **6c**) were not observed or accumulated in the *E. coli* system where *tvaACDEF$_{S-87}$* were overexpressed (**Figure 4B**). This finding is important and indicates that TvaE$_{S-87}$ activity not only accelerates the cyclization step but also coordinates TvaCD$_{S-87}$ and TvaF$_{S-87}$ activities down a path for effective AviCys formation.

TvaE$_{S-87}$ interacts with both TvaF$_{S-87}$ and LP-containing substrate during AviCys formation. To examine whether dehydrated TvaA$_{S-87}$ can serve as the precursor for AviCys formation, we overexpressed and purified TvaE$_{S-}$

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$_{87}$ and TvaF$_{S-87}$ from *E. coli* for *in vitro* assays (**Figure S1**). Similar to previously characterized LanD and LanD-like proteins, purified TvaF$_{S-87}$ appeared light yellow and exhibited an absorbance spectrum characteristic of flavin in oxidized form. The non-covalently bound cofactor, which can be released from TvaF$_{S-87}$ by heating, was confirmed to be flavin mononucleotide (FMN) (**Figure S19**). In terms of substrate preparation, dehydrated TvaA$_{S-87}$ variants appeared to be highly unstable, and only Dhb8-containing **1** was isolated from the *E. coli* system in which *tvaA*$_{S-87}$ was expressed with *tvaCD*$_{S-87}$. Over a 2-hr incubation period at 30℃, this mono-dehydrated variant was completely degraded in buffer solution (**Figure 5**). When LanD-like flavoprotein TvaF$_{S-87}$ was added into this solution, we did observe a distinct − 46 Da product (− [2H + CO_2]) (**Figure 2**), which was then characterized to be **5** by HPLC-HR-MS, HR-MS/MS and IAA derivatization as aforementioned. Thus, **1** can be processed to the AviCys-containing, cyclized variant of TvaA$_{S-87}$ in the presence of TvaF$_{S-87}$, presumedly through the formation of linear enethiol intermediate **6**. Although TvaE$_{S-87}$ alone had no effect on **1** conversion, the incubation of **1** with both TvaE$_{S-87}$ and TvaF$_{S-87}$ led to a remarkable improvement in the production of **5**, with a yield > 5-fold higher than that by incubating **1** with TvaF$_{S-87}$ alone (**Figure 5**). Unambiguously, TvaE$_{S-87}$ activity is not absolutely necessary but indeed facilitates AviCys formation, consistent with the above *in vivo* results obtained from gene inactivation in *S. sp.* NRRL S-87 and co-expression in *E. coli*.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

During *in vitro* assays, we observed that $TvaE_{S\text{-}87}$ can stabilize substrate **1**. With this protein, the rapid degradation of **1** was largely prevented over a 2-hr incubation period at 30℃. This observation led to the hypothesis that $TvaE_{S\text{-}87}$ interacts with the precursor peptide-derived substrate, assumedly by targeting its LP sequence part. To support this hypothesis, we chemically synthesized the LP and CP sequences of $TvaA_{S\text{-}87}$ and measured their individual interactions with $TvaE_{S\text{-}87}$ by microscale thermophoresis (MST) analysis (**Figure S20**).[20] As expected, this analysis revealed a K_d value of 29.3 ± 5.1 nM between the LP sequence and $TvaE_{S\text{-}87}$, indicating a strong interaction. In contrast, interaction with $TvaE_{S\text{-}87}$ was not observed when using a 21-aa sequence containing CP (for solubilization, this sequence contains the additional 8 LP amino acids at the *N*-terminus of CP). In addition, we observed a moderate interaction between $TvaE_{S\text{-}87}$ and LanD-like flavoprotein $TvaF_{S\text{-}87}$ by MST, as supported by a K_d value of 1.1 ± 0.3 μM. These results suggested the notion that $TvaE_{S\text{-}87}$ and $TvaF_{S\text{-}87}$ function together by forming a heterologous complex, which can be engaged by binding tightly with the LP sequence of the precursor peptide to oxidatively decarboxylate the *C*-terminal Cys residue and immediately mediate subsequent Michael addition for macrocyclization and AviCys formation.

DISCUSSION & CONCLUSION

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Recent studies on the biosynthesis of Class III lanthipeptide-related RiPPs demonstrated that previously defined activities for Lan thioether formation is also involved in the formation of a distinct AviCys thioether residue.[10] Based on *in vivo* and *in vitro* investigations into the biosynthesis of TVA non-lanthipeptides in *S. sp.* NRRL S-87, we here provided an alternative path for AviCys formation during the establishment of the *C*-terminal macrocyclic ring system of these thioamide RiPPs. This path is mechanistically similar to that for Lan formation; however, it does not involve any activities of known Lan synthetases, either monofunctional (i.e., LanB and LanC for Class I lanthipeptides) or multifunctional (LanM, LanKC or LanL for Class II, III or IV lanthipeptides).

In *S. sp.* NRRL S-87, effectively forming the AviCys residue of TVAs relies on the activities of four dedicated pathway-specific PTM enzymes. Specifically, phosphotransferase $TvaC_{S-87}$ and lyase $TvaD_{S-87}$ are functionally equivalent to but phylogenetically different from the kinase and lyase domains within Class II, III or IV Lan synthetases (i.e., LanM, LanKC or LanL multifunctional proteins). Most likely, they work together to form a two-component complex catalyzing the dehydration of $TvaA_{S-87}$ (**Figure 6A**), by targeting Thr8 first and Ser1 next to install Dhb8 and Dha1 tandemly into the precursor peptide. Similar to previously characterized LanD-like HFCD-fold flavoproteins, $TvaF_{S-87}$ has oxidative decarboxylation activity and is able to process the *C*-terminal Cys residue of $TvaA_{S-87}$ to provide an enethiol group when functioning alone. We did observe AviCys formation by combining

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the activities of $TvaCD_{S\text{-}87}$ and $TvaF_{S\text{-}87}$ in the conversion of $TvaA_{S\text{-}87}$, albeit with a poor yield. However, competition and a lack of coordination between Thr8 dehydration and Cys13 oxidative decarboxylation complicated the product profile. In particular, the cyclization of **6**, the precursor bearing Dhb8 and the enethiol group derived from Cys13, appears to occur in an inefficient manner, as the accumulation of this unstable intermediate led to spontaneous conversions to produce various derivatives including alcohol **6c**.

The phosphotransferase homolog $TvaE_{S\text{-}87}$ plays a critical role in effective AviCys formation. This protein, which does not have phosphotransferase activity, can interact with both $TvaF_{S\text{-}87}$ and particularly the LP sequence of the precursor peptide. Benefiting from these interactions for trapping unstable substrate/intermediate and minimizing unexpected conversions, it could function with $TvaF_{S\text{-}87}$ to form a two-component, minimum AviCys synthetase complex (**Figure 6**). In this complex, $TvaE_{S\text{-}87}$ can be a dedicated cyclase catalyzing Michael addition between Dhb8 and the *C*-terminal enethiol group; alternatively, it serves as a noncatalytic engaging protein to mediate the putative dual activity of $TvaF_{S\text{-}87}$ (i.e., for Cys13 processing and cyclization). Intriguingly, this phosphotransferase homolog can coordinate $TvaCD_{S\text{-}87}$-catalyzed Thr8 dehydration and $TvaF_{S\text{-}87}$-catalyzed Cys13 processing to avoid their competition in substrate, leading to the further hypothesis that similar to the MicKC-MicD and SpaKC-SpaF complexes in the biosynthesis of microvionin lipolanthines and TVA-like thiosparsoamide, respectively,[10,21] $TvaCD_{S\text{-}}$

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$_{87}$ and TvaEF$_{S-87}$ may form a four component super-complex for introducing an AviCys residue into an unmodified precursor peptide (**Figure 6B**). Indeed, the SpaKC-encoding gene proved to be capable of compensating the loss of *tvaC*$_{S-87}$, *tvaD*$_{S-87}$ or *tvaE*$_{S-87}$ in *S. sp.* NRRL S-87, supporting that this LanKC-type Lan synthetase is functionally equivalent to the combination of TvaC$_{S-87}$, TvaD$_{S-87}$ and TvaE$_{S-87}$ (**Figure S21**). This observation indicates that nature employs a convergent evolution strategy to develop AviCys synthetase activity from different protein ancestors. At the current stage, we do not exclude the possibility that Cys processing precedes Thr dehydration during AviCys formation in this proposed super-complex because enethiol intermediate is unstable and cannot be isolated for *in vitro* assays.

Recently, Tao *et al*. identified lexapeptide, a new RiPP containing both Lan and AviCys thioether residues, and the associated biosynthetic gene cluster in *S. rochei* Sal35.[16] Using a similar heterologous co-expression method in *E. coli*, four dedicated PTM enzymes, i.e., LxmKYXD, were indicated to be necessary for the formation of the AviCys residue in lexapeptide. Remarkably, despite poor homology in primary sequence, secondary structure prediction revealed that LxmKYXD are the counterparts of TvaCDEF$_{S-87}$ in this study. It should be noted that we here reassigned the functions of LxmY and LxmX based on the above analysis in both primary sequence and secondary structure (**Figure S22**). LxmY can be a lyase functioning with phosphotransferase LxmK, which appears to be active and

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

possesses all necessary residues for Ser/Thr activation and phosphorylation with ATP; in contrast, LxmX, which lacks conserved active site residues, is a phosphotransferase homolog proposed to be functionally equivalent to $TvaE_{S-87}$ in *S. sp.* NRRL S-87. Whether the activities associated with LxmKYX are involved in the formation of the Lan residue occurring in lexapeptide remains to be determined. If so, LxmX and its counterpart $TvaE_{S-87}$ in TVA biosynthesis more likely works as a new type, dedicated cyclase functionally corresponding to the cyclase domain of the LanKC-like proteins MicKC and SpaKC (**Figures S23**). However, whether Lan synthetase-encoding gene(s) outside the *lxm* cluster is the involved in the formation of the Lan residue in lexapeptide needs to be clarified. Determining the role of the $TvaE_{S-87}$/LxmX-represented component, either as a cyclase or as an engaging protein, in related AviCys synthetase complex would be extremely interesting. In addition, genome survey revealed a number of genes coding for *tvaE*$_{S-87}$/*lxmKYXD* homologs, which appear to be clustered with different PTM-encoding genes (**Figures S24**), suggesting that the path demonstrated here is widely employed in bacteria for the biosynthesis of diverse AviCys-containing RiPPs. The study presented here provides insights into the biosynthesis of TVA non-lanthipeptides that remains poorly understood. Using chemoenzymatic or synthetic biology strategies, enzymatic AviCys formation has potential for peptide drug development by application in peptide macrocyclization and functionalization, which remain a challenge to current synthesis approaches.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MATRIALS & METHODS

For details of the materials and methods used in this study, please see **Supporting Information**.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: xxxxx

Supplementary materials and methods, results, **Figures S1−S25**, and **Tables S1−S4**.

AUTHOR INFORMATION

# These authors contributed equally to this work.

Corresponding Author

*wliu@sioc.ac.cn

Notes

The authors declare no competing financial interest.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ACKNOWLEDGMENT

We thank Zhijun Tang for discussion during manuscript preparation and Dr. Hao-Yang Wang at the laboratory of Mass Spectrometry Analysis for helps in the mass data collection. This work was supported in part by grants from MST (2019YFA0905400, 2018YFA090103 and 2018ZX091711001-006-010), NSFC (21520102004, 21750004 and 21621002), STCSM (17JC1405200), CAS (XDB20020200) (for L.P.); and NSFC (21750004 and 21520102004), CAS (QYZDJ-SSW-SLH037 and XDB20020200), STCSM (17JC1405100) and K. C. Wang Education Foundation.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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*, 108.

(2) (a) 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*, 6; (b) Izumikawa, M.; Kozone, I.; Hashimoto, J.; Kagaya, N.; Takagi, M.; Koiwai, H.; Komatsu, M.; Fujie, M.; Satoh, N.; Ikeda, H.; Shin-ya, K., Novel thioviridamide derivative--JBIR-140: heterologous expression of the gene cluster for thioviridamide biosynthesis. *J Antibiot (Tokyo)* **2015,** *68*, 533.

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

(4) (a) Kjaerulff, L.; Sikandar, A.; Zaburannyi, N.; Adam, S.; Herrmann, J.; Koehnke, J.; Muller, R., Thioholgamides: Thioamide-Containing Cytotoxic RiPP Natural Products. *ACS Chem. Biol.* **2017,** *12*, 2837; (b) 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*, 2815; (c) Santos-Aberturas, J.; Chandra, G.; Frattaruolo, L.; Lacret, R.; Pham, T. H.; Vior, N. M.; Eyles, T. H.; Truman, A. W., Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. *Nucleic Acids Res.* **2019,** *47*, 4624; (d) Li, Y.; Liu, J.; Tang, H.; Qiu, Y.; Chen, D.; Liu, W., Discovery of New Thioviridamide-Like Compounds with Antitumor Activities. *Chinese J. Chem.* **2019,** *37*, 1015.

(5) Frattaruolo, L.; Fiorillo, M.; Brindisi, M.; Curcio, R.; Dolce, V.; Lacret, R.; Truman, A. W.; Sotgia, F.; Lisanti, M. P.; Cappello, A. R., Thioalbamide, A Thioamidated Peptide from Amycolatopsis alba, Affects Tumor Growth and Stemness by Inducing Metabolic Dysfunction and Oxidative Stress. *Cells* **2019,** *8*.

(6) (a) Sit, C. S.; Yoganathan, S.; Vederas, J. C., Biosynthesis of aminovinyl-cysteine-containing peptides and its application in the production of potential drug candidates. *Acc. Chem. Res.* **2011,** *44*, 261; (b) Tang, J.; Lu, J.; Luo, Q.; Wang, H., Discovery and biosynthesis of thioviridamide-like compounds. *Chinese Chem. Lett.* **2018,** *29*, 1022.

(7) (a) Minami, Y.; Yoshida, K.; Azuma, R.; Urakawa, A.; Kawauchi, T.; Otani, T.; Komiyama, K.; Omura, S.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Structure of Cypemycin, a New Peptide Antibiotic. *Tetrahedron Lett.* **1994,** *35*, 8001; (b) Claesen, J.; Bibb, M., Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 16297.

(8) Ortiz-López, F. J.; Carretero-Molina, D.; Sánchez-Hidalgo, M.; Martín, J.; González, I.; Román-Hurtado, F.; Cruz, M.; García-Fernández, S.; Reyes, F.; Deisinger, J. P.; Müller, A.; Schneider, T.; Genilloud, O., Cacaoidin, First Member of the New Lanthidin RiPP Family. *Angew. Chem. Int. Ed. Engl.* **2020,** *59*, 12654.

(9) (a) Schnell, N.; Entian, K. D.; Schneider, U.; Gotz, F.; Zahner, H.; Kellner, R.; Jung, G., Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings. *Nature* **1988,** *333*, 276; (b) Schnell, N.; Engelke, G.; Augustin, J.; Rosenstein, R.; Ungermann, V.; Gotz, F.; Entian, K. D., Analysis of genes involved in the biosynthesis of lantibiotic epidermin. *Eur. J. Biochem.* **1992,** *204*, 57; (c) Augustin, J.; Rosenstein, R.; Wieland, B.; Schneider, U.; Schnell, N.; Engelke, G.; Entian, K. D.; Gotz, F., Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of Staphylococcus epidermidis. *Eur. J. Biochem.* **1992,** *204*, 1149.

(10) (a) 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*, 652; (b) Wiebach, V.; Mainz, A.; Schnegotzki, R.; Siegert, M. J.; Hugelland, M.; Pliszka, N.; Sussmuth, R., An amphipathic alpha-helix guides maturation of the ribosomally-synthesized lipolanthines. *Angew. Chem. Int. Ed. Engl.* **2020**. DOI:10.1002/anie.202003804.

(11) Repka, L. M.; Chekan, J. R.; Nair, S. K.; van der Donk, W. A., Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. *Chem. Rev.* **2017,** *117*, 5457.

(12) Ortega, M. A.; Cogan, D. P.; Mukherjee, S.; Garg, N.; Li, B.; Thibodeaux, G. N.; Maffioli, S. I.; Donadio, S.; Sosio, M.; Escano, J.; Smith, L.; Nair, S. K.; van der Donk, W. A., Two Flavoenzymes Catalyze the Post-Translational Generation of 5-Chlorotryptophan and 2-Aminovinyl-Cysteine during NAI-107 Biosynthesis. *ACS Chem. Biol.* **2017,** *12*, 548.

(13) Liao, R.; Duan, L.; Lei, C.; Pan, H.; Ding, Y.; Zhang, Q.; Chen, D.; Shen, B.; Yu, Y.; Liu, W., Thiopeptide biosynthesis featuring ribosomally synthesized precursor peptides and conserved posttranslational modifications. *Chem. Biol.* **2009,** *16*, 141.

(14) (a) 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*, 3030; (b) Liu, J.; Lin, Z.; Li, Y.; Zheng, Q.; Chen, D.; Liu, W., Insights into the thioamidation of thiopeptins to enhance the understanding of the biosynthetic logic of thioamide-containing thiopeptides. *Org. Biomol. Chem.* **2019,** *17*, 3727; (c) Mahanta, N.; Szantai-Kis, D. M.; Petersson, E. J.; Mitchell, D. A., Biosynthesis and Chemical Applications of Thioamides. *ACS Chem. Biol.* **2019,** *14*, 142; (d) Dong, S. H.; Liu, A.; Mahanta, N.; Mitchell, D. A.; Nair, S. K., Mechanistic Basis for Ribosomal Peptide Backbone Modifications. *ACS Cent. Sci.* **2019,** *5*, 842.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(15) (a) Kupke, T.; Stevanovic, S.; Sahl, H. G.; Gotz, F., Purification and characterization of EpiD, a flavoprotein involved in the biosynthesis of the lantibiotic epidermin. *J. Bacteriol.* **1992,** *174*, 5354; (b) 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*, 11282; (c) Majer, F.; Schmid, D. G.; Altena, K.; Bierbaum, G.; Kupke, T., The Flavoprotein MrsD Catalyzes the Oxidative Decarboxylation Reaction Involved in Formation of the Peptidoglycan Biosynthesis Inhibitor Mersacidin. *J. Bacteriol.* **2002,** *184*, 1234.

(16) Tao, M.; 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., Functional genome mining reveals a novel class V lanthipeptide containing a D-amino acid introduced by an F420H2-dependent reductase. *Angew. Chem. Int. Ed. Engl.* **2020**. DOI:10.1002/anie.202008035.

(17) Harouz, H.; Rachez, C.; Meijer, B. M.; Marteyn, B.; Donnadieu, F.; Cammas, F.; Muchardt, C.; Sansonetti, P.; Arbibe, L., Shigella flexneri targets the HP1gamma subcode through the phosphothreonine lyase OspF. *EMBO J.* **2014,** *33*, 2606.

(18) Abu-Saleh, A. A. A.; Sharma, S.; Yadav, A.; Poirier, R. A., Role of Asp190 in the Phosphorylation of the Antibiotic Kanamycin Catalyzed by the Aminoglycoside Phosphotransferase Enzyme: A Combined QM:QM and MD Study. *J. Phys. Chem. B* **2020,** *124,* 3494.

(19) 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*, 4676.

(20) (a) Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S., Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **2010,** *1*, 100; (b) Wang, T.; Liang, L.; Xue, Y.; Jia, P. F.; Chen, W.; Zhang, M. X.; Wang, Y. C.; Li, H. J.; Yang, W. C., A receptor heteromer mediates the male perception of female attractants in plants. *Nature* **2016,** *531*, 241.

(21) Lu, J.; Wu, Y.; Li, J.;Li, Y.; Zhang, Y.; Bai, Z.; Zheng, J.; Zhu, J.; Wang, H., Lanthipeptide Synthetases Participate the Biosynthesis of 2-Aminovinyl-Cysteine Motifs in Thioamitides. Unpublished Manuscript.

(22) Rahman, I. R., Acedo, J. Z., Liu, X. R., Zhu, L., Arrington, J., Gross, M. L., van der Donk, W. A., Substrate Recognition by the Class II Lanthipeptide Synthetase HalM2. *ACS Chem. Biol.* **2020,** *15*, 1473.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Legends

Figure 1. Thioether residue-associated structures and forming mechanisms. (**A**) Selected RiPPs containing an AviCys residue (blue). For the nonproteinogenic residues in addition to AviCys and Lan, Dhb, 2,3-didehydrobutyrine; Abu, α-aminobutyric acid; *a*-Ile, *allo*-isoleucine; and Avi, avionin. (**B**) Mechanistic comparison in the formation of thioether residues Lan (top) and AviCys (bottom). Modified residues during AviCys formation are labeled in blue.

Figure 2. Gene cluster and proposed biosynthetic pathway of TVAs in *S. sp.* NRRL S-87. (**A**) Organization of genes responsible for the biosynthesis of TVAs. Gene functions are annotated by colored rectangular blocks. (**B**) Precursor peptide ($TvaA_{S-87}$) and related PTMs in the biosynthesis of TVAs. Residues related to AviCys formation are highlighted in blue. The LP and CP parts of $TvaA_{S-87}$ are indicated. The CP sequence is numbered. PTMs are highlighted by color, as grey for LP removal to form the *N*-terminal latyl and pyruvyl units of TVA-YJ-1 and TVA-YJ-2, respectively; light purple for His12 functionalization (top right) and thioamide formation (bottom left); and light blue for AviCys formation.

Figure 3. *In vivo* examination of TVA production by HPLC-MS. For extracted ion chromatograms (EICs), the ESI

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

m/z M^+ modes for TVA-YJ-1 and TVA-YJ-2 are 1305.5 and 1307.5, respectively. (**A**) Validation of the $tva_{S\text{-}87}$ cluster (cloned from *S. sp.* NRRL S-87) by heterologous expression in *S. laurentii*. For HPLC-MS analysis, **Method I** was used (see **Supporting Information**). i, wild-type *S. laurentii*; ii, *S. laurentii* variant into which $tva_{S\text{-}87}$ was introduced; iii, $\Delta lanL_{SL}$ *S. laurentii* mutant; iv, $\Delta lanL_{SL}$ *S. laurentii* mutant into which $tva_{S\text{-}87}$ was introduced; and v, wild-type *S. sp.* NRRL S-87 (the TVA producer as a control). (**B**) Validation of the necessity of $tvaCDEF_{S\text{-}87}$ for the biosynthesis of TVAs in *S. sp.* NRRL S-87. For HPLC-MS analysis, **Method II** was used (see **Supporting Information**). i, wild-type; ii, $\Delta tvaA_{S\text{-}87}$ mutant; iii, $\Delta tvaC_{S\text{-}87}$ mutant; iv, $\Delta tvaC_{S\text{-}87}$ mutant in which $tvaC_{S\text{-}87}$ was expressed *in trans*; v, $\Delta tvaC_{S\text{-}87}$ mutant in which $tvaC_{S\text{-}87}$*-N241A/D258A* was expressed *in trans*; vi, $\Delta tvaD_{S\text{-}87}$ mutant; vii, $\Delta tvaE_{S\text{-}87}$ mutant; viii, $\Delta tvaE_{S\text{-}87}$ mutant in which the mutated $tva_{S\text{-}87}$ cluster was under the control of the promoter *PermE**; and ix, $\Delta tvaF_{S\text{-}87}$ mutant.

Figure 4. Heterologous expression of AviCys-forming genes in *E. coli*. Target residues for PTMs are labeled in blue. (**A**) Examination of $TvaC_{S\text{-}87}$ and $TvaD_{S\text{-}87}$ activities in the dehydration of $TvaA_{S\text{-}87}$. The genes expressed in *E. coli* include precursor peptide-encoding gene $tvaA_{S\text{-}87}$ alone (i), and the combination of $tvaA_{S\text{-}87}$ with PTM-encoding genes $tvaCD_{S\text{-}87}$ (ii); $tvaDE_{S\text{-}87}$ (iii); mutant $tvaC_{S\text{-}87}$*-N241A/D258A* and wild-type $tvaD_{S\text{-}87}$ (iv); or wild-type $tvaC_{S\text{-}87}$ and mutant $tvaD_{S\text{-}87}$*-H24A/R26A* (v). (**B**) Examination of $TvaE_{S\text{-}87}$ and $TvaF_{S\text{-}87}$ activities in the formation of residue AviCys. The genes expressed in *E. coli* include the combination of $tvaA_{S\text{-}87}$ with PTM-encoding genes $tvaF_{S\text{-}87}$ (i); $tvaCDF_{S\text{-}87}$ (ii); and

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

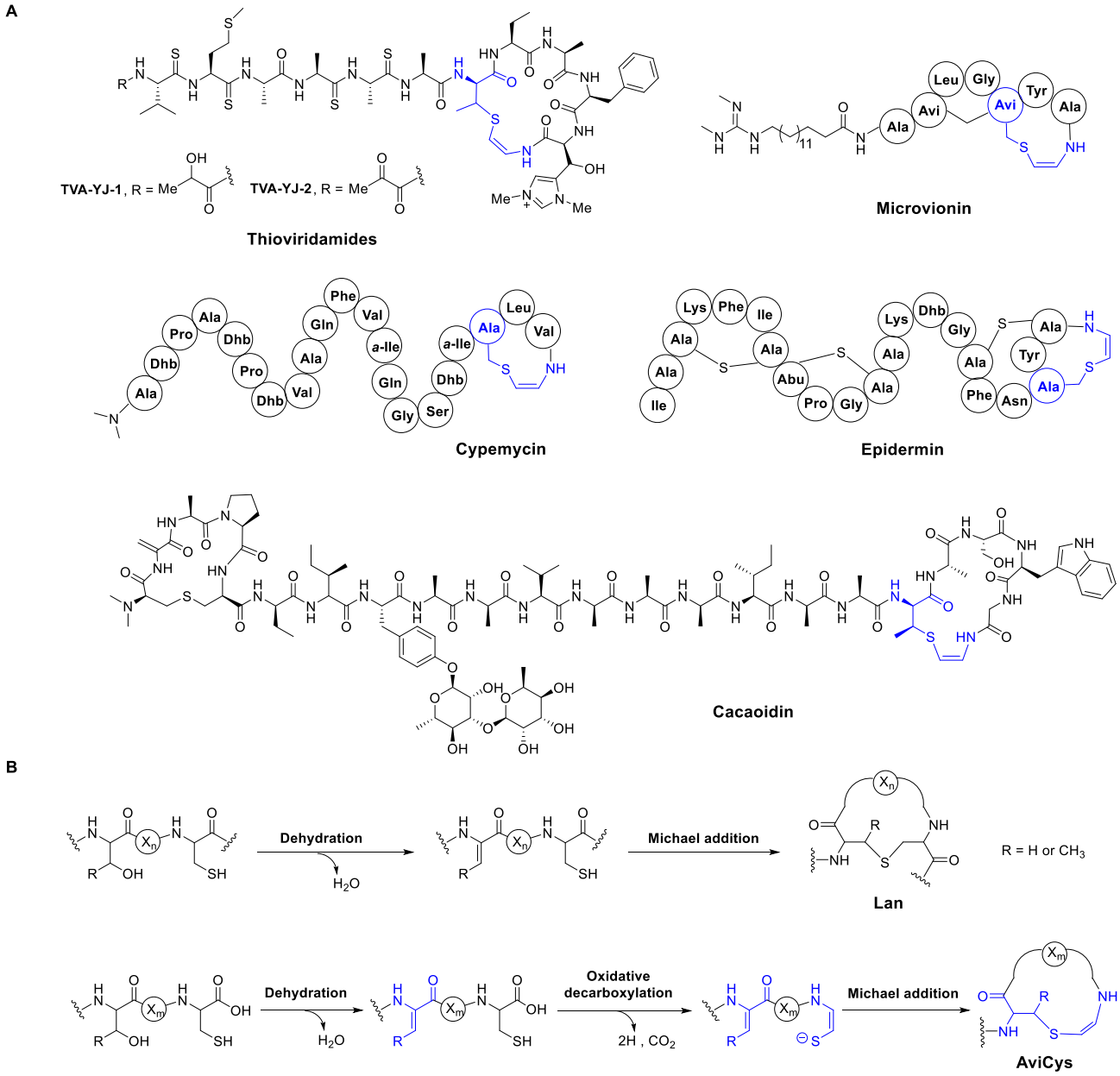
tvaCDEF$_{S\text{-}87}$ (iii).

Figure 5. *In vitro* conversion of substrate **1** (in which residue Thr8 is dehydrated to Dhb8). For HPLC-MS analysis, the ESI *m/z* $[M+H]^{6+}$ modes for **1** and cyclized AviCys-containing product **5** were shown. Tested samples include newly prepared **1** (i), the incubations of **1** alone at 30℃ for 2-hr (ii) or with TvaE$_{S\text{-}87}$ (iii), TvaF$_{S\text{-}87}$ (iv) and TvaEF$_{S\text{-}87}$ (v), respectively, at 30℃ for 2-hr.

Figure 6. Proposed modes for AviCys formation during the biosynthesis of TVAs in *S. Sp.* NRRL S-87. Step 1, Thr phosphorylation; Step 2, H_2O elimination; Step 3, Cys oxidative decarboxylation; and Step 4, cyclization. (**A**) Two component, TvaCD$_{S\text{-}87}$ dehydratase complex and minimum TvaEF$_{S\text{-}87}$ AviCys synthetase complex. In this mode, phosphotransferase TvaC$_{S\text{-}87}$ is proposed to share with phosphotransferase homolog TvaE$_{S\text{-}87}$ the ability for LP binding according to ref. 22. (**B**) Four component TvaCDEF$_{S\text{-}87}$ AviCys synthetase super-complex.

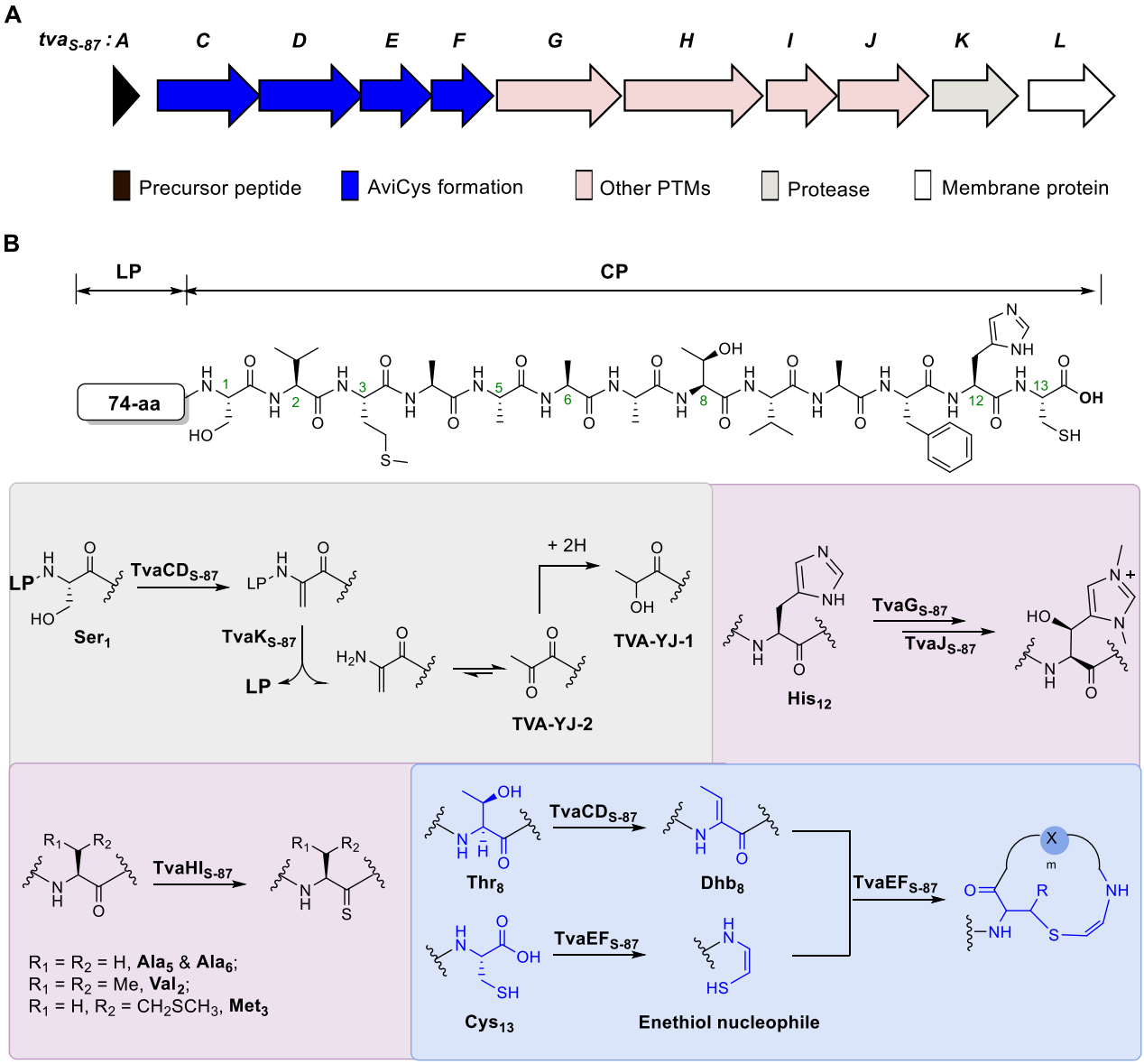
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1.



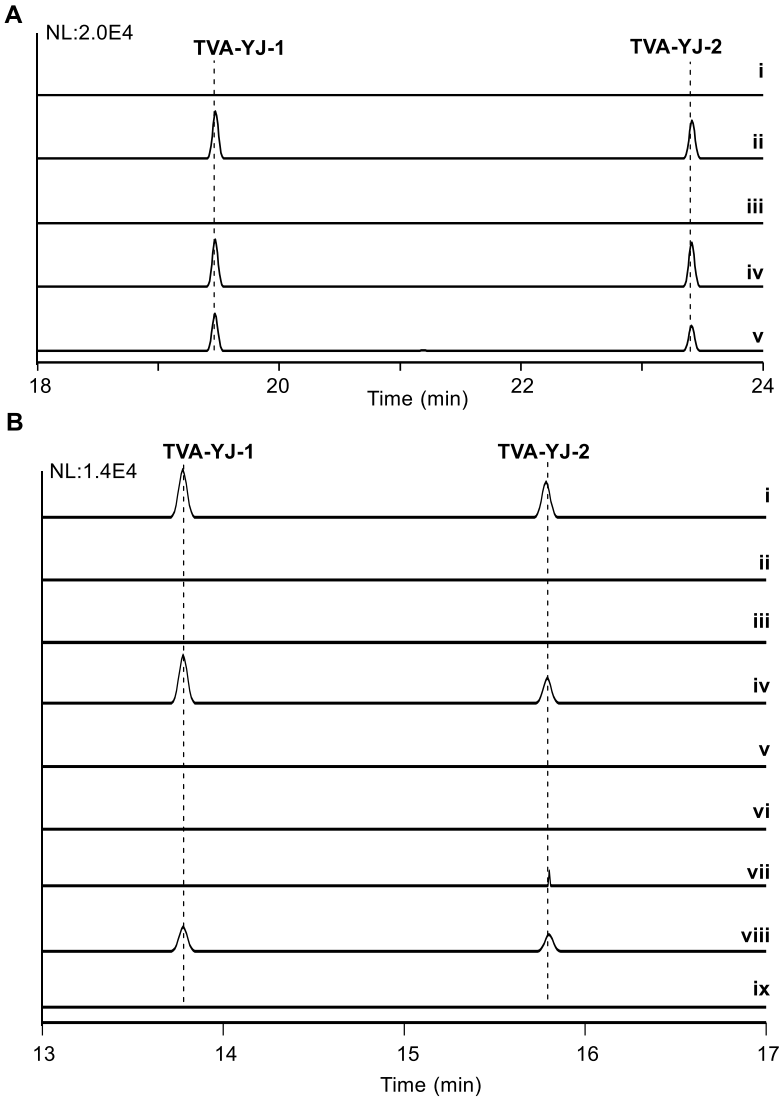
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2.



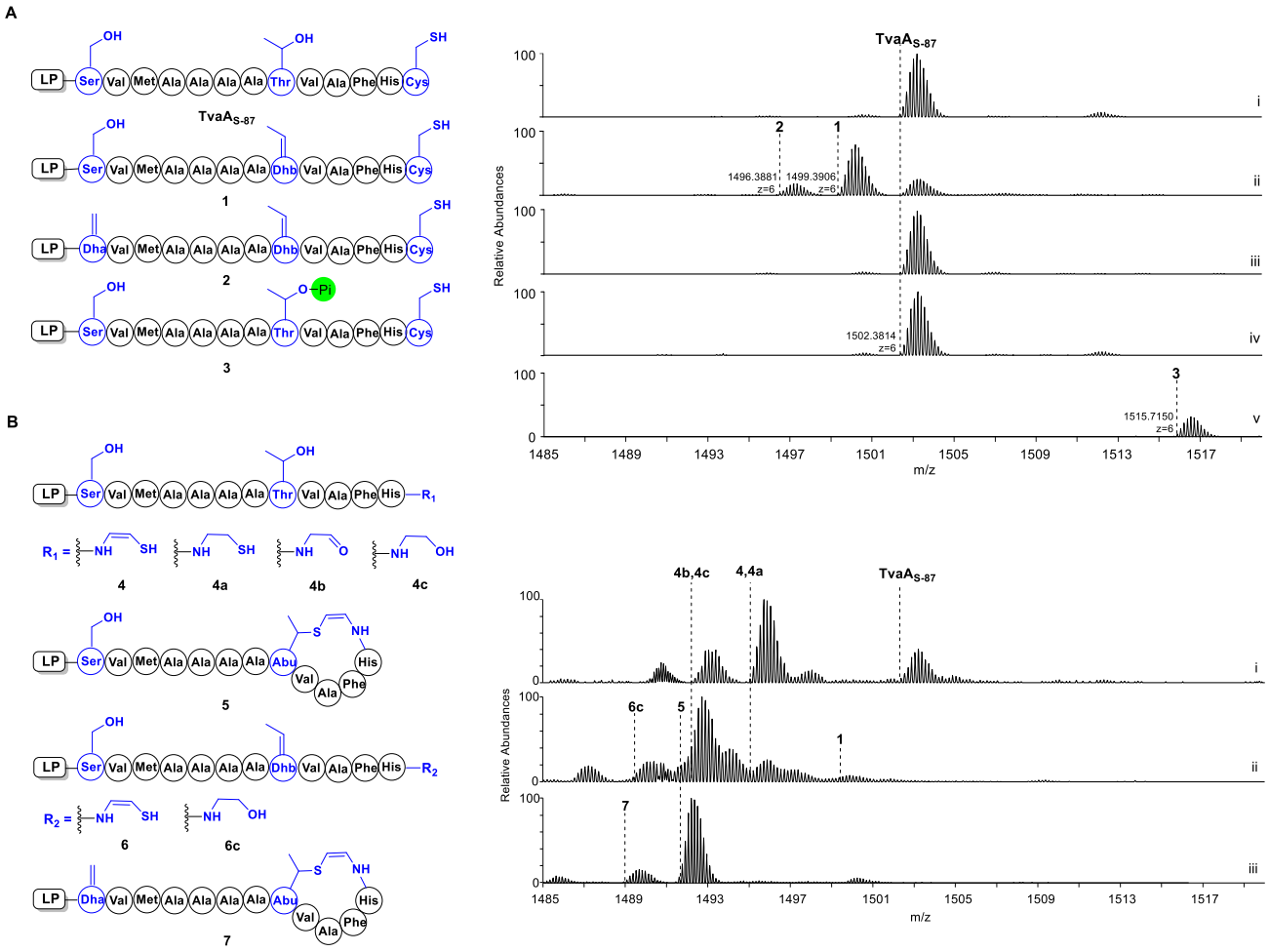
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3.



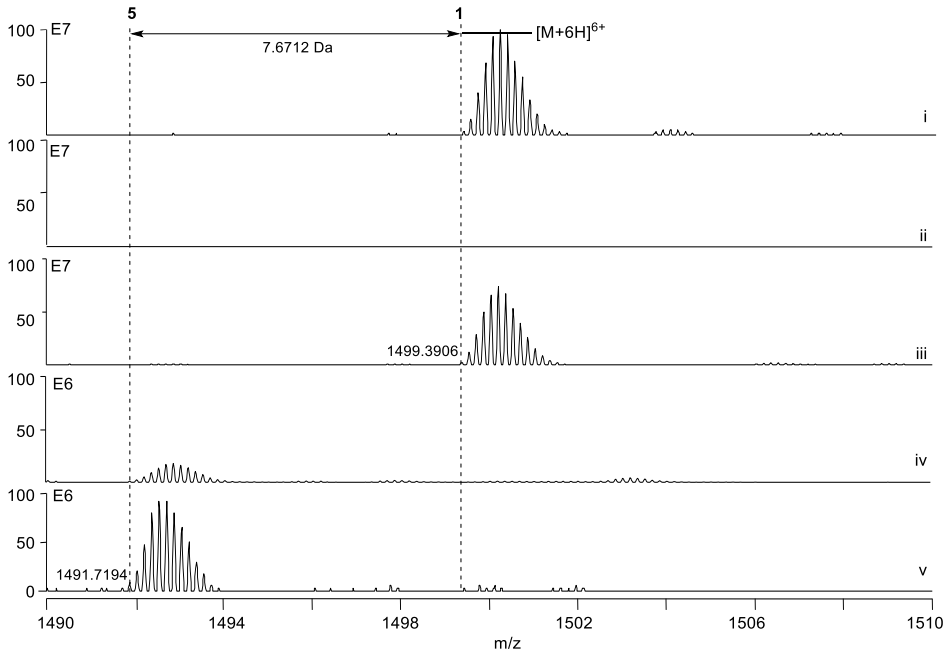
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4.



bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5.



bioRxiv preprint doi: https://doi.org/10.1101/2020.08.21.260521; this version posted August 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6.