ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

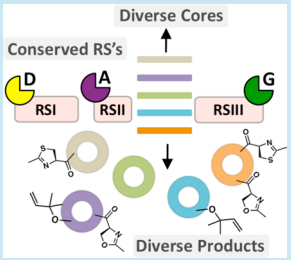
Recognition Sequences and Substrate Evolution in Cyanobactin Biosynthesis

Debosmita Sardar, Elizabeth Pierce, John A. McIntosh,† and Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: Ribosomally synthesized and posttranslationally modified peptide (RiPP) natural products are of broad interest because of their intrinsic bioactivities and potential for synthetic biology. The RiPP cyanobactin pathways *pat* and *tru* have been experimentally shown to be extremely tolerant of mutations. In nature, the pathways exhibit "substrate evolution", where enzymes remain constant while the substrates of those enzymes are hypervariable and readily evolvable. Here, we sought to determine the mechanism behind this promiscuity. Analysis of a series of different enzyme−substrate combinations from five different cyanobactin gene clusters, in addition to engineered substrates, led us to define short discrete recognition elements within substrates that are responsible for directing enzymes. We show that these recognition sequences (RSs) are portable and can be interchanged to control which functional groups are added to the final natural product. In addition to the previously assigned N- and C-terminal proteolysis RSs, here we assign the RS for heterocyclization modification. We show that substrate elements can be swapped *in vivo* leading to successful production of natural products in *E. coli*. The exchangeability of these elements holds promise in synthetic biology approaches to tailor peptide products *in vivo* and *in vitro*.



KEYWORDS: *cyanobactins, posttranslational modifications, ribosomally synthesized and posttranslationally modified peptides (RiPP), recognition sequences, substrate evolution*

Ribosomally synthesized and posttranslationally modified peptides (RiPPs) are ubiquitous natural products found in all branches of life. The resulting compounds are chemically diverse, and their biological activities are equally varied. For example, some RiPPs are cofactors, while others are secreted antibiotics, cytotoxins, or quorum sensing molecules, among many other activities.[1] Despite the many types of possible posttranslational modifications (PTMs) in RiPPs, RiPP pathways share some nearly universal features.[2,3] First, the natural products are encoded on a precursor peptide, which is ribosomally translated. Within the precursor peptide, a core peptide represents the sequence of the active natural product. Recognition sequences (RSs) often flank the core peptide, where they direct enzymes that modify the core, and are often found on a leader peptide, which serves several different purposes.[4] Ultimately, the leader peptide and the RSs are proteolytically cleaved, leaving the core that results in the mature natural product (Figures 1 and Supporting Information S1-A).

It has long been known that RiPP biosynthetic machinery can tolerate mutations in the core sequence, enabling the engineering of novel "unnatural natural products."[5−11] This factor, coupled with the great variety of PTMs, makes the RiPPs good candidates for synthetic biology. The ultimate goal is to be able to gain control of PTMs so that they are portable; leading to the ability to design highly derived peptide materials in organisms and *in vitro*. The field is still a long way from achieving this goal.

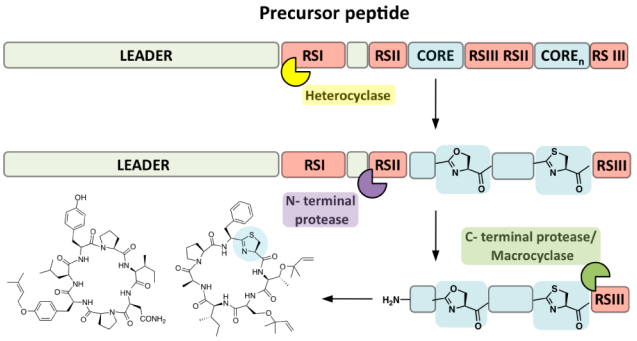


Figure 1. Typical cyanobactin precursor peptide is shown containing leader sequence (green), core peptide sequence (blue), and recognition sequences RS (red) that flank the core. The core sequence eventually matures into the final natural product after proteolytic removal of the rest of the precursor. In a subset of pathways, additional PTMs occur on the core sequence. There are three classes of RS for the PTM enzymes: RSI that directs heterocyclases as defined in this study, and RSII and RSIII that direct the N-terminal and C-terminal protease/macrocyclase, respectively. The structure of the natural products prenylagaramide (left) and trunkamide (right) is shown with the heterocycle modification in trunkamide circled in blue.

We have selected a highly plastic RiPP platform, in which nature already models aspects of this exchange. These are

Received: January 16, 2014
Published: March 13, 2014

Table 1. Sequences of All Peptide Substrates Used in This Study[a]

#	NAME	LEADER	RSI		RSII	CORE	RSIII
1	ThcE4	MDLQNLLPQQSQPIQRATAGQLPTE	LAELTEEAL	NNES	AVLAS	**SCDCSLYGGCESC**	SYEGDEAE
2	TruLy1	MNKKNILPQLGQPVIRLTAGQLSSQ	LAELSEEAL	G---	GVDAS	**-----TLPVPTLC**	SYD-
					GVDAS	**------VCMPCYP**	SYDD
3	TruLy2	MNKKNILPQLGQPVIRLTAGQLSSQ	LAELSEEAL	G---	GVDAS	**-----TFPVPTVC**	SYD-
					GVDAS	**------ACMPCYP**	SYDD
4						**CITFC**	A
5						**AITFC**	AYDGE
6						**CITFC**	AYDGE
7			LAELSEEAL	G---	GVDAS	**------TSIAPFC**	SYD
8	TruLy2 Glu->Ala	MNKKNILPQLGQPVIRLTAGQLSSQ	LAALSAAAL	G---	GVDAS	**-----TFPVPTVC**	SYD-
					GVDAS	**------ACMPCYP**	SYDD
9	Pag/Tru Leu->Ala	MTKKNLKPQQAAPVQREINTTSSES	AAEASEEAA	GTST	GVDAS	**------INPYLYC**	SYDD
10	Pag/Tru ΔRSI	MTKKNLKPQQAAPVQREINTTSSES	---------	GTST	GVDAS	**------INPYLYC**	SYDD
11	Pag/Tru	MTKKNLKPQQAAPVQREINTTSSES	LAELSEEAL	GTST	GVDAS	**------INPYLYC**	SYDD
12	TruE2	MNKKNILPQLGQPVIRLTAGQLSSQ	LAELSEEAL	G---	GVDAS	**-----TFPVPTVC**	SYD-
					GVDAS	**------TSIAPFC**	SYDD
13	PatE1-58	MNKKNILPQQGQPVIRLTAGQLSSQ	LAELSEEAL	GDA-	GLEAS	**-----VTACITFC**	AYD-
					GVEPS		
14	LynE	MDKKNILPHQGKPVRLTTNGKLPSH	LAELSEEAL	GGN-	GVDAS	**------ACMPCYP**	SYD-
					GVDAS	**------VCMPCYP**	SYD-
					GVDAS	**------VCMPCYP**	SYDD
15	PagE6	MTKKNLKPQQAAPVQREINTTSSES	---------	GTST	GLTPH	**------INPYLYP**	FAGDDAE
16	Pag/TruLy	MTKKNLKPQQAAPVQREINTTSSES	LAELSEEAL	GTST	GVDAS	**-----TFPVPTVC**	SYD-
					GVDAS	**------ACMPCYP**	SYDD

[a]Substrates **1** and **12**–**15** are native precursor peptides from different pathways. Substrates **1**–**3** and **8**–**15** were expressed as His-tagged constructs. Substrates **8**–**11** were mutated (green highlights) to probe RS requirements. The shorter substrates **4**–**7** were synthesized chemically on resin. Succeeding lines within a single substrate have been used to represent multiple cores. Where required, dashes have been used to align sequences into the specific regions of RSI (yellow), RSII–RSIII (red) and core (blue). Substrates **16**, **3**, and **8** were used for heterologous production of cyanobactins in *E. coli*.

cyanobactin pathways, often found in symbiotic bacteria within marine animals. For example, there are about 50 known natural products produced by the *pat* cyanobactin pathway.[12] Within this pathway, the enzymes are essentially identical, and yet, the sequences represented in the natural products are highly diverse.[13,14] The resulting compounds are found in great abundance in the animals, and work on their coevolution with animals clearly shows that the selection for these diverse products is at the level of the whole animal.[15] Thus, identical enzymes modify many different hypervariable precursor peptides with diverse sequences, with selection operating at the level of the final product. Previously, we showed that only the core peptides are hypervariable, while the remainder of the precursor peptides remains identical, providing a mechanism by which this diversity can be achieved.[7,14] It should be emphasized that the hypervariable regions are truly variable and do not just have a point mutation here or there. Sequences with 0% identity across the 6- to 8-amino acid lengths of the products can be accepted and processed by the pathway. Here, we have named this phenomenon, where pathways are identical in their native contexts but the substrates and products evolve, "substrate evolution". This name does not imply that the substrates are better (faster) substrates for the enzymes, but instead captures the process observed in nature.

In addition to *pat*, the related *tru* pathway is found in symbionts of animals, where it is responsible for synthesizing an additional 12 known cyanobactin compounds.[14] Beyond the native compounds produced by *pat* and *tru*, we have reported an additional 20 unnatural derivatives that we synthesized in *Escherichia coli* using recombinant technology.[9] In both *pat* and *tru*, the core peptide sequences are hypervariable, while all other elements of the pathway remain identical.

Interestingly, when comparing *pat* and *tru* across multiple samples, there is a position where crossovers occur. *pat* and *tru* are identical in their 5′ and 3′ regions, but there is a swap in the middle section of the pathways that encodes different functionality. For example, *pat* products are heterocyclic at Ser and Thr, while *tru* products are prenylated in these positions.[14] The swap between *pat* and *tru* involves an exchange of genes that encode these different modifying enzymes. The natural swap in function, along with the natural ability to encompass potentially millions of derivatives, makes the cyanobactin pathways ideal for understanding how to achieve the synthetic biology goal of creating a designer posttranslational toolkit.

Cyanobactins are initially encoded on a precursor peptide, represented by PatE. In many cyanobactins, PatE and its homologues are modified by a heterocyclase such as PatD to

dx.doi.org/10.1021/sb500019b | *ACS Synth. Biol.* 2015, 4, 167–176

produce thiazoline and oxazoline rings from Cys and Ser/Thr residues.[16] Subsequently, a PatA-like protease cleaves the N-terminal sequence of PatE, and a PatG-like protease removes the C-terminus and performs a macrocyclization reaction to yield an N−C circular product.[13,17] Further tailoring, such as prenylation by LynF-like enzymes and oxidation by PatG oxidase domain-like enzyme,[18] is also possible (Figures 1 and Supporting Information S1-B). A recent study revealed C-terminal methylation of linear products as a further modification in this compound series.[19] Therefore, when a cyanobactin pathway accepts a new substrate, it efficiently passes through not only one, but through all of the enzymatic steps in the pathway.

Previously, we showed that the macrocyclization of cyanobactins is accomplished using a series of "recognition sequences" (RSs), here named RSII and RSIII.[13,17] When these sequences are present, two proteases recognize them and collaboratively excise and macrocyclize extremely sequence diverse substrates *in vivo* and *in vitro*. We also previously proposed that a third recognition sequence (RSI) element directs the heterocyclization events.[20] Here, we show that this RSI is indeed the primary element responsible for recruiting heterocyclases, which then modify extremely diverse core peptide sequences. More importantly, we show that these RSs act independently of surrounding context, and therefore provide rules for their portability and application in creating new compounds. This result has implications in applying RiPP machinery for advances in synthetic biology.

RESULTS AND DISCUSSION

Expression and Purification of Diverse Cyanobactin Enzymes and Substrates. Cyanobactin gene clusters have been grouped into four distinct genotypes that produce distinct and varied natural products with different PTMs.[19,20] Here, we further examined the previously described *tru* and *pat* pathways, which fall into cyanobactin Genotype I and produce the patellamide and trunkamide cyanobactins. We focused on TruA/PatA N-terminal proteases, which are 95% identical and PatD/TruD heterocyclases, which are 88% identical. To explore biosynthetic pathways that had not been previously analyzed, we expressed precursor peptides from the following pathways: the *lyn* pathway (produces the aestuaramides),[21] which also occupies Genotype I but only exhibits on average 60% protein sequence identity with *pat/tru* pathways; the *pag* pathway (produces the prenylagaramides),[20] which is in Genotype II and is highly divergent from other pathways; and the *thc* pathway,[22] which is in Genotype IV and the chemotype of which was unknown at the time we initiated this study.[20] In addition to *pat/tru* enzymes, we expressed protease *thcA* and heterocyclase *thcD*, both from the *thc* pathway that produces the cyanothecamides. ThcA was expected to perform N-terminal proteolysis, and ThcD was expected to transform Cys and/or Ser/Thr residues into thiazoline and oxazoline, respectively. Multiple precursor peptides are present in each pathway, and only 1 or 2 from each set of precursors were selected for study, as designated by the different numbers. A series of engineered precursor peptides was also made that were a hybrid of two or more pathways (described later), further adding to the substrate diversity of this study. A detailed list of all substrates used in this study is given in Table 1. Proteins were expressed in *E. coli* as His-tagged constructs and purified using standard methods (Supporting Information Figure S2).

Biochemical Analysis of *thc* Pathway Enzymes. The *thc* pathway was characterized to add to the repertoire of the previously studied *pat* and *tru* pathways. ThcE4 (**1**) was chosen as the precursor peptide substrate. Putative heterocyclase ThcD and putative N-terminal protease ThcA were used in enzymatic assays. Using optimized reaction conditions, ThcD was incubated with ThcE4 (**1**) and the reaction was analyzed by MS methods. High-resolution FT-ICR ESI MS/MS (FTMS) has previously been used to localize the position of heterocyclization.[16,23,24] The method identifies sites of dehydration via the 18 Da mass difference, and further, it is possible to determine the type of dehydration (whether reverse-Michael or cyclodehydration) by fragmentation pattern.[25] The ThcD/ThcE4 reaction was digested with chymotrypsin to yield easily detectable fragments. MS/MS analysis showed that all heterocycles were thiazolines derived from Cys and clearly localized them to S**C**D**C**SLYGG**C**ES**C**, where **C** stands for thiazoline ring modification (Figures 2A and Supporting Information S3). Detailed analysis of MS spectra is given in Supporting Information.

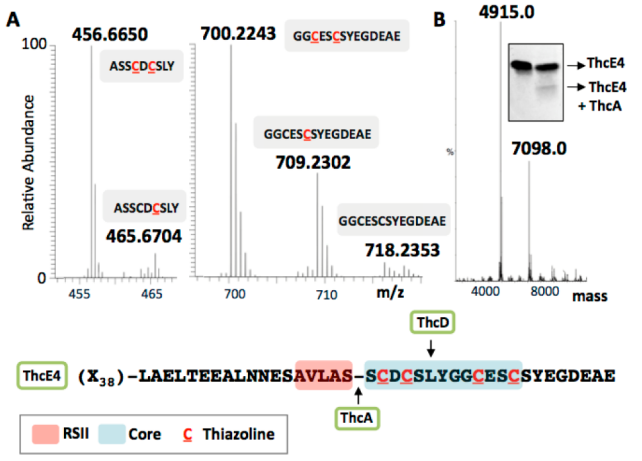


Figure 2. Biochemical characterization of *thc* pathway enzymes. (A) ThcE4/ThcD: FTMS spectra of chymotryptic digests ASSCDCSLY and GGCESCSYEGDEAE of ThcE4 modified by ThcD. Each $[M+2H]^{2+}$ mass peak corresponds to the peptide sequence given in a gray box, and a heterocycle PTM is indicated by a C (in red) within the sequence. (B) ThcE4/ThcA: Deconvoluted ESI-MS spectrum of the ThcE4/ThcA reaction. The $[M-H]^-$ mass peak 7098.0 Da is unmodified ThcE4 (His-tag removed) and 4915.0 Da corresponds to the leader after ThcA proteolysis at the AVLAS RSII site. The inset shows SDS-PAGE visualization of the same reaction, where the left lane is ThcE4 only and the right lane is ThcE4 in presence of ThcA. The smaller band in the right lane indicates the ThcA cleaved product. A schematic representation of results is shown where (X_{38}) represents the 38-residue leader sequence before RSI, RSII (red) and core (blue) sequences.

ThcA was incubated both with the ThcE4 precursor peptide and with the product of the ThcD reaction, which is presumably its native substrate. In both cases, ThcA was active, as observed by SDS-PAGE (Figure 2B inset). In addition, ESI-MS analysis showed that the precursor peptide was cleaved at the RSII encoded by AVLAS (Figure 2B). This RSII is homologous to the previously described PatA RSII sequence of GL(V)E(D)AS.[12] Since the C-terminal core sequence boundary could be easily assigned based on the C-terminal conserved Cys residue,[20] determination of this RSII helped assign the boundary of the ThcE4 core sequence. Figure 2

dx.doi.org/10.1021/sb500019b | ACS Synth. Biol. 2015, 4, 167−176

shows a schematic representation of both of these PTMs of heterocyclization and N-terminal proteolysis of ThcE4.

Efficient Heterocyclase Activity Is Driven by RSI. Before exploring the portability of RS elements, we first needed to identify the heterocyclase RS. It was previously speculated that a LAELSEEAL-like sequence that is conserved in pathways carrying heterocycles is a recognition element for this modification.[19,20] By contrast, when a pathway lacks heterocycles, the corresponding LAELSEEAL-like sequence is absent.[20,26] A similar element is also seen in non-cyanobactin RiPP families such as bottromycins[27−29] and YM-216391[30] (Supporting Information Figure S4). To test this hypothesis, a series of substrates was either synthesized or heterologously expressed that contained intact leader sequence, various fragments thereof, or mutations in the putative RSI (Table 1, Supporting Information Figure S5). These substrates were used in tandem with heterocyclases ThcD, TruD, and PatD. While TruD/PatD share high identity, the newly characterized ThcD is about 60% identical to both enzymes. The native substrates of each enzyme are very different from each other, although in all cases the putative RSI element is well conserved. Reactions were analyzed by a combination of techniques as detailed in Methods. Briefly, precursor peptide modification was followed by SDS-PAGE and ESI-MS. The synthesized peptides **4**−**7** could be analyzed by HPLC, followed by ESI-MS and/or MALDI-MS. In certain cases, FTMS was used to further confirm the identity of products. Substrates **4**−**7** had core sequences from *pat* and *tru* precursors.

Full-length precursor peptides **2** (TruLy1, chimera of TruE1 leader up to its first core cassette fused to LynE core) and **3** (TruLy2, chimera of TruE2 leader up to its first core cassette fused to LynE core) containing all RSs were reactive with ThcD, TruD and PatD (Supporting Information Figure S6). In comparison, a short substrate (**4**) containing only a core peptide sequence was not reactive with any enzyme (data not shown). Short substrates containing RSIII (**5** and **6**) were slowly reactive, with reactions reaching only a little more than 50% completion with ThcD after 24 h, although reactions with TruD were comparatively faster (Supporting Information Figures S7−S9 and Table S2). Reaction progress was judged from comparison of HPLC traces of the substrate and product peaks relative to each other. By contrast, synthetic substrate **7**, which included RSI, was highly reactive (Supporting Information Figure S10), with reactions complete within 15 min with ThcD (Figures 3 and Supporting Information S11). These results showed that although RSI was not absolutely essential for enzymatic reaction, the inclusion of RSI led to a reaction that was >150 times faster with respect to the time taken for reaction completion. Note that modification of **7** was even faster than full-length precursors **2**−**3**, which showed roughly 70% reaction completion after 3 h (Supporting Information Figure S12), but the comparison should be treated with caution for reasons including differences in solubilities and the presence of multiple heterocyclizable residues in the precursor.

We proposed that the slow reactivity of **5** and **6** might be caused by poor enzyme recognition due to the absence of RSI. Obtaining good kinetic data is challenging with this series because of the limits of substrate solubility. Therefore, to test this hypothesis, we used competition experiments (see Methods) in which the fastest substrate **7** underwent reaction with ThcD in competition with substrate **3**, which contains RSI, and **5**-**6**, which lack RSI (Figure 4A). All substrates were

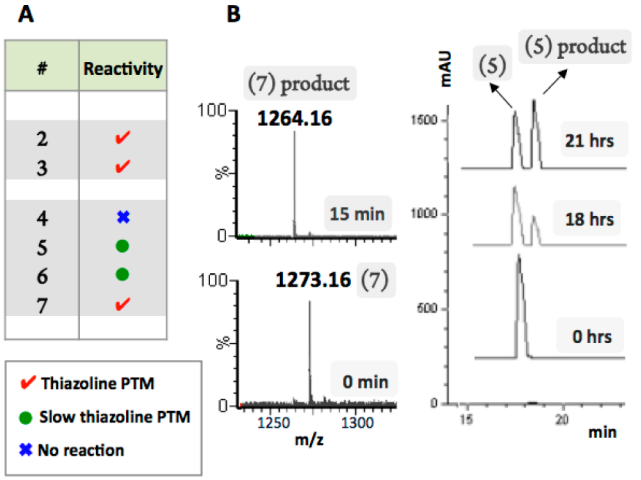


Figure 3. (A) Summary of reactivities of substrates **2**−**7** with ThcD, TruD, and PatD. Full-length precursors **2**−**3** and substrate **7** that lacks most of the leader, but contains RSI, were reactive with all tested enzymes. By contrast, **4** that lacked any RS was unreactive and **5**−**6** that lacked both leader and RSI was a very slow substrate. (B) Using the same assay conditions containing 2 μM ThcD and 200 μM substrate **7** or **5**, a time-course was followed. Substrate **7** was completely modified to give product (1264.16 $[M+2H]^{2+}$) within 15 min of reaction (ESI-MS on left), whereas **5** was very slow, and the reaction was far from complete even after 21 h (HPLC trace on right). Similar results were obtained with substrate **6**, and an expanded reaction time-course is given in Supporting Information Figure S11.

maintained at the same concentrations in these reactions, and it was seen that despite the efficiency of **7** the full-length substrate **3** inhibited modification of **7**, such that no product was seen after the 15-min time point. In contrast, the RSI-lacking substrates **5**−**6** did not inhibit this reaction (Figures 4C and Supporting Information S13). These data indicated that RSI was likely responsible for efficient binding of core peptide to the enzyme.

To confirm that RSI was solely responsible for this effect and not other elements embedded in the substrate or leader peptide, a series of full-length precursors (**8**−**11**) was constructed. RSI motifs were aligned using WebLogo[31] (Supporting Information Figure S14-A), and it was seen that three Glu and three Leu residues were the most conserved. Since the microcin B17 leader peptide is known to adopt a helical conformation,[32] a model of the RSI residues using a helical wheel arrangement was drawn, and it was seen that the conserved Glu and Leu residues were on opposite faces of the helix (Supporting Information Figure S14-B). Hence, the Glu residues were simultaneously converted into Ala residues in triple mutant **8**. The (3[Glu→Ala]) mutant **8** could still be heterocyclized (Supporting Information Figure S15) although with reduced efficiency since only up to double dehydration was observed in comparison to wild-type TruLy2 (**3**) that showed triple dehydration. Since the TruLy2 mutants were harder to purify, the Leu to Ala mutations were made in a different precursor that carried the same leader peptide as TruLy2 but carried *pag* core instead (substrate **9**). In contrast to **8**, the (3[Leu→Ala]) mutant **9** was no longer an efficient substrate (Supporting Information Figure S16). These results indicate that a Leu-rich hydrophobic patch may serve a critical function.

Peptide **10** is a derivative of a natural precursor peptide PagE6 from the non-heterocyclizing *pag* family that lacks the

dx.doi.org/10.1021/sb500019b | ACS Synth. Biol. 2015, 4, 167−176

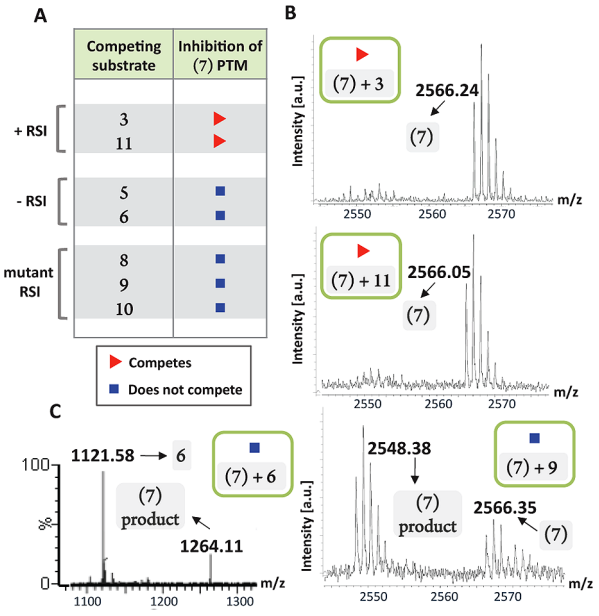


Figure 4. Competition reactions: (A) Substrates **3** and **11** (containing RSI), **5**–**6** (lacking RSI) and **8**–**10** (RSI mutants) were made to compete with ThcD (2 μM) reaction on substrate **7**, which has RSI but lacks the leader before it. Only **3** and **11** inhibited modification of **7**, while **5**–**6** and **8**–**10** did not. All substrates were maintained at concentrations of 50 μM (in reactions with **3**, **11**, and **8**–**10**) or 200 μM (in reactions with **5**–**6**) under the same assay conditions. (B) MALDI-MS of competition with **3** and **11**, which completely inhibit modification of **7**. The mass of 2566 corresponds to $[M+Na]^+$ of **7**. (C) On the left is ESI-MS of competition with **6**, which does not inhibit modification of **7** (product mass 1264.11 $[M+2H]^{2+}$), and substrate **6** is unmodified (1121.58 $[M+H]^+$). On the right is MALDI-MS of competition with **9** that does not affect modification of **7**. The mass peak of 2548.38 corresponds to $[M+Na]^+$ of **7** product, while only a minor peak of unmodified **7** was observed (2566.35 $[M+Na]^+$). Similar results were obtained from competition with **5**, **8**, and **10**, which is shown in Supporting Information Figures S13 and S18 along with necessary controls.

RSI element intrinsically. To make it a potential substrate for heterocyclases, we introduced a Cys residue in the requisite position (with *tru* RSII and RSIII). In addition, peptide **11** was made, in which RSI was inserted into the backbone of peptide **10** (Table 1). Our results showed that **10**, which lacked RSI, was not a substrate for heterocyclases, although it carried a heterocyclizable residue (only a minor product peak was seen that equaled <10% of product even after 18-h reactions), while **11**, which carried RSI insertion, was fully competent for reaction and essentially equal to other full-length precursors (Supporting Information Figure S17). This demonstrates the portability of RSI to import heterocyclization in a completely non-native core peptide sequence.

In addition, competition experiments similar to those described earlier were performed using ThcD and competent substrate **7** versus **8**–**11**. Only substrate **11** with intact RSI competed with **7**, while the RSI mutants **8**–**10** did not (Figures 4 and Supporting Information S18). This further reiterated that RSI is responsible for efficient heterocyclization.

Leader Peptide Coevolves with Cyanobactin Modification Enzymes. To find out the extent of conservation of non-RS elements of the leader sequence, we constructed phylogenetic trees of cyanobactin precursor peptides. We speculated that such an analysis might help us to identify important non-RS elements, if any, which could make portability of PTMs a physiologically and synthetically relevant process. Previously, the hypervariability of core peptides made the tree-building process difficult. This led us to construct trees composed solely of the leader peptide in the absence of the hypervariable cores, both with and without RSI (Figures 5 and

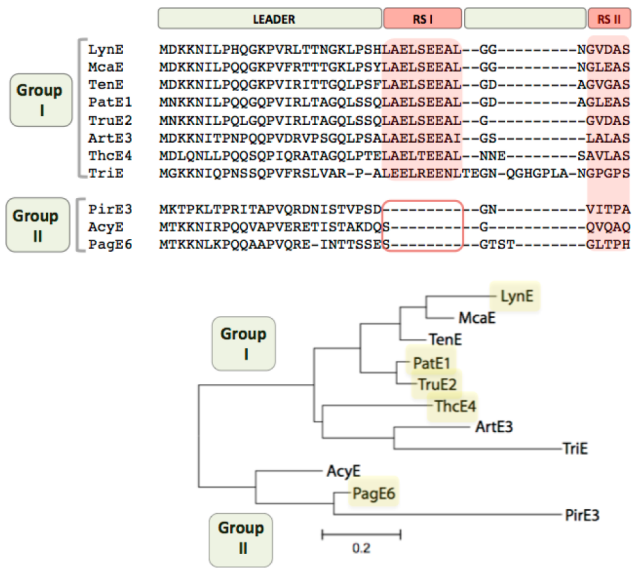


Figure 5. Phylogenetic tree of cyanobactin precursor peptide leader sequences. Group I pathways carry RSI, heterocycle PTM and the corresponding heterocyclase enzyme, whereas Group II pathways lack the same. Only a subset of peptides is shown here, with the ones used in this study highlighted in yellow. See Supporting Information Figure S19 for a complete tree with all known cyanobactin precursors.

Supporting Information S19). Surprisingly, in both cases, the tree topology was identical and split precursor peptides into two groups. Group I peptides contain RSI, and Group II peptides lack RSI. As expected, Group II peptides are derived from pathways that also lack heterocyclases. Note that presence or absence of RSI did not define this branch point since trees made either way were identical.

Furthermore, within Group I, the tree topology is similar to that made using cyanobactin heterocyclase sequences (Supporting Information Figure S20).[20] These observations suggested that the heterocyclases and leader sequences are coevolving, and because of this relationship, we speculated that heterocyclases might prefer their cognate (or native) leader sequences. In that event, not only RSI, but also the remainder of the leader peptide may be needed to enable portability of heterocyclization PTM. To understand this, we took a different approach, where instead of carrying out exhaustive mutations of the leader sequence, we took advantage of the diversity of cyanobactin biochemical components we already had in our hands, as described below.

Noncognate Leaders Do Not Hinder Heterocyclase Reactivity. To test the requirement of cognate leader peptide sequence, ThcD, TruD, and PatD were assayed with a series of both native and non-native substrates from Groups I and II (Figure 6). Group I included the native substrates of these three enzymes, ThcE4 (**1**), TruE2 (**12**), and PatE (**13**), in addition to LynE (**14**) from a different Group I pathway, and TruLy 1 and 2 (**2**) and (**3**), which contain *tru* leaders and hybrid *tru-lyn* cores and were designed in our lab as robust

dx.doi.org/10.1021/sb500019b | *ACS Synth. Biol.* 2015, 4, 167–176

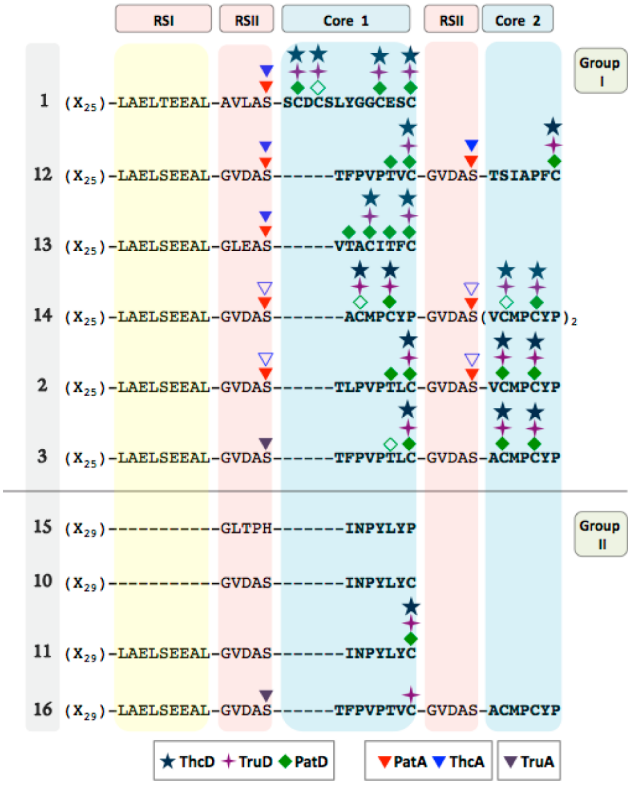


Figure 6. Native leader peptide is not required for cross-selectivity of cyanobactin PTM enzymes. Substrates were grouped based on the type of leader, where Group I are from heterocyclase containing pathways and Group II are from those without. Heterocyclases ThcD/TruD/PatD and proteases PatA/ThcA are reactive with both native and non-native substrates *in vitro* as long as RSI or RSII is present. TruA reactivity was analyzed *in vivo* only on substrates **3** and **16**. (X_n) represents leader peptide where n is the number of residues before RSI. Dashes have been used to align sequences to the color-coded RS and core regions. Regioselectivity of each individual enzyme is shown by specific symbols at the modified residues. Nonfilled symbols represent partially characterized PTMs (Supporting Information).

substrates for TruD. Group II included PagE6 (**15**) from the *pag* pathway, which is divergent from all of the other pathways used in this study. Since the PagE6 core lacked heterocyclizable residues, a Pag/Tru chimera (**11**), which carries a Pro to Cys mutation and RSI, was used.

To our surprise, instead of exhibiting selectivity for cognate leaders as would be expected from the phylogenetic analysis, all heterocyclases were active on all substrates as long as RSI was present (Figure 6). Detailed product characterization is given in Figures S21−24 and Tables S3−S7 (see Supporting Information). As has been reported earlier, TruD was chemoselective for Cys and PatD differed in that it could modify Ser/Thr residues as well.[24] ThcD as characterized in this study resembles TruD in chemoselectivity. Differences in precursor peptide sequences, and even hybrids of different types of precursors, did not alter chemoselectivity or regioselectivity of the resulting products.

Precise kinetic measurements of heterocyclization reactions are complex due to the solubility issues alluded to above and distributive processing by the enzymes, such that the substrate is released from the enzyme after each residue is heterocyclized. Despite these factors, we performed a semiquantitative analysis, wherein the amount of enzyme was varied and reaction completion after an 18-h time-point was monitored based on the SDS-PAGE band-shift (Supporting Information Figure S25). Substrates **14** (LynE), **2** (TruLy1), and **12** (TruE2) were analyzed in this way because of the clear visibility of their band-shifts due to heterocyclization. TruD and ThcD were essentially identical in apparent reactivity, whereas our PatD enzyme preparation was relatively less efficient, such that a higher amount of enzyme was required to allow reaction completion at the same time-point.

Additionally, experiments were performed in which reactions of **14** with TruD and ThcD were competed with substrates **2** (TruLy1), **8** (TruLy2 RSI mutant), and **15** (PagE6), using the same methods that were employed earlier to define RSI. As determined by SDS-PAGE analysis (Supporting Information Figure S26) and confirmed by MS (Supporting Information Figure S27), only **2** carrying intact RSI could compete with **14**. Taken together, these results showed that the native leader peptide was not necessary for reactivity and that leader sequences with <50% sequence identity are recognized by all three heterocyclases as long as RSI is present.

Importance of RSII and Insignificance of the Native Leader Peptide for N-terminal Proteolysis. We wished to see if the above observation with RSI could also be extended to RSII. Beyond PatD-like heterocyclase proteins, the reactivity of PatA and PatG is essential in cyanobactin maturation.[20] PatA and relatives act on full-length heterocyclized precursor peptides. Here, we examined the cross-reactivity of RSII elements with PatA-like proteases. Despite the comparatively lower sequence identity of PatA and ThcA (Supporting Information Figure S2), both could recognize GVDAS-like RSIIs (Figures 6 and Supporting Information S28, Supporting Information Table S8) in native and non-native substrates. The reactivity of TruA was examined *in vivo* (see below). It processed both substrate **3**, which carried native leader, and substrate **16**, which carried a Group II non-native leader. In short, the proteases require RSII for activity, which exhibits the same portability as observed for the heterocyclase RSI.

Finally, PatG and relatives act after N-terminal proteolysis on short substrates carrying the AYDGE-like RSIII. The RSIII-protease interaction has already been extensively studied.[17,33,34] Therefore, with this study and resting on previous work, the selectivity requirements for the major PTM enzymes in the cyanobactin family have been explored. Next, we aimed at testing the portability of all RSs in the context of the complete cyanobactin pathway *in vivo*.

Portability of RSs to Produce Cyanobactins *In Vivo*. The above results with heterocyclases and proteases implied that conserved elements within precursor peptides are virtually interchangeable. Therefore, it should be possible to swap PTMs by importing enzymes and their RSs into preexisting precursor peptides. To test this hypothesis in the context of a whole pathway *in vivo*, we formed hybrid precursor peptides comprising sections of *tru*, *lyn* and *pag* precursors (Supporting Information Figure S29). These were coexpressed with our previously described *tru* production vector in *E. coli* (Supporting Information Figure S29).[9] The resulting cyanobactin products were extracted from the *E. coli* cell pellet and analyzed by ESI-MS (see Methods). Robust expression of patellins from the parent *tru* pathway served as internal controls for cyanobactin expression.

Unfortunately, *lyn* and *pag* core sequences did not lead to detectable products. This could be for any number of reasons, including toxicity to *E. coli*, degradation *in vivo*, etc. Two of the

dx.doi.org/10.1021/sb500019b | *ACS Synth. Biol.* 2015, 4, 167−176

designed precursors did lead to cyanobactin products. The first precursor contained the TruE2 leader peptide (TruLy2; **3**), while the second contained the same TruE2 RSs as **3**, but fused to the PagE6 leader instead (Pag/TruLy; **16**). A TruLy2 mutant (**8**) was also made, carrying a (3[Glu→Ala]) mutated RSI sequence. Our results showed that, irrespective of leader, the *tru* pathway processed both **3** and **16** (Figure 7) to produce

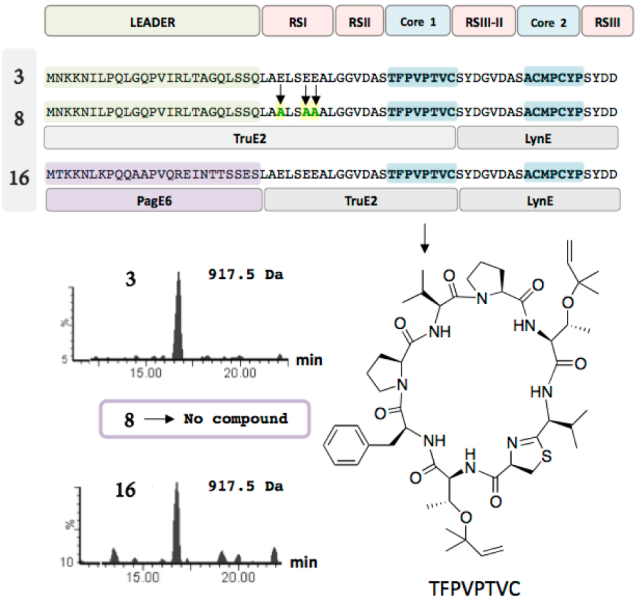


Figure 7. Production of cyanobactins from engineered peptides in *E. coli*. Hybrids of TruE2 and LynE with core peptides TFPVPTVC and ACMPCYP were made. The latter was not detected, but expression of the cyclic TFPVPTVC product was robust. Expression was maintained even when PagE6 leader (purple) replaced the native TruE2 leader. Product formation was abolished when the RSI-mutated precursor was used.

the macrocyclized product TFPVPTVC. In comparison, the mutant **8** failed to produce compounds, as would be expected from the earlier established *in vitro* result that the substrate **8** could not compete with full-length precursors. In the context of the whole *tru* pathway in *E. coli*, where *tru* precursors were present as internal controls (producing patellins), the mutant **8** was not expected to be a substrate. Both nonprenylated and prenylated forms of the compound TFPVPTVC were observed, indicating that the TruF group of prenyltransferase PTM enzymes were also active.

Conclusions. Here, we experimentally define RSI, which we previously proposed to exist based upon extensive sequence analysis. In the presence of an RSI, peptides are efficiently heterocyclized, while in its absence heterocyclization is extremely inefficient or completely abolished. In native cyanobactin precursors lacking RSI, introduction of the element from other pathways enables access to the RS and efficient heterocyclization. Mutations to RSI lead to a greatly diminished enzyme activity when the acidic side of this helical sequence was modified, while activity was abolished on the hydrophobic side of the helix. These results indicate that potentially the hydrophobic helical face is crucial for the interaction with the enzyme. Taken together, we show that RSI is necessary for the efficient reaction required *in vivo*.

It was recently proposed that the sequence identified as RSI here is not responsible for recognition but instead is required just in order to access interior Cys residues.[35] This interpretation does not provide a good evolutionary rationale to maintain RSI because the native substrates of TruD lack interior Cys residues within single cassettes. The results here clearly disprove that idea, showing definitively that RSI is required for efficient synthesis under native conditions. While a slow reaction occurs *in vitro*, this is likely not relevant to the natural reaction.

These add to the previously known RSII and RSIII in cyanobactin biosynthesis. Although these were previously known, here we add to the understanding of RSII function, showing that different sequences can be interchanged. Here, we also show how RSI, RSII, and RSIII function in the context of whole pathways. Previously, precise elements involved in recognition have been determined for individual enzymes and precursor peptides largely through mutagenesis studies in the lanthipeptides and several microcins.[32,36] Additional advances include the *in trans* use of leader peptides and core peptides, rather than the normal fusion products found in nature. This has been especially advantageous, with the covalent fusion of leader peptides with enzymes, providing robust catalysis of, for example, lanthionine bond formation.[37] On the other hand, the lack of importance of the leader has also been demonstrated with the Balh heterocyclases *in vitro*.[38]

Here, we hoped to learn how switching leader peptides would enable us to recruit PTM enzymes. Instead, and to our surprise, we found that quite different leader sequences were still fully functional with noncognate enzymes, and that those leader sequences could be swapped with no impact on activity. The only critical sequences are the three short RS elements, which are necessary to obtain products. The finding that RS elements recruit PTMs has implications for RiPP engineering. Such elements are inherently interchangeable and will thus afford a simple method for mixing different types of PTMs in a rational manner.

Importantly, these noncognate enzymes enabled us to change the pattern of PTMs, leading to the introduction of unnatural modifications *in vitro*. By combining these enzymatic elements at will, novel compounds can be synthesized. Here, we demonstrate the synthesis of some of these compounds *in vitro* by combining PatA and PatD-like enzymes with various substrates. We provide the first example of how this might be applied *in vivo* using a swap where a wholly unnatural precursor peptide still is functionally modified in the context of the *tru* pathway, validating the ability to functionally swap cassettes. On the basis of this technological advance, work is underway in the lab to synthesize libraries of derivatives, which will be reported in due course. The underlying technology, however, should be applicable to other types of RiPP pathways for which recognition elements have been defined.

We used the observation of natural "substrate evolution" and real-time pathway crossovers to inspire the creation of portable synthetic biology tools. Although very precise substrate evolution has still been described only in cyanobactins found within coral reef animals, it is likely that this will be found in many other RiPPs as well, as shown by initial evidence in some other pathway types.[39] Not all RiPPs exhibit this feature, as others that have been studied exhibit fairly narrow substrate selectivity, where more than a small number of mutations are not accepted in the final products. The accumulation of natural RiPP pathways in which substrate evolution is occurring will thus be very valuable in achieving the long-term goal of

dx.doi.org/10.1021/sb500019b | *ACS Synth. Biol.* 2015, 4, 167−176

precisely controlling posttranslational modifications and small peptide design.

METHODS

Genes and Cloning. The ThcA, ThcE4, ThcD, TruLy2, and Pag/Tru genes were obtained from GenScript and cloned into the pET28 vector between NdeI and XhoI sites for protein expression and/or into our previously described pRSF-lac vector[9] between NdeI and KpnI sites for compound expression in *E. coli*. The mutants 3[Glu→Ala], 3[Leu→Ala] and the RSI deletion mutants of TruLy2 and Pag/Tru constructs were made by site-directed PCR mutagenesis.

Protein Expression, Purification, and Synthesis. Precursor peptides ThcE4, LynE, PagE6, TruLy2 and Pag/Tru along with their mutants were made by overexpression (3 h at 37 °C with 1 mM IPTG induction) to drive the protein into inclusion bodies, in R2D-BL21 cells in 2xYT medium. Purification was performed by Ni-NTA affinity chromatography under denaturing conditions. ThcD and ThcA were expressed under similar conditions as the precursor peptides except that milder expression conditions were used (18−21 h at 18 °C with 0.1 mM IPTG induction), and the proteins were purified using native conditions. All proteins were dialyzed, aliquoted, flash-frozen, and stored at −80 °C till used. Enzymes were additionally stored in 10% (v/v) glycerol. The precursor peptides TruE2, PatE, TruLy1, and the enzymes TruD, PatD, and PatA were made as described previously.[15,24] The synthetic substrates **4**−**7** were made at the University of Utah Peptide Synthesis Core Facility.

ThcE4 PTM Assay Conditions. For characterization of *thc* pathway enzymes, ThcE4 (30 μM) was incubated with 0.4 μM of ThcD in optimized additive mixtures containing 50 mM Tris buffer pH 7.5 along with 7.5 mM DTT, 4 mM $MgCl_2$, and 1 mM ATP in a final reaction volume of 50 or 100 μL. The N-terminal cleavage assay was carried out by incubating 40 μM of ThcE4 with 6 μM of ThcA in optimized additive mixtures containing 50 mM Tris buffer pH 7.5 along with 4 mM DTT and 10 mM $CaCl_2$ in a final reaction volume of 50 or 100 μL. In certain cases, cleavage assays were performed after heterocyclization by adding protease to the ThcD/ThcA reaction mixture while maintaining the same final concentration as detailed above. Assays performed in absence of enzyme and/or absence of ATP were used as negative controls. The reactions were incubated at 34 °C for 18 h in a MJ Research Minicycler. Each reaction was done at least in triplicate and was quenched by immediately freezing at −80 °C until used for SDS-PAGE and MS analysis.

Precursor Peptide Assay Conditions and Product Characterization. Reactions were carried out under conditions similar to ThcE4 assays as given above, with 25−50 μM peptide and 0.5−2 μM heterocyclase or 5 μM protease, for 18 h at 34 °C, unless otherwise specified. Products were analyzed in two steps. First, preliminary results were obtained by visualization of modification on SDS-PAGE. Second, reactions were analyzed by ESI-MS, which indicated the number of dehydrations (in case of heterocyclization) or a smaller molecular weight mass (in case of proteolysis). In most cases, identity of the product was further confirmed by high-resolution FTMS and MS/MS fragmentation pattern. See Supporting Information for detailed interpretation of MS data.

Synthetic Substrate Assay Conditions and Product Characterization. Reaction conditions were identical to those for precursor peptide assays, except that up to 200 μM substrate was used with 2 μM heterocyclase, for 18 h at 34 °C, unless otherwise specified. Reactions were quenched with 1 M guanidine hydrochloride, centrifuged, and the supernatant was analyzed by RP-HPLC on a LaChrom Elite System (Hitachi). A 214MS C4 5 μ column (Grace-Vydac) was used and run on a linear gradient from 99% of buffer A (H_2O/0.1% TFA) to 100% ACN over 45 min with a flow rate of 1 mL min^{-1}. Because of the presence of Tyr residues in **5**−**7** a 276 nm UV absorbance maximum was clearly observed for the starting substrate. Thiazoline modification introduces an additional 254 nm shoulder,[40] such that product formation can be easily detected. Further confirmation was provided ESI-MS and/or FTMS. Note that for substrate **7**, it was difficult to attain HPLC separation of substrate and product peaks. Hence, reactions were followed by ESI-MS and/or MALDI-MS. Samples to be analyzed by MALDI were desalted and concentrated using C18 Zip-Tips (Millipore). See Supporting Information for detailed interpretation of MS data.

Competition Reaction Assay Conditions and Product Characterization. Identical reaction conditions were maintained for each competition reaction series. For competition of **7** with precursor peptides **3**, **11**, and **8**−**10**, 50 μM of each substrate was used, since attaining higher concentrations of precursor peptide was difficult, and the reactions were followed by MALDI-MS. For competition of **7** with synthetic peptides **5**−**6**, 200 μM of each substrate was used, and the reactions were followed by ESI-MS. For competition of precursor peptides **14** with other precursor peptides **2**, **8**, and **15**, 25 μM of each substrate was used, and reactions were followed by SDS-PAGE and ESI-MS at increasing time points. All enzymes were used at a concentration of 1 μM, unless otherwise specified.

SDS-PAGE Band-Shift Assay. Comparison of modified peptide with unmodified peptide on SDS-PAGE was used to visualize PTM modification, since heterocyclized peptides showed greater mobility and cleaved peptides resulted in a smaller band in most cases. 18% gels were used to analyzed 15 μL of each reaction by standard SDS-PAGE conditions. In all cases, MS methods confirmed results seen on SDS-PAGE gels.

Mass Spectrometric Methods. ESI-MS and FT-ICR analyses were performed at the University of Utah Mass Spectrometry and Proteomics Core facility. ESI was done using a Micromass Quattro-II (Waters), and analysis of spectra was carried out using predicted masses obtained from Monoisotopic Mass calculator. FT was performed following cleavage of heterocyclized peptide by chymotrypsin or PatA protease, using a LTQ FT Ultra Hybrid Mass Spectrometer (Thermo Scientific) and analyzed both manually using predicted masses from Monoisotopic calculator and using Mascot from Matrix Science. MALDI-TOF samples were mixed with 1-cyano-4-hydroxycinnamic acid (10 mg mL^{-1} in 50:50 water: methanol with 0.1% trifluoroacetic acid) and analyzed using a Micromass MALDI micro MX (Waters) using an automated targeting protocol.

Construction of Phylogenetic Trees. Alignments were done based on ClustalW2 output results. The precursor peptide sequences shown in Supporting Information Figure S19 along with the known/predicted N-terminal protease recognition sequences RSII were obtained from previously published studies[7,16,17,21,23,24,26,41,42] and our present analysis of the *thc* pathway. MEGA5.1 was used to construct the Maximum Likelihood phylogenetic tree using the JTT model.

Expression of Hybrid Precursor Peptides in *E. coli*. This was based on our previously described *tru* pathway-based expression system.[9] A two-plasmid system was coexpressed in *E. coli*, one plasmid carrying the precursors for expression of patellins 2 and 3, and the other carrying the engineered hybrid precursor. The hybrids created for this study were obtained from GenScript before subcloning into our previously described pRSF-lac expression vector. Single colonies were picked to make seed cultures, and at least five seed cultures were pooled the next day and inoculated into optimized 2xYT medium supplemented with antibiotics. Cultures were grown for 5 days, after which the cells were harvested. Their acetone extract were isolated, dried, redissolved in methanol, and analyzed by LC/ESI-MS using an Agilent Eclipse C_{18} column (4.6 mm × 150 mm, 5 μm) on a Waters Micromass ZQ mass spectrometer. Robust expression of patellins 2 and 3 served as internal controls.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details and additional data. This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ews1@utah.edu.

Present Address

[†]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

Author Contributions

All authors have given approval to the final version of the manuscript.

Funding

This work was funded by National Institutes of Health grant GM102602.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Chad Nelson and Krishna Parsawar from the University of Utah Mass Spectrometry Core Facility for MS assistance and Scott Endicott from the Peptide Synthesis Core Facility for synthesis of peptide substrates.

■ ABBREVIATIONS

RS, recognition sequence; PTM, posttranslational modification

■ 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., Göransson, 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., Müller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J. T., Ross, R. P., Sahl, H.-G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T. H., Süssmuth, 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., and van der Donk, W. A (2013) Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat. Prod. Rep. 30*, 108−160.

(2) Dunbar, K. L., and Mitchell, D. A. (2013) Revealing nature's synthetic potential through the study of ribosomal natural product biosynthesis. *ACS Chem. Biol.*, 473−487.

(3) Schmidt, E. W. (2012) Decoding and recoding the ribosomal peptide universe. *Chem. Biol. 19*, 1501−1502.

(4) Oman, T., and van der Donk, W. A. (2010) Follow the leader: The use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol. 6*, 9−18.

(5) Olivera, B. M., Hillyard, D. R., Marsh, M., and Yoshikami, D. (1995) Combinatorial peptide libraries in drug design: Lessons from venomous cone snails. *Trends in Biotechnol. 13*, 422−426.

(6) Widdick, D. A., Dodd, H. M., Barraille, P., White, J., Stein, T. H., Chater, K. F., Gasson, M. J., and Bibb, M. J. (2003) Cloning and engineering of the cinnamycin biosynthetic gene cluster from *Streptomyces cinnamoneus cinnamoneus* DSM 40005. *Proc. Natl. Acad. Sci. U.S.A. 100*, 4316−4321.

(7) Donia, M. S., Hathaway, B. J., Sudek, S., Haygood, M. G., Rosovitz, M. J., Ravel, J., and Schmidt, E. W. (2006) Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. *Nat. Chem. Biol. 2*, 729−735.

(8) Shi, Y., Yang, X., Garg, N., and van der Donk, W. A. (2011) Production of lantipeptides in *Escherichia coli*. *J. Am. Chem. Soc. 133*, 2338−2341.

(9) Tianero, M. D. B., Donia, M. S., Young, T. S., Schultz, P. G., and Schmidt, E. W. (2011) Ribosomal route to small-molecule diversity. *J. Am. Chem. Soc.*, 418−425.

(10) Li, C., Zhangy, F., and Kelly, W. L. (2011) Heterologous production of thiostrepton A and biosynthetic engineering of thiostrepton analogs. *Mol. BioSyst. 7*, 82−90.

(11) Pan, S. J., and Link, A. J. (2011) Sequence diversity in the lasso peptide framework: Discovery of functional microcin J25 variants with multiple amino acid substitutions. *J. Am. Chem. Soc. 133*, 5016−5023.

(12) Schmidt, E. W., Nelson, J. T., Rasko, D. A., Sudek, S., Eisen, J. A., Haygood, M. G., and Ravel, J. (2005) Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. *Proc. Natl. Acad. Sci. U.S.A. 102*, 7315−7320.

(13) Lee, J., McIntosh, J. A., Hathaway, B. J., and Schmidt, E. W. (2009) Using marine natural products to discover a protease that catalyzes peptide macrocyclization of diverse substrates. *J. Am. Chem. Soc. 131*, 2122−2124.

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

(15) Donia, M. S., Fricke, W. F., Partensky, F., Cox, J., Elshahawi, S. I., White, J. R., Phillippy, A. M., Schatz, M. C., Piel, J., Haygood, M. G., Ravel, J., and Schmidt, E. W. (2011) Complex microbiome underlying secondary and primary metabolism in the tunicate−prochloron symbiosis. *Proc. Natl. Acad. Sci. U.S.A. 108*, 1423−1432.

(16) McIntosh, J. A., Donia, M. S., and Schmidt, E. W. (2010) Insights into heterocyclization from two highly similar enzymes. *J. Am. Chem. Soc. 132*, 4089−4091.

(17) McIntosh, J. A., Robertson, C. R., Agarwal, V., Nair, S. K., Bulaj, G. W., and Schmidt, E. W. (2010) Circular logic: Nonribosomal peptide-like macrocyclization with a ribosomal peptide catalyst. *J. Am. Chem. Soc. 132*, 15499−15501.

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

(19) Leikoski, N., Liu, L., Jokela, J., Wahlsten, M., Gugger, M., Calteau, A., Permi, P., Kerfeld, C. A., Sivonen, K., and Fewer, D. P. (2013) Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides. *Chem. Biol. 20*, 1033−1043.

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

(21) McIntosh, J. A., Lin, Z., Tianero, M. D. B., and Schmidt, E. W. (2013) Aestuaramides, a natural library of cyanobactin cyclic peptides resulting from isoprene-derived Claisen rearrangements. *ACS Chem. Biol. 8*, 887−883.
(22) Houssen, W. E., Koehnke, J., Zollman, D., Vendome, J., Raab, A., Smith, M. C. M., Naismith, J. H., and Jaspars, M. (2012) The discovery of new cyanobactins from *Cyanothece* PCC 7425 defines a new signature for processing of patellamides. *ChemBioChem 13*, 2683−2689.
(23) Ghilarov, D., Serebryakova, M., Shkundina, I., and Severinov, K. J. (2011) A major portion of DNA gyrase inhibitor microcin B17 undergoes an N,O-peptidyl shift during synthesis. *J. Biol. Chem. 286*, 26308−26318.
(24) Melby, J. O., Dunbar, K. L., Trinh, N. Q., and Mitchell, D. A. (2012) Selectivity, directionality, and promiscuity in peptide processing from a *Bacillus* sp. Al Hakam cyclodehydratase. *J. Am. Chem. Soc. 134*, 5309−5316.
(25) Igarashi, Y., Kan, Y., Fujii, K., Fujita, T., Harada, K., Naoki, H., Tabata, H., Onaka, H., and Furumai, T. (2001) Goadsporin, a chemical substance which promotes secondary metabolism and morphogenesis in *Streptomycetes*. *J. Antibiot. 54*, 1045−1053.
(26) Leikoski, N., Fewer, D. P., Jokela, J., Alakoski, P., Wahlsten, M., and Sivonen, K. (2012) Analysis of an inactive cyanobactin biosynthetic gene cluster leads to discovery of new natural products from strains of the genus *Microcystis*. *PLoS One 7*, 1−9.
(27) Huo, L., Rachid, S., Stadler, M., Wenzel, S. C., and Muller, R. (2012) Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem. Biol. 19*, 1278−1287.
(28) Hou, Y., Tianero, M. D. B., Kwan, J. C., Wyche, T. P., Michel, C. R., Ellis, G. A., Vazquez-Rivera, E., Braun, D. R., Rose, W. E., Schmidt, E. W., and Bugni, T. S. (2012) Structure and biosynthesis of the antibiotic bottromycin D. *Org. Lett. 14*, 5050−5053.
(29) Gomez-Escribano, J. P., Song, L., Bibb, M. J., and Challis, G. L. (2012) Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chem. Sci. 3*, 3522−3525.
(30) Jian, X., Pan, H., Ning, T., Shi, Y., Chen, Y., Li, Y., Zeng, X., Xu, J., and Tang, G. (2012) Analysis of YM-216391 biosynthetic gene cluster and improvement of the cyclopeptide production in a heterologous host. *ACS Chem. Biol. 7*, 646−651.
(31) Crooks, G. E., Hon, G., Chandonia, C., and Brenner, S. E. (2004) WebLogo: A sequence logo generator. *Genome Res. 14*, 1188−1190.
(32) Roy, R. S., Kim, S., Baleja, J. D., and Walsh, C. T. (1998) Role of the microcin B17 propeptide in substrate recognitibn: Solution structure and mutational analysis of McbA1−26. *Chem. Biol. 5*, 217−228.
(33) Koehnke, J., Bent, A., Houssen, W. E., Zollman, D., Morawitz, F., Shirran, S., Vendome, J., Nneoyiegbe, A. F., Trembleau, L., Botting, C. H., Smith, M. C., Jaspars, M., and Naismith, J. H. (2012) The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nat. Struct. Mol. Biol. 19*, 767−772.
(34) Agarwal, V., Pierce, E., McIntosh, J., Schmidt, E. W., and Nair, S. K. (2012) Structures of cyanobactin maturation enzymes define a family of transamidating proteases. *Chem. Biol. 19*, 1411−1422.
(35) 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., and Naismith, J. H. (2013) The cyanobactin heterocyclase enzyme: A processive adenylase that operates with a defined order of reaction. *Angew. Chem. 52*, 13991−13996.
(36) Zhang, Q., Yu, Y., Velasquez, J. E., and van der Donk, W. A. (2012) Evolution of lanthipeptide synthetases. *Proc. Natl. Acad. Sci. U.S.A. 109*, 18361−18366.
(37) Oman, T. J., Knerr, P. J., Bindman, N. A., Velasquez, J. E., and van der Donk, W. A. (2012) An engineered lantibiotic synthetase that does not require a leader peptide on its substrate. *J. Am. Chem. Soc. 134*, 6952−6955.
(38) Dunbar, K. L., and Mitchell, D. A. (2013) Insights into the mechanism of peptide cyclodehydrations achieved through the chemoenzymatic generation of amide derivatives. *J. Am. Chem. Soc. 135*, 8692−8701.
(39) Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., and van der Donk, W. A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A. 107*, 10430−10435.
(40) Milne, J. C., Eliot, A. C., Kelleher, N. L., and Walsh, C. T. (1998) ATP/GTP hydrolysis is required for oxazole and thiazole biosynthesis in the peptide antibiotic microcin B17. *Biochemistry 37*, 13250−13261.
(41) Sudek, S., Haygood, M. G., Youssef, D. T. A., and Schmidt, E. W. (2006) Structure of trichamide, a cyclic peptide from the bloom-forming cyanobacterium *Trichodesmium erythraeum*, predicted from the genome sequence. *Appl. Environ. Microbiol. 72*, 4382−4387.
(42) Ziemert, N., Ishida, K., Quillardet, P., Bouchier, C., Hertweck, C., de Marsac, N. T., and Dittmann, E. (2008) Microcyclamide biosynthesis in two strains of *Microcystis aeruginosa*: from structure to genes and *vice versa*. *Appl. Environ. Microbiol. 74*, 1791−1797.