Organic & Biomolecular Chemistry

Probing the limits of interrupted adenylation domains by engineering a trifunctional enzyme capable of adenylation, N-, and S-methylation

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-ART-12-2018-002996.R1
Article Type:	Paper
Date Submitted by the Author:	02-Jan-2019
Complete List of Authors:	Lundy, Taylor; University of Kentucky College of Pharmacy, Pharmaceutical Sciences Mori, Shogo; University of Kentucky, Pharmaceutical Sciences Garneau-Tsodikova, Sylvie; University of Kentucky, Pharmaceutical Sciences

Probing the limits of interrupted adenylation domains by engineering a trifunctional enzyme capable of adenylation, *N*-, and *S*-methylation

Taylor A. Lundy,[a] Shogo Mori,[a] and Sylvie Garneau-Tsodikova[a,*]

[a]University of Kentucky, Department of Pharmaceutical Sciences, College of Pharmacy, Lexington, KY 40536-0596, USA. *sylviegtsodikova@uky.edu

KEYWORDS: Enzyme engineering, Methyltransferase, Natural products, Nonribosomal peptide biosynthesis, Thiocoraline.

ABSTRACT

The adenylation (A) domains found in nonribosomal peptide synthetases (NRPSs) exhibit tremendous plasticity. Some A domains have been shown to display the ability to contain within them the catalytic portion of an auxiliary domain, most commonly that of a methyltransferase (M) enzyme. This unique feature of A domains interrupted by M domains allows them to possess bifunctionality, where they can both adenylate and methylate an amino acid substrate. Additionally, these types of inserted M domains are able to selectively carry out either backbone or side chain methylation of amino acids. Interruptions with M domains are naturally found to occur either between the a2-a3 or the a8-a9 of the ten conserved motifs of A domains. Herein, we set out to answer the following question: Can one A domain support two different M domain interruptions occurring in two different locations (a2-a3 and a8-a9) of the A domain and possess the ability to adenylate an amino acid and methylate it on both its side chain and backbone? To answer this question we added a backbone methylating M_{3S} domain from TioS($A_{3a}M_{3S}A_{3b}$) between the a8-a9 region of a mono-interrupted A domain, TioN($A_aM_NA_b$), that already contained a side chain methylating M_N domain between its a2-a3 region. We evaluated the di-interrupted A domain TioN($AM_NAM_{3S}A$) with a series of radiometric and mass spectrometry assays and found that this engineered enzyme was indeed capable of all three activities. These findings show that production of an active trifunctional di-interrupted A domain is possible and represent an exciting new avenue for future nonribosomal peptide (NRP) derivatization.

INTRODUCTION

Nonribosomal peptides (NRPs) are a class of natural products (NPs) that have historically served as inspiration for many therapeutic compounds such as penicillins, vancomycin, and bleomycin.[1] NRPs are biosynthesized by a class of mega-enzymes called nonribosomal peptide synthetases (NRPSs), which work in a linear assembly-line fashion to assemble amino acids or amino acid derivatives into final NRPs. While these NRPs are rich in chemical diversity and biological activity, they often require modifications of the original compound before they can be used as therapeutics.[2] The greatest strengths of NPs, which are also their greatest barriers to further progression through the drug development process include: high structural complexity, unique moieties, abundant stereocenters, and complicated rings, all of which make synthetic modifications difficult and lengthy.[3] Therefore, modification of these NPs during their biosynthesis by enzyme engineering is an alternative approach that could offer targeted site-specific chemical modifications of NPs and thus, enzyme engineering for this purpose warrants substantial investigation.

NRPSs are modular mega-enzymes that can be subdivided into individual catalytic domains that communicate with one another in a synchronized and coordinated fashion.[4] Three core domains: adenylation (A), condensation (C), and thiolation (T) constitute a module, with each domain being responsible for performing a particular chemistry. In short, an A domain activates an amino acid *via* adenylation and loads it onto an active (holo) T domain. The T domain, also referred to as a peptidyl carrier protein (PCP) or simply a carrier protein (CP), is converted into its holo form through the activity of a 4'-phosphopantetheinyltransferase (PPTase), which adds the 4'-phosphopantetheine (Ppant) prosthetic group from coenzyme A (CoA) onto the inactive (apo) T

domain.[5] T domain bound amino acids are covalently linked together through the action of the C domain. C domains can also be dual functional, carrying out condensation and cyclization or epimerization.[6] Auxiliary domains can further modify the NRP with additional chemistries such as epimerization, methylation, and oxidation/reduction. Methylation domains are particularly important in NPs, especially for NRPs as *N*-methylation of peptides can improve oral bioavailability.[7] Additionally, although not in NRPs, di-methylating the nitrogen of aminoglycosides has been shown to improve their potency.[8] Typically, methylated NRPs contain amino acids that have one type of methylation, *N*-, *O*-, or *S*-methylation. However, there are several examples of NPs with amino acids that have been di-methylated resulting in *N*,*O*-dimethyl (*e.g.*, nordolastatin G,[9] didemnin B,[10] and columbamide D[11]), *N*,*S*-dimethyl (*e.g.*, thiochondrilline A,[12] thiocoraline,[13] and echinomycin[14]), and *N*,*N*-dimethyl (*e.g.*, dolastatin 10[15] and symplocin A[16]) amino acid moieties incorporated into the final NP structure (Fig. S1). With the exception of thiochondrilline A/thiocoraline biosynthesis, where *N*- then *S*-methylation by two different interrupted A domains work sequentially on L-Cys covalently tethered to a T domain,[17] the natural di-methylation process has not been studied intensively and the order of methylation not determined.

Since A domains are charged with the task of selecting and activating the amino acids, they dictate the final structure of the NRP and, as a result, have attracted significant attention for bioengineering.[18, 19] A particularly fascinating aspect of A domains is that they possess the ability to contain within them part of an auxiliary domain,[20] most commonly that of a methyltransferase (M) enzyme. These interrupted A domains possess bifunctionality and are able to carry out both adenylation and the activity of the auxiliary domain embedded within them (*e.g.*, methylation for

an M domain). A domains have been found to be interrupted in a few different locations based on the ten conserved amino acid sequence motifs (a1-a10) of A domains.[20, 21] Interestingly, the biosynthetic pathway of the bisintercalator NRP thiocoraline contains representatives of two types of interrupted A domains (Fig. S2). At the time this study was initiated and carried out, there was only one known and characterized interrupted A domain with the interruption occurring between the a2 and a3 motifs.[22] This interrupted A domain in thiocoraline biosynthesis, TioN($A_aM_NA_b$), was shown to activate L-Cys and carry out *S*-methylation (Fig. 1).[17, 22] (Note: when naming the interrupted A domains, we designate the A domain fragments on either side of the interruption as "a" for the first fragment and "b" for the second. Any additional letters or numbers represent the module identity.) Since then, a recent report characterized another interrupted A domain between a2 and a3 that also *S*-methylates L-Cys.[23] Additionally, it was recently reported that interrupted A domains must first activate the amino acid, then load it on the T domain, then methylate it.[17] By far, the most common location for an interruption to occur is between the a8 and a9 conserved motifs, indicating that this could be a favored location for *de novo* insertion of auxiliary domains. Based on a recently published crystal structure of an interrupted A domain from the thiocoraline biosynthetic pathway, TioS($A_{4a}M_{4S}A_{4b}$),[24] this is likely due to the fact that the interruption does not disrupt the normal folding of the A domain. Additionally, the active sites of the A and M domains are in close enough proximity to allow for the Ppant arm of the T domain to easily shuttle between the two active sites.[24]

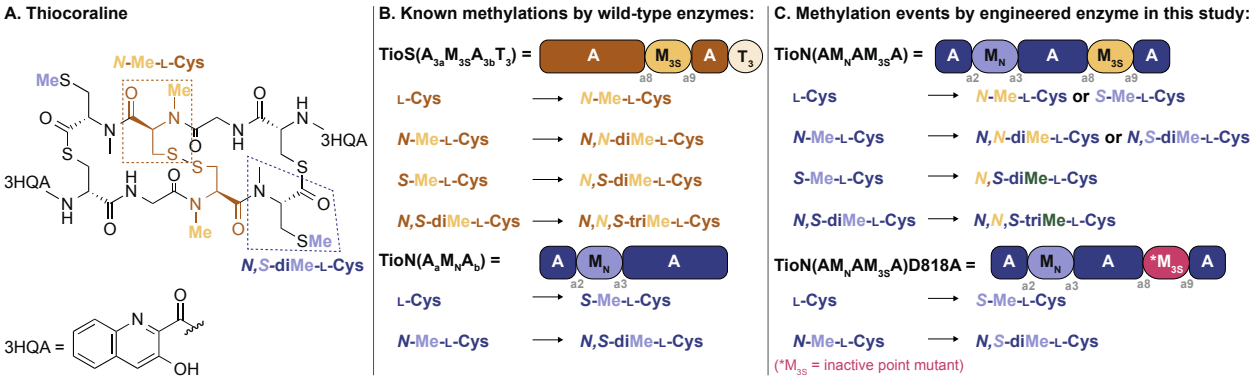

Fig. 1: Overview of the enzymes used in this study. **A.** Thiocoraline structure with some methyl groups incorporated by interrupted A domains used herein. **B.** The wild-type (wt) enzymes and their known methylation activity. **C.** The engineered di-interrupted A domains designed and utilized in this study. *Note*: The methyl (Me) are colored based on the methylation (M) domain that is responsible for their attachment to L-Cys. In the cases where the methylations could result from the action of more than one M domain (*e.g.*, *N*,*S*-diMe-L-Cys and *N*,*N*,*S*-triMe-L-Cys by the action of TioN($AM_NAM_{3S}A$)), we colored the Me in green.

It was recently demonstrated that interrupted A domains have great tolerance to genetic engineering.[25, 26] One study looked at two naturally occurring unrelated interrupted A domains, KtzH($A_{4a}M_HA_{4b}$) from the kutznerides biosynthetic pathway, and TioN($A_aM_NA_b$) from the thiocoraline biosynthetic pathway.[25] Although both of these interrupted A domains have M domain insertions that methylate the side chain of either L-Ser or L-Cys, respectively, they only share 22% amino acid sequence similarity. The M_H domain of KtzH($A_{4a}M_HA_{4b}$) was removed to yield KtzH($A_{4a}\Delta M_HA_{4b}$), and the A domain activity of this deletion mutant was found to be intact. Then, the M_N domain from TioN($A_aM_NA_b$) was used to re-interrupt KtzH($A_{4a}\Delta M_HA_{4b}$) to generate KtzH($A_{4a}M_NA_{4b}$), which demonstrated an increased substrate promiscuity and the ability to methylate L-Ser. Another study showed that it was possible to create fully functional interrupted A domains from uninterrupted A domains.[26] The intact Ecm6(A_1T_1) from the echinomycin biosynthetic pathway was interrupted with either the *O*-methylating M_H domain from KtzH($A_{4a}M_HA_{4b}$)[27] or the *N*-methylating M_{3S} domain from TioS($A_{3a}M_{3S}A_{3b}$)[17] of the thiocoraline biosynthetic pathway, between the a8 and a9 conserved motifs. The engineered

Ecm6($A_{1a}M_HA_{1b}T_1$) and Ecm6($A_{1a}M_{3S}A_{1b}T_1$) retained both the substrate specificity of the A domain and the regiospecificity of the respective M domains. Thus, the a8-a9 region of A domains is a promising location for engineering interrupted A domains.

Based on the above observations that (i) interruptions can occur in two different locations within A domains, (ii) insertion of auxiliary domains can be done between the a8 and a9 region of A domains, and (iii) the A and M domains used in engineering bifunctional mono-interrupted A domains retain the same specificity as their wild-type (wt) counterparts; in this study, we hypothesized that we could construct a di-interrupted A domain containing two different M domains in two different regions of the A domain to generated the first NRPS trifunctional enzyme. To verify our hypothesis, we utilized the naturally occurring mono-interrupted A domain with an a2-a3 interruption, TioN($A_aM_NA_b$), that was previously shown to perform *S*-methylation of L-Cys.[22] We added in between the a8 and a9 motifs, the M_{3S} domain that was characterized to carry out *N*- and *N*,*N*-dimethylation of L-Cys from another mono-interrupted A domain in thiocoraline biosynthesis, TioS($A_{3a}M_{3S}A_{3b}$).[17] By combining these two M domains into one di-interrupted A domain, we set out to create a trifunctional protein capable of adenylating, *N*-methylating, and *S*-methylating L-Cys.

RESULTS AND DISCUSSIONS

Engineering, cloning, heterologous co-overexpression, and co-purification of TioN($AM_NAM_{3S}A$), TioN($AM_NAM_{3S}A$)D818A, and TioN($A_aM_NA_b$)D167A with their MbtH-like protein (MLP) partner TioT.

In order to investigate the hypothesis that one A domain can support two interruptions, we cloned three new proteins for this study: (i) a di-interrupted TioN($AM_NAM_{3S}A$) with two functional M domains; (ii) a di-interrupted TioN($AM_NAM_{3S}A$)D818A point mutant rendering the newly inserted M_{3S} domain inactive to verify that the methylating activity observed was in fact from both methylation domains being active and not just the work of one; and (iii) TioN($A_aM_NA_b$)D167A, a point mutant in the wt mono-interrupted A domain (TioN($A_aM_NA_b$)) to adenylate and load *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys during the methylation assays of those substrates. To generate the point mutants, we used a standard single overlap extension (SOE) method.[28] For the di-interrupted A domain, we determined the insertion site of the second M_{3S} domain between the a8 and a9 motifs by using the crystal structure of a naturally occurring mono-interrupted A domain (TioS($A_{4a}M_{4S}A_{4b}$))[24] as well as amino acid sequence alignments with a previously engineered mono-interrupted A domain (Ecm6($A_{1a}M_{3S}A_{1b}$))[26] (Fig. S3). The insertion point selected was located in a loop region of the A domain according to the structure of TioS($A_{4a}M_{4S}A_{4b}$). We utilized an approach similar to a previously reported method using a series of PCRs and sequential ligations (Fig. S4).[26] Since 2007,[29] MLPs have been shown to play a variety of critical roles in the function of A domains. By structural and biochemical investigations, it was determined that MLPs could influence the expression of A domains as soluble and/or active protein by binding to a specific region of these enzymes,[30-34] alteration/expansion of their substrate profiles and kinetic activity,[33, 35] folding chaperones,[36, 37] etc. Since wt TioN($A_aM_NA_b$) was known to require the MLP TioT for expression of active protein,[22] all TioN($A_aM_NA_b$) derivatives were co-overexpressed and co-purified with TioT in *E. coli* BL21 (DE3) ybdZ::aac(3)IV, with similar yields of purified protein when compared to that reported for the wt TioN($A_aM_NA_b$).[17, 22]

Adenylating activity of TioN($AM_NAM_{3S}A$), TioN($AM_NAM_{3S}A$)D818A, and TioN($A_aM_NA_b$)D167A.

To verify that the adenylating activity remained intact after the addition of a second M domain, M_{3S}, between a8-a9 of TioN($A_aM_NA_b$), we determined the substrate profile of the engineered di-interrupted TioN($AM_NAM_{3S}A$) by using L-Cys, three of its methylated derivatives, as well as the standard 20 amino acids (Fig. 2A). We found that the substrate profile for this di-interrupted A domain displayed a similar trend when compared to that of the wt TioN($A_aM_NA_b$) (Figs. 2C and S5A) with L-Cys and *N*-Me-L-Cys being favored (100% and 93% for TioN($AM_NAM_{3S}A$) and 100% and 102% for wt TioN($A_aM_NA_b$), respectively) over *S*-Me-L-Cys (59% for wt TioN($A_aM_NA_b$); 42% for TioN($AM_NAM_{3S}A$)). The same trend was observed for TioN($AM_NAM_{3S}A$)D818A (Fig. 2B). However, TioN($AM_NAM_{3S}A$)D818A showed lower percent activity for *N*-Me-L-Cys (57%) and *S*-Me-L-Cys (16%) when compared to that observed for the wt TioN($A_aM_NA_b$) and for TioN($AM_NAM_{3S}A$). These data are in agreement with the engineering of the previously reported mono-interrupted Ecm6($A_{1a}M_HA_{1b}T_1$) and Ecm6($A_{1a}M_{3S}A_{1b}T_1$), where inserting a foreign M domain into an uninterrupted A domain did not dramatically alter the substrate profile.[26] Since *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys were not as good of substrates for the di-interrupted A domains (TioN($AM_NAM_{3S}A$) and its D818A point mutant), according to the substrate profiles, we decided to make a point mutant of the mono-interrupted wt TioN($A_aM_NA_b$) as a tool to reduce background noise in the methylation assays described in the next section. We first confirmed that this M_N domain mutant in which the aspartate at residue 167 was mutated to alanine (D167A) retained *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys adenylating activity (Figs. S5 and S6). Then, we established that, as desired, this point mutant rendered the M_N domain inactive (Fig. S8).

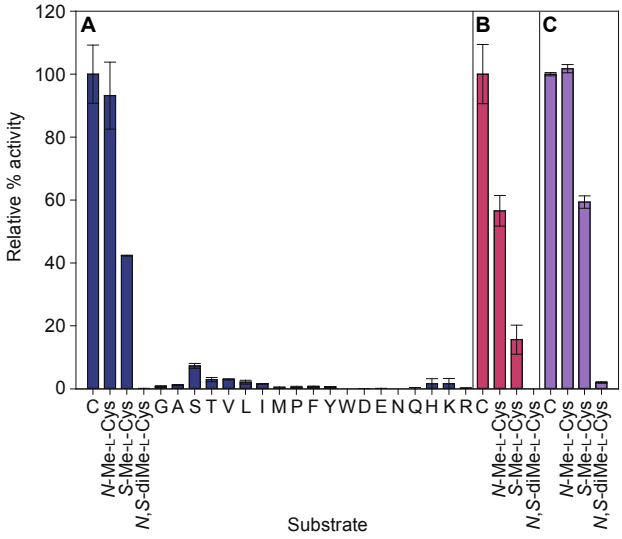

Fig. 2: Substrate profile represented as relative percent activity compared to L-Cys of the di-interrupted **A.** TioN($AM_NAM_{3S}A$) (blue bars), **B.** TioN($AM_NAM_{3S}A$)D818A (pink bars), and **C.** TioN($A_aM_NA_b$) (purple bars) determined *via* ATP-[^{32}P]PP_i exchange assays at a 2-h end point. *Note*: These data are also presented as part of Fig. S5 for comparison purpose.

When we investigated the kinetic parameters of TioN($AM_NAM_{3S}A$) (Table 1 and Fig. S7A-C), the catalytic efficiencies (k_{cat}/K_m) were reduced compared to those of the naturally occurring TioN($A_aM_NA_b$) by 93-fold, 1.6-fold, and 4.8-fold for L-Cys, *N*-Me-L-Cys, and *S*-Me-L-Cys, respectively. A similar trend was observed with TioN($AM_NAM_{3S}A$)D818A (Table 1 and Fig. S7D-F) displaying an 18.8-fold, 7.0-fold, and 3.4-fold decrease in catalytic efficiencies for L-Cys, *N*-Me-L-Cys, and *S*-Me-L-Cys, respectively. However, the k_{cat}/K_m values of the di-interrupted TioN($AM_NAM_{3S}A$) were more in line with those reported for other engineered A domains such as Ecm6($A_{1a}M_HA_{1b}T_1$) (0.43 $mM^{-1}min^{-1}$) and Ecm6($A_{1a}M_{3S}A_{1b}T_1$) (3.7 $mM^{-1}min^{-1}$)[26] and other natural interrupted A domains such as KtzH($A_{4a}M_HA_{4b}$) (0.77 $mM^{-1}min^{-1}$).[25, 26]

Table 1. Steady-state kinetic parameters for AMP derivatization by the di-interrupted A domains engineered in this study.

Protein	Substrate	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m ($mM^{-1}min^{-1}$)
TioN($AM_NAM_{3S}A$)	L-Cys	6.9 ± 1.0	8.6 ± 0.6	1.2 ± 0.2
	N-Me-L-Cys	5.6 ± 0.5	17.9 ± 0.7	3.2 ± 0.3
	S-Me-L-Cys	32 ± 20	7.0 ± 3.2	0.22 ± 0.17
TioN($AM_NAM_{3S}A$)D818A	L-Cys	0.86 ± 0.17	5.0 ± 0.3	5.9 ± 1.2

	N-Me-L-Cys	9.2 ± 0.9	6.5 ± 0.3	0.71 ± 0.08
	S-Me-L-Cys	15 ± 5	4.6 ± 1.0	0.31 ± 0.13
TioN($A_aM_NA_b$)[a]	L-Cys	0.037 ± 0.001	4.13 ± 0.02	111 ± 4
	N-Me-L-Cys	1.1 ± 0.2	5.3 ± 0.3	5.0 ± 0.9
	S-Me-L-Cys	3.5 ± 0.1	3.72 ± 0.05	1.06 ± 0.04

[a] The kinetic parameters for the wt TioN($A_aM_NA_b$) were previously reported in ref [17] and are presented here only for ease of comparison.

Methylation activity of TioN($AM_NAM_{3S}A$) and TioN($AM_NAM_{3S}A$)D818A.

Given that the di-interrupted A domains were still capable of adenylation, we next interrogated their methylation ability. We needed to address two main concerns with regard to the engineered di-interrupted A domain TioN($AM_NAM_{3S}A$): (i) could methylation of an amino acid bound to the T domain still be done by M_N or was the addition of the second M domain, M_{3S}, detrimental to the function of M_N; and (ii) did each M domain present in the di-interrupted A domain retain their original methylation pattern compared to the wt versions of these M domains (*i.e.*, side chain methylation for M_N and backbone methylation for M_{3S}). Since both M domains were previously characterized to methylate L-Cys and *N*-Me-L-Cys, we utilized a radioactive trichloroacetic acid (TCA) precipitation assay to monitor methylation events of these amino acids over time by TioN($AM_NAM_{3S}A$) (Fig. 3A). We found TioN($AM_NAM_{3S}A$) capable of methylating both L-Cys and *N*-Me-L-Cys, indicating that one if not both of the M domains were active. Next, to decipher if the methylation activity observed for TioN($AM_NAM_{3S}A$) with L-Cys and *N*-Me-L-Cys was a result of one or both M domains, we evaluated the same substrates with TioN($AM_NAM_{3S}A$)D818A in which M_{3S} is inactive (Fig. 3B). The observed methylation indicated that the addition of the M_{3S} between a8-a9 did not disrupt the *S*-methylation activity of M_N. It is important to note that while the methylation observed for L-Cys and *N*-Me-L-Cys with both di-interrupted A domains could be a result of just the native M_N domain, TioN($AM_NAM_{3S}A$) showed more cpm than TioN($AM_NAM_{3S}A$)D818A, likely because there was more [^{3}H] on the substrate bound T domain as presumably, there were at least two [^{3}H]methyl groups attached to it.

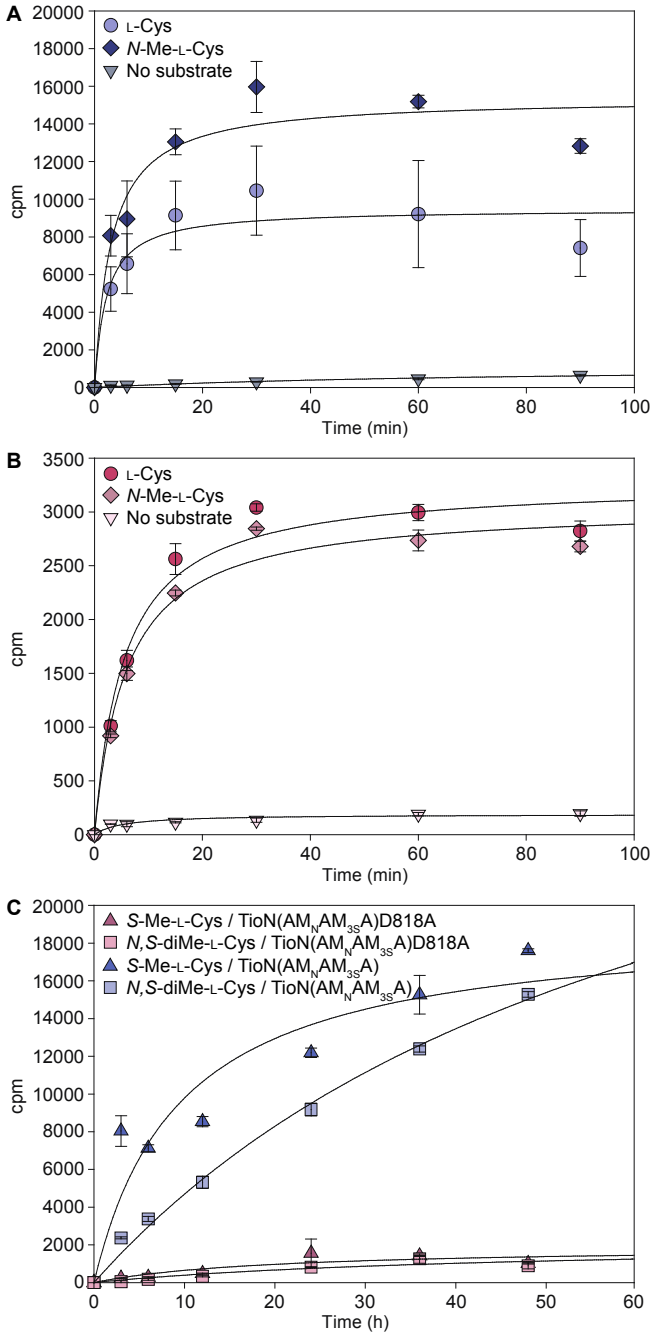

Fig. 3: Methylation of L-Cys (circles), *N*-Me-L-Cys (diamonds), *S*-Me-L-Cys (triangles), and *N*,*S*-diMe-L-Cys (squares) measured by TCA precipitation assays with [methyl-^{3}H]SAM as the methylating agent over time by **A.** TioN(AM$_N$AM$_{3S}$A), **B.** TioN(AM$_N$AM$_{3S}$A)D818A, and **C.** TioN(AM$_N$AM$_{3S}$A) and TioN(AM$_N$AM$_{3S}$A)D818A. Control experiments (no substrate) are presented with inverted triangles in panels **A** and **B**. Experiments with TioN(AM$_N$AM$_{3S}$A) and TioN(AM$_N$AM$_{3S}$A)D818A are presented with different shades of blue and pink, respectively.

To determine if TioN(AM$_N$AM$_{3S}$A) possessed new methylation capability compared to the wt TioN(A$_a$M$_N$A$_b$), we utilized the same radioactive TCA precipitation assays as above with *S*-Me-

L-Cys and *N*,*S*-diMe-L-Cys, because M_N is only known to perform *S*-methylation, which is blocked in these substrates (Fig. 3C). The di-interrupted A domain TioN($AM_NAM_{3S}A$) methylated *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys, indicating that either (i) the inserted M_{3S} retained the same *N*-methylation specificity as that of the M_{3S} domain in the wt TioS($A_{3a}M_{3S}A_{3b}$) or (ii) the addition of M_{3S} to TioN($A_aM_NA_b$) resulted in M_N gaining new *N*-methylation abilities. To assess which of the aforementioned possibilities was true, we performed the same methylation assay of *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys with TioN($AM_NAM_{3S}A$)D818A in which M_{3S} is inactive (Fig. 3C). As TioN($AM_NAM_{3S}A$)D818A did not methylate *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys, we concluded that the *N*-methylation of *S*-Me-L-Cys and *N*,*S*-diMe-L-Cys observed with TioN($AM_NAM_{3S}A$) is a result of the newly inserted M_{3S} domain. It is also important to note that the methylation of *N*,*S*-diMe-L-Cys by TioN($AM_NAM_{3S}A$) pointed to a second *N*-methylation on the nitrogen backbone of this amino acid to generate *N*,*N*,*S*-triMe-L-Cys, which is consistent with what we previously reported for wt TioS($A_{3a}M_{3S}A_{3b}$)[17] and engineered Ecm6($A_{1a}M_{3S}A_{1b}T_1$).[26]

Mass spectrometry (MS) and tandem mass spectrometry (MS^2) evaluation of *N*,*S*-diMe-L-Cys.

All of the above radiometric evaluations for the di-interrupted A domain clearly indicated that the di-interrupted A domain was indeed functional. We can infer from these assays that generation of the di-interrupted A domain yielded an active protein capable of adenylation, *N*-methylation, and *S*-methylation. To confirm the results obtained by radioactive assays that *N*,*S*-diMe-L-Cys was formed enzymatically, we carried out MS analysis of a methylation reaction using the di-interrupted A domain, TioN($AM_NAM_{3S}A$), L-Cys, and TioS(T_3). The product bound to the T domain was then liberated, extracted, and subjected to MS and MS^2 analysis *via* direct sample

injection. The extracted sample was then compared to synthetic (L-Cys, *N*-Me-L-Cys, *S*-Me-L-Cys) and commercially available (*N*,*S*-diMe-L-Cys) standards (Figs. S9-10). The *m*/*z* peak at 150.0570 for the extract from the enzymatic reaction matched that of the synthetic standard *N*,*S*-diMe-L-Cys, which showed a *m*/*z* peak at 150.0566 (calculated $[M+H]^+$ 150.0583). Next, the peaks with *m*/*z* 150.0566 and 150.0570 for both the standard and the extract, respectively, were subjected to MS^2. Fragments with *m*/*z* ratios between 20 to 300 were collected and compared (Fig. S10). We observed that the peak at 150.0570 in the extract sample contained the same major fragmentation peaks as the *N*,*S*-diMe-L-Cys standard at the same voltage indicating *N*,*S*-diMe-L-Cys was indeed produced in the enzymatic reaction.

CONCLUSION

In summary, we have successfully created the first di-interrupted A domain by inserting a backbone *N*-methylating M domain between the a8 and a9 conserved motifs of a mono-interrupted A domain containing a side chain *S*-methylating M domain between the a2 and a3 conserved motifs. This di-interrupted A domain was able to carry out three functions, namely adenylation, *N*-methylation, and *S*-methylation. Additionally, we found that the substrate profile of the newly created trifunctional enzyme remained overall the same. We also found that the two M domains of the engineered enzyme were active and maintained their regiospecificty with regard to *S*- or *N*-methylation. With our recent success in interrupting A domains between the a8-a9 conserved motifs, it is plausible that an artificial interruption between a2-a3 would also work. However, a barrier to engineering this is that there is not a structure of a2-a3 interrupted A domain. Studies are currently underway in our laboratory to both engineer a2-a3 interrupted A domains and structurally characterize a natural or engineered a2-a3 interrupted A domain. This current study lays the

groundwork for adding site-specific modifications to NRPS systems by creating multifunctional A domains. However, the question remains if engineered interrupted A domains can interact with downstream condensation domains in assembly-lines, which is something our laboratory is currently investigating. Additionally, current work underway in our lab aims to expand these interrupted A domains to be more promiscuous and incorporate these types of engineered A domains into biosynthetic assembly-lines.

Supporting information

The Supporting Information is available free of charge. Experimental procedures for all molecular cloning, protein co-expression and co-purification, and biochemical assays are provided. Figures showing examples of di-methylated NPs (Fig. S1), the thiocoraline biosynthetic assembly-line (Fig. S2), amino acid sequence alignments (Fig. S3), the cloning strategy used to engineer TioN($AM_NAM_{3S}A$) (Fig. S4), the A domain activity of TioN($A_aM_NA_b$)D167A (Fig. S5), the time course of *N*,*S*-diMe-L-Cys adenylation by wt TioN($A_aM_NA_b$) and TioN($A_aM_NA_b$)D167A (Fig. S6), the Michaelis-Menten kinetics plots for adenylation of amino acids by TioN($AM_NAM_{3S}A$) and TioN($AM_NAM_{3S}A$)D818A (Fig. S7), a time course plot comparing methylation activity by TioN($A_aM_NA_b$) and TioN($A_aM_NA_b$)D167A (Fig. S8), as well as MS (Fig. S9) and MS^2 (Fig. S10) of standards and enzymatic reaction extract are also provided. A table of primers used in this study (Table S1) is presented.

Author contributions

T.A.L and S.G.-T designed the study. T.A.L. and S.M. performed experiments. T.A.L. and S.G.-T. wrote the manuscript and supporting information. S.M. helped with writing of the supporting

information and proof-reading of the manuscript. All the figures were generated by T.A.L. and S.G.-T.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This project was supported by startup funds from the College of Pharmacy at the University of Kentucky (to S.G.-T.) and by a National Science Foundation CAREER Award MCB-1149427 (to S.G.-T.). T.A.L. was in part supported by a 2018-2019 Pharmaceutical Sciences Excellence in Graduate Achievement Fellowship from the College of Pharmacy at the University of Kentucky. S.M. is a recipient of a 2018 long-term visit fellowship from the Yamada Science Foundation, Japan. We thank Dr. Atefeh Garzan for providing the *N*-Me-L-Cys and *N*,*S*-diMe-L-Cys for which the synthesis was previously published in ref [17]. We thank Jarrod W. Creameans who, under the supervision of T.A.L., constructed the p*tioN(AM$_N$AM$_{3S}$A)*-pET28a and p*tioN(AM$_N$AM$_{3S}$A)D818A*-pET28a plasmids and also expressed and purified the TioN(AM$_N$AM$_{3S}$A)D818A enzyme. We also thank Scott Kinison for training of T.A.L. and his help in developing the MS and MS^2 methods.

REFERENCES

1. R. D. Sussmuth and A. Mainz, *Angew. Chem.*, 2017, **56**, 3770-3821.
2. C. T. Walsh, H. Chen, T. A. Keating, B. K. Hubbard, H. C. Losey, L. Luo, C. G. Marshall, D. A. Miller and H. M. Patel, *Curr. Opin. Chem. Biol.*, 2001, **5**, 525-534.
3. H. Sun, Z. Liu, H. Zhao and E. L. Ang, *Drug Des. Devel. Ther.*, 2015, **9**, 823-833.

4. K. Bloudoff and T. M. Schmeing, *Biochim. Biophys. Acta Proteins Proteom.*, 2017, **1865**, 1587-1604.
5. R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla and C. T. Walsh, *Chem. Biol.*, 1996, **3**, 923-936.
6. C. J. Balibar, F. H. Vaillancourt and C. T. Walsh, *Chem. Biol.*, 2005, **12**, 1189-1200.
7. E. Biron, J. Chatterjee, O. Ovadia, D. Langenegger, J. Brueggen, D. Hoyer, H. A. Schmid, R. Jelinek, C. Gilon, A. Hoffman and H. Kessler, *Angew. Chem.*, 2008, **47**, 2595-2599.
8. R. I. Benhamou, P. Shaul, I. M. Herzog and M. Fridman, *Angew. Chem.*, 2015, **54**, 13617-13621.
9. T. Mutou, T. Kondo, M. Ojika and K. Yamada, *J. Org. Chem.*, 1996, **61**, 6340-6345.
10. J. Rinehart, K. L., J. B. Gloer, J. Cook, J. C., S. A. Mizsak and T. A. Scahill, *J. Am. Chem. Soc.*, 1981, **103**, 1857-1859.
11. J. A. V. Lopez, J. G. Petitbois, C. S. Vairappan, T. Umezawa, F. Matsuda and T. Okino, *Org. Lett.*, 2017, **19**, 4231-4234.
12. T. P. Wyche, Y. Hou, D. Braun, H. C. Cohen, M. P. Xiong and T. S. Bugni, *J. Org. Chem.*, 2011, **76**, 6542-6547.
13. F. Romero, F. Espliego, J. Perez Baz, T. Garcia de Quesada, D. Gravalos, F. de la Calle and J. L. Fernandez-Puentes, *J. Antibiot.*, 1997, **50**, 734-737.
14. A. Dell, D. H. Williams, H. R. Morris, G. A. Smith, J. Feeney and G. C. Roberts, *J. Am. Chem. Soc.*, 1975, **97**, 2497-2502.
15. G. R. Pettit, Y. Kamano, C. L. Herald, A. A. Tuinman, F. E. Boettner, H. Kizu, J. M. Schmidt, L. Bacynskyj, K. B. Tomer and R. J. Bontems, *J. Am. Chem. Soc.*, 1987, **109**, 6883-6885.

16. L.-P. Shao, C.-M. Si, Z.-Y. Mao, W. Zhou, T. F. Molinski, B.-G. Wei and G.-Q. Lin, *Org. Chem. Front.*, 2017, **4**, 995-1004.
17. S. Mori, A. Garzan, O. V. Tsodikov and S. Garneau-Tsodikova, *Biochemistry*, 2017, **56**, 6087-6097.
18. K. T. Nguyen, D. Ritz, J. Q. Gu, D. Alexander, M. Chu, V. Miao, P. Brian and R. H. Baltz, *Proc. Natl. Acad. Sci., U. S. A.*, 2006, **103**, 17462-17467.
19. J. Thirlway, R. Lewis, L. Nunns, M. Al Nakeeb, M. Styles, A. W. Struck, C. P. Smith and J. Micklefield, *Angew. Chem.*, 2012, **51**, 7181-7184.
20. K. J. Labby, S. G. Watsula and S. Garneau-Tsodikova, *Nat. Prod. Rep.*, 2015, **32**, 641-653.
21. A. M. Gulick, *ACS Chem. Biol.*, 2009, **4**, 811-827.
22. A. H. Al-Mestarihi, G. Villamizar, J. Fernandez, O. E. Zolova, F. Lombo and S. Garneau-Tsodikova, *J. Am. Chem. Soc.*, 2014, **136**, 17350-17354.
23. T. Huang, Y. Duan, Y. Zou, Z. Deng and S. Lin, *ACS Chem. Biol.*, 2018, **13**, 2387-2391.
24. S. Mori, A. H. Pang, T. A. Lundy, A. Garzan, O. V. Tsodikov and S. Garneau-Tsodikova, *Nat. Chem. Biol.*, 2018, **14**, 428-430.
25. S. K. Shrestha and S. Garneau-Tsodikova, *ChemBioChem*, 2016, **17**, 1328-1332.
26. T. A. Lundy, S. Mori and S. Garneau-Tsodikova, *ACS Synth. Biol.*, 2018, **7**, 399-404.
27. O. E. Zolova and S. Garneau-Tsodikova, *J. Antibiot.*, 2014, **67**, 59-64.
28. S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen and L. R. Pease, *Gene*, 1989, **77**, 51-59.
29. M. Wolpert, B. Gust, B. Kammerer and L. Heide, *Microbiology*, 2007, **153**, 1413-1423.
30. E. J. Drake, J. Cao, J. Qu, M. B. Shah, R. M. Straubinger and A. M. Gulick, *J. Biol. Chem.*, 2007, **282**, 20425-20434.

31. G. W. Buchko, C. Y. Kim, T. C. Terwilliger and P. J. Myler, *Tuberculosis (Edinb)*, 2010, **90**, 245-251.

32. D. A. Herbst, B. Boll, G. Zocher, T. Stehle and L. Heide, *J. Biol. Chem.*, 2013, **288**, 1991-2003.

33. R. A. Schomer, H. Park, J. J. Barkei and M. G. Thomas, *Biochemistry*, 2018, **57**, 4125-4134.

34. E. A. Felnagle, J. J. Barkei, H. Park, A. M. Podevels, M. D. McMahon, D. W. Drott and M. G. Thomas, *Biochemistry*, 2010, **49**, 8815-8817.

35. S. Mori, K. D. Green, R. Choi, G. W. Buchko, M. G. Fried and S. Garneau-Tsodikova, *ChemBioChem*, 2018, **19**, 2186-2194.

36. M. D. McMahon, J. S. Rush and M. G. Thomas, *J. Bacteriol.*, 2012, **194**, 2809-2818.

37. H. J. Imker, D. Krahn, J. Clerc, M. Kaiser and C. T. Walsh, *Chem. Biol.*, 2010, **17**, 1077-1083.

Image for Table of Content Use Only