BIOCHEMICAL CHARACTERIZATION OF ENZYMES INVOLVED IN THE BIOSYNTHESIS OF THE THIOPEPTIDE THIOMURACIN

BY

ZHENGAN ZHANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Wilfred A. van der Donk, Chair
Professor Douglas A. Mitchell
Professor Satish K. Nair
Professor John A. Gerlt

ABSTRACT

Clinically significant antibiotic resistance has evolved against virtually every antibiotic currently deployed, therefore it is critical for human health to discover new antibacterial agents with novel modes of action. Thiopeptides are a family of ribosomally synthesized and post-translationally modified peptides (RiPPs) that have outstanding biological profile, including high potency against a variety of antibiotic resistant pathogenic strains. Their new modes of action have attracted many efforts to develop chemical syntheses, and to study their biological function and biosynthetic origin. However, the complex architecture and poor physicochemical properties have plagued this otherwise-promising class of antibiotics. To address these challenges, this dissertation used the thiopeptide thiomuracin as a model to establish the reconstitution of the biosynthesis of thiopeptide in vitro, and to characterize the activity of the highly unusual modification enzymes involved in thiopeptide biosynthetic pathway. This work will guide future efforts to improve the properties of thiopeptide natural products.

The biosynthesis of the thiopeptide thiomuracin is a well-orchestrated process involving a multitude of post-translational modifications. Chapter 2 presents the first in vitro biosynthesis of the core scaffold of thiopeptide thiomuracin, while Chapter 3 will provide insights into the substrate specificities, directionality, and timing of catalysis by six proteins involved in the formation of this core scaffold of thiopeptides. Characterization and mechanistic investigation of TbtI, a radical *S*-adenosyl-L-methionine thiazole *C*-methyltransferase involved in the tailoring modification of thiomuracin, are presented in Chapter 4. Lastly, Chapter 5 discusses the characterization of an ATP and tRNA dependent peptidyl transferase involved in a pathway where

a ribosomally synthesized small peptide serves as a catalytic scaffold on which a small-molecule is biosynthesized in *Pseudomonas syringae*.

ACKNOWLEDGEMENTS

Throughout the years of my graduate study, many people have served as an inspiration and motivation to keep me pursuing my dream as a scientist. With their support, encouragement and guidance, I gradually obtained the ability to solve difficult problems, the courage to face different challenges and very importantly, the confidence to trust myself. I would like to thank everyone involved in my five-and-a-half-year journey at the University of Illinois at Urbana-Champaign. I will always be thankful and in great debt to you all.

First of all, my family deserves the greatest thanks of all. I want to express my deepest gratitude to my parents, Guobao Zhang and Junling Zhang. Without them, I would never become who I am right now. I couldn't come up with any words to express even a tiny portion of my great appreciation to their unconditional love for the past twenty eight years. I would also like to give very special thanks to my wife, Hanqing Zhang. All of my successes are, at least to a major extent, because of your love, encouragement and understanding. For everything you have done, for all the sacrifices you have made, and for all of the kindest words you have said, my thanks to you are never enough.

I am very extremely lucky to have become a member of the Prof. van der Donk's group. Since then, Wilfred has provided sage guidance both for my research and my career. His patience and commitment to everyone have allowed us to become better scientists. His thoughtful insights and passionate love for science have been continuously inspiring me to dedicate myself to scientific research. I could not ask for a better advisor than him.

Then, I would like to thank Prof. Douglas A. Mitchell and Prof. Satish K. Nair, for both serving on my committee and for the sincere instructions and helpful suggestions provided during

my research. I will also express my thanks to Prof. John A. Gerlt to serve as a wonderful teacher and for being on my committee.

Next, I will express my gratitude to my collaborators, Graham A. Hudson, Dr. Nilkamal Mahanta, Dillon. P. Cogan and Dr. Jonathan I. Tietz. It is because of their wonderful collaborations that our projects resulted in fruitful achievements. Special thanks to Dr. Nilkamal Mahanta for teaching me how to purify and set up experiments with the air-sensitive radical SAM enzymes. I am very happy to have learned this technique from you.

It is always a pleasure to be part of the van der Donk group. I would like to thank everyone for making the lab such a wonderful place. Specifically, I would first like to thank Prof. Mark C. Walker at the University of New Mexico and Dr. Rebacca A. Splain to serve as my mentors when I joined the lab and to teach me a lot of techniques to perform experiments. Special thanks to Abraham Wang, Kenton Hetrick, Dr. Liujie Huo, Dr. Zedu Huang, Dr. Weixin Tang, Dr. Xiling Zhao, Dr. Xiao Yang, Dr. Gabrielle Thibodeaux, Dr. Manuel Ortega, Dr. Yi Yu, Dr. Qi Zhang, Dr. Chantal Garcia de Gonzalo, Dr. Jiangtao Gao, Dr. Michael Funk, Dr. Julian Hegemann, Dr. Ian Bothwell, Linna An, Chang He, Kuan-Yu Lai, Dr. Chi Pan Ting, Imran Rahman, Silvia Bobeica, Page Daniels, Dr. Graeme Howe and countless others. They are my companions to share joy, and my friends to provide suggestions and encouragement. I also thank Nan Holda for keeping the lab organized.

At last, I would like to express thanks to many of my friends, especially Daokun Sun, Xiaoyu Chang, Dong Xiao, Zhao Wu, Ruopei Feng, Kailiang Li, Sumeng Liu, Liwei Zheng and Ye Shang. Thank you for always being there for me and for your friendship. I will always keep you close to my heart with fond memories.

To my family

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Ribosomally synthesized and post-translationally modified peptide (RiPP) natural products

Natural products have played key roles in human history in advancing our knowledge of chemistry, biology and in the development of medicine (1-3). In the 20th century, four groups of natural products, terpenoids, alkaloids, polyketides, and non-ribosomal peptides, were identified as the most prevalent classes. Then, at the turn of the 21st century, genome sequencing programs revealed the ribosomally synthesized and post-translationally modified peptides (RiPPs) as an additional large class of natural products that are produced in all three domains of life (4). Despite starting from a ribosomal origin, RiPPs are able to reach a similar level of chemical diversity and structural complexity as other groups of natural products through extensive post-translational/co-translational modifications (**Figure 1.1B**). These modifications are chemically diverse and typically restrict conformational flexibility to improve metabolic and chemical stability, and allow better target recognition. Thus, RiPPs may be endowed with a wide variety of highly potent biological activities such as anticancer, antibacterial, antifungal, and antiviral activity (5-9). Indeed, by virtue of their diverse pharmacological properties, this underexplored source of natural products has inspired significant interest in recent years towards the development of novel therapeutics.

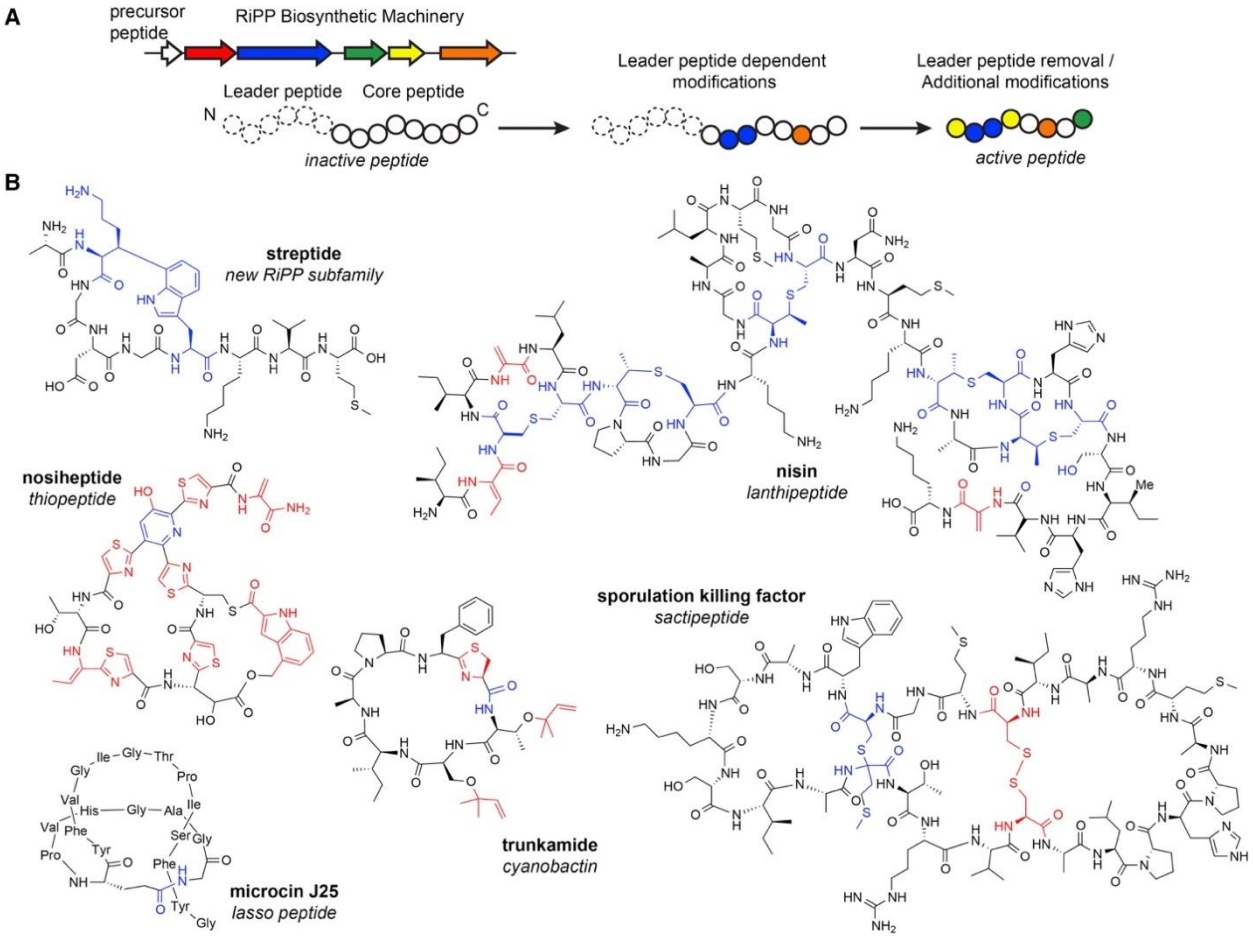


Figure 1.1 (**A**) Schematic RiPP biosynthetic pathway. Representative RiPP biosynthetic gene cluster wherein each gene encodes a biosynthetic enzyme color-coded by the modification it catalyzes. Dashed circles represent amino acids of the leader peptide, which is removed in the late stage of maturation. Full circles represent amino acids within the core peptide. (**B**) Chemical structures of select members of some RiPP subfamilies. Structural features defining a particular RiPP subfamily are shown in blue. Additional post-translational modifications are shown in red. [a]

Unlike non-ribosomal peptide biosynthesis, in which the peptide backbone is constructed by incorporating non-proteinogenic amino acid scaffolds using large multi-modular enzyme complexes, the production of RiPPs starts from a ribosomally synthesized linear precursor peptide

a. The figure was reprinted from: Ortega, M. A., & van der Donk, W. A. (2016). New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem. Biol., 23*, 31-44. Copyright (2016), with permission from Elsevier.

encoded by a structural gene (10, 11). This precursor peptide usually contains an N-terminal leader peptide (LP), which is removed in the late stages of maturation, and a C-terminal core peptide that is converted into the mature product. The leader peptide carries important recognition elements for key biosynthetic enzymes, which in turn allows these enzymes a high degree of tolerance with respect to sequence variation in the core peptide (12). A ribosomally synthesized precursor peptide undergoes modification by a set of enzymes usually encoded in the local genomic region (**Figure 1.1A**), and these modifications include, but are not limited to, dehydrations, cyclizations, methylations, oxidations, and rearrangements. All of these modifications are responsible for endowing the peptide with a rigidified structure and biological activity. The enzymes encoded in a biosynthetic gene cluster govern the posttranslational modifications that a precursor peptide receives, and in some cases, the structure of the final product can be accurately predicted (13, 14). Modification of the core is typically followed by the proteolytic cleavage of the N-terminal leader peptide, and the mature compound is exported. Using this simple strategy, a plethora of chemically diverse natural products can be constructed from a minimal amount of genetic space. In fact, the investigation of RiPPs has overturned the longstanding paradigm that large genomes are required for an organism to produce architecturally complex natural products (15, 16).

The gene-encoded nature of the precursor peptide coupled with the highly promiscuous biosynthetic enzymes of this class of natural products allows for the facile generation of unnatural compound derivatives. Recent studies on many unrelated RiPP families have also revealed a structurally similar precursor peptide binding domain homologous to PqqD — the RiPP precursor peptide recognition element (RRE) (17-21). These findings suggest several future applications, including the engineering of new precursor peptide specificities via RRE swapping, which has garnered a great deal of excitement due to the potential to develop therapeutically relevant

derivatives of this underutilized class of natural products (22-25), and offers the tantalizing opportunity to engineer hybrid RiPPs, not only within RiPP classes (26-28), but also between different RiPP classes (29).

1.2 Thiopeptides and their potential as clinical therapeutics

Thiopeptides, or thiazolyl peptides, are highly modified sulfur-rich peptides belonging to the RiPP family. This class is defined by a series of common chemical moieties that differentiate them from other peptide-derived and/or azole-containing natural products. Their most characteristic feature is a central tri- or tetra-substituted nitrogen-containing six-membered ring, which serves as scaffold to at least one macrocycle and a tail, which both can be decorated with various dehydroamino acids and azoles/azolines (30, 31). All these moieties are formed through dehydration/cyclodehydration of Ser, Thr, and Cys residues present in the core region of the precursor peptide (32-35). Other compound-specific tailoring side-chain modifications, such as oxidations, cyclizations, methylations, and even the presence of another macrocyclic motif are also present in this class of natural product.

Thiopeptides have been classified into different series based on the different oxidation state of the central ring (**Figure 1.2**) (30). Thus, the *a* series possesses a totally reduced central piperidine, whereas the *b* series is oxidized to contain a 1, 2-dehydropiperidine ring. Only one thiopeptide of the *c* series has been isolated to date (36), and its central moiety is somewhat unexpected, as it displays a piperidine ring fused with an imidazoline. All members of the *a*, *b*, and *c* series have a second quinaldic acid containing macrocycle. The *d* series is oxidized further and presents a tri-substituted pyridine ring, the major feature of this subgroup, which is the most prevalent among thiopeptides. The *e* series is oxidized even more and is easily differentiated by

the presence of a hydroxyl group in the central pyridine ring. The *e* series also possesses a very characteristic second macrocycle appending from the main one that is formed by a modified 3,4-dimethylindolic acid moiety.

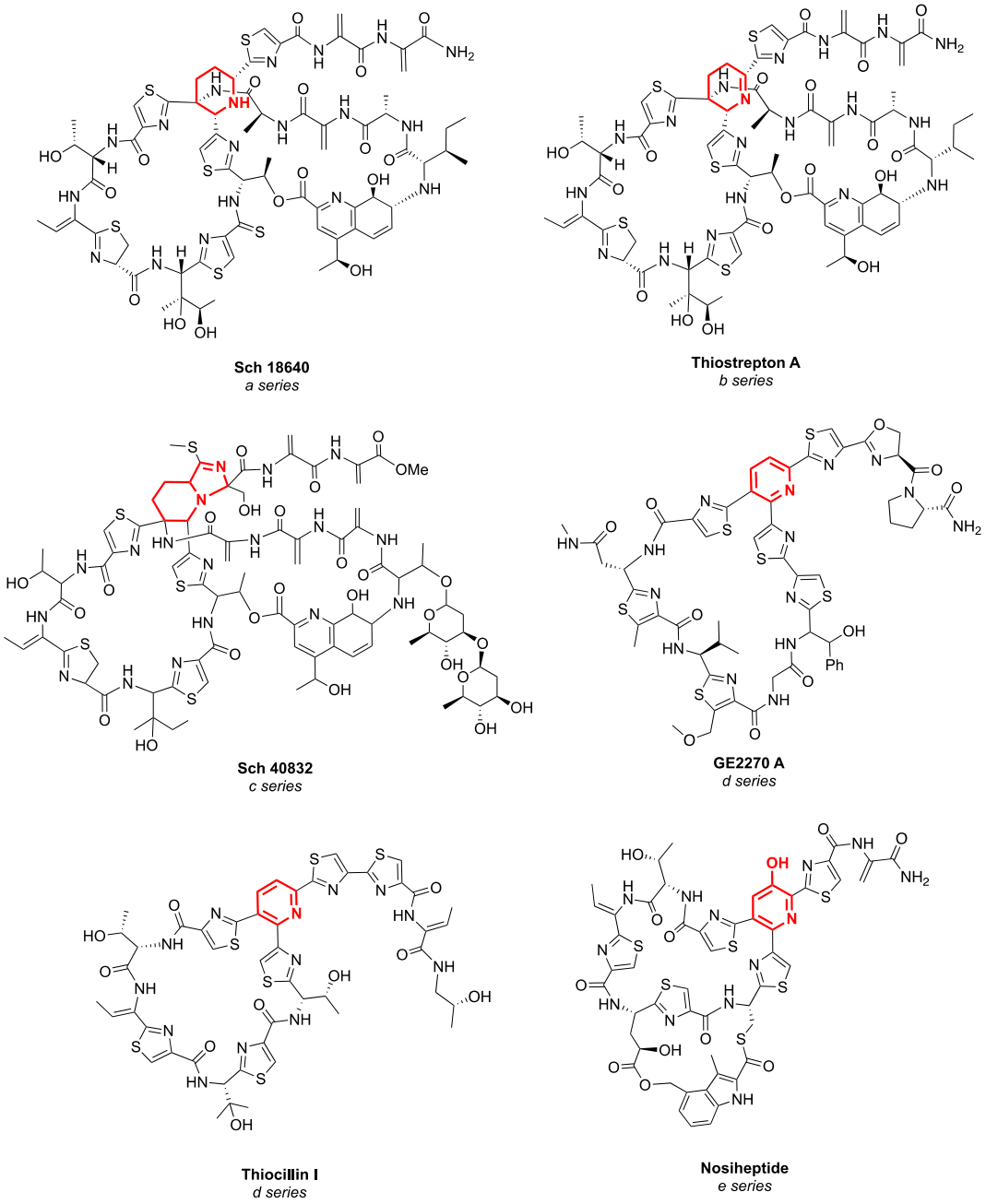


Figure 1.2 Representatives of the five series of thiopeptide classes. Their characteristic central six-member ring is highlighted in red.

The individual chemical identity of structurally distinct compounds isolated from natural sources and classified as thiopeptide antibiotics is relatively diverse. It is an intriguing fact that, despite many similarities in biosynthetic origin, these secondary metabolites have been isolated from a number of different strains of actinomycetes, predominantly soil bacteria but also from marine sources, and elicit a wide range of different biological responses. In 1948, the first known member of the family, micrococcin (37), was isolated from a sample of Oxford's sewage waters. Accounting for the highly diverse origin of thiopeptides, micrococcin P1 was more recently isolated from a completely different source, a French cheese (38). However, more conventional samples, such as from soil, are the main source of most thiopeptides. In fact, thiostrepton, the most famous member of the family, has been isolated from different soil samples (39), including one from Hawaii in 1955 (40), shortly after it was first discovered in 1954 (41). Although a few more thiopeptides were isolated during the following years, it was after the 1980s, especially during the 1990s, that most of the known members were discovered. Many novel entities have also been described during the last decade. Remarkably, the first thiopeptide antibiotics isolated from a marine source were YM-266183 and YM-266184, discovered in 2003 in Japan (42). During the last few years some more thiopeptides have been isolated and characterized (43-47).

Thiopeptides, with their impressive in vitro profile against Gram-positive bacteria, and their diverse mechanisms of action, have gathered the attention of many groups, both in academia and industry. What makes thiopeptides attractive is that they pose an alternative to other antibiotics presently facing resistance by old pathogens. Although thiopeptides are best regarded as antibacterial agents, their therapeutic potential is surprisingly broad as they have been found to possess anticancer (48-52), anti-malarial (53-56), antifungal (57), RNA polymerase inhibition (58), renin inhibitory (59), and even immuno-suppressive activities (60). This wide variety of

biological functions has resulted in a very prolific literature, and suggests that the macrocyclic scaffold of thiopeptides is a privileged structure.

Thiopeptides exert their antibacterial function via the inhibition of ribosomal protein synthesis. However, this antimicrobial activity is the result of different mechanisms of action that depend on macrocycle size. Thiopeptides exhibit macrocycles of three different sizes, 26-, 29-, and 35-membered rings, depending on the number of amino acid residues present. On the one hand, thiopeptides of 26-membered macrocycles, such as micrococcin P1 and the siomycins, are known to bind the GTPase-associated region of the ribosome/L11 protein complex. By doing so, the thiopeptide blocks the binding region of elongation factor G (EF-G) and does not allow translocation of the growing-peptide/tRNA complex in the ribosome to occur (61-63). On the other hand, thiopeptides with a 29-membered ring, such as GE2270A or thiomuracin, bind to elongation factor thermo unstable (EF-Tu), blocking its amino acyl tRNA complex binding site (64-66). As a consequence, the complex would not be able to be delivered into the ribosome and further peptide elongation cannot take place. Thiopeptides with larger macrocycles, the 35-membered rings thioxamycin for example, still possess potent antibacterial activity, but their mode of action has still remained enigmatic.

Somewhat related to antibacterial activity is the *tipA* promoter, which controls transcription of two thiostrepton-induced proteins (Tip), TipAL and TipAS (67). The latter protein, TipAS, serves as a mechanism of defense for bacteria, as it sequesters and covalently binds a thiopeptide molecule, which can no longer inhibit ribosomal protein synthesis. *tipA* promoters have been used to identify thiopeptides in a high-throughput screening program, which detected transcription of the promoter of *tipA* and led to the discovery of several thiopeptides, such as geninthiocin (68), thiotipin (69) and promothiocins (70). Interestingly, the 35-membered thiopeptide radamycin is

completely devoid of antibacterial activity, but is a very strong inducer of *tipA* gene expression. It seems that the induction of the *tipA* promoter is dependent on the presence of a dehydroalanine-containing tail close to the six-membered central scaffold (71, 72).

Thiopeptides known or presumed to bind to EF-Tu, those with 29-membered macrocycles, also possess a very conserved residue in their macrocycle, pointing to a key interaction required to exert their biological activity (73). The GE2270s, GE37468 A, thiomuracin, amythiamicins, as well as baringolin, all belong to the *d* series, and contain an Asn residue in the same position that is either methylated or not. Codon randomization experiments by the Walsh laboratory produced various GE37468 A analogs that substituted Asn with Ala, Cys, His, and Ser; however, none of these analogs retained antibacterial activity, supporting the key role of this residue (74). Except for providing a key contact with EF-Tu, the Asn residue provides a transannular hydrogen bond, thus participating in the stabilization of the bioactive conformation of the macrocycle (75).

One of the biological properties of thiopeptides of major interest is their anticancer activity. In this regard, thiostrepton was found to selectively kill cancer cells without showing any cytotoxicity against healthy tissues. Such a promising effect has been demonstrated to arise from selective inhibition of transcription factor fork-head box M1 (FOXM1) (51, 52). FOXM1 overexpression is associated with the development and progression of cancer and its selective targeting is a very important achievement, as transcription factors have been considered undruggable for a long time (76, 77).

Many thiopeptides have been found to possess anti-malarial activity (53-56). Although the *Plasmodium falciparum* parasite is a eukaryote, thiopeptides target apicoplast protein synthesis (78), which resembles that of prokaryotic organisms. Although it has been demonstrated that thiostrepton binds to the apicoplast 23S rRNA, thiopeptides of the *d* series, such as micrococcin

P1 and amythiamicin A, are much more potent inhibitors of *P. falciparum* growth (79). Recently, the use of thiostrepton semi-synthetic analogs has demonstrated that it targets both the apicoplast ribosomes and the proteasome of *P. falciparum* (56). This dual mode of action could make thiostrepton and similar thiopeptides less likely to develop resistance than drugs with single targets.

A screening program in search of immune-suppressants identified siomycin as inhibitor of antibody production by murine B-cells (60). Comparison with thiostrepton showed the superior behavior of the structurally similar siomycin. Both thiopeptides are thought to possess a different mechanism of action than that of FK506, a common immunosuppressant drug, and to act directly on B-cells.

To date, more than one hundred members of the thiopeptide family have been identified; however, their large molecular size and poor physical-chemical properties have been a major hurdle for introducing them into the clinic. This drawback has become their major limitation and has restricted their use to topical treatments, and thus far, only two agents are in commercial use with both restricted to animal usage (thiostrepton for skin infections; nosiheptide for growth promotion) (31). Impressive efforts in the semi-synthetic derivatization of GE2270A led to LFF571 (80), which recently completed a phase II trial for moderate *Clostridium difficile* infections (81, 82). This trial underscores the opportunities for improving the pharmacological properties of thiopeptides.

1.3 Biosynthesis of thiopeptides

The biosynthetic pathway of thiopeptides has been elusive for a long time. However, recent discoveries have shed light on the biosynthesis of these highly modified natural products. Peptide-based natural products can have two distinct origins depending on how their amino acids are

condensed to form the parent peptides. These can be either synthesized by the ribosome as product of mRNA translation or can be assembled by non-ribosomal peptide synthetases (NRPSs). Although most highly modified peptide-derived natural products are synthesized by NRPSs, there was no evidence of such origin for thiopeptides. In 2009, four different groups demonstrated that the pre-peptide of thiopeptide is ribosomally synthesized and, thus, is genetically encoded. Genome mining efforts have identified genes that encode the precursor peptide and the enzymatic machinery necessary for its subsequent post-translational modifications. The gene encoding the precursor peptide has been discovered for many thiopeptides and, in all cases, there is a perfect agreement of expected amino acid sequence with the final product (32-35, 45, 46, 83-85).

The function of most enzymes found in thiopeptide gene clusters has been established through homology to known enzymes, gene deletion studies, or *in vitro* biochemical analysis (86-92). Through these studies, it has been shown that the characteristic thiazoline/oxazoline rings are initially formed through the cyclodehydration of target Ser, Thr, and Cys residues by YcaO enzymes. Oxidation of these rings to the corresponding thiazole/oxazole structures can then be accomplished through the activity of flavin mononucleotide dependent dehydrogenases (**Figure 1.3A**) (23). In addition, key dehydroalanine (Dha) and dehydrobutyrine (Dhb) moieties are produced in the core peptide by split class I lanthipeptide dehydratases. These enzymes utilize $tRNA^{Glu}$ to first glutamylate target Ser and Thr residues prior to eliminating glutamate in order to form the dehydrated product. This mechanism was recently established through the study of the tRNA-dependent lanthipeptide dehydratase NisB involved in nisin biosynthesis (**Figure 1.3B**) (21). The last step in the biosynthesis of the thiopeptide core scaffold is the proposed intramolecular aza-Diels-Alder-like cycloaddition reaction between two Dha residues to form the central six-membered ring, followed by dehydration and subsequent elimination of the leader

peptide (**Figure 1.3C**) (93, 94). Further side-chain modifications, such as *C*-terminal amidation, oxidation, glycosylation, methylation, and incorporation of indolic or quinaldic acid moieties are thought to occur in later stages of the biosynthetic pathway.

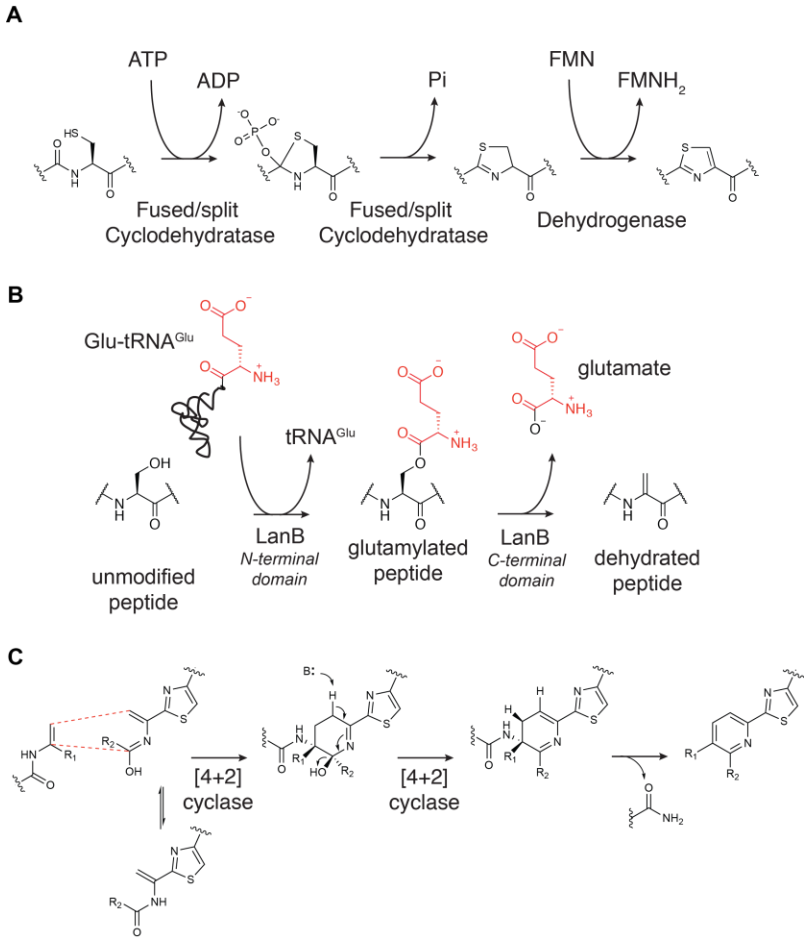


Figure 1.3 (**A**) A general scheme for thiazol(in)e/oxazol(in)e heterocycle biogenesis. Select Cys, Ser, and Thr residues are cyclized by the ATP-dependent cyclodehydratase. Azoline heterocycles can then be oxidized by a flavin-dependent dehydrogenase to afford the aromatic azole. (**B**) Proposed dehydration mechanism employed by Class I lanthipeptide dehydratases. Select Ser and Thr residues are first glutamylated by the N-terminal domain of LanB dehydratases using Glu-tRNA^{Glu}. The elimination of the glutamate yields the dehydrated amino acids. (**C**) Proposed mechanism for the central nitrogen-containing six-membered ring formation. The [4+2] cyclase catalyzes the intramolecular aza-Diels-Alder-like cycloaddition reaction between two distant Dha residues to generate the macrocyclic scaffold of thiopeptides with the elimination of a water molecule and the leader peptide.

1.4 Summary and outlook

Thiopeptides, with their unique structural features and diverse modes of action, have attracted much interest aimed at understanding their biosynthetic origin and pharmacological potential. Their biosynthesis and application have proven to be a challenging pursuit, however, given their complex architecture and poor physicochemical properties. Although prior studies had been successful in providing the general framework, a detailed biochemical and mechanistic understanding of the biosynthetic processes had remained elusive. To address these challenges, in collaboration with the Mitchell laboratory we have reconstituted the biosynthesis of the model thiopeptide, thiomuracin, *in vitro* and characterized the activity of many of the enzymes encoded in its biosynthetic gene cluster. In Chapter 2, I present our findings regarding the reconstitution the activities of six proteins that together biosynthesize the core structure of thiomuracin via 22 chemical transformations; I further discuss in Chapter 3 the substrate specificities, directionality, and timing of catalysis by the six proteins involved in this well-orchestrated set of enzymatic reactions, as well as the mechanistic and crystallographic characterization of the enzyme catalyzing the [4+2] *aza*-cyclization reaction; in Chapter 4, the biochemical and mechanistic investigation of a radial *S*-adenosyl-L-methionine *C*-methyltransferase will be presented. Together, this work lays the foundation for development of a large library of non-natural thiopeptide analogues for pharmacological study. Finally, Chapter 5 discusses a protein related to the thiopeptide glutamylation enzymes but carries out very different chemistry.

1.5 References

1. Carlson, E. E. (2010). Natural Products as Chemical Probes. *ACS Chem. Biol., 5*, 639-653.

2. Newman, D. J., & Cragg, G. M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod., 79*, 629-661.

3. Falconer, S. B., Czarny, T. L., & Brown, E. D. (2011). Antibiotics as probes of biological complexity. *Nat. Chem. Biol., 7*, 415.

4. 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., Rebuffat, S., Ross, R. P., Sahl, H.-G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., 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., & 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.

5. Mohr, K. I., Volz, C., Jansen, R., Wray, V., Hoffmann, J., Bernecker, S., Wink, J., Gerth, K., Stadler, M., & Müller, R. (2015). Pinensins: The First Antifungal Lantibiotics. *Angew. Chem. Int. Ed., 54*, 11254-11258.

6. Essig, A., Hofmann, D., Münch, D., Gayathri, S., Künzler, M., Kallio, P. T., Sahl, H.-G., Wider, G., Schneider, T., & Aebi, M. (2014). Copsin, a Novel Peptide-based Fungal Antibiotic Interfering with the Peptidoglycan Synthesis. *J. Biol. Chem., 289*, 34953-34964.

7. Férir, G., Petrova, M. I., Andrei, G., Huskens, D., Hoorelbeke, B., Snoeck, R., Vanderleyden, J., Balzarini, J., Bartoschek, S., Brönstrup, M., Süssmuth, R. D., & Schols, D. (2013). The Lantibiotic Peptide Labyrinthopeptin A1 Demonstrates Broad Anti-HIV and Anti-HSV Activity with Potential for Microbicidal Applications. *PLoS One, 8*, e64010.

8. Schmidt, E. W., Nelson, J. T., Rasko, D. A., Sudek, S., Eisen, J. A., Haygood, M. G., & 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.

9. Philmus, B., Christiansen, G., Yoshida, W. Y., & Hemscheidt, T. K. (2008). Post-translational Modification in Microviridin Biosynthesis. *Chembiochem, 9*, 3066-3073.

10. Dunbar, K. L., & Mitchell, D. A. (2013). Revealing Nature's Synthetic Potential Through the Study of Ribosomal Natural Product Biosynthesis. *ACS Chem. Biol., 8*, 473-487.

11. Ortega, M. A., & van der Donk, W. A. (2016). New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem. Biol., 23*, 31-44.

12. Oman, T. J., & van der Donk, W. A. (2009). Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol., 6*, 9.

13. Donia, M. S., & Schmidt, E. W. (2011). Linking Chemistry and Genetics in the Growing Cyanobactin Natural Products Family. *Chem. Biol., 18*, 508-519.

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

15. Haft, D. H., Basu, M. K., & Mitchell, D. A. (2010). Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family. *BMC Biol., 8*, 70.

16. Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., Joewono, I., Rusch, D., Chisholm, S. W., & 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.

17. Burkhart, B. J., Hudson, G. A., Dunbar, K. L., & Mitchell, D. A. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol., 11*, 564.

18. Koehnke, J., Mann, G., Bent, A. F., Ludewig, H., Shirran, S., Botting, C., Lebl, T., Houssen, W. E., Jaspars, M., & Naismith, J. H. (2015). Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat. Chem. Biol., 11*, 558.

19. Dunbar, K. L., Tietz, J. I., Cox, C. L., Burkhart, B. J., & Mitchell, D. A. (2015). Identification of an Auxiliary Leader Peptide-Binding Protein Required for Azoline Formation in Ribosomal Natural Products. *J. Am. Chem. Soc., 137*, 7672-7677.

20. Latham, J. A., Iavarone, A. T., Barr, I., Juthani, P. V., & Klinman, J. P. (2015). PqqD Is a Novel Peptide Chaperone That Forms a Ternary Complex with the Radical S-Adenosylmethionine Protein PqqE in the Pyrroloquinoline Quinone Biosynthetic Pathway. *J. Biol. Chem., 290*, 12908-12918.

21. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., van der Donk, W. A., & Nair, S. K. (2015). Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature, 517*, 509-512.

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

23. Melby, J. O., Nard, N. J., & Mitchell, D. A. (2011). Thiazole/oxazole-modified microcins: complex natural products from ribosomal templates. *Curr. Opin. Chem. Biol., 15*, 369-378.

24. Tianero, M. D. B., Donia, M. S., Young, T. S., Schultz, P. G., & Schmidt, E. W. (2012). Ribosomal Route to Small-Molecule Diversity. *J. Am. Chem. Soc., 134*, 418-425.

25. Melby, J. O., Dunbar, K. L., Trinh, N. Q., & 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.

26. van Heel, A. J., Mu, D., Montalban-Lopez, M., Hendriks, D., & Kuipers, O. P. (2013). Designing and producing modified, new-to-nature peptides with antimicrobial activity by use of a combination of various lantibiotic modification enzymes. *ACS Synth. Biol., 2*, 397-404.

27. Houssen, W. E., Bent, A. F., McEwan, A. R., Pieiller, N., Tabudravu, J., Koehnke, J., Mann, G., Adaba, R. I., Thomas, L., Hawas, U. W., Liu, H., Schwarz-Linek, U., Smith, M. C., Naismith, J. H., & Jaspars, M. (2014). An Efficient Method for the In Vitro Production of Azol(in)e-Based Cyclic Peptides. *Angew. Chem. Int. Ed., 53*, 14171-14174.

28. Sardar, D., Lin, Z., & Schmidt, E. W. (2015). Modularity of RiPP Enzymes Enables Designed Synthesis of Decorated Peptides. *Chem. Biol., 22*, 907-916.

29. Burkhart, B. J., Kakkar, N., Hudson, G. A., van der Donk, W. A., & Mitchell, D. A. (2017). Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Cent. Sci., 3*, 629-638.

30. Bagley, M. C., Dale, J. W., Merritt, E. A., & Xiong, X. (2005). Thiopeptide Antibiotics. *Chem. Rev., 105*, 685-714.

31. Just-Baringo, X., Albericio, F., & Álvarez, M. (2014). Thiopeptide Antibiotics: Retrospective and Recent Advances. *Mar. Drugs, 12*, 317.

32. Kelly, W. L., Pan, L., & Li, C. (2009). Thiostrepton Biosynthesis: Prototype for a New Family of Bacteriocins. *J. Am. Chem. Soc., 131*, 4327-4334.

33. Brown, L. C. W., Acker, M. G., Clardy, J., Walsh, C. T., & Fischbach, M. A. (2009). Thirteen posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin. *Proc. Natl. Acad. Sci. U. S. A.*

34. Morris, R. P., Leeds, J. A., Naegeli, H. U., Oberer, L., Memmert, K., Weber, E., LaMarche, M. J., Parker, C. N., Burrer, N., Esterow, S., Hein, A. E., Schmitt, E. K., & Krastel, P.

(2009). Ribosomally Synthesized Thiopeptide Antibiotics Targeting Elongation Factor Tu. *J. Am. Chem. Soc., 131*, 5946-5955.

35. Liao, R., Duan, L., Lei, C., Pan, H., Ding, Y., Zhang, Q., Chen, D., Shen, B., Yu, Y., & Liu, W. (2009). Thiopeptide Biosynthesis Featuring Ribosomally Synthesized Precursor Peptides and Conserved Posttranslational Modifications. *Chem. Biol., 16*, 141-147.

36. Puar, M. S., Chan, T. M., Hegde, V., Patel, M., Bartner, P., Ng, K. J., Pramanik, B. N., & MacFarlane, R. D. (1998). Sch 40832: a novel thiostrepton from Micromonospora carbonacea. *J. Antibiot., 51*, 221-224.

37. Su, T. L. (1948). Micrococcin, an antibacterial substance formed by a strain of Micrococcus. *Br. J. Exp. Pathol., 29*, 473-481.

38. Carnio, M. C., Stachelhaus, T., Francis, K. P., & Scherer, S. (2001). Pyridinyl polythiazole class peptide antibiotic micrococcin P1, secreted by foodborne Staphylococcus equorum WS2733, is biosynthesized nonribosomally. *Eur. J. Biochem., 268*, 6390-6401.

39. Donovick, R., Pagano, J. F., Stout, H. A., & Weinstein, M. J. (1955). Thiostrepton, a new antibiotic. I. In vitro studies. *Antibiot. Ann., 3*, 554-559.

40. Cron, M. J., Whitehead, D. F., Hooper, I. R., Heinemann, B., & Lein, J. (1956). Bryamycin, a New Antibiotic. *Antibiot. Chemother., 6*, 63-67.

41. Dutcher, J. D., & Vandeputte, J. (1955). Thiostrepton, a new antibiotic. II. Isolation and chemical characterization. *Antibiot. Annu., 3*, 560-561.

42. Suzumura, K., Yokoi, T., Funatsu, M., Nagai, K., Tanaka, K., Zhang, H., & Suzuki, K. (2003). YM-266183 and YM-266184, novel thiopeptide antibiotics produced by Bacillus cereus isolated from a marine sponge II. Structure elucidation. *J. Antibiot., 56*, 129-134.

43. Jayasuriya, H., Herath, K., Ondeyka, J. G., Zhang, C., Zink, D. L., Brower, M., Gailliot, F. P., Greene, J., Birdsall, G., Venugopal, J., Ushio, M., Burgess, B., Russotti, G., Walker, A., Hesse, M., Seeley, A., Junker, B., Connors, N., Salazar, O., Genilloud, O., Liu, K., Masurekar, P., Barrett, J. F., & Singh, S. B. (2007). Isolation and Structure Elucidation of Thiazomycin. *J. Antibiot., 60*, 554.

44. Zhang, C., Herath, K., Jayasuriya, H., Ondeyka, J. G., Zink, D. L., Occi, J., Birdsall, G., Venugopal, J., Ushio, M., Burgess, B., Masurekar, P., Barrett, J. F., & Singh, S. B. (2009). Thiazomycins, Thiazolyl Peptide Antibiotics from Amycolatopsis fastidiosa. *J. Nat. Prod., 72*, 841-847.

45. Engelhardt, K., Degnes, K. F., & Zotchev, S. B. (2010). Isolation and Characterization of the Gene Cluster for Biosynthesis of the Thiopeptide Antibiotic TP-1161. *Appl. Environ. Microbiol., 76*, 7093-7101.

46. Engelhardt, K., Degnes, K. F., Kemmler, M., Bredholt, H., Fjærvik, E., Klinkenberg, G., Sletta, H., Ellingsen, T. E., & Zotchev, S. B. (2010). Production of a New Thiopeptide Antibiotic, TP-1161, by a Marine Nocardiopsis Species. *Appl. Environ. Microbiol., 76*, 4969.

47. Zhang, C., Occi, J., Masurekar, P., Barrett, J. F., Zink, D. L., Smith, S., Onishi, R., Ha, S., Salazar, O., Genilloud, O., Basilio, A., Vicente, F., Gill, C., Hickey, E. J., Dorso, K., Motyl, M., & Singh, S. B. (2008). Isolation, Structure, and Antibacterial Activity of Philipimycin, A Thiazolyl Peptide Discovered from Actinoplanes philippinensis MA7347. *J. Am. Chem. Soc., 130*, 12102-12110.

48. Nicolaou, K. C., Zak, M., Rahimipour, S., Estrada, A. A., Lee, S. H., O'Brate, A., Giannakakou, P., & Ghadiri, M. R. (2005). Discovery of a Biologically Active Thiostrepton Fragment. *J. Am. Chem. Soc., 127*, 15042-15044.

49. Radhakrishnan, S. K., Bhat, U. G., Hughes, D. E., Wang, I.-C., Costa, R. H., & Gartel, A. L. (2006). Identification of a Chemical Inhibitor of the Oncogenic Transcription Factor Forkhead Box M1. *Cancer Res., 66*, 9731-9735.

50. Bhat, U. G., Zipfel, P. A., Tyler, D. S., & Gartel, A. L. (2008). Novel anticancer compounds induce apoptosis in melanoma cells. *Cell Cycle, 7*, 1851-1855.

51. Bhat, U. G., Halasi, M., & Gartel, A. L. (2009). Thiazole Antibiotics Target FoxM1 and Induce Apoptosis in Human Cancer Cells. *PLoS One, 4*, e5592.

52. Hegde, N. S., Sanders, D. A., Rodriguez, R., & Balasubramanian, S. (2011). The transcription factor FOXM1 is a cellular target of the natural product thiostrepton. *Nat. Chem., 3*, 725.

53. Schlitzer, M. (2007). Malaria Chemotherapeutics Part I: History of Antimalarial Drug Development, Currently Used Therapeutics, and Drugs in Clinical Development. *ChemMedChem, 2*, 944-986.

54. Goodman, C. D., Su, V., & McFadden, G. I. (2007). The effects of anti-bacterials on the malaria parasite Plasmodium falciparum. *Mol. Biochem. Parasitol., 152*, 181-191.

55. Clough, B., Strath, M., Preiser, P., Denny, P., & Wilson, I. (1997). Thiostrepton binds to malarial plastid rRNA. *FEBS Lett., 406*, 123-125.

56. Schoof, S., Pradel, G., Aminake, M. N., Ellinger, B., Baumann, S., Potowski, M., Najajreh, Y., Kirschner, M., & Arndt, H.-D. (2010). Antiplasmodial Thiostrepton Derivatives: Proteasome Inhibitors with a Dual Mode of Action. *Angew. Chem. Int. Ed., 49*, 3317-3321.

57. Mizuhara, N., Kuroda, M., Ogita, A., Tanaka, T., Usuki, Y., & Fujita, K.-i. (2011). Antifungal thiopeptide cyclothiazomycin B1 exhibits growth inhibition accompanying

morphological changes via binding to fungal cell wall chitin. *Bioorg. Med. Chem., 19*, 5300-5310.

58. Hashimoto, M., Murakami, T., Funahashi, K., Tokunaga, T., Nihei, K.-i., Okuno, T., Kimura, T., Naoki, H., & Himeno, H. (2006). An RNA polymerase inhibitor, cyclothiazomycin B1, and its isomer. *Bioorg. Med. Chem., 14*, 8259-8270.

59. Aoki, M., Ohtsuka, T., Yamada, M., Ohba, Y., Yoshizaki, H., Yasuno, H., Sano, T., Watanabe, J., Yokose, K., & Seto, H. (1991). Cyclothiazomycin, a novel polythiazole-containing peptide with renin inhibitory activity. Taxonomy, fermentation, isolation and physico-chemical characterization. *J. Antibiot., 44*, 582-588.

60. Ueno, M., Furukawa, S., Abe, F., Ushioda, M., Fujine, K., Johki, S., Hatori, H., & Ueda, H. (2004). Suppressive effect of antibiotic siomycin on antibody production. *J. Antibiot., 57*, 590-596.

61. Baumann, S., Schoof, S., Harkal, S. D., & Arndt, H.-D. (2008). Mapping the Binding Site of Thiopeptide Antibiotics by Proximity-Induced Covalent Capture. *J. Am. Chem. Soc., 130*, 5664-5666.

62. Harms, J. M., Wilson, D. N., Schluenzen, F., Connell, S. R., Stachelhaus, T., Zaborowska, Z., Spahn, C. M. T., & Fucini, P. (2008). Translational Regulation via L11: Molecular Switches on the Ribosome Turned On and Off by Thiostrepton and Micrococcin. *Mol. Cell, 30*, 26-38.

63. Walter, J. D., Hunter, M., Cobb, M., Traeger, G., & Spiegel, P. C. (2012). Thiostrepton inhibits stable 70S ribosome binding and ribosome-dependent GTPase activation of elongation factor G and elongation factor 4. *Nucleic Acids Res., 40*, 360-370.

64. Selva, E., Montanini, N., Stella, S., Soffientini, A., Gastaldo, L., & Denaro, M. (1997). Targeted screening for elongation factor Tu binding antibiotics. *J. Antibiot., 50*, 22-26.

65. Heffron, S. E., & Jurnak, F. (2000). Structure of an EF-Tu Complex with a Thiazolyl Peptide Antibiotic Determined at 2.35 Å Resolution: Atomic Basis for GE2270A Inhibition of EF-Tu. *Biochemistry, 39*, 37-45.

66. Parmeggiani, A., & Nissen, P. (2006). Elongation factor Tu-targeted antibiotics: Four different structures, two mechanisms of action. *FEBS Lett., 580*, 4576-4581.

67. Murakami, T., Holt, T. G., & Thompson, C. J. (1989). Thiostrepton-induced gene expression in Streptomyces lividans. *J. Bacteriol., 171*, 1459-1466.

68. Yun, B. S., Hidaka, T., Furihata, K., & Seto, H. (1994). Microbial metabolites with tipA promoter inducing activity. II. Geninthiocin, a novel thiopeptide produced by Streptomyces sp. DD84. *J. Antibiot., 47*, 969-975.

69. Yun, B.-S., Hidaka, T., Furihata, K., & Seto, H. (1994). Thiotipin, a novel thiopeptide with a tipA promoter inducing activity produced by Streptomyces sp. DT31. *Tetrahedron, 50*, 11659-11664.

70. Yun, B. S., Hidaka, T., Furihata, K., & Seto, H. (1994). Promothiocins A and B novel thiopeptides with a tipA promoter inducing activity produced by Streptomyces sp. SF2741. *J. Antibiot., 47*, 510-514.

71. Lentzen, G., Klinck, R., Matassova, N., Aboul-ela, F., & Murchie, A. I. H. (2003). Structural Basis for Contrasting Activities of Ribosome Binding Thiazole Antibiotics. *Chem. Biol., 10*, 769-778.

72. Chiu, M. L., Folcher, M., Katoh, T., Puglia, A. M., Vohradsky, J., Yun, B.-S., Seto, H., & Thompson, C. J. (1999). Broad Spectrum Thiopeptide Recognition Specificity of the Streptomyces lividans TipAL Protein and Its Role in Regulating Gene Expression. *J. Biol. Chem., 274*, 20578-20586.

73. Parmeggiani, A., Krab, I. M., Okamura, S., Nielsen, R. C., Nyborg, J., & Nissen, P. (2006). Structural Basis of the Action of Pulvomycin and GE2270 A on Elongation Factor Tu. *Biochemistry, 45*, 6846-6857.

74. Young, Travis S., Dorrestein, Pieter C., & Walsh, Christopher T. (2012). Codon Randomization for Rapid Exploration of Chemical Space in Thiopeptide Antibiotic Variants. *Chem. Biol., 19*, 1600-1610.

75. Lewis, R. J., Hughes, R. A., Alcaraz, L., Thompson, S. P., & Moody, C. J. (2006). Solution structures of thiopeptide antibiotics. *Chem. Commun.*, 4215-4217.

76. Bauer, R. A., Wurst, J. M., & Tan, D. S. (2010). Expanding the range of 'druggable' targets with natural product-based libraries: an academic perspective. *Curr. Opin. Chem. Biol., 14*, 308-314.

77. Koehler, A. N. (2010). A complex task? Direct modulation of transcription factors with small molecules. *Curr. Opin. Chem. Biol., 14*, 331-340.

78. Chaubey, S., Kumar, A., Singh, D., & Habib, S. (2005). The apicoplast of Plasmodium falciparum is translationally active. *Mol. Microbiol., 56*, 81-89.

79. Clough, B., Rangachari, K., Strath, M., Preiser, P. R., & Iain Wilson, R. J. M. (1999). Antibiotic Inhibitors of Organellar Protein Synthesis in Plasmodium falciparum. *Protist, 150*, 189-195.

80. LaMarche, M. J., Leeds, J. A., Amaral, A., Brewer, J. T., Bushell, S. M., Deng, G., Dewhurst, J. M., Ding, J., Dzink-Fox, J., Gamber, G., Jain, A., Lee, K., Lee, L., Lister, T., McKenney, D., Mullin, S., Osborne, C., Palestrant, D., Patane, M. A., Rann, E. M., Sachdeva, M., Shao, J., Tiamfook, S., Trzasko, A., Whitehead, L., Yifru, A., Yu, D., Yan,

W., & Zhu, Q. (2012). Discovery of LFF571: an investigational agent for Clostridium difficile infection. *J. Med. Chem., 55*, 2376-2387.

81. Mullane, K., Lee, C., Bressler, A., Buitrago, M., Weiss, K., Dabovic, K., Praestgaard, J., Leeds, J. A., Blais, J., & Pertel, P. (2015). Multicenter, randomized clinical trial to compare the safety and efficacy of LFF571 and vancomycin for Clostridium difficile infections. *Antimicrob. Agents Chemother., 59*, 1435-1440.

82. Bhansali, S. G., Mullane, K., Ting, L. S., Leeds, J. A., Dabovic, K., Praestgaard, J., & Pertel, P. (2015). Pharmacokinetics of LFF571 and vancomycin in patients with moderate Clostridium difficile infections. *Antimicrob. Agents Chemother., 59*, 1441-1445.

83. Walsh, C. T., Acker, M. G., & Bowers, A. A. (2010). Thiazolyl Peptide Antibiotic Biosynthesis: A Cascade of Post-translational Modifications on Ribosomal Nascent Proteins. *J. Biol. Chem., 285*, 27525-27531.

84. Li, C., & Kelly, W. L. (2010). Recent advances in thiopeptide antibiotic biosynthesis. *Nat. Prod. Rep., 27*, 153-164.

85. Arndt, H.-D., Schoof, S., & Lu, J.-Y. (2009). Thiopeptide Antibiotic Biosynthesis. *Angew. Chem. Int. Ed., 48*, 6770-6773.

86. Bowers, A. A., Walsh, C. T., & Acker, M. G. (2010). Genetic Interception and Structural Characterization of Thiopeptide Cyclization Precursors from Bacillus cereus. *J. Am. Chem. Soc., 132*, 12182-12184.

87. Zhang, Q., Li, Y., Chen, D., Yu, Y., Duan, L., Shen, B., & Liu, W. (2011). Radical-mediated enzymatic carbon chain fragmentation-recombination. *Nat. Chem. Biol., 7*, 154.

88. Wei, M., Deng, J., Wang, S., Liu, N., & Chen, Y. (2011). A simple reverse genetics approach to elucidating the biosynthetic pathway of nocathiacin. *Biotechnol. Lett., 33*, 585-591.

89. Zhang, Q., Chen, D., Lin, J., Liao, R., Tong, W., Xu, Z., & Liu, W. (2011). Characterization of NocL Involved in Thiopeptide Nocathiacin I Biosynthesis: A [4Fe-4S] cluster and the catalysis of a radical s-adenosylmethionine enzyme. *J. Biol. Chem., 286*, 21287-21294.

90. Pierre, S., Guillot, A., Benjdia, A., Sandström, C., Langella, P., & Berteau, O. (2012). Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes. *Nat. Chem. Biol., 8*, 957.

91. Duan, L., Wang, S., Liao, R., & Liu, W. (2012). Insights into Quinaldic Acid Moiety Formation in Thiostrepton Biosynthesis Facilitating Fluorinated Thiopeptide Generation. *Chem. Biol., 19*, 443-448.

92. Malcolmson, S. J., Young, T. S., Ruby, J. G., Skewes-Cox, P., & Walsh, C. T. (2013). The posttranslational modification cascade to the thiopeptide berninamycin generates linear forms and altered macrocyclic scaffolds. *Proc. Natl. Acad. Sci. U. S. A., 110,* 8483-8488.

93. Mocek, U., Zeng, Z., O'Hagan, D., Zhou, P., Fan, L. D. G., Beale, J. M., & Floss, H. G. (1993). Biosynthesis of the modified peptide antibiotic thiostrepton in Streptomyces azureus and Streptomyces laurentii. *J. Am. Chem. Soc., 115,* 7992-8001.

94. Bycroft, B. W., & Gowland, M. S. (1978). The structures of the highly modified peptide antibiotics micrococcin P1 and P2. *J. Chem. Soc., Chem. Commun.*, 256-258.

CHAPTER 2: RECONSTITUTION OF THE BIOSYNTHESIS OF THE CORE SCAFFOLD OF THE THIOPEPTIDE THIOMURACIN[b]

2.1 Introduction

Thiopeptides are potent antibiotics that inhibit protein synthesis. They are made by a remarkable post-translational modification process that transforms a linear peptide into a polycyclic structure. The biosynthesis of thiopeptides utilizes Ser/Thr dehydratases that produce dehydro-amino acids alongside cyclodehydratases that convert Ser, Thr and Cys residues to the corresponding azolines (1-4). The presence of both types of post-translational modifications facilitated the evolution of new chemistry, including a formal [4+2] cycloaddition enzyme that converts two dehydroalanine (Dha) moieties to the class-defining central pyridine [or (dehydro)piperidine] (5-7). Recent advances in our mechanistic understanding of RiPP cyclodehydratases (8-11) and lanthipeptide dehydratases (12, 13) prompted us to attempt the in vitro biosynthetic reconstitution of this class of natural products. In collaboration with Prof. Douglas Mitchell's lab, we developed the first in vitro biosynthesis system to reconstitute the activities of six proteins that together biosynthesize the core structure of thiomuracin via 22 chemical transformations (**Figure 2.1A**). This chapter describes the insights into the order of events and requirements for catalysis.

[b] Parts of the chapter are reprinted with permission from: Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., & van der Donk, W. A. (2015). In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *J. Am. Chem. Soc.*, *137*, 16012-16015. Copyright © 2015 American Chemical Society. https://pubs.acs.org/doi/abs/10.1021/jacs.5b10194.
The in vitro studies of cyclodehydratase and dehydrogenase TbtF, TbtG and TbtE were performed by Graham A. Hudson.
The in vitro study of the [4+2] cycloaddition enzyme TbtD was a collaboration with Graham Hudson.

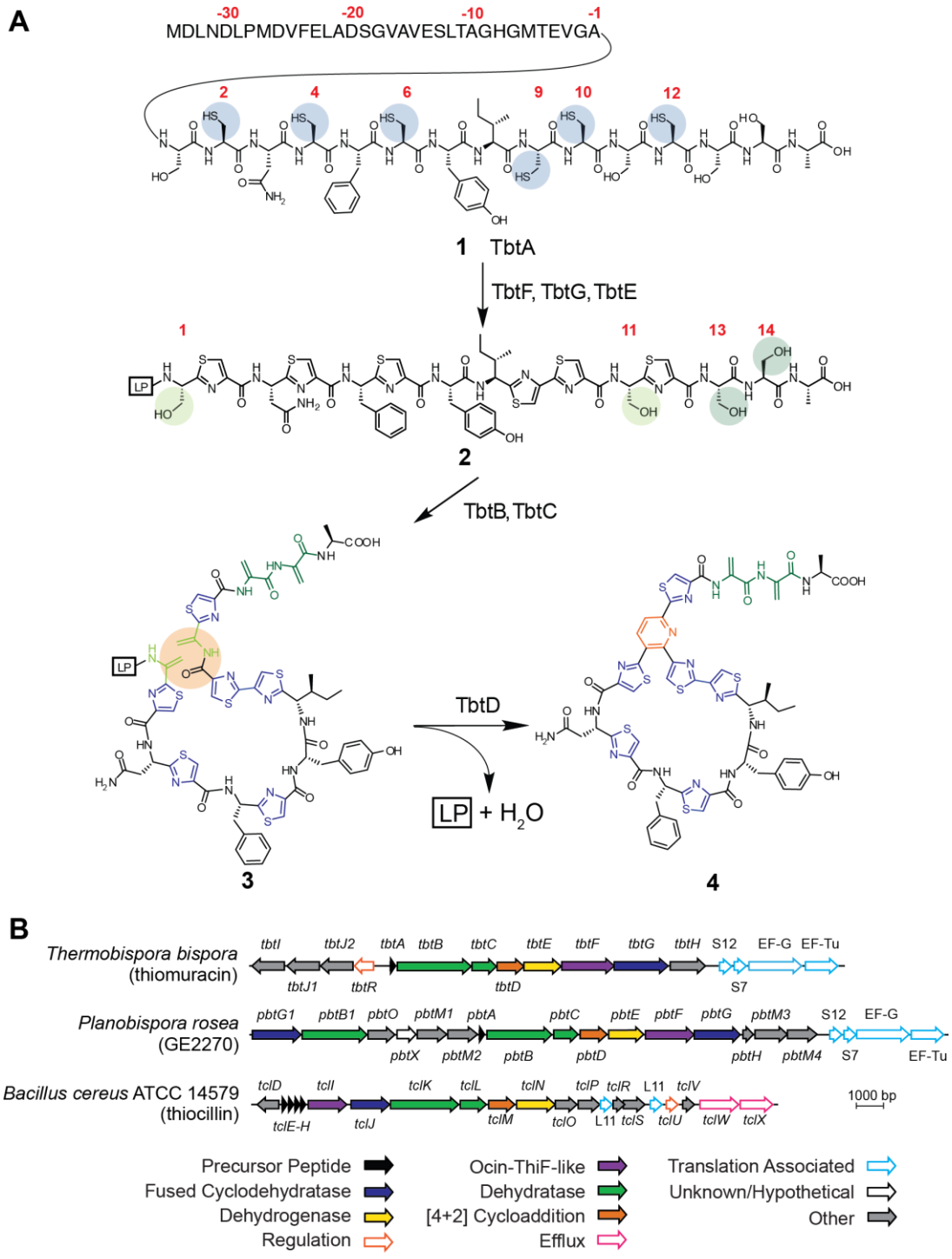


Figure 2.1 (**A**) Biosynthetic route to the thiomuracin core scaffold **4**. (**B**) Gene clusters for the biosynthesis of thiomuracin, GE2270A, and thiocillin. Sequence of the His_6-tagged leader peptide (LP): PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA

The gene-encoded nature of the precursor peptide coupled with the promiscuous biosynthetic enzymes of thiopeptides (5) allows for the facile generation of unnatural compound derivatives, and facilitates the potential production of libraries of unnatural analogues for pharmacological property improvement. Heterologous expression is an appealing option for library construction as it will circumvent the labor invested in enzyme purification for in vitro studies, as well as the time and effort required for organic synthesis, thus making the generation of derivatives faster and more convenient. At the onset of our studies, all heterologous expression of thiopeptides was performed in actinobacteria, such as GE2270 in *Nonomuraea* ATCC39727 (14), berninamycin in *Streptomyces venezuelae* (15), lactazole in *Streptomyces lividans* (16) and thiostrepton in *Streptomyces actuosus* (17). However, the intractability to genetic editing, as well as the instability of the intermediates inside the organism during long growths are disadvantages of these heterologous expression hosts. *Escherichia coli*, with its attractive features such as fast growth rate, ease of manipulation and the ability to introduce nonproteinogenic amino acids (18, 19), could provide an attractive alternative. Although no *E. coli* expression of thiopeptides has been reported, several lanthipeptides, another class of RiPPs family, have been produced in *E. coli* by co-expressing precursor peptides together with their corresponding biosynthetic machinery to afford the fully modified lantibiotics (20-23). Thus in this chapter, in addition to the in vitro studies, I will also discuss the construction of an *E. coli* heterologous producing system to lay the basis for the production of thiopeptide analogues.

2.2 Results and discussion

2.2.1 In vitro biosynthesis of the thiopeptide thiomuracin core scaffold

The thiopeptide family is comprised of >100 members, with many compounds adorned with a second macrocycle (e.g. thiostrepton) (5). To begin our effort, we restricted our thiopeptide in vitro reconstitution targets to mono-macrocyclic compounds whose gene clusters and bioactivities were firmly established. Our top candidates thus became GE2270A and thiomuracin (**Figure 2.1**) (3). Although mono-macrocyclic thiopeptides are more "minimalistic", they harbor high chemical complexity and biological activity. *Planobispora rosea* ATCC53733 and *Nonomuraea* sp. Bp3714-39 are the canonical producers of GE2270A and thiomuracin, respectively (3). We did not have access to a thiomuracin-producing strain of *Nonomuraea*, but genome-mining efforts predicted that *Thermobispora bispora* DSM 43833 would produce an identical molecule owing to the presence of a nearly identical biosynthetic gene cluster (**Figure 2.1B**). Derived from a thermophilic organism, the *T. bispora* proteins were envisioned to have the desired stability for efficient heterologous expression in *E. coli*.

All genes encoding the biosynthetic proteins from *T. bispora* (*tbt*) were codon optimized for *E. coli* expression. Based on previous reports (11), genes for F-protein dependent cyclodehydratase (*tbtF, tbtG*), thiazoline dehydrogenase (*tbtE*), and the [4+2] cycloaddition enzyme (*tbtD*) were subcloned into a previously-described plasmid that fuses maltose-binding protein (MBP) to the N-terminus of the protein of interest. In contrast, genes encoding the precursor peptide (*tbtA*) and the split LanB-type dehydratase (13) (*tbtB, tbtC*) were expressed as N-terminally His_6-tagged proteins. The purity of all proteins used in the study was assessed visually by Coomassie-stained SDS-PAGE analysis (**Figure 2.2**). The TbtE/F/G proteins were expressed and purified by Graham Hudson in the Mitchell laboratory.

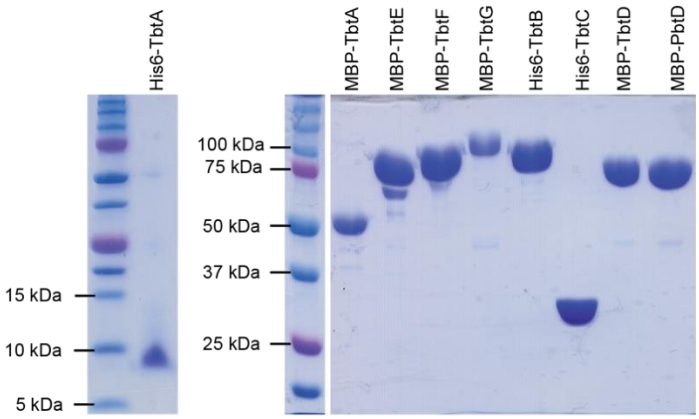


Figure 2.2 SDS-PAGE analysis of proteins used in this study. The minor impurity found at ~45 kDa is the result of endogenous protease activity on the MBP fusion proteins. Cleavage occurs in the linker region between MBP and the protein of interest. Enrichment occurs due to the use of amylose resin for affinity purification.

The precursor peptide TbtA was initially expressed and purified as an N-terminally His_6-tagged fusion protein (**1**) (**Figure 2.3A**). Reaction of TbtA with purified TbtB/C under conditions previously reported for the NisB dehydratase (12, 13) did not result in any chemical transformation (**Figure 2.4**). In contrast, Graham Hudson showed that reaction of **1** with the cyclodehydratase TbtG, the dehydrogenase TbtE and TbtF, an ortholog of HcaF that was recently shown to function as an auxiliary leader peptide binding protein (11), resulted in a mass shift of -120 Da (**Figure 2.3B**). This observation is consistent with successful cyclodehydration and dehydrogenation of the six cysteines in the core peptide to the corresponding thiazoles. When TbtE was omitted, a mass consistent with six thiazoline modifications was observed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Further, when TbtF or ATP were omitted from the reaction, no product formation was detected.

TbtB is an ortholog of the glutamylation domain of the NisB dehydratase that uses glutamyl-tRNA to glutamylate Ser residues (13). TbtC is a member of the SpaB_C family of proteins (PF14028) and has sequence homology to the second domain of NisB that eliminates the

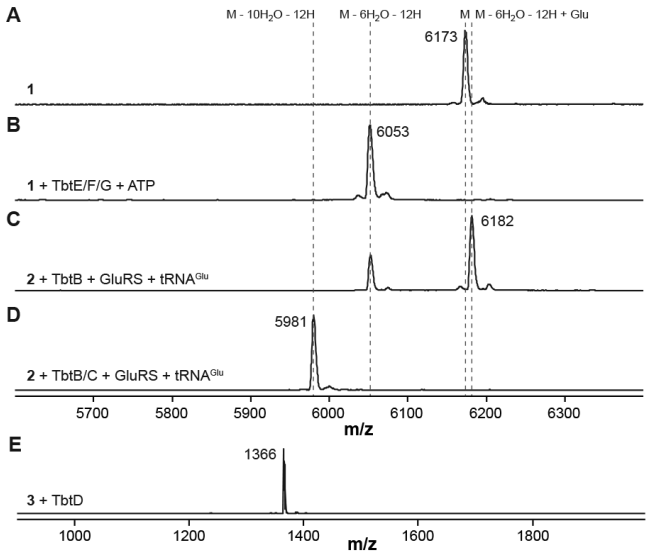
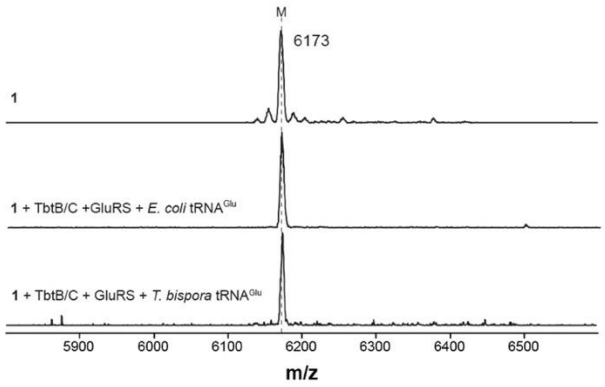


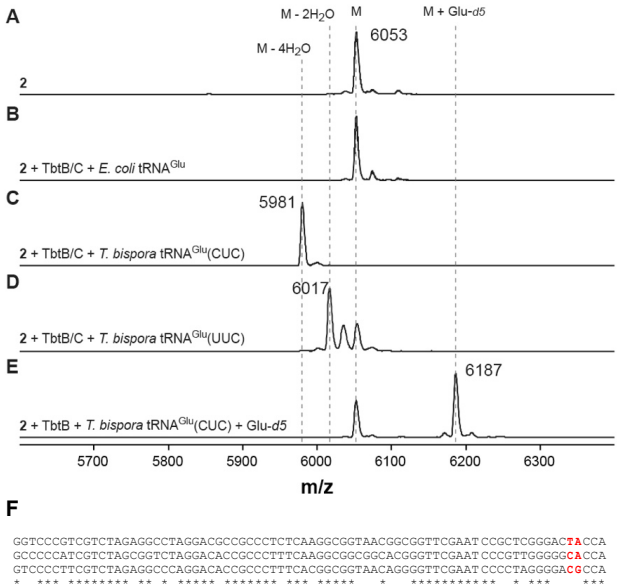
Figure 2.3 MALDI-TOF-MS of (**A**) His_6-TbtA (**1**); calc. *m/z*, 6173; (**B**) hexazole **2**, calc. *m/z*, 6053; (**C**) glutamylated **2**; calc. *m/z*, 6182; (**D**) tetra-dehydrated hexazole **3**, calc. *m/z*, 5981; and (**E**) thiomuracin GZ **4**, calc. *m/z*, 1366. For the reactions shown in panels C and D, ATP, Mg^{2+}, *E. coli* GluRS and *T. bispora* $tRNA^{Glu}$ (CUC) were present.



His_6-TbtA: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

Figure 2.4 TbtB/C do not process unmodified TbtA. The top MALDI-TOF mass spectrum shows unmodified TbtA precursor peptide (**1**). The middle spectrum spectrum shows **1** treated with TbtB/C in the presence of other necessary reaction components [*E. coli* GluRS and *E. coli* $tRNA^{Glu}$]. The bottom spectrum is an equivalent reaction except with *T. bispora* GluRS and *T. bispora* $tRNA^{Glu}$ (CUC). These data suggest that the split LanB dehydratase (TbtB/C) requires a thiazol(in)e-containing substrate, implicating cyclodehydration (TbtF/G) and dehydrogenation (TbtE) as the initial two steps of thiopeptide biosynthesis.

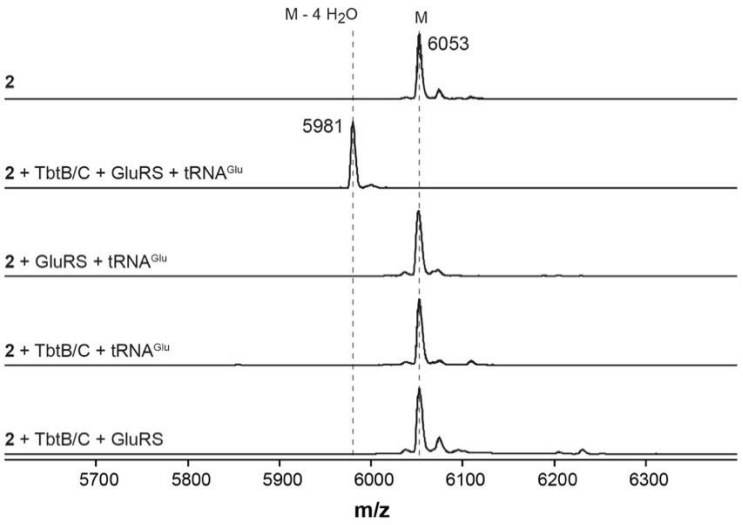
glutamate from the glutamylated Ser to install Dha. Reaction of the purified hexazole-containing intermediate **2** with TbtB in the presence of Glu, ATP, *E. coli* $tRNA^{Glu}$ and *E. coli* glutamate tRNA transferase (GluRS) did not result in any change in mass (**Figure 2.5**). We reasoned that sequence differences between *E. coli* and *T. bispora* $tRNA^{Glu}$ might contribute to the lack of activity. Hence, the two $tRNA^{Glu}$ isoacceptors encoded in the *T. bispora* genome were prepared by in vitro transcription. Reaction of **2** with Glu, ATP, *T. bispora* $tRNA^{Glu}$ (CUC) and *E. coli* GluRS resulted in glutamylation of **2** (**Figure 2.3C**). When TbtC was added to an otherwise identical reaction, four-fold dehydration of **2** was observed (**Figure 2.3D**).



His$_6$-TbtA: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

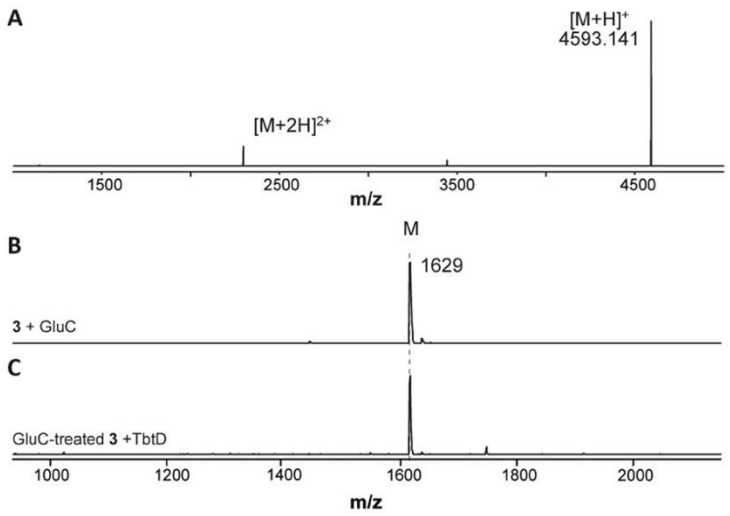
Figure 2.5 TbtB/C is selective for a particular $tRNA^{Glu}$. (**A**) MALDI-TOF mass spectrum of **2** (starting material). (**B**) *E. coli* $tRNA^{Glu}$ (UUC) is not accepted by TbtB/C. (**C**) In contrast, four dehydrations are observed upon treating **2** with TbtB/C and $tRNA^{Glu}$ (CUC) from *T. bispora*. (**D**) An identical reaction with the alternative $tRNA^{Glu}$ (UUC) from *T. bispora* yielded a mixture of dehydrated species, primarily the didehydrated peptide. (**E**) A reaction that omitted TbtC (elimination domain), while employing the optimal tRNA and d_5-Glu, yielded a mono-glutamylated species. (**F**) Sequence alignment of the three pertinent $tRNA^{Glu}$. Top, *T. bispora* $tRNA^{Glu}$ (CUC); middle, *T. bispora* $tRNA^{Glu}$ (UUC); bottom, *E. coli* $tRNA^{Glu}$.

Use of *T. bispora* $tRNA^{Glu}$ (UUC) resulted only in partial dehydration (**Figure 2.5**); this result may either indicate that TbtB preferentially uses *T. bispora* $tRNA^{Glu}$ (CUC) or that *E. coli* GluRS is less efficient in charging *T. bispora* $tRNA^{Glu}$ (UUC). Using the same conditions with **1** did not result in an observable reaction (**Figure 2.4**), demonstrating for the first time that dehydration activity requires prior cyclodehydration. Further, all assay components were required for activity (**Figure 2.6**). These experiments establish the first activity of a split LanB and extend the use of Glu-$tRNA^{Glu}$ to dehydration during thiopeptide biosynthesis.



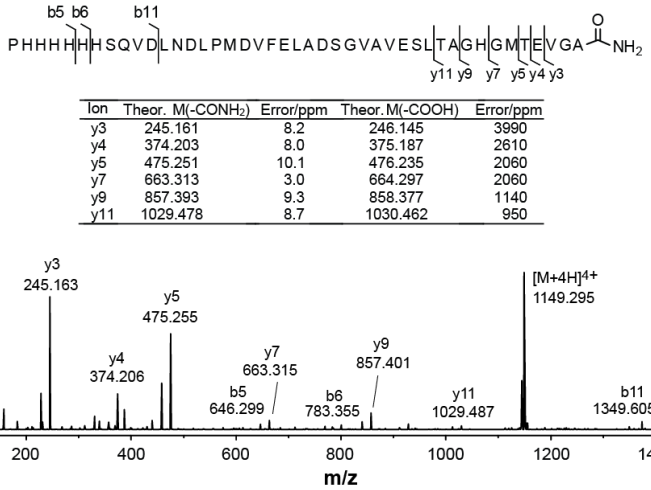
His_6-TbtA: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

Figure 2.6 MALDI-TOF-MS analysis of TbtB/C reaction requirements. Shown in the top spectrum is hexazole **2**, which is the six-fold cyclodehydration and dehydrogenation product of TbtA. Reaction of **2** with the split LanB dehydratase (TbtB/C) in the presence of *E. coli* GluRS and *T. bispora* $tRNA^{Glu}$ (CUC) results in a four-fold dehydrated product (**3**). The formation of **3** was dependent on the presence of all reaction components, as individual omission of TbtB/C, *E. coli* GluRS, or *T. bispora* $tRNA^{Glu}$ (CUC) did not result in any change in mass. Residues modified to a thiazole are shown in blue, and residues that are transformed to dehydroalanine are colored green. All reactions also contained Glu, ATP, and $MgCl_2$ (see methods).



His$_6$-TbtA: PHHHHHHS*Q*VDLNDLPMDVFELADSGVAVESLTAGHGMTE**VGA***SC**N**C**F**C**Y**I**C**C**S**C**S**S**A

Figure 2.7 TbtD requires the leader peptide for activity. (**A**) HR-ESI-MS of the TbtA leader peptide generated by macrocyclization (theoretical $[M+H]^+$, 4593.143 Da, 0.4 ppm). The sequence of the His$_6$-TbtA-derived substrate **3** is given below the figure with modifications present color-coded (blue, thiazoles; green, dehydroalanines). The asterisk designates the end of the leader peptide. Shown italicized are additional N-terminal residues present on the substrate (His-tag and cloning remnants). (**B**) MALDI-TOF-MS of **3** after treatment with endoproteinase GluC. This treatment removes all but three residues (VGA, boldfaced and underlined) of the leader peptide. (**C**) Reaction of GluC-treated **3** with TbtD. No reaction was observed.



Ion	Theor. M(-CONH$_2$)	Error/ppm	Theor. M(-COOH)	Error/ppm
y3	245.161	8.2	246.145	3990
y4	374.203	8.0	375.187	2610
y5	475.251	10.1	476.235	2060
y7	663.313	3.0	664.297	2060
y9	857.393	9.3	858.377	1140
y11	1029.478	8.7	1030.462	950

Figure 2.8 ESI-MS/MS of the His$_6$-TbtA leader peptide (*m/z* 1149) generated by macrocyclization.

Treatment of the hexazole-containing, tetra-dehydrated His_6-TbtA (**3**) with TbtD resulted in the consumption of **3** and appearance of two new masses consistent with the excised leader peptide and the desired macrocycle (**Figure 2.3E & 2.7**). The identity of the leader peptide was confirmed by high-resolution, tandem electrospray ionization MS (HR-ESI-MS/MS, **Figure 2.8**). The data indicated that the leader peptide terminates in a carboxamide rather than a carboxylic acid, providing the first evidence that the cycloaddition enzyme uses elimination chemistry, rather than hydrolysis, to release the leader peptide. The generation of the macrocyclic core scaffold of thiocillin by the TbtD-ortholog TclM using a semi-synthetic substrate was also reported shortly before our work appeared (24, 34). Interestingly, like TbtC, TbtD (and TclM) is a member of the SpaB_C protein family, suggesting that elimination reactions may be common in PF14028 (13). Removal of the N-terminal 31 (of 34) amino acids of the leader peptide from **2** by treatment with endo-proteinase GluC resulted in a peptide that was no longer a substrate for TbtD (**Figure 2.7**). In NisB, where the glutamylation enzyme is covalently linked to the SpaB_C elimination domain, leader peptide binding occurs in the glutamylation domain (13, 25). To probe whether TbtB, which contains this binding domain, would accelerate catalysis by TbtD, TbtB was added to the cycloaddition reaction. Analysis of macrocycle formation suggested that TbtB afforded no rate enhancement (**Figure 2.9**). The same was observed upon addition of TbtF, which also contains a leader peptide-binding domain (11). Thus it appears that TbtD independently engages the leader peptide by a different strategy that will be discussed in Chapter 4.

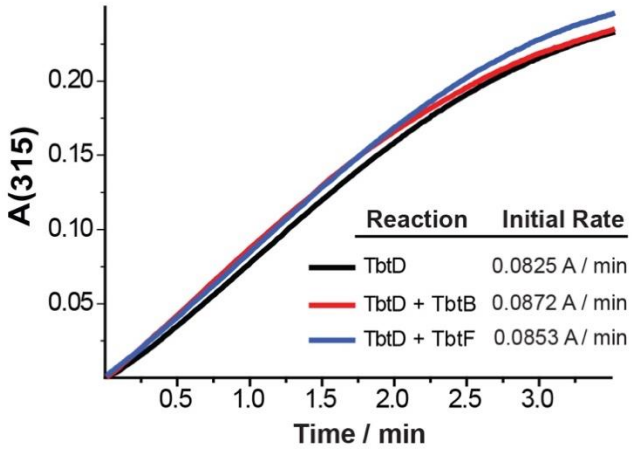


Figure 2.9 TbtD activity is not enhanced by known leader peptide-binding Tbt proteins. Addition of TbtB or TbtF did not increase the rate of macrocycle formation by TbtD as monitored by UV-Vis spectroscopy at 315 nm (owing to the generation of the tri-thiazole-substituted pyridine). Both TbtB and TbtF contain RiPP leader peptide Recognition Elements (RREs) while TbtD does not appear to, despite its activity being leader peptide-dependent (**Figure 2.7**).

To obtain the quantities of **4** required for structure confirmation, the entire in vitro process was optimized and scaled up using MBP-TbtA as the starting material by Graham Hudson. The molecular formula of **4** was deduced by HR-ESI-MS, and its structure was confirmed by Dr. Jonathan Tietz using NMR spectroscopy. The overall chemical shift values were highly consistent with those of the closely-related thiomuracin series reported by Morris et al (3). Correlations obtained via ^{1}H-^{1}H COSY, ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H NOESY, ^{1}H-^{13}C HSQC, and ^{1}H-^{13}C HMBC experiments (**Figure 2.10**) allowed assignment of all ^{1}H and ^{13}C resonances in the molecule, with the exception of the carboxylic acid proton, which was not observed.

To assess the bioactivity of **4**, Graham Hudson determined the minimal inhibitory concentration (MIC) towards a panel of Gram-positive, Gram-negative, and fungal species by microbroth dilution (**Table 2.1**). While **4** was inactive against the tested Gram-negative bacteria and fungi, growth inhibition was found for all Gram-positive bacteria in the panel. The most potent activity was toward vancomycin-resistant *Enterococcus faecium* U503 (VRE), followed by

methicillin-resistant *Staphylococcus aureus* USA300 (MRSA). Despite lacking other posttranslational modifications including C-methylation of thiazole, beta-hydroxylation of Phe, and oxidation/cyclization of Ile, the bioactivity of **4** is on par with fully modified thiomuracin A, which has MICs of 0.25 µg/mL and 0.5 µg/mL against VRE and MRSA, respectively (3, 26). We termed the structure of **4** thiomuracin GZ.

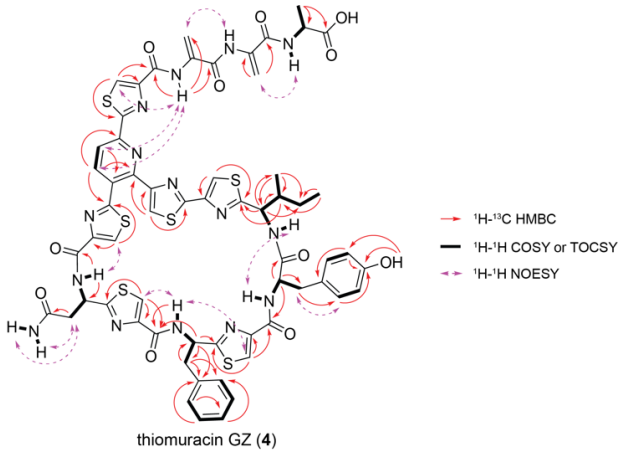
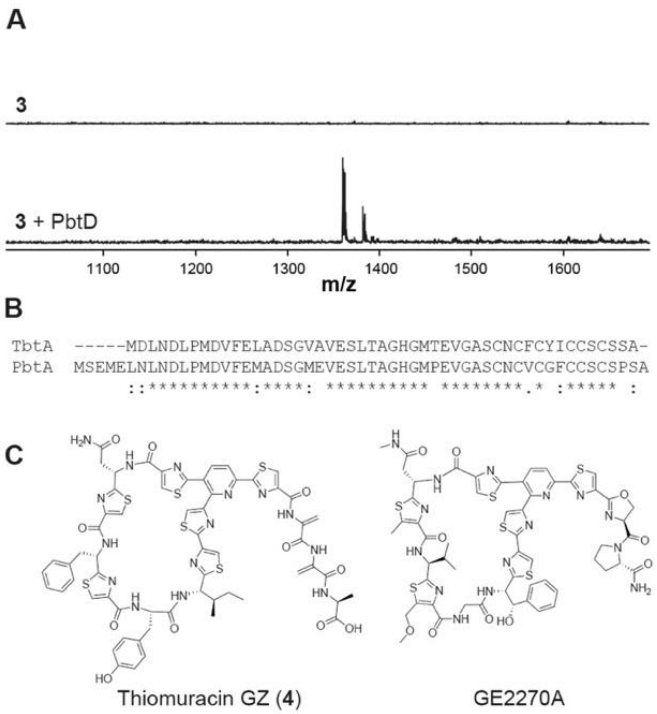


Figure 2.10 NMR correlations used to assign the structure of **4**. Data collected and figure made by Dr. Jonathan Tietz.

Table 2.1 Bioactivity of thiomuracin GZ. [a.] MIC, minimum inhibitory concentration. Data collected and table made by Graham Hudson.

Organism	*MIC[a] (µg/mL)*	*MIC[a] (µM)*
Staphylococcus aureus USA300	*0.25*	*0.18*
Enterococcus faecium U503	*0.063*	*0.046*
Bacillus anthracis str. Sterne	*2*	*1.5*
Escherichia coli MC4100	*>64*	*>47*
Pseudomonas aeruginosa PA01	*>64*	*>47*
Mycobacterium smegmatis B-14616	*>64*	*>47*
Aspergillus niger	*>128*	*>94*
Saccharomyces cerevisiae	*>128*	*>94*
Talaromyces stipitatus	*>128*	*>94*



His$_6$-TbtA: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

Figure 2.11 PbtD accepts **3** as a substrate. (**A**) The amino acid composition of His$_6$-TbtA is shown. In peptide **3** the residues shown in blue are thiazoles and in green are dehydroalanines. In the top spectrum, **3** is not visible owing to the mass window shown. Reaction of **3** with PbtD (homolog of TbtD from the GE2270A biosynthetic gene cluster) shows the formation of *m/z* 1366 Da (compare to **Figure 2.3E**) and the corresponding sodium adduct. (**B**) The aligned amino acid sequences of the thiomuracin GZ and GE2270A precursor peptides, labeled TbtA and PbtA, respectively. (**C**) Structural comparison between thiomuracin GZ and GE2270A. While structurally related, there exist key differences in amino acid composition and post-translational modifications.

GE2270A has a similar macrocyclic structure to thiomuracin (**Figure 2.11**), differing in three of the four amino acids that are present in the macrocycle (3). We therefore wondered whether the TbtD ortholog in the GE2270A biosynthetic gene cluster (PbtD) would also be able to catalyze the non-cognate [4+2] cycloaddition of **3**. Reaction of PbtD with **3** indeed resulted in the clean formation of **4** with concomitant elimination of the leader peptide (**Figure 2.11**). This observation is promising with respect to thiopeptide engineering for structure-activity relationship

efforts, because it illustrates that the [4+2] enzyme is tolerant of substantial changes in the amino acid sequence of the macrocycle (Phe, Tyr, Ile in thiomuracin are Val, Gly, Phe in GE2270A, respectively).

2.2.2 Heterologous expression of thiomuracin in *E. coli*

After the successful in vitro reconstitution of the activities of six proteins that together biosynthesize the core structure of thiomuracin, I further went on to build an *E. coli* heterologous producing system using the knowledge obtained from the in vitro study to lay the basis for the potential production of a library of thiopeptide analogues (**Figure 2.12**).

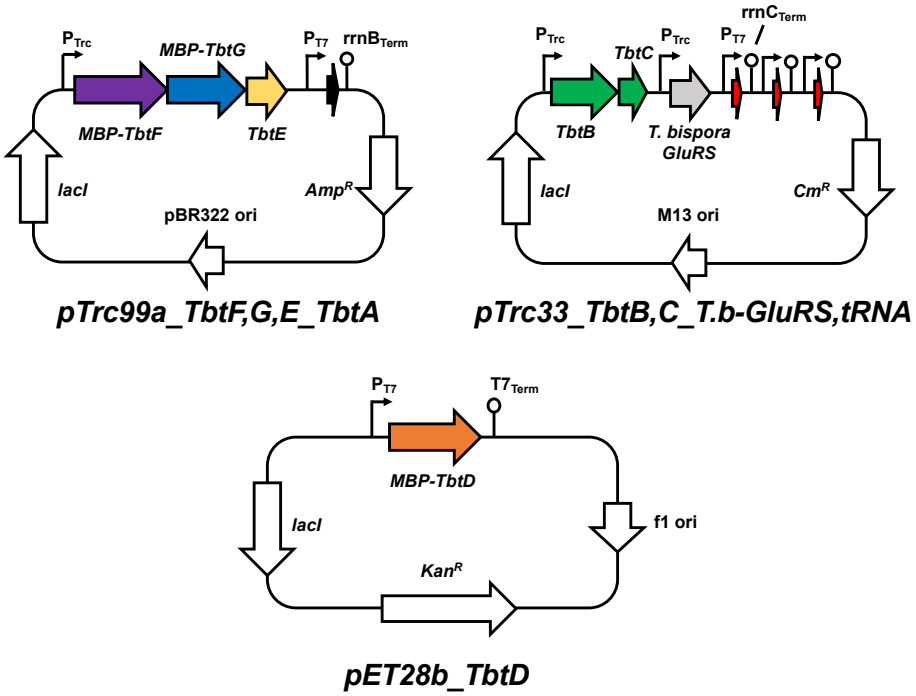


Figure 2.12 Representation of the plasmids used for the heterologous expression experiments in *E. coli* for thiomuracin production.

A plasmid compatible with the T7 expression system in *E. coli* was designed using the *pTrc99a* vector backbone. Based on our observation that the cyclodehydration reactions precede the formation of Dha residues, I first cloned the genes encoding the TbtF, TbtG, and TbtE proteins essential for thiazole installation into the plasmid. Both TbtF and TbtG were cloned as N-terminal MBP fusion proteins, because this should help solve any expression and solubility issues of the modification enzymes while retaining their activity, as our work demonstrated that MBP-tagged TbtF, TbtG, and TbtE can be purified in high yield from *E. coli* without losing their in vitro catalytic ability. The precursor peptide gene *tbtA* was cloned with an N-terminal His_6-tag. The expression of the three modification enzymes was under the control of an IPTG inducible Trc promoter, whose transcription intensity is five times weaker than that of the T7 promoter, while the precursor peptide gene was cloned under control of a T7 promoter. This design was based on the potent activity of the enzymes involved in thiazole formation, and suppressing the transcription of these modification enzymes while increasing the transcription efficiency of the precursor peptide would result in higher yield of the product. To test the formation of six thiazoles in TbtA, a heterologous expression experiment was performed with the newly constructed *pTrc99a_TbtF,G,E_TbtA* plasmid in *E. coli*. A product consistent with six thiazoles installed onto the TbtA precursor peptide was observed by MALDI-TOF MS analysis after immobilized metal affinity chromatography (IMAC; **Figure 2.13B**).

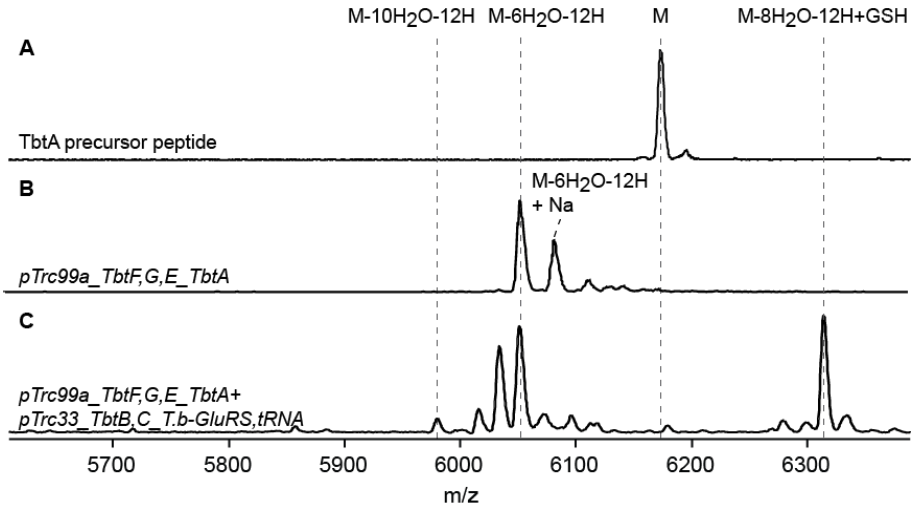


Figure 2.13 MALDI-TOF-MS analysis of (**A**) purified His_6-TbtA, (**B**) product from the expression of the plasmid *pTrc99a_TbtF,G,E_TbtA*, and (**C**) the purified co-expression product of both *pTrc99a_TbtF,G,E_TbtA* and *pTrc33_TbtB,C_T.b-GluRS,tRNA* in *E. coli*.

To introduce the lanthipeptide dehydratases responsible for the formation of the four Dha into the system, another plasmid compatible with the T7 expression system in *E. coli* was designed using the *pTrc33* vector backbone. Similar to the cyclodehydratase and dehydrogenase, the genes for the split lanthipeptide dehydratase TbtB and TbtC were also cloned under the control of a Trc promoter. Since the in vitro studies of TbtB have clearly shown that *E. coli* $tRNA^{Glu}$ is not accepted, the gene for the $tRNA^{Glu}$ (CUC) from the thiomuracin native producer *T. bispora* was introduced into the system to ensure enzymatic activity. Each gene encoding a tRNA copy was flanked with a T7 promoter at the 5' end and an rrnC terminator sequence at the 3'end. To ensure that sufficient level of *T. bispora* glutamyl-$tRNA^{Glu}$ was available during the co-expression experiment, the gene for the *T. bispora* GluRS was also cloned into the plasmid (27). To test the dehydration efficiency of this system, a heterologous expression experiment was performed with both the *pTrc99a_TbtF,G,E_TbtA* and *pTrc33_TbtB,C_T.b-GluRS,tRNA* plasmids in *E. coli*. As shown by MALDI-TOF-MS analysis after IMAC, only a very small amount of the four-fold

dehydrated species could be detected, with the major product carrying two dehydrations and a glutathione (GSH) addition (**Figure 2.13C**). The major product was analyzed by HR-ESI-MS/MS, and the GSH was found attached to the dehydrated Ser14 residue (**Figure 2.14**). I reasoned that the GSH addition may hamper the downstream dehydration reaction, thus resulting in incomplete dehydration. However, attempts to use the *E. coli* SHuffle strain, which overexpresses a disulfide isomerase that helps reduce the GSH concentration inside the cell, did not help preventing the GSH adduct.

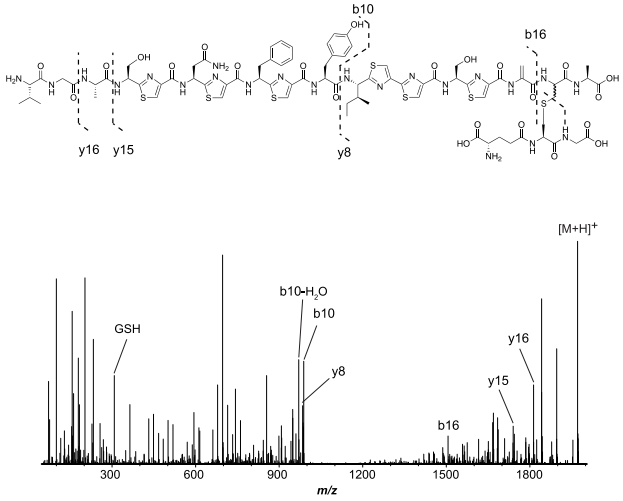


Figure 2.14 HR-ESI-MS/MS analysis of the GluC-digested GSH-added di-dehydrated TbtA hexazole. Tandem MS of the ion of *m/z* 1971 Da gives a fragmentation pattern consistent with GSH addition at the Ser14 position.

In vitro assays with different concentrations of GSH indicated that, even with up to 50 mM of GSH in the reaction, no significant inhibition of the dehydration reaction was detected. The GSH was found to add to the dehydrated product only after all the dehydration steps had finished (**Figure 2.15**). These data suggest that the incomplete dehydration during *E. coli* heterologous expression may result from the poor activity of the dehydratases, which could result from a low effective concentration of *T. bispora* glutamyl-tRNA^{Glu} in the *E. coli* cytoplasm.

Finally, we probed whether the *E. coli* heterologous expression was able to produce the thiomuracin macrocycle or not. The gene for the [4+2] aza-cycloaddition enzyme TbtD was cloned into a *pET28b* vector backbone that is compatible with the other two plasmids. After co-expressing all three plasmids in *E. coli*, the thiomuracin macrocycle was successfully produced, as shown by UPLC-HR-ESI-MS after purification of the peptide from the cell pellet (**Figure 2.16**). However, the yield was very low, possibly because the *E. coli* we used for thiomuracin production was not a resistant strain. We also detected a very small amount of thiomuracin from the growth media.

Many problems still remain to be solved for the construction of an effective heterologous expression system, such as optimization of this system to overcome the GSH addition issue to push the dehydration reaction to completion.

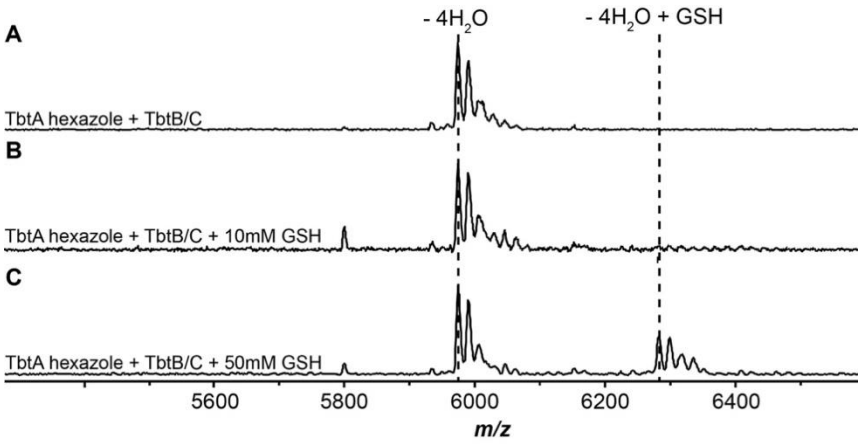


Figure 2.15 MALDI-TOF-MS analysis of TbtB/C dehydration reaction in the absence (**A**) and in the presence of 10 mM (**B**) and 50 mM (**C**) of GSH in the reaction mixture.

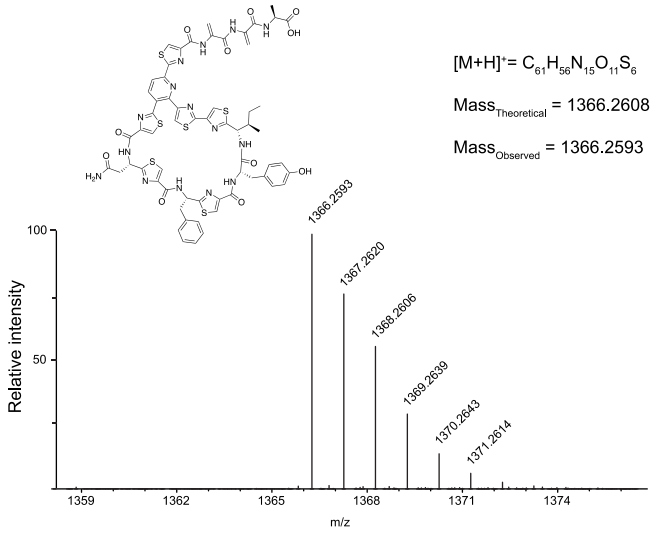


Figure 2.16 HR-ESI-MS analysis of the product formed by co-expression of *pTrc99a_TbtF,G,E_TbtA*, *pTrc33_TbtB,C_T.b-GluRS,tRNA* and *pET28b_TbtD* in *E. coli*.

Table 2.2 Plasmids constructed in this chapter.

plasmid name	marker	origin	restriction site used	insert
His$_6$-TbtA-pRSFDuet-1	Kan	RSF	BamHI HindIII	*tbtA*
His$_6$-TbtB-pET28a	Kan	f1	BamHI XhoI	*tbtB*
His$_6$-TbtC-pET28a	Kan	f1	BamHI XhoI	*tbtC*
His$_6$-TbtD-pET28a	Kan	f1	BamHI XhoI	*tbtD*
His$_6$-PbtD-pET28a	Kan	f1	BamHI XhoI	*pbtD*
pTrc99a_TbtF,G,E_TbtA	Amp	pBR322	NcoI SalI	*tbtF, tbtG, tbtE, tbtA*
pTrc33_TbtB,C_T.b-GluRS,tRNA	Cm	M13	BamHI XhoI	*tbtB, tbtC, T.bispora GluRS, T.bispora tRNA^{Glu}(CUC)*
pET28b_TbtD	Kan	f1	BamHI XhoI	*tbtD*

2.3 Experimental methods

General materials and methods

Reagents used for molecular biology experiments were purchased from New England BioLabs (Ipswich, MA), Thermo Fisher Scientific (Waltham, MA), or Gold Biotechnology Inc. (St. Louis, MO). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). *Escherichia coli* DH5α and BL21 (DE3) strains were used for plasmid maintenance and protein overexpression, respectively. *Thermobispora bispora* DSM 43833 was obtained from DSMZ (Braunschweig, Germany). *Planobispora rosea* ATCC53733 was obtained from the USDA Agricultural Research Service (ARS) Culture Collection. Plasmid inserts were sequenced at ACGT Inc. (Wheeling, IL). MALDI-TOF-MS analysis was performed using a Bruker UltrafleXtreme MALDI TOF-TOF mass spectrometer (Bruker Daltonics) in reflector positive mode at the University of Illinois School of Chemical Sciences Mass Spectrometry Laboratory. ESI-MS/MS analyses were performed using a SYNAPT ESI quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters). Kinetic assays were performed using a Varian Cary 4000 UV-Vis spectrophotometer. HiTrap columns for Ni-NTA affinity chromatography were purchased from GE Healthcare.

Molecular biology techniques

Oligonucleotides were purchased from Integrated DNA Technologies Inc. (Coralville, IA). Primers for *E. coli* $tRNA^{Glu}$ were designed and the dsDNA template for in vitro transcription was prepared as previously described (28). *Thermobispora bispora* was identified as a potential producer of a thiomuracin-like thiopeptide using BLAST to identify homologs of genes from the

relevant biosynthetic gene cluster of thiomuracin producer *Nonomurea* sp. Bp3714-39 (3). Genes optimized for recombinant expression in *Escherichia coli* were synthesized by GenScript (Piscataway, NJ) in *pUC57* (kanamycin, Kan) vectors with BamHI and XhoI sites flanking each gene at the 5′ and 3′ ends (**Table 2.4**), respectively. *E. coli* DH5α were transformed with *pUC57-Kan* vectors containing each gene for replication and subsequent isolation using a QIAprep Spin Miniprep Kit (Qiagen). The isolated DNA was then treated with BamHI-HF and XhoI-HF (New England Biolabs). The digested genes were separated on a 1% (w/v) agarose gel, purified using a QIAQuick gel extraction kit (Qiagen), and ligated into an appropriately endonuclease-digested and gel-purified *pET28a* vectors using T4 DNA ligase (NEB). Ligation reactions were used to transform chemically competent DH5α cells, which were plated on Luria-Bertani (LB) agar plates containing 50 μg/mL kanamycin and grown at 37 °C. Colonies were picked at random and grown in LB broth for 16–20 h prior to plasmid isolation using a QIAprep Spin Miniprep Kit. For TbtA and TbtD, the vector encoded a tobacco etch virus (TEV) cleavable, N-terminal maltose binding protein (MBP) affinity tag, while TbtE, TbtF, and TbtG constructs did not encode a TEV proteolysis site. All recombinant constructs featuring an MBP-tag were sequenced using a custom MBP forward primer, whereas constructs encoding His_6-tagged proteins were sequenced using the T7 forward primer. All constructs were reverse-sequenced using a T7 reverse primer as well as an internal primer if necessary (**Table 2.3**).

The gene encoding PbtD (TbtD ortholog) from *Planobispora rosea* ATCC 23866 was amplified from genomic DNA using the primers listed in **Table 2.3**. The resulting DNA was purified using a QIAprep Spin Miniprep Kit. The purified DNA was then digested using BamHI-HF and XhoI-HF before being ligated into an appropriately digested *pET28b* vector encoding an

N-terminal MBP tag featuring a TEV proteolysis site. The resulting construct was then used to transform *E. coli*, propagated, and sequenced as described above (notebook reference 6/5/2014).

The *tbtA* gene was cloned into *pRSFDuet-1* with an *N*-terminal His_6-tag following standard cloning procedures. In addition, the GlySerSer residues preceding the His_6 tag were mutated to a single Pro to suppress gluconoylation (29). The PCR product was purified by gel extraction on a 1% (w/v) agarose gel using the QIAquick Gel Extraction Kit (Qiagen). The vector *pRSFDuet-1* was digested using BamHI-HF and HindIII-HF (NEB) restriction endonucleases and purified by gel extraction as described above. Insertion of the DNA fragment was achieved by Gibson one-step isothermal DNA assembly as previously described (30). An aliquot (20 μL) of the Gibson assembly reaction was used to transform *E. coli* DH5α cells using the heat shock method. The cells were plated on LB plates supplemented with kanamycin (50 μg/mL) and single colonies were grown at 37 °C for 12–15 h. The plasmid *His_6-TbtA-pRSFDuet-1* was isolated using a QIA Spin Miniprep Kit. Insert integrity was verified by sequencing the plasmids with the appropriate primers (**Table 2.3**) (notebook reference 11/26/2014).

The gene coding for the dehydratase TbtB was cloned into *pET28a* encoding an N-terminal His_6-tag. The *tbtB* gene was amplified by PCR, and purified by gel extraction from a 1% (w/v) agarose gel using the QIAquick Gel Extraction Kit. The vector *pET28a* and the *tbtB* PCR fragment were digested using BamHI-HF and XhoI-HF (NEB) restriction endonucleases, and purified by gel extraction as described above. The DNA fragment was inserted by Gibson one-step isothermal DNA assembly, and an aliquot of 20 μL of the Gibson assembly reaction was used to transform *E. coli* DH5α cells using the heat shock method. The cells were plated on LB plates supplemented with kanamycin (50 μg/mL) and the plates were incubated at 37 °C for 12–15 h. Single colonies were picked and grown in LB supplemented with kanamycin (50 μg/mL) at 37 °C for 12–15 h and

the plasmid *His$_6$-TbtB-pET28a* was isolated using a QIA prep Spin Miniprep Kit. Insert integrity was verified by sequencing the plasmids with the appropriate primers (**Table 2.3**). All the other genes and their corresponding His$_6$-tagged proteins were cloned using the same procedures (notebook reference 11/26/2014).

Expression, and purification of His$_6$-TbtA 1

E. coli BL21 (DE3) cells (50 μL) were electroporated with *His$_6$-TbtA-pRSFDuet-1* (50 ng), plated on LB agar plates supplemented with kanamycin (as above) and grown at 37 °C for 12–15 h. A single colony was used to inoculate 60 mL of LB broth supplemented with kanamycin, grown for 12–15 h at 37 °C, and the culture was used to inoculate 6 L of LB media, supplemented with kanamycin, to an OD_{600} of 0.025. Cultures were grown at 37 °C to an OD_{600} of 1.0. Peptide expression was induced by the addition of IPTG to a final concentration of 1 mM and cultures were grown at 37 °C for 3 h. Peptide purification was performed following a previously described method (31) (notebook reference 12/2/2014).

Expression, and purification of His$_6$-TbtB

E. coli BL21 (DE3) cells (50 μL) were electroporated with *His$_6$-TbtB-pET28a* (50 ng), and cells were plated on LB agar plates supplemented with kanamycin and grown at 37 °C for 12–15 h. A single colony was used to inoculate 20 mL of LB broth supplemented with kanamycin, grown for 12–15 h at 37 °C, and the culture was used to inoculate 2 L of TB media, supplemented with kanamycin, to an OD_{600} of 0.025. Cultures were grown at 37 °C to a final OD_{600} of 0.6–0.8. Protein overexpression was induced by the addition of IPTG to a final concentration of 0.2 mM, and cultures were grown at 18 °C for 18 h. Protein purification was performed following a similar

protocol as described before (12). All other His_6-tagged proteins were purified using the same procedure (notebook reference 1/10/2015).

In vitro transcription of *T. bispora* $tRNA^{Glu}$ (CUC)

Primers for *T. bispora* $tRNA^{Glu}$ were designed according to a previously described method (28). The $tRNA^{Glu}$ dsDNA template was generated from two overlapping synthetic deoxyoligonucleotides (**Table 2.3**). To prepare the dsDNA template for in vitro transcription, 5′ overhangs were filled in using the following conditions: NEB Buffer 2 (1×), primers (4 μM each), dNTP (100 μM each), and DNA polymerase I large (Klenow) fragment (1 U μg^{-1} DNA) in a final volume of 50 μL. The reaction was incubated at 25 °C for 15 min, quenched with EDTA (10 mM) at 75 °C for 25 min, and dsDNA $tRNA^{Glu}$ template was precipitated with cold EtOH overnight. In vitro transcription was performed using a previously described method (32). The transcribed $tRNA^{Glu}$ was then purified by acidic phenol extraction using a previously described method (33) (notebook reference 2/15/2015).

In vitro dehydration of 2 to produce 3

The following reaction conditions were used for dehydration assays: HEPES pH 7.5 (100 mM), $MgCl_2$ (5 mM), KCl (150 mM), glutamate (10 mM), ATP (5 mM), TbtA hexazole (50 μM), *T. bispora* $tRNA^{Glu}$ (10 μM), *E. coli* GluRS (10 μM), TbtB (5 μM), TbtC (5 μM), and thermostable inorganic pyrophosphatase (TIPP; 0.02 U μL^{-1}) in a final volume of 25 μL. The reaction mixture was incubated at 30 °C for 3 h, centrifuged to remove insoluble material (14,000 × *g*, 5 min, 25 °C), and desalted using C18 ZipTips (EMD Millipore). The sample was mixed in a 1:1 ratio with

2,5-dihydroxybenzoic acid matrix, spotted on a Bruker MALDI plate, and analyzed by MALDI-TOF-MS (notebook reference 2/30/2015).

Purification of 3

The reaction mixtures to produce **3** were lyophilized, and the residue was extracted with a volume of 40% aq. MeCN equal to the volume used for the dehydration reaction. The insoluble material was removed by centrifugation at 10,000 × *g* for 10 min, and the supernatant was purified on a Shimadzu Prominence Preparative Liquid Chromatography system equipped with a Phenomenex Luna C18 column (250 × 10 mm, 10 μm particle size, 100 Å pore size). Acetonitrile and 10 mM aq. NH_4HCO_3 were used as the mobile phases, and a gradient of 2-80% aq. MeCN over 45 min at 1 mL/min was used for separation (notebook reference 3/18/2015).

Purification of TbtA-hexazole-tetradehydrated core peptide

HPLC-purified **3** was dissolved in 50 mM Tris-HCl (pH 7.5) to a final concentration of 100 μM. GluC endoproteinase (Worthington Biochemical Corporation) was added to a concentration of 0.02 mg/mL, and the reaction mixture was incubated at 37 °C for 1 h, followed by addition of MeCN to 50% (v/v) to precipitate proteins. The insoluble material was removed by centrifugation at 16,100 × *g* for 10 min, and the supernatant was purified on an Agilent 1260 Infinity HPLC system with a Phenomenex Luna C18 column (250 × 4.6 mm, 10 μm particle size, 100 Å pore size). Acetonitrile and 10 mM NH_4HCO_3 were used as the mobile phases, and a gradient of 2–80% aq. MeCN over 45 min at 1 mL min^{-1} was used for separation (notebook reference 3/27/2015).

In vitro production of 4

Purified **2** was dissolved in DMSO to a concentration of 1 mM. The in vitro production of macrocycle was performed using: HEPES pH 7.5 (100 mM), $MgCl_2$ (5 mM), KCl (150 mM), ATP (5 mM), **2** (100 μM), *T. bispora* $tRNA^{Glu}$ (10 μM), *E. coli* GluRS (10 μM), TbtB (5 μM), TbtC (5 μM), MBP-TbtD (5 μM), and TIPP (0.02 U μL^{-1}). The reaction mixture was allowed to proceed at 30 °C for 4 h (notebook reference 4/16/2015).

UPLC-MS and MS/MS of ejected His_6-TbtA leader

HPLC-Purified ejected His_6-TbtA leader after the macrocylization reaction was dissolved in water, and LC-MS/MS was performed using the same procedure as above. A gradient of 2–100% aq. MeCN with 0.1% formic acid (v/v) over 20 min was used. A ramping of cone voltage of 16–19 kV during the scan was performed to generate peptide fragments for MS/MS analysis (notebook reference 4/22/2015).

Kinetic assay to probe if TbtB/TbtF accelerate the macrocyclization reaction

HPLC-purified **3** was dissolved to a concentration of 200 μM in 50 mM Tris-HCl, 150 mM KCl, pH 7.5. Before starting the reaction, each enzyme and the substrate were incubated at 30 °C separately for 5 min. MBP-TbtF was pre-treated with thrombin in order to release the MBP tag. The reaction was performed with 150 μM of substrate **3**, 5 μM of TbtD and either TbtB or thrombin-treated TbtF protein at 30 °C. Macrocycle formation was monitored by absorbance at 315 nm. Time points were collected every 2 s and initial reaction rates were calculated (notebook reference 9/19/2015).

Construction of the thiomuracin *E. coli* heterologous expression system

The genes coding for the MBP-tagged TbtF, MBP-tagged cyclodehydratase TbtG, dehydrogenase TbtE, dehydratases TbtB/C, *T. bispora* GluRS, *T. bispora* $tRNA^{Glu}$, and [4+2] cyclase TbtD, were amplified by PCR using appropriate primers (**Table 2.3**), and purified by gel extraction from a 1% (w/v) agarose gel using the QIAquick Gel Extraction Kit. The vector *pTrc99a* was digested using NcoI-HF and SalI-HF (NEB) restriction endonucleases, while both the *pTrc33* and *pET28b* plasmids were linearized by digestion using BamHI-HF and HindIII-HF (NEB) restriction endonucleases. All three plasmids were purified by gel extraction as described above. The DNA fragment was inserted by Gibson one-step isothermal DNA assembly, and an aliquot of 20 μL of the Gibson assembly reaction was used to transform *E. coli* DH5α cells using the heat shock method. The cells were plated on LB plates supplemented with ampicillin (100 μg/mL) for *pTrc99a*, chloramphenicol (25 μg/mL) for *pTrc33* and kanamycin (50 μg/mL) for *pET28b*, and the plates were incubated at 37 °C for 12–15 h. Single colonies were picked and grown in LB supplemented with the corresponding antibiotics at 37 °C for 12–15 h and the plasmids were isolated using a QIA prep Spin Miniprep Kit. Insert integrity was verified by sequencing the plasmids with the appropriate primers (**Table 2.3**) (notebook reference 2/13/2015).

Heterologous expression and purification of co-expression products

E. coli BL21 (DE3) cells were co-transformed with the heterologous expression plasmids. Cells were grown for 16 h on LB agar plates containing corresponding antibiotics at 37 °C. Single colonies were used to inoculate LB and the cells were grown at 37 °C for 16 h. This culture was used for 1:100 inoculation into Terrific Broth (24 g/L yeast extract, 12 g/L tryptone, 0.4% glycerol (v/v), 17 mM KH_2PO_4, and 72 mM K_2HPO_4) and grown at 37 °C to an optical density at 600 nm

(OD_{600}) of 0.6 before being placed on ice for 20 min. Production of the peptide of interest was then induced with the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The expression was allowed to proceed for 16-18 h at 18 °C. Cells were harvested by centrifugation at 5,000 × *g* for 10 min. Cell pellets were resuspended in LanA start buffer (20 mM Na_2HPO_4, 150 mM NaCl, 20% glycerol, pH 7.5). Cells were homogenized by sonication (50% amplitude, pulse on: 2.0 s, pulse off: 5.0 s, total pulse time: 5 min). Insoluble debris was removed by centrifugation at 13000 × g for 30 min and the supernatant was then applied to Ni resin pre-equilibrated with LanA start buffer. The column was washed with 20 column volumes (CV) of LanA washing buffer (20 mM Na_2HPO_4, 30 mM imidazole, 300 mM NaCl, pH 7.5). The peptides of interest were eluted with elute buffer (20 mM Na_2HPO_4, 500 mM imidazole, 100 mM NaCl, pH 7.5). Finally, the peptides in the elute buffer were desalted via C18 ZipTip (EMD Millipore) according to the manufacturer's instructions, and was eluted using a saturated solution of sinapinic acid in 50% aq. MeCN. The final product was analyzed by MALDI-TOF MS (notebook reference 5/28/2015).

Table 2.3 Oligonucleotide primers used in this study. All sequences are provided 5′ to 3′. F indicates a forward primer while R indicates the reverse. Lowercase m indicates 2′ *O*-methylation of the following residue; methylation suppresses random addition of bases at the end of the RNA by T7 RNA polymerase.

Primer Name	Oligonucleotide Sequence
mbp F (sequencing)	GAGGAAGAGTTGGCGAAAGATCCACGTA
T7 F (sequencing)	TAATACGACTCACTATAGGG
T7 R (sequencing)	GCTAGTTATTGCTCAGCGG
tbtA Gibson F	GCATCACCATCATCACCACAGCCAGATGGACCTGAATGATCTGCC
tbtA Gibson R	GACTTAAGCATTATGCGGCCGCATTAGGCGCTGCTGCAG
tbtB Gibson F	CTGGTGGACAGCAAATGGGTCGCGGAATGCGTCTGGTGGAAC
tbtB Gibson R	GTGGTGGTGCTCGAGTGCGGCCGCATTATTCACCCAGTTCAACCAG
tbtC Gibson F	CTGGTGGACAGCAAATGGGTCGCGGAATGACCCCGCACGAAG
tbtC Gibson R	GTGGTGGTGCTCGAGTGCGGCCGCATTAAACGGTCACTTCCGCTTC
E. coli gluRS Gibson F	CTGGTGGACAGCAAATGGGTCGCGGAATGAAAATCAAAACTCGCTTC

Table 2.3 (cont.)

Primer Name	Oligonucleotide Sequence
E. coli gluRS Gibson R	GTGGTGGTGCTCGAGTGCGGCCGCTTACTGCTGATTTTCGCG
E. coli tRNA F	AATTCCTGCAGTAATACGACTCACTATAGTCCCCTTCGTCTAGAGGCCCAGGACACC
E. coli tRNA R	mUmGGCGTCCCCTAGGGGATTCGAACCCCTGTTACCGCCGTGAAAGGGCGGTGTCCTGG
T. bispora tRNA(CUC) F	AATTCCTGCAGTAATACGACTCACTATAGGTCCCGTCGTCTAGAGGCCTAGGACGC
T. bispora tRNA(CUC) R	mUmGGTAGTCCCGAGCGGATTCGAACCGCCGTTACCGCCTTGAGAGGGCGGCGTCCTAGG
T. bispora tRNA(UUC) F	AATTCCTGCAGTAATACGACTCACTATAGCCCCCATCGTCTAGCGGTCTAGGACACCGCC
T. bispora tRNA(UUC) R	mUmGGTGCCCCCAACGGGATTCGAACCCGTGCCGCCGCCTTGAAAGGGCGGTGTCC
tbtG Internal	GTGATACCCAAGGCGGTCT
tbtF Internal	TGCACATCACTTTATTCCGCA
tbtB Internal 702	CTTTACGAGCATGGGTGTGG
tbtB Internal 1496	TCCATCGTCACGTTCAGGAA
tbtB Internal 2116	GTTATTCCGCTGCATCTGG
pbtD (BamHI) F	AAGGATCCGTGACCTGGCGACGCTTTGAC
pbtD (XhoI) R	AACTCGAGTCACATGCCTCCTACGGCGATC
MBP-TbtF F	GAGCGGATAACAATTTCACACAGGAGACAGACCATGAAAATCGAAGAAGGTAAAC
MBP-TbtF R	GATTTTCATGGTATATCTCCTTGAATCTTACACATGTGCAGATTC
RBS MBP-TbtG F	GATAACAATTTCACACAGGAAACAGACCATGGATTCAAGGAGATATACCATGAAAATCGAAGAAGGTAAAC
RBS MBP-TbtG R	GGAATGGTCATGGATTATCTCCTTATTAATTACGGAAACGGGTG
RBS TbtE F	GTTTCCGTAATTAATAAGGAGATAATCCATGACCATTCCGCCGG
RBS TbtE R	CAGCCAAGCTTGCATGCCTGCAGGTCGACTTACGATTCTTCAACCGTG
T7_TbtA_F	GCTCCACGGTTGAAGAATCGTAAGGGATCTCGACGCTCTCC
T7_TbtA _R	CTTCTCTCATCCGCCAAAACAGCCATCGACTTAAGCATTATGCGGC
TbtB_F	GAGCGGATAACAATTTCACACGAGCTAATAAGGAGATATACCATGCGTCTGGTGGAAC
TbtB_R	CATGCCTGCAGGTCGACTCTAGAGGATCCTTAAACGGTCACTTCCGCTTC
tRNA_F1	TGCGCTGGACGAAGCGGAAGTG
tRNA_EcoRI_R1	GATCATGAATTCAAATTTGCCGTGCAAATTTAAAAAAAATCCTTAGCTTTCGC
EcoRI_tRNA_F2	TACTTAGAATTCGAGGTTTTAACTACATGTTATTAATACGACTCACTATAGTC
tRNA_HindIII_R2	GAATTTAAGCTTCGAACATATTCGATACATTCAGAATTTAAAAAAAATCCTTAGCTTTCG
HindIII_tRNA_F3	TGTTAGAAGCTTAAATTCAAAACCAATTTGTTAATACGACTCACTATAGTCC
tRNA_pTrc33_R3	CTTCTCTCATCCGCC
T.bispora GluRS_F	GACGAAGCGGAAGTGACCGTTTAAGCACAAAGGAGACAGACCATGATTAGGGTGCGTTTCGCCCCCTC
T.bispora GluRS_R	GGAGACTACATAAAGTAGTTCAGGCGGAGAGCTTGGCGAG
TbtD-F	CCATCACCATCATCACCACAGCCAGATGGCAGCAGGTGAAC
TbtD-R	GTTCGACTTAAGCATTATGCGGCCGCATTAACGGGTCGCAACTTC
M13_pUC	AGCGGATAACAATTTCACACAGG

Table 2.3 (cont.)

Primer Name	Oligonucleotide Sequence
TbtF_775_R	GTTCAACCACGTGCAG
TbtG_600_R	CTGAGAACGCAGGAC
pTrc33_Seq_F	TGATCTGGTTTGACAGC
pTrc33_Seq_R	CTGTATCAGGCTGAAAATC
TbtC_828	CATCCAGGATGAAGCAC

Table 2.4 Sequence of *tbt* biosynthetic genes for optimal *E. coli* expression. All sequences are provided 5′ to 3′. Restriction sites for cloning are underlined (5′ BamHI, 3′ XhoI). These gene constructs were synthesized by GenScript (Piscataway, NJ, USA).

Gene Name	Gene Sequence
tbtB	GGATCCATGCGTCTGGTGGAACGTCGCTTTCCGGTTGCCATGACCTCGACGGCTCCGAAAGTCGCAGTGC GTGAATGCGGTCTGCCTGTGAGCGCGATTGAATCTCTGTGCTGTACCGATAGCTTCGCACTGATCCGTCG CCAGGTTCGCGAAACGGCGTGGCTGAAAGGCGAAGGTAAACGTCTGGCCGTGGATCTGGGCCTGCTGATT GGTGAACGTGGTGACGGTGATGATGGTCTGCGTCCGGTTCTGGTCGGTCTGCGTCGCGCACTGCATACCG GTCGTCTGCCGGATGCACGCGAATGGACGCCGCGTGTGGCATCTGCTCTGCCGGCAGAACTGGCAGCCCG CGTCGCTGATTGGGTGACCCGTATGCGTGCACTGACGCGTGCACGTCGCGAACTGCCGGAACTGTTTGCA GCTGAAGCACGCGTTAAAGAAAAAGTCCTGGCACAGGTGGCAGCCGATCCGGGTTTCCGTCGCGCTCTGA GTCTGGCATCCCCGGAACTGGCAGCTGATCTGGACCGTTGGCTGGCAGAACCGGCACGTCGCCCGAAAAC CCAAAAACTGCTGCGCCTGGCCAAATATGTTGCACGTGCGGCCGTCAAAACGTCACCGTACTCGACCTTT ACGAGCATGGGTGTGGCCGTTTGGGAAAACGGCGAAGATTGGGCAGACGGTGCTATTGTTCGTTTCGCAC CGCGCGAACCGCCGTCAGTCATCCTGGAACCGTCGGGTGAATGGCTGCATGGTGCACTGCGTGCATGGCT GGCACGTCCGGAAAACCTGGTGCGCAGCCGTCTGCGCCTGAATCCGTCTCTGGTTATTCGCGCTGATAAA GCGGAATTTCTGGGCTTCCCGCCGCGCGAACCGATTATCCGTATGGGTCTGACCCCGGTGGTTGCAACGG TTCTGCGTCTGGCAGAACCGGCAGCTGATGCTGACGGCTGGATCGATCCGATGGGTTTTCGTGATCGTCT GGCTCGTGACCTGCCGGCAGAACCGGAACAGGTCGACCGTCTGCTGCGCTCCCTGATTGAAGCCGGCGTG CTGGAAGCACACCCGCTGACCCGTGCGGGTCTGCCGGAAACGGGTGAATGGGCAGAAATCCGTGCAGCAC TGCGTCATGATCCGCACGGCGAAGACCCGGAAGCGTATCGTGTGCGCCTGGCTCGTCTGAAACGCGCGAT GACCATGATGTGGCCGCAAGGTGATACCACGGCGCTGCTGCATGAAACCGCAGTCGTGACGCGTCCGGTT GCAAGCCTGAATCCGACCGCTTGGGGCCGTGGTCTGAGTGATCTGGACGTTGTCCGTCGCTGGCTGTCCG TGTTTGATGGCAAACTGCCGATTCGCATCGTGGTTGCCGAATATCTGCGTGCACGTTACGGTGAACACGC CCGCGTGCCGTTTCTGACCTTCCATCGTCACGTTCAGGAAGAAATTGCAGGTGATGCACCGTCAGGTGCA GACCTGCGCACGTTCGTTGGTCGTTCGGCAGCTATCTGGGCTCCGCCGCTGGCACATAGCCGCCTGCCGC GTCTGCGTGAACTGGCAAAACTGCGTGAAGCAGCCCGCGAACTGGCACTGGGTCGTCCGGAACACGATGG TATTCAGCGTGTGGACCCGGAAGAACTGATCAAACAAATGGCGACCTGGCCGGAATGGATTGTCGTGCCG CGTTCTTGCGCGTGTTATGTCCAGCCGGCACCGGAAGGCCGCCTGGTGCTGAACGTTGTCCATGGCGGTC ACGGTCGTGGTCTGCGTCGCCTGAGTCATCTGATCGGCCGTGTGCGCGGTGAAGCGGTTGATCACCCGAT GGTCGCCGACGAACCGGAAGGTACCGTGTACGCGGAACTGAGCGGCTCTCTGGGTTCAACCCTGAATGTC CATGTGCCGGGCACGCGTTATGAAATTGATTACCCGTTTAGTCCGGGTGATCGTTCCCGCGACCGTCGCC TGCCGCTGTCTGACCTGGAAGTGGTTCTGGCACCGGAAACCGGTCTGGCTGAACTGCGCAGTCGTCGCCT GGGCTTTCGTGTTATTCCGCTGCATCTGGGTATGGCAGCTGAATTCCAGCTGCCGCCGGCAGCACGTTTT CTGGAACGTGCCTTCGGCGTGACCTATCTGCTGCACCCGAGTGCACCGCCGCTGCTGCGTATCGGCGAAG TGCCGCCGCCGCAGGAAGTGACGCGTTACCCGCGCGTTGAAGTCGGTCGTGTCGTGGTTCAACGTCGCCG TTGGCTGGCACCGGCAGGTACCCTGCCGATTCGCGCGAAAGGTGAAGATGACGCCTCCTATCTGCTGCGT CTGGTTGCCTGGACCGATGCAAACGGCATCCCGACGCGTTCATTTGTTCGCGCCTGGCAGGAACGCATGG TCCAGGCTGGTCAAGATAAAGCGCGTAAACCGCTGTTCCTGGACCTGGCGAATCCGTTTCTGGTGAAAGT TTTCGAACGTCAAATTCGCGATTGTGCATTTGTGCTGTTCGAAGAAGCTCTGCCGGACCCGGCAGACGCT CCGCCGCGCGAAGGCAGCGATCTGCCGCGTGTGATCGAATTTCTGGTTGAACTGGGTGAATAACTCGAG

Table 2.4 (cont.)

Gene Name	Gene Sequence
tbtA	GGATCCATGGACCTGAATGATCTGCCGATGGATGTTTTTGAACTGGCAGATAGCGGTGTTGCAGTTGAAA GCCTGACCGCAGGTCATGGTATGACCGAAGTTGGTGCAAGCTGTAATTGCTTTTGTTATATTTGTTGTAG CTGCAGCAGCGCCTAACTCGAG
tbtC	GGATCCATGACCCCGCACGAAGCCGAAGTTATTTGGGTTTCAGCACACATTTTCCATCAAGGCCCGCTGG ACCGTCTGGTTACCGAAGTCGCAGGCCCGCTGTTTGTCGAACTGGCAGCAGCAGGTCTGTCCACCCAGGG CTTTTTCCTGCGCTATTGGGAAGGCGGTCCGCATCTGCGTTTCCGTGTTCGTCTGCTGCCGGGTGCAGAT GCTGCAGAAACGCGTCGCCTGATTCGCGAACGTGCTTGCGCGTATCTGCGTCGCTACCCGAGCCAGGATG TTCTGGACCGTGAAGAATATGCTCGTCTGGCGGAAGTCCTGGCAGCACGCGAAGGTCTGACCGGTCACGC GCGTGAACTGTACCCGAACGATAGTGTGCAATTTATCTCCTATCGCCCGGAACATCACCGTTACGGTCGT GGTCCGAGCCTGGAAGCAGTTGAACGCCATTTCGCAGAAAGCTCTCGTATTGCTCTGGAACTGCTGCACG CAGGTCTGACCATCCCGCAGCGTGGTGCAGTGTGGCTGACCATGCTGATGGCAACGTGGCTGCTGGCAGG CTCTCGTGGTGAAGATGCGTTTCGCCCGTGGCCGGACGCTCGTAAAGAACAAGCGTTCGATCGTGAACGT GAACGTCTGACCGGTCTGGCACGTCATGCACGTCAGCTGACGCTGCGCCCGCCGGACCCGGAAGCACAAG GCGTCCTGGCTGCATGGCACGGTTCAGTGGCACGTCTGGCAGATGCACTGGCAGCATCGGGCTTTACCGC CACGCGTACCGCAACGGTTCTGGACCTGTGTGCTCATCTGCTGGCGAATCGTCTGGGTATTCGCATCCAG GATGAAGCACGTCTGCGCTACCTGGCTAGTCGTGCGCTGGACGAAGCGGAAGTGACCGTTTAACTCGAG
tbtD	GGATCCATGGCAGCAGGTGAACGTTGGTGGCGTTTCCGTGTCGATTATCATGCAGGCCCGATGGACGATC TGATTCTGGATGGTGTGCGCCCGGCATTTGCAGCATTCGCAGCTCAGGCACCGATGGCATATTTTCTGCG TCATTGGCGTCGCGGTCCGCACCTGCGCATTTACGTTAGTACCACGCGTGAAGCACTGGAAGCTGTGGTT CGTCCGGCAATCGAACATGTCGTGGGCGGTTATCTGCGTGCACGTCCGAGTCCGGGTATGGCTGATCCGT CCGCGTTCCTGCCGCTGCATGAACGCCTGGCCGAACTGGAAGGCGAAGATGGTCCGCTGATGCCGTGGTC TCCGGACAACACGATTCACGCAGAAGGTGAACGTCCGGAACCGCTGACCGTGCGTGATGTTCTGCTGGCC GATTTTTATGCAGACACCACGCCGAGCGTTTACCACGCGCTGGAACGTGTCCGCAGTGGTGCTTCCCTGC CGACGATCGCCTTCGATCTGGTTGTCGCTACGGCACATGCACTGAGTACCGGCGGTCTGCCGGTCGCACG TACCTCACTGCGTTCGCACGCAGAAGCTTATCTGGCACGTCGCTCCGATGGTGTGCGTCTGCGTGAACTG TGGCGTGACCATTACGCCCGTAACCGCGAAGCATTTACGGAACGTCTGATTGCGGTGGCCAGCTCTGCGG AAAGCGCCGAAAATGGTGCGCATCTGCCGCACGTCCGTGAATGGGTGCGTCGCCTGCGTCCGATCCGTGA ACGTGCACGTGCTCTGCTGGAATCAGGCGAACTGACCCTGGAATATGCATCGCCGGCTGAAGGTGCGCGC GATCTGCCGAGCCTGGCTGAAGTTTCTGCGTTTCATCGCGAACTGGAAAGCCGTCCGGAATGGGCACGTC TGCGTGACTCTCCGGCATTCGGCGCATATCGTCTGGTTATTAATTGCACCTACCTGCACCTGACGCGTCT GGGTCTGACCCCGCATCAACGTTTCCTGGTCTGTCACCTGGCAGCAGATGCAGCTGCAGACGTGTACGGC ATCGCCGCACATGAAGAAGTTGCGACCCGTTAACTCGAG
tbtE	GGATCCATGACCATTCCGCCGGGCCTGACCGAACGCTATGCCCTGCGTGCGGGTGTTCACTCTGCTGTGC TGCCGGACGGTGTTATGCGCCTGTTTGCGTGGCCGCATGCCGAATCAATTGGTGCGCTGTCGGCCGATGA AACCACGCTGCTGAAAGAACTGGCGGAAGGCCCGCGCGAAATCGCAGATCCGGCTCTGCGTCCGTTTGTG GAACGCCTGTTCCGTGGCGGTTGGCTGAAACGCACCCTGAGCCGTGGCGAACACGACCTGTATACGCTGG ACCCGCTGCGTCGCCCGGGTAGTCGCCCGGCACCGCCGGATGACCCGGTCCTGTCCCGTTTTGCAGCAGT GCGTCGCCGTCCGAGCGGTTTCGTTATTGAATCTCCGCTGGCGTGGTGCGACGTGCATGTTCACGATCCG GCACTGCTGCCGGATCTGCTGGAACCGGCAGGCGGTCGTGCTGGTCGTAGCTCTCTGGCACCGCAGATCC GCCGTCAAGCACTGGCTGATCTGGCATGGGCAGGCCTGGTGGTTCCGCGTGGCGCAGAAGACGGTGCACT GCGTACCCGTCAGTGGGCACCGCATGAACTGGATTTTCACCAACGCTCCCGTCTGTATCATCGCGGCTAC CTGGGCGATGGTTTTGGCGGTACCTTCTGGGCACGTGGTACCTTCGACCCGCCGCAGGCTCGTCCGCAAC GTTATCCGGGCGACCCGATTCCGCTGCACCGTCCGGATCTGAACGCACTGCGTGCAGCTGACCCGCCGCT GACCGCTGTTCTGGAAGATCGCCGTAGCGTCCGTGAATATGATGACGATGCACCGATGACGGTGGAACAG CTGGGTGAACTGCTGTACCGCTCTGCGCGCATTCGTGACGTCAAAGTGATCGATGGCGTGGAATATGTTC GTAAACCGTACCCGTCAGGCGGTTCGGTGTATGAACTGGAAATTTACCCGGTCGTGCGTCATGTTGCAGG TCTGGCACCGGGCATGTATCATTACGATGCGTATGAACACGTTCTGCGTCCGGTCCGTCCGGCAGGTCAT CCGGCTGTGCGCCGTATGCTGACCGTTGCAAGCCACGGCTCTGCAGTGGGTATCCGTCCGCAGCTGCTGC TGGTTGTCAGTGCCCGCGTCGGTCGTGTGATGTGGAAATATGAAGGCATGGGTTACGCACTGATTCTGAA ACATGTTGGTGTCCTGTATCAAACCCTGTACTGCGTTGCAACGGCTATGGGCCTGGCACCGTGTGCTATC GGCAGTGGTGATTCCGCGGCCTTTTCTGAAGCGACCGGTCGCGATCCGCTGGAAGAATGTGCCGTCGGCG ACTTCCTGCTGGGTTCACGTCCGGCGAGTGGCTCCACGGTTGAAGAATCGTAACTCGAG

Table 2.4 (cont.)

Gene Name	Gene Sequence
tbtF	GGATCCGTTATTCCGGCTCGTCCGGTTCTGACCCCGGGCGTTCGTTATGCTCCGACCAGTGATGGCGTTG CTTTCCTGACCCGTGATGGTATTGTGAGTATTACCGGCGTTTCCATCCACCGCTGGATTGACCGTATCGC ACCGCATCTGACCGGTGAACGCAGCGTTGCAGAACTGACGGATGGTCTGGCACCGGAACGTCGCGCATTC GTGCTGCGTCTGCTGCAGGCACTGGCAGATCGCGGTCTGATTATCGACGCAGCCGGCGGTCGTGAACCGC GTCTGGAACGTGGTACCGCTTGCGCACTGCACGTTACCGAAACGTGTGGTCCGTATCGTGAATCACTGCC GAAACTGGCACGTTCGCTGGCAGATGCACTGACCCGTGCCGGTCTGGACGTGCGTCTGGCTAGTGCGTCC GAACCGTGCGATCCGGAACGTTGTCATGCGCTGGTTCACCTGACGGCAGCTGACGAAGTCGGTCTGGCAG CAGCAGAACGTCTGGATCGTCTGTCAGAACGTTGGGGTGTTCCGATTGCACATGTGGTTGTCTGGGGCGG TGAAATCTGGCGTACCGAAGCCGGTACGGTGGGTGAAGACGGTCTGCCGCTGGCTGCAGGTTGGCGTCGC CTGACCGCGCTGCGTCCGGATACGGAAGAACGCGTTCCGATTTCGCCGACCGCAGCAGTGGTTGTCGCCG GTCAGGTCGCTGCAGATGTGCGTGACCGTCTGACCACGGCAACCCGTGCTGGTCGTCGCCCGCGTCTGCA CGTGGTTGAACCGGCTCGTCTGACGTGCCGTGCACATCACTTTATTCCGCATTTCCACACCATGCCGCTG ACGGGTGATGCACCGCTGGCAGAAGAACTGACCGAAGAAGAATTTTCCCGTCGCGCCGCAGCTCTGATGG ATGACCGCACCGGTGTGTTTACGGAAATCGCGGAGGGTGATTTCGGCCAACTGCCGCTGCATGTGGCTGT TACCACGGTTGCCGATCCGGTCGGTCTGCTGGGTCCGGCAGGTGACCGTCCGCGCGTGATTGGCGTTGGT CCGGATTTCGCAACCGCACGTTATCGTACGGCGCGTGCGGCAATCGCACTGTACGGTCTGCTGGCAGTTG ATCCGCGTCGCCTGGTCGATGCAGACGGTGAACCGCTGGCAGGTCCGCGTACCAGCGCAGCTGAACTGGA AGATCTGCTGGGTCGTATTCGCGCGGGCGGTCTGCCGGCATTTGTCCGTGCCGAAGATGTGCGTGGCGGT CCGCCGCGTCTGCTGCCGGCCAGCTCTGTGTTTCCGTTCGCTACCGCGCGTGCCCTGCTGGCCGGTACGG TGCCGACCGGCACGGCAGTTGGTTACTCTGAAGAAGGTGCACTGCGTGCAGCACTGCTGGATCACTGCCG TGAACGCACCCTGGCAGACCTGGCTAGTCGTCGCGTCGCTCGTCTGACCCCGGATACGGCACCGTCCGAC CCGGTGACCGCACGTTATCTGGCTTTTCTGCGCGCGATCGGCCTGCCGTTCGATGTTCTGGACGCAACGG GTCCGCTGGGTGTCCCGGTGTACGTTGGCGTCCTGGATGGTCGCGTCGTGGCAGCTGAAGCAGGTGCTAG CCCGGGTGAAGCGTTTCGTTCTGTGCTGGAAGGTATTCTGCGCGGCCTGCAAGGTATCCCGGCAGCACGT AGTCTGCCGGACACCCTGCCGCCGGTCCTGCCGGAAGCAATGGCAGTGCCGGATGGCAGCCCGCTGTCTG CAGAAGCTCTGGCAGCTGCACTGGCGCGTGCCGGTCTGTCAGTGTCGGTTGTCCGCCTGGATCATGACCG TCGCGTTCACGCCCTGATTCCGCATTGTGTGCGTGTTATGGTCACCGGCGAAGAATCTGCACATGTGTAA CTCGAG
tbtG	GGATCCATGTCCGACGCTCTGGCTGCTGCTATTATTGGTGCTGGTCTGCTGGCCGATGAACTGGCTGCTG CCTTTCGCCCGTATGGTGAAGTCATTCGTCTGGATGAAGTCGAACGCGCTGAAACCGAAGCGCATCGTCT GCGCGCCCTGGTTATGGCATGCGATGGCTGGGACACCAGTATTTACCGTGAAACGCGTGAACGCTGCCGT CGCCTGGGTATCCCGTGGCTGCCGGTTCGTACGGAACATGGCATTGCAGTCATCGGTCCGCTGGAACGTC CGGGTGAACCGGGTTGCGCCCACTGTTTTGAACTGCGTCGCGAACGTGCACGTCCGGATGCAGCCGCATA CCGTGCAATCCTGAACCATCACGGTCCGGAACTGGCAAAACGTCCGTCACCGTATCTGGACGAACTGGCT GCGGCCACCATTGCAGCTCTGGGTGTGCGTATGATCGATCAGGGCTCGGGTTGCCATGTCTGGTATGTGC GTCTGAATGGCCTGACGGTTGAACGTCACGCGTTCCTGCCGGAACCGCTGTGCCCGGAATGTGGTAGCCT GCCGATGGATGACCGCGATAGCGCTGTTATTGTCCTGCGTTCTCAGCCGAAACGCGCCCCGGATGACTAC CGTACCCGCGATATCGTGGACGAACTGGATGCCCTGGTTGCAGACTATGTCGATGGCGAATCTGGTCTGA TTCGCCCGCTGGTGCGTGATACCCAAGGCGGTCTGGTTATCGCAGGTGCAATGCTGCCGCTGCGTTTTGC GGGCGGTTCAGAACCGGGCGTTGGTCGTACGCGCGGTTATCGTACCTCGGAAGTCACGGCTATTCTGGAA GCGCTGGAACGTTGGGGCGGTGTCGAACCGGGCGGTAAACGCACCGTGATTCGTGCGGCCTTCCGCGATA TCGCAGCTGACGCACTGGACCCGCGTACGCTGGGCGTGCATAGCCCGGAAGCATACGCGGCCGAAGACTT TCCGTTCCGCCCGTTTACCGAAGATGAAGAAACGGAATGGGTGTGGGGTTATAGTTTCGCACGTCGCTCC CCGATTCTGGTGCCGGAAACCGTTGCGTATTACTATGTTCACCGTGACCGCCCGCGTCAAGAACGTCCGT TTCTGTATGAAATCAGTAACGGCTGCGCTCTGGGCTCCGGTATGGAAGAAGCGATTCTGTACGGTATCCT GGAAGTGGTTGAACGCGATGCGTTCCTGATGACCTGGTATGGCCGTCTGGGTGTTCCGCGCATTGATCTG GACAGCGCAAAAAATCGTATGGTCCCGCTGCAGGCAGCTGCGATTTCTGCGGAAACCGGCTACCGCATCG AAGTCTATGATACCACGATGGAACATGGTATTCCGAGCGTGTGGGCTCTGGCAATCCGTGGCGGTGAAGA ACCGACGGCAGCACCGCGTCCGCGTATGGTGTGTGCTGCAGCAGCACACCTGGACCCGGAACAAGCAGTT CTGAGCGCACTGTCTGAACTGGGTCCGCTGCTGGCAGACCTGCTGCGTCGCTACCCGGATGAAGCTGAAC GTGCGCGTCGCATGGTCGATGACCCGGGTCTGGTGACCACGATGCATGACCACAGTACCCTGTATGGCGC CGATGAAGCATTTGGTCGCCTGGAATTCCTGACCGGTACGGCATCCACCCGTGACCTGGCTGCAATGCGC GAATCAACGGTGGCTTTTCGTTCGGCGGATCTGCGCGATGACCTGCTGGAAGTGATTGGCCGTATCCTGG CGGAGGGTATGGACGTTATTGTCGTGGATCAGACCACGCCGGAACATCGCGTGCACGGTTTTACCTGTGT GAAAGTTCTGATTCCGGGCATGATCCCGATGACCTTCGGTCATCGTAACCGTCGCCTGGATGGTCTGACG CGTCCGCTGGTTATCCCGTACCGTCTGGGTTATCGTCCGGCACCGCTGACCCCGGAAGCACTGAATCCGC ATCCGCACCCGTTTCCGTAACTCGAG

2.4 References

1. Kelly, W. L., Pan, L., & Li, C. (2009). Thiostrepton biosynthesis: prototype for a new family of bacteriocins. *J. Am. Chem. Soc., 131*, 4327-4334.

2. Wieland Brown, L. C., Acker, M. G., Clardy, J., Walsh, C. T., & Fischbach, M. A. (2009). Thirteen posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin. *Proc. Natl. Acad. Sci. U.S.A., 106*, 2549-2553.

3. Morris, R. P., Leeds, J. A., Naegeli, H. U., Oberer, L., Memmert, K., Weber, E., Lamarche, M. J., Parker, C. N., Burrer, N., Esterow, S., Hein, A. E., Schmitt, E. K., & Krastel, P. (2009). Ribosomally synthesized thiopeptide antibiotics targeting elongation factor Tu. *J. Am. Chem. Soc.*, 5946-5955.

4. Liao, R., Duan, L., Lei, C., Pan, H., Ding, Y., Zhang, Q., Chen, D., Shen, B., Yu, Y., & Liu, W. (2009). Thiopeptide biosynthesis featuring ribosomally synthesized precursor peptides and conserved posttranslational modifications. *Chem. Biol., 16*, 141-147.

5. Bagley, M. C., Dale, J. W., Merritt, E. A., & Xiong, X. (2005). Thiopeptide antibiotics. *Chem. Rev., 105*, 685-714.

6. Walsh, C. T., Malcolmson, S. J., & Young, T. S. (2012). Three ring posttranslational circuses: insertion of oxazoles, thiazoles, and pyridines into protein-derived frameworks. *ACS Chem. Biol., 7*, 429-442.

7. Li, C., & Kelly, W. L. (2010). Recent advances in thiopeptide antibiotic biosynthesis. *Nat. Prod. Rep., 27*, 153-164.

8. Dunbar, K. L., Melby, J. O., & Mitchell, D. A. (2012). YcaO domains use ATP to activate amide backbones during peptide cyclodehydrations. *Nat. Chem. Biol., 8*, 569-575.

9. Dunbar, K. L., & 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.

10. Dunbar, K. L., Chekan, J. R., Cox, C. L., Burkhart, B. J., Nair, S. K., & Mitchell, D. A. (2014). Discovery of a new ATP-binding motif involved in peptidic azoline biosynthesis. *Nat. Chem. Biol., 10*, 823-829.

11. Dunbar, K. L., Tietz, J. I., Cox, C. L., Burkhart, B. J., & Mitchell, D. A. (2015). Identification of an auxiliary leader peptide-binding protein required for azoline formation in ribosomal natural products. *J. Am. Chem. Soc., 137*, 7672-7677.

12. Garg, N., Salazar-Ocampo, L. M., & van der Donk, W. A. (2013). In vitro activity of the nisin dehydratase NisB. *Proc. Natl. Acad. Sci. U. S. A., 110*, 7258-7263.

13. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., van der Donk, W. A., & Nair, S. K. (2015). Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature, 517*, 509-512.

14. Tocchetti, A., Maffioli, S., Iorio, M., Alt, S., Mazzei, E., Brunati, C., Sosio, M., & Donadio, S. (2013). Capturing linear intermediates and C-terminal variants during maturation of the thiopeptide GE2270. *Chem. Biol., 20*, 1067-1077.

15. Malcolmson, S. J., Young, T. S., Ruby, J. G., Skewes-Cox, P., & Walsh, C. T. (2013). The posttranslational modification cascade to the thiopeptide berninamycin generates linear forms and altered macrocyclic scaffolds. *Proc. Natl. Acad. Sci. U.S.A., 110*, 8483-8488.

16. Hayashi, S., Ozaki, T., Asamizu, S., Ikeda, H., Omura, S., Oku, N., Igarashi, Y., Tomoda, H., & Onaka, H. (2014). Genome mining reveals a minimum gene set for the biosynthesis of 32-membered macrocyclic thiopeptides lactazoles. *Chem. Biol., 21*, 679-688.

17. Li, C., Zhang, F., & Kelly, W. L. (2011). Heterologous production of thiostrepton A and biosynthetic engineering of thiostrepton analogs. *Mol. BioSyst., 7*, 82-90.

18. Wang, L., Brock, A., Herberich, B., & Schultz, P. G. (2001). Expanding the genetic code of Escherichia coli. *Science, 292*, 498-500.

19. Liu, C. C., & Schultz, P. G. (2010). Adding new chemistries to the genetic code. *Annu. Rev. Biochem., 79*, 413-444.

20. Shi, Y., Yang, X., Garg, N., & van der Donk, W. A. (2011). Production of Lantipeptides in Escherichia coli. *J. Am. Chem. Soc., 133*, 2338-2341.

21. Kuthning, A., Mösker, E., & Süssmuth, R. D. (2015). Engineering the heterologous expression of lanthipeptides in Escherichia coli by multigene assembly. *Appl. Microbiol. Biotechnol., 99*, 6351-6361.

22. Ökesli, A., Cooper, L. E., Fogle, E. J., & van der Donk, W. A. (2011). Nine post-translational modifications during the biosynthesis of cinnamycin. *J. Am. Chem. Soc., 133*, 13753-13760.

23. Tang, W., & van der Donk, W. A. (2012). Structural characterization of four prochlorosins: a novel class of lantipeptides produced by planktonic marine cyanobacteria. *Biochemistry, 51*, 4271-4279.

24. Wever, W. J., Bogart, J. W., Baccile, J. A., Chan, A. N., Schroeder, F. C., & Bowers, A. A. (2015). Chemoenzymatic synthesis of thiazolyl peptide natural products featuring an enzyme-catalyzed formal [4 + 2] cycloaddition. *J. Am. Chem. Soc., 137*, 3494-3497.

25. Burkhart, B. J., Hudson, G. A., Dunbar, K. L., & Mitchell, D. A. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol., 11*, 564-570.

26. LaMarche, M. J., Leeds, J. A., Dzink-Fox, J., Gangl, E., Krastel, P., Neckermann, G., Palestrant, D., Patane, M. A., Rann, E. M., Tiamfook, S., & Yu, D. (2012). Antibiotic optimization and chemical structure stabilization of thiomuracin A. *J. Med. Chem., 55*, 6934-6941.

27. Ortega, M. A., Hao, Y., Walker, M. C., Donadio, S., Sosio, M., Nair, S. K., & van der Donk, W. A. (2016). Structure and trna specificity of MibB, a lantibiotic dehydratase from actinobacteria involved in NAI-107 biosynthesis. *Cell Chem. Biol., 23*, 370-380.

28. Sherlin, L. D., Bullock, T. L., Nissan, T. A., Perona, J. J., Lariviere, F. J., Uhlenbeck, O. C., & Scaringe, S. A. (2001). Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA, 7*, 1671-1678.

29. Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V., & Stroh, J. G. (1999). Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem., 267*, 169-184.

30. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods, 6*, 343-345.

31. Li, B., Cooper, L. E., & van der Donk, W. A. (2009). In vitro studies of lantibiotic biosynthesis. *Methods Enzymol., 458*, 533-558.

32. Rio, D. C., Ares, M. J., Hannon, G. J., & Nilsen, T. W. (2011). *RNA: A Laboratory Manual*: Cold Spring Harbor Laboratory Press.

33. Walker, S. E., & Fredrick, K. (2008). Preparation and evaluation of acylated tRNAs. *Methods, 44*, 81-86.

34. Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., & van der Donk, W. A. (2015). In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *J. Am. Chem. Soc., 137*, 16012-16015.

CHAPTER 3: CHARACTERIZATION OF ENZYMES INVOLVED IN THE BIOSYNTHESIS OF THE CORE SCAFFOLD OF THIOPEPTIDE THIOMURACIN[c]

3.1 Introduction

As described in the previous chapter, thiopeptides are a structurally complex class of antibiotics that belong to the rapidly expanding group of ribosomally synthesized and post-translationally modified peptides (RiPPs) (1). The in vitro reconstitution of the biosynthesis of thiomuracin GZ (**4**) (2), which only lacks the ancillary tailoring post-translational modifications found in naturally produced thiomuracins (3) has been discussed in Chapter 2. In summary, the inactive, linear precursor peptide TbtA (**1**) is converted into a nanomolar potency antibiotic by cyclodehydration by TbtF/G and dehydrogenation by TbtE of six Cys residues to the corresponding thiazoles (**2**). Subsequently, four alcohol groups are dehydrated in a process that involves glutamylation with Glu-tRNA^{Glu} by TbtB, followed by Glu elimination by TbtC (**3**). Finally, two of the dehydrated residues react in a formal [4+2] cycloaddition catalyzed by TbtD, followed by elimination of water and the leader peptide (LP) to yield **4** (**Figure 2.1A**). After the successful reconstitution of the biosynthetic pathway of thiomuracin, we further characterized the molecular details underlying this well-orchestrated set of enzymatic reactions to provide insights into the substrate specificities, directionality, and timing of catalysis by the six proteins involved in thiomuracin formation.

[c] Parts of the chapter are reprinted with permission from: Zhang, Z., Hudson, G. A., Mahanta, N., Tietz, J. I., van der Donk, W. A., & Mitchell, D. A. (2016). Biosynthetic Timing and Substrate Specificity for the Thiopeptide Thiomuracin. *J. Am. Chem. Soc., 138*, 15511-15514. Copyright © 2016 American Chemical Society. https://pubs.acs.org/doi/abs/10.1021/jacs.6b08987
Parts of the chapter are reprinted with permission from: Cogan, D. P., Hudson, G. A., Zhang, Z., Pogorelov, T. V., van der Donk, W. A., Mitchell, D. A., & Nair, S. K. (2017). Structural insights into enzymatic [4+2] aza-cycloaddition in thiopeptide antibiotic biosynthesis. *Proc. Natl. Acad. Sci. U. S. A., 114*, 12928-12933. http://www.pnas.org/content/114/49/12928
Crystallographic studies of TbtD and PbtD were performed by Dillon P. Cogan and Prof. Satish. K. Nair.

The thiopeptide gene clusters are endowed with a unique enzyme that illustrates nature's amazing ability to perform complex chemical transformations: a long-hypothesized unusual [4+2] cyclase that catalyzes the formation of the central pyridine. Enzymatic [4+2] cycloadditions are relatively rare and therefore of great interest. While other enzymes have been identified as [4+2] cycloaddition catalysts, they frequently serve only to enhance a reaction that can occur spontaneously and/or favor product stereochemistry via conformational restriction of the reactant (4). In contrast, the reaction catalyzed by the thiopeptide pyridine synthase is not spontaneous and synthetically equivalent reactions require either prolonged incubation (~1 week) at high temperatures, or microwave irradiation for non-enzymatic catalysis (5). The thiopeptide-forming pyridine synthases are dedicated [4+2] cycloaddition catalysts whose function is to catalyze the cyclization of two Dha residues within the precursor peptide (6, 7). Notably, pyridine synthases catalyze the unique formation of an *aza*-cyclic product, in contrast to the carbocyclic and dihydropyran products formed in other biosynthetic pathways, such as those for spinosyn A (8, 9), pyrroindomycins (10), *trans*-decalin containing polyketides (11), and leporin B (12). In addition to a detailed characterization of the reaction catalyzed by TbtB/C, this chapter will also discuss a detailed structural and mechanistic analysis of two of such pyridine synthases involved in thiomuracin and GE2270 biosynthesis.

3.2 Results and discussion

3.2.1 Characterization of enzymes involved in thiomuracin thiazole formation

We first focused on the substrate permissiveness of the TbtE/F/G thiazole synthetase. A series of TbtA variants were prepared by Graham Hudson in which each Cys residue was systematically substituted with Ala starting from either the C- or N-terminus. He found that all

Ala-substituted TbtA variants were converted to the pan-thiazole species with incompletely processed peptides typically not observed. When six additional variants of TbtA were prepared that each contained a single Cys at the wild-type locations, TbtE/F/G still processed each variant to the expected monoazole, indicating that the installation of each thiazole is independent of the presence of other thiazoles (6). Graham next investigated the order of thiazole formation. Through high resolution electrospray ionization (HR-ESI) MS/MS analysis of the iodoacetamide trapped intermediates of the TbtE/F/G reaction, he revealed a non-linear processing order: Thz4 -> Thz2 -> Thz12, with the remaining three thiazoles processively introduced under the reaction condition (6).

With the ability to process many variants of the TbtA core peptide, we reasoned that TbtE/F/G relied heavily on the LP to drive substrate binding. Indeed, previous studies on other RiPPs containing thiazole and oxazole heterocycles have shown that the LP is important and that it is bound by a specific RiPP recognition element (RRE) that exhibits homology to PqqD (13-16). For thiomuracin biosynthesis, PqqD-like domains are clearly present in the TbtF protein (**Figure 2.1B**), and previous studies have shown that TbtF strongly binds the TbtA LP (K_d ~70 nM) (13). However, the region(s) of the LP that are specifically recognized were not known. To elucidate which parts of the LP are important for TbtF binding, Graham generated a panel of TbtA variants and performed competition fluorescence polarization (FP) assays with a synthetic TbtA LP derivatized with fluorescein isothiocyanate (FITC) at the N-terminus (FITC-$TbtA_{LP}$) to determine relative IC_{50} values. This experiment revealed a set of residues important for TbtF binding (Leu−29, Asp−26, Phe−24). Interestingly, many thiopeptide LPs are rich in Asp/Glu (17-19); we hypothesize that TbtF will display a binding site to electrostatically complement a polyanionic LP with a hydrophobic pocket to accommodate LP position −29. Graham also found that the N-

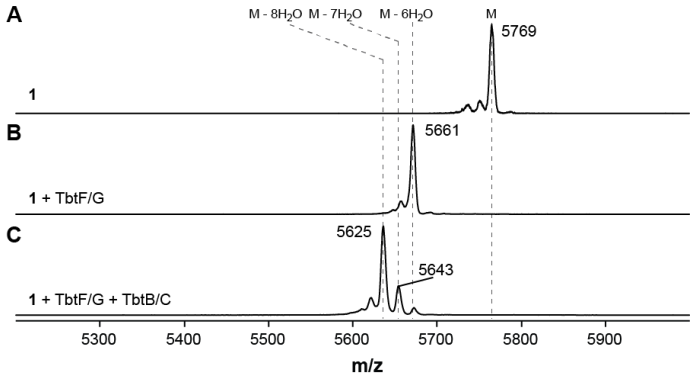
terminal portion of the TbtA LP was bound by TbtF with near wild-type affinity while the C-terminal half showed no binding. Additionally, the RRE of TbtF (residues 1-85) retained most of the binding affinity (K_d ~100 nM), whereas the TbtF protein lacking the RRE (residues 86-628) did not show detectable interaction with FITC-$TbtA_{LP}$ (6).

3.2.2 Characterization of enzymes involved in dehydroalanine formation

I next evaluated the TbtB/C-catalyzed dehydration reaction (**Figure 2.1A**). Dehydration is a two-step process in which TbtB first glutamylates a targeted Ser hydroxyl of TbtA; subsequently, TbtC eliminates the glutamate to yield the alkene. As discussed in Chapter 2, we showed that dehydrations take place at positions 1/11/13/14 of the hexazole **2** (2). To determine if a thiazoline-bearing TbtA would be accepted as a TbtB/C substrate, the dehydrogenase component (TbtE) of the thiazole synthetase was omitted from reactions with TbtA. These reactions cleanly produced the expected hexazoline of TbtA, which upon reaction with TbtB/C, resulted in at most two dehydrations (**Figure 3.1**). These data show that although TbtF/G do not require TbtE for activity, the hexazoline-containing TbtA is incompletely processed by TbtB/C.

I next used the TbtA hexazole **2** as substrate for TbtB in the absence of TbtC, which primarily resulted in one glutamylation event (2). Analysis by ESI MS/MS localized the glutamate to Ser14 (**Figure 3.2**). Similarly, assays that included both TbtB and TbtC that were quenched prior to full tetra-dehydration showed that Ser13/14 were dehydrated significantly faster than Ser1/11 (**Figure 3.3**). Thus, TbtB/C prefer to modify the two isolated Ser residues over those adjacent to a thiazole where the Ser carbonyl group has been removed (**Figure 2.1A**). Alternatively, the dehydration of Ser1/11 sets up the [4+2] cycloaddition reaction, which we previously showed to be very efficient (2). TbtB may have evolved an increased level of substrate

specificity at these positions, such that dehydration at Ser13/14 is a prerequisite for dehydration at Ser1/11.



TbtA: SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*S**C**N**C**F**C**YI**CC**S**C**SSA (**1**)

Figure 3.1 Hexa-thiazoline TbtA is incompletely processed by TbtB/C. MALDI-TOF mass spectra of (**A**) TEV-treated MBP-TbtA precursor peptide **1**, (**B**) TbtA cyclodehydrated by TbtF/G to yield a hexa-thiazoline species (indicated on the primary sequence in brown), and (**C**) the hexa-thiazoline treated with TbtB/C to yield predominately a di-dehydrated species indicating the hexa-thiazoline version of TbtA is not a competent biosynthetic intermediate.

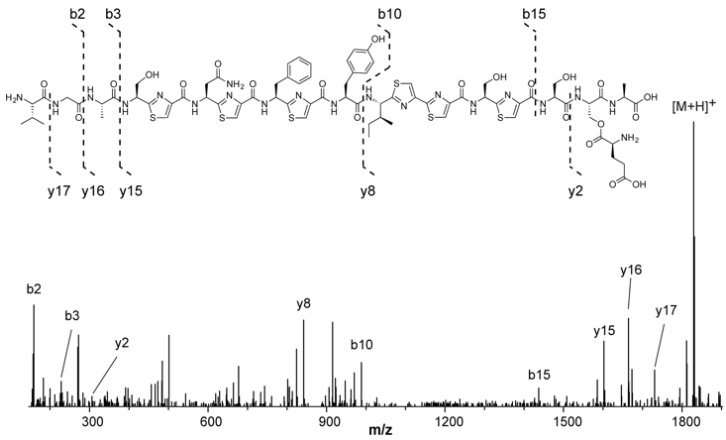


Figure 3.2 MS/MS of mono-glutamylated **2**. The b15 ion localizes glutamylation to either Ser13/14. The low intensity y2 ion supports Ser14 glutamylation.

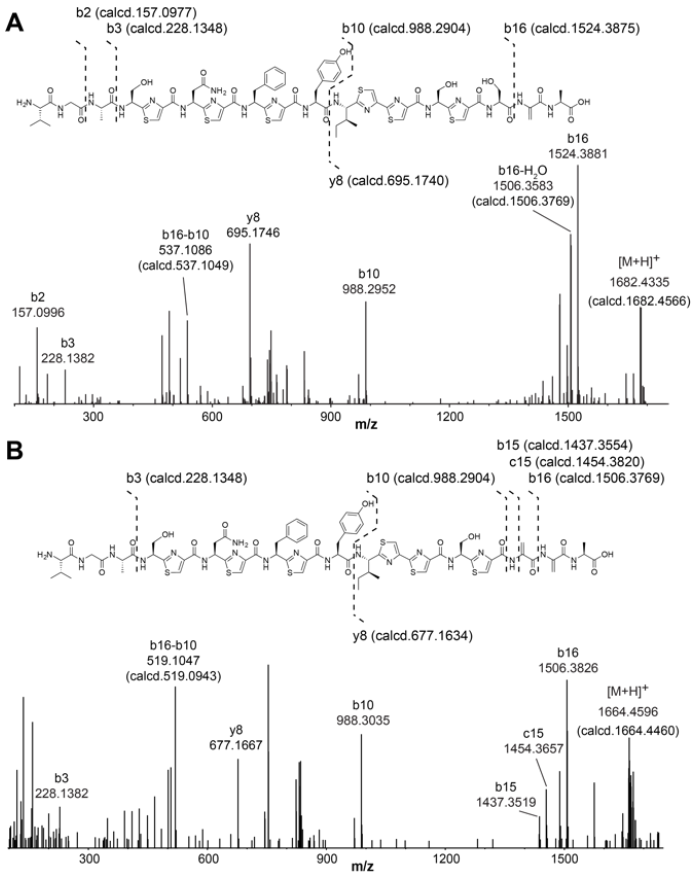
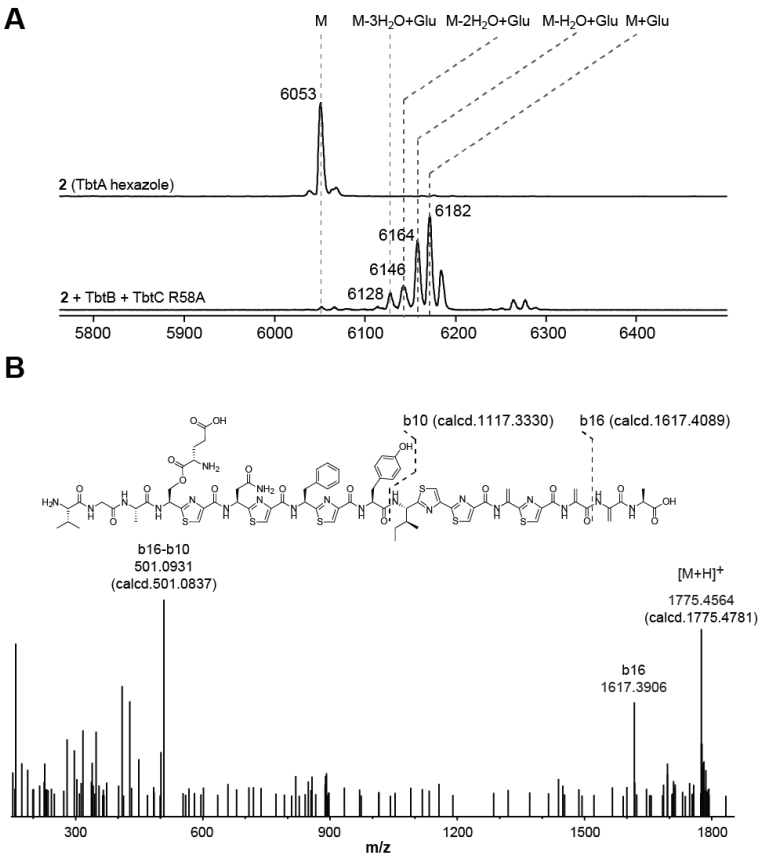


Figure 3.3 ESI-MS/MS analysis of dehydration intermediates of the TbtA hexazole. (**A**) Analysis of the tandem MS data for the mono-dehydrated, TbtA hexa-thiazole reveals that Ser14 is the first position to undergo dehydration by TbtB/C. (**B**) Analogous analysis of the di-dehydrated hexa-thiazole shows that Ser13 is the second residue to be processed by TbtB/C.

I then focused on the glutamate elimination reaction. TbtC is homologous to the C-terminal domain of LanB proteins, which exhibit reduced activity upon mutation of a conserved Arg (20). Therefore, I replaced this residue in TbtC with Ala (R58A) and carried out *in vitro* reactions with **2**. Mass spectral analysis showed a tri-dehydrated and mono-glutamylated intermediate, with the latter localized to position 1 (**Figure 3.4**). Thus, Ser1 is acted on last, and glutamylation (and hence dehydration) progresses in a C-to-N terminal direction.

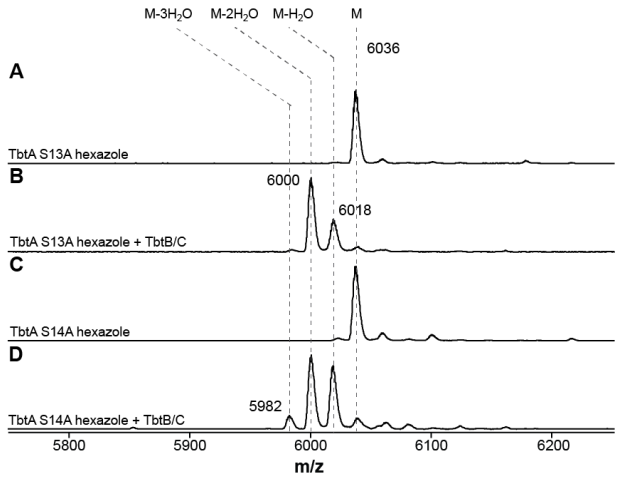


His$_6$-TbtA: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

Figure 3.4 MS analysis of mono-glutamylated and tri-dehydrated intermediate. MS analysis of reactions of **2** with TbtB and TbtC R58A showed an intermediate carrying three dehydrations and one glutamylation, with the latter localized to position Ser1. (**A**) MALDI-TOF MS of **2** treated with TbtB and TbtC-R58A. Replacement of R58 in homologous dehydratases decreases catalytic efficiency (20). (**B**) ESI tandem MS and structure of the tri-dehydrated and mono-glutamylated TbtA hexazole.

These enzymatic preferences do not inform on whether earlier dehydrations are important for subsequent ones. Therefore, a series of TbtA variants were prepared and investigated. Hexazole-bearing TbtA-S13A and -S14A each were dehydrated predominantly twice by TbtB/C, with Ser1/11 largely escaping dehydration (**Figure 3.5-3.6**). This observation suggests that dehydration at Ser1/11 is strongly promoted by prior dehydration at Ser13/14. In contrast, similarly treated hexazole-bearing TbtA-S1A and -S11A peptides each resulted in a tri-dehydrated product,

underscoring that dehydration at Ser1/11 is not important for TbtB/C to process Ser13/S14 (**Figure 3.7**). Taken together, these data are consistent with C-to-N terminal dehydration.



His_6-TbtA S13A: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCASA

His_6-TbtA S14A: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSAA

Figure 3.5 Ser13 and Ser14 are important for dehydration at Ser1/11. (**A**) MALDI-TOF-MS spectrum of the TbtA hexazole with replacement of Ser13 with Ala (indicated with red in the primary sequence). (**B**) Treatment of TbtA hexazole S13A with TbtB/C resulted only in di-dehydration instead of tri-dehydration. (**C**) MALDI-TOF-MS spectrum of TbtA hexazole with replacement of Ser14 with Ala. (**D**) Treatment of TbtA hexazole S14A with TbtB/C also resulted in incomplete dehydration.

For thiomuracin biosynthesis, PqqD-like domains are also clearly present in the TbtB protein (**Figure 2.1B**). Thus, the binding to MBP-tagged TbtB was evaluated using FITC-$TbtA_{LP}$. Despite having a domain homologous to PqqD, TbtB only weakly bound the LP (K_d ~20 μM, **Figure 3.8**). Given the comparatively weak interaction between $TbtA_{LP}$ and TbtB (~100-fold weaker than TbtF), I next investigated whether TbtB/C might act in a LP independent fashion. Peptide **2** was digested with endo-proteinase GluC, providing a hexazole core peptide that retains only three residues of the LP (**Figure 2.1**) (2). Upon treatment with TbtB/C, this peptide was tetra-

dehydrated (**Figure 3.9**). The observed activity on the truncated peptide is in accordance with the abovementioned weak binding to $TbtA_{LP}$ (K_d ~20 μM).

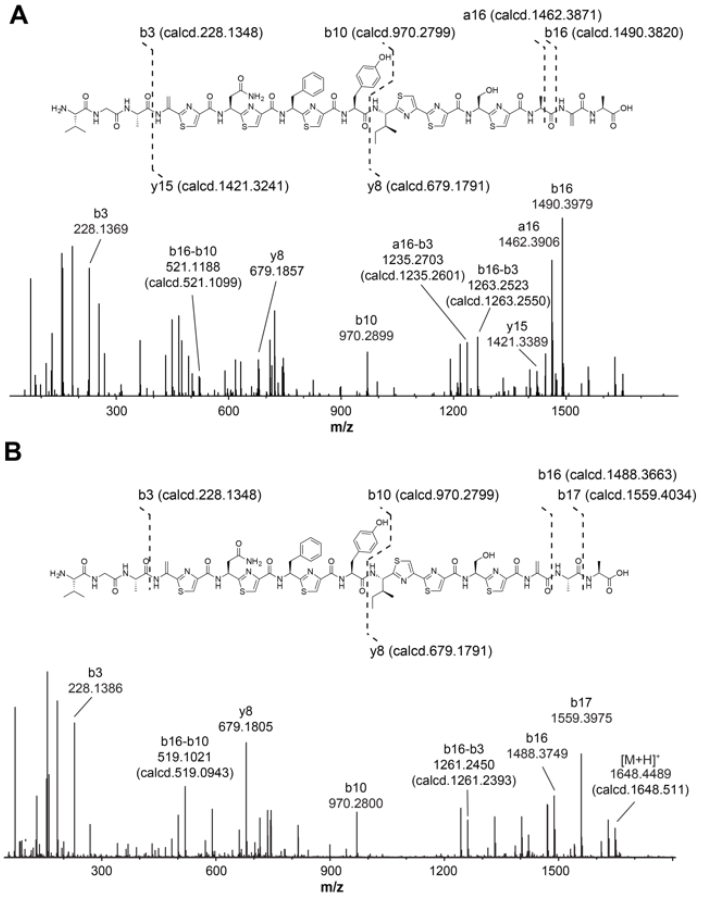
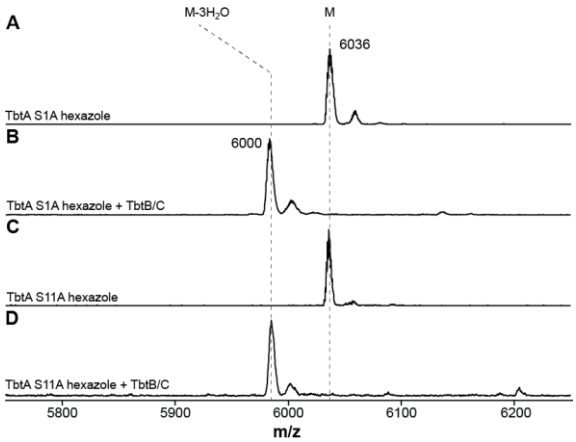


Figure 3.6 ESI-MS/MS analysis of di-dehydrated intermediate of (**A**) TbtA S13A hexazole and (**B**) TbtA S14A hexazole. The fragment ions show that Ser11 escapes dehydration when Ser13 or Ser14 are mutated.

Strikingly, the affinity of TbtB towards FITC-labeled **2** was nearly three orders of magnitude tighter (K_d ~30 nM, **Figure 3.8**). GluC-digested **2** also effectively competed for TbtB binding to FITC-labeled **2**. These data demonstrate that the PqqD-like domain of TbtB is a non-functioning, vestigial RRE, and that TbtB recognizes a specific conformation of the core peptide of TbtA induced by the posttranslational modifications. Ungapped sequence conservation analysis by Graham Hudson confirmed that the defunct RRE of TbtB does not contain conserved motifs,

whereas the RRE of TbtF contains several, including locations known to physically contact the LP (**Figures 3.10**) (13, 14, 20, 21).



His$_6$-TbtA S1A: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*ACNCFCYICCSCSSA

His$_6$-TbtA S11A: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCACSSA

Figure 3.7 Ser1 and Ser11 are dispensable for dehydration at Ser13/14. (**A**) MALDI-TOF-MS spectrum of TbtA hexazole S1A (indicated in red in the primary sequence). (**B**) Upon treatment with TbtB/C, the TbtA hexazole S1A is tri-dehydrated. (**C**) MALDI-TOF-MS spectrum of TbtA hexazole S11A. (**D**) Upon TbtB/C treatment, TbtA hexazole S11A also undergoes tri-dehydration. These results stand in stark contrast to a parallel experiment with mutants of Ser13 and Ser 14 (**Figure 3.6**).

The molecular details of the post-translational modifications that confer TbtB/C activity were next investigated using the aforementioned panel of TbtA variants that systematically replaced Cys with Ala starting from the C- or N-terminus. Each pan-thiazole version of these TbtA variants was tested as a substrate for TbtB/C. Substituting Cys2 and Cys4 with Ala, yielding the corresponding pentazole and tetrazole species, did not greatly affect dehydration at Ser13/14; however, dehydration at Ser11, and especially Ser1, was reduced (**Figures 3.11A**,**B** and **3.12**). Triple replacement of Cys with Ala at positions 2/4/6 (triazole species) abolished dehydration at positions 1/11, but did not affect Ser13/14 (**Figures 3.11C** and **3.13**).

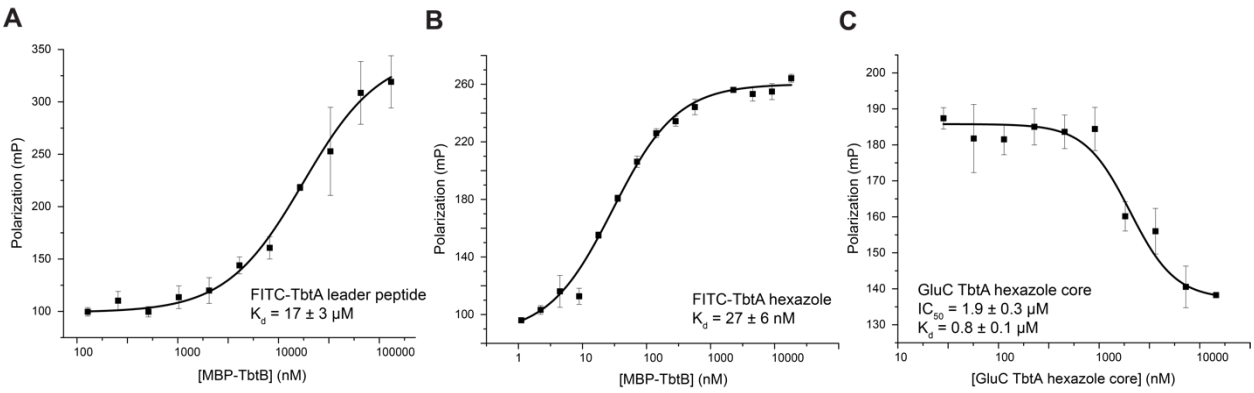


Figure 3.8 Fluorescence polarization analysis of TbtB binding to the TbtA leader peptide, hexazole, and core peptide. TbtB shows considerably greater affinity for the hexazole-containing core peptide of TbtA than the leader peptide. (**A**) FP binding experiments performed using N-terminal FITC-labeled TbtA leader peptide and MBP-TbtB. (**B**) FP binding experiments performed using N-terminal FITC-labeled TbtA hexazole (**2**) and MBP-TbtB (thiazole installation increases affinity by 630 fold). (**C**) FITC-labeled (**2**) displacement from MBP-TbtB by the hexazole-containing core peptide of TbtA (GluC digestion retains 3 residues of leader peptide). Error bars represent s.d. (n=3).

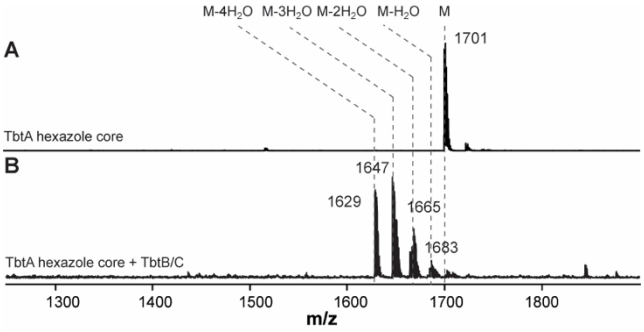


Figure 3.9 MALDI-TOF mass spectra of (**A**) the TbtA core peptide containing six thiazoles and (**B**) treated with TbtB/C.

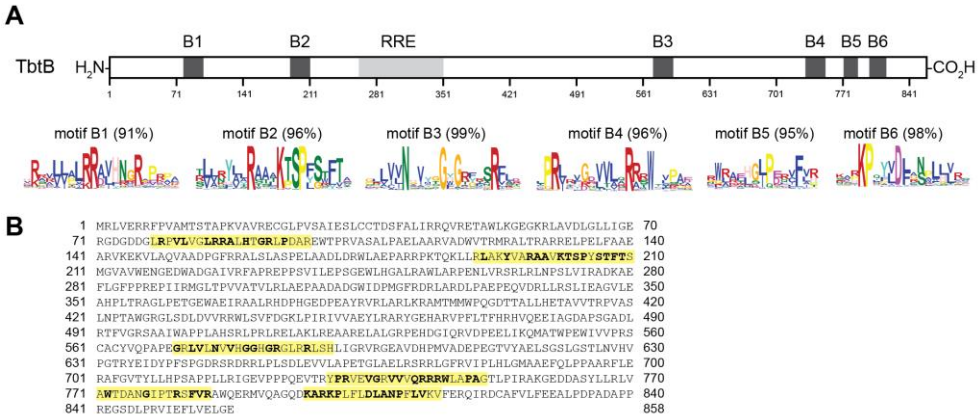


Figure 3.10 Sequence motif analysis of TbtB homologs. (**A**) Diagram of TbtB (part of the split LanB dehydratase) with sites of conserved sequence motifs indicated in dark grey. Frequencies of sequence motifs within TbtB homologs are indicated. The region of PqqD homology (RRE, RiPP recognition element) as described by Burkhart *et al* (13) is indicated in light grey. Note that the RRE of TbtB does not contain any conserved motifs, which is strikingly opposite to that found for TbtF. (**B**) Sequence of TbtB with motif locations highlighted in yellow. Residues in bold indicate amino acids that match the most commonly found residue across the thiopeptide family. Figure made by Graham Hudson.

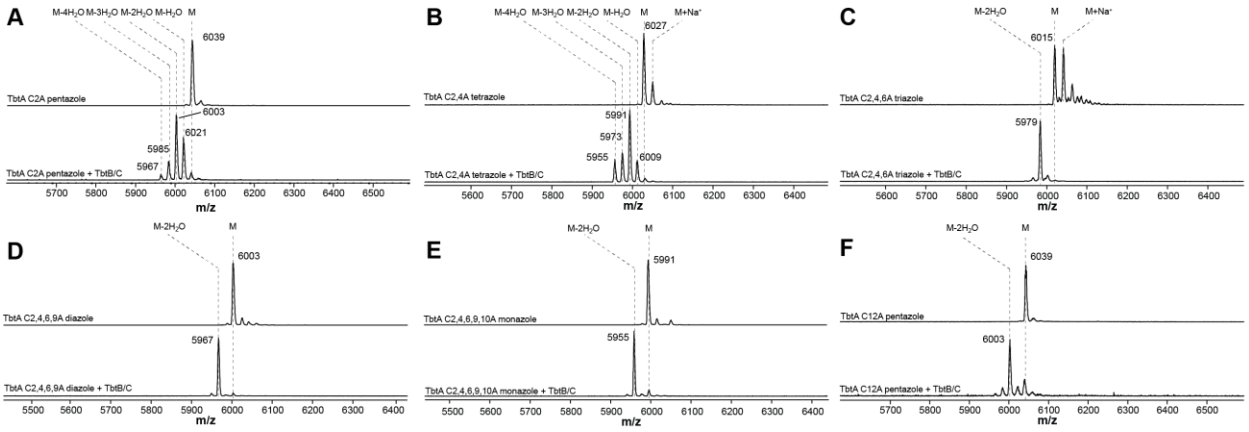


Figure 3.11 MALDI-TOF mass spectra of TbtB/C treatment of (**A**) TbtA C2A pentazole, (**B**) TbtA C2A/C4A tetrazole, (**C**) TbtA C2A/C4A/C6A triazole, (**D**) TbtA C2A/C4A/C6A/C9A diazole, (**E**) TbtA C2A/C4A/C6A/C9A/C10A monoazole, and (**F**) TbtA C12A pentazole.

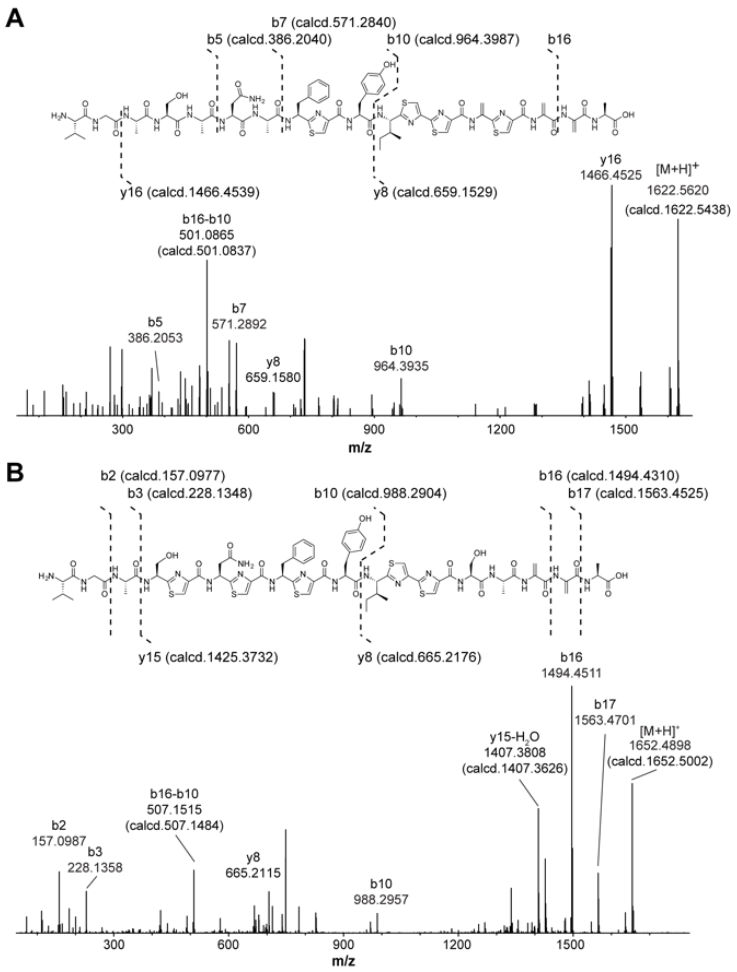


Figure 3.12 ESI-MS/MS analysis of tri-dehydrated TbtA C2A/C4A tetrazole and di-dehydrated C12A pentazole. (**A**) The data show that substituting Cys2 and Cys4 with Ala dramatically reduces dehydration especially at Ser1, but not at Ser13/14. (**B**) Substitution of Cys12 with Ala results only in di-dehydration at Ser13/Ser14.

A similar trend was observed for tetra- and penta-substituted TbtA variants (diazole and monoazole species, respectively, **Figures 3.11D-E**, and **3.13**). This suggests that a thiazole at position 12 might be the gatekeeper modification that turns TbtA into a substrate for TbtB. However, the TbtA-C12A pentazole was still dehydrated twice at Ser13/14 (**Figures 3.11F** and **3.12**). Conversely, replacement of both Cys10/12 with Ala (tetrazole species) abolished processing by TbtB/C and the Cys10 TbtA monoazole was also not a substrate (**Figure 3.14** and **3.15**). Hence,

the thiazole at Cys12 is the gatekeeper modification to elicit TbtB/C activity, although TbtB/C recognize their substrate with a more complicated mechanism.

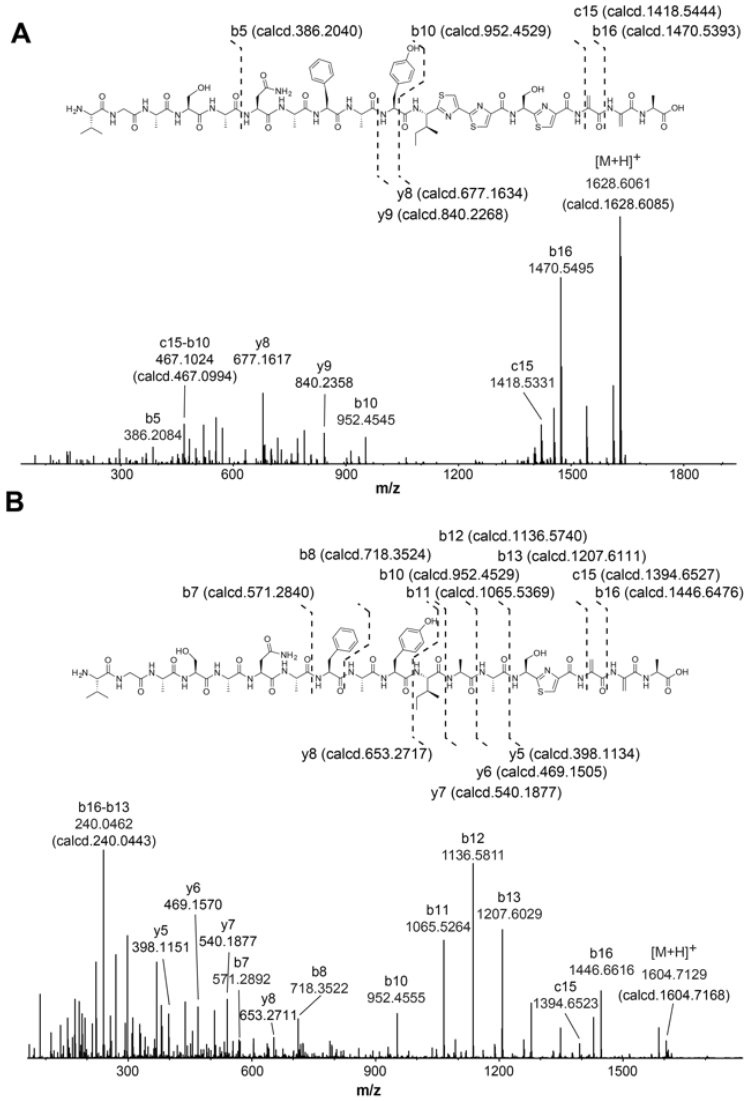
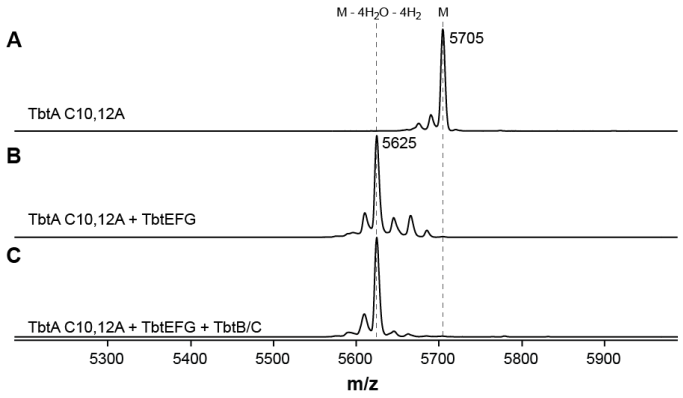
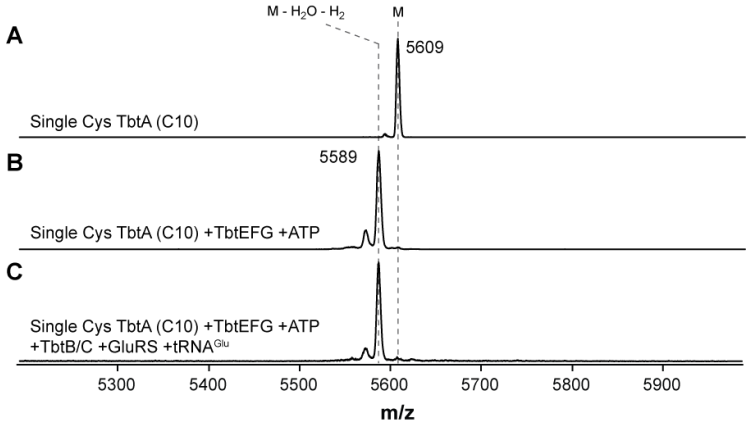


Figure 3.13 ESI-MS/MS analysis of (**A**) di-dehydrated TbtA C2A/C4A/C6A triazole and (**B**) di-dehydrated TbtA C2A/C4A/C6A/C9A/C10A monoazole. Dehydration at Ser1 and Ser11 was abolished, but Ser13/14 were still dehydrated.



TbtA C10A/C12A: SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICASASSA

Figure 3.14 TbtB/C requires thiazole formation at Cys10 and Cys12. (**A**) MALDI-TOF MS spectrum of TbtA with Cys10 and Cys12 both replaced with Ala (indicated in red in the primary sequence). (**B**) Reaction of TbtA C10A/C12A with TbtE/F/G results in the nearly complete production of the expected tetrazole species. (**C**) Reaction of the TbtA C10A/C12A tetrazole with TbtB/C does not result in observable dehydration.



Single Cys TbtA C10: SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SANAFAYIACSASSA

Figure 3.15 TbtB/C cannot dehydrate Cys10 monoazole. (**A**) MALDI-TOF MS spectrum of TbtA bearing only one Cys at residue 10. (**B**) Reaction of single Cys TbtA (Cys10) with TbtE/F/G results in complete conversion to the monoazole species. (**C**) Reaction of the single Cys TbtA (Cys10) monoazole with TbtB/C does not result in observable dehydration.

3.2.3 Biochemical and structural characterization of the [4+2]-*aza*-cycloaddition enzyme in thiopeptide biosynthesis

To characterize the [4+2] cyclase TbtD, we first investigated the importance of conserved residues within the [4+2] cycloaddition clade of the SpaB_C superfamily. A multiple sequence alignment was generated, and twenty-three residues were replaced with alanine by Graham Hudson (**Figure 3.16**). I then tested the activity of these mutant enzymes qualitatively by an endpoint MALDI-TOF-MS assay, and seven of them (W47, E105, H191, H290, S287, R332A and Y319A) were identified to be diminished in activity. The relative activities of these mutants are provided in **Table 3.1**.

```
TbtD              MMAAGERWWRFRVDYHAGPM-DDLILDGVRPAFAAFAAQAPMAYFLRHWRRGPHLRIYVS
gi:242129450      -----MTWTRLRVDYHDGPM-DDLILDALRPAWHEI-----RGYFLRHWVCGPHLRIFVD
gi:552968830      -----MTWRRFDVAYHDPDL-DRLILAA-RPLLS---ESPGRSWFQRHWVRGPHLELWFD
gi:749676247      -----MEWHSLHVVYYANDK-DRLLLEGVRPLLRRLAPEVAATSFTRHWRLGPHVRIHVK
gi:657940957      MSGAAATWHSIHVHHHDESTEPALVLHAVRPAFEAVRASVDGAWFGRHWLRGPHLRLNFR
gi:242129434      -----MSWRRVDVAYHDPDL-DGLILAT-RPLLA---GTPGRGWFQRHWVRGPHLELWFD
                         *  . * ::       *:*   **             * ***  ***:.: .

TbtD              TTREALEAVVRPAIEHVVGGYLRARPSPGMADPSAFLPLHERLAELEGEDGPLMPWSPDN
gi:242129450      GDATPIV--------RAVERHLAACPSAGTTDPQALLPLHERLAELEGERGPLLPWAPDN
gi:552968830      HPEPS-----WERVREVLGTHLRAHPSRTRIDPDRLLPQHRRLALAEQIDEPLLPFYDDN
gi:749676247      CAPAAVPALVRPAAEEIVGGFLARHPSCTHLVPEDHLDEHRRLAELEEDDRPLLPWRPDN
gi:657940957      TTADAWASKVRPQVVAVLETYVRDHPSTVRLDEAALAPVHERLAELEMETGPRHPWVPDN
gi:242129434      AAQPS-----WERIRDVLEPWLRVNPSRARIDRDRLLAQHRHLAAAERIDEPLLPFYADN
                            :   :   **            *.:**  *    *  *:  **

TbtD              TIHAEGERPEPL----TVRDVLLADFYADTTPSVYHALERVRSGASLPTIAFDLVVATAH
gi:242129450      TVTAEP--PGLD----TELDRFLADFYADTTEAAFDALGRVRAGTPLPGIAFDLVVATAH
gi:552968830      TLHRAVPRSRVHVLGSAAAEDLFHDFHAAASTAAFDQLDAVVAGESRLGLAFELMIAAAH
gi:749676247      TIHQAPFERRAHVLGSEEVADLLADFHAATADPAFRMTEAVAGGTPRLGLAFDLMVATAH
gi:657940957      TVLEQPYDHRLPVLGSMRGSEVLADFLSDTNDLAFAMYEHLQGGGALPVLALDLMWTTAS
gi:242129434      TLHRAAPRSRAHVLGGPAAEELFHDFHTTASAVAYDELDAVRAGESRLVMALDLMVAAAH
                  *:                  .: ** : :   .:     : .*     :*::*: ::*

TbtD              A------LSTGGLPVARTSLRSHAEAYLARRSDGVRLRELWRDHYARNREAFTERLIAVA
gi:242129450      D------LSEGGLPTARTSLRSHAEAYLSRLPGGVR--AKWQAHYERNQEPLTARIKALT
gi:552968830      A------HAEGGITGGFVSFRSHAEAFLAGAA---GLRERWEAEYRTRAEALRAQVAAVV
gi:749676247      A------LSGVGITTGFMSFRSHAEAFLTTHPEAHRLRPSWDAHYRGHAASLCDRVGRVT
gi:657940957      VAAVPFDEHSAPIARGFLSLRSHADAFLSRTHDPAAYRAAFDERYRRQADALTTRLRTVE
gi:242129434      A------HAEGGVRGGFVSFRSHAEAFLASAP---GLRERWDAEYAARAGALRARITAVV
                        :  .  *:****:*:*:          :  .*  .   :  ::  :
```

Figure 3.16 (cont.)

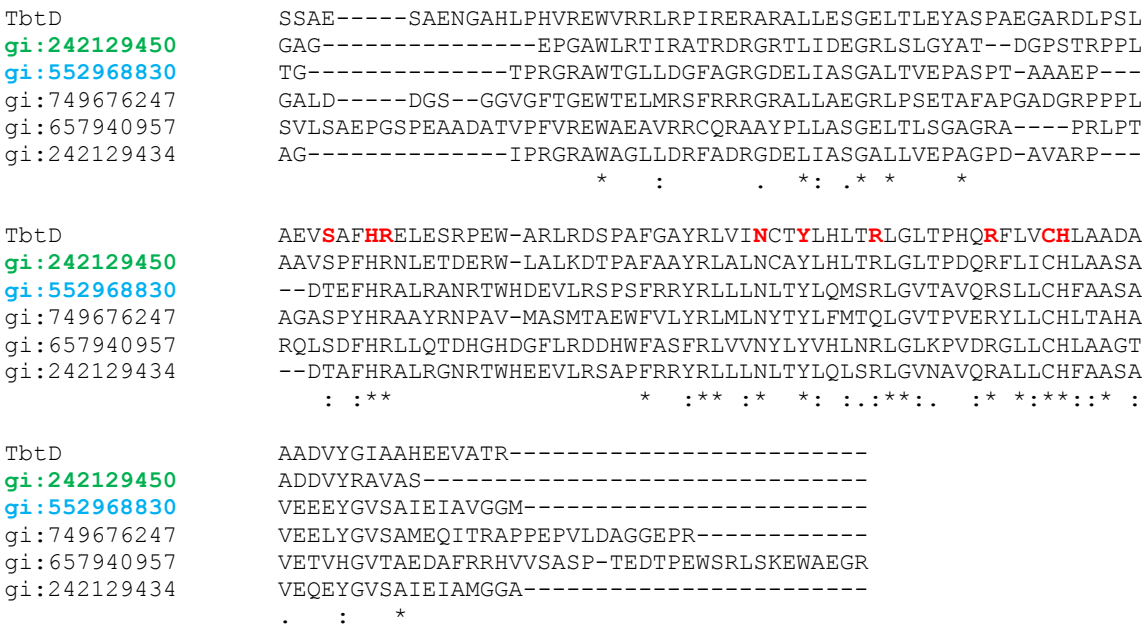


Figure 3.16 Sequence alignment of TbtD homologs. GI numbers are given to the left of the alignment (green indicates the sequence from the canonical thiomuracin producer, *Nonomuraea* sp. Bp3714-39; blue indicates the sequence from the GE2270A producer, *Planobispora rosea*). Shown in red are residues replaced with alanine for enzymatic activity assessment.

Table 3.1 Enzymatic activity of selected TbtD mutant proteins**.** Seven TbtD mutants were selected based on phenotype in the end-point MALDI-TOF MS assay. Conversion efficiency was monitored by comparing EIC intensity to that of wild-type at 90 min (i.e., fully converted). As the wild-type TbtD performed the reaction much faster than the selected mutants, 200 nM enzyme was used for the assay compared to 1 μM for the mutant enzymes. All experiments were performed in duplicate. ND = not detected.

TbtD protein	% conversion 30 min	% conversion 60 min	% conversion 90 min
Wild-type	77 ± 2	87 ± 1	100
W47A	36 ± 3	55 ± 2	64 ± 1
E105A	25 ± 4	47 ± 7	54 ± 11
H191A	36 ± 3	57 ± 1	67 ± 4
S287A	14 ± 4	23 ± 1	39 ± 11
H290A	14 ± 3	33 ± 11	42 ± 4
Y319A	2 ± 1	5.4 ± 3.7	11 ± 1
R332A	ND	ND	ND

Previous studies in our (Chapter 2) and other laboratories (22) have already shown that the [4+2] cycloaddition enzyme is LP-dependent, but a bioinformatically identifiable RRE has not yet

been identified. We thus evaluated binding to MBP-tagged TbtD with FITC-TbtA$_{LP}$, showing that the RRE-less TbtD bound the peptide with a considerably high affinity (K_d ~140 nM, **Figure 3.17**). To elucidate which parts of the LP were important for TbtD binding, Graham Hudson generated a panel of TbtA variants and performed competition FP assays with FITC-TbtA$_{LP}$ to determine relative IC_{50} values. A set of residues important for TbtD binding were revealed (Leu−29 as with TbtF, but also a number of acidic residues: Asp−30/−24/−20 and Glu−14) (6).

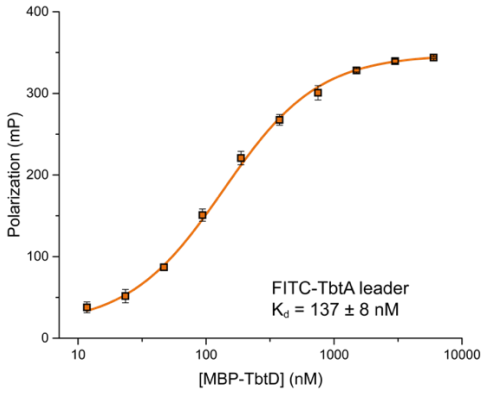
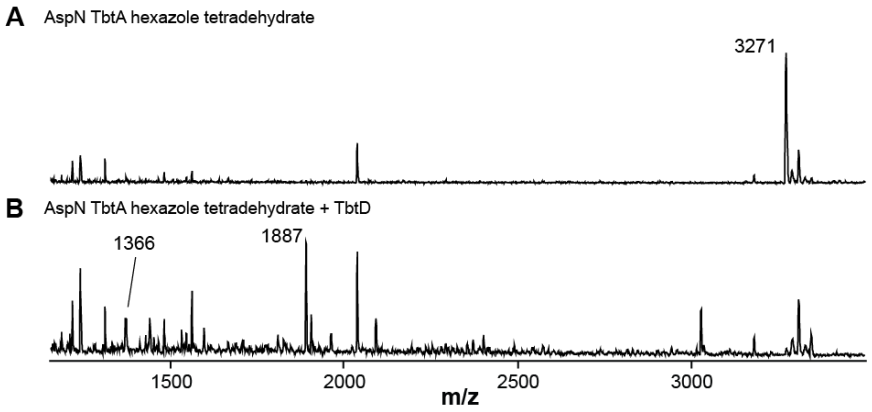


Figure 3.17 FP analysis of binding of TbtD to the TbtA leader peptide. TbtD engages the leader peptide of TbtA with nanomolar affinity as determined by fluorescence polarization. Error bars represent standard deviation of three independent replicates.

Next, I investigated the substrate requirements of TbtD. Since many of the Cys variants precluded dehydration at positions 1/11, which are required for the [4+2] cycloaddition reaction, we focused on further defining the requirements of the leader peptide. As shown in Chapter 2, treatment of TbtA hexazole tetradehydrate **3** with GluC, which removes all but three residues of LP, resulted in loss of TbtD activity (2). I further tested if retaining a larger portion of the LP would result in processing by TbtD. Thus, **3** was treated with endo-proteinase AspN, retaining 20 residues of the LP, and found that TbtD accepted this peptide as a substrate for the [4+2]

cycloaddition reaction (**Figure 3.18**). To further determine which portion of the TbtA LP was required for TbtD activity, we next tested a TbtA variant where Glu(−4) was substituted with Ala. This variant was efficiently converted to the hexazole tetradehydrate species, analogous to **3**, after reaction with TbtE/F/G and TbtB/C. I then treated this peptide with GluC, which owing to the E(−4)A substitution, now retained 13 residues of the LP. This variant was also a substrate for TbtD (**Figure 3.19**), demonstrating that the minimal LP requirement of TbtD is between TbtA residues (−13) and (−3), which is consistent with a recent report showing that ten amino acids from the LP are required for TbtD activity (23).

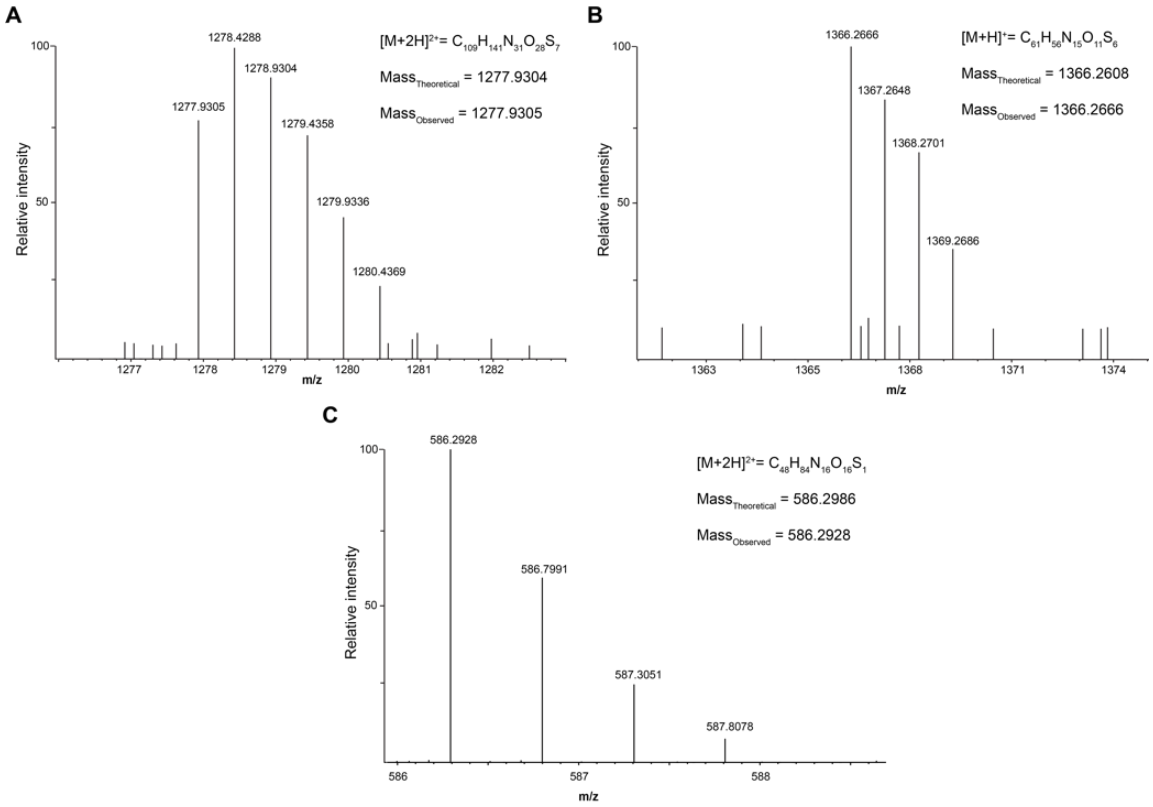


AspN TbtA hexazole tetradehydrate: DSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

Figure 3.18 TbtD tolerates AspN-treated TbtA hexazole tetradehydrate as a substrate. (**A**) MALDI-TOF MS spectrum of endo-proteinase AspN-digested TbtA hexazole tetradehydrate (Calcd $[M+H]^+$ 3271). (**B**) Reaction of AspN-digested TbtA hexazole tetradehydrate with TbtD results in the formation of thiomuracin core (calcd. $[M+H]^+$ 1366) and the ejected leader peptide (calcd. $[M+H]^+$ 1887).

These findings may appear to contradict the binding data of FITC-$TbtA_{LP}$ to MBP-TbtD, but it is important to note that the binding experiments were performed with linear TbtA variants whereas the TbtD activity experiments with truncated TbtA variants were performed with

cyclodehydrated and dehydrated core peptide. We hypothesize that TbtD recognizes both the aforementioned N-terminal residues of the LP and the post-translationally modified core peptide. To test this hypothesis, I performed binding measurements using fluorescently tagged TbtA hexazole, and demonstrated that TbtD binds with nanomolar affinity (**Figure 3.20**). TbtD was also shown to have micromolar affinity with a fluorescently tagged TbtA hexazole treated with endo-proteinase GluC, containing only three leader peptide residues, Val(−3) through Ala(−1) (**Figure 3.21**).



GluC TbtA E(−4)A hexazole tetradehydrate: SLTAGHGMTAVGA*SCNCFCYICCSCSSA

Figure 3.19 TbtD accepts GluC-treated TbtA E(−4)A hexazole tetradehydrate as a substrate. (**A**) ESI-MS spectrum of endo-proteinase GluC-digested TbtA E(−4)A hexazole tetradehydrate. Reaction of GluC-digested TbtA E(−4)A hexazole tetradehydrate with TbtD results in the formation of (**B**) thiomuracin core and (**C**) the ejected leader peptide.

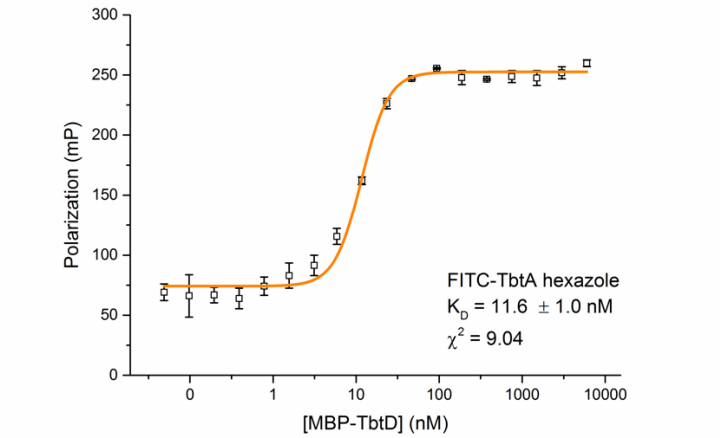


Figure 3.20 FP binding studies of MBP-tagged TbtD towards the FITC-TbtA hexazole. Error bars represent the standard deviation of three replicate measurements.

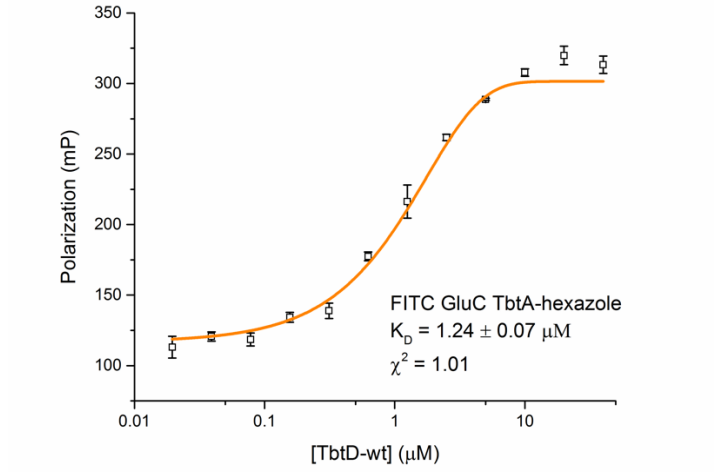


Figure 3.21 FP binding studies wild-type TbtD towards the FITC-labeled GluC-treated TbtA-hexazole. Error bars represent the standard deviation of three replicate measurements.

Our collaborator Prof. Satish Nair and his graduate student Dillon Cogan were able to obtain crystal structures of both TbtD alone, and in complex with the N-terminal region of the leader peptide to define its LP binding site. TbtD is structurally homologous to the core of the

lanthipeptide dehydratase elimination domain (Pfam identifier PF14028; abbreviated as Lant_Dehydr_C) elucidated in NisB (20) and MibB (24), but contains numerous secondary structure insertions (**Figure 3.22**). Notable examples include a helical insertion in the region spanning Pro86 through Asp116 (insertion region 1), a reorganization of the region encompassing Pro183 through Pro241 (insertion region 2), and a large insertion spanning Gly264 through Arg301 (insertion region 3), which is a short connection loop in the Lant_Dehydr_C domains. Importantly, each of the insertions help to configure the binding sites for the substrates specific to the pyridine synthases, and insertion 3 is only visible in structures with bound pyridine core (**Figure 3.23C**). The co-crystal structure with the N-terminal 16 residues of the TbtA leader peptide shows electron density corresponding to residues Leu(−32) through Leu(−22) in the vicinity of insertion region 2 (**Figure 3.22**). Residues Ser200 through Ala234 of TbtD form a large, bent helical insert that abuts the enzyme core and form the binding pocket for the leader peptide. The equivalent region is abridged in both MibB (Gly990 through Glu1035) and NisB (Asp882 through Gly919) and is largely devoid of secondary structure (**Figure 3.22**).

Binding experiments between TbtD and site-directed alanine mutants of TbtA identified several residues within the leader peptide that are important for binding (25). Their identity can now be rationalized in light of the structural data. For example, an Ala mutation at Leu(−29) severely impairs TbtA binding, and in the co-crystal structure, this residue is docked into a hydrophobic pocket that is made from residues from insertion 3. Similar hydrophobic docking interactions are important for leader peptide engagement by enzymes involved in the biosynthesis of lanthipeptides (20) and cyanobactins (26). Additionally, binding by TbtD is compromised by mutations at one of several acidic residues in TbtA, including Asp(−30), Asp(−26), and Glu(−23) (25), and the structural data reveal that TbtD engages some of these residues through electrostatic

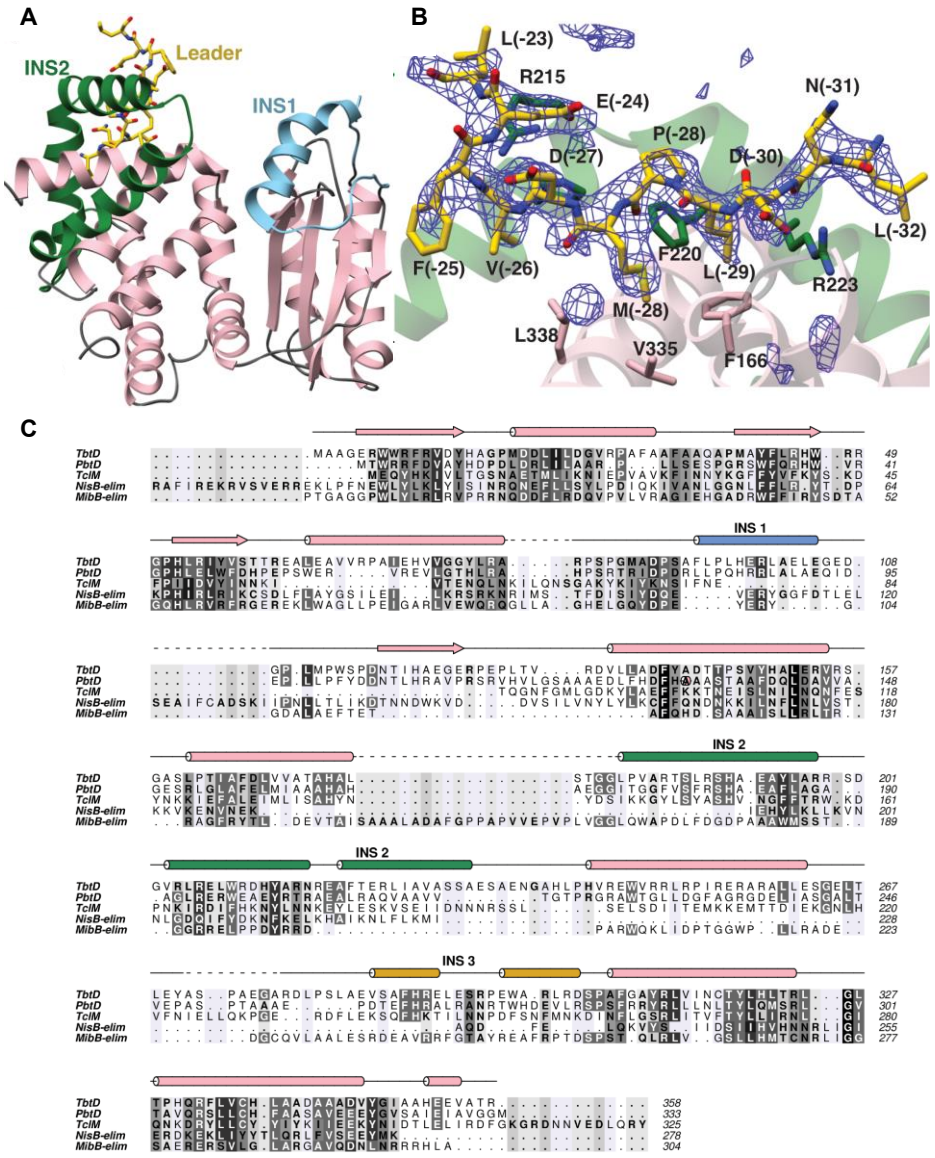


Figure 3.22 (**A**) Overall structure of TbtD in complex with the N-terminal leader fragment, showing insertions 1 (cyan) and 2 (green) that distinguish the [4+2] enzymes from the LanB elimination domains. (**B**) Simulated annealing difference Fourier map ($F_{obs} - F_{calc}$) contoured at 2.5σ showing residues from the leader peptide that bind to a region near insertion 2 (green) in TbtD. (**C**) Secondary structure based multiple sequence alignment (MSA). Included in the MSA are the pyridine synthases TbtD, PbtD, and TclM in addition to the class I LanB dehydratase elimination domains (-elim) NisB-elim, and MibB-elim generated using ALINE software (27). A majority of secondary structural elements are conserved in these functionally divergent enzymes and notable differences are observed in the insertion regions (INS 1, INS 2, and INS 3) found mainly in the pyridine synthases. Figure was made by Dillon Cogan and Prof. Satish Nair.

complementarity. Mutational analysis of the binding site in TbtD was performed by Dillon Cogan to establish the importance of residues in mediating interactions with TbtA, using a shorter peptide corresponding to the N-terminal 12 residues visualized in the structure. Binding of this peptide by the pyridine synthase is compromised upon mutations at enzyme residues that define the hydrophobic pocket for Leu(−29), including Phe166, Phe220, and Val335. Mutation of Arg223, which is involved in additional hydrophobic contacts with Leu(−29) and hydrogen-bonding interactions with main chain amides of the leader peptide to Ala, also diminishes TbtA binding (**Table 3.2**). These structural and mutational data establish the validity of a leader peptide-binding site in TbtD, and demonstrate that the C-terminal region of the leader peptide is dispensable for substrate binding.

Table 3.2 Summary of binding parameters in FP binding studies of WT and site-directed variants of TbtD where a fluorescein labeled, full-length TbtA leader (FITC-TbtA leader) was competed off with an N-terminal leader peptide fragment (N12-mer: LNDLPMDVFELA). Data obtained and table made by Dillon Cogan.

TbtD variant	K_I (μM)	IC_{50} (μM)	$K_{D(labeled)}$ (μM)
WT	2.4 ± 0.1	8.9 ± 0.3	0.26 ± 0.02
F166A	11.7 ± 3.0	42.3 ± 4.5	2.17 ± 0.15
H212A	1.4 ± 0.1	5.9 ± 0.3	0.22 ± 0.01
R215A	5.9 ± 0.3	23.5 ± 0.7	0.35 ± 0.02
F220A	21.5 ± 1.6	96.3 ± 5.3	0.44 ± 0.02
R223A	40.0 ± 5.5	136.9 ± 12.4	2.15 ± 0.06
L334A	4.3 ± 0.2	18.5 ± 0.6	0.25 ± 0.02
V335A	15.8 ± 2.2	71.9 ± 7.6	0.76 ± 0.04
L338A	7.2 ± 09	42.7 ± 1.3	0.32 ± 0.03

Repeated attempts by Dillon Cogan to crystallize a complex of hexazole-containing tetradehydrated TbtA bound to catalytically deficient variants of TbtD proved recalcitrant. The high affinity between the fluorescently tagged TbtA hexazole and TbtD (**Figure 3.20**) ruled out weak binding as the cause of crystallization issues. Although TbtD also binds with micromolar

affinity to the modified TbtA hexazole core peptide (**Figure 3.21**), crystallization attempts using this substrate were not fruitful either. Likewise, attempts at co-crystallization of PbtD with intact GE2270A failed. Rationalizing that flexible regions of the modified peptide substrate may interfere with lattice packing, we focused efforts on studies using analogs that retained the central pyridine moiety (**Figure 3.23A**). Prior studies have shown that acid hydrolysis of GE2270A produced a tri-thiazole-substituted pyridine (TSP) (28). Thus, I treated GE2270A (obtained from Novartis) with 6 M HCl and purified the resultant TSP (**Figure 3.23A**). The binding affinity of PbtD pyridine synthase towards TSP was measured by Dillon Cogan using Förster resonance energy transfer (FRET), yielding a K_D of 2.1 x 10^{-5} M (**Figure 3.23B**).

To discern active site details of the pyridine synthases, the structure of PbtD both alone and in complex with the GE2270A TSP was determined by Dillon. The overall structure of PbtD recapitulates the architecture observed in the homologous TbtD, including the lack of interpretable electron density for residues Glu248 through His273 within insertion 3, despite the much higher resolution of this structure. In contrast, the structure with bound TSP reveals that these residues organize and form two short α-helices that define a periphery of the binding pocket for the ligand (**Figure 3.23C**). The helix from insertion 2 forms a base at the bottom of the binding groove. Ala-scanning mutational analysis of TbtD implicated several residues in catalysis. Indeed, substitutions at His191 (equivalent PbtD numbering is given in parentheses; His182), Ser287 (Thr260), His290 (His263), Phe308 (Phe282), Tyr319 (Tyr293), and Arg332 (Arg306) diminish activity (2). Each of these residues is near the bound TSP, but are located ~5-7 Å from each other and are not all spatially clustered (**Figure 3.23D**). Rather, these residues line alternate sides of a shallow cleft that is located at the juncture of the leader peptide-binding site (on TbtD) and the location of bound TSP. It should be noted that some of these variants may carry out the [4+2] cyclization and only

be deficient in dehydration activity, but are refractory to identification as there would be no net change in mass.

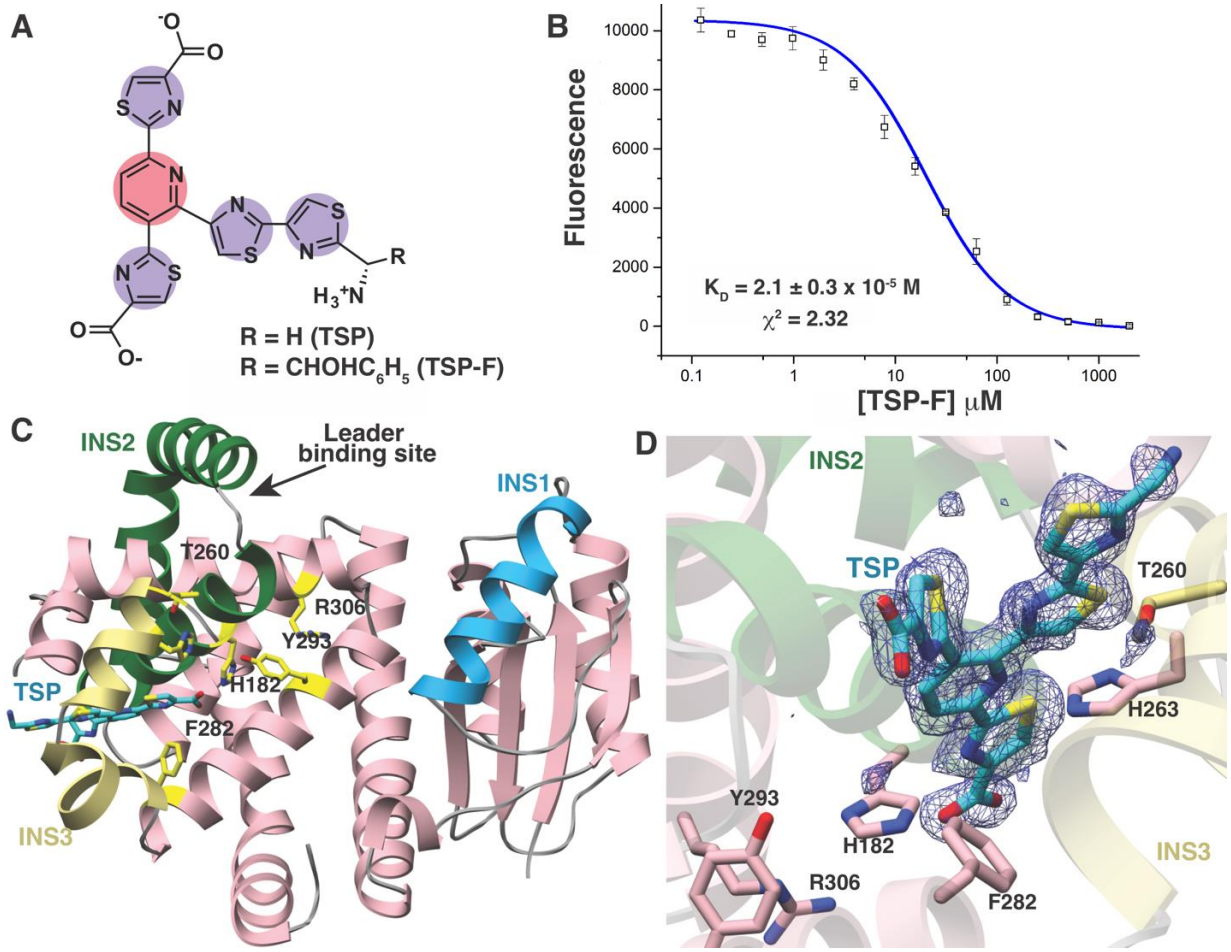


Figure 3.23 Structural and binding studies of the PbtD-TSP complex. (**A**) Chemical structure of the tri-substituted pyridine (TSP) core produced by acid hydrolysis of GE2270A. (**B**) Binding of TSP to PbtD as measured by the changes in FRET upon complex formation. (**C**) Overall structure of PbtD in complex with TSP-F shows that residues in insertion 3 (yellow) become ordered upon ligand binding. (**D**) Simulated annealing difference Fourier map (F_{obs} – F_{calc}) contoured at 2.5σ showing the TSP core binds to a region near insertion 3 (yellow) in PbtD. Active site residues implicated in catalysis, without influencing leader peptide binding, are shown in stick figures. Figure made by Dillon Cogan and Prof. Satish Nair.

A feasible model of pyridine formation involves a looped substrate embedded within the shallow cleft in PbtD, where the aforementioned residues may play roles in catalyzing iminol formation, water elimination, and/or leader peptide ejection (**Figure 3.23D**) as the peptide transverses along the binding cleft. In this model, the anchored N-terminus of the leader peptide would allow a flexibly tethered core peptide to form this looped structure within the cleft, orienting the C-terminus of the substrate away from its N-terminus and at the interface of insertion 1 and the periphery of the cleft (**Figure 3.24**). To test this hypothesis, docking of the [4+2] substrate, as well as various intermediates, was carried out by Prof. Taras V. Pogorelov. The stability of each state over a 1 ns time-course molecular dynamics (MD) simulation was evaluated. Among the lowest energy poses obtained, we observed the maturing substrate and product docking results to be consistent with the predicted substrate-binding site based on the TSP co-crystal structure (i.e. C-terminus and macrocycle orientations). Residues within insertion 3 that are specific to the [4+2] enzymes are largely disordered in the absence of substrate, and the proximity of this region to the active site cleft suggests that these residues may serve a role in active site capping and/or modified peptide core binding. Indeed, ordering of insertion 3 upon binding of TSP is consistent with this proposal. Recruitment of the true substrate may lead to additional ordering of insertion 3 to facilitate substrate threading, whereupon the resultant macrocycle becomes solvent exposed and is directed away from the binding cleft.

Mapping the primary sequences of nine pyridine synthases onto the structure of the PbtD-TSP co-crystal structure reveals a pattern of conservation at the active site (**Figure 3.25A**). As noted, this active site of the [4+2] enzyme is distinct from that of other Lant_Dehydr_C domains, which are built upon a common scaffold observed in the LsrG epimerase (31). LsrG contains a ferredoxin-like domain, a prevalent $\alpha+\beta$ fold observed in numerous metabolic enzymes (32)

(**Figure 3.25B**). For example, the architecture of allantoate amidohydrolase, an enzyme involved in purine degradation in both plants and bacteria (33), is formed through the fusion of the LsrG scaffold with a second domain containing a metal-binding active site (34) (**Figure 3.25C**). Likewise, the elimination active site of NisB is formed by the insertion of an additional subdomain onto the LsrG ferredoxin-like fold (**Figure 3.25D**). The data collected by Dillon Cogan demonstrate that the thiopeptide [4+2] cycloaddition enzymes amend the Lant_Dehydr_C fold with functionally relevant insertion regions to achieve gain-of-function.

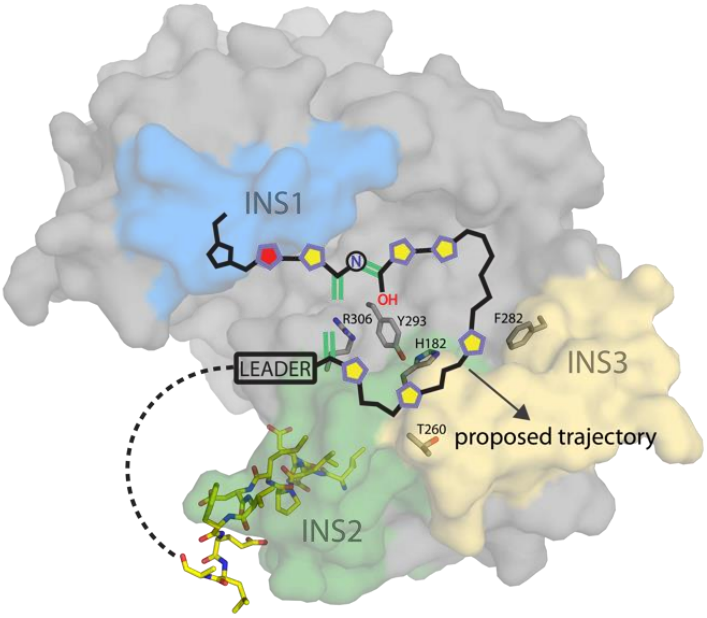


Figure 3.24 Model of pyridine formation catalyzed by PbtD. The PbtA [4+2] substrate is superimposed onto the structure of PbtD, with active site residues drawn as grey sticks. Five-membered heterocycles along the substrate backbone are represented as pentagons (purple-outline/yellow-center = thiazole; purple-outline/red-center = oxazoline; black-outline/blue-center = proline). Dehydroalanine and iminol double bonds involved in a formal [4+2] cycloaddition are drawn in green. The N-terminal leader peptide fragment (N12-mer: LNDLPMDVFELA) observed in the TbtD co-crystal structure has also been superimposed and is represented as yellow sticks. Figure was made by Dillon Cogan and Prof. Satish Nair.

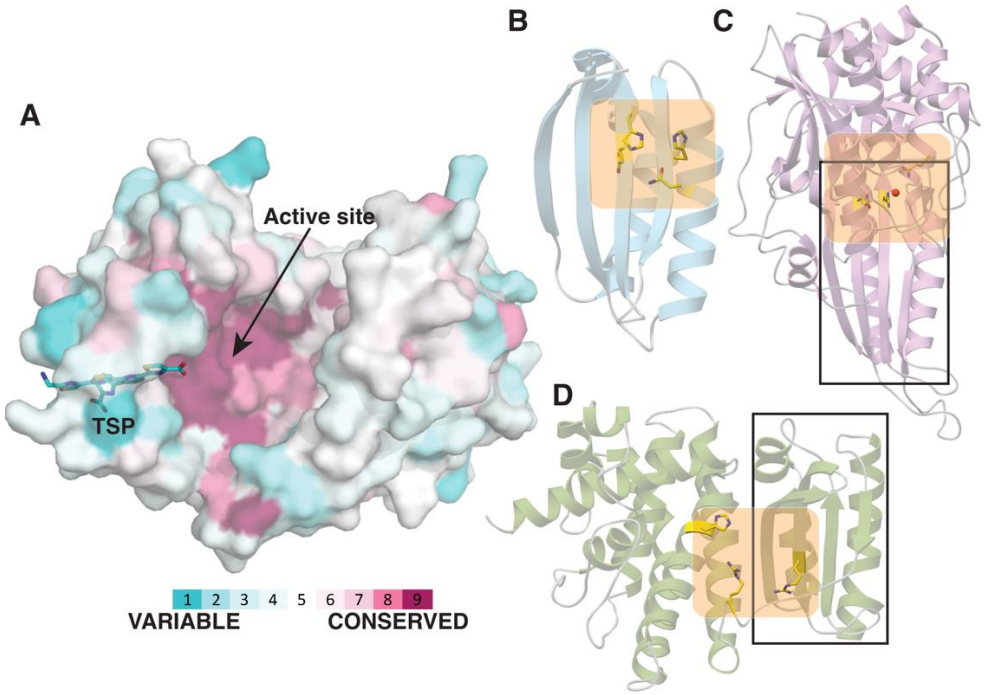


Figure 3.25 Pyridine synthase scaffold is conserved in enzymes with divergent functions. (**A**) PbtD surface rendering with the bound TSP shown as a stick figure in cyan. Using the ConSurf server, the surface has been color-coded according to conservation in primary sequence based upon the alignment of nine sequences of pyridine synthases and colored according to regions of most variability (cyan), modest (white) or highest conservation (magenta). The strongest sequence conservation occurs near the active site where the [4+2] cycloaddition occurs. (**B**) Ribbon diagram of the LsrG epimerase (PDB entry 3QMQ) with active site residues shown as yellow sticks. (**C**) Ribbon diagram of the allantoate amidohydrolase AAH (PDB entry 4PXD) with active site residues shown as yellow sticks and a manganese ion shown as a red sphere. (**D**) Ribbon diagram of the elimination domain of the lanthipeptide dehydratase MibB (PDB entry 5EHK) with active site residues shown as yellow sticks. The ferredoxin-like α+β folds of AAH and MibB are demarcated within the rectangle, and the active sites are colored in orange. Figure was made by Dillon Cogan and Prof. Satish Nair.

A comparison of the TbtD/PbtD structures with that of MibB and LsrG also provides an unexpected occurrence of proteins that contain active sites located at different regions of the polypeptide, illustrating how a common scaffold can accommodate new chemistry at a new active

site region (**Figure 3.25**). Recent *de novo* protein design efforts by Baker and colleagues describe methodologies for engineering proteins with entirely new primary sequences through considerations of long-range interactions observed in prevalent folds such as the ferredoxin-like domain (35). These studies predict the possibility of engineering custom-designed large assemblies from such robust building blocks. The structure of the pyridine synthase provides a conspicuous example of how nature has already employed this strategy to yield new catalysts, based on an ancestral ferredoxin-like fold.

3.3 Experimental methods

General materials and methods

Reagents used for molecular biology experiments were purchased from New England BioLabs (Ipswich, MA), Thermo Fisher Scientific (Waltham, MA), or Gold Biotechnology Inc. (St. Louis, MO). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). *Escherichia coli* DH5α and BL21 (DE3) strains were used for plasmid maintenance and protein overexpression, respectively. Plasmid inserts were sequenced at ACGT Inc. (Wheeling, IL). MALDI-TOF-MS analysis was performed using a Bruker UltrafleXtreme MALDI TOF-TOF mass spectrometer (Bruker Daltonics) in reflector positive mode at the University of Illinois School of Chemical Sciences Mass Spectrometry Laboratory. ESI-MS/MS analyses were performed using a SYNAPT ESI quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters). HiTrap columns for Ni-NTA affinity chromatography were purchased from GE Healthcare.

Molecular biology techniques

Oligonucleotides were purchased from Integrated DNA Technologies Inc. (Coralville, IA). Primers for *T. bispora* $tRNA^{Glu}$ were designed and the dsDNA template for in vitro transcription was prepared as previously described (36). Genes optimized for recombinant expression in *Escherichia coli* were synthesized by GenScript (Piscataway, NJ) as previously described (**Table 2.4**) (2). Site-directed mutagenesis on TbtA was performed using the QuikChange method (Agilent) as per the manufacturer's instructions using PfuTurbo DNA polymerase. The primers for each mutant are listed in **Table 3.3**. The mutations were verified by sequencing using the forward primer or the T7 reverse primer.

In vitro dehydration of TbtA mutants

The following reaction conditions were used for dehydration assays: HEPES pH 7.5 (100 mM), $MgCl_2$ (5 mM), KCl (150 mM), glutamate (10 mM), ATP (5 mM), TbtA hexazole (50 μM), *T. bispora* $tRNA^{Glu}$ (10 μM), *E. coli* GluRS (10 μM), TbtB (5 μM), TbtC (5 μM), and thermostable inorganic pyrophosphatase (TIPP; 0.02 U μL^{-1}) in a final volume of 25 μL. The reaction mixture was incubated at 30 °C for 1.5 h, centrifuged to remove insoluble material (14,000 × *g*, 5 min, 25 °C), and desalted using C18 ZipTips (EMD Millipore). The sample was mixed in a 1:1 ratio with 2,5-dihydroxybenzoic acid matrix, spotted on a Bruker MALDI plate, and analyzed by MALDI-TOF-MS (notebook reference 1/20/2016).

In vivo cyclodehydration and purification of progressive N- to C-terminal Cys to Ala substitutions of TbtA

E. coli BL21 (DE3) cells were co-transformed with a *pTrc99a_TbtF,G,E* plasmid encoding the MBP-tagged TbtF, MBP-tagged TbtG, and TbtE and a *His_6-TbtA-pRSFDuet-1* plasmid encoding the His_6-TbtA mutant peptide of interest. Cells were grown for 16 h on LB agar plates containing 25 μg/mL kanamycin and 50 μg/mL ampicillin at 37 °C. Single colonies were used to inoculate LB containing 25 μg/mL kanamycin and 50 μg/mL ampicillin and the cells were grown at 37 °C for 16 h. This culture was used for 1:100 inoculation into Terrific Broth (24 g/L yeast extract, 12 g/L tryptone, 0.4% glycerol (v/v), 17 mM KH_2PO_4, and 72 mM K_2HPO_4) and grown at 37 °C to an optical density at 600 nm (OD_{600}) of 0.6 before being placed on ice for 20 min. Production of thiazole-installed TbtA mutant peptides was then induced with the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The expression was allowed to proceed for 16-18 h at 18 °C. Cells were harvested by centrifugation at 5,000 × *g* for 15 min. Cell pellets were resuspended in LanA start buffer (20 mM Na_2HPO_4, 150 mM NaCl, 20% glycerol, pH 7.5). Cells were further homogenized by sonication (50% amplitude, pulse on: 2.0 s, pulse off: 5.0 s, total pulse time: 5 min). Insoluble debris was removed by centrifugation at 13,000 × g for 30 min and the supernatant was then applied to Ni resin pre-equilibrated with LanA start buffer. The column was washed with 20 column volumes (CV) of LanA washing buffer (20 mM Na_2HPO_4, 30 mM imidazole, 300 mM NaCl, pH 7.5). The thiazole-installed TbtA mutant peptides were eluted with elute buffer (20 mM Na_2HPO_4, 500 mM imidazole, 100 mM NaCl, pH 7.5). The combined eluted fractions were dialyzed with buffer (20 mM HEPES, 150 mM KCl, pH 7.5) to remove imidazole, and the supernatant was purified on a Shimadzu Prominence Preparative Liquid Chromatography system equipped with a Phenomenex Luna C18 column (250 × 10 mm, 10 μm particle size, 100

Å pore size). Acetonitrile and 10 mM aq. NH_4HCO_3 were used as the mobile phases, and a gradient of 2-80% aq. MeCN over 45 min at 1 mL min^{-1} was used for separation (notebook reference 5/6/2016).

HPLC Purification of TbtA-hexazole core peptide

HPLC-purified TbtA-hexazole (compound **2**) was dissolved in 50 mM Tris-HCl (pH 7.5) to a final concentration of 100 μM. GluC endo-proteinase was added to a concentration of 0.02 mg min^{-1}, and the reaction mixture was incubated at 37 °C for 1 h, followed by addition of MeCN to 20% (v/v) and TFA to 2% (v/v) to precipitate proteins. The insoluble material was removed by centrifugation at 16,100 × *g* for 10 min, and the supernatant was purified on an Agilent 1260 Infinity HPLC system with a Phenomenex Luna C18 column (250 × 4.6 mm, 10 μm particle size, 100 Å pore size). Acetonitrile and 10 mM NH_4HCO_3 were used as the mobile phases, and a gradient of 2–80% aq. MeCN over 45 min at 1 mL min^{-1} was used for separation. TbtA-hexazole core eluted at 38-40% MeCN (notebook reference 8/12/2015).

UPLC-MS and MS/MS of TbtA mutants

GluC endo-proteinase was added to a concentration of 0.02 mg/mL to His_6-TbtA variants after treatment by TbtE/F/G and dehydration by TbtB/C to remove the N-terminal leader peptide. After 30 min of incubation at 37 °C, MeCN was added to 20% (v/v), followed by addition of TFA to 2% (v/v) to precipitate proteins. The insoluble material was removed by centrifugation at 16,100 × *g* for 10 min, and the supernatant was used for analysis. LC-MS was performed using a Waters SYNAPT mass spectrometer outfitted with an ACQUITY UPLC, an ACQUITY Bridged Ethyl Hybrid C8 column (2.1 × 50 mm, 1.7 μm particle size, 200 Å; Waters), an ESI ion source, and a

quadruple TOF detector. A gradient of 2–100% aq. MeCN with 0.1% formic acid (v/v) over 20 min was used. A ramping of cone voltage of 20-25 kV during the scan was performed to generate peptide fragments for MS/MS analysis (notebook reference 7/17/2016).

Labeling of TbtA hexazole (2) with fluorescein isothiocyanate (FITC)

HPLC purified TbtA hexazole (**2**) was dissolved in PBS buffer (pH 8.0) to a final concentration of 50 μM. Fluorescein isothiocyanate (FITC; Sigma-Aldrich) was dissolved in a 10 mg min^{-1} DMSO stock and was added drop wise to the reaction (by adding a drop or two and vortex, repeating this procedure until all FITC was added to a final concentration of 250 μM. This procedure was used to prevent any type of localized precipitation). Labeling was allowed to proceed for 18 h at 25 °C, and then the reaction mixture was lyophilized. The residue was extracted with 20% aq. MeCN. The insoluble material was removed by centrifugation at 16,100 × *g* for 10 min, and the supernatant was purified on an Agilent 1260 Infinity HPLC system with a Phenomenex Luna C18 column (250 × 4.6 mm, 10 μm particle size, 100 Å pore size). Acetonitrile and 10 mM NH_4HCO_3 were used as the mobile phases, and a gradient of 2–80% aq. MeCN over 45 min at 1 mL min^{-1} was used for separation. FITC labeled TbtA-hexazole eluted at 30-31% MeCN (notebook reference 3/15/2016).

Fluorescence polarization (FP) binding assay

The interaction between fluorescently labeled peptides and TbtF, TbtB and TbtD was quantified using an FP assay. To maximize the polarization signal, all proteins remained MBP tagged. In general, protein was serially diluted into binding buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 2.5% (v/v) glycerol, 0.5 mM TCEP) and mixed with 25 nM of the appropriate FITC

labeled peptides. Binding assays were carried out in nonbinding-surface, 384-black-well polystyrene microplates (Corning) and measured using a Synergy H4 Hybrid plate reader (BioTek) with λ_{ex} = 485 nm and λ_{em} = 538 nm and data were recorded with Gen5 software. Prior to measurement, the dilutions were equilibrated with shaking for 30 min at 25 °C. Dissociation constant (K_d) values were calculated from the 50% saturation point using a dose-response curve fit in Origin Pro 9.1 (OriginLab) with three independent titrations. Background fluorescence from the proteins alone was subtracted from the fluorescence polarization signal obtained with the fluorophore (notebook reference 2/8/2016).

Competition FP binding assay

In general, 300 nM protein was mixed with 25 nM of the appropriate FITC-labeled peptide in binding buffer. MBP-tagged peptides for FP studies with TbtF and TbtD were first cleaved by TEV protease at 1:10 loading for 30 min prior to serial dilution. The peptides were serially diluted into binding buffer. Binding assays were carried out on the same instrument and analyzed with the same software as above. IC_{50} values were calculated from the 50% saturation point using a dose-response curve fit in Origin Pro 9.1 (OriginLab) with three independent titrations. K_d values for unlabeled ligands were calculated using the following equation:

$$\mathbf{K_d = IC_{50} / (1 + [L] / K_{d,\ labeled})}$$

Where:

K_d = Dissociation constant for the unlabeled peptide

[L] = Concentration of labeled peptide

$K_{d,\ labeled}$ = Dissociation constant for the labeled peptide

Enzymatic activity of selected TbtD mutant proteins

Seven TbtD mutants were analyzed in more detail by monitoring the peak area of the EIC (extracted-ion chromatogram) of the macrocycle (ESI, *m/z* 1366.2608) produced over time. Experiments with macrocycle standard solutions under the same conditions of the enzymatic activity assay showed that the peak area of the EIC of the macrocycle responded linearly to the macrocycle concentration in a 0.1-25 μM range. The following reaction conditions were used for the assay: HEPES pH 7.5 (50 mM), 15 μM **3**, and 1 μM MBP-TbtD mutant enzyme. The reaction was allowed to proceed at 30 °C, and was quenched by addition of MeCN to 50% (v/v) at different time points. The insoluble material was removed by centrifugation at 16,100 × *g* for 5 min, and the supernatant was subjected to LC-MS using a Waters SYNAPT mass spectrometer outfitted with an ACQUITY UPLC, an ACQUITY Bridged Ethyl Hybrid C18 column (2.1 × 50 mm, 1.7 μm particle size, 200 Å; Waters), an ESI ion source, and a quadruple TOF detector. A gradient of 3–100% aq. MeCN with 0.1% formic acid (v/v) over 20 min was used, and the EIC was recorded (notebook reference 11/30/2015).

Production and purification of TSP and TSP-F

Tri-thiazole-substituted pyridines (TSP and TSP-F) were generated by acid hydrolysis of GE2270A (a gift from Novartis). Hydrolysis was carried out at 110-120 °C for 12 h in the presence of 6 M HCl. The reaction mixture was subsequently lyophilized and extracted with 10% (*v/v*) MeCN. Insoluble material was removed by centrifugation at 10,000 × g for 15 min, and the supernatant was purified on a Shimadzu Prominence Preparative Liquid Chromatography system equipped with a Phenomenex Luna C18 column (250 × 10 mm, 10 μm particle size, 100 Å pore size). MeCN and 10 mM aq. NH_4HCO_3 were used as the mobile phases with a gradient of 2 - 80%

aq. MeCN over 45 min at 7 mL min^{-1} for separation. TSP and TSP-F eluted from the column with a retention time around 13 min and 16 respectively (notebook reference 11/28/2016).

Methyl-esterification of TSP and TSP-F

In a typical reaction, 1 mg of TSP and TSP-F was dissolved in 1 mL of 2 M methanolic HCl (prepared by careful dropwise addition of 160 μL of acetyl chloride to 1 mL of ice-cold methanol while stirring, with the reaction vessel on ice) and then incubated at 25 °C with shaking for 1h. The sample was lyophilized and re-dissolved in 1 mL of 20% MeCN (v/v) and employed for purified on an Agilent 1260 Infinity HPLC system with a Phenomenex Luna C18 column (250 × 4.6 mm, 10 μm particle size, 100 Å pore size). Acetonitrile and 10 mM NH_4HCO_3 were used as the mobile phases, and a gradient of 2–80% aq. MeCN over 45 min at 1 mL min^{-1} was used for separation. Methyl-esterified TSP and TSP-F eluted from the column with a retention time around 15 min and 18 respectively (notebook reference 2/15/2017).

Table 3.3 Oligonucleotides used in this study.

Primer	Oligonucleotide Sequence
tbtA C12A F	TTTGTTGTAGCGCCAGCAGCGCCTAACTCGAGCACCACCAC
tbtA C12A R	TAGGCGCTGCTGGCGCTACAACAAATATAACAAAAGCAATTAC
tbtA C12A, C10A F	GTTATATTTGTGCGAGCGCCAGCAGCGCCTAACTCGAGCAC
tbtA C12A, C10A R	CTGCTGGCGCTCGCACAAATATAACAAAAGCAATTACAGCTTG
tbtA C12A, C10A, C9A F	TTTGTTATATTGCGGCGAGCGCCAGCAGCGCCTAACTCGAG
tbtA C12A, C10A, C9A R	CTGGCGCTCGCCGCAATATAACAAAAGCAATTACAGCTTGC
tbtA C12A, C10A, C9A, C6A F	GTAATTGCTTTGCGTATATTGCGGCGAGCGCCAGCAGCGCC
tbtA C12A, C10A, C9A, C6A R	GCCGCAATATACGCAAAGCAATTACAGCTTGCACCAACTTC
tbtA C12A, C10A, C9A, C6A F	CTGCGTTTCATGCCGAACTGGAAAGCCGTCCGGAATGGGCAC
tbtA C12A, C10A, C9A, C6A R	CTTTCCAGTTCGGCATGAAACGCAGAAACTTCAGCCAGGCTC

Table 3.3 (cont.)

Primer	Oligonucleotide Sequence
tbtA C12A, C10A, C9A, C6A, C4A F	CAAGCTGTAATGCGTTTGCGTATATTGCGGCGAGCGCCAGC
tbtA C12A, C10A, C9A, C6A, C4A R	ATATACGCAAACGCATTACAGCTTGCACCAACTTCGGTCATAC
tbtA C12A, C10A, C9A, C6A, C4A, C2A F	TTGGTGCAAGCGCGAATGCGTTTGCGTATATTGCGGCGAGC
tbtA C12A, C10A, C9A, C6A, C4A, C2A R	GCAAACGCATTCGCGCTTGCACCAACTTCGGTCATACCATG
tbtA S1A F	GGTGCAGCGTGTAATTGCTTTTGTTATATTTGTTGTAGC
tbtA S1A R	CAAAAGCAATTACACGCTGCACCAACTTCGGTCATAC
tbtA S11A F	GTTATATTTGTTGTGCTTGCAGCAGCGCCTAATG
tbtA S11A R	CTGCTGCAAGCACAACAAATATAACAAAAGCAATTACAGCTTG
MBP-tbtB F	GAGAACCTGTACTTCCAATCCATGCGTCTGGTGGAAC
MBP-tbtB R	GGTGCTCGAGTGCGGCCGCATTATTCACCCAGTTCAACCAG
tbtC R60A F	CATCTGCGTTTCGCTGTTCGTCTGCTGCCG
tbtC R60A R	CAGACGAACAGCGAAACGCAGATGCGGAC
tbtA S13A F	GTTGTAGCTGCGCAAGCGCCTAATGCGGC
tbtA S13A R	CATTAGGCGCTTGCGCAGCTACAACAAATATAACAAAAGCAATTAC
tbtA S14A F	GTAGCTGCAGCGCTGCCTAATGCGGCCGC
tbtA S14A R	GCATTAGGCAGCGCTGCAGCTACAACAAATATAACAAAAGC
tbtA C2A F	GTGCAAGCGCGAATTGCTTTTGTTATATTTGTTGTAGC
tbtA C2A R	CAAAAGCAATTCGCGCTTGCACCAACTTCG
tbtA C2A, C4A F	GCGCGAATGCATTTTGTTATATTTGTTGTAGCTGCAG
tbtA C2A, C4A R	CAAATATAACAAAATGCATTCGCGCTTGCACC
tbtA C2A, C4A, C6A F	GAATGCATTTGCGTATATTTGTTGTAGCTGCAGCAG
tbtA C2A, C4A, C6A R	CAAATATACGCAAATGCATTCGCGCTTGCACC
tbtA C2A, C4A, C6A, C9A F	ATCGAAGGTCTCATGCGTATATTGCATGTAGCTGCAGCAGCG
tbtA C2A, C4A, C6A, C9A R	TTAGCTGGTCTCACGCAAATGCATTCGCGCTTGCACCAAC
tbtA C2A, C4A, C6A, C9A, C10A F	ATCGAAGGTCTCATGCGTATATTGCAGCTAGCTGCAGCAGCGC
tbtA C2A, C4A, C6A, C9A, C10A R	TTAGCTGGTCTCACGCAAATGCATTCGCGCTTGCACCAAC
mbp F	GAGGAAGAGTTGGCGAAAGATCCACGTA
T7 F	TAATACGACTCACTATAGGG
T7 R	GCTAGTTATTGCTCAGCGG
tbtD H46A F	ATTTTCTGCGTGCTTGGCGTCGCGGTCCGCACCTGCGCATTTAC
tbtD H46A R	CCGCGACGCCAAGCACGCAGAAAATATGCCATCGGTGCCTG

Table 3.3 (cont.)

Primer	Oligonucleotide Sequence
tbtD W47A F	TTCTGCGTCATGCGCGTCGCGGTCCGCACCTGCGCATTTAC
tbtD W47A R	GGACCGCGACGCGCATGACGCAGAAAATATGCCATCGGTGC
tbtD R49A F	GTCATTGGCGTGCCGGTCCGCACCTGCGCATTTACGTTAG
tbtD R49A R	CAGGTGCGGACCGGCACGCCAATGACGCAGAAAATATGCCATC
tbtD H52A F	GTCGCGGTCCGGCCCTGCGCATTTACGTTAGTACCACGCGTG
tbtD H52A R	GTAAATGCGCAGGGCCGGACCGCGACGCCAATGACGCAG
tbtD R54A F	GTCCGCACCTGGCCATTTACGTTAGTACCACGCGTGAAGCAC
tbtD R54A R	CTAACGTAAATGGCCAGGTGCGGACCGCGACGCCAATGACG
tbtD S85A F	GTGCACGTCCGGCTCCGGGTATGGCTGATCCGTCCGCGTTC
tbtD S85A R	GCCATACCCGGAGCCGGACGTGCACGCAGATAACCGCCCAC
tbtD H98A F	GTTCCTGCCGCTGGCTGAACGCCTGGCCGAACTGGAAGGCGAAG
tbtD H98A R	GCCAGGCGTTCAGCCAGCGGCAGGAACGCGGACGGATCAGC
tbtD E105A F	CTGGCCGAACTGGCAGGCGAAGATGGTCCGCTGATGCCGTGG
tbtD E105A R	CCATCTTCGCCTGCCAGTTCGGCCAGGCGTTCATGCAGCGG
tbtD N118A F	GGTCTCCGGACGCCACGATTCACGCAGAAGGTGAACGTCCGG
tbtD N118A R	GCGTGAATCGTGGCGTCCGGAGACCACGGCATCAGCGGACC
tbtD T144A F	TTTATGCAGACGCCACGCCGAGCGTTTACCACGCGCTGGAAC
tbtD T144A R	ACGCTCGGCGTGGCGTCTGCATAAAAATCGGCCAGCAGAAC
tbtD R189A F	GTACCTCACTGGCTTCGCACGCAGAAGCTTATCTGGCACGTC
tbtD R189A R	TCTGCGTGCGAAGCCAGTGAGGTACGTGCGACCGGCAGACC
tbtD S190A F	CCTCACTGCGTGCGCACGCAGAAGCTTATCTGGCACGTCGC
tbtD S190A R	GCTTCTGCGTGCGCACGCAGTGAGGTACGTGCGACCGGCAG
tbtD H191A F	CACTGCGTTCGGCCGCAGAAGCTTATCTGGCACGTCGCTCC
tbtD H191A R	TAAGCTTCTGCGGCCGAACGCAGTGAGGTACGTGCGACCGG
tbtD E193A F	GTTCGCACGCAGCAGCTTATCTGGCACGTCGCTCCGATGGTG
tbtD E193A R	GCCAGATAAGCTGCTGCGTGCGAACGCAGTGAGGTACGTGC
tbtD S287A F	TGGCTGAAGTTGCTGCGTTTCATCGCGAACTGGAAAGCCGTC
tbtD S287A R	CGATGAAACGCAGCAACTTCAGCCAGGCTCGGCAGATCGCG
tbtD H290A F	GTTTCTGCGTTTGCTCGCGAACTGGAAAGCCGTCCGGAATGG
tbtD H290A R	TCCAGTTCGCGAGCAAACGCAGAAACTTCAGCCAGGCTCGG
tbtD R291A F	CTGCGTTTCATGCCGAACTGGAAAGCCGTCCGGAATGGGCAC
tbtD R291A R	CTTTCCAGTTCGGCATGAAACGCAGAAACTTCAGCCAGGCTC
tbtD N316A F	GTCTGGTTATTGCTTGCACCTACCTGCACCTGACGCGTCTG
tbtD N316A R	CAGGTAGGTGCAAGCAATAACCAGACGATATGCGCCGAATGC
tbtD Y319A F	TTAATTGCACCGCCCTGCACCTGACGCGTCTGGGTCTGACC

Table 3.3 (cont.)

Primer	Oligonucleotide Sequence
tbtD Y319A R	GTCAGGTGCAGGGCGGTGCAATTAATAACCAGACGATATGC
tbtD R324A F	CTGCACCTGACGGCTCTGGGTCTGACCCCGCATCAACGTTTC
tbtD R324A R	GTCAGACCCAGAGCCGTCAGGTGCAGGTAGGTGCAATTAATAAC
tbtD R332A F	CCCCGCATCAAGCTTTCCTGGTCTGTCACCTGGCAGCAGATG
tbtD R332A R	CAGACCAGGAAAGCTTGATGCGGGGTCAGACCCAGACGCGTC
tbtD C336A F	GTTTCCTGGTCGCTCACCTGGCAGCAGATGCAGCTGCAGAC
tbtD C336A R	GCTGCCAGGTGAGCGACCAGGAAACGTTGATGCGGGGTCAG
tbtD H337A F	TCCTGGTCTGTGCCCTGGCAGCAGATGCAGCTGCAGACGTG
tbtD H337A R	TCTGCTGCCAGGGCACAGACCAGGAAACGTTGATGCGGGGTC

3.4 References

1. Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G. L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., & van der Donk, W. A. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep., 30*, 108-160.

2. Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., & van der Donk, W. A. (2015). In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *J. Am. Chem. Soc., 137*, 16012-16015.

3. Morris, R. P., Leeds, J. A., Naegeli, H. U., Oberer, L., Memmert, K., Weber, E., Lamarche, M. J., Parker, C. N., Burrer, N., Esterow, S., Hein, A. E., Schmitt, E. K., & Krastel, P. (2009). Ribosomally Synthesized Thiopeptide Antibiotics Targeting Elongation Factor Tu. *J. Am. Chem. Soc.*, 5946-5955.

4. Jeon, B. S., Wang, S. A., Ruszczycky, M. W., & Liu, H. W. (2017). Natural [4 + 2]-Cyclases. *Chem. Rev., 117*, 5367-5388.

5. Hughes, R. A., Thompson, S. P., Alcaraz, L., & Moody, C. J. (2005). Total synthesis of the thiopeptide antibiotic amythiamicin D. *J. Am. Chem. Soc., 127*, 15644-15651.

6. Zhang, Z., Hudson, G. A., Mahanta, N., Tietz, J. I., van der Donk, W. A., & Mitchell, D. A. (2016). Biosynthetic Timing and Substrate Specificity for the Thiopeptide Thiomuracin. *J. Am. Chem. Soc., 138*, 15511-15514.

7. Wever, W. J., Bogart, J. W., Baccile, J. A., Chan, A. N., Schroeder, F. C., & Bowers, A. A. (2015). Chemoenzymatic Synthesis of Thiazolyl Peptide Natural Products Featuring an Enzyme-Catalyzed Formal [4 + 2] Cycloaddition. *J. Am. Chem. Soc., 137*, 3494-3497.

8. Fage, C. D., Isiorho, E. A., Liu, Y., Wagner, D. T., Liu, H. W., & Keatinge-Clay, A. T. (2015). The structure of SpnF, a standalone enzyme that catalyzes [4 + 2] cycloaddition. *Nat. Chem. Biol., 11*, 256-258.

9. Kim, H. J., Ruszczycky, M. W., Choi, S. H., Liu, Y. N., & Liu, H. W. (2011). Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature, 473*, 109-112.

10. Tian, Z., Sun, P., Yan, Y., Wu, Z., Zheng, Q., Zhou, S., Zhang, H., Yu, F., Jia, X., Chen, D., Mandi, A., Kurtan, T., & Liu, W. (2015). An enzymatic [4+2] cyclization cascade creates the pentacyclic core of pyrroindomycins. *Nat. Chem. Biol.*, *11*, 259-265.

11. Li, L., Yu, P., Tang, M. C., Zou, Y., Gao, S. S., Hung, Y. S., Zhao, M., Watanabe, K., Houk, K. N., & Tang, Y. (2016). Biochemical characterization of a eukaryotic decalin-forming diels-alderase. *J. Am. Chem. Soc.*, *138*, 15837-15840.

12. Ohashi, M., Liu, F., Hai, Y., Chen, M., Tang, M. C., Yang, Z., Sato, M., Watanabe, K., Houk, K. N., & Tang, Y. (2017). SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature*, *549*, 502-506.

13. Burkhart, B. J., Hudson, G. A., Dunbar, K. L., & Mitchell, D. A. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.*, *11*, 564-570.

14. Koehnke, J., Mann, G., Bent, A. F., Ludewig, H., Shirran, S., Botting, C., Lebl, T., Houssen, W. E., Jaspars, M., & Naismith, J. H. (2015). Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat. Chem. Biol.*, *11*, 558-563.

15. Dunbar, K. L., Tietz, J. I., Cox, C. L., Burkhart, B. J., & Mitchell, D. A. (2015). Identification of an Auxiliary Leader Peptide-Binding Protein Required for Azoline Formation in Ribosomal Natural Products. *J. Am. Chem. Soc.*, *137*, 7672-7677.

16. Latham, J. A., Iavarone, A. T., Barr, I., Juthani, P. V., & Klinman, J. P. (2015). PqqD is a novel peptide chaperone that forms a ternary complex with the radical s-adenosylmethionine protein pqqe in the pyrroloquinoline quinone biosynthetic pathway. *J. Biol. Chem.*, *290*, 12908-12918.

17. Li, J., Qu, X., He, X., Duan, L., Wu, G., Bi, D., Deng, Z., Liu, W., & Ou, H.-Y. (2012). ThioFinder: A Web-Based Tool for the Identification of Thiopeptide Gene Clusters in DNA Sequences. *PLoS ONE*, *7*, e45878.

18. Zhang, Q., & Liu, W. (2013). Biosynthesis of thiopeptide antibiotics and their pathway engineering. *Nat. Prod. Rep.*, *30*, 218-226.

19. Just-Baringo, X., Albericio, F., & Alvarez, M. (2014). Thiopeptide antibiotics: retrospective and recent advances. *Mar. Drugs*, *12*, 317-351.

20. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., van der Donk, W. A., & Nair, S. K. (2015). Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature*, *517*, 509-512.

21. Bailey, T. L., Williams, N., Misleh, C., & Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.*, *34*, W369-W373.

22. Wever, W. J., Bogart, J. W., Baccile, J. A., Chan, A. N., Schroeder, F. C., & Bowers, A. A. (2015). Chemoenzymatic synthesis of thiazolyl peptide natural products featuring an enzyme-catalyzed formal [4 + 2] cycloaddition. *J. Am. Chem. Soc., 137*, 3494-3497.

23. Wever, W. J., Bogart, J. W., & Bowers, A. A. (2016). Identification of Pyridine Synthase Recognition Sequences Allows a Modular Solid-Phase Route to Thiopeptide Variants. *J. Am. Chem. Soc., 138*, 13461-13464.

24. Ortega, M. A., Hao, Y., Walker, M. C., Donadio, S., Sosio, M., Nair, S. K., & van der Donk, W. A. (2016). Structure and trna specificity of mibB, a lantibiotic dehydratase from actinobacteria involved in nai-107 biosynthesis. *Cell Chem. Biol., 23*, 370-380.

25. Zhang, Z., Hudson, G. A., Mahanta, N., Tietz, J. I., van der Donk, W. A., & Mitchell, D. A. (2016). Biosynthetic timing and substrate specificity for the thiopeptide thiomuracin. *J. Am. Chem. Soc., 138*, 15511-15514.

26. Koehnke, J., Mann, G., Bent, A. F., Ludewig, H., Shirran, S., Botting, C., Lebl, T., Houssen, W. E., Jaspars, M., & Naismith, J. H. (2015). Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat. Chem. Biol., 11*, 558-563.

27. Bond, C. S., & Schuttelkopf, A. W. (2009). ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr. D Biol. Crystallogr., 65*, 510-512.

28. Kettenring, J., Colombo, L., Ferrari, P., Tavecchia, P., Nebuloni, M., Vekey, K., Gallo, G. G., & Selva, E. (1991). Antibiotic GE2270 a: a novel inhibitor of bacterial protein synthesis. II. Structure elucidation. *J. Antibiot., 44*, 702-715.

29. Burkhart, B. J., Schwalen, C. J., Mann, G., Naismith, J. H., & Mitchell, D. A. (2017). YcaO-dependent posttranslational amide activation: biosynthesis, structure, and function. *Chem. Rev., 117*, 5389-5456.

30. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., & Whalen, K. L. (2015). Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, 1854*, 1019-1037.

31. Marques, J. C., Lamosa, P., Russell, C., Ventura, R., Maycock, C., Semmelhack, M. F., Miller, S. T., & Xavier, K. B. (2011). Processing the interspecies quorum-sensing signal autoinducer-2 (AI-2): characterization of phospho-(S)-4,5-dihydroxy-2,3-pentanedione isomerization by LsrG protein. *J. Biol. Chem., 286*, 18331-18343.

32. Ma, B. G., Chen, L., Ji, H. F., Chen, Z. H., Yang, F. R., Wang, L., Qu, G., Jiang, Y. Y., Ji, C., & Zhang, H. Y. (2008). Characters of very ancient proteins. *Biochem. Biophys. Res. Commun., 366*, 607-611.

33. Werner, A. K., Romeis, T., & Witte, C. P. (2010). Ureide catabolism in Arabidopsis thaliana and Escherichia coli. *Nat. Chem. Biol., 6*, 19-21.

34. Shin, I., Percudani, R., & Rhee, S. (2012). Structural and functional insights into (S)-ureidoglycine aminohydrolase, key enzyme of purine catabolism in Arabidopsis thaliana. *J. Biol. Chem., 287*, 18796-18805.

35. Koga, N., Tatsumi-Koga, R., Liu, G., Xiao, R., Acton, T. B., Montelione, G. T., & Baker, D. (2012). Principles for designing ideal protein structures. *Nature, 491*, 222-227.

36. Sherlin, L. D., Bullock, T. L., Nissan, T. A., Perona, J. J., Lariviere, F. J., Uhlenbeck, O. C., & Scaringe, S. A. (2001). Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA, 7*, 1671-1678.

37. Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J. B., Blin, K., de Bruijn, I., Chooi, Y. H., Claesen, J., Coates, R. C., Cruz-Morales, P., Duddela, S., Dusterhus, S., Edwards, D. J., Fewer, D. P., Garg, N., Geiger, C., Gomez-Escribano, J. P., Greule, A., Hadjithomas, M., Haines, A. S., Helfrich, E. J., Hillwig, M. L., Ishida, K., Jones, A. C., Jones, C. S., Jungmann, K., Kegler, C., Kim, H. U., Kotter, P., Krug, D., Masschelein, J., Melnik, A. V., Mantovani, S. M., Monroe, E. A., Moore, M., Moss, N., Nutzmann, H. W., Pan, G., Pati, A., Petras, D., Reen, F. J., Rosconi, F., Rui, Z., Tian, Z., Tobias, N. J., Tsunematsu, Y., Wiemann, P., Wyckoff, E., Yan, X., Yim, G., Yu, F., Xie, Y., Aigle, B., Apel, A. K., Balibar, C. J., Balskus, E. P., Barona-Gomez, F., Bechthold, A., Bode, H. B., Borriss, R., Brady, S. F., Brakhage, A. A., Caffrey, P., Cheng, Y. Q., Clardy, J., Cox, R. J., De Mot, R., Donadio, S., Donia, M. S., van der Donk, W. A., Dorrestein, P. C., Doyle, S., Driessen, A. J., Ehling-Schulz, M., Entian, K. D., Fischbach, M. A., Gerwick, L., Gerwick, W. H., Gross, H., Gust, B., Hertweck, C., Hofte, M., Jensen, S. E., Ju, J., Katz, L., Kaysser, L., Klassen, J. L., Keller, N. P., Kormanec, J., Kuipers, O. P., Kuzuyama, T., Kyrpides, N. C., Kwon, H. J., Lautru, S., Lavigne, R., Lee, C. Y., Linquan, B., Liu, X., Liu, W., Luzhetskyy, A., Mahmud, T., Mast, Y., Mendez, C., Metsa-Ketela, M., Micklefield, J., Mitchell, D. A., Moore, B. S., Moreira, L. M., Muller, R., Neilan, B. A., Nett, M., Nielsen, J., O'Gara, F., Oikawa, H., Osbourn, A., Osburne, M. S., Ostash, B., Payne, S. M., Pernodet, J. L., Petricek, M., Piel, J., Ploux, O., Raaijmakers, J. M., Salas, J. A., Schmitt, E. K., Scott, B., Seipke, R. F., Shen, B., Sherman, D. H., Sivonen, K., Smanski, M. J., Sosio, M., Stegmann, E., Süssmuth, R. D., Tahlan, K., Thomas, C. M., Tang, Y., Truman, A. W., Viaud, M., Walton, J. D., Walsh, C. T., Weber, T., van Wezel, G. P., Wilkinson, B., Willey, J. M., Wohlleben, W., Wright, G. D., Ziemert, N., Zhang, C., Zotchev, S. B., Breitling, R., Takano, E., & Glockner, F. O. (2015). Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol., 11*, 625-631.

38. Duan, L., Wang, S., Liao, R., & Liu, W. (2012). Insights into quinaldic acid moiety formation in thiostrepton biosynthesis facilitating fluorinated thiopeptide generation. *Chem. Biol., 19*, 443-448.

39. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol., 215*, 403-410.

40. Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Comput. Biol., 7*, e1002195.

41. Tietz, J. I., Schwalen, C. J., Patel, P. S., Maxson, T., Blair, P. M., Tai, H.-C., Zakai, U. I., & Mitchell, D. A. (2017). A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol., 13*, 470.

42. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., & Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res., 37*, W202-208.

43. Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics, 27*, 1017-1018.

CHAPTER 4: BIOCHEMICAL AND MECHANISTIC CHARACTERIZATION OF A RADICAL *S*-ADENOSYL-L-METHIONINE (SAM) THIAZOLE *C*-METHYLTRANSFERASE IN THIOMURACIN BIOSYNTHESIS [d]

4.1 Introduction

Radical SAM enzymes catalyze unique chemical reactions in primary metabolism and natural product biosynthesis (1-4). A subset of radical SAM proteins catalyze methylation of non-nucleophilic acceptors, such as sp^2 or sp^3 carbons (5-7). Currently, four distinct classes of radical SAM methyltransferases (MTs) have been identified that differ in their cofactor/co-substrate requirements and their proposed reaction mechanisms (7-10). Class A includes the RNA-modifying enzymes RlmN and Cfr (11, 12). Class B comprises cobalamin-dependent MTs, such as TsrM (13, 14), Fom3 (15-19), GenK (20), CysS (21), and PoyC (22). Class C are cobalamin-independent enzymes and found in a variety of natural product biosynthetic gene clusters (BGCs) (23, 24). Although class A and B enzymes have been extensively studied (17-19, 25-30), class C MTs have received much less attention (5, 9). The only examples of partially characterized class C MTs are ChuW and NosN (31, 32). Class D includes a methylenetetrahydrofolate-dependent enzyme involved in methanopterin biosynthesis (33).

[d] Parts of the chapter are reprinted with permission from: Mahanta, N., Zhang, Z., Hudson, G. A., van der Donk, W. A., & Mitchell, D. A. (2017). Reconstitution and Substrate Specificity of the Radical SAM Thiazole *C*-Methyltransferase in Thiomuracin Biosynthesis. *J. Am. Chem. Soc.*, *139*, 4310-4313. Copyright © 2017 American Chemical Society. https://pubs.acs.org/doi/abs/10.1021/jacs.7b00693
Parts of the chapter are reprinted with permission from: Zhang, Z., Mahanta, N., Hudson, G. A., Mitchell, D. A., & van der Donk, W. A. (2017). Mechanism of a Class C Radical *S*-Adenosyl-L-methionine Thiazole Methyl Transferase. *J. Am. Chem. Soc.*, *139*, 18623-18631. Copyright © 2017 American Chemical Society. https://pubs.acs.org/doi/abs/10.1021/jacs.7b10203

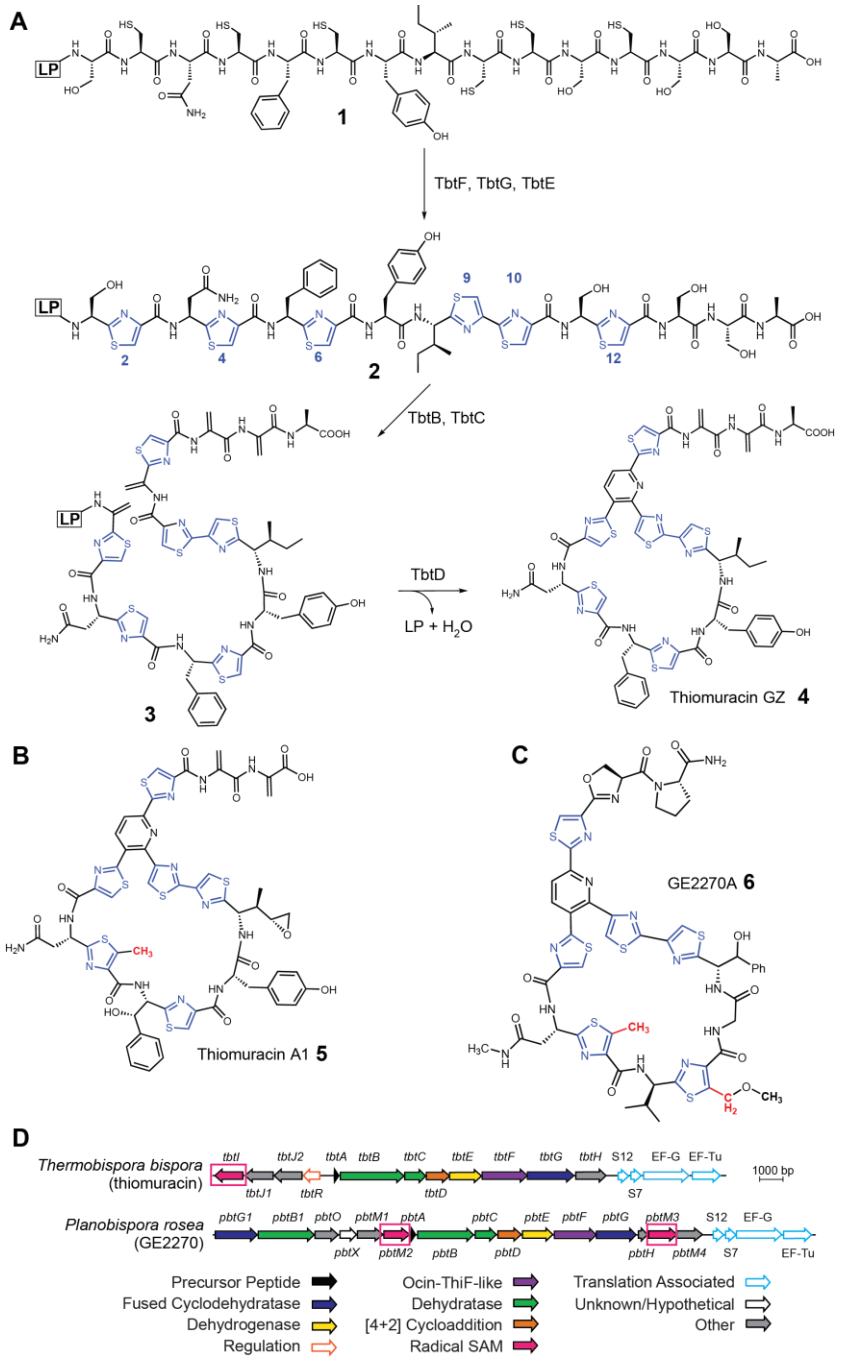


Figure 4.1 (**A**) Thiomuracin GZ (**4**) biosynthesis. (**B**) Structure of thiomuracin A1 (**5**). (**C**) Structure of GE2270A (**6**). LP, leader peptide. Red, *C*-methylation sites. (**D**) Biosynthetic gene clusters from *Thermobispora bispora* and *Planobispora rosea*, respectively. Shown in pink and boxed are the class C radical SAM MTs (e.g. TbtI, the subject of the current chapter). The BGC of **6** additionally encodes two additional "classic" SAM-MTs, PbtM1 and PbtM4. These MTs act upon nucleophilic centers and proceed via a polar mechanism unlike the radical SAM enzymes.

In both thiomuracin A1 (**5**) and GE2270A (**6**) the thiazole (Thz) arising from Cys4 is *C*-methylated; in **6**, the Thz at position 6 is also *C*-methylated (**Figure 4.1**) (34). These transformations are remarkable in that they occur on unactivated sp^2 carbon centers. Inspection of the BGC of **5** identifies a single gene, *tbtI*, annotated as a class C radical SAM MT enzyme (**Figure 4.1D**) (35). Previous studies (36) on the biosynthesis of **6** suggested that TbtI is involved in thiazole *C*-methylation of **5**. The BGC for **6** encodes a nearly identical radical SAM protein, PbtM2, in addition to a paralog, PbtM3 (**Figure 4.1D**). Gene deletion experiments demonstrated that PbtM2 and PbtM3 *C*-methylate Thz4 and Thz6 of **6**, respectively (37). In this chapter, we present the investigation of the timing and substrate scope of thiazole *C*-methylation by TbtI involved in the biosynthesis of the thiopeptide thiomuracin, as well as a detailed mechanistic study of this remarkable transformation.

4.2 Results and discussion

4.2.1 Reconstitution and substrate specificity of the radical SAM thiazole *C*-methyltransferase TbtI

The radical SAM thiazole *C*-methyltransferase TbtI was first recombinantly expressed and purified from *Escherichia coli* as an N-terminal maltose binding protein (MBP) fusion (**Figure 4.2A**) (38). Using the ferrozine and methylene blue assays, Dr. Nilkamal Mahanta found that the expressed enzyme contained 1.6 ± 0.2 Fe and 1.9 ± 0.3 sulfides per monomer. Fe-S cluster reconstitution with recombinant cysteine desulfurase (IscS), Fe (II), Cys, and dithiothreitol yielded 3.7 ± 0.3 Fe and 4.1 ± 0.3 sulfides per protein, near the expected values for a single [4Fe-4S] cluster. UV-Vis analysis of TbtI supported the proper incorporation of the Fe-S cluster (**Figure 4.2B**).

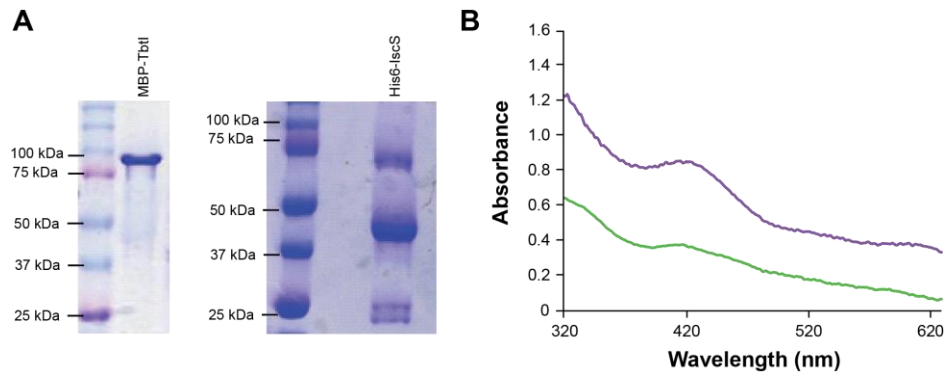


Figure 4.2 (**A**) SDS-PAGE analysis of proteins used in this chapter. (**B**) UV-Vis spectrum of TbtI. The UV-Vis spectrum of MBP-tagged TbtI (30 μM) collected on a Cary 4000 spectrophotometer. Shown are spectra of MBP-TbtI (green) as purified from *E. coli* and MBP-TbtI after in vitro Fe-S cluster reconstitution with IscS (purple). Figure made by Dr. Nilkamal Mahanta

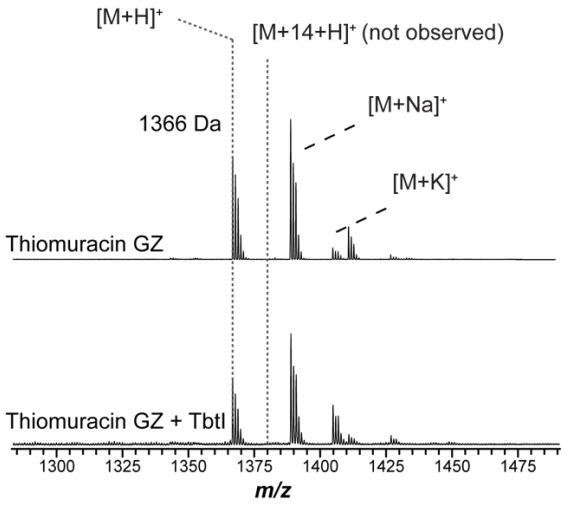


Figure 4.3 TbtI reaction with thiomuracin GZ (**4**). Shown are MALDI-TOF-MS spectra of thiomuracin GZ (top) and after reaction with TbtI supplemented with SAM and sodium dithionite (bottom). No methylation was observed even after extended reaction times. *m/z* assignments: 1366 Da, thiomuracin GZ (**4**).

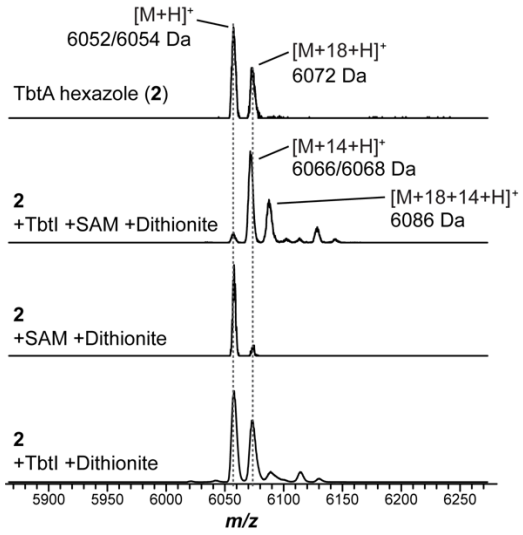
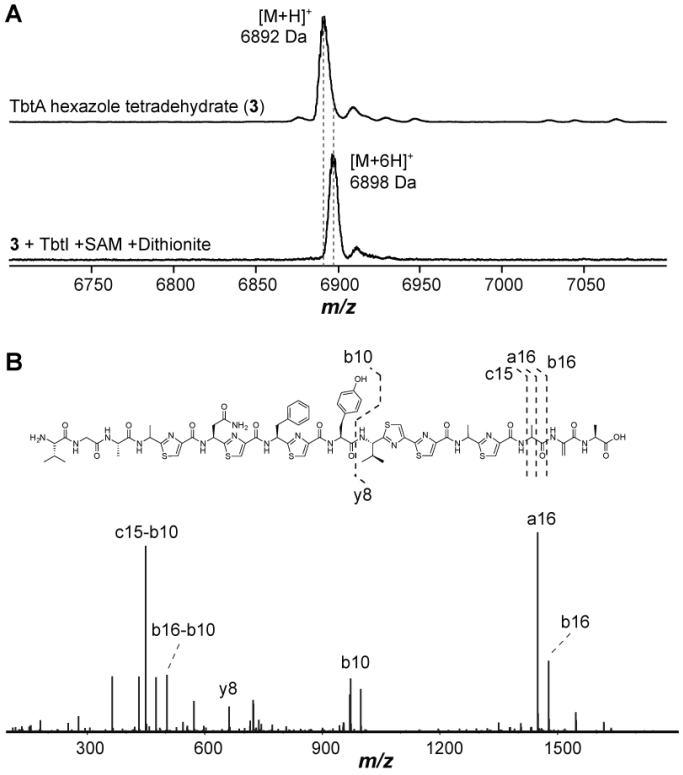


Figure 4.4 TbtI reaction with TbtA hexazole (**2**). TbtA hexazole (**2**) was only methylated (+14 Da) in the presence of TbtI, SAM, and dithionite as monitored by MALDI-TOF-MS. Omission of TbtI, SAM, or dithionite resulted in no reaction. *m/z* assignments: 6052 Da, TbtA hexazole (**2**) which at this resolution, forms a single peak with a minor product at *m/z* 6054, corresponding to TbtA pentazole/monoazoline, due to incomplete oxidation of one of the thiazolines; 6072 Da, TbtA pentazole with a hydrolyzed thiazoline (+18 Da); 6068 Da, methylated TbtA hexazole; 6086 Da, TbtA methylated pentazole with a hydrolyzed thiazoline.

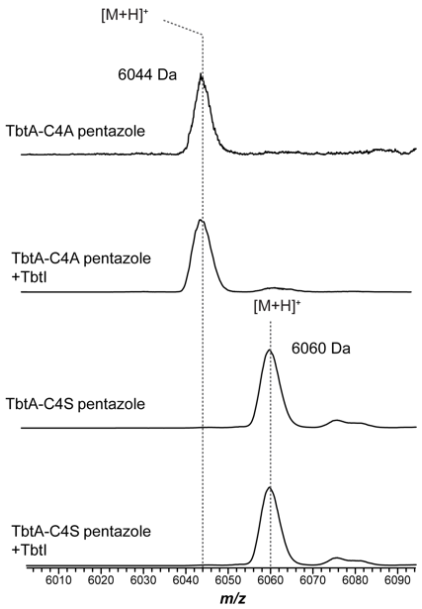
To determine the timing of thiazole *C*-methylation, Dr. Nilkamal Mahanta first evaluated whether **4** was a substrate. Under anaerobic conditions, **4** was treated with TbtI in the presence of SAM and dithionite. The reaction mixture was analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) and no mass shift (+14 Da) for methylation was observed (**Figure 4.3**). To investigate if methylation occurred at an earlier point in the biosynthesis, I tested the TbtA hexazole **2** and the downstream tetradehydrate **3** as substrates for TbtI. Analysis by MALDI-TOF-MS showed a +14 Da species consistent with a single methylation after the reaction with **2** (**Figure 4.4**). Control experiments lacking SAM, TbtI, or dithionite did not result in a mass shift. Under our reaction conditions, **3** was not a substrate for TbtI (**Figure 4.5**). Based on the structures of **5** and **6** (**Figure 4.1**) and previous gene deletion

experiments (37), methylation of **2** likely occurred at Thz4. To test this hypothesis, the pentazole-containing versions of TbtA-C4A and TbtA-C4S were prepared and tested as substrates for TbtI. No mass shift was observed in either case, implicating Thz4 as a critical moiety for TbtI (**Figure 4.6**).



His$_6$-TEV-TbtA: PHHHHHHSQENLYFQSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

Figure 4.5 TbtI reaction with TbtA hexazole tetradehydrate (**3**). (**A**) TbtA tetradehydrate was tested as a substrate for TbtI. The amino acid sequence shows the site of thiazoles (blue) and dehydrated Ser (green). No methylation (+14 Da) was observed in the presence of TbtI, SAM, and dithionite as monitored by MALDI-TOF-MS. Instead, a +6 Da species was observed. *m/z* assignments: 6892 Da, TbtA hexazole tetradehydrate (**3**); 6898 Da, off-pathway product of TbtI reaction with **3**. (**B**) High-resolution tandem MS analysis of the +6 Da species indicates that **3** underwent reduction at the three most N-terminal alkene groups. This is a known reaction for dithionite (39-41), which is required for the *in vitro* TbtI reactions.



His_6-TbtA (C4A): PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNAFCYICCSCSSA

His_6-TbtA (C4S): PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNSFCYICCSCSSA

Figure 4.6 TbtI reaction with TbtA-C4A and TbtA-C4S pentazoles. TbtA pentazoles carrying the C4A or C4S substitutions were not substrates for TbtI, as visualized by MALDI-TOF-MS. Sites of thiazole installation are blue and the Ala/Ser substitutions are shown in red. *m/z* assignments: 6044 Da, TbtA-C4A pentazole; 6060 Da, TbtA-C4S pentazole.

A prominent feature of RiPP biosynthesis is the role of the LP in substrate recognition by the modifying enzymes (42). TbtI lacks an identifiable RiPP Recognition Element (RRE) (43) which is normally an indicator of LP dependence. However, TbtD also lacks an RRE yet is LP-dependent as discussed in Chapter 2 (44); further, the RRE-containing TbtB is LP-independent as discussed in Chapter 3 (45). To assess the LP-dependence of TbtI, **2** was treated with the endo-proteinase GluC to generate **7** (**Figure 4.7**), which retains Val-Gly-Ala from the LP. In testing **7** as a TbtI substrate, we observed full conversion to the methyl-thiazole **9** (**Figure 4.7**). Control experiments omitting TbtI, SAM, or dithionite did not result in observable modification of **7**. The

site of methylation in **9** was confirmed as Thz4 by high-resolution electrospray ionization tandem MS analysis (HR-ESI-MS/MS; **Figure 4.8**). These data indicate that most of the LP (31 out of 34 residues) is dispensable for TbtI activity and that the enzyme recapitulates in vitro the specificity observed in the natural product.

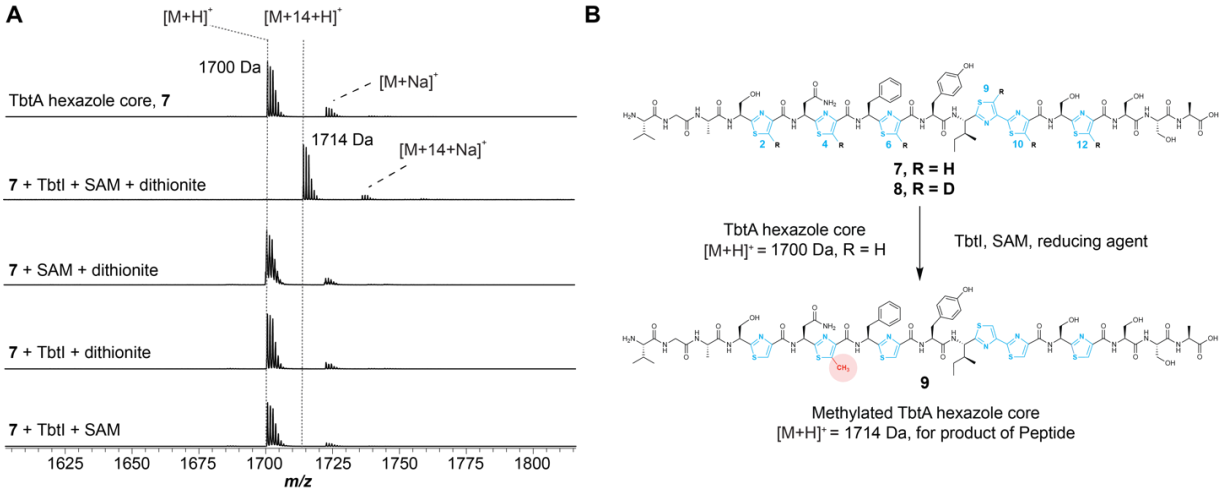


Figure 4.7 (**A**) MALDI-TOF mass spectra of **7** treated as indicated. (**B**) Reaction catalyzed by TbtI to convert **7** to **9** (for the mechanistic study of TbtI discussed below, the ^{15}N-, ^{13}C-depleted **7** is annotated as **7'**).

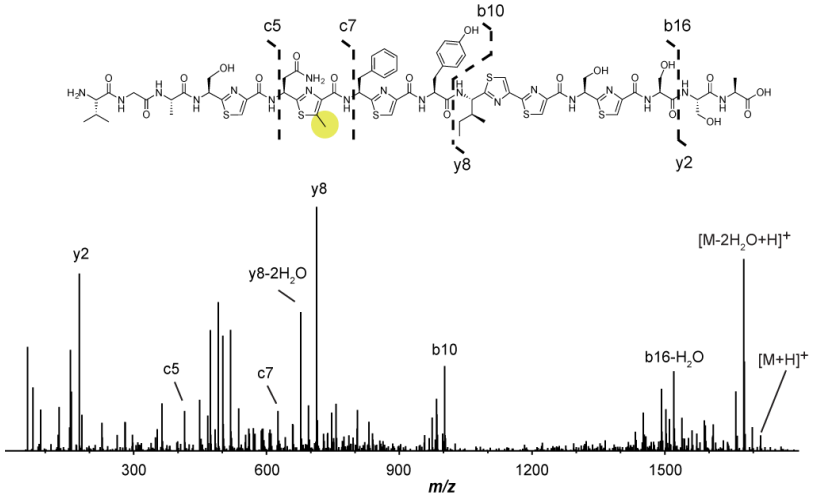


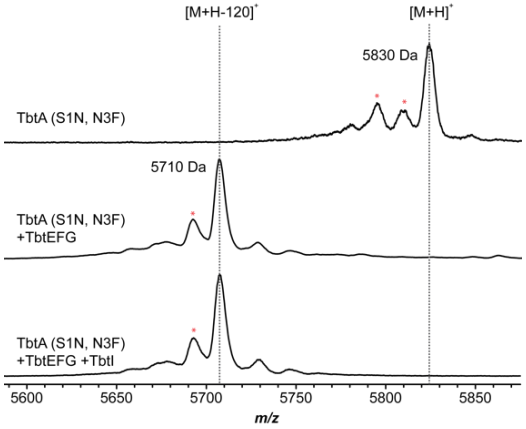
Figure 4.8 MS/MS of the TbtI reaction product **9**. The c5, c7, and y8 ions localize the site of C-methylation to the Thz4.

We next addressed the determinants of regioselectivity in the TbtI reaction, given that Thz4 undergoes selective *C*-methylation despite the presence of five additional thiazoles. We hypothesized that TbtI recognizes a substructure of **2** adjacent to Thz4. Inspection of **5** and **6** shows that the methylthiazole is preceded in both cases by an Asn3, which in **6** undergoes additional *N*-methylation (37). It has been reported that Asn3 forms a key contact with elongation factor thermos unstable (EF-Tu) and that substitution of Asn3 results in loss of antibiotic activity (46, 47). To investigate whether Asn3 governs the regioselectivity of TbtI, I replaced this residue in TbtA with Ala, Asp, and Gln, followed by hexazole production in *E. coli*. Dr. Nilkamal Mahanta found that no methylation was observed with any Asn3 variants, demonstrating that Asn3 in TbtA is critical for TbtI activity.

We hypothesized that introducing an Asn N-terminal to one of the other thiazoles might redirect the location of methylation. Thus, four TbtA double-substituted variants of **7** were constructed by inserting Asn at a position preceding Thz 2, 6, 9 and 12 and replacing Cys4 with Ala. These four GluC-treated pentazole variants (S1N/C4A, C4A/F5N, C4A/I8N, and C4A/S11N) were then tested as TbtI substrates. Still no methylation was observed, indicating that other factors contribute to the regioselectivity. Further experiments performed by Dr. Nilkamal Mahanta and Graham Hudson have shown that TbtA with a single thiazole alone at Cys4 was not a substrate, and that only the 4/6 diazole was able to be fully converted to the methylated product. These data suggested that -Asn-Thz-Phe-Thz- may serve as the minimal substrate recognition motif for TbtI.

To test this hypothesis, the motif was installed at positions 1-4 by Graham Hudson. This double-substituted variant, TbtA-S1N/N3F, was reacted with TbtE/F/G and subsequently TbtI. Thiazole formation was observed, but methylation was not (**Figure 4.9**), suggesting that while an N-terminal flanking Asn and a downstream thiazole are necessary, they are not sufficient for TbtI

methylation. Other, currently unknown structural aspects may further contribute to the substrate selectivity of TbtI. All results are summarized in **Table 4.1**.



TbtA S1N/N3F: SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*NCFCFCYICCSCSSA

Figure 4.9 TbtI reaction with TbtA S1N/N3F double variant. TbtA hexazole featuring the substitutions Ser1Asn and Asn3Phe was not a substrate for TbtI as visualized by MALDI-TOF-MS. Sites of thiazole installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk. *m/z* assignments: 5830 Da, unreacted TbtA (S1N, N3F) precursor peptide; 5710 Da, TbtA (S1N, N3F) hexazole.

Table 4.1 TbtA precursor peptide variants tested for methylation by TbtI. TbtA sequence: Leader-SCNCFCYICCSCSSA.

Variant	Outcome
wild-type hexazole	**Fully processed**
wild-type hexazole core (three res. leader)	**Fully processed**
C4A pentazole	Not processed
C4S pentazole	Not processed
N3A hexazole	Not processed
N3D hexazole	Not processed
N3Q hexazole	Not processed
S1N/C4A pentazole	Not processed
C4A/F5N pentazole	Not processed

Variant	Outcome
C4A/I8N pentazole	Not processed
C4A/S11N pentazole	Not processed
Cys4 monoazole	Not processed
Cys2/4 diazole	Not processed
Cys4/6 diazole	**Fully processed**
Cys4/9 diazole	**Partially processed**
Cys4/10 diazole	Not processed
Cys4/12 diazole	Not processed
S1N/N3F hexazole	Not processed

4.2.2 Mechanistic study of the class C radical SAM thiazole *C*-methyltransferase TbtI

Sequence analysis of TbtI did not reveal significant homology to class A radical SAM MTs outside of the conserved [4Fe-4S] and SAM-binding motifs. TbtI is clearly distinguished from the cobalamin-dependent class B MTs (9). Class C MTs are homologous to HemN/HemZ, mechanistically unrelated radical SAM decarboxylases from heme biosynthesis (5).

To start our mechanistic study of TbtI, we first assessed whether TbtI might employ a similar mechanism to class A radical SAM MT, even though this was inconsistent with sequence analysis. This class of enzymes utilizes an unusual mechanism in which a methyl group from SAM is first transferred to a Cys on the protein. Then, reductive cleavage of a second SAM molecule occurs, mediated by a reduced iron-sulfur [4Fe-4S] cluster, which generates a 5'-deoxyadenosyl radical (5'-dA•). The 5'-dA• species then abstracts a hydrogen atom from the methylcysteine to initiate the methyl transfer process (**Figure 4.10**). The sequence of TbtI and its orthologs from various thiopeptide gene clusters was first analyzed by Dr. Nilkamal Mahanta (48) and he identified five conserved Cys residues (Cys14, Cys18, Cys21, Cys254 and Cys343; **Figure 4.11**). Three of these appear in a canonical motif (CxxxCxxC), which typically supply the ligands for a characteristic radical SAM [4Fe-4S] cluster (49). TbtI variants in which these Cys residues were replaced by Ala (C14A, C18A and C21A) were inactive (**Figure 4.12**), but substitution of the remaining Cys residues (C89A, C254A, and C343A) did not adversely affect the activity of TbtI (**Figure 4.12**), suggesting that they are functionally dispensable.

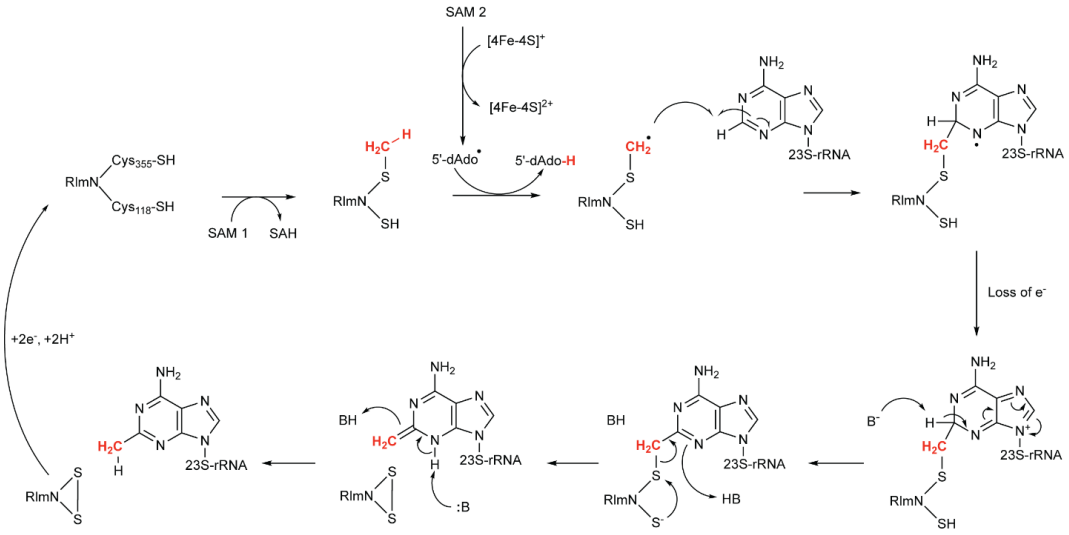


Figure 4.10 Mechanism of class A radical SAM MT RlmN (50). RlmN methylates an adenine in 23S-tRNA by first methylating a Cys in the enzyme, and then uses 5'-dA• derived from a second SAM to abstract a hydrogen atom from the methylated Cys. The resulting radical is added to the adenine, and oxidized to a cation. Deprotonation generates a neutral intermediate that fragments via disulfide bond formation. A tautomerization then affords the methylated adenine.

```
TbtI       ---------------------------------------------------MTRPLLLYVNIP
TpdI       ---------------------------------------------------MTKPLLIYANVP
PbtM2      ------------------------------------MSTTLRGTGGAVAEQPLLIYVNVP
PbtM3      ----------------------------MTT-EAITNALPLTGPKTDEPLLLYVNVP
HemN       MSVQQIDWDLALIQKYNYSGPRYTSYPTALEFSEDFGEQAFLQAVARYPERPLSLYVHIP
                                                                .** :*.::*

TbtI       FCNSKCHFCDWVVQIPVRDLRLDQQAPGRIAYLDAVRAQIRGQAPVLRE-HYQPAIIYWG
TpdI       FCNSKCHFCDWVVQVPVRDLRLGEQSPGRVAYLEAIRTQIRVQAPALRE-HYHPDIVYWG
PbtM2      FCNSKCHFCDWVTEVPLADLRLTPDSSPRQRYIAALVQQIETHAPVLTGFGYRPEIMYWG
PbtM3      FCNSKCHFCDWVVDVPVSDLRLTPVSAGRIDYLEALRTQIRIHAPALREAGYRSEVMYWG
HemN       FCHKLCYFCGCNKIVT-------RQQHKADQYLDALEQEIVHRAPLFAG--RHVSQLHWG
           **:. *:**.    :                *: *:  :*  :** :     :   ::**

TbtI       GGTASILGESEIESLYTCLRSEFDLSHVRETTIEGSPESLTPQKLRRLRELGFDRISIGV
TpdI       GGTASILGPHEIESLYGTLSAEFDLSTVREATIEGSPESLDRDKLRLLRELGFNRISIGV
PbtM2      GGTASILSIDEIEAVAGALSSRFDMSGLTEATIEGSPESLDPDKLKLFRAAGFNRISIGV
PbtM3      GGTATVLTEREIEQTYACLAAELDLSSLAEATIEGSPESVDLAKLTFMRDLGFDRVSLGV
HemN       GGTPTYLNKAQISRLMKMLRENFQFNADAEISIEVDPREIELDVLDHLRAEGFNRLSMGV
           *** : *   :*.     *  .::::.   * :** .*..:    *  :*  **:*:*:**

TbtI       QSFDDQRLRRLGRAHSAEQAVEVVKNAHAAGFRNINIDLIVGFPGQTDAEVAESVRTALT
TpdI       QSFDDARLRRLGRAHAADQALTAVQDAHEAGFTNINIDLIVGFPGQTGQEVADSVRTALT
PbtM2      QAFDDARLRRIGRVHSAEQAVRAVEMAAEAGFDNINIDLIVGFPGQEVDEVSHMIQRAVT
PbtM3      QSFDETRLRRIGRAHSAGQAVQSVEAAHAAGFDNINIDLIVGFPDQSLEEVEEMIRRALD
HemN       QDFNKEVQRLVNREQDEEFIFALLNHAREIGFTSTNIDLIYGLPKQTPESFAFTLKRVAE
           * *:.   * :.* :     .  :: *   ** . ***** *:* *   ..   :: .
```

Figure 4.11 (cont.)

Figure 4.11 Sequence alignment of select class C radical SAM MTs. Proteins shown include TbtI from the hypothetical thiomuracin producer *Thermobispora bispora* (WP_013130805.1), TpdI from known thiomuracin producer *Nonomuraea* sp. Bp3714-39 (ACS83777.1), PbtM2/M3 from GE2270A producer *Nonomuraea* sp. WU8817 (AGY49586.1/AGY49595.1), and HemN from *Escherichia coli* (WP_001502804.1). Cys residues involved in [4Fe-4S] cluster formation are highlighted yellow. TbtI residues chosen for Ala substitution are shown in red. HemN residues within 5 Å of SAM2 are shown in blue.

Thus, these data show that unlike class A radical SAM MTs, which require at least two additional Cys other than the three used to bind the [4Fe-4S] cluster (28), TbtI does not require additional Cys residues for catalysis. This conclusion is further supported by the fact that a TbtI homolog involved in nocathiacin biosynthesis (NocN) contains only one additional Cys other than the three involved in [4Fe-4S] cluster coordination (8). We considered that a methyl group from SAM could be transferred to an amino acid residue other than Cys.

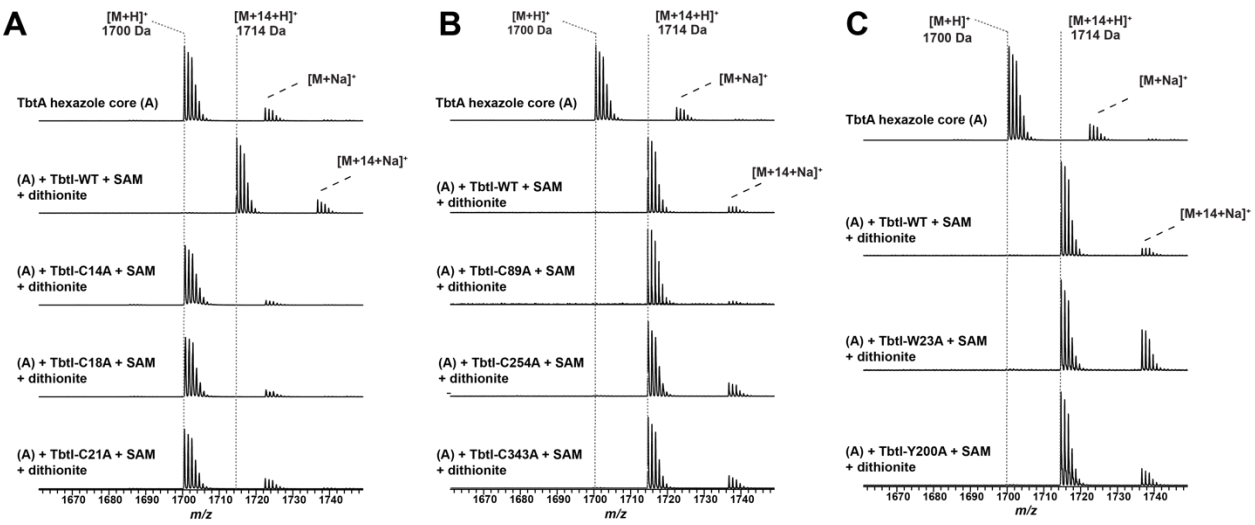
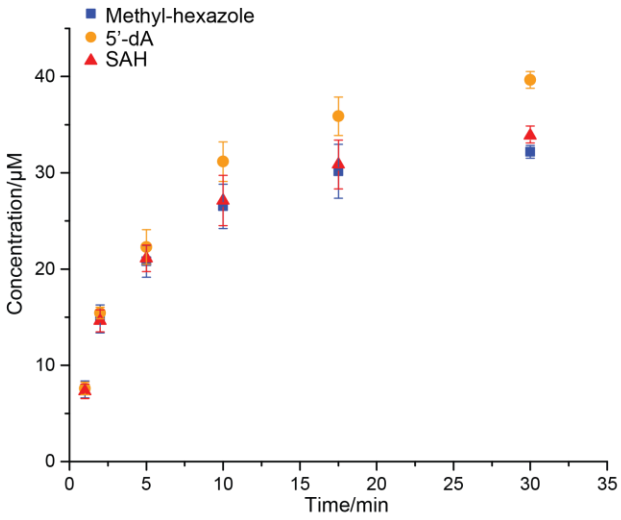


Figure 4.12 MALDI-TOF MS analysis of the products generated by TbtI variants. (**A**) Spectra of TbtA hexazole treated with wild type (wt) TbtI or TbtI variants in which the Cys residues expected to be involved in binding to the [4Fe-4S] cluster were substituted with Ala. (**B**) Spectra of TbtA hexazole treated with wt TbtI or TbtI variants in which the remaining Cys residues were substituted with Ala. (**C**) Spectra of TbtA hexazole treated with wt TbtI or variants in which conserved Tyr/Trp residues were individually substituted with Ala without loss of activity. Very similar results were obtained for the Cys variants when DTT was removed from the assays (Experiment performed by Dr. Nilkamal Mahanta).

To test this possibility, I incubated the enzyme with $^{14}CH_3$-SAM and the peptide substrate without addition of reducing agents. The enzyme was isolated and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie staining or autoradiographic imaging. We did not observe a radioactive band for the protein, suggesting that the methyl group is not initially transferred to the enzyme. This finding could also be explained by a methyl group already being present on the enzyme upon isolation from *E. coli*; however, single turnover experiments presented below suggest against this possibility. We cannot completely rule out that transfer of the methyl group from SAM to an amino acid residue only occurs in the presence of all reaction components including reductants. TbtI lacks the cobalamin-binding motif

characteristic of class B radical SAM MTs (49, 51-58), as well as the tetrahydrofolate-binding domain observed in a recently discovered class D radical SAM MT (59). All these factors point towards a distinct mechanism for TbtI.



	1 min	2 min	5 min	10 min	17.5 min	30 min
5'-dA	7.6 ± 0.7	15,4 ± 0.6	22.3 ± 1.8	31.2 ± 2.1	35.9 ± 2.0	39.7 ± 0.9
SAH	7.3 ± 0.8	14.6 ± 1.1	21.1 ± 1.4	26.9± 2.6	30.3 ± 2.5	33.9 ± 1.0
Methyl-hexazole	7.5 ± 0.9	14.8 ± 1.5	20.8 ± 1.7	26.3 ± 2.3	29.6 ± 2.8	32.0 ± 0.7

Figure 4.13 Time-dependent formation of the methylated TbtA peptide, 5'-dA, and SAH. Reaction mixtures included the following components: TbtA hexazole (50 μM), [4Fe-4S] cluster reconstituted TbtI (10 μM), SAM (1 mM), flavodoxin (10 μM), flavodoxin reductase (10 μM), NAPDH (2 mM) in reaction buffer (50 mM Tris-HCl, pH 7.5).

To gain insight into the mechanism of methyl transfer, I first determined the products of the reaction that are derived from SAM. Out of a concern that the strong reductant dithionite might react with the initial products of the enzymatic reaction, we replaced dithionite in our original assay (48) with flavodoxin/flavodoxin reductase and NADPH to deliver the reducing equivalents to TbtI (see Experimental Section). As anticipated based on other radical SAM proteins (60), 5'-deoxyadenosine (5'-dA) was detected in slightly more than one equivalent compared to the methylated TbtA product (**Figure 4.13**). Additionally, *S*-adenosyl-L-homocysteine (SAH) was

detected in near stoichiometric quantities, and 5'-dA and SAH were formed at very similar rates (**Figure 4.13**). When the reaction was conducted with CD_3-SAM (96% CD_3 by ESI-MS; **Figure 4.14**), ~90% of the 5'-dA contained a single deuterium (**Figure 4.14B**), suggesting that the 5'-dA• generated from SAM1 may abstract a hydrogen atom from another SAM molecule (SAM2) during catalysis. It also suggests that the small amount of unlabeled 5'-dA produced was the result of uncoupled SAM cleavage, as has been observed in other radical SAM enzymes (60).

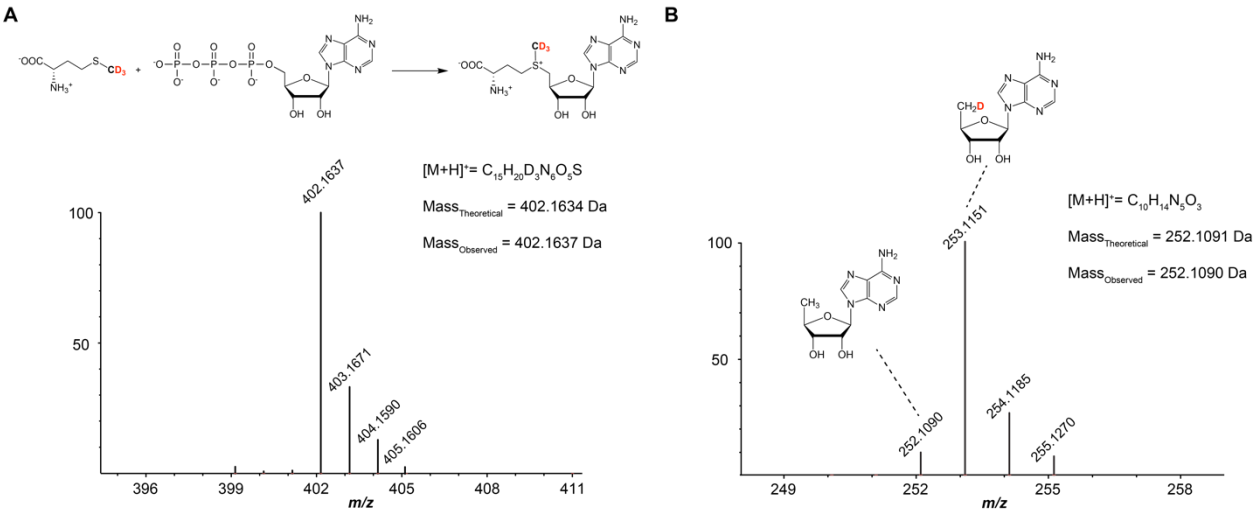


Figure 4.14 (**A**) ESI-MS of the CD_3-SAM starting material. Error, 0.7 ppm. (**B**) ESI-MS analysis of 5'-dA produced during the methylation of TbtA by TbtI in the presence of CD_3-SAM. Error, 0.4 ppm. The deuterium content of CD_3-SAM was calculated as follows: (Peak intensity$_{m/z\ =\ 402}$)/ [(Peak intensity$_{m/z\ =\ 402}$) + (Peak intensity$_{m/z\ =\ 399}$) + (Peak intensity$_{m/z\ =\ 400}$) + (Peak intensity$_{m/z\ =\ 401}$)].

Control experiments showed that production of SAH and 5'-dA was dependent on the presence of both TbtI and the reducing system (**Figure 4.15**). It was recently reported that another class C radical SAM MT involved in nosiheptide biosynthesis (NosN) produces thioadenosine in its reaction (61). We therefore investigated if TbtI generated thioadenosine; however, I was unable to observe its production by LC-MS whereas an authentic standard was readily detected at the

concentrations expected to be produced (**Figure 4.15**). Another recent study on NosN also demonstrated that thioadenosine is not a product of the reaction (62). Based on the products we observed, we formulated a series of working hypotheses for the enzymatic mechanism of TbtI (**Figure 4.16**).

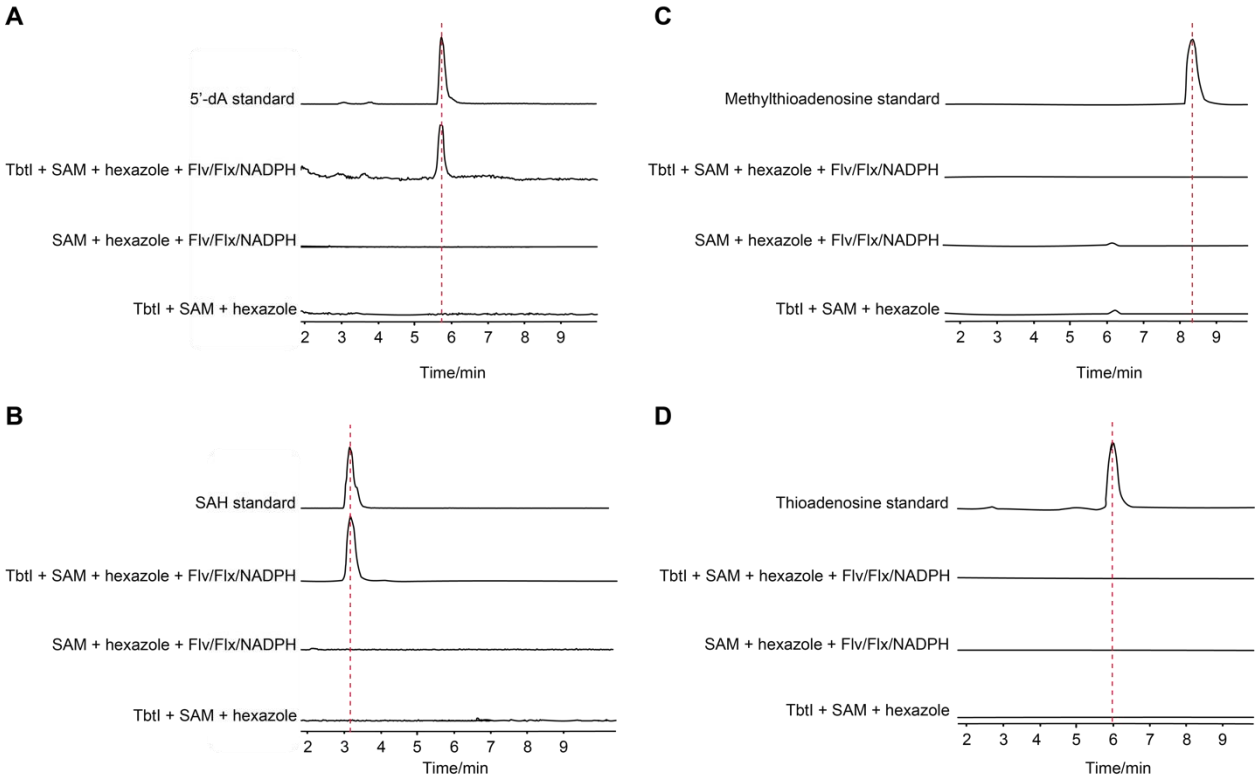


Figure 4.15 Products of the TbtI reaction. (**A**-**B**) Extracted ion chromatograms to analyze production of 5'-dA and SAH showed their generation was dependent on the presence of both TbtI and the reducing system. (**C**-**D**) Extracted ion chromatograms to analyze potential production of methylthioadenosine and thioadenosine did not show the presence of these products under the reaction conditions.

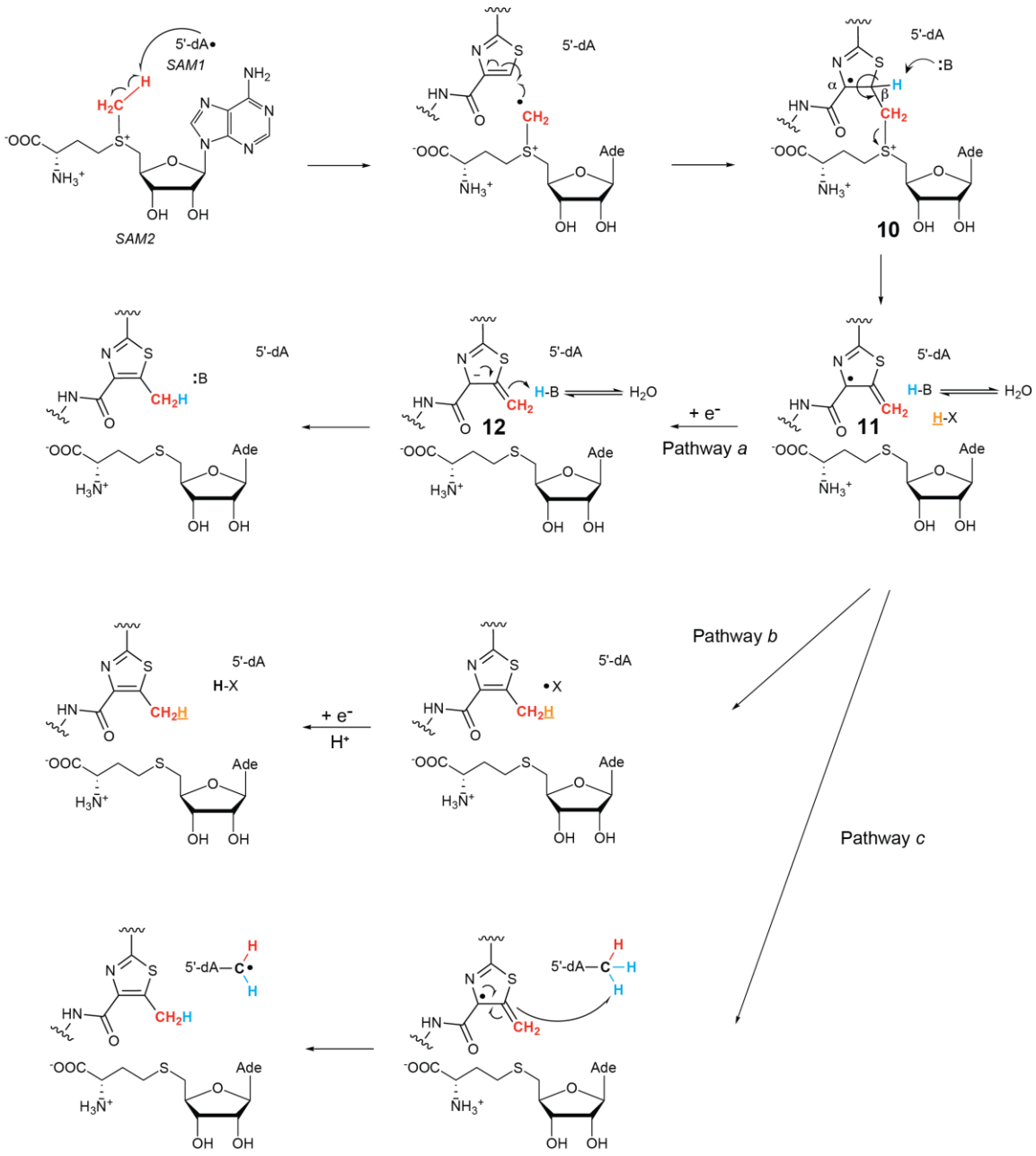


Figure 4.16 Proposed potential reaction mechanisms for TbtI. 5'-dA• is generated from SAM1 in a process mediated by the reduced [4Fe-4S] cluster. The 5'-dA• then abstracts a hydrogen atom from the methyl group of SAM2, and the resulting radical adds to the thiazole. An active site base deprotonates radical **11** leading to the elimination of SAH, yielding radical **11**. Three different pathways (*a*-*c*) can provide the product and reset the enzyme. In pathway *a*, the radical is reduced to the anion **12** and protonated. In pathway *b*, radical **11** abstracts a hydrogen atom from an active site amino acid (X-H). X-H could be the protonated base B-H. In pathway *c*, radical **11** abstracts a hydrogen atom from the methyl group of 5'-dA. Ade = adenine.

To evaluate the proposed mechanisms, we first investigated whether thiazole methylation involves transfer of the intact methyl group of SAM or a methylene group. Therefore, reactions with TbtI were carried out using CD_3-SAM. Owing to the large molecular weight of the TbtA hexazole core substrate even after removal of most of the leader peptide (three residues remain, Val-Gly-Ala) with endo-proteinase GluC ($C_{71}H_{85}N_{19}O_{19}S_6$, $[M+H]^+$ = 1700 Da), the numerous isotopologues visible by mass spectrometry rendered isotope labeling studies difficult to interpret. Therefore, we prepared the TbtA hexazole substrate from *E. coli* grown in minimal media using ^{13}C-depleted glucose and ^{15}N-depleted $(NH_4)_2SO_4$ as the sole carbon and nitrogen sources, and again removed the majority of the leader peptide by treatment with endo-proteinase GluC. The resulting TbtA hexazole core peptide **7'** (**Figure 4.7B**) was analyzed by electrospray ionization mass spectrometry (ESI-MS) and exhibited a much-simplified spectrum (**Figure 4.17**).

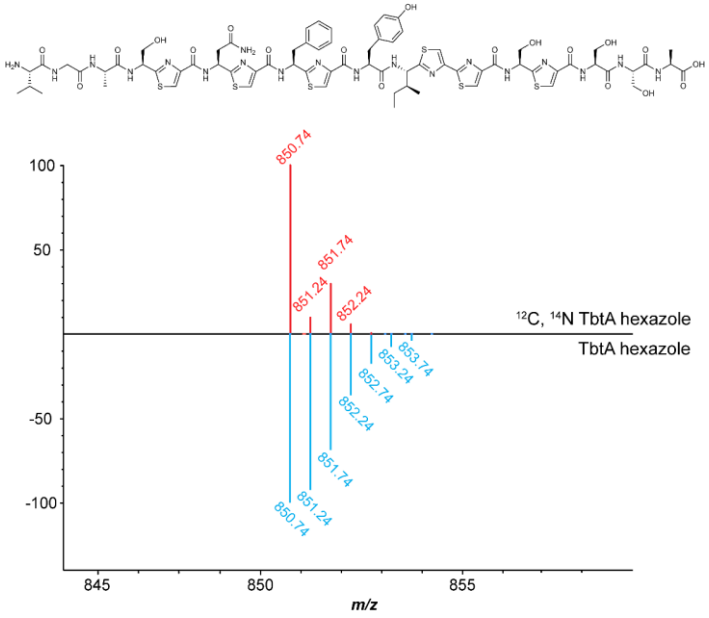


Figure 4.17 ESI-MS analysis of the doubly charged endo-proteinase GluC digested (^{13}C, ^{15}N)-depleted TbtA hexazole **7'** (top, red) and natural abundance TbtA hexazole (bottom, blue).

Use of CD_3-SAM in a 16 h methylation assay with TbtI and peptide **7'** resulted in an unexpectedly complicated mass spectrum for the product in which ions were observed consistent with the addition of CH_3, CH_2D, and CHD_2 groups (**Figure 4.18A**). These findings indicated that there was either an intermediate that underwent considerable solvent exchange or the product exchanged with solvent after initial formation (or both). To test the latter possibility, we repeated the experiment with CD_3-SAM and quenched the reaction after only 1 h. The product was analyzed by ESI-MS and an ion was observed that was consistent with the transfer of a CD_2H group (**Figure 4.18A**, blue spectrum). This product was purified by high performance liquid chromatography (HPLC), exposed to the assay conditions for the extended reaction time (16 h), re-isolated, and then again analyzed by ESI-MS. In agreement with the original findings, a complicated mixture of product peptide was observed that corresponded to apparent transfer of CH_3, CH_2D, and CHD_2 (**Figure 4.19**). This exchange was dependent on the presence of TbtI and SAM; replacing SAM with SAH did not result in exchange. Thus, these experiments show that the initial product of the reaction of TbtI with CD_3-SAM and **7'** mainly contains a CD_2H group, but that the enzyme can exchange the deuterium labels with protium from solvent. Unfortunately, the large amounts of 5'-dA produced from the abortive cleavage of SAM during the 16 h reactions precluded our ability to determine if deuterium from the methyl group is transferred to the 5'-dA during the exchange process.

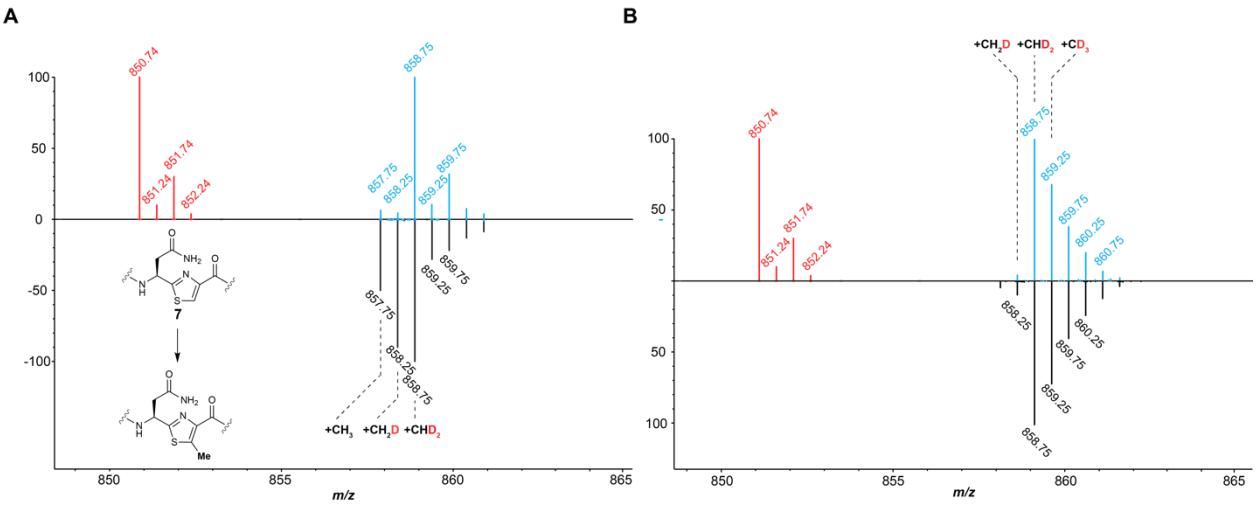


Figure 4.18 ESI mass spectra to investigate the origin of the methyl protons. ^{13}C- and ^{15}N-depleted substrate peptides were obtained as described (see Experimental Section). (**A**) Spectra showing the doubly charged ion for the hexazole-containing core peptide substrate **7'** (red) and the methylated product obtained with CD_3-SAM in H_2O after 1 h (blue) or 16 h (black) reactions. (**B**) Spectra showing substrate (red) and product obtained with CD_3-SAM in D_2O (blue) and CD_3-5',5',4',3'-D_4-SAM in D_2O (black) after 1 h.

Because of the solvent exchange observed at extended reaction times, all subsequent assays were performed for 1 h. We first investigated whether the generation of the product containing a CHD_2 group when using CD_3-SAM in H_2O involves the incorporation of a proton from solvent. Thus, the reaction was performed using CD_3-SAM in D_2O. Analysis of the product by ESI-MS (**Figure 4.18B**) showed that the major product still contained CHD_2, although the relative amount of CD_3 transfer increased. This observation suggests that a hydrogen atom in an initially non-exchangeable position migrates to the methyl group, possibly with the intermediacy of a partially solvent exchangeable intermediate. The mechanisms in **Figure 4.16** provide three possible sources for this hydrogen. First, the thiazole proton on the β-carbon of **7'** that must be removed during the methylation process could migrate via an active site base to the methyl group of the product (pathway *a*). Alternatively, a hydrogen that originates from the protein that does not exchange

could be transferred to the methyl group (pathway *b*). Based on literature precedent, an example of the latter could be the α-protons of Gly. Lastly, the hydrogen could originate from the methyl group of 5'-dA (pathway *c*).

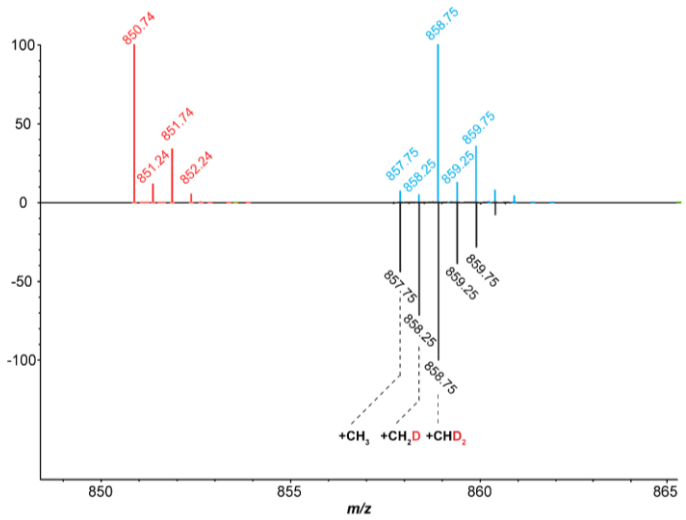


Figure 4.19 ESI-MS spectra to investigate methyl group solvent exchange. Spectra showing the doubly charged core peptide of the TbtA hexazole substrate **7'** (red) and product obtained with CD_3-SAM in H_2O after a 1 h reaction (blue) or the product from this reaction re-exposed to the same reaction for 16 h (black).

We first considered 5'-dA as the source for the additional hydrogen in the methyl group of the product. Abstraction of a hydrogen atom from 5'-dA by radical **11** (pathway *c*, **Figure 4.16**) would explain transfer of CHD_2 when employing CD_3-SAM. We therefore prepared SAM from Met-CD_3 and 5',5',4',3'-D_4-ATP (~ 92% CD_3-5',5',4',3'-D_4-SAM by ESI-MS, **Figure 4.20**). The resulting CD_3-5',5',4',3'-D_4-SAM was then reacted with TbtI and the TbtA hexazole **7'** in D_2O. The reaction products were then analyzed by ESI-TOF MS. The amount of product that contained CHD_2 was similar to that observed when the 5' position of SAM was unlabeled (**Figure 4.18B**).

Hence, the hydrogen in this product does not originate from the 5’ position of the adenosine, ruling out pathway *c*.

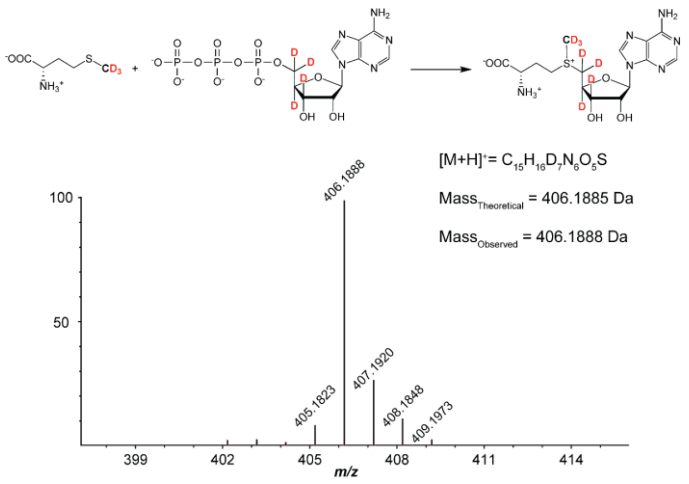


Figure 4.20 Synthesis and ESI-MS analysis of CD_3-5',5',4',3'-D_4-SAM. Error, 0.7 ppm. The deuterium content of CD_3-5’,5’,4’,3’-D_4-SAM was calculated as follows taking into account that some of the ion at 406 might arise from a ^{13}C-contaning isotopologue of the ion at m/z 405: [(Peak intensity$_{m/z\,=\,406}$) − (0.3 × Peak intensity$_{m/z\,=\,405}$)/ [(Peak intensity$_{m/z\,=\,406}$) + (Peak intensity$_{m/z\,=\,405}$) + (Peak intensity$_{m/z\,=\,404}$) + (Peak intensity$_{m/z\,=\,403}$) + (Peak intensity$_{m/z\,=\,402}$)].

We next considered pathway *b* to account for the product hydrogen, in which a non-exchangeable hydrogen on the protein is transferred to the product methyl group in the first turnover; this hydrogen would be replaced by a solvent-derived hydrogen in subsequent turnovers. This behavior has previously been observed, such as the transfer of Cα hydrogens from a Gly residue (63, 64). In principle, hydrogen transfer could also originate from other amino acids or very tightly hydrogen bonded positions that would only exchange upon reaction. Because TbtI contains several conserved Gly residues (**Figure 4.11**), we reacted a 20-fold excess of substrate **7’** and CD_3-SAM with TbtI in buffered D_2O and analyzed the product by ESI-MS (**Figure 4.21**). The

major product again contained a CHD_2 group, arguing against a protein-derived hydrogen that is transferred to the methyl group.

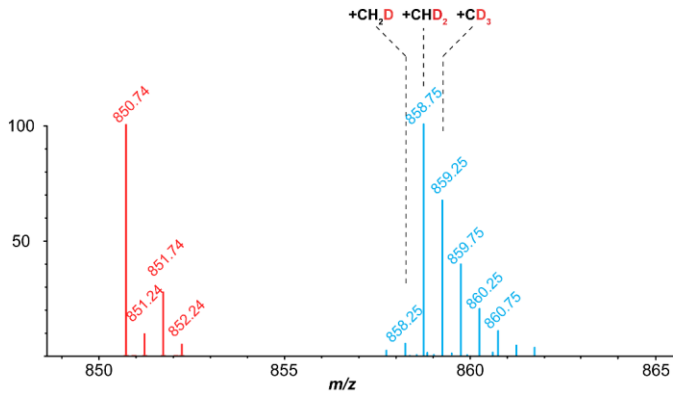


Figure 4.21 ESI-MS spectra ruling out transfer of a non-exchangeable hydrogen derived from the protein. Spectra showing the doubly charged core peptide of the TbtA hexazole substrate **7'** (red) and product obtained with a 20-fold excess of substrate and CD_3-SAM with TbtI in buffered D_2O.

Having ruled out pathways *b* and *c*, we next replaced the Cβ proton of the thiazole with deuterium to interrogate pathway *a* (**Figure 4.16**). We prepared TbtA hexazole using an *E. coli* strain that is auxotrophic for Cys and provided β-D_2-Cys to the $^{13}C/^{15}N$-depleted minimal media. Unfortunately, the expression level of the peptide under these conditions was very low. During optimization of the expression conditions, we noted that improved production could be achieved upon addition of a limited number of unlabeled amino acids to the culture media (see Experimental Section for details). Analysis by tandem ESI-MS confirmed the desired deuterium labeling of the six Cys residues in the TbtA precursor (**Figure 4.22**). The resulting peptide (~ 95% D_2 at Cys4 by ESI-MS-MS) was then converted to the corresponding hexazole-containing core peptide **8** *in vitro* as previously reported (65). Owing to the need to add additional amino acids (which were not ^{13}C- and ^{15}N-depleted), the monoisotopic purity of the hexazole **8** was lower than that of the substrate

7'. Nevertheless, an ion consistent with the hexa-deuterated TbtA hexazole core peptide as the major product was observed by ESI-MS.

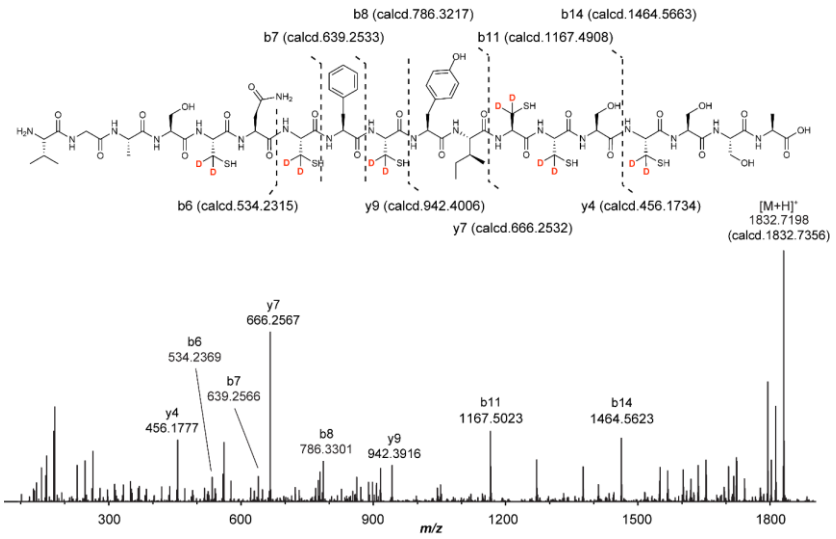


Figure 4.22 Tandem ESI-MS on TbtA core peptide containing six D_2-Cys residues. The tandem MS analysis was performed on the precursor peptide pre-treated with endo-proteinase GluC to remove most of the leader peptide (65). Analysis of the tandem MS data reveals that six D_2-Cys were incorporated into the peptide. The deuterium content at Cys4 was estimated by determining the deuterium content of the b7 ion in comparison to the deuterium content of the b6 ion: [peak intensity of the b_7 ion (m/z = 639)/ (peak intensity of the b_7 ion (m/z = 639) + the peak intensity of b7 − 1 ion (m/z = 638))] – [peak intensity of the b_6 ion (m/z = 534)/ (peak intensity of the b_6 ion (m/z = 534) + the peak intensity of b6 − 1 ion (m/z = 533))].

With peptide **8** in hand, we conducted a series of single turnover experiments. If the proton in the transferred CHD_2 was protein-derived in Figure 3b, then a single turnover experiment in D_2O with substrate **8** and CD_3-SAM should still result in CHD_2 transfer. However, when the experiment was performed with a three-fold excess of TbtI over substrate **8**, CD_3 transfer was observed as the major product (**Figure 4.23A**). This result provided additional data against pathway *b*: a non-exchangeable hydrogen from the enzyme was not the source of hydrogen in the

methyl group of the product. We tentatively propose that the small amount of CHD_2 transfer observed may originate in part because of incomplete deuteration of the reagents. We then repeated the single turnover experiment with CD_3-SAM in D_2O but employed substrate **7'**. A considerable amount of CD_3 transfer was observed, indicating that solvent exchange occurred in the single turnover experiments with excess enzyme (**Figure 4.23B**), but importantly, the relative amount of CD_3 transfer declined considerably compared to experiments conducted with substrate **8**. Collectively, these data provide strong support for pathway *a* being operative during TbtI catalysis. These experiments also argue against TbtI already containing a covalently bound methyl group introduced during expression in *E. coli*; this methyl group would have been unlabeled and therefore, we would have observed a considerable amount of product containing CH_3 and CH_2D.

To identify residues that may mediate the transfer process, Dr. Nilkamal Mahanta substituted conserved residues Tyr8 and Glu105 of TbtI with Ala (**Figure 4.11**). The corresponding residues are near the SAM-binding site in HemN (66, 67), a radical SAM decarboxylase from the heme biosynthetic pathway that is homologous to TbtI. Replacement of either residue resulted in TbtI variants with severely decreased activity (**Figure 4.24**). Two additional conserved residues (Trp23 and Tyr200) were targeted for Ala substitution given that they could participate in electron or hydrogen atom transfer; however, the activity of the resulting TbtI variants was not decreased, suggesting that they are not involved in catalysis (**Figure 4.12C**).

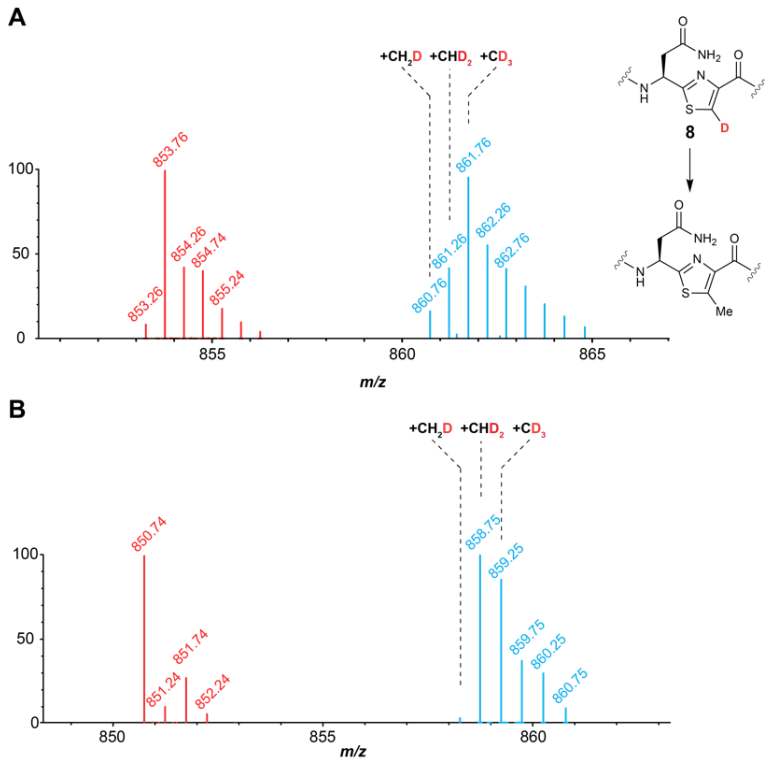


Figure 4.23 A solvent exchangeable proton migrates from the β-position of thiazole 4 to the methyl group in the product. (**A**) Spectra showing the doubly charged core peptide of the TbtA hexazole substrate **8** that is deuterated at the β-carbon of each thiazole (red) and the corresponding product obtained with CD_3-SAM in D_2O (blue) in a single turnover reaction. (**B**) Spectra showing the doubly charged core peptide of the TbtA hexazole substrate **7'** (red) and the corresponding product obtained with CD_3-SAM in D_2O (blue) in a TbtI-catalyzed single turnover reaction.

A combination of product analysis and isotope labeling studies has allowed us to formulate a mechanism for TbtI that is consistent with all of our available data. The observation of 5'-dA as a product suggests that the TbtI-catalyzed reaction is initiated in a manner similar to canonical radical SAM proteins: electron transfer from the reduced [4Fe-4S] cluster initiates the reductive cleavage of SAM1 to form methionine and 5'-dA• (**Figure 4.16**) (60). We propose that the 5'-dA• then abstracts a hydrogen atom from the methyl group of SAM2. This hypothesis is supported by the observation that use of CD_3-SAM resulted in the incorporation of one deuterium into 5'-dA.

This reaction appears irreversible since we never observed multiple deuteriums in 5'-dA (**Figure 4.14**). Use of two molecules of SAM is also consistent with *(i)* the observed stoichiometry of 5'-dA and SAH to the methylated product, *(ii)* the previously reported crystal structure of HemN with two bound SAM molecules (66, 67), and (*iii*) recent results reported for the class C radical SAM MT NosN (62). Although HemN and TbtI catalyze different reactions (decarboxylation and methylation, respectively), the use two molecules of SAM might be a general property of class C radical SAM enzymes.

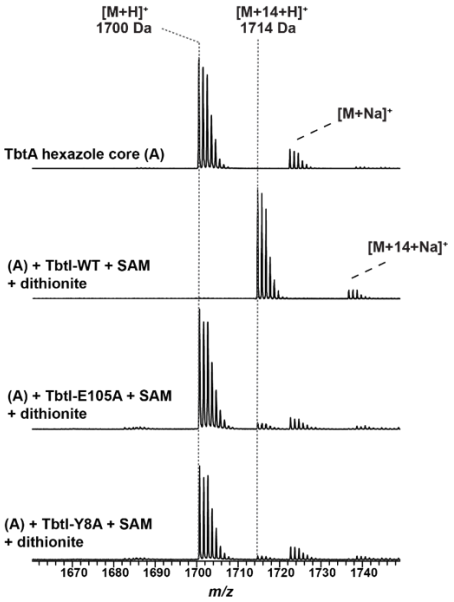


Figure 4.24 MALDI-TOF MS analysis of the products generated by TbtI variants. Spectra of TbtA treated with wt TbtI or TbtI variants in which the class C rSAM-conserved Glu and Tyr residues were substituted with Ala. According to the crystal structure of HemN, these residues are near the SAM-binding site (Experiment performed by Dr. Nilkamal Mahanta).

Upon hydrogen atom abstraction from the methyl group of SAM2, we propose that the resulting methylene radical adds to the electrophilic C=C bond of the thiazole, generating resonance-stabilized radical **10** (**Figure 4.16**). This may trigger deprotonation from the β-position

by an active site base **B** to eliminate SAH and form resonance-stabilized radical **11** that partially restores thiazole aromaticity (**Figure 4.16**). This proton is adjacent to an enoxy-radical and previous studies on 2-hydroxyglutaryl-CoA dehydratase (68) have suggested that the pK_a of such protons are on the order of ~14. One candidate residue for the active site base is Glu105, which upon substitution with Ala severely decreases TbtI activity (**Figure 4.24**). The equivalent Glu in HemN (Glu145) is near the SAM-binding site (66, 67); thus, an alternative role for Glu105 could be to bind SAM. Reduction of radical **11**, possibly by the [4Fe-4S] cluster, would produce resonance-stabilized anion **12**, which can then be protonated by the conjugate acid of the active site base, possibly Glu105. This pathway would explain the observed deuterium transfer from the β position of the thiazole to the methyl group in the product (**Figure 4.13**).

A variation of this mechanism features abstraction by radical **11** of a hydrogen atom from a redox active amino acid in the active site (e.g. Tyr/Trp/Gly; Cys can be ruled out since their replacement with Ala had no effect on activity, **Figure 4.12**). The amino acid radical would then be reduced and protonated to reset the enzyme for another turnover (**Figure 4.16**, pathway *b*). The single turnover experiments argue against the possibility that this hydrogen is in a non-exchangeable position in the first turnover but becomes a solvent-derived hydrogen in subsequent turnovers. Hence, the involvement of Gly radicals is unlikely, but other redox active amino acids cannot yet be ruled out. For instance, a tyrosinate could serve as the base that deprotonates radical **10** and the resulting phenol moiety could then provide a hydrogen atom to radical **11**. Tyr8 is conserved in the class C radical SAM proteins (**Figure 4.11**). In HemN, the equivalent Tyr56 π-stacks with the adenine of SAM2. Substitution of TbtI Tyr8 with Ala severely decreased activity (**Figure 4.24**), and hence Tyr8 might be a hydrogen atom donor that generates the final product. We ruled out the possibility, at least *in vitro*, that the final hydrogen atom donor is 5'-dA by

conducting experiments with SAM that was deuterated at both the methyl group and the 5' position (**Figure 4.16**, pathway *c*). Use of this deuterated SAM analogue did not result in an increase in the deuterium content of the final product. However, from an efficiency perspective, hydrogen atom abstraction from 5'-dA would be attractive as it would directly regenerate the 5'-dA• and would not require the additional electron required for pathways *a* and *b*. Furthermore, as shown by the SAM-dependent solvent exchange of the methyl group in the product, 5'-dA• may abstract a hydrogen atom from the methyl group in the product, which would be the microscopic reverse of the last step of pathway *c*. Hence, while pathway *c* clearly is not consistent with our *in vitro* findings, perhaps under the right conditions pathway *c* might be operational *in vivo*.

The proposed mechanism accounts for the observed solvent exchange upon prolonged reaction times by realization that the methylated peptide product, SAM1, and SAH can likely still bind to the active site since these three molecules can occupy nearly identical positions as the substrate peptide, SAM1, and SAM2 at the start of catalysis (the only difference is that the methyl group of SAM2 is transferred to the peptide). Reductive cleavage of SAM1 would generate 5'-dA• and the methyl group of the product peptide would be in a very similar position as the methyl group of SAM2 during productive turnover. Thus, hydrogen atom abstraction could again generate radical **11** and either pathway *a* or *b* would result in hydrogen transfer by an active site residue, thus affecting the observed exchange (**Figure 4.25A**). Alternatively, the exchange could occur by a heterolytic deprotonation/reprotonation sequence that requires bound SAM (**Figure 4.25B**).

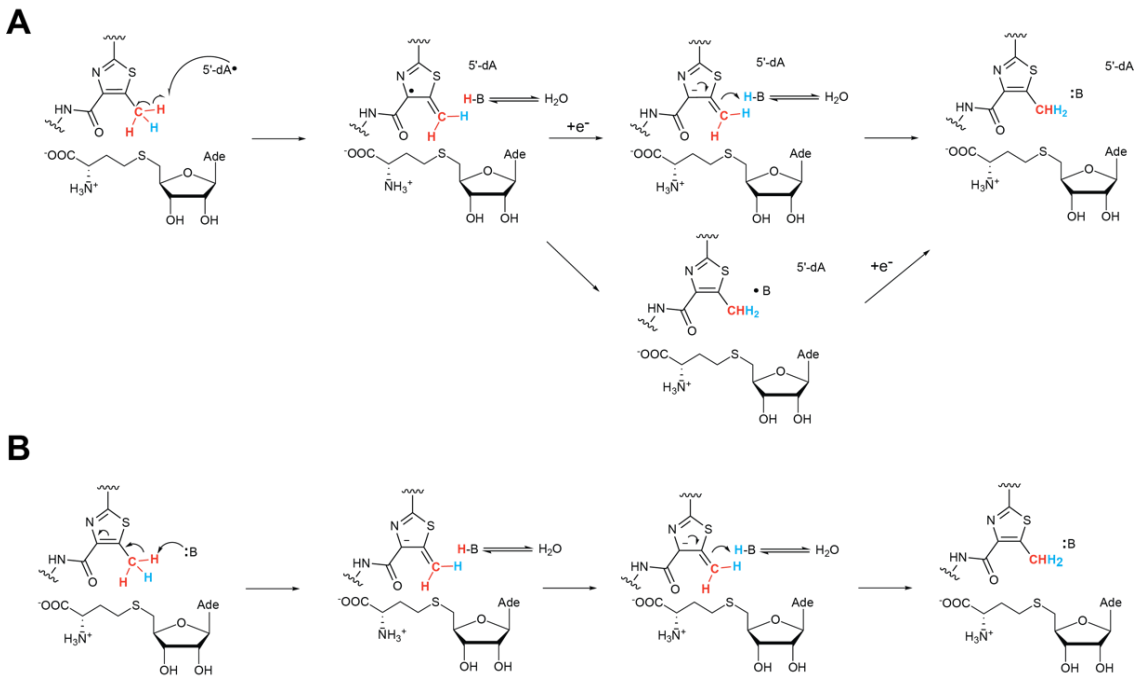


Figure 4.25 Proposed mechanisms for exchange of the methyl protons of the product peptide. (**A**) The methylated product and SAH can bind to the enzyme, SAM1 can be cleaved to initiate the hydrogen atom abstraction from the methylated peptide. Its methyl group would be expected to occupy a very similar position as the methyl group of SAM2 during normal turnover. Reduction of the resulting radical to the anion and protonation results in exchange (analogous to pathway *a* in **Figure 4.16**), or hydrogen atom abstraction from a redox active residue in the active site (analogous to pathway *b*, **Figure 4.16**) would result in exchange. (**B**) An alternative mechanism involves deprotonation of the methyl group, the reverse of the protonation step that provides the product in pathway *a*. The mechanism in panel (**A**) requires SAM cleavage whereas the mechanism in (**B**) does not. Unfortunately, we were not able to determine whether a hydrogen is transferred from the methyl group to 5'-dA during prolonged reactions because of extensive abortive SAM cleavage. Ade = adenine.

In summary, our study suggests that methyl transfer catalyzed by TbtI involves two SAM molecules. Our data are consistent with a mechanism in which one SAM is reductively cleaved to 5'dA• as in other radical SAM enzymes. The 5'dA• then abstracts a hydrogen atom from a second SAM, and the resulting carbon-based methylene radical attacks the thiazole at position 4 of the

peptide substrate. This SAM-derived radical is unprecedented in radical SAM enzymology. We note that like the class A radical SAM MTs and a recent report on the class C MT NosN (62), TbtI formally transfers a methylene group from SAM to the substrate and hence might not be considered a true methyl transferase. Regardless, the manner by which this is accomplished adds a fundamentally different use of SAM in this class of enzymes, and highlights their versatility and catalytic diversity in natural product biosynthesis. Additional biochemical and structural studies on TbtI are underway to gain further insights to the mechanism of this remarkable transformation.

4.3 Experimental methods

General materials and methods

Reagents used for molecular biology experiments were purchased from New England BioLabs (Ipswich, MA), Thermo Fisher Scientific (Waltham, MA), or Gold Biotechnology Inc. (St. Louis, MO). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). *Escherichia coli* DH5α and BL21 (DE3) strains were used for plasmid maintenance and protein overexpression, respectively. *E. coli* JW3582-2 that is auxotrophic for cysteine was purchased from the Coli Genetic Stock Center at Yale University (https://cgsc2.biology.yale.edu/Strain.php?ID=108920). Plasmid inserts were sequenced at the University of Illinois core sequencing facility and ACGT Inc. (Wheeling, IL). MALDI-TOF-MS analysis was performed using a Bruker UltrafleXtreme matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer (Bruker Daltonics) in reflector positive mode at the University of Illinois School of Chemical Sciences Mass Spectrometry Laboratory. ESI-MS/MS analyses were performed using a SYNAPT electrospray ionization (ESI) quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid

Chromatography (UPLC) system (Waters). HiTrap columns for Ni-NTA affinity chromatography were purchased from GE Healthcare.

Molecular biology techniques

Oligonucleotides were purchased from Integrated DNA Technologies Inc. (Coralville, IA). Genes optimized for recombinant expression in *E. coli* were synthesized by GenScript (Piscataway, NJ) in pUC57 (kanamycin, Kan) vectors with BamHI and XhoI sites flanking each gene at the 5' and 3' ends, respectively. *E. coli* DH5α were transformed with pUC57-Kan vectors containing each gene for replication and subsequent isolation using a QIAprep Spin Miniprep Kit (Qiagen). The isolated DNA was then treated with BamHI-HF and XhoI-HF (New England Biolabs; NEB). The digested genes were separated on a 1% (*w/v*) agarose gel, purified using a QIAQuick gel extraction kit (Qiagen), and ligated into an appropriately endonuclease-digested and gel-purified pET28 vector with maltose binding protein (MBP) 5' to the multiple cloning site using T4 DNA ligase (NEB). Ligation reactions were used to transform chemically competent DH5α cells, which were plated on Luria-Bertani (LB) agar plates containing 50 μg/mL kanamycin and grown at 37 °C. Colonies were picked at random and grown in LB broth for 16–20 h prior to plasmid isolation using a QIAprep Spin Miniprep Kit. The construct encoding for *tbtA* with either one Cys at position 4 or two Cys at positions 2 and 4 of the core as well as all Cys replaced with Ala was previously described and the latter was used for downstream site-directed mutagenesis (SDM) (45). SDM was performed using the QuikChange method (Agilent) as per the manufacturer's instructions using PfuTurbo DNA polymerase. The primers for each mutant are listed in **Table 4.2**. The mutations were verified by sequencing using a custom MBP forward primer or the T7 reverse primer (**Table 4.2**).

MBP-tagged TbtI overexpression

E. coli BL21(DE3) cells were co-transformed with a pET28 plasmid encoding the MBP-tagged TbtI and the pSUF plasmid encoding the Fe-S biosynthetic operon (38). Cells were grown for 24 h on LB agar plates containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37 °C. Single colonies were used to inoculate 10 mL of LB containing the same concentration of antibiotics and grown at 37 °C for 16–18 h. This culture was used to inoculate 1 L of LB (5 g/L yeast extract, 10 g/L tryptone and 10 g/L NaCl) supplemented with the same concentration of antibiotics and grown to an optical density at 600 nm (OD_{600}) of 0.6 before being placed on ice for 15 min. Protein expression was then induced with the addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and supplemented with 100 mg/L of ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2$) and 100 mg/L of cysteine. Expression was allowed to proceed for 12–16 h at 15 °C. Cells were harvested by centrifugation at 3,000 × *g* for 20 min, washed with phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4), and harvested by centrifugation. The cells were flash-frozen and stored at –80 °C for a maximum of one week before use (notebook reference 5/5/2016).

MBP-tagged TbtI purification

Protein purification was performed in a Coy anaerobic chamber. All buffers were degassed and stored for 24–48 h in the anaerobic chamber before use. Cells were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 2.5% glycerol (*v/v*), and 0.1% Triton X-100 (*v/v*)) containing 4 mg/mL lysozyme, 2 μM leupeptin, 2 μM benzamidine, and 2 μM E64 to inhibit proteolysis. Cells were further lysed by sonication (3 × 45 s with 10 min agitation periods at 4 °C). Insoluble debris was removed by centrifugation at 20,000 × *g* for 40 min at 4 °C. The supernatant

was then applied to lysis-buffer pre-equilibrated amylose resin (NEB; 15 mL of resin per L of initial cell culture). The column was washed with 10 column volumes (CV) of lysis buffer followed by 10 CV of wash buffer (lysis buffer with 400 mM NaCl and lacking Triton X-100). The MBP-tagged proteins were eluted using elution buffer (lysis buffer with 300 mM NaCl, 10 mM maltose, but omitting Triton X-100) until the eluent no longer contained protein detectable with the Bradford reagent. Eluent was concentrated using a 30 kDa molecular weight cut-off (MWCO) Amicon Ultra centrifugal filter (EMD Millipore). A buffer exchange with 10× volume of protein storage buffer [50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 300 mM NaCl, 2.5% glycerol (*v/v*)] was performed prior to final concentration, freezing in liquid nitrogen, and storage at –80 °C. Protein concentrations were determined using both absorbance at 280 nm (theoretical extinction coefficients were calculated using the ExPASy ProtParam tool; http://web.expasy.org/protparam/protpar-ref.html) and a Bradford colorimetric assay. Purity was assessed visually by analysis of Coomassie-stained SDS-PAGE gels (**Figure 4.2A**). All wash, elution, and storage buffers were supplemented with 0.5 mM tris-(2-carboxyethyl)-phosphine (TCEP) (notebook reference 5/6/2016).

His_6-tagged cysteine desulfurase (IscS) overexpression

E. coli BL21 (DE3) cells were co-transformed with a *His_6-IscS-pET15* plasmid encoding His_6-tagged IscS (plasmid obtained from Prof. Tadhg Begley group at Texas A&M University). Cells were grown for 24 h on LB agar plates containing 100 μg/mL ampicillin at 37 °C. Single colonies were used to inoculate 10 mL of LB containing 100 μg/mL ampicillin and grown at 37 °C for 16–18 h. This culture was used to inoculate 1 L of LB (5 g/L yeast extract, 10 g/L tryptone and 10 g/L NaCl) and grown to an optical density at 600 nm (OD_{600}) of 0.6 before being placed

on ice for 15 min. Protein expression was then induced with the addition of 0.5 mM IPTG. Expression was allowed to proceed for 12–16 h at 15 °C. Cells were harvested by centrifugation at 3,000 × *g* for 20 min, washed with PBS and harvested by centrifugation. The cells were flash-frozen and stored at –80 °C for a maximum of one week before use (notebook reference 6/3/2016).

His_6-tagged IscS purification

Cells were resuspended in lysis buffer (20 mM Na_2HPO_4, 500 mM NaCl, 10 mM imidazole, pH 7.4, 2.5% glycerol (v/v), and 0.1% Triton X-100) containing 4 mg/mL lysozyme, 2 μM leupeptin, 2 μM benzamidine, and 2 μM E64. Cells were further lysed by sonication (3 × 45 s with 10 min agitation periods at 4 °C). Insoluble debris was removed by centrifugation at 20,000 × *g* for 40 min. The supernatant was then applied to a pre-equilibrated Ni-NTA resin (Thermo Scientific; 15 mL of resin per L of cells). The column was washed with 10 CV of lysis buffer followed by 10 CV of wash buffer (lysis buffer with 30 mM imidazole and lacking Triton X-100). The His_6-tagged proteins were eluted using elution buffer (lysis buffer with 300 mM NaCl, 300 mM imidazole and lacking Triton X-100) until the eluent no longer contained protein detectable with the Bradford reagent. Eluent was further buffer exchanged with storage buffer (20 mM Tris-HCl , 300 mM NaCl, 2.5% glycerol (v/v), pH 7.4) using PD-10 column, and the protein solution was concentrated and stored and the protein concentration was assayed as described for TbtI (*vide supra*). Purity and possible truncation were assessed by Coomassie-stained SDS-PAGE gel (**Figure 4.2A**). All wash, elution, and storage buffers were supplemented with 0.5 mM tris-(2-carboxyethyl)-phosphine (TCEP) (notebook reference 6/5/2016).

In vitro reconstitution of the Fe-S cluster in TbtI

TbtI (50 μM) was equilibrated with 1 mM ferrous ammonium sulfate [$Fe(NH_4)_2(SO_4)_2$], cysteine (2 mM), dithiothreitol (DTT, 2 mM) and IscS (2 μM) in 50 mM Tris-HCl pH 7.5 buffer in the anaerobic chamber. The reaction mixture was buffer exchanged using a 10 kDa MWCO gel filtration column (EMD Millipore) before concentration and use for assays. The reactions were allowed to proceed for 4–6 h (notebook reference 9/5/2016).

In vitro methylation of TbtA by TbtI

Reaction mixtures generally included the following components: substrate (50 μM), purified TbtI from *E. coli* heterologous expression after in vitro reconstitution of the Fe-S cluster (2 μM), *S*-adenosyl-L-methionine (SAM, 1 mM) and sodium dithionite (5 mM) in reaction buffer (50 mM Tris-HCl pH 7.5). The reaction was allowed to proceed for 6–12 h at 25 °C in an anaerobic chamber. Reaction mixtures were desalted via C18 ZipTip (EMD Millipore) per the manufacturer's instructions, and the product was eluted using a saturated solution of sinapinic acid in 60% aq. MeCN. Reactions were monitored using MALDI-TOF-MS.

The in vitro methylation reaction was also carried out using flavodoxin/flavodoxin reductase/NADPH as the reducing system. Briefly, under this condition, the *in vitro* methylation reaction mixtures generally included the following components: TbtA hexazole substrate (50–100 μM), purified TbtI from *E. coli* heterologous expression after *in vitro* reconstitution of the [4Fe-4S] cluster (10 μM), SAM (1 mM), flavodoxin (10 μM), flavodoxin reductase (10 μM), reduced nicotinamide adenine dinucleotide phosphate (NAPDH, 2 mM) in reaction buffer (50 mM Tris-HCl, pH 7.5). The reaction was allowed to proceed for 1 h at 25 °C in an anaerobic chamber (Coy laboratory) (notebook reference 9/5/2016).

For reactions performed in D_2O, all reagents that participate in the reaction were first dissolved in D_2O, lyophilized and re-dissolved in D_2O followed by adjusting of the pD. The buffer containing the proteins was exchanged with D_2O storage buffer (50 mM Tris, 300 mM NaCl, pD 7.5) by repeated concentration by ultrafiltration and dilution (four times) before addition into the reaction. The reaction was allowed to proceed in D_2O reaction buffer (50 mM Tris, pD 7.5) for 1 h at 25 °C in the anaerobic chamber (notebook reference 6/3/2017).

For single turnover reactions, a reaction mixture generally included the following components: TbtA hexazole substrate (20 μM), purified TbtI after *in vitro* reconstitution of the [4Fe-4S] cluster (60 μM), SAM (1 mM), flavodoxin (30 μM), flavodoxin reductase (30 μM), reduced nicotinamide adenine dinucleotide phosphate (NAPDH, 2 mM) in reaction buffer (50 mM Tris-HCl, pH 7.5). The reaction was allowed to proceed for 10 min at 25 °C in the anaerobic chamber.

To determine the specific activity of the TbtI used, the enzyme (10 μM) was incubated with the hexazole substrate peptide (100 μM), SAM (1 mM), flavodoxin (20 μM), flavodoxin reductase (20 μM), reduced nicotinamide adenine dinucleotide phosphate (NAPDH, 2 mM) in reaction buffer (50 mM Tris-HCl, pH 7.5). The reaction was allowed to proceed at 25 °C in an anaerobic chamber, and aliquots were removed at set time points (30, 60, 90, and 120 s). At 120 s the conversion was 15%. Product formation over time was linear with a velocity of 15.6×10^{-3} μmol/min/mg of TbtI (i.e. 0.72 min^{-1}). To assure that under these conditions the reaction catalyzed by TbtI was rate limiting, the amounts of flavodoxin and flavodoxin reductase were doubled, and no change in velocity was observed. Similarly, to assure that all substrates were saturating, the concentrations of SAM, NADPH and peptide were doubled without observing a change in velocity (notebook reference 11/2/2017).

Generation of ^{13}C- and ^{15}N-depleted TbtA hexazole

Isotopically depleted TbtA hexazole was produced in *E. coli* cultured in modified M9 minimal media (69). *E. coli* BL21(DE3) cells containing a *pTrc99a_TbtF,G,E* plasmid encoding MBP-tagged TbtF, MBP-tagged TbtG, and TbtE and a *His_6-TbtA-pRSFDuet-1* plasmid encoding the His_6-TbtA were grown overnight at 37 ºC in lysogeny broth (LB) containing 25 μg/mL kanamycin and 50 μg/mL ampicillin. After harvesting the cells, the supernatant was discarded and the cells were washed with 5 mL of wash buffer (22 mM KH_2PO_4, 42 mM Na_2HPO_4, 8.5 mM NaCl, pH 7.4). Following washing, the cells were resuspended in wash buffer and used to inoculate (1:100) the modified M9 minimal media, with the following composition per 100 mL: 10 mL of a 10× stock of minimal media (220 mM KH_2PO_4, 420 mM Na_2HPO_4, 85 mM NaCl, pH 7.4), 0.3 mL of 40% aqueous $(^{14}NH_4)_2SO_4$ (Sigma-Aldrich), 2 mL of 20% aqueous ^{13}C-depleted glucose (Cambridge Isotope Laboratories), 0.1 mg of $FeSO_4$, 10 μg of thiamine, 200 μL of 1 M $MgSO_4$, 10 μL of 1 M $CaCl_2$, and 75 μL of a trace element solution (5 mM $CaCl_2$, 1.25 mM $ZnCl_2$, 260 μM $CuCl_2•H_2O$, 252 μM $CoCl_2•6H_2O$, 250 μM $Na_2MoO_4•2H_2O$, pH 7.4). The media also contained 50 μg/mL ampicillin and 25 μg/mL kanamycin. The cells were grown at 37 °C and induced at OD_{600} = 0.6-0.8 by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and grown for an additional 16-18 h at 18 °C before harvesting. Purification of TbtA hexazole was carried out as previously described (45) (notebook reference 3/4/2017).

Generation of TbtA hexazole deuterium labeled at each thiazole

β-D_2-cysteine labeled MBP-tagged TbtA was generated using *E. coli* strain JW3582-2. A lysogenization step was performed to JW3582-2 before production of MBP-tagged TbtA so that

the host strain could be used to express target genes cloned in T7 expression vectors. Lysogenization was performed using the λDE3 Lysogenization Kit (Novagen) as per the manufacturer's instructions. The lysogenized JW3582-2 was transformed with a *pET28a_MBP-TbtA* plasmid containing the ampicillin resistance gene and encoding MBP-tagged TbtA.

Expression of β-D_2-cysteine labeled MBP-TbtA was performed in modified M9 minimal media as described above, except that amino acids were added into the culture (final concentrations: 0.5 mM Arg, 1 mM His, 1 mM Leu, 1 mM Lys, 0.5 mM Met, 2 mM Pro, 3 mM Thr, 0.4 mM Trp and 1 mM of β-D_2-Cys). Purification and *in vitro* cyclodehydration of the β-D_2-cysteine-labeled MBP-tagged TbtA was carried out as previously described (65) (notebook reference 6/12/2017).

Preparation of SAM

E. coli strain DM22-(pK8), which overexpresses *E. coli* SAM synthetase, was constructed by G. D. Markham *et al.* (Fox Chase Cancer Center, Philadelphia, PA) (70). DM22-(pK8) was grown in LB containing 30 μg/mL tetracycline at 37 °C while being shaken and harvested by centrifugation 12 h after inoculation. The cells (~12 g) were resuspended in 36 mL of 0.1 M Tris-HCl (pH 8.0) containing 1 mM ethylenediaminetetraacetic acid (EDTA) and 50 μg/mL lysozyme and gently stirred at room temperature for 30 min. Phenylmethylsulfonyl fluoride (PMSF) was added to a final concentration of 1 mM, and cells were lysed with a MultiFlex C3 Homogenizer (Avestin). The cell lysate was centrifuged for 30 min at 26,500 × g at 4 °C. The cleared lysate was dialyzed against 4 L of 100 mM Tris-HCl (pH 8.0), 1 mM EDTA, in 2 × 30 mL 20,000 molecular weight cut-off (MWCO) Slide-A-Lyzer dialysis cassettes (Pierce) at 4 °C overnight, and was divided into three portions, frozen in liquid nitrogen, and stored at –80°C.

SAM was generated enzymatically by incubating an aliquot of the lysate in a buffered solution (0.1 L) containing 0.1 M Tris-HCl (pH 8.0), 50 mM KCl, 1 mM EDTA, 20% acetonitrile, 26 mM $MgCl_2$, 13 mM adenosine triphosphate (ATP), and 10 mM methionine. After the mixture had been gently stirred for 5 h at room temperature, the reaction was quenched by adjusting the pH to 5.0 with HCl. The mixture was placed on ice for 15 min and centrifuged at 4,000 × g for 30 min at 4 °C to remove precipitated material. The supernatant was diluted to 1 L with 1 mM sodium acetate (pH 5.0) and loaded on a CM-52 cation exchange resin (Whatman) that had been charged with 0.2 M sodium acetate (pH 5.0) and equilibrated in 1 mM sodium acetate (pH 5.0). The column was rinsed with 1 L of 1 mM sodium acetate (pH 5.0), and SAM was eluted with 0.5 L of 40 mM H_2SO_4. The presence of SAM was monitored by UV-Vis absorption at 260 nm, and fractions containing SAM were combined. The pH of the solution was adjusted to 5.2 with $Ba(OH)_2$ and then filtered to remove the $BaSO_4$ precipitate. The filtered solution was concentrated by rotovap to a volume of 4 mL, and was divided into several fractions and stored at −80 °C. Deuterium-labeled SAM was prepared using the same procedure but with labelled methionine or ATP (notebook reference 2/18/2017).

Radioactive assay with $^{14}CH_3$-SAM

The assay to determine whether the SAM methyl group is transferred onto TbtI was performed with $^{14}CH_3$-SAM (0.02 mCi/mL, 300 μM, PerkinElmer). The *in vitro* reconstitution of the [4Fe-4S] cluster of MBP-TbtI was performed as previously described, followed by addition of TbtA hexazole inside an anaerobic chamber. After incubation at room temperature for 1 h, the reaction mixture was taken out of the anaerobic chamber and moved into an anaerobic bag filled with nitrogen gas, and 1:1 $^{14}CH_3$-SAM/unlabeled CH_3-SAM mixture was then added into the

reaction inside the anaerobic bag. The reaction mixture included the following components: MBP-TbtI (10 μM), TbtA hexazole (100 μM), 1:1 $^{14}CH_3$-SAM/unlabeled CH_3-SAM mixture (300 μM) in a final volume of 100 μL. The reaction was incubated at room temperature for 1 h, and 30 μL of the reaction mixture was loaded on a 4-20% polyacrylamide gel. Electrophoresis was performed for 45 min at 150 V (constant voltage), and then the gel was stained with Coomassie blue solution (40% methanol, 10% acetic acid, 0.1 % Coomassie blue R250 in H_2O) for 30 min and destained in destain buffer (20% methanol, 10% acetic acid in H_2O) for 3 h. After drying the gel overnight with a gel-drying kit (Promega) in accordance with the manufacturer's procedure, the dried gel was exposed on a phosphorimager cassette for 4 days. The film was visualized with a STORM840 Phosphorimager Scanner (Amersham Biosciences). Whereas the Coomassie stained gel clearly showed the presence of the protein, no radioactivity was observed in the gel (notebook reference 11/5/2017).

Table 4.2 Oligonucleotides primers used in this study. All sequences are provided 5′ to 3′. F indicates a forward primer while R indicates the reverse primer.

Primer Name	Oligonucleotide Sequence
tbtI C14A F	ATATCCCGTTCGCTAACTCTAAATGCCACTTCTGCGACTGG
tbtI C14A R	CATTTAGAGTTAGCGAACGGGATATTCACATACAGCAGCAG
tbtI C18A F	GTAACTCTAAAGCCCACTTCTGCGACTGGGTTGTCCAAATC
tbtI C18A R	TCGCAGAAGTGGGCTTTAGAGTTACAGAACGGGATATTCAC
tbtI C21A F	AATGCCACTTCGCCGACTGGGTTGTCCAAATCCCGGTGCGT
tbtI C21A R	ACAACCCAGTCGGCGAAGTGGCATTTAGAGTTACAGAACGG
tbtI C89A F	CTCTGTATACGGCCCTGCGTTCAGAATTCGATCTGTCGCAC
tbtI C89A R	TCTGAACGCAGGGCCGTATACAGAGATTCGATTTCGCTTTC
tbtI C254A F	GTGCACCGCGTGCCGAAGCAGATCAGGCATATTACCGCCTG
tbtI C254A R	TGATCTGCTTCGGCACGCGGTGCACCGAAGTAGCTCATCGC
tbtI C343A F	TGCGTGTCGCAGCTGAAGAACCGGAAGTGCAAGCCTACCTG
tbtI C343A R	TCCGGTTCTTCAGCTGCGACACGCAGCGGCGTACCGGTACGT
tbtI W23A F	ACTTCTGCGACGCGGTTGTCCAAATCCCGGTGCGTGATCTG

Table 4.2 (cont.)

Primer Name	Oligonucleotide Sequence
tbtI W23A R	ATTTGGACAACCGCGTCGCAGAAGTGGCATTTAGAGTTACA
tbtI Y200A F	ATTTTAGTATCGCTCCGTACCGTGCATCCCCGGGTACCATT
tbtI Y200A R	GCACGGTACGGAGCGATACTAAAATGGTTAATCGGCAGCGT
tbtI E105A F	AAACCACGATTGCAGGTAGTCCGGAATCCCTGACCCCGCAG
tbtI E105A R	TCCGGACTACCTGCAATCGTGGTTTCGCGGACGTGCGACAG
tbtI Y8A F	CGCTGCTGCTGGCTGTGAATATCCCGTTCTGTAACTCTAAA
tbtI Y8A R	GGGATATTCACAGCCAGCAGCAGCGGGCGCGTCAT
tbtA N3A F	GTGCAAGCTGTGCTTGCTTTTGTTATATTTGTTGTAGCTGC
tbtA N3A R	TAACAAAAGCAAGCACAGCTTGCACCAACTTCGGTCATACC
tbtA N3D F	GTGCAAGCTGTGATTGCTTTTGTTATATTTGTTGTAGCTGC
tbtA N3D R	TAACAAAAGCAATCACAGCTTGCACCAACTTCGGTCATACC
tbtA N3Q F	GTGCAAGCTGTCAATGCTTTTGTTATATTTGTTGTAGCTGC
tbtA N3Q R	TAACAAAAGCATTGACAGCTTGCACCAACTTCGGTCATACC
tbtA C4A F	CAAGCTGTAATGCCTTTTGTTATATTTGTTGTAGCTGCAGC
tbtA C4A R	ATATAACAAAAGGCATTACAGCTTGCACCAACTTCGGTCAT
tbtA C4S F	CAAGCTGTAATAGCTTTTGTTATATTTGTTGTAGCTGCAGC
tbtA C4S R	ATATAACAAAAGCTATTACAGCTTGCACCAACTTCGGTCAT
tbtA S1N/C4A F	AAGTTGGTGCAAACTGTAATGCATTTTGTTATATTTGTTGT
tbtA S1N/C4A R	AATGCATTACAGTTTGCACCAACTTCGGTCATACCATGACC
tbtA F5N/C4A F	GCTGTAATGCAAATTGTTATATTTGTTGTAGCTGCAGCAGC
tbtA F5N/C4A R	CAAATATAACAATTTGCATTACAGCTTGCACCAACTTCGGT
tbtA I8N/C4A F	CATTTTGTTATAATTGTTGTAGCTGCAGCAGCGCCTAATGC
tbtA I8N/C4A R	CAGCTACAACAATTATAACAAAATGCATTACAGCTTGCACC
tbtA S11N/C4A F	ATATTTGTTGTAACTGCAGCAGCGCCTAATGCGGCCG
tbtA S11N/C4A R	GCGCTGCTGCAGTTACAACAAATATAACAAAATGCATTACA
tbtA C4,6 F	CGAATTGTTTTTGTTATATTGCGGCGAGCGCCAGCAGCGCC
tbtA C4,6 R	GCCGCAATATAACAAAAACAATTCGCGCTTGCACCAACTTC
tbtA C4,9 F	TTGCGTATATTTGTGCGAGCGCCAGCAGCGCCTAACTCGAG
tbtA C4,9 R	GCTGGCGCTCGCACAAATATACGCAAAACAATTCGCGCTTG
tbtA C4,10 F	CGTATATTGCGTGTAGCGCCAGCAGCGCCTAACTCGAGCAC
tbtA C4,10 R	CTGCTGGCGCTACACGCAATATACGCAAAACAATTCGCGC
tbtA C4,12 F	TTGCGGCGAGCTGTAGCAGCGCCTAACTCGAGCACCACCAC
tbtA C4,12 R	TAGGCGCTGCTACAGCTCGCCGCAATATACGCAAAACAATTC
tbtA S1N/N3F F	AAGTTGGTGCAAACTGTTTTTGCTTTTGTTATATTTGTTGTAGCTGCAGC
tbtA S1N/N3F R	AACAAAAGCAAAAACAGTTTGCACCAACTTCGGTCATACCATGACCTGCGG

4.4 References

1. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res., 29*, 1097-1106.

2. Vey, J. L., & Drennan, C. L. (2011). Structural insights into radical generation by the radical SAM superfamily. *Chem. Rev., 111*, 2487-2506.

3. Broderick, J. B., Duffus, B. R., Duschene, K. S., & Shepard, E. M. (2014). Radical S-adenosylmethionine enzymes. *Chem. Rev., 114*, 4229-4317.

4. Mehta, A. P., Abdelwahed, S. H., Mahanta, N., Fedoseyenko, D., Philmus, B., Cooper, L. E., Liu, Y., Jhulki, I., Ealick, S. E., & Begley, T. P. (2015). Radical S-adenosylmethionine (SAM) enzymes in cofactor biosynthesis: a treasure trove of complex organic radical rearrangement reactions. *J. Biol. Chem., 290*, 3980-3986.

5. Zhang, Q., van der Donk, W. A., & Liu, W. (2012). Radical-mediated enzymatic methylation: A tale of two SAMS. *Acc. Chem. Res., 45*, 555-564.

6. Chan, K. K. J., Thompson, S., & O'Hagan, D. (2013). The mechanisms of radical SAM/cobalamin methylations: an evolving working hypothesis. *ChemBioChem, 14*, 675-677.

7. Fujimori, D. G. (2013). Radical SAM-mediated methylation reactions. *Curr. Opin. Chem. Biol., 17*, 597-604.

8. Zhang, Q., van der Donk, W. A., & Liu, W. (2012). Radical-Mediated Enzymatic Methylation: A Tale of Two SAMS. *Acc. Chem. Res., 45*, 555-564.

9. Bauerle, M. R., Schwalm, E. L., & Booker, S. J. (2015). Mechanistic Diversity of Radical S-Adenosylmethionine (SAM)-dependent Methylation. *J. Biol. Chem., 290*, 3995-4002.

10. Ding, Y., Yu, Y., Pan, H., Guo, H., Li, Y., & Liu, W. (2010). Moving posttranslational modifications forward to biosynthesize the glycosylated thiopeptide nocathiacin I in *Nocardia* sp. ATCC202099. *Mol. Biosyst., 6*, 1180-1185.

11. Grove, T. L., Benner, J. S., Radle, M. I., Ahlum, J. H., Landgraf, B. J., Krebs, C., & Booker, S. J. (2011). A radically different mechanism for S-adenosylmethionine-dependent methyltransferases. *Science, 332*, 604-607.

12. Grove, T. L., Livada, J., Schwalm, E. L., Green, M. T., Booker, S. J., & Silakov, A. (2013). A substrate radical intermediate in catalysis by the antibiotic resistance protein Cfr. *Nat. Chem. Biol., 9*, 422-427.

13. Pierre, S., Guillot, A., Benjdia, A., Sandstroem, C., Langella, P., & Berteau, O. (2012). Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes. *Nat. Chem. Biol., 8*, 957-959.

14. Blaszczyk, A. J., Silakov, A., Zhang, B., Maiocco, S. J., Lanz, N. D., Kelly, W. L., Elliott, S. J., Krebs, C., & Booker, S. J. (2016). Spectroscopic and Electrochemical Characterization of the Iron-Sulfur and Cobalamin Cofactors of TsrM, an Unusual Radical S-Adenosylmethionine Methylase. *J. Am. Chem. Soc., 138*, 3416-3426.

15. Allen, K. D., & Wang, S. C. (2014). Initial characterization of Fom3 from Streptomyces wedmorensis: The methyltransferase in fosfomycin biosynthesis. *Arch. Biochem. Biophys., 543*, 67-73.

16. Woodyer, R. D., Li, G., Zhao, H., & van der Donk, W. A. (2007). New insight into the mechanism of methyl transfer during the biosynthesis of fosfomycin. *Chem. Commun.*, 359-361.

17. Wang, B., Blaszczyk, A. J., Knox, H. L., Zhou, S., Blaesi, E. J., Krebs, C., Wang, R. X., & Booker, S. J. (2018). Stereochemical and mechanistic investigation of the reaction catalyzed by Fom3 from streptomyces fradiae, a cobalamin-dependent radical S-adenosylmethionine methylase. *Biochemistry, 57*, 4972-4984.

18. McLaughlin, M. I., & van der Donk, W. A. (2018). Stereospecific radical-mediated B12-dependent methyl transfer by the fosfomycin biosynthesis enzyme Fom3. *Biochemistry, 57*, 4967-4971.

19. Sato, S., Kudo, F., Kuzuyama, T., Hammerschmidt, F., & Eguchi, T. (2018). C-methylation catalyzed by Fom3, a cobalamin-dependent radical S-adenosyl-l-methionine enzyme in fosfomycin biosynthesis, proceeds with inversion of configuration. *Biochemistry, 57*, 4963-4966.

20. Kim, H. J., McCarty, R. M., Ogasawara, Y., Liu, Y.-n., Mansoorabadi, S. O., LeVieux, J., & Liu, H.-w. (2013). GenK-catalyzed C-6' methylation in the biosynthesis of gentamicin: isolation and characterization of a cobalamin-dependent radical SAM enzyme. *J. Am. Chem. Soc., 135*, 8093-8096.

21. Wang, Y., Schnell, B., Baumann, S., Muller, R., & Begley, T. P. (2017). Biosynthesis of branched alkoxy groups: iterative methyl group alkylation by a cobalamin-dependent radical SAM enzyme. *J. Am. Chem. Soc., 139*, 1742-1745.

22. Parent, A., Guillot, A., Benjdia, A., Chartier, G., Leprince, J., & Berteau, O. (2016). The B12-radical SAM enzyme poyC catalyzes valine Cβ-methylation during polytheonamide biosynthesis. *J. Am. Chem. Soc., 138*, 15515-15518.

23. Huang, W., Xu, H., Li, Y., Zhang, F., Chen, X.-Y., He, Q.-L., Igarashi, Y., & Tang, G.-L. (2012). Characterization of yatakemycin gene cluster revealing a radical S-

adenosylmethionine dependent methyltransferase and highlighting spirocyclopropane biosynthesis. *J. Am. Chem. Soc., 134*, 8831-8840.

24. Hiratsuka, T., Suzuki, H., Kariya, R., Seo, T., Minami, A., & Oikawa, H. (2014). Biosynthesis of the structurally unique polycyclopropanated polyketide-nucleoside hybrid jawsamycin (FR-900848). *Angew. Chem., Int. Ed., 53*, 5423-5426.

25. Yan, F., LaMarre, J. M., Rohrich, R., Wiesner, J., Jomaa, H., Mankin, A. S., & Fujimori, D. G. (2010). RlmN and Cfr are radical SAM enzymes involved in methylation of ribosomal RNA. *J. Am. Chem. Soc., 132*, 3953-3964.

26. Boal, A. K., Grove, T. L., McLaughlin, M. I., Yennawar, N. H., Booker, S. J., & Rosenzweig, A. C. (2011). Structural basis for methyl transfer by a radical SAM enzyme. *Science, 332*, 1089-1092.

27. Yan, F., & Fujimori, D. G. (2011). RNA methylation by radical SAM enzymes RlmN and Cfr proceeds via methylene transfer and hydride shift. *Proc. Natl. Acad. Sci. USA, 108*, 3930-3934.

28. Grove, T. L., Benner, J. S., Radle, M. I., Ahlum, J. H., Landgraf, B. J., Krebs, C., & Booker, S. J. (2011). A radically different mechanism for S-adenosylmethionine-dependent methyltransferases. *Science, 332*, 604-607.

29. Schwalm, E. L., Grove, T. L., Booker, S. J., & Boal, A. K. (2016). Crystallographic capture of a radical S-adenosylmethionine enzyme in the act of modifying tRNA. *Science, 352*, 309-312.

30. Boal, A. K., Grove, T. L., McLaughlin, M. I., Yennawar, N. H., Booker, S. J., & Rosenzweig, A. C. (2011). Structural basis for methyl transfer by a radical SAM enzyme. *Science, 332*, 1089-1092.

31. Ding, W., Li, Y., Zhao, J., Ji, X., Mo, T., Qianzhu, H., Tu, T., Deng, Z., Yu, Y., Chen, F., & Zhang, Q. (2017). The catalytic mechanism of the Class C radical S-adenosylmethionine methyltransferase NosN. *Angew. Chem. Int. Ed. Engl*., 129, 3915-3919.

32. LaMattina, J. W., Nix, D. B., & Lanzilotta, W. N. (2016). Radical new paradigm for heme degradation in Escherichia coli O157:H7. *Proc. Natl. Acad. Sci. USA, 113*, 12138-12143.

33. Allen, K. D., Xu, H., & White, R. H. (2014). Identification of a unique radical S-adenosylmethionine methylase likely involved in methanopterin biosynthesis in Methanocaldococcus jannaschii. *J. Bacteriol., 196*, 3315-3323.

34. Bagley, M. C., Dale, J. W., Merritt, E. A., & Xiong, X. (2005). Thiopeptide antibiotics. *Chem. Rev., 105*, 685-714.

35. Morris, R. P., Leeds, J. A., Naegeli, H. U., Oberer, L., Memmert, K., Weber, E., LaMarche, M. J., Parker, C. N., Burrer, N., Esterow, S., Hein, A. E., Schmitt, E. K., & Krastel, P. (2009). Ribosomally synthesized thiopeptide antibiotics targeting elongation factor Tu. *J. Am. Chem. Soc., 131*, 5946-5955.

36. Tocchetti, A., Maffioli, S., Iorio, M., Alt, S., Mazzei, E., Brunati, C., Sosio, M., & Donadio, S. (2013). Capturing linear intermediates and C-terminal variants during maturation of the thiopeptide GE2270. *Chem. Biol., 20*, 1067-1077.

37. Tocchetti, A., Maffioli, S., Iorio, M., Alt, S., Mazzei, E., Brunati, C., Sosio, M., & Donadio, S. (2013). Capturing Linear Intermediates and C-Terminal Variants during Maturation of the Thiopeptide GE2270. *Chem. Biol., 20*, 1067-1077.

38. Jang, S., & Imlay, J. A. (2010). Hydrogen peroxide inactivates the Escherichia coli Isc iron-sulphur assembly system, and OxyR induces the Suf system to compensate. *Mol. Microbiol., 78*, 1448-1467.

39. Camps, F., Coll, J., Guerrero, A., Guitart, J., & Riba, M. (1982). Reduction of conjugated dienoic carboxylic-acids and esters with sodium dithionite. *Chem. Lett.*, 715-718.

40. Camps, F., Coll, J., & Guitart, J. (1986). Regiospecific reduction of unsaturated conjugated ketones with sodium dithionite under phase-transfer catalysis. *Tetrahedron, 42*, 4603-4609.

41. Louisandre, O., & Gelbard, G. (1985). Exclusive 1-4 reduction of conjugated ketones by sodium dithionite. *Tetrahedron Lett., 26*, 831-832.

42. 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., Goeransson, 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., Mueller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J. T., Rebuffat, S., Ross, R. P., Sahl, H.-G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Suessmuth, R. D., Tagg, J. R., Tang, G.-L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., & van der Donk, W. A. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep., 30*, 108-160.

43. Burkhart, B. J., Hudson, G. A., Dunbar, K. L., & Mitchell, D. A. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol., 11*, 564-570.

44. Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., & van der Donk, W. A. (2015). In vitro biosynthesis of the core scaffold of the thiopeptide thiomuracin. *J. Am. Chem. Soc., 137*, 16012-16015.

45. Zhang, Z., Hudson, G. A., Mahanta, N., Tietz, J. I., van der Donk, W. A., & Mitchell, D. A. (2016). Biosynthetic timing and substrate specificity for the thiopeptide thiomuracin. *J. Am. Chem. Soc., 138*, 15511-15514.

46. Just-Baringo, X., Albericio, F., & Álvarez, M. (2014). Thiopeptide antibiotics: retrospective and recent advances. *Mar. Drugs, 12*, 317.

47. Young, T. S., Dorrestein, P. C., & Walsh, C. T. (2012). Codon randomization for rapid exploration of chemical space in thiopeptide antibiotic variants. *Chem. Biol., 19*, 1600-1610.

48. Mahanta, N., Zhang, Z., Hudson, G. A., van der Donk, W. A., & Mitchell, D. A. (2017). Reconstitution and substrate specificity of the radical *S*-adenosyl-methionine thiazole *C*-methyltransferase in thiomuracin biosynthesis. *J. Am. Chem. Soc., 139*, 4310-4313.

49. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res., 29*, 1097-1106.

50. Grove, T. L., Livada, J., Schwalm, E. L., Green, M. T., Booker, S. J., & Silakov, A. (2013). A substrate radical intermediate in catalysis by the antibiotic resistance protein Cfr. *Nat. Chem. Biol., 9*, 422-427.

51. van der Donk, W. A. (2006). Rings, radicals, and regeneration: the early years of a bioorganic laboratory. *J. Org. Chem., 71*, 9561-9571.

52. Werner, W. J., Allen, K. D., Hu, K., Helms, G. L., Chen, B. S., & Wang, S. C. (2011). In vitro phosphinate methylation by PhpK from *Kitasatospora phosalacinea*. *Biochemistry, 50*, 8986-8988.

53. Pierre, S., Guillot, A., Benjdia, A., Sandstrom, C., Langella, P., & Berteau, O. (2012). Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes. *Nat. Chem. Biol., 8*, 957-959.

54. Marous, D. R., Lloyd, E. P., Buller, A. R., Moshos, K. A., Grove, T. L., Blaszczyk, A. J., Booker, S. J., & Townsend, C. A. (2015). Consecutive radical S-adenosylmethionine methylations form the ethyl side chain in thienamycin biosynthesis. *Proc. Natl. Acad. Sci. USA, 112*, 10354-10358.

55. Parent, A., Guillot, A., Benjdia, A., Chartier, G., Leprince, J., & Berteau, O. (2016). The B12-radical SAM enzyme PoyC catalyzes valine C-beta-methylation during polytheonamide biosynthesis. *J. Am. Chem. Soc., 138*, 15515-15518.

56. Zhou, S., Alkhalaf, L. M., de Los Santos, E. L., & Challis, G. L. (2016). Mechanistic insights into class B radical-S-adenosylmethionine methylases: ubiquitous tailoring enzymes in natural product biosynthesis. *Curr. Opin. Chem. Biol., 35*, 73-79.

57. Wang, Y., Schnell, B., Baumann, S., Muller, R., & Begley, T. P. (2017). Biosynthesis of branched alkoxy groups: iterative methyl group alkylation by a cobalamin-dependent radical SAM enzyme. *J. Am. Chem. Soc., 139*, 1742-1745.

58. Bridwell-Rabb, J., Zhong, A., Sun, H. G., Drennan, C. L., & Liu, H. W. (2017). A B12-dependent radical SAM enzyme involved in oxetanocin A biosynthesis. *Nature, 544*, 322-326.

59. Allen, K. D., Xu, H., & White, R. H. (2014). Identification of a unique radical S-adenosylmethionine methylase likely involved in methanopterin biosynthesis in *Methanocaldococcus jannaschii*. *J. Bacteriol., 196*, 3315-3323.

60. Broderick, J. B., Duffus, B. R., Duschene, K. S., & Shepard, E. M. (2014). Radical S-adenosylmethionine enzymes. *Chem. Rev., 114*, 4229-4317.

61. Ding, W., Li, Y., Zhao, J., Ji, X., Mo, T., Qianzhu, H., Tu, T., Deng, Z., Yu, Y., Chen, F., & Zhang, Q. (2017). The catalytic mechanism of the class C radical S-adenosylmethionine methyltransferase NosN. *Angew. Chem. Int. Ed., 56*, 3857-3861.

62. LaMattina, J. W., Wang, B., Badding, E. D., Gadsby, L. K., Grove, T. L., & Booker, S. J. (2017). NosN, a Radical S-adenosylmethionine methylase, catalyzes both C1 transfer and formation of the ester linkage of the side-ring system during the biosynthesis of nosiheptide. *J. Am. Chem. Soc., 139*, 17438-17445.

63. Shisler, K. A., & Broderick, J. B. (2014). Glycyl radical activating enzymes: structure, mechanism, and substrate interactions. *Arch. Biochem. Biophys., 546*, 64-71.

64. Kamat, S. S., & Raushel, F. M. (2015). PhnJ – A novel radical SAM enzyme from the C–P lyase complex. *Perspectives in Science, 4*, 32-37.

65. Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., & van der Donk, W. A. (2015). In vitro biosynthesis of the core scaffold of the thiopeptide thiomuracin. *J. Am. Chem. Soc., 137*, 16012-16015.

66. Layer, G., Moser, J., Heinz, D. W., Jahn, D., & Schubert, W. D. (2003). Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J., 22*, 6214-6224.

67. Layer, G., Grage, K., Teschner, T., Schunemann, V., Breckau, D., Masoumi, A., Jahn, M., Heathcote, P., Trautwein, A. X., & Jahn, D. (2005). Radical S-adenosylmethionine enzyme coproporphyrinogen III oxidase HemN: functional features of the [4Fe-4S] cluster and the two bound S-adenosyl-L-methionines. *J. Biol. Chem., 280*, 29038-29046.

68. Smith, D. M., Buckel, W., & Zipse, H. (2003). Deprotonation of enoxy radicals: theoretical validation of a 50-year-old mechanistic proposal. *Angew. Chem. Int. Ed. Engl., 42*, 1867-1870.

69. Miller, L. M., Chatterjee, C., van der Donk, W. A., & Kelleher, N. L. (2006). The dehydration activity of lacticin 481 synthetase is highly processive. *J. Am. Chem. Soc., 128*, 1420-1421.

70. Markham, G. D., DeParasis, J., & Gatmaitan, J. (1984). The sequence of metK, the structural gene for S-adenosylmethionine synthetase in *Escherichia coli*. *J. Biol. Chem., 259*, 14505-14507.

CHAPTER 5: CHARACTERIZATION OF THE ATP AND TRNA DEPENDENT PEPTIDYL TRANSFERASE INVOLVED IN THE BIOSYNTHESIS OF 3-THIAGLUTAMATE

5.1 Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a large and diverse class of natural products (1). Included in this family of natural products is the lanthipeptide nisin, which has been used for decades against food-borne pathogens with no evidence of stable resistance (2). During the maturation of the compound, Ser and Thr residues are first glutamylated by the LanB dehydratase, NisB, in a glutamyl-tRNA dependent mechanism (3, 4). Subsequently, the glutamate is eliminated to generate dehydroamino acids (**Figure 5.1A**). Our recent studies in thiopeptide biosynthesis (4), as well as studies by Onaka's laboratory on goadsporin maturation (5) have further revealed that the glutamylation and elimination steps for dehydroamino acids formation can also be carried out by a split LanB dehydratase enzyme. Our survey of the available genomes revealed a large number of genes encoding LanB-like proteins lacking the putative elimination domain. In these clusters, the elimination domain is absent entirely from the biosynthetic gene cluster and the entire genome. This observation suggests that these stand-alone glutamylation domains of LanB dehydratases (sLanB) might add an amino acid in a tRNA-dependent fashion without subsequent elimination.

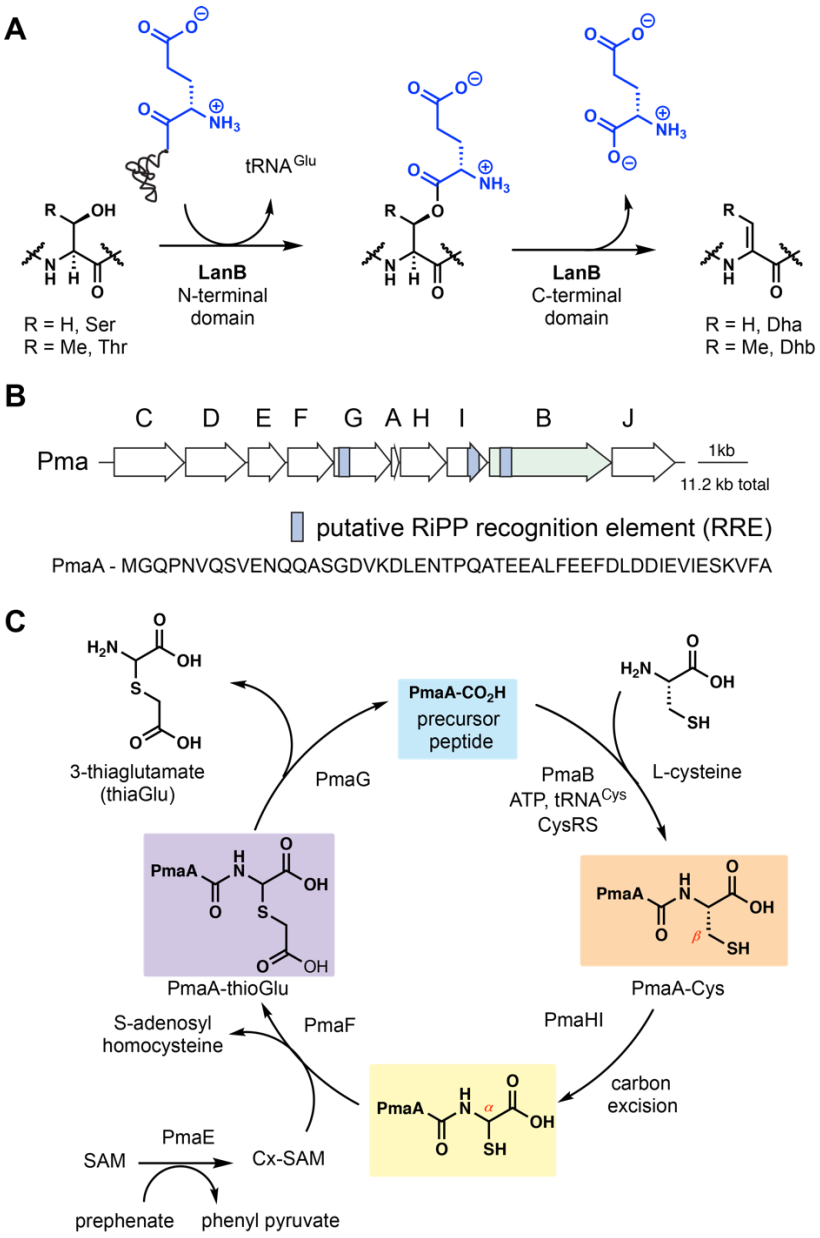


Figure 5.1 (**A**) LanB enzymes glutamylate Ser/Thr residues and subsequently eliminate the glutamate to form dehydroamino acids. sLanB proteins lack the elimination domain. Dha, dehydroalanine; Dhb, dehydrobutyrine. (**B**) Biosynthetic gene cluster in *P. syringae* encoding a sLanB. (**C**) Inferred biosynthetic pathway towards 3-thiaglutamate.

In *Pseudomonas syringae* such a sLanB (PmaB) is encoded near an open reading frame of a 50-amino acid peptide (PmaA; **Figure 5.1B**). Through the joint effort of Dr. Michael Funk and

Dr. Chi Ting, reconstitution of the activity of enzymes in the neighborhood of the PmaA and PmaB revealed a biosynthetic pathway where a small-molecule is biosynthesized with a ribosomally synthesized small peptide serving as a catalytic scaffold (**Figure 5.1C)**. First, a cysteine residue is transferred from $tRNA^{Cys}$ to the C-terminus of the peptide, a reaction that replaces ribosomal protein synthesis. Then, a translocation of the cysteine thiol from the β-carbon to the α-carbon is catalyzed by an oxidase that removes the β-carbon as formate. The resulting thiol is carboxymethylated and proteolysis releases 3-thiaglutamate, in the process regenerating the peptide scaffold. This gene cluster is intriguing in that the biosynthetic pathway is fundamentally different from other known pathways. Unlike RiPPs, the final product in this pathway is not encoded in the genome since the pathway does not contain a gene-encoded core peptide (6). But unlike non-ribosomal peptides, or other ribosome-independent peptide ligases (7-9), the current pathway utilizes a ribosomal scaffold and aminoacyl-tRNA to extend a peptide sequence. sLanB-encoding genes are abundant in bacterial genomes, with some clusters encoding multiple such proteins. Thus, this mechanism of peptide bond formation appears to be widespread and is likely driven by the catalytic use of leader peptides as scaffolds for small molecule production. Thus, it is very appealing to characterize this aminoacyl-tRNA dependent peptide ligase.

Previous work by Dr. Michael Funk has shown that co-expression of His_6-PmaA and PmaB in *Escherichia coli* and subsequent affinity purification of the peptide resulted in an increase in mass by 103 Da (**Figure 5.2A**). Further high resolution tandem mass spectrometry analysis of the peptide suggested that the adduct was attached to the C-terminal alanine (**Figure 5.2B**), through a linkage of a terminal amide. Michael was also able to demonstrate that if PmaB catalyzes the reaction in buffer made with $H_2{}^{18}O$, the product contains one ^{18}O atom (**Figure 5.2C**). In this

chapter, I present a detailed biochemical and mechanistic investigation of the aminoacyl-tRNA dependent peptide ligase PmaB.

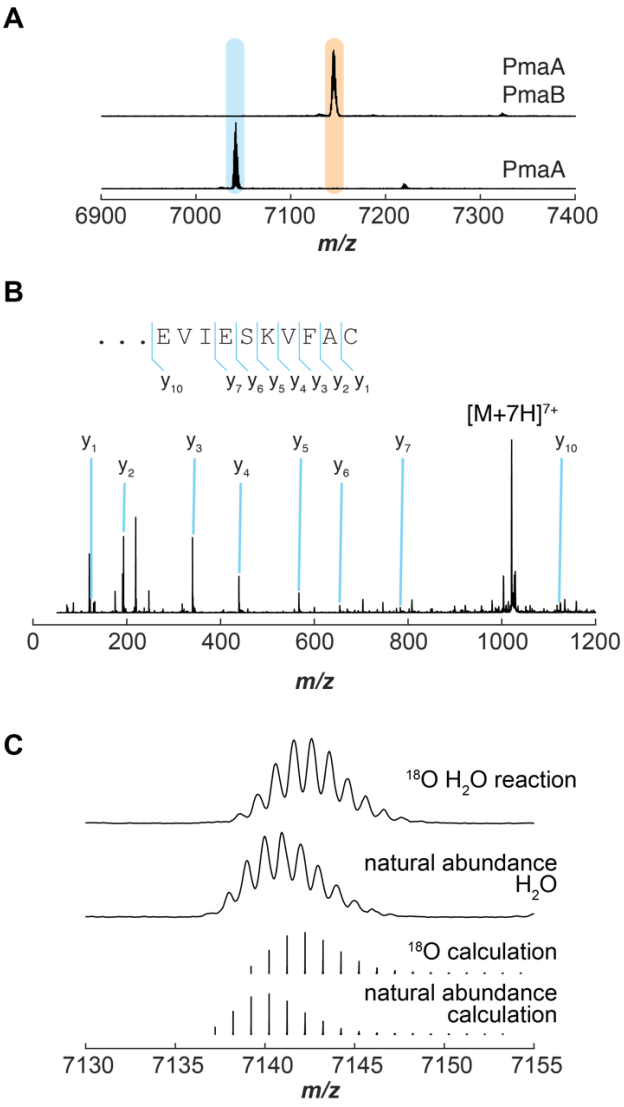


Figure 5.2 (**A**) Matrix-assisted laser desorption ionization with time-of-flight (MALDI-TOF) mass spectra of PmaA coexpressed with PmaB. (**B**) Analysis of the PmaB product by tandem electrospray ionization (ESI) mass spectrometry. (**C**) MALDI-TOF MS of PmaA-Cys prepared by *in vitro* PmaB reaction in 85% $H_2{}^{18}O$ results in ^{18}O enrichment in the product.

5.2 Results and discussion

5.2.1 Biochemical characterization of PmaB

I first identified the minimal substrate for PmaB. For this purpose, the PmaA precursor peptide was systematically truncated from the N-terminus, and the remaining peptides were used in the in vitro cysteine addition reaction. Matrix-assisted laser desorption ionization with time-of-flight (MALDI-TOF) mass analysis of the reaction suggested that removal of the first 30 out of 50 amino acids at the N-terminus of the PmaA peptide does not affect PmaB activity (**Figure 5.3A**). However, when PmaA was treated with endo-proteinase AspN to remove the N-terminal 39 residues, the resulting peptide could be partially converted (**Figure 5.3B**). When one more amino acid residue was retained on the peptide, then complete activity of PmaB was detected again (**Figure 5.3C**), suggesting a PmaA 12mer possesses enough information for PmaB recognition. This is find suggests that most part of the peptide is dispensable for the amino acyl ligation reaction, and further implies a division of precursor peptide recognition region amongst the various biosynthetic enzymes to ensure efficient downstream modifications.

We next evaluated the substrate specificity of PmaB towards the last amino acid of PmaA. A series of PmaA variants were prepared in which the last Ala residue was systematically substituted with seven other amino acids, ranging from polar, charged to aromatic (**Figure 5.4A**). These PmaA variants were co-expressed with PmaB in *E. coli*. All the variants were accepted by PmaB except for the charged Glu or Lys residues. An interesting observation was that, as long as the charge is blocked, the PmaA variant was converted into the cysteinylated product (see product formation using the A50N variant). However, if the last Ala residue is removed from the PmaA peptide, no product formation would be detected (**Figure 5.4B**). These data suggest that PmaB is relatively permissive towards the change to the last amino acid of its peptide substrate, and the

active site of the protein is large enough to accommodate the different size of amino acid side chains. If this last amino acid is removed from the substrate PmaA, the C-terminal carboxylate group may not be able to reach into the active site for cysteine addition.

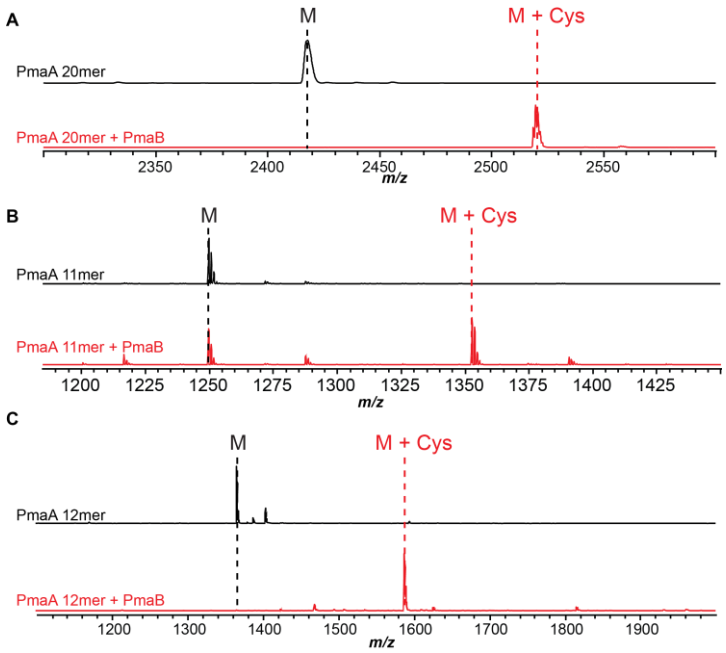


Figure 5.3 Identification of the minimal substrate for PmaB. MALDI-TOF mass spectra of (**A**) PmaA N-terminal 20-mer and its product for in vitro cysteine addition reaction with PmaB; (**B**) PmaA 11-mer and its product for in vitro cysteine addition reaction with PmaB; (**C**) PmaA 12-mer and its product for in vitro cysteine addition reaction with PmaB.

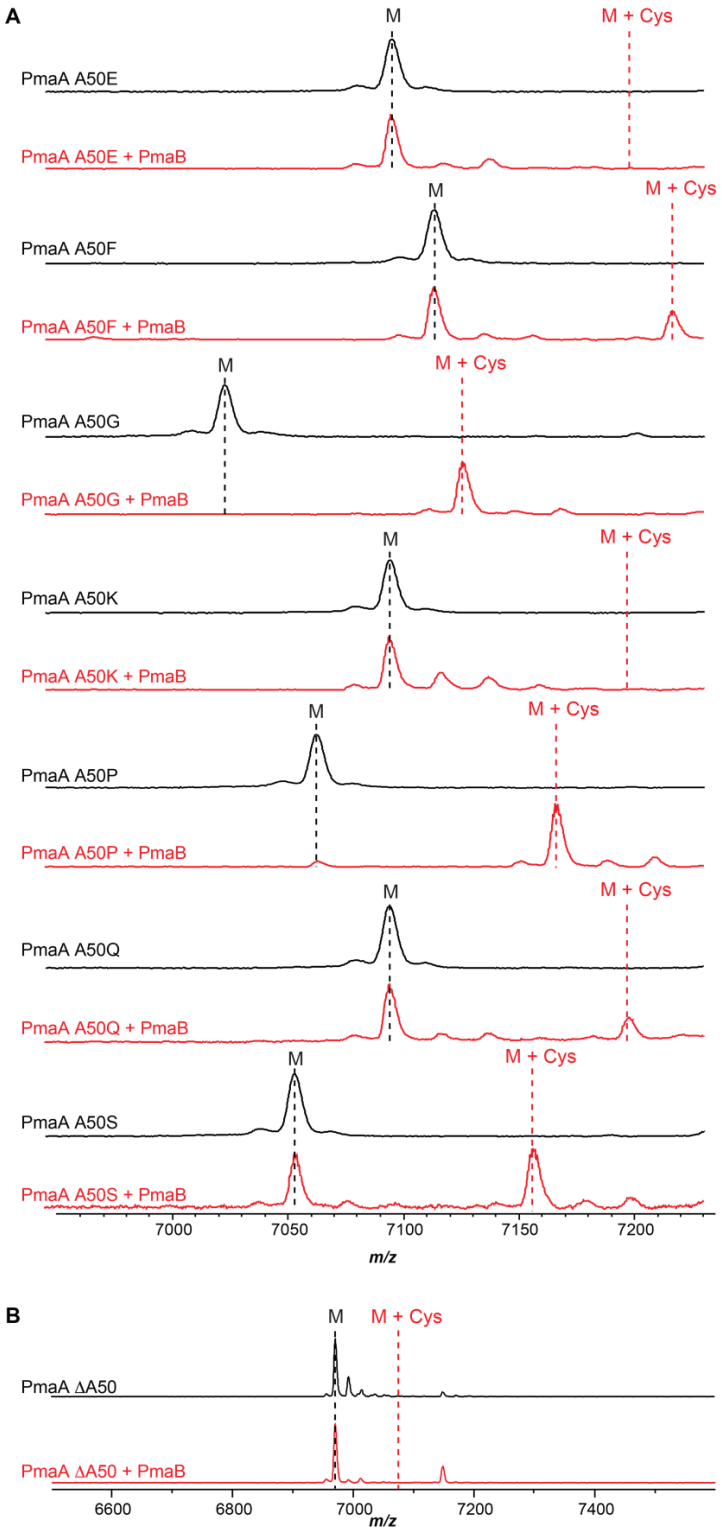


Figure 5.4 PmaB is relatively promiscuous towards a change of the last amino acid of PmaA. MALDI-TOF mass spectra of (**A**) PmaA A50 variants and their co-expression products with PmaB in *E. coli*. (**B**) PmaA ΔA50 variant and its co-expression product with PmaB in *E. coli*.

The previous in vitro reactions were all performed using *E. coli* cell extract, and the complicated nature of the cell extract may hamper mechanistic study of PmaB. As lanthipeptide dehydratases glutamylate Ser and Thr residues in a glutamyl-tRNA dependent manner (3, 4), the PmaB was anticipated to perform the cysteinylation reaction by utilizing cysteinyl-tRNA^{Cys} as a substrate. Hence, the *P. syringae* tRNA^{Cys} was prepared by in vitro transcription. Reaction of PmaA with Cys, ATP, *P. syringae* tRNA^{Cys} and *P. syringae* GluRS resulted in clean conversion of PmaA. A previous group member Dr. Rebecca Splain showed in another sLanB system that the sLanB enzyme requires the presence of ATP for activity. Thus, a cysteinylation assay with purified cysteinyl-tRNA^{Cys} was performed with or without ATP. Subsequent MALDI-TOF-MS analysis confirmed that the amino acid ligation reaction catalyzed by PmaB is indeed ATP dependent (**Figure 5.5**).

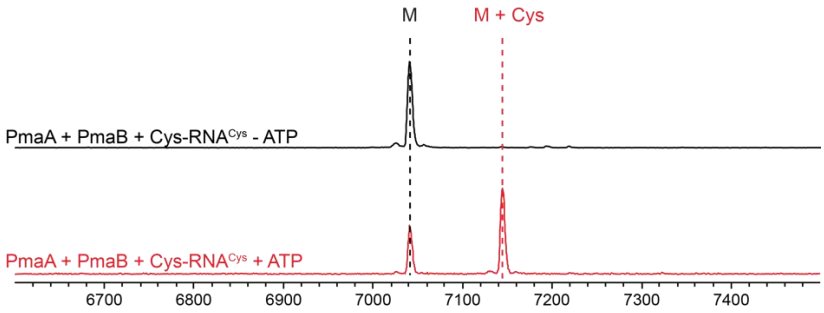


Figure 5.5 In vitro cysteinylation reaction by PmaB requires the presence of ATP. MALDI-TOF mass spectra of in vitro cysteine addition reaction with purified cysteinyl-tRNA^{Cys} in the absence and presence of ATP.

5.2.2 Mechanistic study of the PmaB catalyzed amino acyl ligation reaction

To investigate whether PmaB adopts a similar ping-pong mechanism involving the formation of acyl-enzyme intermediate as the cyclodipeptide synthases (10), we first analyzed the sequence of PmaB and its orthologs from various sLanB gene clusters and identified four

conserved Ser or Thr residues (Ser 11, Ser 180, Thr 540, and Ser 780; **Figure 5.6**). PmaB variants were generated in which these Ser or Thr residues were replaced by Ala and co-expressed with PmaA in *E. coli*. All of the PmaB mutants remained active (**Figure 5.7**), suggesting that no conserved Ser or Thr residues in sLanB families anchor an amino acid to generate the corresponding acyl-enzyme intermediate. Further mutagenesis of the conserved Tyr in PmaB to Phe ruled out the possibility of a Tyr to serve the purpose of forming an acyl-enzyme intermediate (**Table 5.1**), and no conserved Cys residue was identified during the alignment. These data indicate that PmaB does not utilize a ping-pong mechanism to generate an acyl-enzyme intermediate for catalysis.

Further multiple sequence alignment identified sixteen more residues that are conserved. Each of the residues was replaced with Ala and co-expressed with PmaA in *E. coli* (**Table 5.1**). A set of nine residues (R10A, R173A, K177A, D179A, D542A, H566A, K752A, E801A and R803A) was identified as possibly involved in substrate binding or catalysis. To confirm that the diminished activity did not result from improper folding of the enzyme, the nine PmaB variants were individually expressed and purified from *E. coli* (**Figure 5.8**), and tested for in vitro activity (**Figure 5.9**). The consistency between the in vivo and in vitro experimental data further corroborated important roles of the seven residues in catalysis.

```
WP_010898202.1      MLSVKQSSINEDHNNWDFFGHFVLRSTGFPYEWMKEL-----------------------
WP_095098125.1      ---------------MRCADYFWLRSAGFPADQLMQATSFPSLPAF----TSLMALHQLR
WP_048368865.1      ---------------MESSQYFWLRSTGFPIHYLTDLGRLSELPACRAFEQDFRAFQALK
PmaB                ---------------MESSHYFWLRSTGFAVHHLTRLGKMAELPLLKDFETDYRSLNTLR
WP_117166554.1      ---------------MESSLYFWLRSTGFPIHHLTQLGKMDELPLLKDFEDSYRCVAVLR
                                          :* ***:**.   :

WP_010898202.1      -----------------------KMQQTFDLIFQN------------AKWEQVESKFN
WP_095098125.1      GRLLQQFEQQMAVVGEVQCRKFARKLAAQQAVSVSDLPLPLRDVLQQPLDEWHNVNAKIV
WP_048368865.1      AMMLEEANSHSLLT----CRKLIRKLNEGQSLLLSDLPQALREPLNTAVLKWNDLLKRLV
PmaB                DSLLEKSISHSSQA----CRKLIRKLNENLPLQVSDLPEALRDDAVDELAHWNERLGRLK
WP_117166554.1      SKLLQKSMSHSAQA----SRKLIRKLNENLPLQVSDLPEALRESSREDLERWNEQLARLA
                                       *:     : ..:              * :   .:
```

Figure 5.6 (cont.)

Figure 5.6 (cont.)

```
WP_010898202.1      FSKPMLLHVPISSGKHTPRIVIDDVVYQRERWFFYTKQLVDLFHQLQGPLLLKKVEEWRQ
WP_095098125.1      AGDVLPRHEP-------RRIRVNRTLYKRRTWAFNAGGWPEAVS--DEFMAFIQWRDWQQ
WP_048368865.1      ADSAVAKHEY-------RRIWLGRTLYKRASWLFTREQLPEPKGTIEELEYTMQWRAWAV
PmaB                AGSVVAKNDR-------RRIGLGKTLYKRASWWFSPEQLPCSEFSVDKLDDVLAWRAWAV
WP_117166554.1      AGSAVARGDR-------RRIWLGKTLYKRASWWFTQEQFPCPEFSVDKLDDTLAWRAWAQ
                      .. :             ** :. .:*:*  * *           :         *

WP_010898202.1      AEGIPEVVYIKGSDVRKPYWVDFKNYFSLELMQQILLENNEITIEEMLPDPHHLWLKSRK
WP_095098125.1      RQALPRWVFIKCDSEPKPLFIDFDNPLSLDALATALKKARVIQVSEMLPTPDELWFNDAR
WP_048368865.1      SQGLPRYAFIKIDTEPKPLFLDFDNPLSLDGIINALKSAGHVKFSEMRPGPDELWLEDAR
PmaB                EHGLPRYVFAKIDIEPKPIFIDFDNPLSLDGVSNSMKKAGHVKFSEMCPAPDQLWLEETR
WP_117166554.1      EHGLPRYVFAKIDIEPKPIFIDFDNPLSFDGVSNSMKKATHVKFSEMCPSPDELWLEEAR
                      .:*  .: * .   ** ::**.* :*:: :   : .   : ..** * *  **::. .

WP_010898202.1      GSHSCELRMSVYKLGIKEVSEHA
WP_095098125.1      GRVCCEIRTTFSPIKQETTENAE
WP_048368865.1      GRFCCEIRTTFSINKALAT----
PmaB                GHFCCEIRTTFRDNGVTRDE---
WP_117166554.1      GRFCCEIRTTFTSDGVVPDEQ--
                    *  .**:* :.
```

Figure 5.6 Sequence alignment of PmaB homologs (WP_010898202.1 is from genome of *Bacillus halodurans*; WP_095098125.1 is from genome of *Serratia ficaria*; WP_048368865.1 is from genome of *Pseudomonas helleri*; WP_117166554.1 is from genome of *Pseudomonas cichorii*). Shown in red are residues replaced with alanine for enzymatic activity assessment in this study.

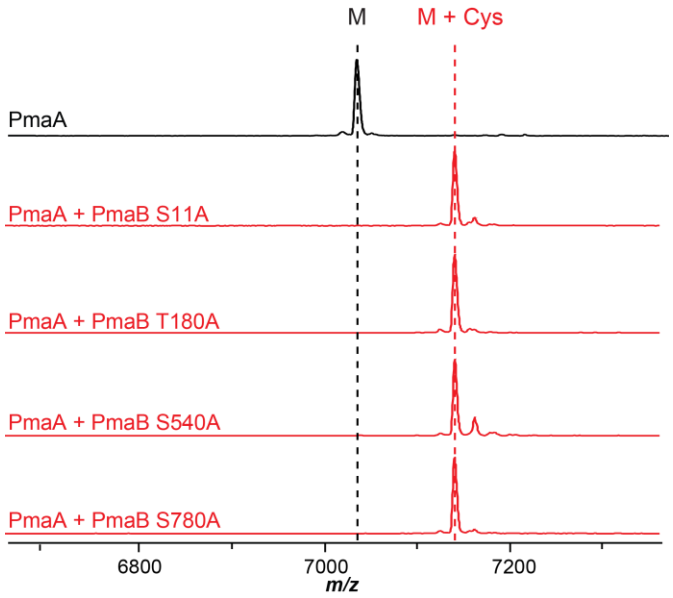


Figure 5.7 MALDI-TOF mass spectra of PmaA co-expressed in *E. coli* with PmaB variants in which selective Ser or Thr residues were replaced with Ala.

Table 5.1 Activity of alanine-substituted PmaB proteins. Sixteen positions of PmaB were targeted for alanine replacement by site-directed mutagenesis using the alignment of **Figure 5.6**. These proteins were assayed for enzymatic activity by co-expressing with PmaA and analyzed using a MALDI-TOF-MS. +++ indicates enzyme activity roughly equal to wild-type (full conversion to product); ++ indicates modestly reduced enzyme activity (~50% conversion to product); + indicates severely reduced enzyme activity (low but detectable product formation); - indicates no detectable enzyme activity (no observed product formation).

PmaB Protein	Relative Product	PmaB Protein	Relative Product
Wild-type	+++	D404A	+++
R10A	-	**D542A**	-
K161A	+++	**H566A**	++
Y170F	+++	R704A	+++
R173A	-	W707A	+++
K177A	-	K746A	+++
D179A	++	**K752A**	+
R397A	+++	**E801A**	-
E403A	+++	**R803A**	-

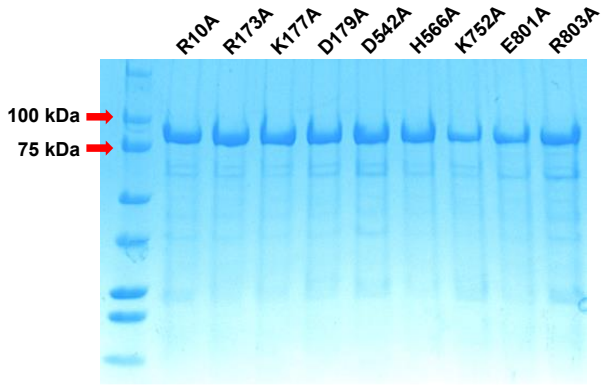


Figure 5.8 SDS-PAGE analysis of PmaB variants which have diminished activity. All proteins were expressed as N-terminus His_6-tag fusions.

The requirement for ATP by PmaB seems counterintuitive since cysteinyl-$tRNA^{Cys}$ is a molecule already at high energy state with an activated amino acyl-ester, and it alone would be sufficient to catalyze the reaction. To further probe the role of ATP in the reaction, ^{31}P NMR

analysis was performed using purified cysteinyl-tRNA^{Cys} and ATP. Phosphate and ADP were detected as the products of the reaction (**Figure 5.10**), suggesting the presence of a phosphorylation step for to activate the substrate. Quenching the assay by adding NH_2OH into the reaction mixture suggests that it is the C-terminal carboxylate of PmaA that is activated, as a PmaA-NH_2OH adduct was detected by MALDI-TOF-MS (**Figure 5.11A**). The identity of this PmaA-NH_2OH adduct was confirmed by electrospray ionization high resolution (ESI-HR) MS and MS/MS analysis of the endo-proteinase GluC digested peptide (**Figure 5.11B, C)**, and the NH_2OH addition was located at the very C-terminal carboxylate of PmaA.

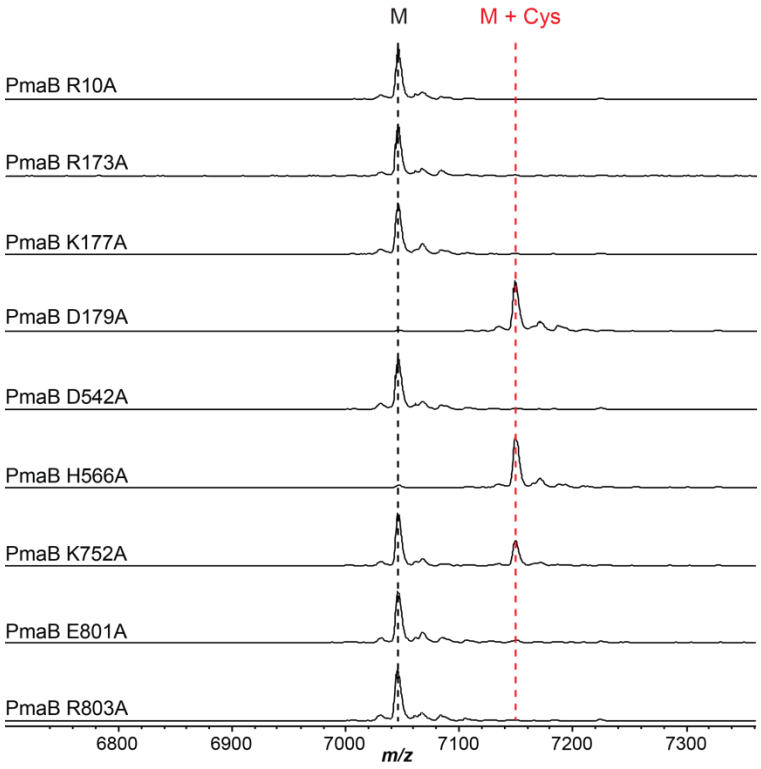


Figure 5.9 MALDI-TOF mass spectra of in vitro cysteinylation reaction of PmaA with PmaB variants which have diminished activity.

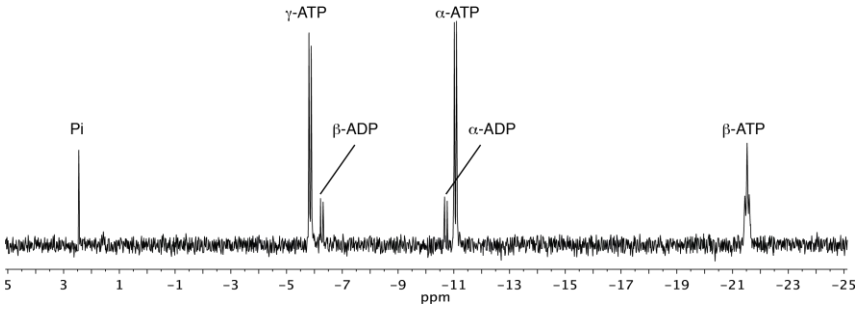


Figure 5.10 ^{31}P NMR analysis shows ADP and phosphate as products of the PmaB reaction.

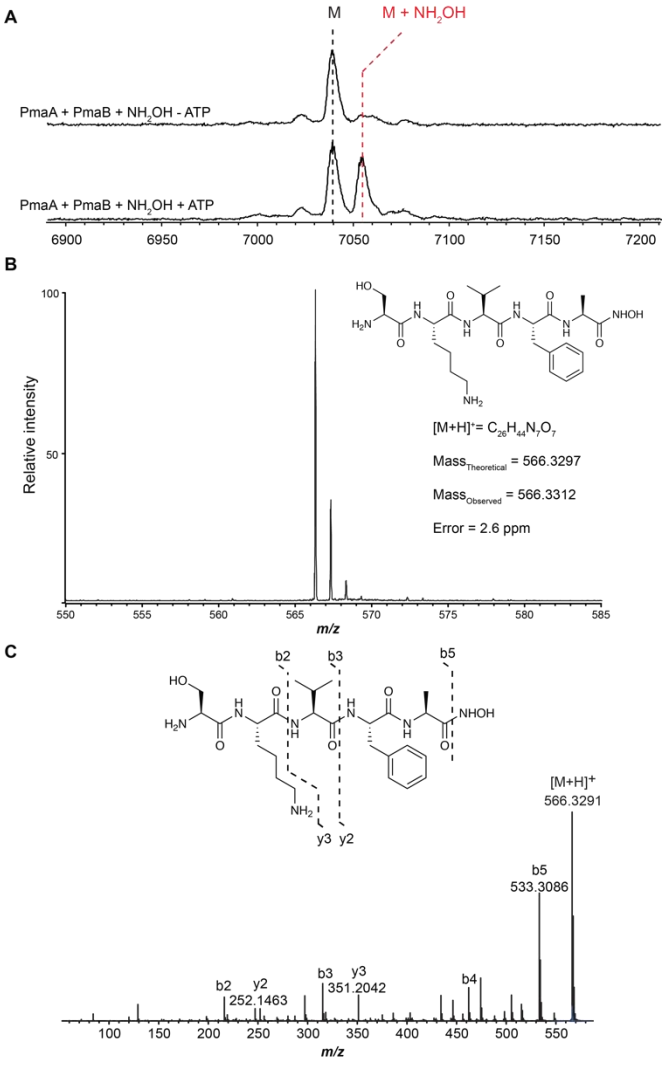


Figure 5.11 NH_2OH interception of C-terminal carboxylate activated PmaA. (**A**) MALDI-TOF mass spectra of NH_2OH treated assay with PmaA and PmaB in the absence and presence of ATP. (**B** and **C**) ESI-HR-MS and MS/MS analysis of the endo-proteinase GluC digested PmaA-NH_2OH adduct.

To provide direct evidence that the C-terminal carboxylate of PmaA was activated by phosphorylation, we labeled one of the C-terminal carboxylate oxygen with an ^{18}O through hydrolyzing a PmaA thioester in $H_2{}^{18}O$ (see experimental methods). According to ESI-MS analysis of the endo-proteinase GluC digested PmaA, ~85% of the C-terminal carboxylate was labeled with an ^{18}O (**Figure 5.12A**). ^{31}P NMR analysis of a multiple turnover assay indicates a clear ^{18}O incorporation into phosphate (^{18}O-Pi), which could only originate from the C-terminal carboxylate of PmaA (**Figure 5.12B**). However, this NMR spectrum still could not confirm direct phosphorylation of the carboxylate, as two mechanisms could explain the ^{18}O incorporation into the generated phosphate (**Figure 5.13**) in a multi-turnover assay.

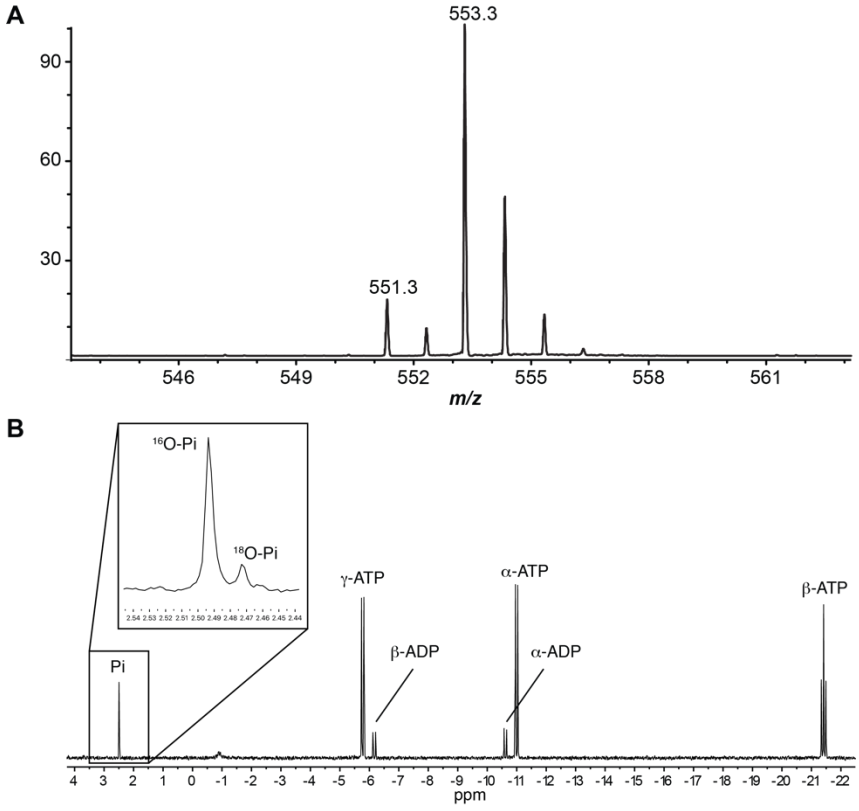


Figure 5.12 (**A**) ESI-MS analysis of the GluC digested ^{18}O labeled PmaA. (**B**) ^{31}P NMR analysis shows ^{18}O incorporation into phosphate when using ^{18}O labeled PmaA for the multiple turnover in vitro cysteinylation reaction.

To test the feasibility of both mechanisms, ^{31}P NMR analysis of a single turnover assay was performed (see experimental methods). If PmaB utilized the proposed mechanism shown in **Figure 5.13A**, no ^{18}O-Pi would be detected in the first turnover. However, ^{18}O-Pi was found to account for ~50% of the ^{16}O-Pi, thus ruling out the possibility of the intermediacy of a PmaA-PmaB anhydride (**Figure 5.13A** and **Figure 5.14**). Integration of the peak corresponding to ADP and phosphate (^{16}O-Pi + ^{18}O-Pi) indicated some discrepancies, as the peak area of labeled and unlabeled phosphate combined is more than that of ADP, and the amount of ^{18}O-Pi to ^{16}O-Pi deviates from the theoretical ratio (~100 : 74; **Figure 5.14**). These discrepancies probably resulted from the prolonged time for ^{31}P NMR recording, and a portion of ADP was hydrolyzed to generate ^{18}O-Pi. If the difference between the peak area of phosphate and ADP is subtracted from that of the ^{16}O-Pi, the ratio of ^{18}O-Pi to ^{16}O-Pi would reach ~71%, which is very close to the theoretical ratio of the single turnover reaction. These data strongly support the mechanism **B** in **Figure 5.13,** in which the C-terminal carboxylate of PmaA is directly activated by phosphorylation during reaction.

The PmaB-catalyzed peptidyl-transfer reaction uses a similar activation strategy as the ATP-grasp domain-containing enzymes from the ADP-forming carboxylate-amine/thiol ligase superfamily (11-14). Recent mechanistic analysis of PglC, a prototypic dual domain phosphor-glycosyl transferase from *Campylobacter concisus*, revealed a ping-pong mechanism with a covalent aspartyl-phosphate intermediate (15). To probe whether PmaB has a similar acyl-phosphate intermediate during catalysis (**Figure 5.15**), a chemical trapping approach was adopted by following protocols used in the study of phosphotransferases (16). In this method, experiments were carried out by reacting PmaB with or without ATP, recovering the protein by precipitation, followed by treatment with sodium borohydride ($NaBH_4$). In this case, the acyl-phosphate

intermediate, if generated, would be reduced to the corresponding protein-based alcohol, which would be amenable to protease digestion and MS analysis. However, LC-ESI-MS analysis of the endo-proteinase chymotrypsin or GluC digested samples did not reveal a modified peptide, suggesting PmaB may not generate an acyl-phosphate intermediate when catalyzing the peptidyl-transfer reaction.

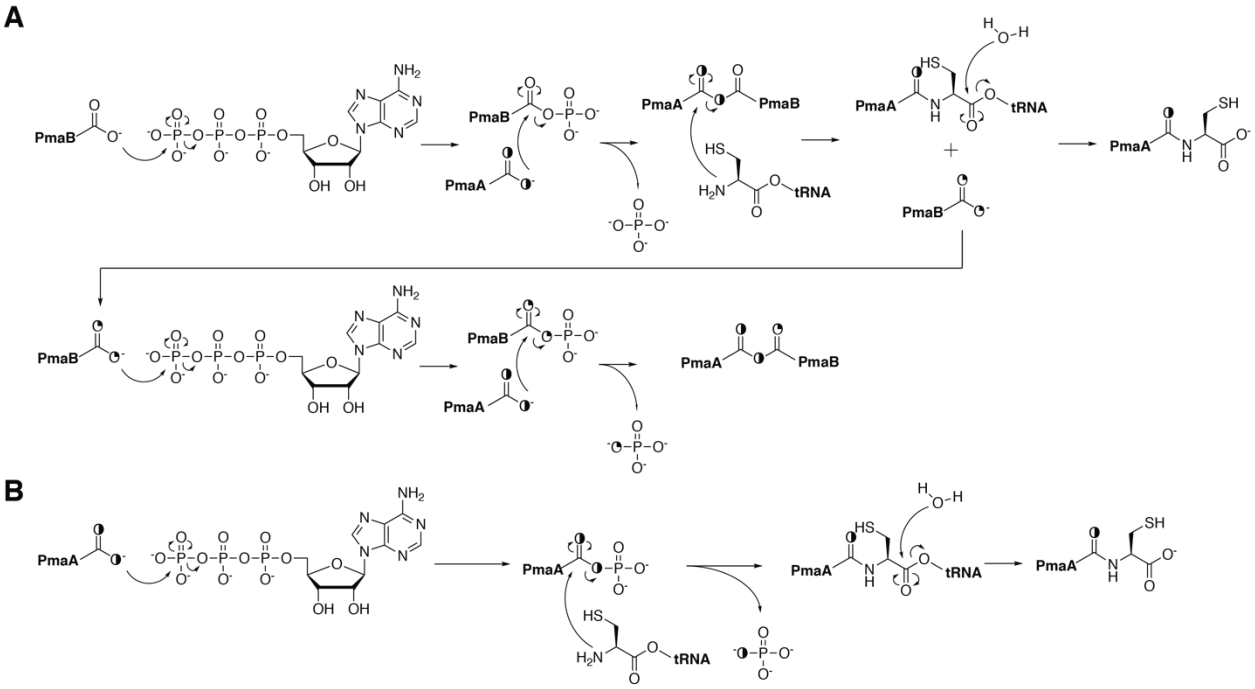


Figure 5.13 Two proposed mechanisms that would result in ^{18}O incorporation into phosphate in multi-turnover reaction. In mechanism **A**, the activation happens on a carboxylic residue on PmaB enzyme, and involves the intermediacy of a PmaA-PmaB anhydride. This would result in ^{18}O-labeled PmaB which could deposit the label into product in subsequent turnover. In mechanism **B**, it is the C-terminal carboxylate of PmaA that is directly activated during reaction.

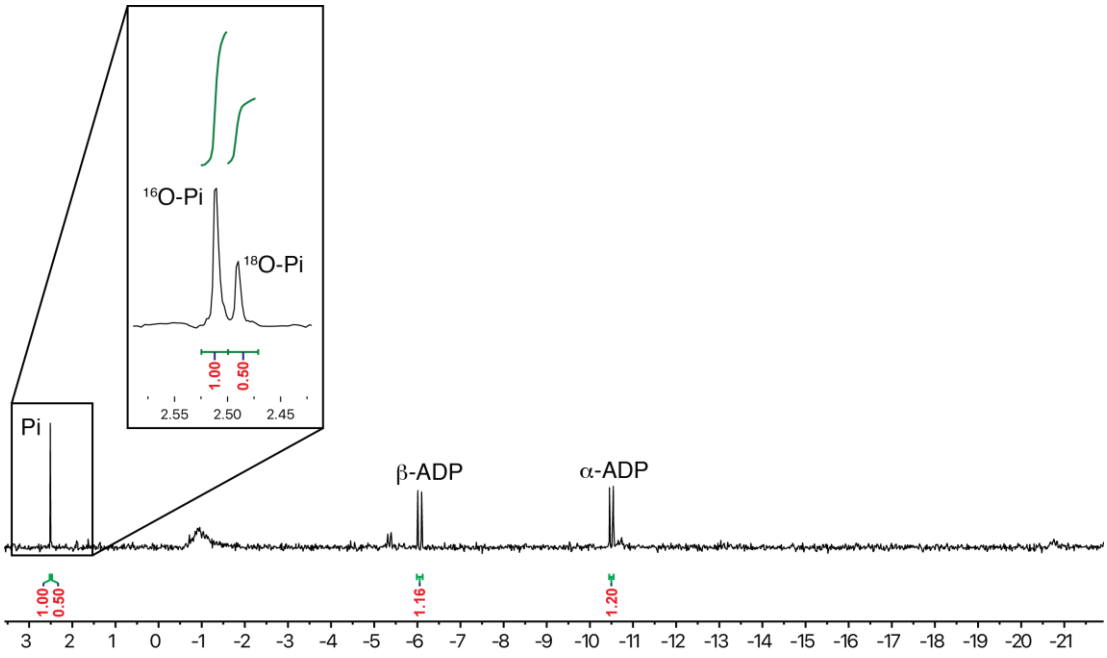


Figure 5.14 ^{31}P NMR analysis of single turnover PmaB reaction using ^{18}O labeled PmaA.

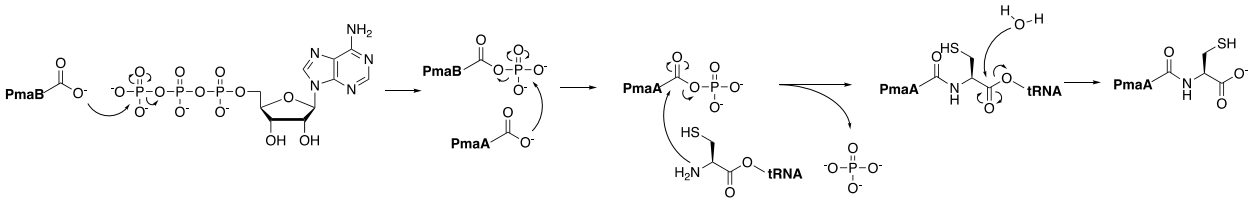


Figure 5.15 A potential ping-pong mechanism with a covalent acyl-phosphate intermediate.

Repeated attempts to crystallize PmaB proved recalcitrant. Thus, a predicted model of PmaB was generated using Phyre2 web portal (**Figure 5.16**) (17). As shown in the PmaB model structure, an Arg/Lys-rich region suitable for interacting with the negatively charged phosphor-diester groups of amino acyl-tRNA was revealed. We were also able to identify a helix-turn-helix structural motif, which is predicted to be the leader peptide binding element (RRE) engaged in precursor peptide recognition. Moreover, the nine residues identified to have diminished activity through Ala substitution assays all clustered at the putative active site, which further corroborates

a role of these residues during catalysis. NH_2OH quenching assays with the PmaB mutants that had no detectable enzyme activity suggests that R10, R173 and K177 do not participate in the phosphorylation of PmaA during catalysis, as mutation of these residues did not affect the NH_2OH addition to the PmaA, whereas residues D542, E801 and R803 could potentially be involved in the phosphorylation step of the reaction (**Figure 5.17**), either through ATP binding or catalyzing the acyl-phosphate formation on PmaA, since these residues did affect NH_2OH addition.

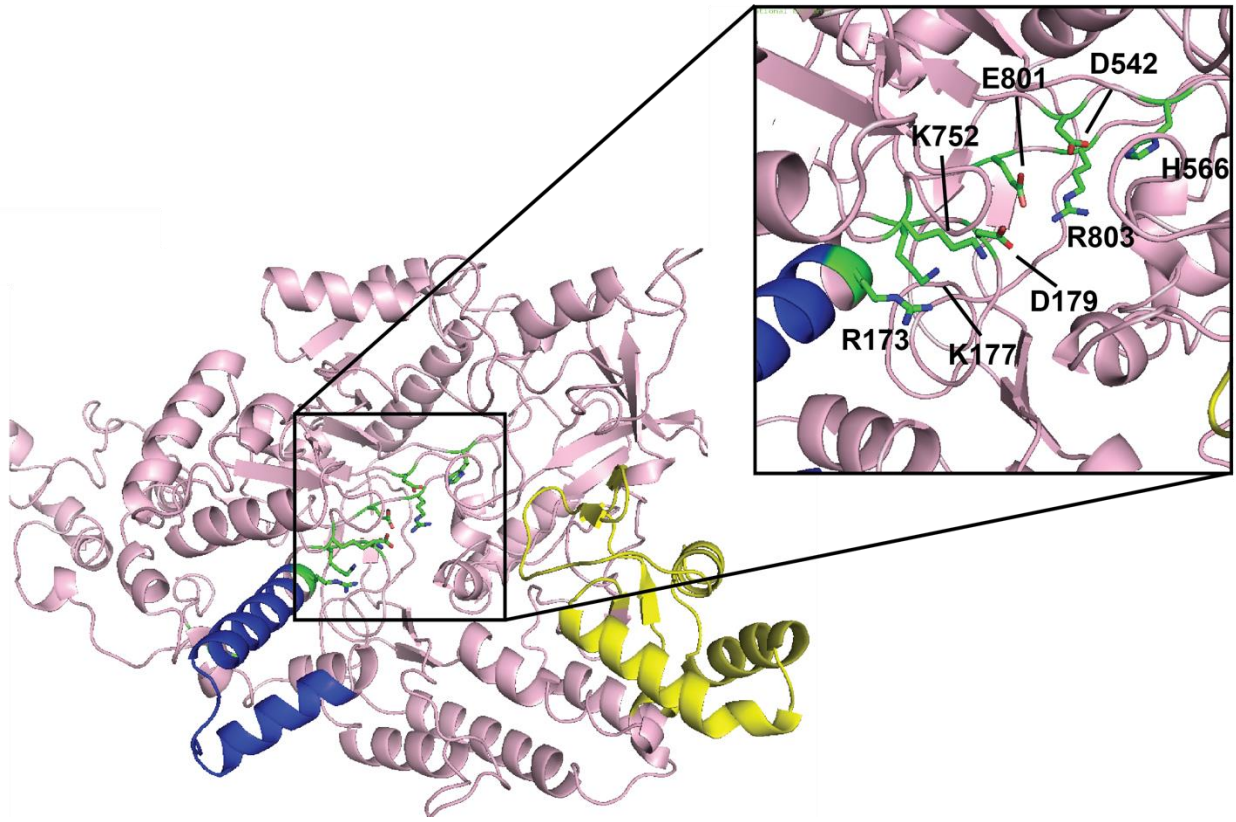


Figure 5.16 Phyre2 model of PmaB. The overall structure is shown in pink. The Arg/Lys-rich region putatively involved in tRNA binding is colored in navy. The RRE motif engaged in precursor peptide recognition is colored in yellow. Highlighted are the residues identified to have diminished activity.

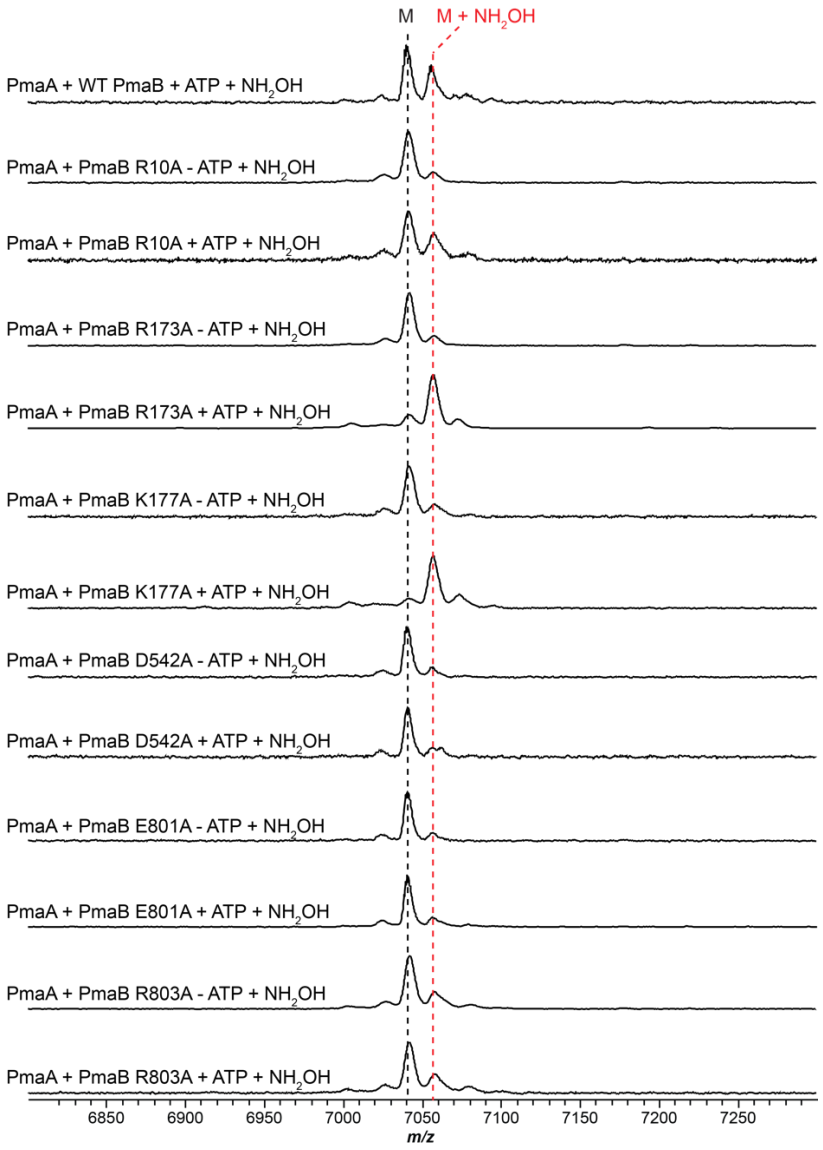


Figure 5.17 MALDI-TOF mass spectra of NH_2OH quenching assays with PmaA and PmaB variants in the absence and presence of ATP.

In summary, we performed a detailed biochemical and mechanistic study of the PmaB enzyme involved in the biosynthesis of 3-thiaglutamate. We found that PmaB catalyzed complete amino acyl ligation reaction using only 12 out of 50 residues of the precursor peptide PmaA. With the exception of charged residues, PmaB is very promiscuous towards the last amino acid of the substrate peptide. We also found that PmaB performs its activity in an ATP and tRNA dependent

manner, which is fundamentally different from other known pathways. ^{18}O labeling assays suggested the role of ATP in the reaction, and provided strong evidence for the direct activation of the PmaA C-terminal carboxylate. Further mutagenesis study of PmaB identified several residues important for catalysis, and their potential function in catalysis was probed by NH_2OH quenching assays and Phyre2 model prediction. Combining all the data we obtained, the mechanism in **Figure 5.14B** is consistent with all of our observations, and is thus favored by this study.

5.3 Experimental methods

General materials and methods

Reagents used for molecular biology experiments were purchased from New England BioLabs (Ipswich, MA), Thermo Fisher Scientific (Waltham, MA), or Gold Biotechnology Inc. (St. Louis, MO). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). *Escherichia coli* DH5α and BL21 (DE3) strains were used for plasmid maintenance and protein overexpression, respectively. Co-expression vector *pRSFDuet-1* was obtained from Novagen. MALDI-TOF-MS analysis was performed using a Bruker UltrafleXtreme MALDI TOF-TOF mass spectrometer (Bruker Daltonics) in reflector positive mode at the University of Illinois School of Chemical Sciences Mass Spectrometry Laboratory. ESI-MS/MS analyses were performed using a SYNAPT ESI quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters). ^{31}P NMR spectra were recorded on an Agilent 600 MHz spectrometer in 10% D_2O. HiTrap columns for Ni-NTA affinity chromatography were purchased from GE Healthcare.

Molecular biology techniques

DNA was prepared using MiniPrep kits (Qiagen) using the manufacturer's instructions from *E. coli* DH10b cells (New England Biolabs) made chemically competent by the KCM method (18). Genomic DNA from *P. syringae* pv. maculicola ES4326 (Pma) was prepared using an UltraClean Microbial DNA isolation kit (Mo Bio Laboratories) according to the manufacturer's instructions; Pma cells were grown in liquid KB medium with vigorous aeration. Plasmids were constructed with type-II restriction enzymes using the New England Biolabs Golden Gate Assembly Tool (http://goldengate.neb.com/editor) or manually designed and prepared by Gibson assembly. For plasmids assembled by Golden Gate Assembly, two primer pairs were designed for PCR to prepare both the gene and the vector with complementary sticky ends for type IIs restriction digest. BsaI and T4 DNA Ligase were used for single step DNA ligation. For plasmids assembled by Gibson Assembly, one primer pair was designed for PCR amplification of the gene with an overhang complementary to the vector. The vector was linearized by restriction digest with NdeI and XhoI and ligated with the gene of interest using NEBuilder® HiFi DNA Assembly Master Mix. Restriction enzymes and PCR polymerases were obtained from New England Biolabs. Genes for each of the *pma* gene cluster products and the *P. syringae* cysteine tRNA synthetase (CysRS) were identified using the Joint Genome Initiative Integrated Microbial Genomes and Microbiomes webtool (http://img.jgi.doe.gov). Primers for cloning and site-directed mutagenesis were obtained from Integrated DNA Technologies. Primers for *P. syringae* $tRNA^{Cys}$ were designed and the dsDNA template for in vitro transcription was prepared as previously described (19). Mutagenesis was accomplished using the QuikChange II Site-Directed Mutagenesis Kit according to manufacturer's instructions. Mutagenesis primers were designed using the QuikChange Primer Design webtool. Primers are listed in **Table 5.2**.

In vitro transcription of *P. syringae* tRNA^{Cys}

Primers for *P. syringae* tRNA^{Cys} were designed according to a previously described method (19). The tRNA^{Glu} dsDNA template was generated from two overlapping synthetic deoxyoligonucleotides (**Table 5.2**). To prepare the dsDNA template for in vitro transcription, 5′ overhangs were filled in using the following conditions: NEB Buffer 2 (1×), primers (4 μM each), dNTP (100 μM each), and DNA polymerase I large (Klenow) fragment (1 U μg^{-1} DNA) in a final volume of 50 μL. The reaction was incubated at 25 °C for 15 min, quenched with EDTA (10 mM) at 75 °C for 25 min, and dsDNA tRNA^{Cys} template was precipitated with cold EtOH overnight. In vitro transcription was performed using a previously described method (20). The transcribed tRNA^{Cys} was then purified by acidic phenol extraction using a previously described method (21) (notebook reference 6/4//2018).

Peptide expression and purification

For peptide expression, *E. coli* BL21(DE3) cells (New England Biolabs) expressing N-terminal His_6-PmaA from *His_6-PmaA-pRSF-Duet-1* were grown with 50 μg/mL kanamycin in autoinduction (AI) media (for 1 L total: 10 g bacto tryptone, 5 g yeast extract, 5 g NaCl, 3 g KH_2PO_4, 6 g Na_2HPO_4) with 1X AI sugar solution containing 0.5% (vol/vol) glycerol, 0.05% (w/vol) glucose, and 0.2% (w/vol) lactose (20 mL of 50X stock)). Cultures were shaken at 37 ºC for 6-7 h following inoculation with 1 mL of saturated culture. Cells were harvested and lysed by sonication. Peptides were purified by immobilized metal affinity chromatography (IMAC). The lysate was clarified by centrifugation at 29,000 rcf and applied to a 5 mL NiNTA-agarose column (GE Healthcare) using a peristaltic pump. The immobilized peptide was washed with 5 column volumes (CV) of 90% lysis buffer (50 mM HEPES, 100 mM NaCl, pH 7.5), 10% elution buffer

(50 mM HEPES, 100 mM NaCl, 500 mM imidazole, pH 7.5; final imidazole concentration in wash: 50 mM) and eluted with 100% elution buffer. The elution fraction was concentrated using a 3 kDa MWCO Amicon spin filter and washed with 10-20 CV of deionized water to remove imidazole. Crude peptide was desalted with a VYDAC® Bioselect C4 cartridge. Peptide elution fractions were lyophilized (notebook reference 11/20/2017).

Protein expression and purification for PmaB and CysRS.

The general expression protocol for the *pma* cluster enzymes PmaB is as follows: *E. coli* BL21 (DE3) cells (50 µL) were electroporated with *His_6-PmaB-pET28a* (50 ng), and cells were plated on LB agar plates supplemented with kanamycin and grown at 37 °C for 12–15 h. A single colony was used to inoculate 20 mL of LB broth supplemented with kanamycin, grown for 12–15 h at 37 °C, and the culture was used to inoculate 2 L of TB media, supplemented with kanamycin, to an OD_{600} of 0.025. Cultures were grown at 37 °C to a final OD_{600} of 0.6–0.8. Protein overexpression was induced by the addition of IPTG to a final concentration of 0.2 mM, and cultures were grown at 18 °C for 18 h. Cells were harvested by centrifugation, collected in 50 mL tubes and frozen in liquid nitrogen. For purification of His_6-PmaB and His_6-CysRS, cells were thawed, resuspended in lysis buffer (50 mM HEPES, 100 mM NaCl, pH 7.5; 30 mL per 10 g wet cell paste) and lysed by treatment with lysozyme (100 µg/mL) and sonication (3 min active time; 1 s pulse, 2 s rest at 60 % max amplitude using a 1 cm tip). Proteins were purified by immobilized metal affinity chromatography (IMAC). Phenylmethylsulfonyl fluoride (PMSF) was added to prevent degradation by serine proteases and 1 mM TCEP was included to maintain reduced thiols. The eluate was concentrated to 2.5 mL in a 30 kDa MWCO centrifuge filter and desalted on a PD-

10 size-exclusion column (GE Healthcare Life Sciences). Protein was separated into aliquots and stored at –70 ºC. Protein purity was judged by SDS-PAGE (notebook reference 4/4/2018).

HPLC peptide purification and LC-MS

Following IMAC and C4 desalting, PmaA was purified using an Agilent 1200 HPLC. HPLC buffers were 0.1% TFA in water (HPLC buffer A) and 0.1% TFA in acetonitrile (HPLC buffer B). The desalted material was lyophilized, resuspended in buffer A, and applied to a Phenomenenx Luna C18 column equilibrated with 95% HPLC buffer A, 5% HPLC buffer B. A linear gradient to 100% HPLC buffer B was run over 20 min at 0.1 mL min^{-1}. PmaA eluted near the end of the gradient. PmaA-Cys behaved similarly and eluted slightly later. Full-length and AspN or trypsin-digested peptides were analyzed by LC-MS/MS on a Waters Synapt Q-TOF equipped with a Waters UPLC and Phenomenex C18 Luna or C4 Jupiter columns. Buffers for LC-MS were 0.1% formic acid in water (LC-MS buffer A) or 0.1% formic acid in LC-MS grade acetonitrile (LC-MS buffer B).

In vitro assay of PmaB

In a 0.5 mL centrifuge tube, reaction components were combined to the following optimized final concentrations: PmaA (50 μM), PmaB (5 μM), CysRS (10 μM), ATP (5 mM), L-cysteine (5 mM), TCEP (1 mM), $tRNA^{Cys}$ (10 μM), *P. syringae* CysRS (10 μM), and HEPES assay buffer (100 mM HEPES pH 7.5; 5 mM $MgCl_2$, 100 mM NaCl). The assay mixture was incubated at 30 ºC for 30 min, desalted and concentrated by ZipTip, eluted in 50% acetonitrile, 50% water, 0.1% TFA, and analyzed by MALDI-TOF MS. For ^{31}P NMR analysis, reaction components were combined to final concentrations as follows: PmaA (50 μM), PmaB (5 μM),

Cys-tRNA^{Cys} (50 μM), ATP (0.5 mM), Tris assay buffer (50 mM Tris pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl). Cys-tRNA^{Cys} was obtained using a similar procedure for preparation of Glu-tRNA^{Glu} reported previously (4). For single turnover reaction prepared for ^{31}P NMR analysis, the reaction reaction components were combined to final concentrations as follows: PmaA (10 μM), PmaB (50 μM), Cys-tRNA^{Cys} (50 μM), ATP (10 μM), Tris assay buffer (50 mM Tris pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl). The reaction was incubated at 30 ºC for 15 min, then Chelex resin was added to remove Mg^{2+}. After using Amicon Ultra 3K MWCO (3,000 Da cut-off) tube to remove proteins and tRNA, the sample was added into NMR tube for analysis (notebook reference 7/6/2018).

LC-MS analysis anhydride intermediate

Assays were performed by reacting 25 μM PmaB in the presence or absence of ATP (5 mM) for 5 min at 30 ºC. Then assays were quenched with 20% ice-cold trichloroacetic acid and kept on ice for 20 min to precipitate the protein. The precipitate was separated by centrifugation (10,000 × g for 10 min) at 4 °C, washed with 3 × 400 μL of 10 mM ice-cold HCl, dried under vacuum, and dissolved in 20 μL of DMSO. To this solution, 15 μL of 0.1 M $NaBH_4$ in DMSO was added and incubated at 30 °C for 10 min, followed by addition of 1 mL of ice-cold perchloric acid (0.44 M). The mixture was incubated on ice for another 30 min, followed by centrifugation at 10,000 × g for 20 min at 4 °C. The resulting precipitate was washed with 3 × 400 μL of 10 mM ice-cold HCl, and dried under vacuum. The dried protein was re-dissolved in 6 M guanidine HCl and dilute to with H_2O to 1 M guanidine HCl in the solution. The protein was digested with 0.02 mg min^{-1} chymotrypsin at 37 °C overnight and analyzed by LC-MS (notebook reference 8/31/2018).

NH_2OH quenching assay

Assays were performed by reacting 50 μM PmaB and 50 μM PmaA in the presence or absence of ATP (5 mM) for 5 min at 30 ºC. Then assays were quenched with NH_2OH to final concentration of 1 M and incubate at 30 ºC for another 20 min. The assay was then desalted by ZipTip and analyzed by MALDI-TOF-MS (notebook reference 8/29/2018).

Preparation of plasmid *pTXB1-PmaA-intein-chitin*

The genes coding for PmaA was amplified by PCR using appropriate primers (**Table 5.2**), and purified by gel extraction from a 1% (w/v) agarose gel using the QIAquick Gel Extraction Kit. The vector *pTXB1* was digested using NdeI and SapI (NEB) restriction endonucleases and purified by gel extraction as described previously. The DNA fragment was inserted by Gibson one-step isothermal DNA assembly, and an aliquot of 20 μL of the Gibson assembly reaction was used to transform *E. coli* DH5α cells using the heat shock method. The cells were plated on LB plates supplemented with ampicillin (100 μg/mL) and the plates were incubated at 37 °C for 12–15 h. Single colonies were picked and grown in LB supplemented with ampicillin (100 μg/mL) at 37 °C for 12–15 h and the plasmids were isolated using a QIA prep Spin Miniprep Kit. Insert integrity was verified by sequencing the plasmids with the appropriate primers (**Table 5.2**).

Preparation of C-terminal ^{18}O labeled PmaA

E. coli BL21 (DE3) cells (50 μL) were electroporated with *pTXB1-PmaA-intein-chitin* (50 ng), and cells were plated on LB agar plates supplemented with kanamycin and grown at 37 °C for 12–15 h. A single colony was used to inoculate 20 mL of LB broth supplemented with ampicillin, grown for 12–15 h at 37 °C, and the culture was used to inoculate 2 L of TB media,

supplemented with ampicillin, to an OD_{600} of 0.025. Cultures were grown at 37 °C to a final OD_{600} of 1.0. Protein overexpression was induced by the addition of IPTG to a final concentration of 1 mM, and cultures were grown at 37 °C for 3 h. The cells were harvested by centrifugation, and then frozen in liquid nitrogen and stored at –80 ^{o}C.

For the purification of the PmaA peptide, the cell paste was placed into a metal sonication cup using a minimal amount of resuspension buffer (20 mM Tris-HCl, 500 mM NaCl, 1mM EDTA, pH 7.5) to transfer all of the cell paste. The cell paste was lysed by sonication (65% amplitude with 3 s pulses and 9 s delays) for 16 min. The sonicated cells were centrifuged for 30 min at 13,000 × g. The supernatant was poured into a column containing 20 mL of chitin beads pre-equilibrated with resuspension buffer. The supernatant and beads were gently shaken/rocked at 4 ^{o}C for 2 h to allow binding of the PmaA-intein-chitin binding domain fusion protein to the resin. The column was drained and washed with 600 mL of resuspension buffer to remove all impurities. The chitin beads were then added with 40 mL of cleavage buffer (50 mM HEPES, 200 mM NaCl, 1mM EDTA, 20 mM 2-mercaptoethanesulfonate, pH 7.7) and the column was capped and sealed with parafilm. The column was shaken at 4 ^{o}C for 6 h, and the solution was drained from the column directly into an Amicon Ultra 3K MWCO (3,000 Da cut-off) tube. The Amicon was then centrifuged until the protein solution had been concentrated to ~10 mL. The concentrated peptide solution was frozen in liquid nitrogen and lyophilized till dry.

The lyophilized peptide was re-dissolved in basic $H_2{}^{18}O$ buffer (100 mM Na_2CO_3, pH 10.0) and incubated at 30 °C for 1 h, and was purified on a Shimadzu Prominence Preparative Liquid Chromatography system equipped with a Phenomenex Luna C18 column (250 × 10 mm, 10 μm particle size, 100 Å pore size). Acetonitrile and 10 mM aq. NH_4HCO_3 were used as the mobile

phases, and a gradient of 2-80% aq. MeCN over 45 min at 5 mL min^{-1} was used for separation (notebook reference 7/10/2018).

Table 5.2 Oligonucleotides used in this study. All sequences are provided 5′ to 3′. Lowercase m indicates 2′ *O*-methylation of the following residue; methylation suppresses random addition of bases at the end of the RNA by T7 RNA polymerase.

Primer	Oligonucleotide Sequence
PmaA (-Ala)_F	GCAAGGTCTTTTGAAGCTTGCGGCCGCATAATG
PmaA (-Ala)_R	CGCAAGCTTCAAAAGACCTTGCTCTCGATGACTTCAATGTC
P. syringae $tRNA^{Cys}$ F	AATTCCTGCAGTAATACGACTCACTATAGGCCGAGTAGCAAAATGGTTATGCAGC
P. syringae $tRNA^{Cys}$ R	mUmGGAGGCCGAGGTCGGAATCGAACCGGCGTAGGCGGATTTGCAATCCGCTGCATAACC
PmaB_Seq_623	CTGGATCAGAGAACG
PmaB_Seq_1549	GGCTGATTGTCGAGC
PmaA_ 40mer_F	GAGAACCTGTACTTCCAATCCAACCAGCAAGCGTCC
PmaA_ 30mer_F	GAGAACCTGTACTTCCAATCCCTCGAAAACACTCCGCAG
PmaA_ 20mer_F	GAGAACCTGTACTTCCAATCCGCGTTGTTTGAAGAGTTTGACC
PmaA R	GGATTGGAAGTACAGGTTCTCCGGATCCTGGCTGTG
PmaB_S11A_F	CTGGCTTCGCGCTACCGGGTTTGCGGTGC
PmaB_S11A_R	CAAACCCGGTAGCGCGAAGCCAGAAATAGTGTG
PmaB_T180A_F	CAAAAAATGATGCATCGAGTTTTTTCGGCCC
PmaB_T180A_R	GAAAAAACTCGATGCATCATTTTTTGAACAGAATCGCTG
PmaB_S540A_F	CACTTTCACGCTCCGGATTTCCTGATTTCCAG
PmaB_S540A_R	GAAATCCGGAGCGTGAAAGTGGCTGCCAAATAC
PmaB_S780A_F	GTCAAGTTTGCTGAGATGTGTCCTGCTCC
PmaB_S780A_R	CACATCTCAGCAAACTTGACGTGCCCC
PmaB_R10A_F	CTATTTCTGGCTTGCATCGACCGGGTTTGCG
PmaB_R10A_R	CGGTCGATGCAAGCCAGAAATAGTGTGAGCTTTCC
PmaB_K161A_F	CAGCCGCAAGGCTCAGAAACTGCGCCTGG
PmaB_K161A_R	CAGTTTCTGAGCCTTGCGGCTGTCAGTTC
PmaB_Y170F_F	CCTGGAGCTTCGCACAGCGATTCTGTTC
PmaB_Y170F_R	GAATCGCTGTGCGAAGCTCCAGGCCAGG
PmaB_R173A_F	CTACGCACAGGCATTCTGTTCAAAAAATGATACGTCGAG
PmaB_R173A_R	GAACAGAATGCCTGTGCGTAGCTCCAG
PmaB_K177A_F	GATTCTGTTCAGCTAATGATACGTCGAGTTTTTTCGG
PmaB_K177A_R	CGTATCATTAGCTGAACAGAATCGCTGTGC
PmaB_D179A_F	GTTCAAAAAATGCTACGTCGAGTTTTTTCGGC

Table 5.2 (cont.)

Primer	Oligonucleotide Sequence
PmaB_D179A_R	CTCGACGTAGCATTTTTTGAACAGAATCGCTGTG
PmaB_R397A_F	CATGTATGTCGGTGCATACCCGGTCTACGAGGATTG
PmaB_R397A_R	CGGGTATGCACCGACATACATGGCGC
PmaB_E403A_F	GGTCTACGCAGATTGTTCGCGCAATATCGATATC
PmaB_E403A_R	GCGAACAATCTGCGTAGACCGGGTAGCG
PmaB_D404A_F	GGTCTACGAGGCATGTTCGCGCAATATCGATATCAG
PmaB_D404A_R	GCGAACATGCCTCGTAGACCGGGTAG
PmaB_D542A_F	CACTCTCCGGCATTCCTGATTTCCAGTACCTCG
PmaB_D542A_R	GAAATCAGGAATGCCGGAGAGTGAAAGTGGC
PmaB_H566A_F	GAGAGGTCGCTCCGGGCGTGCACAC
PmaB_H566A_R	CGCCCGGAGCGACCTCTCCCAGAATGATTGAGTAGTCG
PmaB_R704A_F	CCTTGTACAAAGCTGCCTCGTGGTGGTTCAG
PmaB_R704A_R	CGAGGCAGCTTTGTACAAGGTCTTGCCCAAG
PmaB_W707A_F	GCCTCGGCATGGTTCAGTCCAGAGCAAC
PmaB_W707A_R	GACTGAACCATGCCGAGGCGCGTTTGTAC
PmaB_K746A_F	GTTCGCCGCTATCGATATCGAACCCAAGCC
PmaB_K746A_R	CGATATCGATAGCGGCGAACACATAACGGG
PmaB_K752A_F	CGAACCCGCTCCGATTTTCATTGATTTCGATAACCC
PmaB_K752A_R	CAATGAAAATCGGAGCGGGTTCGATATCGATCTTGGC
PmaB_E801A_F	CATTTCTGTTGTGCTATACGTACAACCTTTAGGGACAATGG
PmaB_E801A_R	GTTGTACGTATAGCACAACAGAAATGCCCGC
PmaB_R803A_F	CTGTTGTGAAATAGCAACAACCTTTAGGGACAATGGAG
PmaB_R803A_R	GGTTGTTGCTATTTCACAACAGAAATGCCCG
PmaA_ A50F_F	CAAGGTCTTTTTCTGAAGCTTGCGGCCG
PmaA_ A50F_R	CAAGCTTCAGAAAAAGACCTTGCTCTCGATGACTTC
PmaA_ A50K_F	CAAGGTCTTTAAATGAAGCTTGCGGCCG
PmaA_ A50K_R	CAAGCTTCATTTAAAGACCTTGCTCTCGATGACTTC
PmaA_ A50E_F	CAAGGTCTTTGAATGAAGCTTGCGGCCG
PmaA_ A50E_R	CAAGCTTCATTCAAAGACCTTGCTCTCGATGACTTC
PmaA_ A50Q_F	CAAGGTCTTTCAGTGAAGCTTGCGGCCG
PmaA_ A50Q_R	CAAGCTTCACTGAAAGACCTTGCTCTCGATGACTTC
PmaA_ A50P_F	CAAGGTCTTTCCATGAAGCTTGCGGCCG
PmaA_ A50P_R	CAAGCTTCATGGAAAGACCTTGCTCTCGATGACTTC
PmaA_ A50S_F	CAAGGTCTTTTCTTGAAGCTTGCGGCCG
PmaA_ A50S_R	CAAGCTTCAAGAAAAGACCTTGCTCTCGATGACTTC
pTXB1-PmaA_F	GGTCTTTGCCTGCATCACGGGAGATG

Table 5.2 (cont.)

Primer	Oligonucleotide Sequence
pTXB1-PmaA_R	GTTGGGTTGTCCCATATGTATATCTCCTTCTTAAAGTTAAAC

5.4 References

1. Ortega, M. A., & van der Donk, W. A. (2016). New insights into the biosynthetic logic of ribosomally synthesized and post-translationally modified peptide natural products. *Cell Chem. Biol., 23*, 31-44.

2. Lubelski, J., Rink, R., Khusainov, R., Moll, G. N., & Kuipers, O. P. (2008). Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. *Cell. Mol. Life Sci., 65*, 455-476.

3. Ortega, M. A., Hao, Y., Zhang, Q., Walker, M. C., van der Donk, W. A., & Nair, S. K. (2015). Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature, 517*, 509-512.

4. Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., & van der Donk, W. A. (2015). In vitro biosynthesis of the core scaffold of the thiopeptide thiomuracin. *J. Am. Chem. Soc., 137*, 16012-16015.

5. Ozaki, T., Kurokawa, Y., Hayashi, S., Oku, N., Asamizu, S., Igarashi, Y., & Onaka, H. (2016). Insights into the biosynthesis of dehydroalanines in goadsporin. *Chembiochem, 17*, 218-223.

6. 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., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., 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., & 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.

7. Maruyama, C., Toyoda, J., Kato, Y., Izumikawa, M., Takagi, M., Shin-ya, K., Katano, H., Utagawa, T., & Hamano, Y. (2012). A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. *Nat. Chem. Biol., 8*, 791.

8. Kadi, N., Oves-Costales, D., Barona-Gomez, F., & Challis, G. L. (2007). A new family of ATP-dependent oligomerization-macrocyclization biocatalysts. *Nat. Chem. Biol., 3*, 652.

9. Giessen, T. W., & Marahiel, M. A. (2012). Ribosome-independent biosynthesis of biologically active peptides: Application of synthetic biology to generate structural diversity. *FEBS Lett., 586*, 2065-2075.

10. Moutiez, M., Schmitt, E., Seguin, J., Thai, R., Favry, E., Belin, P., Mechulam, Y., & Gondry, M. (2014). Unravelling the mechanism of non-ribosomal peptide synthesis by cyclodipeptide synthases. *Nat. Comm., 5*, 5141.

11. Noike, M., Matsui, T., Ooya, K., Sasaki, I., Ohtaki, S., Hamano, Y., Maruyama, C., Ishikawa, J., Satoh, Y., Ito, H., Morita, H., & Dairi, T. (2014). A peptide ligase and the ribosome cooperate to synthesize the peptide pheganomycin. *Nat. Chem. Biol., 11*, 71.

12. Fawaz, M. V., Topper, M. E., & Firestine, S. M. (2011). The ATP-grasp enzymes. *Bioorg. Chem., 39*, 185-191.

13. Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J. i., & Katsube, Y. (1993). Three-dimensional structure of the glutathione synthetase from Escherichia coli B at 2·0 Å Resolution. *J. Mol. Biol, 229*, 1083-1100.

14. Fan, C., Moews, P. C., Shi, Y., Walsh, C. T., & Knox, J. R. (1995). A common fold for peptide synthetases cleaving ATP to ADP: glutathione synthetase and D-alanine:d-alanine ligase of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A., 92*, 1172-1176.

15. Das, D., Kuzmic, P., & Imperiali, B. (2017). Analysis of a dual domain phosphoglycosyl transferase reveals a ping-pong mechanism with a covalent enzyme intermediate. *Proc. Natl. Acad. Sci. U. S. A., 114*, 7019-7024.

16. Collet, J.-F., Stroobant, V., & Van Schaftingen, E. (2002). Evidence for phosphotransferases phosphorylated on aspartate residue in N-terminal DXDX(T/V) motif. *Methods Enzymol., 354*, 177-188.

17. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc., 10*, 845.

18. Chung, C. T., Niemela, S. L., & Miller, R. H. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U. S. A., 86*, 2172-2175.

19. Sherlin, L. D., Bullock, T. L., Nissan, T. A., Perona, J. J., Lariviere, F. J., Uhlenbeck, O. C., & Scaringe, S. A. (2001). Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA, 7*, 1671-1678.

20. Rio, D. C., Ares, M. J., Hannon, G. J., & Nilsen, T. W. (2011). *RNA: A Laboratory Manual*: Cold Spring Harbor Laboratory Press.

21. Walker, S. E., & Fredrick, K. (2008). Preparation and evaluation of acylated tRNAs. *Methods, 44*, 81-86.