ADENYLATION AND TAILORING ACTIVITIES IN THE NONRIBOSOMAL PEPTIDE SYNTHESIS OF THE SIDEROPHORE PYOCHELIN

BY

Trey A. Ronnebaum

Submitted to the graduate degree program in the Department of Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson Audrey L. Lamb

_____

Robert C. Dunn

_____

Timothy A. Jackson

_____

James D. Blakemore

_____

Michael J. Hageman

Date Defended: 08/17/2018

The Dissertation Committee for Trey A. Ronnebaum

certifies that this is the approved version of the following dissertation:

ADENYLATION AND TAILORING ACTIVITIES IN THE NONRIBOSOMAL PEPTIDE SYNTHESIS OF THE SIDEROPHORE PYOCHELIN

Chairperson Audrey L. Lamb

_____

Date Defended: 08/17/2018

Abstract

Pathogenic bacteria are becoming increasingly resistant to antibiotics. In response to this alarming trend, the scientific community must determine therapeutic targets. Essential nutrients, such as iron, are necessary for pathogens to survive and become virulent within a host system. One mechanism used by pathogenic bacteria to acquire iron from its surrounding environment is to produce low-molecular weight compounds which have a high affinity towards ferric iron. These compounds are called siderophores, and studies have shown their production to be essential for growth and virulence of some pathogens in iron-limited environments, such as the human host. Siderophores are often biosynthesized by nonribosomal peptide synthetases (NRPSs), which rarely have human homologs, making them attractive targets for novel therapeutics.

NRPSs are enzymes utilized by bacteria, fungi, and plants to generate bioactive peptides. These bioactive peptides are not only used as secondary metabolites (toxins, pigments, siderophores) but have also found their way into the clinic as antibiotics, anticancer drugs, and immunosuppressants. To elicit their unique bioactivity, these peptides are tailored, making the compound chemically unique. Natural product chemists, metabolic engineers, and researchers in biochemistry and biotechnology work to exploit NRPS biosynthesis to generate new compounds for clinical use.

This dissertation describes mechanistic and structural analyses of the adenylation and tailoring domains of the NRPS biosystem responsible for the production of pyochelin, a siderophore produced by antibiotic resistant *Pseudomonas aeruginosa*. A large portion of this work aims to better understand adenylation and "stuffed" tailoring didomains which lack

structural characterization and have limited mechanistic understanding yet are ubiquitous in NRPS bioassembly. Pyochelin biosynthesis employs an adenylation-epimerase stuffed didomain in PchE and an adenylation-methyltransferase stuffed didomain in PchF. Substrate and product analogs were synthesized and steady-state adenylation, epimerase, and methyltransferase assays, along with onium chalcogen effects of the methyltransferase reaction, were used to characterize the adenylation-tailoring stuffed domains in pyochelin bioassembly. Similarly, substrate analogs were generated and used in steady-state kinetic and crystallography experiments with the stand-alone tailoring, NADPH-dependent reductase, PchG, and homolog, Irp3, of yersiniabactin biosynthesis. Finally, a steady-state adenylation assay was developed for the stand-alone salicylate adenylation enzyme, PchD, and potential warhead inhibitors were synthesized and co-crystallized laying the groundwork for future inhibitor design.

Acknowledgements

This work would not have been possible without my family, friends, and wonderful wife and best friend, Julia. I am very grateful for their continued support and presence in my life.

I would like to acknowledge and thank the efforts of those who have assisted in my graduate education and research. First and foremost, I would like to especially thank the NIH Dynamic Aspects of Chemical Biology Fellowship for its unique scholarship, career training, and funding for conference gatherings and research rotations. I was very fortunate to get to spend time in Dr. Geoff Horsman's lab at Wilfrid Laurier University in Waterloo, ON, Canada, and want to give a special thank you to Dr. Horsman and his lab who assisted with chiral chromatography experiments. I want to also give a special thank you to Dr. Squire Booker and his lab at Pennsylvania State University and Howard Hughes Medical Institute. I would like to thank Dr. Booker and his lab team for hosting me and assisting in the biosynthesis and purification of AdoMet and onium chalcogen analogs on my visit. Another esteemed thank you to Dr. Thomas Prisinzano and the Prisinzano lab, who contributed to discussions and assisted with the synthetic work presented in this dissertation and provided me with a welcoming fume hood on west campus.

I would like to recognize and thank Dr. T. Christopher Gamblin and the Gamblin lab for use of their equipment, enjoyable discussions, and the hundreds of Keurig cups of coffee! I would like to acknowledge and thank the Medicinal Chemistry Department and the Wolfe and Altman lab for letting me use their UPLC and assisting with its setup. I would also like to acknowledge the Stanford Synchrotron Radiation Laboratory for the use of their facilities which were used for the crystal structures of Irp3 and PchD.

I would like to acknowledge and thank colleagues from my previous lab whom aided in my growth and development as a scientific thinker: Aaron Rudeen, HeeJung Moon, Mason Lantz, and Joel Finney. And I would like to thank the Lamb lab undergraduates who have assisted with my projects in one way or another: Cara Davis, Kathryn Brewer, Easton Mickelson, and Kristen Khoang. A very special thank you to current and former lab members, Dr. Catherine Shelton, Dr. Kathleen Meneely, Annemarie Chilton, Nikola Kenjic, and Jeffrey McFarlane for experiment assistance, countless fruitful discussions, and long-lasting friendship.

I would like to acknowledge and thank current and past committee members that have contributed to my scientific growth: Dr. Paul Hanson, Dr. David Weis, Dr. Alexander Moise, Dr. Robert Dunn, Dr. Michael Hageman, Dr. Timothy Jackson, and Dr. James Blakemore.

Finally, I would like to express my deepest gratitude to Dr. Audrey Lamb whose professional advice, scientific expertise, and criticism is always welcomed. Without her support, encouragement, and advocacy, this work would not have been possible. I am deeply indebted.

Table of Contents

List of Figures

List of Schematics

List of Tables

Chapter 1

Introduction:

Nonribosomal peptides for iron acquisition: pyochelin biosynthesis as a case study

*This chapter was published in Ronnebaum, T.A., and Lamb, A.L. (2018) Nonribosomal peptides for iron acquisition: pyochelin biosynthesis as a case study, *Curr. Opin. Struct. Biol. 53,* 1-11.

Iron uptake: Siderophores

Iron is an essential nutrient that is needed by microbes to perform critical biological processes necessary for survival.[1] Due to the paucity of free iron in aerobic environments, microbes have developed intricate systems to acquire iron from their surroundings.[1] One such system is the biosynthesis of small molecules known as siderophores that have a high affinity for ferric iron.[2] When iron availability is low, microbes synthesize and secrete siderophores, and selectively reimport the iron-loaded form from the surrounding environment. Siderophore biosynthesis in bacteria is accomplished by nonribosomal peptide synthetase (NRPS) enzymes, polyketide synthase (PKS) enzymes, and/or by NRPS independent siderophore (NIS) synthetase enzymes.[3-5] Typically these enzymes do not have human homologues and knockdown experiments of siderophore biosynthesis have shown partial or complete suppression of pathogenicity for many different bacteria, making these enzymes attractive drug targets.[6] This introduction will provide an overview of the structural biology of NRPS enzymes and then focus on the structural and catalytic aspects of the enzymes involved in the biosynthesis of pyochelin, a

siderophore from *Pseudomonas aeruginosa*, highlighting the involved accessory and tailoring activities that generate the unique chemical structure of pyochelin.

Nonribosomal peptide synthetases (NRPS): chemical logic of peptide chain formation

NRPS enzymes are large multidomain and multifunctional enzymes that display a chemical logic in which each module is responsible for the addition of a single amino acid to a growing peptide chain, including non-proteinogenic amino acids and hydroxy acids. An NRPS module consists of a condensation (C) domain, an adenylation (A) domain, and a peptidyl carrier protein (P) domain, except for the first initiation module, which lacks a condensation domain. The P-domain acts as a tethering system for the growing peptide chain and must be post-translationally modified with coenzyme A on a conserved serine residue thereby generating a 4'phosphopantetheinyl (Ppant) swinging arm. A-domains have been named the "gate keepers" of the NRPS assembly line as they selectively activate and incorporate the appropriate amino acids into the growing peptide chain.[7] The A-domains activate an amino acid by catalyzing the formation of an aminoacyl-AMP, using ATP and releasing inorganic phosphate (**Figure 1-1A**). The Ppant thiol of the P-domain performs a nucleophilic attack on the carboxyl group of the aminoacyl-AMP, eliminating AMP and generating an aminoacyl-thioester bond, thereby priming the module. Once primed, the thioester tether on the P-domain transfers its amino acid cargo to subsequent domains, similar to a mechanical crane.

Peptide bond formation, or elongation, is catalyzed between the amino acids of two primed modules within the C-domain of the downstream module. The upstream thioester bond is broken as the peptide bond is formed, thereby transferring the elongating peptide to the downstream module. This process is repeated with each subsequent module until the full chain is

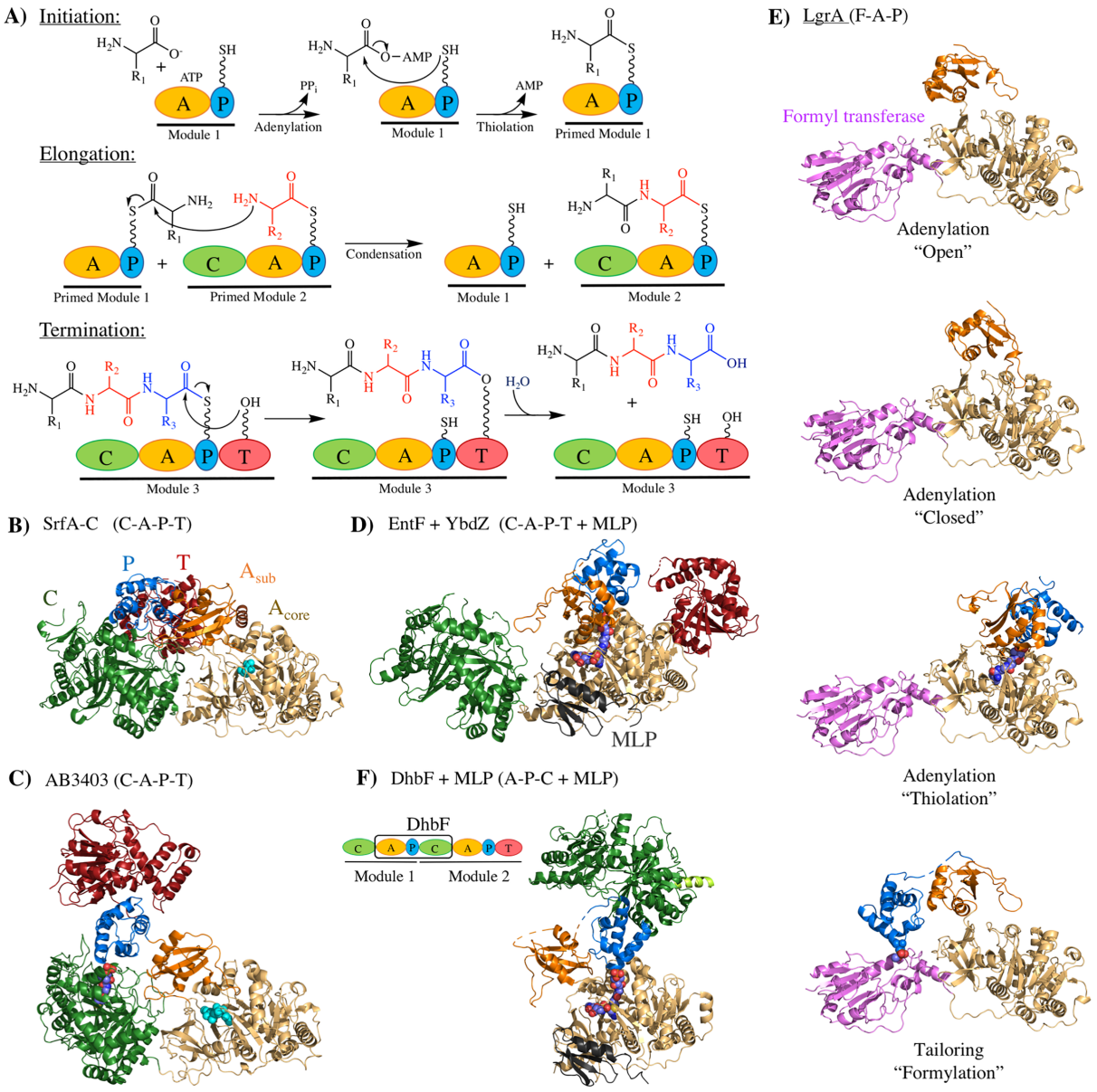


Figure 1-1: NRPS assembly line and crystal structures highlighting intra- and inter-domain contacts and movements. A) NRPS natural product biosynthesis takes place in three phases: initiation, condensation, and termination. Adenylation and subsequent thiolation of the amino acid onto the Ppant tether prime a module. A condensation reaction is performed between amino acids tethered to primed modules to generate a peptide bond, thereby elongating the peptide chain. Termination and release of the mature peptide occurs by hydrolysis within a terminal thioesterase domain after transfer of the product to a conserved serine. Throughout this review: yellow = adenylation (A); blue = peptidyl carrier domain (P); green = condensation (C); red = thioesterase (T). **B)** The first crystal structure of a complete NRPS module, SrfA-C (PDB: 2VSQ). Note that the A_{core}-domains (light yellow) are aligned in parts B thru F of this figure to highlight domain movements. The substrate leucine (cyan spheres) is in the A_{core}-domain. The A_{sub}-domains are in dark yellow. **C)** The full module of AB3403 (PDB: 4ZXI) with the Ppant

tether (purple carbon and red oxygen spheres) in the closed C-domain ready to accept an upstream peptide. The substrates AMP and Gly (cyan spheres) are in the A_{core}-domain active site. **D)** A full-module crystal structure of EntF with its MLP, YbdZ (black), bound to the A_{core} (PDB: 5JA1). The mechanistic-based inhibitor, Ser-AVS, is covalently attached to the Ppant tether (spheres) which is trapped in the A_{core}-domain. **E)** Domain movements of the A_{sub}- and P-domain are highlighted by the multi-domain structures of LgrA. These snap-shots show the movement of the A_{sub}-domain in "open" and "closed" (PDB: 5ES5) conformations during adenylation, and the movement of the A_{sub}- and P-domains during thioester formation (Val-NH-Ppant analog shown in spheres) (PDB: 5ES8) and for tailoring by the formyltransferase domain (pink) (PDB: 5ES9). **F)** The cross-module structure of DhbF containing an A- and P-domain from module 1 and C-domain from module 2. The cartoon emphasizes the cross-module construct. The MLP from *Geobacillus sp*. is bound to the A_{core}-domain and shown in black. The mechanistic-based inhibitor, Gly-AVS (spheres), is covalently bound to the Ppant arm in the A_{core}-domain (PDB: 5U89).

made. Chain elongation is terminated by a thioesterase (T) domain in the terminal module.[8] The peptide is transferred from the thioester of the P-domain to a conserved serine residue of the T-domain, generating an amino ester, and allowing for hydrolysis and release of the mature peptide.[8]

Nonribosomal peptide synthetases (NRPS): structural biology of an assembly line

The C-, A-, and P-domains have been well-studied, with both structural and kinetic characterizations.[9-13] The first structure elucidating the organization of a full NRPS module was the termination module (C-A-P-T) of surfactin biosynthesis, SrfA-C, from *Bacillus subtilis* (**Figure 1-1B**).[14] Crystals of the full module were obtained after mutating the serine of the P-domain that is post-translationally modified with Ppant to alanine. The A-domain contains two subdomains, a larger N-terminal A core (A_{core}) domain and a smaller C-terminal A sub (A_{sub}) domain. The structure revealed a large interface between the C-domain and A_{core}-domain creating a platform for the A_{sub}- and P-domains. In this structure, the P-domain was oriented so that the tether would be directed towards the acceptor site of the C-domain.

Structures harboring the Ppant tether were documented with *holo*-AB3403 (C-A-P-T), a NRPS termination module from an uncharacterized biosynthetic pathway from *Acinetobacter baumanni* (AB3403, **Figure 1-1C**) and EntF from *E. coli* for the production of enterobactin (**Figure 1-1D**).[15] In AB3403, the C-domain adopted a closed conformation and the P-domain was found to be rotated ~30˚ in comparison to the *apo*-SrfA-C structure thereby promoting insertion of the Ppant into the C-domain.[15] In EntF, the Ppant tether was directed into the adenylation domain, causing a shift in domain orientation of both the A_{sub}- and P-domains. Therefore, during the two different catalytic steps (adenylation and thiolation), the A_{sub}-domain

assumes different orientations relative to the A_{core}-domain to promote the appropriate chemistries (**Figure 1-1 C,D**).[16] This has been named "domain alternation".[17] Importantly, when the Ppant tether is directed into the condensation domain for peptide bond formation, the A_{sub}-domain is in an adenylate-forming conformation, supporting hypotheses that while the peptide bond is being formed in the condensation domain for one biosynthetic product molecule, the A_{core}-domain adenylates an amino acid for the next catalytic cycle, suggesting NRPS enzymes work in true assembly line fashion.[15]

Domain alternation of A_{sub}- and P-domains were captured in multiple structures of LgrA, an NRPS of gramicidin synthase, representing the different steps of the catalytic cycle.[16] LgrA is an initiation module containing a formyltransferase tailoring domain, an A-, and a P-domain.[16] Structures of A_{sub} in the open and closed adenylation conformations were determined (**Figure 1-1E**) in which the P-domains were disordered. An isosteric analog mimicked the natural Ppant tether, replacing the hydrolyzable thioester with an amide (Val-NH-Ppant). With this analog, structures in the thiolation and formyltransferase conformations were determined in which the P-domain is ordered.[16, 18] In the thiolation state, the A_{sub}-domain was oriented so that the P-domain inserts the Ppant mimic into the adenylation domain extending evidence that A_{sub}-domain positioning is altered by the P-domain conformation. The differing thiolation and formyltransferase conformations of LgrA demonstrate the remarkable variability in module architecture.[16]

A cross-module structure consisting of an A- and P-domain from the first module of DhbF and a C-domain from the second module was determined (**Figure 1-1F**).[19] DhbF is a dimodular NRPS involved in the biosynthesis of the *Bacillus* siderophore, bacillibactin.[19] Unlike the observed interactions between A-domains and upstream C-domains, the downstream C-

domain had no interaction with the upstream A-domain. In this structure, the C-domain only interacts with the "back" face of the P-domain as the Ppant tether was trapped in the thiolation state of the upstream adenylation domain by a mechanistic inhibitor.[19] These data in addition to cryo-EM structures and modeling suggest a lack of conserved intermodular surface interactions.[19] Additional higher-order structures examining intermodular interfaces may reveal specific structural interactions necessary for chain elongation.

MbtH-like proteins (MLPs) are auxiliary proteins sometimes found in NRPS gene clusters that have been hypothesized to have chaperone properties or to augment the catalytic activity of adenylation domains.[20, 21] Crystal structures have been determined of MLPs alone and bound to A_{core}-domains.[22-24] The DhbF structure described above was determined with an MLP bound to the A_{core}-domain (**Figure 1-1F**), and the MLP was shown to be essential for substrate binding and adenylation activity.[19] The full-length EntF NRPS module was co-crystallized with MLPs that enhance its adenylation activity: the MLP encoded within its operon, YbdZ (**Figure 1-1D**), and an MLP from *P. aeruginosa*, PA2412.[24, 25] These structures exhibit no significant change in secondary structure of the A-domain upon MLP binding suggesting MLP binding may alter protein dynamics, though the mechanism by which MLPs enhance solubility or catalytic activity of A-domains remain elusive.[19, 24]

Pyochelin from *Pseudomonas aeruginosa*

The chemical logic and structural biology of NRPS modules is key for understanding how the peptide bonds are made by these nanomachines. However, the bioactive peptides produced by these enzymes are far more diverse than the 20 amino acids found in proteins. The diversity is incorporated by accessory enzymes that produce unusual amino and hydroxy acid

substrates for adenylation enzymes or by tailoring enzymes that alter amino acids already incorporated into the growing chain. Here, we use the biosynthesis of pyochelin, a small salicylate-capped, siderophore produced by *P. aeruginosa*, as an example to describe some of the accessory enzymes and tailoring domains involved in the biosynthesis of nonribosomal peptides.

Pyochelin is derived from three precursor molecules: one molecule of salicylate, a hydroxy acid, and two molecules of cysteine. Therefore, biosynthesis requires three NRPS modules: an initiation module that lacks a C-domain, an elongation module, and a termination module with a T-domain (**Figure 1-2A**). Biosynthesis of pyochelin is initiated by two accessory enzymes.[26-28] Chorismate is converted to isochorismate by the isochorismate synthase (ICS), PchA,[29] and the isochorismate is subsequently converted to salicylate by the isochorismate pyruvate lyase (IPL), PchB (**Figure 1-2B**).[30] The salicylate is activated by the stand-alone A-domain, PchD, and transferred to the N-terminal P-domain of the NRPS, PchE (the initiation module).[31] The elongation module of PchE integrates an L-cysteine into the growing peptide chain, cyclizes the L-cysteine to a thiazoline ring, and alters the stereochemistry with an epimerase tailoring domain.[32] The termination module of PchF adds and cyclizes a second L-cysteine to generate a second thiazoline. The tailoring protein PchG reduces the thiazoline to thiazolidine.[27] The stuffed methyltransferase domain of PchF performs an *S*-adenosylmethionine (AdoMet) dependent *N*-methylation of the thiazolidine.[33] The mature pyochelin is released by the T-domain of PchF.

Accessory enzymes in pyochelin biosynthesis: PchA and PchB

PchA, isochorismate synthase. PchA catalyzes the first step in pyochelin biosynthesis, isomerizing chorismate to isochorismate (**Figure 1-3A**).[29, 30] PchA is part of the menaquinone,

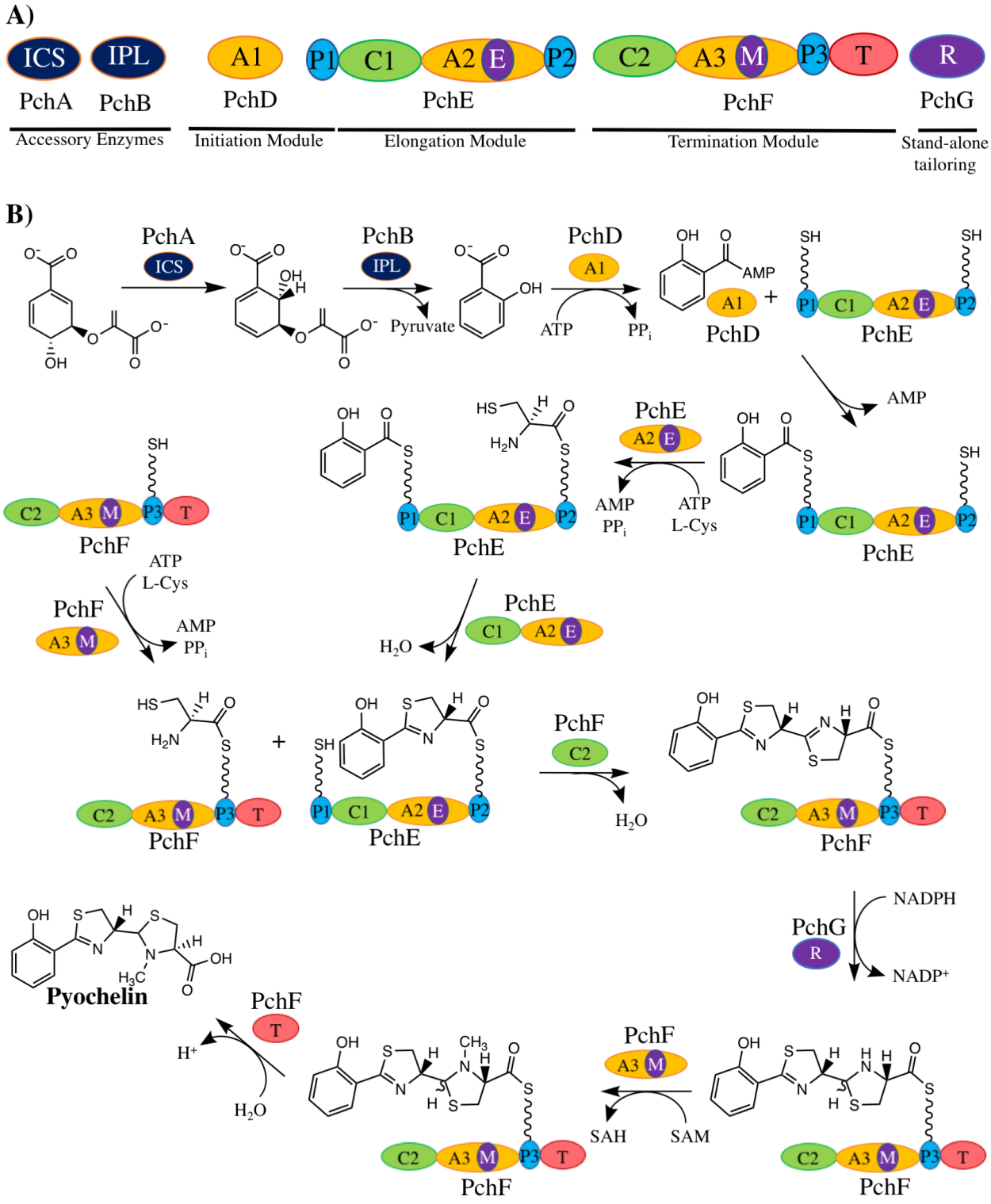


Figure 1-2: Pyochelin biosynthesis. A) Three NRPS enzymes consisting of eleven different domains, two accessory enzymes, and one stand-alone tailoring enzyme are used to generate pyochelin, a siderophore from *P. aeruginosa*. Domain abbreviations are as in Figure 1-1, with the exception that domain types are numbered in the assembly line (A1 = first adenylation

domain), and these additions: ICS = isochorismate synthase, IPL = isochorismate-pyruvate lyase, R = reductase. **B)** Biosynthesis is initiated by PchA and PchB, which generate the hydroxy acid, salicylate, for the N-terminal cap. PchA converts chorismate to isochorismate and PchB converts isochorismate to salicylate and pyruvate. Salicylate is activated by PchD and delivered to the Ppant tether of the first P-domain in PchE. Adenylation and thiolation of L-cysteine occur in the second module. Condensation with the upstream salicylate and cyclization and epimerization of L-cysteine to D-thiazoline occurs in the elongation module of PchE. The termination module in PchF, primed with an L-cysteine, condenses and cyclizes the cysteine with the upstream peptide chain to form a hydroxyphenyl-bis-thiazoline intermediate. PchG, the stand-alone tailoring domain, reduces the second thiazoline to a thiazolidine in an NADPH-dependent manner. Subsequent AdoMet-dependent methylation of the thiazolidine nitrogen is catalyzed by the stuffed tailoring methyltransferase domain of PchF. Mature pyochelin is hydrolyzed and released by the T-domain.

siderophore, and tryptophan (MST) family of enzymes that are Mg^{2+}-dependent chorismate utilizing enzymes.[34-36] Although the structure of PchA remains unsolved, the structure is hypothesized to be homologous to the isochorismate synthases, EntC [37, 38] and MenF [35] (both *E. coli* enzymes) and the salicylate synthases Irp9 [39] and MbtI [34, 40, 41] (**Figure 1-3B**). Interestingly, Irp9 and MbtI are the initiating enzymes to generate salicylate-capped siderophores for *Yersinia spp.* and *Mycobacterium tuberculosis*. These enzymes generate salicylate, first performing the isomerase activity of PchA and then the lyase activity, described below for PchB, in a single active site. The active site amino acids for all four isochorismate and salicylate synthase enzymes are highly conserved,[42] catalyzing the isomerase activity using a general acid – general base mechanism.[43] The active site is found at the interface of two structural domains capped shut by the binding of a catalytically required magnesium ion, preventing substrates entering or products leaving the active site. Magnesium ions at high concentrations are found to be inhibitory for isochorismate synthase enzymes, but promote lyase activity in the salicylate synthase enzymes.[39] Interestingly, ferrous ions bind to the catalytic metal binding site with nanomolar affinity for both isochorismate and salicylate synthase enzymes inhibiting catalytic turnover and suggesting a potentially physiologically relevant negative feedback mechanism for siderophore production.[39]

PchB, isochorismate-pyruvate lyase. PchB is the second enzyme in the pyochelin biosynthetic pathway, cleaving isochorismate to salicylate and pyruvate, a lyase activity (**Figure 1-3C**).[30] PchB has no structural homology to the salicylate synthase enzymes that perform the lyase activity (described above), and is instead a structural homologue of *E. coli* chorismate mutase

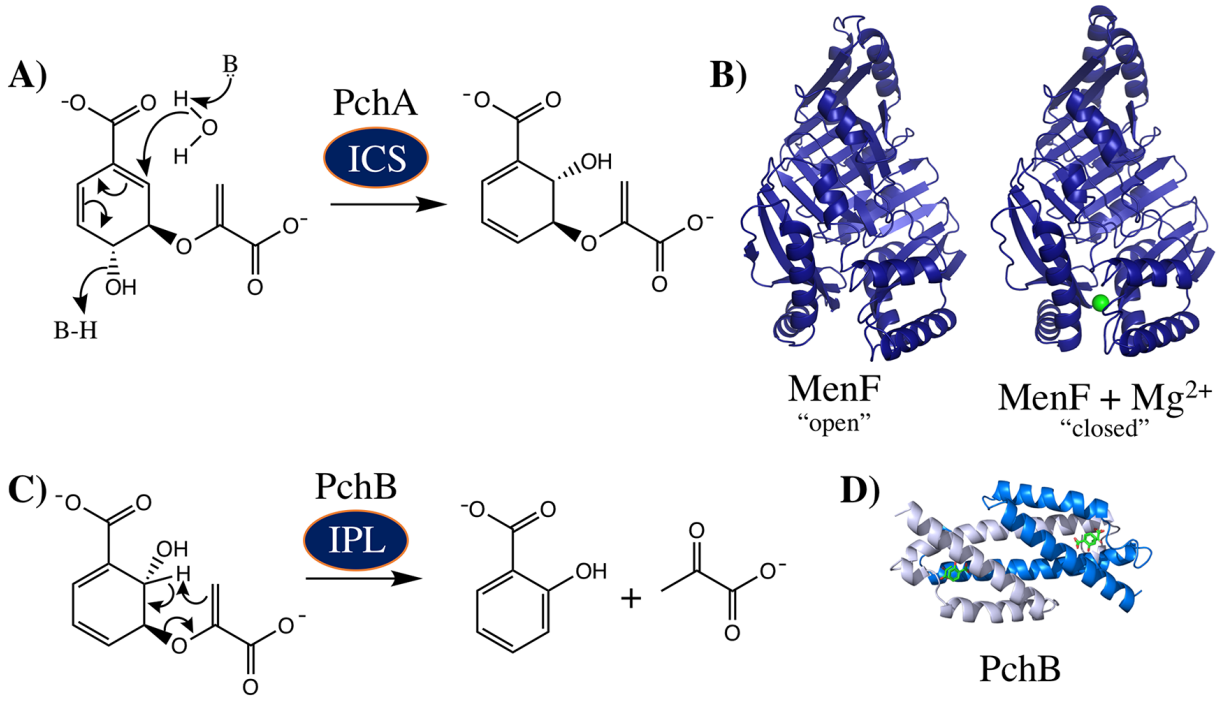


Figure 1-3: Accessory enzymes in pyochelin biosynthesis. A) PchA is an isochorismate synthase (ICS) converting chorismate to isochorismate. **B)** PchA structural homolog MenF in the "open" (PDB: 3BZM) and "closed" (PDB: 3BZN) structures, showing the magnesium ion (green) that creates a button closing the active site. **C)** PchB is an isochorismate pyruvate-lyase (IPL) converting isochorismate to salicylate and pyruvate. **D)** Crystal structure of dimeric PchB (PDB: 3REM). Monomers are shown in blue and gray, with the products (green sticks) in the active sites.

(EcCM) of the AroQ class (**Figure 1-3D**). Interestingly, PchB has adventitious chorismate mutase activity, albeit with much lower catalytic efficiency. The chorismate mutase reaction is a pericyclic Claisen rearrangement converting chorismate to prephenate.[44, 45] The mechanism of the lyase reaction is an asynchronous and sigmatropic pericylic reaction, with a quantitative hydrogen transfer from the ring to the pyruvylenol tail, substantiated by NMR experiments [46] and further supported computationally.[47] While this mechanism has also been proposed for the lyase-active salicylate synthases,[48] a second hypothesis for the salicylate synthases is that the reaction is performed by acid-base chemistry.[49] PchB is as an intertwined dimer, in which each monomer consists of three helices.[44, 50, 51] The active site loop is between helix 1 and helix 2 and is fully ordered when products are bound, whereas in the apo-structure the loop is disordered suggesting an open state for substrate entry or product egress, with closure during catalysis.[50] Lys42 is part of the mobile site loop and is hydrogen bonded to the pyruvate product.[44] Mutating Lys42 to an Ala lowers lyase activity 100-fold although active site architecture is maintained suggesting that the chemical nature of Lys42 is key for efficient catalysis.[44]

Stuffed tailoring domains in pyochelin biosynthesis: PchE and PchF

Most commonly, tailoring domains are incorporated into an NRPS module as an independent domain following the P-domain. However, in some cases tailoring domains are inserted within the A-domain, considered "interrupted" adenylation domains.[52] The tailoring domains of both PchE and PchF are stuffed into interrupted adenylation domains between the core sequence motifs, A8 and A9, of the A_{sub}-domain (**Figure 1-4A**).[33, 52] The adenylation-tailoring didomains lack structural characterization and have limited mechanistic data.[53, 54]

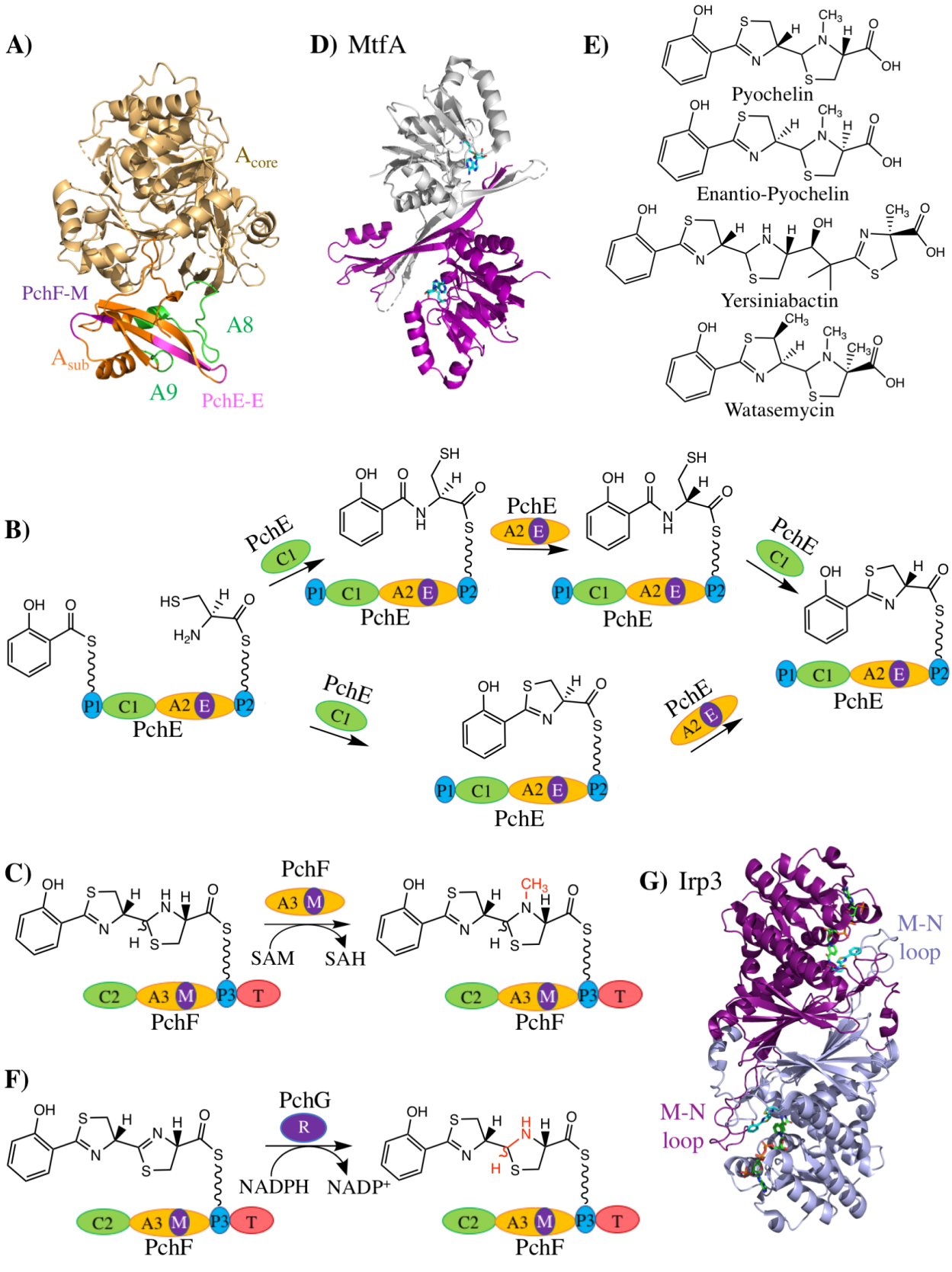

A)
A_core
PchF-M
A8
A_sub
A9
PchE-E
D) MtfA
E)
Pyochelin
Enantio-Pyochelin
Yersiniabactin
Watasemycin
B)
PchE
C1
A2
E
P1
P2
C)
PchF
A3
M
SAM
SAH
C2
P3
T
F)
PchG
R
NADPH
NADP+
G) Irp3
M-N loop
M-N loop

Figure 1-4: Tailoring enzymes. A) The A-domain of AB3403 (PDB: 4ZXI) highlighting the core adenylation sequences A8 and A9 (purple), between which tailoring domains are stuffed. Also highlighted is a sequence-based prediction of where the PchE epimerase-domain (blue) and PchF M-domain (green) are stuffed into the A_{sub}-domain. **B)** PchE first condenses the upstream salicylate to the L-cysteine within its condensation domain. From here, two hypotheses arise. The first (top) suggests that the peptide tether moves to the epimerase domain to epimerize the L-cysteine to D-cysteine, and then moves back to the condensation/cyclization domain to cyclize the D-cysteine to D-thiazoline. The second hypothesis (bottom) suggests that after condensation to the upstream salicylate, L-cysteine cyclizes to L-thiazoline while in the condensation domain and then moves to the epimerization domain to generate D-thiazoline, resulting in less domain movement. **C)** The "stuffed" methyltransferase domain of PchF performs a AdoMet-dependent methyl transfer to the nitrogen of the thiazolidine ring. **D)** MtfA (PDB: 3G2O), a stand-alone AdoMet-dependent nitrogen methyltransferase of vancomycin biosynthesis from *Amycolatopsis orientalis*, with AdoMet (cyan sticks) bound. **E)** Pyochelin-like natural products made by NRPSs: Pyochelin – *Pseudomonas aeruginosa*. Enantio-pyochelin – *Pseudomonas fluorescens*. Yersiniabactin – *Yersinia pestis*. Watasemycin – *Streptomyces venezuelae*. **F)** PchG, the stand-alone tailoring enzyme, is a NADPH-dependent reductase converting the second thiazoline of the PchF intermediate to a thiazolidine. **G)** Irp3 (PDB: 5KVS), the PchG homologue for the production of yersiniabactin, is dimeric. The M-N loop closes the active site of an opposing monomer, with $NADP^+$ (green sticks) and substrate analog bound (cyan sticks).

Epimerase domain of PchE. PchE incorporates the first L-cysteine into the growing chain, which is cyclized and epimerized to a D-thiazoline in the final pyochelin product. The condensation domain is specialized, also being called a cyclization domain, because it cyclizes the cysteine to make the thiazoline ring. Cyclization domains are common in NRPS assembly lines generating natural products with thiazoline or oxazoline derivatives.[9, 55] The substrate for the epimerase domain is not a free cysteine amino acid, but is instead the product of the condensation reaction of L-cysteine with the upstream salicylic acid.[32] Experiments designed to determine if the conjugated L-cysteine was cyclized before or after epimerization were inconclusive, although the authors lean toward epimerase activity preceding cyclization (**Figure 1-4B, upper path**).[32] However, this would require an additional domain movement. If the cysteine is epimerized before cyclization, the peptide bond would be formed in the condensation domain, epimerized in the stuffed tailoring domain, and then move back to the condensation domain for cyclization. The peptide would only have to visit the condensation domain once if both condensation and cyclization occur concurrently or sequentially, before epimerization (**Figure 1-4B, lower path**).

Currently there is no determined structure for any stuffed-epimerase domains, but structures of the independent epimerase domain from TycA (A-P-E)[56] of tyrocidine biosynthesis and an apo- and holo-P-epimerase didomain structure from GrsA (A-P-E)[57] of gramicidin S biosynthesis have been determined, both from *Bacillus brevis*. The canonical epimerase domains of TycA and GrsA have similar active site motifs and share secondary structure homology to C-domains.[58] However, the epimerase domain of PchE (PchE-E) does not share any sequence similarity with TycA or GrsA epimerase domains and is considered a noncanonical epimerase domain.[32]

Methyltransferase domain of PchF. PchF adds the second L-cysteine to the growing peptide chain. After condensation to the upstream peptide and cyclization, the thiazoline is reduced to thiazolidine by NADPH-dependent PchG, described below. The stuffed methytransferase domain (M-domain) of PchF catalyzes an *S*-adenosyl methionine (AdoMet)-dependent methyl transfer to the nitrogen of the thiazolidine ring (**Figure 1-4C**).[31, 33] The M-domain is a Class I methyltransferase possessing the consensus (GxG) AdoMet-binding motif.[59] MtfA, a AdoMet-dependent *N*-methyltransferase accessory protein of vancomycin biosynthesis from *Amycolatopsis orientalis*, shares a 32% sequence identity with the M-domain of PchF. MtfA transfers the AdoMet methyl group to the backbone nitrogen of leucine in the vancomycin-type antibiotics.[60] Grouped into five different classes, methyltransferases exhibit diverse tertiary structure and different binding modes of the cofactor and substrate. AdoMet, the methyltransferase product *S*-adenyosylhomocysteine (AdoHCys), and the inhibitor sinefungin were separately crystallized with MtfA (**Figure 1-4D**).[60] AdoMet and AdoHCys bound in the consensus AdoMet-binding motif of the Rossmann fold, corresponding to binding modes seen in other Class I AdoMet-dependent methyltransferases.[60]

Stuffed epimerase as a defunct methyltransferase. Interestingly, the epimerase domain of PchE and the methyltransferase domain of PchF share ~21% sequence identity and ~55% sequence similarity, suggesting the epimerase domain of PchE is more similar to a dysfunctional AdoMet-dependent M-domain than a canonical epimerase domain. This can be rationalized mechanistically, since the first step of the methyltransferase activity is removal of a proton to generate an enol that is randomly reprotonated in the absence of AdoMet, racemizing the

position. Only the appropriate epimer is a substrate for subsequent steps. Indeed, the PchE-epimerase possesses a Gly to Arg mutation in the conserved AdoMet-binding motif (RxG), suggesting interruption of AdoMet binding.[32, 52] This hypothesis can be applied more broadly to other pyochelin-like natural products. The biosynthesis of enantio-pyochelin (**Figure 1-4E**), the enantiomer of pyochelin generated by *Pseudomonas fluorescens* under iron-limiting conditions, lacks the stuffed epimerase domain.[61] A stuffed epimerase of the multi-module NRPS called high molecular weight protein 2 (HMWP2) in yersiniabactin biosynthesis may also be a defunct methyltransferase, sharing the same substrate as PchE and possessing a Gly to Arg mutation in the AdoMet-binding motif (GxRxG).[62] Watasemycin, an NRPS natural product from *Streptomyces venezuelae*, possesses a stuffed methyltransferase domain with neither methyltransferase nor epimerase activity, though it has the same domain architecture and substrates as PchE.[63] Further structure-function studies of the aforementioned enzymes will be necessary to elucidate the differences between the methyltransferase, epimerase, and non-functional stuffed domains of pyochelin-like natural products. Stuffed domains have also been with other, unrelated catalytic activities, such as the oxidase activity of epothilone biosynthesis[64] and the monooxygenase activity of myxothiozole[65] and thuggacin[66] biosynthesis.

Stand-alone tailoring domain: PchG

The terminal cyclized cysteine must be reduced to a thiazolidine while tethered to the Ppant tail on PchF, prior to final methylation and release of the mature pyochelin by the thioesterase domain (**Figure 1-4F**).[27] PchG, a stand-alone tailoring enzyme, is proposed to perform the reduction by proton donation from a general acid and subsequent hydride transfer from NADPH. Structures of a functionally homologous enzyme of yersiniabactin biosynthesis,[67] Irp3, have been

determined: an apo-form, $NADP^+$-bound form, and with $NADP^+$- and substrate analog-bound (**Figure 1-4G**).[68, 69] Irp3 is structurally similar to sugar oxidoreductases and to biliverdin reductase and possesses an N-terminal Rossmann fold domain for binding NADPH cofactor and a C-terminal α/β domain for substrate binding and dimerization.[69] Upon binding of a substrate analogue, a long loop creates a lid over the active site of the opposing monomer leaving a tunnel that is the proposed entrance of the Ppant tether.[68] As the authors point out, the binding of the substrate analog observed in this study may not be physiologically relevant, because the substrate analog binds in a position that would promote incorrect stereochemistry if hydride transfer were to take place.[68]

Conclusion

Siderophores are small, iron-chelating molecules generated by NRPS systems. NRPS structures have provided insight into how these nanomachines form peptide bonds outside of the ribosome. Recently, several full-module structures have elucidated the different domain interactions and domain alternations that occur during peptide assembly. Additional full-module and cross-module structures are necessary to understand domain interactions between modules. While NRPS enzymes generate peptide assembly, accessory and tailoring enzymes provide the chemical diversity that make each natural product unique. Using pyochelin biosynthesis as a model system, we have provided examples of the differing accessory and tailoring activities required to generate even a small natural product. In pyochelin biosynthesis, accessory proteins are required to make the salicylate that caps the N-terminal end. Tailoring domains in the generation of pyochelin fall into two categories. Epimerase and methyltransferase domains are stuffed within the A_{sub}-domain. This understudied class of tailoring domains may provide a

unique assembly line logic. Pyochelin biosynthesis also requires a stand-alone tailoring reductase. While NRPS enzymes have been the subject of considerable investigation and the field is considered mature, there are still many unanswered questions regarding peptide bond formation and the decoration of these peptides to generate the chemical structures that provide biodiversity.

References

[1] Hood, M. I., and Skaar, E. P. (2012) Nutritional immunity: transition metals at the pathogen-host interface, *Nature Reviews Microbiology 10*, 525-537.

[2] Drechsel, H., and Jung, G. (1998) Peptide siderophores, *J Pept Sci 4*, 147-181.

[3] Oves-Costales, D., Kadi, N., and Challis, G. L. (2009) The long-overlooked enzymology of a nonribosomal peptide synthetase-independent pathway for virulence-conferring siderophore biosynthesis, *Chem Commun (Camb)*, 6530-6541.

[4] Gulick, A. M. (2017) Nonribosomal peptide synthetase biosynthetic clusters of ESKAPE pathogens, *Nat Prod Rep 34*, 981-1009.

[5] Lamb, A. L. (2015) Breaking a pathogen's iron will: Inhibiting siderophore production as an antimicrobial strategy, *Biochim Biophys Acta 1854*, 1054-1070.

[6] Quadri, L. E. (2007) Strategic paradigm shifts in the antimicrobial drug discovery process of the 21st century, *Infect Disord Drug Targets 7*, 230-237.

[7] Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem Biol 6*, 493-505.

[8] Keating, T. A., Ehmann, D. E., Kohli, R. M., Marshall, C. G., Trauger, J. W., and Walsh, C. T. (2001) Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis, *Chembiochem 2*, 99-107.

[9] Bloudoff, K., and Schmeing, T. M. (2017) Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity, *Biochim Biophys Acta 1865*, 1587-1604.

[10] Lai, J. R., Koglin, A., and Walsh, C. T. (2006) Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis, *Biochemistry 45*, 14869-14879.

[11] Koglin, A., and Walsh, C. T. (2009) Structural insights into nonribosomal peptide enzymatic assembly lines, *Nat Prod Rep 26*, 987-1000.

[12] Miller, B. R., and Gulick, A. M. (2016) Structural Biology of Nonribosomal Peptide Synthetases, *Methods Mol Biol 1401*, 3-29.

[13] Gulick, A. M. (2009) Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase, *ACS Chem Biol 4*, 811-827.

[14] Tanovic, A., Samel, S. A., Essen, L. O., and Marahiel, M. A. (2008) Crystal structure of the termination module of a nonribosomal peptide synthetase, *Science 321*, 659-663.

[15] Drake, E. J., Miller, B. R., Shi, C., Tarrasch, J. T., Sundlov, J. A., Allen, C. L., Skiniotis, G., Aldrich, C. C., and Gulick, A. M. (2016) Structures of two distinct conformations of holo-non-ribosomal peptide synthetases, *Nature 529*, 235-238.

[16] Reimer, J. M., Aloise, M. N., Harrison, P. M., and Schmeing, T. M. (2016) Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase, *Nature 529*, 239-242.

[17] Drake, E. J., Nicolai, D. A., and Gulick, A. M. (2006) Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain, *Chem Biol 13*, 409-419.

[18] Liu, Y., and Bruner, S. D. (2007) Rational manipulation of carrier-domain geometry in nonribosomal peptide synthetases, *Chembiochem 8*, 617-621.

[19] Tarry, M. J., Haque, A. S., Bui, K. H., and Schmeing, T. M. (2017) X-Ray Crystallography and Electron Microscopy of Cross- and Multi-Module Nonribosomal Peptide Synthetase Proteins Reveal a Flexible Architecture, *Structure 25*, 783-793 e784.
[20] Heemstra, J. R., Jr., Walsh, C. T., and Sattely, E. S. (2009) Enzymatic tailoring of ornithine in the biosynthesis of the Rhizobium cyclic trihydroxamate siderophore vicibactin, *J Am Chem Soc 131*, 15317-15329.
[21] Zhang, W., Heemstra, J. R., Jr., Walsh, C. T., and Imker, H. J. (2010) Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins, *Biochemistry 49*, 9946-9947.
[22] Drake, E. J., Cao, J., Qu, J., Shah, M. B., Straubinger, R. M., and Gulick, A. M. (2007) The 1.8 A crystal structure of PA2412, an MbtH-like protein from the pyoverdine cluster of Pseudomonas aeruginosa, *J Biol Chem 282*, 20425-20434.
[23] Herbst, D. A., Boll, B., Zocher, G., Stehle, T., and Heide, L. (2013) Structural basis of the interaction of MbtH-like proteins, putative regulators of nonribosomal peptide biosynthesis, with adenylating enzymes, *J Biol Chem 288*, 1991-2003.
[24] Miller, B. R., Drake, E. J., Shi, C., Aldrich, C. C., and Gulick, A. M. (2016) Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture, *J Biol Chem 291*, 22559-22571.
[25] Felnagle, E. A., Barkei, J. J., Park, H., Podevels, A. M., McMahon, M. D., Drott, D. W., and Thomas, M. G. (2010) MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases, *Biochemistry 49*, 8815-8817.
[26] Reimmann, C., Serino, L., Beyeler, M., and Haas, D. (1998) Dihydroaeruginoic acid synthetase and pyochelin synthetase, products of the pchEF genes, are induced by extracellular pyochelin in Pseudomonas aeruginosa, *Microbiology 144 (Pt 11)*, 3135-3148.
[27] Reimmann, C., Patel, H. M., Serino, L., Barone, M., Walsh, C. T., and Haas, D. (2001) Essential PchG-dependent reduction in pyochelin biosynthesis of Pseudomonas aeruginosa, *J Bacteriol 183*, 813-820.
[28] Serino, L., Reimmann, C., Visca, P., Beyeler, M., Chiesa, V. D., and Haas, D. (1997) Biosynthesis of pyochelin and dihydroaeruginoic acid requires the iron-regulated pchDCBA operon in Pseudomonas aeruginosa, *J Bacteriol 179*, 248-257.
[29] Gaille, C., Reimmann, C., and Haas, D. (2003) Isochorismate synthase (PchA), the first and rate-limiting enzyme in salicylate biosynthesis of Pseudomonas aeruginosa, *J Biol Chem 278*, 16893-16898.
[30] Gaille, C., Kast, P., and Haas, D. (2002) Salicylate biosynthesis in Pseudomonas aeruginosa. Purification and characterization of PchB, a novel bifunctional enzyme displaying isochorismate pyruvate-lyase and chorismate mutase activities, *J Biol Chem 277*, 21768-21775.
[31] Quadri, L. E., Keating, T. A., Patel, H. M., and Walsh, C. T. (1999) Assembly of the Pseudomonas aeruginosa nonribosomal peptide siderophore pyochelin: In vitro reconstitution of aryl-4, 2-bisthiazoline synthetase activity from PchD, PchE, and PchF, *Biochemistry 38*, 14941-14954.
[32] Patel, H. M., Tao, J., and Walsh, C. T. (2003) Epimerization of an L-cysteinyl to a D-cysteinyl residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from Pseudomonas aeruginosa, *Biochemistry 42*, 10514-10527.

[33] Patel, H. M., and Walsh, C. T. (2001) In vitro reconstitution of the Pseudomonas aeruginosa nonribosomal peptide synthesis of pyochelin: characterization of backbone tailoring thiazoline reductase and N-methyltransferase activities, *Biochemistry 40*, 9023-9031.
[34] Zwahlen, J., Kolappan, S., Zhou, R., Kisker, C., and Tonge, P. J. (2007) Structure and mechanism of MbtI, the salicylate synthase from Mycobacterium tuberculosis, *Biochemistry 46*, 954-964.
[35] Kolappan, S., Zwahlen, J., Zhou, R., Truglio, J. J., Tonge, P. J., and Kisker, C. (2007) Lysine 190 is the catalytic base in MenF, the menaquinone-specific isochorismate synthase from Escherichia coli: implications for an enzyme family, *Biochemistry 46*, 946-953.
[36] He, Z., Stigers Lavoie, K. D., Bartlett, P. A., and Toney, M. D. (2004) Conservation of mechanism in three chorismate-utilizing enzymes, *J Am Chem Soc 126*, 2378-2385.
[37] Sridharan, S., Howard, N., Kerbarh, O., Blaszczyk, M., Abell, C., and Blundell, T. L. (2010) Crystal structure of Escherichia coli enterobactin-specific isochorismate synthase (EntC) bound to its reaction product isochorismate: implications for the enzyme mechanism and differential activity of chorismate-utilizing enzymes, *J Mol Biol 397*, 290-300.
[38] Parsons, J. F., Shi, K. M., and Ladner, J. E. (2008) Structure of isochorismate synthase in complex with magnesium, *Acta Crystallogr D Biol Crystallogr 64*, 607-610.
[39] Meneely, K. M., Sundlov, J. A., Gulick, A. M., Moran, G. R., and Lamb, A. L. (2016) An Open and Shut Case: The Interaction of Magnesium with MST Enzymes, *J Am Chem Soc 138*, 9277-9293.
[40] Harrison, A. J., Yu, M., Gardenborg, T., Middleditch, M., Ramsay, R. J., Baker, E. N., and Lott, J. S. (2006) The structure of MbtI from Mycobacterium tuberculosis, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate synthase, *J Bacteriol 188*, 6081-6091.
[41] Chi, G., Manos-Turvey, A., O'Connor, P. D., Johnston, J. M., Evans, G. L., Baker, E. N., Payne, R. J., Lott, J. S., and Bulloch, E. M. (2012) Implications of binding mode and active site flexibility for inhibitor potency against the salicylate synthase from Mycobacterium tuberculosis, *Biochemistry 51*, 4868-4879.
[42] Lamb, A. L. (2011) Pericyclic reactions catalyzed by chorismate-utilizing enzymes, *Biochemistry 50*, 7476-7483.
[43] Meneely, K. M., Luo, Q., Dhar, P., and Lamb, A. L. (2013) Lysine221 is the general base residue of the isochorismate synthase from Pseudomonas aeruginosa (PchA) in a reaction that is diffusion limited, *Arch Biochem Biophys 538*, 49-56.
[44] Luo, Q., Olucha, J., and Lamb, A. L. (2009) Structure-function analyses of isochorismate-pyruvate lyase from Pseudomonas aeruginosa suggest differing catalytic mechanisms for the two pericyclic reactions of this bifunctional enzyme, *Biochemistry 48*, 5239-5245.
[45] Gustin DJ, M. P., Kast P, Wiest O, Lee L, Cleland WW, Hilvert D. (1999) Heavy Atom Isotope Effects Reveal a Highly Polarized Transition State for Chorismate Mutase, *J. Am. Chem. Soc. 121*, 1756-1757.
[46] DeClue, M. S., Baldridge, K. K., Kunzler, D. E., Kast, P., and Hilvert, D. (2005) Isochorismate pyruvate lyase: a pericyclic reaction mechanism?, *J Am Chem Soc 127*, 15002-15003.
[47] Marti, S., Andres, J., Moliner, V., Silla, E., Tunon, I., and Bertran, J. (2009) Mechanism and plasticity of isochorismate pyruvate lyase: a computational study, *J Am Chem Soc 131*, 16156-16161.

[48] Kerbarh, O., Chirgadze, D. Y., Blundell, T. L., and Abell, C. (2006) Crystal structures of Yersinia enterocolitica salicylate synthase and its complex with the reaction products salicylate and pyruvate, *J Mol Biol 357*, 524-534.
[49] Culbertson, J. E., Chung, D., Ziebart, K. T., Espiritu, E., and Toney, M. D. (2015) Conversion of aminodeoxychorismate synthase into anthranilate synthase with Janus mutations: mechanism of pyruvate elimination catalyzed by chorismate enzymes, *Biochemistry 54*, 2372-2384.
[50] Zaitseva, J., Lu, J., Olechoski, K. L., and Lamb, A. L. (2006) Two crystal structures of the isochorismate pyruvate lyase from Pseudomonas aeruginosa, *J Biol Chem 281*, 33441-33449.
[51] Olucha, J., Ouellette, A. N., Luo, Q., and Lamb, A. L. (2011) pH Dependence of catalysis by Pseudomonas aeruginosa isochorismate-pyruvate lyase: implications for transition state stabilization and the role of lysine 42, *Biochemistry 50*, 7198-7207.
[52] Labby, K. J., Watsula, S. G., and Garneau-Tsodikova, S. (2015) Interrupted adenylation domains: unique bifunctional enzymes involved in nonribosomal peptide biosynthesis, *Nat Prod Rep 32*, 641-653.
[53] Zolova, O. E., and Garneau-Tsodikova, S. (2014) KtzJ-dependent serine activation and O-methylation by KtzH for kutznerides biosynthesis, *J Antibiot (Tokyo) 67*, 59-64.
[54] Al-Mestarihi, A. H., Villamizar, G., Fernandez, J., Zolova, O. E., Lombo, F., and Garneau-Tsodikova, S. (2014) Adenylation and S-methylation of cysteine by the bifunctional enzyme TioN in thiocoraline biosynthesis, *J Am Chem Soc 136*, 17350-17354.
[55] Konz, D., Klens, A., Schorgendorfer, K., and Marahiel, M. A. (1997) The bacitracin biosynthesis operon of Bacillus licheniformis ATCC 10716: molecular characterization of three multi-modular peptide synthetases, *Chem Biol 4*, 927-937.
[56] Samel, S. A., Czodrowski, P., and Essen, L. O. (2014) Structure of the epimerization domain of tyrocidine synthetase A, *Acta Crystallogr D Biol Crystallogr 70*, 1442-1452.
[57] Chen, W. H., Li, K., Guntaka, N. S., and Bruner, S. D. (2016) Interdomain and Intermodule Organization in Epimerization Domain Containing Nonribosomal Peptide Synthetases, *ACS Chem Biol 11*, 2293-2303.
[58] Keating, T. A., Marshall, C. G., Walsh, C. T., and Keating, A. E. (2002) The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains, *Nat Struct Biol 9*, 522-526.
[59] Schubert, H. L., Blumenthal, R. M., and Cheng, X. (2003) Many paths to methyltransfer: a chronicle of convergence, *Trends Biochem Sci 28*, 329-335.
[60] Shi, R., Proteau, A., Wagner, J., Cui, Q., Purisima, E. O., Matte, A., and Cygler, M. (2009) Trapping open and closed forms of FitE: a group III periplasmic binding protein, *Proteins 75*, 598-609.
[61] Youard, Z. A., Mislin, G. L., Majcherczyk, P. A., Schalk, I. J., and Reimmann, C. (2007) Pseudomonas fluorescens CHA0 produces enantio-pyochelin, the optical antipode of the Pseudomonas aeruginosa siderophore pyochelin, *J Biol Chem 282*, 35546-35553.
[62] Gehring, A. M., DeMoll, E., Fetherston, J. D., Mori, I., Mayhew, G. F., Blattner, F. R., Walsh, C. T., and Perry, R. D. (1998) Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by Yersinia pestis, *Chem Biol 5*, 573-586.
[63] Inahashi, Y., Zhou, S., Bibb, M. J., Song, L., Al-Bassam, M. M., Bibb, M. J., and Challis, G. L. (2017) Watasemycin biosynthesis in Streptomyces venezuelae: thiazoline C-methylation by a type B radical-AdoMet methylase homologue, *Chem Sci 8*, 2823-2831.

[64] Schneider, T. L., Shen, B., and Walsh, C. T. (2003) Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation, *Biochemistry 42*, 9722-9730.
[65] Perlova, O., Fu, J., Kuhlmann, S., Krug, D., Stewart, A. F., Zhang, Y., and Muller, R. (2006) Reconstitution of the myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in Myxococcus xanthus, *Appl Environ Microbiol 72*, 7485-7494.
[66] Buntin, K., Irschik, H., Weissman, K. J., Luxenburger, E., Blocker, H., and Muller, R. (2010) Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity, *Chem Biol 17*, 342-356.
[67] Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of Yersinia pestis, *Chem Biol 9*, 333-344.
[68] Meneely, K. M., Ronnebaum, T. A., Riley, A. P., Prisinzano, T. E., and Lamb, A. L. (2016) Holo Structure and Steady State Kinetics of the Thiazolinyl Imine Reductases for Siderophore Biosynthesis, *Biochemistry 55*, 5423-5433.
[69] Meneely, K. M., and Lamb, A. L. (2012) Two structures of a thiazolinyl imine reductase from Yersinia enterocolitica provide insight into catalysis and binding to the nonribosomal peptide synthetase module of HMWP1, *Biochemistry 51*, 9002-9013.

Chapter 2

Stuffed Methyltransferase Catalyzes Penultimate Step of Pyochelin Biosynthesis

*This chapter has been submitted to Biochemistry and includes guided instruction and assistance from Prisinzano, T. E., Booker, S.J., and Lamb, A. L.

Introduction

Bacteria, fungi, and plants generate siderophores, small molecules with high affinity for ferric iron, to scavenge the surrounding environment during times of iron starvation. Indeed, many siderophores have been identified as virulence factors in pathogenic bacteria due to the metal-limiting environment of the host species. Siderophores are classified by the chemical nature of the oxygen involved in ferric iron coordination.[1] Pyochelin is a phenolate siderophore from *Pseudomonas aeruginosa*, a commensal and opportunistic gram-negative pathogen, notorious for causing infections in immunocompromised patients, burn-wound victims, and lung infections of cystic fibrosis patients.[2-5]

The precursors of pyochelin are chorismate and two cysteine molecules (**Figure 2-1A**).[6-8] Chorismate, from the shikimate pathway, is converted to salicylate in two steps, and is the initiating hydroxy acid. The cysteine molecules are added in succession by nonribosomal peptide synthetases (NRPSs), forming peptide bonds.[6] Encoded within the *pch* operon are tailoring domains, which perform the epimerization, reduction and methyltransferase activities required to make the final bioactive compound.[6, 9-11] Pyochelin is a tricyclic acid: the 2'-hydroxyphenyl (HP) ring derived from salicylate attached to a thiazoline (T) ring from one

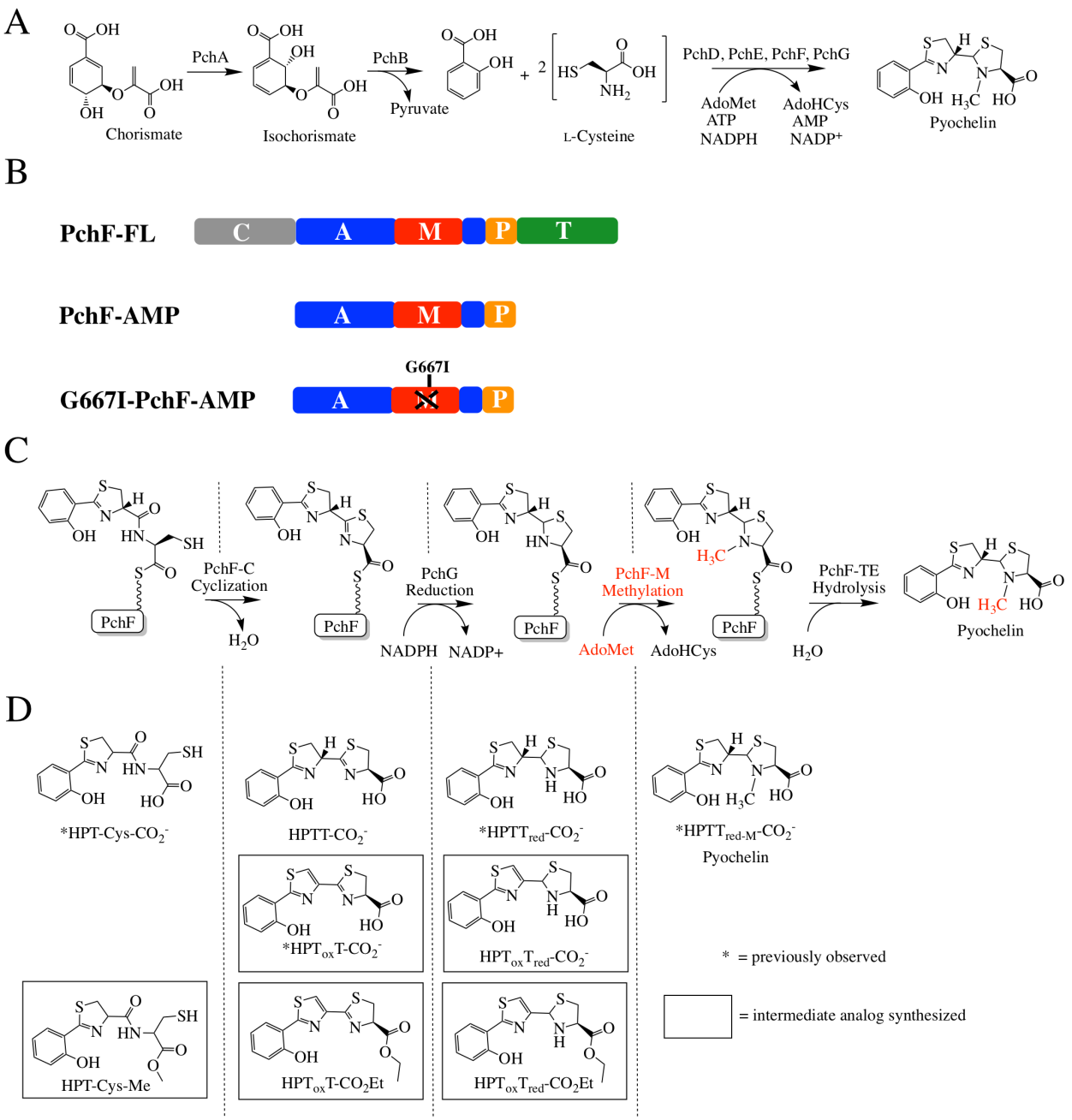


Figure 2-1: Synthesis of pyochelin. A) Biosynthetic pathway for pyochelin production. PchA, an isochorismate synthase, and PchB, an isochorismate pyruvate lyase, convert chorismate to salicylate. PchD-G consists of 12 different domains constituting three NRPS modules with 3 tailoring domains. Together, they assemble pyochelin from salicylate and two L-cysteines, with co-substrates AdoMet, ATP, and NADPH. AdoMet = *S*-adenosylmethionine, AdoHcys = *S*-adenosylhomocysteine. **B)** PchF is a full-NRPS module consisting of 5 different domains: a condensation domain (C; grey), adenylation domain (A; blue), stuffed methyltransferase tailoring domain (M; red), peptidyl carrier protein domain (P; yellow), and thioesterase domain (T; green).

PchF-FL refers to the "full-length" protein, whereas PchF-AMP only includes the stuffed adenylation methyltransferase-didomain and the peptidyl carrier domain. G667I-PchF-AMP has a mutation to the proposed AdoMet binding domain of the variant PchF-AMP. **C)** The proposed order of chemistries performed by PchF. PchF incorporates L-cysteine to the growing peptide chain. The C-domain of PchF continues chain elongation by generating a peptide bond between L-cysteine and the upstream hydroxyphenyl-D-thiazoline moiety of PchE (an intermediate attached to the Ppant of the P-domain of PchE). The C-domain then cyclizes the cysteine to form a hydroxyphenyl-(D)-thiazoline-(L)-thiazoline intermediate. PchG (green) reduces the terminal thiazoline of the hydroxyphenyl-bis-heterocycle to thiazolidine using NADPH. The stuffed methyltransferase domain of PchF then catalyzes the AdoMet-dependent methyl transfer (highlighted in red) onto the nitrogen of the newly reduced thiazolidine ring before transferring the complete heterocyclic siderophore to the T-domain releasing the fully mature pyochelin by hydrolysis in the T-domain. The alternate synthetic schemes where methyl transfer occurs at a different position in the sequence of biosynthesis are shown in **Figure 2-2**. **D)** During reconstitution assays, premature hydrolysis releases peptidyl intermediates during biosynthesis. Intermediates previously recovered by Walsh and colleagues are noted with an asterisk. Intermediate analogs used as potential substrates in this study are surrounded by a square. Nomenclature of intermediates is as follows: HP, hydroxyphenyl; T_{ox}, thiazole; T, thiazoline; T_{red}, thiazolidine; CO_2- carboxyl; Et, ethyl; Me/M, methyl.

cyclized and epimerized cysteine, followed by a methylated 4'-thiazolidine ($T_{red\text{-}M}$) heterocycle from the second cyclized cysteine.

NRPSs are multi-domain, multi-catalytic proteins arranged in modules to activate and add one substituent each to the growing chain, with each module containing a condensation domain, an adenylation domain and a carrier domain. The NRPS assembly line uses a 4'phosphopantetheinyl (Ppant) thiotemplate tethering system reminiscent of polyketide and fatty acid biosynthesis. This means that the peptide bond forming and tailoring chemistries are performed while the intermediates are covalently attached to a carrier domain. The NRPS tailoring domains are affiliated with particular modules in several geometries. The most common and best characterized geometry has the tailoring domain included at the end of a module following the carrier domain, as an independent domain that works *in cis* with the peptide bond forming domains of the module. Pyochelin biosynthesis lacks this most common type of tailoring domain. Some tailoring domains are stand-alone: independently expressed proteins acting *in trans*, such as the reductase of pyochelin biosynthesis (PchG).[11] Finally, there are tailoring domains that are incorporated within the adenylation domain.[12] Adenylation domains activate the substrate and attach it to the thiotemplate tethering system. The "stuffed" tailoring domains replace a loop within an adenylation domain, making a true di-domain (**Figure 2-1B**).[13] Pyochelin biosynthesis uses two stuffed tailoring domains. The first is an adenylation-epimerase didomain in the protein named PchE, and the second is an adenylation-methyltransferase didomain in PchF.[10, 12-14]

Examples of stuffed tailoring domains, also known as "interrupted adenylation domians",[12] participating in natural product biosynthesis include oxygenases (i.e. epothilone,[15] myxothiazol,[16] thuggacins)[17], ketoreductases (i.e. cereulide)[18], methyltransferases (i.e.

thiocoraline,[1, 19] cyclosporine,[20] echinomycin,[21] enniatins,[22] micacocidin,[23] microcystin,[24] pyochelin,[9, 10] yersiniabactin[9]), and epimerases (i.e. pyochelin,[9, 10] yersiniabactin[9]). Importantly, some of the natural products synthesized using stuffed tailoring domains have been isolated and experimentally show anticancer (i.e. epothilone, thiocoraline, echinomycin), antifungal (i.e. myxothiazol), antibacterial (i.e. myxothiazol, thuggacin, micacocidin), or immunosuppressive (i.e. cyclosporine) activity. However, the mechanistic and structural characterization of stuffed tailoring domains has only just begun.

Recently, mechanistic and structural studies of the stuffed methyltransferase domains from the biosynthesis of the microbial depsipeptides thiocoraline and kutzneride have been reported.[19, 25-27] A main goal of these studies was to determine when in the biosynthetic pathway methylation occurs, with four possibilities: methylation of 1) the free amino acid, 2) the aminoacyl-AMP intermediate (after the amino acid is activated by ATP in the adenylation domain), 3) after thiolation (after the amino acid is attached to the Ppant tether of the peptidyl carrier protein domain), or 4) after condensation to the upstream peptide. Evidence suggests that the stuffed methyltransferase does not use the free amino acid as a substrate, but rather the substrate is either the aminoacyl-AMP intermediate or the amino acid attached to the carrier domain (after thiolation).[19, 25-27] The possibility remains that methyl transfer occurs after elongation with the upstream peptide; however, thiocoraline and kutzneride are complex natural products, making intermediate analog synthesis to test the hypothesis of late stage methylation difficult. Contradictory to the proposed model, isolates observed during reconstitution assays of epothilone biosynthesis and mechanistic studies of EpoB suggest tailoring occurs after condensation to the upstream peptide chain.[14, 15] Epothilone biosynthesis employs a hybrid NRPS/PKS assembly line in which EpoB, a protein encoding a single NRPS module, has a

stuffed oxidase domain. The EpoB substrate is simply a 2-methylthiazoline attached to the Ppant of the carrier domain.

Here, using the termination module of pyochelin biosynthesis found in the protein PchF, we test the opposing hypotheses for the order of chemistry. An order of catalytic events has been proposed in the literature (**Figure 2-1C**). The PchF adenylation domain activates the cysteine substrate and attaches the cysteine to the thiotemplate of the carrier domain.[6] The condensation domain of PchF links the nascent chain donated from the PchE carrier domain (HPT) to the cysteine on the carrier domain of PchF (HPT-Cys), and subsequently cyclizes this cysteine generating a thiazoline ring (HPTT). While still tethered to the carrier domain of PchF, PchG reduces the second thiazoline ring generating a thiazolidine ($HPTT_{red}$). The *N*-methyltransferase stuffed domain in PchF modifies the ring to generate an *N*-methylated thiazolidine ($HPTT_{red\text{-}M}$), the final proposed synthetic step before release of the mature pyochelin ($HPTT_{red\text{-}M}CO_2^-$) by the thioesterase. Two intermediates have been hydrolyzed from the thiotemplate tether, isolated and characterized during pyochelin reconstitution assays: HPT-Cys and $HPTT_{red}$-CO_2^- (**Figure 2-1D**).[6, 11] The chemical release and isolation process also produced a more stable, off-pathway, product hypothesized to be the result of air oxidation: $HPT_{ox}T$- CO_2^-.[6, 11] Since none of these isolated compounds contain the *N*-methyl on the thiazolidine, the order of the steps proposed for PchF is: adenylation, thiolation, condensation, cyclization, reduction, methylation and hydrolysis. However, there are some contradictory data that suggest that methylation may proceed before reduction: labeled methyl groups, from either ^{14}C-*S*-adenosylmethionine (AdoMet) or ^{3}H-AdoMet were incorporated into the product in the absence of the reductase or NADPH.[10] Altogether, the order of catalysis, and in particular when methyl transfer occurs remains elusive, leaving four reasonable pathways: 1) before condensation, 2) before cyclization,

3) before reduction, or 4) before hydrolysis (**Figure 2-2**). Herein, we synthesized a series of substrate analogues and developed kinetic analyses for the adenylation and methyltransferase activities of PchF and PchF variants to establish the order of catalytic steps performed by PchF in pyochelin biosynthesis. Additionally, onium chalcogen effects were used to define the mechanism of PchF's methyltransferase reaction.

Materials and Methods

Preparation of full-length PchF (pchf-fl) Overexpression Plasmid. The *pchF* overexpression construct was a gift from the laboratory of Christopher Walsh.[6] The construct received was the kanamycin resistant, pET28b plasmid (Novagen).

Preparation of Overexpression Plasmid containing PchF adenylation, methyltransferase, and peptidyl carrier protein domains (pchf-amp). The variant plasmid was produced by site directed mutagenesis using Herculase polymerase supplemented with 0-10% DMSO and the genomic DNA from *Pseudomonas aeruginosa* PAO1, which was purified using the DNeasy® Blood & Tissue Kit (Qiagen) as per the manufacturer's instructions. The forward primer (5'-ATT AGA CAT ATG GTC GAG GCG CCG CCG CAG G-3') including an *Nde*I site (underlined) and reverse primer (5'-TA ATA CTC GAG TCA GGT TCC GGC GCG CTG CGC AG-3') including an *Xho*I site (underlined) were used to amplify the gene of interest. The amplified sequence was ligated into the kanamycin resistant pET28b plasmid (Novagen) digested with the same restriction enzymes. Upon sequencing, it was determined that there were an extra 2 nucleotides inserted before the stop codon. Therefore, using QuikChange XL (Agilent), a forward primer (5'-GCA GCG CGC CGG AAT GAC TCG AGC AC-3') and reverse primer (5'-

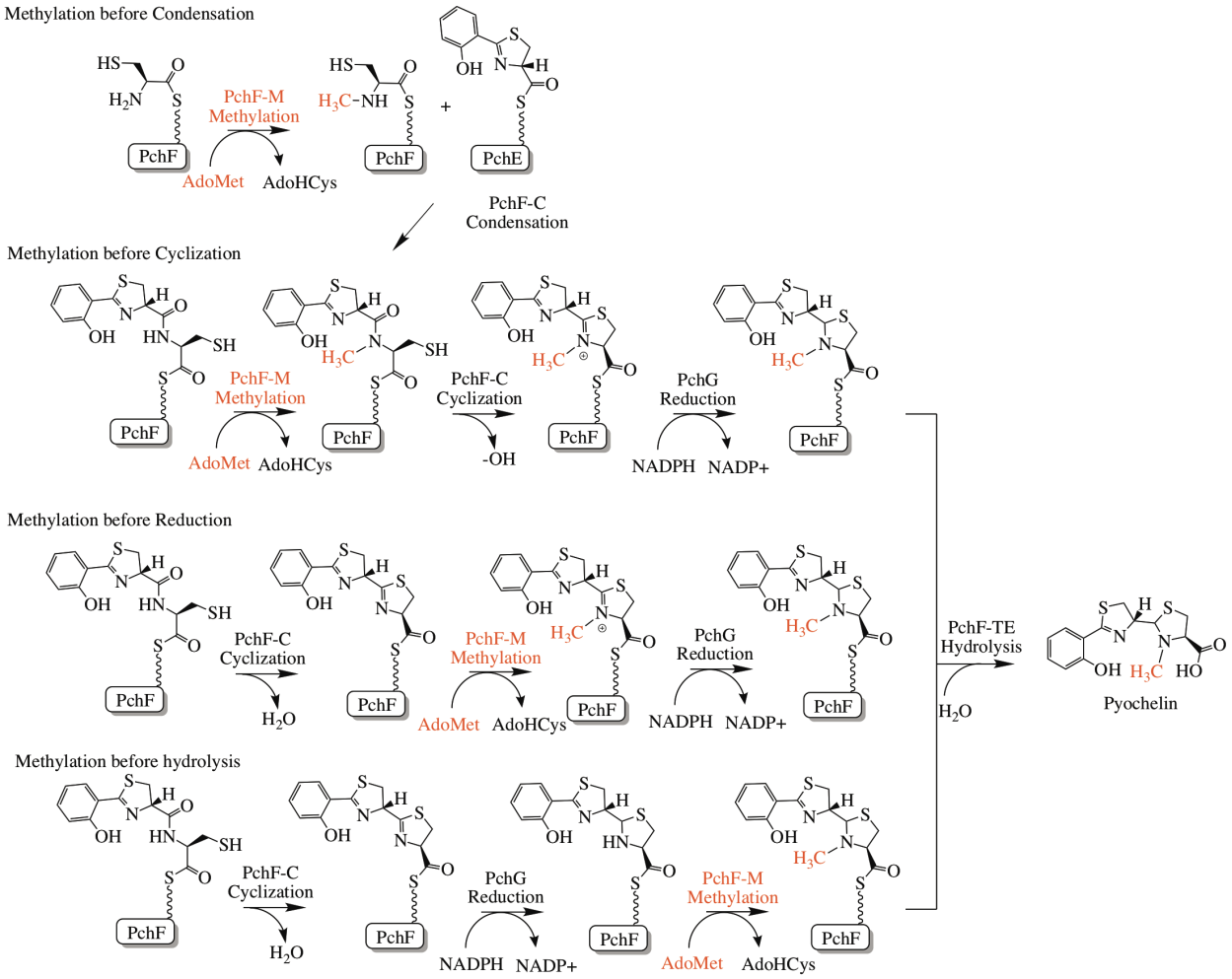


Figure 2-2: Possible biosynthetic strategies for PchF. AdoMet dependent methylation is proposed to take place as the last tailoring step before hydrolysis and the release of the mature peptide; however previous evidence is contradictory. Four potential biosynthetic pathways are shown for methyl transfer: before condensation, before cyclization, before reduction, or directly before hydrolysis. The methyltransferase chemistry is highlighted in red.

GTC CTC GAG TCA TTC CGG CGC GCT GC-3') were designed to remove the 2-nucleotide insertion before the stop codon (underlined). The isolated plasmid sequence encodes the entire adenylation domain, methyltransferase domain, and peptidyl carrier protein domain, residues 502 to 1483 of PchF-FL.

g667i-pchf-amp Overexpression plasmid. The variant plasmid was produced by using the kanamycin resistant *pchF-AMP* plasmid and the QuikChange XL (Agilent) as per the manufacturer's instructions. The forward primer (5'-GCGGCGGCGGTGATGGCGCCAAGCTC-3') and reverse primer (5'-GAGCTTGGCGCCATCACCGCCGCCGC-3') were used to mutate the desired nucleotides (underlined). The isolated plasmid was sequenced showing that only the desired mutations were present.

pa2412 Overexpression Plasmid. The *pa2412* overexpression plasmid was a gift from the laboratory of Andrew Gulick.[28] PA2412 is an MbtH-like protein (MLP) that increases *P. aeruginosa* NRPS protein expression including PchF (data not shown).[29-32] The *pa2412* gene is in a modified pET15b plasmid, conferring ampicillin resistance, containing a 5XHis affinity tag and a TEV protease recognition site in place of the thrombin site.[28]

Methanococcus jannaschii (Mj) S-Adenosyl Methionine (AdoMet) Synthetase Overexpression Plasmid was a gift from Doug Markham from the Fox Chase Cancer Institute.[33]

PchF-FL protein overexpression and purification. BL21 (DE3) *E. coli* containing the *pchf-fl* and *pa2412* expression plasmids were grown in LB broth containing 50 μg/mL kanamycin and 200 μg/mL ampicillin at 37 °C with shaking (250 rpm). Protein expression was induced when the OD_{600} reached ~0.6 by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 200 μM. The temperature was reduced to 22 °C following induction. The cells were harvested by centrifugation (4 230 x g, 10 min, 4 °C) after ~21 hours. The cell pellet was resuspended in 25 mM Tris-HCl pH 8, 500 mM NaCl, 50 mM imidazole (buffer A). Cells were disrupted by French press (35 000 psi), and cellular debris was removed by centrifugation (23 430 x g, 45 min, 4 °C). The supernatant was applied to a chelating Sepharose fast-flow column (Amersham Biosciences) charged with nickel chloride and pre-equilibrated in buffer A. Two column volumes of buffer A were followed by a 15-column volume linear gradient of 0-500 mM imidazole in buffer A. PchF-FL and PA2412 proteins eluted together in 1 peak at ~150 mM imidazole. Fractions containing the proteins of interest were pooled and concentrated using an Amicon stirred cell with YM30. The concentrated fractions were applied to a Superdex 200 size-exclusion column (Amersham Biosciences) equilibrated with 50 mM Tris pH 8, 150 mM sodium citrate, 10% (v/v) glycerol (buffer B). PchF-FL eluted separately from PA2412 and the PchF-FL fractions were concentrated as above to 1.4 mg/mL as determined by the general Beer-Lambert law and the predicted A_{280} of $\varepsilon = 276\ 620\ M^{-1}\ cm^{-1}$ (ProtParam).[34] A total of 14.7 mg was obtained from the 2 L of cell growth. The protein was flash cooled and stored at -80 °C.

PchF-AMP and G667I-PchF-AMP protein overexpression and purification. BL21 (DE3) *E. coli* containing the *pchf-amp or g667i-pchf-amp* and *pa2412* expression plasmids were grown in LB broth containing 50 μg/mL kanamycin and 200 μg/mL ampicillin at 37 °C with shaking (225

rpm). When the OD_{600} reached ~0.6, protein expression was induced by the addition of IPTG to a final concentration of 200 μM and the temperature was reduced to 15°C. The cells were harvested by centrifugation (4 230 x g, 10 min, 4 °C) after ~26 hours. The cell pellet was resuspended in 25 mM Tris-HCl pH 8, 500 mM NaCl, 5 mM imidazole, 10% glycerol (buffer C). Cells were disrupted by use of a French pressure cell (35 000 psi), and cellular debris was removed by centrifugation (11 950 x g, 60 min, 4 °C). The supernatant was applied to a chelating Sepharose fast-flow column (Amersham Biosciences) charged with nickel chloride and pre-equilibrated in buffer C. After injection, the column was washed with four column volumes of buffer C, followed by four column volumes of 50 mM imidazole in buffer C. A linear gradient from 50-500 mM imidazole in buffer C over 6 column volumes allowed elution of PchF-AMP or G667I-PchF-AMP and PA2412 proteins in a single peak around 200 mM imidazole. The pooled fractions were dialyzed using SnakeSkin® Dialysis Tubing (10 kDa cutoff) against 50 mM Tris pH 8, 1 mM DTT, 10% (v/v) glycerol (buffer D). The dialysate was changed for fresh buffer twice: at 1 and 2 hours before exchanging overnight. The protein solution was loaded onto a Source 30Q (GE Healthcare Life Sciences) anion exchange column pre-equilibrated in buffer D. After injection, two column volumes of buffer D were followed by a 0-500 mM NaCl linear gradient over 15 column volumes in buffer D. PchF-AMP or G667I-PchF-AMP eluted at ~150 mM NaCl. The collected fractions were concentrated and applied to a Superdex 200 size-exclusion column (Amersham Biosciences) equilibrated with 50 mM Tris pH 8, 100 mM NaCit, 1 mM DTT, 10% (v/v) glycerol (buffer C). The fractions containing PchF-AMP or G667I-PchF-AMP were pooled and concentrated using a 30 kDa cutoff Amicon ultracentrifugal spin filter. The final concentration of PchF-AMP was 7.8 mg/mL as determined by the general Beer-Lambert law and the predicted A_{280} of $\varepsilon = 149\ 310\ M^{-1}\ cm^{-1}$ (ProtParam).[34] A total of 13.3 mg of

PchF-AMP was obtained from the 3 L cell growth. The protein was flash cooled and stored at -80 °C. G667I-PchF-AMP was concentrated to 1.4 mg/mL, also determine by Beer-Lambert's law and the predicted A_{280} of $\varepsilon = 149\ 310\ M^{-1}\ cm^{-1}$ (ProtParam).[34] A total of 0.621 mg was obtained from the 3 L cell growth. The protein was flash cooled and stored at -80˚C.

Methanococcus jannaschii (Mj) AdoMet Synthetase overexpression and purification. BL21 (DE3) *E. coli* containing the pMJ1208-1 expression plasmid were grown in LB broth containing 100 μg/mL ampicillin at 37 °C with shaking (225 rpm). When the OD_{600} reached ~0.6, protein expression was induced by the addition of IPTG to a final concentration of 500 μM. Cultures were incubated for an additional 5 hours. The cells were harvested by centrifugation (10 820 x g, 10 min, 4 °C). The cell paste (50 g) was resuspended in 50mM HEPES pH 7.5, 10 mM β-mercapto-ethanol, 300 mM NaCl, 5 mM imidazole, supplemented with 1 mg/mL lysozyme, 0.1 mg/mL DNAse I, 1 mM PMSF, and 0.1% Triton X (Buffer E). Cells were lysed by sonication (6 cycles of 45 seconds with 7 minutes between cycles, 4˚C) and subsequently heat denatured (50 min, 75˚C). Cell debris and denaturants were removed by centrifugation (50 000 x g, 60 min, 4˚C). The supernatant was loaded onto a Ni-NTA agarose column equilibrated with Buffer E. The column was washed with five column volumes of Buffer E and eluted with Buffer E supplemented with 500 mM imidazole. Fractions containing protein were concentrated using a 10 kDa cutoff Amicon ultracentrifugal spin filter. The protein was loaded onto a G-25 column equilibrated in 50 mM Tris-HCl pH 8, 50 mM KCl, 0.1% BME, 10% (v/v) glycerol (Buffer F). Protein containing fractions determined by Bradford assay and were collected and concentrated using a 30-kDa cutoff Amicon ultracentrifugal spin filter. The concentrated protein was determined by the general Beer-Lambert law and the predicted A_{280} of $\varepsilon = 24\ 870\ M^{-1}\ cm^{-1}$

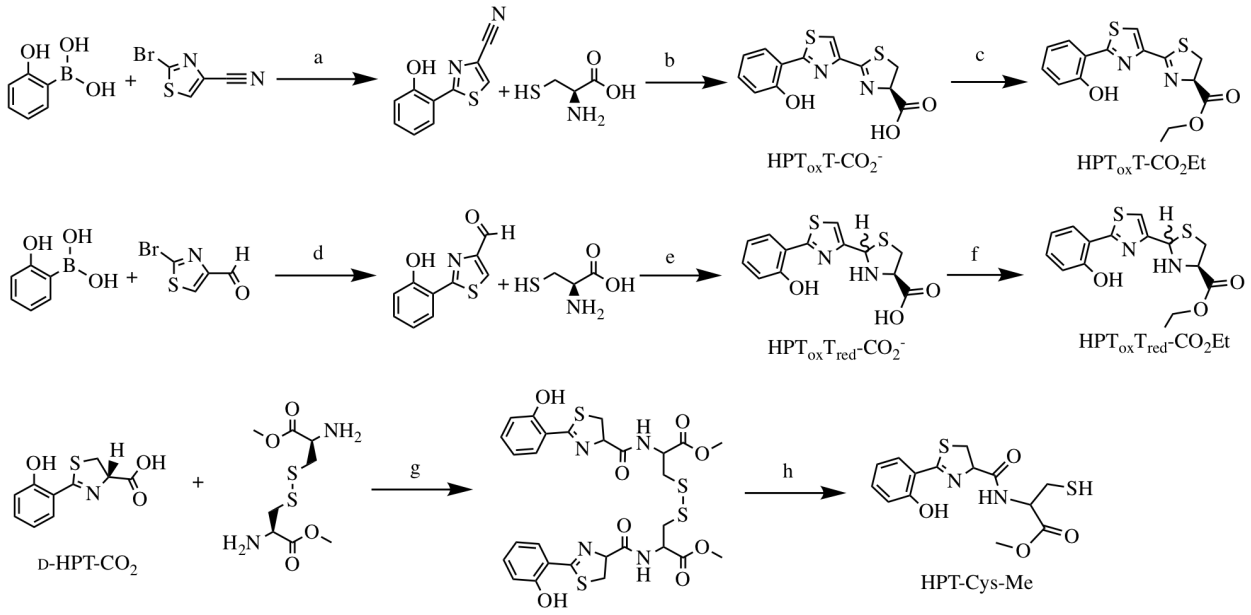
(ProtParam) before being flash cooled and stored at -80 ˚C.[34] A total of 211 mg of *Mj* AdoMet synthetase was obtained from the 10 L of cell growth.

*Preparation of Substrate Analogues (**Scheme 2-1**).*

General experimental procedures. Chemical reagents purchased from commercial suppliers were not further purified prior to use. Flash column chromatography was performed using silica gel (4-63 mm) purchased from Sorbent Technologies. Other chemical separations were performed using a Teledyne Isco CombiFlash Rf. A Biotage microwave reactor was used when microwaving was necessary. ^{1}H and ^{13}C NMR were recorded using a 500 MHz Bruker AVIII spectrometer equipped with a cryogenically-cooled carbon observe probe. Tetramethylsilane was used as an internal standard and chemical shifts (δ) are reported in ppm and coupling constants (*J*) are reported in Hz. High-resolution mass spectometry (HRMS) was performed using an electrospray ion source in either positive or negative mode on an LCT Premier (Micromass Ltd., Manchester UK). Melting points were measured with a Thomas Capillary Melting Point Apparatus and are uncorrected. Chiral analysis was carried out using HPLC, 1 ml/min hexane:isopropanol (90:10) on a CHIRALCEL® OD-H column, monitored at 320 nm unless otherwise specified.

2-(2-hydroxyphenyl)thiazole-4-carbaldehyde

A microwave vial was charged with 2-bromothiazole-4-carbaldehyde (2.08 mmol), 2-hydroxyphenyl boronic acid (2.71 mmol), and $Pd(PPh_3)_4$ (0.104 mmol). The vial was sealed and flushed with argon for 5 minutes. Dimethoxyethane (16 mL) and a 2 M potassium carbonate solution (2.0 mL, 4.17 mmol) were added through the septum, the reaction was stirred at room



Scheme 2-1: Substrate analog synthesis of $HPT_{ox}T\text{-}CO_2^-$, $HPT_{ox}T\text{-}CO_2Et$, $HPT_{ox}T_{red}\text{-}CO_2^-$, $HPT_{ox}T_{red}\text{-}CO_2Et$, and HPT-Cys-Me. Substrate analogs were generated to assess the methyltransferase capabilities of PchF. Reagents and conditions: a) $Pd(PPh_3)_4$, K_2CO_3/DME, H_2O, 150 °C, 49%;[42] b) MeOH, K_2HPO_4(aq) (pH 6.5), 60 °C, 70%;[42] c) $SOCl_2$/EtOH, 40 °C, 49%;[42] d) $Pd(PPh_3)_4$, K_2CO_3/DME, H_2O, 100˚C, 70%; e) EtOH/KCO_2CH_3 (aq), RT, 95%; f) $SOCl_2$/EtOH, 40˚C, 55%; g) DIEA, COMU/DMF, 0˚C 1hr, RT 3hr, 12%; h) generated in the presence of 5mM TCEP *in vitro*.

temperature for an additional 5 minutes and heated to 100˚C for 45 minutes using a microwave reactor. The reaction was cooled to room temperature and filtered through a pad of celite and sepharose. The reaction was concentrated under reduced pressure and purified using flash chromatography (0:1-1:6 EtOAc/Hexane gradient) to yield a yellow solid (300 mg, 70%). The purified product was in agreement with literature values for melting point, ^{1}H NMR, and ^{13}C NMR.[35]

(2RS-4R)-2-(2-(2-hydroxyphenyl)thiazol-4-yl)thiazolidine-4-carboxylic acid ($HPT_{ox}T_{red}$-CO_2^-)

2-(2-Hydroxyphenyl)thiazole-4-carbaldehyde (1.46 mmol) and potassium acetate (9.71 mmol) were dissolved in a solution of EtOH/H_2O (10:3, 104 mL) in a round bottom flask. L-Cysteine-HCl (5.40 mmol) was added and the reaction was stirred at room temperature for 1 hour. The reaction was washed with hexanes (50 mL), diluted with H_2O and acidified with citric acid to ~pH 2. The aqueous layer was extracted with CH_2Cl_2 (3 x 75 mL). The organic layers were combined and dried over Na_2SO_4. The solvent was removed in vacuo to afford a yellow/white solid (430 mg, 95%). The product was purified as a mixture of epimers and was in agreement with literature values for ^{1}H NMR and ^{13}C NMR.[36] HRMS: $[M + H]^+$ 309.0289 (calcd), 309.0363 (found).

Ethyl (2RS-4R)-2-(2-(2-hydroxyphenyl)thiazol-4-yl)thiazolidine-4-carboxylate ($HPT_{ox}T_{red}$-CO_2Et).

Thionyl chloride (1.75 mmol) was added dropwise to anhydrous ethanol (3.25 mL) with vigorous stirring in a -10˚C ice water bath. The mixture was warmed to room temperature. $HPT_{ox}T_{red}$-CO_2^- (0.649 mmols) was added and the reaction was heated to 40 ˚C and stirred for 16

hours. The mixture was cooled to room temperature and the solvent was removed in vacuo. The residue was reconstituted in saturated Na_2CO_3 (30 mL) and extracted with CH_2Cl_2 (4 x 50 mL). The organic layers were combined, and the mixture was concentrated under reduced pressure and purified using the CombiFlash (0:1-1:2 EtOAc/Hexane gradient) to yield a clear oil (119 mg, 55%). The product was purified as a diastereomer (38:62) as determined by chiral chromatography (**Figure 2-3**). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 11.50 (d, J = 764.2 Hz, 2H), 8.37 (s, 1H), 8.05 (dd, J = 7.9, 1.7 Hz, 1H), 8.02 (dd, J = 7.9, 1.7 Hz, 1H), 7.69 (s, 1H), 7.57 (s, 1H), 7.34 – 7.21 (m, 2H), 7.00 (ddd, J = 8.3, 5.4, 1.2 Hz, 2H), 6.91 (dddd, J = 8.1, 7.2, 5.7, 1.1 Hz, 2H), 5.84 (d, J = 9.9 Hz, 1H), 5.72 (d, J = 12.1 Hz, 1H), 4.41 (ddd, J = 8.7, 6.9, 5.1 Hz, 1H), 4.25 – 4.10 (m, 4H), 4.03 – (ddd, J = 12.1, 8.8, 7.0 Hz, 1H), 3.94 (t, J = 9.4 Hz, 1H), 3.59 (t, J = 12.2 Hz, 1H), 3.42 – 3.30 (m, 1H), 3.15 – 3.03 (m, 1H), 1.21 (td, J = 7.1, 4.7 Hz, 6H) **^{13}C NMR** 126 MHz, DMSO-d_6) δ 171.30, 170.84, 164.90, 163.92, 163.59, 155.35, 154.87, 152.37, 131.22 – 131.01 (m), 127.43, 127.34, 119.27, 118.95, 118.82, 116.64, 116.51, 115.14, 67.06, 65.23, 64.96, 61.08, 60.80, 38.03, 37.94, 14.08, 14.04. **HRMS**: $[M + H]^+$ 337.0675 (calcd), 337.0708 (found).

(S)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylic acid (HPT-CO_2^-)

Synthesis was performed as previously described in Bergeron *et al* and Zamri *et al* except starting materials differed by using D-cysteine instead of L-cysteine.[37, 38] Product analysis agreed with previously reported ^{1}H NMR and ^{13}C NMR. HRMS: $[M + H]^+$ 224.0303 (calcd), 224.0378 (found). Chiral analysis was performed using HPLC and a CHIRALCEL® OD-H column. Elution was performed using a hexane:isopropanol (90:10) mixture supplemented with 0.1% trifluoroacetic acid at a flow rate of 0.500 mL/min, R_f = 14.7 minutes.

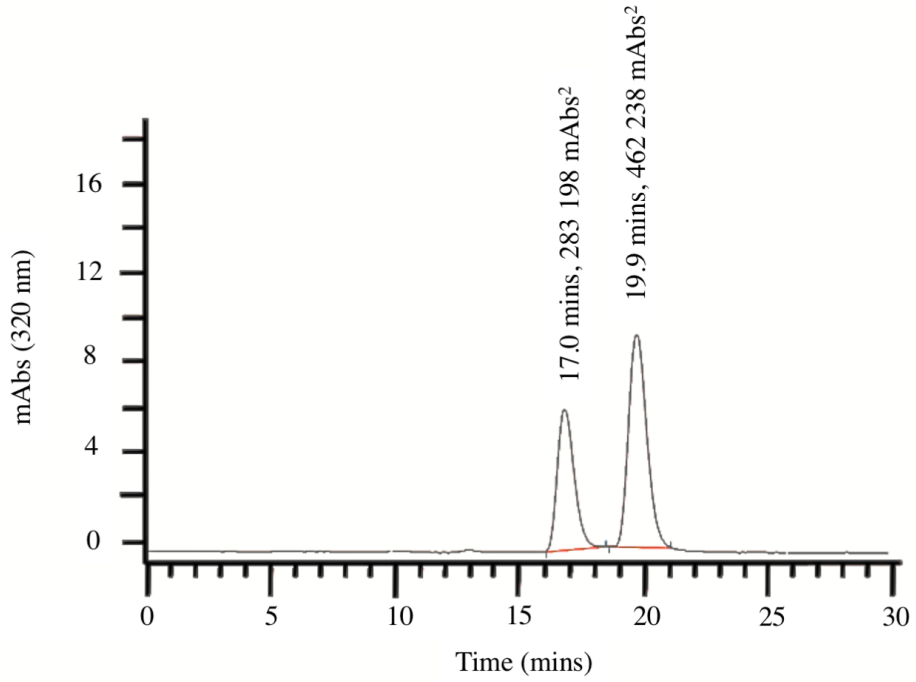


Figure 2-3: Separation of $HPT_{ox}T_{red}$-CO_2Et diastereomers on CHIRALCEL OD-H column. Separation conditions: 1 mL/min, Hexane:Isopropanol (90:10), absorbance monitored at 320 nm. $HPT_{ox}T_{red}$-CO_2Et exists as a diastereomer (38:62) eluting at 17.0 mins and 19.9 mins.

dimethyl 3,3'-disulfanediyl(2R,2'R)-bis(2-((S)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxamido)propanoate)

D-HPT-CO_2^-, L-cystine dimethyl ester dichloride (Sigma), and 1-cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate (Sigma) were purged under argon in a 25 mL round bottom flask before being dissolved in DMF (10 mL) at 0˚C. Diisopropylethylamine was then added dropwise and the reaction was stirred vigorously for 1 hour at 4˚C and 3 hours at room temperature. Ethyl acetate (50 mL) was then added to the reaction. The organic fraction was washed with 1M HCl (2 x 50 mL), water (2 x 50 mL), and brine (2 x 50 mL) before being concentrated under reduced pressure. The residue was purified using the CombiFlash (DCM from 0-5 minutes before a gradient to 5% MeOH from 5-8 minutes and holding the 5% MeOH in DCM for 5 more minutes) eluting a peak at 10 minutes. An additional purification step was performed using a Waters Acquity UPLC with a photodiode array UV detector and an LCT premiere TOF mass spectrometer. The UPCL/HRMS used a gradient elution increasing 20% CH_3CN in water over 5 minutes. Injections of a 1 mL DMSO stock were made onto a Waters (T3 C18) 19 x 150 mm, 5 µm column with a flow rate of 20 mL/min to yield a white solid (35 mg, 12%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.09 (s, 1H), 8.86 (d, J = 8.0 Hz, 1H), 7.45 (ddd, J = 15.5, 7.8, 1.7 Hz, 2H), 7.02 – 6.92 (m, 2H), 5.38 (t, J = 9.0 Hz, 1H), 4.68 (td, J = 8.7, 4.8 Hz, 1H), 3.66 (s, 3H), 3.55 (dd, J = 11.1, 8.4 Hz, 1H), 3.23 (dd, J = 13.9, 4.9 Hz, 1H), 3.07 (dd, J = 13.9, 9.1 Hz, 1H). **^{13}C NMR** 126 MHz, DMSO-d_6) δ 173.03, 171.05, 169.83, 158.61, 134.75, 131.09, 119.79, 117.35, 116.24, 77.99, 54.07, 52.32, 33.93. **HRMS**: $[M + H]^+$ 679.0946 (calcd), 679.1008 (found).

Preparation of Co-substrate Analogues.

Se-Adenosyl selenomethionine (SeAdoMet).

SeAdoMet was biosynthesized and purified by a modification of known methods.[39] 10 mM L-methionine was added to a reaction mixture containing 100 mM Tris pH 8, 50 mM KCl, 40 mM $MgCl_2$, 15 mM ATP, 8% 2-mercaptoethanol, 1.3 U inorganic pyrophosphatase (Sigma), 130 µM *Mj* AdoMet Synthetase in a final volume of 10 mL. The reaction was gently rocked for 4 hours, quenched with 1 mL 1 M H_2SO_4, and neutralized with 1 mL 2.5 M Ammonium Acetate. The mixture was centrifuged (3 082 x g for 10 minutes at 4˚C), flash cooled, and stored as 800 µL aliquots at -80˚C for further purification. 400 µL was injected onto a Waters Semi-Prep C18 column (10 x 250 mm, 5µM) previously equilibrated in 95% 83 mM triethyl ammonium amine pH 5 (solvent A) and 5% methanol running at 5 mL/min. SeAdoMet was isolated at 3.7 minutes. A linear gradient of 0-50% acetonitrile and 5-50% methanol from 8-18 minutes was used to flush the column. A gradient back to 95% solvent A and 5% methanol was performed at 19 minutes to re-equilibrate the column for subsequent runs. Immediately after collection, SeAdoMet was flash cooled and lyophilized overnight. The lyophilized sample was resuspended in 250 µL of 100 mM H_2SO_4 and the concentration was determined spectrophotometrically (AdoMet, ε_{260} = 15 400 $M^{-1}cm^{-1}$)[40] to be 35.3 mM.

Te-Adenosyl telluromethionine (TeAdoMet).

The TeAdoMet was synthesized as previously described.[39]

Steady-state Adenylation Assay

Reactions contained 50 mM Tris pH 7.5, 1 mM $MgCl_2$, 200 µM 2-amino-6-mercapto-7-methylpurine ribonucleoside (MesGR), 0.002 U purine nucleoside phosphorylase (PNP, Sigma), 0.001 U inorganic pyrophosphatase (IPP, Sigma), 2 mM ATP, 150 mM hydroxylamine, 0.600 µM enzyme, and varying concentrations of L-cysteine. Stock concentrations (200 mM – 12.5 µM) of L-cysteine were prepared by serial dilution in 25 mM Tris pH 8 buffer supplemented with 5 mM dithiothreitol. ATP, L-cysteine, and hydroxylamine were made fresh for each triplicate reaction. A parent mixture containing all reaction components except L-cysteine were incubated for 15 minutes at room temperature. The reaction (final volume, 100µL) was initiated by the addition of the parent mixture to substrate in a 96-well flat bottom plate (Corning Cat. #9017) and production of 2-amino-6-mercapto-7-methylpurine (MesG) was monitored at A_{360} on a Cary 50 UV-VIS Spectrophotometer with a plate reader attachment. Steady-state turnover was converted from *AU/min* to *µM/min* using the general Beer-Lambert law using 0.29 cm as the path-length and 11,000 $cm^{-1}M^{-1}$ as the molar extinction coefficient of MesG.[41] The calculated rate was divided in half because each pyrophosphate released from the adenylation activity undergoes catalytic hydrolysis by inorganic pyrophosphatase creating two inorganic phosphates, both used individually by purine nucleoside phosphorylase to convert MesGR to MesG. Negative controls (without enzyme, without L-cysteine, and without ATP) were carried out with each triplicate reaction. The largest rate from the multiple negative control reactions was subtracted as background from all of the calculated rates. Triplicate trials were performed twice, on separate days for each enzyme. The initial rates were fit to the Michaelis-Menten equation using the nonlinear regression function of KaleidaGraph (Synergy Software). Kinetic parameters

were calculated for each of the six individual trials per enzyme. The reported values represent the average and standard deviation of these six sets of kinetic parameters.

Methyltransferase Activity Assay

Steady-state kinetics reactions varying $HPT_{ox}T_{red}$-CO_2Et substrate analog. Methyltransferase activity was observed with $HPT_{ox}T_{red}$-CO_2Et, but not observed with $HPT_{ox}T_{red}$-CO_2^-, $HPT_{ox}T$-CO_2Et,[42] or HPT-Cys-Me at 50 μM during a 1-hour incubation. Product formation of $HPT_{ox}T_{red\text{-}M}$-CO_2Et was monitored at A_{320} by HPLC and confirmed by selected ion recording mass spectrometry. Each assay mixture (100 μL) contained 100 mM Tris-HCl pH 8.0, 45 nM enzyme, 1 mM AdoMet, varying amounts of $HPT_{ox}T_{red}$-CO_2Et (1 μM – 30 μM) in DMSO (final DMSO concentration ≤ 2%), and 1 mM tryptophan (an internal standard). A parent mixture of all reaction components, excluding $HPT_{ox}T_{red}$-CO_2Et, was incubated at room temperature for 5 minutes before initiating the reaction with $HPT_{ox}T_{red}$-CO_2Et. A time-course assay was performed for both low (1 μM) and high (30 μM) $HPT_{ox}T_{red}$-CO_2Et to determine that 15 minutes was within the linear steady-state turnover for PchF-FL and PchF-AMP. Therefore, an endpoint assay was employed after a 15-minute incubation at room temperature. Reactions were quenched with 5 μL of 1 M H_2SO_4, vortexed, and neutralized with 10 μL of 1M sodium citrate pH 5.5.

Steady-state kinetics varying Onium Chalcogen co-substrates. Each assay mixture (100 μL) contained 100 mM Tris-HCl pH 8.0, 45 nM enzyme, 50 μM $HPT_{ox}T_{red}$-CO_2Et, varying amounts of AdoMet, SeAdoMet or TeAdoMet (0-2 mM), and 1 mM tryptophan (an internal standard). A parent mixture of all reaction components, excluding AdoMet, SeAdoMet, or TeAdoMet, was incubated at room temperature for 5 minutes before initiating the reaction with AdoMet,

SeAdoMet or TeAdoMet. Reactions were quenched with 5 µL of 1 M H_2SO_4, vortexed, and neutralized with 10 µL of 1 M sodium citrate pH 5.5. Linear steady-state turnover was assessed for low (5 µM) and high (1 mM) concentrations of AdoMet and SeAdoMet. The reaction was quenched at 10 minutes, within the linear steady-state for both AdoMet and SeAdoMet. TeAdoMet reactions, also within steady-state, were incubated for 4 hours and contained twice the concentration of enzyme. AdoMet was varied in reactions from 0 – 2000 µM, SeAdoMet was varied from 0 – 650 µM, whereas TeAdoMet was varied from 0 – 800 µM.

Methyltransferase steady-state kinetic analysis. Quenched and neutralized methyltransferase reaction samples were centrifuged (10 600 x g, 10 mins, 4 ˚C) to remove denatured enzyme, and 100 µL of supernatant was injected onto an IRIS Technologies LLC IProSIL C18 column (4.6 X 250 mm, 5 µM) previously equilibrated in 92% solvent B (0.016% formic acid, 0.02% triethylamine) and 8% acetonitrile. A linear gradient of 8-80% acetonitrile from 0-15 mins, an 80-100% linear gradient from 15-20 minutes, and a 100-8% gradient from 20-21 minutes at 1 mL/min gave good separation of $HPT_{ox}T_{red}$-CO_2Et (19.9 min) and methylated $HPT_{ox}T_{red}$-CO_2Et (21.6 min). A standard curve for product formation was determined by monitoring assay completion of 0-10 µM $HPT_{ox}T_{red}$-CO_2Et reactions with PchF-AMP performed in triplicate (9 000 ± 100 mV/µM, **Figure 2-4**). Due to substrate concentrations being less than 100 times enzyme concentration, Equation 1 was used to calculate steady-state parameters.[43] The reported kinetic parameters were calculated by finding the average and standard deviation of each individual assay performed in triplicate.

$$\boldsymbol{v = V_{max}\left(\frac{([E]_t+[S]_t + K_S) - \sqrt{(([E]_t) + ([E]_t) + K_S)^2 - 4([E]_t)([S]_t)}}{2([E]_t)}\right) + m_4 x}$$

Equation 1.

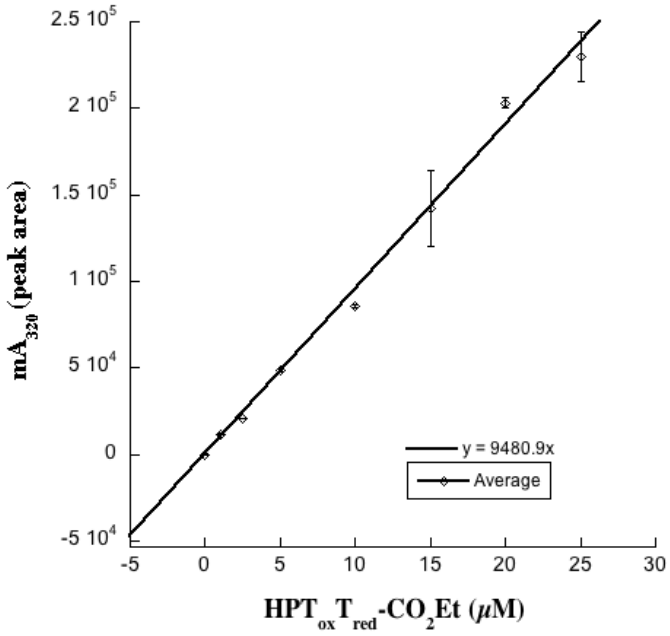


Figure 2-4: $HPT_{ox}T_{red\text{-}M}$-CO_2Et standard curve. $HPT_{ox}T_{red\text{-}M}$-CO_2Et standard curve was produced by running methyltransferase assays to completion at concentrations varying from 0-25 μM $HPT_{ox}T_{red}$-CO_2Et. Reactions were performed in triplicate. The slope, 9480 mAU/μM, was used to convert mAU to concentration of $HPT_{ox}T_{red\text{-}M}$-CO_2Et.

Results

Cloning, sequencing, and purification of PchF-FL, PchF-AMP, and G667I-PchF-AMP.

Three variants of PchF were generated heterologously in an *E. coli* expression system. The first is the full-length protein as produced by *P. aeruginosa*, called PchF-FL herein (**Figure 2-1B**). The second consists of the adenylation-methyltransferase-didomain with the peptidyl-carrier-protein domain, and is thus abbreviated PchF-AMP. This construct removes the N-terminal condensation domain and the C-terminal thioesterase domain to eliminate confounding activities. Previously, Patel and colleagues developed a pyochelin reconstitution assay using recombinant PchD, PchE, PchF, PchG with substrates and cofactors, monitoring the formation of mature pyochelin by HPLC analysis.[10] Patel *et al* reported that mutation of the second glycine of a core AdoMet binding motif (GxG) to an arginine (G1167R) in PchF resulted in undetectable amounts of pyochelin.[10] The authors suggested that pyochelin formation was inhibited due to a lack of methyl transfer from AdoMet to the $HPTT_{red}$ intermediate (**Figure 2-1 B,D**). We attempted to generate G667R-PchF-AMP, but overexpression yielded negligible amounts of protein. We therefore proposed that substitution of the same glycine with a bulky hydrophobic residue would likewise inhibit AdoMet binding, and designed a glycine to isoleucine mutation, G667I-PchF-AMP, which proved easier to purify in necessary amounts for kinetic analysis.

Each of the PchF variants was co-expressed with PA2412, an MbtH-like protein (MLP) from the pyoverdin biosynthetic cluster of *P. aeruginosa,* which increases the amount of the PchF variants purified. This is consistent with other reports that have shown that MLPs greatly enhance NRPS solubility and are sometimes even necessary for expression.[28, 29, 44, 45] Recently, MLPs were co-crystallized with NRPS modules showing complex formation with the adenylation domains.[28, 45-48] While each PchF variant was co-expressed with PA2412 and the

proteins co-eluted by nickel chelating chromatography, the proteins were separated by gel filtration chromatography. Therefore, the final purified PchF variants did not contain PA2412. The PchF-AMP and G667I-PchF-AMP purification protocols included an additional anion exchange step to achieve >95% purity, estimated by polyacrylamide gel electrophoresis.

Steady-state kinetics of adenylation activity. The adenylation domains of NRPS modules catalyze two activities. First, an aminoacyl-AMP is produced. Second, the amino acid is transferred to the thiol of a phosphopantethienyl-modified serine residue (the Ppant tether) of the carrier domain.[49] Once the amino acid is attached to the Ppant tether, the next enzymatic steps are carried out by downstream condensation, tailoring, or thioesterase domains. For an adenylation assay to enter the steady state, only the first reaction can be catalyzed: the amino acid cannot be transferred to the Ppant tether of the carrier domain, which would lead to a single turnover. Furthermore, some adenylation domains do not release pyrophosphate when the second reaction is not catalyzed.[50] Therefore, a nucleophilic surrogate may be required. The addition of hydroxylamine allows for nucleophilic attack of the aminoacyl-AMP, generating an aminoacyl-hydroxymate and AMP allowing for the release of the bound intermediate from the active site, promoting continuous turnover. Using this premise, Aldrich *et al* previously described a continuous assay to measure activity of adenylation domains by linking pyrophosphate production to an absorbance change (**Figure 2-5)**.[51] The coupled assay uses inorganic pyrophosphatase (IPP) to hydrolyze the pyrophosphate generated by the adenylation reaction to two inorganic phosphate ions. Subsequently, purine nucleoside phosphorylase (PNP) converts 7-methylthioguanosine (MesGR) to ribose 1-phosphate and the chromogenic 7-methythioguanine (MesG), with the change measured as an increase in absorbance at 360 nm. This assay was optimized to measure

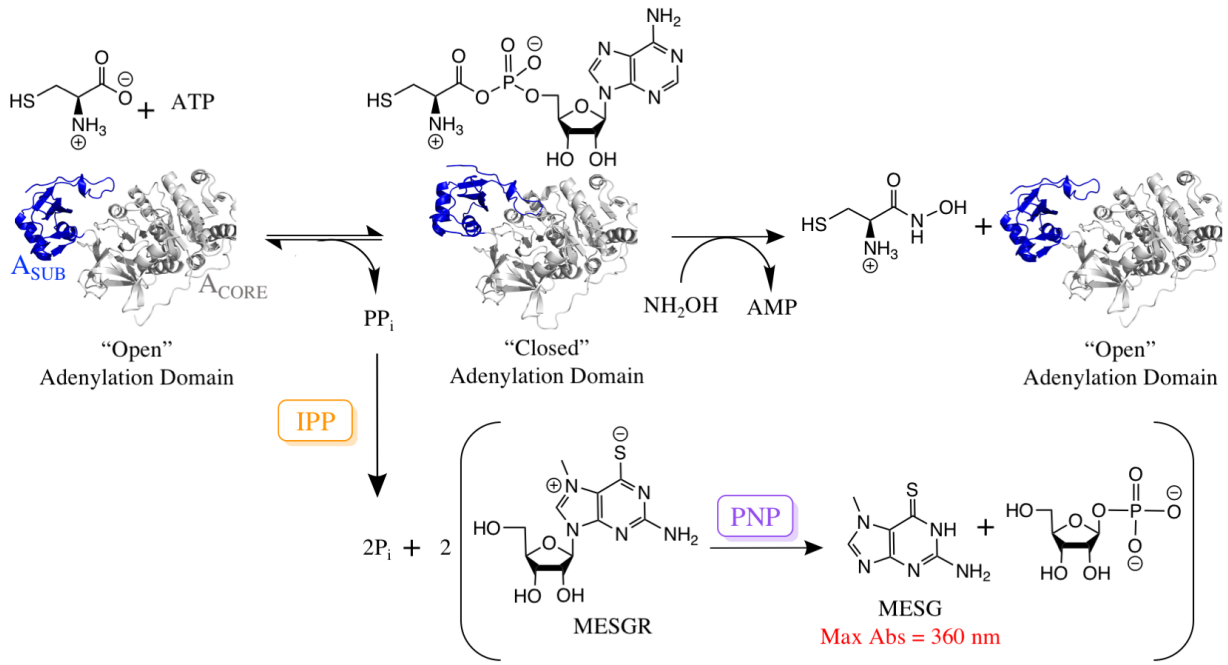


Figure 2-5: Adenylation assay (based on the assay designed by the Aldrich laboratory[1, 2]). The adenylation domain of PchF is specific for L-cysteine, forming an aminoacyl-AMP bond and releasing pyrophosphate (PP_i). The native protein is post-translationally modified with coenzyme A, generating a 4'phosphopantetheinyl (Ppant) on a conserved serine in the peptidyl carrier protein (P) domain. The high energy aminoacyl-adenylate bond is used to link the cysteine substrate to the Ppant arm of the P-domain. In this assay, the recombinant PchF variants are not post-translationally modified with the Ppant stalling the enzyme after the formation of the aminoacyl-AMP intermediate, with the adenylation domain trapped in the "closed" conformation. Opening of the active site and release of pyrophosphate (PP_i) can be promoted by addition of the nucleophile surrogate hydroxylamine, allowing the adenylation activity to enter the steady state. Inorganic pyrophosphatase (IPP) converts one pyrophosphate (PP_i) molecule to two inorganic phosphate (P_i) molecules. Purine nucleoside phosphorylase (PNP) catalyzes the phosphorylation of 7-methylthioguanosine (MesGR), generating ribose 1-phosphate and the chromogenic 7-methylthioguanine (MesG), which is monitored at 360 nm (bottom). Two molecules of MESG are generated for every one aminoacyl-AMP formation. Structures of LgrA[3] from gramicidin biosynthesis are used to illustrate the open (PDB: 5ES5) and closed (PDB: 5ES5) A-domains. The A_{core} subdomain is gray and the smaller C-terminal A_{sub} is blue.

adenylation of L-cysteine in the stuffed adenylation-methyltransferase didomain of PchF-FL, PchF-AMP, and G667I-PchF-AMP. Without hydroxylamine addition, detection of the downstream chromogenic MesG was within error of negative controls through 30 minutes, whereas the inclusion of hydroxylamine gave reliable steady state turnover **(Figure 2-6)**, suggesting that the adenylation domain remains in a closed conformation and inorganic pyrophosphate is not released until after attachment of the amino acid to the Ppant tether (mimicked here by the hydroxylamine surrogate). A representative assay with and without hydroxylamine can be found in **Figure 2-6**, along with Michaelis-Menten curves for the three PchF variants. Steady state kinetic parameters for adenylation of all three constructs can be found in **Table 2-1**. Note that k_{cat} values are within 1.2-fold for all three constructs. The K_M values are within 2-fold of each other, which accounts for the similar variance in k_{cat}/K_M. These values indicate that all three constructs have similar adenylation parameters, suggesting differences in catalytic efficiency for the methyltransferase domains are due to specific changes within the methyltransferase domains. It is interesting to note that kinetic parameters were found to be an entire magnitude different for PchF-FL than previously reported by Walsh *et al* (**Table 2-1**), highlighting discrepancies between the reverse assay measuring the incorporation of radioactive [^{32}P]-pyrophosphate into ATP used by Walsh and the assay described herein. The disparities between these two assays were formerly reported by Aldrich *et al*.[51]

Substrate analog synthesis. Kinetic characterization of NRPS domains can be difficult because the biological substrate is attached to the PCP-domain through the Ppant tether. Previously, substrate analogues, the peptidyl *N*-acetylcysteamine thioesters (peptidyl-SNACs), have been synthesized

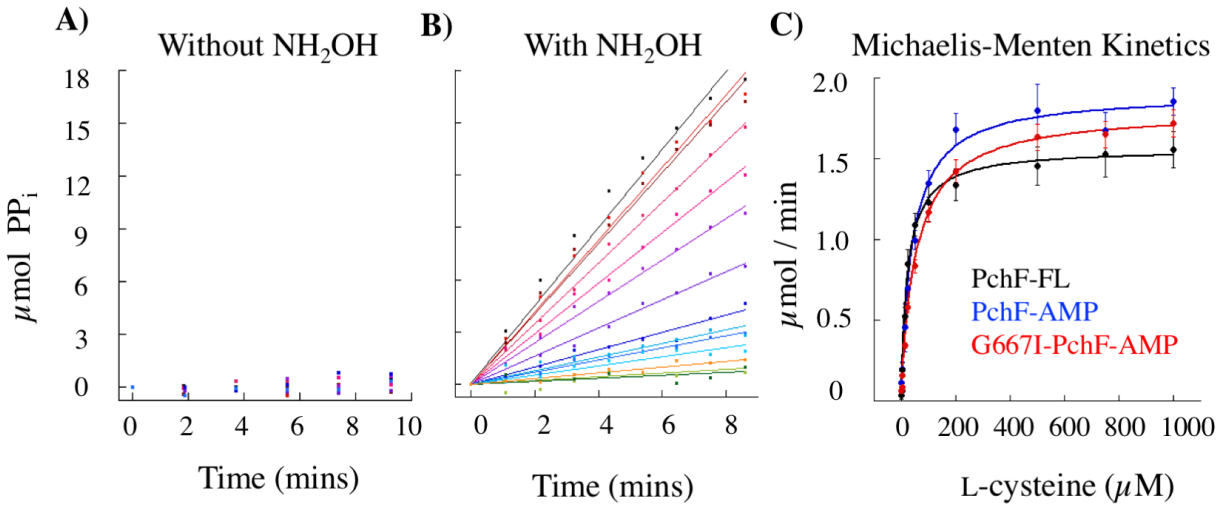


Figure 2-6: Hydroxylamine is a necessary addition for the adenylation assay. PchF catalyzes the adenylation of L-cysteine by ATP forming an aminoacyl-adenylate bond, releasing pyrophosphate. In the adenylation assay (**Figure 2-5**), pyrophosphate is converted to two inorganic phosphates by inorganic pyrophosphatase. 7-methylguanosine is then phosphorylated by purine nucleoside phosphorylase to generate 7-methylguanine, which absorbs at 360 nm. **A)** PchF variants do not turnover with L-cysteine and ATP (**A**), but do with the addition of hydroxylamine as a nucleophilic surrogate (**B**). Michaelis-Menten steady-state kinetics (**C**) were performed comparing the L-cysteine adenylation activity of PchF-FL, PchF-AMP, and the methyltransferase null mutant of G667I-PchF-AMP.

Table 2-1: L-Cysteine adenylation assay – kinetic parameters

	PchF-FL[a]	PchF-FL	PchF-AMP	G667I-PchF-AMP
k_{cat} (min^{-1})	415	2.6 ± 0.2	3.2 ± 0.2	3.0 ± 0.3
K_M (L-Cys μM)	537	25 ± 2	41 ± 8	54 ± 6
k_{cat}/K_M ($M^{-1}s^{-1}$)	13300	1800 ± 200	1300 ± 300	900 ± 100

[a] Quadri, Keating, Patel, Walsh. *Biochemistry* 1999, 38, 14941-14954. Quadri *et al* used an exchange assay measuring the reverse reaction that incorporates ^{32}P-pyrophosphate into ATP.

and used to elucidate substrate specificity and kinetic parameters of NRPS domains.[52] For example, Ehmann and colleagues revealed the enantioselectivity of the upstream substrate in the condensation domain of TycB from tyrocidine biosynthesis.[52] Schneider *et al* used peptidyl-SNAC mimics to study the kinetics of the oxidase activity of EpoB from epothilone biosynthesis and BlmIII of bleomycin biosynthesis.[15] Tseng *et al* also generated peptidyl-SNACs, peptidyl *N*-acetylethanolamine esters (ONACs) and peptidyl *N*-acetylethylenediamine amides (NNACs) to probe thioesterase cyclization and hydrolysis capabilities of the thioesterase domain of Srf from surfactin biosynthesis.[53] However, there are several examples where aminoacyl-SNACs or analogs have not been able to replace carrier-protein-bound thioesters or a substrate covalently bonded to CoA, and likely several more examples exist, though have not been published.[54-58] As natural product biosynthesis advances through the NRPS assembly line, intermediates become more complex, often including non-proteogenic amino acids, special tailoring, and enantioselectivity, further complicating organic synthetic routes to generate intermediate analogs.

During reconstitution assays, the growing peptide chain tethered by the phosphopantetheinyl (Ppant) thioester is susceptible to hydrolysis, releasing a carboxylate derivative of the intermediate of the biosynthetic pathway. HPT-Cys-Me, $HPT_{ox}T\text{-}CO_2^-$ and $HPTT_{red}\text{-}CO_2^-$ (**Figure 2-1D**) had been detected during pyochelin reconstitution assays.[6, 10] As a starting point for synthesis of potential substrate analogs to probe the methyltransferase kinetics of PchF, we generated these compounds, and derivatives altering the carboxyl to a terminal ester generating, HPT-Cys-Me, $HPT_{ox}T\text{-}CO_2Et$ and $HPT_{ox}T_{red}\text{-}CO_2Et$, to more closely mimic the natural thioester linkage of the Ppant tether.

Synthesis of $HPT_{ox}T\text{-}CO_2^-$ and $HPT_{ox}T\text{-}CO_2Et$ was previously reported.[42] For $HPT_{ox}T_{red}\text{-}CO_2^-$ and $HPT_{ox}T_{red}\text{-}CO_2Et$, 2-hydroxyphenyl boronic acid and 2-bromothiazole-4-carbaldehyde

were coupled by a Suzuki reaction to generate the 2-(2-hydroxyphenyl)thiazole-4-carbaldehyde product providing a 70% yield (**Scheme 2-1**). Condensation of the aldehyde with L-cysteine furnished the desired $HPT_{ox}T_{red}$-CO_2^- substrate analog as a diastereomer in 95% yield. The terminal carboxylic acid was transformed to an ethyl ester by generation of an acyl chloride before nucleophilic addition of the solvent ethanol providing $HPT_{ox}T_{red}$-CO_2Et in 55% yield also as a diastereomer **(Figure 2-3)**. HPT-Cys-Me was furnished by first synthesizing D-HPT-CO_2^- before performing an amino acid coupling with L-cystine using COMU as the coupling reagent. Generation of the free thiol HPT-Cys-Me, was performed *in vitro* by TCEP reduction and was confirmed by LC-HRMS.

Steady-state kinetics of methyltransferase activity. Using the synthesized substrate analogs, the steady-state capabilities of the stuffed methyltransferase domain of PchF-FL, PchF-AMP, and G667I-PchF-AMP were analyzed by a discontinuous assay performed in the presence of excess AdoMet. Neither HPT-Cys-Me, $HPT_{ox}T$-CO_2Et, nor $HPT_{ox}T_{red}$-CO_2^- demonstrated methyl transfer after 1-hour incubation. Methylation of $HPT_{ox}T_{red}$-CO_2Et by both PchF-FL and PchF-AMP was observed, generating $HPT_{ox}T_{red\text{-}M}$-CO_2Et, and the substrate and product were separated by HPLC (**Figure 2-7)**. Product formation was confirmed by secondary ion mass spectrometry. G667I-PchF-AMP did not catalyze methyltransferase activity, even after extending the assay to 10 hours. Kinetic parameters from the methyltransferase assay can be found in **Table 2-2** and Michaelis-Menten plots can be seen in **Figure 2-8**. It is important to note that PchF-FL and PchF-AMP provided comparable kinetic parameters using $HPT_{ox}T_{red}$-CO_2Et as a substrate analog, indicating that the three-domain fragment is a viable substitute for the full-length protein. The k_{cat}

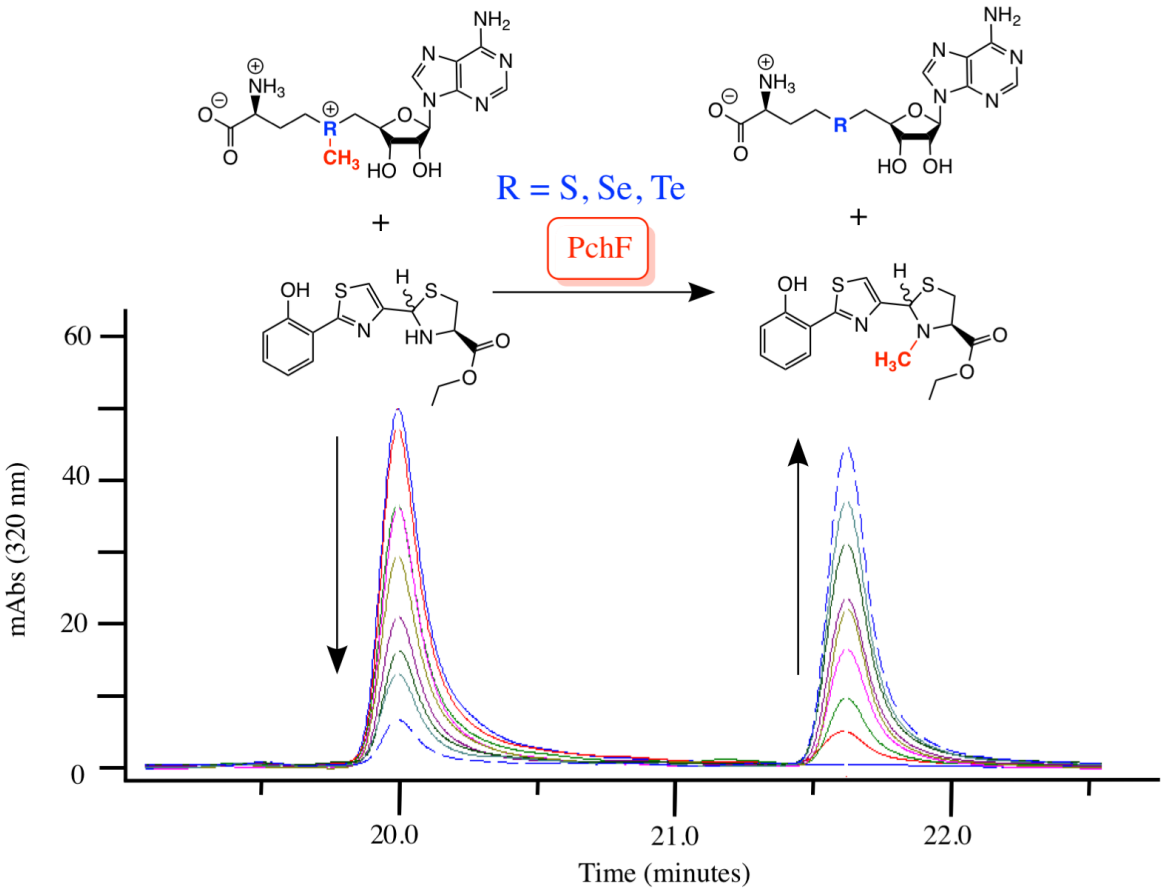


Figure 2-7: Methyltransferase Assay. PchF catalyzes methyl transfer from its natural co-substrate AdoMet to the synthesized substrate analog, $HPT_{ox}T_{red}$-CO_2Et. Steady-state Michaelis-Menten kinetics were performed by a discontinuous assay. Substrate analog (eluting at 20 min) and product analog (eluting at 21.6 min) were separated by a C18 column using HPLC.

Table 2-2: Methyltransferase assay varying substrate analog $HPT_{ox}T_{red}$-CO_2Et – kinetic parameters			
	PchF-FL	PchF-AMP	G667I-PchF-AMP
k_{cat} (min^{-1})	1.96 ± 0.04	2.9 ± 0.1	Not Detected
K_M (AdoMet μM)	3.6 ± 0.2	7.5 ± 0.2	Not Detected
k_{cat}/K_M ($M^{-1}s^{-1}$)	9100 ± 500	6500 ± 300	Not Detected

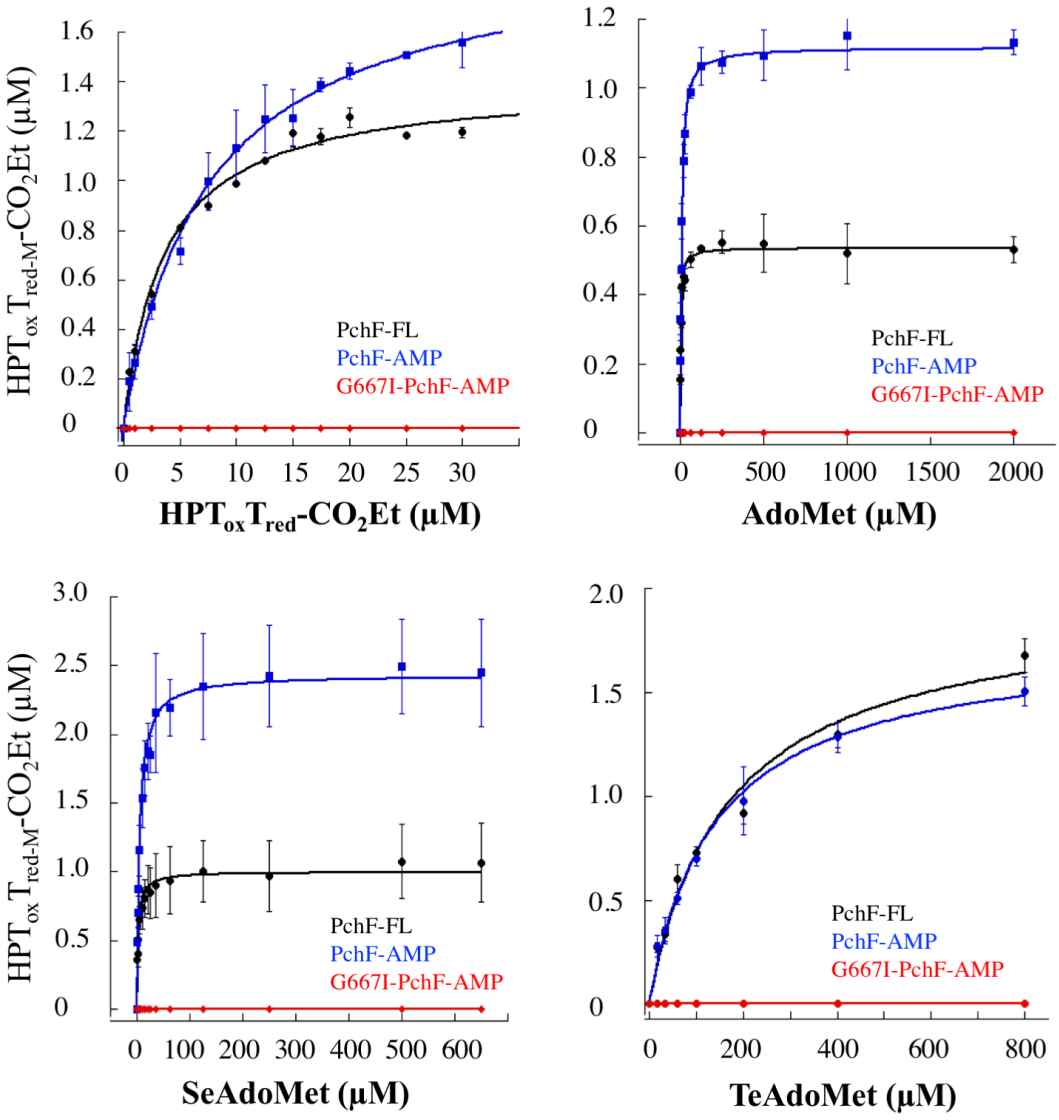


Figure 2-8: Methyltransferase Assays Michaelis-Menten Curves. Methyltransferase assay of PchF-FL, PchF-AMP, and G667I-PchF-AMP. Michaelis-Menten steady-state kinetic plots of (A) $HPT_{ox}T_{red}$-CO_2Et, (B) AdoMet, (C) SeAdoMet, (D) TeAdoMet.

was 1.5-fold greater and the K_M was 2.1-fold greater for PchF-AMP than that of PchF-FL. This resulted in PchF-AMP having a 1.4-fold lower catalytic efficiency compared to PchF-FL.

Onium chalcogen effects on the methyltransferase kinetic parameters of PchF variants. Replacement of the sulfonium of AdoMet with selenonium (SeAdoMet) and telluronium (TeAdoMet) has been used to investigate the mechanism of cyclopropane fatty acid (CFA) synthase. Because the mechanism has been well-defined experimentally, steady-state kinetics of catechol *O*-methyltransferase (COMT) using AdoMet, SeAdoMet, and TeAdoMet were determined and kinetic parameters were used as a control.[59] The primary (k_{cat}) and secondary (k_{cat}/K_M) rate constants for methyltransferase activity of CFA synthase and COMT follow the trend SeAdoMet > AdoMet > TeAdoMet, parallel to the electrophilicities determined for the onium chalcogen.[39, 59] Additional kinetic isotope effects (KIEs) using *S*-adenosyl-L-[*methyl-d_3*]methionine have been interpreted to support a tight S_N2 transition state for CFA and COMT. The methyltransferase assay described above was performed with varying concentrations of AdoMet and co-substrate analogs, SeAdoMet and TeAdoMet, while keeping $HPT_{ox}T_{red}$-CO_2Et in excess. The effect of AdoMet, SeAdoMet, and TeAdoMet on PchF-FL, PchF-AMP, and G667I-PchF-AMP is displayed in **Table 2-3** and kinetic ratios are reported in **Table 2-4**. Michaelis-Menten curves are in **Figure 2-8**. When AdoMet was used as the co-substrate for PchF-FL k_{cat} and K_M were 1.2 ± 0.1 min^{-1} and 3.3 ± 0.8 μM. SeAdoMet acted as a better co-substrate displaying a turnover number 1.9-fold greater and the K_M was within error compared to AdoMet. Substitution of AdoMet with TeAdoMet required the assay to be run 36-times longer (6 hours) with twice the enzyme concentration for both PchF-FL and PchF-AMP to detect product formation. Comparatively, TeAdoMet demonstrated a k_{cat} that was 21-fold slower and a K_M that

Table 2-3: Onium chalcogen effects for methyltransferase activity of PchF-FL and PchF-AMP.

Enzyme	Substrate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
PchF-FL	AdoMet	1.2 ± 0.1	3.3 ± 0.8	6000 ± 2000
	SeAdoMet	2.3 ± 0.6	3 ± 1	13000 ± 5000
	TeAdoMet	0.056± 0.003	160 ± 10	6.0 ± 0.5
PchF-AMP	AdoMet	2.48 ± 0.04	7.3 ± 0.8	5700 ± 600
	SeAdoMet	5.4 ± 0.8	5.4 ± 0.5	17000 ± 3000
	TeAdoMet	0.054± 0.003	140 ± 20	7 ± 1

was 47-fold greater in comparison to AdoMet. When performing the same assay with PchF-AMP, the kinetic values were comparable, yielding k_{cat}/K_m values within error to those for the full-length protein. Overall, the primary and secondary rate constants for the different onium chalcogen co-substrate and analogs follow the trend for an s_N2 mechanism.

Discussion

Peptides generated by nonribosomal peptide synthetases are not simply peptides but have been modified by tailoring domains to provide these natural products with their structural diversity and unique bioactive properties. Tailoring domains can be divided into three categories: independent, stand-alone, and stuffed. Independent tailoring domains are commonly incorporated into NRPS modules following the carrier domain, modifying the nascent chain *in cis*. Stand-alone tailoring domains, like PchG of pyochelin biosynthesis, are separate enzymes that modify the natural product biosynthetic intermediates *in trans*. Intriguingly, an increasing number of tailoring domains have been identified that are stuffed within the NRPS adenylation domain, replacing loops between signature adenylation domain sequences. [12, 13] Natural products built using stuffed tailoring domains have been isolated and used in the clinic as anticancer, antifungal, or antibiotic drugs. However, investigation of stuffed tailoring domains has been limited, leaving the structural and biochemical characterization of these domains and their distinctive integration into the NRPS bioassembly as an emerging area of study.

A mechanistic and structural characterization of the stuffed methyltransferase domains involved in thiocoraline and kutzneride biosynthesis, microbial depsipeptides, has been previously reported.[12, 19, 25-27] These stuffed methyltransferase domains perform *O*-methylation of a serine residue (kutzeride) and *N*- and *S*-methylation of cysteine residues (thiocoraline). An

adenylation domain consists of two subdomains, an A_{core} subdomain, which includes the catalytic machinery, and the A_{sub} subdomain that works with the carrier domain to orient the substrate attached to the Ppant tether into each active site of the NRPS module.[46, 49, 60] The methyltransferase domain of the adenylation-methyltransferase didomain of the kutzneride assembly line for *O*-methylation of serine is in the second module of the protein KtzH and is stuffed into a loop of the A_{sub} subdomain between the A8 and A9 adenylation core sequence motifs.[27] The methyltransferase domains of the adenylation-methyltransferase didomains of the thiocoraline assembly line for *N*-methylation are found in the two modules of TioS and are likewise stuffed into loops of the A_{sub} subdomain between the A8 and A9 adenylation core sequence motifs. The didomain for *S*-methylation is in the standalone protein TioN, and the methyltransferase domain is stuffed into loops of the A_{core} subdomain between the A2 and A3 adenylation sequence motifs.[19, 25, 26] The premise of the previous mechanistic work on these didomains was that methylation could occur on the free amino acid, the aminoacyl-AMP intermediate, the amino acid after thiolation (attachment to the Ppant tether), or as part of the condensed chain before being passed to the next module.[19, 25-27] Experiment has ruled out methylation of free serine or cysteine by these didomains.[19, 25-27] The authors maintain that the substrate of the methyltransferase domain is either the aminoacyl-AMP intermediate, or the amino acid after thiolation but before condensation with the growing peptide chain.[19, 25-27]

The first structure of a stuffed adenylation domain from the second module of TioS, co-crystallized with the MbtH-like protein (MLP), TioT, was recently reported to 2.90 Å.[26] The authors describe the structure as a dumbbell with the *N*-terminal A_{core} and methyltransferase domains being the two balls of the dumbbell, with the discontinuous A_{sub} domain connecting these two domains as the dumbbell bar. By comparison to the structures for the initiation

module involved in biosynthesis of the antibiotic gramicidin (LgrA) and to the termination module of the enterobactin biosynthetic pathway (EntF), this new structure is in the "thiolation" conformation[13, 46, 61]. In other words, the *C*-terminal A_{sub} domain is in position for thiolation chemistry by the adenylation A_{core} subdomain (the conformation for attachment of the amino acid substrate to the Ppant tether), even though the carrier domain is not present in the TioS construct. The structure is quite valuable, showing for the first time an A_{sub} subdomain interrupted by a tailoring domain and yet maintaining the expected secondary and tertiary structure. [26] However, it does not answer mechanistic questions such as the order of events for peptide formation and tailoring; for example, whether methylation occurs before or after condensation to the nascent peptide.

Adenylation domains are considered the 'gate-keepers' of the growing chain,[62] selective for the amino acid that gets attached to the carrier domain, and in the case of adenylation-methyltransferase stuffed didomains, successively methylated. If the amino acid attached to the Ppant tether is the substrate for the tailoring domain, then the methylated amino acid could subsequently condense with the upstream growing peptidyl chain. Conversely, the substrate for the stuffed methyltransferase domain could be the condensed product of the module, a last step before interaction with the downstream module. Thiocoraline and kutzneride are complex natural products, and probing the methyltransferase activities that occur late in the biosynthesis becomes increasingly difficult as intermediates become increasingly complex. Thus, studying the stuffed methyltransferase didomain of PchF, also interrupting the A8 and A9 core motifs in the A_{sub} of the adenylation domain, is ideal: pyochelin is a much simpler natural product and intermediate analogs can be easily synthesized.

Previous studies of PchF and pyochelin biosynthesis do not support the model proposed for thiocoraline or kutzneride biosynthesis in which the stuffed methyltransferase didomains act upon the amino acid tethered to the carrier domain prior to condensation with the upstream peptidyl chain. From reconstitution assays Walsh and colleagues isolated and characterized HPT-Cys-Me, $HPT_{ox}T\text{-}CO_2^-$, and $HPTT_{red}\text{-}CO_2^-$, but not their methylated analogues, indicating condensation to the upstream nascent peptide occurs before methylation (**Figure 2-1 C,D**).[6, 10, 11] However, pyochelin biosynthesis is further complicated by the cyclization reaction that generates the thiazoline, which is hypothesized to be catalyzed by the condensation domain. Methylation may occur after condensation and before cyclization, after cyclization and before reduction, or after cyclization and reduction directly prior to release of the mature peptide (**Figure 2-2**).[6, 10, 11] To decipher when on the pathway methylation occurs, intermediate analogs were synthesized mimicking isolates from the reconstitution assays (**Figure 2-1D**): HPT-Cys-Me is an ester analog mimicking the growing peptide chain after condensation and before cyclization; $HPT_{ox}T\text{-}CO_2Et$ is an ethyl ester mimicking the growing peptide chain after cyclization but before reduction; and $HPT_{ox}T_{red}\text{-}CO_2Et$ mimics the growing chain after cyclization and reduction. Importantly, methyl transfer was only observed when $HPT_{ox}T_{red}\text{-}CO_2Et$ was used as the substrate analog, indicating that condensation, cyclization, and reduction by PchG are necessary events prior to methylation by PchF. It should also be noted that methyl transfer was not observed for the carboxylic acid analogue, $HPT_{ox}T_{red}\text{-}CO_2^-$, suggesting that the previously isolated acid is not a substrate. The ester, $HPT_{ox}T_{red}\text{-}CO_2Et$, which more closely resembles the natural thioester, serves as a substrate analogue for methyl transfer.

Steady-state kinetic parameters were measured for PchF-FL, PchF-AMP, and G667I-PchF-AMP, for both the adenylation methyltransferase activity. The adenylation activity assay

showed that all three constructs, PchF-FL, PchF-AMP, and G667I-PchF-AMP, had similar adenylation parameters, indicating that the methyltransferase variant did not negatively impact the larger didomain (**Table 2-1**). While the catalytic efficiency was an order of magnitude less than that previously reported, the assays are significantly different, as documented in the Results section, and similar differences are well documented.[51, 63] A newly defined, discontinuous HPLC methyltransferase assay gave reliable kinetic parameters, with the G667I-PchF-AMP variant showing complete loss of methyltransferase activity.

In 1976, Schowen and colleagues proposed that catechol O-methyltransferase (COMT) used a symmetrical and "tight" S_N2 transition state, one in which the bonds being formed and broken were both short, based on kinetic isotope effects measured with deuterated and ^{13}C-labeled AdoMet.[64] More recently (2016), Luo and colleagues have used similar experiments to probe the transition state of SET8, a lysine methyltransferase, and showed a "tight" yet early asymmetrical transition state, with a *C-N* distance of ~2.4 Å and a *C-S* distance of ~2.0 Å.[64, 65] In addition to measuring KIEs and BIEs, the authors also measured steady state kinetic parameters with SeAdoMet and found the values to be in agreement with the slightly weaker *Se-C* bond expected for the leaving group due to the intrinsic chemical properties of the chalcogen.[39] Indeed, onium chalcogen effects have been used to probe methyltransferase activity. By monitoring the formation of 5'-methylthioadenosine (MTA), 5'-methylselenoadenosine (SeMTA), and 5'-methyltelluroadenosine (TeMTA), a model for relative electrophilicities was observed and followed the trend: SeAdoMet > AdoMet > TeAdoMet.[39] Interestingly, substitution of AdoMet with its onium chalcogen analogs used in steady-state analysis with cyclopropane fatty acid synthase and COMT resulted in primary (k_{cat}) and secondary (k_{cat}/K_M) rate constants following the same trend.[59] In this study, AdoMet and its onium congeners, SeAdoMet and

TeAdoMet, were used as co-substrate and co-substrate analogs in the PchF methyltransferase assay. PchF-FL and PchF-AMP exhibited primary and secondary rate constants that followed the SeAdoMet > AdoMet > TeAdoMet trend. By comparison of onium chalcogen steady state kinetics for methyltransferases of determined transition state (**Table 2-4**), these results allow us to hypothesize that PchF methyl transfer proceeds by an S_N2 reaction mechanism for this secondary amine.

Conclusion

In summary we have characterized the stuffed methyltransferase didomain of PchF in pyochelin biosynthesis. Potential substrate analogs were synthesized based on previously recovered biosynthesis intermediates. The stuffed methyltransferase didomain of PchF catalyzes AdoMet dependent methyl transfer, the penultimate step in pyochelin biosynthesis, after condensation of the cysteine onto the upstream peptide, cyclization to form the thiazoline, and reduction to form the thiazolidine, but before release of the mature peptide. Onium chalcogen effects suggest that the stuffed methyltransferase domain proceeds through an S_N2 mechanism.

Table 2-4: Kinetic parameter comparison between AdoMet, SeAdoMet, and TeAdoMet

Enzyme	Substrate Analog	k_{cat} AdoMet / (Se/Te)AdoMet	k_{cat}/K_M AdoMet / (Se/Te)AdoMet
CFA [a] Synthase	SeAdoMet	1.8	2.8
	TeAdoMet	0.50	0.059
COMT [a]	SeAdoMet	1.1	1.2
	TeAdoMet	0.16	0.0089
SET8 [b]	SeAdoMet	1.9	1.9
PchF-FL	SeAdoMet	1.9	2.2
	TeAdoMet	0.047	0.0010
PchF-AMP	SeAdoMet	2.2	3.0
	TeAdoMet	0.022	0.0012

[a] Iwig, Grippe, McIntyre, Booker. *Biochemistry*, 2004, 43, 13510-13524. [b] Linscott, Kapilashrami, Wang, Senevirathne, Bothwell, Blum, Luo. *Proc. Natl. Acad. Sci.*, 2016, 113, E8369-E8378.

References:

[1] Miethke, M., and Marahiel, M. A. (2007) Siderophore-based iron acquisition and pathogen control, *Microbiol Mol Biol Rev 71*, 413-451.

[2] Cox, C. D., Rinehart, K.L., Moore, M.L. Carter Cook Jr., J. (1981) Pyochelin: Novel structure of an iron-chelating growth promoter for Pseudomonas aeruginosa, *Proc Natl Acad Sci U S A 78*, 4256-4260

[3] Church, D., Elsayed, S., Reid, O., Winston, B., and Lindsay, R. (2006) Burn wound infections, *Clin Microbiol Rev 19*, 403-434.

[4] Bhagirath, A. Y., Li, Y., Somayajula, D., Dadashi, M., Badr, S., and Duan, K. (2016) Cystic fibrosis lung environment and Pseudomonas aeruginosa infection, *BMC Pulm Med 16*, 174.

[5] Sadikot, R. T., Blackwell, T. S., Christman, J. W., and Prince, A. S. (2005) Pathogen-host interactions in Pseudomonas aeruginosa pneumonia, *Am J Respir Crit Care Med 171*, 1209-1223.

[6] Quadri, L. E., Keating, T. A., Patel, H. M., and Walsh, C. T. (1999) Assembly of the Pseudomonas aeruginosa nonribosomal peptide siderophore pyochelin: In vitro reconstitution of aryl-4, 2-bisthiazoline synthetase activity from PchD, PchE, and PchF, *Biochemistry 38*, 14941-14954.

[7] Reimmann, C., Serino, L., Beyeler, M., and Haas, D. (1998) Dihydroaeruginoic acid synthetase and pyochelin synthetase, products of the pchEF genes, are induced by extracellular pyochelin in Pseudomonas aeruginosa, *Microbiology 144 (Pt 11)*, 3135-3148.

[8] Serino, L., Reimmann, C., Visca, P., Beyeler, M., Chiesa, V. D., and Haas, D. (1997) Biosynthesis of pyochelin and dihydroaeruginoic acid requires the iron-regulated pchDCBA operon in Pseudomonas aeruginosa, *J Bacteriol 179*, 248-257.

[9] Patel, H. M., Tao, J., and Walsh, C. T. (2003) Epimerization of an L-cysteinyl to a D-cysteinyl residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from Pseudomonas aeruginosa, *Biochemistry 42*, 10514-10527.

[10] Patel, H. M., and Walsh, C. T. (2001) In vitro reconstitution of the Pseudomonas aeruginosa nonribosomal peptide synthesis of pyochelin: characterization of backbone tailoring thiazoline reductase and N-methyltransferase activities, *Biochemistry 40*, 9023-9031.

[11] Reimmann, C., Patel, H. M., Serino, L., Barone, M., Walsh, C. T., and Haas, D. (2001) Essential PchG-dependent reduction in pyochelin biosynthesis of Pseudomonas aeruginosa, *J Bacteriol 183*, 813-820.

[12] Labby, K. J., Watsula, S. G., and Garneau-Tsodikova, S. (2015) Interrupted adenylation domains: unique bifunctional enzymes involved in nonribosomal peptide biosynthesis, *Nat Prod Rep 32*, 641-653.

[13] Ronnebaum, T. A., and Lamb, A. L. (2018) Nonribosomal peptides for iron acquisition: pyochelin biosynthesis as a case study, *Curr Opin Struct Biol 53*, 1-11.

[14] Schneider, T. L., and Walsh, C. T. (2004) Portability of oxidase domains in nonribosomal peptide synthetase modules, *Biochemistry 43*, 15946-15955.

[15] Schneider, T. L., Shen, B., and Walsh, C. T. (2003) Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation, *Biochemistry 42*, 9722-9730.

[16] Perlova, O., Fu, J., Kuhlmann, S., Krug, D., Stewart, A. F., Zhang, Y., and Muller, R. (2006) Reconstitution of the myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in Myxococcus xanthus, *Appl Environ Microbiol 72*, 7485-7494.

[17] Buntin, K., Irschik, H., Weissman, K. J., Luxenburger, E., Blocker, H., and Muller, R. (2010) Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity, *Chem Biol 17*, 342-356.

[18] Sattely, E. S., Fischbach, M. A., and Walsh, C. T. (2008) Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways, *Natural product reports 25*, 757-793.

[19] Al-Mestarihi, A. H., Villamizar, G., Fernandez, J., Zolova, O. E., Lombo, F., and Garneau-Tsodikova, S. (2014) Adenylation and S-methylation of cysteine by the bifunctional enzyme TioN in thiocoraline biosynthesis, *J Am Chem Soc 136*, 17350-17354.

[20] De Crecy-Lagard, V., Marliere, P., and Saurin, W. (1995) Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation, *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie 318*, 927-936.

[21] Watanabe, K., Hotta, K., Praseuth, A. P., Koketsu, K., Migita, A., Boddy, C. N., Wang, C. C., Oguri, H., and Oikawa, H. (2006) Total biosynthesis of antitumor nonribosomal peptides in Escherichia coli, *Nat Chem Biol 2*, 423-428.

[22] Sussmuth, R., Muller, J., von Dohren, H., and Molnar, I. (2011) Fungal cyclooligomer depsipeptides: from classical biochemistry to combinatorial biosynthesis, *Natural product reports 28*, 99-124.

[23] Kreutzer, M. F., Kage, H., Gebhardt, P., Wackler, B., Saluz, H. P., Hoffmeister, D., and Nett, M. (2011) Biosynthesis of a complex yersiniabactin-like natural product via the mic locus in phytopathogen Ralstonia solanacearum, *Appl Environ Microbiol 77*, 6117-6124.

[24] Christiansen, G., Fastner, J., Erhard, M., Borner, T., and Dittmann, E. (2003) Microcystin biosynthesis in planktothrix: genes, evolution, and manipulation, *Journal of bacteriology 185*, 564-572.

[25] Mori, S., Garzan, A., Tsodikov, O. V., and Garneau-Tsodikova, S. (2017) Deciphering Nature's Intricate Way of N,S-Dimethylating l-Cysteine: Sequential Action of Two Bifunctional Adenylation Domains, *Biochemistry 56*, 6087-6097.

[26] Mori, S., Pang, A. H., Lundy, T. A., Garzan, A., Tsodikov, O. V., and Garneau-Tsodikova, S. (2018) Structural basis for backbone N-methylation by an interrupted adenylation domain, *Nat Chem Biol 14*, 428-430.

[27] Zolova, O. E., and Garneau-Tsodikova, S. (2014) KtzJ-dependent serine activation and O-methylation by KtzH for kutznerides biosynthesis, *J Antibiot (Tokyo) 67*, 59-64.

[28] Drake, E. J., Cao, J., Qu, J., Shah, M. B., Straubinger, R. M., and Gulick, A. M. (2007) The 1.8 A crystal structure of PA2412, an MbtH-like protein from the pyoverdine cluster of Pseudomonas aeruginosa, *J Biol Chem 282*, 20425-20434.

[29] Boll, B., Taubitz, T., and Heide, L. (2011) Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis, *J Biol Chem 286*, 36281-36290.

[30] Felnagle, E. A., Barkei, J. J., Park, H., Podevels, A. M., McMahon, M. D., Drott, D. W., and Thomas, M. G. (2010) MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases, *Biochemistry 49*, 8815-8817.

[31] Ochsner, U. A., Wilderman, P. J., Vasil, A. I., and Vasil, M. L. (2002) GeneChip expression analysis of the iron starvation response in Pseudomonas aeruginosa: identification of novel pyoverdine biosynthesis genes, *Mol Microbiol 45*, 1277-1287.

[32] Zhang, W., Heemstra, J. R., Jr., Walsh, C. T., and Imker, H. J. (2010) Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins, *Biochemistry 49*, 9946-9947.

[33] Lu, Z. J., and Markham, G. D. (2002) Enzymatic properties of S-adenosylmethionine synthetase from the archaeon Methanococcus jannaschii, *J Biol Chem 277*, 16624-16631.

[34] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, *John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press*, 571-607

[35] Cuppels, D. A., Stipanovic, R. D., Stoessl, A., and Stothers, J. B. (1987) C-13mr Studies .132. The Constitution and Properties of a Pyochelin Zinc Complex, *Can J Chem 65*, 2126-2130.

[36] Mislin, G. L., Burger, A., and Abdallah, M. A. (2004) Synthesis of new thiazole analogues of pyochelin, a siderophore of Pseudomonas aeruginosa and Burkholderia cepacia. A new conversion of thiazolines into thiazoles, *Tetrahedron 60*, 12139-12145.

[37] Zamri, A., and Abdallah, M. A. (2000) An improved stereocontrolled synthesis of pyochelin, siderophore of *Pseudomonas aeruginosa* and *Burkholderia cepacia*, *Tetrahedron 56*, 249-256.

[38] Bergeron, R. J., Wiegand, J., Dionis, J. B., Egli-Karmakka, M., Frei, J., Huxley-Tencer, A., and Peter, H. H. (1991) Evaluation of desferrithiocin and its synthetic analogues as orally effective iron chelators, *J Med Chem 34*, 2072-2078.

[39] Iwig, D. F., and Booker, S. J. (2004) Insight into the polar reactivity of the onium chalcogen analogues of S-adenosyl-L-methionine, *Biochemistry 43*, 13496-13509.

[40] Schlenk, F., and Depalma, R. E. (1957) The formation of S-adenosylmethionine in yeast, *J Biol Chem 229*, 1037-1050.

[41] Webb, M. R. (1992) A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems, *Proc Natl Acad Sci U S A 89*, 4884-4887.

[42] Meneely, K. M., Ronnebaum, T. A., Riley, A. P., Prisinzano, T. E., and Lamb, A. L. (2016) Holo Structure and Steady State Kinetics of the Thiazolinyl Imine Reductases for Siderophore Biosynthesis, *Biochemistry 55*, 5423-5433.

[43] Segel, I. H. (1975) *Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems*, Wiley, New York.

[44] Heemstra, J. R., Jr., Walsh, C. T., and Sattely, E. S. (2009) Enzymatic tailoring of ornithine in the biosynthesis of the Rhizobium cyclic trihydroxamate siderophore vicibactin, *J Am Chem Soc 131*, 15317-15329.

[45] Miller, B. R., Drake, E. J., Shi, C., Aldrich, C. C., and Gulick, A. M. (2016) Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture, *J Biol Chem 291*, 22559-22571.

[46] Drake, E. J., Miller, B. R., Shi, C., Tarrasch, J. T., Sundlov, J. A., Allen, C. L., Skiniotis, G., Aldrich, C. C., and Gulick, A. M. (2016) Structures of two distinct conformations of holo-non-ribosomal peptide synthetases, *Nature 529*, 235-238.

[47] Herbst, D. A., Boll, B., Zocher, G., Stehle, T., and Heide, L. (2013) Structural basis of the interaction of MbtH-like proteins, putative regulators of nonribosomal peptide biosynthesis, with adenylating enzymes, *J Biol Chem 288*, 1991-2003.

[48] Tarry, M. J., Haque, A. S., Bui, K. H., and Schmeing, T. M. (2017) X-Ray Crystallography and Electron Microscopy of Cross- and Multi-Module Nonribosomal Peptide Synthetase Proteins Reveal a Flexible Architecture, *Structure 25*, 783-793 e784.

[49] Drake, E. J., Nicolai, D. A., and Gulick, A. M. (2006) Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain, *Chem Biol 13*, 409-419.

[50] Kadi, N., and Challis, G. L. (2009) Chapter 17. Siderophore biosynthesis a substrate specificity assay for nonribosomal peptide synthetase-independent siderophore synthetases involving trapping of acyl-adenylate intermediates with hydroxylamine, *Methods Enzymol 458*, 431-457.

[51] Wilson, D. J., and Aldrich, C. C. (2010) A continuous kinetic assay for adenylation enzyme activity and inhibition, *Anal Biochem 404*, 56-63.

[52] Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases, *Chem Biol 7*, 765-772.

[53] Tseng, C. C., Bruner, S. D., Kohli, R. M., Marahiel, M. A., Walsh, C. T., and Sieber, S. A. (2002) Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase, *Biochemistry 41*, 13350-13359.

[54] Piasecki, S. K., Taylor, C. A., Detelich, J. F., Liu, J., Zheng, J., Komsoukaniants, A., Siegel, D. R., and Keatinge-Clay, A. T. (2011) Employing modular polyketide synthase ketoreductases as biocatalysts in the preparative chemoenzymatic syntheses of diketide chiral building blocks, *Chem Biol 18*, 1331-1340.

[55] Haynes, S. W., Ames, B. D., Gao, X., Tang, Y., and Walsh, C. T. (2011) Unraveling terminal C-domain-mediated condensation in fungal biosynthesis of imidazoindolone metabolites, *Biochemistry 50*, 5668-5679.

[56] Franke, J., and Hertweck, C. (2016) Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges, *Cell Chem Biol 23*, 1179-1192.

[57] Foulke-Abel, J., and Townsend, C. A. (2012) Demonstration of starter unit interprotein transfer from a fatty acid synthase to a multidomain, nonreducing polyketide synthase, *Chembiochem 13*, 1880-1884.

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

[59] Iwig, D. F., Grippe, A. T., McIntyre, T. A., and Booker, S. J. (2004) Isotope and elemental effects indicate a rate-limiting methyl transfer as the initial step in the reaction catalyzed by Escherichia coli cyclopropane fatty acid synthase, *Biochemistry 43*, 13510-13524.

[60] Reimer, J. M., Aloise, M. N., Harrison, P. M., and Schmeing, T. M. (2016) Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase, *Nature 529*, 239-242.

[61] Bloudoff, K., and Schmeing, T. M. (2017) Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity, *Biochim Biophys Acta 1865*, 1587-1604.

[62] Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem Biol 6*, 493-505.

[63] Duckworth, B. P., Wilson, D. J., and Aldrich, C. C. (2016) Measurement of Nonribosomal Peptide Synthetase Adenylation Domain Activity Using a Continuous Hydroxylamine Release Assay, *Methods Mol Biol 1401*, 53-61.

[64] Hegazi, M. F., Borchard, R. T., and Schowen, R. L. (1976) Letter: SN2-like transition state for methyl transfer catalyzed by catechol-O-methyl-transferase, *J Am Chem Soc 98*, 3048-3049.

[65] Linscott, J. A., Kapilashrami, K., Wang, Z., Senevirathne, C., Bothwell, I. R., Blum, G., and Luo, M. (2016) Kinetic isotope effects reveal early transition state of protein lysine methyltransferase SET8, *Proc Natl Acad Sci U S A 113*, E8369-E8378.

Chapter 3

Investigating the Cyclization, Adenylation, and Stuffed Epimerase Domains of PchE from Pyochelin Biosynthesis

Introduction

Bacteria, fungi, and plants employ nonribosomal peptide synthetases (NRPSs) to generate a chemically and functionally diverse set of natural products. Aside from peptide bonds, other hallmarks of nonribosomal peptides (NRPs) include azoles/azolines/azolidines resulting from threonine (Thr), serine (Ser), or cysteine (Cys) cyclodehydration reactions, and inclusion of D-amino acids in the mature NRP. Introduction of heterocyclic azole derivatives and/or D-amino acids to NRPs are often important or necessary to induce their biological activity.[1-3] Additionally, the chemical uniqueness given by the altered stereochemistry of peptide α-carbons and/or heterocyclic azole derivatives creates NRPs that are less susceptible to proteolytic cleavage.[4, 5] These reasons likely contribute to the evolution of natural product biosynthesis to include heterocyclic azole derivatives which are found in effective secondary metabolites, such as siderophores, (i.e. pyochelin,[6] yersiniabactin,[7] vibriobactin,[8] mycobactin[9]) and some of which have been isolated with antibacterial (i.e. myxothiazole,[10, 11] bacitracin A,[12] zelkovamycin[13]), antitumor (i.e. bleomycin,[14] epothilone[15]), and immunosuppressant (i.e. argyrin[16]) properties. Similarly, NRPs isolated with D-amino acid stereochemistry include secondary metabolites such as siderophores (i.e. pyochelin,[6] yersiniabactin,[7] vibriobactin,[8] mycobactin[9]), and have also been found to possess antibacterial (i.e. tyrocidine A[17], bacitracin A[12]), antitumor (i.e. thiocoraline[18]), and immunosuppressant (i.e. cyclosporin A[19]) properties. Overall, many of these NRPs have

therapeutic implications for a variety of diseases. Understanding the mechanistic and structural aspects of NRPSs involved in heterocycle formation and inclusion of, or altered stereochemistry, continues to be a major importance in novel drug design and development.

The formation of heterocycles and inclusion of D-stereocenters are normally catalyzed by cyclization and epimerase domains. Interestingly, NRPS condensation domains, cyclization domains, and canonical epimerase domains share overall tertiary similarity. Condensation, cyclization, and epimerase domains are pseudodimers with each monomer structurally related to monomeric chloramphenicol acetyltransferase (CAT), although CATs exist as homotrimers and share low sequence identity with condensation, cyclization, and epimerase domains (**Figure 3-1**).[4, 20] Structural and mechanistic characterization of condensation, cyclization, and epimerase domains has provided a better understanding of their distinctive active site residues involved in catalysis. However, many questions regarding the molecular mechanisms remain.[4, 20-22] This chapter aims to further understand cyclization and epimerase domains in NRPS biosynthesis by studying PchE of pyochelin biosynthesis, a siderophore from *Pseudomonas aeruginosa*.

Pyochelin is a tricyclic NRP consisting of a hyrdroxyphenyl, D-thiazoline, and an *N*-methylated L-thiazolidinic acid (**Figure 3-2**). The biosynthesis of pyochelin employs two cyclization domains, for the cyclization of Cys residues to generate the thiazoline and thiazolidine rings, and a non-canonical stuffed epimerase domain, hypothesized to convert the L-thiazoline to the D-configuratrion.[23-25] PchE is proposed to be responsible for the addition and modification of the D-thiazoline, consisting of 5 different domains: an *N*-terminal aryl carrier protein domain, a cyclization domain, an adenylation and stuffed epimerase didomain, and a *C*-terminal peptidyl carrier protein domain (**Figure 3-3C**).[24]

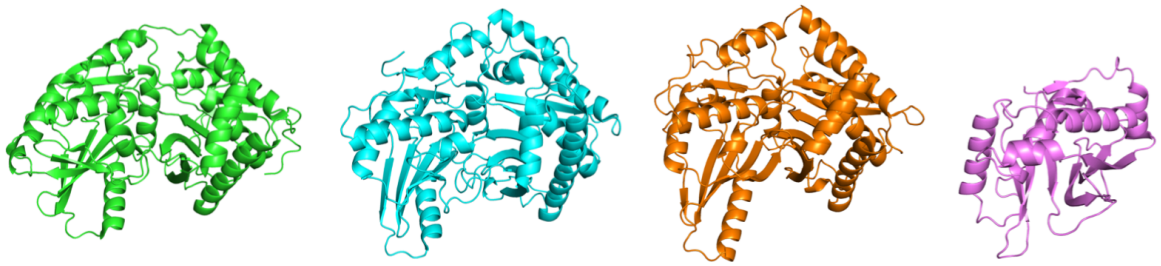


Figure 3-1: Condensation, cyclization, and epimerase domains share tertiary structure. NRPS condensation (green, PDB: 4ZXH, an NRPS from unknown pathway in *Acinetobacter baumanii*)[26], epimerase (cyan, 5ISW, from gramicidin S synthetase of gramicidin A biosynthesis from *Brevibacillus brevis*)[22], and cyclization (orange, 5T3E, from bacillamide synthetase from bacillamide biosynthesis in *Thermoactinomyces vulgaris*)[4] domains share a pseudo-dimeric tertiary fold similar to the tertiary structure of chloramphenicol acetyltransferase shown here as a monomer but naturally exists as a trimer (violet, PDB: 3CLA, from *Escherichia coli*).[27]

Previously, Walsh and colleagues performed *in vitro* experiments to elucidate the order and biochemical steps performed by PchE in pyochelin biosynthesis.[24, 25] The authors showed the adenylation domain was specific to L-Cys, not D-Cys, and that L-Cys retained its stereochemistry once loaded upon the Ppant tether.[24] Other experimental results aimed at understanding condensation, cyclization, and epimerization were inconclusive. However, the authors proposed the following model (**Figure 3-3A, top**). After L-Cys is primed to the Ppant tether (Cys_L-Ppant), a condensation reaction ensues with the upstream salicylate in the cyclization domain forming a hydroxyphenyl-L-Cys-S-Ppant (HP-Cys_L-Ppant). Next, HP-Cys_L-Ppant travels from the cyclization domain to the epimerase domain where the α-carbon of the Cys is abstracted and randomly re-protonated creating a new HP-Cys_D-Ppant intermediate or regenerating the HP-Cys_L-Ppant intermediate. The peptide then returns to the cyclization domain for cyclodehydration of the Cys creating a thiazoline (T) or $HPT_{L/D}$-Ppant biosynthetic intermediate. Walsh and colleagues propose that the downstream cyclization domain of PchF, the next biosynthetic enzyme, is a gatekeeper, only continuing chain elongation with HPT_D-Ppant and not HPT_L-Ppant. Indeed, condensation domains have displayed stereospecificity towards their substrates stereochemistry and epimerase activity of the acyclic HP-Cys_L-Ppant to HP-Cys_D-Ppant may be necessary to catalyze cyclodehydration to the HPT_D-Ppant intermediate, although this does not support the formation of its racemate, HPT_L-Ppant.[17, 28] Also, to our knowledge, this would be the only cyclization domain not performing condensation and cyclization chemistries back-to-back, suggesting PchE may have a novel cyclization domain.

Alternatively, we propose PchE sequentially condenses the salicylate to L-Cys-Ppant and immediately following, cyclizes the Cys to a thiazoline forming HPT_L-Ppant (**Figure 3-3A, bottom**). The peptide would then migrate to the stuffed epimerase domain altering the

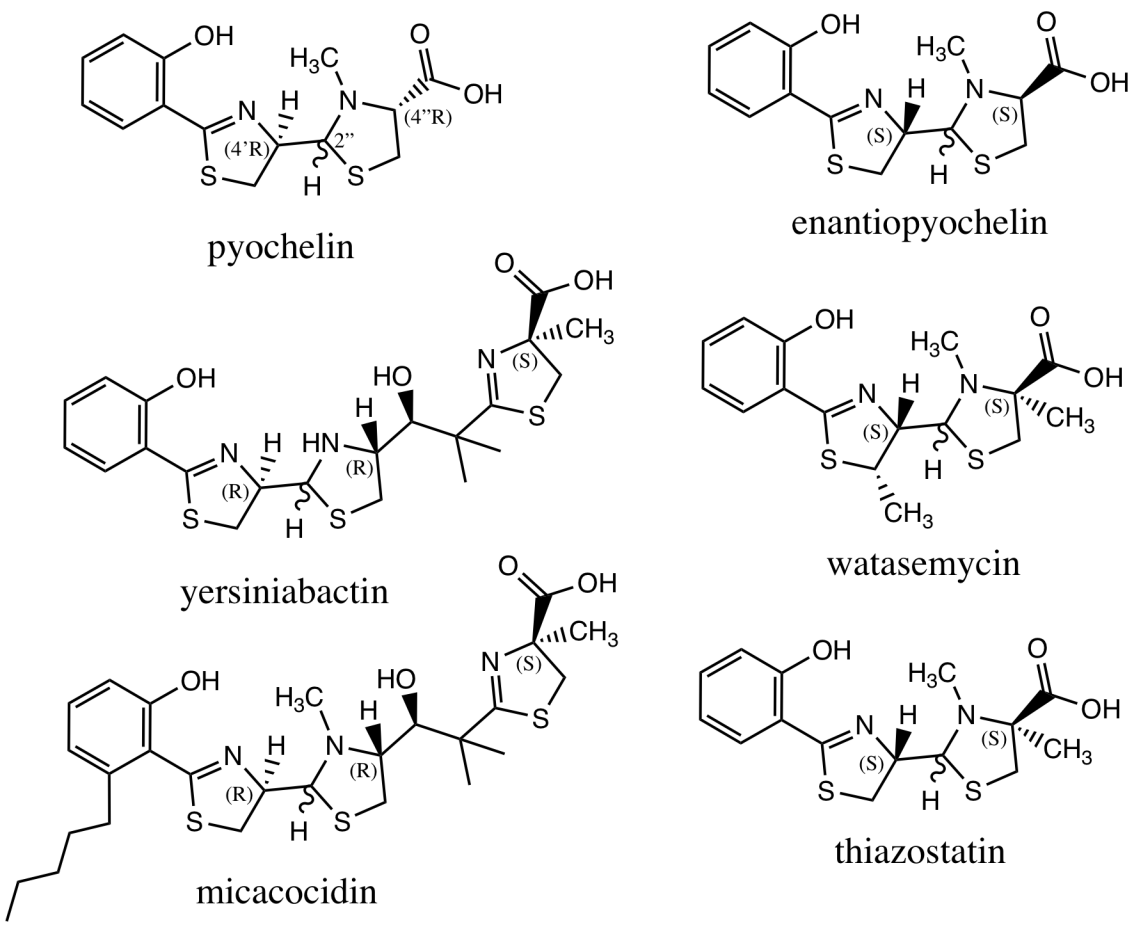


Figure 3-2: Different 2-hydroxyphenylthiazoline-4-thiazolidine natural products from bacteria. Pyochelin, yersiniabactin, micacocidin, enantiopyochlin, watasemycin, and thiazostatin are all secondary metabolites each having at least a tri-cyclic hydroxyphenyl, thiazoline, and thiazolidine ring. Interestingly, the stereochemistry of the thiazoline, thiazolidine ring system is either *R*, *R* or *S*, *S* for each natural product.

stereochemistry to HPT_D-Ppant before condensing with the downstream module. There are a few attributes supporting this hypothesis. This biosynthetic route would be more energetically favorable; only requiring two domain movements as opposed to three. Epimerization would also be more likely to occur on the thiazoline rather than the acyclic Cys as the α-carbon of a thiazoline would have a much lower pKa (**Figure 3-3A**).

Intriguingly, PchE's epimerase domain shares sequence homology with canonical methyltransferase domains but not canonical epimerase domains.[24, 25] However, the GxG *S*-adenosyl methionine (AdoMet) binding motif is RxG in PchE which would inhibit AdoMet binding to the stuffed tailoring domain. Yersiniabactin, a siderophore from *Yersinia pestis*, has a *C*-methyltransferase (M2 domain of Irp1) proposed to methylate the α-carbon of a thiazoline (**Figure 3-3B**).[29] The proposed mechanism for methylation of the thiazoline includes an active site base that abstracts a proton forming an enol-intermediate and prompting nucleophilic attack of the electrophilic *methyl* of AdoMet. The nucleophilic attack towards the *methyl* group occurs on the opposing side of the enol altering the stereochemistry of the thiazoline carbon to its D-isoform. PchE's stuffed epimerase domain may act similarly. An active site base abstracts a proton from the L-thiazoline forming an enol intermediate. When AdoMet is not present, the α-carbon becomes re-protonated by an active site acid forming a HPT_D-Ppant which condenses with the downstream peptide. Alternatively, if no active site acid exists, solvent re-protonation occurs until the correct stereochemistry is achieved.[24, 30]

This chapter aims to further elucidate the cyclization and epimerase chemistry performed by PchE. Herein we purify recombinant PchE (PchE-FL) and a shortened variant only consisting of the stuffed epimerase and adenylation didomain (PchE-AE) (**Figure 3-3C**). We synthesize

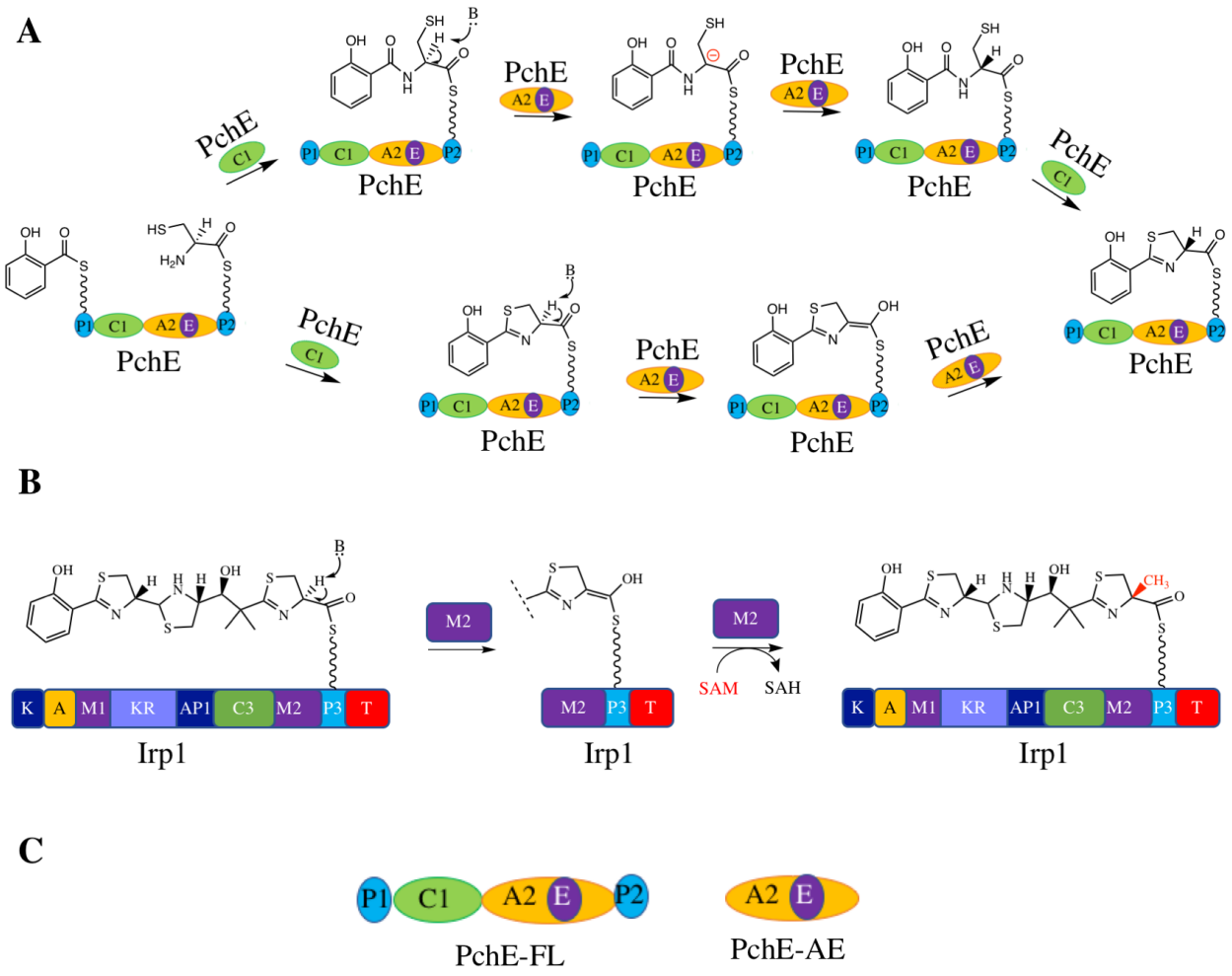


Figure 3-3: Proposed biosynthetic pathways of PchE in pyochelin biosynthesis, Irp1 methyltransferase proposed intermediate in yersiniabactin biosynthesis, and the PchE constructs used for experimentation. A) The proposed biosynthetic pathways of PchE in pyochelin biosynthesis. The currently accepted pathway in the literature is depicted on top. PchE's cyclization domain condenses the L-Cys to the upstream salicylate forming HP-Cys$_L$-Ppant. The α-carbon of the Cys is then abstracted by an active site base in the stuffed epimerase domain forming a carbanion. The carbanion is randomly re-protonated in succession before traveling back to the cyclization domain to be cyclized to form HPT$_{L/D}$-Ppant. The condensation domain of PchF is proposed to be specific only condensing the downstream aminoacyl-Ppant with HPT$_D$-Ppant. A second pathway is proposed on the bottom. Within the cyclization domain domain, the L-Cys is condensed to the upstream salicylate and cyclized to form HPT$_L$-Ppant. The peptidyl chain then travels to the stuffed epimerase domain where a proton is abstracted forming an enol intermediate before random re-protonation occurs generating HPT$_D$-Ppant. The bottom pathway requires only two domain movements compared to the top pathway requiring three movements between domains. **B)** The proposed enol intermediate of the cyclized thiazoline would be a shared intermediate with the proposed substrate of the independent tailoring domain

(M2) from Irp1 of yersiniabactin biosynthesis. The stuffed epimerase domain in PchE is sequentially similar to methyltransferase domains, although the AdoMet binding domain is mutated, suggesting AdoMet is unable to bind to PchE. We propose the mechanism for epimerase activity of the stuffed epimerase domain is similar to the mechanism of NRPS *C*-methyltransferases sharing the enol intermediate. **C)** PchE-FL, consisting of the full-NRPS module with two carrier protein domains, a cyclization domain, and an adenylation and stuffed epimerase didomain. In this work, we purified a full-length enzyme (PchE-FL) and a variant enzyme consisting of only the adenylation and stuffed epimerase didomian (PchE-AE).

substrate and product analogs and attempt assays to further define the order in which PchE performs its chemical steps during pyochelin biosynthesis.

Materials and Methods

pche-fl Overexpression Plasmid

The overexpression construct for pche-fl was gifted to the lab from the laboratory of Christopher Walsh.[1] The construct received was in pET28b, a kanamycin resistant plasmid (Novagen).

pche-ae Overexpression Plasmid

The pche-ae variant plasmid was generated by site directed mutagenesis using Herculase polymerase supplemented with 4-10% DMSO and the genomic DNA from *Pseudomonas aeruginosa* PAO1 which was purified using the DNeasy® Blood & Tissue Kit (Qiagen) as per the manufacturer's instructions. The forward primer (5'-ATT AGA CAT ATG CCG TTG CCC TGG GCG CAG CAG-3') including a NdeI site (underlined) and reverse primer (5'-ATA CTC GAG TCA AAG TGC GCG GGT CAT GCT C-3') including an XhoI site (underlined) were used to amplify the gene of interest. The amplified sequence was ligated into the kanamycin resistant pET28b plasmid (Novagen) digested with the same restriction enzymes. The isolated plasmid sequence was sequenced and encodes the entire adenylation domain and proposed epimerase domain, residues 532 to 1339 of PchE-FL.

pa2412 Overexpression Plasmid. The pa2412 overexpression plasmid was a gift from the laboratory of Andrew Gulick.[31] PA2412 is an MbtH-like protein (MLP) that has been shown to increase *P. aeruginosa* NRPS protein expression including PchE and PchF (data not shown).[32-35]

The *pa2412* gene is in a modified pET15b plasmid with ampicillin resistance and contains a 5XHis affinity tag in addition to a TEV protease cleavage recognition site for thrombin proteolysis. [31]

PchE-FL protein overexpression and purification

BL21 (DE3) pLysS *E. coli* cells containing the *pa2412* and *pche-fl* expression plasmids were grown in LB broth containing 50 µg/mL kanamycin and 200 µg/mL ampicillin at 37°C with shaking (250 rpm). Protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 200 µM after the OD_{600} reached 0.6. The temperature of the culture was lowered to 30˚C following induction. After 3 hours, the cells were harvested by centrifugation (4 230 x g, 10 min, 4˚C). The pellet was resuspended in 25 mM potassium phosphate pH 8.0, 50 mM sodium citrate, 500 mM NaCl, 5 mM imidazole, 1 mM n-octyl-β-D-glucopyranoside, and 10% glycerol (buffer A). The resuspended pellet was flash frozen at -80˚C until further use. While thawing, 10 µL of 75 mg/mL lysozyme was added to the pellet. The cells were disrupted using a French pressure cell (35 000 psi), and cellular debris were removed by centrifugation (11 950 x g, 40 min, 4˚C). The supernatant was injected onto a nickel charged Sepharose fast-flow column (Amersham Biosciences) pre-equilibrated in buffer A. After 8 column volumes of buffer A, the protein eluted by linear gradient of buffer B (buffer A supplemented with 250 mM imidazole) from 0-100% over 10 column volumes. The fractions containing PchE-FL and pa2412 were pooled together and dialyzed using SnakeSkin® Dialysis Tubing (10 kDa cutoff) at 4˚C against 25 mM potassium phosphate pH 8.0, 1 mM n-octyl-β-D-glucopyranoside, 1 mM dithiothreitol, 1 mM ethylenediaminetetraacetic acid, and 10% glycerol (buffer C). The dialysate was exchanged twice after incubating one hour and once more before

incubating overnight. The dialyzed sample was concentrated using a stirred cell with YM30 membrane filter and applied to a Source 30Q (GE Healthcare Life Sceinces) anion exchange column pre-equilibrated in buffer C. After injection, the column was washed with 4 column volumes of buffer C and followed with a 50% step gradient of buffer D (buffer C supplemented with 500mM NaCl and 1 mM n-octyl-β-D-glucopyranoside) for 4 column volumes followed by 100% buffer D for 5 column volumes. The fractions containing PchE-FL were identified by SDS-PAGE, collected, and concentrated as above. The concentrated sample was subjected to a Superdex 200 size-exclusion column (Amersham Biosciences) equilibrated in 50 mM Tris pH 8.0, 50 mM sodium citrate, 1 mM n-octyl-β-D-glucopyranoside, 1 mM dithiothreitol, and 10% glycerol (buffer E). PchE-FL fractions were collected and concentrated as above. The final concentration of PchE-FL was determined by the general Beer-Lambert law and the predicted A_{280} of ε = 190 675 M^{-1} cm^{-1} (ProtParam).[36] A total of 3.8 mg of PchE-FL was obtained from the 3 L cell growth. The protein was flash cooled and stored at -80˚C until further used.

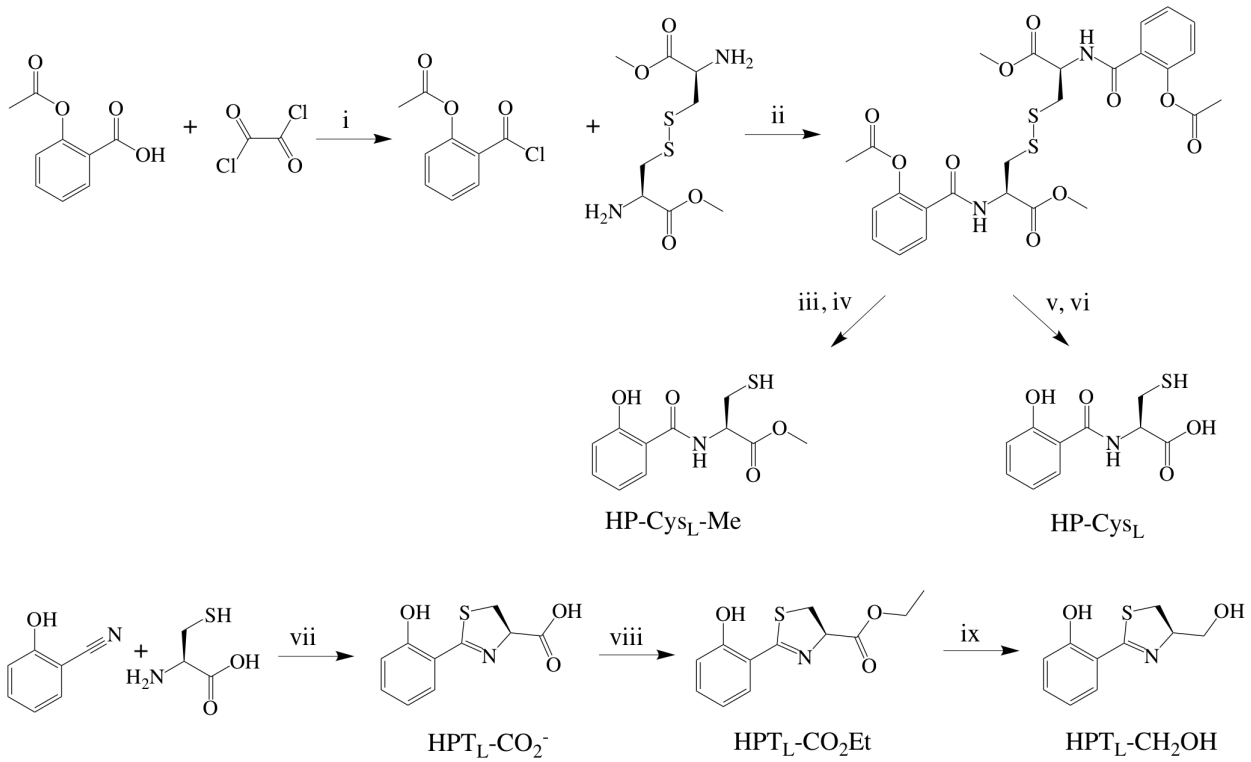
PchE-AE protein overexpression and purification

BL21 (DE3) E. coli containing the *pche-ae* and *pa2412* expression plasmids were grown in LB broth containing 50 μg/mL kanamycin and 200 μg/mL ampicillin at 37 °C with shaking (250 rpm). Protein expression was induced when the OD_{600} reached ~0.7 by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 200 μM. The temperature was lowered to 15 °C following induction. The cells were harvested by centrifugation (4 230 x g, 10 min, 4 °C) after ~20 hours. The cell pellet was resuspended in 25 mM potassium phosphate buffer pH 7.5, 500 mM NaCl, 50 mM sodium citrate, 1 mM n-octyl-β-D-glucopyranoside, 5 mM imidazole, and 10% glycerol (buffer A). Cells were disrupted by French press (35 000 psi), and

cellular debris was removed by centrifugation (23 430 x g, 60 min, 4 °C). The supernatant was applied to a chelating Sepharose fast-flow column (Amersham Biosciences) charged with nickel chloride and pre-equilibrated in buffer A. Four-column volumes of buffer A were followed by a 4-column volume wash with 10% buffer B (buffer A supplemented with 500 mM imidazole). An 8-column volume linear gradient to 100% buffer B was used to elute the PchE-AE and PA2412 proteins together in 1 peak at ~125 mM imidazole. The pooled fractions were dialyzed using SnakeSkin® Dialysis Tubing (10 kDa cutoff) at 4˚C against 50 mM potassium phosphate buffer pH 7.5, 1 mM DTT, 50 mM sodium citrate, 1 mM n-octyl-β-D-glucopyranoside, and 10% (v/v) glycerol (buffer C). The dialysate was exchanged for fresh buffer two times after stirring for an hour each, before being exchanging to dialyze overnight. The dialyzed protein was concentrated using a 30 kDa cutoff Amicon ultracentrifugal spin filter. The concentrated protein was applied to a Superdex 200 size-exclusion column (Amersham Biosciences) equilibrated with buffer C. PchE-AE eluted separately from PA2412 and the PchE-AE fractions were concentrated as above to 1.62 mg/mL as determined by the general Beer-Lambert law and the predicted A_{280} of $\varepsilon = 106\,505\ M^{-1}\ cm^{-1}$ (ProtParam).[36] A total of 10.7 mg was obtained from the 3 L of cell growth. The protein was flash cooled and stored at -80 °C. For the adenylation assay, 1 mL of purified PchE-AE was subjected to a PD-10 column to exchange the buffer to 50 mM Tris pH 8.0, 1 mM DTT, 50 mM sodium citrate, and 10% (v/v) glycerol (buffer D) and was concentrated to 1.70 mg/mL, determined analogously, and was flash cooled and stored at 80˚C until further use.

*Preparation of Substrate Analogues (**Scheme 3-1**).*

General experimental procedures. Initial compounds used in synthesis were purchased from chemical suppliers are used without any further purification methods. Chromatic separations



Scheme 3-1: Substrate analog synthesis of HP$_L$-Cys, HP-Cys$_L$-Me, HPT$_L$-CO$_2^-$, HPT$_L$-CO$_2$Et, HPT$_L$-CH$_2$OH. Substrate analogs were generated to assess the epimerase capabilities of PchE. Product analogs were synthesized by replacing the starting reagent with its D-isomer. Reagents and conditions: i) DMF/CH_2Cl_2, RT; ii) pyridine/CH_2Cl_2, 0˚C; iii) dithiothreitol/3N HCl in MeOH; iv) 1M NH_4OH, pH 9.0/EtOH; v) dithiothreitol, NH_4OH, pH 9.0/EtOH, RT; vi) NaOH/MeOH, RT; vii) phosphate buffer, potassium carbonate, pH 8/MeOH, 50˚C; viii) $SOCl_2$/EtOH, 40 °C; ix) $Ca(BH_4)_2$/EtOH.

were performed with a Teledyne Isco CombiFlash Rf. or by flash column chromatography using silica gel (4-63 mm) purchased from Sorbent Technologies. Chromatography was also performed with a Teledyne Isco CombiFlash Rf. ^{1}H and ^{13}C NMR were recorded using a 500 MHz Bruker AVIII spectrometer equipped with a cryogenically-cooled carbon observe probe using tetramethyl silane as an internal standard. Chemical shifts (δ) are reported in ppm and coupling constants (*J*) are reported in Hz. High-resolution mass spectrum (HRMS) was performed on a LCT Premier (Micromass Ltd., Manchester UK) time of flight mass spectrometer with an electrospray ion source in either positive or negative mode. Melting points were measured with a Thomas Capillary Melting Point Apparatus and are uncorrected. Chiral analysis to determine racemic ratio was carried out using either a Chiralcel-OD, Lux Amylose-1 or Lux Cellulose-1 column from Phenomenex using a Waters Acquity UPLC/UHPLC. Peaks were monitored at 254 nm and 320 nm. The conditions for each racemic pair are included below.

ethyl (*R*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylate (HPT_L-CO_2Et).

Thionyl chloride (10 mmol) was added dropwise at -10˚C to EtOH (10 mL) before the addition of (*R*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylic acid (HPT_L-CO_2^-,[37] 2.25 mmol). The reaction was warmed to 40˚C and stirred overnight. The mixture was concentrated to a black sludge under reduced pressure and reconstituted in Na_2CO_3 (100 mL). After filtering through a pad of celite the solution was extracted with CH_2Cl_2 (3 x 50 mL). The organic phase was concentrated under reduced pressure and the residue was further purified by Combiflash (0:1 to 1:3 MeOH in CH_2Cl_2). The major peak was collected and concentrated under high pressure overnight yielding an light orange oil (1.6 mmol, 71%). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.39 (s, 1H), 7.46 (t, *J* = 7.6 Hz, 2H), 7.03 – 6.92 (m, 2H), 5.57 (dd, *J* = 9.7, 7.5 Hz, 1H), 4.20 (q, *J* =

7.1 Hz, 2H), 3.76 (dd, *J* = 11.4, 9.7 Hz, 1H), 3.64 (dd, *J* = 11.4, 7.5 Hz, 1H), 1.24 (t, *J* = 7.1 Hz, 3H). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 173.21, 169.84, 158.32, 133.80, 130.56, 119.37, 116.87, 115.53, 76.09, 61.36, 33.33, 13.96. **HRMS**: $[M + H]^+$ 252.0716 (calcd) 252.0702 (found). Ethyl (*S*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylate was generated by replacing the starting HPT_L-CO_2^-,with HPT_D-CO_2^-,. The final ethyl (*S*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylate (HPT_D-CO_2Et) product produced matching values for ^{1}H NMR, ^{13}C NMR, and HRMS.

(R)-2-(4-(hydroxymethyl)-4,5-dihydrothiazol-2-yl)phenol (HPT_L-CH_2OH).

Calcium chloride (0.9 mmol) and sodium borohydride (0.9 mmol) were added to a solution of EtOH (5 mL) at -20˚C. A solution of ethyl (*R*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylate (0.9 mmol) in EtOH (5 mL) was added dropwise to the calcium borohydride solution at -20˚C with vigorous spinning. The reaction was left at -20˚C and stirred overnight. The mixture was concentrated under reduced pressure before being quenched with deionized H_2O (50 mL). The mixture was extracted with CH_2Cl_2 (3 x 50 mL) and the combined organic phase was dried with Na_2SO_4. The product was purified by Combiflash (0:1 to 1:3 EtOAc gradient in hexanes over 30 minutes) yielding a yellow solid (0.16 mmol, 18%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.65 (s, 1H), 7.41 (td, *J* = 7.2, 1.6 Hz, 2H), 7.01 – 6.90 (m, 2H), 5.11 (d, *J* = 6.6 Hz, 1H), 4.80 (ddt, *J* = 8.9, 7.5, 5.4 Hz, 1H), 3.68 – 3.58 (m, 2H), 3.50 (dd, *J* = 10.9, 8.9 Hz, 1H) 3.31 (dd, *J* = 10.9, 7.5 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 171.06, 158.39, 133.21, 130.32, 119.11, 116.69, 115.79, 77.82, 62.31, 32.56. **HRMS**: $[M + H]^+$ 210.0610 (calcd) 210.0599 (found). *(S)*-2-(4-(hydroxymethyl)-4,5-dihydrothiazol-2-yl)phenol (HPT_D-CH_2OH)

was generated by replacing the starting HPT_L-CO_2Et with HPT_D-CO_2Et. The final HPT_D-CH_2OH product provided matching values for ^1H NMR, ^{13}C NMR, and HRMS.

Steady-state Adenylation Assay

The steady-state adenylation assay is thoroughly described in the PchF chapter and similar conditions are briefly described here. Each reaction contained 1mM $MgCl_2$, 200 μM 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESGR), 0.002 U purine nucleoside phosphorylase (PNP, Sigma), 0.001 U inorganic pyrophosphatase (IPP, Sigma), 2 mM ATP, 150 mM hydroxylamine, 0.375 μM enzyme, and varying concentrations of L-Cys. Reactions were also attempted with D-Cys, but no activity was observed (**Figure 3-4**). Stock concentrations of L-Cys, ATP, and hydroxylamine were prepared the day of the reaction. L-Cys stock concentrations were serially dilute in 25 mM Tris pH 8 and 5 mM dithiothreitol. A parent mixture containing all components of the assay except substrate were incubated for 15 minutes at room temperature while the substrate was aliquoted into a 96-well flat bottom plate (Corning Cat. #9017). A multichannel pipette was then used to inject the reaction mixture into individual wells to initiate the 100 μL reaction. The production of 2-amino-6-mercapto-7-methylpurine (MesG) was monitored by an increase in absorbance at 360 nm with a Cary 50 UV-VIS Spectrophotometer with a plate reader attachment. The path length of the plate reader (0.29 cm) and the extinction coefficient of MesG (11,000 $cm^{-1}M^{-1}$)[38] were used in the general Beer-Lambert's law to convert AU/min to μM/min. The rate was then halved due to the catalytic hydrolysis of pyrophosphate to two inorganic phosphates by inorganic pyrophosphatase (IPP) which are then successively used by purine nucleoside phosphorylase (PNP) to convert MESGR to MESG. Negative controls were performed with each triplicate reaction: without enzyme, without L-Cys, and without ATP. The

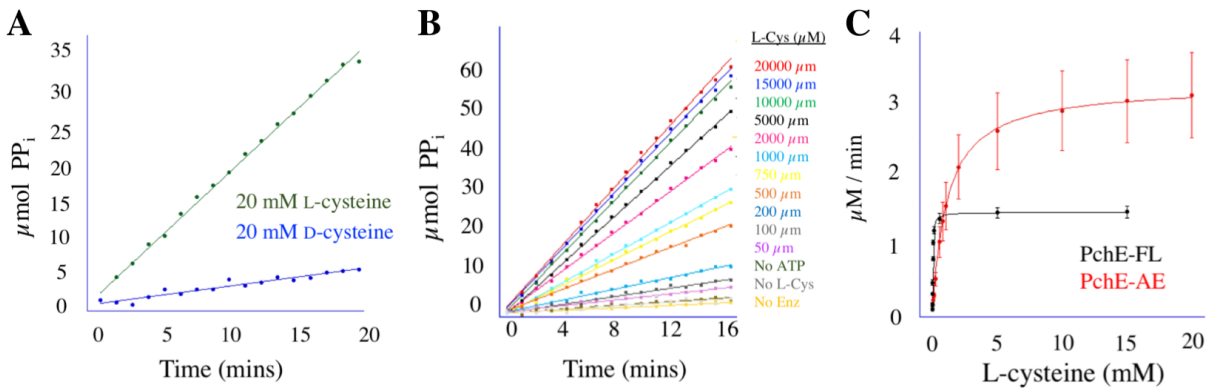


Figure 3-4: Adenylation assay. The adenylation domain of PchE catalyzes the formation of an aminoacyl-AMP bond from L-Cys and ATP, releasing pyrophosphate. A full description of the assay can be found in Chapter 2, Figure 2-5. Briefly, the assay to measure the adenylation capability of PchE-FL and PchE-AE is a coupled assay. Pyrophosphate released from the aminoacyl-AMP bond formation is hydrolyzed to two inorganic phosphates by inorganic pyrophosphatase. Each inorganic phosphate is then used to catalyze the phosphorylation of 7-methylthioguanosine (MesGR) by purine nucleoside phosphorylase, producing ribose 1-phosphate and 7-methylthioguanine (MesG), which absorbs at 360 nm. Because the natural substrate is replaced in the adenylation domain by the nucleophilic attack of the Ppant tether, NH_2OH is used as a nucleophilic surrogate to allow for steady-state turnover. **A)** An adenylation assay showing the difference in activity with 20 mM L-Cys (green) and 20 mM D-Cys (blue) using PchE-FL. D-Cys is within error of experimental negative controls and does not appear to be a substrate for PchE. **B)** PchE-AE turnover with various amounts of L-Cys and ATP. **C)** Michaelis-Menten steady state analysis of PchE-FL and PchE-AE with L-Cys. Two separate sets of triplicate trials were performed on different days for a total of six trials.

greatest rate generated by a negative control was subtracted from all of the calculated reaction rates. Two separate triplicate trials were performed on successive days for each enzyme. The initial calculated rates were fit to the Michaelis-Menten equation using the nonlinear regression function in KaleidaGraph (Synergy Software). Kinetic parameters were generated for each of the six individual trials for each enzyme. The reported kinetic parameters were calculated by finding the average and the standard deviation of each parameter was considered the error.

Epimerase Assay by Circular Dichroism

Circular dichroism (CD) spectra using ultraviolet (UV) absorbance of reactions with and without PchE-FL or PchE-AE were recorded on a Jasco J-815 CD spectrometer at 20.0˚C in 50 mM Tris pH 8.0 and 5mM TCEP. For each spectrum, three measurements were taken for every 2 nm and averaged from 240 nm to 360 nm. Spectra are presented in millidegrees of ellipticity. A solution of 100 μM substrate in 50 mM Tris pH 8.0 and 5 mM tris(2-carboxyethyl)phosphine) (TCEP) was made from a DMSO stock for a final volume of 2 mL. A CD measurement was taken as before addition of enzyme and directly after the reaction was initiated with 2.0 μM PchE-FL or PchE-AE. Subsequent measurements were taken at 5, 10, 15, 30, 45, and 60 minutes. Reactions using HPT_L-CO_2Et were extended to 4 hours. The reactions were performed at 20.0˚C. An additional endpoint assay was performed after a reaction incubated at room temperature for 8 hours.

Epimerase Assay by Chiral Chromatography

Separation of synthesized racemic pairs was screened using Lux Amylose-1 and Lux Cellulose-1 columns with 4.6 mm x 100 mm dimensions and 5 μm beads purchased from

Phenomenex (Torrance, CA). Each racemic pair was subjected to a series of isocratic trials to optimize separation conditions. This was performed by varying the eluent by 10% ranging from 10-90%. Standards were diluted using EtOH at a concentration of 1 mg/mL and 3 µL was injected onto the column. The mobile phase was a variation of H_2O/ACN/TFA running at 0.5 mL/min or 1 mL/min. An Acquity UPLC from Waters (Millford, MA) was used to perform the chromatography. HP-Cys_L-Me (rt 7.49 mins) and HP-Cys_D-Me (rt 7.05 mins) were separated using the Lux Cellulose-1 column with a mobile phase of H_2O/ACN/TFA (60:40:0.1). HPT_L-CO_2^- (rt 7.35 mins) and HPT_D-CO_2^- (rt 6.80 mins) was also separated using the Lux Cellulose-1 column with a mobile phase of H_2O/ACN/TFA (70:30:0.1). HP-Cys_L (rt 8.73 mins) and HP-Cys_D (rt 7.84 mins) were separated using the Lux Cellulose-1 column with a mobile phase of H_2O/ACN/TFA (85:15:0.1). HPT_L-CO_2Et (rt 3.77 mins) and HPT_D-CO_2Et (rt 4.45 mins) were separated using the Lux Amylose-1 column with a mobile phase of H_2O/ACN/TFA (40:60:0.1). HPT_L-CH_2OH (rt 2.87 mins) and HPT_D-CH_2OH (rt 1.89 mins) were separated using the Lux Amylose-1 column with a mobile phase of H_2O/ACN/TFA (40:60:0.1). Each enzyme reaction was performed in 50 mM Tris pH 7.5 with 1 mM L-substrate analog, except HPT_L-CO_2Et (300 µM), and 1 µM PchE-FL or 1 µM PchE-AE. At 0 hr., 1 hr., 2hr., and 4hr., 100 µL of the reaction was quenched with 5 µL of 1 mM H_2SO_4. Quenched reactions were centrifuged to rid the enzyme before introduction to the column (10 600 x g, 10 mins, 4 ˚C). 3 µL of the reaction was injected onto the column. A control reaction without enzyme was performed to monitor non-enzymatic racemization.

Cyclization Assay

Incubation of HP-Cys_L-Me or HP-Cys_D-Me with PchE-AE or PchE-FL resulted in an increased absorbance from 310 – 350 nm with a peak max at 335 nm and an increased fluorescence at 413 nm by excitation at 310 nm. A 250 μM solution of HP-Cys_L-Me and HP-Cys_D-Me was made in 50 mM Tris pH 8.0 buffer from a 100 mM DMSO stock and was sequentially diluted for reaction concentrations. PchE-AE was added to a final concentration of 1.88 μM and fluorescence was measured from 350 – 550 nm by excitation at 310 nm very 2 minutes for 34 minutes. Fluorescent measurements were carried out on a Varian Cary Eclipse fluorescence spectrophotometer with an excitation and emission slit set at 5 nm. Full progress curves were performed, and a model was generated by KinTek explorer using steady-state approximations.[39, 40]

Results

Purification of PchE-FL and PchE-AE

Two variants of PchE were purified by heterologous expression in *E. coli* cells. The first was the full-length enzyme, PchE-FL, consisting of two peptidyl carrier domains, a cyclization domain, and a stuffed epimerase-adenylation didomain. The focus of this study was the stuffed tailoring epimerase domain. Therefore, a second construct, called PchE-AE, was made consisting of only the stuffed epimerase-adenylation didomain, removing the possibility of catalytic contribution coming from the carrier and cyclization domains while attempting to assesses specific catalytic activities provided by the tailoring domain **(Figure 3-1C)**. Each construct was co-expressed with the MbtH-like protein (MLP), PA2412, from the pyoverdin biosynthetic cluster from *P. aeruginosa*, to enhance NRPS solubility.[31, 32, 41, 42] Discussed in the PchF chapter,

MLPs are often necessary for expression and are frequently observed to increase solubility or catalytic efficiency of adenylation activity. Similarly, the stability of recombinantly expressed PchF, PchE, and variant constructs of PchF and PchE is improved when co-expressed with PA2412. PchE-FL and PchE-AE were purified by a N-terminal His-tag. Size exclusion chromatography was used to separate PchE-FL and PchE-AE from PA2412. An additional purification step of anion exchange chromatography was used during purification of PchE-FL.

Steady-state kinetics of adenylation activity.

Adenylation activity of PchE-AE and PchE-FL assessed proper folding of the variants. PchE is responsible for the incorporation of the Cys that becomes the first thiazoline of pyochelin. In the final pyochelin product, the α-carbon has a stereochemistry that must result either from the incorporation of D-Cys or the epimerization of an L-Cys. Walsh and colleagues previously reported that the adenylation domain of PchE was specific to L-Cys and did not catalyze the adenylation of D-Cys, suggesting another catalytic event was responsible for epimerization.[23] Therefore, the adenylation domain of PchE is hypothesized to catalyze the formation of an aminoacyl-AMP bond from L-Cys and ATP with release of pyrophosphate. The next step in NRPS assembly would be for the Ppant tether of the peptidyl carrier domain to perform a nucleophilic attack binding the amino acid to the thiotemplate and releasing AMP from the active site. In the experiments described in this chapter, PchE-FL and PchE-AE were not post-translationally modified with a Ppant tether to allow for steady-state turnover of the adenylation reaction, as described in the previous chapter. Hydroxylamine was used as a nucleophilic surrogate in place of the Ppant tether to release the aminoacyl-AMP intermediate promoting adenylation activity in the steady-state. The pyrophosphate is converted to two

inorganic phosphates by inorganic pyrophosphatase which are subsequently used by purine nucleoside phosphorylase to generate a chromogenic nucleoside. In our assessment, PchE-FL and PchE-AE were able to catalyze turnover of L-Cys but not D-Cys as previously observed by Walsh **(Figure 3-4A)**.[23] Steady-state kinetic parameters of adenylation activity were determined with L-Cys for PchE-FL and PchE-AE and can be found in **Figure 3-4** and **Table 3-1**. The k_{cat} for PchE-FL was almost 2.5-fold less than for PchE-AE and, surprisingly, the K_m was nearly 2 magnitudes less for PchE-FL than PchE-AE. The variants of PchF from the previous chapter were within 2-fold of one another for k_{cat}, K_m, and k_{cat}/K_m. The variant PchF-AMP still contained the peptidyl carrier domain whereas PchE-AE only consists of the stuffed tailoring and adenylation didomain.

Substrate and product analog synthesis.

Analyzing mechanisms of catalytic steps during NRPS biosynthesis becomes increasingly difficult as the peptide becomes elongated. Intermediates within the biosynthesis are tethered to a Ppant chain further increasing the complexity of the natural substrate. Thus, intermediate analogs have been synthesized in which the terminal thioester linkage is altered to elucidate catalytic mechanisms of different domains in nonribosomal peptide and polyketide biosynthesis.[15, 43-46] Our lab has observed and reported catalytic efficiency with substrate analogs designed for the pyochelin biosynthetic pathway. PchG, discussed in the following chapter, was found to catalyze the NADPH-dependent reduction of an intermediate analog replacing the thioester linkage with a terminal carboxylic acid, but was unable to catalyze reduction of compounds replacing the terminal acid with an ester or alcohol. Alternatively, PchF catalyzed *S*-

Table 3-1: L-Cysteine Adenylation Assay of PchE-FL and PchE-AE			
	PchE-FL*	PchE-FL	PchE-AE
k_{cat} (min^{-1})	38	3.7 ± 0.3	9 ± 2
K_M (L-Cys μM)	110	23 ± 3	1180 ± 70
k_{cat}/K_M ($M^{-1}s^{-1}$)	6700	2700 ± 100	120 ± 30

* - Quadri, Keating, Patel, Walsh. *Biochemistry* 1999, 38, 14941-14954. Walsh *et al* used an assay to measure adenylation activity by the incorporation of ^{32}pyrophosphate into ATP through the reverse reaction, a different assay than the coupled assay used herein.

adenosyl methionine (AdoMet)-dependent methyl transfer to an intermediate analog with a terminal ethyl ester, but not with a similar compound replacing the ester with a carboxylic acid.

Proposed intermediate substrates for the stuffed epimerase domain of PchE were generated to determine when in biosynthesis epimerization occurs: 1) after condensation to the upstream salicylate but before cyclization of the Cys to a thiazoline, or 2) after cyclization to the thiazoline but before condensation to the downstream peptide. To elucidate the substrate of the stuffed epimerase domain, multiple substrate analogs of each proposed pathway were synthesized **(Scheme 3-1)**. Both substrate (α-carbons with L-stereocenters) and product analogs (with D-stereocenters) were synthesized.

Synthesis of analogs representing the acyclic hydroxyphenyl-Cys (HP-Cys) were performed in accordance to Bergeron *et al*.[47] Synthesis was initiated by nucleophilic addition of L-cystine dimethylester (or D-cystine dimethylester) to acetylsalicoyl chloride providing the diamide in 35% yield, similar to the 31% yield previously reported.[47] The diamide was subsequently used to generate both HP-Cys and HP-Cys-Me analogs. HP-Cys was generated by breaking the disulfide linkage of the diamide with dithiothreitol and a subsequent saponification step by NaOH to provide a 49% yield. The synthesis of HP-Cys-Me was continued from the diamide by selective deprotection of the hydroxyphenyl acetyl followed by reduction of the disulfide with dithiothreitol to provide a 54% yield.[47]

Synthesis of HPT-CO_2^- was performed in alignment with a protocol provided by Bergeron *et al*.[37] L-Cys (or D-Cys) was added to 2-hydroxybenzonitrile allowing for nucleophilic addition and cyclization of the Cys to a thiazoline retaining the stereochemistry at the α-carbon providing a comparable 74% yield. HPT-CO_2^- was altered to an ethyl ester by converting the carboxylic acid to an acyl chloride promoting nucleophilic addition of the ethanol solvent

providing HPT-CO_2Et in 71% yield. The ethyl ester was further modified to generate HPT-CH_2OH by a calcium borohydride reduction to afford an alcohol, HPT-CH_2OH, in 18% yield.

A summary of each substrate and product analog used in this study can be found in **Figure 3-5**. Chiral purity was assessed for each synthetic substrate and product analog by chiral chromatography and enantiomeric ratios are summarized in **Table 3-2**.

Cyclization assay.

The proposed catalytic pathway by PchE in pyochelin biosynthesis suggests L-Cys is activated in the adenylation domain and is then added to the Ppant tether of the peptidyl carrier domain. The Cys then condenses with the upstream salicylate in the cyclization domain before moving to the stuffed epimerase domain to alter the stereochemistry of the Cys α-carbon from *(S)* to *(R)*. The altered peptide must then travel back to the condensation domain to be cyclized before condensation with the peptide from the downstream module. An experiment to test the proposed pathway is to identify if PchE is able to catalyze cyclization of the synthesized analogs: arrangement of the substrate is necessary for PchE catalyzed cyclization, L- and D- analogs could be used to substantiate the order of chemistry in the biochemical pathway. If cyclization takes place prior to epimerization, we would assume PchE would be more specific towards cyclizing L-analogs over the D-analogs. Conversely, if epimerization occurs on the acyclic intermediate before cyclization, PchE cyclization may be specific for the D-analogs and unable to catalyze cyclization of the L-analogs.

After synthesis, stock concentrations (100 mM) of the potential substrate analogs, HP-Cys_L, HP-Cys_D, HP-Cys_L-Me, or HP-Cys_D-Me, were made in dimethylsulfoxide (DMSO) and stored at room temperature. Addition of the potential substrates to a solution containing PchE-FL

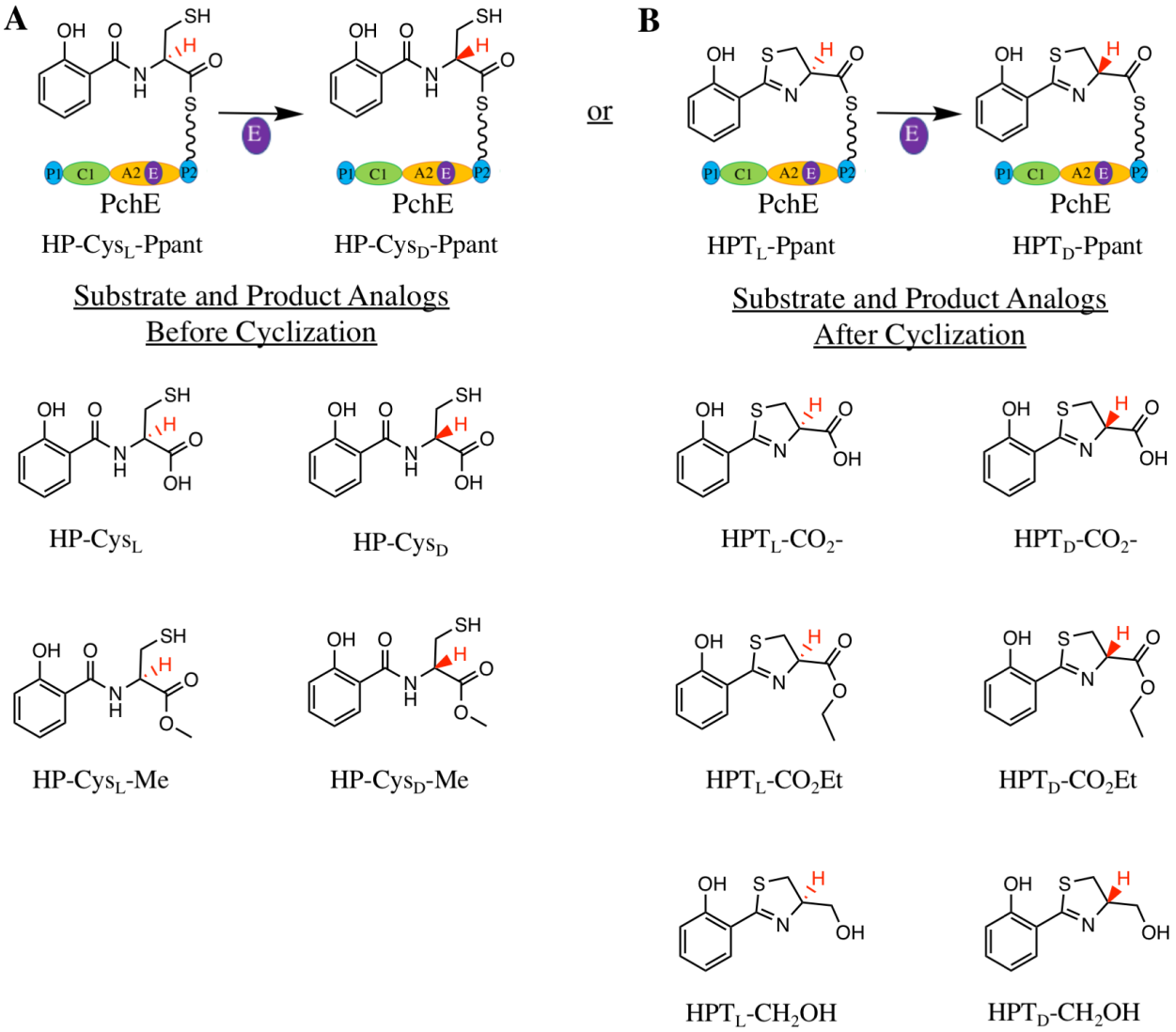


Figure 3-5: Summary of substrate and product analogs generated for PchE epimerase assays. The natural substrate and product intermediates during peptide elongation are tethered to the Ppant tail covalently linked to the peptidyl carrier domain. Therefore, substrate and product analogs of the naturally thiolated intermediates were generated. **A)** The substrate and product analogs of proposed PchE epimerase intermediates before cyclization: HP$_L$-Cys and HP$_D$-Cys, and HP$_L$-Cys-Me and HP$_D$-Cys-Me. **B)** Substrate and product analogs synthesized mimicking the proposed epimerase intermediate after cyclization: HPT$_L$-CO$_2^-$ and HPT$_D$-CO$_2^-$, HPT$_L$-CO$_2$Et and HPT$_D$-CO$_2$Et, and HPT$_L$-CH$_2$-OH and HPT$_D$-CH$_2$-OH.

Table 3-2: Synthesized substrate and product analog ratios determined by chiral chromatography

	Retention time		L-substrate analog		D-product analog	
	L	D	L (%)	D (%)	L (%)	D (%)
HP-Cys	8.73	7.84	99.1	0.9	0	~100
HP-Cys-Me	7.05	7.49	99.6	0.4	0	~100
HPT-CO_2^-	7.97	7.44	94.0	6.0	6.1	93.9
HPT-CO_2Et	3.77	4.45	87.1	12.9	91.1	8.9
HPT-CH_2OH	2.87	1.89	81.7	18.3	77.9	22.1

L-HP-Cys, D-HP-Cys, L-HP-Cys-Me, or D-HP-Cys-Me. Assuming correct stereochemical or PchE-AE resulted in a time-dependent increase in fluorescence at 413 nm (ex. 310 nm) while no change in fluorescence was observed with only substrate present (**Figure 3-6 A, B**). Cyclization of an acyclic Cys to a thiazoline would increase conjugation through the hydroxyphenyl ring changing the absorbance and fluorescence properties of the molecule. However, separation methods and identification of ions representative of cyclized products by mass spectrometry were irreproducible and often not detected. When fresh substrate stocks were made (still in DMSO), the addition of PchE-FL or PchE-AE did not result in a time-dependent increase in fluorescence (**Figure 3-6C**). Further investigation into the observed phenomenon suggests HP-Cys_L, HP-Cys_D, HP-Cys_L-Me, and HP-Cys_D-Me become dimeric by a disulfide bond while being stored in DMSO. The absorbance profile of freshly made stocks has a peak maximum at 325 nm where greater than 1-week-old DMSO stocks have an absorbance maximum at 300 nm. The addition of 50 μM dithiothreitol to the fresh stock does not alter the absorbance over a 1-hour period. However, the addition of 50 μM dithiothreitol to the greater than 1-week-old stocks causes a shift in absorbance towards the peak maximum of the monomeric substrate at 325 nm (**Figure 3-6 D, E**). The addition of 50 μM dithiothreitol to the greater than 1-week-old stocks also causes an increase in fluorescence at 413 nm (ex. 310 nm) overtime. The enzyme buffer provided a final concentration of 50 μM dithiothreitol to the fluorescent reactions suggesting that the observed spectral changes are the result of dithiothreitol oxidation or substrate reduction, but not substrate cyclization. Enzyme assays with fresh stock in the presence of 5 mM dithiothreitol or 5 mM tris(2-carboxyethyl)phosphine did not result in detectable amounts of cyclized product analogs over a 4-hour period as analyzed by LC-HRMS.

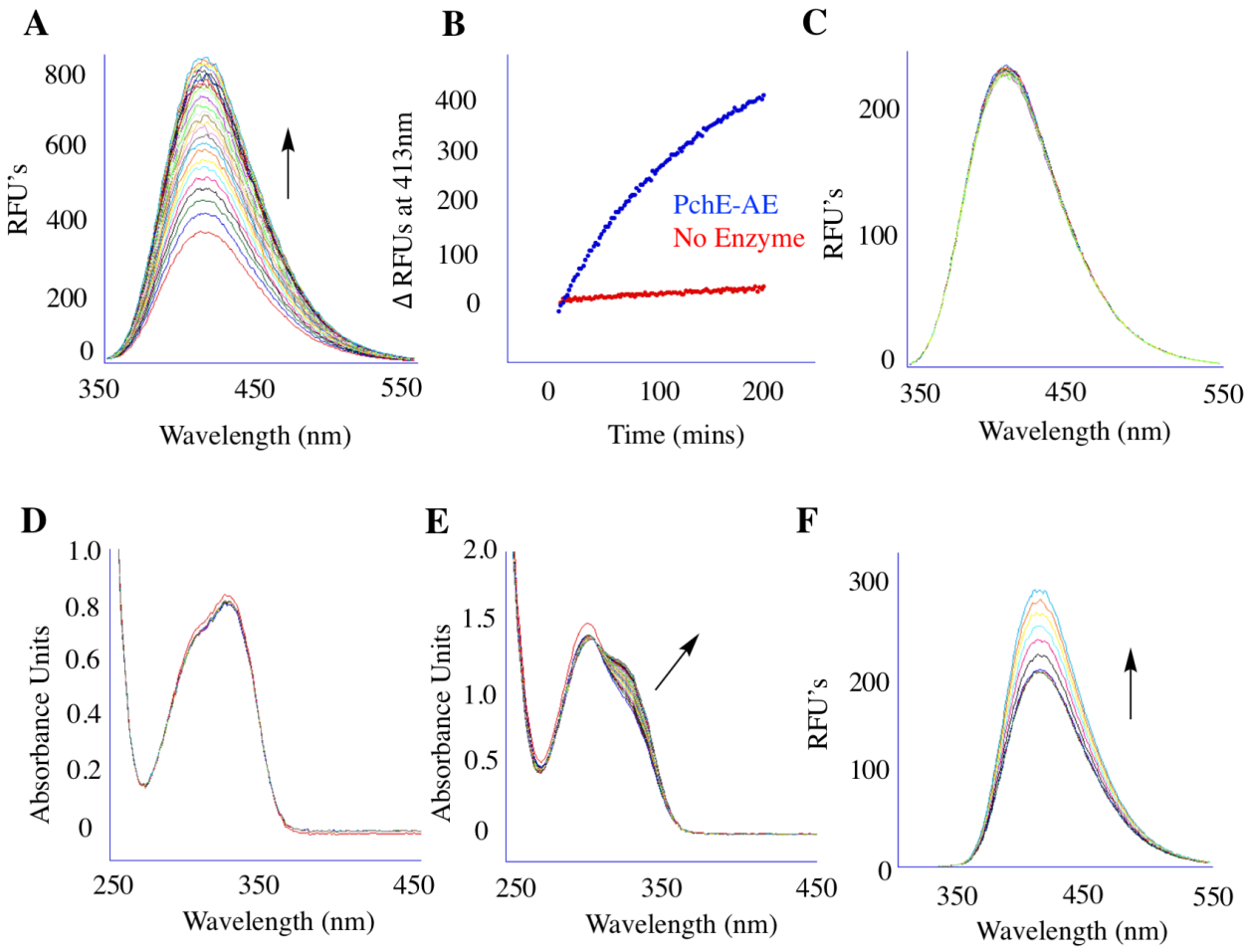


Figure 3-6: Absorbance and fluorescence changes observed with HP-Cys-Me are due to DTT reduction and not PchE catalyzed cyclization. PchE-FL and PchE-AE are stored in buffer with 1 mM DTT eventually determined to be the responsible reagent for changes in absorbance and fluorescence of HP-Cys and HP-Cys-Me from 1-week old DMSO stocks. **A)** A representative time course of HP-Cys_L-Me or HP-Cys_D-Me and PchE-AE or PchE-FL in 50 mM Tris pH 8.0 resulted in increased fluorescence overtime with an excitation wavelength of 310 nm and fluorescence peak maximum of 413 nm when diluted from 1-week old DMSO stock. **B)** The change in relative fluorescence units at 413 nm is plotted over a 240-minute time course of 225 µM HP-Cys_L-Me diluted form old DMS0 stock in the presence of PchE-AE (blue) or absence of PchE-AE (red). **C)** A 20-minute fluorescence time course of freshly diluted HP-Cys_L-Me in Tris pH 8.0 and PchE-AE with excitation at 310 nm. **D)** A 1-hour time course reaction of freshly diluted HP-Cys_L-Me in Tris pH 8.0 and 50 µM DTT. The absorbance was taken every 2 minutes. **E)** A 1-hour time course reaction of HP-Cys_L-Me from a 1-week old DMSO stock in Tris pH 8.0 and 50 µM DTT. The arrow is showing the direction of absorbance change overtime. A spectrum was captured every 2 minutes. **F)** A 20-minute fluorescence time course of HP-Cys_L-Me from a 1-week old DMSO stock in Tris pH 8.0 and 50 µM DTT with excitation at 310 nm.

Circular dichroism epimerase assay.

Previously, Walsh *et al* reported *in vitro* reconstitution assays with PchD and PchE in the presence of ^{14}C labelled L-Cys, ATP, and benzoate, replacing the natural salicylate substrate of PchD.[24] Without the downstream PchF module, PchE derived intermediates accumulate on the thiotemplate of the second carrier domain of PchE. Walsh and colleagues added an external thioesterase or strong base to catalyze hydrolysis of the PchE derived intermediates and identified proposed intermediates based on HPLC alignment of synthetic standards. To assess the epimerase activity of PchE, we used the full length, PchE-FL, and a variant enzyme, PchE-AE, only consisting of the stuffed epimerase and adenylation didomain. PchE-FL and PchE-AE incubated with substrate (HP-Cys_L, HP-Cys_L-Me, HPT_L-CO_2^-, HPT_L-CO_2Et, or HPT_L-CH_2OH) and product analogs (HP-Cys_D, HP-Cys_D-Me, HPT_D-CO_2^-, HPT_D-CO_2Et, or HPT_D-CH_2OH) for up to 8 hours and a circular dichroism spectrum was collected from 240 – 360 nm to observe if there was a change in stereochemistry of the analog. Experiments without enzyme present were used as negative controls. Because synthesized substrate and product analogs are racemic pairs of one another and were synthesized with the starting material in enantiomeric excess, the synthesized analogs with opposing stereochemistry display an equal but opposite circular dichroism spectral footprint that can therefore be traced if stereochemical changes were to occur (**Figure 3-7**). Eight-hour endpoint assays indicated that PchE-FL or PchE-AE were unable to catalyze the racemization of any of the substrate or product analogs that were synthesized (**Figure 3-7**). One-hour time course reactions of HP-Cys_L, HP-Cys_L-Me, HPT_L-CO_2^-, HPT_L-CO_2Et, and HPT_L-CH_2OH were also performed with PchE-FL and slight changes were observed for HPT_L-CO_2Et. Therefore, the time course was extended to 4 hours and spectra were overlaid

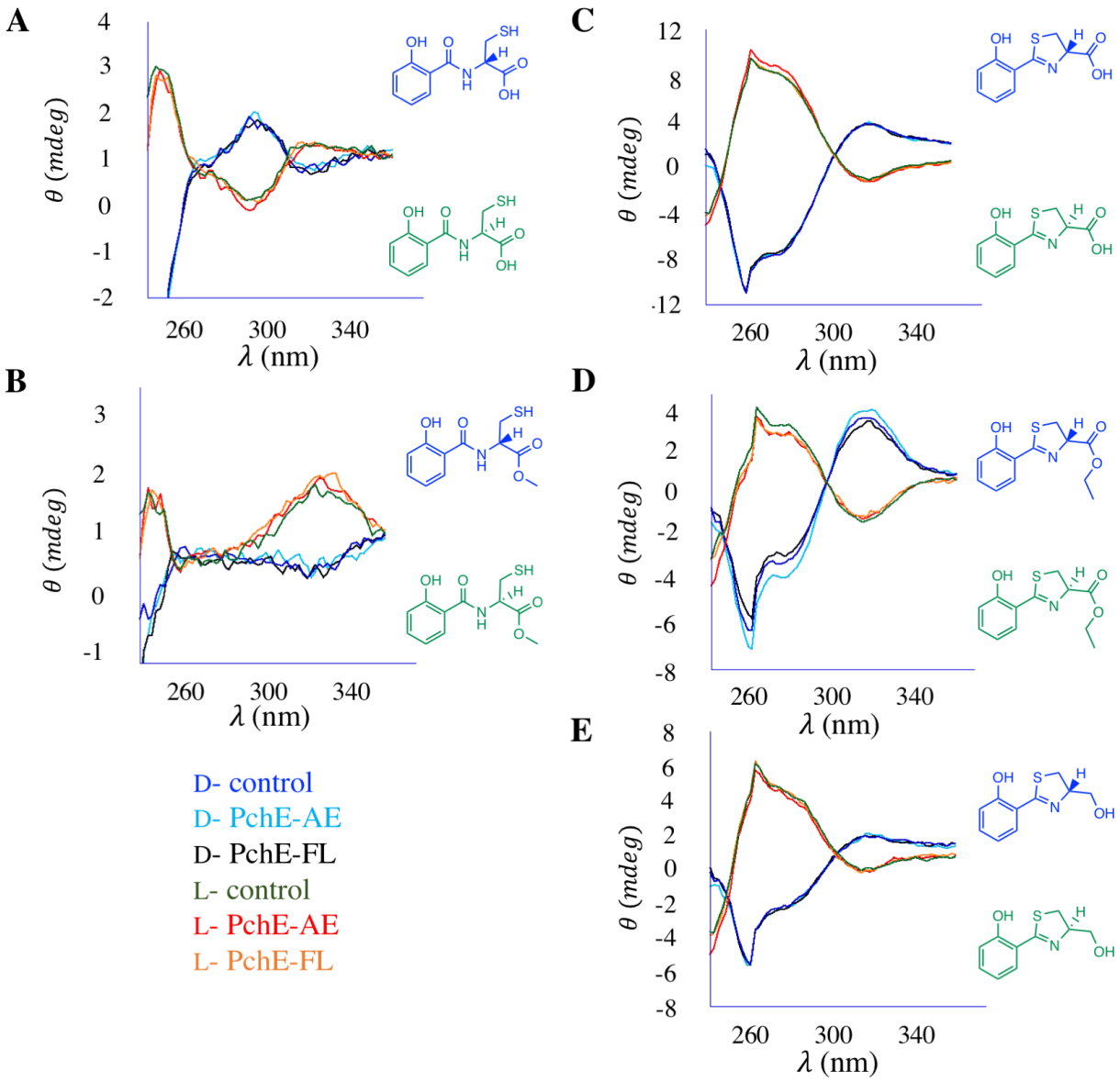


Figure 3-7: Circular dichroism reactions with substrate and product analogs to probe the epimerase activity of PchE. The synthesized racemic pairs of substrate and product analogs exhibit mirrored circular dichroism spectra. The degree of elliptical rotation is plotted over the wavelength range of 240 – 360 nm. The spectra are shown as endpoint data after an 8-hour incubation period at room temperature in 50 mM Tris pH 8.0 and 5 mM TCEP. Both L- and D-analogs were used as substrates for PchE-FL and PchE-AE. Control experiments were performed without the presence of enzyme. Spectra for each experiment are assigned by color: D-control (blue), L-control (green), PchE-FL with D-analog (black), PchE-FL with L-analog (red), PchE-AE with D-analog (light blue), PchE-FL with L-analog (orange). Spectra are combined for each racemic pair: A) HP-Cys; B) HP-Cys-Me; C) HPT-CO_2^-; D) HPT-CO_2Et; E) HPT-CH_2OH.

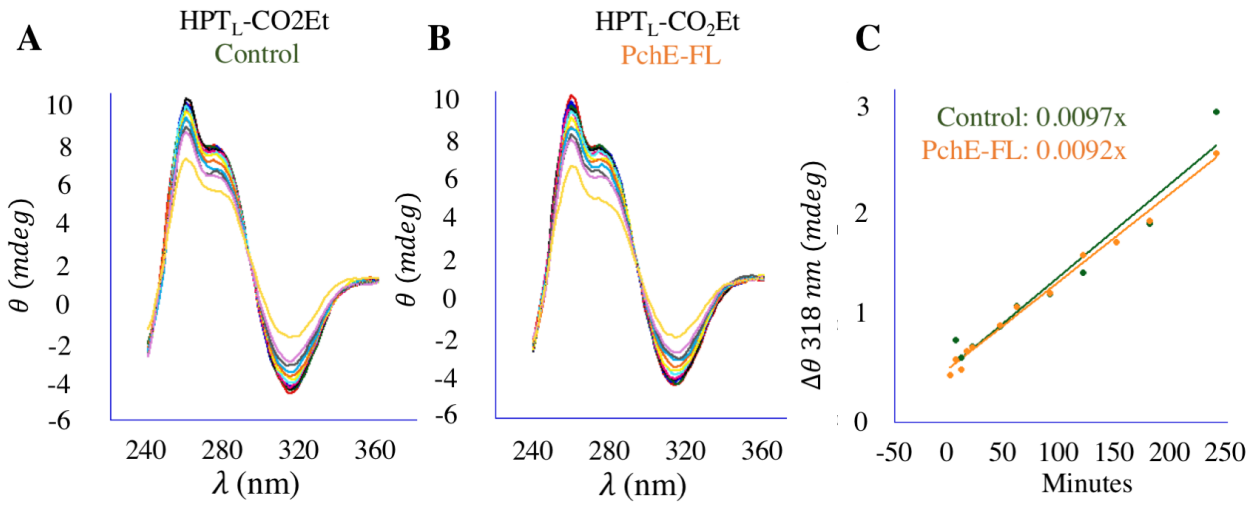


Figure 3-8: Circular dichroism time course with HPT$_L$-CO$_2$Et. Racemization of HPT$_L$-CO$_2$Et observed over 4 hours in 50 mM Tris pH 8.0 and 5 mM TCEP **A)** without enzyme and **B)** with PchE-FL. **C)** A linear regression of the change in ellipticity (mdeg) at 318 nm over the 4-hour period. The control without enzyme is in green and the reaction with PchE-FL is in orange.

(**Figure 3-8**). The negative control without enzyme also racemized over the 4-hour period. The change in ellipticity at a local minimum (318 nm) was plotted for HPT_L-CO_2Et with and without enzyme present. The observed racemization is spontaneous at pH 8.0 and not enzymatically driven.

Chiral chromatography epimerase assay.

Separation protocols for the substrate and product analogues were determined using chiral chromatography, allowing determination of the enantiomeric ratio of each analog synthesized **(Figure 3-9)**. Additionally, the methodology to separate the racemic pair was further used as a more sensitive technique to determine if the stuffed epimerase domain of PchE expresses activity towards the substrate analogs. In agreement with the circular dichroism experiments, enzyme catalyzed racemization was not detected through a 4-hour reaction for any substrate analog. Control reactions with HPT_L-CO_2Et slightly racemized at pH 7.5, similar to what was observed during the circular dichroism experiment at pH 8.0.

Discussion

Pyochelin is a small siderophore produced by *P. aeruginosa* whose chemical uniqueness is mostly attributed to its NRPS cyclization and tailoring domains. The cyclization domains, each in the second (PchE) and third (PchF) module of the NRPS system, first perform condensation of the upstream peptide to the module's activated L-Cys-Ppant, which is tethered to its peptidyl carrier domain. Conventional understanding of NRPS cyclization domains suggest cyclodehydration of the Cys to a thiazoline would occur directly after condensation. However, Walsh and colleagues advocate for a different path in the second NRPS module, which is found

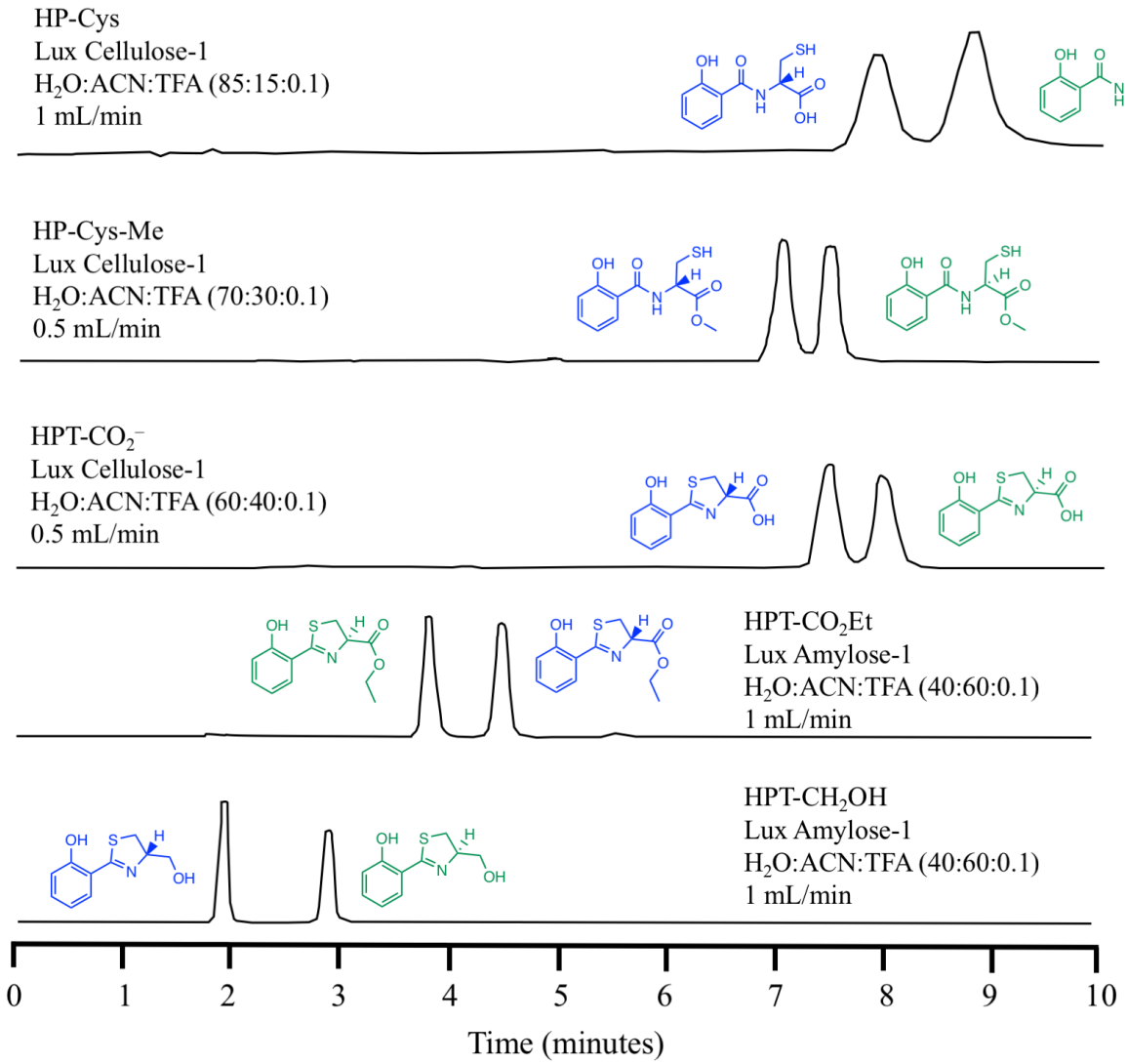


Figure 3-9: UPLC profile of racemic separation of substrate and product analogs. Substrate and product analogs were synthesized and separated via chiral chromatography using UPLC. Substrate analogs have L-stereochemistry and are depicted in green whereas product analogs have D-stereochemistry and are portrayed in blue. Respective analogs are next to the peak where they elute. The column and elution conditions used for chiral separation are listed in the figure. Each trace is representative of the absorbance at 320 nm from injecting a racemic mixture (0.5 mg/mL of each substrate and product analog). Ratios of enantiomeric excess of each substrate and product analog can be found in **Table 3-2**. Enzyme assays with PchE-FL (1 µM) or PchE-AE (1 µM) were performed separately with 1 mM HP-Cys$_L$, 1mM HP-Cys$_L$-Me, 1 mM HPT$_L$-CO_2^-, 1 mM HPT$_L$-CH_2OH, and 300 µM HPT$_L$-CO_2Et for 4 hours. Control experiments without enzyme were assayed simultaneously. No apparent racemization was observed with any substrate analog.

in PchE.[24] They propose after condensation with the upstream salicylate, the newly generated HP-Cys_L-Ppant swings to the stuffed noncanonical epimerase domain where an active site base abstracts a proton form the α-carbon of the Cys, forming a carbanion, followed by random re-protonation, creating HP-Cys_D-Ppant or regenerating the HP-Cys_L-Ppant intermediate.[24] The peptide chain then travels back to the cyclization domain for cyclodehydration forming $HPT_{L/D}$-Ppant intermediates.[24] Walsh *et al* hypothesize the following cyclization domain of PchF selectively condenses with HPT_D-Ppant and not HPT_L-Ppant.[24] Indeed, canonical epimerase domains are often followed by selective downstream condensation domains, only condensing with the isomer exhibiting the correct stereochemistry.[17]

To our knowledge, this would be the only cyclization domain in NRPS chemistry not performing condensation and cyclodehydration chemistries back-to-back. The proposed pathway by Walsh *et al* also seems inefficient, traveling from the cyclization domain to the stuffed epimerase domain, and back to the cyclization domain. Additionally, Walsh *et al* also propose epimerization of the carbanion occurs through random re-protonation by the solvent and is a reversible process. If the incorrect stereochemistry is loaded and cannot condense with the downstream peptide, the tethered peptide can travel back to the stuffed epimerase domain to alter stereochemistry until HPT_D-Ppant is generated.

Interestingly, the noncanonical epimerase domain of PchE shares sequence homology with methyltransferase domains.[24, 25, 30] However, upon sequence alignment, the canonical AdoMet binding motif of PchE's methyltransferase-like domain is mutated.[24] Micacocidin, yersiniabactin, watasemycin, and thiazostatin (**Figure 3-2**) each have methyltransferase tailoring domains catalyzing methyltransfer to the α-carbon of a cyclized thiazoline, a substrate that is similar to what we propose the substrate is of the noncanonical epimerase domain of PchE. The

methylated α-carbon of micacocidin,[48] yersiniabactin, [48] watasemycin[49] and thiostatin[49] are proposed to have D-stereochemistry although their adenylation domain activates L-Cys. The proposed mechanism of the *C*-methyltransferase suggests that cyclodehydration of L-Cys to L-thiazoline occurs before methyl transfer.[50] Next, an active site base is proposed to abstract the α-carbon proton of the thiazoline forming an enol intermediate, and based on stereochemical assignment, nucleophilic attack of the enol to the electrophilic methyl of AdoMet occurs on the opposite side of the substrate, altering the stereochemistry of the carbon center to a D-stereoisomer.

We propose a more efficient route, with condensation and cyclodehydration occurring in succession to form the HPT_L-Ppant. Next, the peptide travels to the noncanonical epimerase domain where it encounters an active site base that abstracts a proton forming an enol intermediate, which subsequently accepts a proton from solvent or an active site acid to alter the stereochemistry to HPT_D-Ppant before condensing with the downstream amino acid in PchF's cyclization domain. Overall, this model presents a substrate for the epimerase domain that has a lower pKa and less domain movements are required, providing an energetically more favorable path.

To identify and assess the chemistries catalyzed by the different domains of PchE we first generated recombinant variants: a full length PchE (PchE-FL) and a variant consisting of only the stuffed epimerase adenylation didomain (PchE-AE, **Figure 3-1C**). A continuous steady-state adenylation assay using hydroxylamine as a nucleophilic surrogate was used to determine the activity of the adenylation domain of both recombinant enzymes.[51, 52] The adenylation domain of each variant was found to be active towards L-Cys but not D-Cys, consistent with previous studies.[25] The catalytic efficiency was only 3-fold less for PchE-FL than that reported by Walsh

et al, although k_{cat} and K_M were nearly a magnitude different.[25] This is likely attributed to the difference in assays where similar differences have previously been documented.[51, 52] In this study, the shortened variant, PchE-AE, had a k_{cat} that was nearly 2.5-fold greater than PchE-FL and a K_M that was more than 50-fold greater, creating a k_{cat}/K_M that was a magnitude less than that of PchE-FL. A similar trend of increased k_{cat} and K_M was found in the previous chapter with a PchF variant missing its cyclization and thioesterase domains (PchF-AMP), although the difference was less drastic, having a k_{cat} and K_M within 2-fold of the full-length enzyme. This suggests the peptidyl carrier domain, as well as the other domains, may play an important role in adenylation efficiency. Importantly, the adenylation domain of PchE-FL and PchE-AE was active suggesting the other domain(s) in each variant are also likely properly folded and potentially active.

Walsh *et al* previously attempted to characterize the biosynthetic steps of PchE by isolating intermediates using an external thioesterase and/or by strong base hydrolysis during reconstitution assays with PchD.[24] While suggesting cyclodehydration occurs after epimerization, the results were inconclusive.[24] Thus, the order and substrate of cyclodehydration and epimerization remain to be determined.[24] To assess the possible substrate and order of catalytic steps of PchE's cyclization and epimerase domain, we synthesized potential substrate and product analogs representative of each domain's proposed chemistries (**Figure 3-3 and 3-5**).

PchE cyclization domain

Cyclization domains are commonly located upstream of a Thr, Ser, or Cys specific adenylation domain and perform two catalytic reactions.[4, 7, 10, 12, 20, 21, 53, 54] The first reaction parallels the condensation domain performing a condensation reaction between the upstream

peptide and the downstream aminoacyl-S-Ppant. The second catalytic reaction is a two-step cyclodehydration. The hydroxyl (Thr, Ser) or thiol (Cys) performs a nucleophilic attack of the previously formed amide forming a hydroxylated azolidine. The amide oxygen is then released by a dehydration reaction forming the azoline heterocycle. The azoline loaded thiotemplate is now ready to move forward to the next step of the natural product bioassembly. It is understood that the cyclodehydration reaction is dependent on the condensation reaction taking place first.[4, 12, 53] Cyclization domains can be distinguished by a canonical DxxxxDxxS sequence, replacing the canonical HHxxxDG sequence found in condensation domains.[12] Importantly, the second His of the condensation domain's canonical sequence, proposed to be the active site base for condensation chemistry, is not conserved in the cyclization domain suggesting chemistry in the cyclization domain is likely different.[4, 12, 21]

Recently, Schmeing *et al* and Drennan *et al* presented structures of cyclization domains in conjunction with mutational analyses.[4, 21] Drennan *et al* published structures of EpoB from epothilone biosynthesis and Schmeing *et al* published structures of the second cyclization domain of bacillamide synthetase from bacillamide E biosynthesis.[4, 21] Each structure is quite similar and the mutational analyses between the two cyclization domains are in agreement with one another. Previous mutation experiments have shown that mutating residues of the conserved DxxxxDxxS motif can lessen or abolish condensation and cyclization activity.[53-55] Interestingly, both cyclization domain structures show the conserved motif as a loop with the Asp residues and Ser pointed away from the active site. The structures of the cyclization domains suggest that mutations to the conserved sequence is likely responsible to orient the substrate or provide structural stability to the domain, but not due to being part of the catalytic mechanism.

Schmeing *et al* presented mutational data distinguishing residues that are separately important for cyclization chemistry and not for condensation chemistry.[4] They identified T1196 and D1226 as essential residues for cyclodehydration but not condensation chemistry and proposed an acid/base catalytic mechanism by these two residues. This matched the analysis presented by Drennan and co-workers who provided evidence that the same spatial Asp (D449 of EpoB) is necessary for catalysis (although they did not distinguish which catalytic step was affected).[21] Schmeing and colleagues also performed a bioinformatic investigation identifying seven core sequences of cyclization domains. Importantly, the bioinformatics also found that nearly 7% of predicted cyclization domains have an alanine mutation to the Thr seen to be essential for cyclodehydration chemistry, bringing their proposed mechanism into question. Schmeing *et al* suggest cyclization domains with Ala replacing Thr use a nearby Glu to assist in the acid/base chemistry instead but provided no evidence to support this claim.

Intriguingly, PchE's cyclization domain is part of the 7% of cyclization domains that have the Ala mutation to the conserved and proposed catalytic Thr, while the downstream cyclization domain of PchF alternatively possesses the conserved Thr. Experiments attempted herein to identify cyclodehydration of substrate analogs by PchE-FL or PchE-AE were unsuccessful. One possibility is that the substrate analogs were not similar enough to the actual substrate which is tethered to the P-domain by a Ppant linkage. Previously, aminoacyl-*N*-acetylcysteamine (aminoacyl-SNACs) were synthesized and used as biosynthetic intermediates to characterize the substrate specificity of condensation domains.[43] While condensation occurred with an aminoacyl-SNAC analog being used as either a downstream acceptor or upstream donor to a natural Ppant substrate, condensation was not observed when both donor and acceptor SNAC-analogs were used together, suggesting the necessity of a Ppant tether with either the

donor or acceptor substrate or a peptidyl carrier domain interaction with the condensation domain.[43] Interestingly the condensation domain was specific to the correct stereochemistry of the aminoacyl-donor.[43] Although we were hopeful that PchE's cyclization domain could catalyze cyclodehydration of a carboxy (HP-Cys_L, HP-Cys_D) or ester (HP-Cys_L-CO_2Me, HP-Cys_D-CO_2Me) analog, no cyclized product was detected. Future experiments will be necessary to further elucidate the catalytic mechanism of the unique cyclization domain of PchE.

PchE noncanonical epimerase domain

Substrate and product analogs of the proposed cyclized or acyclic substrates were synthesized to examine the epimerase activity of PchE (**Figure 3-5**). In these experiments, analogs varied by a terminal hydroxy, carboxy, or ester group replacing the natural thioester linkage of the Ppant tether. Unfortunately, circular dichroism (**Figure 3-7 and 3-8**) and chiral separation (**Figure 3-9**) experiments suggested PchE was not able to alter the stereochemistry of any of the five substrate analogs. While we have observed catalysis of substrate analogs with a carboxy or ester terminus for other tailoring domains in pyochelin's biosynthetic pathway (PchF and PchG)[56], failure to observe catalysis with substrate analogs is not rare for NPRS enzymes.[57-61] Likely, other examples of failed substrate analogs have not been reported. The inability of PchE to turnover these particular analogs could, again, be due to the necessity of the Ppant tether interactions or domain-domain interactions between the peptidyl carrier domain and stuffed tailoring domain.

Another possible explanation is that the stuffed epimerase domain is simply a non-functional methyltransferase domain and does not have any catalytic activity. While the acyclic substrate analogs (HP-Cys_L, HP-Cys_L-CO_2Me) and carboxyl- and hydroxy-cyclized analogs

(HPT_L-CO_2^-, HPT_L-CH_2OH) were relatively stable isomers, the closest related analog to the natural substrate, HPT_L-CO_2Et, racemized overtime at pH 7.5 and pH 8.0 (**Figure 3-8**). The transformation from the acyclic Cys to the thiazoline increases the acidity of the α-carbon proton. Generating potential analogs that are more similar to the thioester linkage of the Ppant tether are also likely to racemize at the α-carbon of the 2,4-substituted thiazoline. In fact, a similar analog, phenylthiazolinyl-SNAC, has only a 2-hour half-life in D_2O before becoming completely racemized.[15] It seems possible that HPT-Ppant is naturally in a racemic state alternating between L- and D-isomers.

While pyochelin isomers with alternating stereocenters have been synthesized (i.e. neopyochelin)[62], no evidence of pyochelin-like natural products retaining opposing 4'*R*, 4"*S* or 4'*S*, 4"*R* have been reported to our knowledge since watasemycin's and thiazostatin's stereochemistry has recently been reassigned (**Figure 3-2**).[49] The 2-carbon of a 2,4 disubstituted thiazolidine ring rapidly alternates between stereocenters and is well documented.[63] Pyochelin is isolated as a diastereomer due to the rapid racemization of the 2" carbon but in the presence of a divalent metal, the 2" carbon shifts and locks into an *R* stereocenter.[63, 64] If, in fact, the stuffed domain of PchE is defunct in catalytic activity, there may be an undetermined enzymatic or chemical driving force responsible for promoting the 4'*R*, 4"*R* stereocenters of pyochelin which may also be the cause for the 4'*R*, 4"*R* or 4'*S*, 4"*S* stereochemistry of 2-hydroxyphenylthiazoline-4-thiazolidine pyochelin-like natural products.

Walsh and colleagues suggested the downstream PchF cyclization domain may discriminate between the L- and D-isoform.[24] In a later study, Walsh *et al* generated a chimeric enzyme, replacing the stuffed domain of PchE with the oxidase domain of MtaD from myxothiazol biosynthesis.[65] The chimeric enzyme was able to oxidize the thiazoline ring to a

thiazole and condense with the downstream L-Cys of PchF, suggesting condensation of PchF's cyclization domain is not stereospecific. However, Walsh and colleagues reported no detection of cyclization of the second added Cys implying a stalled system after condensation.[65] Interestingly, we were able to catalyze PchF methyltransferase activity with a substrate analog containing a thiazole replacing the D-thiazoline, suggesting the stereochemistry of the thiazoline α-carbon was not necessary for PchF methyltransferase activity (Chapter 2).

Recently, the stereochemistry of watasemycin and thiazostatin, two metabolic products from *Streptomyces venezuelae*, were reassigned.[49] Watasemycin and thiazostatin have a very similar biosynthetic gene cluster to pyochelin and are part of the 2-hydroxyphenylthiazoline-4-thiazolidine family of iron-chelating NRPs (**Figure 3-2**).[49] Challis and colleagues determined the PchE homolog in watasemycin biosynthesis has a stuffed methyltransferase-like domain that is non-functional as a methyltransferase or epimerase, retaining the L-isoform configuration of the thiazoline α-carbon from the activated L-Cys. The following module, similar to PchF, also generates a second cyclized thiazoline from an activated L-Cys and is predicted to alter its stereochemistry during methyltransferase activity by the stuffed methyltransferase domain of Sven0517. The proposed mechanism is similar to the *C*-methylation in yersiniabactin biosynthesis: proton abstraction, enol intermediate formation, and nucleophilic attack on the opposite side to the electrophilic methyl of AdoMet. Based on isolated intermediates, Challis *et al* observed methylation of the α-carbon occurred before reduction or any other methylation during watasemycin biosynthesis.[49] Overall the stereocenter of each azole-like heterocycle in watasemycin is an *S*-stereocenter. The first activated L-Cys retains its stereochemistry due to the defunct stuffed methyltransferase-like domain while the stereochemistry of the second activated L-Cys is altered to its D-isoform during methyltransferase activity.

Intriguingly, *Pseudomonas fluorescens*, a non-pathogenic pseudomonad, generates enantiopyochelin, the optical antipode of pyochelin (**Figure 3-2**).[3] The noncanonical stuffed epimerase domain found in PchE from *P. aeruginosa* is missing from the PchE homolog of *P. fluorescens*, likely accounting for the retained L-isoform of the first thiazoline in enantiopyochelin. However, the origin of the altered stereochemistry of the second heterocycle by the PchF homolog to the D-isoform remains unknown. Youard *et al* identified a few possibilities for the change in stereochemistry.[3, 66] The PchF homolog could activate D-Cys instead of L-Cys in its adenylation domain. Or it may be possible that the stuffed methyltransferase domain in the PchF homolog of *P. fluorescens* can catalyze methyl transfer to the thiazolidine *N* and epimerization of the neighboring α-carbon. In *P. fluorescens*, no sequence homology of PchG is present, although PchK, a predicted putative thiazolinyl reductase, appears to replace the *pchg* gene in the operon. Based on sequence homology, this is the only major difference between the two operons that encode pyochelin and enantiopyochelin. Upon a BLAST alignment, it is interesting to find PchK shares 28% sequence homology over 85 amino acids with the stuffed E-domain of PchE from pyochelin biosynthesis.[67]

Future purification and characterization of the NRPS enzymes involved in enantiopyochelin biosynthesis is necessary to identify how the altered stereochemistry is generated in comparison to pyochelin. However, based on our work and the stereochemical assignments of other 2-hydroxyphenylthiazoline-4-thiazolidine natural products, we propose noncanonical stuffed epimerase domains are not catalytically active and are likely an evolutionary artifact. It may be necessary to perform experiments in the presence and absence of metal to determine if stereochemistry change is catalyzed by metal binding. Future studies are necessary to further examine this hypothesis.

Conclusion

In this study we examined PchE, the second module of pyochelin biosynthesis, and began a detailed characterization of the adenylation, cyclization, and epimerase domains. The PchE module's adenylation domain is specific towards activating L-Cys. Potential substrate and product analogs were synthesized to gain further insight into the cyclization and epimerization chemistry performed by PchE. These chemistries were not observed to be catalyzed by PchE during *in vitro* assays. While this may be due to the substrate analog being too dissimilar to the natural Ppant tethered substrate, recent stereochemical assignments of other 2-hydroxyphenylthiazoline-4-thiazolidine natural products bring into question whether the stuffed noncanonical epimerase domain of PchE is a functional epimerase domain, or simply a non-functional methyltransferase-like domain. Future experiments aimed to elucidate the differences between the biosynthesis of pyochelin and enantiopyochelin are necessary to better understand the unique stereochemical alteration of these siderophores.

References

[1] Veatch, W. (1976) The structure of the gramicidin A transmembrane channel, *J Supramol Struct 5*, 431(435)-451(455).

[2] Kawai, Y., Ishii, Y., Arakawa, K., Uemura, K., Saitoh, B., Nishimura, J., Kitazawa, H., Yamazaki, Y., Tateno, Y., Itoh, T., and Saito, T. (2004) Structural and functional differences in two cyclic bacteriocins with the same sequences produced by lactobacilli, *Appl Environ Microbiol 70*, 2906-2911.

[3] Youard, Z. A., Mislin, G. L., Majcherczyk, P. A., Schalk, I. J., and Reimmann, C. (2007) Pseudomonas fluorescens CHA0 produces enantio-pyochelin, the optical antipode of the Pseudomonas aeruginosa siderophore pyochelin, *J Biol Chem 282*, 35546-35553.

[4] Bloudoff, K., Fage, C. D., Marahiel, M. A., and Schmeing, T. M. (2017) Structural and mutational analysis of the nonribosomal peptide synthetase heterocyclization domain provides insight into catalysis, *Proc Natl Acad Sci U S A 114*, 95-100.

[5] Bessalle, R., Kapitkovsky, A., Gorea, A., Shalit, I., and Fridkin, M. (1990) All-D-magainin: chirality, antimicrobial activity and proteolytic resistance, *FEBS Lett 274*, 151-155.

[6] Cox, C. D., Rinehart, K.L., Moore, M.L. Carter Cook Jr., J. (1981) Pyochelin: Novel structure of an iron-chelating growth promoter for Pseudomonas aeruginosa, *Proc Natl Acad Sci U S A 78*, 4256-4260

.

[7] Gehring, A. M., DeMoll, E., Fetherston, J. D., Mori, I., Mayhew, G. F., Blattner, F. R., Walsh, C. T., and Perry, R. D. (1998) Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by Yersinia pestis, *Chem Biol 5*, 573-586.

[8] Griffiths, G. L., Sigel, S. P., Payne, S. M., and Neilands, J. B. (1984) Vibriobactin, a siderophore from Vibrio cholerae, *J Biol Chem 259*, 383-385.

[9] De Voss, J. J., Rutter, K., Schroeder, B. G., Su, H., Zhu, Y., and Barry, C. E., 3rd. (2000) The salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages, *Proc Natl Acad Sci U S A 97*, 1252-1257.

[10] Duerfahrt, T., Eppelmann, K., Muller, R., and Marahiel, M. A. (2004) Rational design of a bimodular model system for the investigation of heterocyclization in nonribosomal peptide biosynthesis, *Chem Biol 11*, 261-271.

[11] Gerth, K., Irschik, H., Reichenbach, H., and Trowitzsch, W. (1980) Myxothiazol, an antibiotic from Myxococcus fulvus (myxobacterales). I. Cultivation, isolation, physico-chemical and biological properties, *J Antibiot (Tokyo) 33*, 1474-1479.

[12] Konz, D., Klens, A., Schorgendorfer, K., and Marahiel, M. A. (1997) The bacitracin biosynthesis operon of Bacillus licheniformis ATCC 10716: molecular characterization of three multi-modular peptide synthetases, *Chem Biol 4*, 927-937.

[13] Tabata, N., Tomoda, H., Zhang, H., Uchida, R., and Omura, S. (1999) Zelkovamycin, a new cyclic peptide antibiotic from Streptomyces sp. K96-0670. II. Structure elucidation, *J Antibiot (Tokyo) 52*, 34-39.

[14] Shen, B., Du, L., Sanchez, C., Edwards, D. J., Chen, M., and Murrell, J. M. (2002) Cloning and characterization of the bleomycin biosynthetic gene cluster from Streptomyces verticillus ATCC15003, *J Nat Prod 65*, 422-431.

[15] Schneider, T. L., Shen, B., and Walsh, C. T. (2003) Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation, *Biochemistry 42*, 9722-9730.

[16] Vollbrecht, L., Steinmetz, H., Hofle, G., Oberer, L., Rihs, G., Bovermann, G., and von Matt, P. (2002) Argyrins, immunosuppressive cyclic peptides from myxobacteria. II. Structure elucidation and stereochemistry, *J Antibiot (Tokyo) 55*, 715-721.

[17] Luo, L., Kohli, R. M., Onishi, M., Linne, U., Marahiel, M. A., and Walsh, C. T. (2002) Timing of epimerization and condensation reactions in nonribosomal peptide assembly lines: kinetic analysis of phenylalanine activating elongation modules of tyrocidine synthetase B, *Biochemistry 41*, 9184-9196.

[18] Mori, S., Garzan, A., Tsodikov, O. V., and Garneau-Tsodikova, S. (2017) Deciphering Nature's Intricate Way of N,S-Dimethylating l-Cysteine: Sequential Action of Two Bifunctional Adenylation Domains, *Biochemistry 56*, 6087-6097.

[19] Quimby, B. B., Alano, A., Almashanu, S., DeSandro, A. M., Cowan, T. M., and Fridovich-Keil, J. L. (1997) Characterization of two mutations associated with epimerase-deficiency galactosemia, by use of a yeast expression system for human UDP-galactose-4-epimerase, *Am J Hum Genet 61*, 590-598.

[20] Keating, T. A., Marshall, C. G., Walsh, C. T., and Keating, A. E. (2002) The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains, *Nat Struct Biol 9*, 522-526.

[21] Dowling, D. P., Kung, Y., Croft, A. K., Taghizadeh, K., Kelly, W. L., Walsh, C. T., and Drennan, C. L. (2016) Structural elements of an NRPS cyclization domain and its intermodule docking domain, *Proc Natl Acad Sci U S A 113*, 12432-12437.

[22] Chen, W. H., Li, K., Guntaka, N. S., and Bruner, S. D. (2016) Interdomain and Intermodule Organization in Epimerization Domain Containing Nonribosomal Peptide Synthetases, *ACS Chem Biol 11*, 2293-2303.

[23] Quadri, L. E., Keating, T. A., Patel, H. M., and Walsh, C. T. (1999) Assembly of the Pseudomonas aeruginosa nonribosomal peptide siderophore pyochelin: In vitro reconstitution of aryl-4, 2-bisthiazoline synthetase activity from PchD, PchE, and PchF, *Biochemistry 38*, 14941-14954.

[24] Patel, H. M., Tao, J., and Walsh, C. T. (2003) Epimerization of an L-cysteinyl to a D-cysteinyl residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from Pseudomonas aeruginosa, *Biochemistry 42*, 10514-10527.

[25] Patel, H. M., and Walsh, C. T. (2001) In vitro reconstitution of the Pseudomonas aeruginosa nonribosomal peptide synthesis of pyochelin: characterization of backbone tailoring thiazoline reductase and N-methyltransferase activities, *Biochemistry 40*, 9023-9031.

[26] Drake, E. J., Miller, B. R., Shi, C., Tarrasch, J. T., Sundlov, J. A., Allen, C. L., Skiniotis, G., Aldrich, C. C., and Gulick, A. M. (2016) Structures of two distinct conformations of holo-non-ribosomal peptide synthetases, *Nature 529*, 235-238.

[27] Leslie, A. G. (1990) Refined crystal structure of type III chloramphenicol acetyltransferase at 1.75 A resolution, *J Mol Biol 213*, 167-186.

[28] Balibar, C. J., Vaillancourt, F. H., and Walsh, C. T. (2005) Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains, *Chem Biol 12*, 1189-1200.

[29] Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of Yersinia pestis, *Chem Biol 9*, 333-344.

[30] Labby, K. J., Watsula, S. G., and Garneau-Tsodikova, S. (2015) Interrupted adenylation domains: unique bifunctional enzymes involved in nonribosomal peptide biosynthesis, *Nat Prod Rep 32*, 641-653.

[31] Drake, E. J., Cao, J., Qu, J., Shah, M. B., Straubinger, R. M., and Gulick, A. M. (2007) The 1.8 A crystal structure of PA2412, an MbtH-like protein from the pyoverdine cluster of Pseudomonas aeruginosa, *J Biol Chem 282*, 20425-20434.

[32] Boll, B., Taubitz, T., and Heide, L. (2011) Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis, *J Biol Chem 286*, 36281-36290.

[33] Felnagle, E. A., Barkei, J. J., Park, H., Podevels, A. M., McMahon, M. D., Drott, D. W., and Thomas, M. G. (2010) MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases, *Biochemistry 49*, 8815-8817.

[34] Ochsner, U. A., Wilderman, P. J., Vasil, A. I., and Vasil, M. L. (2002) GeneChip expression analysis of the iron starvation response in Pseudomonas aeruginosa: identification of novel pyoverdine biosynthesis genes, *Mol Microbiol 45*, 1277-1287.

[35] Zhang, W., Heemstra, J. R., Jr., Walsh, C. T., and Imker, H. J. (2010) Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins, *Biochemistry 49*, 9946-9947.

[36] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, *John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press*, 571-607

.

[37] Bergeron, R. J., Wiegand, J., Dionis, J. B., Egli-Karmakka, M., Frei, J., Huxley-Tencer, A., and Peter, H. H. (1991) Evaluation of desferrithiocin and its synthetic analogues as orally effective iron chelators, *J Med Chem 34*, 2072-2078.

[38] Webb, M. R. (1992) A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems, *Proc Natl Acad Sci U S A 89*, 4884-4887.

[39] Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data, *Anal Biochem 387*, 30-41.

[40] Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data, *Anal Biochem 387*, 20-29.

[41] Heemstra, J. R., Jr., Walsh, C. T., and Sattely, E. S. (2009) Enzymatic tailoring of ornithine in the biosynthesis of the Rhizobium cyclic trihydroxamate siderophore vicibactin, *J Am Chem Soc 131*, 15317-15329.

[42] Miller, B. R., Drake, E. J., Shi, C., Aldrich, C. C., and Gulick, A. M. (2016) Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture, *J Biol Chem 291*, 22559-22571.

[43] Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases, *Chem Biol 7*, 765-772.

[44] Tseng, C. C., Bruner, S. D., Kohli, R. M., Marahiel, M. A., Walsh, C. T., and Sieber, S. A. (2002) Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase, *Biochemistry 41*, 13350-13359.

[45] Holzbaur, I. E., Harris, R. C., Bycroft, M., Cortes, J., Bisang, C., Staunton, J., Rudd, B. A., and Leadlay, P. F. (1999) Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase, *Chem Biol 6*, 189-195.

[46] Gokhale, R. S., Hunziker, D., Cane, D. E., and Khosla, C. (1999) Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase, *Chem Biol 6*, 117-125.

[47] Bergeron, R. J., Wollenweber, M., and Wiegand, J. (1994) An investigation of desferrithiocin metabolism, *J Med Chem 37*, 2889-2895.

[48] Miller, M. C., Parkin, S., Fetherston, J. D., Perry, R. D., and Demoll, E. (2006) Crystal structure of ferric-yersiniabactin, a virulence factor of Yersinia pestis, *J Inorg Biochem 100*, 1495-1500.

[49] Inahashi, Y., Zhou, S., Bibb, M. J., Song, L., Al-Bassam, M. M., Bibb, M. J., and Challis, G. L. (2017) Watasemycin biosynthesis in Streptomyces venezuelae: thiazoline C-methylation by a type B radical-SAM methylase homologue, *Chem Sci 8*, 2823-2831.

[50] Miller, D. A., Walsh, C. T., and Luo, L. (2001) C-methyltransferase and cyclization domain activity at the intraprotein PK/NRP switch point of yersiniabactin synthetase, *J Am Chem Soc 123*, 8434-8435.

[51] Duckworth, B. P., Wilson, D. J., and Aldrich, C. C. (2016) Measurement of Nonribosomal Peptide Synthetase Adenylation Domain Activity Using a Continuous Hydroxylamine Release Assay, *Methods Mol Biol 1401*, 53-61.

[52] Wilson, D. J., and Aldrich, C. C. (2010) A continuous kinetic assay for adenylation enzyme activity and inhibition, *Anal Biochem 404*, 56-63.

[53] Marshall, C. G., Hillson, N. J., and Walsh, C. T. (2002) Catalytic mapping of the vibriobactin biosynthetic enzyme VibF, *Biochemistry 41*, 244-250.

[54] Kelly, W. L., Hillson, N. J., and Walsh, C. T. (2005) Excision of the epothilone synthetase B cyclization domain and demonstration of in trans condensation/cyclodehydration activity, *Biochemistry 44*, 13385-13393.

[55] Keating, T. A., Miller, D. A., and Walsh, C. T. (2000) Expression, purification, and characterization of HMWP2, a 229 kDa, six domain protein subunit of Yersiniabactin synthetase, *Biochemistry 39*, 4729-4739.

[56] Meneely, K. M., Ronnebaum, T. A., Riley, A. P., Prisinzano, T. E., and Lamb, A. L. (2016) Holo Structure and Steady State Kinetics of the Thiazolinyl Imine Reductases for Siderophore Biosynthesis, *Biochemistry 55*, 5423-5433.

[57] Piasecki, S. K., Taylor, C. A., Detelich, J. F., Liu, J., Zheng, J., Komsoukaniants, A., Siegel, D. R., and Keatinge-Clay, A. T. (2011) Employing modular polyketide synthase ketoreductases as biocatalysts in the preparative chemoenzymatic syntheses of diketide chiral building blocks, *Chem Biol 18*, 1331-1340.

[58] Haynes, S. W., Ames, B. D., Gao, X., Tang, Y., and Walsh, C. T. (2011) Unraveling terminal C-domain-mediated condensation in fungal biosynthesis of imidazoindolone metabolites, *Biochemistry 50*, 5668-5679.

[59] Franke, J., and Hertweck, C. (2016) Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges, *Cell Chem Biol 23*, 1179-1192.

[60] Foulke-Abel, J., and Townsend, C. A. (2012) Demonstration of starter unit interprotein transfer from a fatty acid synthase to a multidomain, nonreducing polyketide synthase, *Chembiochem 13*, 1880-1884.

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

[62] Ankenbauer, R. G., and Cox, C. D. (1988) Isolation and characterization of Pseudomonas aeruginosa mutants requiring salicylic acid for pyochelin biosynthesis, *J Bacteriol 170*, 5364-5367.

[63] Schlegel, K., Taraz, K., and Budzikiewicz, H. (2004) The stereoisomers of pyochelin, a siderophore of Pseudomonas aeruginosa, *Biometals 17*, 409-414.

[64] Schlegel, K., Lex, J., Taraz, K., and Budzikiewicz, H. (2006) The X-ray structure of the pyochelin Fe3+ complex, *Z Naturforsch C 61*, 263-266.

[65] Schneider, T. L., and Walsh, C. T. (2004) Portability of oxidase domains in nonribosomal peptide synthetase modules, *Biochemistry 43*, 15946-15955.

[66] Youard, Z. A., Wenner, N., and Reimmann, C. (2011) Iron acquisition with the natural siderophore enantiomers pyochelin and enantio-pyochelin in Pseudomonas species, *Biometals 24*, 513-522.

[67] Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res 25*, 3389-3402.

Chapter 4

Substrate analog and inhibitor synthesis of stand-alone thiazolinyl reductase and salicylate adenylation domains, respectively, in NRPS biosynthesis

*Portions of this chapter were published in Meneely, K. M., Ronnebaum, T. A., Riley, A. P., Prisinzano, T. E., and Lamb, A. L. (2016) Holo Structure and Steady State Kinetics of the Thiazolinyl Imine Reductases for Siderophore Biosynthesis, *Biochemistry 55*, 5423-5433.

Introduction

The human body has evolved over time to sequester and store essential elements from bacterial intruders making it difficult for bacteria to grow and survive.[1, 2] However, bacterial pathogens have also evolved possessing creative mechanisms to scavenge essential elements needed for growth, virulence, and pathogenicity. One of these intricate systems is the production of secondary metabolites called siderophores, or iron scavenging molecules, that are generated by means of nonribosomal peptide synthetases (NRPSs) in times of iron deficiency.[3-5] *Pseudomonas aeruginosa*, a multidrug-resistant *ESKAPE* pathogen, *Mycobacterium tuberculosis*, the etiological agent of tuberculosis, and *Yersinia pestis*, the causative agent of plague, hold serious global threats to human health and are bacterial pathogens that produce and utilize salicylate-capped siderophores for iron acquisition in human hosts (**Figure 4-1**).[6-8] Many reports have suggested that salicylate-capped siderophores from these systems are essential for bacterial growth and virulence in iron limiting environments, making the siderophore biosynthetic pathway a potential target for creation of novel antibacterials.[9-16]

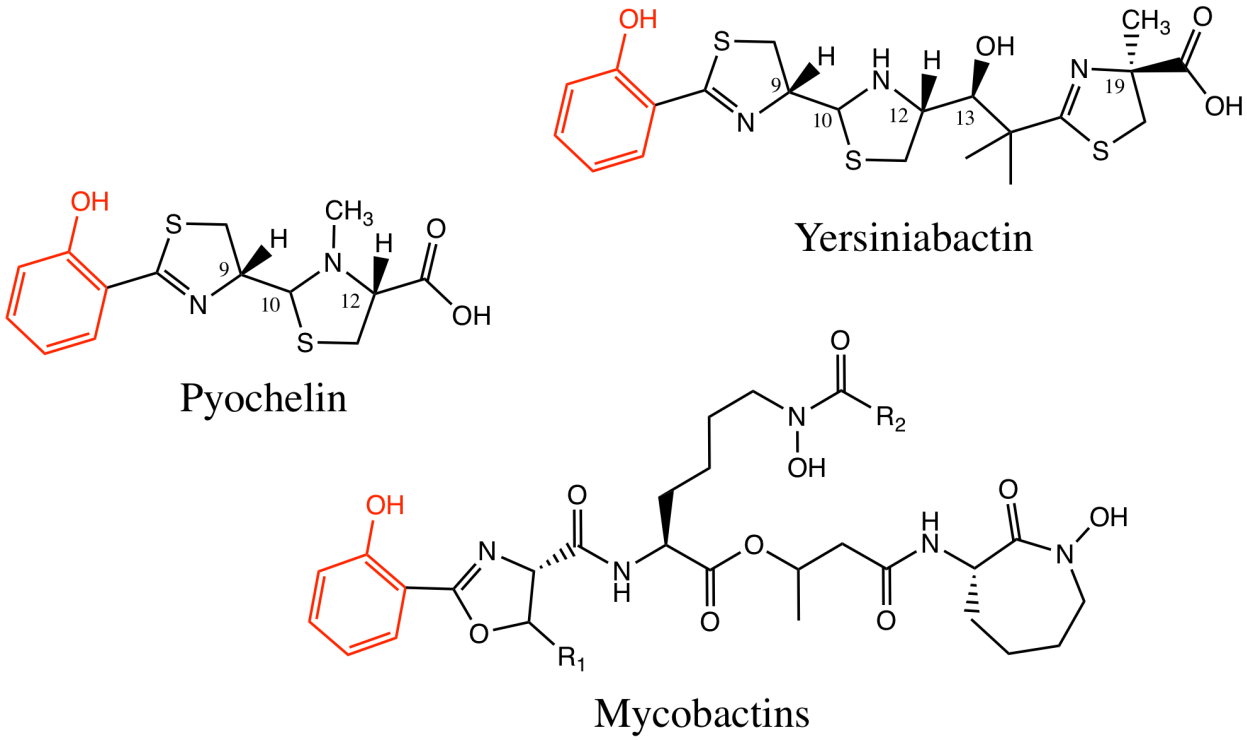


Figure 4-1: Salicylate capped siderophores. Pyochelin, yersiniabactin, and mycobactins are salicylate capped siderophores. The biosynthesis of these siderophores is initiated by a stand-alone adenylation domain predicted to be structurally homologous based on sequence identity. Importantly, these stand-alone adenylation domains are share conserved active site residues. The numbers in the yersiniabactin and pyochelin compounds are indicative of carbons used in the text. In the mycobactin structure, R_1 is sometimes methylated and R_2 is lipophilic chains that differ in length.

NRPS assembly lines are not found in humans them attractive drug targets. The beginning ideal steps in drug design targeting enzymes is to understand 1) the detailed mechanism by which the enzyme performs its catalysis and 2) understand the three-dimensional space occupied by the enzyme in which you are trying to inhibit.[17, 18] Thoughtful experiments detailing the structural and mechanistic details of an enzyme allows for rational drug design. NRPS biosynthesis has been finely detailed in the previous chapters and a great deal highlights the difficulty in deciphering the mechanistic details and structural identity of NRPSs during peptide assembly. Briefly, peptide bond formation and elongation occur by transfer between 4'phosphantetheinyl (Ppant) thiotemplates covalently attached to carrier domains of NRPS modules. This creates a complex substrate making it difficult to perform kinetic experiments in attempt to decipher intermediate mechanisms. Additionally, the large and flexible modular structures of NRPSs make it difficult to structurally characterize by traditional methods. Because of these difficulties, smaller stand-alone domains, such as stand-alone adenylation and stand-alone tailoring domains, from NRPS biosynthetic pathways become more manageable drug targets because of their comparatively smaller size allowing for easier structural determination.

Pyochelin (from *P. aeruginosa*), yersiniabactin (from *Y. pestis*), and mycobactins (from *M. tuberculosis*), are all salicylate capped siderophores (**Figure 4-1**). The biosynthesis of each of these siderophores is initiated by structurally homologous stand-alone salicylate adenylation domains (PchD from pyochelin biosynthesis, YbtE from yersiniabactin biosynthesis, and MbtA from mycobactin biosynthesis). These enzymes catalyze the formation of salicyl-AMP with release of pyrophosphate from salicylic acid and ATP (**Figure 4-2**). The next step in natural product biosynthesis is nucleophilic attack of the Ppant tether from the downstream carrier

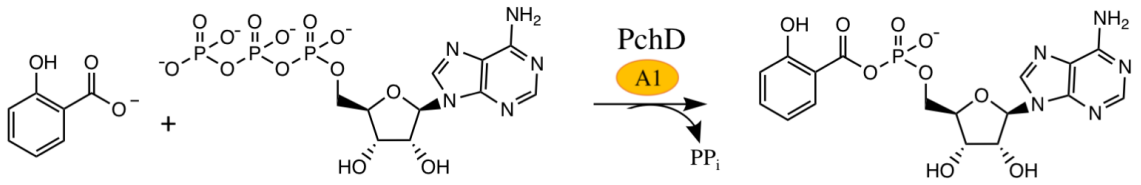


Figure 4-2: PchD catalyzes the formation of salicyl-AMP. PchD is a stand-alone adenylation domain that catalyzes the adenylation of salicylic acid with ATP forming salicyl-AMP and releasing pyrophosphate. PchD initiates the NRPS assembly-line in the biosynthesis of pyochelin.

domain, binding the amino acid and releasing AMP. Over a decade ago, Ferreras et al presented salicyl-AMS, a tight-binding and non-hydrolyzable substrate mimic that inhibits salicylate adenylation domains.[19] Salicyl-AMS exhibits nanomolar IC_{50} and K_i values towards PchD, MbtA, and YbtE, displays growth inhibition of *M. tuberculosis* and *Y. pestis* culture in micromolar concentrations, and has shown decreased rates of growth of *M. tuberculosis* in murine models.[19, 20] Over the years, several alterations to salicyl-AMS have been performed in hopes to increase efficacy, half-life, and bioavailability of the compound.[21-28] However, rigorous modifications and experimentation have not resulted in more efficacious inhibitors. Recently a crystal structure of apo-MbtA was published, the first published salicylate adenylation domain.[29] Unfortunately, we were scooped as we have synthesized salicyl-AMS and novel salicyl-AMS derivatives which have provided holo-structures of PchD (not published), the first holo-structures of salicylate adenylation domains. This breakthrough provides a platform for rational drug design of new salicylate adenylation inhibitors.

The biosynthesis of pyochelin (PchG) and yersiniabactin (YbtU of *Y. Pestis*, Irp3 of *Y. enterocoletica*) also employ a stand-alone tailoring domain acting *in trans* with the NRPS system. PchG and Irp3 are proposed to NADPH-dependent thiazolinyl reductases.[30, 31] Walsh *et al* provided evidence through reconstitution assays that PchG reduces the second L-thiazoline ring attached to the Ppant tether of PchF during pyochelin biosynthesis (**Figure 4-3**).[31] Our work, presented in Chapter 2, agrees with Walsh's hypothesis that reduction of the substrate is necessary for *N*-methyltransfer, subsequent hydrolysis, and release of mature pyochelin.[31, 32] While there has been evidence supporting the intermediate substrate and order of PchG's role in pyochelin biosynthesis, the role of Irp3 in yersiniabactin biosynthesis remains undefined.[30]

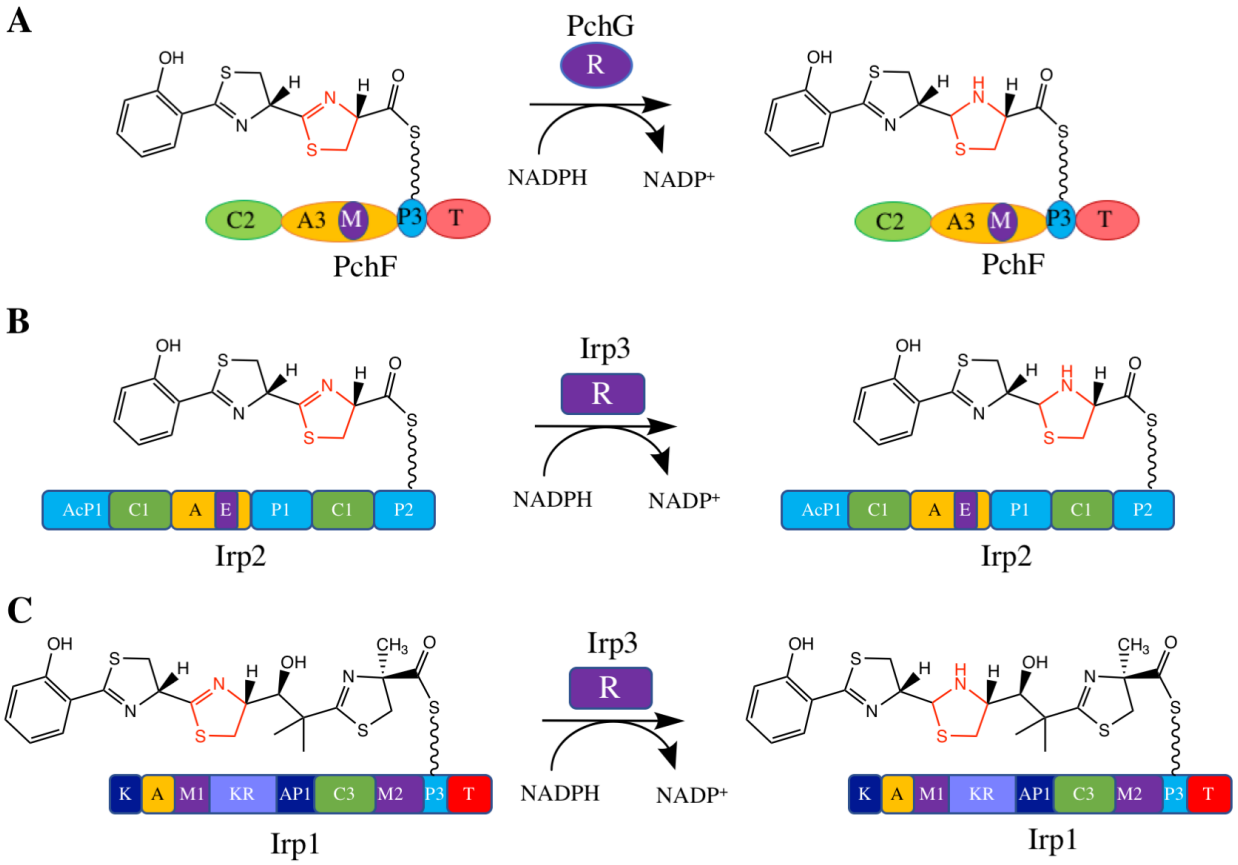


Figure 4-3: Proposed roles of PchG in pyochelin biosynthesis and Irp3 in yersiniabactin biosynthesis. PchG and Irp3 are NADPH dependent thiazolinyl reductases and structural homologs based on sequence identity. A) Proposed role of PchG in pyochelin biosynthesis. PchG and Irp3 share 26% structural identity and may have the same substrate. B and C are possible substrate candidates for Irp3 reduction. B) If PchG and Irp3 share the same substrate, the growing peptide chain would be tethered to P2 of Irp2. C) Irp3 may also be the last step of yersiniabactin biosynthesis before release of the mature peptide.

Mutations to the YbtU gene inhibits siderophore production, downregulates yersiniabactin biosynthetic genes, and inhibits growth of yersiniabactin in iron-deficient media.[33] During reconstitution assays, Walsh *et al* observed a mass representative of an oxidized mature yersiniabactin in the absence of YbtU suggesting two possibilities: 1) reduction by YbtU is the last step of yersiniabactin biosynthesis, or 2) reduction occurs elsewhere in the assembly line but is not required for subsequent steps.[31] Irp3 share a 26% sequence identity with PchG.[34] PchG and Irp3 may share a similar substrate. However, the order in which Irp3 performs reduction could change which carrier domain the growing peptide is attached. If the substrate is shared with PchG, the growing peptide is linked to Irp2 (high-molecular weight protein 2 of *Y. pestis*), but if reduction occurs later in the biosynthesis, the peptide would be tethered to Irp1 (high-molecular weight protein 1 of *Y. pestis*) (**Figure 4-3**). Previously, our lab has reported apo- and $NADP^+$-bound structures of Irp3 in an open confirmation.[35] Irp3 is a homodimer and structurally homologous to sugar oxidoreductases and biliverdin reductase.[35] While these structures provide a significant step towards better understanding the role of Irp3 in yersiniabactin biosynthesis, the order of interaction in the biosynthetic pathway and thus, the substrate remain elusive.

This chapter will focus on the structural biology and mechanistic enzymology of the stand-alone salicylate adenylation domain, PchD, from the biosynthesis of pyochelin from *P. aeruginosa* and stand-alone thiazolinyl reductase domains involved in the biosynthesis of pyochelin from *P. aeruginosa*, PchG, and yersiniabactin from *Y. pestis*, Irp3. My contribution to this project was the synthesis of substrate analogs and potential inhibitors. Drs, Kathleen M. Meenely and Catherine Shelton used these compounds to perform steady-state kinetic assays and co-crystallization experiments. Substrate analog synthesis aided in determination of the first closed-holo-structure of a thiazolinyl reductase, Irp3, at a high resolution of 1.28 Å, and provided

steady-state kinetic parameters of NADPH-dependent PchG. Synthesis of potential salicylate adenylation inhibitors were co-crystallized with PchD and an adenylation assay was performed providing steady-state kinetic parameters of salicylate for PchD.

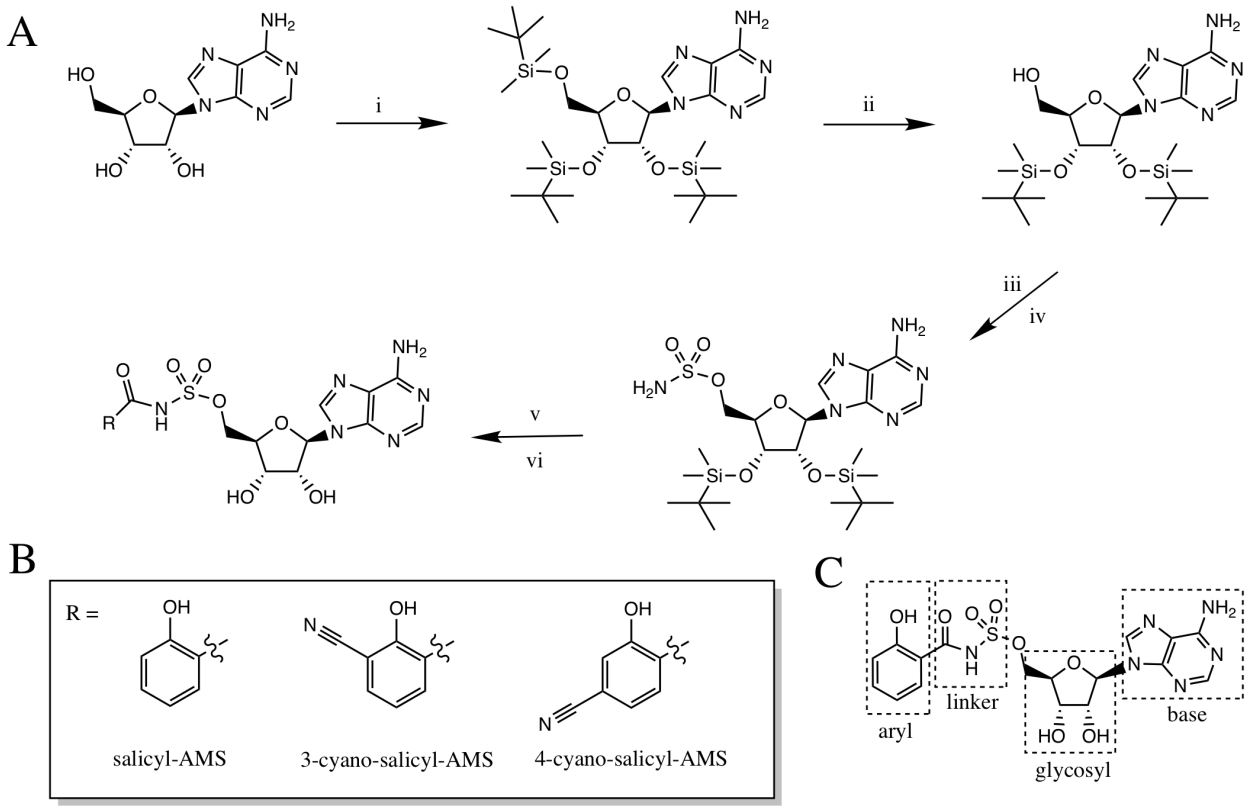
Materials and Methods

Preparation of Substrate Analogue and Inhibitors.

General experimental procedures. All chemical reagents were purchased from commercial suppliers and used without further purification. Flash column chromatography was performed on silica gel (4-63 mm) from Sorbent Technologies. Separation was also performed with a Teledyne Isco CombiFlash Rf. Microwaved reactions took place in a Biotage Microwave reactor. ^{1}H and ^{13}C NMR were recorded using a 500 MHz Bruker AVIII spectrometer equipped with a cryogenically-cooled carbon observe probe using tetramethyl silane as an internal standard. Chemical shifts (δ) are reported in ppm and coupling constants (*J*) are reported in Hz. High-resolution mass spectrum (HRMS) was performed on a LCT Premier (Micromass Ltd., Manchester UK) time of flight mass spectrometer with an electrospray ion source in either positive or negative mode. Melting points were measured with a Thomas Capillary Melting Point Apparatus and are uncorrected. Chiral analysis was carried out on an Agilent 1200 RRLC with photo diode array on a Chiralpak ID column.

*Preparation of PchD warhead inhibitors (**Scheme 4-1**)*

((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl (4-cyano-2-hydroxybenzoyl)sulfamate (4-cyano-salicyl-AMS).



Scheme 4-1: Synthetic route to potential warhead inhibitors of salicylate adenylation enzymes. A) Reagents and conditions: i) TBSCl, imidazole/DMF, RT, 20 hr.; ii) TFA-H_2O-THF (1:1:4), 0˚C, 5 hr.; iii) $(Bu_3Sn)_2O$, PhH, reflux 2 hr.; iv) H_2NSO_2Cl, dioxane, 5˚C, 0.5 hr.; v) salicylic acid, CDI, 60˚C, 2hr./DBU, CH_3CN, 60˚C, 0.5 hr. vi) TBAF, THF, rt. 2hr. **B)** Salicyl-AMS, 3-cyano-salicyl-AMS, and 4-cyano-salicyl-AMS were generated by the synthetic scheme illustrated above. **C)** Salicyl-AMS can be divided into four motifs as defined by Aldrich *et al*.: an aryl, linker, glycosyl, and base.[26]

4-cyano-2-hydroxybenzoic acid (0.24 mmol, ArkPharma) and 1,1-carbonyldiimidazole (0.575 mM) were added to anhydrous acetonitrile (5 mL) under argon and stirred at 60˚C for 2 hours. 1,8-diazabicyclo[5.4.0]undec-7-ene (0.24 mmol) and ((2*R*,3*R*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-bis((*tert*-butyldimethylsilyl)oxy)tetrahydrofuran-2-yl)methyl sulfamate [37] (0.16 mmol) in 1 mL of anhydrous acetonitrile were then added to the reaction dropwise and the reaction stirred at 60˚C for an additional 30 minutes. The reaction was subsequently quenched and extracted with EtOAc (4 x 20 mL). The organic layer was washed with 1M HCl (1 x 20 mL), $NaHCO_3$ (1 x 20 mL), and brine (1 x 20 mL). After being dried with $MgSO_4$, the organic phase was concentrated under reduced pressure and the TBS-protected product was purified by Combiflash (0-30% MeOH gradient in CH_2Cl_2 over 40 minutes). Fractions containing the TBS-protected product were collected and concentrated in vacuo and left on high vacuum overnight to afford a red-orange solid (25 mg). The protected product was dissolved in anhydrous THF (1.6 mL) and 1M TBAF in THF (100 µL) was added dropwise and the reaction stirred at room temperature for 30 minutes. The product was purified by Combiflash (0-35% MeOH gradient in CH_2Cl_2) to yield a red-orange solid (13 mg, 16%). The product was only 59% pure based on the UV-VIS trace from LC-HRMS. **HRMS:** $[M + H]^+$ 492.0859 (calcd), 492.0940 (found).

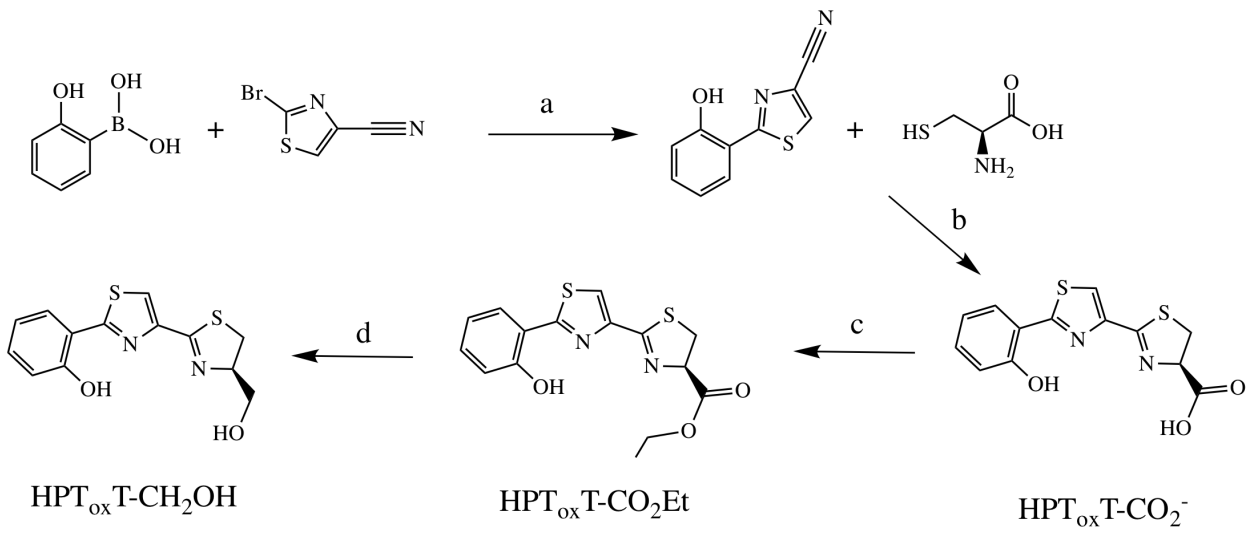
<u>((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl (4-cyano-2-hydroxybenzoyl)sulfamate</u> (3-cyano-salicyl-AMS).

A mixture of 3-cyano-2-hydroxybenzoic acid (0.72 mmol) and 1,1-carbonyldiimidazole (0.86 mM) were in anhydrous acetonitrile (9 mL) was stirred at 60˚C for 2 hours under an argon atmosphere. After cooling to room temperature, 1,8-diazabicyclo[5.4.0]undec-7-ene (0.36 mmol) and ((2*R*,3*R*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-bis((*tert*-

butyldimethylsilyl)oxy)tetrahydrofuran-2-yl)methyl sulfamate [37] (0.24 mmol) in anhydrous acetonitrile (1 mL) was added in a dropwise manner. The resulting mixture was heated to 60˚C and stirred for an additional 30 minutes. The reaction was subsequently quenched and extracted with EtOAc (4 x 30 mL). The organic layer was washed with 1M HCl (1 x 30 mL), $NaHCO_3$ (1 x 30 mL), and brine (1 x 30 mL). After being dried ($MgSO_4$), the organic phase was concentrated under reduced pressure to afford a crude oil that was purified by flash column chromatography (0-20% MeOH step gradient in CH_2Cl_2). Fractions containing the TBS-protected product were collected and concentrated in vacuo and left on high vacuum overnight to afford a yellow-orange solid (41 mg). The protected product was dissolved in anhydrous THF (1.6 mL) and 1M TBAF in THF (100 µL) was added dropwise and the reaction stirred at room temperature for 30 minutes. The product was purified by Combiflash (0-50% MeOH gradient in CH_2Cl_2 over 25 minutes) to yield a yellow-orange solid (24.9 mg, 21%). The product was only 40% pure based on the UV-VIS trace from LC-HRMS. **HRMS:** $[M + H]^+$ 492.0859 (calcd), 492.0940 (found).

Steady-state Adenylation Assay

The coupled steady-state adenylation assay is thoroughly described in the previous two chapters. The adenylation assay was performed with PchD and salicylic acid. PchD was used at 750 nM enzyme concentration and salicylic acid ranged from 0 – 20 µM. All other reaction components and concentrations remain the same as previously described. The assay was performed in triplicate.



Scheme 4-2: Substrate analog synthesis for PchG and Irp3.
Reagents and conditions: a) $Pd(PPh_3)_4$, K_2CO_3/DME, H_2O, 150 °C, 49%; b) MeOH, K_2HPO_4(aq) (pH 6.5), 60 °C; c) $SOCl_2$/EtOH, 40 °C, 70%; d) Ca(BH4)2/EtOH, −20 °C, 49%.

*Preparation of PchG and Irp3 substrate analogs (**Scheme 4-2**).*

2-(2-Hydroxyphenyl)thiazole-4-carbonitrile

A microwave vial was charged with 2-bromo-4-cyano thiazole (1.06 mmol), 2-hydroxyphenyl boronic acid (1.38 mmol), and $Pd(PPh_3)_4$ (0.0529 mmol). The vial was sealed and flushed with argon for 5 minutes. Dimethoxyethane (8 mL) and a 2 M potassium carbonate solution (1.06 mL, 2.12 mmol) were added through the septum, and the reaction was stirred at room temperature for an additional 5 minutes, then heated to 150˚C for 45 minutes using a microwave reactor. The reaction was cooled to room temperature and diluted with EtOAc/H_2O (2:1, 30 mL). The aqueous layer was extracted with EtOAc (2 x 30 mL). The organic layers were combined, washed with brine and dried over Na_2SO_4. The solvent was removed in vacuo and the resulting residue was purified by flash column chromatography (1:5 EtOAc/Pentane) to yield a pale yellow solid (104 mg, 49%). **Mp**: 169-172˚C. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 11.42 (s, 1H), 8.84 (s, 1H), 8.17 (dd, J = 7.9, 1.7 Hz, 1H), 7.39 (ddd, J = 8.6, 7.2, 1.7 Hz, 1H), 7.07 (dd, J = 8.3, 1.2 Hz, 1H), 7.00 (td, J = 7.8, 1.1 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 164.13, 155.12, 133.20, 132.06, 127.36, 124.01, 119.60, 118.24, 116.30, 114.91.

2'-(2-Hydroxyphenyl)-4,5-dihydro-[2,4'-bithiazole]-4-carboxic acid (*$HPT_{ox}T$-CO_2^-)*

L-Cysteine (0.546 mmol) and 2-(2-hydroxyphenyl)thiazole-4-carbonitrile (0.273 mmol) were combined in a round bottom flask. Methanol (5 mL) and 0.1 M potassium phosphate buffer pH 6.5 (5 mL) were added and the reaction was heated to 60˚C and stirred for 14 hours. The reaction was cooled to room temperature, diluted with H_2O (20 mL) and acidified with citric acid to ~pH 2. The mixture was filtered through a pad of celite and extracted with CH_2Cl_2 (3 x 20 mL). The

organic layers were combined and dried over Na_2SO_4. The solvent was removed in vacuo to afford a light beige solid (58 mg, 70%). The purified product was in agreement with literature values for melting point, **^{1}H NMR** and **^{13}C NMR**.[36] Enantiopurity was determined by chiral HPLC using a mobile phase of (A) water containing 0.05% formic acid (B) acetonitrile, gradient 2% to 80% B over 12 minutes, and isocratic at 80% B from 12 to 15 minutes (*R:S*, 98:2).

Ethyl 2'-(2-hydroxyphenyl)-4,5-dihydro-[2,4'-bithiazole]-4-carboxylate ($HPT_{ox}T$-CO_2Et)

Thionyl chloride (0.414 mmol) was added to ethanol (8.56 mmol) with vigorous stirring in a -10˚C ice water bath. The mixture was warmed to room temperature. *$HPT_{ox}T$-CO_2^-* (0.153 mmol) was added and the reaction was the heated to 40˚C for 4 hours. The mixture was cooled to room temperature and the solvent was removed in vacuo. The residue was reconstituted in saturated Na_2CO_3 (10 mL) and extracted with CH_2Cl_2 (4 x 10 mL). The organic layers were combined and the solvent was removed in vacuo to afford a yellow oil (24.8 mg, 49%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 11.24 (s, 1H), 8.34 (s, 1H), 8.14 (dd, J = 8.0, 1.7 Hz, 1H), 7.35 (ddd, J = 8.5, 7.2, 1.8 Hz, 1H), 7.05 (dd, J = 8.2, 1.1 Hz, 1H), 6.99 (ddd, J = 8.2, 7.2, 1.2 Hz, 1H), 5.38 (dd, J = 9.7, 8.4 Hz, 1H), 4.20 (qd, J = 7.1, 2.4 Hz, 2H), 3.64 (ABX, J_{AX} 8.4 Hz, J_{BX} 9.8, J_{AB} 11.4, 2H), 1.25 (t, J = 7.1 Hz, 3H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 170.96, 164.92, 163.34, 155.50, 147.07, 131.95, 127.81, 122.74, 120.08, 119.25, 116.92, 78.60, 61.61, 34.71, 14.52. **HRMS**: $[M + H]^+$ 335.0449 (calcd), 335.0516 (found).

2-(4-(Hydroxymethyl)-4,5-dihydro-[2,4'-bithiazol]-2'-yl)phenol ($HPT_{ox}T$-CH_2OH)

$NaBH_4$ (0.236 mmol) was added to a mixture of $HPT_{ox}T$-CO_2Et (0.236mmol), $CaCl_2 \bullet 2H_2O$ (0.109 mmol) and EtOH (1.5 mL) at -20˚C. The reaction was stirred at that temperature

overnight. The solution was raised to room temperature and saturated NH_4Cl (1 mL) was added. The reaction was then concentrated under reduced pressure leaving a residue. Deionized H_2O (8 mL) was added and the mixture was extracted with CH_2Cl_2 (3 x 15 mL). The combined organic layer was dried over Na_2SO_4 and the solvent was removed in vacuo to afford a crude oil. The oil was purified using the CombiFlash (4:1-5:1 EtOAc/CH_2Cl_2 gradient) to afford an off-white solid (22 mg, 31%). Mp: 159-160˚C. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 11.23 (s, 1H), 8.23 (s, 1H), 8.13 (dd, J = 8.0, 1.7 Hz, 1H), 7.34 (ddd, J = 8.6, 7.2, 1.7 Hz, 1H), 7.04 (dd, J = 8.3, 1.1 Hz, 1H), 7.02 – 6.96 (m, 1H), 4.99 (t, J = 5.7 Hz, 1H), 4.72 (dtd, J = 8.8, 7.2, 4.4 Hz, 1H), 3.69 (ddd, J = 10.9, 5.5, 4.4 Hz, 1H), 3.55 (ddd, J = 10.9, 6.9, 5.9 Hz, 1H), 3.45 (dd, J = 11.0, 9.0 Hz, 1H), 3.31 (dd, J = 10.9, 7.4 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 163.25, 162.04, 155.48, 147.74, 131.88, 127.82, 121.85, 120.08, 119.26, 116.93, 79.99, 62.77, 34.41. **HRMS**: $[M + H]^+$ 293.0340 (calcd), 293.0419 (found).

Results and Discussion

Salicyl-AMS

NRPS adenylation domains assist in catalyzing two distinct reactions. First, an amino acid, including nonproteinogenic acids and hydroxy acids, is activated at the expense of ATP forming an aminoacyl-AMP bond and releasing pyrophosphate. Second, the free thiol of a Ppant tether, covalently attached to a carrier domain, performs a nucleophilic attack at the acyl carbon of the aminoacyl-AMP intermediate, releasing AMP, and binding the amino acid priming the NRPS module. The first part of the reaction forming the aminoacyl-AMP intermediate, is similarly catalyzed by tRNA synthetases. Although tRNA synthetases and adenylation domains are not structurally or sequentially related, both can be inhibited by non-hydrolyzable analogs of

their shared aminoacyl-AMP intermediate, such as derivatives of 5'-*O*-[*N*-(aminoacyl)-sulfamoyl]adenosine (AMS).[61-64] In 2003, Marahiel and colleagues published the first use of aminoacyl-AMS inhibitors to inhibit NRPS adenylation domains of gramicidin S-synthetase A (GrsA) and surfactin synthetase C (SrfA-C) *in vitro*.[63] At first glance, aminoacyl-AMS inhibitors did not appear as viable candidates for treatment of pathogenic bacteria and fungi through inhibition of NRPS-derived virulence factors due to their ability to also inhibit tRNA synthetases. However, many NRPS-derived natural products include building blocks of nonproteinogenic amino acids and hydroxy acids of which are ideal candidates for aminoacyl-AMS inhibitor targets. Adenylation domains do not have human homologs and aminoacyl-AMS inhibitors of adenylation domains specific to nonproteinogenic amino acids or hydroxy acids will not be specific to tRNA synthetases.

In 2005, Ferreras *et al* introduced the first biochemically confirmed inhibitor of siderophore biosynthesis, salicyl-AMS (**Scheme 4-1**), targeting salicylate adenylation domains of siderophore biosynthesis pathways, specifically MbtA from mycobactin biosynthesis of *Mycobacterium tuberculosis* (the causative agent of tuberculosis) YbtE from yersiniabactin biosynthesis of *Yersinia pestis* (the causative agent of the black plague) and PchD from the opportunistic pathogen, *P. aeruginosa*.[19] Salicyl-AMS provided nanomolar IC_{50} and K_i values for MbtA, YbtE, and PchD and was effective in inhibiting *M. tuberculosis* and *Y. pestis* culture growth at low micromolar levels, 2.2 ± 0.3 μM and 51.2 ± 4.7 μM respectively.[19] Since then, Aldrich and coworkers have established an extensive SAR profile of salicyl-AMS, altering the linker,[21, 23, 24] glycosyl,[22, 25] aryl,[26] and base[27, 28] regions of the molecule (**Scheme 4-1C**).[65, 66] Indeed, various and often combined modifications have only slightly improved growth inhibition in comparison to the original salicyl-AMS.[22, 26]

Quadri, Tan, and Bishai recently provided the first *in vivo* proof-of-concept for using salicyl-AMS as an antibacterial showing its effectiveness in slowing *M. tuberculosis* growth of a murine lung model in 2013.[20] However, authors suggest further modifications to salicyl-AMS are ideal to lower toxicity and/or increase half-life, efficacy, and bioavailability of the compound before moving to primate models. Up until this point, modifications to salicyl-AMS have been driven by using crystal structures of DhbE, a dihydroxybenzoic acid adenylation domain, as a model for salicylate adenylation active sites. Although sharing 40-50% sequence identity, DhbE, has varying residues in the aminoacyl-AMP binding pocket in comparison to PchD, MbtA, and YbtE which likely contributes to the ability of the domain to differentiate between different hydroxy acids.[19, 67] Recently, our lab successively solved the structure of PchD, the salicylate adenylation domain from *P. aeruginosa*, with salicyl-AMS bound in the active site (**Figure 4-4**, unpublished) – the first *holo*-structure of salicylate adenylation domains. Using this structure, our goal is to rationally design novel salicyl-AMS warhead inhibitors increasing its efficacy by covalently binding to the enzyme.

Inhibitor design and synthesis of novel salicyl-AMS inhibitors

Initial analysis of the active site of PchD presented a Cys residue near the aryl moiety of salicyl-AMP. The Cys residue is found to be in similar position to a Ser residue of DhbE. We hypothesized that adding an electrophile off the salicyl ring and in the vicinity of the Cys could induce a reversible or irreversible binding event from the Cys to the inhibitor. Therefore, we followed a similar synthesis protocol by Ferreras *et al* to generate new inhibitors (**Scheme 4-1**).[19] Starting with adenosine, each hydroxyl first undertook a silyation protection, before selective deprotection of the 5'-*O*-TBS group giving a 2'3'-bis-*O*-TBS-adenosine. Aminochloro-λ^6-silanedione was separately synthesized to perform sulfamoylation on the unprotected 5'-*O*-

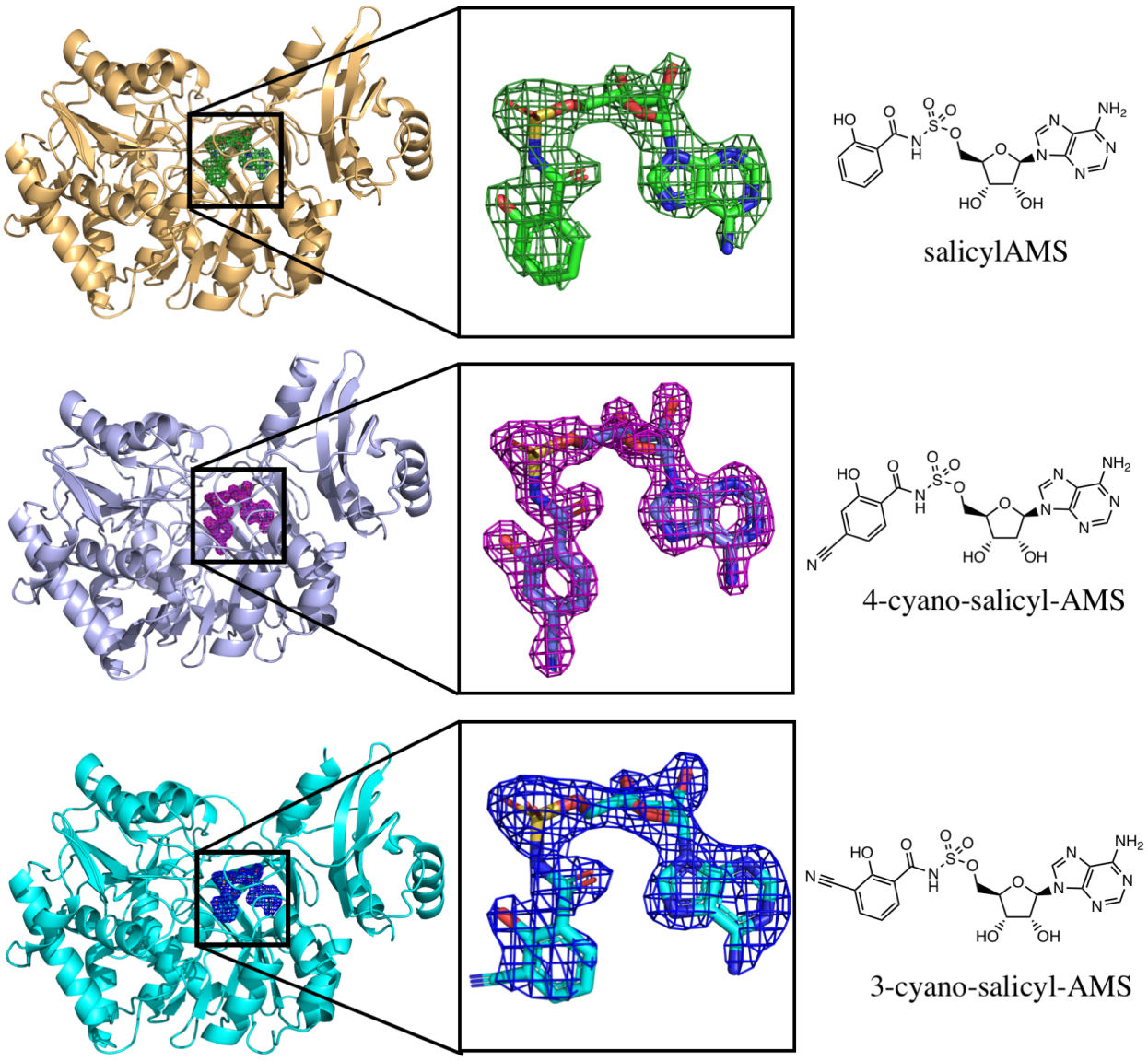


Figure 4-4: PchD co-crystallized with salicyl-AMS, 4-cyano-salicyl-AMS, and 3-cyano-salicyl-AMS. Unpublished crystal structures of PchD co-crystallized with salicyl-AMS (2.11 Å, wheat with green mesh), 4-cyano-salicylAMS (1.69 Å, light blue with purple mesh), and 3-cyano-salicylAMS (2.03 Å, cyan with blue mesh) (left). The box inserts (middle) display a $2F_o$-F_c omit map (mesh) contoured at 2σ with the inhibitor modelled into the electron density. The linear structure of the inhibitor is on the right side of the panel. Protein structure figures were generated using Pymol.[68]

position of the silylated adenosine providing a silylated sulfamate. A hydroxy acid (3- or 4-cyano-2-hydroxybenzoic acid) was charged with 1,1'-carbonyldiimidazole and coupled with the sulfamoyl amine of the silylated sulfamate. Deprotection of the silylated sugar hydroxyls wascarried out by TBAF to yield the desired cyano-salicyl-AMS. Unfortunately, careful steps to ensure purity of each intermediate step during the synthesis was not performed (i.e. ^{1}H NMR and ^{13}C NMR) leaving the final purity of 3-cyano-salicyl-AMS and 4-cyano-salicyl-AMS to 40% and 59%, respectively, as determined by LC-HRMS UV-VIS. Aldrich and colleagues have generated more robust protocols for synthesis of salicyl-AMS derivatives, including protection of the base amine and salicyl hydroxyl with a final global deprotection step. Future synthesis including protecting these groups is desired to maximize yield, reduce side products, and increase purity.

PchD crystal structures with bound salicyl-AMS inhibitors and steady-state adenylation parameters.

The crude inhibitors were co-crystallized with PchD to a resolution between 1.69 Å – 2.11 Å (**Figure 4-4**). The $2F_o$-F_c maps contoured to 2σ in **Figure 4-4** clearly show the presence of each inhibitor in the active site of PchD. The targeted Cys of PchD, not depicted in the figure, did not appear to be covalently linked to the inhibitor. Both the 3- and 4-cyano-salicyl-AMS inhibitors aiming to generate a thioimidate with the Cys were approximately 2 – 3 Å away from the rotamer. New salicyl-AMS derivatives extending the length of the electrophile are currently being synthesized. Additional modifications altering the identity of the electrophile, including various Michael acceptors, are also a future goal of the lab. Knowing the crystallization condition to produce inhibitors bound to PchD is a major benefit to rationally design more effective inhibitors for salicylate adenylation domains.

I performed steady-state kinetics of the stand-alone adenylation domain of PchD (**Figure 4-5**). The kinetic parameters can be found in **Table 4-1**. This assay was performed in a plate-reader format and provides a proof-of-concept for generating *in vitro* IC_{50} values. 3-cyano-salicyl-AMS and 4-cyano-salicyl-AMS were not used in this assay due to their crude nature.

Substrate analog synthesis for PchG and Irp3.

Mechanistic analysis of downstream tailoring domains in NRPS biosynthesis is difficult due to the complexity of the growing peptide covalently linked to the Ppant tether of the carrier domain, as previously discussed.[39-43] Thus, we synthesized proposed intermediate analogs to use in mechanistic and structural studies of PchG, from pyochelin biosynthesis, and Irp3 of yersiniabactin biosynthesis. PchG and Irp3 are NADPH-dependent thiazolinyl imine reductases acting on the growing peptide *in trans* as stand-alone tailoring domains.[30, 31, 44] While the two natural products are distinctive of one another, the upstream peptidyl motif of the proposed substrate for PchG and Irp3 is shared consisting of a three-ring system, a 2-hydroxyphenyl, a D-thiazoline, and a 4-L-thiazoline (HPTT-S-Enz). Reduction occurs on the 4-L-thiazoline. During reconstitution assays, the Ppant tether is susceptible to hydrolysis releasing a carboxylate intermediate. Walsh *et al* reported isolation of $HPT_{ox}T\text{-}CO_2^-$ from pyochelin full-reconstitution assays in the absence of AdoMet, PchG, or NADPH.[31] Additionally, we provide evidence in Chapter 2 suggesting the reduction of the L-thiazoline to thiazolidine is necessary for *N*-methylation identifying the likely order of catalysis in pyochelin biosynthesis. Conversely, the order of Irp3 catalyzed reduction of the L-thiazoline in yersiniabactin biosynthesis is much less defined. During yersiniabactin full-reconstitution assays, experiments lacking YbtU (Irp3 homolog from *Y. pestis,* sharing 98% identity) produced a product with a mass indicative of a

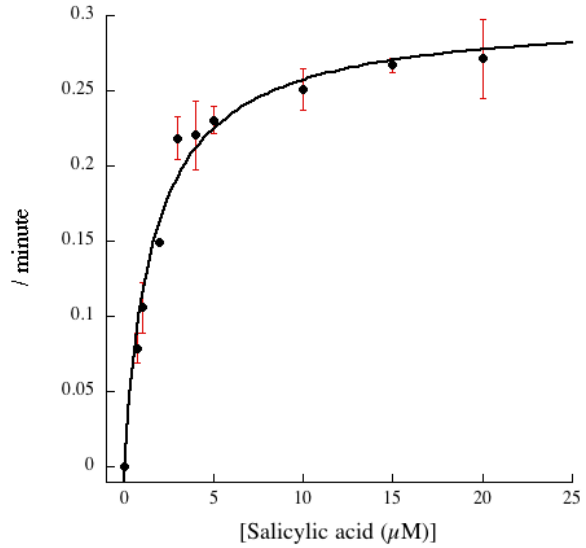


Figure 4-5: Michaelis-Menten plot of PchD and salicylic acid from steady-state adenylation assay. The coupled adenylation assay is described in detail in previous chapters. Steady-state parameters were generated for PchD and salicylic acid establishing a viable assay to perform IC_{50} experiments with potential inhibitors of PchD's adenylation activity.

Table 4-1: Salicylic Acid Adenylation Assay of PchD		
	PchD*	PchD
k_{cat} (min^{-1})	74	0.30 ± 0.02
K_m (Salicylic acid μM)	2.8	1.7 ± 0.2
k_{cat}/ K_m ($M^{-1}s^{-1}$)	45000	3000 ± 400

* - Quadri *et al*. *Biochemistry* 1999, 38, 14941-14954. A different assay used to measure adenylation activity. Activity quantitated by measuring the reverse incorporation of ^{32}pyrophosphate into ATP by the reverse reaction. As discussed in chapter 2, Aldrich and coworkers have found considerable differences between the ^{32}pyrophosphate assay and the assay found herein.[69, 70]

non-reduced thiazoline.[30] Indeed, $HPT_{ox}T\text{-}CO_2^-$ was also isolated during reconstitution assays, but this occurred with and without YbtU; suggesting reduction may occur at any point in the biosynthesis after cyclization of the L-cysteine to L-thiazoline as YbtU acts *in trans* during peptide assembly.[30] However, isolation of the non-reduced product in the absence of YbtU suggests reduction of the thiazoline is not a precursor to the rest of yersiniabactin bioassembly, unlike PchG in pyochelin biosynthesis. Since $HPT_{ox}T\text{-}CO_2^-$ was isolated from each pathway it remains a possibility that they share a common substrate. Therefore, we decided to use $HPT_{ox}T\text{-}CO_2^-$ as a starting template to generate substrate analogs.

We provide a simpler synthetic method to $HPT_{ox}T\text{-}CO_2^-$ (**Scheme 4-2**) than the route previously reported by Mislin *et al*.[45] The synthesis requires two steps. First, 2-(2-hydroxyphenyl)thiazole-4-carbonitrile was produced by a Suzuki coupling reaction between 2-hyrdoxyphenyl boronic acid and 2-bromo-4-cyano thiazole providing a 49% yield. Second, L-cysteine was added to the nitrile prompting nucleophilic addition and cyclization to a thiazoline affording a 70% yield of the product. $HPT_{ox}T\text{-}CO_2^-$ retained a 98:2 (*R*:*S*) ratio as determined by chiral chromatography. Additional analogs were also generated varying at the carboxyl terminus. The terminal carboxyl was first altered to an ethyl ester more closely resembling the thioester linkage to the Ppant tether. The carboxyl of $HPT_{ox}T\text{-}CO_2^-$ was transformed to an acyl chloride promoting nucleophilic addition of the solvent ethanol providing $HPT_{ox}T\text{-}CO_2Et$ in 49% yield. The ester was successively reduced by a calcium borohydride catalyzed reduction giving a 31% yield of $HPT_{ox}T\text{-}CH_2OH$. Each of the substrate analogs, $HPT_{ox}T\text{-}CO_2^-$, $HPT_{ox}T\text{-}CO_2Et$, and $HPT_{ox}T\text{-}CH_2OH$ were used in crystallization and kinetic experiments by other researchers in the lab. The following Irp3 and PchG sections highlight our findings. This work was published in Meneely *et al*.[46]

Crystallization and structural analysis of Irp3 with substrate analog $HPT_{ox}T\text{-}CO_2^-$.

Previously, our lab reported two structures of Irp3: an apo- (1.85 Å, PDB: 4GMF) and $NADP^+$-bond form (2.31 Å, PDB: 4GMG).[35] The overall structure of Irp3 is similar to sugar oxidoreductases and biliverdin reductase but shares no structural similarity with other imine reductases. Irp3 is a homodimer in solution and in its crystal forms. Each monomer consists of two domains: a *N*-terminal Rossman fold domain responsible for NADP(H)-binding and a *C*-terminal α/β domain responsible for dimerization (**Figure 4-6 A, B**).[35] A new crystallization condition in a lower symmetry lattice garnered a $NADP^+$-bound Irp3 structure at a higher resolution (1.45 Å) with a root-mean-square deviation (r.m.s.d.) of 0.54 – 0.75 Å for 336-347 Cα residues depending on the monomer being compared.[46] The apo- and NADPH-bound structures were missing a loop connecting the M-N strands in the *C*-terminal dimerization domain (residues 250-277) whereas the NADPH-bound and holo-structures in the new crystal lattice provided a fully ordered loop forming a cap over the active site of the opposing monomer (**Figure 4-6 C**).[46]

A holo-Irp3 structure with the substrate analog, $HPT_{ox}T\text{-}CO_2^-$, and $NADP^+$ was also determined in the new crystal lattice at a high resolution of 1.28 Å.[46] The structure was very similar to the new $NADP^+$-bound structure having a r.m.s.d. of 0.15-0.49 Å for 353 Cα residues depending on the monomer. The $NADP^+$ was fully occupied in the active site although the $HPT_{ox}T\text{-}CO_2^-$ was only 70-85% occupied depending on the monomer. $HPT_{ox}T\text{-}CO_2^-$ bound in a seemingly hydrophobic pock with no apparent binding interactions (i.e., hydrogen bonding, pi stacking) with Irp3. The carboxyl of $HPT_{ox}T\text{-}CO_2^-$ is within hydrogen bonding distance of the nicotinamide amine and tyrosine 128. It is important to note the actual substrate contains an additional methylated malonyl group and final methylated thiazoline connected to a Ppant tether.

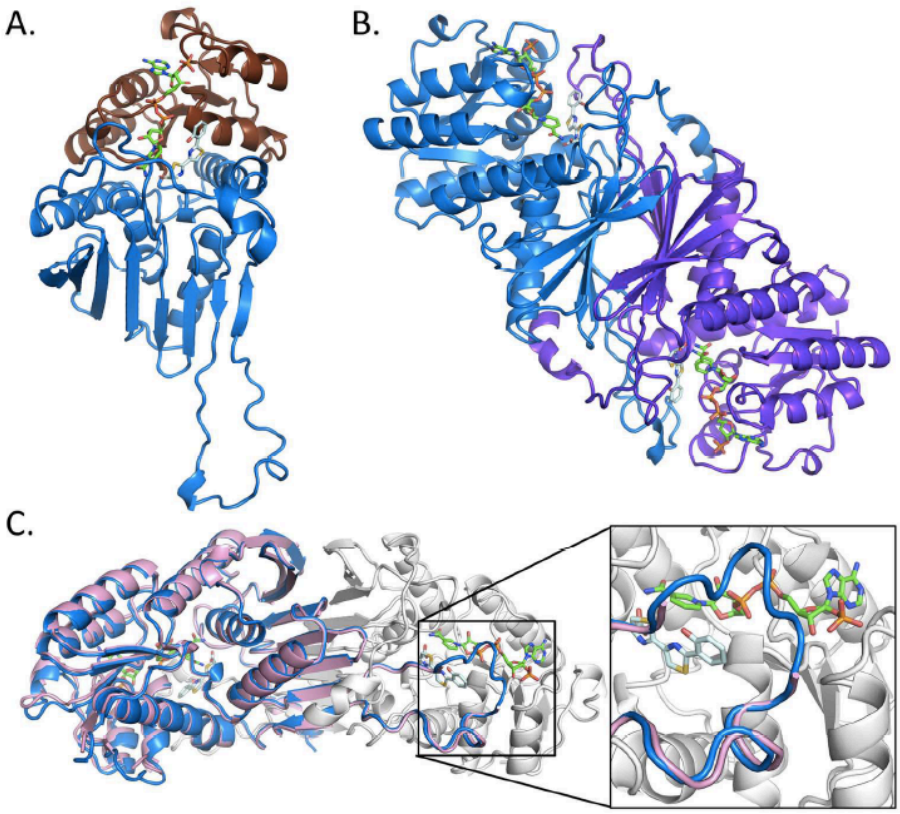


Figure 4-6: Crystal structure of Irp3. A) The tertiary structure of Irp3 represents two domains. A *N*-terminal NADP(H) binding domain (brown) and a *C*-terminal dimerization domain (blue). **B)** The quaternary structure of Irp3 is a dimer: monomer A (blue), monomer B (purple). **C)** The new crystal structure of Irp3 (blue and white) was co-crystallized with NADP(H) (green sticks) and substrate analog, $HPT_{ox}T\text{-}CO_2H$ (cyan sticks), elucidating the M-N loop closing over top of the active site (inset – blue loop). This figure was taken directly from Meneely *et al.* *Biochemistry* 2016, 55: 5423-5433.[46]

A tunnel past the carboxylate of $HPT_{ox}T\text{-}CO_2^-$ extends to the surface of the protein and is the likely cavity of the natural substrate. Further space is also found in the substrate binding cavity past the hydroxyphenyl of $HPT_{ox}T\text{-}CO_2^-$. However, it is unlikely that natural substrate would protrude further into that void as it would move the substrate beyond the nicotinamide ring of NADPH, making hydride transfer unlikely. The wideness of the cavity and few interactions with the protein likely allow the analog lacking the stereochemistry of the natural D-thiazoline (analog was altered to an oxidized thiazole) to bind in the active site. The natural substrate would provide a slightly angled spatial arrangement opposed to the planar orientation of the analog in the active site.

Donation of a hydride from the nicotinamide ring of $NADP^+$ and a proton from a general acid residue is the proposed catalytic mechanism of the structurally homologous sugar oxidoreductases. [47-50] The general acid is hypothesized to be the lysine of the consensus sequence EKP (EHP in thiazolinyl reductases) or the final residue of the $GGX_3DX_3(Y/H)$ consensus sequence.[47-50] The $GGX_3DX_3(Y/H)$ consensus sequence is not conserved in thiazolinyl reductases although tyrosine 128 is in the same three dimensional space as the Y/H in sugar oxidoreductases. The likely general acid candidate in Irp3 is tyrosine 128 which is 4.4 Å from the nitrogen of the substrate analog, $HPT_{ox}T\text{-}CO_2^-$, whereas the EHP histidine is 9.0 Å. Therefore, we propose a mechanism similar to dihyrofolate reductase where tyrosine 128 is the general acid donating a proton to the nitrogen as hydride transfer occurs on the electropositive carbon.[32]

In the $NADP^+$-bound structures, the nicotinamide ring is pi-stacking with phenylalanine 17 of Irp3. The nicotinamide and Phe17 are shifted in the holo-Irp3 to align the nicotinamide

with the second thiazoline ring of $HPT_{ox}T\text{-}CO_2^-$. Yersiniabactin and pyochelin are isolated from culture as diastereomers altering stereochemistry at the C-10 carbon (**Figure 4-1**).[51-53] Although, when binding to iron or other transition metals, the labile carbon becomes locked in the (*R*) confirmation for both pyochelin and yersiniabactin.[52, 54-60] Excellent electron density at this high resolution provides a greater amount of density around the sulfur atoms compared to the carbon and nitrogen atoms distinguishing the orientation of the analog (**Figure 4-7**). Interestingly, the substrate analog is positioned in a confirmation that would promote reduction to the stereochemistry of the non-metal binding diastereomer of yersiniabactin. The current position of the hydride transfer and proton donation from the proposed general acid would promote (*S*) stereochemistry at the C-10 carbon. It may be possible that the substrate analog is bound in a catalytically incompetent manor. If this were the case, the entire molecule or the just the terminal thiazoline ring would need to be rotated 180˚ to regain catalytic competence. This would better position the carboxyl towards the proposed Ppant tunnel but would also move the nitrogen further way from the proposed general acid residues. Further investigation is necessary to ascertain the catalytically competent position of the substrate. It should also be noted that attempts were made to co-crystallize or soak in $HPT_{ox}T\text{-}CO_2Et$ or $HPT_{ox}T\text{-}CH_2OH$ with $NADP^+$ but density for these analogs was not observed in the active site.

Steady-state kinetics with $HPT_{ox}T\text{-}CO_2^-$.

Each of the substrate analogs ($HPT_{ox}T\text{-}CO_2^-$, $HPT_{ox}T\text{-}CO_2Et$ and $HPT_{ox}T\text{-}CH_2OH$) were used to measure reductase activity of PchG and Irp3. Common experiments of NADPH-dependent enzymes measure the decrease of absorbance at 340 nm as NADPH is oxidized to $NADP^+$. However, each of the substrate analogs absorb in the same region prohibiting this

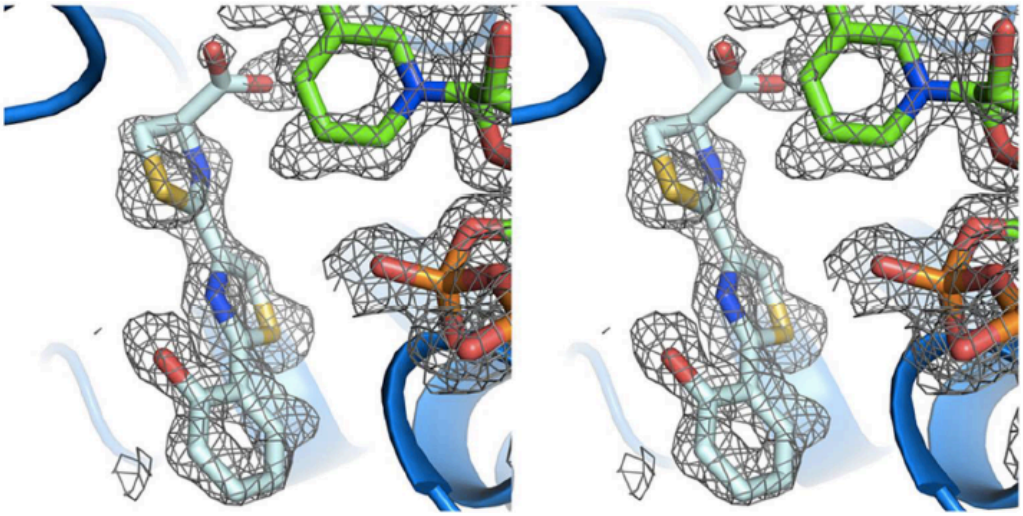


Figure 4-7: $HPT_{ox}T\text{-}CO_2H$ bound in the active site of Irp3. A stereoview of a simulated annealing map contoured to 1.5σ (gray cages) showing the substrate analog, $HPT_{ox}T\text{-}CO_2H$ (cyan sticks), in the active site of Irp3. This figure was taken directly from Meneely *et al*. *Biochemistry* 2016, 55: 5423-5433.[46]

simple experimental setup. Therefore, an HPLC method was created to separate the analogs from NADPH and $NADP^+$ monitoring the decrease in absorbance of the NADPH at 340 nm. Detection of NADPH oxidation did not occur with Irp3 and any of the substrate analogs. Steady-state kinetics of PchG with $HPT_{ox}T\text{-}CO_2^-$ was observed albeit at non-physiological rates. The kinetic parameters were determined: K_m of 80 ± 20 μM, a k_{cat} of 0.020 ± 0.001 min^{-1}, and a k_{cat}/K_m of 4 ± 1 $M^{-1}s^{-1}$.[46] It is important to remember the natural substrate for PchG and Irp3 is attached to a Ppant tether covalently linked to the carrier domain of another NRPS. Several explanations can be given as to why the substrate analogs were poorly or not at all compatible with each enzyme but two should be highlighted. First, the terminus of the analogs may not have mimicked the thioester linkage closely enough for PchG or Irp3. On the other hand, Irp3 reduction may occur later in the biosynthesis and require the additional hydroxylated malonyl and methylated thiazoline of the mature yersiniabactin. Second, the domain interactions with the Ppant carrier domain may be necessary for catalysis or product release. Development of a series of analogs more closely mimicking the natural substrate of each enzyme may be necessary to further elucidate the mechanism.

Conclusion

In summary we have presented structural and kinetic studies of stand-alone adenylation domains and stand-alone tailoring thiazolinyl reductases of siderophore biosynthesis. First, by generating substrate analogs, we provide the first closed-holo-structure of a thiazolinyl reductase, Irp3, from yersiniabactin biosynthesis in addition to an $NADP^+$-bound structure each determined at high resolution. The substrate analog bound in the active site provides support that tyrosine 128 is the general acid for catalysis. Additional steady-state kinetics with the substrate analog provided kinetic parameters for NADPH-dependent reduction by PchG, albeit at physiologically

irrelevant rates. Second, co-crystallization of salicyl-AMS with PchD, a salicylate adenylation domain from pyochelin biosynthesis, allowed rational design for more effective inhibitors. Two additional inhibitors were presented here and co-crystallized in the active site of PchD. Adenylation steady-state parameters of PchD were also presented as a proof-of-concept for measuring future *in vitro* inhibition of the enzyme in the presence of novel inhibitors.

References

[1] Hood, M. I., and Skaar, E. P. (2012) Nutritional immunity: transition metals at the pathogen-host interface, *Nature Reviews Microbiology 10*, 525-537.

[2] Skaar, E. P. (2010) The battle for iron between bacterial pathogens and their vertebrate hosts, *PLoS Pathog 6*, e1000949.

[3] Cornelis, P., and Matthijs, S. (2002) Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines, *Environ Microbiol 4*, 787-798.

[4] Crosa, J. H., and Walsh, C. T. (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria, *Microbiol Mol Biol Rev 66*, 223-249.

[5] Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis, *Chemical reviews 97*, 2651-2674.

[6] Organization, W. H. (2018) Tuberculosis, World Health Organization.

[7] Perry, R. D., and Fetherston, J. D. (1997) Yersinia pestis--etiologic agent of plague, *Clin Microbiol Rev 10*, 35-66.

[8] Tacconelli, E. C., Magrini, N.C., . (2017) Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics, World Health Organization.

[9] Bearden, S. W., Fetherston, J. D., and Perry, R. D. (1997) Genetic organization of the yersiniabactin biosynthetic region and construction of avirulent mutants in Yersinia pestis, *Infect Immun 65*, 1659-1668.

[10] Gobin, J., and Horwitz, M. A. (1996) Exochelins of Mycobacterium tuberculosis remove iron from human iron-binding proteins and donate iron to mycobactins in the M. tuberculosis cell wall, *J Exp Med 183*, 1527-1532.

[11] De Voss, J. J., Rutter, K., Schroeder, B. G., Su, H., Zhu, Y., and Barry, C. E., 3rd. (2000) The salicylate-derived mycobactin siderophores of Mycobacterium tuberculosis are essential for growth in macrophages, *Proc Natl Acad Sci U S A 97*, 1252-1257.

[12] Smith, I. (2003) Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence, *Clin Microbiol Rev 16*, 463-496.

[13] Lopez-Medina, E., Fan, D., Coughlin, L. A., Ho, E. X., Lamont, I. L., Reimmann, C., Hooper, L. V., and Koh, A. Y. (2015) Candida albicans Inhibits Pseudomonas aeruginosa Virulence through Suppression of Pyochelin and Pyoverdine Biosynthesis, *PLoS Pathog 11*, e1005129.

[14] Takase, H., Nitanai, H., Hoshino, K., and Otani, T. (2000) Requirement of the Pseudomonas aeruginosa tonB gene for high-affinity iron acquisition and infection, *Infect Immun 68*, 4498-4504.

[15] Takase, H., Nitanai, H., Hoshino, K., and Otani, T. (2000) Impact of siderophore production on Pseudomonas aeruginosa infections in immunosuppressed mice, *Infect Immun 68*, 1834-1839.

[16] Quadri, L. E. (2007) Strategic paradigm shifts in the antimicrobial drug discovery process of the 21st century, *Infect Disord Drug Targets 7*, 230-237.

[17] Blundell, T. L. (2017) Protein crystallography and drug discovery: recollections of knowledge exchange between academia and industry, *IUCrJ 4*, 308-321.

[18] Holdgate, G. A., Meek, T. D., and Grimley, R. L. (2018) Mechanistic enzymology in drug discovery: a fresh perspective, *Nat Rev Drug Discov 17*, 115-132.

[19] Ferreras, J. A., Ryu, J.-S., Di Lello, F., Tan, D. S., and Quadri, L. E. (2005) Small-molecule inhibition of siderophore biosynthesis in Mycobacterium tuberculosis and Yersinia pestis, *Nat Chem Biol 1*, 29-32.

[20] Lun, S., Guo, H., Adamson, J., Cisar, J. S., Davis, T. D., Chavadi, S. S., Warren, J. D., Quadri, L. E., Tan, D. S., and Bishai, W. R. (2013) Pharmacokinetic and in vivo efficacy studies of the mycobactin biosynthesis inhibitor salicyl-AMS in mice, *Antimicrobial agents and chemotherapy 57*, 5138-5140.

[21] Somu, R. V., Boshoff, H., Qiao, C., Bennett, E. M., Barry, C. E., 3rd, and Aldrich, C. C. (2006) Rationally designed nucleoside antibiotics that inhibit siderophore biosynthesis of Mycobacterium tuberculosis, *J Med Chem 49*, 31-34.

[22] Somu, R. V., Wilson, D. J., Bennett, E. M., Boshoff, H. I., Celia, L., Beck, B. J., Barry, C. E., 3rd, and Aldrich, C. C. (2006) Antitubercular nucleosides that inhibit siderophore biosynthesis: SAR of the glycosyl domain, *J Med Chem 49*, 7623-7635.

[23] Vannada, J., Bennett, E. M., Wilson, D. J., Boshoff, H. I., Barry, C. E., 3rd, and Aldrich, C. C. (2006) Design, synthesis, and biological evaluation of beta-ketosulfonamide adenylation inhibitors as potential antitubercular agents, *Org Lett 8*, 4707-4710.

[24] Engelhart, C. A., and Aldrich, C. C. (2013) Synthesis of chromone, quinolone, and benzoxazinone sulfonamide nucleosides as conformationally constrained inhibitors of adenylating enzymes required for siderophore biosynthesis, *J Org Chem 78*, 7470-7481.

[25] Dawadi, S., Viswanathan, K., Boshoff, H. I., Barry, C. E., 3rd, and Aldrich, C. C. (2015) Investigation and conformational analysis of fluorinated nucleoside antibiotics targeting siderophore biosynthesis, *J Org Chem 80*, 4835-4850.

[26] Qiao, C., Gupte, A., Boshoff, H. I., Wilson, D. J., Bennett, E. M., Somu, R. V., Barry, C. E., 3rd, and Aldrich, C. C. (2007) 5'-O-[(N-acyl)sulfamoyl]adenosines as antitubercular agents that inhibit MbtA: an adenylation enzyme required for siderophore biosynthesis of the mycobactins, *J Med Chem 50*, 6080-6094.

[27] Gupte, A., Boshoff, H. I., Wilson, D. J., Neres, J., Labello, N. P., Somu, R. V., Xing, C., Barry, C. E., and Aldrich, C. C. (2008) Inhibition of siderophore biosynthesis by 2-triazole

substituted analogues of 5'-O-[N-(salicyl)sulfamoyl]adenosine: antibacterial nucleosides effective against Mycobacterium tuberculosis, *J Med Chem 51*, 7495-7507.

[28] Neres, J., Labello, N. P., Somu, R. V., Boshoff, H. I., Wilson, D. J., Vannada, J., Chen, L., Barry, C. E., 3rd, Bennett, E. M., and Aldrich, C. C. (2008) Inhibition of siderophore biosynthesis in Mycobacterium tuberculosis with nucleoside bisubstrate analogues: structure-activity relationships of the nucleobase domain of 5'-O-[N-(salicyl)sulfamoyl]adenosine, *J Med Chem 51*, 5349-5370.

[29] Vergnolle, O., Xu, H., Tufariello, J. M., Favrot, L., Malek, A. A., Jacobs, W. R., Jr., and Blanchard, J. S. (2016) Post-translational Acetylation of MbtA Modulates Mycobacterial Siderophore Biosynthesis, *J Biol Chem 291*, 22315-22326.

[30] Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002) Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of Yersinia pestis, *Chem Biol 9*, 333-344.

[31] Reimmann, C., Patel, H. M., Serino, L., Barone, M., Walsh, C. T., and Haas, D. (2001) Essential PchG-dependent reduction in pyochelin biosynthesis of Pseudomonas aeruginosa, *J Bacteriol 183*, 813-820.

[32] Patel, H. M., and Walsh, C. T. (2001) In vitro reconstitution of the Pseudomonas aeruginosa nonribosomal peptide synthesis of pyochelin: characterization of backbone tailoring thiazoline reductase and N-methyltransferase activities, *Biochemistry 40*, 9023-9031.

[33] Geoffroy, V. A., Fetherston, J. D., and Perry, R. D. (2000) Yersinia pestis YbtU and YbtT are involved in synthesis of the siderophore yersiniabactin but have different effects on regulation, *Infect Immun 68*, 4452-4461.

[34] Meneely, K. M., and Lamb, A. L. (2012) Two structures of a thiazolinyl imine reductase from Yersinia enterocolitica provide insight into catalysis and binding to the nonribosomal peptide synthetase module of HMWP1, *Biochemistry 51*, 9002-9013.

[35] Meneely, K. M., and Lamb, A. L. (2012) Two Structures of a Thiazolinyl Imine Reductase from Yersinia enterocolitica (Irp3) Provide Insight for Catalysis and Binding to the Nonribosomal Peptide Synthetase Module of HMWP1, *Biochemistry 51*, 9002-9013.

[36] Mislin, G. L., Burger, A., and Abhallah, M., A,. (2004) Synthesis of new thiazole analogues of pyochelin, a siderophore of Pseudomonas aeruginosa and Burkholderia cepacia. A new conversion of thiazolines into thiazoles, *Tetrahedron 60*, 12139-12145.

[37] Tan, D. S., Quadri, L.E.N, Ryu, J., Cisar, J.S., Ferreras J.A., Lu, X. . (2006) Inhibitors of salicylate adenylation enzymes as antimicrobial agents for treatment of infections, In *PCT Int. Appl.* (Appl., P. I., Ed.), Sloan-Kettering Institute for Cancer Research, Cornell Research Foundation, International.

[38] Geisler, J., Schneider, F., Lopez Holguin, F., Lovis, K. (2006) Industrially applicable process for the sulfamoylation of alcohols and phenols, In *United States Patent Office*, Schering, AG., United States.

[39] Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases, *Chem Biol 7*, 765-772.

[40] Schneider, T. L., Shen, B., and Walsh, C. T. (2003) Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation, *Biochemistry 42*, 9722-9730.

[41] Tseng, C. C., Bruner, S. D., Kohli, R. M., Marahiel, M. A., Walsh, C. T., and Sieber, S. A. (2002) Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase, *Biochemistry 41*, 13350-13359.

[42] Holzbaur, I. E., Harris, R. C., Bycroft, M., Cortes, J., Bisang, C., Staunton, J., Rudd, B. A., and Leadlay, P. F. (1999) Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase, *Chem Biol 6*, 189-195.

[43] Gokhale, R. S., Hunziker, D., Cane, D. E., and Khosla, C. (1999) Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase, *Chem Biol 6*, 117-125.

[44] Ronnebaum, T. A., and Lamb, A. L. (2018) Nonribosomal peptides for iron acquisition: pyochelin biosynthesis as a case study, *Curr Opin Struct Biol 53*, 1-11.

[45] Mislin, G. L., Burger, A., and Abdallah, M. A. (2004) Synthesis of new thiazole analogues of pyochelin, a siderophore of Pseudomonas aeruginosa and Burkholderia cepacia. A new conversion of thiazolines into thiazoles, *Tetrahedron 60*, 12139-12145.

[46] Meneely, K. M., Ronnebaum, T. A., Riley, A. P., Prisinzano, T. E., and Lamb, A. L. (2016) Holo Structure and Steady State Kinetics of the Thiazolinyl Imine Reductases for Siderophore Biosynthesis, *Biochemistry 55*, 5423-5433.

[47] Kubiak, R. L., and Holden, H. M. (2011) Combined structural and functional investigation of a C-3"-ketoreductase involved in the biosynthesis of dTDP-L-digitoxose, *Biochemistry 50*, 5905-5917.

[48] Thoden, J. B., and Holden, H. M. (2011) Biochemical and structural characterization of WlbA from Bordetella pertussis and Chromobacterium violaceum: enzymes required for the biosynthesis of 2,3-diacetamido-2,3-dideoxy-D-mannuronic acid, *Biochemistry 50*, 1483-1491.

[49] Dambe, T. R., Kuhn, A. M., Brossette, T., Giffhorn, F., and Scheidig, A. J. (2006) Crystal structure of NADP(H)-dependent 1,5-anhydro-D-fructose reductase from Sinorhizobium morelense at 2.2 A resolution: construction of a NADH-accepting mutant and its application in rare sugar synthesis, *Biochemistry 45*, 10030-10042.

[50] Kingston, R. L., Scopes, R. K., and Baker, E. N. (1996) The structure of glucose-fructose oxidoreductase from Zymomonas mobilis: an osmoprotective periplasmic enzyme containing non-dissociable NADP, *Structure 4*, 1413-1428.

[51] Rinehart, K. L., Jr., Staley, A. L., Wilson, S. R., Ankenbauer, R. G., and Cox, C. D. (1995) Stereochemical assignment of the pyochelins, *J. Org. Chem. 60*, 2786-2791.

[52] Ino, A., and Murabayashi, A. (2001) Synthetic studies of thiazoline and thiazolidine-containing natural products. Part 3: Total synthesis and absolute configuration of the siderophore yersiniabactin, *Tetrahedron 57*, 1897-1902.

[53] Ankenbauer, R. G., and Cox, C. D. (1988) Isolation and characterization of Pseudomonas aeruginosa mutants requiring salicylic acid for pyochelin biosynthesis, *J Bacteriol 170*, 5364-5367.

[54] Schlegel, K., Taraz, K., and Budzikiewicz, H. (2004) The stereoisomers of pyochelin, a siderophore of Pseudomonas aeruginosa, *Biometals 17*, 409-414.

[55] Chambers, C. E., McIntyre, D. D., Mouck, M., and Sokol, P. A. (1996) Physical and structural characterization of yersiniophore, a siderophore produced by clinical isolates of Yersinia enterocolitica, *Biometals 9*, 157-167.

[56] Cobessi, D., Celia, H., and Pattus, F. (2005) Crystal structure at high resolution of ferric-pyochelin and its membrane receptor FptA from Pseudomonas aeruginosa, *J Mol Biol 352*, 893-904.

[57] Klumpp, C., Burger, A., Mislin, G. L., and Abdallah, M. A. (2005) From a total synthesis of cepabactin and its 3:1 ferric complex to the isolation of a 1:1:1 mixed complex between iron (III), cepabactin and pyochelin, *Bioorg Med Chem Lett 15*, 1721-1724.

[58] Hayen, H., and Volmer, D. A. (2006) Different iron-chelating properties of pyochelin diastereoisomers revealed by LC/MS, *Anal Bioanal Chem 385*, 606-611.

[59] Schlegel, K., Lex, J., Taraz, K., and Budzikiewicz, H. (2006) The X-ray structure of the pyochelin Fe3+ complex, *Z Naturforsch C 61*, 263-266.

[60] Tseng, C. F., Burger, A., Mislin, G. L., Schalk, I. J., Yu, S. S., Chan, S. I., and Abdallah, M. A. (2006) Bacterial siderophores: the solution stoichiometry and coordination of the Fe(III) complexes of pyochelin and related compounds, *J Biol Inorg Chem 11*, 419-432.

[61] Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs, *Nature 347*, 203-206.

[62] Weber, T., and Marahiel, M. A. (2001) Exploring the domain structure of modular nonribosomal peptide synthetases, *Structure 9*, R3-9.

[63] Finking, R., Neumuller, A., Solsbacher, J., Konz, D., Kretzschmar, G., Schweitzer, M., Krumm, T., and Marahiel, M. A. (2003) Aminoacyl adenylate substrate analogues for the inhibition of adenylation domains of nonribosomal peptide synthetases, *Chembiochem 4*, 903-906.

[64] Ueda, H., Shoku, Y., Hayashi, N., Mitsunaga, J., In, Y., Doi, M., Inoue, M., and Ishida, T. (1991) X-ray crystallographic conformational study of 5'-O-[N-(L-alanyl)-

sulfamoyl]adenosine, a substrate analogue for alanyl-tRNA synthetase, *Biochim Biophys Acta 1080*, 126-134.

[65] Labello, N. P., Bennett, E. M., Ferguson, D. M., and Aldrich, C. C. (2008) Quantitative three dimensional structure linear interaction energy model of 5'-O-[N-(salicyl)sulfamoyl]adenosine and the aryl acid adenylating enzyme MbtA, *J Med Chem 51*, 7154-7160.

[66] Nelson, K. M., Viswanathan, K., Dawadi, S., Duckworth, B. P., Boshoff, H. I., Barry, C. E., 3rd, and Aldrich, C. C. (2015) Synthesis and Pharmacokinetic Evaluation of Siderophore Biosynthesis Inhibitors for Mycobacterium tuberculosis, *Journal of medicinal chemistry 58*, 5459-5475.

[67] May, J. J., Kessler, N., Marahiel, M. A., and Stubbs, M. T. (2002) Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases, *Proc. Natl. Acad. Sci. U S A 99*, 12120-12125.

[68] DeLano, W. (2002) The PyMOL Molecular Graphics System, Inc., DeLano Scientific.

[69] Duckworth, B. P., Wilson, D. J., and Aldrich, C. C. (2016) Measurement of Nonribosomal Peptide Synthetase Adenylation Domain Activity Using a Continuous Hydroxylamine Release Assay, *Methods Mol Biol 1401*, 53-61.

[70] Wilson, D. J., and Aldrich, C. C. (2010) A continuous kinetic assay for adenylation enzyme activity and inhibition, *Anal Biochem 404*, 56-63.

Chapter 5

Conclusion

Iron acquisition by siderophores

Pseudomonas aeruginosa is a Gram-negative bacterium that can exist in a variety of environments and is regularly found in soil, water, plants, animals, and humans. This commensal bacterium is considered an opportunistic pathogen, often infecting humans with immunodeficiency. *P. aeruginosa* is one of the leading nosocomial, or hospital-acquired, pathogens. According to the Center for Disease Control (CDC), in the United States, nearly 8% of all nosocomial infections are caused by *P. aeruginosa* and nearly 13% of these infections are multidrug resistant.[1] Thus, the world health organization has named *P. aeruginosa* as a critical priority for expanded research.[2]

The ability of *P. aeruginosa* to acquire metals needed for primary metabolic processes is not only necessary for its survival but is also a critical step in its pathogenesis. One of these essential elements is ferric iron. Eukaryotes sequester iron, leaving freely available ferric iron in limited concentrations ranging from 10^{-18} – 10^{-24} M.[3, 4] While bacteria need micromolar concentrations to grow, pathogenic bacteria have developed many intricate systems to scavenge iron from its surrounding environment.[3-7] One of the main strategies used by pathogenic bacteria to scavenge ferric iron is by the production of siderophores, or low-molecular weight molecules possessing a high affinity for ferric iron.[4] *P. aeruginosa* utilizes two siderophores, pyoverdine and pyochelin, which have been shown to be virulence factors in different models of infection.[8-10] The enzymes responsible for the production of siderophores rarely have human homologues

making them ideal antimicrobial targets.[11] This dissertation focuses on better understanding the nonribosomal peptide synthetase (NRPS) system responsible for the biosynthesis of pyochelin.

Pyochelin biosynthesis and stuffed domains

Pyochelin biosynthesis employs two accessory enzymes (PchA and PchB), one stand-alone adenylation enzyme (PchD), one stand-alone tailoring enzyme (PchG), and two modular NPRS enzymes (PchE and PchF), each consisting of 5 different catalytic domains.[12-15] While pyochelin is considered a small siderophore (324 Da),[16-18] its biosynthesis includes many facets which are ubiquitous in NRPS bioassembly, yet remain poorly understood. One of the aspects is the inclusion of structurally unique tailoring domains which are stuffed within the adenylation domain in NRPS modules.[19, 20] We consider these to be "stuffed" tailoring domains.[20] PchE has a proposed stuffed non-canonical epimerase domain and PchF has a stuffed methyltransferase domain. A large portion of this work aimed to better understand each stuffed tailoring domain's catalytic mechanism in addition to the order in which these domains perform their chemistry during pyochelin biosynthesis.

Although mechanistic and structural studies of stuffed tailoring domains is in its infancy, the majority of the published work has been targeted towards stuffed methyltransferase domains, specifically in the biosynthesis of depsipeptides, thiocoraline and kutzneride.[21-24] In these studies, Garneau-Tsodikova *et al* provide evidence and strongly advocate for a global mechanism of methyl transfer by stuffed methyltransferase domains. They suggest that methyl transfer occurs to either the aminoacyl-AMP intermediate or the amino acid tethered to the Ppant of the carrier domain, but before condensation with the upstream peptide.[21-24] This would suggest stuffed

domains are not specific to the growing peptide but could generally be inserted into any adenylation domain to tailor the amino acid before condensing with the nascent peptide.

Methyl transfer and epimerization in pyochelin biosynthesis

The order of bioassembly, and in particular at which step the methyl transfer occurs during pyochelin biosynthesis, was previously not clearly defined. However, isolates of biosynthetic intermediates during reconstitution assays suggested methyl transfer transpires sometime after condensation with the upstream peptide, contradictory to the hypothesis presented by Garneau-Tsodikova *et al*. To elucidate the order of methyl transfer in pyochelin biosynthesis we purified recombinant variants of PchF, synthesized a series of substrate analogs representing different biosynthetic intermediates, and developed kinetic assays for adenylation and methyltransferase activities. Overall, our data provides evidence that methyl transfer is the penultimate step of pyochelin biosynthesis. After condensation of L-Cys with the upstream peptide, the L-Cys is cyclized to a thiazoline, which is subsequently reduced by NADPH-dependent PchG, and then *N*-methylated by PchF's stuffed methyltransferase domain, before being released through hydrolysis as mature pyochelin. This work provides a foundation that stuffed tailoring domains may be specific towards the nascent peptide and not simply to the aminoacyl-AMP or aminoacyl-Ppant intermediates. Onium chalcogen effects were also used to support that the methyltransferase reaction proceeds through an S_N2 mechanism.

Intriguingly, the stuffed domain of PchE shares sequence homology with methyltransferase domains but is considered to be a non-canonical epimerase domain. Currently, the accepted mechanism in the literature suggests the following order of events.[25] PchE first condenses the upstream salicylate with the downstream Cys_L-Ppant in its cyclization domain

forming HP-Cys_L-Ppant. The peptide then travels to the stuffed epimerase domain where an active site base abstracts a proton from the Cys α-carbon forming a carbanion. Random re-protonation from the solvent occurs altering the stereochemistry to HP-Cys_D-Ppant or regenerating HP-Cys_L-Ppant. The nascent peptide then travels back to the cyclization domain for cyclodehydration of the acyclic Cys to a thiazoline forming a mixture of $HPT_{L/D}$-Ppant. The downstream cyclization domain of PchF is proposed to selectively condense its Cys_L-Ppant with only the D-isomer before moving forward with peptide assembly.

Herein, we proposed a more conventional and energy efficient pathway. Like other known cyclization domains, we hypothesize that condensation and cyclodehydration occur in chronological order forming HPT_L-Ppant. The growing peptide then moves to the stuffed epimerase domain where the chemistry is similar to other *C*-thiazolinyl methyltransferases. An active site acid abstracts a proton from the thiazoline α-carbon forming an enol intermediate. When AdoMet is not present in the active site for methyl transfer, the enol becomes re-protonated by an active site acid or by solvent to generate HPT_D-Ppant, which moves forward in the biosynthetic pathway. This pathway entails only two domain movements (cyclization-epimerase) to perform three chemistries (condensation-cyclization-epimerase), whereas the pathway proposed by Walsh *et al* suggests three domain movements (cyclization-epimerase-cyclization) to perform the same three chemistries.[25]

If stuffed epimerase domains are simply defunct methyltransferase domains due to their inability to bind AdoMet, one could imagine the possibility of engineering the methyltransferase domain into an epimerase domain. Experiments discussed in this dissertation lay a groundwork to continue studies examining the evolution and relation between stuffed methyltransferase and epimerase domains. To examine these hypotheses, potential substrate and product analogs were

synthesized representing the acyclic (HP-Cys_L-Ppant and HP-Cys_D-Ppant) and cyclized (HPT_L-Ppant and HPT_D-Ppant) proposed biosynthetic intermediates. Unexpectedly, especially considering the success of similar analogues in the methyltransferase assays, *in vitro* assays with PchE variants directed towards examining cyclodehydration and epimerization chemistry by the epimerase domains did not yield detectable product. This may be due to the analogs not being similar enough to the natural Ppant substrate but may also indicate the stuffed epimerase domain is non-functional. Other similar stuffed methyltransferase-like domains in other 2-hydroxyphenylthiazoline-thiazolidine natural products biosynthetic pathways are non-functional as methyltransferases or epimerases.[26]

Substrate analogues in the study of NRPS assembly lines

NRPSs are fascinating enzymes generating many diverse natural products. Their impressive chemical logic works in an assembly line-like method, tailoring the nascent peptide and passing it to the next module by means of each post-translationally modified carrier domain through its Ppant tether. While the thiotemplate allows for this methodology, it makes it difficult to study each NRPS module or domain with its native substrate. In this work, we synthesized substrate analogs altering the natural thioester of the Ppant tether to a hydroxyl, carboxyl, and/or ester derivative. While PchF was able to use an ethyl ester mimic, it would not accept the carboxyl as a substrate analog for methyl transfer. Alternatively, PchG was able to reduce a carboxyl analog, but not an ester or hydroxyl analog.[27] If the stuffed epimerase domain of PchE is functional, it was unable to perform epimerase activity with carboxyl, hydroxyl, or ester analogs. Indeed, analogs more similar to the Ppant tether have been generated to study substrate specificity or kinetics of NRPS enzymes such as peptidyl *N*-acetylcysteiamine (SNACs),

peptidyl *N*-acetylethanolamine esters (ONACs), and *N*-acetylethylenediamine amides (NNACs).[28-30] However, several examples exist where peptidyl-SNACs or analogs do not work as substrate mimics and most likely, several more examples have gone unpublished.[31-34] Thus, a more robust systematic approach to studying the chemistry of different NRPS domains is still in need.

Burkart *et al* previously reported a synthetic method to generate pantetheine derivatives and successfully performed a one-pot chemo-enzymatic synthesis attaching the pantetheine derivatives to VibB, a carrier protein in vibriobactin biosynthesis.[35] Coenzyme A (CoA) is biosynthesized from vitamin B5, or pantothenate, by 5 different enzymes, CoAA-E.[35] Previously, Burkart and colleagues discovered CoAA was still able to phosphorylate the 4'-hyroxyl group of pantetheine derivatives *in vivo* and *in vitro*.[35-37] Using this methodology, analogs mimicking nearly the entire Ppant tether can be generated (**Figure 5-1A**). While synthesis of these derivatives would be more similar to the actual substrate, Burkart *et al* also show the ability to post-translationally modify the carrier domain, VibB. Using the synthesized derivatives and CoA biosynthetic enzymes (CoAA, CoAD-E), Burkart and colleagues, generated CoA derivatives with terminal fluorescent probes.[35] Once the CoA derivative is generated, Sfp, a commonly used phosphopantetheinyl transferase (PPTase), was used to post translationally modify the carrier domain of VibB. This methodology can be used to generate intermediate analogs of NRPS biosynthetic pathways, in turn, directing and unifying a module to view a specific chemical step in the pathway (**Figure 5-1B**).[35, 37]

Future work related to the PchE and PchF modules

An important question to answer is whether the lack of activity detected by the stuffed epimerase domain is the result of an evolutionary defunct domain or due to not generating substrate analogues that are effective mimics of the physiologically important substrate. An amino acid coupling reaction can be used with the carboxyl derivatives synthesized and presented in this dissertation to generate more comparable analogs for intermediates in pyochelin biosynthesis, differing only by an amide linkage instead of the thioester linkage of a Ppant tether (**Figure 5-1 A,B**). For example, substrate and product analogs of PchE (HP-Cys_L, HP-Cys_D, HPT_L-CO_2^-, and HPT_D-CO_2^-) can be coupled with the amine-containing, p-methoxyphenyl protected pantothenic acid. After deprotection, newly generated substrate (HP-Cys_L-*N*-pant and HPT_L-*N*-pant) and product (HP-Cys_D-*N*-pant and HPT_D-*N*-pant) analogs can be used in the circular dichroism and chiral chromatography assays described in chapter 3. If epimerase activity still cannot be detected, each *N*-Ppant analog can be further modified by CoAA, CoAD, and CoAE to generate a CoA-derivative that can be transferred to PchE's P2-domain by Sfp, a commonly used Ppant transferase (**Figure 5-1B**). To ensure epimerase activity is being monitored in PchE's stuffed domain, and that Sfp does not transfer the CoA-derivative to PchE's P1-domian, a new PchE variant consisting of the stuffed tailoring-adenylation didomain and the P2-domain should be used (PchE-AEP). After PchE-AEP is modified, it can be purified by size-exclusion chromatography, removing the other enzymes. The modified PchE-AEP's can be incubated with PchF-FL and Cys_L-*N*-pant to transfer the substrate analog to the Cys_L-*N*-pant through PchF's C-domain, which can be isolated, analyzed, and compared to synthesized standards to determine if epimerase activity occurred on the acyclic or cyclic analog (**Figure 5-**

1C). This last proposed experiment may have its pitfalls as post-translational modification and detection after a single turnover experiment may be too hopeful and provide difficulties.

Similar modifications can be used in crystallography experiments. Specifically, PchF can be modified with $HPT_{ox}T_{red}$-*N*-Ppant to unify the recombinant enzyme into promoting methyltransferase chemistry. Crystallography screens, including experiments along with other modular substrates (L-Cys, ATP) and products (AdoHCys) can be used. Successful structure determination would be the first holo-structure of an NRPS module with a stuffed tailoring domain in the conformation for tailoring chemistry. The structure would be important, because it would provide evidence for how the A_{sub}-domain and peptidyl carrier domain migrate in accordance to the Ppant tether to promote the tailoring chemistry. Another interesting crystallography experiment could be co-crystallization of PchG and PchF after modifying PchF's peptidyl carrier domain with $HPT_{ox}T_{ox}$-*N*-Ppant and incubating with $NADP^+$. Structural elucidation of a holo-stand-alone tailoring domain in complex with its *in trans* NRPS partner would also be the first of its kind. If crystallization experiments prove to be too difficult, these large multimodular enzymes would be suitable for cryoEM experiments.

Stand-alone domains of pyochelin biosynthesis

In addition to PchE and PchF experimentation, structural and kinetic studies aimed towards the stand-alone adenylation domain, PchD, and stand-alone tailoring thiazolinyl reductases, PchG and Irp3, were presented in this dissertation. Substrate analogs were synthesized and used for kinetic assays with PchG and allowed determination of the first closed-holo-structure of Irp3 with $NADP^+$. Steady-state kinetics of the adenylation domain of PchD were performed with salicylate. Synthesis of potential warhead inhibitors were co-crystallized

with PchD providing means for rational design of more effective inhibitors of salicylate adenylation domains. This project is currently ongoing in our lab.

The benefit of studying pyochelin biosynthesis

Many natural products biosynthesized by NRPS assembly lines are larger and more complex than pyochelin. Intermediate analogs of pyochelin biosynthesis are relatively easy to synthesize, allowing biochemical and structural analyses of downstream domains more feasible. Therefore, pyochelin bioassembly was exploited as an ideal model to study the complexities involved in NRPS biosynthesis. Overall, the experiments and results in this dissertation provide a groundwork for better understanding structural and biosynthetic aspects in pyochelin biosynthesis, and thus, NRPS enzymes as a whole. Continued studies aim to further elucidate these intricacies.

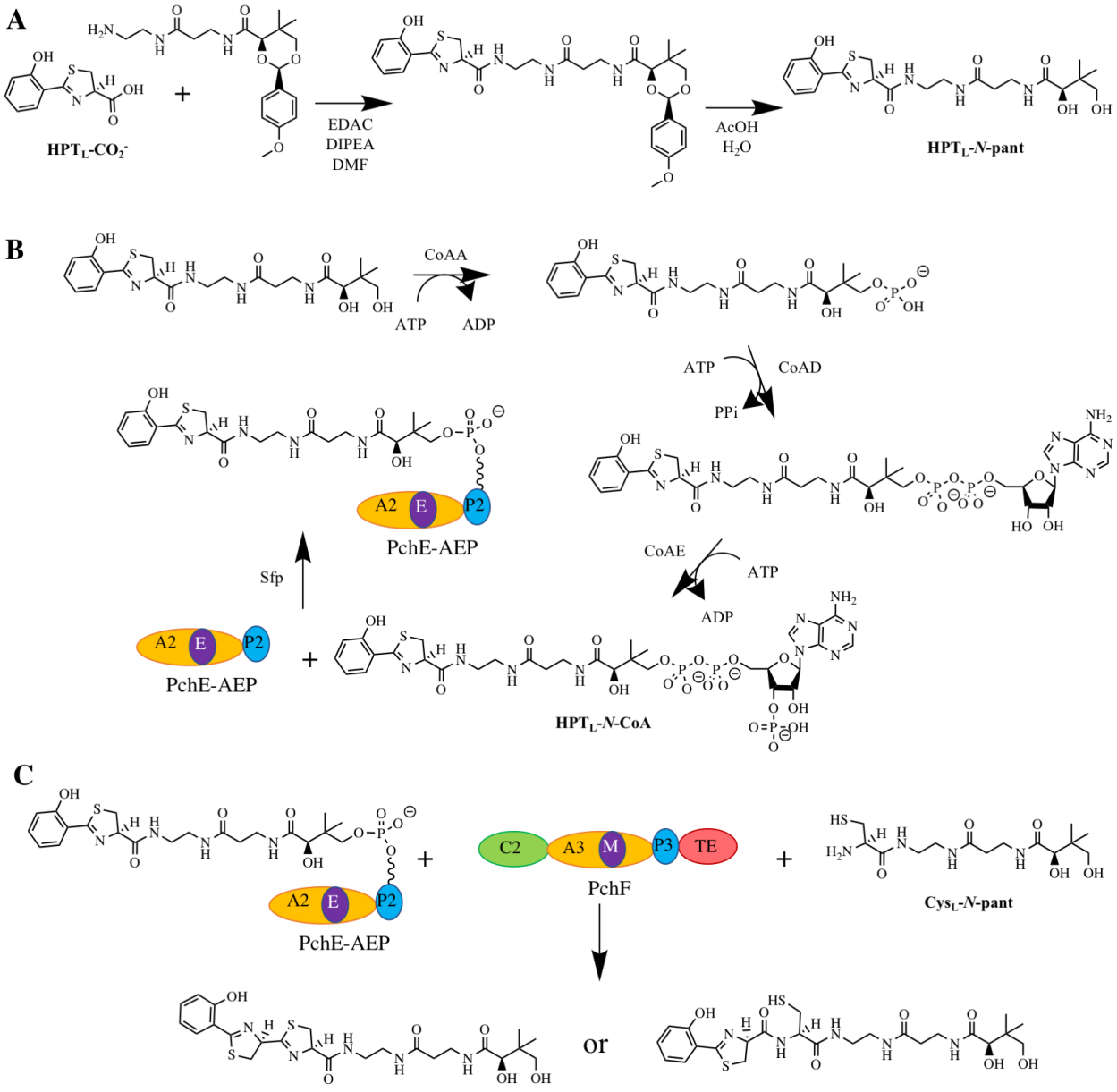


Figure 5-1: Pantetheine and CoA derivatives and uses during in vitro pyochelin NRPS experiments.[35-37] **A)** Synthetic scheme of PchE pantetheine substrate analog, HPT_L-*N*-pant. **B)** Biosynthetic scheme of HPT_L-CoA derivative and post-translational modification PchE-AEP providing a potential intermediate analog of pyochelin biosynthesis. CoAA phosphorylates the 4' hydroxyl of the pantetheine analog. CoAD, a phosphopantetheine adenylyltransferase, and CoAE, a dephosphocoenzyme A kinase, can be used to create the CoA derivative. Sfp, a commonly used phosphopantetheinyl transferase, can catalyze transfer of the Ppant derivative by post-translationally modifying a serine in PchE-AEP's peptidyl carrier domain. **C)** If epimerase activity is not catalyzed using pantetheine substrate analogs, PchE-AEP modified with HPT_L-*N*-Ppant or other substrate analogs can be incubated and added to PchF-FL and Cys_L-*N*-pant to promote condensation and possibly cyclization. The PchF product can be isolated and analyzed

to determine the substrate of the epimerase domain or if chirality is important for downstream condensation.

References

[1] (2013) Antibiotic Resistant Threats in the United States, 2013, Center for Disease Control.

[2] Tacconelli, E. C., Magrini, N.C., . (2017) Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics, World Health Organization.

[3] Miethke, M., and Marahiel, M. A. (2007) Siderophore-based iron acquisition and pathogen control, *Microbiol Mol Biol Rev 71*, 413-451.

[4] Poole, K., and McKay, G. A. (2003) Iron acquisition and its control in Pseudomonas aeruginosa: many roads lead to Rome, *Front Biosci 8*, d661-686.

[5] Otto, B. R., Verweij-van Vught, A. M., and MacLaren, D. M. (1992) Transferrins and heme-compounds as iron sources for pathogenic bacteria, *Crit Rev Microbiol 18*, 217-233.

[6] Wooldridge, K. G., and Williams, P. H. (1993) Iron uptake mechanisms of pathogenic bacteria, *FEMS Microbiol Rev 12*, 325-348.

[7] Braun, V. (2001) Iron uptake mechanisms and their regulation in pathogenic bacteria, *Int J Med Microbiol 291*, 67-79.

[8] Lopez-Medina, E., Fan, D., Coughlin, L. A., Ho, E. X., Lamont, I. L., Reimmann, C., Hooper, L. V., and Koh, A. Y. (2015) Candida albicans Inhibits Pseudomonas aeruginosa Virulence through Suppression of Pyochelin and Pyoverdine Biosynthesis, *PLoS Pathog 11*, e1005129.

[9] Takase, H., Nitanai, H., Hoshino, K., and Otani, T. (2000) Impact of siderophore production on Pseudomonas aeruginosa infections in immunosuppressed mice, *Infect Immun 68*, 1834-1839.

[10] Takase, H., Nitanai, H., Hoshino, K., and Otani, T. (2000) Requirement of the Pseudomonas aeruginosa tonB gene for high-affinity iron acquisition and infection, *Infect Immun 68*, 4498-4504.

[11] Quadri, L. E. (2007) Strategic paradigm shifts in the antimicrobial drug discovery process of the 21st century, *Infect Disord Drug Targets 7*, 230-237.

[12] Reimmann, C., Serino, L., Beyeler, M., and Haas, D. (1998) Dihydroaeruginoic acid synthetase and pyochelin synthetase, products of the pchEF genes, are induced by extracellular pyochelin in Pseudomonas aeruginosa, *Microbiology 144 (Pt 11)*, 3135-3148.

[13] Serino, L., Reimmann, C., Visca, P., Beyeler, M., Chiesa, V. D., and Haas, D. (1997) Biosynthesis of pyochelin and dihydroaeruginoic acid requires the iron-regulated pchDCBA operon in Pseudomonas aeruginosa, *J Bacteriol 179*, 248-257.

[14] Quadri, L. E., Keating, T. A., Patel, H. M., and Walsh, C. T. (1999) Assembly of the Pseudomonas aeruginosa nonribosomal peptide siderophore pyochelin: In vitro

reconstitution of aryl-4, 2-bisthiazoline synthetase activity from PchD, PchE, and PchF, *Biochemistry 38*, 14941-14954.

[15] Reimmann, C., Patel, H. M., Serino, L., Barone, M., Walsh, C. T., and Haas, D. (2001) Essential PchG-dependent reduction in pyochelin biosynthesis of Pseudomonas aeruginosa, *J Bacteriol 183*, 813-820.

[16] Ankenbauer, R. G., and Cox, C. D. (1988) Isolation and characterization of Pseudomonas aeruginosa mutants requiring salicylic acid for pyochelin biosynthesis, *J Bacteriol 170*, 5364-5367.

[17] Cox, C. D., Rinehart, K.L., Moore, M.L. Carter Cook Jr., J. (1981) Pyochelin: Novel structure of an iron-chelating growth promoter for Pseudomonas aeruginosa, *Proc Natl Acad Sci U S A 78*, 4256-4260

.

[18] Cox, C. D., and Graham, R. (1979) Isolation of an iron-binding compound from Pseudomonas aeruginosa, *J Bacteriol 137*, 357-364.

[19] Labby, K. J., Watsula, S. G., and Garneau-Tsodikova, S. (2015) Interrupted adenylation domains: unique bifunctional enzymes involved in nonribosomal peptide biosynthesis, *Nat Prod Rep 32*, 641-653.

[20] Ronnebaum, T. A., and Lamb, A. L. (2018) Nonribosomal peptides for iron acquisition: pyochelin biosynthesis as a case study, *Curr Opin Struct Biol 53*, 1-11.

[21] Al-Mestarihi, A. H., Villamizar, G., Fernandez, J., Zolova, O. E., Lombo, F., and Garneau-Tsodikova, S. (2014) Adenylation and S-methylation of cysteine by the bifunctional enzyme TioN in thiocoraline biosynthesis, *J Am Chem Soc 136*, 17350-17354.

[22] Mori, S., Garzan, A., Tsodikov, O. V., and Garneau-Tsodikova, S. (2017) Deciphering Nature's Intricate Way of N,S-Dimethylating l-Cysteine: Sequential Action of Two Bifunctional Adenylation Domains, *Biochemistry 56*, 6087-6097.

[23] Mori, S., Pang, A. H., Lundy, T. A., Garzan, A., Tsodikov, O. V., and Garneau-Tsodikova, S. (2018) Structural basis for backbone N-methylation by an interrupted adenylation domain, *Nat Chem Biol 14*, 428-430.

[24] Zolova, O. E., and Garneau-Tsodikova, S. (2014) KtzJ-dependent serine activation and O-methylation by KtzH for kutznerides biosynthesis, *J Antibiot (Tokyo) 67*, 59-64.

[25] Patel, H. M., Tao, J., and Walsh, C. T. (2003) Epimerization of an L-cysteinyl to a D-cysteinyl residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from Pseudomonas aeruginosa, *Biochemistry 42*, 10514-10527.

[26] Inahashi, Y., Zhou, S., Bibb, M. J., Song, L., Al-Bassam, M. M., Bibb, M. J., and Challis, G. L. (2017) Watasemycin biosynthesis in Streptomyces venezuelae: thiazoline C-methylation by a type B radical-SAM methylase homologue, *Chem Sci 8*, 2823-2831.

[27] Meneely, K. M., Ronnebaum, T. A., Riley, A. P., Prisinzano, T. E., and Lamb, A. L. (2016) Holo Structure and Steady State Kinetics of the Thiazolinyl Imine Reductases for Siderophore Biosynthesis, *Biochemistry 55*, 5423-5433.

[28] Schneider, T. L., Shen, B., and Walsh, C. T. (2003) Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation, *Biochemistry 42*, 9722-9730.

[29] Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases, *Chem Biol 7*, 765-772.

[30] Tseng, C. C., Bruner, S. D., Kohli, R. M., Marahiel, M. A., Walsh, C. T., and Sieber, S. A. (2002) Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase, *Biochemistry 41*, 13350-13359.

[31] Haynes, S. W., Ames, B. D., Gao, X., Tang, Y., and Walsh, C. T. (2011) Unraveling terminal C-domain-mediated condensation in fungal biosynthesis of imidazoindolone metabolites, *Biochemistry 50*, 5668-5679.

[32] Franke, J., and Hertweck, C. (2016) Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges, *Cell Chem Biol 23*, 1179-1192.

[33] Foulke-Abel, J., and Townsend, C. A. (2012) Demonstration of starter unit interprotein transfer from a fatty acid synthase to a multidomain, nonreducing polyketide synthase, *Chembiochem 13*, 1880-1884.

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

[35] Worthington, A. S., and Burkart, M. D. (2006) One-pot chemo-enzymatic synthesis of reporter-modified proteins, *Organic & biomolecular chemistry 4*, 44-46.

[36] Clarke, K. M., Mercer, A. C., La Clair, J. J., and Burkart, M. D. (2005) In vivo reporter labeling of proteins via metabolic delivery of coenzyme A analogues, *J Am Chem Soc 127*, 11234-11235.

[37] Worthington, A. S., Rivera, H., Torpey, J. W., Alexander, M. D., and Burkart, M. D. (2006) Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis, *ACS Chem Biol 1*, 687-691.