Probing the nonribosomal peptide synthetase epimerization domain via rationally designed suicide inhibitors

Wuo, Michael

2011

Peer reviewed|Thesis/dissertation
Probing the Nonribosomal Peptide Synthetase Epimerization Domain via Rationally Designed Suicide Inhibitors

A Thesis submitted in partial satisfaction of the requirements for the degree
Master of Science

in
Chemistry

by
Michael Wuo

Committee in Charge:

Professor Michael Burkart, Chair
Professor William Gerwick
Professor Yitzhak Tor

2011
The Thesis of Michael Wuo is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

__________________________________________

__________________________________________

__________________________________________

__________________________________________

Chair

University of California, San Diego

2011
DEDICATION

To my family and friends for their support and guidance.
EPIGRAPH

Simplicity is embarrassing when you have to work for months to achieve it.

*Kary Mullis*
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature Page</td>
<td>iii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Epigraph</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>List of Schemes</td>
<td>viii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>ix</td>
</tr>
<tr>
<td>Abstract</td>
<td>x</td>
</tr>
<tr>
<td><strong>Chapter One</strong> Modular Biosynthesis of Nonribosomal Peptides</td>
<td>1</td>
</tr>
<tr>
<td><strong>Chapter Two</strong> Chemical Probes of the Epimerization Domain</td>
<td>14</td>
</tr>
<tr>
<td><strong>Chapter Three</strong> Biochemical Assays for Inhibition</td>
<td>27</td>
</tr>
<tr>
<td>References</td>
<td>34</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Notable nonribosomal peptide natural products........................................ 2
Figure 2. Monomer unit loading in yersiniabactin biosynthesis.................................. 3
Figure 3. Catalytic domains of NRP biosynthesis...................................................... 4
Figure 4. Different conformers of the peptidyl carrier protein.................................. 5
Figure 5. Reaction of the adenylation domain.......................................................... 6
Figure 6. Modular biosynthesis of NRPs................................................................. 7
Figure 7. Reaction of the epimerization domain....................................................... 8
Figure 8. Mechanism of diketopiperazine formation.................................................. 9
Figure 9. X-ray crystal structure of the epimerization domain.................................... 10
Figure 10. Domain swapping in combinatorial biosynthesis...................................... 11
Figure 11. Variety of chemical probes................................................................. 15
Figure 12. Covalent crosslinking with chloroacrylate............................................. 16
Figure 13. Inhibition mechanism of alanine racemase with β-fluoroalanine.............. 17
Figure 14. Inhibition mechanism of alanine racemase with β-chlorovinylglycine....... 18
Figure 15. Inhibition mechanism of E domain with β-chlorovinylglycine................... 19
Figure 16. Inhibition mechanism of E domain with β,γ-phenylethynylglycine.......... 20
Figure 17. Inhibition of diketopiperazine formation.................................................. 28
Figure 18. Adenylation domain kinetics................................................................... 30
LIST OF SCHEMES

Scheme 1. Synthesis of β-chlorovinylglycine ........................................... 20
Scheme 2. Retrosynthetic route of β,γ-phenylethynylglycine .................. 22
Scheme 3. Model synthesis of β,γ-phenylethynylglycine ....................... 23
Scheme 4. Synthesis of β,γ-phenylethynylglycine ................................. 23
Scheme 5. ATP-pyrophosphate exchange assay ...................................... 29
ACKNOWLEDGMENTS

First and foremost, I would like to thank Professor Mike Burkart for providing me with the opportunity to engage in both exciting research and insightful discussion during my undergraduate and graduate career. He along with all the members of the Burkart lab have provided assistance and support over the course of my study. I would especially like to thank Dr. Joris Beld who has been more than a friend, but also a mentor, in and out of lab; I cannot thank him enough for laughing at me during my first group meeting and laughing with me during my last.

I have had the opportunity of taking chemistry courses from three of the Burkart lab’s members. I would like to thank Jill Blatti (143b), Gene Hur (143c), and Bob Haushalter (114a) for their efforts in providing an enjoyable learning experience during my undergraduate studies. To the other members of the Burkart lab, I also thank you for everything. From Okan to tennis, from failed experiments to reddit, thank you!

I would like to mention the Department of Chemistry and Biochemistry for providing financial support throughout my graduate studies. I am especially appreciative of Jeff Rances for providing me with both academic guidance and friendship.

Thank you to my family and friends who have been there with love and support along every step of the way. And last, but certainly not least, thank you Patty Lee for barking with me through the darkest and brightest of times.
Probing the Nonribosomal Peptide Synthetase Epimerization Domain
via Rationally Designed Suicide Inhibitors

by

Michael Wuo

Master of Science in Chemistry

University of California, San Diego, 2011

Professor Michael Burkart, Chair

Nonribosomal peptides (NRPs) represent a class of biologically active natural products that contain a great deal of functional and therapeutic diversity. They are produced by both prokaryotes and eukaryotes from large, multidomain megaenzymes called nonribosomal peptide synthetases (NRPSs). Many of these NRPs contain D-amino acids that provide a great deal of physiological importance to their functionality.
Incorporation of D-amino acids into the NRP natural product often occurs through integrated epimerization (E) domains that catalyze racemization of the Cα position of a PCP-tethered L-amino acid. Several small molecule probes have been rationally designed in order to investigate the mechanism as well as gain insight into intradomain communication of E domains. β-Chlorvinylglycine 10 and β,γ-phenylethynylglycine 15 are proposed to form a reactive allene intermediate in situ upon catalysis of the E domain that can be rapidly attacked by an active-site base. Utilizing the diketopiperazine condensation assay, we have demonstrated that probe 10 acts as an inhibitor of the E domain of the initiation module, grsA, of the gramicidin S synthetase. Inhibition of E domain activity suggests covalent crosslinking of the E domain and its proximal PCP partner, allowing insight into protein interactions that regulate NRP biosynthesis.
Natural products are considered a large source of inspiration for drug discovery and design due to their unique chemical moieties and diversity. The National Cancer Institute reports 61% of newly discovered drugs from 1981-2002 are derived from natural products.\textsuperscript{[1]} Specifically, these bioactive molecules offer functional diversity that has allowed natural products to thrive in the competitive field of drug discovery. Within the vast family of medicinal natural products exists a sub-class known as the peptidyl natural products. These peptide-based natural products are mostly nonribosomally synthesized, despite exceptions that utilize the ribosomal machinery such as those of the lantibiotic family.\textsuperscript{[2]} Biosynthetically, nonribosomal peptides (NRPs) are unbound by the laws that dictate ribosomal peptide formation as exhibited by their prolific incorporation of nonproteinogenic amino acids. Stemming additionally from the proteinogenic amino acid building blocks found within their backbones, NRPs achieve functional and therapeutic diversity ranging from the antibiotic penicillin, vancomycin, and surfactin families to the iron-chelating siderophores of the enterobactins and yersiniabactins. Additionally, the NRP cyclosporin acts as an immunosuppressant in post-allogeneic organ transplant to reduce rejection (Fig. 1).
Various pathways have been discovered that reveal the method in which these metabolites are biosynthesized\textsuperscript{[3]}\textsuperscript{[3]}\textsuperscript{[3]}. Modular synthase biosynthetic clusters such as the polyketide (PKS), fatty acid (FAS), and nonribosomal peptide synthases\textsuperscript{[4]} (NRPS) share the homologous, assembly-line-like construction of elongated monomer units, either acetate or amino acid, to afford small molecule natural products; in each instance, the substrates are loaded onto large multifunctional modular enzymes, which are further divided into discrete domains that catalyze substrate condensation and modification (Fig. 2). Previously, the catalytic domains within NRPSs were believed to be covalently-linked forming large megaenzymes (type I), however, work by Shen and coworkers...
revealed the existence of a type II architecture in which each catalytic domain is located within individual proteins.\textsuperscript{[5]}

\textbf{Figure 2.} Chain elongation during yersiniabactin biosynthesis. The covalently tethered substrate, cysteiny1-S-P CP, is channeled to the corresponding upstream acyl acceptor, salicyl-S-A r CP forming salicyl-cysteiny1-S-P CP. A ry1 carrier protein; Cy: cyclization domain; A: adenylation domain PCP: peptidyl carrier protein

Winnick and coworkers provided preliminary insights into how these NRPs are biosynthesized by extracting gramicidin S, a cyclic decapptide antibiotic, out of crude cell lysates of \textit{Bacillus brevis} with ethanolic HCl. The cell medium was spiked with \textsuperscript{14}C-labeled amino acids found within the gramicidin S backbone and incubated over the course of 44 hours, which yielded 2-4\% radiolabeled gramicidin, thereby providing the first evidence for whole amino acid incorporation into small molecules. Nearly a decade later, the modular biosynthesis of these NRPs was first suggested as a polyenzyme template modeled after the PKS and FAS pathways\textsuperscript{[6]}, whose biosynthesis had been previously investigated.\textsuperscript{[7]} Lipmann and coworker's previous work with the amino acid activation by the tRNA synthase provided facile recognition of similarities between the activation exhibited in ribosomal protein synthesis and nonribosomal peptide synthesis.
They soon discovered that a 70-75 kDa aggregate of enzymes could activate different amino acids within each NRP. Stemming from this pioneering research, each catalytic domain within the NRPS biosynthetic cluster was elucidated. (Fig. 3).

**Figure 3.** Catalytic domains found within the NRPS biosynthetic cluster. Essential domains for chain elongation and substrate release are highlighted in black. Auxiliary domains are highlighted in white.

Granicidin S synthetase became a template for NRPS research. The mechanisms for amino acid activation and substrate channeling were elucidated on the N-terminus initiation module, PheATE (grsA), of the gramicidin S synthetase. Similarities between modular synthesis of the FAS and NRPS allowed for the proposal of similar substrate transfer. The 4'-phosphopantetheine linker attached to a conserved serine residue on the acyl carrier protein (ACP) of FAS biosynthetic pathway is integral in elongation and modification. The 4'-phosphopantetheine moiety is post-translationally transferred from the enzyme cofactor, CoenzymeA, by a phosphopantetheinyll
transferase. Upon loading the phosphopantetheine prosthetic arm, the ACP moves from a state known as apoCP to holoCP. Similarly, NRPS systems have a substrate guide known as the peptidyl carrier protein (PCP). The PCP structure is highly homologous to that of ACP and additionally offers different structural conformations in the apo and holo state. However, within the PCP domain lies another structural conformation known as the Apo/Holo (A/H) state\[10]\ (Fig. 4).

![Apo State, A/H State, Holo State](image)

**Figure 4.** NMR solution structures of tycC3-PCP conformations in the A, A/H, and H states.

The A/H state exists as an equilibrium structural conformation between the apo and holo state. Moreover, detailed analysis of \[^{15}\text{N}, 1\text{H}\] Transverse relaxation optimized spectroscopy (TROSY) and nuclear Overhauser effect spectroscopy (NOESY) allowed for the determination of this flexible yet highly stabilized A/H state. This newly discovered A/H state is believed to be that in which intradomain substrate shuttling occurs, however, it is only when the PCP is in its holo state can substrate attachment occur via the phosphopantetheine’s terminal sulfhydryl group.\[^{11}\] Since a thioester linkage is formed between the loaded substrate and the phosphopantetheine functionality, the PCP is
often dubbed the thiolation domain. Thioester formation is a low-energy transition from the activated adenosine monophosphate mixed anhydride amino acid constructed by the adenylation (A) or activation domain[12] (Fig 5). The aminoacyl-AMP ester exists as a thermodynamically activated and kinetically

\[ \text{Figure 5. Initiation module PheATE of the gramicidin S synthetase serving as a template for ATP-Mg}^{2+} \text{-dependent adenylation and formation of L-phenylalanyl-S-PCP.} \]

labile species in the activate site of the A-domain promoting a favorable transformation for substrate loading. This low energy transition of adenylation is analogous to the mechanism of activation exhibited by the tRNA synthase, which activates the carboxylate end of amino acids with the utility of adenosine triphosphate for a covalent tether to the 2' or 3'-OH of the terminal A residue of the cognate tRNA.[13]

After, each amino acid substrate is activated and loaded onto the PCP, the upstream (N-terminal) module is then primed to undergo peptide formation by a downstream (C-terminal) C domain with the utility of its cognate amino acid. Although NRPS have catalytically facilitated peptide synthesis, it is the additional modifying domains that make these NRPs structurally unique (Fig. 6). Each auxillary domain introduces specific alterations such as cyclized amino acids, methylations, formylations,
Figure 6. Biosynthesis of the branched, cyclic, dodecapeptide bacitracin A. In this Type I NRPS, twelve modules are responsible for the production of this antibiotic natural product.

halogenations, glycosylations, and epimerizations that add to the functional diversity and sustainability of each NRP.

The 50 kDa epimerization (E) domains are auxiliary proteins responsible for altering the stereochemistry at the Cα position of an amino acid (Fig. 6). E domains can be found in two places within an NRPS and have thus been divided into two families, those found within the N-terminus initiation modules Ei and those encoded downstream in the biosynthetic assembly line within elongation modules, Ee. Although their genetic identities are homologous, they differ in that Ee epimerizes the covalently attached substrate only after upstream condensation has gone to completion. Moreover, as E
domains generally act as a natural racemases, their substrate D and L isomers exist nearly 2:1, respectively (Fig. 7). The isomer composition was determined using

![Figure 7. Epimerization at the Cα position of phenylalanyl-S-PCP.](image)

radiolabeled $^{14}$C L-Phe to spike the gramicidin S synthetase and incubated for 30 minutes. The megaenzyme was subsequently denatured by boiling and treated with D-amino acid oxidase to convert D-Phe to phenylpyruvic acid. A 2,4-dinitrophenylhydrazine derivative of the radiolabeled phenylpyruvic acid; subsequent analysis afforded the ratio 7:3 of D-Phe to L-Phe, which demonstrated general preference for the isomer incorporated in its native structure. Despite the E domains stereoselective nature, the downstream condensation domain discriminates against the non-native isomer of the growing peptidyl chain.

Selectivity of the C domain and activity of the E domain can be assayed by pairing the intiation module grsA with a homologous downstream ProCAT module of the tyrocidine biosynthetic pathway. In this experiment, a diketopiperazine (DKP) of D-Phe and L-Pro is formed stereospecifically; The stereopreference for D,L-DKP over L,L-DKP is observed at about 60:1, respectively. Experimentally, each cognate structure is incubated with its respective module and incubated with ATP separately for three minutes, assuring complete substrate loading. Combining the two reaction mixtures
would yield the DKP product over a kinetic time course. In order to assemble DKP, the C domain must first condense the two amino acids by forming a peptide linkage between the N-terminus of L-Pro and the C terminus of D-Phe (Fig. 8). Subsequent deprotonation yields the free-base amino acid, which is primed to cyclize and autocatalytically release\textsuperscript{[17]}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8}
\caption{Mechanism of D,L Diketopiperazine (DKP) formation and release.}
\end{figure}

In general, identification of a D-isomer in a NRP backbone allows for facile prediction and location of E domains within a biosynthetic gene cluster. Their ease of identification coupled with their interesting bioactivity has allowed them to be the target of structural and functional research. Although many efforts have been made on the tridomain, grsA, to determine the mechanism of epimerization, a conclusive mode of action has yet to be reached. Extensive biochemical and structural analysis ranging from mutational analysis to deducing substrate specificity have been utilized to study the mechanism of epimerization.\textsuperscript{[18]} Sequence blasting of E domains demonstrates that 22 residues across the approximately 450 amino acid E domains are conserved. However, only a few residues have been identified to be involved in the catalytic mechanism of epimerization. Previous work surmised that the mechanism involves the deprotonation
and reprotonation by a single cysteine residue\textsuperscript{[19]}, however, this cysteine residue was not conserved in the known racemase region and was discounted thereafter for single-base mechanisms were generally found within the family of PLP dependent racemases.\textsuperscript{[20]}

Vater and coworkers later suggested that several highly conserved bases in the putative racemase region were possible including the arginine, histidine, and tyrosine residues found within several conserved motifs of the E domain.\textsuperscript{[21]} Subsequent structure-function analysis was performed by Walsh and coworkers. They identified several residues that dramatically affect epimerization when mutated. These residues are highly conserved across E domains and contain the sequence HHxxxDxxSW (Fig. 9).

**Figure 9.** Crystal structure of homologous E domain within the initiation module TycA. The catalytic residues His753, D757, and Tyr976 are highlighted within the active site.

Furthermore, while the histidine and aspartate residues of this "histidine motif" showed reduced activity when mutated, the proximal tyrosine residue initially proposed by Vater and coworkers did not exhibit a dramatic decrease in catalytic activity when mutated, suggesting that acid-base charge transfer occurs via a one base mechanism. This result
largely disproves the notion that non-PLP dependent epimerization follows a two-base mechanism.

Mechanistic studies on modular synthases have seen a revolution within the last 25 years. Understanding the assembly-line-like fashion of monomer unit manipulation and condensation showed promise in the engineering of novel natural products. Previously, the field relied upon the total synthesis of natural products that were often financially unproductive and inefficient. With the advent of bioengineering, however, manipulation of the biosynthetic gene clusters could yield not just existing natural products at a catalytic rate, but could biosynthesize novel molecules with increased potency and broader bioactivity. Combinatorial biosynthesis began with a mix-and-match approach, integrating pieces of separate assembly lines (Fig. 10).

Figure 10. Combinatorial Biosynthesis. Swapping in heterologous A domain into biosynthetic pathway in order to generate a new peptidyl natural product.
However, this burgeoning field was soon hit with the realization that modules could not be easily manipulated.\textsuperscript{[24]} Leadlay and coworkers reported the swapping of the acyltransferase (AT) domain of the first module of DEBS1-TE with a non-native AT domain of the rapamycin producing PKS. This experiment yielded a novel polyketide molecule, however, was not met with appreciable yield. It was later deduced that protein-protein interactions were an integral part in modular biosynthesis.

Communication-mediating (COM) domains\textsuperscript{[25]} were recognized as small 15-20 amino acid sequence region existing between the C terminus of the donating NRPS and the N terminus of its cognate downstream partner.\textsuperscript{[26]} Stachelhaus and coworkers reported that deletion of random residues found within COM domains leads to no product formation. This research was later supported by Burkart and coworkers who utilized the copper catalyzed azid-alkyne cycloaddition to demonstrate COM domain interaction; cognate domain interaction demonstrated a covalently cross-linked protein, whereas domains without their cognate pair lacked the covalent tether.\textsuperscript{[27]} Uncovering these COM domains became an integral portion of understanding the protein-protein interactions necessary for intermodular communication.

Since the discovery that protein-protein interactions dictate natural product biosynthesis, genetic engineering has been repeatedly demonstrated.\textsuperscript{[28]} Baltz and coworkers were able to successfully increase the antibacterial activity of the NRP, daptomycin, by swapping out CAT modules within its native NRPS.\textsuperscript{[29]} Despite ongoing advancements in the field of combinatorial biosynthesis, however, engineering of E domains into NRPSs has yet to be shown. The introduction of non-cognate E domains into NRPSs could not only increase the potency and selectivity of the pathway’s cognate natural product, but also help combat proteolysis \textit{in vivo}.\textsuperscript{[30]} However, only upon full understanding of the E domain can NRPSs be effectively engineered. The mechanism of
epimerization has yet to be solidified, and along with the E domain's functional
significance lacks proper biochemical analysis on the protein-protein interactions of the
PCP and E domains. Therefore, our lab has since developed a panel of inhibitors in an
ongoing effort to understand the mechanism of epimerization through covalent
crosslinking. Several suicide inhibitors have been rationally designed that probe the
interdomain communication of the PCP and E domains.
Chapter 2

Chemical Probes of the Epimerization Domain

Chemical probes have been integral in the structural and mechanistic elucidation of NRPS domains. They have been used extensively to examine interdomain communication, substrate specificity and selectivity, as well as structural aspects of modular NRPSs.\[^{31}\] These probes can be introduced into modular synthases several ways, one of which requires the use of the Coenzyme A (CoA) biosynthetic pathway to convert synthetic pantetheine analogues into modified CoA derivatives.\[^{32}\] CoA is prokaryotically produced by five enzymes, CoAA-CoAE, however, in eukaryotes, CoA biosynthesis occurs through a fused CoAD and CoAE enzyme that combines the phosphopantetheine adenylyl transferase and dephosphoenzyme A kinase activities into a discrete protein. Taking advantage of the biosynthetic pathway’s substrate promiscuity, small reactive moieties can be synthetically attached to pantetheine to afford a modified probe, which in turn can be introduced into the CoA biosynthetic pathway. Furthermore, the longevity of such probes can be increased by substituting the terminal sulfhydryl group with a primary amine to afford an amide linkage upon modification.\[^{33}\] Subsequent usage of a phosphopantetheinytransferase (PPTase) enzyme, provides promiscuous transfer of pantetheine analogues to the carrier protein’s
active site serine; the attachment of 4'-phosphopantetheine analogues provides the conformational change from apoCP to "crypto"CP.\textsuperscript{[34]} Our lab has demonstrated attachment of numerous chemical probes with various functionality ranging from fluorescence to suicide inhibition\textsuperscript{[35]} (Fig. 11).

\begin{figure}
\centering
\includegraphics[width=\textwidth]{chemical_probes.png}
\caption{Various chemical probes used to identify structural and functional properties of catalytic domains.}
\end{figure}

As carrier proteins are actively involved in substrate shuttling and interdomain communication, they are a viable tool for biochemical assays. Choroacrylic-pantetheine 1 has been integral in deducing protein-protein interactions between ACPs and its partner ketosynthase (KS) domains involved in chain elongation within FAS and PKS
modular biosynthesis. The α,β unsaturation and chlorine leaving group found at the reactive end of 1 allows for a suitable nucleophilic 1,4 Michael addition by an active site cysteine residue of the KS domain; subsequent elimination of chlorine yields a cross-linked KS-ACP species. This covalently cross-linked species allows for facile isothermal titration calorimetry (ITC) between PCP and cognate KSs and non-native partners, demonstrating high binding affinity or protein-protein interactions between cognate partners[54] (Fig. 12).

![Diagram](image)

**Figure 12.** Covalent crosslink of the ACP and KS domains by chloracryllic pantetheine.

In a separate experiment, the functional warhead, 2, was covalently attached to the PCP domain in the same chemoenzymatic fashion probing NRPS non-native domain communication, however, no cross-linked PCP-KS species was observed. Presumably, the non-native PCP does not have the specific recognition residues that are integral for protein-protein interactions.

Since the discovery that modular biosynthesis is avidly dependent upon protein communication, small molecule probes have been consistently used to investigate interdomain crosstalk. Understanding protein interactions has advanced the field of combinatorial biosynthesis into generating novel natural products with increased potency.[28] Despite success in genetic engineering, however, furnishing E domains into
modular NRPSs has yet to be shown. Rationally designed chemical probes could reveal mechanism of epimerization as well as interdomain protein communication that would give insights into genetically engineering E domains into non-native NRPSs. The installation of non-cognate E domains into NRPSs could not only increase the therapeutic potential of engineered natural products, but also help combat proteolytic digestion of NRPs.

Previously, our lab synthesized β-fluorophenylalanine 1 in order to investigate the mechanism of epimerization. The PCP of the initiation module, grsA, was covalently labeled with 2 in efforts to covalently cross-link the PCP and E domains, however, no cross-linked species was observed. Moreover, utilizing the DKP assay, 1 showed negligible inhibition of DKP formation, demonstrating that 1 is not a suitable inhibitor of the E domain. Thornberry and coworkers previously surmised that β-fluoro amino acids did not exhibit inhibitory effects on the pyridoxal phosphate-dependent (PLP) (Fig. 13).

![Figure 13. Inhibition mechanism of PLP-dependent alanine racemase with β-fluoroalanine.](image-url)
alanine racemases due in part to the minimal reactivity of the nucleophilic acrylate that is generated in situ. They reported that inhibition occurs 1 out of 800 times, assuming that diffusing of the aminoacrylate moiety to the enzyme surface is more favorable than attacking pyridoxal-imino moiety in which the β-halo amino acids are covalently attached. Thornberry and coworkers further described the use of β-chlorovinylglycines 10 as mechanism-based inhibitors of alanine racemases. The mechanism of inhibition is similar to that of the β-fluoroalanine, except differs in the reactivity of the intermediate. Following deprotonation of the Cα proton, alanine racemase eliminates the chloride leaving group of 10, affording the enol-carbanion stabilized allene intermediate. Furthermore, the allene motif is a highly electrophilic species that readily reacts with a proximal enzyme base to form a covalent linkage in alanine racemases (Fig. 14).

Figure 14. Inhibition mechanism of PLP-depedent alanine racemase with β-chlorovinylglycine.
Being that β-halo amino acids are unsuitable inhibitors for both alanine racemases and E domains, we proposed that if β-halo amino acids are unsuitable inhibitors for alanine racemases, they will in turn show inhibitory effects for E domains. Although E domains within the NRPS biosynthetic pathways have been found to be PLP independent, a similar method of racemization is proposed.\textsuperscript{[15]} Substrate of E domains exist generally in a 2:1 mixture of entatiomers, and as this is finding is in agreement with an equilibrium mixture of isomers observed in natural racemases, we proposed that if 10 is a suicide inhibitor of alanine racemase, then it should also be an inhibitor of the E domain. Analogous deprotonation at the \( \text{C}_\alpha \) position would yield an allene intermediate that would undergo subsequent attack by a nucleophilic base. This suicide inhibitor could then be regarded as a tool to investigate both mechanism of epimerization and protein-protein interactions of the E and PCP domains (Fig. 15).

**Figure 15.** Inhibition mechanism of E domain with β-chlorvinylglycine.

Synthesis of 10 has been previously described in detail.\textsuperscript{[37]} The synthesis of 10 begins with Z-methionine-OH 3. Subsequent methyl protection at the carboxylate moiety with a 3.5% HCl-methanol mixture afforded the methyl ester 4. Compound 4 was then subjected to oxidative conditions (NaIO\(_4\) in methanol), converting the sulfur functional group to the corresponding sulfoxide 5. Afterwards, Kugelrohr distillation was utilized
for pyrolysis of compound 5 to afford a highly viscous yellow residue containing several isomers as described by Rapaport and coworkers. However, the pyrolysis yielded a crude mixture of varying amounts isomers including the cis and trans β,γ olefins and α,β olefin that were not in correspondence with the percentage reported. The Z-vinylglycine-methylester 6 was subsequently ordered as pure starting material. Dropwise addition of phenylselenylchloride across the olefin motif afforded 7. Subsequently, oxidation of the selenium ether yielded the corresponding selenoxide 8; the di-oxo species was observed under LC-MS analysis of the reaction. Compound 8 was then refluxed in 1,2 dichloroethane to afford the eliminated species and form the protected β-chlorovinylglycine 9. Subsequent reflux in HCl provided the deprotected amino acid 10 (Scheme 1).

Scheme 1 Synthetic route for β-chlorovinylglycine.
Since probe 10 relies on the reactivity of the allene motif to form a covalent crosslink between the PCP and E domains, another probe was designed that mimics more closely grsA's natural substrate, phenylalanine. We began the design and synthesis of the amino acid probe, β-phenylethynylglycine 15. Compound 15 was designed after previous chemical probes synthesized in our lab.[39] Prior work investigating the dehydratase (DH) domain of the FAS biosynthetic pathway required the design of probes to examine the role that DH domains play in isomerizing unsaturated fatty acids. In this work, the mechanism of isomerization of FabA, the prototypical DH enzyme, was examined using suicide substrates that covalently crosslink ACP-DH partner domains. A panel of small molecule probes were designed with β,γ-alkyne motifs that generate an electrophilic allene upon deprotonation by a His residue within the DH's active site. This reactive allene mechanism is analogous to our proposed mechanism for the inhibition of E domains with 10. (Fig. 16).

**Figure 16.** Inhibition mechanism of E domain with β-phenylethynylglycine.

Therefore, utilizing the the same functional motif, we designed β-phenylethynylglycine 15 (Scheme 2).
Synthesis of 15 in its free amino acid state had previously never been achieved.\textsuperscript{[40]} The reactivity of the conjugated alkyne with the aromatic ring system, affording racemization and tautomerization to reactive α,β-dehydro amino acids has often complicated the synthesis of this small molecule probe. We therefore proposed a model synthesis of 15 beginning with the dibenzyl-protected-tartaric acid 11. Starting material 11 was subjected to a lemieux-johnson oxidative cleavage between the 1,2-diol motif affording the benzyl-protected glyoxylate 12. We then used the primary amine of p-anisidine to form an imine with the terminal aldehyde of 12 pushing equilibrium towards imine formation and generating the α-imino ester 13. Utilizing the silver-catalyzed addition of phenylacetylenes developed by Chan and coworkers, the di-protected ethynylglycine 14 was procured\textsuperscript{[41]} (Scheme 3). However, efforts to deprotect the following species and obtain 15 were not examined for we were determining whether the chemistry to obtain the protected species from available and inexpensive resources was possible.

We previously assessed that deprotection of the benzyl group requires acidic
Scheme 3. Model synthetic route for β-phenylethynylglycine

conditions that would harm the fragile alkyne motif. Subsequently, we designed a new synthetic route that began with L-(+)-tartaric acid (Scheme 2.3).

Scheme 4. Synthetic route for β-phenylethynylglycine
Acid-catalyzed allyl alcohol addition would then afford the diallyl-protected tartaric acid 18. Under the same previously employed oxidative conditions, we cleaved the carbon-carbon bond between the 1,2 diol of the protected tartaric acid 19 affording the allyl glyoxylate 20. The next step involves the convergent use of allyl carbamate 17 prepared from the addition of ammonium hydroxide to allyl chloroformate 16. Subsequent imine formation occurred by kinetically controlling the equilibrium of 17 and 20 to afford the \( \alpha \)-imino ester 21. Silver nitrate catalytic addition phenylacetylene yields the diallyl-protected ethnylglycine 22.

Materials and Methods

Synthesis of \( \beta, \gamma \)-Phenylethynylglycine model

*Synthesis of benzyl glyoxylate.* To a stirring solution of dibenzyltartrate 11 (100 mg, 0.3 mmol, 1 eq) in 5mL THF, a catalytic amount of OsO\(_4\) (30 \( \mu \)L, 0.3 \( \mu \)mol, 0.01 eq) was added to the reaction flask. Subsequently, NaIO\(_4\) (257 mg, 1.2 mmol, 4 eq) was dissolved in 5mL of water and added dropwise over 30 min at 25 °C to the stirring dibenzyltartrate solution. The reaction was allowed to stir at 25 °C for an additional hour. The resulting solution was extracted with EtOAc/Hex and dried over Na\(_2\)SO\(_4\) and concentrated under reduced pressure to afford 12 as a white solid (Yield: 97 mg, 98%) R\(_f\)=0.38 (1:1 EtOAc:Hex) \(^1\)H NMR (500 MHz, CDCl\(_3\)): \( \delta = 9.39 \) (s, 1H, CHO), 7.26-7.35 (m, 5H, C\(_5\)H\(_5\)), 5.24 (s, 1H, CH\(_2\))
Synthesis of N-PMP-protected α-imino benzyl glyoxylate. To a stirring solution of benzyl glyoxylate 12 (97 mg, 0.6 mmol, 1 eq) in 2 mL dry CH$_2$Cl$_2$, a solution of p-anisidine (74 mg, 0.6 mmol, 1 eq) in 2mL dry CH$_2$Cl$_2$ was added. The reaction mixture was stirred under argon at 25 °C for 1 hour, and then 4 Å molecular sieves (1 g) were added. After stirring for an additional 2 hours at 40 °C, the reaction mixture was filtered under argon and concentrated under reduced pressure affording 13 as a nearly pure viscous red oil (Yield: 158 mg, 98%) $^1$H NMR (500 MHz, CDCl$_3$): δ = 7.98 (s, 1H, N=CH), 7.46 (d, 2H), 6.95 (d, 2H), 5.27 (s, 2H, CH$_2$), 3.74 (s, 3H, OCH$_3$)

Synthesis of N-PMP-protected α-phenylethynyl benzyl glycine. To a stirring solution of α-imino ester 13 (79 mg, 0.3 mmol, 1 eq) in 5 mL of dry toluene, AgNO$_3$ (2.5 mg, 0.015 mmol, 0.05 eq) camphorsulfonic acid (7mg, 0.03 mmol, 0.1 eq), and phenylacetylene (65 μL, 0.6 mmol, 2eq) were added and stirred under argon at 25 °C overnight. The reaction was filtered through celite and concentrated under reduced pressure to afford the crude produced 14 as a brownish oil (Yield: 11mg, 10%) $^1$H NMR (500 MHz, CDCl$_3$): δ = 7.36 (m, 5H, C$_5$H$_5$), 6.69 (d, 2H), 5.39 (s, 2H, CH$_2$), 4.61 (s, 1H, CH), 3.75 (s, 3H, OCH$_3$)

Synthesis of β,γ-Phenylethynylglycine

Synthesis of allyl carbamate. To a stirring solution of allyl chloroformate (8 ml, 75 mmol, 1 eq) in 50 mL of toluene at 0 °C, 32% aq. NH$_3$ (3 mL, 75 mmol, 1 eq) was added dropwise over 6 minutes. The reaction was stirred at 25 °C for an additional hour. The organic phase was then decanted and dried over Na$_2$SO$_4$. Subsequent concentration under reduced pressure afforded allyl carbamate 17 as a colorless liquid (Yield: 6 g,
80% \textsuperscript{1}H NMR (500 MHz, CDCl\textsubscript{3}): \( \delta = 7.35 (b, 2H, NH\textsubscript{2}) \), 5.84-5.91 (m, 1H, CH=\text{C} \)), 5.16-5.19 (dd, 1H, CH=\text{C} \)), 4.52 (d, 1H, CH\textsubscript{2})
Chapter 3

Biochemical assays for Inhibition

The initiation module, grsA, of the gramicidin S synthetase contains the necessary catalytic domains to assay for amino acid activation and carrier protein attachment. A domain activity is examined using the general ATP-pyrophosphate exchange assay that checks for activity and specificity in addition to proper folding of recombinant proteins.\[^{31}\] For functional characterization of the grsA's C-terminal E domain, the elongation module, TycB1, can be examined. The TycB1 elongation module interacts with the C-terminal end of grsA specifically as a non-native-downstream partner to catalytically synthesize the dipeptide Phe-Pro-S-PCP. This formation terminates \textit{in vitro} autocatalytically releasing the stereoperfered D,L-DKP.\[^{16}\] Subsequent, substrate extraction followed by high pressure liquid chromatography (HPLC) analysis allows for facile determination of E domain activity. We suspected that inhibition of E domain activity using these established assays allows for identification of active probes. We examined the inhibition of the DKP condensation assay by first incubating grsA at 37°C with probe 10 and ATP for a 3, 6, and 9 hour time course. The incubation period would allow ample time for substrate loading on the 4'phosphopantetheine arm of holoPCP. Subsequently, we ran the the DKP assay by
using each module's natural substrate. L-Phe and L-Pro were incubated in their respective modules for 5 minutes and the reaction was quenched. Probe 10 showed notable inhibitory effect, suggesting inhibition of E domain activity (Fig. 17).

Figure 17. Notable inhibition diketopiperazine formation upon incubating with the suicide substrate β-chlorovinylglycine.

Further support for suicide substrate inhibition was determined utilizing a modified ATP-pyrophosphate exchange assay in which pre-incubation with inhibitor was proposed to yield a similar set of kinetic parameters due to the A domain's ability to activate an amino acid substrate to an aminoacyl-AMP ester independently from PCP interaction. Therefore, incubation with the suicide substrate 10 should have little effect on the kinetic values for L-phenylalanine loading. The assay measures the amount of radioactive ATP
formed from the reverse direction of the catalytic adenylation reaction when an excess of radioactive pyrophosphate (\(^{32}\)PPP) is supplied (Scheme 5).

**Scheme 5.** ATP-pyrophosphate exchange assay measuring the activation kinetics of the A domain in addition to the effects of crosslinking on A domain activity.

Utilizing the protoypical module, grsA, we investigated the kinetic parameters of the A domain with its cognate substrate, L-phenylalanine, and with preincubation with probe 10. L-phenylalanine demonstrated a \(K_m\) value of 0.07 ± 0.01 mM and catalytic turnover, \(k_{cat}\), of 84 ± 7 min\(^{-1}\), which are comparable with previously reported results.\(^{[32]}\) Subsequently, preincubation with probe 10 yielded repeated results of adenylation
activity, however, individual kinetic parameters without the presence of L-Phe demonstrated that 10 was a poor substrate for the A domain of grsA (Fig. 18).

**A**

![Graphs showing reaction rates for L-Phenylalanine and β-Chlorovinylglycine](image)

**B**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$k_{cat}/K_m$ (mM$^{-1}$ min$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Phe</td>
<td>$7.2 \times 10^{-3} \pm 0.01$</td>
<td>$8.4 \times 10^{-1} \pm 7$</td>
<td>$1.2 \times 10^3 \pm 400$</td>
</tr>
<tr>
<td>L-CVG</td>
<td>$1.3 \pm 0.8$</td>
<td>$9.3 \pm 2.1$</td>
<td>$7.2 \pm 0.9$</td>
</tr>
<tr>
<td>L-PEG</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Figure 18.** (A) Adenylation domain reversible reaction kinetics with its cognate amino acid and probe 10. (B) Michaelis-Menten variables of respective substrates.
Probe 10 showed a $K_m$ value of $1.3 \pm 0.8$ mM and a $k_{cat}$ of $9 \pm 2$, therefore showing a weak binding affinity and a low rate of adenylation.

Kinetic parameters for $\beta,\gamma$-Phenylethynylglycine 15 were not calculated as the synthesis of it has yet to be completed. However, preliminary results for 10 show promise for inhibitory effects for 15 due to their mechanistic similarity.

**Conclusion**

Two probes have been rationally designed to target the E domain of NRPSs and investigate the mechanism of epimerization. We began our efforts to thoroughly understand the E domain by examining an analogous system, the natural racemases. We identified a known inhibitor of the alanine racemase and mechanistically predicted that through a reactive allene intermediate, epimerization could be inhibited as well forming a covalently crosslinked species with its partner PCP domain.

Our modified diketopiperazine assay confirmed our hypothesis of inhibition by limiting the amount of diketopiperazine formed, however, the ATP-pyrophosphate exchange assay demonstrated that probe 10 is not a suitable substrate for the A domain. Moreover, probe 10 did not exhibit a deleterious effect on L-phenylalanine activation when incubated with grsA prior to initiation of the assay. The next goal lies in completing the synthesis of probe 15 and examining its inhibitory characteristics using our biochemical assays.

Further biochemical analysis could prove insightful in understanding the domain-domain interaction between E domains and their partner PCP domains. Recent release of the X-ray crystal structure of the E domain from the initiation module tycA in tyrocidine synthetase revealed the previously proposed active site of the E domain, however, no
substrate ligand was bound in its active site providing little evidence for the mechanism of epimerization.\cite{45} Through the use of crystallographic studies, insights into different structural conformations as well as mechanism of epimerization can be revealed with our generated crosslinked species.

**Materials and Methods**

**Biochemical Assays for Inhibition**

*Diketopiperazine Inhibition Assay.* Inhibition of the epimerization domain of grsA by 10 was examined using the grsA and tycB1 condensation assay, monitoring the D,L-Phe,Pro diketopiperazine formation by HPLC. Reaction mixture A contained the following (final volume of 250 μL): assay buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM MgCl₂, 8 mM ATP) and 2.8 μM grsA. Reaction mixture B contained the following (final volume 250 μL): assay buffer and 2.8μM tycB1. The condensation assay was initiated by addition of 10mM L-Phe and 10mM L-Pro (final concentration) to mixtures A and B, respectively. The mixture was incubated for 5 mins at 37 °C and subsequently combined and incubated further at 37 °C for 1hr. The dipeptide was extracted using 500μL of a butanol/chloroform(4:1) mixture (2x) and the organic phase was subsequently washed with 500 μL of 0.1 M NaCl. The organic solvent was removed under reduced pressure and the resulting residue was resuspended in 100 μL of 30% methanol. The suspension was then subjected to a C18 reversed-phase column (Honeywell Burdick and Jackson 4.6 x 250 mm) following equilibration with assay buffer and analyzed using an HP Agilent 1100 series HPLC system (λ = 220, 256 nm) using previously reported HPLC conditions. (citation). At a retention time of 10.7 mins, a
chromatographic trace appeared corresponding to the D,L-Phe, Pro DKP. MS analysis (HRMS (El) (m/z): [M+Na]⁺ calculated for C₁₄H₁₆N₂O₂Na yielded a mass of 267.1107.

Inhibition of DKP formation was assessed by preincubation of reaction mixture A with 0.8 mM 10 at 37 °C for a time course of 3, 6 and 9 hours. The condensation assay was then applied as described previously. Using the Aardvark® Software Program, HPLC traces were analyzed by determining the peak area of each corresponding chromatogram. Percent DKP formation was determined by comparing the DKP peak area of the cognate condensation assay with each time coursed assay.

*Kinetic Measurements of Adenylation Activity Using the ATP-Pyrophosphate Exchange Assay.* The kinetic parameters for the substrates, L-phenylalanine and β-chlorovinylglycine, were determined using the well-described ATP-pyrophosphate exchange assay. This assay utilized the prototypical initiation module, grsA, to determine A domain activity. Reaction mixtures contained the following (final volume of 100μL): 50 mM Hepes, pH 7.5, 100 mM NaCl, 10 mM MgCl₂, 2 mM DTT, 20-200 μM L-Phe or 0.01 - 50 mM 10, and 1.4 μM grsA. The reaction was initiated by adding 2 mM ATP, 0.2 mM tetrasodium pyrophosphate, and 0.1 μCi (1-60 Ci/mmol) of tetrasodium [³²P]pyrophosphate (PerkinElmer Inc.) at 25 °C. The reaction was quenched with 1.6% (w/v) activated charcoal and wash buffer (0.1 M tetrasodium pyrophosphate, and 3.5% perchloric acid) at different time points over a timecourse from 10 to 3600 seconds. The charcoal was pelleted by centrifugation, washed with wash buffer (3x), and resuspended in 0.5 mL of water. The suspension was then transferred to scintillation vials containing 3.5 mL of liquid scintillation fluid. A Beckman Coulter LS 6500 scintillation counter was used to measure the charcoal-bound radioactivity in each vial.

PRISM® was used to calculate Michaelis-Menten parameters for L-phenylalanine and 10 from nonlinear regression analysis.
References


