Published in final edited form as:

Chembiochem. 2019 January 18; 20(2): 147–152. doi:10.1002/cbic.201800439.

Mechanistic Probes for the Epimerization Domain of Nonribosomal Peptide Synthetases

Woojoo E. Kim^{‡,a}, Ashay Patel^{‡,a,b}, Gene H. Hur^a, Peter Tufar^a, Michael G. Wuo^a, J. Andrew McCammon^{a,b}, and Michael D. Burkart^a

^[a]W.E. Kim, Dr. A. Patel, Dr. G. H. Hur, Dr. P. Tufar, Dr. M.G. Wuo, Prof. Dr. J.A. McCammon, Prof. Dr. M.D. Burkart, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0358, United States

^[b]Dr. A. Patel, Prof. Dr. J.A. McCammon, Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0636, United States

Abstract

Nonribosomal peptide synthetases (NRPSs) are responsible for the synthesis of a variety of bioactive natural products with clinical and economic significance. Interestingly, these large multimodular enzyme machineries incorporate nonproteinogenic D-amino acids through the use of auxiliary epimerization domains, converting L-amino acids into D-amino acids, which impart into the resulting natural products unique bioactivity and resistance to proteases. Due to the large and complex nature of NRPSs, several questions remain unanswered about the mechanism of the catalytic domain reactions. Herein, we investigate the use of mechanism-based crosslinkers to probe the mechanism of an epimerization domain in gramicidin S biosynthesis. In addition, MD simulations were performed, showcasing the possible roles of catalytic residues within the epimerization domain.

Keywords

Nonribosomal peptide synthetase (NRPS); epimerization; enzyme inhibitors; molecular dynamic simulations; natural product biosynthesis

Natural products produced by microorganisms (i.e. bacteria and fungi) have had a tremendous impact on human health even before Alexander Fleming's discovery of penicillin's activity as an antibiotic in 1928[1]. Today, many of these microbial metabolites serve as antibiotics, antivirals, immunosuppressants, and anticancer agents[2]. Prominent among these therapeutics are a large class of peptide-based natural products called nonribosomal peptides, which are assembled by large multimodular enzymes, called nonribosomal peptide synthetases (NRPSs). NRPSs are comprised of modules, which are in turn composed of catalytic domains that each perform a single chemical reaction either to

mburkart@ucsd.edu.

^[‡]These authors contributed equally to this work.

Dedication to Chi-Huey Wong on the event of his 70th birthday.

Supporting information for this article is given via a link at the end of the document.

incorporate or modify a monomer unit into the final natural product. Central in mediating NRP synthesis is the peptidyl carrier protein (PCP), which acts as a scaffold, tethering the amino acid building blocks and peptidyl intermediates through a 4'-phosphopantetheine prosthetic arm as they are modified and condensed by other domains in the module[3]. The past decade has seen remarkable progress in the structural and mechanistic elucidation of NRPS domains through the use of biochemical, structural, and genetic methods[4]. Despite these achievements, new strategies to explore the enzymology of these complex, natural product-producing assembly lines are in constant development. One such approach that has gained considerable attention is the design, synthesis and application of small molecule probes to investigate the functional workings of biological systems, a field called chemical biology[5]. Such tools have been shown as having great promise in elucidating the physiochemical mechanisms of type II fatty acid biosynthesis[6], a modular process that is similar in synthetic logic to NRPS. The development of new chemical biology approaches will extend researchers' ability to interrogate the mechanisms of NRPS catalytic domains as well as uncover the proper protein interactions regulating efficient processing of biosynthetic intermediates in NRP synthesis with the ultimate goal of reengineering their biosynthetic machinery for the biocombinatorial synthesis of novel peptides with improved potency.

At present, chemical probes have been developed to investigate the protein interactions governing communication-mediating (COM) domains^[7], determining the substrate specificities of the core thiolation (PCP), adenylation (A), and condensation (C) domains^[8] as well as a few tailoring domains in NRP systems. In continuing with this approach, we have designed mechanism-based probes to investigate the epimerization (E) domains of NRPSs, which are responsible for the incorporation of D-amino acids into NRPs. The E domain producing D-amino acids occurs through the *in situ* epimerization of the C $_{\alpha}$ center of the S-aminoacyl-PCP, few biochemical studies have been performed on the E domain since its discovery. Much regarding the mechanism of action of this enzyme^[9] has yet to be elucidated. Mutational studies of the E domain in the initiation module, GrsA, of the gramicidin S synthetase revealed active site residues H753 and E892 to be vital for epimerization activity, suggesting a deprotonation/reprotonation mechanism^[9]. Sequence alignment and secondary structure comparisons of the E domain of GrsA and the C domain of VibH^[10], reveal the structural similarity of the two domains, including a similar channel present at the face of their active site. Consequently, it was proposed that the 4'-phosphopantetheine prosthetic arm of the peptidyl carrier protein (PCP) binds within this channel from the N-terminal face of the E domain, presenting the covalently attached L-amino acyl substrate at the active site in order for the catalytic residues to epimerize the C $_{\alpha}$ center of the amino acid. It was also shown that the E domain catalyzes racemization of the C $_{\alpha}$ center of the amino acid only when it is tethered to the 4'-phosphopantetheine prosthetic arm of the PCP and not in its free form^[11].

In contrast to E domains of NRPSs, amino acid racemases, best exemplified by the bacterial alanine racemase, require a cofactor, pyridoxal 5'-phosphate (PLP), to catalyze racemization^[12]. PLP is a coenzyme form of vitamin B6, stabilizes the anionic intermediate formed when L- or D-alanine is deprotonated at its C $_{\alpha}$ center by capturing this intermediate covalently as a quinoid adduct^[13]. In addition, mutational studies of the bacterial alanine racemase demonstrated that two catalytic residues (K39 and Y265) were responsible for proton

abstraction and addition at the C $_{\alpha}$ -center of the amino acid substrate^[14]. Haloalanines and halovinylglycines have been reported to be potent mechanism-based inhibitors for the bacterial alanine racemase^[15,16]. Given the mechanistic similarities between the alanine racemase and NRPS epimerization domain, we reasoned that chlorovinylglycine (CVG) **1** and β -fluorophenylalanine **2** (Figure 2) would serve as promising candidates for the design of mechanism-based probes of the E domain. To accomplish this, we first synthesized probes **1** and **2**. In the synthesis of **1** and **2**, we followed previously published procedures^[16–18].

We utilized the tridomain (A-PCP_{Phe}-E) initiation module, GrsA, of the gramicidin S synthetase (Figure 1) to determine the efficacy of probes **1** and **2** as a mechanism-based crosslinkers. We chose this module because it contained the necessary domains to activate (adenylation domain), load (peptidyl carrier protein), and undergo crosslinking (epimerization domain) with the probes to show the inactivation of the E domain. To determine the activity of probes **1** and **2** toward the E domain, we employed a previously developed *in vitro* condensation assay to indirectly show the inhibition of E domain. This assay monitors for dipeptide formation between the initiation module GrsA and the second tridomain (C-A-PCP_{Pro}) module of the tyrocidine system, TycB₁^[19]. In this assay, L-Phe and L-Pro are first activated via ATP-dependent adenylation by the A domains of GrsA and TycB₁, respectively. Once activated, they are loaded onto their corresponding PCPs, forming L-Phe-S-GrsA and L-Pro-S-TycB₁. L-Phe-S-GrsA is modified to D-Phe-S-GrsA by E domain. Due to the stereoselectivity of the TycB₁ C domain, condensation only occurs once D-Phe is produced via epimerization by GrsA. The C domain of TycB₁ then, catalyzes peptide bond formation to form the dipeptidyl intermediate D-Phe-L-Pro-S-TycB₁. This dipeptide undergoes intramolecular cyclization releasing the D-Phe-L-Pro diketopiperazine (DKP) product, which can be detected by HPLC analysis. DKP formation cannot happen when E domain is inactivated using mechanism-based probes. Thus, DKP production can be used as a measure of the probe activity. We carried out the condensation assay after preincubating GrsA with the probe to prevent competitive adenylation of the natural substrate, L-Phe. We first preincubated GrsA at 37°C with the probe and ATP for 1, 3, 6, 9, 12 and 24 hrs. This would activate and load the probe onto the PCP so that they could then react with and inactivate the E domain. Afterwards, we ran the condensation assay using the natural substrates L-Phe and L-Pro with the preincubated probe bearing GrsA and TycB₁. HPLC analysis of the assay showed no decrease in DKP formation when GrsA was preincubated with β -fluorophenylalanine but showed a decrease in DKP formation with chlorovinylglycine (**1**), suggesting that it was partially inactivating the E domain (Figure 3).

Based on these results, we next preincubated GrsA with **1** in the absence of ATP and then performed the condensation assay to verify that inhibition of DKP formation was dependent on the A domain's ability to recognize, activate, and load **1** onto the PCP domain prior to its reaction with the E domain. HPLC analysis of this assay showed no decrease in DKP formation when ATP was omitted in the preincubation period, signifying that that **1** alone could not inhibit the E domain, and thus DKP formation, as a free amino acid. Instead, the probe must be activated by the A domain and tethered onto the PCP in order to inactivate the E domain. Probe **2** showed full DKP formation meaning no inhibition of E domain. We presume that because probe **2** is very similar to the natural substrate and product, L- and D-

phenylalanine, that once loaded onto PCP, it undergoes E domain-catalyzed epimerization and is subsequently offloaded.

We propose a mechanism for the chemical crosslinking between **1**-PCP and the GrsA E domain that is consistent with the mechanism proposed for the epimerization of L-Phe. The X-ray crystal structures of the TycA epimerization domain^[20], a homolog of GrsA's E domain, also responsible for the epimerization of L-Phe, suggests that the catalytic histidine – based upon the local environment – exists in its doubly protonated (imidazolium) form. Computational predictions of pKa's (using propKa as implemented on the H++ webserver) the catalytic histidines of both the TycA and GrsA E domain supports this conclusion. Consequently, it is proposed that the histidine acts not as an acid or a base, but rather promotes epimerization by stabilizing the enolate species that is generated as a reaction intermediate. Given this inability of His753 to act as an acid or base, Glu892 (in the case of GrsA's E domain), is proposed to assume this role. Thus, we propose that the Glu892 deprotonates **1** at C_α, generating an enolate intermediate that collapses to form the reactive allene intermediate. As shown in Figure 4, the allene species is activated by protonation by His753, which simultaneously enhances the nucleophilicity of His753 enabling it to undergo Michael addition with **1** (Figure 4). Further work is necessary to definitely demonstrate that the catalytic histidine is targeted by **1**-PCP. Nonetheless, we contend that of the two active site residues, only His753 could react with **1**-PCP to form a stable enzyme-probe adduct.

In order to further investigate the probe activity on epimerization domain, we performed molecular dynamics (MD) simulations to better rationalize the behaviour of **1** as probe and substrate, respectively. To do so, we generated models of L-phenylalanyl-PCP-E, D-phenylalanyl-PCP-E, chlorovinylglycyl-PCP-E didomain using the holo-PCP-E didomain crystallized (PDB accession code: 5ISX) by Bruner et al.^[21] Our analysis of the MD simulations was performed with the aforementioned mechanistic information in mind.

Simulations suggest that the L-phenylalanyl substrate must undergo internal rotation in order to present the hydrogen at C_α to Glu892. For this to occur, the salt bridge between Glu892 and the (protonated) amino group of the substrate must be disrupted. The reorganization of the substrate in this manner can be monitored using the S-C(O)-C_α-C_β dihedral (Figure 5). Analysis of this dihedral for L-phenylalanyl-PCP-E indicates L-Phe-PCP assumes one of two rotamers, (–) anticlinal and (+) synclinal conformations. The (+) synclinal conformer disposes the substrate's phenyl substituent in tryptophan-rich region of the active, while simultaneously maintaining a salt bridge between protonated amine of the substrate and the Glu892. It is the (–) anticlinal conformation of L-Phe disposes the substrate alpha hydrogen in proximity to Glu892, but substrate thioester is no longer hydrogen bonding to His753. Simulations of D-phenylalanyl-PCP-E indicate that the D-phenylalanyl substrate adopts one of three conformations about its S-C(O)-C_α-C_β dihedral in the E domain active site, the (–) synclinal and (+) synclinal, or (+) anticlinal, conformation. In both the (–) synclinal and (+) synclinal substrate rotamer dispose the substrate towards Glu892, consistent with the E domain's ability to convert D-phenylalanyl-PCP into the starting material. The difference in the S-C(O)-C_α-C_β dihedral angles the two epimeric phenylalanyl substrates samples can be attributed to the shape complementary of the GrsA E domain towards L-phenylalanyl-PCP, namely that the substrate can dispose its aromatic side chain toward a region rich in

tryptophan (aromatic) residues while forming salt bridge (Figure 5). Since no such rotamer exists for the D substrate transitions between S-C(O)-C α -C β conformers are likely more facile.

Data collected from simulations of L- and D-phenylalanyl-PCP-E provide a context by which the reactivity of the probe **1** can be better understood. The chlorovinylglycyl substrate is able to sample a board range of S-C(O)-C α -C β dihedral angles similar those sampled by the cognate L-Phe substrate. This presumably allows the probe to reorient itself within the E domain active site so that its reactive warhead is presented to the active site nucleophile. Further analysis and a complete description of the computational methodology employed to perform simulations described herein can be found in the Supporting Information.

In summary, we propose new insight into the mechanistic role of the catalytic residues in the epimerization domain of NRPSs. Mechanism-based crosslinker, chlorovinylglycine – previously shown to undergo reaction with alanine racemase – demonstrates partial inhibition towards the GrsA's epimerization domain. MD simulations suggest that substrate organization with the E domain active site is central to reactivity of both cognate and noncognate aminoacyl substrates.

Experimental Section

Expression and Purification of GrsA and TycB₁

GrsA and TycB₁ were expressed and purified using a modified version of previously published protocols^[19,22]: pQE60 plasmids (QIAGEN) harboring the genes were transformed into BL21(DE3) cells. *grsA* was co-transformed with *sfp* (the gene for the phosphopantetheine transferase from the surfactin A NRPS) in pREP4, and *tycB₁* with *gsp* (the gene for the phosphopantetheine transferase from the gramicidin S NRPS) in pREP4.

For protein production, 500 mL two times yeast and tryptone media (2xYT: 16 g/L tryptone, 10 g/L yeast extract and 5 g/L NaCl) with 30 mg/L kanamycin sulfate and 100 mg/L ampicillin were inoculated with 5 mL overnight culture of the respective expression strain. The cultures were grown at 37°C/120 rpm to an optical density at 600 nm (OD₆₀₀) of 0.3. At this point, the temperature was reduced to 30°C, the expression was induced at an OD₆₀₀ of 0.6–0.8 by addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM, and the cells were subsequently grown for an additional 2 h.

Cells were harvested by centrifugation, resuspended in buffer A (25 mM HEPES, 200 mM NaCl, 10 mM imidazole, pH 8.0), and lysed by passing through a French Press twice. The lysate, clarified by centrifugation, was loaded onto a NiNTA column (QIAGEN), washed with 15 column volumes buffer A, and eluted with 3 CV buffer B (same as buffer A, but with 250 mM imidazole). The eluate was dialyzed against 50 mM HEPES, 300 mM NaCl, 10 mM MgCl₂, 5 mM dithiothreitol (DTT) (pH 7.5) overnight at 4°C.

To ensure quantitative phosphopantetheinylation, the proteins were incubated with 10.0 eq coenzyme A (CoA) and 0.05 eq Sfp for 1 h at 20°C before they were concentrated and loaded onto a Sephacryl S300 size exclusion column (GE Healthcare) equilibrated with 50

mM HEPES, 300 mM NaCl, 10 mM MgCl₂ (pH 7.5). Monomeric fractions were again concentrated, and aliquots were frozen in liquid nitrogen and stored at –80°C.

Running the DKP formation assay with non-frozen enzymes directly after the size exclusion chromatography step and enzymes that were stored at –80°C resulted in virtually identical yields, indicating no loss of activity upon freezing/thawing.

Diketopiperazine (DKP) Formation Assay: Positive Control

The assay was performed following a previously published protocol^[19] with some modifications: mixtures of GrsA (final concentration 2.8 μM), L-Phe (8.0 mM), and ATP (8.0 mM) and TycB1 (2.8 μM), L-Pro (8.0 mM), and ATP (8.0 mM), respectively, were made in a 200 μL scale in 50 mM HEPES, 300 mM NaCl, 10 mM MgCl₂ (pH 7.5). DKP formation was started by combining the two mixtures and carried out for 1 h at 37°C. The reaction was stopped by extracting with 400 μL n-butanol/chloroform (4:1) twice. The combined organic phases were washed with 400 μL NaCl (0.1 M) and dried under reduced pressure.

The dried extract was dissolved in 50 μL 10%(v/v) methanol, and 20 μL were loaded onto Ascentis Express Peptide ES C18 column (15 cm × 4.6 mm; 2.7 μm particle size; Sigma-Aldrich) for quantification. Solvents A was 0.05% trifluoroacetic acid (TFA) in water, solvent B 0.05% TFA in acetonitrile. Prior to loading of a sample, the column was equilibrated with 2% B, and the following gradient was used at a flow rate of 1 mL/min: 0–2 min, 2% phase B; 2–10 min, 2–30% phase B; 10–12 min, 30–95% phase B; 12–15 min, 95% phase B; 15–17 min, 95–2% phase B; 17–20 min, 2% phase B (UV absorbance at 280, 254, 214 nm). Peaks were collected to screen for the mass of DKP ([M+H]⁺ theoretical: 245.13, observed: 245.18). The yield of DKP was determined by the peak intensity of its absorbance at 214 nm.

Diketopiperazine (DKP) Formation Assay: Inhibition by mechanism-based crosslinkers

In 200 μL scale reactions in 50 mM HEPES, 300 mM NaCl, and 10 mM MgCl₂ (pH 7.5), GrsA (2.8 μM final concentration), ATP (8.0 mM), and probe (0.8 mM from a stock in DMSO) was pre-incubated for 1 – 24 h at 18°C. Following the pre-incubation, DKP formation was started by addition of 200 μL of a mixture of TycB₁ (2.8 μM), ATP (10.0 mM), L-Phe (8.0 mM), and L-Pro (8.0 mM). After 1 h at 37°C, DKP was extracted and quantified as described for the positive control. In control reactions with 24 h pre-incubation, ATP was omitted in the pre-incubation step. All of the described reactions were also performed with substituting chlorovinylglycine (CVG) with DMSO.

Acknowledgments

This work was supported by NIH RO1 GM095970 (to M.D.B.) and NIH RO1 GM21749 (to J.A.M.). National Biomedical Computing Resource (NBCR) and San Deigo Computing Resource are acknowledged for providing computing resources. We thank Drs. Anthony Mrse and Yongxuan Su for assistance with NMR and MS data acquisition, and Drs. Joris Beld and Heriberto Rivera with synthetic advice and discussions.

References

- [1]. Fleming A, Br. J. Exp. Pathol 1929, 10, 226–236.
- [2]. Streit JM, Steenbergen JN, Thorne GM, Alder J, Jones RN, J. Antimicrob. Chemother 2005, 55, 574–578. [PubMed: 15722390]
- [3]. Lai JR, Koglin A, Walsh CT, Biochemistry (Mosc.) 2006, 45, 14869–14879.
- [4]. Hill AM, Nat. Prod. Rep 2006, 23, 256–320. [PubMed: 16572230]
- [5]. Meier JL, Burkart MD, Chem. Soc. Rev 2009, 38, 2012–2045. [PubMed: 19551180]
- [6]. Nguyen C, Haushalter RW, Lee DJ, Markwick PRL, Bruegger J, Caldara-Festin G, Finzel K, Jackson DR, Ishikawa F, O'Dowd B, et al., Nature 2014, 505, 427–431. [PubMed: 24362570]
- [7]. Hur GH, Meier JL, Baskin J, Codelli JA, Bertozzi CR, Marahiel MA, Burkart MD, Chem. Biol 2009, 16, 372–381. [PubMed: 19345117]
- [8]. Belshaw PJ, Walsh CT, Stachelhaus T, Science 1999, 284, 486–489. [PubMed: 10205056]
- [9]. Stachelhaus T, Walsh CT, Biochemistry (Mosc.) 2000, 39, 5775–5787.
- [10]. Keating TA, Marshall CG, Walsh CT, Keating AE, Nat. Struct. Mol. Biol 2002, 9, 522–526.
- [11]. Stein T, Kluge B, Vater J, Franke P, Otto A, Wittmann-Liebold B, Biochemistry (Mosc.) 1995, 34, 4633–4642.
- [12]. Watanabe A, Yoshimura T, Mikami B, Hayashi H, Kagamiyama H, Esaki N, J. Biol. Chem 2002, 277, 19166–19172. [PubMed: 11886871]
- [13]. Eliot AC, Kirsch JF, Annu. Rev. Biochem 2004, 73, 383–415. [PubMed: 15189147]
- [14]. Watanabe A, Yoshimura T, Mikami B, Esaki N, J. Biochem. (Tokyo) 1999, 126, 781–786. [PubMed: 10502689]
- [15]. Wang E, Walsh C, Biochemistry (Mosc.) 1978, 17, 1313–1321.
- [16]. Thornberry NA, Bull HG, Taub D, Greenlee WJ, Patchett AA, Cordes EH, J. Am. Chem. Soc 1987, 109, 7543–7544.
- [17]. Tsushima T, Kawada K, Nishikawa J, Sato T, Tori K, Tsuji T, Misaki S, J. Org. Chem 1984, 49, 1163–1169.
- [18]. Okonya JF, Johnson MC, Hoffman RV, J. Org. Chem 1998, 63, 6409–6413. [PubMed: 11672280]
- [19]. Stachelhaus T, Mootz HD, Bergendahl V, Marahiel MA, J. Biol. Chem 1998, 273, 22773–22781. [PubMed: 9712910]
- [20]. Samel SA, Czodrowski P, Essen L-O, Acta Crystallogr. D Biol. Crystallogr 2014, 70, 1442–1452. [PubMed: 24816112]
- [21]. Li W-H, Li K, Sandya SN, Bruner SD, ACS Chem. Biol 2016, 11, 2293–2303. [PubMed: 27294598]
- [22]. Mootz HD, Schwarzer D, Marahiel MA, Proc. Natl. Acad. Sci 2000, 97, 5848–5853. [PubMed: 10811885]

Figure 1.
Biosynthetic pathway of gramicidin S.

Figure 2. Mechanism-based crosslinking probes chlorovinylglycine 1 and β-fluorophenylalanine 2 for the NRPS epimerization domain.

Figure 3.
HPLC analysis of % DKP formation after preincubation of probe 1 and 2.

Figure 4.
Schematic diagram of probe **1** inhibition on epimerization domain.

Figure 5.

Distribution of the $S-C(O)-C_{\alpha}-C_{\beta}$ of the substrate dihedral angle of L-Phe-PCP-E (A), D-Phe-PCP-E (B), and 1-PCP-E (C). Newman projections and 3D representations of representative structures of each $S-C(O)-C_{\alpha}-C_{\beta}$ rotamer sampled using conventional MD simulation. Catalytic residues, His753 and Glu892, are shown in orange. The substrate and phosphopantetheine are shown as sticks and coloured white. Trp632, Trp644, and Trp911 (aromatic binding pocket) are shown as spheres in light blue. Note that hydrogen, oxygen, nitrogen atoms are coloured white, red, and blue.