Chemistry & Biology
Article

Directed Evolution of the Nonribosomal Peptide Synthetase AdmK Generates New Andrimid Derivatives In Vivo

Bradley S. Evans,[1,2] Yunqiu Chen,[3,4] William W. Metcalf,[2,5] Huimin Zhao,[2,6] and Neil L. Kelleher[3,7,4,*]
[1]Department of Biochemistry
[2]Institute for Genomic Biology
University of Illinois, Urbana, IL 61801, USA
[3]Department of Chemistry
[4]The Chemistry of Life Processes Institute
Northwestern University, Evanston, IL 60208, USA
[5]Department of Microbiology
[6]Department of Chemical and Biomolecular Engineering
University of Illinois, Urbana, IL 61801, USA
[7]Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208, USA
*Correspondence: n-kelleher@northwestern.edu
DOI 10.1016/j.chembiol.2011.03.008

SUMMARY

Many lead compounds in the search for new drugs derive from peptides and polyketides whose similar biosynthetic enzymes have been difficult to engineer for production of new derivatives. Problems with generating multiple analogs in a single experiment along with lack of high-throughput methods for structure-based screening have slowed progress in this area. Here, we use directed evolution and a multiplexed assay to screen a library of >14,000 members to generate three derivatives of the antibacterial compound, andrimid. Another limiting factor in reengineering these mega-enzymes of secondary metabolism has been that commonly used hosts such as *Escherichia coli* often give lower product titers, so our reengineering was performed in the native producer, *Pantoea agglomerans*. This integrated in vivo approach can be extended to larger enzymes to create analogs of natural products for bioactivity testing.

INTRODUCTION

Nonribosomal peptides are a proven source of antibacterial (Recktenwald et al., 2002), antifungal (Denning, 1997), and anticancer drugs (Du et al., 2000). These peptides often contain nonproteogenic amino acids as well as fatty acid, polyketide, carbohydrate, and isoprenoid appendages. The enzymes that build these modified peptides, nonribosomal peptide synthetases (NRPSs), are often large enzymes with multiple active sites and are characterized by keeping intermediates covalently tethered by a series of phosphopantetheine (PPant) linked thioester bonds at the thiolation (T) domains during biosynthesis. The gatekeepers of these molecular "assembly lines" are the adenylation (A) domains which select an amino acid building block from the pool available within producing cells. The modular structure of NRPSs make them attractive targets for protein engineering, but realization of this goal over the last few decades has been far more difficult than many had hoped. Here we screen a library of ~14,330 clones for derivatives of the antibiotic andrimid (**1**). Focusing on the site of valine incorporation, three new chemical derivatives have been created with similar or improved biological activity.

Even though amino acid substrates for adenylation domains may be predicted with some confidence via the "NRPS Code" (Challis et al., 2000; Rausch et al., 2005; Stachelhaus et al., 1999), engineering of these domains to accept noncognate substrates has been elusive. Often referred to as combinatorial biosynthesis (Khosla et al., 1999; Khosla and Zawada, 1996; Tsoi and Khosla, 1995; Walsh, 2002; Walsh et al., 2003), a main approach has employed chimeric synthetases where an adenylation domain of one NRPS is replaced by an adenylation domain from another NRPS which activates a different substrate. This so-called "domain swapping" strategy has proved successful in terms of generating the desired product, but has been plagued by decreased product yields and outright dysfunction presumably due to interruption of protein-protein interactions caused by this cut and paste approach (Baltz et al., 2006; Doekel and Marahiel, 2000; Mootz et al., 2000, 2002; Nguyen et al., 2006; Schneider et al., 1998; Stachelhaus and Marahiel, 1995a, 1995b). The activity of chimeric synthetases could be restored in the first reported use of directed evolution for reengineering NRPS systems (Fischbach et al., 2007). Domain swapping, however, is still not ideal for true combinatorial biosynthesis in that at most one variant of an NRP can be generated per swap. Here, we present the second example of using directed evolution to alter thiotemplated biosynthetic systems. Instead of rescuing activity of a chimeric synthetase using error-prone PCR, we directly targeted active site residues for evolution using saturation mutagenesis (Chockalingam et al., 2005). This orthogonal approach to NRPS reengineering focuses on modifying the sites of substrate specificity,

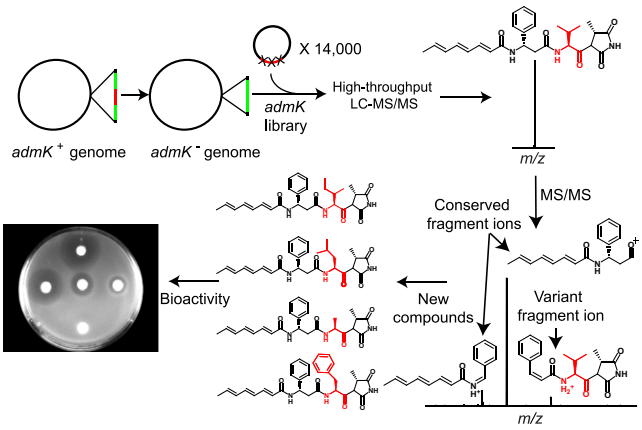

Figure 1. Workflow for Directed Evolution of AdmK
After deletion of *admK* from the *P. agglomerans* chromosome and transformation with a library of *admK* mutants, the library is screened by high-throughput LC-MS/MS. New compounds are identified and assayed for bioactivity. See also Figure S1.

rather than modifying the protein-protein interactions of the A- and T-domains as in a previous study (Fischbach et al., 2007). By targeting mutations to the substrate binding site and generating hundreds of enzyme variants in parallel, we created a large library of mutants, each with the potential to generate a natural product derivative. Further, we used mass spectrometry as a structure-based readout for the library. Ultimately, four mutants that produce robust levels of new derivatives were found. The workflow for this study is illustrated in Figure 1.

The system under study here is the andrimid biosynthetic pathway (Figure 2). Andrimid (**1**) is a hybrid NRP/PK molecule produced by *P. agglomerans* (Fredenhagen et al., 1987; Jin et al., 2006; Singh et al., 1997). Andrimid is a broad spectrum antibiotic and acts by inhibiting fatty acid biosynthesis at the acetyl CoA carboxylase step, preventing formation of malonyl CoA, the precursor for fatty acids and some polyketides (Freiberg et al., 2004). Previous structure-activity relationship (SAR) studies identified a portion of andrimid, the valine subunit, that could be exchanged to create more potent and specific compounds (Freiberg et al., 2006). Domain swapping experiments on the corresponding NRPS gene, *admK*, showed that the valine subunit could be substituted for isoleucine or aminobutyrate and that there was no apparent editing by downstream condensing enzymes (Belshaw et al., 1999), allowing generation of andrimid derivatives in vivo (Fischbach et al., 2007).

RESULTS

Library Design and Construction

We sought to express a library of mutant enzymes and monitor the effectiveness of the mutations by screening for end products produced by the pathway in vivo. To accomplish this, removal of the wild-type (WT) enzyme activity was necessary. We employed an allele replacement strategy based on homologous recombination that has proved successful in a number of organisms, namely the streptomycin resistance counterselection (Russell and Dahlquist, 1989). Using this strategy, we made a scarless deletion of the WT *admK* gene in *P. agglomerans*. The resulting strain served as a background strain for expression of an *admK* mutant library (see Figure S1A available online). This result was validated by PCR (Figure S1B). Also, we found that the *E. coli* expression plasmid pQE60 was capable of supplying *admK* and restoring andrimid production levels comparable to WT in *P. agglomerans* (Figure S1C).

In order to select residues in *admK* for mutagenesis, we turned to sequence alignments of AdmK with A-domains known to activate substrates different than the valine found in andrimid (Figure S3). We chose A-domains that activate nonpolar amino acids because SAR studies (Freiberg et al., 2006) showed that this type of residue could substitute for valine and increase the potency and selectivity of the andrimid scaffold. From the sequence alignment with the ten residues used for substrate prediction of NRPS A-domains (Stachelhaus et al., 1999), we chose the three most highly variant of the ten for saturation mutagenesis (Miyazaki and Arnold, 1999) (Figure 3). Our three-site saturation mutagenesis library was designed to create 1404 distinct mutants, and was limited by selection of codons for synthetase residues likely to select nonpolar substrates (Figure 3) for incorporation into the andrimid backbone. The mutant library was constructed by utilizing sequential megaprimer and overlap extension PCRs (Horton, 1997; Sarkar and Sommer, 1990). In

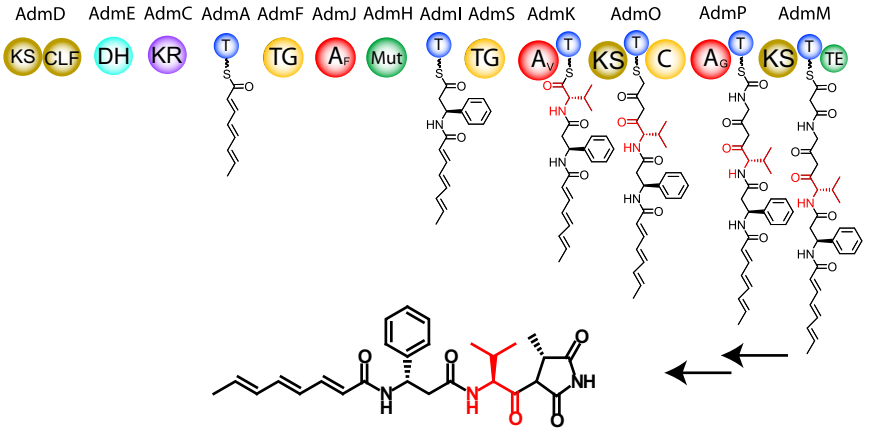

Figure 2. Andrimid Biosynthetic Pathway
Andrimid is biosynthesized by a highly disconnected and hybrid NRPS/PKS pathway. Adenylation domain substrate specificity is indicated by single letter amino acid abbreviation subscript. AdmK incorporates the valyl subunit (highlighted) of andrimid and has been targeted for mutagenesis. KS- ketosynthase, CLF- chain length factor, DH-dehydrogenase, KR- ketoreductase, T-thiolation, TG-transglutaminase, A-adenylation, Mut- aminomutase, C- condensation, TE-thioesterase.

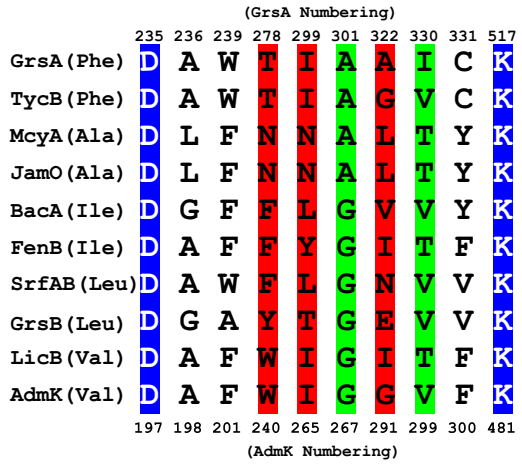

Figure 3. Extracted Nonribosomal Code Residues

Nonribosomal code residues extracted from adenylation domains from GrsA, AdmK and other systems with consensus sequences for nonpolar amino acids (shown in parentheses). Strictly conserved residues are highlighted in blue, highly conserved residues are highlighted in green and highly variant residues targeted for limited saturation mutagenesis are highlighted in red. See also Figure S3.

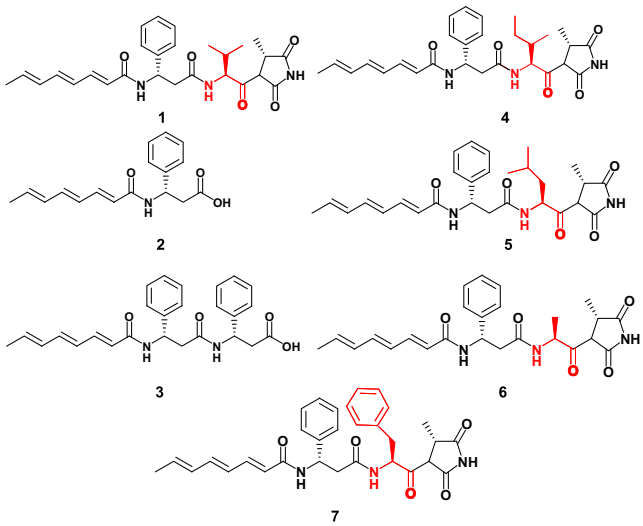

Figure 4. Structures of Andrimid and Derivatives Produced in This Study

1 andrimid, **2** octatrienyl-β-phenylalnine, **3** octatrienyl-β-phenylalanyl-β-phenylalanine, **4** isoleucine-andrimid, **5** leucine-andrimid, **6** alanine-andrimid, **7** phenylalanine-andrimid. See also Figure S4.

order to make a library that maximized the amino acid substitutions we desired, while minimizing unwanted substitutions truncations, PCR primers with limited degeneracy were used. Position 240 of AdmK was mutagenized using a WHK codon in place of the native TGG codon substituting the native Trp with Met, Thr, Ser, Tyr, Leu, Phe, Ile, Asn, and Lys and only one possible stop codon. For position 265 of AdmK, a WHW codon replaced the native ATC codon thereby substituting Leu, Phe, Thr, Tyr, Asn, Lys, and one possible stop codon for Ile. Position 291 of AdmK was mutagenized by using an RHK codon in place of the WT GGT codon in order to substitute Val, Ile, Asn, Ala, Glu, Met, Thr, Lys, and Asp for Gly with no possibility for a truncation at this position. Together, these three degenerate codons allow for 1404 codon combinations. To ensure 95% coverage of 1404 mutants, we screened 14,330 clones. Statistical analysis using the program GLUE (Patrick et al., 2003) estimated that the library was 99.99% complete with a 94.9% probability that the library contained all possible variants.

Assay Development and Library Screening

To screen our library of andrimid producing clones, we employed a structure-based assay. We reasoned that using an antibiotic bioassay we would not be able to distinguish between andrimid and derivatives of andrimid. Also, we would be incapable of identifying which precise andrimid derivative had been produced because most derivatives would likely be active antibiotics (Freiberg et al., 2006). During evaluation, mass spectrometric detection of andrimid showed ~23,000-fold greater sensitivity than a bioassay, and we noticed a fragmentation pattern that cleanly dissected the molecule into regions to be conserved or variant (Figure 1, lower right panel). This calculation of relative sensitivity for MS versus classic halo assays is based on the limit of detection (LOD) during MS of andrimid, which is 20 fmol. Given that 96 samples are mixed per LC-MS injection, this translates to a LOD of 1.9 pmol for the pooled assay. The MIC data for andrimid against *E. coli* show that 0.47 nmol can be detected in a 10 μl spot in the agar overlay assay. If this assay were diluted 96-fold as in the LC-MS assay this translates to a 45.1 nmol LOD. These attributes of the MS-based screen allowed us to identify which andrimid derivative was produced and conduct the screening of 96 clones in a single LC-MS run. Single clones arrayed into 96-well plates, were inoculated into 2×YT microbial medium supplemented only with antibiotic and grown for 24 hr prior to pooled sampling of each 96-well plate. Each pooled sample representing 96 individual clones constituted one LC-MS/MS injection. In this way, over 14,330 clones from 150 individual 96-well plates of the three site library were screened and four clones producing andrimid derivatives were isolated. Twenty-four percent of assay plates (i.e., 96 samples combined into one) still showed production of andrimid while 41% of plates showed no production beyond the octatrienyl-β-phenylalanine precursor to andrimid (a truncated product, **2**). Curiously, almost all of the plates assayed showed production of a species at *m/z* 433.21 that was determined to be octatrienyl-β-phenylalanyl-β-phenylalanine (**3**), a result of apparent "stuttering" by the transglutaminase AdmF, the enzyme responsible for incorporating β-phenylalanine into the growing ketide-peptide chain.

Characterization of Mutants and Derivative Compounds

The four mutant clones were identified subsequently (by pooled -row and -column searching of assay plates), sequenced and designated A2 (W240L, G291V), A7 (W240S, I265T, G291E), A9 (W240T, I265T, G291E), and B12 (R235K, W240L, I265L, G291E). Together, these clones produced three new andrimid derivatives and another that has been described previously (Figure 4; Figure S4A, i-iv). Clones A2 and B12 were found to

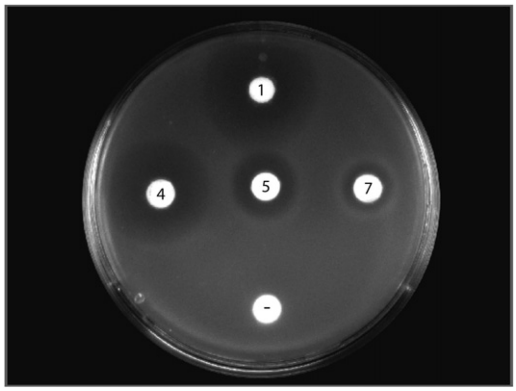

Figure 5. Andrimid and Derivatives Bioassay
Agar-overlay bioassay of andrimid and analogs generated against *E. coli* imp ASR. Aliquots of each compound (numbered) and a solvent negative control (–) were spotted onto paper disks atop a layer of soft agar seeded with *E. coli*.

produce an andrimid related peak at *m/z* 494.26, corresponding to either isoleucine- or leucine-substituted andrimid (**4, 5**) (Figure S4A, i, ii). Clones A7 and A9 produced a mixture of alanine- and phenylalanine-substituted andrimid (**6, 7**) (Figure S4A, iii, iv). In order to resolve the L-Ile versus L-Leu ambiguity for clones A2 and B12, an isotope feeding study was conducted. Growing clones A2 and B12 in media supplemented with a 50:50 mixture of $^{13}C_2$-Leu and $^{13}C_6$-Ile demonstrated that both clones A2 and B12 have preference for Ile over Leu, but at different levels of discrimination (Figure S4B). Clone A2 showed a 4:1 preference for Ile to Leu, whereas clone B12 showed a 20:1 preference for Ile to Leu. Control experiments where only one isotope labeled amino acid was supplemented clearly showed that when one amino acid is supplied in 50 mM excess, incorporation of that amino acid is efficient (Figure S4C). For that reason, we employed a mutasynthesis strategy where the desired amino acid to be incorporated was supplied in excess for scale up of fermentations in order to determine yield and bioactivity. We found that applying this mutasynthesis strategy also improved the yield for clones A7 and A9. Compounds **1**, **4**, **5**, and **7** were all found to be bioactive in disk diffusion assays (Figure 5). When using the mutasynthesis, we were able to significantly improve the yields, in some cases approaching WT levels (Table 1; Figure S5). MICs were determined for several representative bacterial strains (Table 2).

Table 1. Mutations in *admK* and Corresponding Products Compared with WT Production Levels

Clone	Mutations	Compound Produced	Medium	Production per Liter (mg)
WT	N/A	**1**	2xYT	378.16
A2	W240L, G291V	**4**	2xYT + Ile	4.94
		5	2xYT + Leu	0.99
B12	R235K, W240L, I265L, G291E	**4**	2xYT + Ile	466.39
		5	2xYT + Leu	93.83
A7	W240S, I265T, G291E	**6**	2xYT + Ala	0.19
		7	2xYT + Phe	0.36
A9	W240T, I265T, G291E	**6**	2xYT + Ala	0.20
		7	2xYT + Phe	0.49

DISCUSSION

This work takes on the challenges associated with engineering enzymes of secondary metabolism with a technological approach that applies generally to metabolites detectable by tandem mass spectrometry. When combined with high level expression in the native host, *P. agglomerans*, the reduced metabolic production from active site mutation of *admK* (largely from poor solubility of variants), did not prevent four successes from being found in a library pool where ~80% were unaffected, ~20% were dead, and ~0.03% had altered production.

Creating a mutant library of *admK* rather than performing several domain-swapping experiments allowed us to generate four andrimid derivatives in the same overall experiment while examining many solutions for each derivative simultaneously. The key aspect of our work that allowed for this combinatorial biosynthesis experiment is the structure-based assay that provided a direct readout of not just the AdmK variants' activity, but also the combined output of downstream enzymes in the assembly line. Unlike many indirect assays of enzyme activity (bioactivity or coupled assays), direct structural readout of mutant activity in the integrated pathway tightly connects the overall impact of the mutagenesis with a main goal of the exercise: production of new derivatives.

We found four mutants that exhibited altered substrate specificity, A2, A7, A9, and B12. All showed a broadening of substrate specificity relative to the WT, with increased preference to substrates other than the cognate substrate, valine. To our surprise, the overall best performing mutant was actually a serendipitous quadruple mutant. The fourth mutation in B12, R235K is likely an artifact of PCR; however, it is unknown if this stray mutation is required for robust production of the Ile/Leu analogs. Structures for NRPS A-domains available exhibit low sequence identity to AdmK, 25% for 1AMU and 24% for 2VSQ. However, based on alignment of AdmK to the structures above, R235 corresponds to N697 of 2VSQ and E272 of 1AMU. These positions in the respective crystal structures are remote from the active site and on the surface of the protein. We can only speculate that R235K may change a protein-protein interaction surface or induce a structural change to the active site through an outer sphere effect. The relatively higher activity and solubility of the B12 mutant over the others suggests that mutations

Table 2. MIC Data for Andrimid and Derivatives against Representative Gram-Negative and Gram-Positive Bacterial Strains

	MIC (nmol)		
	1	**4**	**5**
E. coli imp ASR	0.47	2.72	0.73
K. pneumoniae ATCC 13883	9.05	2.93	NT[a]
S. aureus ATCC 25923	7.71	0.39	0.28
E. faecalis ATCC 19433	36.35	42.76	NT[a]

[a] Not tested.

outside the selectivity conferring code should be considered for future A-domain mutant libraries. A second round of mutagenesis and screening was performed on the four isolated clones using an error-prone PCR approach in order to increase the activity of the first round mutants, but no improvement was observed. MIC values for the derivative compounds **4** and **5** showed increased bioactivity toward *Staphylococcus aureus* relative to **1**, and unchanged or decreased bioactivity toward *E. coli* consistent with previous SAR studies (Freiberg et al., 2006) (Table 2). Prediction of A-domain substrates is fairly robust (Rausch et al., 2005). Despite this, prescription of what active site mutations to make for a particular metabolite structure is far less developed. This study provides a large-scale survey of what amino acid substitutions could be tolerated in the A-domain active site. While many of the triple mutants that still yielded andrimid were not sequenced here, one can envisage a large-scale study specifically aimed at determining empirically what sites are stable to mutation. This would effectively map functional outcomes for many positions and sharply increase our predictive ability for creation of new NRPS derivatives.

One reason the andrimid system was chosen is because its biosynthetic enzymes are mostly < < 100 kDa (i.e., relatively small for NRPS or PKS proteins). There is no reason why our methods cannot be extended to larger multimodule NRPSs, provided the host is genetically tractable, or the pathway is amenable to heterologous expression with robust yields. We envision that the general approach used here could be extended to other domains within NRPS systems to probe substrate specificity of condensation domains. We estimate that a ~100-fold improvement in screening speed could be achieved, leading to an ability to screen 4 or even 5 site mutant libraries with similar coverage within the same approximate scope of time and resources used in this work.

Peptide natural products from NRPS systems draw high interest because of their medical uses, but harnessing their rather predictable correlation between enzyme primary structure and metabolite structure has proved elusive over the past two decades (Khosla et al., 1999). Reengineering strategies to achieve combinatorial biosynthesis (Walsh, 2002) compete conceptually with organic synthesis as a strategy for enhancing the diversity of natural molecular scaffolds. Challenges to attaining the goal of true combinatorial biosynthesis include not only the general challenges of enzyme engineering, but also the methods used to evaluate the modified enzymes and pathways. The field of NRPS reengineering is making strides through better understanding of protein-protein recognition motifs (Li and Vederas, 2009) and the use of directed evolution (Fischbach et al., 2007) to generate libraries of A-domains capable of activating a variety of substrates. Alternative approaches to making mutant libraries such as family shuffling (Crameri et al., 1998) could be employed in the future to generate even larger libraries of large active site fragments, perhaps leading to more stable and active mutant enzymes.

SIGNIFICANCE

Combinatorial biosynthesis is an attractive alternative to chemical synthesis for creating analogs of complex natural products. A major hurdle to realizing the potential of combinatorial biosynthesis includes the lack of selections or screens to process mutant libraries of enzymes in the context of their entire biosynthetic pathway. This work demonstrates a general method for rapidly generating and screening mutant libraries of nonribosomal peptide synthetases. The unbiased nature of structure-based screening allows for evaluation of the output from entire mutant pathways on a basis largely disconnected from total yield or specific bioactivity. Using the high-throughput multiplexed LC-MS/MS screening method outlined here, four derivatives of the antibiotic andrimid were generated, three of them not previously reported. The relative bioactivities of the derivatives were measured and it was found that two of the derivatives exhibited a shift in antibacterial specificity with concomitant decreases in their MIC values, consistent with previous synthetic chemistry driven SAR studies. Mutagenesis combined with LC-MS/MS for a structure-based readout is a general approach that can be readily extended to other nonribosomal peptide synthetase and polyketide synthase pathways as well as to other biosynthetic pathway paradigms.

EXPERIMENTAL PROCEDURES

Chemicals and Reagents

Restriction enzymes were from Invitrogen (Carlsbad, CA), T4 DNA ligase was from New England Biolabs (Ipswich, MA), Phusion polymerase from Finnzymes (Woburn, MA) was used for PCR. Growth medium was from BD Biosciences (San Jose, CA). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and are listed in Supplemental Experimental Procedures. HPLC solvents were from Fisher (Pittsburgh, PA). Formic Acid was from Acros Organics (Geel, Belgium). PCR cleanup, plasmid miniprep and gel extraction kits were from QIAGEN (Valencia, CA). Stable-isotope labeled amino acids were from Cambridge Isotopes (Andover, MA). All other reagents, chemical and consumables were from Sigma unless otherwise specified.

Strains and Plasmids

E. coli DH5α λ pir was used for all cloning steps for plasmids containing *oriR6K*. *E. coli* DH5α and BL21(DE3) were obtained from the UIUC Cell Media Facility. *E. coli* imp ASR, *Klebsiella pneumoniae* ATCC 13883, *Staphylococcus aureus* ATCC 25923, and *Enterococcus faecalis* ATCC 19433 were used as the indicator strains for agar overlay assays. The andrimid producer, *P. agglomerans* Eh 335 and *E. coli* imp ASR were a gift from Christopher T. Walsh. *K. pneumoniae* ATCC 13883, *S. aureus* ATCC 25923, and *E. faecalis* ATCC 19433 were obtained from ATCC (Manassas, VA). All strains were cultured in LB or 2xYT at 37°C (*E. coli, K. pneumoniae, S. aureus*, *and E. faecalis*) or 28°C (*P. agglomerans*). Plasmid pQE60 (QIAGEN) was used for expression of *admK* in *P. agglomerans.* Plasmids pACYC-Duet, pET-Duet, pET28a, and pUC-19 were obtained from Novagen (Darmstadt, Germany). The suicide plasmid pMQ118 was obtained from Presque Isle Cultures (Erie, PA). All strains and plasmids are listed with relevant characteristics in Supplemental Experimental Procedures.

Antibiotic Bioassays

LB agar plates (100 mm diameter) were overlaid with 4 ml molten LB agar inoculated with 75 μl of an overnight culture of the indicator strain. The inoculated top agar was allowed to solidify before sterile paper disks were laid onto the assay plates. Onto each disk 10–20 μl of sterile filtered culture supernatant or control was spotted before incubation at 37°C overnight or until zones of inhibition were visible.

DNA Sequencing

DNA sequencing was carried out at the University of Illinois core sequencing facility on an ABI 3730XL capillary sequencer on reactions using BigDye

Version 3.0 terminator/enzyme mix and standard protocols. Sequence outputs were assembled and analyzed using Sequencher 4.6 software (Gene Codes, Ann Arbor, MI).

Mass Spectrometry

A 7 Tesla LTQ-FT equipped with a Surveyor autosampler and a MS pump was used for all mass spectrometric analysis (Thermo-Fisher Scientific). Andrimid and its derivatives were analyzed using a Jupiter 4.6 × 150 mm C18 HPLC column (Phenomenex) and a gradient of 0%–75% ACN over 35 min. LC solvents (water and acetonitrile) contained 0.1% formic acid. Initial screening was performed at low (ion-trap) resolution with a selected ion monitoring (SIM) window from *m/z* 435–535 as the full scan and data-dependent fragmentation of the top 10 peaks from the full scan. Confirmation of derivative production was carried out using the same scans but in high (FT, 100,000) resolution. Mass spectrometric data were analyzed using Qual Browser software (Thermo-Fisher Scientific).

Molecular Genetics

Please refer to Supplemental Experimental Procedures for detailed methods for genetic manipulation of *P. agglomerans*.

Alignment of AdmK to Other Adenylation Domains and Extraction of Binding Pocket Residues

Several adenylation domains with specificity for the desired activity in the AdmK (GenBank accession number AAO39105.1) were aligned: GrsA (GenBank accession number CAA33603.1, phenylalanine specific), TycB M3 (GenBank accession number AAC45929.1, phenylalanine specific), McyA M2 (GenBank accession number CAO90227.1, alanine specific), JamO_ (GenBank accession number AAS98786.1, alanine specific), BacA M1 (GenBank accession number AAC06346.1, isoleucine specific), FenB (GenBank accession number AAB00093.1, isoleucine specific), SrfAB M3 (GenBank accession number BAA08983.1, leucine specific), GrsB M4 (GenBank accession number CAA43838.1, leucine specific) and LicB M1 (GenBank accession number AAD04758.1, valine specific). Adenylation domains were aligned using the program CLUSTALX (Thompson et al., 1997) (Figure S3). Positions aligning to D235, A236, W239, T278, I299, A301, A322, I330, C331, and K517 of GrsA were extracted as the putative binding pocket residues (Stachelhaus et al., 1999). The binding pocket residues were inspected for their variation across the aligned sequences and three positions corresponding to T278, I299, and A322 in GrsA were chosen for mutagenesis.

Construction and Screening of admK Mutant Library

Please refer to Supplemental Experimental Procedures for detailed methods for construction and screening of the admK mutant library.

Isolation and Quantification of Andrimid and Analogues

Andrimid was produced from liquid culture of *P. agglomerans* in 2xYT supplemented with antibiotics and amino acids as required for 28°C for 24 hr. Cells were removed by centrifugation followed by filtration through a 0.45 μm membrane (Millipore, Billerica, MA). Filtered culture supernatant was extracted three times with one-half volume ethyl acetate. Ethyl acetate fractions were combined and dried under vacuum. The crude extract was dissolved in 50:50 water: methanol and extracted with 9 volumes of hexane. The water:methanol fraction was then extracted with one half volume methylene chloride. The methylene chloride fraction was dried under vacuum and then dissolved in 10 ml 20% acetonitrile and applied to a 12 cc C_{18} SPE column (Whatman, Piscataway, NJ) equilibrated with 20% acetonitrile. The column was washed with 2 volumes 40% acetonitrile and eluted with one volume 80% acetonitrile. The eluate was dried and dissolved in 20% acetonitrile then loaded onto a 10 × 150 mm Eclipse C_{18} column (Agilent, Santa Clara, CA). Andrimid and analogs were eluted from the column using a linear gradient 0%–70% acetonitrile over 30 min. Andrimid and analogs eluted as pure peaks from 25 to 32 min. Peak purity and yield were determined by a LC-MS assay using the same column, gradient and mass spectrometer used in screening the mutant library calibrated with standards at known concentrations.

Stable Isotope Feeding Study

Mutants A2 and B12 were grown in M9 medium supplemented with 0.1% yeast extract, 0.1% casamino acids and 50 mM each of $^{13}C_2$ leucine and/or $^{13}C_6$ isoleucine (Cambridge Isotopes) for 24 hr at 30°C. The culture supernatants were syringe filtered and analyzed using the LC-MS/MS method outlined above.

MIC Determination

LB agar plates (140 mm diameter) were overlaid with 12 ml molten LB agar inoculated with 225 μl of OD_{600} = 0.8 indicator strain. The inoculated top agar was allowed to solidify before sterile paper disks were laid onto the assay plates. Ten microliters of antibiotic was spotted onto the disks and plates were incubated 8 hr at 37°C. Zones of inhibition were defined as the radius and measured using calipers. Zones of inhibition were plotted against Log_{10} of the concentration of spotted antibiotic. MICs were defined as the minimum concentration required to yield a 1 mm zone of inhibition and calculated by nonlinear regression using the software package GraphPad Prism (La Jolla, CA).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and five figures and can be found with this article online at doi:10.1016/j.chembiol.2011.03.008.

ACKNOWLEDGMENTS

We gratefully acknowledge the gift of *P. agglomerans* Eh 335 and *E. coli* imp ASR from Christopher T. Walsh. This work also would not have been possible without financial support from the National Institute of General Medical Sciences under grants number R01 GM 067725-08 (N.L.K.) and P01 GM 077596-03 (N.L.K., H.Z., and W.W.M.), and a Molecular Biophysics Training Grant 5T32GM008276 (B.S.E.).

Received: November 20, 2010
Revised: February 2, 2011
Accepted: March 7, 2011
Published: May 26, 2011

REFERENCES

Baltz, R.H., Brian, P., Miao, V., and Wrigley, S.K. (2006). Combinatorial biosynthesis of lipopeptide antibiotics in *Streptomyces roseosporus*. J. Ind. Microbiol. Biotechnol. *33*, 66–74.

Belshaw, P.J., Walsh, C.T., and Stachelhaus, T. (1999). Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. Science *284*, 486–489.

Challis, G.L., Ravel, J., and Townsend, C.A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. *7*, 211–224.

Chockalingam, K., Chen, Z., Katzenellenbogen, J.A., and Zhao, H. (2005). Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. Proc. Natl. Acad. Sci. USA *102*, 5691–5696.

Crameri, A., Raillard, S.A., Bermudez, E., and Stemmer, W.P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature *391*, 288–291.

Denning, D.W. (1997). Echinocandins and pneumocandins–a new antifungal class with a novel mode of action. J. Antimicrob. Chemother. *40*, 611–614.

Doekel, S., and Marahiel, M.A. (2000). Dipeptide formation on engineered hybrid peptide synthetases. Chem. Biol. *7*, 373–384.

Du, L., Sanchez, C., Chen, M., Edwards, D.J., and Shen, B. (2000). The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. Chem. Biol. *7*, 623–642.

Fischbach, M.A., Lai, J.R., Roche, E.D., Walsh, C.T., and Liu, D.R. (2007). Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. Proc. Natl. Acad. Sci. USA *104*, 11951–11956.

Fredenhagen, A., Tamura, S.Y., Kenny, P.T.M., Komura, H., Naya, Y., Nakanishi, K., Nishiyama, K., Sugiura, M., and Kita, H. (1987). Andrimid, a New Peptide Antibiotic Produced by an Intracellular Bacterial Symbiont Isolated from a Brown Planthopper. J. Am. Chem. Soc. *109*, 4409–4411.

Freiberg, C., Brunner, N.A., Schiffer, G., Lampe, T., Pohlmann, J., Brands, M., Raabe, M., Habich, D., and Ziegelbauer, K. (2004). Identification and characterization of the first class of potent bacterial acetyl-CoA carboxylase inhibitors with antibacterial activity. J. Biol. Chem. *279*, 26066–26073.

Freiberg, C., Pohlmann, J., Nell, P.G., Endermann, R., Schuhmacher, J., Newton, B., Otteneder, M., Lampe, T., Habich, D., and Ziegelbauer, K. (2006). Novel bacterial acetyl coenzyme A carboxylase inhibitors with antibiotic efficacy *in vivo*. Antimicrob. Agents Chemother. *50*, 2707–2712.

Horton, R.M. (1997). In vitro recombination and mutagenesis of DNA. SOEing together tailor-made genes. Methods Mol. Biol. *67*, 141–149.

Jin, M., Fischbach, M.A., and Clardy, J. (2006). A biosynthetic gene cluster for the acetyl-CoA carboxylase inhibitor andrimid. J. Am. Chem. Soc. *128*, 10660–10661.

Khosla, C., and Zawada, R.J.X. (1996). Generation of polyketide libraries via combinatorial biosynthesis. Trends Biotechnol. *14*, 335–341.

Khosla, C., Gokhale, R.S., Jacobsen, J.R., and Cane, D.E. (1999). Tolerance and specificity of polyketide synthases. Annu. Rev. Biochem. *68*, 219–253.

Li, J.W., and Vederas, J.C. (2009). Drug discovery and natural products: end of an era or an endless frontier? Science *325*, 161–165.

Miyazaki, K., and Arnold, F.H. (1999). Exploring nonnatural evolutionary pathways by saturation mutagenesis: Rapid improvement of protein function. J. Mol. Evol. *49*, 716–720.

Mootz, H.D., Schwarzer, D., and Marahiel, M.A. (2000). Construction of hybrid peptide synthetases by module and domain fusions. Proc. Natl. Acad. Sci. USA *97*, 5848–5853.

Mootz, H.D., Schwarzer, D., and Marahiel, M.A. (2002). Ways of assembling complex natural products on modular nonribosomal peptide synthetases. ChemBioChem *3*, 490–504.

Nguyen, K.T., Ritz, D., Gu, J.Q., Alexander, D., Chu, M., Miao, V., Brian, P., and Baltz, R.H. (2006). Combinatorial biosynthesis of novel antibiotics related to daptomycin. Proc. Natl. Acad. Sci. USA *103*, 17462–17467.

Patrick, W.M., Firth, A.E., and Blackburn, J.M. (2003). User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. Protein Eng. *16*, 451–457.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D.H. (2005). Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res. *33*, 5799–5808.

Recktenwald, J., Shawky, R., Puk, O., Pfennig, F., Keller, U., Wohlleben, W., and Pelzer, S. (2002). Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. Microbiol. *148*, 1105–1118.

Russell, C.B., and Dahlquist, F.W. (1989). Exchange of chromosomal and plasmid alleles in Escherichia coli by selection for loss of a dominant antibiotic sensitivity marker. J. Bacteriol. *171*, 2614–2618.

Sarkar, G., and Sommer, S.S. (1990). The "megaprimer" method of site-directed mutagenesis. Biotechniques *8*, 404–407.

Schneider, A., Stachelhaus, T., and Marahiel, M.A. (1998). Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. Mol. Gen. Genet. *257*, 308–318.

Singh, M.P., Mroczenski-Wildey, M.J., Steinberg, D.A., Andersen, R.J., Maiese, W.M., and Greenstein, M. (1997). Biological activity and mechanistic studies of andrimid. J. Antibiot. (Tokyo) *50*, 270–273.

Stachelhaus, T., and Marahiel, M.A. (1995a). Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. FEMS Microbiol. Lett. *125*, 3–14.

Stachelhaus, T., and Marahiel, M.A. (1995b). Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. J. Biol. Chem. *270*, 6163–6169.

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.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876–4882.

Tsoi, C.J., and Khosla, C. (1995). Combinatorial Biosynthesis of Unnatural Natural-Products - the Polyketide Example. Chem. Biol. *2*, 355–362.

Walsh, C.T. (2002). Combinatorial biosynthesis of antibiotics: Challenges and opportunities. ChemBioChem *3*, 125–134.

Walsh, C.T., O'Connor, S.E., and Schneider, T.L. (2003). Polyketide-nonribosomal peptide epothilone antitumor agents: the EpoA, B, C subunits. J. Ind. Microbiol. Biotechnol. *30*, 448–455.