How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 22172–22177
International Edition: doi.org/10.1002/anie.202104874
German Edition: doi.org/10.1002/ange.202104874

Antibiotic Analogues

Metagenome-Guided Analogue Synthesis Yields Improved Gram-Negative-Active Albicidin- and Cystobactamid-Type Antibiotics

*Zongqiang Wang[+], Amanda Kasper[+], Rabia Mehmood, Melinda Ternei, Shaogang Li, Joel S. Freundlich, and Sean F. Brady**

Abstract: *Natural products are a major source of new antibiotics. Here we utilize biosynthetic instructions contained within metagenome-derived congener biosynthetic gene clusters (BGCs) to guide the synthesis of improved antibiotic analogues. Albicidin and cystobactamid are the first members of a new class of broad-spectrum ρ-aminobenzoic acid (PABA)-based antibiotics. Our search for PABA-specific adenylation domain sequences in soil metagenomes revealed that BGCs in this family are common in nature. Twelve BGCs that were bio-informatically predicted to encode six new congeners were recovered from soil metagenomic libraries. Synthesis of these six predicted structures led to the identification of potent antibiotics with changes in their spectrum of activity and the ability to circumvent resistance conferred by endopeptidase cleavage enzymes.*

In many instances a characterized natural product represents only one example of a larger family of congener structures. These natural congeners may arise as a result of selective pressures to have different potencies, responses to resistance, and spectra of activity. Soil metagenomes contain large collections of bacteria, including many difficult-to-culture species, thus making them rich sources of congener biosynthetic gene clusters (BGCs). In this study, we combine soil metagenomics with bio-informatic natural product structure prediction methods to intelligently guide the synthesis of albicidin and cystobactamid analogues with improved antibacterial activity and the ability to overcome naturally occurring resistance, two likely outcomes of the diverse selective pressures that drive the evolution of natural antibiotic congeners (Figure 1).

Albicidin and cystobactamid are closely related antibiotics that contain ρ-aminobenzoic acid (PABA) and target DNA gyrase.[1,2] Their potent Gram-negative activity and unique chemical structures make them appealing leads for the potential development of novel therapeutics.[1,3] However, the identification of an endopeptidase (AlbD) that confers resistance by antibiotic cleavage suggests that this resistance could eventually threaten their therapeutic utility.[4,5] Here we searched for PABA-specific adenylation domain sequences to identify albicidin and cystobactamid BGCs in soil metagenomes. Although they were originally isolated from traditionally underexplored bacterial taxa (*Xanthomonas albilineans* and *Cystobacter* spp, respectively),[2,6] we found that congener BGCs are common in soil metagenomes and that natural structural variations encoded by these BGCs are different from those explored in previous medicinal chemistry studies (Figure 2a). The synthesis of albicidin and cystobactamid analogues predicted from metagenomic BGCs (i.e. synthetic bio-informatic natural products or syn-BNPs)[7] led to our identification of antibiotics with improved potency as well as the ability to circumvent AlbD cleavage without a loss of antibiosis.

To identify albicidin and cystobactamid congener BGCs, we screened previously archived metagenomic libraries for PABA-specific adenylation (A) domain sequences. In total, these libraries contain about 720 million unique cosmid clones. Their construction and arraying have been described previously.[8] Briefly, environmental DNA (eDNA) extracted directly from diverse soils was cloned into a cosmid vector and then introduced into *E. coli* using lambda phage.[9] To facilitate the recovery of specific BGCs of interest, each library of $>2\times10^7$ cosmid clones was arrayed into collections of subpools each containing about 25000 unique cosmid clones. Library subpools were screened using subpool-specific barcoded A domain degenerate primers. PCR amplicons were then sequenced using Illumina MiSeq technology and the resulting reads from each library subpool were clustered at 95% identity to generate natural product sequence tags (NPSTs) that can be used to guide the discovery of BGCs of interest. Using our environmental surveyor of natural product diversity (eSNaPD) software package,[10] library-derived NPSTs were compared to PABA-specific A domain sequences from albicidin and cystobactamid BGCs. NPSTs that returned eValues of $<10^{-25}$ were considered potential PABA-specific A domains and used to construct a PABA A domain phylogenetic tree (Figure 1). NPSTs that grouped most closely with PABA-specific A domains from albicidin and cystobactamid biosynthesis were considered markers for congener BGCs. Sets of overlapping cosmids associated with predicted PABA NPSTs were isolated from the appropriate library subpools and sequenced to reveal 12 complete and partial cystobactamid or albicidin-like BGCs (Figure 2).

[*] Dr. Z. Wang,[+] Dr. A. Kasper,[+] R. Mehmood, M. Ternei, Prof. S. F. Brady
Laboratory of Genetically Encoded Small Molecules
The Rockefeller University
1230 York Avenue, New York, NY 10065 (USA)
E-mail: sbrady@rockefeller.edu
Dr. S. Li, Prof. J. S. Freundlich
Department of Medicine, Center for Emerging and Re-emerging Pathogens
Rutgers University—New Jersey Medical School
Newark, NJ 07103 (USA)
[+] These authors contributed equally to this work.
Supporting information and the ORCID identification numbers for some of the authors of this article can be found under: https://doi.org/10.1002/anie.202104874.

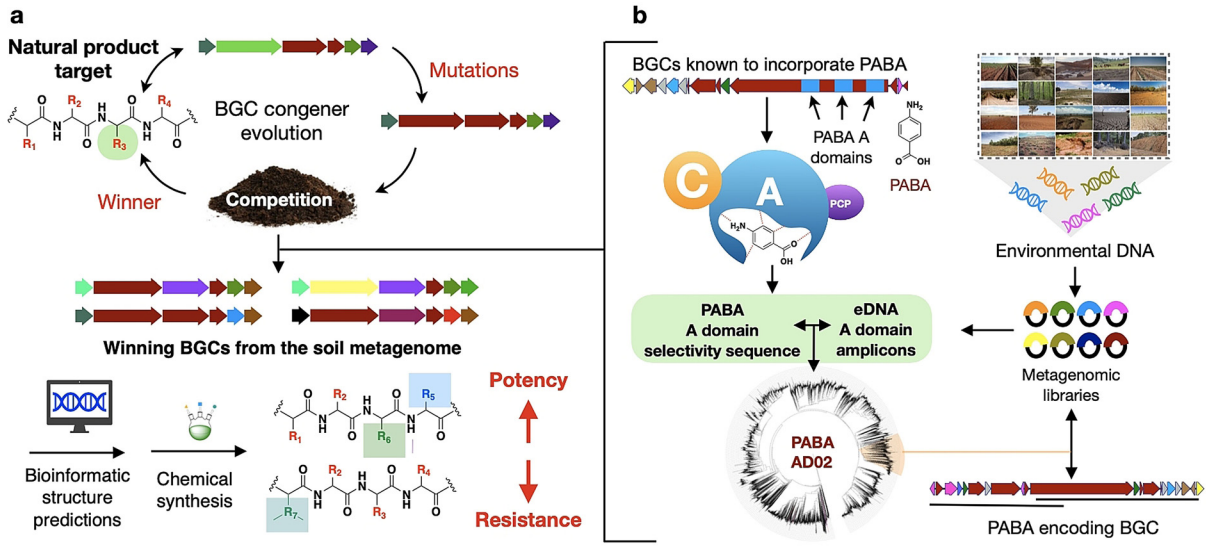


Figure 1. a) Overview of metagenome-guided chemical synthesis. Natural selection is predicted to result in congeners with varied potency, spectrum of activity, and response to resistance. Bio-informatic analysis of BGCs that encode these congeners can guide the synthesis of improved antibiotics. b) To identify albicidin and cystobactamid congener BGCs, A domain sequences amplified from soil metagenomic libraries were searched for sequences that grouped with domains known to incorporate PABA. These sequences were used to guide the recovery of clones containing albicidin and cystobactamid congener BGCs.

The structure of the natural product encoded by each metagenomic BGC was predicted based on the presence or absence of a PKS module, the types of PABA modification enzymes it encodes, and—most importantly—the predicted substrate specificity of each NRPS A domain. For this analysis, 10 conserved amino acids in the A domain substrate binding pocket were used to determine the substrate specificity (positions 235, 236, 239, 278, 299, 301, 322, 330, 331, and 517).[11] We compared the A domain substrate binding pockets found in each metagenomic NRPS with characterized A domain substrate binding pockets (Tables S9–S16). Signature sequences for PABA as well as modified PABAs [4-amino-2-hyroxy-3-isopropoxybenzoic acid (AHIBA) or 4-amino-2-hydroxy-3-methoxylbenzoic acid (AHMBA)] were derived from A domains found in the albicidin and cystobactamid BGCs (Tables S9 and S10). An analysis of the tailoring genes that encode the differential functionalization of PABA allowed us to distinguish between the presence of AHIBA and AHMBA in the predicted product of each BGC.

Among the seven cystobactamid-like BGCs we identified, we could predict three different structural analogues, which are embodied by BGCs PABA48, PABA70, and PABA57. Each unique predicted cystobactamid congener BGC encodes 6 NRPS modules. A domain substrate specificity analysis of the 6 A domains in each BGC together with an analysis of the tailoring genes in these BGCs indicates that the central four residues (positions B, C, D, E) are conserved across this family of congeners (Figures 2a and 3a): PABA, β-methoxyasparagine (MO-Asn), PABA, and AHIBA. In place of the PABA seen at the N-terminus of known cystobactamids, our analysis predicted that congener BGCs incorporate either tyrosine or phenylalanine at this position. The C-terminal residue is predicted to be an AHIBA in two cystobactamid congeners (PABA48 and PABA70) and PABA in the third (PABA57). As would be expected for a BGC that encodes a metabolite containing AHIBA, cystobactamid congener BGCs are predicted to encode a B12-dependent radical SAM enzyme (CysS homologue) that introduces the *t*-butyl functionality onto AHIBA.[12]

Key differences between cystobactamid- and albicidin-like BGCs include the presence of a specific polyketide synthase module in albicidin BGCs and genes for the biosynthesis of AHIBA in cystobactamid BGCs (Figure 3). Among the five hybrid PKS/NRPS BGCs identified, we could predict three distinct albicidin congeners. BGCs PABA34, PABA157, and PABA95 are representative of these three new structures. Three residues are positionally conserved across these congeners. This includes a central β-L-cyanoalanine (CN-Ala), an AHMBA at the E position, and a PABA at the F position. Positions B and D contain different arrangements of PABA or AHMBA. The N-terminal methylcoumaric acid (MCA) seen in albicidin is introduced by the PKS module that contains a ketosynthase (KS), dehydratase (DH), ketoreductase (KR), and methyltransferase (MT) domain. The corresponding PKS module in the PABA34 and PABA95 BGCs contain the same collection of domains (Figures S4 and S6); however, in BGC PABA157 this module is missing the MT domain, which suggests the biosynthesis of coumaric acid (CA) in place of MCA (Figure S5). This led to our prediction that the products of PABA34 and PABA95 contain MCA at position A, while PABA157 incorporates a CA as the N-terminal building block.

Interestingly, no residue remains constant across these six predicted natural congeners. Although a growing number of analogues have been produced in synthetic efforts to improve the potency and resistance profile of this class of antibiotics, none of them match the structures of the naturally selected congeners encoded by these metagenomic BGCs.[13,14] Based

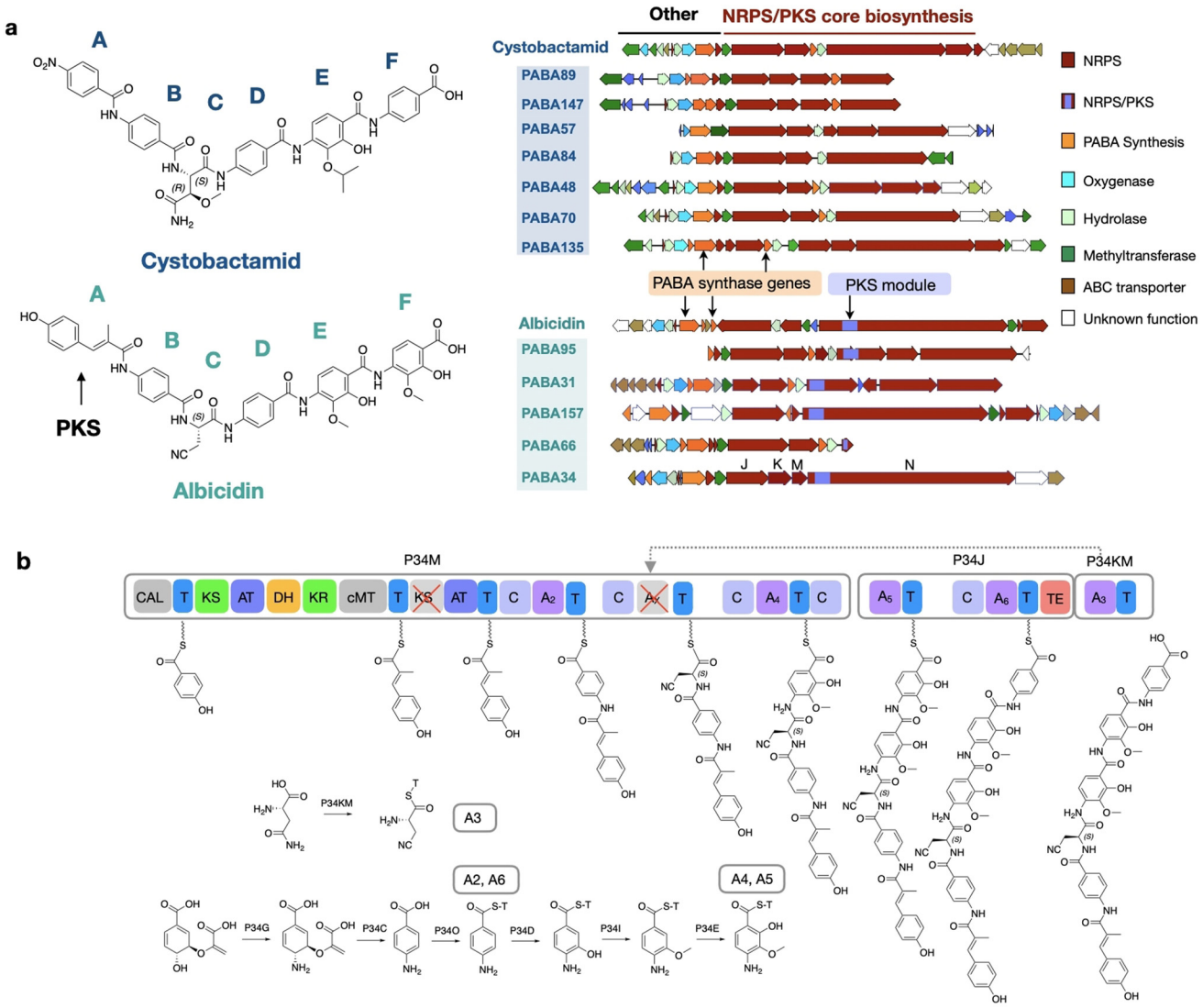
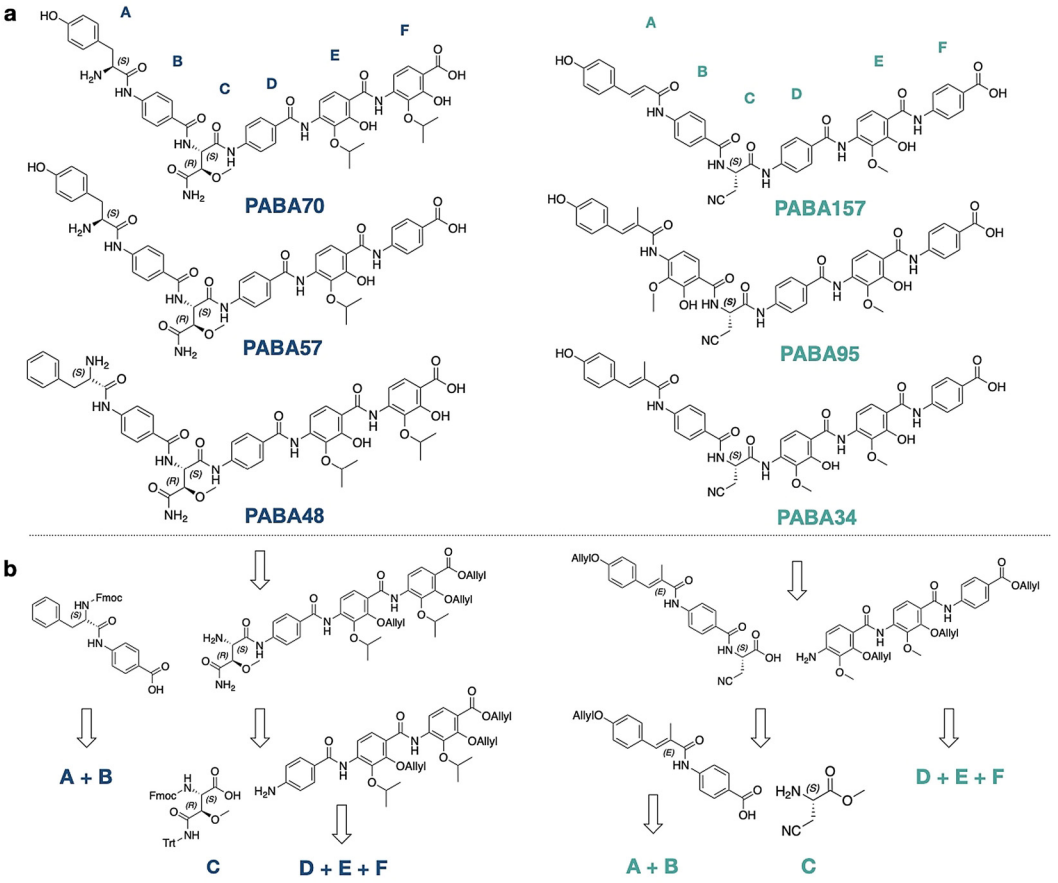


Figure 2. a) Albicidin and cystobactamid congener BGCs isolated from soil metagenomic libraries. b) Proposed biosynthetic pathway for PABA34. Representative analysis used to predict syn-BNP targets.

on our bio-informatic predictions, the total synthesis of each predicted congener was undertaken. These structures arise almost exclusively from the coupling of a shared set of substituted and unsubstituted PABA monomers. A convergent synthesis was therefore envisioned that would enable facile access to all six structures from a minimum number of monomer building blocks. In keeping with previous synthetic routes,[15] each congener was disconnected retrosynthetically into three fragments: the N-terminal dipeptide, the central MO-Asn or CN-Ala unit, and the C-terminal tripeptide. Figure 3b shows a representative retrosynthetic analysis of both a predicted albicidin and cystobactamid congener. Utilization of this general strategy allowed us to easily incorporate different predicted PABA building blocks at positions B, D, E, and F as well as incorporate diverse building blocks at the N-terminal position. Integration of the central CN-Ala or MO-Asn units occurred strategically in each synthesis to minimize opportunities for racemization. The final deprotected products, which we have generically called synthetic bio-informatic natural products (syn-BNPs) were purified by HPLC, and their structures were confirmed by HRMS as well as ^{1}H and ^{13}C NMR spectroscopy.

All syn-BNPs, as well as the parent antibiotic albicidin were assayed for antimicrobial activity against the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter species*), which are a collection of bacteria frequently associated with antibacterial resistant nosocomial infections. Compared to albicidin, syn-BNPs showed improved potency against a number of pathogens (Tables 1, S1, and S18) and differences in spectrum of activity, as might be expected for congeners that have evolved under different natural selective pressures. PABA34 showed the broadest potent Gram-negative activity with an MIC of $\leq 2\ \mu g\,mL^{-1}$ against *A. baumannii*, *P. aeruginosa*, *E. cloacea*, and *E. coli* (Table 1). PABA48 was the most potent broad-spectrum antibiotic we identified. It had the same MIC as albicidin against *A. baumannii*, but was 2-, 16-, and >64-



c **BGC specific substrate predictions**

	A	B	C	D	E	F
PABA48	Phe	PABA	MO-Asn	PABA	AHIBA	AHIBA
PABA70	Tyr	PABA	MO-Asn	PABA	AHIBA	AHIBA
PABA57	Tyr	PABA	MO-Asn	PABA	AHIBA	PABA

	A	B	C	D	E	F
PABA34	MCA	PABA	CN-Ala	AHMBA	AHMBA	PABA
PABA157	CA	PABA	CN-Ala	PABA	AHMBA	PABA
PABA95	MCA	AHMBA	CN-Ala	PABA	AHMBA	PABA

Figure 3. Bio-informatically predicted structures and representative retrosynthetic analyses. a) Bio-informatically predicted congener structures. b) Representative synthetic analysis. c) BGC-specific substrate predictions.

Table 1: Syn-BNP MIC values [μg mL^{-1}] against ESKAPE pathogens and *E. coli*.

PABA compounds	Gram-positive		Gram-negative				
	E. faecium com15	*S. aureus* SH1000	*K. pneumoniae* 10031	*A. baumannii* 17978	*P. aeruginosa* PA01	*E. cloacae* 13047	*E. coli* 25922
albicidin	2	0.5	>8	0.125	1	8	0.015
PABA48	0.125	0.25	0.125	0.125	>8	>8	0.125
PABA70	1	1	1	4	>8	>8	2
PABA57	>8	>8	1	>8	>8	>8	2
PABA34	>8	>8	>8	2	2	2	0.0075
PABA157	>8	0.5	>8	0.125	1	8	0.03
PABA95	8	4	>8	0.5	8	8	0.25
ciprofloxacin	0.5	0.5	0.0019	0.125	0.125	0.06	0.0075

fold more potent than albicidin against *S. aureus*, *E. faecium*, and *K. pneumoniae*, respectively. Interestingly, unlike albicidin, all three cystobactamid analogues (PABA48, PABA70, and PABA57) showed potent activity against *K. pneumoniae*, with MIC values ranging from 0.125 to 1 μg mL^{-1}. Based on a time-dependent killing analysis against *E. coli*, all six syn-BNPs caused rapid cell death (Figure S9).

The most variable position among these BGC predicted analogues was the A position, which involved four different building blocks (Figure 3). PABA48, PABA70, and PABA57 are the first examples of analogues with a proteinogenic amino acid appearing at this position. This arrangement places a positively charged amine at the N-terminus of the antibiotic. In contrast, all previously characterized congeners contain a neutral PABA

or MCA at this position. The difference in potency between PABA48 and PABA70, which differ only by the proteinogenic amino acid at their N-termini, suggests this position is particularly important for potency. Traditional structure–activity relationship (SAR) studies have found that a small H-bond acceptor at the *para* position of the phenyl ring in this position is important for antibacterial activity.[16] As this is not found in the most potent broad-spectrum analogue PABA48, it appears that nature has found an orthogonal strategy for increasing potency in this family of antibiotics.

DNA gyrase resistance: Consistent with cystobactamid and albicidin, all syn-BNPs inhibited DNA gyrase in vitro (Figure S8). As mutations in DNA gyrase commonly confer resistance to DNA gyrase inhibitors, we tested each syn-BNP against *E. coli* containing two different GyrA mutations (S83L, D87G) that are commonly seen in fluoroquinolone (i.e. ciprofloxacin) resistant clinical isolates.[17] Neither mutation conferred cross resistance to any syn-BNP (Figure 4a). Although other mutations would likely arise with clinical introduction of this class of antibiotics, it is appealing that our current analogues retain activity against existing problematic DNA gyrase variants.

AlbD-encoded resistance: One key naturally occurring resistance mechanism for this class of antibiotics is cleavage of the amide bond between the D and E monomers by AlbD-like endopeptidases.[4,18] As circumventing common antibiotic resistance mechanisms is likely to play a key role in the natural selection of antibiotic congeners, we reasoned that some congeners might have evolved to evade AlbD-mediated cleavage. To determine the susceptibility of our BGC-inspired analogues to AlbD-encoded resistance, the MIC of each syn-BNP against *E. coli* engineered to express AlbD was compared to that of *E. coli* containing an empty expression vector. All syn-BNPs, with the exception of PABA34, showed at least a 100-fold increase in MIC against the AlbD-expressing strain (Figure 4b, Table S2), which suggests that PABA34 was resistant to proteolytic cleavage by AlbD. To test this hypothesis in vitro, AlbD was overexpressed in *E. coli* BL21 and purified as a 6-His protein. Recombinant AlbD was then used to digest each albicidin-like syn-BNP. The amount of AlbD-mediated cleavage after 20 minutes—beyond which we saw little additional cleavage—was determined for each syn-BNP by HPLC with UV detection (Figure 4c). The extent to which AlbD cleaved each substrate was compared to the extent it cleaved albicidin to give a relative cleavage ratio (RCR) for all three albicidin-like syn-BNPs. Both PABA95 and PABA157 had RCRs close to 1. PABA34, on the other hand, had an RCR of 0.38, which indicates, as anticipated from the in vivo resistance experiment, that PABA34 was less susceptible than albicidin to AlbD cleavage.

PABA34 is unique in that it contains the modified PABA (AHMBA) at the D position directly adjacent to the endopeptidase cleavage site, which suggests that this modification prevents hydrolysis by AlbD, thereby circumventing AlbD resistance (Figure 4d). To test this hypothesis, we synthesized a version of PABA95 (PABA95-2) where the PABA at the D position was replaced with AHMBA. As seen with PABA34, this structure was not susceptible to AlbD-encoded resistance (Figure 4e). Against *E. coli*, the MIC for

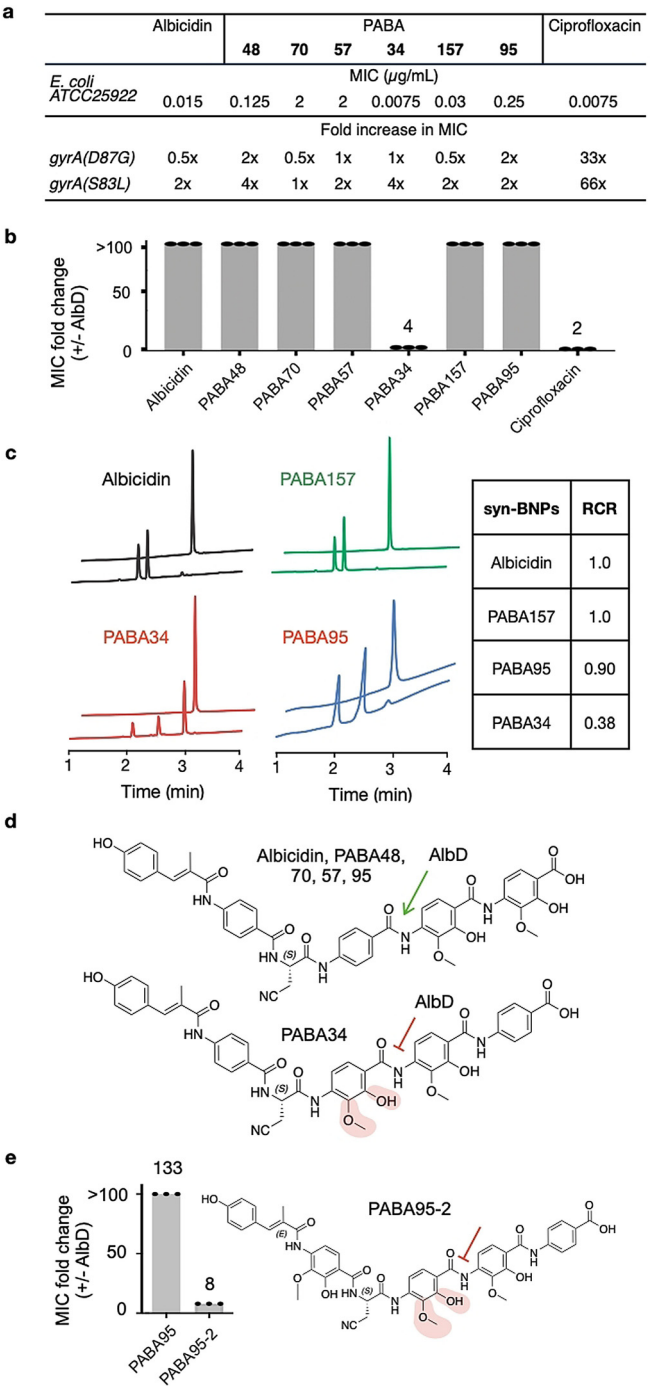


Figure 4. Syn-BNPs have different resistance profiles. a) Activity of syn-BNPs against ciprofloxacin-resistant *E. coli*. b) MIC fold difference between *E. coli* that either expresses or does not express AlbD. c) HPLC analysis of PABA34 and albicidin digested by AlbD. d) Schematic representation of AHMBA protection from AlbD cleavage. e) A PABA95 analogue containing AHMBA is protected from AlbD cleavage.

PABA95 increased by 133-fold upon expression of the *albD* gene. In the case of PABA95-2, the MIC only increased by 8-fold ((i.e. 0.06 $\mu g\,mL^{-1}$ to 0.5 $\mu g\,mL^{-1}$); Table S2).

In summary, bio-informatic analysis of PABA-encoding BGCs cloned from soil metagenomes guided our synthesis of

albicidin and cystobactamid analogues that demonstrated increased potency and different spectra of activity as well as the ability to circumvent resistance conferred by AlbD endopeptidase degradation. Traditional synthetic efforts to address AlbD resistance have focused primarily on replacing the amide bond at the AlbD cleavage site with bioisosteres.[5,13,16] Although this provided resistance to AlbD cleavage, these structures often showed reduced antibiosis against a number of Gram-negative pathogens.[13,16] In contrast, PABA34 was not only resistant to AlbD cleavage, it also retained a high affinity for DNA gyrase (Figure S8) and showed potent activity against most Gram-negative ESKAPE pathogens. Future synthetic efforts to further optimize resistance to AlbD cleavage while maximizing activity against Gram-negative pathogens will likely benefit from mimicking the natural incorporation of a highly substituted PABA monomer (such as AHMBA) at the D position. Our results highlight the potential utility of metagenome-derived BGCs to guide the synthesis of improved natural product variants, especially variants that overcome clinically relevant resistance while maintaining potent antibiosis.

Acknowledgements

This work was supported by the National Institutes of Health (grants 1U19AI142731 and 5R35GM122559)

Conflict of Interest

The authors declare no conflict of interest.

Keywords: albicidin · antibiotics · cystobactamid · endopeptidase cleavage enzymes · metagenomes

[1] S. Cociancich, A. Pesic, D. Petras, S. Uhlmann, J. Kretz, V. Schubert, L. Vieweg, S. Duplan, M. Marguerettaz, J. Noell, I. Pieretti, M. Hugelland, S. Kemper, A. Mainz, P. Rott, M. Royer, R. D. Sussmuth, *Nat. Chem. Biol.* **2015**, *11*, 195 – 197.

[2] S. Baumann, J. Herrmann, R. Raju, H. Steinmetz, K. I. Mohr, S. Huttel, K. Harmrolfs, M. Stadler, R. Muller, *Angew. Chem. Int. Ed.* **2014**, *53*, 14605 – 14609; *Angew. Chem.* **2014**, *126*, 14835 – 14839.

[3] S. M. Hashimi, M. K. Wall, A. B. Smith, A. Maxwell, R. G. Birch, *Antimicrob. Agents Chemother.* **2007**, *51*, 181 – 187.

[4] L. Vieweg, J. Kretz, A. Pesic, D. Kerwat, S. Gratz, M. Royer, S. Cociancich, A. Mainz, R. D. Sussmuth, *J. Am. Chem. Soc.* **2015**, *137*, 7608 – 7611.

[5] T. Planke, K. Cirnski, J. Herrmann, R. Muller, A. Kirschning, *Chem. Eur. J.* **2020**, *26*, 4289 – 4296.

[6] R. Birch, S. Patil, *Phytopathology* **1983**, *73*, 1368 – 1374.

[7] J. Chu, X. Vila-Farres, D. Inoyama, M. Ternei, L. J. Cohen, E. A. Gordon, B. V. Reddy, Z. Charlop-Powers, H. A. Zebroski, R. Gallardo-Macias, M. Jaskowski, S. Satish, S. Park, D. S. Perlin, J. S. Freundlich, S. F. Brady, *Nat. Chem. Biol.* **2016**, *12*, 1004 – 1006.

[8] S. F. Brady, *Nat. Protoc.* **2007**, *2*, 1297 – 1305; F. Y. Chang, M. A. Ternei, P. Y. Calle, S. F. Brady, *J. Am. Chem. Soc.* **2015**, *137*, 6044 – 6052; H. S. Kang, S. F. Brady, *Angew. Chem. Int. Ed.* **2013**, *52*, 11063 – 11067; *Angew. Chem.* **2013**, *125*, 11269 – 11273.

[9] J. G. Owen, B. V. Reddy, M. A. Ternei, Z. Charlop-Powers, P. Y. Calle, J. H. Kim, S. F. Brady, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11797 – 11802; J. G. Owen, Z. Charlop-Powers, A. G. Smith, M. A. Ternei, P. Y. Calle, B. V. Reddy, D. Montiel, S. F. Brady, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 4221 – 4226.

[10] B. V. Reddy, A. Milshteyn, Z. Charlop-Powers, S. F. Brady, *Chem. Biol.* **2014**, *21*, 1023 – 1033.

[11] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493 – 505.

[12] Y. Wang, B. Schnell, R. Muller, T. P. Begley, *Methods Enzymol.* **2018**, *606*, 199 – 216; Y. Wang, T. P. Begley, *J. Am. Chem. Soc.* **2020**, *142*, 9944 – 9954.

[13] I. Behroz, P. Durkin, S. Gratz, M. Seidel, L. Rostock, M. Spinczyk, J. B. Weston, R. D. Sussmuth, *Chem. Eur. J.* **2019**, *25*, 16538 – 16543.

[14] D. Kerwat, S. Gratz, J. Kretz, M. Seidel, M. Kunert, J. B. Weston, R. D. Sussmuth, *ChemMedChem* **2016**, *11*, 1899 – 1903.

[15] W. A. M. Elgaher, M. M. Hamed, S. Baumann, J. Herrmann, L. Siebenburger, J. Krull, K. Cirnski, A. Kirschning, M. Bronstrup, R. Muller, R. W. Hartmann, *Chem. Eur. J.* **2020**, *26*, 7219 – 7225; S. Huttel, G. Testolin, J. Herrmann, T. Planke, F. Gille, M. Moreno, M. Stadler, M. Bronstrup, A. Kirschning, R. Muller, *Angew. Chem. Int. Ed.* **2017**, *56*, 12760 – 12764; *Angew. Chem.* **2017**, *129*, 12934 – 12938; M. Moeller, M. D. Norris, T. Planke, K. Cirnski, J. Herrmann, R. Muller, A. Kirschning, *Org. Lett.* **2019**, *21*, 8369 – 8372; T. Planke, M. Moreno, S. Huttel, J. Fohrer, F. Gille, M. D. Norris, M. Siebke, L. Wang, R. Muller, A. Kirschning, *Org. Lett.* **2019**, *21*, 1359 – 1363.

[16] G. Testolin, K. Cirnski, K. Rox, H. Prochnow, V. Fetz, C. Grandclaudon, T. Mollner, A. Baiyoumy, A. Ritter, C. Leitner, *Chem. Sci.* **2020**, *11*, 1316 – 1334.

[17] S. Bansal, V. Tandon, *Int. J. Antimicrob. Agents* **2011**, *37*, 253 – 255.

[18] M. J. Walker, R. G. Birch, J. M. Pemberton, *Mol. Microbiol.* **1988**, *2*, 443 – 454.

Manuscript received: April 8, 2021
Revised manuscript received: August 5, 2021
Accepted manuscript online: August 5, 2021
Version of record online: September 7, 2021

