

Original citation:

Schwalen, Christopher J., Hudson, Graham A., Kosol, Simone, Mahanta, Nilkamal, Challis, Gregory L. and Mitchell, Douglas A.. (2017) In vitro biosynthetic studies of bottromycin expand the enzymatic capabilities of the YcaO superfamily. *Journal of the American Chemical Society*, 139 (50). pp. 18154-18157.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/96377>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

This document is the Accepted Manuscript version of a Published Work that appeared in final form in *Journal of the American Chemical Society*, copyright © American Chemical Society after peer review and technical editing by the publisher.

To access the final edited and published work see <http://dx.doi.org/10.1021/jacs.7b09899>

A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP url' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

In vitro biosynthetic studies of bottromycin expand the enzymatic capabilities of the YcaO superfamily

Christopher J. Schwalen[†], Graham A. Hudson[†], Simone Kosol[§], Nilkamal Mahanta^{†‡}, Gregory L. Challis^{§, #}, Douglas A. Mitchell^{†‡*}

[†]Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA. [§]Department of Chemistry and Warwick Integrative Synthetic Biology Center, University of Warwick, Coventry CV4 7AL, UK. [‡]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 West Gregory Drive, Urbana, Illinois 61801, USA. [#]Department of Biochemistry and Molecular Biology and ARC Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Victoria 3800, Australia

Supporting Information Placeholder

ABSTRACT: The bottromycins belong to the ribosomally synthesized and posttranslationally modified peptide (RiPP) family of natural products. Bottromycins exhibit unique structural features, including a hallmark macrolactamidine ring and thiazole heterocycle for which divergent members of the YcaO superfamily have been biosynthetically implicated. Here we report the in vitro reconstitution of two YcaO proteins, BmbD and BmbE, responsible for the ATP-dependent cyclodehydration reactions that yield thiazoline- and macrolactamidine-functionalized products, respectively. We also establish the substrate tolerance for BmbD and BmbE and systematically dissect the role of the follower peptide, which we show serves a purpose similar to canonical leader peptides in directing the biosynthetic enzymes to the substrate. Lastly, we leverage the expanded capabilities of YcaO proteins to conduct an extensive bioinformatic survey to classify known YcaO chemistry. This analysis predicts new functions remain to be uncovered within the superfamily.

Ribosomally synthesized and posttranslationally modified peptides (RiPPs) are a growing natural product class with remarkable structural diversity.¹ RiPP precursor peptides are direct gene products that undergo various posttranslational modifications by locally encoded biosynthetic enzymes.² RiPP precursor peptides are functionally bipartite with the modifications typically being installed on the C-terminal “core peptide” region of the precursor. The N-terminal “leader peptide” region frequently contains recognition sequences that are bound by the biosynthetic enzymes. The bottromycin RiPP class possesses an unusually high density of posttranslational modifications on a small (8-mer) core peptide³⁻⁶ and inverts the RiPP precursor paradigm by displaying an N-terminal core peptide and C-terminal “follower peptide” (Figure 1). The biosynthetic gene cluster for bottromycin was reported in 2012 by four independent groups, which demonstrated that the sequences of both the precursor peptide and the tailoring genes are strongly conserved in several *Streptomyces* species (Figure S1 and Table S1).⁷⁻¹⁰

Despite considerable progress in characterizing several posttranslational modifications present in bottromycin,^{6,8-11} the formation of the macrolactamidine, which involves Gly1 and Val4, as well as a thiazoline derived from Cys8, have remained enigmatic (Figure 1).

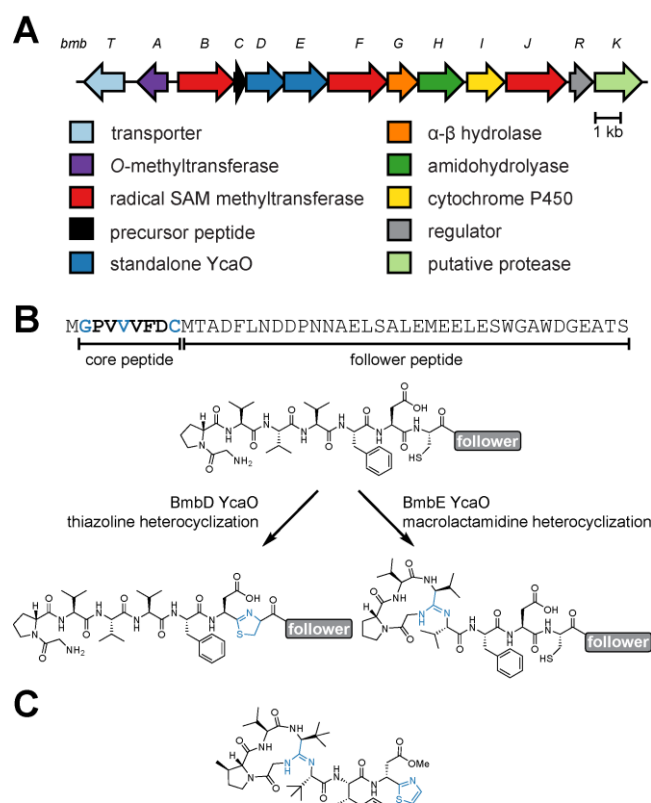


Figure 1 (A) Organization of the bottromycin biosynthetic gene cluster in *Streptomyces bottropensis*. (B) Precursor peptide sequence and reactions catalyzed by the two YcaO proteins (BmbD and BmbE). (C) Structure of bottromycin A2.

Deletion of three genes encoding class B radical-S-adenosylmethionine (SAM) methyltransferases produced bottromycin analogs lacking methyl groups on the β -carbons of Pro2, Val3/Val4 and Phe6.⁷⁻⁹ However, two proteins belonging to the YcaO superfamily (BmbD and BmbE, protein family PF02624)¹² lacked clearly defined functions, but were hypothesized to be involved in modifications requiring activation of the peptide backbone.^{7,8} Work from our group has demonstrated that members of the YcaO superfamily utilize ATP to directly *O*-phosphorylate backbone amides resulting in the formation of azoline heterocycles using the nucleophilic side chains of Cys, Ser, and Thr residues.^{13,14} Individual deletion of

bmbE and *bmbH* homologs in *Streptomyces scabies* were capable of producing a species consistent with a methylated, thiazoline-containing linear peptide, suggesting BmbE and BmbH cooperate to form the macrolactamidine. However, the precise roles of these proteins, as well as the origin of the thiazoline, has remained unresolved.¹¹

All previous *in vitro* characterization of the YcaO superfamily has focused on biosynthetic gene clusters that encode a YcaO partner protein from either the E1 ubiquitin-activating enzyme^{15,16} or ocin-ThiF-like protein family (PF00899).¹⁴ These partner proteins are found either as N-terminal fusions to the YcaO or as discrete proteins. Both varieties contain a ~90-residue domain known as the RiPP precursor peptide recognition element (RRE) that physically engages the peptide substrate and delivers the core region to the biosynthetic enzymes.¹⁷ In contrast, bottromycin gene clusters contain RRE domains on the radical-SAM methyltransferases (**Figure 1**) and the BmbD and BmbE proteins bear minimal sequence similarity to canonical YcaOs outside of the conserved ATP-binding site (**Figure S2**).¹⁴ Intrigued by the putative involvement of RRE-independent YcaOs in macrolactamidine and thiazoline formation, we sought to characterize the catalytic activity of BmbD and BmbE *in vitro*.

BmbC:GPVVVFDCMTADFLNDDPNNNAELSALEMEELESWGAWDGEATS

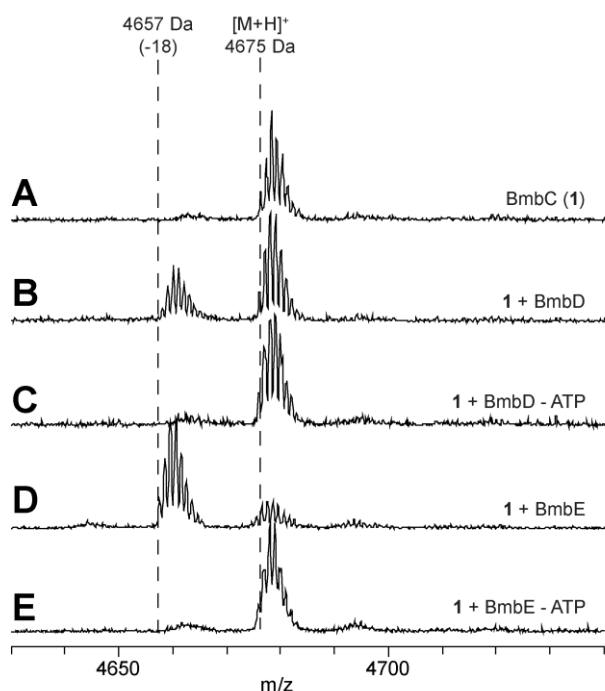


Figure 2 MALDI-TOF-MS analysis of (A) precursor peptide BmbC calc. m/z 4675 (B) Reaction with BmbD yields ions with an m/z value consistent with dehydration. (C) BmbD reaction omitting ATP. (D) Same as panel B but with BmbE. (E) BmbE reaction omitting ATP.

The *bmbD* and *bmbE* genes were cloned into an expression vector that affords an N-terminal fusion to maltose-binding protein (**Tables S2-S3**). The corresponding proteins were overproduced in *Escherichia coli* and purified by amylose affinity chromatography (**Figure S3**). The precursor peptide, BmbC, was overproduced in a form that lacks the N-terminal Met residue. Treatment of BmbC with BmbD resulted in the ATP-dependent loss of 18 Da, consistent with dehydration as assessed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS, **Figure 2**). An identical result was obtained when treating BmbC with BmbE. These findings are in accord with the known chemistry catalyzed by YcaO cyclodehydratases.² Furthermore, the Glu-rich,

ATP-binding motifs conserved amongst YcaOs were essential for BmbD and BmbE processing, as substitution with Ala abolished activity (**Figures S2, S4**). Unlike canonical E1/ocin-ThiF-dependent, azoline-forming YcaOs,¹³ BmbD and BmbE do not contain Pro-rich C-termini; instead, their C-termini are relatively basic. Modest truncations ($\Delta 20$) to the C-termini did not abolish activity for either enzyme, which stands in contrast to YcaOs with Pro-rich C-termini (**Figure S5**).

To determine which dehydration reaction was catalyzed by which protein, we exploited the fact that the thiazoline product contains a primary amino group (Gly1), whereas the macrolactamidine product contains a thiol group (Cys8). Thus, the products of the BmbD and BmbE reactions were subjected to an amine-specific reductive amination with formaldehyde and borane and thiol-specific alkylation with 2-bromoethylamine and subsequently analyzed by MALDI-TOF-MS. As expected, the thiazoline product underwent dimethylation on the Gly1 N-terminus while the macrolactamidine product was alkylated on the Cys8 side chain with no cross-products observed. Moreover, a BmbC-C8A variant was processed by BmbE, but not BmbD, which together implicated BmbD in thiazoline formation and BmbE in macrolactamidine formation (**Figure S6, Table S4**). High-resolution electrospray ionization tandem MS (HR-ESI-MS/MS) confirmed the locations of these posttranslational modifications (**Figure S7, Table S5**).

NMR spectroscopic analysis of a trypsin-treated BmbC-N15K variant was performed to further characterize the BmbC starting material and the BmbD/BmbE reaction products. ¹H chemical shifts and the absence of characteristic nuclear Overhauser effect correlations in DMSO suggested the absence of secondary structure in unmodified BmbC. Comparison of the spectra for BmbC before and after BmbD processing showed that the only significant differences were in the region around Cys8, including: the amide protons for Asp7, Met9, Thr10, and Ala11; the C α protons for Asp7 and Met9; the C β protons of Met9 (**Figure S8**). Moreover, the signals for the C α and C β protons of Cys8 were shifted to values typically observed for thiazoline-containing peptides (**Table S6**).¹⁸ The same analysis was performed for BmbE-treated BmbC, where the C β and C δ protons for Val3-5, as well as the C α , C β , and C γ protons of Pro2 were strongly shifted in comparison with unmodified BmbC (**Figure S8**). More modest changes were detected for the Phe6 and Asp7 protons. Collectively, the chemical labeling, processing of BmbC-C8A, HR-ESI-MS/MS, and NMR data allow the unambiguous assignment of thiazoline formation to BmbD and macrolactamidine formation to BmbE.

Typical YcaO cyclodehydratases bind their substrates via an RRE-containing partner protein;¹⁴ however, BmbD and BmbE were catalytically competent *in vitro* in the absence of any partner protein. For BmbE, this finding is in contrast to a previous *in vivo* metabolomics study that implicated BmbH in macrolactamidine formation.¹¹ To assess if BmbH or BmbG (another functionally unassigned protein, **Figure 1**) enhanced BmbE activity *in vitro*, the proteins were individually supplied to the YcaO reaction mixture. Neither addition of BmbH or BmbG resulted in more product formation. Identical results were obtained when BmbD reactions were supplemented with BmbG/H (**Figure S9**).

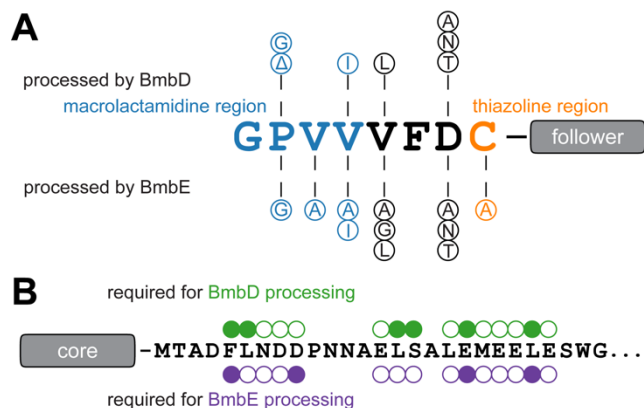


Figure 3 Substrate tolerance of BmbD and BmbE (A) Core peptide variants processed by each YcaO. Δ indicates amino acid deletion at this position. (B) Activity of BmbC variants with Ala substitution in the follower peptide. The Phe13-Asp17 and Glu22-Glu32 regions were targeted based on their similarity to the lanthipeptide and cyanobactin recognition sequences, respectively. Circles indicate sites of Ala substitution. Filled circles indicate variants that inhibited BmbD or BmbE processing.

A previous heterologous expression study¹⁰ showed that the bottromycin biosynthetic pathway was largely intolerant towards accepting BmbC variants. To determine if the same was true *in vitro*, we evaluated 21 BmbC variants as BmbD/BmbE substrates using a MALDI-TOF-MS endpoint assay (Figures 3, S7). As observed for other RiPPs,¹⁹ the variants were more widely tolerated as substrates *in vitro* as opposed to *in vivo* (Table S4). However, BmbD was still only able to process Cys8 and substitution of distal residues could abolish thiazoline formation. Similarly, BmbE exclusively formed a macrolactamidine with the fourth residue of BmbC and attempts to alter macrocycle size were unsuccessful (Table S2).

We next investigated if the BmbC follower peptide played a role in substrate recognition. Stop codons (denoted as asterisks) were introduced to produce C-terminal truncations of BmbC at intervals of five to identify the minimal sequence compatible with BmbD/BmbE processing. Consistent with follower peptide sequence conservation (Figure S1), BmbC-G35* was processed while E30* was not (Figure S10). Analysis of additional truncations confirmed that residues 1-34 of BmbC represents the minimum substrate for BmbD and BmbE.

The follower peptide of BmbC contains regions that are reminiscent of lanthipeptide and cyanobactin recognition sequences (Figure 3).^{20,21} To determine the relevance of these regions, 15 residues were individually replaced with Ala and assessed as BmbD/BmbE substrates. Despite their sequence divergence and the lack of a recognizable RRE, many of the same BmbC residues were important for both BmbD and BmbE processing (Figures 3, S11). To directly measure the binding affinity towards BmbD/BmbE, we prepared double variant BmbC-C8S/S43C to allow for the selective installation of fluorescein near the C-terminus of the peptide (Figure S12). Using a fluorescence polarization assay, the dissociation constants (K_D) for BmbD and BmbE were estimated to be greater than 50 μ M (Figure S13). As mentioned above, the bottromycin biosynthetic gene cluster does indeed encode proteins with RREs, but these domains are fused to the C-termini of the rSAM methyltransferases. To evaluate their role in BmbC binding, we excised the RRE domains from BmbB and BmbF and subjected the proteins to the same fluorescence polarization assay. The binding affinities for BmbC were significantly tighter relative to the YcaOs with K_D values of 876 ± 67 nM (BmbB) and 597 ± 25 nM (BmbF). Further studies are warranted to reconcile how these binding and activity data fit into the yet-unresolved order of events during bottromycin biosynthesis.

Having confirmed a new activity for an RRE-independent YcaO *in vitro*, we wondered how often YcaOs are encoded in genomic contexts consistent with macrolactamidine and azoline formation. For this, we employed RODEO (Rapid ORF Description and Evaluation Online)²² to catalog the co-occurrence frequency of protein families for the ~13,000 YcaO proteins deposited in GenBank (Figure S14). Less than one-third of all YcaOs co-occur with a E1/ocin-ThiF-like member, which is the arrangement found in several RiPP classes, including linear azol(in)e-containing peptides, thiopeptides, and cyanobactins. A separate subset of YcaO proteins (~8% of the total) co-occur with “TfuA”, a protein of unknown function. Recent work from our group has shown that the TfuA-YcaO pair found in methanogenic archaea is responsible for installing a thio-glycine moiety in methyl-coenzyme M reductase.²³ This result suggests that the TfuA-YcaO present in the thioviridamide biosynthetic gene cluster would similarly be responsible for thioamide formation.²⁴

To shed light on the potential function of the remaining 60+% of YcaOs, we analyzed the local genomic region for Pfam co-occurrence (Table S7) as well as generated sequence similarity networks to assess phylogenetic and relatedness trends (Figure S15). We then created custom profile hidden Markov models²⁵ based on known BmbD and BmbE homologs to discriminate between azoline- and amidine-forming YcaOs and reanalyzed the networks based on the resulting expectation values (Figures S16-17). While very few of the uncharacterized YcaOs were similar to BmbD (azoline-forming YcaO), one-third of the sequences showed significant similarity to BmbE, suggesting that amidine-forming YcaOs may be more prevalent than currently appreciated.²⁶ The data further predict that RRE-independent azoline-forming YcaOs may be relatively rare (Figure S18).

In summary, our work has delineated the roles of BmbD, BmbE, and the follower peptide in bottromycin biosynthesis. BmbE has been confirmed to carry out a new enzymatic reaction for the YcaO superfamily. Our bioinformatic analysis not only implicates YcaO involvement in a diversity of new natural product biosynthetic gene clusters, but also a significant potential for new chemistry that would expand upon the established roles in azoline, thioamide, and amidine formation.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information, including experimental methods and supporting figures, is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author

* douglasm@illinois.edu

Notes

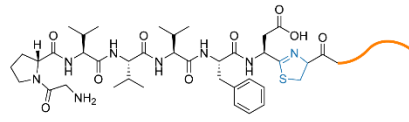
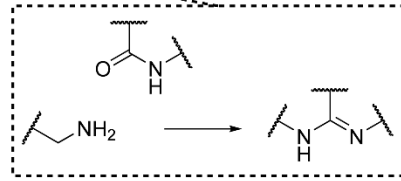
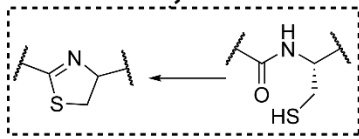
The authors declare no competing financial interests.

ACKNOWLEDGMENT

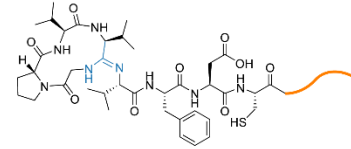
This work was supported by the National Institutes of Health (GM097142 to D.A.M.), the Chemistry–Biology Interface Training Program (GM070421 to C.J.S.), and the Seemon Pines Fellowship from the Department of Chemistry at the University of Illinois at Urbana-Champaign (to G.A.H.), Wolfson Research Merit Award (WM130033) from the Royal Society (to G.L.C.).

REFERENCES

- (1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; et al. *Nat. Prod. Rep.* **2013**, *30*, 108.
- (2) Burkhardt, B. J.; Schwalen, C. J.; Mann, G.; Naismith, J. H.; Mitchell, D. A. *Chem. Rev.* **2017**, *117*, 5389.
- (3) Schipper, D. *J. Antibiot.* **1983**, *36*, 1076.
- (4) Kaneda, M. *J. Antibiot.* **2002**, *55*, 924.
- (5) Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Omura, S. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 914.
- (6) Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Muller, R.; Koehnke, J. *Chembiochem* **2016**, *17*, 2286.
- (7) Crone, W. J. K.; Leeper, F. J.; Truman, A. W. *Chem. Sci.* **2012**, *3*, 3516.
- (8) Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L. *Chem. Sci.* **2012**, *3*, 3522.
- (9) Huo, L.; Rachid, S.; Stadler, M.; Wenzel, Silke C.; Müller, R. *Chem. Biol.* **2012**, *19*, 1278.
- (10) Hou, Y.; Tianero, M. D.; Kwan, J. C.; Wyche, T. P.; Michel, C. R.; Ellis, G. A.; Vazquez-Rivera, E.; Braun, D. R.; Rose, W. E.; Schmidt, E. W.; Bugni, T. S. *Org. Lett.* **2012**, *14*, 5050.
- (11) Crone, W. J.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 9639.
- (12) Finn, R. D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonhammer, E. L.; Tate, J.; Punta, M. *Nucleic Acids Res.* **2014**, *42*, D222.
- (13) Dunbar, K. L.; Chekan, J. R.; Cox, C. L.; Burkhardt, B. J.; Nair, S. K.; Mitchell, D. A. *Nat. Chem. Biol.* **2014**, *10*, 823.
- (14) Dunbar, K. L.; Tietz, J. I.; Cox, C. L.; Burkhardt, B. J.; Mitchell, D. A. *J. Am. Chem. Soc.* **2015**, *137*, 7672.
- (15) Dunbar, K. L.; Melby, J. O.; Mitchell, D. A. *Nat. Chem. Biol.* **2012**, *8*, 569.
- (16) Cox, C. L.; Doroghazi, J. R.; Mitchell, D. A. *BMC Genomics* **2015**, *16*, 1.
- (17) Burkhardt, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. *Nat. Chem. Biol.* **2015**, *11*, 564.
- (18) Mocek, U.; Beale, J. M.; Floss, H. G. *J. Antibiot.* **1989**, *42*, 1649.
- (19) Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. *J. Am. Chem. Soc.* **2016**, *138*, 15511.
- (20) Ortega, M. A.; Velasquez, J. E.; Garg, N.; Zhang, Q.; Joyce, R. E.; Nair, S. K.; van der Donk, W. A. *ACS Chem. Biol.* **2014**, *9*, 1718.
- (21) Sardar, D.; Pierce, E.; McIntosh, J. A.; Schmidt, E. W. *ACS Synth. Biol.* **2015**, *4*, 167.
- (22) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. *Nat. Chem. Biol.* **2017**, *13*, 470.
- (23) Nayak, D. D.; Mahanta, N.; Mitchell, D. A.; Metcalf, W. W. *eLife* **2017**, *6*, e29218.
- (24) Izawa, M.; Kawasaki, T.; Hayakawa, Y. *Appl. Environ. Microbiol.* **2013**, *79*, 7110.
- (25) Finn, R. D.; Clements, J.; Arndt, W.; Miller, B. L.; Wheeler, T. J.; Schreiber, F.; Bateman, A.; Eddy, S. R. *Nucleic Acids Res.* **2015**, *43*, W30.
- (26) Metelev, M.; Osterman, I. A.; Ghilarov, D.; Khabibullina, N. F.; Yakimov, A.; Shabalin, K.; Utkina, I.; Travin, D. Y.; Komarova, E. S.; Serebryakova, M.; Artamonova, T.; Khodorkovskii, M.; Konevega, A. L.; Sergiev, P. V.; Severinov, K.; Polikanov, Y. S. *Nat. Chem. Biol.* **2017**, *13*, 1129.



thiazoline



macrolactamidine