Macroamidine formation in bottromycins is catalyzed by a divergent YcaO enzyme

Laura Franz,[†,‡] Sebastian Adam,[†,‡] Javier Santos-Aberturas,[§] Andrew W. Truman,[§] and Jesko Koehnke[†,*]

[†]Workgroup Structural Biology of Biosynthetic Enzymes, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research, Saarland University, Campus Geb. E8.1, 66123 Saarbrücken, Germany

§ Department of Molecular Microbiology, John Innes Centre, Colney Lane, Norwich, NR4 7UH, UK

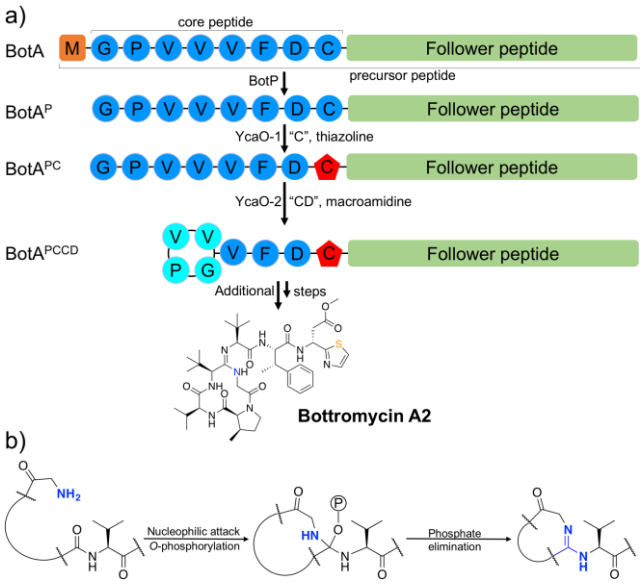
Supporting Information Placeholder

ABSTRACT: The YcaO superfamily of proteins catalyzes the phosphorylation of peptide backbone amide bonds, which leads to the formation of azolines and azoles in ribosomally synthesized and post-translationally modified peptides (RiPPs). Bottromycins are RiPPs with potent antimicrobial activity, and their biosynthetic pathway contains two divergent, stand-alone YcaO enzymes, IpoC and PurCD. From an untargeted metabolomics approach, it had been suggested that PurCD acts with a partner protein to form the 12-membered macroamidine unique to bottromycins. Here we report the biochemical characterization of IpoC and PurCD. We demonstrate that IpoC installs a cysteine-derived thiazoline, while PurCD alone is sufficient to create the macroamidine structure. Both enzymes are catalytically promiscuous and we generated 10 different macroamidines. Our data provide important insights into the versatility of YcaO enzymes, their ability to utilize different nucleophiles and provide a framework for the creation of novel bottromycin derivatives with enhanced bioactivity.

Bottromycins (Scheme 1) are ribosomally synthesized and post-translationally modified peptide (RiPP) natural products that bind to the A site of the 50S ribosome, thus inhibiting prokaryotic protein synthesis and represent a novel class of antibiotics.[1] They are derived from the precursor peptide BotA, which is unique amongst the RiPPs family in that it has a C-terminal "follower" rather than the canonical N-terminal "leader" peptide.[2-5] BotA is tailored by a series of enzymatic (and possibly non-enzymatic) chemical transformations and their order has recently been proposed based on data from an untargeted metabolomics approach (Scheme 1).[6] The bottromycin biosynthetic gene cluster contains two unusual, stand-alone YcaO-domain enzymes, BotC and BotCD (sequence identity of 26%, Figure S2).

YcaO-domain proteins were enigmatic until they were linked to the formation of azol(in)e heterocycles in the biosynthesis of microcin B17.[7-8] Biosynthetic pathways responsible for the production of RiPPs frequently contain enzymes with a YcaO domain, which catalyze azoline formation in an ATP-dependent cyclodehydration reaction.[9]

Scheme 1. a) Bottromycin biosynthesis with the proposed functions of the two YcaO enzymes. b) Use of a terminal amine by a YcaO enzyme as a nucleophile to achieve macroamidine formation.



Vital for this process in all RiPPs systems studied to date is the presence of RiPP recognition element (RRE) domains.[10] Without the RRE domain of a specific system, which can either be present on a stand-alone E1 ubiquitin activating-like superfamily protein or part of a multidomain heterocyclase, the heterocyclase activity is minimal.[11] The YcaO domain of the responsible enzyme catalyzes the nucleophilic attack of a Cys/Ser/Thr sidechain on the preceding backbone carbonyl to form a

hemiorthoamide.[12] This intermediate is then *O*-phosphorylated[12] (or pyrophosphorylated[13]) followed by an elimination reaction to yield the azoline (Figure S1a). It has been proposed (and one example reported[14]) that other nucleophiles may also be used by this enzyme family. Despite the presence of two YcaO enzymes, bottromycins contain only one azole and the gene cluster encodes no E1-like protein or known RRE domain. An untargeted metabolomic approach using mass spectral networking identified the order of reactions in bottromycin biosynthesis.[6] This study revealed that BotC was responsible for the formation of the thiazoline, while BotCD (in cooperation with a hydrolase) catalyzed the macrocyclization reaction to yield the macroamidine.[6] Macroamidine formation can be viewed as analogous to azoline formation, with the amino-terminus of the precursor peptide BotA functioning as the nucleophile (Scheme 1b and Figure S1b).

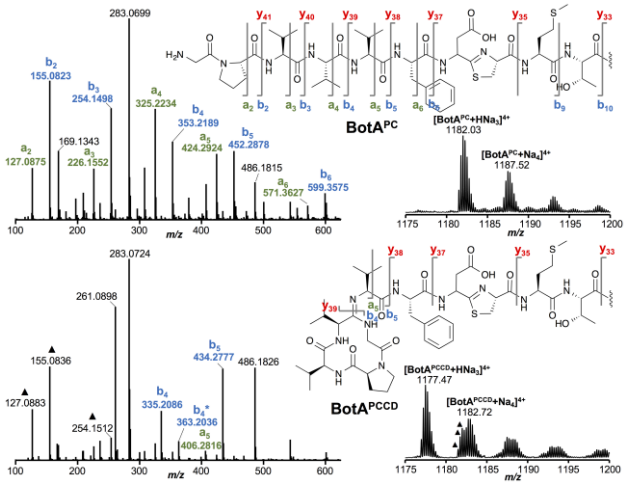
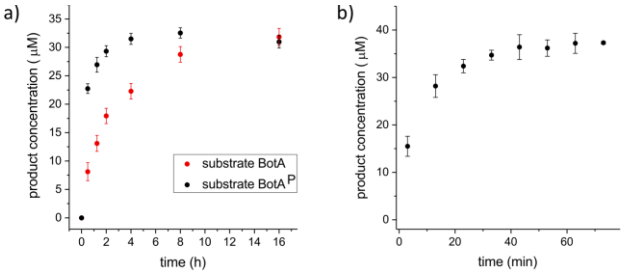


Figure 1. MS^2 characterization of BotA^P after incubation with IpoC and BotA^{PC} after incubation with PurCD. Full details of the MS^2 fragmentation are provided in Figures S4-S6 and Tables S1-S5. Signals in the BotA^{PCCD} spectra that likely come from residual BotA^{PC} are labelled with triangles.

To investigate this potentially novel YcaO-domain catalysis, we sought to characterize BotC and BotCD *in vitro*, but neither protein (from *Streptomyces* sp BC16019) could be expressed in a soluble form. In contrast, the homologs IpoC (76% sequence identity to BotC) from *Streptomyces ipomoeae* and PurCD (81% sequence identity to BotCD) from *Streptomyces purpureus* were tractable. Both proteins originate from gene clusters that are homologous and highly similar to those reported to produce bottromycins (Figure S3). When BotA was incubated with IpoC and ATP/Mg^{2+}, we observed a loss of 18 Da (Figure S4a), and the conversion of Cys9 to thiazoline was confirmed by tandem MS (MS^2) (Figure S4b and Table S2). We treated BotA with BotP (BotA^P) to make the glycine amino group available as a nucleophile for macrocyclization and repeated the experiments.[15] The reaction with IpoC also resulted in thiazoline formation at Cys9 (Figure 1 and S5, Table S4). For PurCD incubated with BotA^P we also observed a loss of 18 Da, but the yield was too low for MS^2 analysis (Figure S5a). Based on metabolomic data, we expected the heterocyclized intermediate (BotA^{PC}) to be a better PurCD substrate.[6] When we reacted purified BotA^{PC} with PurCD and ATP/Mg^{2+} we observed macroamidine formation (BotA^{PCCD}), which was confirmed by MS^2, where characteristic[6] macroamidine fragments of *m/z* 335.21, 363.20 and 434.28 were observed (Figures 1 and S6, Table S5). Neither BotA nor heterocyclized BotA^C were substrates for PurCD. These data established IpoC as the heterocyclase, while PurCD functions as the macrocyclase. PurCD is the first reported case of a YcaO enzyme able to a) utilize nitrogen as a nucleophile and b) form larger ring systems. Our data indicates that for efficient catalysis removal of the N-terminal methionine precedes heterocyclization, which is followed by macrocylization.

A detailed analysis of the IpoC-catalyzed heterocyclization showed that 50 µM BotA^P required 8 h to reach completion when it was incubated with 5 µM IpoC and 5 mM ATP/Mg^{2+} at 37 °C (Figure 2a). In contrast, the reaction using unprocessed precursor peptide BotA required 16 h (Figure 2a). The heterocyclization rate was calculated to be 0.1 heterocycles per enzyme per minute with BotA^P, while the reaction using BotA was roughly half this rate (Figure S7), indicating that while IpoC can process both versions of the precursor peptide, BotA^P is preferred. Changes in the reaction conditions (temperature, pH, salt concentration) did not lead to accelerated heterocycle formation (data not shown).



c)

	IpoC (T_m = 33 °C)		PurCD (T_m = 24 °C)	
	T_m (°C)	Conversion of peptide	T_m (°C)	Conversion of IpoC-cyclized peptide
BotA^P	45	✓	32	✓
BotA^P-5AA	42	✓	29	✓
BotA^P-10AA	35	✓	23	✗
BotA^P-15AA	31	✗	23	✗
BotA-F	41	not possible	31	not possible
BotA-F+C9	50	not possible	33	not possible

Figure 2. a) Time-course of IpoC-catalyzed heterocyclization of BotA (red) and BotA^P (black) at pH 7.4. b) Time-course of PurCD-catalyzed macroamidine formation at pH 9.5. c) Effects of BotA^P truncations and follower peptide on IpoC and PurCD stability (pH 8.5) and turnover (pH 7.4 and 9.5, respectively).

The macrocyclization reaction of PurCD (10 µM) with BotA^{PC} (50 µM) was approximately 15 % complete after 90 min at pH 7.5 (Figure S8a). We hypothesized that the reaction may be accelerated by increasing the pH of the reaction as increasing the pH results in a larger propor-

tion of deprotonated substrate N-terminus, which is then available as the nucleophile for macroamidine formation. We analyzed PurCD activity between pH 7 and 11, which is more than 1 pH unit above the pK_a of the amino group of glycine (pK_a 9.6), so most of the substrate should be available in its neutral (nucleophilic) form. We observed a steady increase in product formation with a maximum at pH 9.5 (Figure S8b), with ~85% turnover after 1 h (Figures 2b and S8c). We wondered if the decrease in enzyme activity above pH 9.5 was due to enzyme degradation. When we analyzed PurCD stability by measuring its melting temperature (T_m) in thermal shift assays (TSAs), we observed a steady decrease of enzyme stability with increasing pH and were unable to record a T_m at pH 11 (Figure S9a). We therefore believe that pH 9.5 represents the best balance between substrate availability and enzyme stability. Interestingly, we the enzyme and ATP-dependent reopening of the macroamidine after extended incubations of the reactions, which was accelerated at lower pH (Figure S8). After 12 h no $BotA^{PCCD}$ could be observed in the samples. It appears possible, that the hydrolase identified as essential for macroamidine formation *in vivo*[6] influences this reaction's equilibrium and thus prevents re-opening of the ring. IpoC and PurCD appear to be tractable for the use in *in vitro* systems, but the surprising enzyme-induced hydrolysis of the macroamidine by PurCD must be taken into account in such systems to accomplish satisfactory yields.

A hallmark of YcaO domains characterized to date is their requirement to bind their substrate's leader peptide for efficient catalysis via an RRE domain. The bottromycin gene cluster does not encode a protein with homology to known RREs, and we investigated if IpoC and PurCD were able to bind the follower peptide of BotA. Based on published precursor peptide – enzyme structures, binding to the leader appears to stabilize the corresponding enzyme.[13,16] This feature would allow TSAs to be used as a convenient tool to probe these critical interactions. In TSAs, one can measure the T_m of a protein in the absence and presence of ligands and stabilizing interactions between protein and ligand lead to an increase in T_m. IpoC and PurCD had T_ms of 33 and 24 °C, respectively, which increased dramatically after the addition of the substrate $BotA^P$/$BotA^{PC}$ (Figure S9b and Table S6). To verify that this effect was not exerted by the core peptide, we repeated the experiment using follower peptide BotA-F. The effect of the follower on IpoC was almost as pronounced as for $BotA^P$ (Figure S10a). In the case of PurCD, the differences between $BotA^{PC}$ and BotA-F were within the experimental error (Figures 2c and S10b, Table S6). We asked if the presence of the substrate cysteine had an effect on IpoC stability and observed a drastic increase in T_m for IpoC + BotA-F with an additional N-terminal cysteine (Figures 2c and S10a, Table S6). To identify the follower region responsible for binding we produced a series of systematic truncations (Figures 2c and S10, Tables S7 and S8). $BotA^P$ with the 5 C-terminal residues removed ($BotA^P$-5AA) had a slightly less stabilizing effect on IpoC and PurCD when compared to $BotA^P$ and was processed with similar efficiency. Removal of the next 5 residues ($BotA^P$-10AA) abolished the stabilizing effect on both enzymes, severely hampered processing by IpoC and abolished PurCD activity. The next truncation, $BotA^P$-15AA, had no stabilizing effect on either enzyme and was not processed. These data indicated that the region between residues 30 and 39 is critical for the interaction between BotA and both enzymes. We produced a series of point mutants covering this area (Figure S11, Tables S7 and S8), which all showed a reduced stabilizing effect in TSAs, with E31R and W35A reducing the effect to within experimental error. However, all point mutants could be processed by both enzymes (Figure S12). Only $BotA^P$E28R, a mutation slightly outside the suggested range, abolished stabilization and processing, highlighting the importance of this residue. The observation that impairment of substrate binding (implied via T_ms) does not abolish processing has been observed for other YcaO enzymes. A deeper understanding will require the determination of C and CD crystal structures in complex with BotA-F. However, we have demonstrated that the stabilizing effect of the follower or leader peptides observed in TSAs may be used to group tailoring enzymes into early (requiring leader/follower) and late (acting after leader/follower removal) enzymes, to reduce the complexity of reconstituting RiPPs pathways *in vitro*.

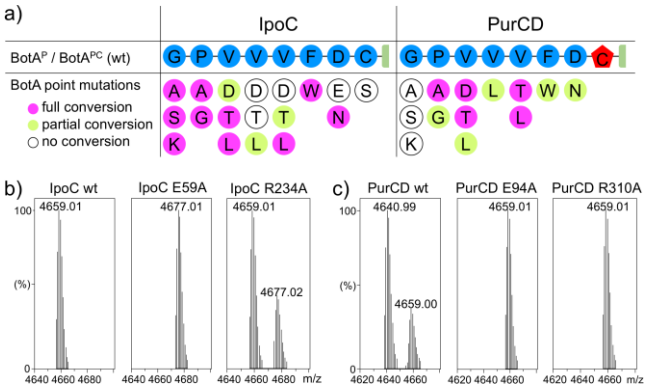


Figure 3. a) Tolerance of IpoC and PurCD to changes in the core peptide of BotA. Conversion rate given in relation to wt substrate. Effect of b) IpoC and c) PurCD ATP/Mg^{2+} binding residue mutations on enzyme activity.

The YcaO enzymes studied to date have shown a great tolerance for changes in their respective substrates[9] and we wondered if IpoC and PurCD would also display catalytic promiscuity. To this end we tested a set of 18 core peptide mutants (Figure S13, Tables S7 and S8). IpoC was unable to produce oxazoline from a C9S mutant, but processed 13 other point mutants (Figures 3a and S14). The most sensitive position appeared to be V5. The 13 heterocyclized peptides were tested for reactivity with PurCD. We observed macroamidine formation for 10 of these (Figures 3a and S15, Table S8) with varying conversion rates. Our data indicates that Gly2 cannot be

modified, and the requirement of a glycine at this position was likely, since aside from D-amino acids[17] it introduces the least strain into the ring system and is the nucleophile in macrocyclization.

To date, only two YcaO domain proteins have been crystallized in complex with a nucleotide cofactor. Both are very distantly related to IpoC and PurCD (pairwise protein sequence identity < 20%), but sequence alignment nevertheless revealed that key ATP and Mg^{2+} binding residues are mostly conserved (Figure S16). To explore IpoC and PurCD ATP binding residues, we designed a series of point mutants of both enzymes based on the sequence alignments (Figure S16, Tables S7 and S8). In IpoC we mutated the putative Mg^{2+} binding residue E59 to A or R, while we swapped the putative ATP (phosphate) binding residue R234 to A or E. All four mutants could be expressed, purified and stabilized by $BotA^P$ (TSAs), indicating correct protein folding (Figure S17 and Table S6). The stabilizing effect of ATP/Mg^{2+} remained as observed for wild-type IpoC, within experimental error. Mutations of E59 reduced activity to less than 5% of wt enzyme, while mutations of R234 reduced enzymatic activity (Figure 3b and S18). Equivalent *in vivo* mutants (Table S9) were consistent with this *in vitro* data. In PurCD we selected E94 (to A or R) and R310 (to A or E) for mutations. Here, only the conservative alanine substitutions could be expressed. Both mutant proteins were stabilized by $BotA^P$ and ATP/Mg^{2+} as observed for wt PurCD, but were unable to process substrate (Figures 3c, S19 and S20, Tables S7 and S8). These deleterious *in vitro* results for PurCD were confirmed when the same mutations were tested *in vivo* (Table S9), which all provided the same metabolite profile as a gene deletion. The conservation of key ATP/Mg^{2+} binding residues in highly divergent enzymes from this family may be used to validate predicted YcaO domain enzymes *in vitro* or *in vivo* and offers a way to selectively abolish enzyme function without disrupting protein folding.

We have demonstrated that the catalytic repertoire of YcaO enzymes also includes the formation of macroamidines. This enzyme family has also been implied in the formation of thioamide bonds,[18] and it will be fascinating to discover their true chemical scope. The promiscuity of these two key biosynthetic enzymes *in vitro* is an important cornerstone towards the development of an efficient system to generate bottromycin variants with enhanced stability and activity.

ASSOCIATED CONTENT

Supporting Information.

The Supporting Information is available free of charge on the ACS Publications website.

Detailed experimental procedures and supporting information (PDF).

AUTHOR INFORMATION

Corresponding Author

*jesko.koehnke@helmholtz-hzi.de

ORCID

Jesko Koehnke: 0000-0002-7153-1365

Author Contributions

‡L.F. and S.A. contributed equally.

Funding Sources

J.K. thanks the DFG for an Emmy Noether Fellowship (KO 4116/3-1). This work was also supported by BBSRC grant BB/M003140/1 (A.W.T. and J.S.-A.) and a Royal Society University Research Fellowship (A.W.T.).

Notes

No competing financial interests have been declared.

ACKNOWLEDGMENT

We thank Dr. David Auerbach, Dr. Gerhard Saalbach and Daniel Sauer for assistance with the LC-MS measurements. We are grateful to Dr. Rafael Guimaraes da Silva and Prof. Rolf Müller for helpful discussions.

REFERENCES

(1) Otaka, T.; Kaji, A., *J Biol Chem* 1976, *251* (8), 2299-306.
(2) Crone, W. J. K.; Leeper, F. J.; Truman, A. W., *Chemical Science* 2012, *3* (12), 3516-3521.
(3) Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., *Chemical Science* 2012, *3* (12), 3522-3525.
(4) 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* (19), 5050-3.
(5) Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Muller, R., *Chem Biol* 2012, *19* (10), 1278-87.
(6) 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* (33), 9639-43.
(7) Li, Y. M.; Milne, J. C.; Madison, L. L.; Kolter, R.; Walsh, C. T., *Science* 1996, *274* (5290), 1188-93.
(8) Bayer, A.; Freund, S.; Nicholson, G.; Jung, G., *Angewandte Chemie International Edition in English* 1993, *32* (9), 1336-1339.
(9) Burkhart, B. J.; Schwalen, C. J.; Mann, G.; Naismith, J. H.; Mitchell, D. A., *Chem Rev* 2017, *117* (8), 5389-5456.
(10) Dunbar, K. L.; Tietz, J. I.; Cox, C. L.; Burkhart, B. J.; Mitchell, D. A., *J Am Chem Soc* 2015, *137* (24), 7672-7.
(11) Dunbar, K. L.; Chekan, J. R.; Cox, C. L.; Burkhart, B. J.; Nair, S. K.; Mitchell, D. A., *Nat Chem Biol* 2014, *10* (10), 823-9.
(12) Dunbar, K. L.; Melby, J. O.; Mitchell, D. A., *Nat Chem Biol* 2012, *8* (6), 569-75.
(13) Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W.; Jaspars, M.; Naismith, J. H., *Nat Chem Biol* 2015, *11* (8), 558-563.
(14) Koehnke, J.; Morawitz, F.; Bent, A. F.; Houssen, W. E.; Shirran, S. L.; Fuszard, M. A.; Smellie, I. A.; Botting, C. H.; Smith, M. C.; Jaspars, M.; Naismith, J. H., *Chembiochem* 2013, *14* (5), 564-7.
(15) Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Muller, R.; Koehnke, J., *Chembiochem* 2016, *17* (23), 2286-2292.
(16) Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van der Donk, W. A.; Nair, S. K., *Nature* 2015, *517* (7535), 509-12.
(17) Yongye, A. B.; Li, Y.; Giulianotti, M. A.; Yu, Y.; Houghten, R. A.; Martinez-Mayorga, K., *J Comput Aided Mol Des* 2009, *23* (9), 677-89.
(18) Izawa, M.; Kawasaki, T.; Hayakawa, Y., *Appl Environ Microbiol* 2013, *79* (22), 7110-3.