The Biochemistry of Bottromycin Biosynthesis
and
Ribosomal Natural Product Engineering

Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

von

Laura Franz

Saarbrücken

2021

*"I want to know why. Why everything. I don't know the answers,
but a few days ago I didn't know there were questions."*

Terry Pratchett

Meinen Eltern zum 60. Geburtstag

Danksagung

Allen voran möchte ich mich bei meinem Doktorvater Jun.-Prof. Jesko Köhnke bedanken, der mir das Vertrauen gegeben hat meine Promotion in seiner Arbeitsgruppe durchzuführen und die Möglichkeit geschaffen hat spannende Projekte unter exzellenten Bedingungen zu bearbeiten. Ich möchte mich insbesondere für seine immerwährende Unterstützung und gute wissenschaftliche Betreuung bedanken. Gleichzeitig habe ich mich stets herausgefordert gefühlt, wodurch ich mich Wissenschaftlerin immens weiterentwickeln konnte.

Zudem möchte ich mich bei Prof. Dr. Andriy Luzhetskyy für die wissenschaftliche Begleitung und die Übernahme des Zweitgutachtens bedanken.

Ein großer Dank geht auch an Dr. Sebastian Adam, Andreas Klein und Dr. Asfandyar Sikandar aus der Arbeitsgruppe Strukturbiologie biosynthetischer Enzyme für die Zusammenarbeit im Labor und an Projekten.

Darüber hinaus möchte ich mich bei Ellen Merckel und Christina Decker, Mark Caspari und Michael Roth, Frank Jacob und Christian Zeuner, sowie bei den Mitgliedern der Arbeitsgruppe Mikrobielle Naturstoffe aus dem HIPS für die die organisatorische Unterstützung und Hilfe bei wissenschaftlichen Fragestellungen bedanken. Im Speziellen danke ich den Mitgliedern des Analytik-Teams für die gute Zusammenarbeit, sowie für die Einführung in die Bedienung von Massenspektrometern und Hilfe bei Fragen im Bereich der MS.

Ganz besonders danke ich meinen Kollegen und Freunden Chantal Bader, Dr. Alexander Kiefer, Patrick Haack, Yvonne Kaiser, Dr. Sebastian Adam, Nicolas Frank, Christine Walt, Viktoria George, Kathrin Andres, Florian Häckel, Lisa Schmidt, Sabrina Stabler und den Basketballerinnen des TV St. Ingbert Saints für die schöne Zeit in den letzten 4 Jahren, inner- und außerhalb des Labors.

Zuletzt möchte ich mich bei den wichtigsten Personen, meinen Eltern und meiner Schwester, für ihr Verständnis, die stete Unterstützung und Ermutigung bedanken. Ohne euch hätte ich das nicht geschafft.

Zusammenfassung

Ribosomal synthetisierte und post-translational modifizierte Peptide (RiPPs) sind eine große Klasse von Naturstoffen mit vielfältigen Bioaktivitäten. Trotz ihres ribosomalen Ursprungs weisen RiPPs eine hohe strukturelle Diversität auf, die das Resultat post-translational agierender Enzyme sind. Ein besseres Verständnis der unterliegenden Biosynthese-Mechanismen ermöglicht die Entdeckung von neuen RiPPs und ihre Weiterentwicklung.

In dieser Arbeit wurde die Biosynthese des Grundgerüstes der Bottromycine mit einem *in vitro* Ansatz aufgeklärt. Die Charakterisierung der an der Synthese dieses antibiotisch wirkenden RiPP beteiligten Enzyme brachte einige Überraschungen hervor. So konnte zum Beispiel gezeigt werden, dass eine α/β-Hydrolase als Aspartat-Epimerase fungiert. Die Ergebnisse dieser Arbeit haben zu einer Revision des für Bottromycine vorgeschlagenen Biosyntheseweges geführt und unser biochemisches Verständnis der RiPP Biosynthese grundsätzlich erweitert. Zusätzlich kann die *in vitro* Biosynthese der Bottromycin Grundstruktur nun zur Herstellung von Derivaten und weiterer Entwicklung dieser Naturstoffe genutzt werden.

RiPP Enzyme sind in der Lage verschiedenste, synthetisch schwer zugängliche Modifikationen in Peptide einzufügen. Die Kombination dieser Enzyme aus unverwandten Biosynthesewegen ist jedoch nur sehr eingeschränkt möglich. In dieser Arbeit wurde mit biochemischen Methoden die Machbarkeit eines neuen Konzeptes zur Herstellung von Neonaturstoff-Bibliotheken gezeigt.

Summary

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a major class of natural products with a wide variety of biological activities. They owe their structural diversity to extensive post-translational modifications of a ribosomal peptide through RiPP biosynthetic enzymes. A better understanding of these enzymes facilitates the discovery, production and engineering of RiPPs.

In this thesis, the biosynthesis of the bottromycin core scaffold was elucidated with an *in vitro* approach. The characterization of the enzymes involved in the biosynthesis of this antibiotic RiPP revealed several unexpected enzymatic activities and dependencies. We could, for example, demonstrate that a predicted α/β-hydrolase functioned as an aspartate epimerase. The results of this thesis have led to a revision of the proposed biosynthetic route for bottromycins and given new insights into RiPP biosynthesis in general. In addition, ready access to the bottromycin core scaffold can now be used to produce bottromycin derivatives and thus develop these natural products.

RiPP enzymes are able to install a wide variety of modifications into peptides that are difficult to access synthetically. The combination of these enzymes from unrelated pathways is very challenging. In this thesis biochemical methods were employed to provide proof-of-concept for a new approach that is going to enable the production of neo-natural product libraries.

Veröffentlichungen der Dissertation

Teile dieser Dissertation wurden mit Genehmigung der entsprechenden Verlage sowie der Naturwissenschaftlich-Technischen Fakultät III, vertreten durch den Mentor der Arbeit, in folgenden Journalbeiträgen veröffentlicht oder sind in im Prozess zur Veröffentlichung.

Franz, L.,[+] Adam, S.,[+] Santos-Aberturas, J., Truman, A. W., and Koehnke, J., Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme. *Journal of the American Chemical Society* **2017,** *139* (50), 18158-18161.

Sikandar, A.,[+] **Franz, L.**,[+] Melse, O., Antes, I., and Koehnke, J., Thiazoline-Specific Amidohydrolase PurAH Is the Gatekeeper of Bottromycin Biosynthesis. *Journal of the American Chemical Society* **2019,** *141* (25), 9748-9752.

Sikandar, A.,[+] **Franz, L.**,[+] Adam, S., Santos-Aberturas, J., Horbal, L., Luzhetskyy, A., Truman, A. W., Kalinina, O. V., and Koehnke, J., The bottromycin epimerase BotH defines a group of atypical α/β-hydrolase-fold enzymes. *Nature Chemical Biology* **2020**, 16, 1013-1018.

Adam, S.,[+] **Franz, L.**,[+] Milhim, M., Bernhardt, R., Kalinina, O. V., and Koehnke, J., Characterization of the Stereoselective P450 Enzyme BotCYP Enables the In Vitro Biosynthesis of the Bottromycin Core Scaffold. *Journal of the American Chemical Society* **2020,** *142* (49), 20560-20565.

Franz, L., and Koehnke, J., Leader peptide exchange to produce hybrid, new-to-nature ribosomal natural products. *Manuscript in revision*.

Franz, L., Kazmaier, U., Truman, A. W., and Koehnke, J., Bottromycins - Biosynthesis, Synthesis and Activity. *Manuscript accepted for publication*.

[+] shared first authorship

Table of Contents

1 Introduction

1.1 Natural products as a source for drugs

Natural products (NP) are defined as substances produced by living organisms.[1] Some definitions narrow this term down to secondary metabolites[2] secreted for defense or communication,[3] or metabolites with biological activity[4]. They originate from plants, animals, fungi, archaea or bacteria and comprise an enormous chemical and structural diversity showing a plethora of diverse biological activities (Figure 1-1).[5] As a result, NPs have been used by humankind for medicinal purposes for thousands of years and continue to be a rich source for drugs.[6, 7] They are often used in medicine e.g. as anti-infectives, immunosuppressants or for cancer treatment, but they also find application in agriculture or the food industry.[6-9]
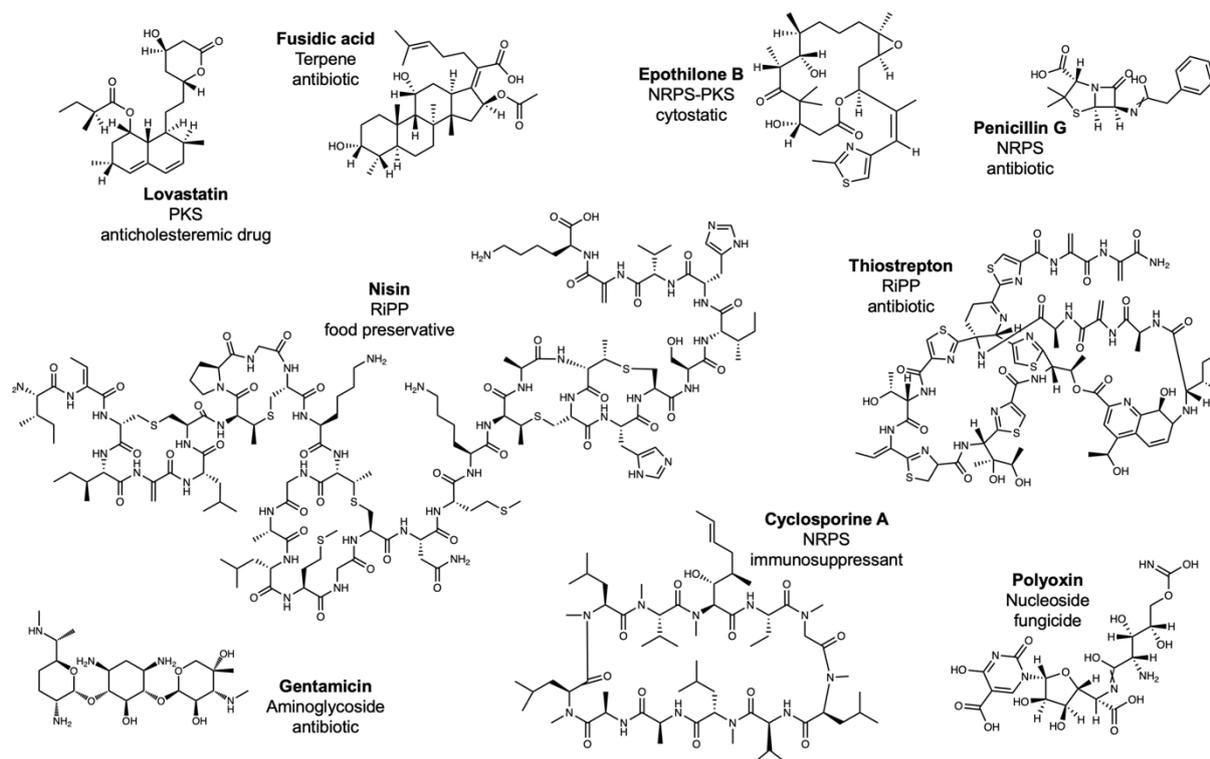
Figure 1-1 | Examples of natural products used in medicine, agriculture or food industry.

The beginning of NP use by humans is considered as evolutionary-based and medicinal plants were used instinctively. The behavior of self-medication is even observed for animals (zoopharmacognosy), which eat or apply selected bioactive plants.[10] The knowledge of medicinal plant usage has grown with input from different cultures.[11-16] The earliest evidence for the medicinal use of NPs was found on clay cuneiforms from Mesopotamia, which are approximately 5000 years old.[3, 16] They describe the use of plant oils or extracts e.g. from

1

cypress, myrrh or poppy juice for the treatment of various diseases including coughs and inflammation.[17]

The 19th century constituted a turning point in human NP use and the beginning of scientific pharmacy, as developments in the field of chemistry allowed the isolation and application of single active ingredients from plants, which could be related to specific pharmaceutical effects.[3, 16] This transition was initiated in 1804, when Friedrich Sertürner isolated morphine from the opium poppy of *Papaver somniferum*, which became the first commercially used pure NP marketed by Merck in 1826.[3, 5, 18, 19] Further examples of isolated plant NPs during this period of time are the anti-malarial alkaloid drug quinine from the bark of *Cinchona* species (1820) and salicin from the bark of the willow tree *Salix alba* (1828).[3, 17, 19] Salicin or rather its active metabolite salicylic acid, further developed to acetylsalicylic acid by Charles Frédéric Gerhardt in 1853, represents the first semi-synthetic NP. Acetylsalicylic acid was first marketed by Bayer as the world-wide known and still intensively used painkiller Aspirin in 1899.[19-21]

One of the most important events in NP research was the discovery of the antibiotic penicillin, produced by the fungus *Penicillium chrysogenum*, by Alexander Fleming in 1928.[3, 22] This discovery led to a transformation of NP discovery, promoting the use of microorganisms as a source for novel drugs.[3, 17] The discovery of penicillin initiated the so-called 'Golden Age' of anti-infective drug discovery, which lasted from the 1940s to the 1970s and led to the discovery of several new classes of antibiotics of microbial origin.[17, 21, 23]

Despite its initial success, the interest in NP research declined after this first boom and several NP departments of big pharmaceutical companies were closed as the threat from bacterial pathogens seemed to be under control. Additionally, bioactivity-guided NP research remained a very time-consuming and resource-intensive process and suffered from a high rediscovery rate of already known molecules.[3, 24] Uncertainties regarding the collection of biological samples resulting from the 1992 Rio Convention on Biological Diversity, held to secure natures diversity, and the decreasing interest of pharmaceutical companies in the not very profitable anti-infective research, almost ended the most successful area for NP research.[25, 26] In parallel, advances in combinatorial chemistry and introduction of high-throughput screenings (HTS) for defined molecular targets led many companies to move from NP extract libraries to the use of synthetic chemical libraries for drug discovery.[25] Nevertheless, NPs and NP-derived or -inspired drugs remained to be an important source for new drugs (Figure 1-2), although the proportion of these compounds in screening efforts constitutes less than 1%.[7, 27] From 1981 to 2019, 67% of all small-molecule approved drugs are NPs, NP derivatives, NP mimics or synthetic drugs with a NP pharmacophore.[7] For anti-bacterial drugs, the proportion of NPs is

even higher.[7] These values emphasize the continued relevance of NP research for the discovery of novel drugs and raises the question: Why are NPs so successful compared to the combinatorial libraries?
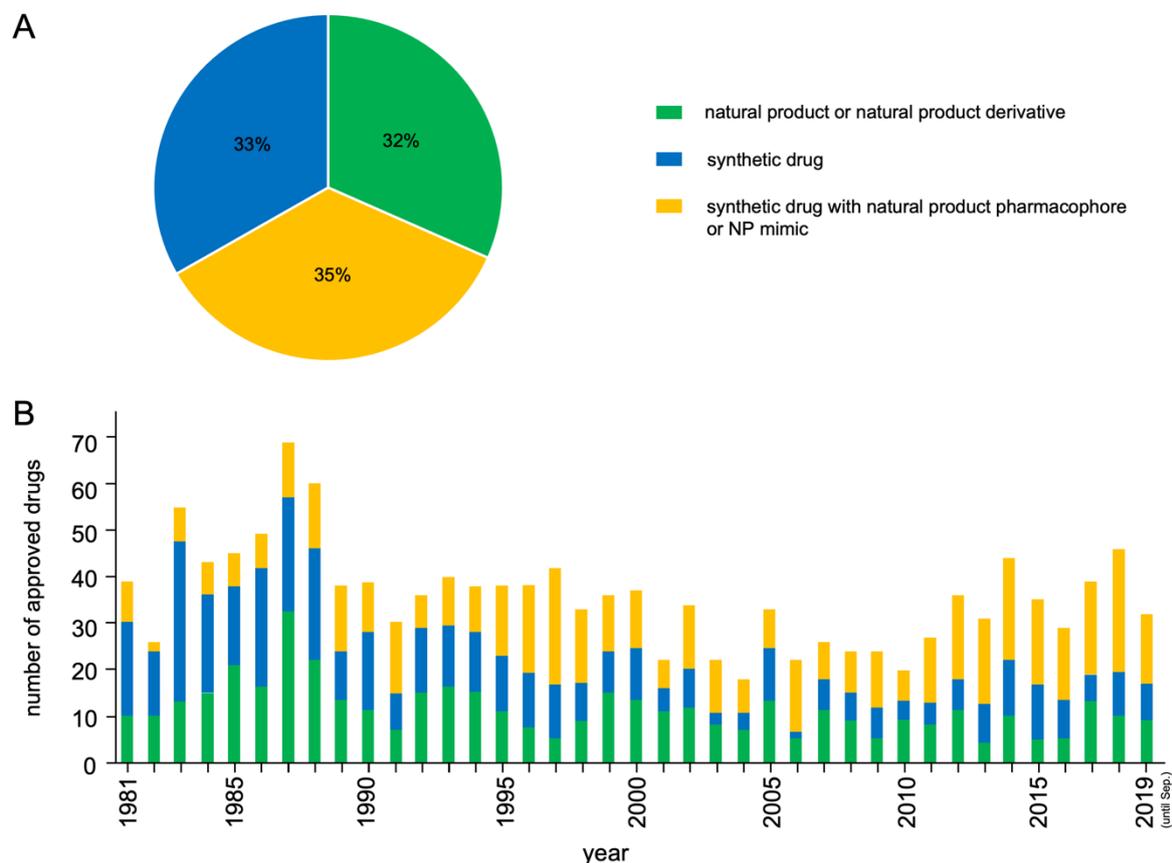
Figure 1-2 | Proportion of small molecule approved drugs from 1981 until (September) 2019 by source (green: NP or NP derivative; blue: synthetic drug; yellow: synthetic drug with NP pharmacophore or NP mimic). Data extracted from Newman *et* Cragg, 2020[7].

NPs are often referred to as 'privileged scaffolds' because they have gone through evolutionary optimization to be biologically active and provide the organisms with a competitive advantage in their natural habitat.[5, 28] While the generation of synthetic combinatorial libraries is usually mainly influenced by the availability of reagents and suitable reactions, NP 'libraries' from plants and bacteria were generated over millions of years under the aspects of available precursors and biosynthetic reactions, but in particular under the aspect of biological utility.[29, 30]

NPs differ from synthetic drug-like molecules as they often show higher structural complexity as well as a larger structural and physicochemical diversity.[31-33] The exact differences of compounds derived from combinatorial chemistry and NPs were determined by analyzing the

structural properties or the so-called 'chemical space' in various studies in the early 2000s. In general, NPs were found to contain less nitrogen, sulfur or halogen but more oxygen atoms.[30, 34] They also show a larger diversity of ring systems, contain more chiral centers, sp^3-hybridized bridgehead atoms and are more rigid. Surprisingly, the fraction of structures with Lipinski's 'rule of 5' violations is comparable between NP and combinatorial drugs.[30, 34, 35] The intrinsic properties of NP scaffolds enable them to strongly and selectively bind to specific molecular targets, which the less stereospecific and more flexible combinatorically synthesized molecules often fail to do.[29, 30] These investigations highlight why NPs and NP core scaffolds are and will continue to be an invaluable sources for new molecules in drug research.[36] This especially holds true when it comes to combating the increasing emergence of anti-microbial resistance (AMR), which is estimated to cause 10 million deaths per year by 2050, surpassing the number of deaths caused by cancer.[37] These projections highlight the urgency for the discovery of novel anti-microbial drugs with novel scaffolds and unprecedented mechanisms of action (MoAs) to treat infectious diseases.[38, 39] Even today, more people die of MRSA infections in the USA than of tuberculosis and HIV/AIDS combined.[38, 40] The therapeutic options are already limited when multidrug- or even pan-antibiotic-resistant pathogens need to be treated, forcing the use of drugs such as colistin, which were previously avoided due to high toxicity.[38] NPs are a fruitful source for finding molecules featuring the properties very much needed to overcome the AMR crisis. However, the motifs responsible for their strong and highly selective binding to specific targets are often difficult to access by synthetic chemistry, which is why there is growing interest in engineering of biosynthetic pathways and the biocatalytic use of NP biosynthetic enzymes for their synthesis. Recent technological advances in bioinformatic, genomic, biotechnological and metabolomic approaches made such approaches feasible to deliver new drugs to fill the antibiotic discovery gap.[27, 31, 41-43]

1.1.1 Peptide natural products

NPs can be divided into different structural classes ranging from peptides, polyketides, terpenoids and alkaloids to amino acid, carbohydrate, and nucleic acid derivatives. Among these classes, peptidic NPs make up a huge fraction of therapeutic NPs and cover the largest chemical space. Prominent examples of approved peptide NPs are the antibiotics penicillin (human drug), thiostrepton (used in veterinary medicine), nisin (used as food preservative) or the immunosuppressant cyclosporine A (Figure 1-1).[7, 24]

Peptide-based therapeutics in general fill the gap between small molecules and biologics, combining the advantages of both compound classes. They are known to be selective, efficient,

well tolerated and safe.[44-47] But there are also major obstacles for peptides such as short plasma half-life due to degradation by proteases and they often show poor oral bioavailability and membrane permeability due to their relatively high polarity and molecular weight.[45, 47-50]

Peptide NPs often contain non-canonical amino acids, backbone modifications and cyclic structures, which improve the pharmacological properties of these molecules. In comparison to their linear and unmodified counterparts, they have increased proteolytic stability as a result of their restrained and unusual backbone dihedral angles. They often show better membrane permeability due to the absence of ionized termini caused by their cyclisation and defined intramolecular hydrogen bonds that keep hydrophilic groups from the molecular surface.[51-53] Furthermore, cyclic peptides bind to their targets very selectively and exhibit better biological activities. Their conformational rigidity decreases the entropy of the molecule and enhances the binding affinity to the molecular target.[50, 54-56] These medium-sized molecules have also been shown to be able to interfere with biologically relevant protein-protein interactions (PPI). PPI usually involve larger shallow protein surface areas without defined binding pockets, which are difficult to address with small molecules, but are targetable by peptides in a manner similar to antibodies.[57-59] The advantageous properties of NP-like peptides led to an increased interest in peptide therapeutics in both, academia and the pharmaceutical industry.[59] NP peptides derive from highly complex biosynthetic machineries, which are also an intriguing starting point for NP engineering and identification of biosynthetic enzymes with exceptional, unprecedent enzyme chemistry for biocatalytical use.[60-65]

1.2 Biosynthesis of peptide natural products

Peptide NPs are either biosynthesized by ribosomal synthesis followed by subsequent post-translational modification or by large multi-enzyme protein complexes, termed non-ribosomal peptide synthases (NRPSs). In bacteria, the genes that encode for the biosynthetic machinery of a NP are, together with those involved in the regulation and export, usually clustered and are referred to as biosynthetic gene clusters (BGCs). The underlying biosynthetic machineries of non-ribosomal peptide (NRP) and ribosomal (RiPP) NPs will be described in detail in the following sections.

1.2.1 Non-ribosomal peptides (NRPs)

NRP synthetases have a modular organization and represent template and biosynthetic machinery at the same time.[66] Each module is responsible for the incorporation of a specific monomeric building block into the growing polypeptide chain. The number and order of the modules usually corresponds to the number and order of the building blocks of the final peptide.[67] These modules are further divided into domains. One minimal elongation module is composed of an adenylation (A) domain, a peptidyl carrier protein (PCP) domain and a condensation (C) domain.[68] The adenylation domain activates its cognate amino acid substrate by adenylation to generate an aminoacyl-AMP derivate, which is then attacked by the 4′-phosphopantetheinyl (PPE) arm of the PCP domain, resulting in a PCP-tethered aminoacyl thioester (Figure 1-3).[69] The condensation (C) domain catalyzes the peptide bond formation between an amino acyl or peptidyl-S-PCP intermediate from the upstream module (module n-1, Figure 1-3) to the amino acyl moiety of the current module (module n, Figure 1-3) and thereby elongates the peptide chain by one amino acid building block.[69, 70] The resulting peptide is elongated in this assembly-line fashion until it reaches the NRPS termination module, which additionally contains a thioesterase (TE) domain that releases the peptide product in linear or cyclized form.[71, 72] NRPS have shown to be able to accept more than 500 unique amino acid building blocks into their peptide scaffolds.[73] Further modifications can be introduced into the substrates by additional tailoring domains, such as methylation, epimerization, and cyclization domains or by independent tailoring enzymes, such as glycosyltransferases.[74-80]

R$_1$ = H or peptide chain
R$_2$, R$_3$ = amino acid side chain
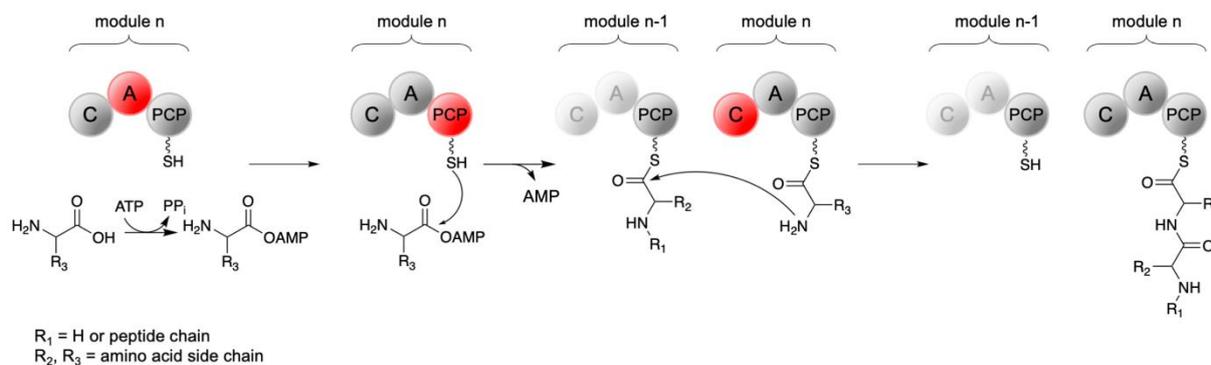
Figure 1-3 | Schematic of peptide chain elongation in NRP biosynthesis. A cycle of peptide chain elongation by a minimal elongation module is initiated by the activation of the respective amino acid by the adenylation (A) domain using ATP. The activated substrate is then attached to the 4′-phosphopantetheinyl group on the peptidyl carrier protein (PCP) domain. Finally, the condensation (C) domain catalyzes the peptide bond formation between the amino acid substrate and the peptide attached to the upstream PCP domain.

1.2.2 Ribosomally and post-translationally modified peptides (RiPPs)

In contrast to NRPs, ribosomally and post-translationally modified peptides (RiPPs) are, at least in nature (see section 1.3.2), limited to the 20 canonical amino acids, as their biosynthesis proceeds via the ribosomal pathway. However, RiPPs possess a similar degree of chemical diversity as NRPs due to an unprecedented array of post-translational modifications (PTMs) introduced by RiPPs biosynthetic enzymes.

The biosynthesis of a RiPP (Figure 1-4) begins with a so-called precursor peptide of ribosomal origin. This precursor peptide, with a length of usually 20-110 amino acids, is in most cases comprised of a N-terminal leader peptide region and a core peptide region.[81] The core peptide is where PTMs are introduced and it becomes the mature RiPP (Figure 1-4). The leader peptide is required for binding by many of the PTM-introducing enzymes to the precursor peptide and can also act as an allosteric effector for the biosynthetic enzymes.[82] In addition, the leader can keep the peptide inactive inside the producing cell or ensures that the PTMs are introduced in the correct order.[83-85] In eukaryotic RiPPs, the leader peptide also contains the signal sequence.[81] Some RiPP precursor peptides also contain a C-terminal follower peptide for the substrate recognition by modifying enzymes. In the unique example of bottromycins (see section 1.2.3), the bottromycin precursor peptide contains no leader peptide but only a follower peptide.
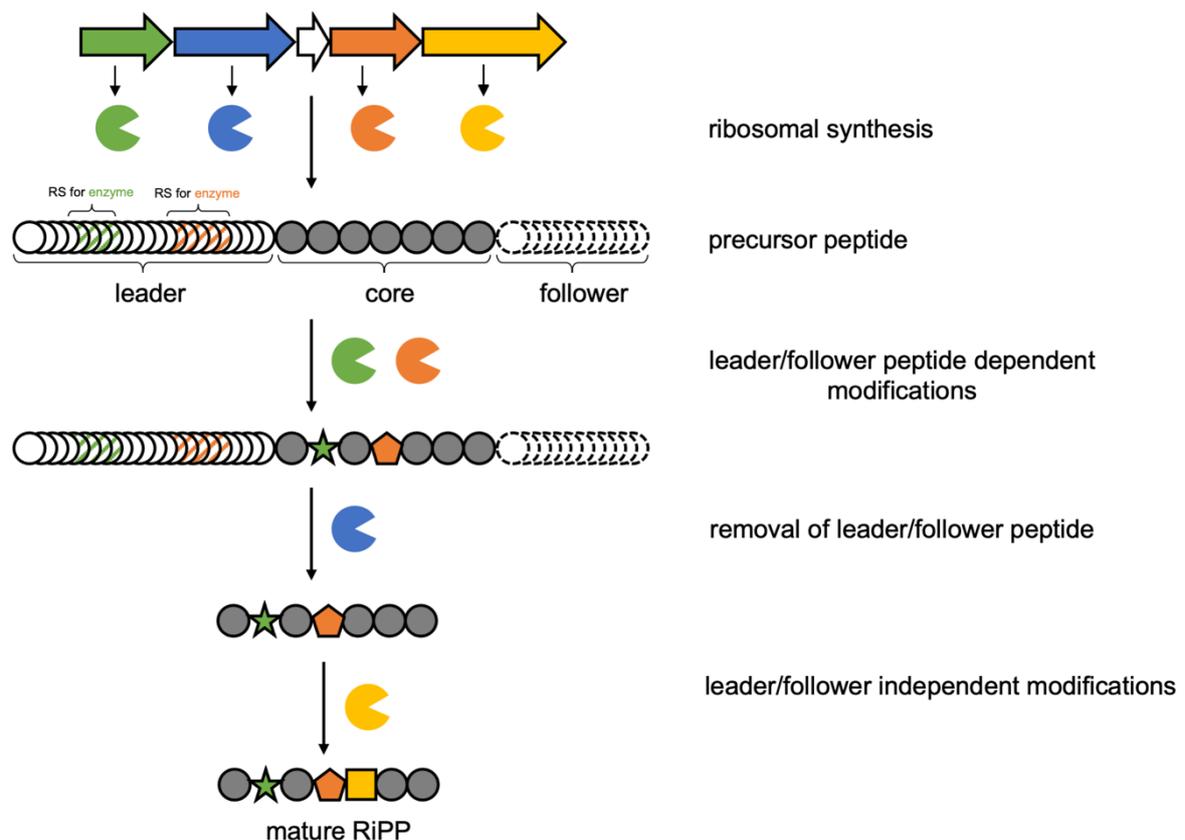
Figure 1-4 | Scheme of RiPP biosynthesis. The genetically encoded precursor peptide is ribosomally synthesized and is comprised of a core peptide region (grey circles) and a leader (and follower) peptide region (white circles). The leader (or follower) peptide contains enzyme specific, so-called, recognition sequences (RS) that faciliate enzyme binding to the peptide. Primary, leader-dependent, enzymes (Pac-Man symbol) introduce characteristic, RiPP class-defining modifications (symbols, e.g. star, pentagon) into the core peptide. After proteolytic removal of the leader (or follower) peptide, the core peptide can be further modified by tailoring enzymes to yield the mature RiPP.

The biosynthetic enzymes encoded in the RiPP BGC play a central role in RiPP biosynthesis, introducing the PTMs that endow RiPPs with structurally diverse modifications crucial for their biological activity by improving proteolytic resistance, affinity/selective binding to the cognate target and membrane permeability (see section 1.1.1). RiPP biosynthetic enzymes can be classified into primary and secondary enzymes. Primary enzymes introduce characteristic, RiPP class-defining PTMs and are leader-dependent. They bind to their respective precursor peptide through specific parts in the leader (or follower) peptide, the so-called recognition sequences (RS).[86] Many primary biosynthetic enzymes recognize and bind to their cognate precursor peptide through a PqqD-like domain, known as RiPP recognition element (RRE), that can exist as a discrete protein or be fused to the biosynthetic enzyme.[87] This element is present in over 50% of the known RiPP classes produced by prokaryotes. As class-independent feature of RiPP BGCs, the RRE represents a common genetic feature that can also be used for the bioinformatic detection of new RiPP classes (RRE-Finder).[88]

After introduction of the PTMs by the leader/follower-dependent enzymes, the leader/follower peptide is removed by proteases. Sometimes this leader/follower peptide removal is associated with N-to-C cyclisation, e.g. during thiopeptide or patellamide biosynthesis.[89-93] In some RiPP classes, the modified core peptide is further altered by secondary (leader-independent) tailoring enzymes. This modification happens after leader/follower removal to yield the mature biologically active RiPP, which is then exported by transporters encoded in the pathway. There are also RiPP classes in which the protease is localized outside of the producing cells or is part of the transporter, keeping the modified peptide inactive inside the host.[81, 86]

Post-translational enzymatic modifications are capable of producing complicated peptide scaffolds that may contain inter alia azol(in)e heterocycles, different macrocyclic scaffolds, thioether bridges, substituted pyridines, C-C crosslinks, thioamide backbones and D-amino acids.[81, 83, 94] Based on their structural features and biosynthetic machinery, RiPPs are grouped in different classes (e.g. lanthipeptides, lasso peptides and thiopeptides). The structural diversity of the mature RiPP scaffolds ranges from small organic molecules such as pyrroloquinoline quinone (PQQ) to large peptides such as nisin or polytheonamides. RiPPs are found in all domains of life and their biological functions are diverse: They include signaling molecules, anti-bacterial, -viral, -cancer compounds and bacterial cofactors. In fact, the function of many RiPPs is yet unknown.[81, 83, 95-99]

The different RiPP PTMs and RiPP classes were recently compiled by Merwin *et al.*[94] and Montalbán-López *et al.*.[83] The number of known post-translational modifications (PTMs) introduced by RiPP biosynthetic enzymes has increased steadily during recent years due the availability of genomic sequences and mining thereof using new bioinformatic tools.[88, 100] While in 2013, 24 RiPP classes had been identified,[81] this number grew to 41 RiPP classes in 2020.[83]

Thousands of unique unknown RiPPs have been found in sequenced bacterial genomes using bioinformatic tools.[94, 101, 102] A recent analysis of the BiG-FAM database that contains 1,225,071 predicted BGCs from a set of 188,622 microbial genomes revealed that 22% of all BGCs encode for RiPPs, 31% for NRPs, 18% for polyketides and 9% for terpenes (Figure 1-5).[103] This analysis likely still underrepresents RiPPs because of the bioinformatic challenges associated with their discovery.[101, 104] The data nevertheless confirms that RiPPs are more widespread than originally anticipated and emphasize the importance of RiPP research for the discovery of novel NPs.
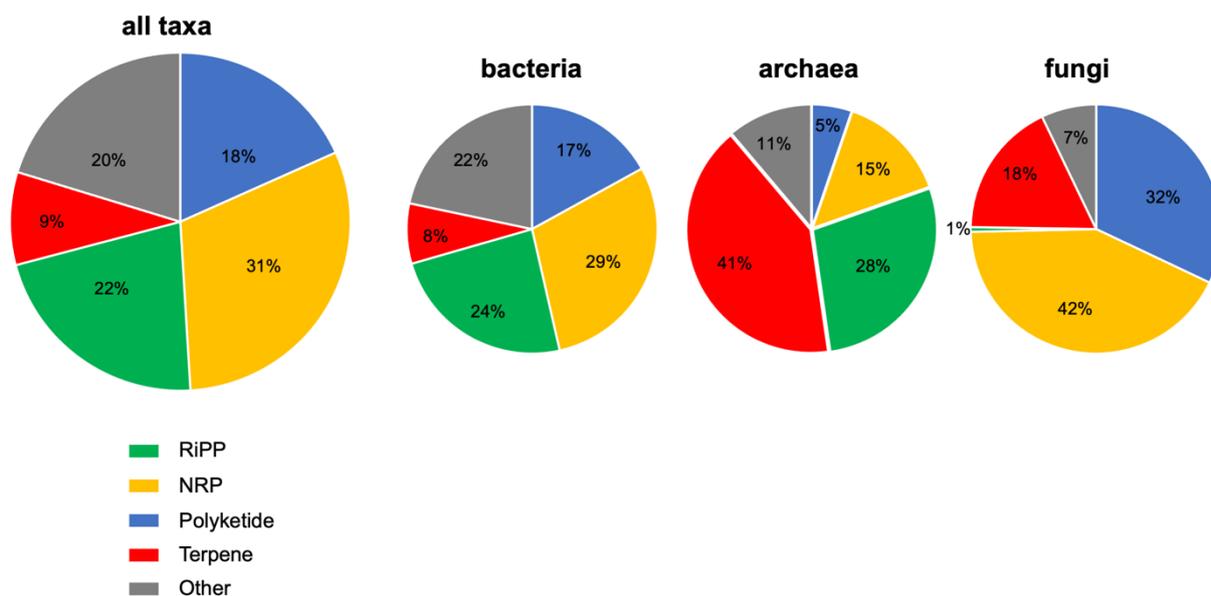
Figure 1-5 | Proportion of BGCs predicted to encode for the major biosynthetic classes from a set of 1,224,563 BGCs (bacteria 1,096,473, archaea 3,998, fungi 124,092). Data from Kautsar *et al*., 2021.[103]

The simple biosynthetic logic of RiPPs (Figure 1-4) and the promiscuity and versatility of RiPP biosynthetic enzymes make them attractive targets for pathway engineering and a source for novel biocatalysts.[45, 82] Different approaches for engineering RiPP pathways are presented in section 1.3.2.

In this work, the in many ways unique RiPP class of bottromycins (section 1.2.3) was studied *in vitro* to elucidate the biosynthetic pathway and unprecedented enzyme functions to enable the generation of bottromycin analogues.

1.2.3 Bottromycins

Bottromycins are a class of macrocyclic peptide NPs that were first isolated in 1955 from *Streptomyces bottropensis*.[105] They show antibacterial activity against Gram-positive bacteria such as the multidrug-resistant human pathogens methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE).[106, 107] Mechanism of action studies revealed that bottromycins inhibit protein synthesis by binding to the A-site of the prokaryotic 50S ribosome.[108-110]
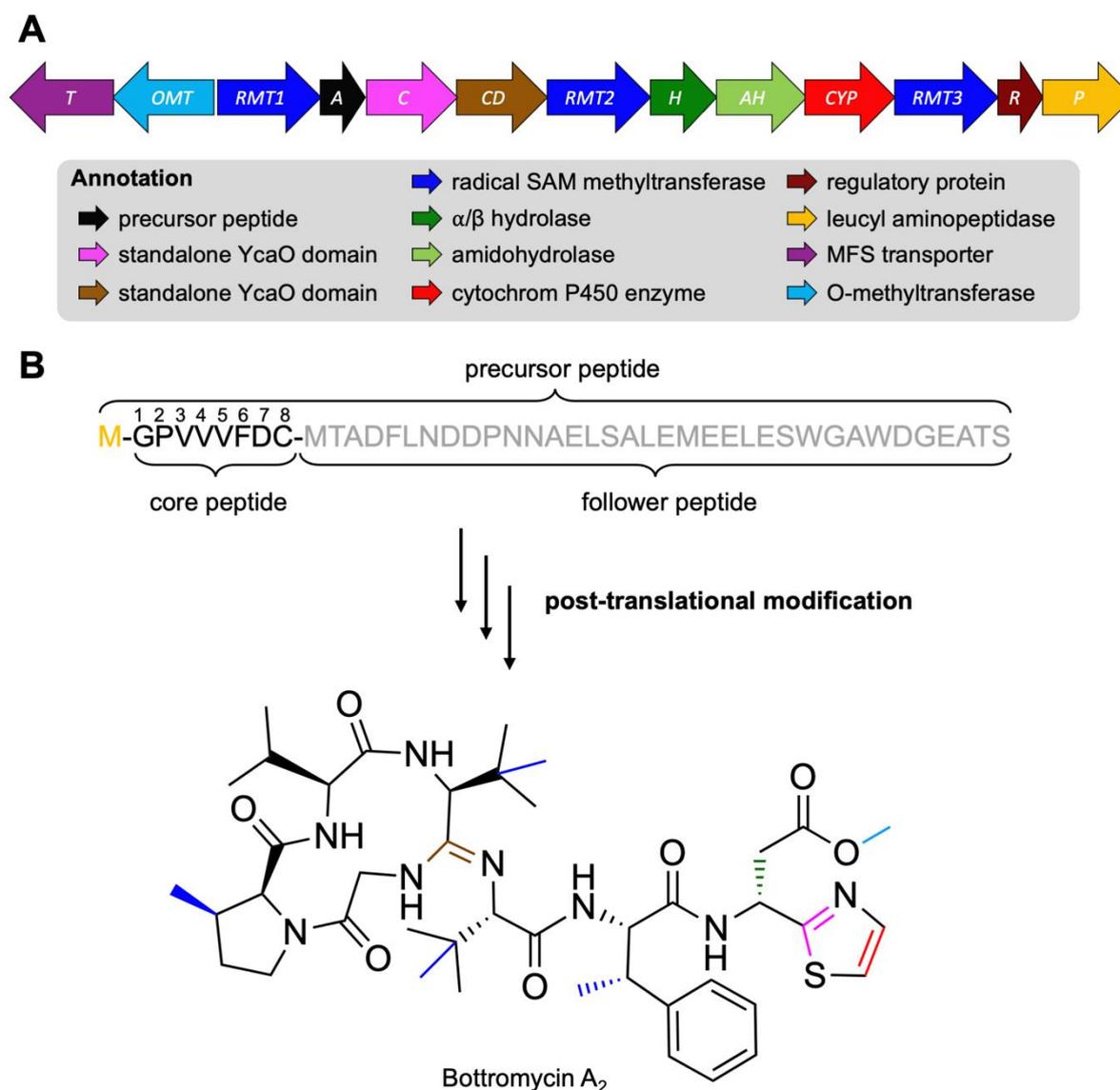
Figure 1-6 | Bottromycin gene cluster, biosynthesis, and structure. **A** *S. sp*. BC16019 bottromycin BGC and protein annotation. Genes are not drawn to scale. **B** Conversion of the bottromycin precursor peptide BotA into bottromycin A_2 by post-translational modification. Modifications introduced by the biosynthetic enzymes are colored.

Despite the efforts of several groups, the final structure of bottromycins was not elucidated until its total synthesis in 2009.[111] The total synthesis of bottromycin A_2 is challenging and requires 37 steps with unsatisfying yields. The bottromycin structure contains a four-amino acid macrocycle formed via a unique amidine linkage, several β-methylated amino acids, an O-methylated D-aspartate and a C-terminal thiazole (Figure 1-6B). Bottromycins A_2-C_2 differ by their methylation pattern (see Figure 8-2).[106, 111]

The biosynthetic origin of bottromycins was long suspected to be an NRPS system. In 2012, four independent groups identified the bottromycin biosynthetic gene cluster in different

Streptomyces species, which revealed that bottromycins are in fact RiPPs.[112-115] The genetic organization of the BGCs is basically identical in all *Streptomyces* species.

The *S. sp.* BC16019 bottromycin BGC contains 13 genes and the gene cluster organization and annotation of the predicted protein functions based on sequence homology are shown in Figure 1-6A.

The gene *botA* encodes for the precursor peptide (BotA), which is extensively modified by biosynthetic enzymes to become the biologically active bottromycin (Figure 1-6B). BotA is unique in the RiPP family because it has a C-terminal follower rather than a N-terminal leader peptide. The follower peptide and the core peptide sequence (GPVVVFDC) are highly conserved; only in *S. sp.* WMMB272 the core peptide sequence differs (GPAVVFDC), which leads to the production of the analogue bottromycin D.[115]

Based on data from gene deletion studies and an untargeted metabolomics approach, a biosynthetic scheme that included the order of steps was proposed (Figure 1-7, as of beginning of 2017):[112, 113, 116]
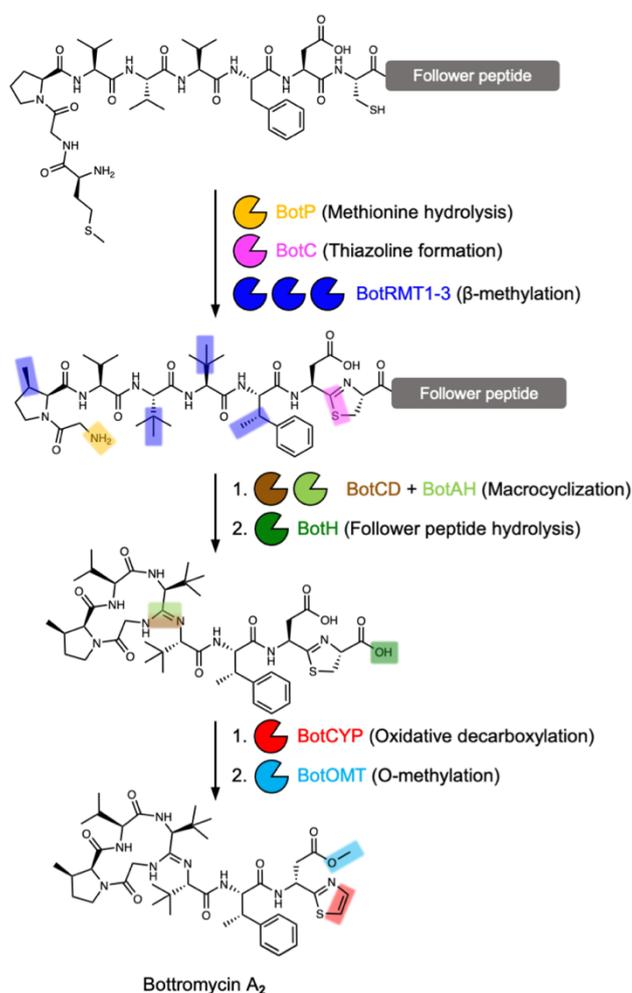
Figure 1-7 | Proposed bottromycin biosynthetic pathway[116] based on the data available in the beginning of 2017.

The aminopeptidase BotP first removes the N-terminal methionine, which is followed by heterocyclization of Cys8 by the standalone YcaO domain enzyme BotC and radical β-methylation of the amino acids Phe6, Val4, Val5 and Pro2 by the radical SAM methyltransferases BotRMT1-3.[113, 116, 117] Gene deletion of the RMTs leads to a severely reduced or abolished production of bottromycins.[113]

Macrocyclization was proposed to be catalyzed by the second YcaO domain enzyme BotCD together with the amidohydrolase BotAH, based on the observation that only linear peptides were found after deletion of one of the two genes.[116] The remaining hydrolase of the BGC, BotH, was suggested to be responsible for follower peptide removal.[116]

Further, the cytochrome P450 enzyme BotCYP and the O-methyltransferase BotOMT were proposed to catalyze the oxidative decarboxylation of the thiazoline to form a thiazole and the O-methylation of Asp7. Their function was confirmed by gene inactivation/deletion and analysis of the metabolites from the mutant strains confirmed that these reactions have to be the last steps of the biosynthesis. It was posited that the epimerization of Asp7 occurs non-enzymatically and spontaneously happens before the formation of the thiazole.[113, 116]

Not involved in the biosynthesis are the two remaining genes of the BGC that encode for a putative MFS (major facilitator superfamily) transporter BotT and the putative regulatory protein BotR. BotT is probably involved in the export of the matured bottromycin and self-resistance, as overexpression resulted in an increase of production.[113] BotR was recently shown to regulate the expression of *botA*, but it does not modulate the other genes.[118]

More detailed information about the discovery, structure elucidation, activity, mode-of-action, chemical synthesis, gene cluster and heterologous expression of bottromycins can be found in the bottromycin review manuscript (chapter 8), which the thesis author recently co-wrote.

The proposed biosynthetic pathway (Figure 1-7) was mainly based on *in vivo* data (gene deletion experiments and metabolomics). Until the beginning of 2017, only the aminopeptidase BotP was studied and its function reconstituted *in vitro*. To understand bottromycin biosynthesis in detail and to learn more about the function and the promiscuity of each enzyme towards amino acid changes in the core peptide, we set out to elucidate the bottromycin biosynthesis *in vitro*.

1.3 Natural product pathways as a source for the development of new-to-nature natural products and biocatalytic tools

As discussed above, NPs show an unprecedented structural and chemical diversity and are a rich source for new drugs (section 1.1). They are produced by dedicated enzymes of the respective biosynthetic pathways (section 1.2). In this section, the engineering of peptide NP pathways by combinatorial biosynthesis for the generation of new-to-nature NPs and the use of biosynthetic enzymes as biocatalytic tools will be discussed. While engineering approaches have been applied to all NP classes, I am going to restrict the discussion to NRPS and RiPP pathways.

Combinatorial biosynthesis is the modification of biosynthetic pathways towards the production of new-to-nature NPs. The ultimate goal is to produce novel molecules with desired biological activities and properties by the combination of enzymes from different biosynthetic pathways.[119] Besides combinatorial biosynthetic approaches, the use of single biosynthetic enzymes as biocatalytic tools in the synthesis of novel drugs is an attractive opportunity.[61, 65, 120, 121] These biocatalysts can introduce complex structural moieties that are inaccessible or extremely challenging to obtain by chemical synthesis.[61, 64, 120, 121] The use of biocatalysts for chemical synthesis has several advantages compared to chemocatalysis. Enzymatic reactions are usually performed in aqueous solution, while chemocatalysis requires organic solvents and expensive, as well as commonly toxic metal catalysts, with often uncertain supply due to geopolitical turmoil.[122] However, the limited stability of enzymes (heat, pH, salt) in aqueous solution needs to be considered. Enzymes are genetically encoded, which enables their straightforward production, predictable access and provides greater diversity from nature or by genetic engineering. Their chemo-, regio-, and stereoselectivity allows the production of stereochemically pure products, shortening of synthetic routes and a higher atom efficiency.[122] Over the past decades, biocatalysis has slowly become a relevant tool for synthesis approaches and is recognized by many companies, which have integrated this more sustainable, efficient, and less polluting method for the production of fine chemicals and pharmaceuticals. Nonetheless the current number of well-characterized biocatalysts remains insufficient and must be increased for biocatalysts to become everyday tools for the synthesis of molecules.[122-124]

The following chapters will discuss engineering of peptide NP pathways by combinatorial biosynthesis for the generation of new-to-nature NPs and the use of biosynthetic enzymes as biocatalytic tools covering NRP and RiPP biosynthetic pathways.

1.3.1 NRPS engineering

Non-ribosomal peptides are synthesized by large, modular non-ribosomal peptide synthetases (NRPS). Each module incorporates a specific building block into the polypeptide structure and the number and order of modules corresponds usually directly to the number and order of the building blocks, as the synthesis proceeds in most cases in a colinear fashion (see section 1.2.1). The high number of building blocks that can be incorporated by NRPS modules is a substantial reason for the high structural diversity of NRPs. Early studies to generate new-to-nature NRP analogues focused on precursor directed biosynthesis or mutasynthesis (Figure 1-8B) and assumed that the NRPS substrate specificity is flexible enough to allow the incorporation of different building blocks. In precursor directed biosynthesis, non-native building blocks are fed to cultures of the NRP producer (Figure 1-8B).[125] This approach requires a limited understanding of the biosynthesis and no genetic modification of the producing strain. It was for example used to successfully generate cyclosporin analogues that exhibit high immunosuppressive effects.[126] However, the yields of the new derivatives are often low because the non-native building blocks compete with the native building blocks, which likely remain the preferred substrates. This issue is addressed in mutasynthesis approaches. Enzyme(s) for the biosynthesis of a specific native building block are deleted in the NP producing organism, wherefore the non-native building block is incorporated more efficiently. This approach was for example used to generate novel calcium-dependent antibiotics (CDAs) by deletion of a gene involved in the biosynthesis of the precursor for 4-hydroxyphenylglycine, which is incorporated into the CDA structure.[125, 127]

Due to the limited substrate flexibility of the NRPS enzymes, however, building block exchanges must be very conservative and are limited. The first specificity gatekeeper in the selection of the substrates is the adenylation (A) domain. Elucidation of the first crystal structures of adenylation domains led to an understanding of critical residues involved in their substrate recognition and enabled engineering of their substrate binding pockets, as well as prediction of the substrate specificities of A domains ("non-ribosomal code").[67, 68, 78, 128-132] The mutation of single or multiple amino acids in the adenylation domain binding pocket re-programmed the domain and facilitated the incorporation of non-native building blocks into

NRPs in multiple studies (Figure 1-8C),[125, 133-135] but was also used to narrow down the substrate tolerance of promiscuous A domains to mainly yield the desired product.[125, 136] However, the change of the building blocks by altering the A domain specificity may not be accepted by downstream domains or modules.[70]
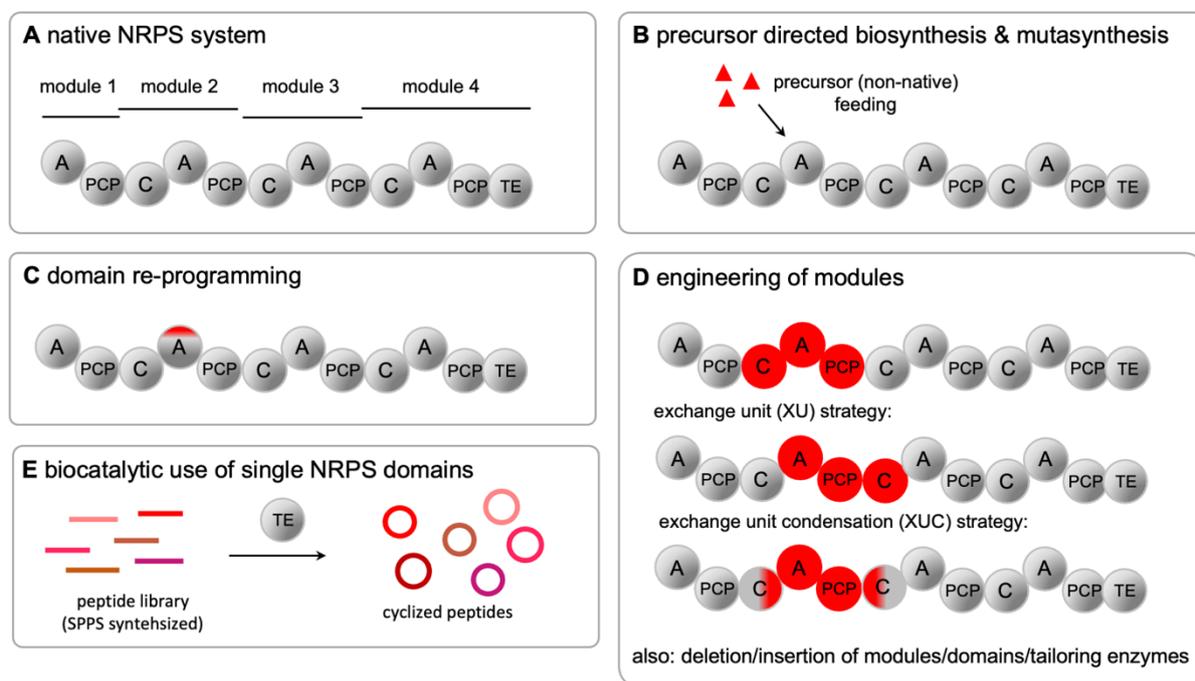
Figure 1-8 | Strategies for NRPS engineering. **A** Native NRPS system. **B** Introduction of non-native building blocks into NRPs by precursor directed biosynthesis or mutasynthesis. **C** Engineering of (adenylation, A) domain substrate specificity. **D** Exhange (or deletion/intersion) of modules. **E** Biocatalytic use of single NRPS domains, e.g. TE domains for peptide cyclization.

Further derivatization of NRPs can be achieved by the heterologous expression of additional or different tailoring enzymes derived from other NP pathways. Following this strategy, a differently chlorinated enduracidin analogue was generated by inactivating the natural chlorinase gene *end30* of the enduracidin pathways and its complementation by chlorinase *ram20* from the ramoplanin gene cluster (Figure 1-8D).[120, 125]

Instead of the exchange of tailoring enzymes, more advanced engineering approaches exchange subunits or modules from the multi-modular NRPS to produce novel peptidic structures (Figure **1-8**D). A prominent example where NRPS subunit exchange was applied is daptomycin, a cyclic 13 amino acid lipopeptide. The NRPS subunit DptD, responsible for incorporating the final two amino acids (3mGlu and kynurenine (Kyn)), peptide cyclisation and release, was deleted and then complemented with either CdaPS3 or LptD from CDA and A54145 biosynthesis, respectively. This led to the production of daptomycin analogues with Trp or Ile at position 13 instead of the native kynurenine (Kyn). These two subunits had the advantage of

having an initial Glu/3mGlu-specific A domain like DptD, which were found to be sufficiently similar to interact with the upstream PCP-domain. As DptD is the final subunit, no interaction with downstream domains could be affected.[125, 137] A similar approach was also applied for the exchange of module 8 (incorporates D-Ala) from DptBC, which was exchanged with the highly homologous module 11 (incorporates D-Ser) from DptBC, leading to the production of a daptomycin analogue with D-Ser8 and D-Ala11. Albeit successful for the production of some new-to-nature analogues, those approaches were accompanied by severely reduced or even abolished production levels.[125, 138, 139] An alternative strategy is the alteration of the number of building blocks, which can be achieved by deletion or insertion of NRPS modules. This strategy was used to generate a novel octapeptide variant of balhimycin by introducing an additional module into the balhimycin NRPS gene *bpsB*.[140, 141] Critical for the exchange of subunits or modules is to ensure a functional interaction of the native and non-native domains. This point was recently addressed by the Bode group with the development of the so-called exchange unit strategy. Following their approach, instead of C-A-PCP units, A-PCP-C or C_{Asub}-A-PCP-C_{Dsub} units are fused at a specific point within the conserved C-A linker or at the linker region between the N-terminal sub-domain of a C domain and the C-terminal sub-domain of a C domain (Figure 1-8D).[142, 143]

In the past decades, the understanding of NRPS biosynthesis increased significantly and guided researchers to genetically redesign NRPSs to generate new NRPs. Since one NRPS module incorporates only a single building block, NRP BGCs can be large (>100 kb), which make them difficult to manipulate and hampers a faster exploration of artificial NRPS constructs. Improvements of DNA assembly and editing techniques allow the faster generation of mutant or chimeric NRPS constructs that will help to further understand the NRPS architecture and improve methods to exchange NRPS modules and domains without a massive drop in production levels. Further developments as for example the recently introduced DNA-templated NRPS (DT-NRPS) approach[144] will also contribute to the success of combinatorial NRPS biosynthesis. At this point the approaches of module exchange are far from optimal and not broadly applicable to different systems. The exchange of modules to achieve functional chimeric NRPSs is often only successful when modules are phylogenetically closely related. However, the ultimate goal would be to achieve a 'plug-and-play' approach, were distantly related NRPS modules can be assembled by universal linker regions to generate non-ribosomal peptide scaffolds with new chemistry and functionalities.

In addition to combinatorial NRPS biosynthesis approaches, the biocatalytic use of single NRPS enzymes is of great interest. Especially thioesterase (TE) domains have been explored as macrocyclization catalysts to generate cyclic peptides. One example where a TE domain was successfully used to cyclize linear peptides obtained from solid phase peptide synthesis (SPPS) is the TE domain TycTE from the tyrocidine pathway. This TE domain is capable of tolerating various simultaneous side chain alterations, changes to the ring size and incorporation of nonpeptidic elements (Figure 1-8E). TycTE therefore enables the generation of cyclized peptide libraries with improved or novel biological activity. Macrocyclization of linear peptide precursors containing a RGD sequence by TycTE for example yielded potent integrin inhibitors.[72, 145-149]

1.3.2 RiPP engineering

The biosynthesis of RiPPs starts with the ribosomal synthesis of a genetically encoded precursor peptide that is subsequently modified by a series of biosynthetic enzymes. The RiPP precursor peptide usually has a bipartite nature, being composed of a N-terminal leader peptide that facilitates binding of the biosynthetic enzyme to the substrate and is discarded during the biosynthesis, and a C-terminal core peptide where modifications are introduced. Hence, the amino acid sequence of the core peptide substantially defines the structure of the mature RiPP (see section 1.2.2). Due to the physical separation of substrate recognition/binding in the leader peptide and site of modification in the core peptide, RiPP biosynthetic enzymes are often tolerant to changes of amino acids in the core peptide region.[84, 150-155] Indeed, variable core peptide sequences are also observed in nature for some RiPP pathways, e.g. in the prochlorosin or cyanobactin pathways, while the leader peptide is usually highly conserved.[155-160] Because of their modularity, the substrate promiscuity of the involved biosynthetic enzymes, gene-encoded nature, and reduced size of their BGCs in comparison to NRPSs, RiPPs are attractive targets for engineering and combinatorial biosynthesis. In the last two decades, several RiPP engineering studies have applied different *in vivo* and *in vitro* strategies to generate RiPP analogues or new-to-nature peptides. These strategies include the use of artificial precursor peptides, leader peptide complementation, and generation of artificial RiPP pathways (Figure 1-9).
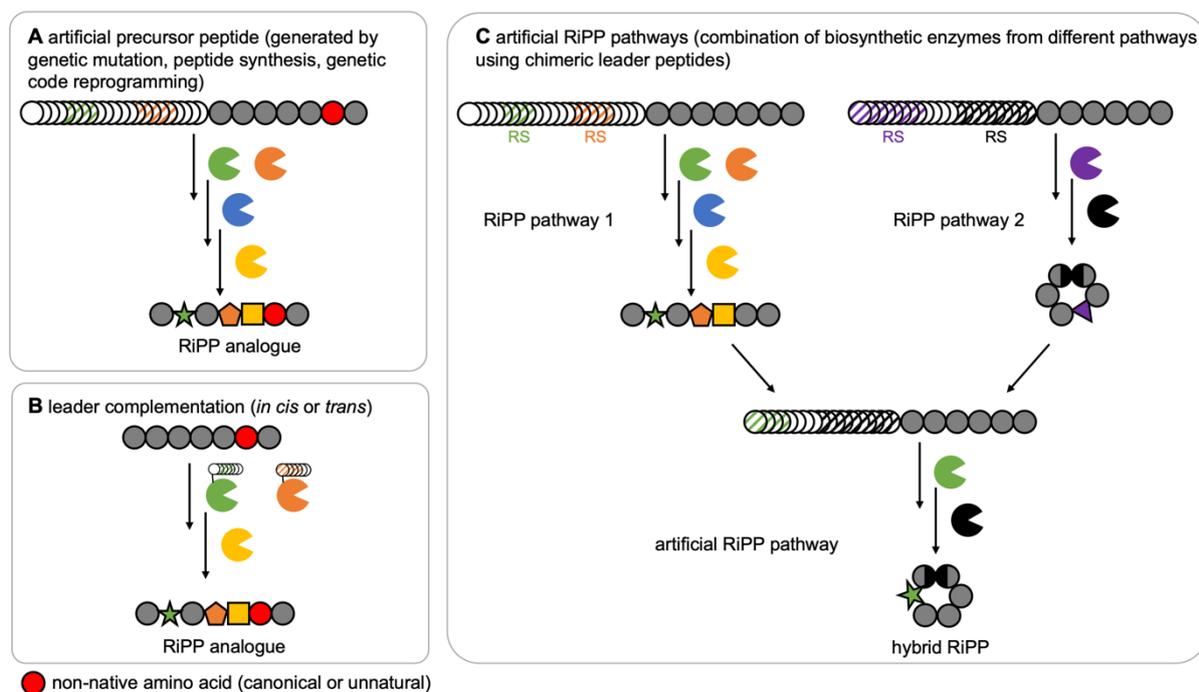
Figure 1-9 | Strategies for RiPP engineering. **A** Introduction of a non-native amino acid into the core peptide leads to the prodction of a RiPP analogue. The non-native amino acid can be a different canonial amino acid or unnatural amino acid and can be introduced into the core peptide by different methods. **B** Complementation of the leader peptide *in cis* (fused to the enzyme, shown here) or *trans*, enables enzymatic modification of peptides without the requirement of leader peptide sequences. **C** Generation of artifical precursor peptides by combination of recogntion sequences (RS) from biosynthetic RiPP enzymes from different pathways enables the new-to-nature combination of PTM and production of hybrid RiPPs.

Artificial precursor peptides

RiPP analogues can easily be generated by mutation of single or multiple codons encoding for amino acids in the core peptide (Figure 1-9A).[81, 83] Several mutagenesis studies were executed in the last two decades,[81, 83] aiming to produce RiPP analogues with improved solubility, stability and activity against their molecular targets or to focus or extend the spectrum of their biological activities. For example, mutagenesis studies of the thiopeptide GE37468 yielded several novel RiPP analogues. By codon randomization 133 precursor peptide variants were generated, in which one amino acid in the core peptide was altered. 29 yielded matured thiopeptide analogues, whereof 12 retained biological activity. One analogue (Thr2Cys) even showed an improved *in vitro* potency compared to the native GE37468 (Figure 1-10A).[151, 161] Extensive mutagenesis studies were also applied to the well-studied lantibiotic nisin and several analogues with improved properties were generated.[153, 162, 163] Engineering and screening of lasso peptides and lantibiotics, which are very tolerant to multiple amino acid substitutions in the core peptide, led to molecules with new biological activities. Incorporation of the RGD

epitope into the lasso peptide microcin J25, yielded a potent αvβ3 integrin inhibitor (Figure 1-10B).[164, 165] Screening of a genetically encoded library of 10^6 lanthipeptide precursor peptides produced a lantibiotic with a new biological activity, which inhibits a critical protein-protein-interaction for HIV budding from infected cells.[166]
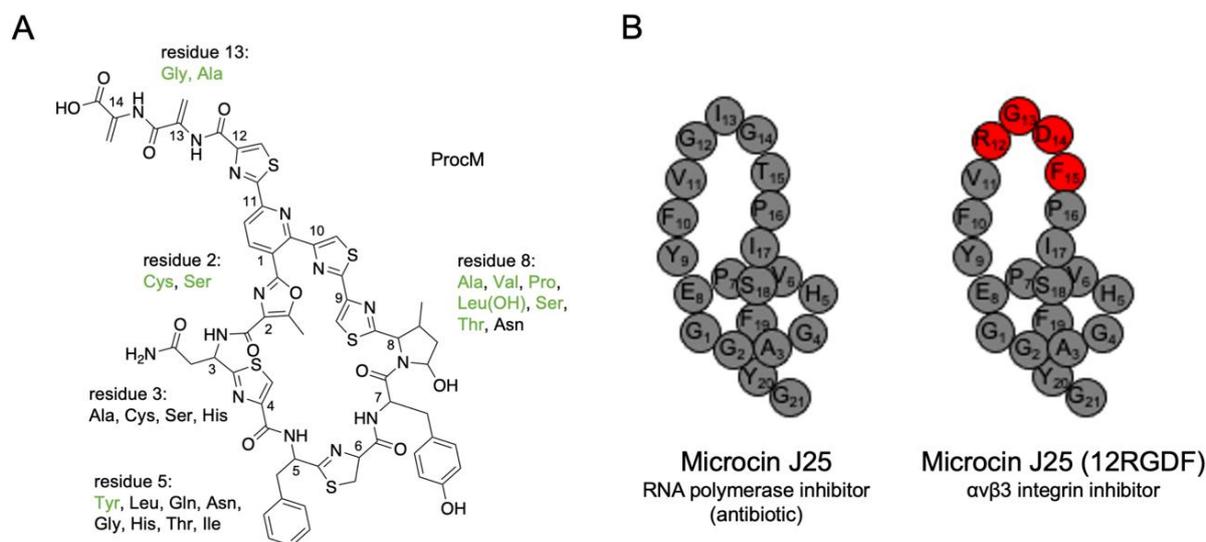
Figure 1-10 | Generation of new-to-nature RiPPs by using a non-native precursor peptide. **A** Mutasynthesis study of the thiopeptide GE37468 generated novel GE37468 analogues, whereof several remained biological active (green). **B** Introduction of the RGD motif in the antibiotic lasso peptide microcin J25, yields a potent integrin inhibitor.

Due to the ribosomal origin of RiPPs, incorporated amino acid building blocks are limited to the 20 canonical amino acids. However, methods have been applied to introduce unnatural amino acids to increase the chemical diversity of RiPPs for improved characteristics and the introduction of bioorthogonal groups that allow post-biosynthetic modification.[83, 158, 167] Unnatural amino acids can be introduced into the precursor peptide by amino acid supplementation or amber stop codon suppression (both *in vivo*), using synthetic peptides (*in vitro* mutasynthesis), and genetic code reprogramming and *in vitro* translation (flexible *in vitro* translation (FIT) system[168]).[83, 158]

In vivo approaches (amino acid supplementation or amber stop codon suppression) have been applied to generate lanthipeptide and lasso peptide analogues with unnatural amino acids.[167, 169] They have the advantage of utilizing the cellular machineries (e.g. the translation system, cofactor supply), reactions occur in the native environment and cultures can be scaled up for the production of larger quantities. In addition, the generation of *in vivo* RiPP libraries can be coupled with cell-based assays, such as reverse two-hybrid systems, surface display technologies or the nanoFleming inhibition assay platform for high-throughput activity

screening.[166, 170-174]

In vitro mutasynthesis approaches require soluble and stable biosynthetic enzymes, the expression of which can be difficult or tedious. However, they do not only allow a better understanding of RiPP pathways, but also the use of synthetic precursor peptides or even substrates with non-peptide moieties that cannot be produced *in vivo*.[83, 158] *In vitro* produced RiPPs moreover can be used without purification for bioactivity high-throughput screening.[175] One example for successful *in vitro* mutasynthesis is the engineering of the lantibiotic lacticin 481 leading to the generation of analogues incorporating α-amino acids with unnatural side chains, β-amino acids or peptoids. Some of these analogues showed an improved activity compared to the native lacticin 481.[158, 176, 177]

The flexible *in vitro* translation (FIT) system introduced by Goto *et al.* is a custom-made cell-free translation system, that enables the *in vitro* synthesis of precursor peptides with unnatural amino acids from synthetic DNA templates by genetic code reprogramming and the utilization of versatile tRNA-acylating ribozymes, called flexizymes.[158, 168] Combination of the FIT-system with an *in vitro* constituted RiPP pathway in one pot allows the fast study of the substrate promiscuity of a RiPP pathway from DNA-templated precursor peptide variants.[158, 178-182] Recently, this system was impressively used to dissect the biosynthesis of the thiopeptide lactazole A and demonstrated the promiscuity of the lactazole biosynthetic pathway by producing thiopeptides with 14- to 62-membered rings, different tail lengths that are very tolerant to amino acids changes and the incorporation of non-natural amino acids ().[181] The FIT system approach enables rapid, robust and flexible RiPP engineering, but in contrast to other methods only small quantities are produced, which may not be sufficient for biological activity screening.[183]
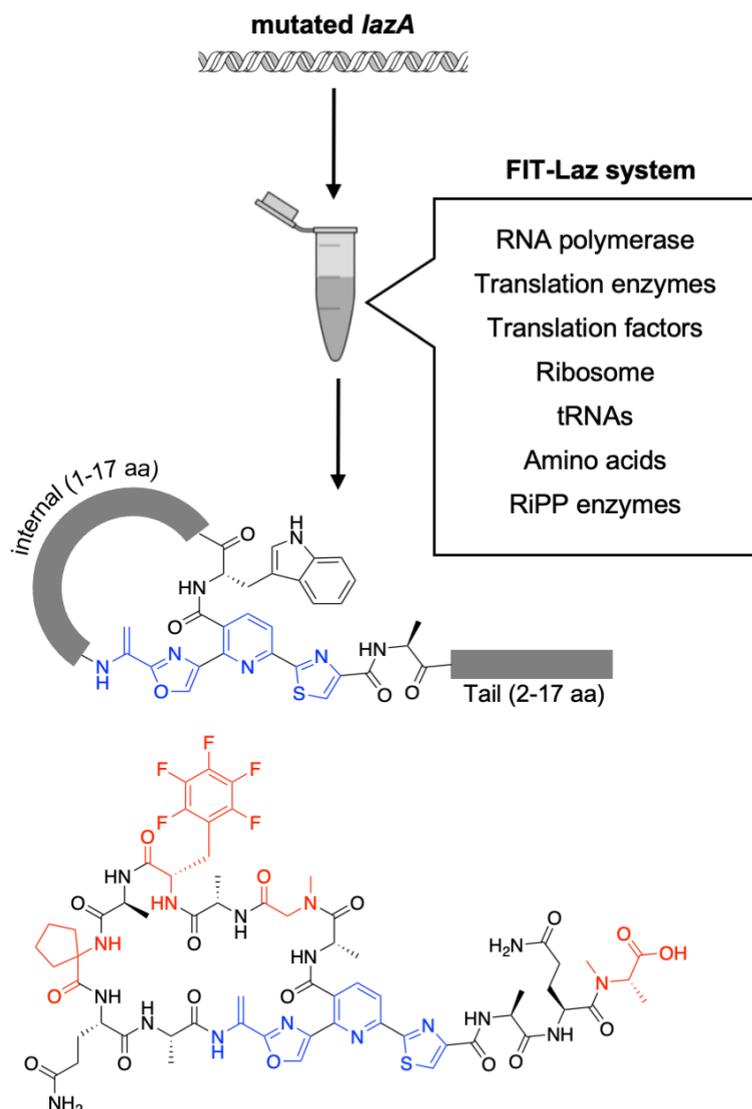
Figure 1-11 | Combination of the flexible *in vitro* translation (FIT) system with lactazole biosynthetic enzymes (FIT-Laz system) for thiopeptide engineering. Synthetic DNA templates encode for the lactazole precursor peptide (LazA) or its variants and are *in vitro* transcribed and translated to generate precursor peptides (with unnatural amino acids), which are subsequently modified by the lactazole biosynthetic enzymes. The study led to the identification of a minimal lactazole scaffold and generation of lactazole analogs with various ring and tail sizes and containing unnatural amino acids. One example of such an analogue is shown at the bottom. blue: peptide modifications by lactazole biosynthetic enzymes, red: unnatural amino acids.

Leader complementation

Peptide libraries for *in vitro* mutagenesis studies are synthesized by SPPS (solid-phase peptide synthesis) that occurs in C- to N-direction. In the case of RiPPs, this is suboptimal for the generation of precursor peptide libraries, as the N-terminal part of the peptide (the leader) is consistent, and the C-terminal part, which is synthesized first in SPPS, is variable. One *in vitro* mutasynthesis studies of lacticin 481 described above tried to minimize the peptide synthesis

efforts by using Cu(I)-catalyzed 1,3-dipolar cycloaddition of an alkyne-functionalized leader peptide and an array of azide-modified core peptides to synthesize the precursor peptide.[176, 177] The other lacticin 481 *in vitro* mutagenesis study produced lacticin analogues by incubating synthetic core peptides with a constitutively active leader-lacticin synthetase (LctM) fusion enzyme.[184, 185] As discussed in section 1.2.2, the leader peptide is crucial for primary biosynthetic enzymes. However, studies have shown that the leader peptide can be separated from the core peptide for enzyme activation. The leader peptide can therefore be complemented *in cis* (fused to the enzyme) or *in trans* (addition of the LP *in vitro* or co-expression *in vivo*) (Figure 1-9B).[83] These approaches have been applied to several enzymes, e.g. the above mentioned LctM, the cyanobactin heterocyclase LynD, and the microviridin cyclases MvdC/MvdD.[83, 185-187] Reconstitution of the full microviridin biosynthesis and fusion of the leader peptide to the natively leader-dependent biosynthetic enzymes allowed the generation of bicyclic peptide libraries from synthetic core peptides and yielded novel biologically active microviridins.[187] This approaches reduces the length of the peptides that have to be synthesized and facilitates the screening of peptide libraries.[83] In addition, this method may allow the generation of new-to-nature RiPPs as theoretically, leader-independent, meaning not requiring a leader peptide covalently attached to the core peptide, primary enzymes can be generated and enzymes from different pathways could be combined to modify a core peptide. However, leader peptide complementation is not applicable for all biosynthetic enzymes.

Artificial RiPP pathways

The above-mentioned methods to generate novel RiPPs aim to alter the native RiPP structure by modifying the amino acid sequence of the core peptide. However, the ultimate goal is to produce peptides with desired features by combining RiPP biosynthetic enzymes in a 'plug-and-play' system. The modularity of RiPP biosynthesis enables the synthesis of hybrid RiPPs by combining enzymes from different RiPP pathways. This was first shown for the generation of novel lantibiotics and cyanobactins, where enzymes of related families were mixed.[62, 159] To combine RiPP enzymes from non-related RiPP pathways, the chimeric leader peptide approach was introduced.[155] In this approach, the recognition sequences (RS) of biosynthetic enzymes from different biosynthetic pathways are fused to yield a chimeric leader peptide (Figure 1-9C), enabling the new-to-nature combination of RiPP modifications, e.g. thiazoline-lanthipeptide and thiazoline-sactipeptide hybrids (Figure 1-12).[155] This approach can be used *in vivo* and *in vitro*, but requires extensive knowledge of RSs of the biosynthetic enzymes and consideration of the distance between RS and core peptide. Furthermore, the number of RSs that can be combined is limited.
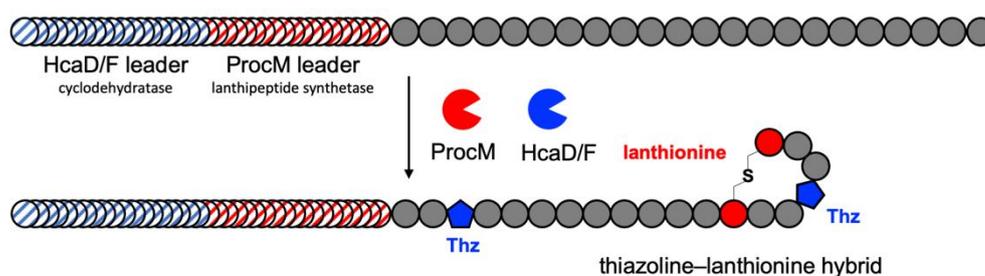
Figure 1-12 | Combination of the HcaD/F and the ProcM leader peptide to an artifical precursor peptide enables to modification of a core peptide by enzymes from unrelated RiPP pathways and production of a new-to-nature thiazoline-lanthionine hybrid.

Methods for RiPP engineering can also be combined, as demonstrated in the study enabling the benchtop biosynthesis of thiopeptide scaffolds.[183] Using cell free expression and flexizyme-based codon reprogramming, precursor peptides with chimeric leader peptides, containing LynD (cyclodehydratase from the aesturamide pathway) and TclM (pyridine synthase from thiocillin biosynthesis) recognition sequences (RSs), and a core peptide with the unnatural amino acid *Se*-phenylselenocysteine (SecPh) were produced. LynD and the leader-independent azoline-oxidase TbtE (from thiomuracin biosynthesis) introduced thiazole moieties into the peptides. Instead of a dehydratase to introduce dehydroalanines (Dha), the SecPh residues can undergo oxidative elimination with H_2O_2 to produce Dhas synthetically for the pyridine-forming cycloaddition by TclM.[183]

Biocatalysis with RiPPs enzymes

Besides the malleability of RiPP pathways, the involved biosynthetic enzymes are also attractive for biocatalytic or chemoenzymatic approaches. Butelase-1, responsible for cyclotide backbone cyclization, is a promising peptide and protein ligase.[188-191] The macrocyclases PatG$_{mac}$, from the patellamide pathway, and PCY1, involved in plant orbitide biosynthesis, have been shown to macrocyclize various substrates of different lengths and substrates that contained non-peptidic structures (Figure 1-13).[192-194] One further example of interesting RiPP enzymes as biocatalytic tools is the N-methyltransferase OphA from omphalotin biosynthesis, that has been shown to allow N-methylation of the amide backbone of the peptide of choice (here cyclosporin A and dictyonamide A) and may be a useful tool for peptide N-methylation in general.[195]
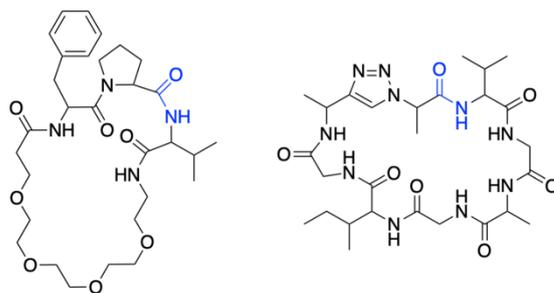
Figure 1-13 | The promicsous macrocyclase PatG$_{mac}$ can macrocyclize (blue) diverse substrates containing unnatural or non-peptide building blocks.

In the recent years, several approaches were introduced for RiPP engineering and many studies demonstrated the successful production of novel RiPP analogues with improved properties and biological activities. But most studies were applied to well-studied RiPP classes (e.g. lasso peptides, lanthipeptides) whose enzymes are very promiscuous and tolerate changes in the core peptide. Not all approaches are transferable for all RiPP pathways. For 'plug-and-play' like systems, the not well understood importance and prevalence of protein-protein interactions of RiPP biosynthetic enzymes[196] may be a setback. RiPP engineering is, despite first successful examples, still in its infancy. At the moment the commercial use of RiPPs is underdeveloped and only a few RiPP (derivatives) are in clinical trials. The future will show if RiPP engineering approaches (e.g. from companies such as Lassogen or GyreOx) will produce novel peptide drugs for commercial use.

1.4 Outline of this thesis

The main focus of this work was the elucidation of the biochemical steps in bottromycin biosynthesis (Figure 1-14). The RiPP antibiotic bottromycin has a unique scaffold that contains a four amino acid amidine macrocycle and a C-terminal (R)-thia-β-Ala-OMe. Over the course of this thesis, chapter 2 to 5, the functional and structural characterization of the bottromycin biosynthetic enzymes that install these interesting post-translational modifications into the peptide is presented. The results led to a partially revised biosynthetic pathway and *in vitro* reconstitution of the enzymes required to produce the core scaffold, which enables the production of bottromycin derivatives.
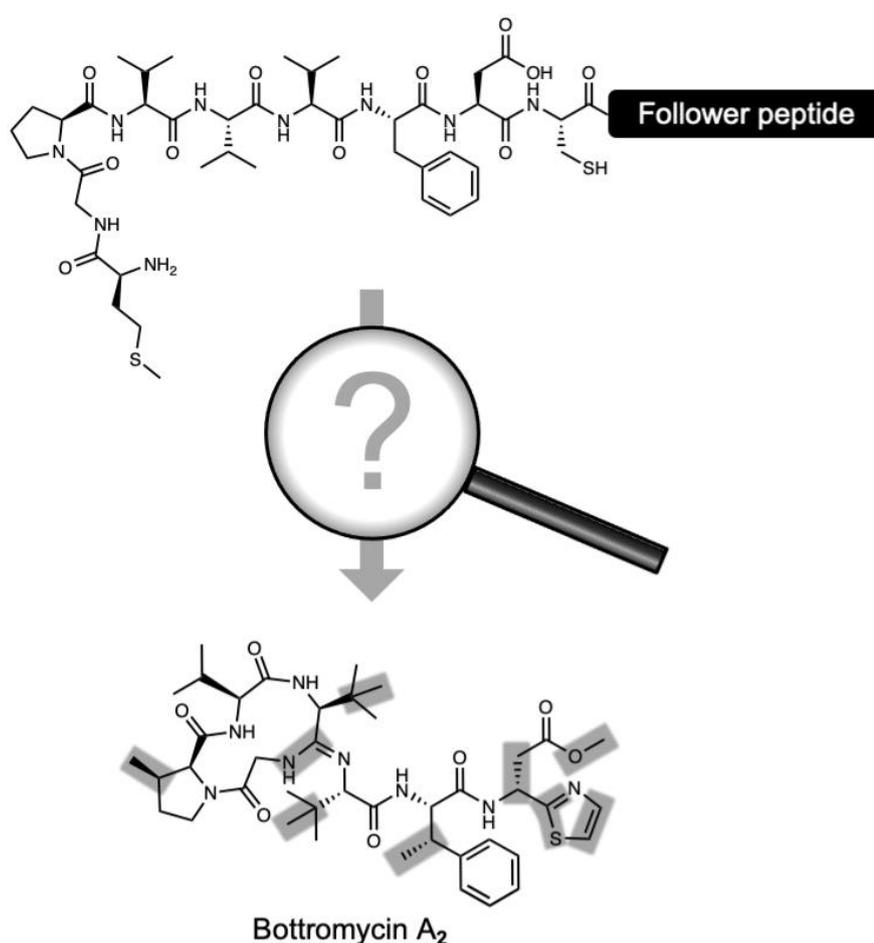
Figure 1-14 | The main focus of this work is to decipher the bottromycin biosynthesis.

In chapter 2, the biochemical study of the two YcaO domain enzymes encoded in the bottromycin BGC is reported. It demonstrates that they are sufficient to install the thiazoline and the macroamidine into the precursor peptide. While thiazoline formation is an established transformation for YcaO domain enzymes, the study presents the first characterization of an amidine-forming YcaO and we propose a mechanism analogous to YcaO-catalyzed azoline

formation. Surprisingly, the macroamidine formation was reversible. However, the characterization of the bottromycin amidohydrolase in chapter 3 demonstrates that this enzyme can act as a YcaO accessory protein by specifically removing the follower peptide of macroamidine-containing intermediates. With these results, the function of the putative hydrolase in the bottromycin BGC, a predicted α/β-hydrolase, became enigmatic, because it had been proposed to catalyze follower peptide removal. The biochemical investigation of this enzyme in chapter 4established it as an aspartate epimerase and the crystal structure in complex with the substrate led to an understanding of the enzymatic mechanism. Finally, the *in vitro* investigation of the bottromycin cytochrome P450 enzyme of the bottromycin BGC demonstrated that this enzyme catalyzes the oxidative decarboxylation of the thiazoline to yield the bottromycin core scaffold. The study of this enzyme in chapter 5 revealed the stereospecificity for the D-configured intermediate to provide stereochemical resolution for the pathway.

In addition to the elucidation of the bottromycin biosynthesis, a novel approach to combine RiPP enzymes from unrelated pathways is presented in this work (chapter 6). In the proof-of-concept study, Sortase A-mediated leader peptide exchange enabled the modification of a peptide by a heterocyclase from a cyanobactin pathway and an ATP-grasp ligase from a graspetide pathway. This strategy will allow the exploration of the cross-pathway combinatorial potential of RiPPs and the production of neo-natural product libraries.

1.5 References

1.   Hanson, J. R., Natural products the secondary metabolites. **2003**.

2.   Williams, D. H.; Stone, M. J.; Hauck, P. R.; Rahman, S. K., Why Are Secondary Metabolites (Natural Products) Biosynthesized? *Journal of Natural Products* **1989,** *52* (6), 1189-1208.

3.   Bernardini, S.; Tiezzi, A.; Laghezza Masci, V.; Ovidi, E., Natural products for human health: an historical overview of the drug discovery approaches. *Natural Product Research* **2018,** *32* (16), 1926-1950.

4.   Baker, D. D.; Chu, M.; Oza, U.; Rajgarhia, V., The value of natural products to future pharmaceutical discovery. *Nat Prod Rep* **2007,** *24* (6), 1225-44.

5.   Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. *Metabolites* **2012,** *2* (2), 303-36.

6.   Newman, D. J.; Cragg, G. M.; Snader, K. M., Natural Products as Sources of New Drugs over the Period 1981−2002. *Journal of Natural Products* **2003,** *66* (7), 1022-1037.

7.   Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *Journal of Natural Products* **2020,** *83* (3), 770-803.

8.   Cantrell, C. L.; Dayan, F. E.; Duke, S. O., Natural Products As Sources for New Pesticides. *Journal of Natural Products* **2012,** *75* (6), 1231-1242.

9.   Katz, L.; Baltz, R. H., Natural product discovery: past, present, and future. *Journal of Industrial Microbiology & Biotechnology* **2016,** *43* (2), 155-176.

10.   Clayton, D. H.; Wolfe, N. D., The adaptive significance of self-medication. *Trends Ecol Evol* **1993,** *8* (2), 60-3.

11.   Ebers, G. M.; Stern, L. C., *Papyrus Ebers. Das hermetische Buch über die Arzneimittel der alten Ägypter in hieratischer Schrift*. Engelmann: 1875.

12.   Chang, H. M.; But, P. P.-H.; Wang, L.-L.; Yao, S.-C.; Yeung, S. C.-S., *Pharmacology and Applications of Chinese Materia Medica*.

13.   Kapoor, L. D., *Handbook of Ayurvedic medicinal plants: Herbal reference library*. 2017; p 1-416.

14.   Asfaw, N.; Demissew, S., Handbook of African medicinal plants. *Economic Botany* **1994,** *48* (1), 64-64.

15.   Eshbaugh, W. H., Medicinal Plants of the West Indies. *Economic Botany* **1984,** *38* (1), 133-133.

16.   Petrovska, B. B., Historical review of medicinal plants' usage. *Pharmacogn Rev* **2012,** *6* (11), 1-5.

17.   Cragg, G. M.; Newman, D. J., Biodiversity: A continuing source of novel drug leads. *Pure and Applied Chemistry* **2005,** *77* (1), 7.

18.   Krishnamurti, C.; Rao, S., The isolation of morphine by Serturner. *Indian Journal of Anaesthesia* **2016,** *60*, 861.

19.   Kaiser, H., Von der Pflanze zur Chemie – Die Frühgeschichte der „Rheumamittel". *Zeitschrift für Rheumatologie* **2008,** *67* (3), 252-262.

20.   Jeffreys, D., Aspirin: The Remarkable Story of a Wonder Drug. **2005**.

21.   Newman, D. J.; Cragg, G. M.; Snader, K. M., The influence of natural products upon drug discovery. *Natural Product Reports* **2000,** *17* (3), 215-234.

22.   Fleming, A., On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. *Br J Exp Pathol* **1929,** *10* (3), 226-236.

23.   Hutchings, M. I.; Truman, A. W.; Wilkinson, B., Antibiotics: past, present and future. *Current Opinion in Microbiology* **2019,** *51*, 72-80.

24.   Cragg, G. M.; Newman, D. J., Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta* **2013,** *1830* (6), 3670-95.

25.   Koehn, F. E.; Carter, G. T., The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* **2005,** *4* (3), 206-20.

26.   Li, J. W.-H.; Vederas, J. C., Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009,** *325* (5937), 161-165.

27.   Wessjohann, L. A.; Ruijter, E.; Garcia-Rivera, D.; Brandt, W., What can a chemist learn from nature's macrocycles?--a brief, conceptual view. *Mol Divers* **2005,** *9* (1-3), 171-86.

28.   Schneider, P.; Schneider, G., Privileged Structures Revisited. *Angewandte Chemie International Edition* **2017,** *56* (27), 7971-7974.

29.   Reymond, J.-L.; Awale, M., Exploring chemical space for drug discovery using the chemical universe database. *ACS Chem Neurosci* **2012,** *3* (9), 649-657.

30.   Feher, M.; Schmidt, J. M., Property Distributions: Differences between Drugs, Natural Products, and Molecules from Combinatorial Chemistry. *Journal of Chemical Information and Computer Sciences* **2003,** *43* (1), 218-227.

31.   Medina-Franco, J. L., New Approaches for the Discovery of Pharmacologically-Active Natural Compounds. *Biomolecules* **2019,** *9* (3), 115.

32.   Bajorath, J., Chemoinformatics methods for systematic comparison of molecules from natural and synthetic sources and design of hybrid libraries. *Mol Divers* **2002,** *5* (4), 305-13.

33. Henkel, T.; Brunne, R. M.; Müller, H.; Reichel, F., Statistical Investigation into the Structural Complementarity of Natural Products and Synthetic Compounds. *Angewandte Chemie International Edition* **1999,** *38* (5), 643-647.

34. Lee, M.-L.; Schneider, G., Scaffold Architecture and Pharmacophoric Properties of Natural Products and Trade Drugs: Application in the Design of Natural Product-Based Combinatorial Libraries. *Journal of Combinatorial Chemistry* **2001,** *3* (3), 284-289.

35. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G., Counting on natural products for drug design. *Nature Chemistry* **2016,** *8* (6), 531-541.

36. Ntie-Kang, F.; Nyongbela, K. D.; Ayimele, G. A.; Shekfeh, S., "Drug-likeness" properties of natural compounds. *Physical Sciences Reviews* **2019,** *4* (11), 20180169.

37. O'Neill, J., Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. London: Review on Antimicrobial Resistance. **2014**.

38. Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J., Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clinical Infectious Diseases* **2009,** *48* (1), 1-12.

39. Spellberg, B.; Guidos, R.; Gilbert, D.; Bradley, J.; Boucher, H. W.; Scheld, W. M.; Bartlett, J. G.; Edwards, J., Jr; Tagg, J. R., The Epidemic of Antibiotic-Resistant Infections: A Call to Action for the Medical Community from the Infectious Diseases Society of America. *Clinical Infectious Diseases* **2008,** *46* (2), 155-164.

40. Klevens, R. M.; Edwards, J. R.; Tenover, F. C.; McDonald, L. C.; Horan, T.; Gaynes, R.; System, N. N. I. S., Changes in the Epidemiology of Methicillin-Resistant Staphylococcus aureus in Intensive Care Units in US Hospitals, 1992–2003. *Clinical Infectious Diseases* **2006,** *42* (3), 389-391.

41. Cragg, G. M.; Grothaus, P. G.; Newman, D. J., New Horizons for Old Drugs and Drug Leads. *Journal of Natural Products* **2014,** *77* (3), 703-723.

42. Wright, G. D., Opportunities for natural products in 21st century antibiotic discovery. *Natural Product Reports* **2017,** *34* (7), 694-701.

43. Rossiter, S. E.; Fletcher, M. H.; Wuest, W. M., Natural Products as Platforms To Overcome Antibiotic Resistance. *Chemical Reviews* **2017,** *117* (19), 12415-12474.

44. Fosgerau, K.; Hoffmann, T., Peptide therapeutics: current status and future directions. *Drug Discov Today* **2015,** *20* (1), 122-8.

45. Vogt, E.; Künzler, M., Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics. *Appl Microbiol Biotechnol* **2019,** *103* (14), 5567-5581.

46. Hummel, G.; Reineke, U.; Reimer, U., Translating peptides into small molecules. *Molecular BioSystems* **2006,** *2* (10), 499-508.

47. Loffet, A., Peptides as Drugs: Is There a Market? *Journal of Peptide Science* **2002,** *8* (1), 1-7.

48. Lau, J. L.; Dunn, M. K., Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry* **2018,** *26* (10), 2700-2707.

49. Beutler, J. A., Natural Products as a Foundation for Drug Discovery. *Current Protocols in Pharmacology* **2009,** *46* (1), 9.11.1-9.11.21.

50. Joo, S. H., Cyclic peptides as therapeutic agents and biochemical tools. *Biomol Ther (Seoul)* **2012,** *20* (1), 19-26.

51. Roxin, Á.; Zheng, G., Flexible or fixed: a comparative review of linear and cyclic cancer-targeting peptides. *Future Med Chem* **2012,** *4* (12), 1601-18.

52. Hewitt, W. M.; Leung, S. S. F.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S., Cell-Permeable Cyclic Peptides from Synthetic Libraries Inspired by Natural Products. *Journal of the American Chemical Society* **2015,** *137* (2), 715-721.

53. Kwon, Y.-U.; Kodadek, T., Quantitative Evaluation of the Relative Cell Permeability of Peptoids and Peptides. *Journal of the American Chemical Society* **2007,** *129* (6), 1508-1509.

54. Edman, P., Chemistry of amino acids and peptides. *Annu Rev Biochem* **1959,** *28*, 69-96.

55. Horton, D. A.; Bourne, G. T.; Smythe, M. L., Exploring privileged structures: the combinatorial synthesis of cyclic peptides. *J Comput Aided Mol Des* **2002,** *16* (5-6), 415-30.

56. Namjoshi, S.; Benson, H. A. E., Cyclic peptides as potential therapeutic agents for skin disorders. *Peptide Science* **2010,** *94* (5), 673-680.

57. Cardote, T. A. F.; Ciulli, A., Cyclic and Macrocyclic Peptides as Chemical Tools To Recognise Protein Surfaces and Probe Protein–Protein Interactions. *ChemMedChem* **2016,** *11* (8), 787-794.

58. Cary, D.; Ohuchi, M.; Reid, P.; Masuya, K., Constrained Peptides in Drug Discovery and Development. *Journal of Synthetic Organic Chemistry, Japan* **2017,** *75*, 1171-1178.

59. Vinogradov, A. A.; Yin, Y.; Suga, H., Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *Journal of the American Chemical Society* **2019,** *141* (10), 4167-4181.

60. Alexandru-Crivac, C. N.; Umeobika, C.; Leikoski, N.; Jokela, J.; Rickaby, K. A.; Grilo, A. M.; Sjö, P.; Plowright, A. T.; Idress, M.; Siebs, E.; Nneoyi-Egbe, A.; Wahlsten, M.; Sivonen, K.; Jaspars, M.; Trembleau, L.; Fewer, D. P.; Houssen, W. E., Cyclic peptide production using a macrocyclase with enhanced substrate promiscuity and relaxed recognition determinants. *Chemical Communications* **2017,** *53* (77), 10656-10659.

61. Scott, T. A.; Piel, J., The hidden enzymology of bacterial natural product biosynthesis. *Nat Rev Chem* **2019,** *3* (7), 404-425.

62. Sardar, D.; Lin, Z.; Schmidt, E. W., Modularity of RiPP Enzymes Enables Designed Synthesis of Decorated Peptides. *Chem Biol* **2015,** *22* (7), 907-16.

63.   Yu, X.; Sun, D., Macrocyclic drugs and synthetic methodologies toward macrocycles. *Molecules* **2013,** *18* (6), 6230-68.

64.   Tibrewal, N.; Tang, Y., Biocatalysts for natural product biosynthesis. *Annu Rev Chem Biomol Eng* **2014,** *5*, 347-66.

65.   Hubert, C. B.; Barry, S. M., New chemistry from natural product biosynthesis. *Biochem Soc Trans* **2016,** *44* (3), 738-44.

66.   Sieber, S. A.; Marahiel, M. A., Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics. *Chemical Reviews* **2005,** *105* (2), 715-738.

67.   Stachelhaus, T.; Marahiel, M. A., Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiology Letters* **1995,** *125* (1), 3-14.

68.   Mootz, H. D.; Schwarzer, D.; Marahiel, M. A., Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *Chembiochem* **2002,** *3* (6), 490-504.

69.   Drake, E. J.; Miller, B. R.; Shi, C.; Tarrasch, J. T.; Sundlov, J. A.; Leigh Allen, C.; Skiniotis, G.; Aldrich, C. C.; Gulick, A. M., Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **2016,** *529* (7585), 235-238.

70.   Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T., Aminoacyl-CoAs as Probes of Condensation Domain Selectivity in Nonribosomal Peptide Synthesis. *Science* **1999,** *284* (5413), 486-489.

71.   Schneider, A.; Marahiel, M. A., Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in Bacillus subtilis. *Archives of Microbiology* **1998,** *169* (5), 404-410.

72.   Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T., Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **2000,** *407* (6801), 215-8.

73.   Walsh, C. T.; O'Brien, R. V.; Khosla, C., Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angewandte Chemie International Edition* **2013,** *52* (28), 7098-7124.

74.   Imani, A. S.; Freeman, M. F., RiPPing apart the rules for peptide natural products. *Synth Syst Biotechnol* **2018,** *3* (2), 81-82.

75.   Caboche, S.; Pupin, M.; Leclère, V.; Fontaine, A.; Jacques, P.; Kucherov, G., NORINE: a database of nonribosomal peptides. *Nucleic Acids Research* **2008,** *36* (suppl_1), D326-D331.

76.   Flissi, A.; Ricart, E.; Campart, C.; Chevalier, M.; Dufresne, Y.; Michalik, J.; Jacques, P.; Flahaut, C.; Lisacek, F.; Leclère, V.; Pupin, M., Norine: update of the nonribosomal peptide resource. *Nucleic Acids Research* **2020,** *48* (D1), D465-D469.

77. Weckwerth, W.; Miyamoto, K.; Iinuma, K.; Krause, M.; Glinski, M.; Storm, T.; Bonse, G.; Kleinkauf, H.; Zocher, R., Biosynthesis of PF1022A and related cyclooctadepsipeptides. *J Biol Chem* **2000,** *275* (23), 17909-15.

78. Stachelhaus, T.; Walsh, C. T., Mutational Analysis of the Epimerization Domain in the Initiation Module PheATE of Gramicidin S Synthetase. *Biochemistry* **2000,** *39* (19), 5775-5787.

79. Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; Marshall, C. G.; Miller, D. A.; Patel, H. M., Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Current Opinion in Chemical Biology* **2001,** *5* (5), 525-534.

80. Süssmuth, R. D.; Mainz, A., Nonribosomal Peptide Synthesis-Principles and Prospects. *Angew Chem Int Ed Engl* **2017,** *56* (14), 3770-3821.

81. 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.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* **2013,** *30* (1), 108-60.

82. Ortega, M. A.; van der Donk, W. A., New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem Biol* **2016,** *23* (1), 31-44.

83. Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A., New developments in RiPP discovery, enzymology and engineering. *Natural Product Reports* **2020**.

84. Yang, X.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: new insights into the role of leader and core peptides during biosynthesis. *Chemistry* **2013,** *19* (24), 7662-77.

85. Plat, A.; Kuipers, A.; Rink, R.; Moll, G. N., Mechanistic aspects of lanthipeptide leaders. *Curr Protein Pept Sci* **2013,** *14* (2), 85-96.

86.   Oman, T. J.; van der Donk, W. A., Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat Chem Biol* **2010,** *6* (1), 9-18.

87.   Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A., A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat Chem Biol* **2015,** *11* (8), 564-70.

88.   Kloosterman, A. M.; Shelton, K. E.; van Wezel, G. P.; Medema, M. H.; Mitchell, D. A., RRE-Finder: a Genome-Mining Tool for Class-Independent RiPP Discovery. *mSystems* **2020,** *5* (5), e00267-20.

89.   Truman, A. W., Cyclisation mechanisms in the biosynthesis of ribosomally synthesised and post-translationally modified peptides. *Beilstein J Org Chem* **2016,** *12*, 1250-1268.

90.   Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A., Chemoenzymatic synthesis of thiazolyl peptide natural products featuring an enzyme-catalyzed formal [4 + 2] cycloaddition. *Journal of the American Chemical Society* **2015,** *137* (10), 3494-3497.

91.   Koehnke, J.; Bent, A.; Houssen, W. E.; Zollman, D.; Morawitz, F.; Shirran, S.; Vendome, J.; Nneoyiegbe, A. F.; Trembleau, L.; Botting, C. H.; Smith, M. C.; Jaspars, M.; Naismith, J. H., The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nat Struct Mol Biol* **2012,** *19* (8), 767-72.

92.   Lee, J.; McIntosh, J.; Hathaway, B. J.; Schmidt, E. W., Using marine natural products to discover a protease that catalyzes peptide macrocyclization of diverse substrates. *J Am Chem Soc* **2009,** *131* (6), 2122-4.

93.   Luo, H.; Hong, S. Y.; Sgambelluri, R. M.; Angelos, E.; Li, X.; Walton, J. D., Peptide macrocyclization catalyzed by a prolyl oligopeptidase involved in α-amanitin biosynthesis. *Chem Biol* **2014,** *21* (12), 1610-7.

94.   Merwin, N. J.; Mousa, W. K.; Dejong, C. A.; Skinnider, M. A.; Cannon, M. J.; Li, H.; Dial, K.; Gunabalasingam, M.; Johnston, C.; Magarvey, N. A., DeepRiPP integrates multiomics data to automate discovery of novel ribosomally synthesized natural products. *Proceedings of the National Academy of Sciences* **2020,** *117* (1), 371-380.

95.   Hegemann, J. D.; Bobeica, S. C.; Walker, M. C.; Bothwell, I. R.; van der Donk, W. A., Assessing the Flexibility of the Prochlorosin 2.8 Scaffold for Bioengineering Applications. *ACS Synth Biol* **2019,** *8* (5), 1204-1214.

96.   Essig, A.; Hofmann, D.; Münch, D.; Gayathri, S.; Künzler, M.; Kallio, P. T.; Sahl, H. G.; Wider, G.; Schneider, T.; Aebi, M., Copsin, a novel peptide-based fungal antibiotic interfering with the peptidoglycan synthesis. *J Biol Chem* **2014,** *289* (50), 34953-64.

97.   Férir, G.; Petrova, M. I.; Andrei, G.; Huskens, D.; Hoorelbeke, B.; Snoeck, R.; Vanderleyden, J.; Balzarini, J.; Bartoschek, S.; Brönstrup, M.; Süssmuth, R. D.; Schols, D., The lantibiotic peptide labyrinthopeptin A1 demonstrates broad anti-HIV and anti-HSV activity with potential for microbicidal applications. *PLoS One* **2013,** *8* (5), e64010.

98.   Mohr, K. I.; Volz, C.; Jansen, R.; Wray, V.; Hoffmann, J.; Bernecker, S.; Wink, J.; Gerth, K.; Stadler, M.; Müller, R., Pinensins: the first antifungal lantibiotics. *Angew Chem Int Ed Engl* **2015,** *54* (38), 11254-8.

99.   Zheng, Q.; Wang, Q.; Wang, S.; Wu, J.; Gao, Q.; Liu, W., Thiopeptide Antibiotics Exhibit a Dual Mode of Action against Intracellular Pathogens by Affecting Both Host and Microbe. *Chem Biol* **2015,** *22* (8), 1002-7.

100.   Russell, A. H.; Truman, A. W., Genome mining strategies for ribosomally synthesised and post-translationally modified peptides. *Computational and Structural Biotechnology Journal* **2020,** *18*, 1838-1851.

101.   Skinnider, M. A.; Johnston, C. W.; Edgar, R. E.; Dejong, C. A.; Merwin, N. J.; Rees, P. N.; Magarvey, N. A., Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining. *Proceedings of the National Academy of Sciences* **2016,** *113* (42), E6343-E6351.

102.   Amiri Moghaddam, J.; Jautzus, T.; Alanjary, M.; Beemelmanns, C., Recent highlights of biosynthetic studies on marine natural products. *Organic & Biomolecular Chemistry* **2020**.

103.   Kautsar, S. A.; Blin, K.; Shaw, S.; Weber, T.; Medema, M. H., BiG-FAM: the biosynthetic gene cluster families database. *Nucleic Acids Research* **2020**.

104.   Kloosterman, A. M.; Medema, M. H.; van Wezel, G. P., Omics-based strategies to discover novel classes of RiPP natural products. *Curr Opin Biotechnol* **2020,** *69*, 60-67.

105.   Elgersma, J. N.; Rusting, N.; van Berne, J. H. Werkwijze voor de bereiding van een antibiotische stof met behulp van actinomyceten van het geslacht Streptomyes 79749, 1955.

106.   Nakamura, S.; Yajima, T.; Lin, Y.; Umezawa, H., Isolation and characterization of bottromycins A2, B2, C2. *J Antibiot (Tokyo)* **1967,** *20* (1), 1-5.

107.   Kobayashi, Y.; Ichioka, M.; Hirose, T.; Nagai, K.; Matsumoto, A.; Matsui, H.; Hanaki, H.; Masuma, R.; Takahashi, Y.; Omura, S.; Sunazuka, T., Bottromycin derivatives: efficient chemical modifications of the ester moiety and evaluation of anti-MRSA and anti-VRE activities. *Bioorg Med Chem Lett* **2010,** *20* (20), 6116-6120.

108.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2. Release of aminoacyl or peptidyl tRNA from ribosomes. *The Journal of biological chemistry* **1976,** *251*, 2299-306.

109.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2: Effect of bottromycin A2 on polysomes. *FEBS Letters* **1983,** *153* (1), 53-59.

110.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect on peptide bond formation. *FEBS Letters* **1981,** *123* (2), 173-176.

111.   Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Ōmura, S., Structure Determination and Total Synthesis of Bottromycin A2: A Potent Antibiotic against MRSA and VRE. *Angewandte Chemie International Edition* **2009,** *48* (5), 914-917.

112. Crone, W. J. K.; Leeper, F. J.; Truman, A. W., Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chemical Science* **2012,** *3* (12), 3516-3521.

113. Huo, L.; Rachid, S.; Stadler, M.; Wenzel, Silke C.; Müller, R., Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis. *Chemistry & Biology* **2012,** *19* (10), 1278-1287.

114. Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chemical Science* **2012,** *3* (12), 3522-3525.

115. Hou, Y.; Tianero, M. D. B.; 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., Structure and Biosynthesis of the Antibiotic Bottromycin D. *Organic Letters* **2012,** *14* (19), 5050-5053.

116. Crone, W. J. K.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angewandte Chemie International Edition* **2016,** *55* (33), 9639-9643.

117. Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Müller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *ChemBioChem* **2016,** *17* (23), 2286-2292.

118. Vior, N. M.; Cea-Torrescassana, E.; Eyles, T. H.; Chandra, G.; Truman, A. W., Regulation of Bottromycin Biosynthesis Involves an Internal Transcriptional Start Site and a Cluster-Situated Modulator. *Frontiers in Microbiology* **2020,** *11* (495).

119. Zhang, W.; Tang, Y., Combinatorial Biosynthesis of Natural Products. *Journal of Medicinal Chemistry* **2008,** *51* (9), 2629-2633.

120. Zhou, H.; Xie, X.; Tang, Y., Engineering natural products using combinatorial biosynthesis and biocatalysis. *Curr Opin Biotechnol* **2008,** *19* (6), 590-6.

121. Ogasawara, Y., New enzymes for peptide biosynthesis in microorganisms. *Bioscience, Biotechnology, and Biochemistry* **2019,** *83* (4), 589-597.

122. Adams, J. P.; Brown, M. J. B.; Diaz-Rodriguez, A.; Lloyd, R. C.; Roiban, G.-D., Biocatalysis: A Pharma Perspective. *Advanced Synthesis & Catalysis* **2019,** *361* (11), 2421-2432.

123. Abdelraheem, E. M. M.; Busch, H.; Hanefeld, U.; Tonin, F., Biocatalysis explained: from pharmaceutical to bulk chemical production. *Reaction Chemistry & Engineering* **2019,** *4* (11), 1878-1894.

124. Bommarius, A. S., Biocatalysis: A Status Report. *Annual Review of Chemical and Biomolecular Engineering* **2015,** *6* (1), 319-345.

125. Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J., Recent advances in engineering nonribosomal peptide assembly lines. *Nat Prod Rep* **2016,** *33* (2), 317-47.

126. Traber, R.; Hofmann, H.; Kobel, H., Cyclosporins--new analogues by precursor directed biosynthesis. *J Antibiot (Tokyo)* **1989,** *42* (4), 591-7.

127. Hojati, Z.; Milne, C.; Harvey, B.; Gordon, L.; Borg, M.; Flett, F.; Wilkinson, B.; Sidebottom, P. J.; Rudd, B. A. M.; Hayes, M. A.; Smith, C. P.; Micklefield, J., Structure, Biosynthetic Origin, and Engineered Biosynthesis of Calcium-Dependent Antibiotics from Streptomyces coelicolor. *Chemistry & Biology* **2002,** *9* (11), 1175-1187.

128. Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A., The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **1999,** *6* (8), 493-505.

129. von Döhren, H.; Dieckmann, R.; Pavela-Vrancic, M., The nonribosomal code. *Chemistry & biology* **1999,** *6* (10), R273-R279.

130. Stachelhaus, T.; Marahiel, M. A., Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J Biol Chem* **1995,** *270* (11), 6163-9.

131. Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P., Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J* **1997,** *16* (14), 4174-4183.

132. Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H., Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* **2005,** *33* (18), 5799-808.

133. Evans, B. S.; Chen, Y.; Metcalf, W. W.; Zhao, H.; Kelleher, N. L., Directed evolution of the nonribosomal peptide synthetase AdmK generates new andrimid derivatives in vivo. *Chem Biol* **2011,** *18* (5), 601-7.

134. Villiers, B.; Hollfelder, F., Directed evolution of a gatekeeper domain in nonribosomal peptide synthesis. *Chem Biol* **2011,** *18* (10), 1290-9.

135. Eppelmann, K.; Stachelhaus, T.; Marahiel, M. A., Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* **2002,** *41* (30), 9718-26.

136. Kaljunen, H.; Schiefelbein, S. H.; Stummer, D.; Kozak, S.; Meijers, R.; Christiansen, G.; Rentmeister, A., Structural Elucidation of the Bispecificity of A Domains as a Basis for Activating Non-natural Amino Acids. *Angew Chem Int Ed Engl* **2015,** *54* (30), 8833-6.

137. Miao, V.; Coëffet-Le Gal, M. F.; Nguyen, K.; Brian, P.; Penn, J.; Whiting, A.; Steele, J.; Kau, D.; Martin, S.; Ford, R.; Gibson, T.; Bouchard, M.; Wrigley, S. K.; Baltz, R. H., Genetic engineering in Streptomyces roseosporus to produce hybrid lipopeptide antibiotics. *Chem Biol* **2006,** *13* (3), 269-76.

138. Nguyen, K. T.; Ritz, D.; Gu, J. Q.; Alexander, D.; Chu, M.; Miao, V.; Brian, P.; Baltz, R. H., Combinatorial biosynthesis of novel antibiotics related to daptomycin. *Proc Natl Acad Sci U S A* **2006,** *103* (46), 17462-7.

139. Calcott, M. J.; Ackerley, D. F., Genetic manipulation of non-ribosomal peptide synthetases to generate novel bioactive peptide products. *Biotechnol Lett* **2014,** *36* (12), 2407-16.

140. Butz, D.; Schmiederer, T.; Hadatsch, B.; Wohlleben, W.; Weber, T.; Süssmuth, R. D., Module extension of a non-ribosomal peptide synthetase of the glycopeptide antibiotic balhimycin produced by Amycolatopsis balhimycina. *Chembiochem* **2008,** *9* (8), 1195-200.

141. Wu, M. C.; Law, B.; Wilkinson, B.; Micklefield, J., Bioengineering natural product biosynthetic pathways for therapeutic applications. *Curr Opin Biotechnol* **2012,** *23* (6), 931-40.

142. Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B., De novo design and engineering of non-ribosomal peptide synthetases. *Nature Chemistry* **2018,** *10* (3), 275-281.

143. Bozhüyük, K. A. J.; Linck, A.; Tietze, A.; Kranz, J.; Wesche, F.; Nowak, S.; Fleischhacker, F.; Shi, Y.-N.; Grün, P.; Bode, H. B., Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nature Chemistry* **2019,** *11* (7), 653-661.

144. Huang, H. M.; Stephan, P.; Kries, H., Engineering DNA-Templated Nonribosomal Peptide Synthesis. *Cell Chem Biol* **2020**.

145. Trauger, J. W.; Kohli, R. M.; Walsh, C. T., Cyclization of Backbone-Substituted Peptides Catalyzed by the Thioesterase Domain from the Tyrocidine Nonribosomal Peptide Synthetase. *Biochemistry* **2001,** *40* (24), 7092-7098.

146. Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T., Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry* **2001,** *40* (24), 7099-108.

147. Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T., Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **2000,** *407* (6801), 215-218.

148. Kohli, R. M.; Takagi, J.; Walsh, C. T., The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides. *Proc Natl Acad Sci U S A* **2002,** *99* (3), 1247-52.

149. Grünewald, J.; Marahiel, M. A., Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides. *Microbiol Mol Biol Rev* **2006,** *70* (1), 121-146.

150. Pavlova, O.; Mukhopadhyay, J.; Sineva, E.; Ebright, R. H.; Severinov, K., Systematic structure-activity analysis of microcin J25. *The Journal of biological chemistry* **2008,** *283* (37), 25589-25595.

151. Just-Baringo, X.; Albericio, F.; Álvarez, M., Thiopeptide engineering: a multidisciplinary effort towards future drugs. *Angew Chem Int Ed Engl* **2014,** *53* (26), 6602-16.

152. Maksimov, M. O.; Pan, S. J.; James Link, A., Lasso peptides: structure, function, biosynthesis, and engineering. *Nat Prod Rep* **2012,** *29* (9), 996-1006.

153. Field, D.; Cotter, P. D.; Ross, R. P.; Hill, C., Bioengineering of the model lantibiotic nisin. *Bioengineered* **2015,** *6* (4), 187-92.

154. Ruffner, D. E.; Schmidt, E. W.; Heemstra, J. R., Assessing the combinatorial potential of the RiPP cyanobactin tru pathway. *ACS Synth Biol* **2015,** *4* (4), 482-92.

155. Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A., Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Cent Sci* **2017,** *3* (6), 629-638.

156. Cubillos-Ruiz, A.; Berta-Thompson, J. W.; Becker, J. W.; van der Donk, W. A.; Chisholm, S. W., Evolutionary radiation of lanthipeptides in marine cyanobacteria. *Proceedings of the National Academy of Sciences* **2017,** *114* (27), E5424-E5433.

157. Donia, M. S.; Hathaway, B. J.; Sudek, S.; Haygood, M. G.; Rosovitz, M. J.; Ravel, J.; Schmidt, E. W., Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. *Nature Chemical Biology* **2006,** *2* (12), 729-735.

158. Goto, Y.; Suga, H., Engineering of RiPP pathways for the production of artificial peptides bearing various non-proteinogenic structures. *Curr Opin Chem Biol* **2018,** *46*, 82-90.

159. Sardar, D.; Pierce, E.; McIntosh, J. A.; Schmidt, E. W., Recognition Sequences and Substrate Evolution in Cyanobactin Biosynthesis. *ACS Synthetic Biology* **2015,** *4* (2), 167-176.

160. Zhang, Q.; Yang, X.; Wang, H.; van der Donk, W. A., High Divergence of the Precursor Peptides in Combinatorial Lanthipeptide Biosynthesis. *ACS Chemical Biology* **2014,** *9* (11), 2686-2694.

161. Young, T. S.; Dorrestein, P. C.; Walsh, C. T., Codon randomization for rapid exploration of chemical space in thiopeptide antibiotic variants. *Chemistry & biology* **2012,** *19* (12), 1600-1610.

162. Rollema, H. S.; Kuipers, O. P.; Both, P.; de Vos, W. M.; Siezen, R. J., Improvement of solubility and stability of the antimicrobial peptide nisin by protein engineering. *Applied and Environmental Microbiology* **1995,** *61* (8), 2873-2878.

163. Li, Q.; Montalban-Lopez, M.; Kuipers, O. P., Increasing the Antimicrobial Activity of Nisin-Based Lantibiotics against Gram-Negative Pathogens. *Applied and environmental microbiology* **2018,** *84* (12), e00052-18.

164. Knappe, T. A.; Manzenrieder, F.; Mas-Moruno, C.; Linne, U.; Sasse, F.; Kessler, H.; Xie, X.; Marahiel, M. A., Introducing lasso peptides as molecular scaffolds for drug design: engineering of an integrin antagonist. *Angew Chem Int Ed Engl* **2011,** *50* (37), 8714-7.

165. Hegemann, J. D.; De Simone, M.; Zimmermann, M.; Knappe, T. A.; Xie, X.; Di Leva, F. S.; Marinelli, L.; Novellino, E.; Zahler, S.; Kessler, H.; Marahiel, M. A.,

Rational Improvement of the Affinity and Selectivity of Integrin Binding of Grafted Lasso Peptides. *Journal of Medicinal Chemistry* **2014,** *57* (13), 5829-5834.

166. Yang, X.; Lennard, K. R.; He, C.; Walker, M. C.; Ball, A. T.; Doigneaux, C.; Tavassoli, A.; van der Donk, W. A., A lanthipeptide library used to identify a protein-protein interaction inhibitor. *Nature chemical biology* **2018,** *14* (4), 375-380.

167. Al Toma, R. S.; Kuthning, A.; Exner, M. P.; Denisiuk, A.; Ziegler, J.; Budisa, N.; Süssmuth, R. D., Site-Directed and Global Incorporation of Orthogonal and Isostructural Noncanonical Amino Acids into the Ribosomal Lasso Peptide Capistruin. *ChemBioChem* **2015,** *16* (3), 503-509.

168. Goto, Y.; Katoh, T.; Suga, H., Flexizymes for genetic code reprogramming. *Nat Protoc* **2011,** *6* (6), 779-90.

169. Kakkar, N.; Perez, J. G.; Liu, W. R.; Jewett, M. C.; van der Donk, W. A., Incorporation of Nonproteinogenic Amino Acids in Class I and II Lantibiotics. *ACS Chemical Biology* **2018,** *13* (4), 951-957.

170. Bosma, T.; Kuipers, A.; Bulten, E.; de Vries, L.; Rink, R.; Moll, G. N., Bacterial display and screening of posttranslationally thioether-stabilized peptides. *Applied and environmental microbiology* **2011,** *77* (19), 6794-6801.

171. Urban, J. H.; Moosmeier, M. A.; Aumüller, T.; Thein, M.; Bosma, T.; Rink, R.; Groth, K.; Zulley, M.; Siegers, K.; Tissot, K.; Moll, G. N.; Prassler, J., Phage display and selection of lanthipeptides on the carboxy-terminus of the gene-3 minor coat protein. *Nat Commun* **2017,** *8* (1), 1500.

172. Hetrick, K. J.; Walker, M. C.; van der Donk, W. A., Development and Application of Yeast and Phage Display of Diverse Lanthipeptides. *ACS Central Science* **2018,** *4* (4), 458-467.

173. Schmitt, S.; Montalbán-López, M.; Peterhoff, D.; Deng, J.; Wagner, R.; Held, M.; Kuipers, O. P.; Panke, S., Analysis of modular bioengineered antimicrobial lanthipeptides at nanoliter scale. *Nature Chemical Biology* **2019,** *15* (5), 437-443.

174. Si, T.; Tian, Q.; Min, Y.; Zhang, L.; Sweedler, J. V.; van der Donk, W. A.; Zhao, H., Rapid Screening of Lanthipeptide Analogs via In-Colony Removal of Leader Peptides in Escherichia coli. *J Am Chem Soc* **2018,** *140* (38), 11884-11888.

175. Liu, R.; Zhang, Y.; Zhai, G.; Fu, S.; Xia, Y.; Hu, B.; Cai, X.; Zhang, Y.; Li, Y.; Deng, Z.; Liu, T., A Cell-Free Platform Based on Nisin Biosynthesis for Discovering Novel Lanthipeptides and Guiding their Overproduction In Vivo. *Adv Sci (Weinh)* **2020,** *7* (17), 2001616-2001616.

176. Levengood, M. R.; Knerr, P. J.; Oman, T. J.; van der Donk, W. A., In Vitro Mutasynthesis of Lantibiotic Analogues Containing Nonproteinogenic Amino Acids. *Journal of the American Chemical Society* **2009,** *131* (34), 12024-12025.

177. Levengood, M. R.; Kerwood, C. C.; Chatterjee, C.; van der Donk, W. A., Investigation of the Substrate Specificity of Lacticin 481 Synthetase by Using Nonproteinogenic Amino Acids. *ChemBioChem* **2009,** *10* (5), 911-919.

178. Ozaki, T.; Yamashita, K.; Goto, Y.; Shimomura, M.; Hayashi, S.; Asamizu, S.; Sugai, Y.; Ikeda, H.; Suga, H.; Onaka, H., Dissection of goadsporin biosynthesis by in vitro reconstitution leading to designer analogues expressed in vivo. *Nature Communications* **2017,** *8* (1), 14207.

179. Goto, Y.; Ito, Y.; Kato, Y.; Tsunoda, S.; Suga, H., One-Pot Synthesis of Azoline-Containing Peptides in a Cell-free Translation System Integrated with a Posttranslational Cyclodehydratase. *Chemistry & Biology* **2014,** *21* (6), 766-774.

180. Goto, Y.; Suga, H., In Vitro Biosynthesis of Peptides Containing Exotic Azoline Analogues. *ChemBioChem* **2020,** *21* (1-2), 84-87.

181. Vinogradov, A. A.; Shimomura, M.; Goto, Y.; Ozaki, T.; Asamizu, S.; Sugai, Y.; Suga, H.; Onaka, H., Minimal lactazole scaffold for in vitro thiopeptide bioengineering. *Nature Communications* **2020,** *11* (1), 2272.

182. Goto, Y.; Suga, H., Artificial In Vitro Biosynthesis Systems for the Development of Pseudo-Natural Products. *Bulletin of the Chemical Society of Japan* **2018,** *91* (3), 410-419.

183. Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A., Flexizyme-Enabled Benchtop Biosynthesis of Thiopeptides. *Journal of the American Chemical Society* **2019,** *141* (2), 758-762.

184. Knerr, P. J.; Oman, T. J.; Garcia De Gonzalo, C. V.; Lupoli, T. J.; Walker, S.; van der Donk, W. A., Non-proteinogenic Amino Acids in Lacticin 481 Analogues Result in More Potent Inhibition of Peptidoglycan Transglycosylation. *ACS Chemical Biology* **2012,** *7* (11), 1791-1795.

185. Oman, T. J.; Knerr, P. J.; Bindman, N. A.; Velásquez, J. E.; van der Donk, W. A., An Engineered Lantibiotic Synthetase That Does Not Require a Leader Peptide on Its Substrate. *Journal of the American Chemical Society* **2012,** *134* (16), 6952-6955.

186. Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W.; Jaspars, M.; Naismith, J. H., Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat Chem Biol* **2015,** *11* (8), 558-563.

187. Reyna-González, E.; Schmid, B.; Petras, D.; Süssmuth, R. D.; Dittmann, E., Leader Peptide-Free In Vitro Reconstitution of Microviridin Biosynthesis Enables Design of Synthetic Protease-Targeted Libraries. *Angew Chem Int Ed Engl* **2016,** *55* (32), 9398-401.

188. Nguyen, G.; Kam, A.; Loo, S.; Jansson, A.; Pan, L.; Tam, J., Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. *Journal of the American Chemical Society* **2015,** *137*.

189. Hemu, X.; Qiu, Y.; Nguyen, G. K. T.; Tam, J. P., Total Synthesis of Circular Bacteriocins by Butelase 1. *Journal of the American Chemical Society* **2016,** *138* (22), 6968-6971.

190. Cao, Y.; Nguyen, G.; Chuah, S.; Tam, J.; Liu, C.-F., Butelase-Mediated Ligation as an Efficient Bioconjugation Method for the Synthesis of Peptide Dendrimers. *Bioconjugate chemistry* **2016,** *27*.

191. Nguyen, G.; Wang, S.; Qiu, Y.; Hemu, X.; Lian, Y.; Tam, J., Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nature chemical biology* **2014,** *10*.

192. Oueis, E.; Stevenson, H.; Jaspars, M.; Westwood, N.; Naismith, J., Bypassing the proline/thiazoline requirement of the macrocyclase PatG. *Chemical Communications* **2017,** *53*.

193. Oueis, E.; Nardone, B.; Jaspars, M.; Westwood, N.; Naismith, J., Synthesis of Hybrid Cyclopeptides through Enzymatic Macrocyclization. *ChemistryOpen* **2017,** *6*, 11 - 14.

194. Ludewig, H.; Czekster, C.; Oueis, E.; Munday, E.; Arshad, M.; Synowsky, S.; Bent, A.; Naismith, J., Characterization of the Fast and Promiscuous Macrocyclase from Plant PCY1 Enables the Use of Simple Substrates. *ACS Chemical Biology* **2018,** *13*.

195. van der Velden, N. S.; Kälin, N.; Helf, M. J.; Piel, J.; Freeman, M. F.; Künzler, M., Autocatalytic backbone N-methylation in a family of ribosomal peptide natural products. *Nature Chemical Biology* **2017,** *13* (8), 833-835.

196. Sikandar, A.; Koehnke, J., The role of protein–protein interactions in the biosynthesis of ribosomally synthesized and post-translationally modified peptides. *Natural Product Reports* **2019,** *36* (11), 1576-1588.

Chapter 2

Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme

Laura Franz,[+] Sebastian Adam,[+] Javier Santos-Aberturas, Andrew W. Truman, and Jesko Koehnke

[+] shared authorship

Author contributions:

Laura Franz contributed to the manuscript by performing and analyzing the *in vitro* biochemical experiments, demonstrating the function of the YcaO domain enzymes and providing biochemical characterization of the enzymes. She cloned, expressed and purified the (mutant) precursor peptides. Sebastian Adam contributed by cloning the YcaO (mutant) enzymes and establishing enzyme expression and purification protocols for the two featured YcaO proteins. He also helped with the purification of mutant precursor peptides. Javier Santos-Aberturas and Andrew W. Truman contributed by designing, performing and evaluating the featured *in vivo* experiments and performed and analyzed the MS fragmentation experiments. Jesko Köhnke supervised and guided the project. Jesko Köhnke, Laura Franz and Sebastian Adam wrote the manuscript.

All authors contributed by designing figures for their respective experiments, provided critical feedback and discussed the results of the manuscript.

2 Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme

2.1 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.
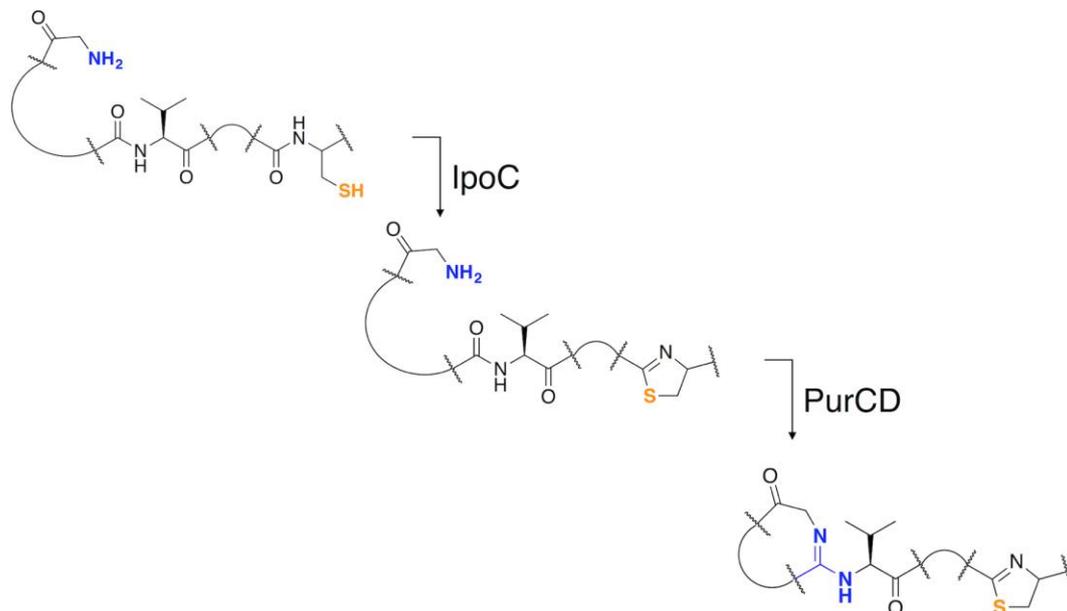


Figure 2-1 | Graphical abstract.

2.2 Main Text

Bottromycins (Figure 2-1) 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 (Figure 2-2).[6] The bottromycin biosynthetic gene cluster contains two unusual, stand-alone YcaO-domain enzymes, BotC and BotCD (sequence identity of 26%, Figure 2-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]
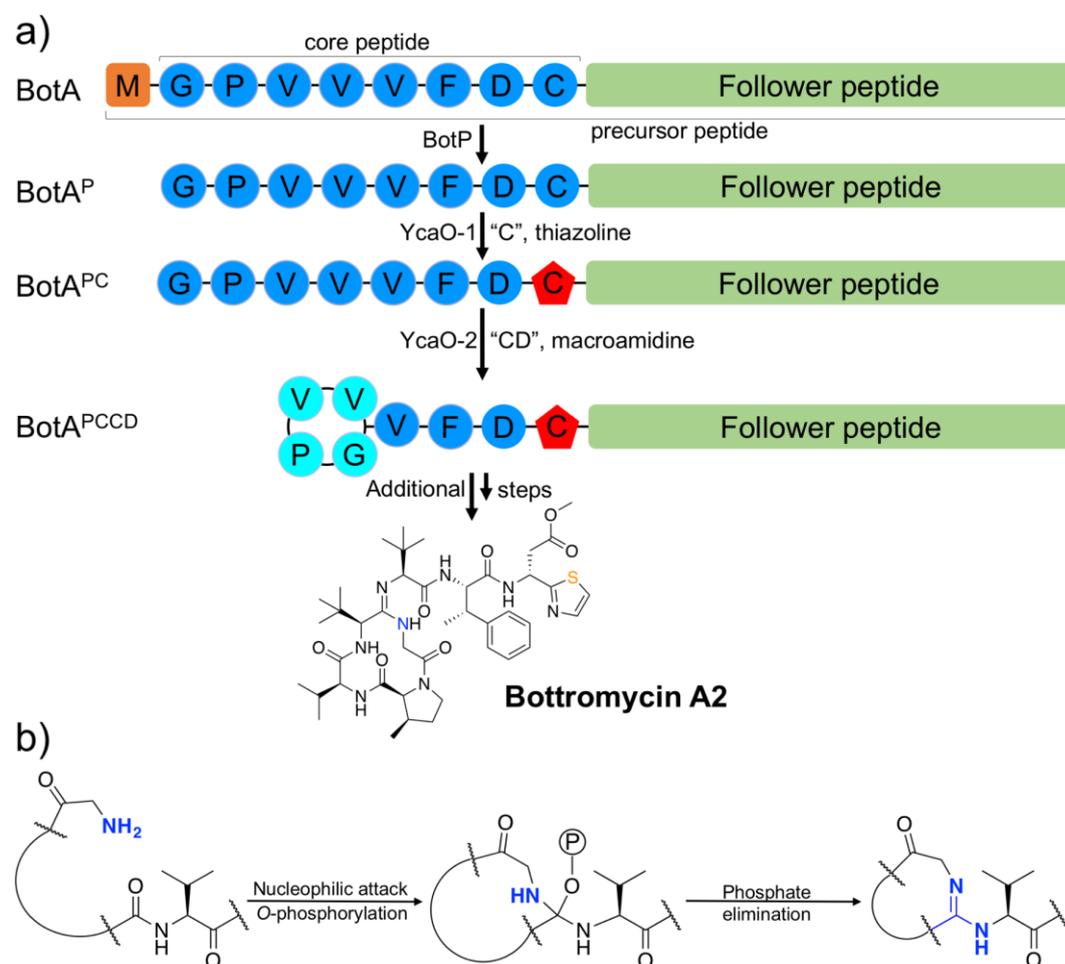
Figure 2-2 | a) Bottromycin biosynthesis with the proposed functions of the two YcaO enzymes. **b)** Use of a terminal amine by a YcaO enzyme as a nucleophile to achieve macroamidine formation.

45

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 multi-domain heterocyclase, the heterocyclase activity is minimal.[11] The YcaO domain of the responsible enzyme catalyzes the nucleophilic attack of a Cys/Ser/Thr side-chain 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 2-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 (Figure 2-2b and Figure 2-S1b).
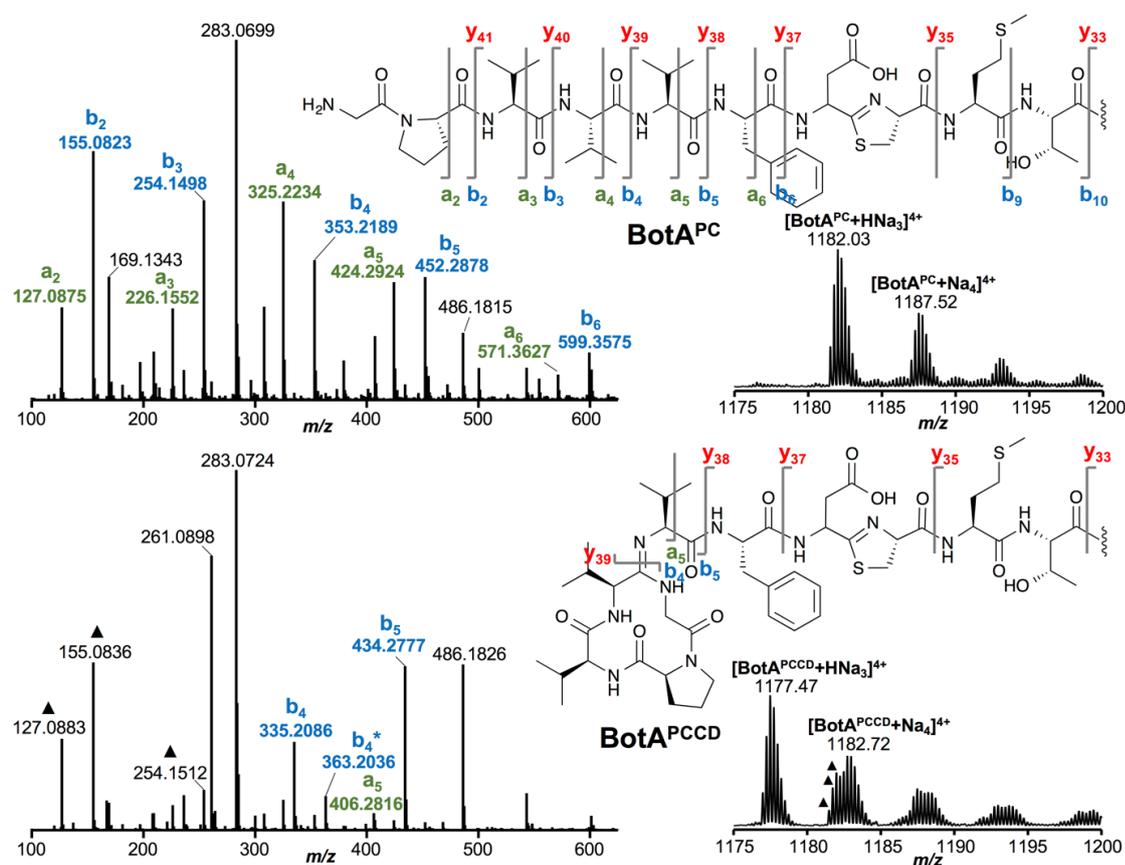
Figure 2-3 | MS^2 characterization of BotA^P after incubation with IpoC and BotA^{PC} after incubation with PurCD. Full details of the MS^2 fragmentation are provided in Figure 2-S4-S6 and Table 2-S1-S5. Signals in the BotA^{PCCD} spectra that likely come from residual BotA^{PC} are labelled with triangles.

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 2-S3). When BotA was incubated with IpoC and ATP/Mg^{2+}, we observed a loss of 18 Da (Figure 2-S4a), and the conversion of Cys9 to thiazoline was confirmed by tandem MS (MS^2) (Figure 2-S4b and Table 2-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 2-3 and Figure 2-S5, Table 2-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 2-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 (Figure 2-3 and Figure 2-S6, Table 2-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 2-4a). In contrast, the reaction using unprocessed precursor peptide BotA required 16 h (Figure 2-4a). 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 2-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).
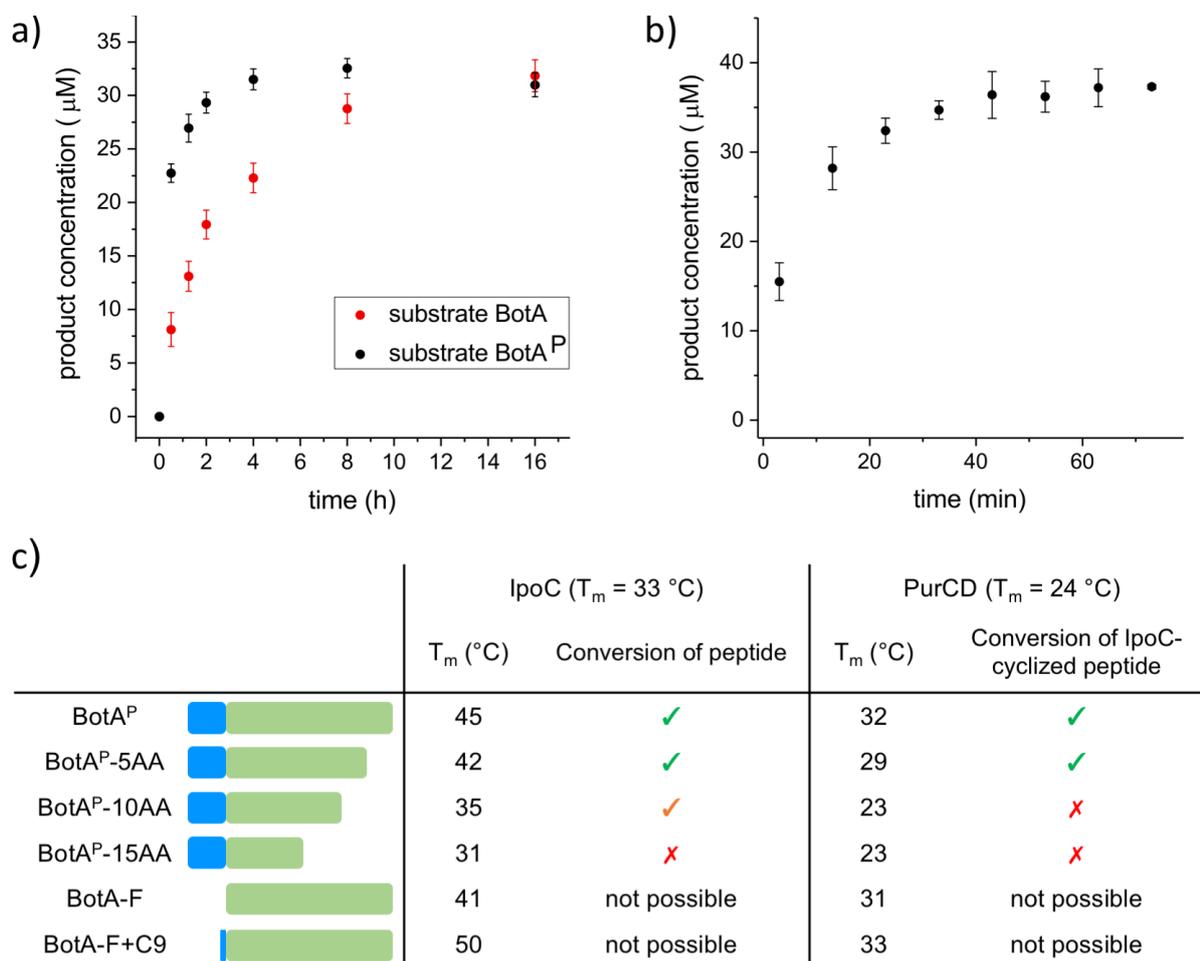
Figure 2-4 | a) Time-course of IpoC-catalyzed heterocyclization of BotA (red) and BotA^P (black) at pH 7.4. **b)** Time-course of PurCD-catalyzed macroamidine formation at pH 9.5. **c)** Effects of BotA^P truncations and follower peptide on IpoC and PurCD stability (pH 8.5) and turnover (pH 7.4 and 9.5, 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 2-S8). We hypothesized that the reaction may be accelerated by increasing the pH of the reaction as increasing the pH results in a larger proportion 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 2-S8b), with ~85% turnover after 1 h (Figure 2-4b and Figure 2-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 2-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 2-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 2-S9b and Table 2-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 2-S10a). In the case of PurCD, the differences between BotA^{PC} and BotA-F were within the experimental error (Figure 2-4c and Figure 2-S10b, Table 2-S5). 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 (Figure 2-4c and Figure 2-S10a, Table 2-S6). To identify the follower region responsible for binding we produced a series of systematic truncations (Figure 2-4c and Figure 2-S10, Table 2-S7 and Table 2-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 2-S11, Table 2-S7 and Table 2-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 2-

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*.
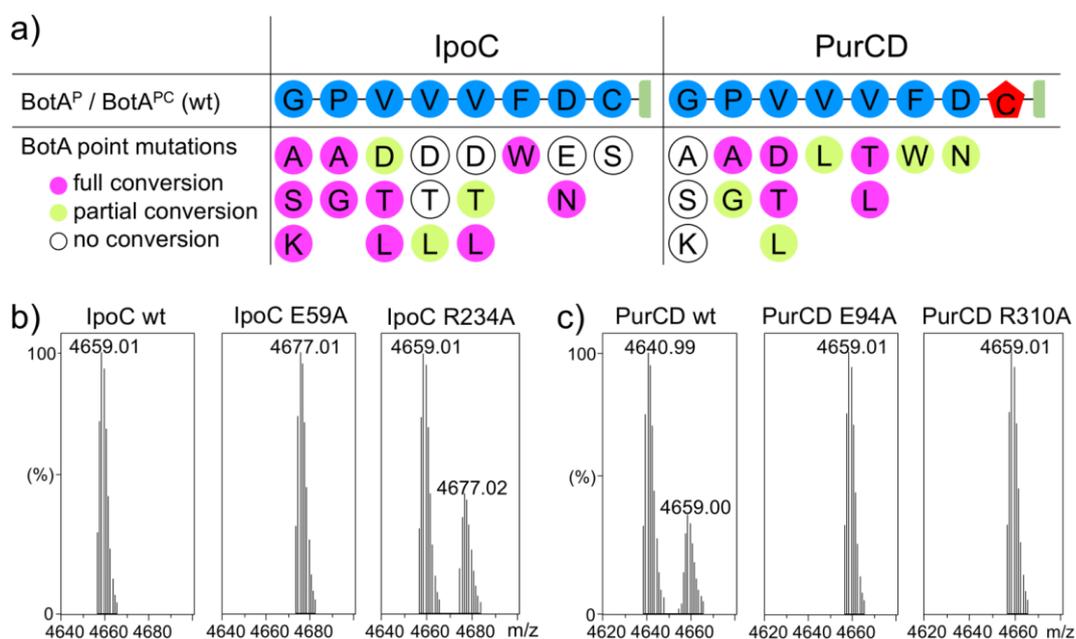
Figure 2-5 | a) Tolerance of IpoC and PurCD to changes in the core peptide of BotA. Conversion rate given in relation to wt substrate. Effect of **b)** IpoC and **c)** PurCD ATP/Mg^{2+} binding residue mutations on enzyme activity.

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 2-S13, Table 2-S7 and Table 2-S8). IpoC was unable to produce oxazoline from a C9S mutant, but processed 13 other point mutants (Figure 2-5a and Figure 2-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 (Figure 2-5a and Figure 2-S15, Table 2-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 2-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 2-S16, Table 2-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 2-S17 and Table 2-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 2-5b and Figure 2-S18). Equivalent *in vivo* mutants (Table 2-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 (Figure 2-5c, Figure 2-S19 and Figure 2-S20, Table 2-S7 and Table 2-S8). These deleterious *in vitro* results for PurCD were confirmed when the same mutations were tested *in vivo* (Table 2-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.

2.3 Supporting Information

a)

X = O, S, Se

b)
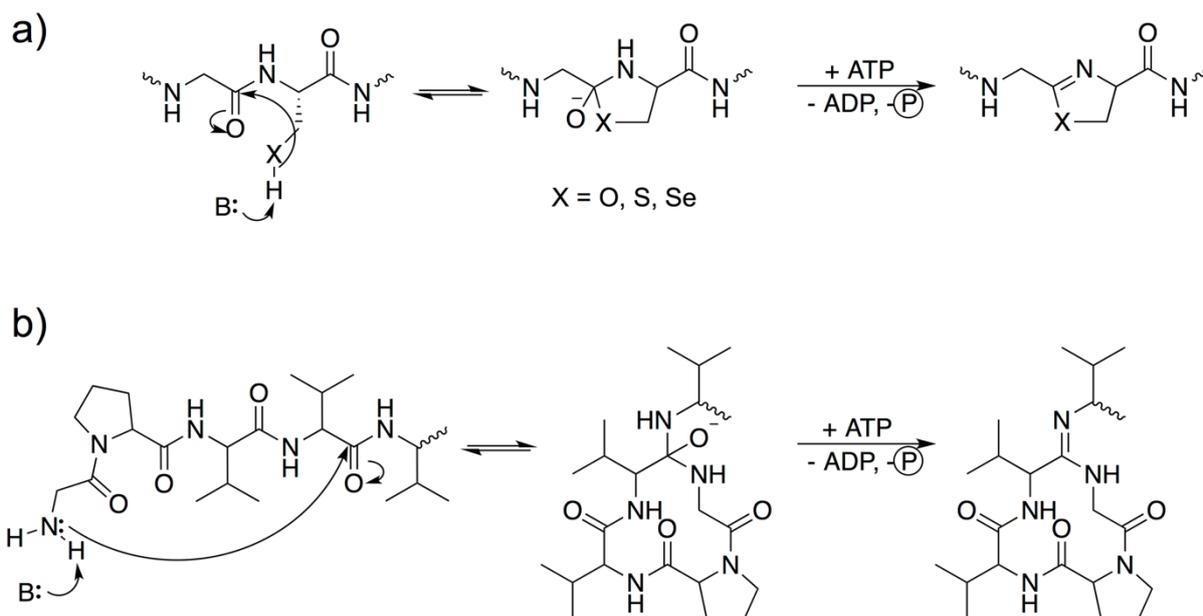
Figure 2-S1 | a) Proposed mechanism of heterocyclization by YcaO domain enzymes. ATP is used to phosphorylate the amide oxygen. Phosphate is eliminated from the hemiorthoamide intermediate afterwards and the azoline is formed. Nucleophiles reported to be utilized to date are O, S and Se. **b)** Possible route to macroamidine formation using N as a nucleophile in a reaction analogous to a).

```
BotC    --MLEATATECELR---EVVHRSY-----PSDRTVTVRCTVRPAEGTAQ-ADGYGTATTE    49
BotCD   MRLTDAEGALCSRSSGKGLGHRSMASALFESAEHYHLDWRRDPRASEAEFLPGREIAG-Q    59
          : :* .: *.      : ***      *  .  :     *  . *:  *   *  :

BotC    AVARAKALSEAVERLVACTPFATV-ARPPTTPAPGS-GSGP------VPPFP-AAGVRTA   100
BotCD   DTARRSALLRRLGGIMPDAPVLTRTYRPVEETLTGRADHGPAGRPHRHPVFLRDGGYRNW   119
         .** .** .:  ::  :*. *   **       *  .**       * *   .* *.

BotC    PDGCASREYRPLSGDGPRRVPLYWSSPWVAGEELRAGTLSAAEARLSSTVGWAVAPTPEA   160
BotCD   PHPDDDR-----------SFQPLWHYT-------------------SSAGYAAGATVHE   148
        *.   .*           . *                        *:.*:*.. * .

BotC    ALRGALLELTELIDYGVFLHRRLAGPAR----PRSAGD-------------------ET   196
BotCD   ALVHAVNELIERDAWSYQLARSYFGLPGEGPELRVVDHDTLPAELRELTGTVEEVREAPV   208
        **  *: ** *   :.  * *   *         * ...                   .

BotC    LVVPLGGAVRTPTVLAVAYGRGRLMPATGLGCGASRAEATDRALLELAQAETMWRSNPTA   256
BotCD   LIVDITCDTGVPAYVVCDARSREEVRLIGSGASPVRTYALQRALLEYLQVRTMFEHGPVD   268
        *:* :    . .*: :.       . :   * *.. *: * :***** *..**:. .*.

BotC    EPAERF-FLRRFERWPLLTRCATLDFELSDLSGLDGLSGLDGLSRLDGLSRLDHPSDLND   315
BotCD   ADTEAGQIATALARYPRHLAAAR--FDVGQLPHER-----RAFDADDGLPAAASPEQL--   319
         :*   :    : *:*    .*   *::.:*        .:.   ***    *.:*

BotC    PAGRPAPYDDGPHPASPLEELEAGGITVWADSGAVDISGPDIPRTRLCFAHVVSDPQPLL   375
BotCD   -------------LRHLVDRLRDLGVDVHH-------RVLTPPGSVTVVDVVAPGLEMF   358
                     ::.*.  *: *       *    :  ...**:    ::

BotC    GLVRAGIPVFDTGEVRRTLDPSRRDAHRDTPRGAPRDAARDRSADRGPAGRGRQGRRSA   434
BotCD   DKARAGYPVLPTGRLGERLRPRKDHE-----RGHGVDH----GKDRGGAPR-------   400
        . .*** **: **.: . * * :.      **  *   . *** * *
```

Figure 2-S2 | Alignment of BotC (top) and BotCD (bottom) created using CLUSTAL Omega.
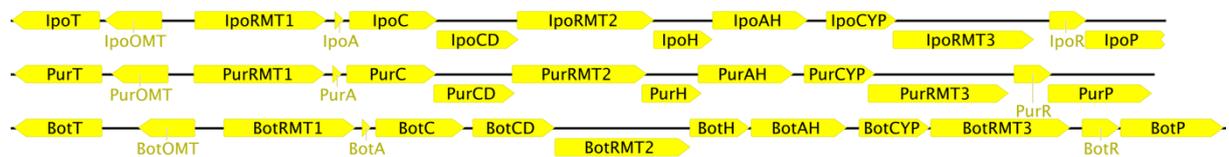
Figure 2-S3 | Gene clusters highly homologous in sequence and organisation to the bottromycin biosynthetic gene cluster identified in *Streptomyces sp.* BC16019 were identified in *Streptomyces ipomoeae* and *Streptomyces purpureus*. BotC homolog IpoC and BotCD homolog PurCD were used for this study.
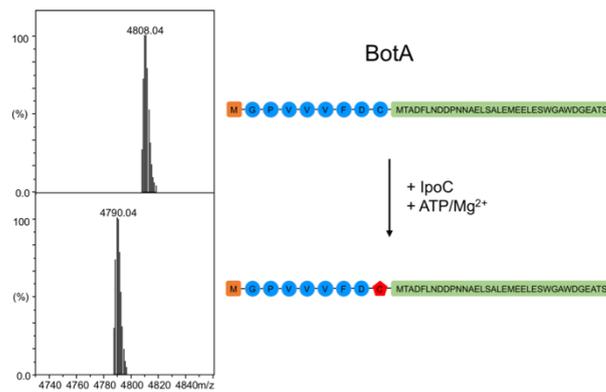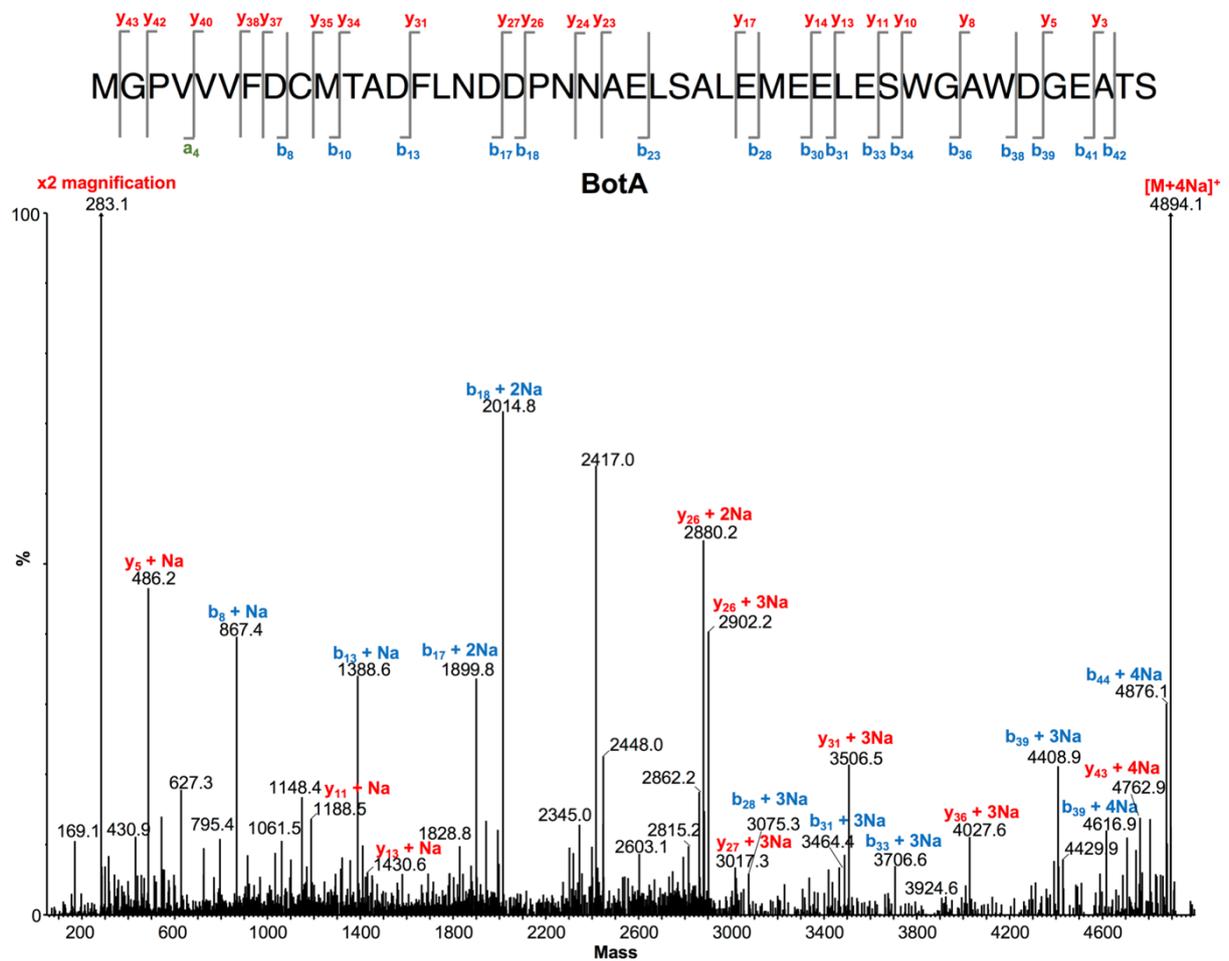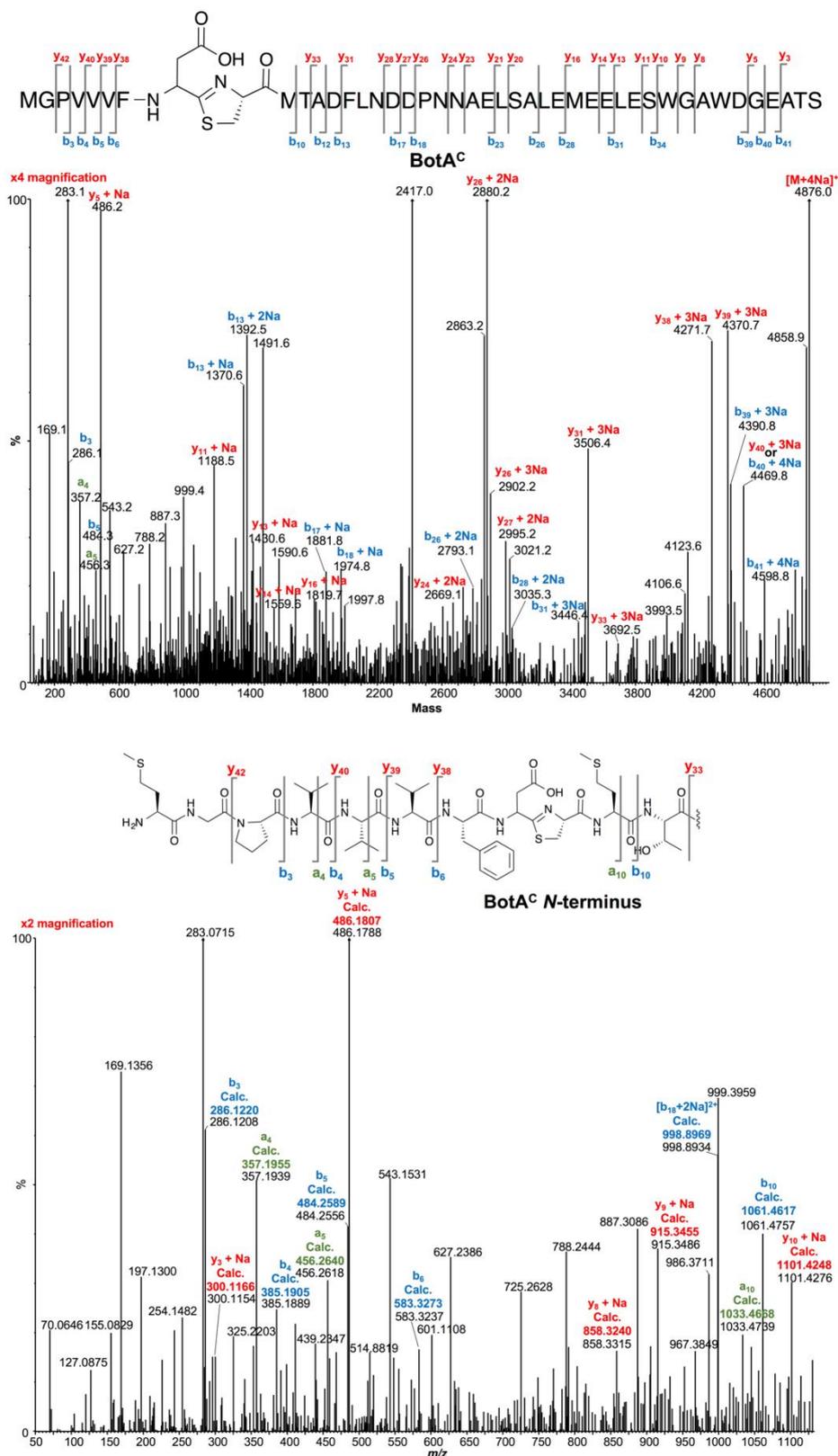
Figure 2-S4 | a) MS data showing BotA before and after incubation with IpoC and ATP/MgCl$_2$. The product, BotA^C, shows the characteristic loss of 18 Da, indicating loss of one water molecule **b)** MS^2 analysis of unmodified BotA and BotA^C. The fragments visible in the MS^2 are indicated on the peptide sequence and these are detailed in Table 2-S1 and Table 2-S2. The deconvoluted mass spectrum is displayed and is annotated with peptide fragments where possible. The presence of sodium salts meant that most fragments contained at least one sodium atom.

Figure 2-S5 | a) MS data showing BotA[P] before and after incubation with IpoC/PurCD and ATP/MgCl$_2$. The products, BotA[PC] and BotA[PCD], show the characteristic loss of 18 Da, indicating loss of one water molecule **b)** MS^2 analysis of BotA treated with aminopeptidase BotP (BotA[P]) and of BotA[P] treated with YcaO-domain enzyme IpoC (BotA[PC]). The fragments visible in the MS^2 are indicated on the peptide sequence and these are detailed in Table 2-S3 and Table 2-S4. The deconvoluted mass spectrum is displayed and is annotated with peptide fragments where possible. The presence of sodium salts meant that most fragments contained at least one sodium atom.

a)

BotA^{PC}

G-P-V-V-V-F-D-⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS

+ PurCD
+ ATP/Mg^{2+}

V-V-V-F-D-⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS
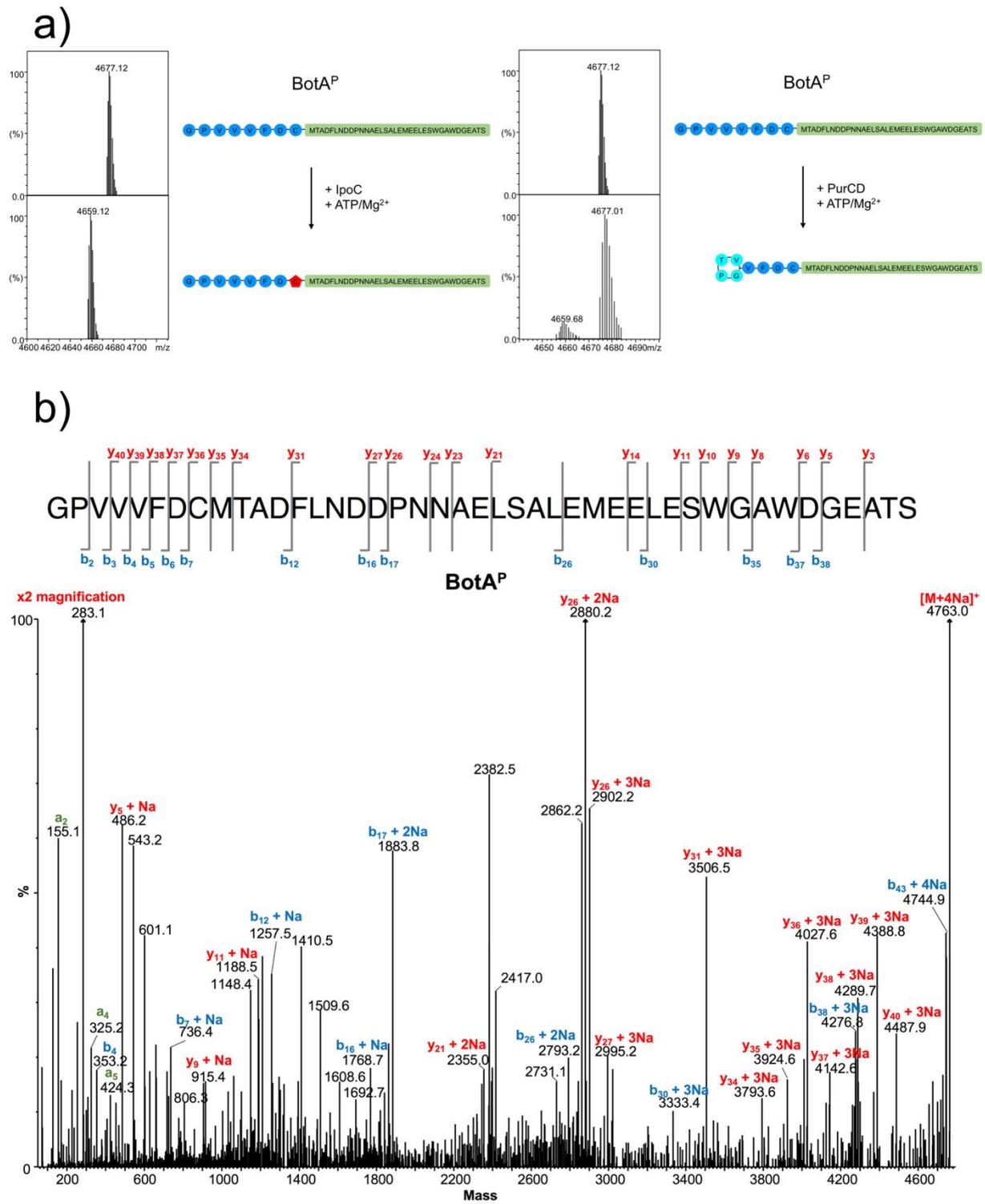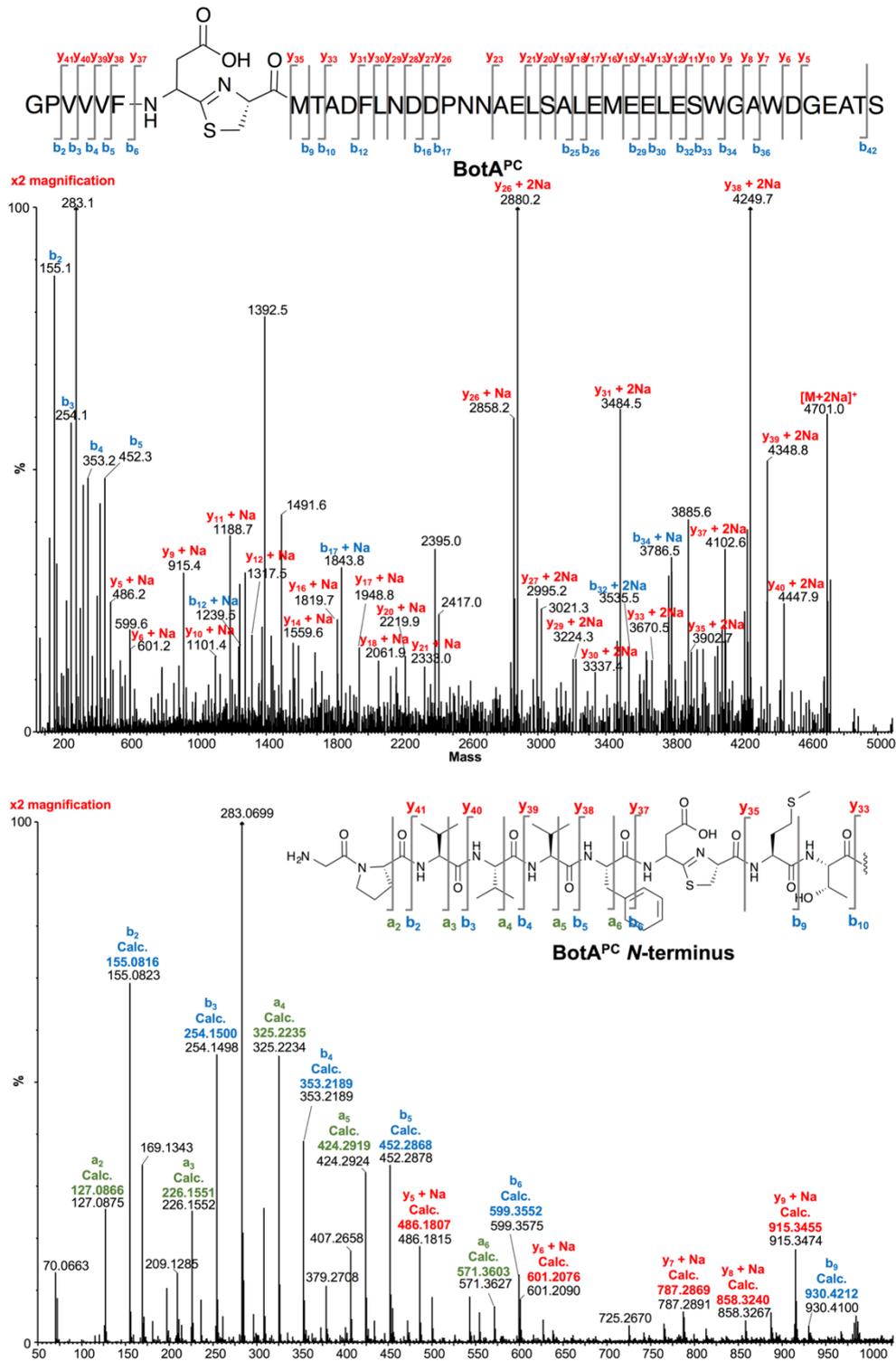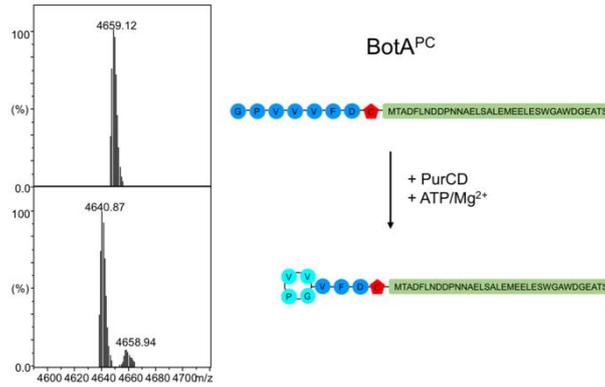P-G

b)
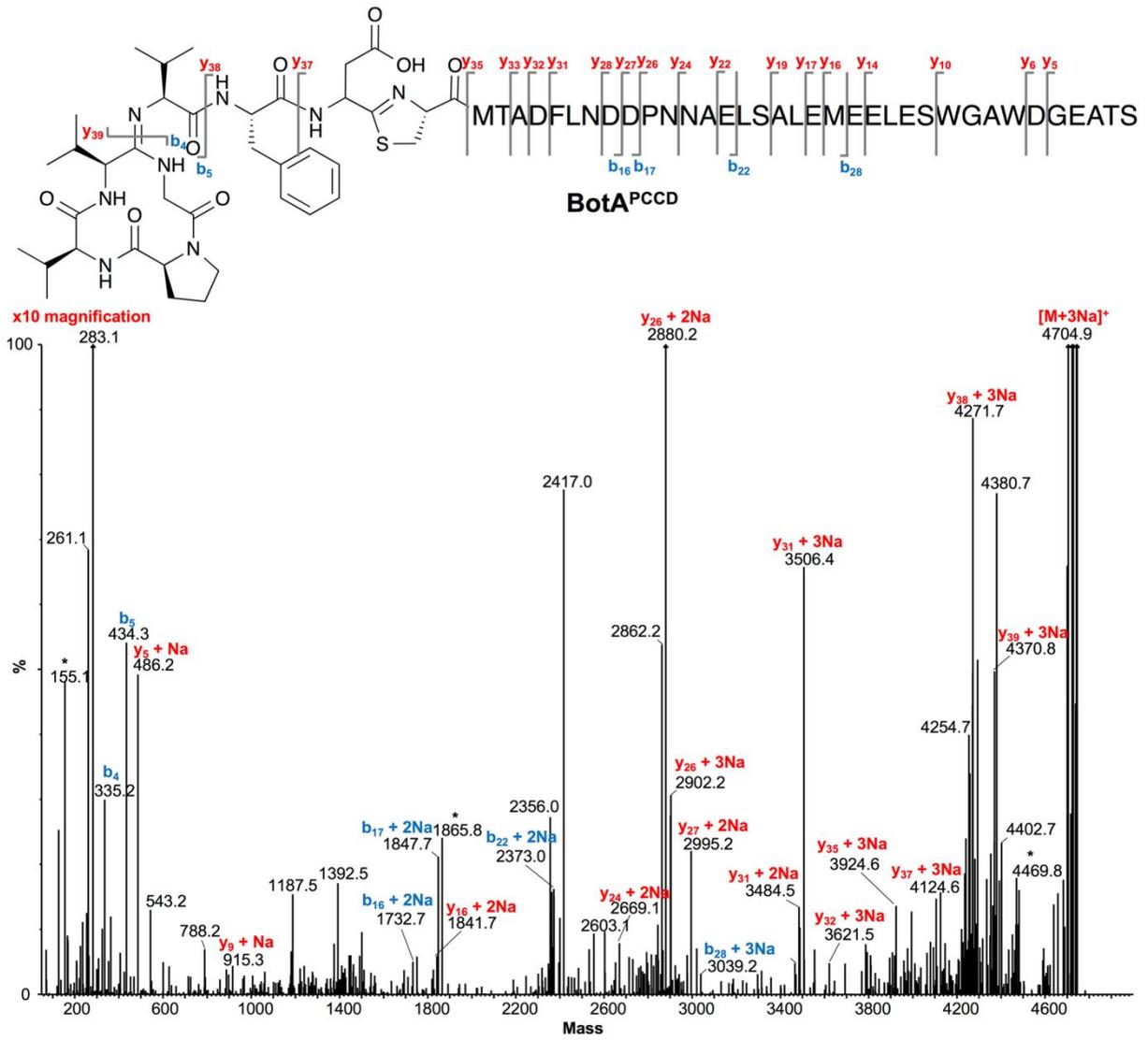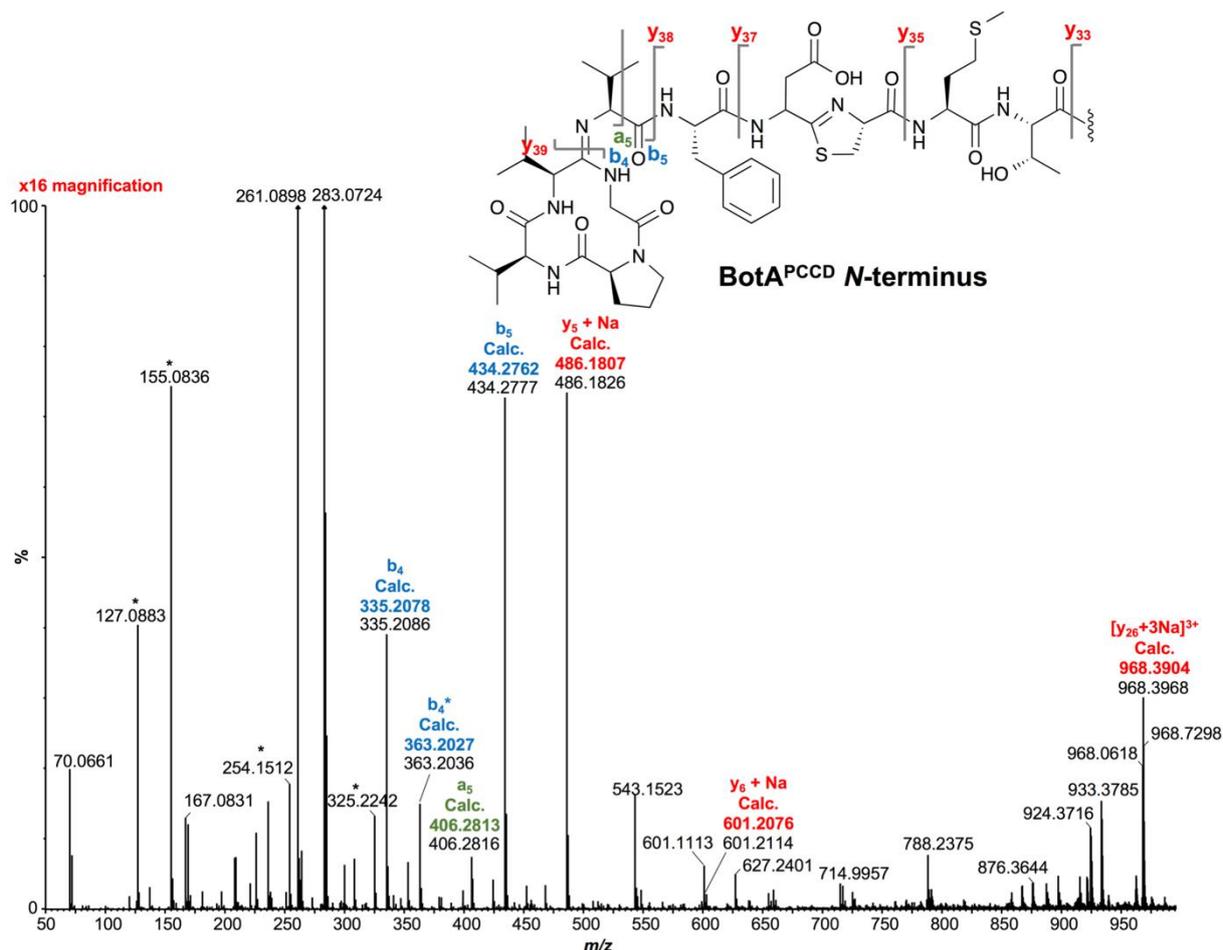
BotA^{PCCD}

Figure 2-S6 | a) MS data showing BotA[PC] before and after incubation with PurCD and ATP/MgCl$_2$. The product, BotA[PCCD], shows the characteristic loss of 18 Da, indicating loss of one water molecule **b)** MS^2 analysis of BotA[PC] treated with YcaO-domain enzyme PurCD (BotA[PCCD]). The fragments visible in the MS^2 are indicated on the peptide sequence and these are detailed in Table 2-S5. The deconvoluted mass spectrum is displayed and is annotated with peptide fragments where possible. It was not possible to fully separate BotA[PC] and BotA[PCCD] for MS^2 analysis, and the mass window for selecting peaks for fragmentation could not be limited to exclude fragmentation of BotA[PC]. Consequently, some fragments relate to BotA[PC] and these are highlighted by stars in the mass spectrum. The presence of sodium salts meant that most fragments contained at least one sodium atom. Fragment b$_4$* relates to a fragment that is characteristically associated with the bottromycin macrocycle, and analogous fragments had previously been observed for all bottromycins and related macrocyclic peptides. Starred peaks indicate fragments that are likely to derive from the BotA[PC] present in this sample.
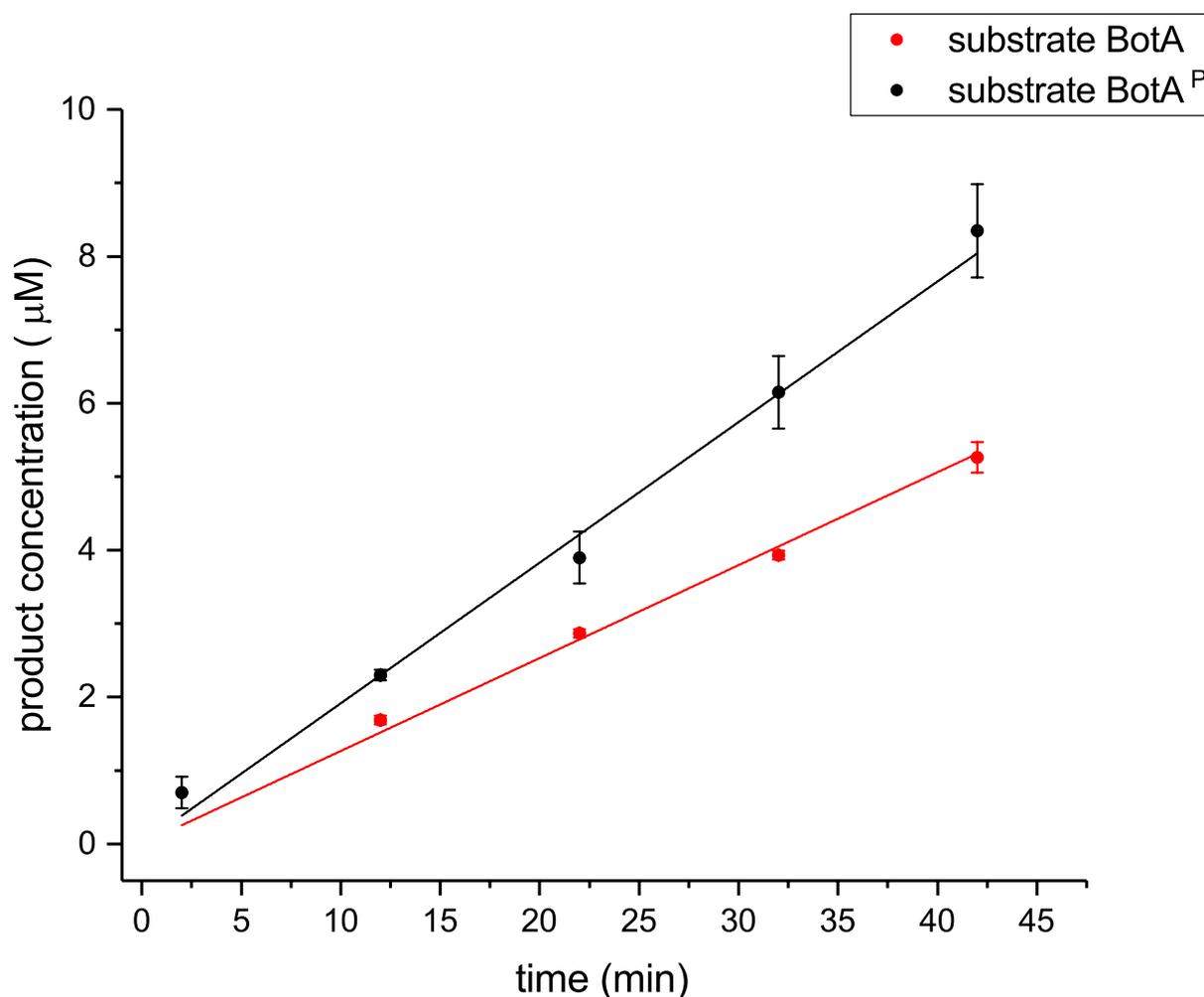
Figure 2-S7 | Analysis of the turnover rate of IpoC using BotA and BotA[P] as substrates. Using 2 μM enzyme and 100 μM substrate led to a linear increase in product concentration during the first 45 min of the reaction. This linear phase was then used to estimate the turnover rate. Data was fitted using y(BotA[P]) = 0.19x and y(BotA) = 0.12x. The reactions were performed at pH 7.4 and contained 5 mM ATP/Mg^{2+}. Data are given as mean ± SD.
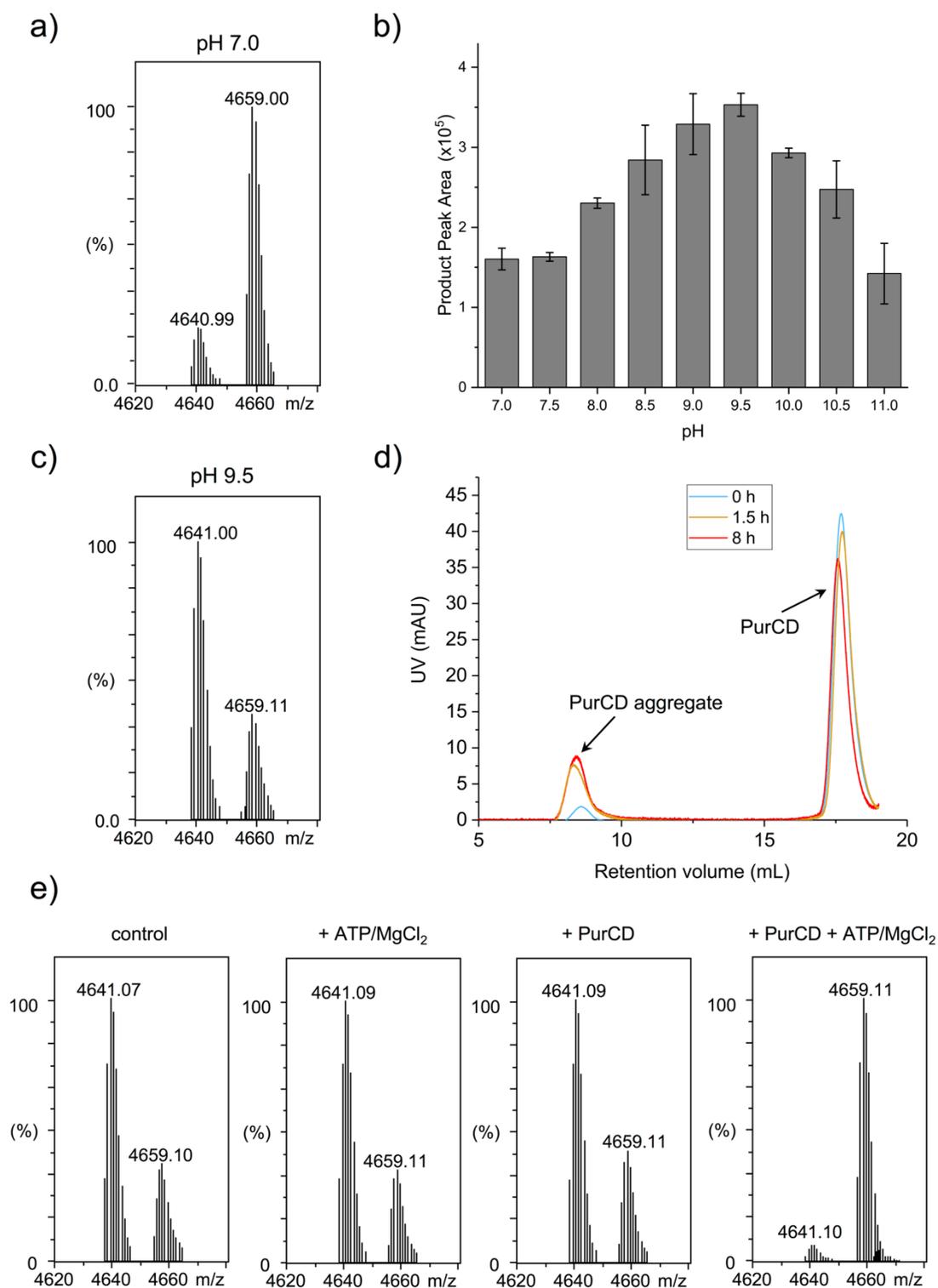
Figure 2-S8 | a) Turnover of PurCD (10 μM) when incubated with BotA^{PC} (50 μM) and 10 mM ATP/MgCl$_2$ for 90 min at pH 7.0. **b)** pH dependence of PurCD activity with a maximum at pH 9.5. Data are given as mean ± SD. **c)** Repeat of the experiment displayed in a) at pH 9.5. **d)** PurCD is stable at RT in macrocyclisation buffer supplemented with 10 mM ATP/MgCl$_2$. **e)** BotA^{PCCD} was purified after 90 min from a reaction of BotA^{PC} with PurCD (and ATP/MgCl$_2$) and then incubated for 12 h at RT in macrocyclisation buffer with 1. no further additive (control) 2. 10 mM ATP/MgCl$_2$ 3. 10 μM PurCD 4. 10 μM PurCD and 10 mM ATP/MgCl$_2$. The macroamidine is unstable when left in solution with PurCD and 10 mM ATP/MgCl$_2$ (reaction solution) for an extended period of time.
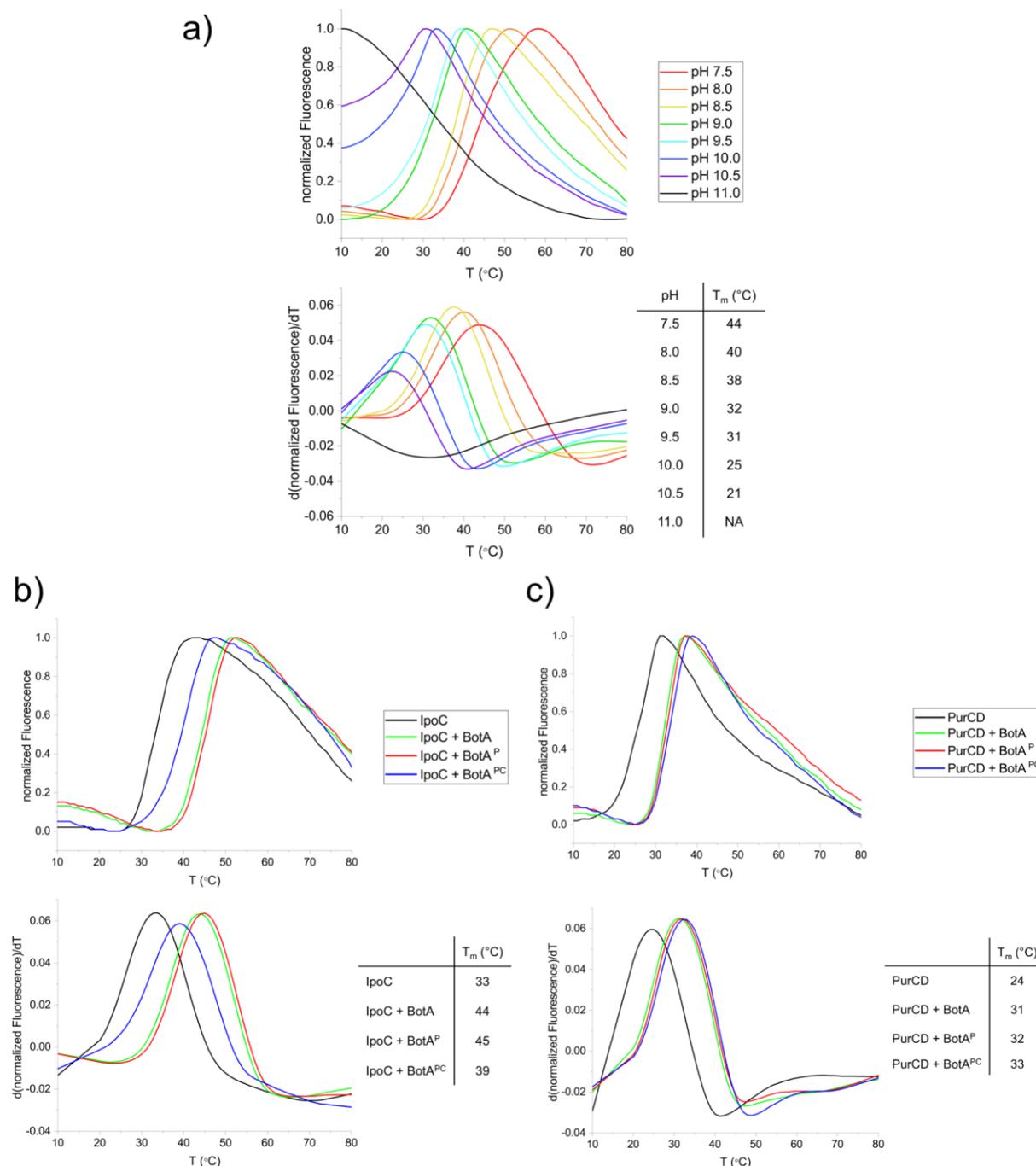
Figure 2-S9 | TSAs of IpoC and PurCD. **a)** pH dependence of PurCD stability. PurCD was incubated with ATP/MgCl$_2$ before the assay to better reflect reaction conditions. The T$_m$ decreases steadily with increasing pH. **b)** Effect of substrates and product on the T$_m$ of IpoC. The product, BotA^{PC}, does not stabilize the enzyme to the same extent as both substrates as expected. **c)** Effect of different modified precursor peptides on the T$_m$ of PurCD. These melting curves were recorded at pH 8.5 without the addition of ATP. These data indicate that ATP has a marked stabilizing effect on the enzyme. Melting points (T$_m$s) were calculated as described in the materials and methods. A complete list of all T$_m$s can be found in Table 2-S6.
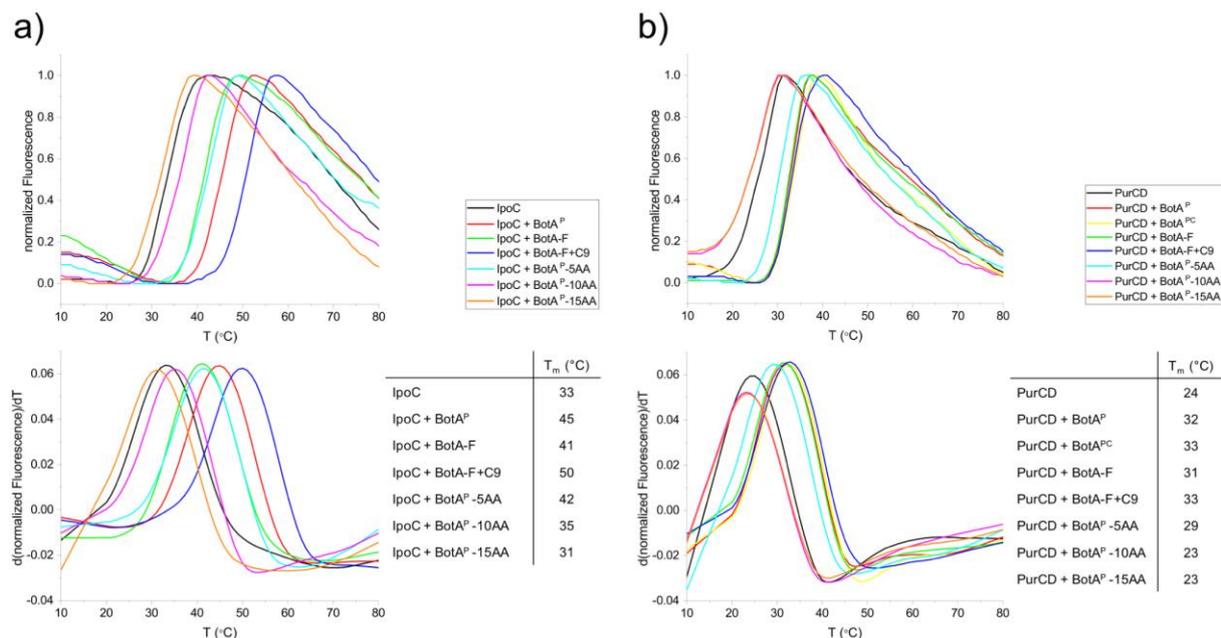
Figure 2-S10 | TSAs of IpoC **(a)** and PurCD **(b)** and their complexes with BotA^P, follower peptide (BotA-F) and BotA^P truncations. Interestingly, for IpoC the addition of the final core peptide residue to the follower peptide, the substrate cysteine, had a dramatic stabilizing effect. This would be unexpected for a standard, processive, RiPPs heterocyclase and may indicate specific substrate interactions that processive YcaOs lack. PurCD stability is not affected to the same extent as IpoC by substrate binding. For both enzymes, the stabilizing interactions appear to be mostly contributed by residues between -5 and -10. Melting curves were recorded at pH 8.5 and melting points (T_ms) were calculated as described in the materials and methods. A complete list of all T_ms can be found in Table 2-S6.
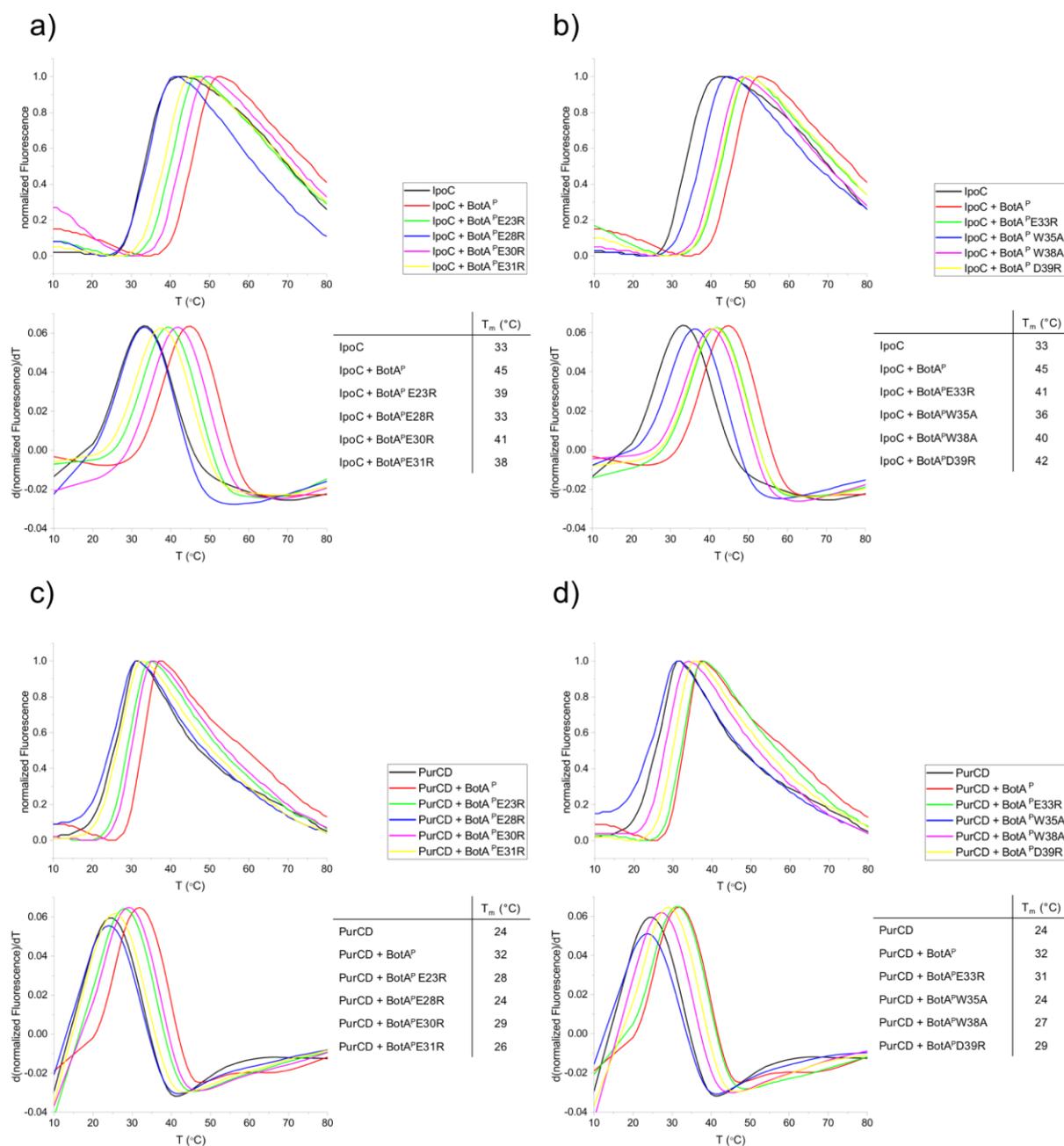
Figure 2-S11 | TSAs of IpoC (**a** and **b**) and PurCD (**c** and **d**) with all follower peptide point mutants reported in this study. While all point mutations lead to a reduced T_m, only E28R failed to stabilize the enzymes at all. Since this residue lies outside the region identified in Figure 2-S9, it is probable that this residue is normally not involved in substrate binding. Instead, the radical mutation may prevent substrate binding by preventing an off-site steric hinderance. Melting curves were recorded at pH 8.5 and melting points (T_ms) were calculated as described in the materials and methods. A complete list of all T_ms can be found in Table 2-S6.
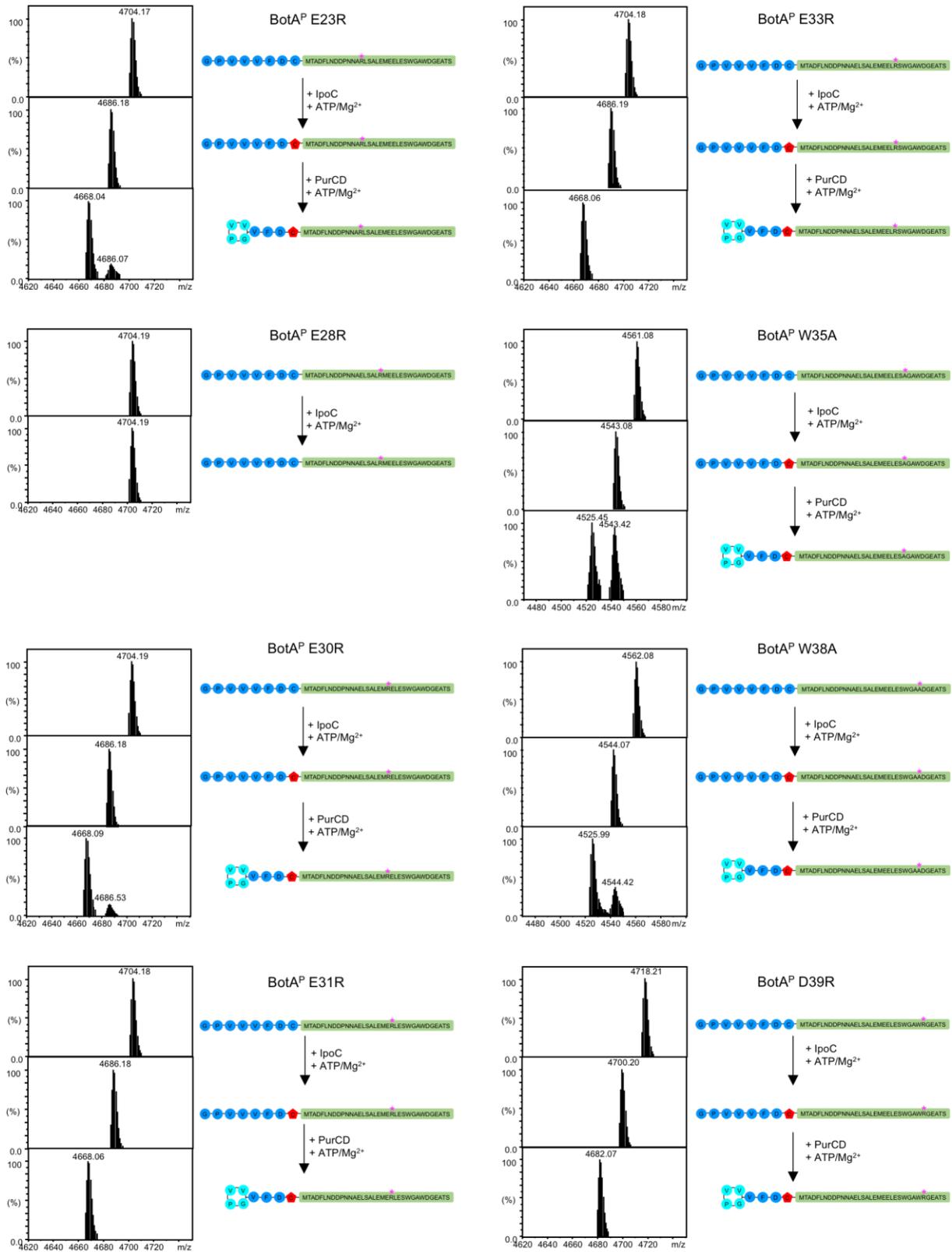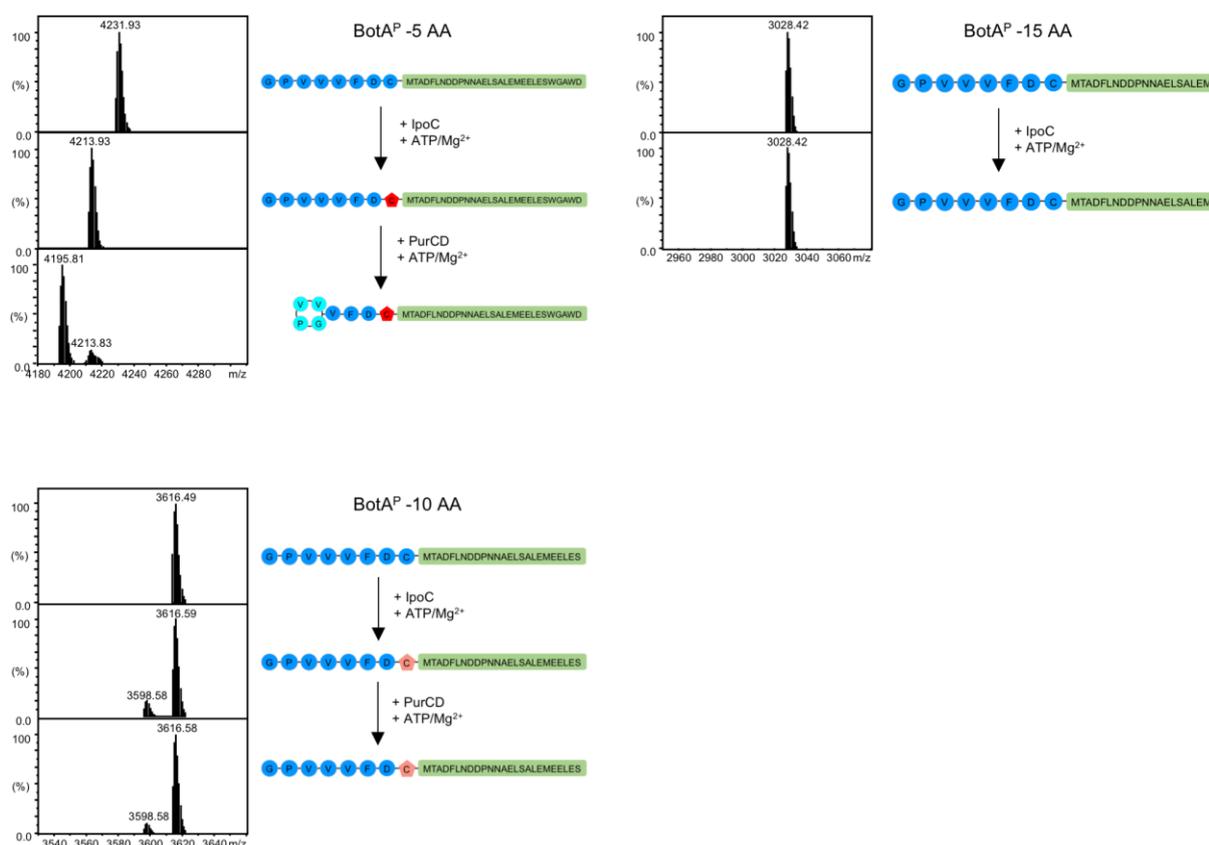
Figure 2-S12 | MS data showing the reactions of IpoC and PurCD with all BotA^{P}s that carry point mutations in their follower peptides. Heterocyclization and macroamidine formation result in successive losses of 18 Da each. Overall, IpoC appears to be much more tolerant to point mutations than PurCD, which is most affected by the W35A mutation. These data, when combined with the TSA data, may indicate that both enzymes have different modes of substrate binding. A summary of the results can be found in Table 2-S7 and Table 2-S8.
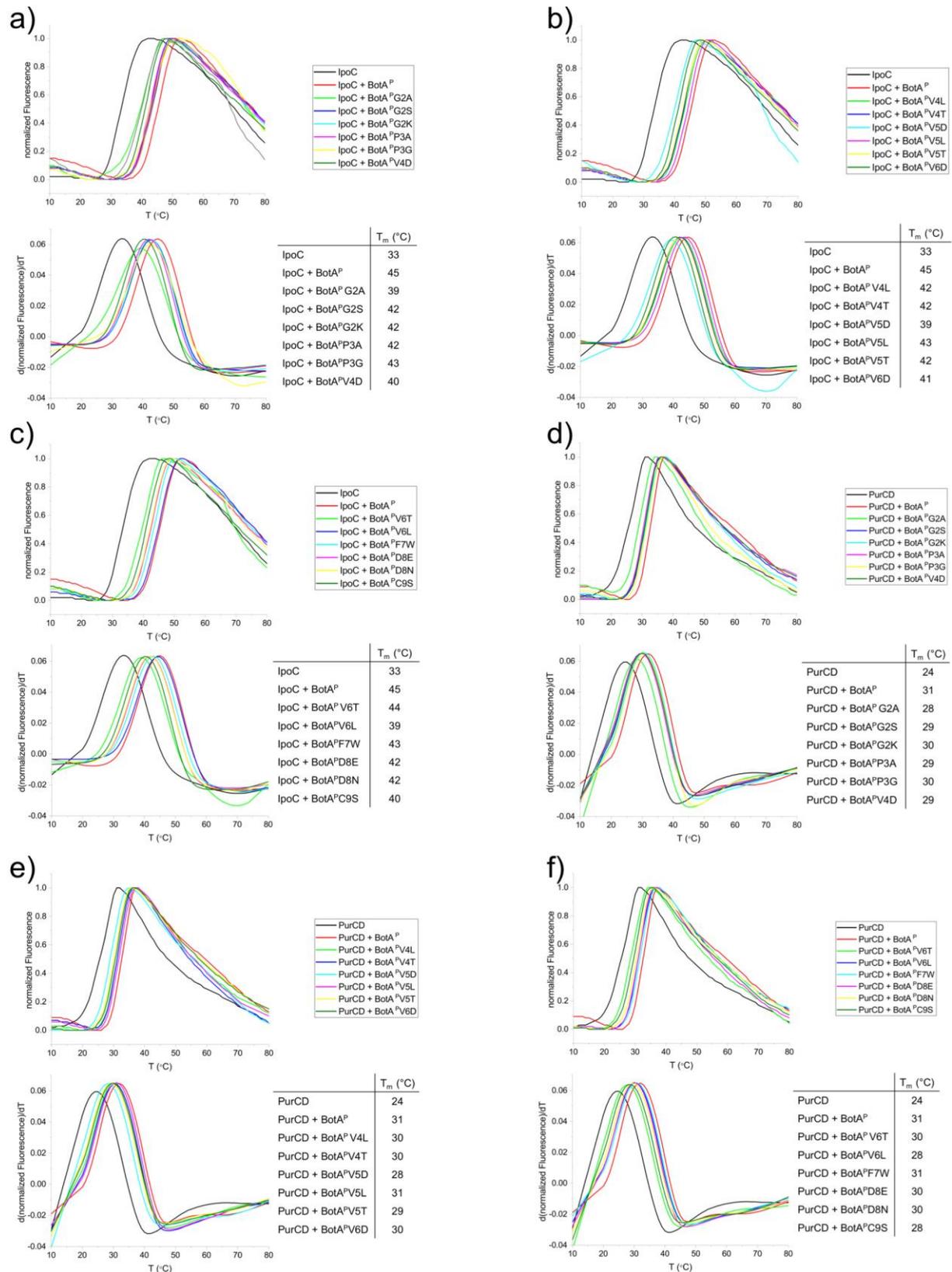
Figure 2-S13 | TSAs of all core peptide mutants with IpoC and PurCD. Variations of the core peptide caused variations of T_ms in a very narrow range only when compared to wt BotA. This indicates that while not all mutant BotAs are substrates, they are all still able to bind to the enzymes, echoing the spatial separation between substrate recognition and catalysis paradigm. Melting curves were recorded at pH 8.5 and melting points (T_ms) were calculated as described in the materials and methods. A complete list of all T_ms can be found in Table 2-S6.
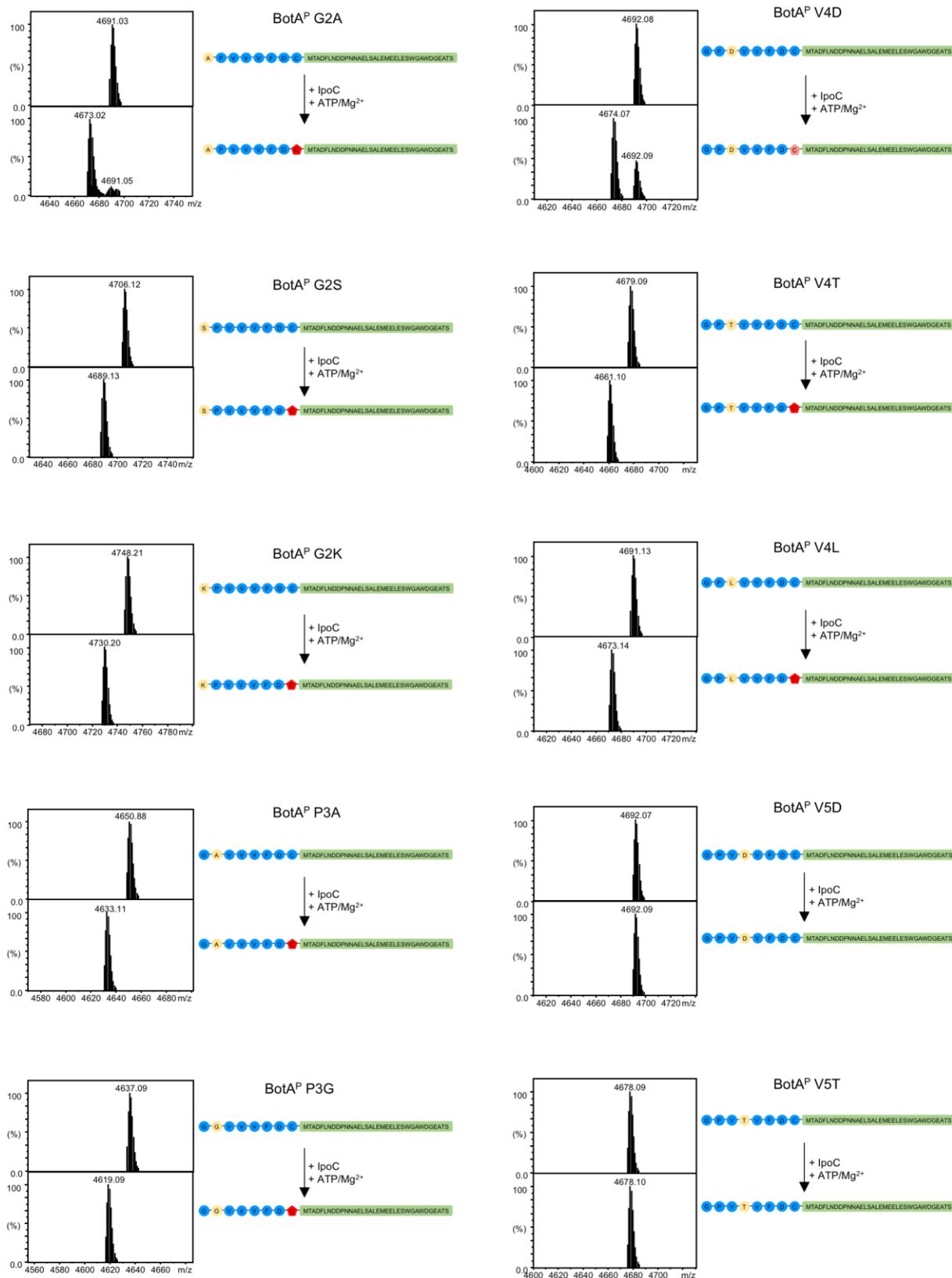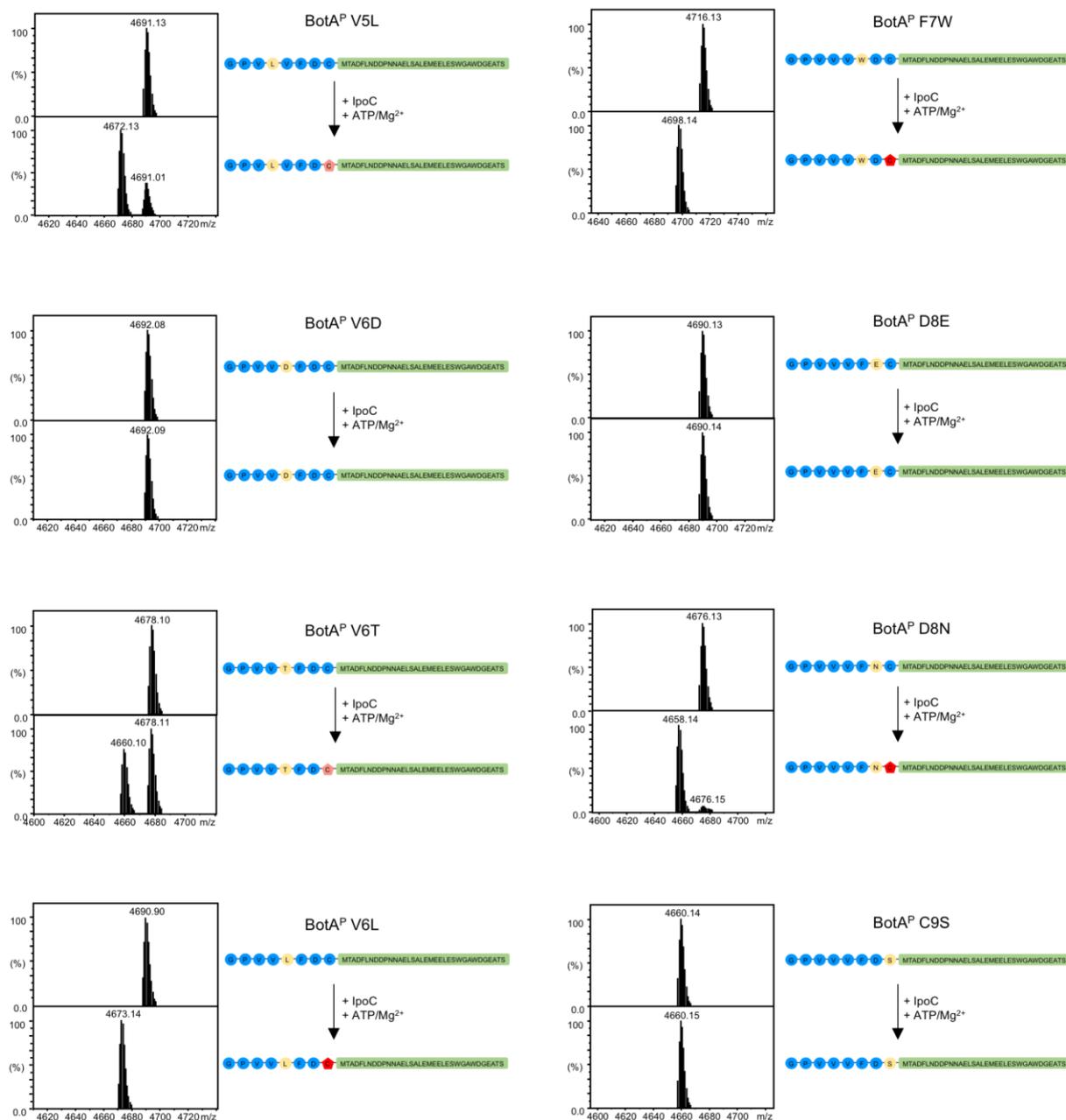
Figure 2-S14 | MS data showing the reaction of IpoC with all BotA[P]s with mutations in the core peptide used in this study. Thiazoline formation results in a loss of 18 Da. A summary of the results can be found in Table 2-S7.
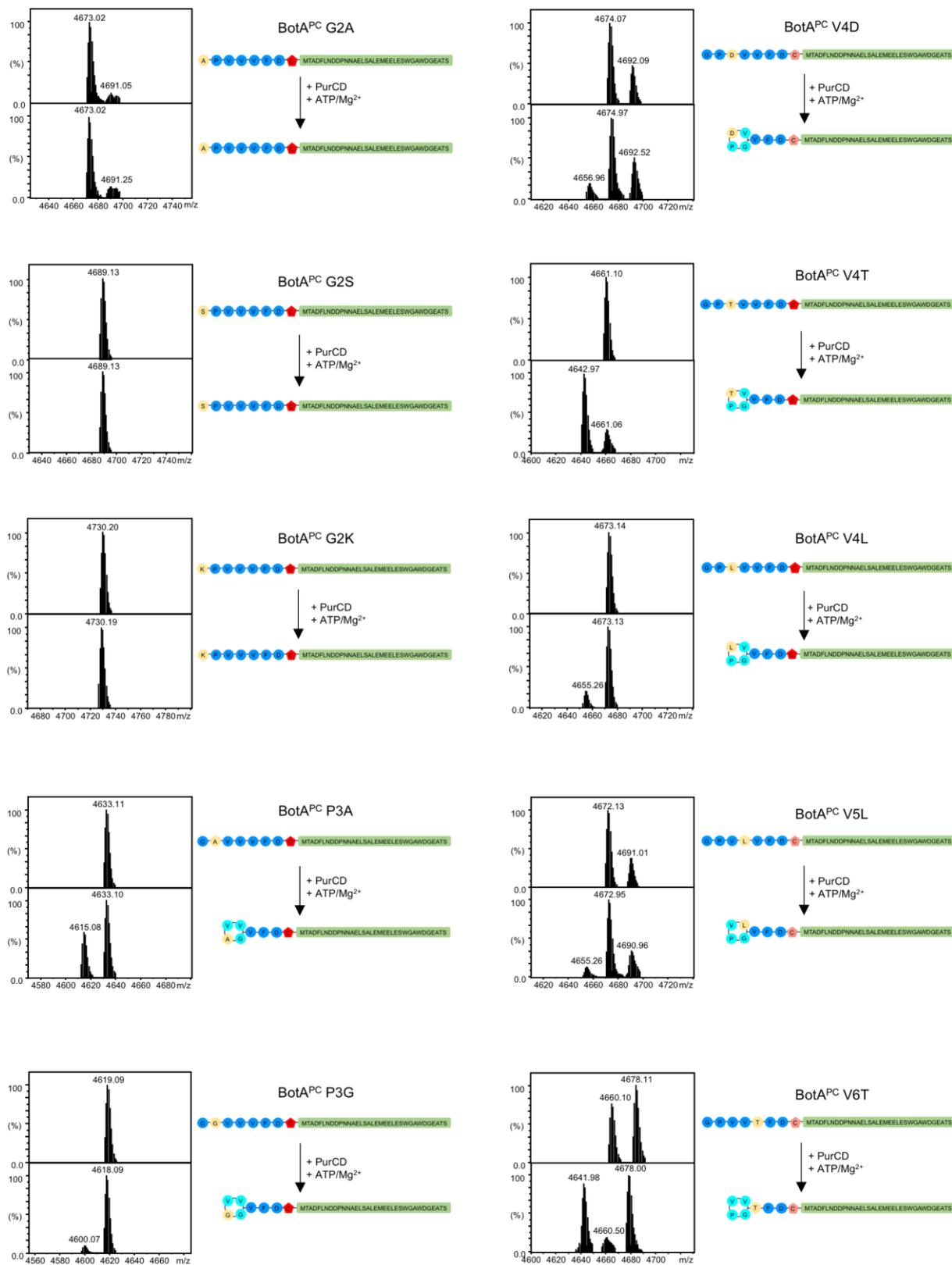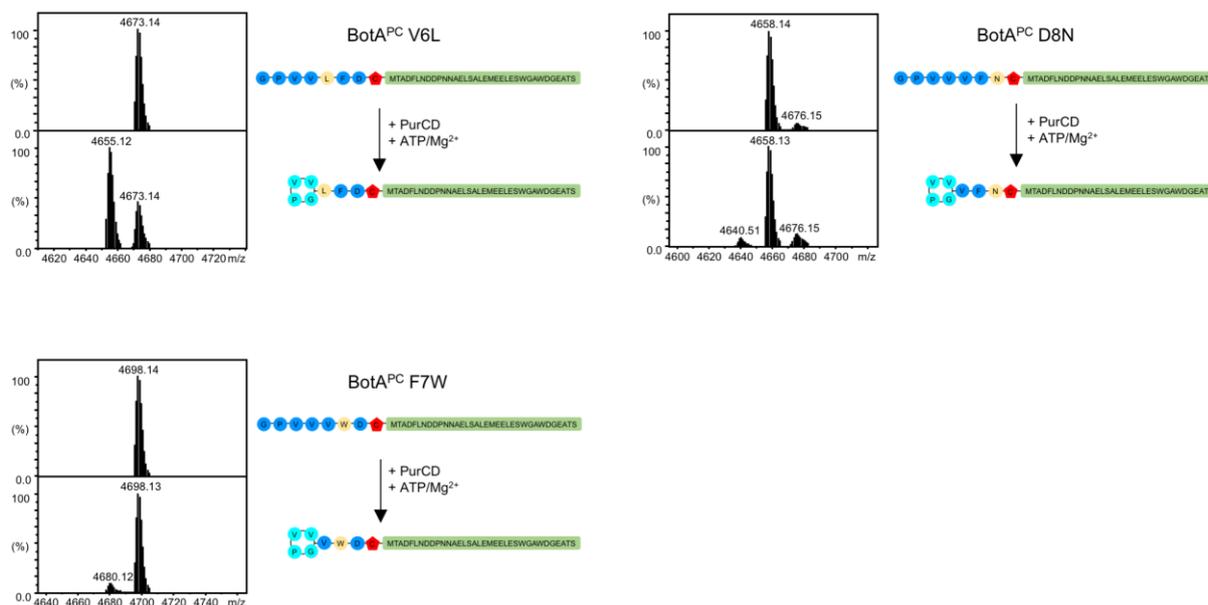
Figure 2-S15 | MS data showing the reaction of PurCD with all BotA^PCs with mutations in the core peptide used in this study. Only those mutant BotAs that could be processed with IpoC (contained a heterocycle) could be tested as substrates, since the heterocycle is required for PurCD to catalyze formation of the macroamidine (-18 Da). A summary of the results can be found in Table 2-S8.

```
4q85    ------------------MTQTF-IPG--KDAALEDSIARFQQKLSDLGFQIEEASWLN-          38
4v1t    QHRGFEPLKLESRPKQ-----FTSDGGHRGTTPEQTVQKYQHLISPVTGVVT--ELVRI          372
CD      ---------MTRPAENDLLTAALLDGTLKEDVPPERETPLGKALAAVG------DWLA-          43
C       ----------------MLEATATECALRE-----------------------------          13
                                                         :

4q85    -----PV--------PNVWSVHIRDKECALCF---TNGKGATKKAALASALGEYFERLST          82
4v1t    TDPANPLVHTYRAGHSFGSATSLRGLRNTLKHKSSGKGK--TDSQSKASGLCEAVERYSG          430
CD      AERLTADIRKYGIDSAPTYEVQLKDADGALCSRSSGKGLGHRSLAS---ALFGSAEHYHL          100
C       -----VVYRSYPSERTVTVRCTVRPAEGSA--RADGYGTAATKAVARLKALSTAVERL--          64
                      ::    :        *    .  :    .* * * *:

4q85    NYFFADFWLGETIANGPFVHYPNEKWFPLTEN--------------------DDV----          117
4v1t    IFQG-----DEP------------------RKRATLAELGDLAIHPEQ-CLCFSDGQYAN          446
CD      DWRR-----DPRASDTVFLSGREIARQPTARRSALLQRLGGIAADAPVLTREYRP----V          151
C       ------------VACTPF---ASVVRPPTTSR-------GSGAVPPFPASGVRTVPDGC          101
                                                     .

4q85    -------PEGLLDDRLRAFYDPENELTGSMLIDLQSGNEDRGICGLPFTRQSDNQTVYIP          170
4v1t    RETLNEQAT-VAHDWIPQRFDASQAIEWT----------------PVW-SLTEQTHKYLP          508
CD      AEALAEPADGAAPHRQPVFLRDGGYRNWP----------------HPA-DDGSFRT----          190
C       ASRVYRPLTGGGPRRVPLYW--------S---------------SPW-TAGEELRAAV-          135
                                                            .

4q85    -----------MNIIGNLYVSNGMSAGNTRNEARVQGLSEVFERYVKNRIIAESI-----          214
4v1t    TALCYYHYPLPPEHRFARGDSNGNAAGNTLEEAILQGFMELVERDGVALWWYNRL-----          563
CD      --------------LWHYTSSAGYAAGATLHEALVHAVNELVERDAWSYQLARSYFGLAD          236
C       ----------LTAPEARLSSTVGWAVAPTPEAALHGALLELTELLNHGVFLHRSLA----          181
                        : * :.. * . *    .. *: *        .

4q85    SLPEIPADVLARYPAVVEA--IETLEAEGFPIFAYDGSLGGQYPVICVVLFNP--ANGTC          270
4v1t    ---RRPAVDLGSFNEPYFVQLQQFYRENDRDLWVLDLTADLGIPAFAGVSNRKTGSSERL          620
CD      GGPALRVVDPATLPAELRQLARRVEAVRDAPVLVVDVTCDTDVPAYVVCDAESRE--DVR          294
C       -GPRRPAA-------------------GDETLVLPLGGPVRTPTVLAVAYGRGR--RMP          218
               .                 .   .            *.

4q85    FASFGAHPDFGVALERTVTELLQGRGLKDLDVFTPPTFDDEEVAEHTNLETHFIDSSGLI          330
4v1t    ILGFGAHLDPTIAILRAVTEVNQIGLELDKVPDENLKSD----------ATDWLITEKLA          670
CD      LIGSGASPVRTYALQRALLEYLQVRTMFEHGPVD---AD----------TEARQIGTALA          341
C       ATGLGCGATRAEATDRALLELAQAETMWRSNPTA---VP----------AERFF-LRRFE          264
            . *.        *  *:: *  *                                :

4q85    SWDLFKQDADYPFVDWNFSGT-----TEEEFATLMAIFNKEDKEVYI-------------          372
4v1t    DHPYLLPDTTQPLKTAQDYPKRWSDDIYTDVMTCVNIAQQAGLETLV-------------          717
CD      RYPRHLAAARFDIHELPH-----E---------QVAFASDGGLSAAASPEDLLRHLVDRL          387
C       RWPLLGRCATLDFDLSGH-GTPYDD--ERSVASPLEELEAGGISVWADSGA---------          312
                  :   :                    :   .  . ..

4q85    ----------ADYEHLGVYACRIIVPGMSDIYPAEDLWLANNSMGSHLRETILSLPGSEW          422
4v1t    ----IDQTR----PDIGLNVVKVTVPGMRHF----------------------------          740
CD      RAVRVDVHYRVLTRPGAVTVVDVVAPGLEMLDKA-------------------------          421
C       ----VDISGPD-TPRTRLCFAHVVSDPQPLLGLV-------------------------          341
                    :      :           :

4q85    EKEDYLNLIEQLDEEGFDDFTRVRELLGLATGSDNGWYTLRIGELKAMLALAGGDLEQAL          482
4v1t    -----------------------------------WSRFGEGRLYDVPV-------KL          756
CD      -------------------------------RAGYPVLPTGRL---------------          433
C       -------------------------------RAGIPVFDTGEV--------------          353
                                             :  *.:

4q85    VWTEWTMEFNSSVFSPERANYYRCLQTLLLLAQEEDRQPLQYLNAFV-RMYGADAVEAAS          541
4v1t    GWLDEPLTE--------------------AQMNPTPMPF------------------          775
CD      ---AERLRS----------------------RGSER-A--------------------          445
C       ---RRTLDP--------------------PRRDR-PAHHARSADHRAAGRG------          380
           :                          :

4q85    AAMSGEAAFYGLQPVDSDLHAFAAHQSLLKAYEKLQRAKAAFWAK        586
4v1t    -----------------------------------------        775
CD      -----------------------------------------        445
C       --RQGRRTA--------------------------------        387
```
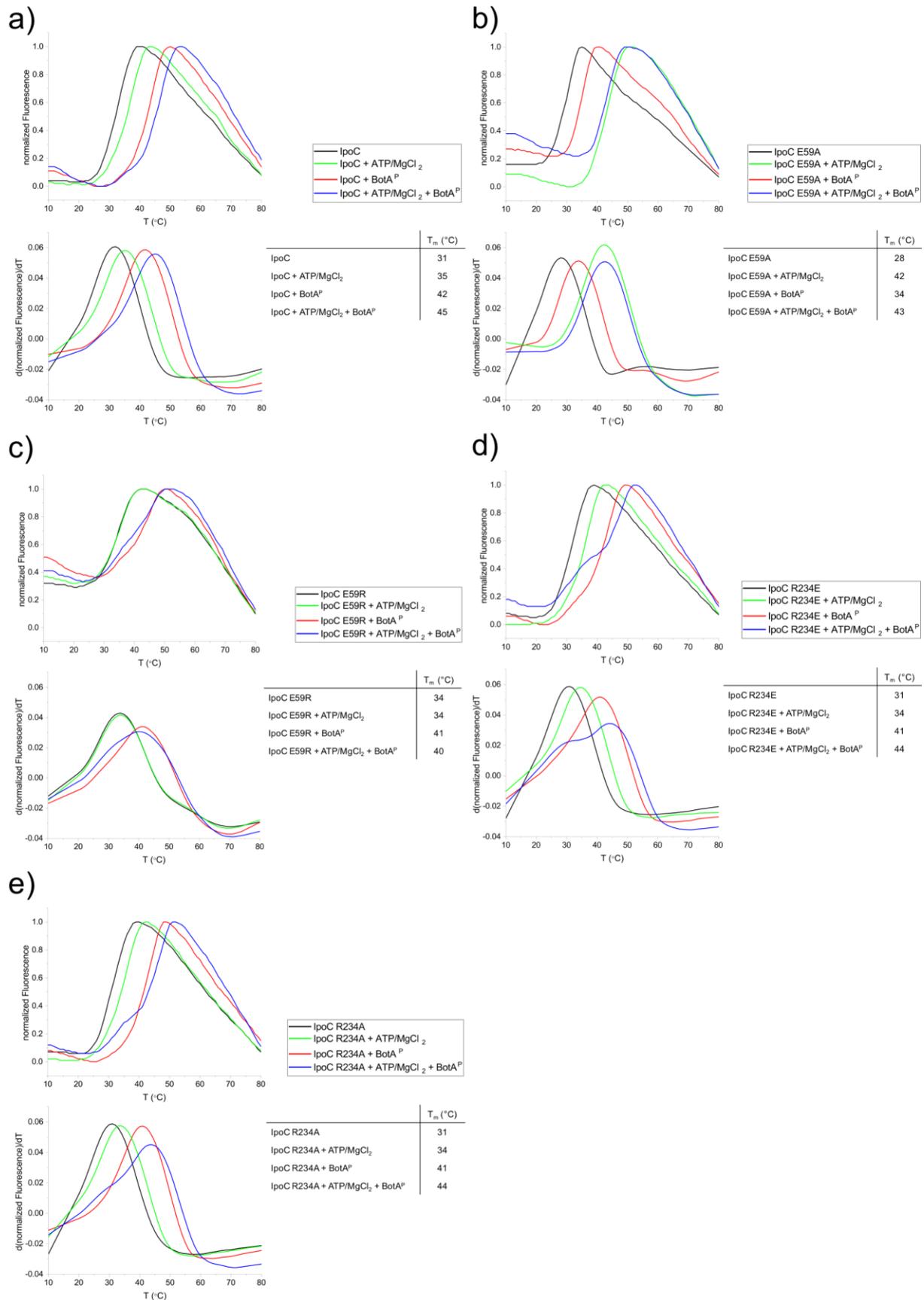
Figure 2-S16 | Sequence alignment of the only two YcaO enzymes for which structures in complex with nucleotide cofactor have been reported to date (PDB IDs 4v1t and 4q85) with IpoC and PurCD. Residues directly involved ATP or Mg^{2+} binding in the two protein structures are highlighted in yellow. The two residues chosen for mutation in this study are among those highlighted in yellow and have been highlighted in red for clarity. Non-conserved residues are marked in grey.

a)

	T_m (°C)
IpoC	31
IpoC + ATP/MgCl$_2$	35
IpoC + BotA^P	42
IpoC + ATP/MgCl$_2$ + BotA^P	45

b)

	T_m (°C)
IpoC E59A	28
IpoC E59A + ATP/MgCl$_2$	42
IpoC E59A + BotA^P	34
IpoC E59A + ATP/MgCl$_2$ + BotA^P	43

c)

	T_m (°C)
IpoC E59R	34
IpoC E59R + ATP/MgCl$_2$	34
IpoC E59R + BotA^P	41
IpoC E59R + ATP/MgCl$_2$ + BotA^P	40

d)

	T_m (°C)
IpoC R234E	31
IpoC R234E + ATP/MgCl$_2$	34
IpoC R234E + BotA^P	41
IpoC R234E + ATP/MgCl$_2$ + BotA^P	44

e)

	T_m (°C)
IpoC R234A	31
IpoC R234A + ATP/MgCl$_2$	34
IpoC R234A + BotA^P	41
IpoC R234A + ATP/MgCl$_2$ + BotA^P	44

Figure 2-S17 | TSAs of all IpoC constructs used in this study using 1. apo enzyme, or enzyme in complex with 2. ATP/MgCl$_2$, 3. BotA^P, 4. + ATP/MgCl$_2$ and BotA^P. Melting points (T_mS) were calculated as described in the materials and methods. A complete list of all T_ms can be found in Table 2-S6.
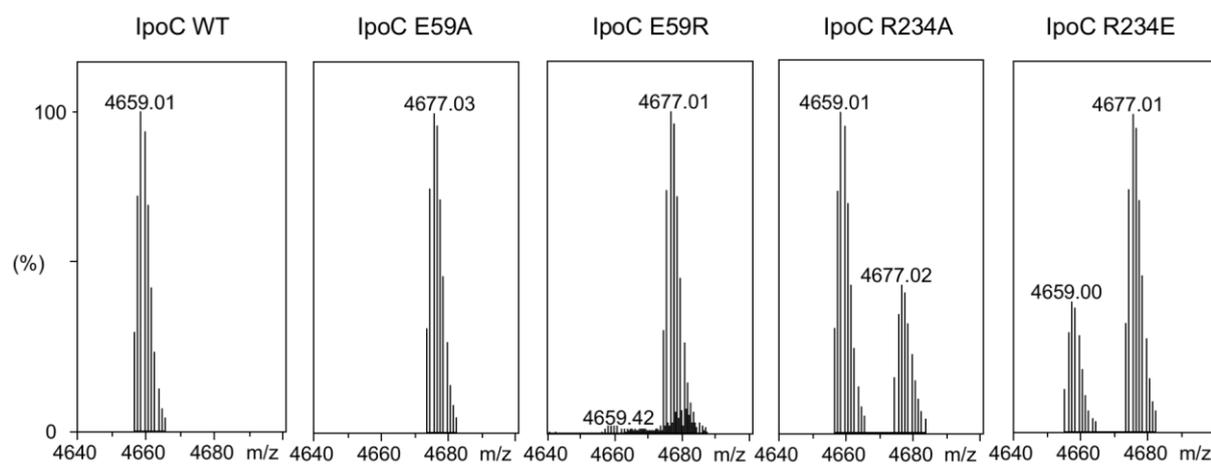
Figure 2-S18 | MS data comparing the activity of IpoC with that of IpoC mutants. Enzymatic activity of the E59 mutant was almost completely abolished (trace amounts of product could be found, data not shown), while R234 mutants retained significant activity.
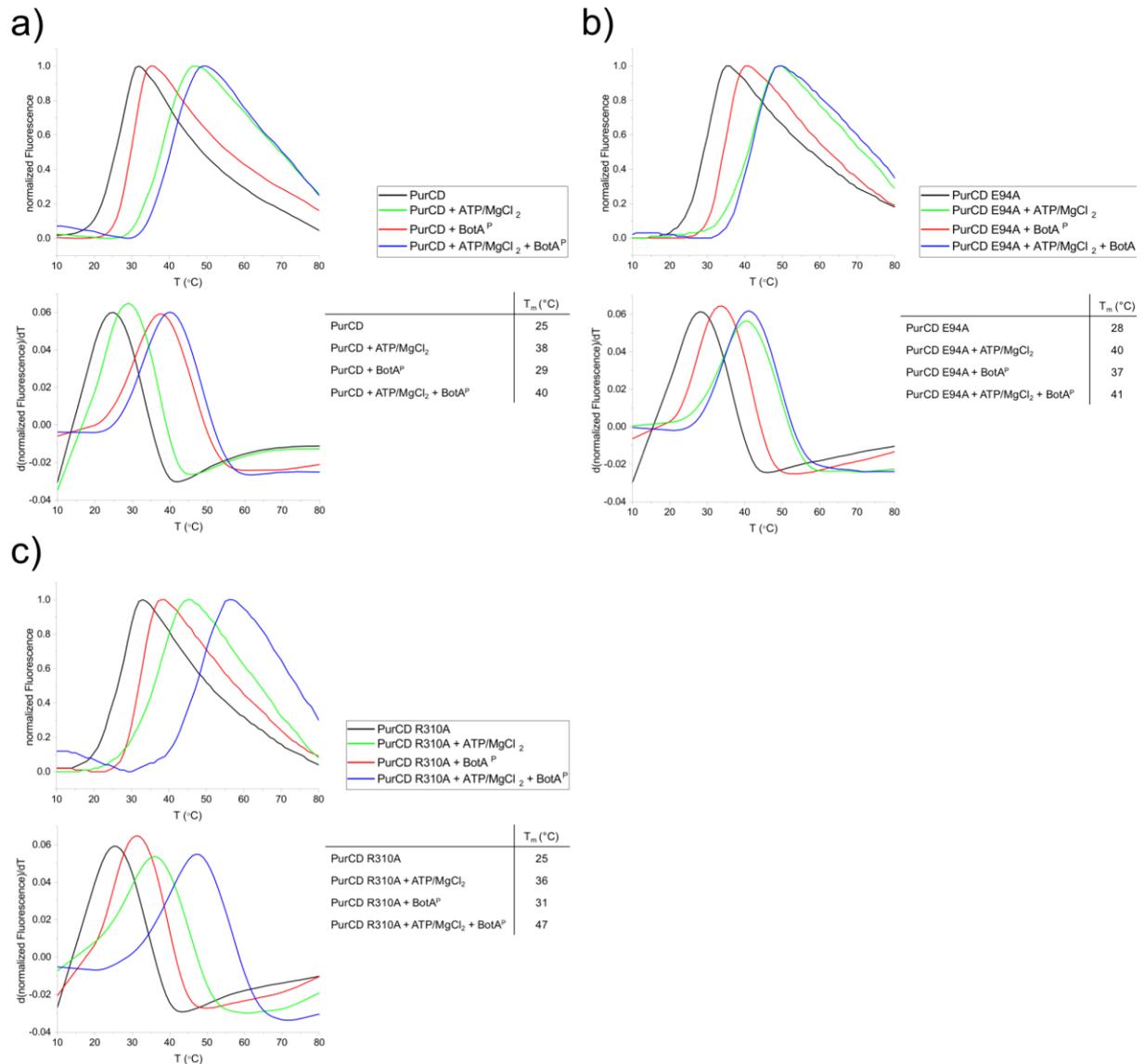
Figure 2-S19 | TSAs of all PurCD constructs used in this study using 1. apo enzyme, or enzyme in complex with 2. ATP/MgCl$_2$, 3. BotA^P, 4. + ATP/MgCl$_2$ and BotA^P. Interestingly, the E94A mutation appears to stabilize the enzyme when it is incubated with BotA^P, but the effect was not additive, as the co-incubation with ATP/MgCl$_2$ and BotA^P corresponds to the wt enzyme. PurCD R310A on the other hand is stabilized more than wt when coincubated with nucleotide and BotA^P. Answering how these effects are exerted will have to await structure determination. Melting points (T$_m$s) were calculated as described in the materials and methods. A complete list of all T$_m$s can be found in Table 2-S6.
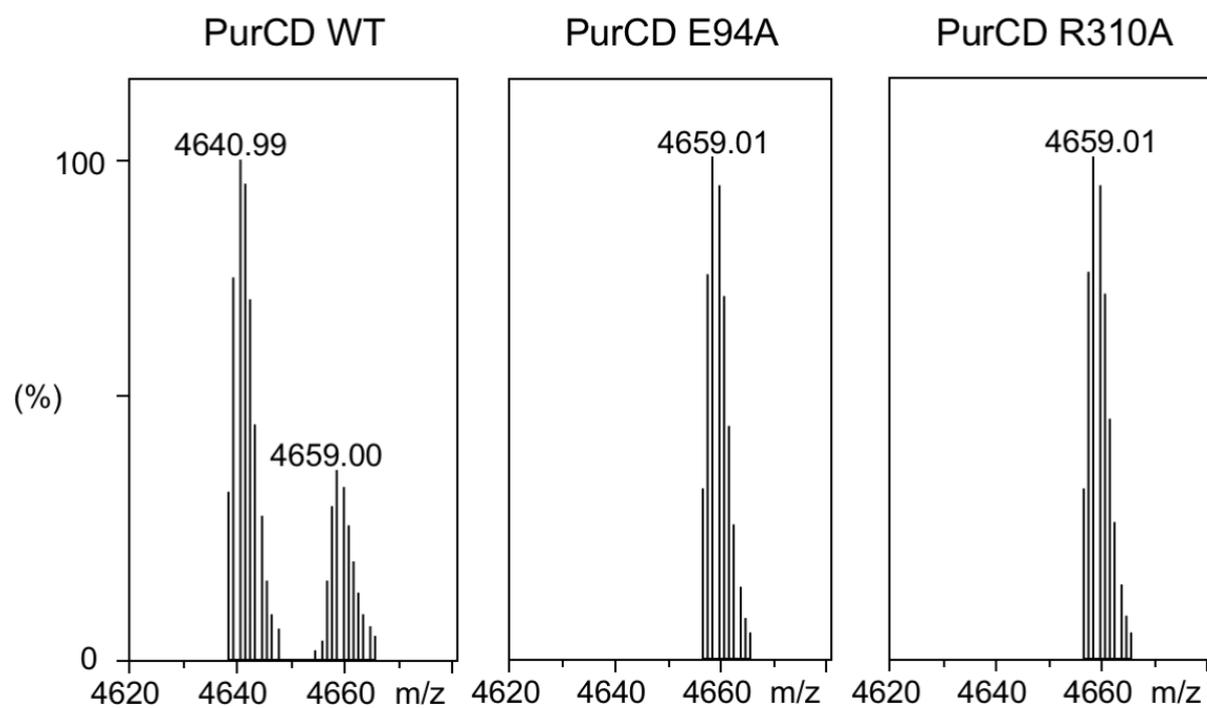
Figure 2-S20 | MS data comparing the activity of PurCD with that of PurCD mutants. Only two mutants could be expressed and purified. Both were completely inactive.

Table 2-S1 | Peptide fragments observed in MS^2 analysis of BotA.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
M	1	-	-	44	4894.1	4894.0 (+4Na)
G	2	-	-	43	4762.9	4762.9 (+4Na)
P	3	-	-	42	4706.0	4705.9 (+4Na)
V	4	-	-	41	-	-
V	5	-	-	40	4509.9	4509.8 (+4Na)
V	6	-	-	39	-	-
F	7	-	-	38	4311.8	4311.7 (+4Na)
D	8	867.4105	867.4045 (+Na)	37	4142.6	4142.6 (+3Na)
C	9	-	-	36	4027.6	4027.6 (+3Na)
M	10	1101.4	1101.5 (+Na)	35	3924.6	3924.6 (+3Na)
T	11	-	-	34	3793.6	3793.5 (+3Na)
A	12	-	-	33	-	-
D	13	1388.6	1388.6 (+Na)	32	-	-
F	14	-	-	31	3506.5	3506.4 (+3Na)
L	15	-	-	30	-	-
N	16	-	-	29	-	-
D	17	1877.8, 1899.8	1877.8 (+Na), 1899.8 (+2Na)	28	-	-
D	18	1992.8, 2014.8	1992.8 (+Na), 2014.8 (+2Na)	27	3017.3	3017.2 (+3Na)
P	19	-	-	26	2880.2, 2902.2	2880.2 (+2Na), 2902.2 (+3Na)
N	20	-	-	25	-	-
N	21	-	-	24	2669.1	2669.1 (+2Na)
A	22	-	-	23	2555.1	2555.1 (+2Na)
E	23	2540.0	2540.0 (+2Na)	22	-	-
L	24	-	-	21	-	-
S	25	-	-	20	-	-
A	26	-	-	19	-	-
L	27	-	-	18	-	-
E	28	3075.3	3075.3 (+3Na)	17	-	-
M	29	-	-	16	-	-
E	30	3335.4	3335.4 (+3Na)	15	-	-
E	31	3464.4	3464.4 (+3Na)	14	1559.6, 1581.6	1559.6 (+Na)
L	32	-	-	13	1430.6	1430.6 (+Na)
E	33	3706.6	3706.5 (+3Na)	12	-	1317.5 (+Na)
S	34	3793.6	3793.6 (+3Na)	11	1188.5	1188.5 (+Na)
W	35	-	-	10	1101.4	1101.4 (+Na)
G	36	4035.7	4036.7 (+3Na)	9	-	-
A	37	-	-	8	858.3565	858.3240 (+Na)
W	38	4293.8	4293.8 (+3Na)	7	-	-
D	39	4408.9	4408.8 (+3Na)	6	-	-
G	40	-	-	5	486.1849	486.1807 (+Na)
E	41	4616.9	4616.9 (+4Na)	4	-	-
A	42	4688.0	4687.9 (+4Na)	3	300.1185	300.1166 (+Na)
T	43	-	-	2	-	-
S	44	4876.1	4876.0 (+4Na)	1	-	-

Table 2-S2 | Peptide fragments observed in MS^2 analysis of BotA^C.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
M	1	-	-	44	4876.0	4876.0 (+4Na)
G	2	-	-	43	-	-
P	3	286.1208	286.122	42	4687.9	4687.9 (+4Na)
V	4	385.1889	385.1905	41	-	-
V	5	484.2556	484.2589	40	4469.8^a	4469.8 (+3Na)
V	6	583.3237	583.3273	39	4370.7	4370.7 (+3Na)
F	7	-	-	38	4271.7	4271.7 (+3Na)
D^b	8	-	-	37	-	-
C^b	9	-	-	36	-	-
M	10	1061.5	1061.5	35	-	-
T	11	-	-	34	-	-
A	12	1277.5	1277.5 (+2Na)	33	3692.5	3692.5 (+3Na)
D	13	1370.6	1370.6 (+Na)	32	-	-
F	14	-	-	31	3506.4	3506.4 (+3Na)
L	15	-	-	30	-	-
N	16	-	-	29	-	-
D	17	1859.8, 1881.8	1859.8 (+Na), 1881.8 (+2Na)	28	3110.2	3110.2 (+2Na)
D	18	1974.8, 1996.8	1974.8 (+Na), 1996.8 (+2Na)	27	2995.2	2995.2 (+2Na)
P	19	-	-	26	2880.2, 2902.2	2880.2 (+2Na), 2902.2 (+3Na)
N	20	-	-	25	-	-
N	21	-	-	24	2669.1	2669.1 (+2Na)
A	22	-	-	23	2533.0	2533.1 (+Na)
E	23	2522.0	2522.0 (+2Na)	22	-	-
L	24	-	-	21	2333.0	2333.0 (+Na)
S	25	-	-	20	2219.9	2219.9 (+Na)
A	26	2793.1	2793.2 (+2Na)	19	-	-
L	27	-	-	18	-	-
E	28	3035.3	3035.3 (+2Na)	17	-	-
M	29	-	-	16	1819.7	1819.7 (+Na)
E	30	-	-	15	-	-
E	31	3446.4	3446.4 (+3Na)	14	1559.6	1559.6 (+Na)
L	32	-	-	13	1430.6	1430.6 (+Na)
E	33	-	-	12	-	-
S	34	3775.5	3775.6 (+3Na)	11	1188.5	1188.5 (+Na)
W	35	-	-	10	1101.4	1101.4 (+Na)
G	36	-	-	9	915.3486	915.3455 (+Na)
A	37		-	8	858.3315	858.3240 (+Na)
W	38	-	-	7	-	-
D	39	4390.8	4390.8 (+3Na)	6	-	-
G	40	4469.8^a	4469.8 (+4Na)	5	486.1788	486.1807 (+Na)
E	41	4598.8	4598.8 (+4Na)	4	-	-
A	42	-	-	3	300.1156	300.1166 (+Na)
T	43	-	-	2	-	-
S	44	-	-	1	-	-

a. 4469.8 fragment could be either b_{40}+4Na or y_{40}+3Na.
b. Cys9 heterocyclization with Asp8 carbonyl.

Table 2-S3 | Peptide fragments observed in MS^2 analysis of BotA^P.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G	1	-	-	43	4763.0	4762.9 (+4Na)
P	2	155.0824	155.0816	42	-	-
V	3	254.1499	254.1500	41	-	-
V	4	353.2191	353.2184	40	4487.9	4487.8 (+3Na)
V	5	452.2879	452.2868	39	4388.8	4388.7 (+3Na)
F	6	599.3584	599.3552	38	4289.7	4289.7 (+3Na)
D	7	736.3655	736.3641 (+Na)	37	4142.6	4142.6 (+3Na)
C	8	-	-	36	4027.6	4027.6 (+3Na)
M	9	-	-	35	3924.6	3924.6 (+3Na)
T	10	-	-	34	3793.6	3793.5 (+3Na)
A	11	-	-	33	-	-
D	12	1257.5	1257.5 (+Na)	32	-	-
F	13	-	-	31	3506.4	3506.4 (+3Na)
L	14	-	-	30	-	-
N	15	-	-	29	-	-
D	16	1768.7	1768.7 (+2Na)	28	-	-
D	17	1861.8, 1883.8	1861.8 (+Na), 1883.8 (+2Na)	27	2995.2	2995.2 (+2Na)
P	18	-	-	26	2880.2, 2902.2	2880.2 (+2Na), 2902.2 (+3Na)
N	19	-	-	25	-	-
N	20	-	-	24	-	-
A	21	-	-	23	-	-
E	22	-	-	22	-	-
L	23	-	-	21	2355.0	2355.0 (+2Na)
S	24	-	-	20	-	-
A	25	-	-	19	-	-
L	26	2793.2	2793.2 (+2Na)	18	-	-
E	27	-	-	17	-	-
M	28	-	-	16	-	-
E	29	-	-	15	-	-
E	30	3333.4	3333.4 (+3Na)	14	1559.6	1559.6 (+Na)
L	31	-	-	13	-	-
E	32	-	-	12	-	-
S	33	-	-	11	1188.5	1188.5 (+Na)
W	34	-	-	10	1101.4254	1101.4 (+Na)
G	35	3905.6	3905.6 (+3Na)	9	915.3485	915.3455 (+Na)
A	36	-	-	8	858.3399	858.3240 (+Na)
W	37	4184.7	4184.7 (+4Na)	7	-	-
D	38	4277.7, 4299.8	4277.8 (+3Na), 4299.7 (+4Na)	6	601.2091	601.2076 (+Na)
G	39	-	-	5	486.1823	486.1807 (+Na)
E	40	-	-	4	-	-
A	41	-	-	3	300.1170	300.1166 (+Na)
T	42	-	-	2	-	-
S	43	4744.9	4744.9 (+4Na)	1	-	-

Table 2-S4 | Peptide fragments observed in MS^2 analysis of BotA^{PC}.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G	1	-	-	43	4701.0	4701.0 (+2Na)
P	2	155.0823	155.0816	42	-	-
V	3	254.1498	254.1500	41	4546.9	4546.9 (+2Na)
V	4	353.2189	353.2184	40	4447.9	4447.8 (+2Na)
V	5	452.2878	452.2868	39	4348.8	4348.8 (+2Na)
F	6	599.3575	599.3552	38	4249.7	4249.7 (+2Na)
D^a	7	-	-	37	4102.6	4102.6 (+2Na)
C^a	8	-	-	36	-	-
M	9	930.4138	930.4212	35	3902.7	3902.6 (+2Na)
T	10	1031.5	1031.5	34	-	-
A	11	-	-	33	3670.5	3670.5 (+2Na)
D	12	1239.5	1239.5 (+Na)	32	-	-
F	13	-	-	31	3484.5	3484.4 (+2Na)
L	14	-	-	30	3338.4	3337.4 (+2Na)
N	15	-	-	29	3224.3	3224.3 (+2Na)
D	16	1728.7	1728.7 (+Na)	28	3110.3	3110.2 (+2Na)
D	17	1843.8	1843.8 (+Na)	27	2995.2	2995.2 (+2Na)
P	18	-	-	26	2858.2, 2880.2	2858.2 (+Na), 2880.2 (+2Na)
N	19	-	-	25	-	-
N	20	-	-	24	-	-
A	21	-	-	23	2533.1	2533.1 (+Na)
E	22	-	-	22	-	-
L	23	-	-	21	2333.0	2333.0 (+Na)
S	24	-	-	20	2219.9	2219.9 (+Na)
A	25	2662.1	2662.1 (+2Na)	19	2132.9	2132.9 (+Na)
L	26	2775.2	2775.2 (+2Na)	18	2061.9	2061.8 (+Na)
E	27	-	-	17	1948.8	1948.8 (+Na)
M	28	-	-	16	1819.7	1819.7 (+Na)
E	29	3164.3	3164.3 (+2Na)	15	1688.7	1688.7 (+Na)
E	30	3293.4	3293.4 (+2Na)	14	1559.6	1559.6 (+Na)
L	31	-	-	13	1430.6	1430.6 (+Na)
E	32	3535.5	3535.5 (+2Na)	12	1317.5	1317.5 (+Na)
S	33	3622.5	3622.5 (+2Na)	11	1188.5	1188.5 (+Na)
W	34	3786.5	3786.6 (+Na)	10	1101.4	1101.4 (+Na)
G	35	-	-	9	915.3567	915.3455 (+Na)
A	36	3936.7	3936.7 (+2Na)	8	858.3267	858.3240 (+Na)
W	37	-	-	7	765.3071, 787.2891	765.3050, 787.2869 (+Na)
D	38	-	-	6	601.2090	601.2076 (+Na)
G	39	-	-	5	486.1815	486.1807 (+Na)
E	40	-	-	4	-	-
A	41	-	-	3	-	-
T	42	4595.9	4595.9 (+2Na)	2	-	-
S	43	4683.0	4683.0 (+2Na)	1	-	-

a. Cys8 heterocyclization with Asp7 carbonyl.

Table 2-S5 | Peptide fragments observed in MS^2 analysis of $BotA^{PCCD}$.

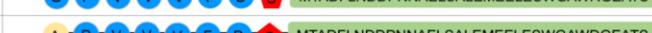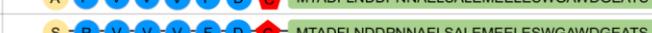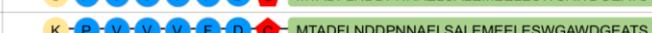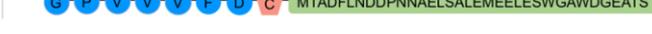| Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G^a	1	-	-	43	4704.9	4704.9 (+3Na)
P	2	-	-	42	-	-
V	3	-	-	41	-	-
V^a	4	335.2086	335.2078	40	-	-
V	5	434.2777	434.2762	39	4370.8	4370.7 (+3Na)
F	6	-	-	38	4271.7	4271.7 (+3Na)
D^b	7	-	-	37	4124.6	4124.6 (+3Na)
C^b	8	-	-	36	-	-
M	9	-	-	35	3924.6	3924.6 (+3Na)
T	10	-	-	34	-	-
A	11	-	-	33	3692.5	3692.5 (+3Na)
D	12	-	-	32	3621.5	3621.4 (+3Na)
F	13	-	-	31	3484.5, 3506.4	3484.4 (+2Na), 3506.4 (+3Na)
L	14	-	-	30	-	-
N	15	-	-	29	-	-
D	16	1732.7	1732.7 (+2Na)	28	3131.2	3132.2 (+3Na)
D	17	1825.7, 1847.7	1825.8 (+Na), 1847.7 (+2Na)	27	2995.2	2995.2 (+2Na)
P	18	-	-	26	2880.2, 2902.2	2880.2 (+2Na), 2902.2 (+3Na)
N	19	-	-	25	-	-
N	20	-	-	24	2669.1	2669.1 (+2Na)
A	21	-	-	23	-	-
E	22	2373.0	2373.0 (+2Na)	22	2484.0	2484.0 (+2Na)
L	23	-	-	21	-	-
S	24	-	-	20	-	-
A	25	-	-	19	2154.9	2154.9 (+2Na)
L	26	-	-	18	-	-
E	27	-	-	17	1970.7	1970.7 (+2Na)
M	28	3039.2	3039.3 (+3Na)	16	1819.7, 1841.7	1819.7 (+Na), 1841.7 (+2Na)
E	29	-	-	15	-	-
E	30	-	-	14	1559.6	1559.6 (+Na)
L	31	-	-	13	-	-
E	32	-	-	12	-	-
S	33	-	-	11	-	-
W	34	-	-	10	1101.4	1101.4 (+Na)
G	35	-	-	9	-	-
A	36	-	-	8	-	-
W	37	-	-	7	-	-
D	38	-	-	6	601.2114	601.2076 (+Na)
G	39	-	-	5	486.1826	486.1807 (+Na)
E	40	-	-	4	-	-
A	41	-	-	3	-	-
T	42	-	-	2	-	-
S	43	-	-	1	-	-

a. Gly1 macrocyclization with Val4 carbonyl.
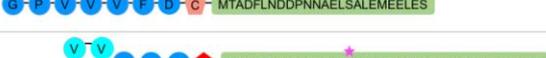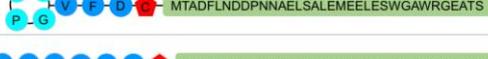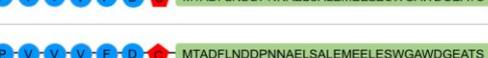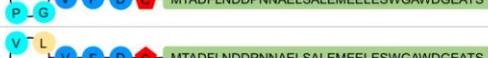b. Cys8 heterocyclization with Asp7 carbonyl.

Table 2-S6 | Summary of all TSA data presented in this manuscript.

Ligand	IpoC			PurCD		
		T_m (°C)	ΔT_m (°C)		T_m (°C)	ΔT_m (°C)
-	wt	33	0	wt	24	0
ATP/Mg^{2+}	wt	35	2	wt	38	14
BotA	wt	44	11	wt	31	7
BotA^P	wt	45	12	wt	32	8
BotA^{PC}	wt	39	6	wt	33	9
ATP/Mg^{2+} + BotA^P	wt	45	12	wt	40	16
BotA^P-5 AA	wt	42	9	wt	29	5
BotA^P-10 AA	wt	35	2	wt	23	-1
BotA^P-15 AA	wt	31	-2	wt	23	-1
BotA^P E23R	wt	39	6	wt	28	4
BotA^P E28R	wt	33	0	wt	24	0
BotA^P E30R	wt	41	8	wt	29	5
BotA^P E31R	wt	38	5	wt	26	2
BotA^P E33R	wt	41	8	wt	31	7
BotA^P W35A	wt	36	3	wt	24	0
BotA^P W38A	wt	40	7	wt	27	3
BotA^P D39R	wt	42	9	wt	29	5
BotA^P G2A	wt	39	6	wt	28	4
BotA^P G2S	wt	42	9	wt	29	5
BotA^P G2K	wt	42	9	wt	30	6
BotA^P P3A	wt	42	9	wt	29	5
BotA^P P3G	wt	43	10	wt	30	6
BotA^P V4D	wt	40	7	wt	29	5
BotA^P V4T	wt	42	9	wt	30	6
BotA^P V4L	wt	42	9	wt	30	6
BotA^P V5D	wt	39	6	wt	28	4
BotA^P V5T	wt	42	9	wt	29	5
BotA^P V5L	wt	43	10	wt	31	7
BotA^P V6D	wt	41	8	wt	30	6
BotA^P V6T	wt	44	11	wt	30	6
BotA^P V6L	wt	39	6	wt	28	4
BotA^P F7W	wt	43	10	wt	31	7
BotA^P D8E	wt	42	9	wt	30	6
BotA^P D8N	wt	42	9	wt	30	6
BotA^P C9S	wt	40	7	wt	28	4
BotA-F (AA 10-44)	wt	41	8	wt	31	7
BotA-F+C9 (AA 9-44)	wt	50	17	wt	33	9
-	E59A	28	0	E94A	28	0
ATP/Mg^{2+}	E59A	42	14	E94A	40	12
BotA^P	E59A	34	6	E94A	37	9
ATP/Mg^{2+} + BotA^P	E59A	43	15	E94A	41	13
-	E59R	34	0	R310A	25	0
ATP/Mg^{2+}	E59R	34	0	R310A	36	11
BotA^P	E59R	41	7	R310A	31	6
ATP/Mg^{2+} + BotA^P	E59R	40	6	R310A	47	22
-	R234A	31	0			
ATP/Mg^{2+}	R234A	34	3			
BotA^P	R234A	41	10			
ATP/Mg^{2+} + BotA^P	R234A	44	13			
-	R234E	31	0			
ATP/Mg^{2+}	R234E	34	3			
BotA^P	R234E	41	10			
ATP/Mg^{2+} + BotA^P	R234E	44	13			

Table 2-S7 | Summary of all MS data of IpoC reactions presented in this manuscript.

IpoC	BotA	Schematic representation of the detected products	Heterocycle formation
wt	BotA	M-G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ-5 AA	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWD	yes
wt	BotAᴾ-10 AA	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELES	partial
wt	BotAᴾ-15 AA	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEM	no
wt	BotAᴾ E23R	G-P-V-V-V-F-D-C MTADFLNDDPNNARLSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ E28R	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALRMEELESWGAWDGEATS	no
wt	BotAᴾ E30R	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMRELESWGAWDGEATS	yes
wt	BotAᴾ E31R	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMERLESWGAWDGEATS	yes
wt	BotAᴾ E33R	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELRSWGAWDGEATS	yes
wt	BotAᴾ W35A	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESAGAWDGEATS	yes
wt	BotAᴾ W38A	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAADGEATS	yes
wt	BotAᴾ D39R	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWRGEATS	yes
wt	BotAᴾ G2A	A-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ G2S	S-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ G2K	K-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ P3A	G-A-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ P3G	G-G-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ V4D	G-P-D-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotAᴾ V4T	G-P-T-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ V4L	G-P-L-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ V5D	G-P-V-D-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotAᴾ V5T	G-P-V-T-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotAᴾ V5L	G-P-V-L-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotAᴾ V6D	G-P-V-V-D-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotAᴾ V6T	G-P-V-V-T-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotAᴾ V6L	G-P-V-V-L-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ F7W	G-P-V-V-V-W-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ D8E	G-P-V-V-V-F-E-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotAᴾ D8N	G-P-V-V-V-F-N-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotAᴾ C9S	G-P-V-V-V-F-D-S MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
E59A	BotAᴾ	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
E59R	BotAᴾ	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	trace
R234A	BotAᴾ	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
R234E	BotAᴾ	G-P-V-V-V-F-D-C MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial

Table 2-S8 | Summary of all MS data of PurCD reactions presented in this manuscript.

PurCD	BotA	Schematic representation of the detected products	Macrocycle formation
wt	BotA^P	T–V / P–G –V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotA^{PC}	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC}-5 AA	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWD	yes
wt	BotA^{PC}-10 AA	G–P–V–V–V–F–D–C– MTADFLNDDPNNAELSALEMEELES	no
wt	BotA^{PC} E23R	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNARLSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC} E30R	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMRELESWGAWDGEATS	yes
wt	BotA^{PC} E31R	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMERLESWGAWDGEATS	yes
wt	BotA^{PC} E33R	V–V / P–G –V–F–D–C– MTADFLNDDPNNAELSALEMEELRSWGAWDGEATS	yes
wt	BotA^{PC} W35A	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELSAGAWDGEATS	yes
wt	BotA^{PC} W38A	V–V / P–G –V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAADGEATS	yes
wt	BotA^{PC} D39R	V–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWRGEATS	yes
wt	BotA^{PC} G2A	A–P–V–V–V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotA^{PC} G2S	S–P–V–V–V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotA^{PC} G2K	K–P–V–V–V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
wt	BotA^{PC} P3A	V–V / A–G –V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC} P3G	V–V / G–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotA^{PC} V4D	D–V / P–G –V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC} V4T	T–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC} V4L	L–V / P–G –V–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotA^{PC} V5L	V–L / P–G –V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotA^{PC} V6T	V–V / P–G –T–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC} V6L	V–V / P–G –L–F–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	yes
wt	BotA^{PC} F7W	V–V / P–G –V–W–D–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
wt	BotA^{PC} D8N	V–V / P–G –V–F–N–⬠ MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	partial
E94A	BotA^{PC}	G–P–V–V–V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no
R310A	BotA^{PC}	G–P–V–V–V–F–D–C– MTADFLNDDPNNAELSALEMEELESWGAWDGEATS	no

Table 2-S9 | Bottromycin related metabolites observed in *S. scabies btm* mutants. Compounds containing the macrocycle were only detected in *btmE* R239 mutants.

Strain[a]	Bottromycin-related compounds[b] (*m/z*)
S. scabies Δ*btmF*	406.27, 606.30, 723.32, 737.34, 873.45, 887.47
S. scabies Δ*btmF* pIB139-RBS-*btmF* E117A	406.27, 606.30, 723.32, 737.34, 873.45, 887.47
S. scabies Δ*btmF* pIB139-RBS-*btmF* E117R	406.27, 606.30, 723.32, 737.34, 873.45, 887.47
S. scabies Δ*btmF* pIB139-RBS-*btmF* R341A	406.27, 606.30, 723.32, 737.34, 873.45, 887.47
S. scabies Δ*btmF* pIB139-RBS-*btmF* R341E	406.27, 606.30, 723.32, 737.34, 873.45, 887.47
S. scabies Δ*btmE*	406.27
S. scabies Δ*btmE* pIB139-RBS-*btmE* E59A	406.27
S. scabies Δ*btmE* pIB139-RBS-*btmE* E59R	406.27
S. scabies Δ*btmE* pIB139-RBS-*btmE* R239A	406.27, 795.42, 809.44, 823.45[c]
S. scabies Δ*btmE* pIB139-RBS-*btmE* R239E	406.27, 795.42, 809.44, 823.45[c]

a. The genes *btmF*, *btmE* and *bmtD* from *S. scabies* encode proteins that are homologs to BotCD/PurCD, BotC/IpoC and BotA, respectively, and the mutated residues are equivalent to the residues mutated for PurCD and IpoC.
b. See Crone *et al.* 2016[19] for further details on these bottromycin-related compounds. No detected compound contains the macrocycle.
c. 795.42, 809.44 and 823.45 are differentially methylated forms of mature bottromycin. This indicates that the BtmE mutants are active, which is consistent with *in vitro* activity of IpoC mutants.

Table 2-S10 | Primers used to generate *in vivo* mutations in *S. scabies*.

Primer	Sequence[a]
BtmF start	aggagga<u>catatg</u>acccgtcccgc
BtmF end	acagg<u>gaattc</u>cagggcggtgccac
BtmF E117A F	gcatggcgagcgcgctgttcGCCagcgccgagcactaccac
BtmF E117A R	gtggtagtgctcggcgctGGCgaacagcgcgctcgccatgc
BtmF E117R F	gcatggcgagcgcgctgttcCGCagcgccgagcactaccac
BtmF E117R R	gtggtagtgctcggcgctGCGgaacagcgcgctcgccatgc
BtmF R341A F	cggacgtacgcggtgcagGCCgcgctgctcgagtacctc
BtmF R341A R	gaggtactcgagcagcgcGGCctgcaccgcgtacgtccg
BtmF R341E F	cggacgtacgcggtgcagGAGgcgctgctcgagtacctc
BtmF R341E R	gaggtactcgagcagcgcCTCctgcaccgcgtacgtccg
BtmE start	gggagaggg<u>catatg</u>cgcgaagcg
BtmE end	ctgccgg<u>aattc</u>cgtctcgcgctc
BtmE E59A F	cgggcgaaagccctgtcgGCCgccgtggagcggctggtg
BtmE E59A R	caccagccgctccacggcGGCcgacagggctttcgcccg
BtmE E59R F	cgggcgaaagccctgtcgCGGgccgtggagcggctggtg
BtmE E59R R	caccagccgctccacggcCCGcgacagggctttcgcccg
BtmE R239A F	gtcgcggaggcgaccggcGCCgcgctgctggagctggcg
BtmE R239A R	cgccagctccagcagcgcGGCgccggtcgcctccgcgac
BtmE R239E F	gtcgcggaggcgaccggcGAGgcgctgctggagctggcg
BtmE R239E R	cgccagctccagcagcgcCTCgccggtcgcctccgcgac

a. Underlined nucleotides indicate a restriction site while capital letters indicate mutated codons.

Materials and Methods

Cloning of native and mutant PurCD and IpoC

Synthetically produced, codon-optimized genes of the genes with accession numbers WP_019887079 and WP_048819674.1 were obtained from Eurofins Genomics and named "PurCD" and "IpoC", respectively. Using the built-in N-terminal *NcoI* as well as C-terminal *HindIII* restriction sites, the genes were cut out of the provided pEX-K4 cloning vector and ligated with standard T4 DNA ligase protocols using a 1:3 vector to plasmid molar ratio into the expression vectors pHisTEV (PurCD) and pHisSUMOTEV (IpoC), both of which were a gift from Dr. Huanting Liu, St. Andrews University.[20] Clones were sequenced using T7 promotor as well as T7 terminator sequencing (LGC genomics), and yielded the expression constructs pHisTEV-PurCD and pHisSUMOTEV-IpoC. Point mutations were designed using the overlap extension method,[21] after which the PCR products were cloned into the respective plasmids for each gene.

Cloning of native BotA and BotA point mutants

Two different vectors were used for BotA precursor peptide expression. In both vectors the restriction sites *NcoI* and *HindIII* were used for cloning:

For the production of wt BotA and follower peptide mutants, codon optimized *botA* was cloned into a modified pHisSUMOTEV vector, from which the TEV protease site had been removed. *botA* was amplified using a 5' primer that added a TEV site before the N-terminal methionine of BotA, leading to a TEV site with the amino acid sequence ENLYFQM, where the "M" is the N-terminal methionine of BotA. After TEV cleavage, BotA is released with its native N-terminus.

For the production of BotA^P core peptide mutants, the follower peptide (BotA-F) and follower+C9 (BotA-F+C9), we used unmodified pHisSUMOTEV. Expression of BotA from pHisSUMOTEV when using the *NcoI*-site for cloning of botA, would have resulted in peptides with N-terminal cloning artefacts after TEV cleavage (GAMA<u>GPVVV</u>… (BotA residues are underlined)). To make the native N-terminus of BotA^P available, we modified the 5' primers to give the N-terminus GAMAGK<u>GPVVV</u>… (BotA residues underlined). This would allow proteolytic processing with Trypsin during BotA^P purification and liberate the native N-terminus of the peptide.

Expression of native and mutant PurCD and IpoC

Expression vectors with verified sequence were transformed into chemically competent *E. coli* Lemo21(DE3) cells (New England Biolabs) using a standard heat shock procedure. Using double selection plates with both Kanamycin (50 µg/mL) as well as Chloramphenicol (34 µg/mL), colonies with both the expression as well as the Lemo21-plasmid were obtained. For expression, a fresh colony was used to inoculate 100 mL of LB Broth starter culture using the Luria-Miller formulation with both antibiotics and was incubated at 37 °C and 200 rpm for 16 h.

Large scale expression of pHisTEV-PurCD was carried out in LB Broth whereas large scale expression of pHisSUMOTEV-IpoC was carried out in modified Terrific Broth. Modified Terrific Broth was obtained by adding 1 mM $MgSO_4$ and NPS buffer[22] to a final concentration of 1x after autoclaving instead of the usual potassium phosphate buffer system. Large scale expression for both proteins was carried out by inoculating each medium with both antibiotics at the aforementioned concentrations with the starter culture using a 1/100 dilution. Cultures were grown at 37 °C and 200 rpm until the OD_{600} reached 0.8, at which point the agitation was stopped and the culture equilibrated for 1 h while reducing the temperature of the shaker to 16 °C and 18 °C for pHisTEV-PurCD and pHisSUMOTEV-IpoC, respectively. While reaching the desired temperature, protein expression was induced by adding IPTG to a final concentration of 0.1 mM. Cultures were then incubated at 200 rpm for 16 h and 24 h for pHisTEV-PurCD and pHisSUMOTEV-IpoC, respectively.

Expression of native and mutant BotA

Expression vectors were transformed into chemically competent *E. coli* Lemo21(DE3) cells (New England Biolabs) using a standard heat shock procedure. Using double selection plates with both Kanamycin (50 µg/mL) as well as Chloramphenicol (34 µg/mL), colonies with both the expression as well as the Lemo21-plasmid could be obtained. For expression, a fresh colony was added into 100 mL of LB Broth starter culture using the Luria-Miller formulation with both antibiotics and was incubated at 37 °C and 200 rpm for 16 h.

Large scale expression of both native and mutant BotAs was carried out in LB Broth. Large scale expression was carried out by inoculating LB Broth with both antibiotics at the aforementioned concentrations with the starter culture using a 1:100 dilution. Cultures were grown at 37 °C and 200 rpm until the OD_{600} reached 0.8. At this point, protein expression was induced by adding IPTG to a final concentration of 0.5 mM and dropping the temperature to 18 °C. Cultures were incubated at 200 rpm for 16 h.

<u>Purification of native and mutant PurCD and IpoC</u>

Cells were harvested by centrifuging the culture at 4,000 rpm and 4 °C for 10 min. The pellet was retrieved and frozen at -80 °C until further use.

Two different lysis buffers were used for purification. The pellet of pHisTEV-PurCD was resuspended in PurCD lysis buffer (200 mM NaCl, 50 mM Tris pH 8.0, 20 mM Imidazole pH 8.0, 3 mM β-ME, 10% Glycerol, final pH 8.7 at RT). Pellets of pHisSUMOTEV-IpoC were resuspended in IpoC lysis buffer (500 mM NaCl, 20 mM Bis-Tris pH 6.8, 20 mM Imidazole pH 8.0, 3 mM β-ME, 10% Glycerol). In both cases, 100 mL of lysis buffer was used per 25 g of wet cell pellet. Only for IpoC, the lysis buffer was supplemented with cOmplete EDTA-free protease inhibitor tablets (Roche), while DNase (0.4 mg/g wet cells, Sigma) was added to lysis buffers of both proteins. The cell suspension was lysed via passage through a cell disruptor (30 kpsi, Microfluidics Corp.), and the cell debris was removed by centrifugation (40,000x g, 4 °C, 20 min). The supernatant was filtered through a 0.45 μm filter and passed over a 5 mL Ni-NTA superflow cartridge (Qiagen) at a flow rate of 5 mL/min pre-equilibrated in lysis buffer. After application of the lysate and an extensive column wash (30 CV lysis buffer), the bound protein was eluted in one step using lysis buffer supplemented with 250 mM Imidazole. Imidazole was removed by passing the eluate over a desalting column (16/10 GE Healthcare) at a flow rate of 10 mL/min pre-equilibrated in lysis buffer. The protein was then incubated with TEV protease for 14 h at 4 °C at a 1:10 protein : TEV ratio to cleave the expression and purification tags. Digested protein was reapplied to a 5 mL Ni-NTA-Superflow cartridge equllibrated in lysis buffer and the flow-through was collected at a flow rate of 5 mL/min. Finally, proteins were polished by size exclusion chromatography (Superdex 200 16/600, GE Healthcare) pre-equilibrated in the corresponding gel filtration buffers (PurCD: 200 mM NaCl, 50 mM Tris, 10% Glycerol, 0.5 mM TCEP, pH 8.7 at RT; IpoC: 300 mM NaCl, 10 mM HEPES, 10% Glycerol, 0.5 mM TCEP, final pH 7.4 at RT) at a flow rate of 1 mL/min. IpoC and PurCD and the respective enzyme point mutants eluted at a retention volume of 88 mL and 85 mL, respectively. Protein purity and integrity was verified by SDS-PAGE as well as intact protein mass spectrometry using a previously described method[23]. The enzymes were concentrated using a 30 kDa cutoff filter (Sartorius AG). The protein concentration was determined at a Spectrophotometer (Thermo Fisher Scientific) using the extinction coefficients calculated from the amino acid sequence by the ExPASY ProtParam Server[24] (http://web.expasy.org/protparam/). In general, all purification steps were performed at 4 °C. Enzymes at a concentration of 120 μM were flash frozen in single-use aliquots and stored at -80 °C until further use.

Purification of native and mutant BotA

All BotA purifications were carried out in the same lysis buffer (500 mM NaCl, 20 mM Tris pH 8.0, 20 mM imidazole pH 8.0, 3 mM β-ME). For every 25 g of wet cell pellet, 100 mL of lysis buffer was added. The lysis buffer was supplemented with cOmplete EDTA-free protease inhibitor tablets (Roche) and Dnase (0.4 mg/g wet cell pellet, Sigma). The cell suspension was lysed via passage through a cell disruptor (30 kpsi, Microfluidics Corp.), and the cell debris was removed by centrifugation (40,000x g, 4 °C, 20 min). The supernatant was filtered through a 0.45 μm filter and applied to a His-Trap HP 5 mL column (GE Healthcare) pre-equilibrated in lysis buffer at a flow rate of 5 mL/min. After application of the lysate and an extensive column wash (30 CV lysis buffer), the bound protein was eluted using a lysis buffer supplemented with 250 mM imidazole. The proteins were passed over a desalting column (16/10 GE Healthcare) at a flow rate of 10 mL/min pre-equilibrated in lysis buffer. The proteins were subsequently incubated with TEV protease for 2 h at room temperature at a 1:10 ratio to remove the His$_6$-tagged SUMO. Digested protein was passed over a 5 mL His Trap HP column at a flow rate of 5 mL/min to remove the expression and purification tag and the flow-through was collected. Peptides expressed using the pHisSUMOTEV vector, were subsequently incubated with trypsin at a ratio of 1:100 for 1 h at 37 °C to generate the native N-terminus of the peptides. Finally, the proteins were passed over a Superdex 30 16/60 size exclusion chromatography column (GE Healthcare) pre-equilibrated in gel filtration buffer (150 mM NaCl, 10 mM HEPES, 0.5 mM TCEP, pH 7.4) at a flow rate of 1 mL/min. This purification step yielded a pure and homogenous sample that eluted at a retention volume of 64 mL. Following the size exclusion chromatography, BotA follower mutants were incubated with 2 μM BotP and 200 μM CoCl$_2$ for 1 h at 37 °C to remove the N-terminal methionine. Afterwards, they were passed over a Superdex Peptide 10/300 GL to remove the BotP. Protein purity and integrity was verified by SDS-PAGE as well as mass spectrometry.

LC-MS analysis

All measurements to analyse the mass of processed or unprocessed BotA peptides were performed on a Dionex Ultimate 3000 RSLC system using a BEH C18, 100 x 2.1 mm, 3.5 μm dp column equipped with a C18 precolumn (Waters). Samples of 1 μL were separated by a gradient from (A) H$_2$O + 0.1% formic acid to (B) ACN + 0.1% formic acid at a flow rate of 800 μL/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step at 5% B, followed by an increase to 35% B in 2.5 min, to 42.5% B in 3 min and to 95% B in 0.5 min. After a 2 min

step at 95% B the system was re-equilibrated to the initial conditions (5% B). UV spectra were recorded by a DAD in the range from 200 to 600 nm.

For MS measurements on maXis-2 UHR-TOF mass spectrometer (Bruker Daltonics), the LC flow was split 1:8 before entering the mass spectrometer using the Apollo ESI source. In the source region, the temperature was set to 250 °C, the capillary voltage was 4000 V, the dry-gas flow was 10.0 L/min and the nebulizer was set to 30 psi. After the generated ions passed the quadrupole with a low cut-off at 150 m/z they were trapped in the collision cell for 100 μs and then transferred within 10 μs through the hexapole into the ICR cell. Data were recorded in the mass range from 250 to 2500 m/z. Full length BotA and BotA^P peptides showed a distinct peak at a retention time of 4.6 min. Peptides with a thiazoline, introduced by IpoC, had a retention time of 4.3 min, while peptides with a macroamidine cycle had a retention time of 4.2 min. Due to the size, BotA truncations showed different retention times and eluted later from the column. For quantification of processed BotAs in enzyme reactions the peak area of $[M+H]^{3+}$ was integrated and concentrations were calculated using a calibration curve obtained from serial dilutions of the respective purified BotA using the TASQ 1.1 software (Bruker Daltonics).

Thermal shift assay (TSA)

Protein melting temperatures (T_m) were determined by monitoring protein unfolding using SYPRO orange as fluorescence probe. Enzymes were diluted to 5 μM in buffer containing 5x SYPRO orange (Sigma). For the determination of stabilising effects of the enzymes by peptides or other additives, buffer containing 50 mM HEPES, 200 mM NaCl and 10% glycerol (pH 8.5) was used. Final peptides concentrations of 30 and 50 μM for IpoC and PurCD, respectively, were used. To investigate the effects of ATP and MgCl$_2$, final concentrations of 5 mM each were used. To determine the stability of PurCD at different pH, buffers containing 50 mM buffering agent, 200 mM NaCl and 10% glycerol were used. For pH 7.0 to 9.0 Tris and pH 9.5 to 11.0 CHES were used as buffering agents. Samples (40 μL/well) were analysed in 96 well plates (qPCR semi-skirted 96well PCR plates (peqlab)) that were sealed (adhesive qPCR film (peqlab)). Measurements were carried out with a realtime PCR machine (peqSTAR 96Q) using the SYBR Green I filter and a temperature gradient from 10 to 80 °C with stepwise increments of 1 °C and 1 min hold. After each temperature step the fluorescence intensity was measured. The melting temperature (T_m) were obtained using the derivative method. All conditions were tested in triplicates and mean values were calculated for the graphic presentation.

<u>o-Phthalaldehyd peptide concentration assay</u>

Generally, peptide concentrations were determined by measuring the absorbance at 280 nm using a spectrophotometer (Thermo Scientific) and calculated using extinction coefficients (see above). For truncated BotA variants without tryptophan residues (BotA-10 and BotA-15AA), o-phthalaldehyd (OPA)[25] was used as fluorogenic reagent, which reacts in presence of 2-mercaptoethanol with primary amino groups (e.g. of peptides) and forms a blue coloured fluorescent product.

OPA reagent solution was freshly made by dissolving 10 mg OPA (Fluka) in 250 µL ethanol, 9.8 mL PBS and 20 µL 2-mercaptoethanol.

The OPA peptide concentration assay was performed using 96-well flat-bottom plates. Full length BotA was used as the standard in the concentration range 0-100 µM. 200 µL OPA reagent was added to 30 µL of sample/well and incubated 5 min at room temperature. Fluorescence intensity was measured at 360 nm excitation and 430 nm emission using an Infinite 200 Pro (Tecan) plate reader. Samples and standards were tested in triplicates.

<u>Heterocyclization of BotA with IpoC</u>

For heterocyclization reactions, in general, 50 µM substrate (BotA, BotA^P, or BotA^P core or follower mutants) was incubated with 5 µM wt IpoC and 5 mM ATP/MgCl$_2$ in gel-filtration buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 1 mM TCEP) for 12 h at 37 °C. Turnover of the different BotA variants by IpoC was analyzed by LC-MS.

For further use (e.g. enzyme reactions with PurCD or MS^2-analysis), large scale reactions of heterocyclized BotA (BotA^{PC} or BotA^C) were purified by size-exclusion chromatography on a Superdex Peptide 10/300 GL at 0.5 mL/min using gel-filtration buffer. Peptides eluted between 9.5 and 11.5 mL, were concentrated using 3 kDa cutoff filters (Sartorius AG) and peptide concentrations were determined spectrophotometrically as described above.

To determine the time until complete turnover was reached, enzyme reactions using BotA and BotA^P were set up as described above using wt IpoC. Samples were taken and directly frozen in liquid nitrogen after 0, 0.5, 1, 2, 4, 8 and 16 h. The conversion of BotA and BotA^P after the different incubation times was subsequently analyzed by LC-MS (see above for details).

To test the enzymatic activity of IpoC mutants (IpoC E59A, E59R, R234A, R234E), BotA^P was used as the substrate and reactions were set up and analyzed as describe above.

For the determination of the initial conversion rates of BotA and BotA^P by IpoC, 100 µM substrate was incubated with 2 µM wt IpoC and 5 mM ATP/MgCl$_2$ in gel filtration buffer. Samples were incubated in the LC sample holder at 37 °C and 1 µl aliquots were automatically

taken and analyzed every 10 min by the LC-MS system in the time range from 2 to 42 min. All reactions (different substrates, reaction conditions, IpoC mutants) were tested in triplicates.

Macrocyclization reaction by PurCD

For macrocyclization reactions by PurCD, 50 µM BotA peptide (BotA^P, BotA^{PC}) was incubated with 10 µM enzyme and 10 mM ATP/MgCl$_2$ in macrocyclization buffer (50 mM Tris pH 9.5, 200 mM NaCl, 10% glycerol) for 90 min at room temperature. Macroamidine formation was determined by LC-MS analysis as described above. Product (BotA^{PCCD}) was purified for MS^2-analysis and further experiments by size-exclusion chromatography on a Superdex peptide 10/300 GL column using gel filtration buffer at a flow rate of 0.5 mL/min. BotA^{PCCD} eluted eluted between 10.0 and 12.0 mL. Peptides were concentrated using 3 kDa cutoff filters (Sartorius AG) and peptide concentrations were determined spectrophotometrically as described above.

To determine the effect of different pHs on PurCD turnover, buffers containing 50 mM buffering agent, 200 mM NaCl and 10% glycerol were used. For pHs 7.0 to 8.5 TRIS and pHs 9.0-11.0 CHES were used as buffering agents. Reactions were set up as described above (50 µM BotA^{PC}, 10 µM PurCD, 10 mM ATP/MgCl$_2$) and analyzed by LC-MS.

The ability of PurCD to use different BotA^P follower and core mutants as substrate was tested. As PurCD strongly favored IpoC-cyclyzed BotA^P peptides, 50 µM substrate was first incubated with 5 µM IpoC and 10 mM ATP/MgCl$_2$ in macrocyclization buffer for 12 h at 37 °C. Samples were cooled down to room temperature, 10 µM PurCD was added and samples were incubated for a further 90 min at room temperature. The conversion of the BotA^P mutants was analyzed by LC-MS.

To determine the conversion rate of BotA^{PC} by PurCD, 50 µM substrate was incubated with 10 µM PurCD and 10 mM ATP/MgCl$_2$. Samples were incubated in the LC sample holder at 20 °C and 1 µl aliquots were automatically taken and analyzed every 10 min by the LC-MS system in the time range from 2 to 122 min.

To test the enzymatic activity of PurCD mutants (PurCD E94A, R310A), BotA^{PC} was used as substrate (50 µM BotA^{PC}, 10 µM PurCD mutant, 10 mM ATP/MgCl$_2$ in macrocyclization buffer). The enzymatic activity of the PurCD mutants was analyzed by LC-MS as described above.

We observed macroamidine ring opening after extended incubation time of the reactions. To test the enzyme dependence of the ring opening, 50 µM purified BotA^{PCCD} was incubated for 12 h at RT in macrocyclization buffer with 1. no additive, 2. 10 mM ATP/MgCl$_2$, 3. 10 µM

PurCD, 4. 10 µM PurCD + 10 mM ATP/MgCl$_2$. Macroamidine ring opening was analyzed by LC-MS-analysis as described above. To determine if PurCD was stable in the macrocyclization buffer for at least 8 h and thus catalyze re-opening of the macroamidine, we performed analytical gel filtration experiments. 10 µM PurCD was incubated at room temperature in macrocyclization buffer (200 mM NaCl, 50 mM Tris, 10% Glycerol, pH 9.5) with addition of 10 mM ATP and 10 mM MgCl$_2$. After 0, 1.5 and 8 h, the full reaction mixture was injected onto a Superose 6 Increase 10/300 GL column at a flow rate of 0.5 mL/min. The resulting chromatogram was analyzed by plotting the absorption at 280 nm against the retention volume. All reactions (different substrates, reaction conditions, PurCD mutants) were tested in triplicates.

MS^2 analysis of BotA peptides

Tandem MS data were acquired on a Synapt G2-Si mass spectrometer (Waters) operated in electrospray ionization positive ion mode with a scan time of 1 s (inter scan time of 0.015 s) in a mass range of *m/z* 50 to 5000. Samples were diluted into 50% methanol/water containing 0.1% formic acid and infused into the mass spectrometer at 5-10 µL/min over a period of 2 min using a Harvard Apparatus syringe pump. Data were collected in continuum mode and the following parameters were used: collision gas = argon; capillary voltage = 2.60 kV; cone voltage = 40 V; lockspray capillary voltage = 3.5 kV; source temperature = 100 °C; desolvation temperature = 250 °C; cone gas flow = 10 L/h; desolvation gas flow = 500 L/h. For each peptide, the most intense multiply charged peak in the MS1 was selected for MS^2 fragmentation, where the trap collision energy was ramped stepwise from 10 V until the precursor was fully fragmented, which usually occurred in the range of 40-45 V. Acquisition and analysis of data was carried out using MassLynx 4.1 (Waters). The reported MS^2 spectra are the average of approximately 100 scans, and MS^2 spectra were deconvoluted using the MaxEnt 3 tool in MassLynx 4.1.

Construction of *btmF* point mutants in *Streptomyces scabies*

To assess the *in vivo* effect of mutations to a BotCD-like protein in a native producer, four different point mutants were generated in *btmF* (E117A, E117R, R341A, R341E), the gene encoding the BotCD homolog in the *S. scabies* bottromycin gene cluster. These mutations are equivalent to the ones tested for *in vitro* activity of PurCD. These were generated by fusion PCR, using internal overlapping primers containing the desired codon change with external primers covering the start and end of *btmF* (see Table 2-S10 for primers). The final fusion PCR

products were digested and ligated into the NdeI and EcoRI sites of the pIB139-RBS vector[6] to generate pIB139-RBS-*btmF* E117A, pIB139-RBS-*btmF* E117R, pIB139-RBS-*btmF* R341A and pIB139-RBS-*btmF* R341E. These four constructs were introduced by intergeneric conjugation from *E. coli* ET12567/pUZ8002 into *S. scabies* Δ*btmF*[6] to study the effect of each *btmF* point mutation in the complementation of the chromosomal deletion (note: pIB139-RBS-*btmF* had previously been determined to fully restore bottromycin production in *S. scabies* Δ*btmF*).[6] Following selection with apramycin (50 μg/mL) and nalidixic acid (30 μg/mL) in SFM medium, three different exconjugants for each point mutation were independently inoculated into 5 mL of TSB medium and grown for 48 h. 500 μL of each culture were employed as seed cultures to inoculate 5 mL of bottromycin production medium (BPM)[6] in 50 mL tubes capped with foam bungs. After 96 h of fermentation (250 rpm, 30 °C), 500 μL of each culture was mixed with 500 μL methanol, shaken for 30 min at RT and centrifuged at 13,000 rpm for 30 min before analysis of the resulting supernatant by LC-MS as previously described.[6] Bottromycin-related metabolites were manually identified using Browser and Postrun analysis software (Shimadzu). Untargeted metabolomic analysis using Profiling Solution (Shimadzu) indicated that there were no significant differences between the total metabolome of *S. scabies* Δ*btmF* and the metabolomes of any of the *btmF* point mutants.

Construction of *btmE* point mutants in *Streptomyces scabies*

To assess the *in vivo* effect of mutations to a BotC-like protein in a native producer, four different point mutants were generated in *btmE* (E59A, E59R, R239A, R239E), the gene encoding the BotC homolog in the *S. scabies* bottromycin gene cluster. These mutations are equivalent to the ones tested for *in vitro* activity of IpoC. These were generated by fusion PCR, using internal overlapping primers containing the desired codon change with external primers covering the start and end of *btmE* (see Table 2-S10 for primers). The final fusion PCR products were digested and ligated into the NdeI and EcoRI sites of pIB139-RBS to generate pIB139-RBS-*btmE* E59A, pIB139-RBS-*btmE* E59R, pIB139-RBS-*btmE* R239A and pIB139-RBS-*btmE* R239E. These four constructs were introduced by intergeneric conjugation from *E. coli* ET12567/pUZ8002 into *S. scabies* Δ*btmE*[6] (note: pIB139-RBS-*btmE* had previously been determined to fully restore bottromycin production in *S. scabies* Δ*btmE*).[6] Following selection with apramycin (50 μg/mL) and nalidixic acid (30 μg/mL) in SFM medium, three different exconjugants for each point mutation were independently inoculated into 5 mL of TSB medium and grown for 48 h. 500 μL of each culture were employed as seed cultures to inoculate 5 mL of BPM in 50 mL tubes capped with foam bungs. After 96 h of fermentation (250 rpm, 30 °C),

500 µL of each culture was mixed with 500 µL methanol, shaken for 30 min at RT and centrifuged at 13,000 rpm for 30 min before analysis of the resulting supernatant by LC-MS as previously described.[6] Bottromycin-related metabolites were manually identified using Browser and Postrun analysis software (Shimadzu). Untargeted metabolomic analysis using Profiling Solution (Shimadzu) indicated that there were no significant differences between the total metabolome of *S. scabies* Δ*btmE* and the metabolomes of any of the *btmE* E59 point mutants, whereas bottromycin was detected in the R239 mutants (Table 2-S9). This is consistent with the *in vitro* activity of IpoC mutants (Figure 2-S18).

2.4 References

1. Otaka, T.; Kaji, A., Mode of action of bottromycin A2. Release of aminoacyl- or peptidyl-tRNA from ribosomes. *J Biol Chem* **1976,** *251* (8), 2299-306.

2. Crone, W. J. K.; Leeper, F. J.; Truman, A. W., Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chemical Science* **2012,** *3* (12), 3516-3521.

3. Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., Posttranslational [small beta]-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *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., Structure and biosynthesis of the antibiotic bottromycin D. *Org Lett* **2012,** *14* (19), 5050-3.

5. Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Muller, R., Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *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., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *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., From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* **1996,** *274* (5290), 1188-93.

8. Bayer, A.; Freund, S.; Nicholson, G.; Jung, G., Posttranslational Backbone Modifications in the Ribosomal Biosynthesis of the Glycine-Rich Antibiotic Microcin B17. *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., YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function. *Chem Rev* **2017,** *117* (8), 5389-5456.

10. Dunbar, K. L.; Tietz, J. I.; Cox, C. L.; Burkhart, B. J.; Mitchell, D. A., Identification of an Auxiliary Leader Peptide-Binding Protein Required for Azoline Formation in Ribosomal Natural Products. *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., Discovery of a new ATP-binding motif involved in peptidic azoline biosynthesis. *Nat Chem Biol* **2014,** *10* (10), 823-9.

12. Dunbar, K. L.; Melby, J. O.; Mitchell, D. A., YcaO domains use ATP to activate amide backbones during peptide cyclodehydrations. *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., Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *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., An enzymatic route to selenazolines. *Chembiochem* **2013,** *14* (5), 564-7.

15. Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Muller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *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., Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature* **2015,** *517* (7535), 509-12.

17. Yongye, A. B.; Li, Y.; Giulianotti, M. A.; Yu, Y.; Houghten, R. A.; Martinez-Mayorga, K., Modeling of peptides containing D-amino acids: implications on cyclization. *J Comput Aided Mol Des* **2009,** *23* (9), 677-89.

18. Izawa, M.; Kawasaki, T.; Hayakawa, Y., Cloning and heterologous expression of the thioviridamide biosynthesis gene cluster from Streptomyces olivoviridis. *Appl Environ Microbiol* **2013,** *79* (22), 7110-3.

19. Panjikar, S.; Parthasarathy, V.; Lamzin, V. S.; Weiss, M. S.; Tucker, P. A., On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr D Biol Crystallogr* **2009,** *65* (Pt 10), 1089-97.

20. Liu, H.; Naismith, J. H., A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr Purif* **2009,** *63* (2), 102-11.

21. Higuchi, R.; Krummel, B.; Saiki, R. K., A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res* **1988,** *16* (15), 7351-67.

22. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **2005,** *41* (1), 207-34.

23. von Tesmar, A.; Hoffmann, M.; Pippel, J.; Fayad, A. A.; Dausend-Werner, S.; Bauer, A.; Blankenfeldt, W.; Müller, R., Total Biosynthesis of the Pyrrolo[4,2]benzodiazepine Scaffold Tomaymycin on an In Vitro Reconstituted NRPS System. *Cell chemical biology* **2017,** *24* (10), 1216-1227.e8.

24. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S. e.; Wilkins, M. R.; Appel, R. D.; Bairoch, A., Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, Walker, J. M., Ed. Humana Press: Totowa, NJ, 2005; pp 571-607.

25. Roth, M., Fluorescence reaction for amino acids. *Analytical Chemistry* **1971,** *43* (7), 880-882.

Chapter 3

Thiazoline-Specific Amidohydrolase PurAH Is the Gatekeeper of Bottromycin

Asfandyar Sikandar,[+] **Laura Franz**,[+] Okke Melse, Iris Antes, and Jesko Koehnke

[+] shared authorship

Author contributions:

Laura Franz contributed to the manuscript by performing the *in vitro* biochemical experiments and mass spectrometry measurements, which established the function of the bottromycin amidohydrolase PurAH. She expressed and purified all biosynthetic enzymes used in this study for the biochemical experiments. In addition, she cloned, expressed and purified the enzyme and peptide mutants. Asfandyar Sikandar cloned the wt enzyme and purified the enzyme for protein crystallization experiments. He crystallized the enzyme and determine protein structure. Okke Melse and Ires Antes developed a novel docking approach for highly modified peptides that was instrumental in determining the binding mode and proposing a mechanism for PurAH. Jesko Köhnke supervised and guided the project and wrote the manuscript. All authors contributed to the manuscript writing, designed figures for their respective experiments, provided critical feedback and discussed the results of the manuscript.

3 The Thiazoline-Specific Amidohydrolase PurAH is the Gatekeeper of Bottromycin Biosynthesis

3.1 Abstract

The ribosomally synthesized and post-translationally modified peptide (RiPP) bottromycin A2 possesses potent antimicrobial activity. Its biosynthesis involves the enzymatic formation of a macroamidine, a process previously suggested to require the concerted efforts of a YcaO enzyme (PurCD) and an amidohydrolase (PurAH) *in vivo*. *In vitro*, PurCD alone is sufficient to catalyze formation of the macroamidine, but the process is reversible. We set out to probe the role of PurAH in macroamidine formation *in vitro*. We demonstrate that PurAH is highly selective for macroamidine-containing precursor peptides and cleaves C-terminal of a thiazoline, thus removing the follower peptide. After follower cleavage, macroamidine formation is irreversible, indicating PurAH as the gatekeeper of bottromycin biosynthesis. The structure of PurAH suggests residues involved in catalysis, which were probed through mutagenesis.
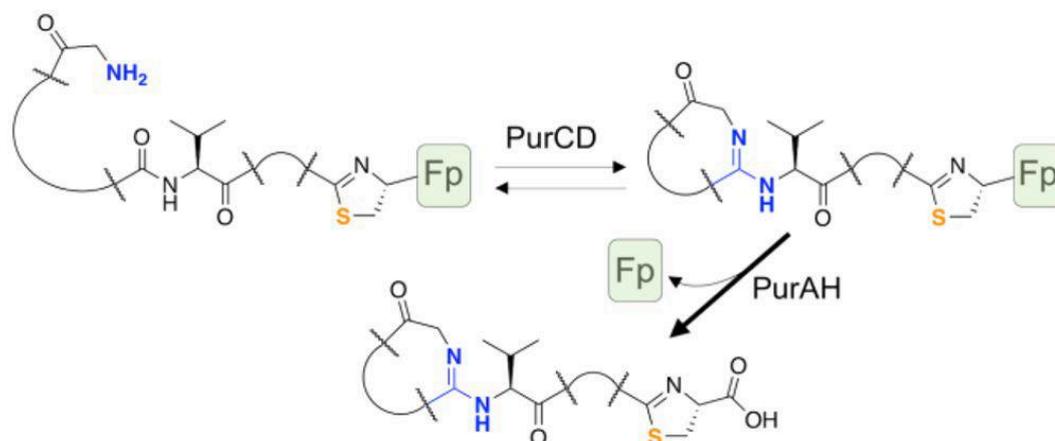
Figure 3-1 | Graphical abstract.

3.2 Main Text

Bottromycins[1, 2] are natural product antibiotics with activity against problematic human pathogens such as Methicillin-resistant Staphylococcus aureus (MRSA)[3, 4]. They bind to the A-site of the prokaryotic 50S ribosome, which is a novel antibiotic target[5-7]. Bottromycins belong to the growing family of ribosomally synthesized and post-translationally modified peptides (RiPPs), and their biosynthesis and total synthesis have received increasing attention[8-12]. They are derived from the precursor peptide BotA, which undergoes a series of post-translational enzymatic tailoring steps, the order of which has been proposed based on an untargeted metabolomics approach using mass spectral networking (Figure 3-2, we will use *Streptomyces sp.* BC16019 nomenclature)[13]: First, the N-terminal methionine is removed by a leucyl-amino peptidase (BotP), followed by heterocyclization of the BotA cysteine residue to thiazoline by the YcaO enzyme BotC and Cβ-methylation by radical methyl transferases 1-3. Next, a second YcaO enzyme, BotCD, was reported to act together with the metallo-dependent amidohydrolase BotAH in macroamidine formation[13]. Removal of the follower peptide by the α/β-hydrolase BotH, successive oxidative decarboxylation of the thiazoline to a thiazole (BotCYP) and O-methylation of an aspartate (BotOMT) complete bottromycin biosynthesis. *In vitro* work has confirmed the assigned functions for BotP[13, 14], BotC[8, 9] and BotCD[8, 9]. It was demonstrated that the BotCD homologs BmbE and PurCD alone are sufficient for macroamidine formation *in vitro*[8, 9]. For PurCD, macroamidine formation was reversible: It catalyzes both, macroamidine formation and its reopening[8].

To probe the role of the amidohydrolase in bottromycin biosynthesis *in vitro*, we attempted to express BotAH (insoluble). The close homolog PurAH (72% sequence identity to BotAH) from *Streptomyces purpureus* (Figure 3-2) could be expressed and purified. We first tested possible substrates and incubated PurAH with precursor peptide BotA, BotA[P] (N-terminal methionine removed), BotA[PC] (BotA[P] with heterocyclized cysteine) and BotA[PCCD] (macrocyclized BotA[PC]) (Figure 3-3a).

All samples were analyzed by liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS). We observed no activity on BotA or BotA[P]. For the BotA[PC] reaction, we observed two new masses, which corresponded to the heterocyclized BotA core peptide (**1**, Figure 3-S1A) (m/z calc$_{mono.}$: 816.3840 Da, observed: 816.3814 Da, error -3.2 ppm) and the follower peptide (m/z calc$_{mono.}$: 3857.62 Da, observed: 3857.63 Da, error 2.6 ppm), but the reaction was incomplete. For the BotA[PCCD] reaction, the substrate was consumed completely and two new masses appeared: One corresponded to the heterocyclized, macroamidine-
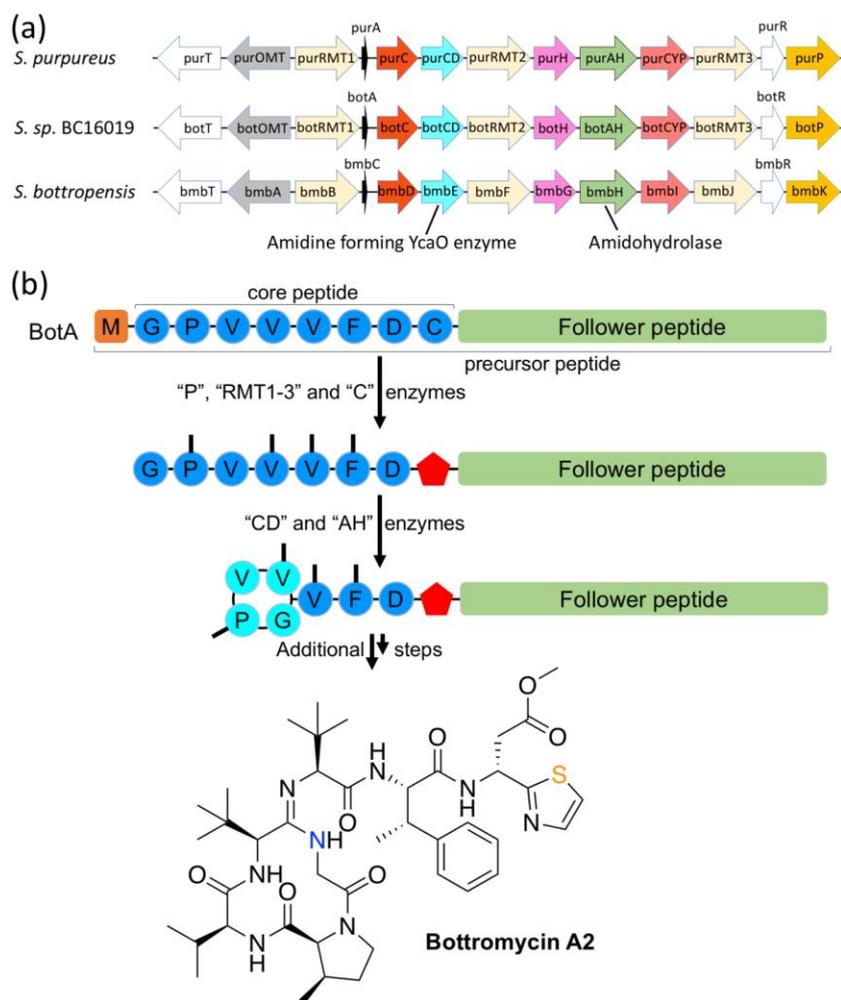
Figure 3-2 | (a) A gene cluster highly homologous in sequence and organization to those of confirmed bottromycin producers *Streptomyces sp.* BC16019 and *S. bottropensis* was found in *S. purpureus* (top). Intergenic regions not drawn to scale. **(b)** Bottromycin biosynthesis as proposed by untargeted metabolomics[13]. Methylation by radical methyl transferases (RMT1-3) is shown as black lines, heterocyclized cysteine as a red pentagon and the macroamidine in cyan.

containing BotA core peptide (**2**, Figure 3-S1b), while the other corresponded to the follower peptide (m/z calc$_{mono.}$: 798.3734 Da, observed: 798.3706 Da, error -3.5 ppm and m/z calc$_{mono.}$: 3857.62 Da, observed: 3858.63 Da, error 2.6 ppm), respectively. The identity of **1** and **2** was confirmed by MS^2 (Figure 3-S2, Table 3-S1 and Table 3-S2). These data demonstrate that PurAH is responsible for removing the follower peptide during bottromycin biosynthesis. Since the protein is annotated as a metal-dependent amidohydrolase, we dialyzed purified PurAH extensively with EDTA to remove endogenous metal ions and set up reactions of PurAH with BotA^{PCCD} using different divalent metal ions (Figure 3-S3). No activity was observed for CdCl$_2$ and NiCl$_2$, while addition of FeCl$_2$ or MgCl$_2$ accelerated the background rate slightly. MnCl$_2$ << ZnCl$_2$ < CoCl$_2$ each led to appreciable turnover, with Co^{2+} giving best activity. A time course experiment with 0.2 µM PurAH and 50 µM BotA^{PC} or BotA^{PCCD} revealed that the reaction using BotA^{PCCD} was complete after 80 min, while no product formation was detectable

in the same time frame using BotA^{PC} (Figure 3-3b). PurAH is thus highly selective for the heterocyclized, macrocyclized intermediate BotA^{PCCD} and removes the follower peptide.
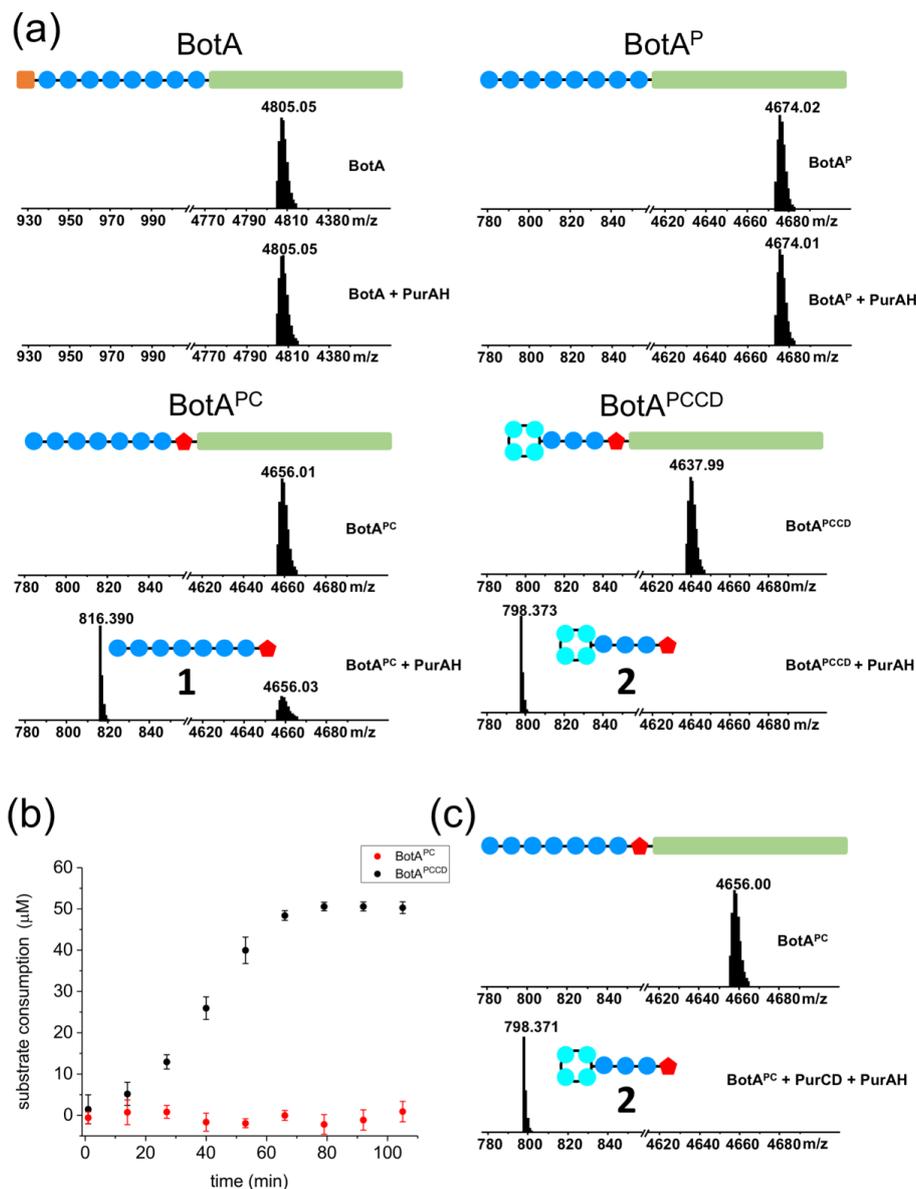
Figure 3-3 | LC-ESI-MS characterization of PurAH reactions. **(a)** Incubation of PurAH with BotA, BotA^P, BotA^{PC}, and BotA^{PCCD} (5 µM enzyme, 50 µM substrate, 37 °C, 16 h). Only BotA^{PC} and BotA^{PCCD} were substrates of PurAH. Colors correspond to scheme 1. **(b)** Time-course of BotA^{PC} and BotA^{PCCD} (50 µM) cleavage by PurAH (0.2 µM) in the presence of Co^{2+} (100 µM). **(c)** When PurAH is added to a macrocyclization reaction, the reaction goes to completion and **2** is formed.

With the function of the amidohydrolase established we investigated its involvement in macroamidine formation. The macroamidine forming YcaO enzyme PurCD converts BotA^{PC} into BotA^{PCCD} in an ATP/Mg^{2+}-dependent reaction for which the BotA follower peptide is essential[8, 9]. We have shown that this reaction is reversible – purified BotA^{PCCD} incubated with PurCD and ATP/MgCl$_2$ will be converted back to BotA^{PC}, now lacking the macrocycle (Figure 3-S4)[8]. Given the very strong preference of PurAH for BotA^{PCCD} and the dependence of PurCD

activity on the presence of the follower peptide that is removed by PurAH, we hypothesized that PurAH may prevent re-opening of the macrocycle. First, we purified **2** and incubated it with PurCD and ATP/MgCl$_2$. Even after extensive incubation times at 37 °C, we did not observe reopening of the macroamidine to yield **1** (Figure 3-S4). BotA^{PCCD} on the other hand converted readily back to BotA^{PC} (Figure 3-S4)When BotA^{PC} is incubated with PurCD under optimized conditions, the turnover to BotA^{PCCD} does not exceed 70%.[8] However, when PurAH is added to the reaction, we observe complete conversion of BotA^{PC} to **2**, indicating complete conversion of BotA^{PC} (macroamidine formation) and subsequent removal of the follower peptide (Figure **3-3**c). These data rationalize why the amidohydrolase was reported as essential for macroamidine formation *in vivo*[13] – knocking out PurAH likely leads to reopening of the macroamidine and prevented downstream processing. This places PurAH as the gatekeeper of bottromycin biosynthesis, that removes the follower peptide once all enzymes requiring the follower have acted on the precursor peptide.

To understand how PurAH selects for the macrocycle in BotA^{PCCD}, we determined the PurAH crystal structure to 1.73 Å resolution. All data collection and refinement statistics can be found in Table 3-S3. The refined model contained one protomer in the asymmetric unit and includes residues 9 – 74 and 78 – 460 (Figure 3-S5). Two Zn^{2+} ions are coordinated at the active site: Zn1 by His210 and His229, Zn2 by His94, His96 and Asp348. Lys183 has been carboxylated as commonly observed in amidohydrolases[15], with each oxygen coordinating one Zn^{2+} (Figure 3-S5). The distorted trigonal bipyramidal coordination of the two Zn^{2+} is completed by ordered water molecules, one of which bridges the two Zn^{2+} ions. This bridging water may be an activated hydroxyl known to facilitate catalysis in amidohydrolases[16-19]. We generated a sequence similarity network (SSN) for PurAH (Figure 3-S6a) and found it to be part of a small node that exclusively contained PurAH homologs from all known bottromycin biosynthetic gene clusters. Mapping the sequences of this node onto the PurAH structure using ConSurf[20-23], we found residues surrounding the active site to be highly conserved (Figure 3-4a). In contrast, the closest structural homologs identified by a DALI[24] search show virtually no sequence conservation at the active site (Figure 3-4b). PurAH contains an extended, wide binding site that may have evolved specifically to accept branched cyclic peptide substrates (Figure 3-S6b).

To rationalize PurAH's selectivity, we sought to determine the complex crystal structure of PurAH with its substrate, but extensive screening failed in delivering suitable conditions. In most RiPP systems the core peptide sequence is at least to some degree variable, but in all bottromycin biosynthetic gene clusters identified to date the core peptide sequence is fully

conserved[25-27] (with the exception of a single-amino acid change found in one cluster[28]). This is at odds with the substrate promiscuity observed for bottromycin biosynthetic enzymes *in vitro* thus far[8, 9, 14]. To probe the promiscuity of PurAH and the fidelity for macrocyclized intermediate, we generated a series of mutant precursor peptides (Figure 3-4c). The degree of conversion was estimated by monitoring the consumption of substrate (see SOI for details). Data for the mutations are provided in Figure 3-S7 – Figure 3-S20. BotA mutants Pro2Ala, Val4Leu, Phe6Trp, and Asp7Asn were found to be good substrates of PurAH (Figure 3-4c). The mutants Val3Leu, Val5Ala, and Asp7Ala were fair and Val4Ala, Val5Glu, and Phe6Ala were poor substrates of PurAH. Asp7Val could not be processed, which may implicate Asp7 in metal coordination. While presence of a heterocycle within the S1 site (Using standard protease nomenclature) of PurAH was essential for activity, the S1' site is more flexible, as demonstrated by partial conversion of BotA mutants Met9Ala and Met9Phe. In all cases where turnover was observed, only the mass for heterocyclized, macrocyclized product could be detected. These data demonstrate that PurAH is promiscuous in the context of single amino-acid changes, while retaining a high fidelity for macrocyclized substrate.
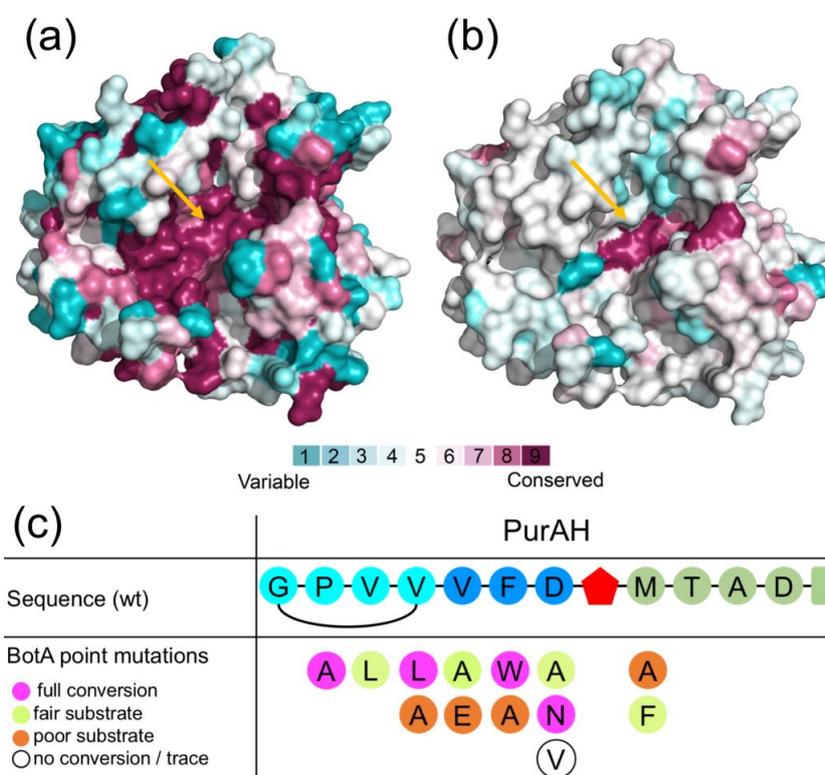
Figure 3-4 | (a) and **(b)** Consurf maps showing the conservation of residues around the PurAH active site (arrow) for (a) PurAH homologs from other bottromycin biosynthetic gene clusters (PurAH containing node, Figure 3-S6(a) and (b) the closest structural homologs identified by a DALI search (Figure 3-S6b). **(c)** Schematic view of the conversion of 13 BotA mutants designed to test PurAH promiscuity.
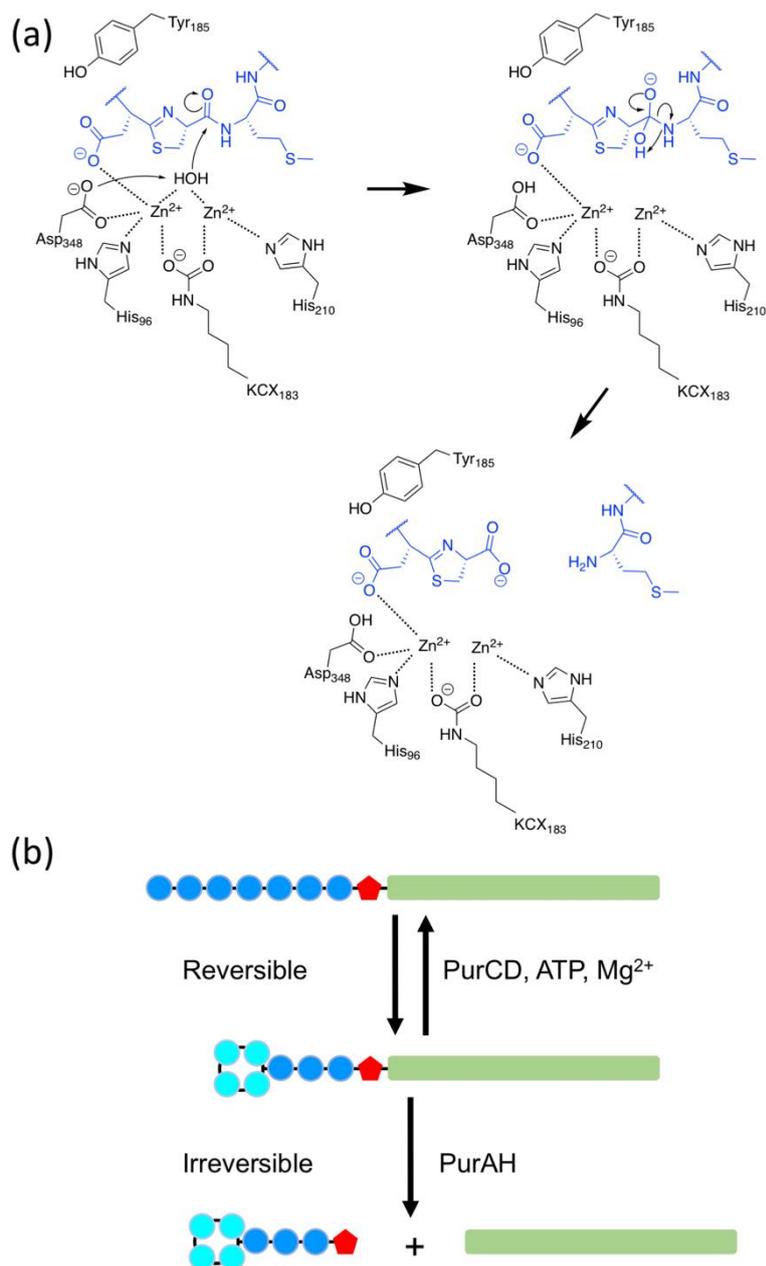
Figure 3-5 | (a) Proposed mechanism of follower peptide removal by PurAH. It is unclear which residue stabilizes the oxyanion. Zn^{2+}-coordinating residues His94 and His229 have been omitted for clarity **(b)** Proposed role of PurAH in bottromycin biosynthesis. Removal of the follower peptide (green) by PurAH prevents reopening of the macroamidine (cyan) and thus drives biosynthesis.

It has been demonstrated for amidohydrolases distantly related to PurAH that the aspartate residue involved in metal coordination at the active site (Asp348 in PurAH) and a tyrosine residue in proximity of the active site were involved in catalysis (e.g.[29]). In PurAH, the mutation Asp348Asn resulted in a loss of catalytic activity (Figure 3-S21). These data may implicate Asp348 as the base in catalysis, but we cannot exclude disruption of metal binding. The only

tyrosine residue in the immediate vicinity of the active-site Zn^{2+}-ions was Tyr185, and in PurAH the mutant Tyr185Phe showed impaired substrate processing (Figure 3-S21).

We have demonstrated that removal of the follower peptide in bottromycin biosynthesis is catalyzed by PurAH. Since the preceding macroamidine formation is reversible, this enzyme has a profound effect on bottromycin biosynthesis by cooperating with the YcaO enzyme PurCD to ensure efficient macrocyclization. This resolves the apparent disparity between *in vitro*[8, 9] and *in vivo*[13] data and places PurAH as the gatekeeper between primary and secondary modification steps (Figure 3-5). It also adds to our understanding of the YcaO superfamily: Thioamide-forming YcaO enzymes appear to require a TfuA protein for activity[30], while YcaO enzymes installing heterocycles in linear azolic peptides require an E1-like protein[31, 32] and their activity appears in some cases coupled to an FMN-dependent oxidase[33, 34]. Despite the cooperation between PurCD and PurAH, we were unable to detect complex formation *in vitro* using various methods (*data not shown*). This raises questions regarding the potential colocalization of RiPP enzymes within the producing organism, which have only been addressed in a very limited number of RiPP systems and require intensive further study. The selectivity of PurAH for macrocyclized precursor peptides coupled with its substrate promiscuity will be important in future efforts to derivatize bottromycins.

3.3 Supporting Information

A

B

C
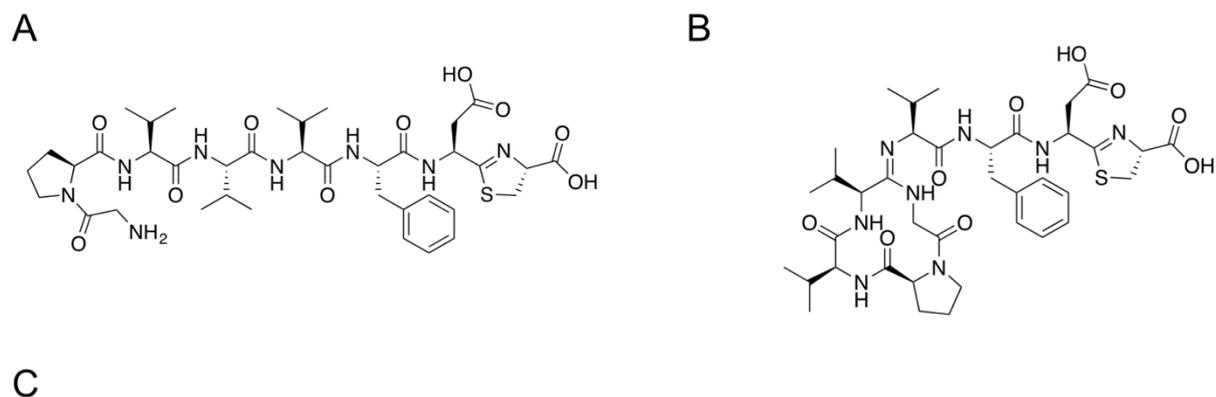
MTADFLNDDPNNAELSALEMEELESWGAWDGEATS

Figure 3-S1 | A Chemical structure of **1** (MW: 816.38 Da)- the modified BotA core peptide after methionine removal (BotP), heterocyclization (IpoC) and follower removal (PurAH). **B** Chemical structure of **2** (MW: 798.37 Da)- the modified BotA core peptide after methionine removal (BotP), heterocyclization (IpoC), macroamidine formation (PurCD) and follower removal (PurAH). **C** Single-letter amino acid sequence of the BotA follower peptide (MW: 3857.62 Da).
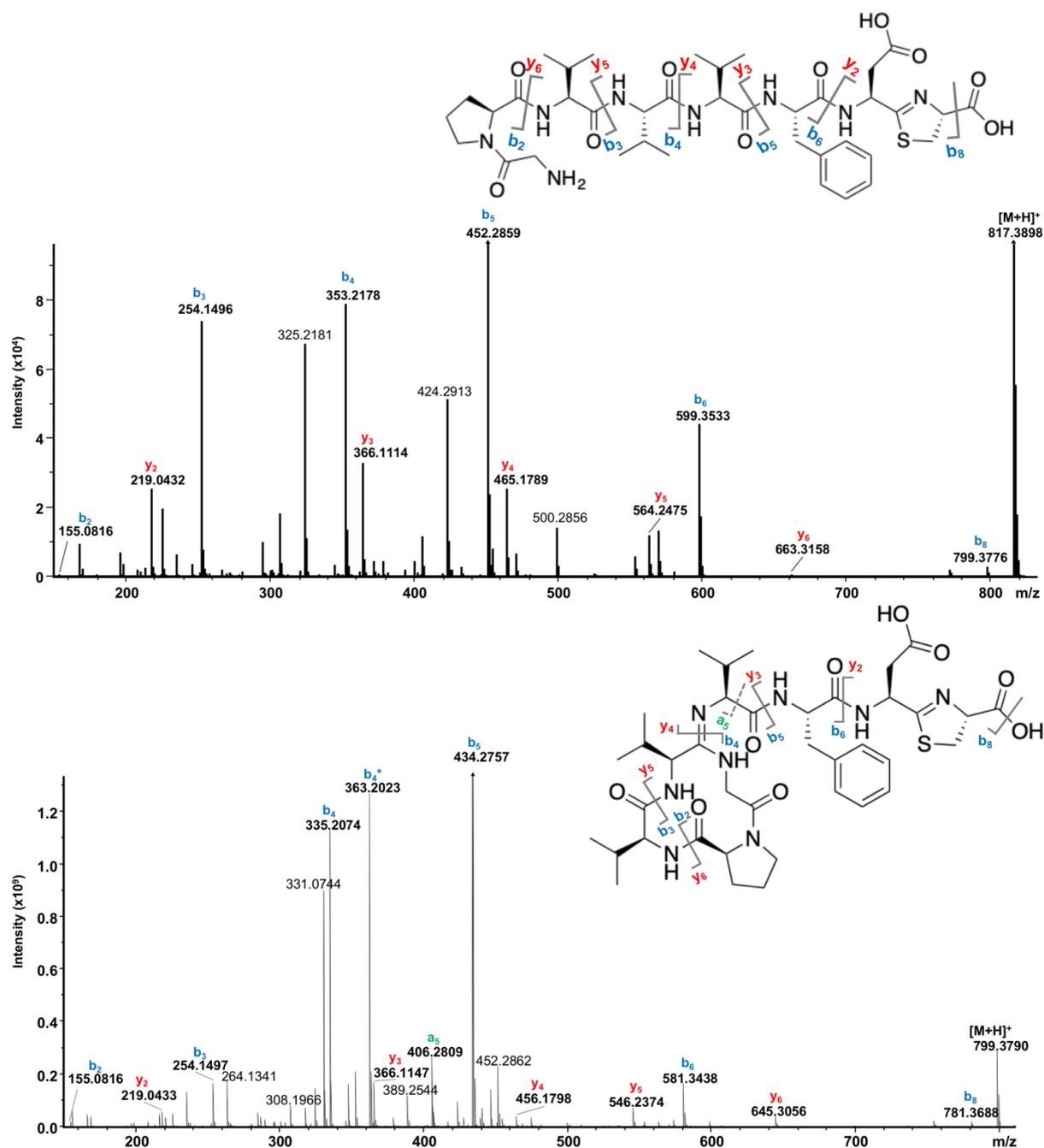
Figure 3-S2 | MS^2 fragmentation spectra for **1** (top) and **2** (bottom). A list of the associated fragments can be found in Table 3-S1 and Table 3-S2, respectively.
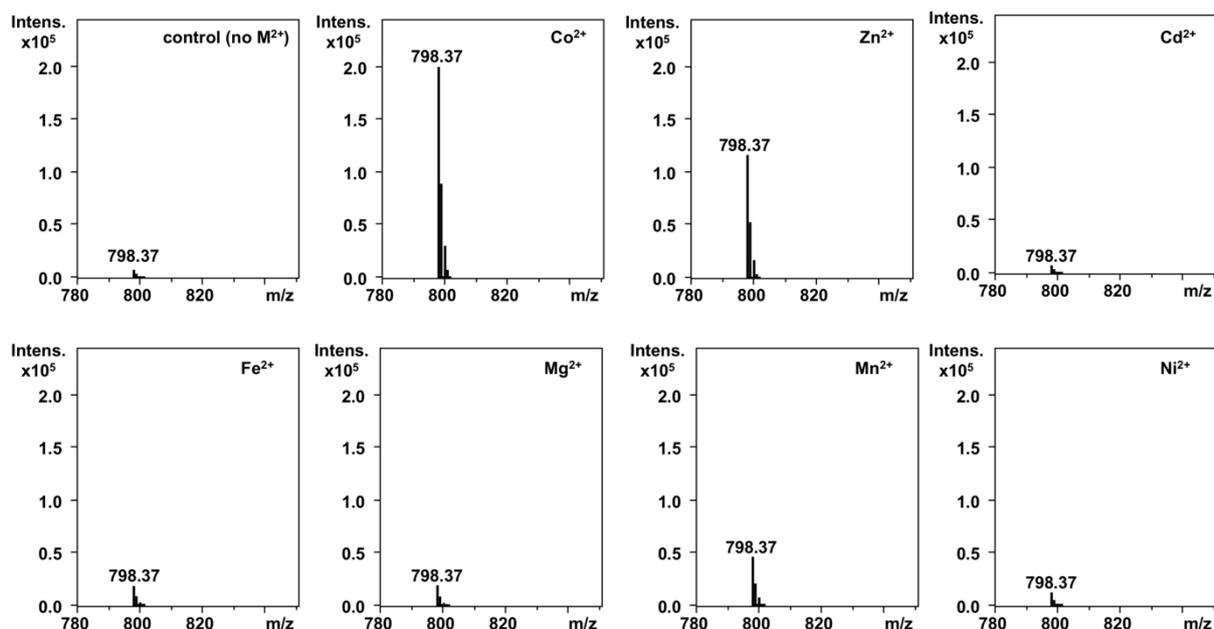
Figure 3-S3 Metal-dependency of PurAH-catalyzed follower removal analysed by LC-MS. As can be seen, addition of Co^{2+} or Zn^{2+} gives greatest activity.
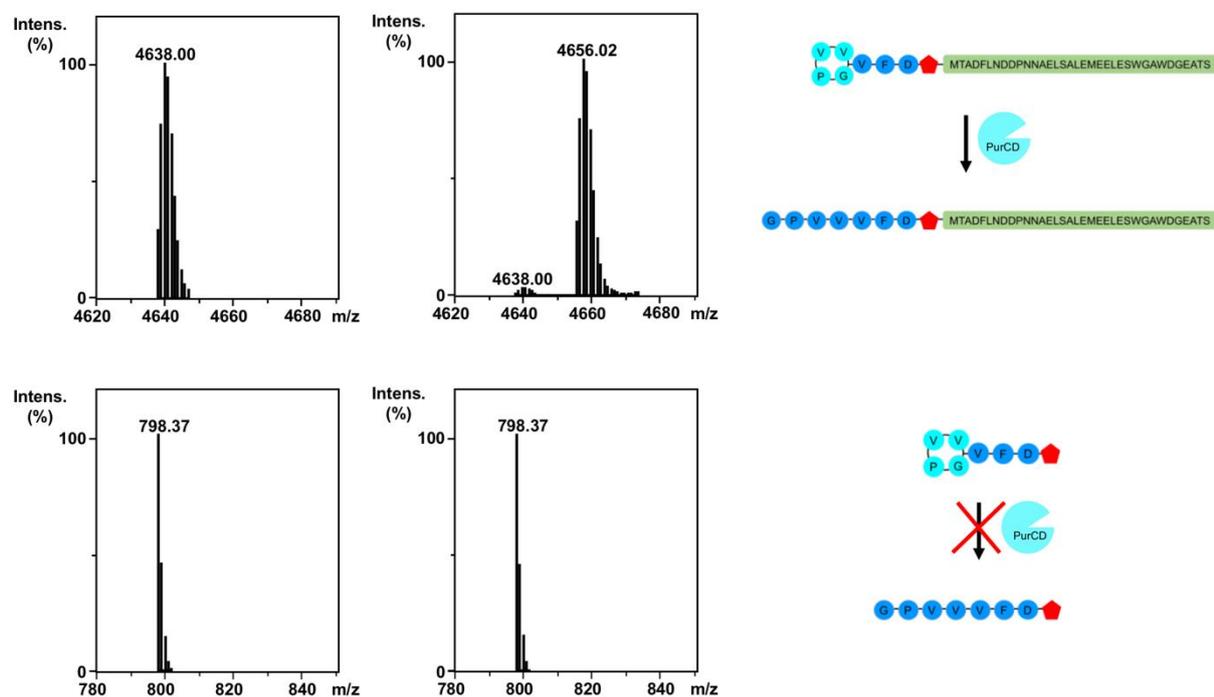
Figure 3-S4 | The PurCD, Mg^{2+} and ATP-dependent reopening of the macroamidine in BotA is follower-dependent. Analysis by LC-MS.
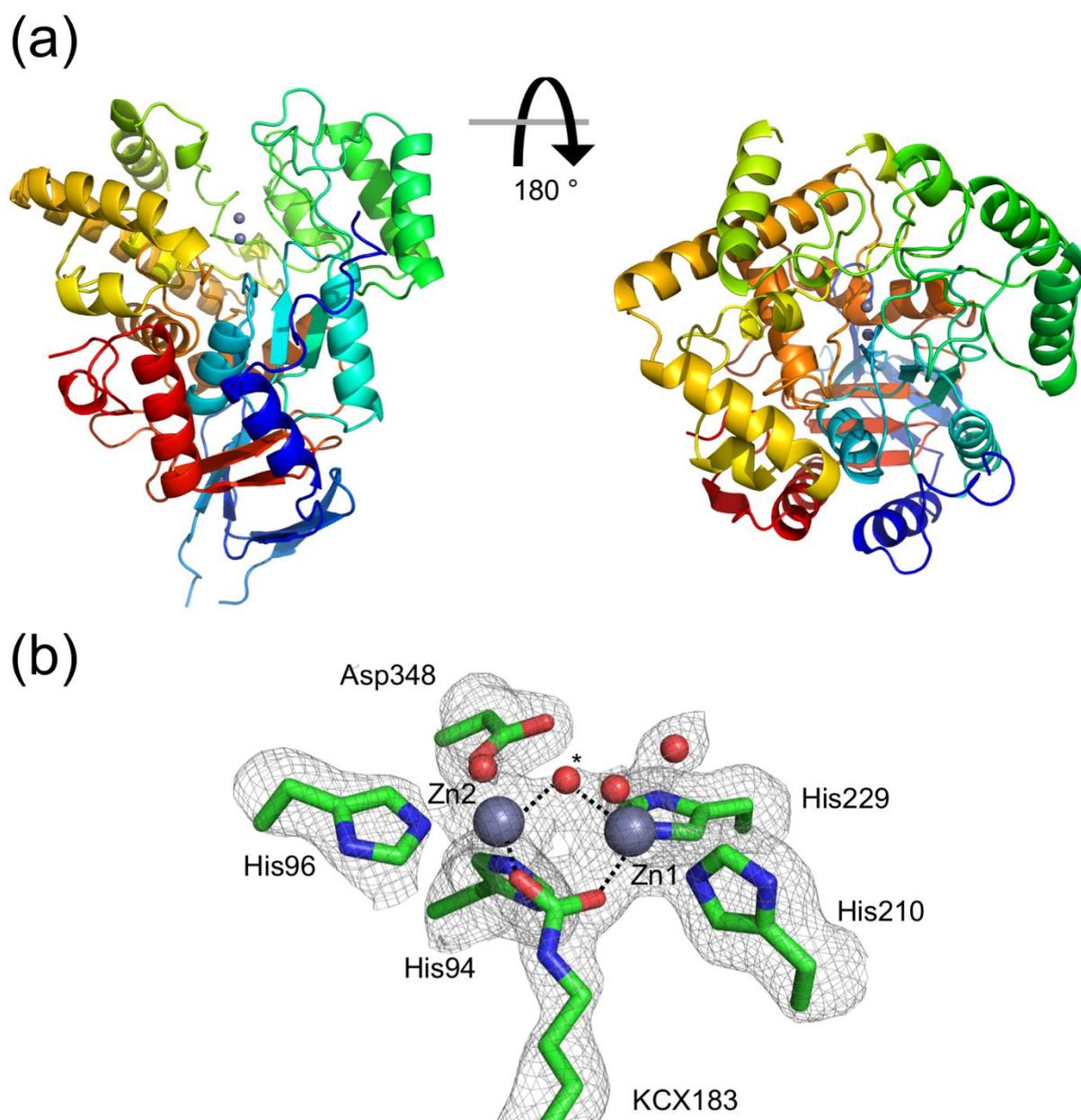
Figure 3-S5 | (a) Overall structure of PurAH. Protein is shown as a cartoon representation using the standard rainbow color scheme, Zn^{2+} ions are shown as grey spheres. **(b)** Coordination of Zn^{2+} ions at the active site. Water molecules are shown as red spheres, selected hydrogen bonds are shown as dashed lines. The bridging water molecule between the two Zn^{2+} ions is marked by an asterix. The difference electron density (Fo – Fc contoured at 3σ with phases calculated from a model which was refined after deletion of the shown residues, ions and waters) is shown as a grey isomesh.
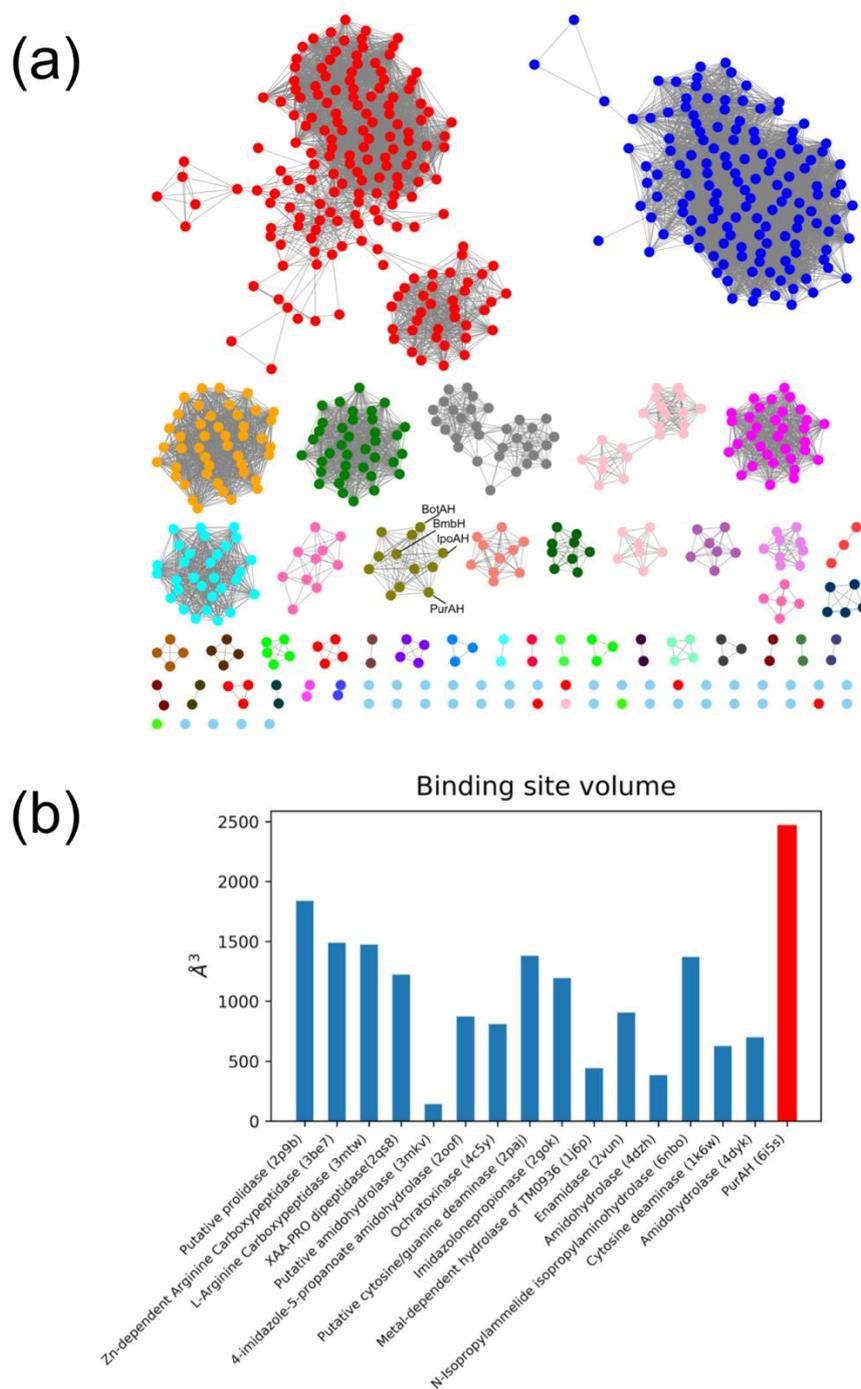
Figure 3-S6 | (a) Sequence similarity network (SSN) generated using the PurAH amino acids sequence. After one iteration of PSI-BLAST[35] the closest 1000 sequences were downloaded (e-value cut-off of 2e-10). The SSN was generated using these sequences on the server of the enzyme function initiative using the enzyme similarity tool[36]. The cut-offs used were sequence maximum/minimum lengths of 650/250 and an alignment score of 80. The SSN was visualized using Cytoscape[37]. PurAH and other sequences from bottromycin gene clusters form a node. No experimentally validated function has been assigned to any of the proteins that are members of other nodes (or singlets). **(b)** Binding pocket volumes determined by POVME 2.0[38] for 15 proteins identified as closest structural homologs of PurAH from a DALI[24] search. PurAH is colored red. The pocket volumes were determined using a gridspacing of 1 Å and applying the 'ConvexHullExclusion' function to remove points outside the binding pocket.
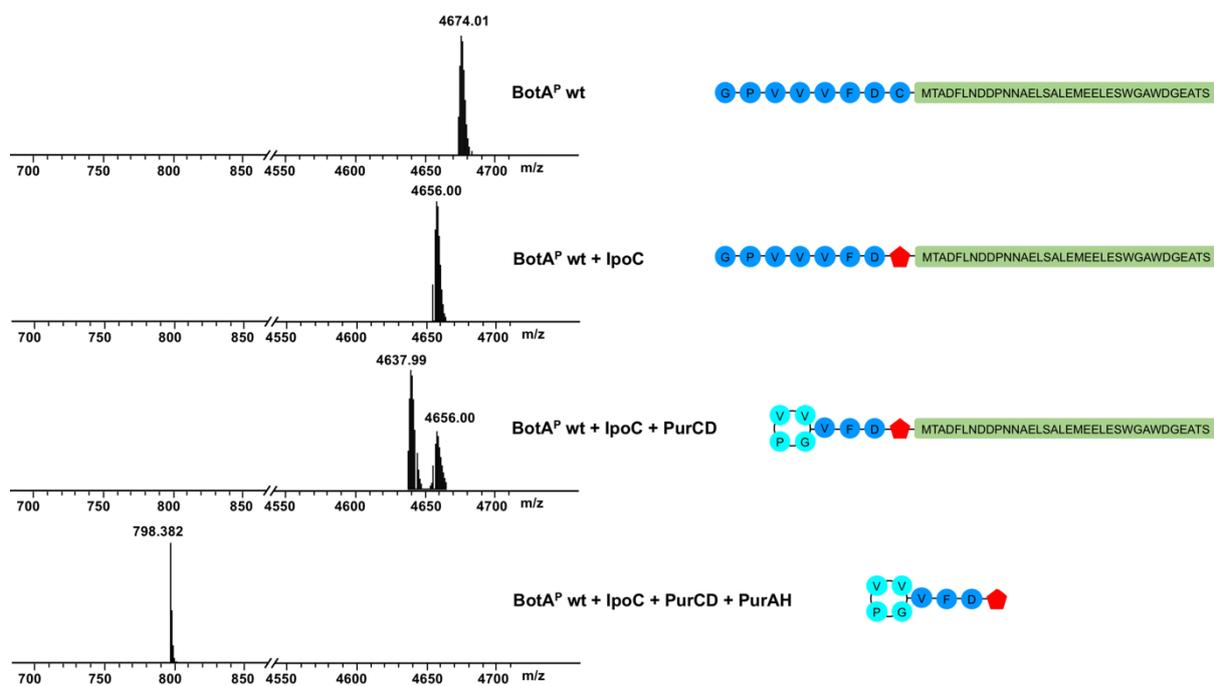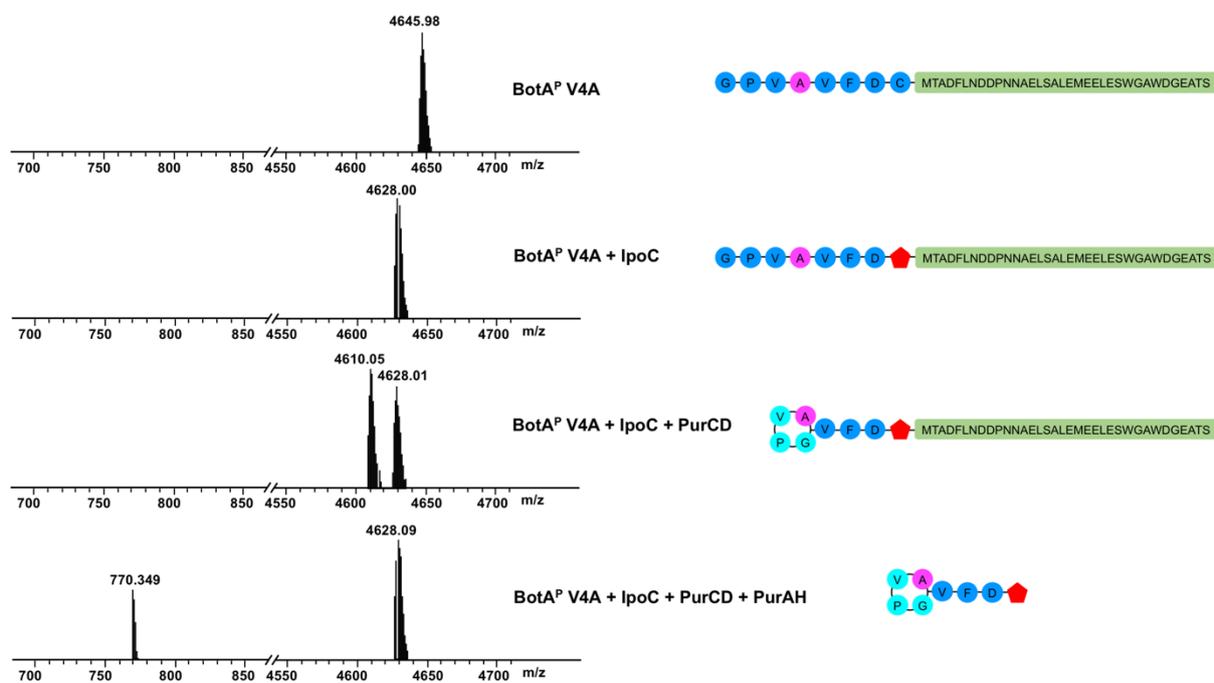
Figure 3-S7 | Processing of BotA^P by IpoC, PurCD and PurAH. To heterocyclize the bottromycin core peptide cysteine to thiazoline we previously reported using the close BotC homolog IpoC from *S. ipomoeae*. The PurCD reaction is not unidirectional but an equilibrium, in which PurCD catalyzes both, macroamidine formation and its reopening. Addition of PurAH cleaves the follower and drives macroamidine formation to completion. Analysis by LC-MS.
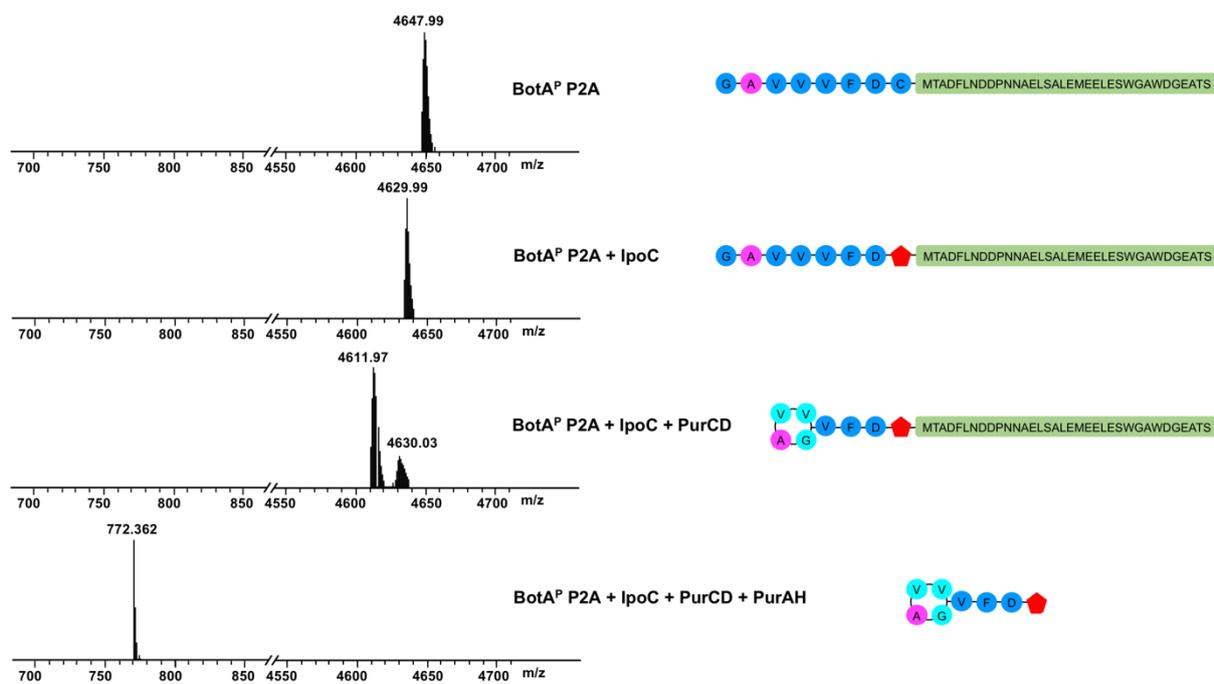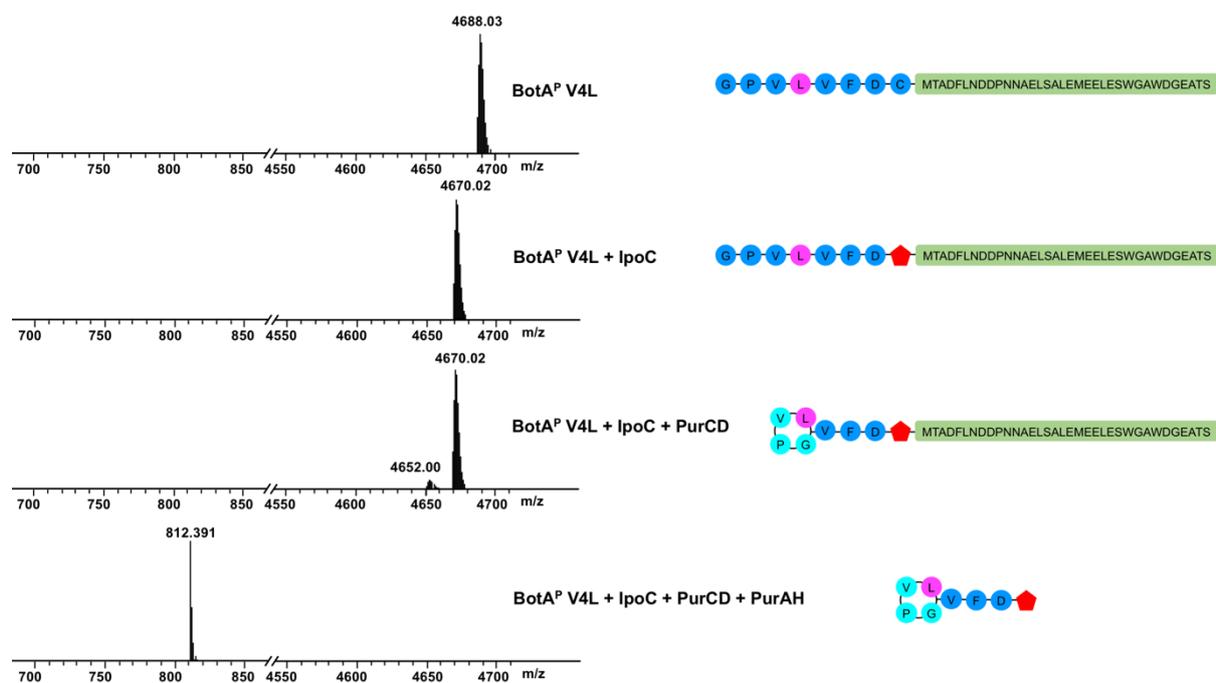

Figure 3-S8 | Processing of BotA^{P,P2A} by IpoC, PurCD and PurAH. Turnover goes complete when PurAH is added (as seen for the wild-type BotA^P). Analysis by LC-MS.
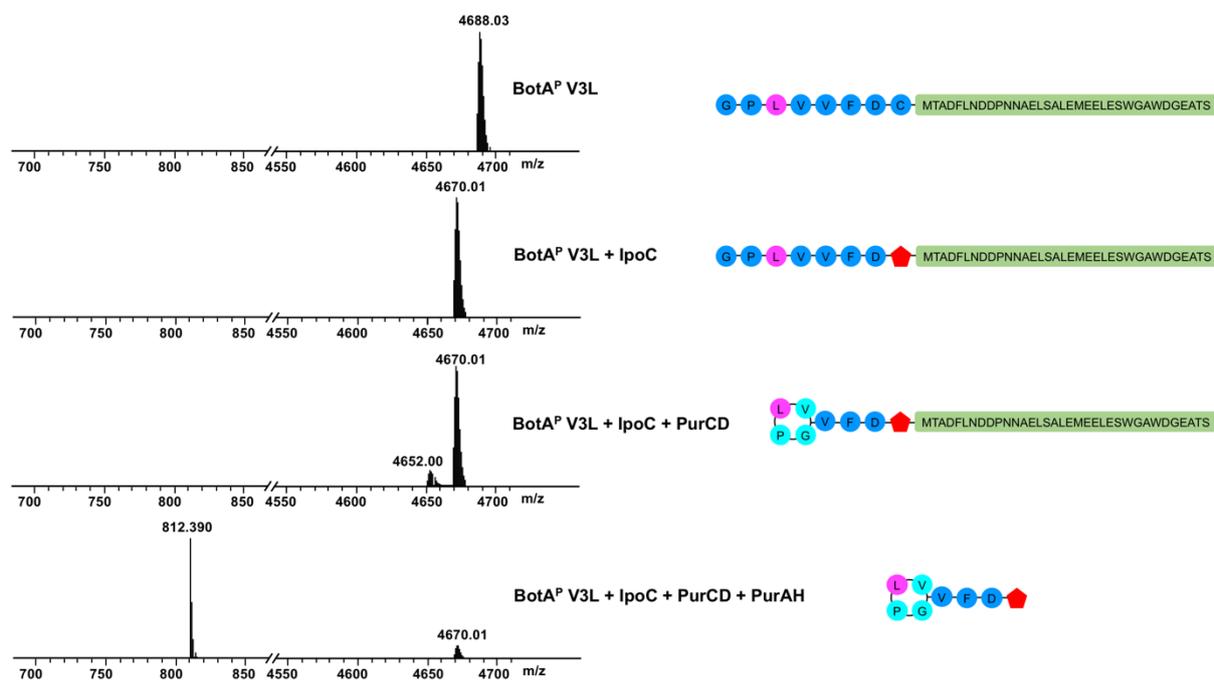
Figure 3-S9 | Processing of BotA^{P,V3L} by IpoC, PurCD and PurAH. Turnover of heterocyclized BotA^{P,V3L} by PurCD is smaller comparted to the wild-type BotA^{PC}. Addition of PurAH leads to the follower cleavage from BotA^{PCCD,V3L}, but turnover by PurAH is not complete. Although turnover of V3L and V4L (see Figure 3-S11) by PurCD are comparable, PurAH is able to drive the BotA^{V4L} reaction to completion. Analysis by LC-MS.
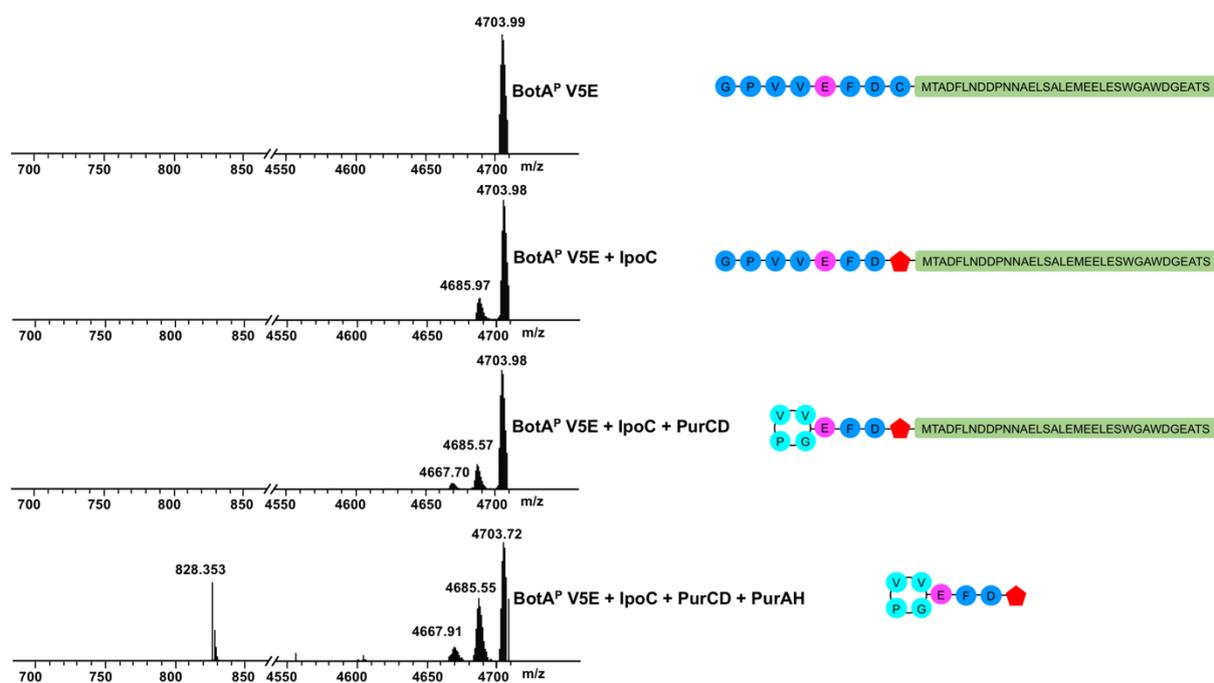

Figure 3-S10 | Processing of BotA^{P,V4A} by IpoC, PurCD and PurAH. Although turnover by PurCD is reasonable, after addition of PurAH only small amounts of processed peptide without the follower are detected. Altering Val4 to Ala seems to slow down PurAH turnover. Analysis by LC-MS.
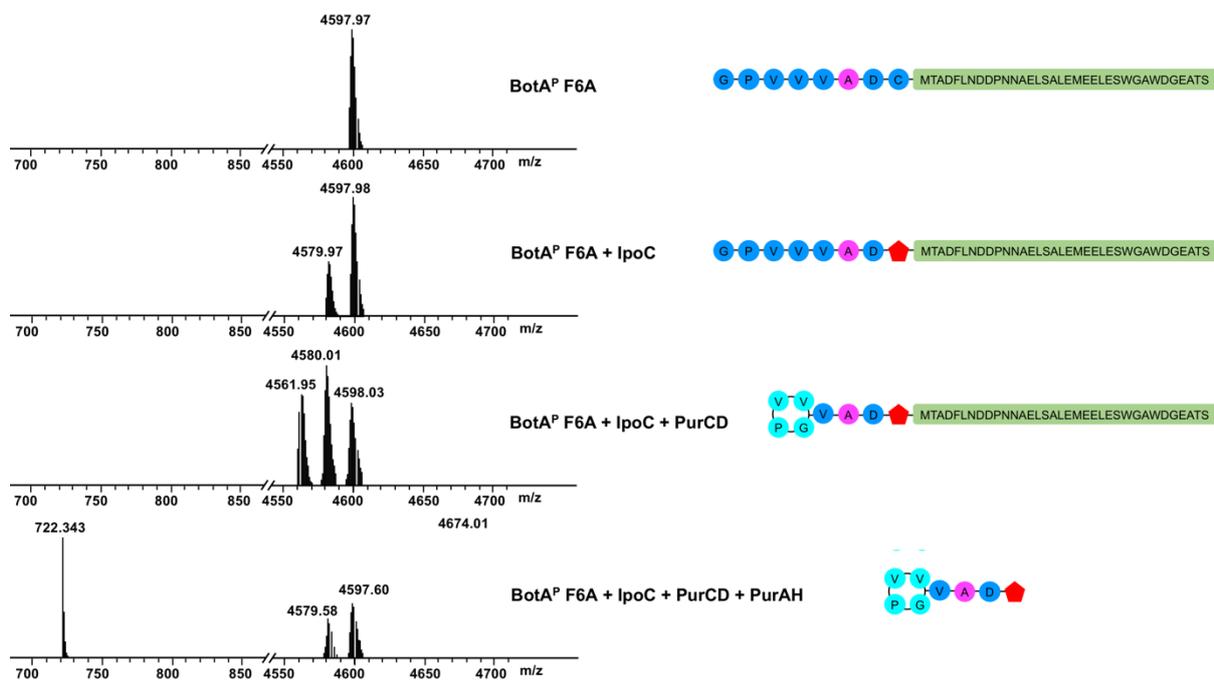
Figure 3-S11 | Processing of BotA[P,V4L] by IpoC, PurCD and PurAH. BotA[P,V4L] is a substrate for IpoC, but the turnover of this heterocyclized BotA variant by PurCD is very small compared to the wild type BotA. Addition of PurAH cleaves the follower and drives macroamidine formation to completion. Analysis by LC-MS.
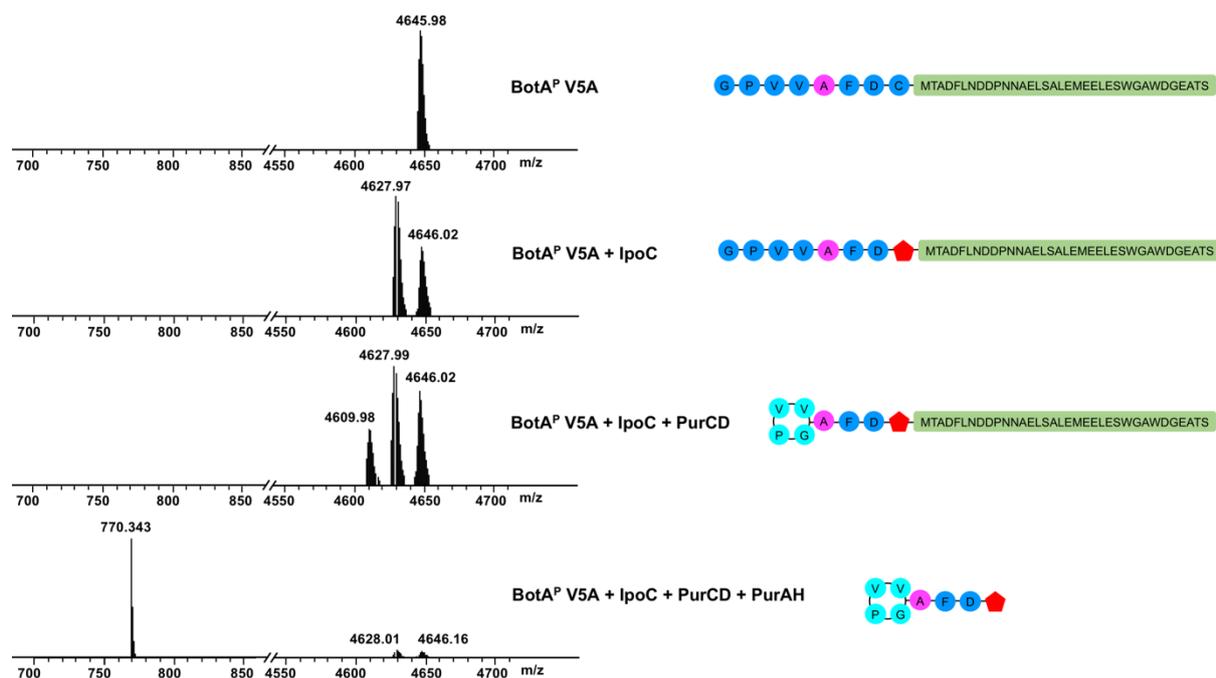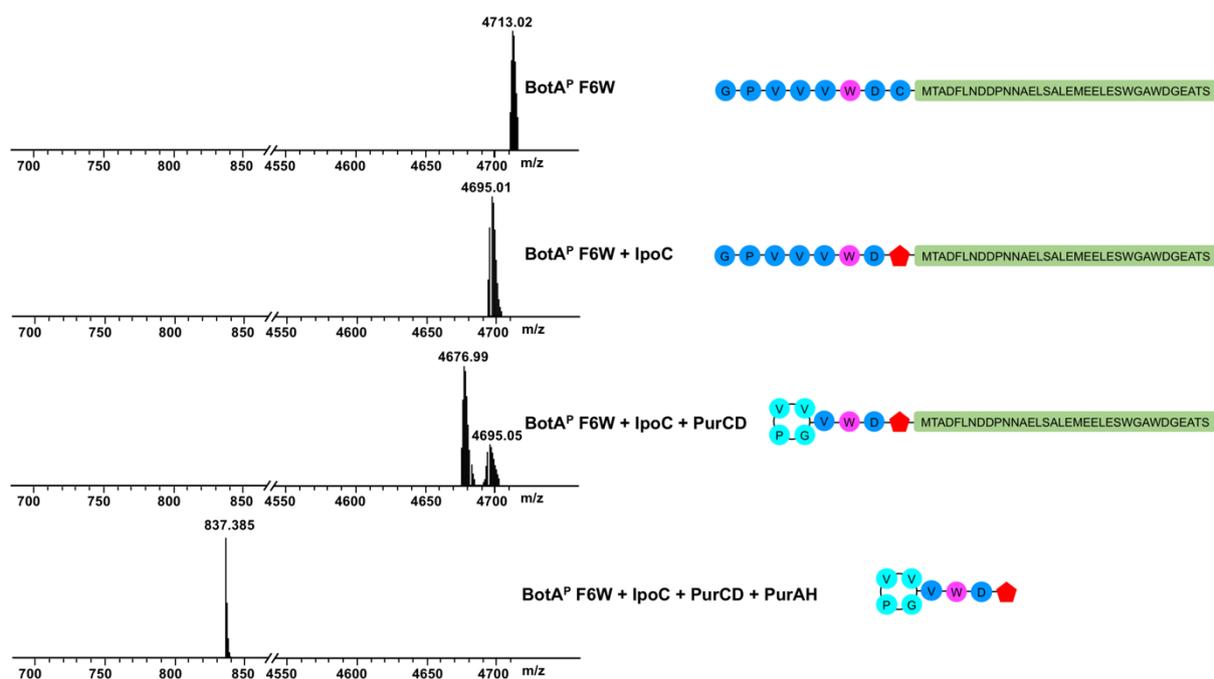

Figure 3-S12 | Processing of BotA[P,V5A] by IpoC, PurCD and PurAH. The turnover of BotA[P,V5A] by IpoC is not complete. As seen for the wild type, the PurCD reaction is not unidirectional and an equilibrium between the processed peptide with and without the macroamidine is observed. Addition of PurAH cleaves the follower, but reaction is not complete. Analysis by LC-MS.
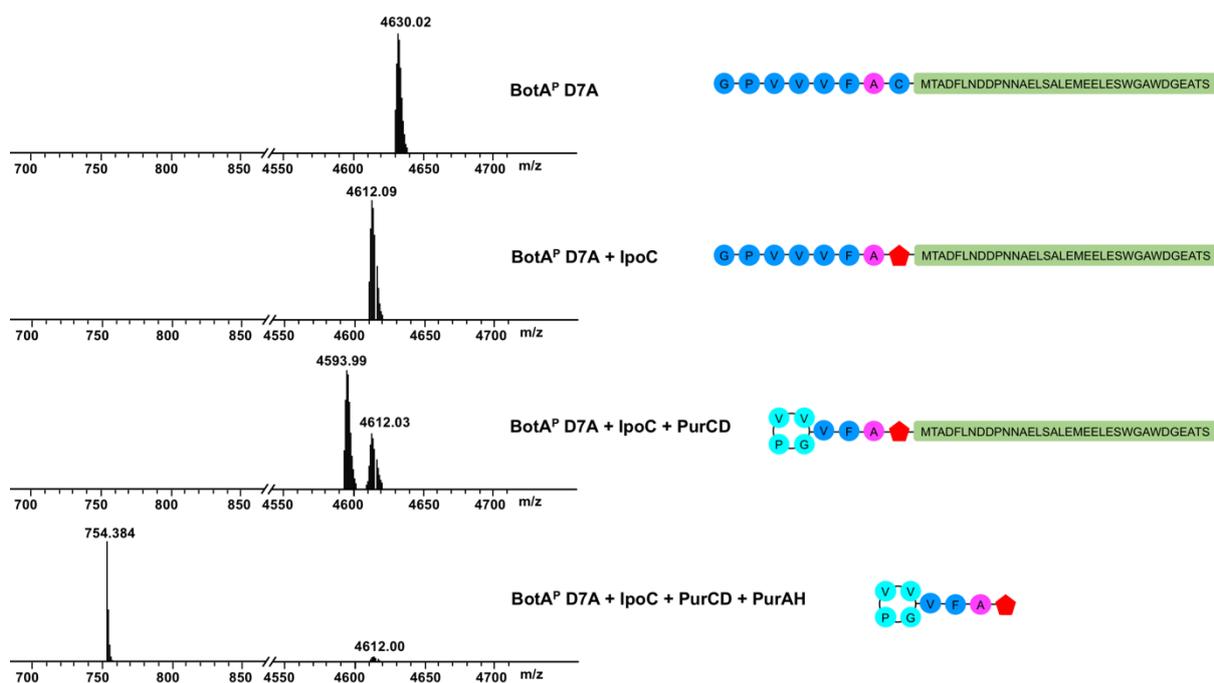
Figure 3-S13 | Processing of BotA[P,V5E] by IpoC, PurCD and PurAH. Turnover of BotA[P,V5E] by IpoC is not complete. After addition of PurAH, only very small amounts of follower-cleaved and processed BotA are detected. Analysis by LC-MS.
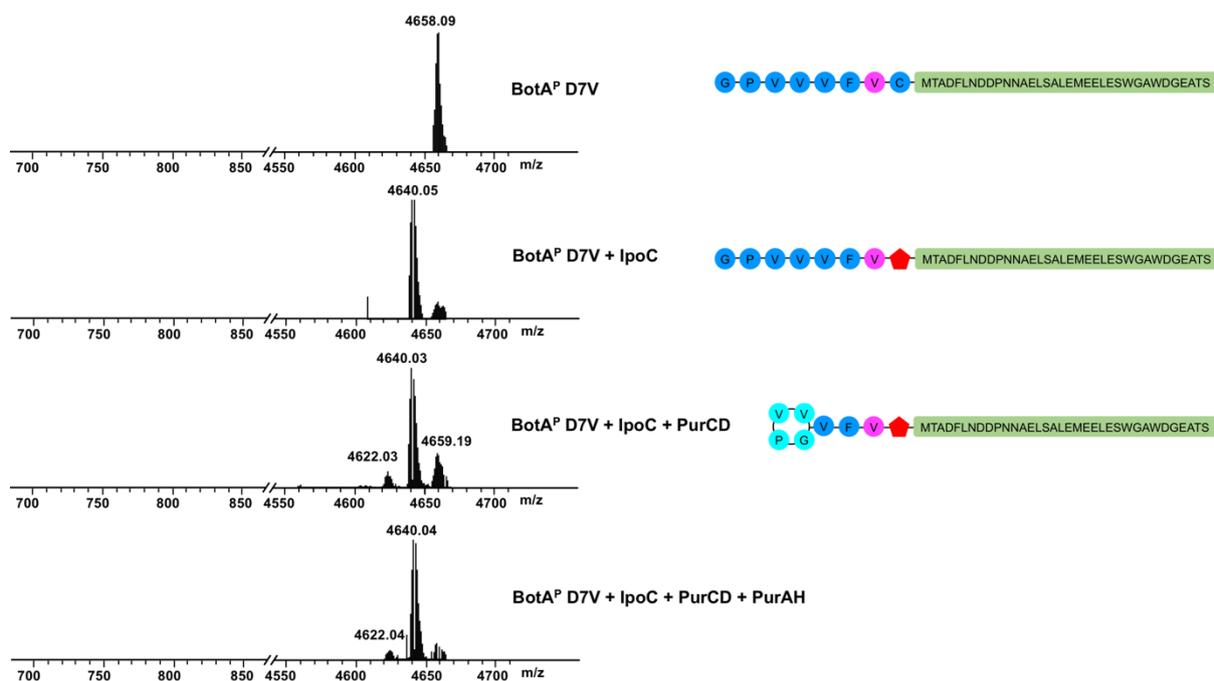

Figure 3-S14 | Processing of BotA[P,F6A] by IpoC, PurCD and PurAH. Turnover of BotA[P,F6A] by IpoC is not complete. After addition of PurAH, only small amounts of follower-cleaved and processed BotA are detected. Analysis by LC-MS.
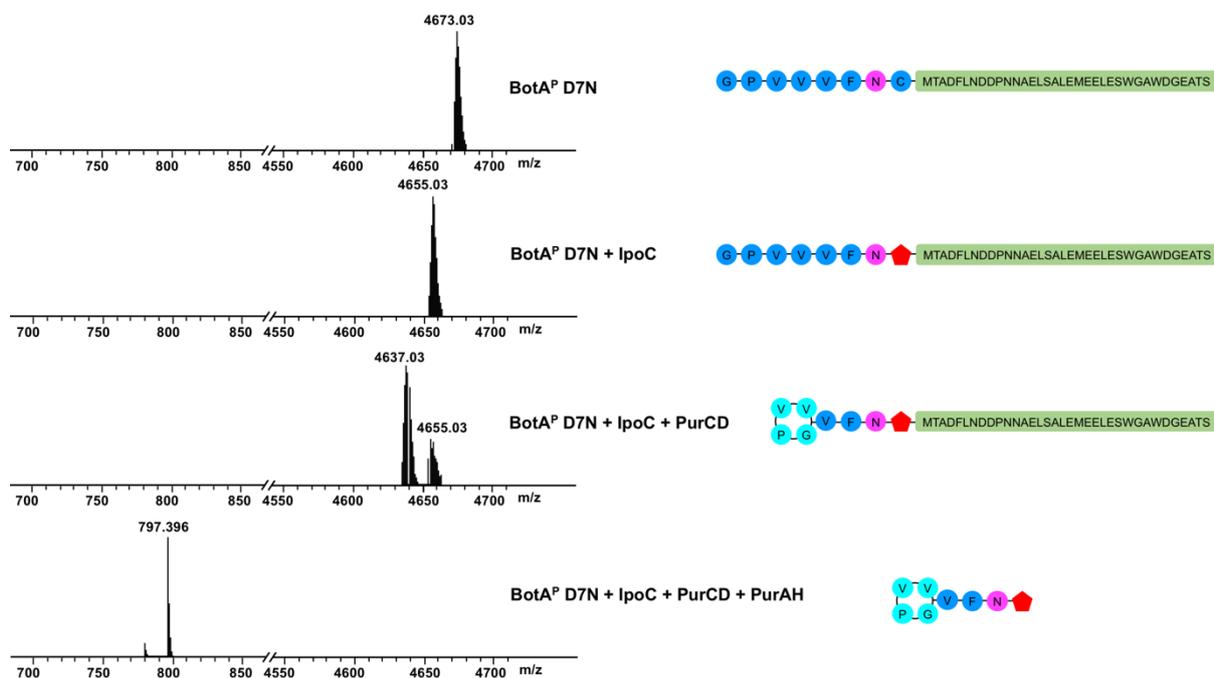
Figure 3-S15 | Processing of BotA[P,F6W] by IpoC, PurCD and PurAH. The PurCD reaction is not unidirectional but an equilibrium, in which PurCD catalyzes both, macroamidine formation and its reopening. Addition of PurAH cleaves the follower and drives macroamidine formation to completion. Analysis by LC-MS.

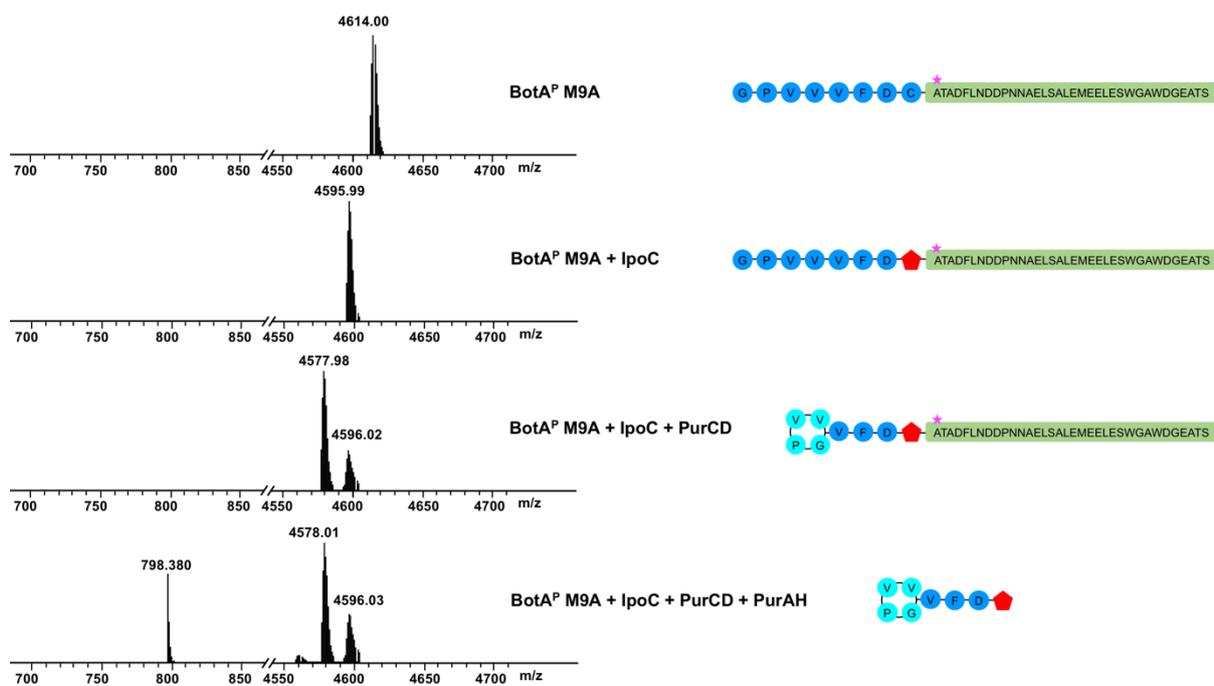**Figure 3-S16** | Processing of BotA[P,D7A] by IpoC, PurCD and PurAH. BotA[P,D7A] is good substrate for IpoC and PurCD (comparable with the wild type BotA[P]). Addition of PurAH cleaves the follower, but does not drive the macroamidine formation to completion (as seen for the wild type BotA[P]). The turnover rate of PurAH for the variant D7A seems be slower than for the wild type. Analysis by LC-MS.

Figure 3-S17 | Processing of BotA^{P,D7V} by IpoC, PurCD and PurAH. Turnover of BotA^{P,D7V} by IpoC is not complete and the turnover of PurCD is lower than for the wild type peptide. After addition of PurAH, no follower-cleaved product was detected. PurAH seems not to be able to use the D7V variant as substrate. Analysis by LC-MS.


Figure 3-S18 | Processing of BotA^{P,D7N} by IpoC, PurCD and PurAH. Turnover goes complete when PurAH is added (as seen for the wild-type BotA). Analysis by LC-MS.

Figure 3-S19 | Processing of BotA^(P,M9A) by IpoC, PurCD and PurAH. BotA^(P,M9A) is a substrate for IpoC and PurCD (comparable to the wild type BotA^P). When PurAH is added, only small amounts of follower-cleaved peptide are detected. The turnover rate of PurAH for the processed variant is much lower than those for the wild type peptide. Analysis by LC-MS.
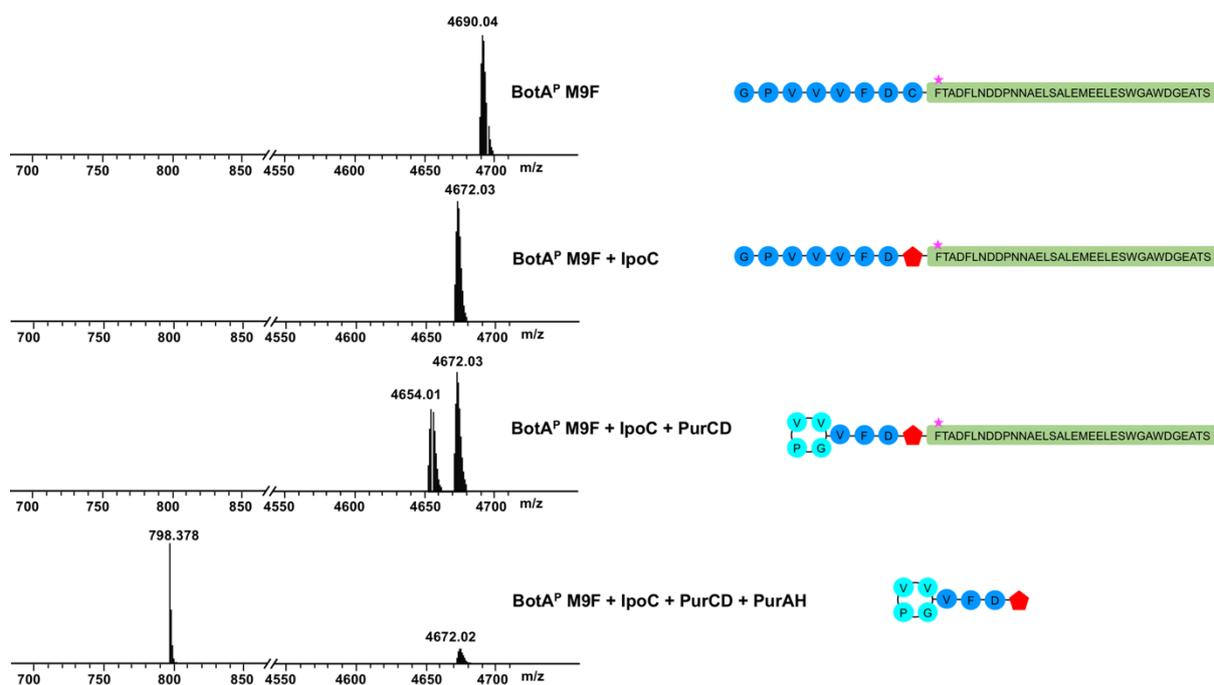

Figure 3-S20 | Processing of BotA^(P,M9F) by IpoC, PurCD and PurAH. BotA^(P,M9F) is a substrate for IpoC and PurCD (comparable to the wild type BotA^P). Addition of PurAH cleaves the follower and drives macroamidine formation to nerly to completion. The turnover rate of PurAH for the processed BotA^(P,M9F) is slower than thos wor the wild type BotA^P. Analysis by LC-MS.

119
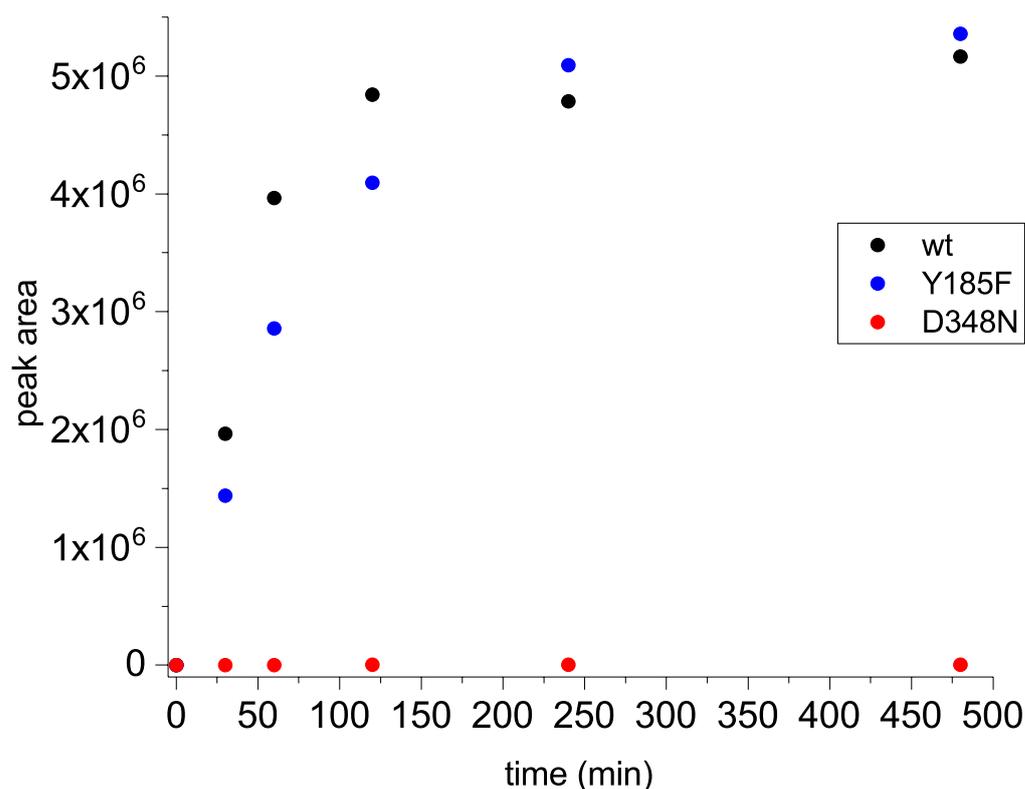
Figure 3-S21 | Time dependent formation of **2** (integrated peak areas of $[M+H]^+=799.3807 \pm 0.05$ Da) by PurAH wt and mutants (Y185F, D348N) measured by LC-ESI-MS.
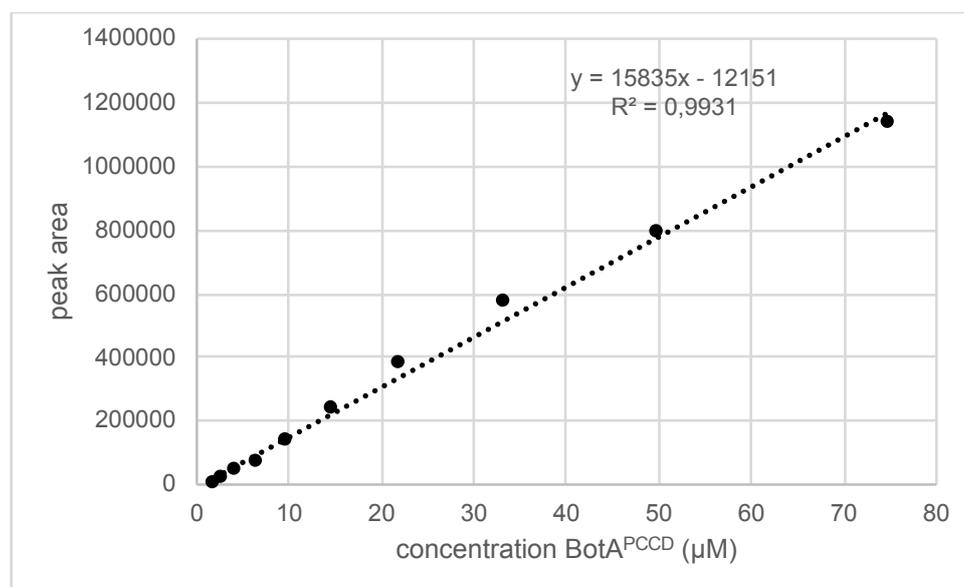
Figure 3-S22 | Calibration curve for BotA^{PCCD} for the concentration range between 2 and 75 µM. Peak areas for the $[M+H]^{4+}$,$[M+H]^{5+}$ and $[M+H]^{6+}$ ions (EIC width $\pm$ 5 ppm) of BotA^{PCCD} (see Table 3-S4), measured by LC-ESI-MS, were plotted against the concentration of purified BotA^{PCCD} peptide.
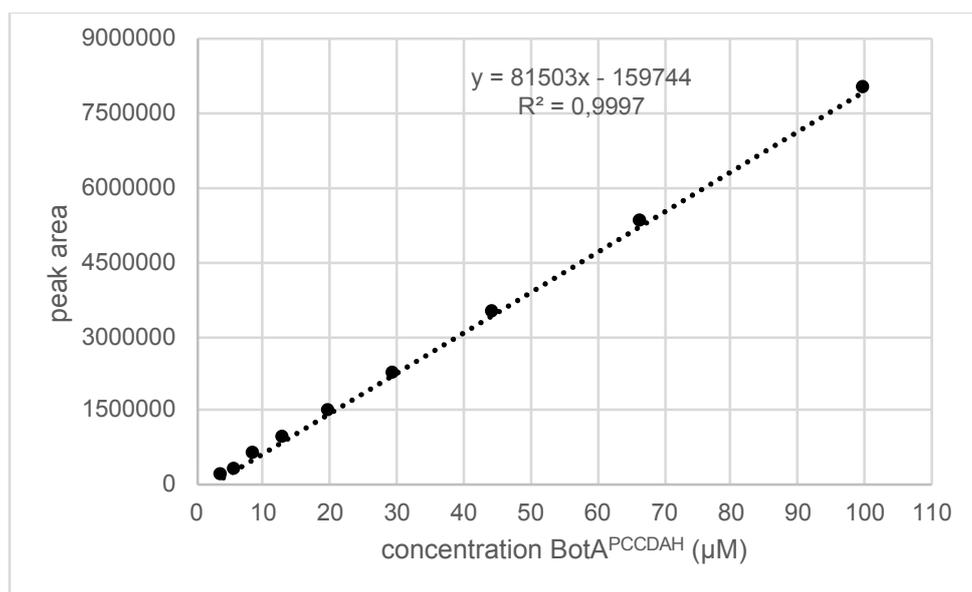
Figure 3-S23 | Calibration curve for **2** for the concentration range between 2.6 and 100 µM. The peak area for the $[M+H]^+$ ion of **2** ($[M+H]^+=799.3815 \pm 0.05$ Da), measured by LC-ESI-MS, was plotted against the concentration of purified **2**.

Table 3-S1 | MS^2 fragmentation data for **1**, corresponding to Figure 3-S2.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G	1	-	-	8	817.3898	817.3918
P	2	155.0816	155.0821	7	-	-
V	3	254.1496	254.1605	6	663.3158	663.3176
V	4	353.2178	353.2189	5	564.2475	564.2492
V	5	452.2859	452.2873	4	465.1789	465.1808
F	6	599.3533	599.3557	3	366.1114	366.1124
D^b	7	-	-	2	219.0432	219.0439
C^b	8	799.3776	799.3813	1	-	-

    b.   Cys8 heterocyclization with Asp7 carbonyl

Table 3-S2 | MS^2 fragmentation data for **2**, corresponding to Figure 3-S2.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G^a	1	-	-	8	799.3790	799.3807
P	2	155.0816	155.0821	7	-	-
V	3	254.1497	254.1605	6	645.3056	645.3070
V^a	4	335.2074	335.2083	5	546.2374	546.2386
V	5	434.2757	434.2767	4	465.1798	465.1808
F	6	581.3438	581.3451	3	366.1147	366.1124
D^b	7	-	-	2	219.0433	219.0439
C^b	8	781.3688	781.3707	1	-	-

    a.   Gly1 macrocyclisation with Val4 carbonyl
    b.   Cys8 heterocyclisation with Asp7 carbonyl

Additionally, fragment a_5 (obs. 406.2809; calc. 406.2818) and b_4* (obs. 363.2023; calc. 363.2032), which is characteristically associated with the bottromycin macrocycle and has been previously observed for bottromycins and macrocyclic peptides, was detected.

Table 3-S3 | Data collection and refinement statistics.

	PurAH high red	PurAH high res
Data collection		
Space group	C 2 2 21	C 2 2 21
Cell dimensions		
a, b, c (Å)	100.98 110.93 84.50	100.97 110.73 84.90
α, β, γ (°)	90.00 90.00 90.00	90.00 90.00 90.00
Resolution (Å)	43.34 - 2.50	74.61 - 1.73
	(2.60 – 2.50)	(1.76 - 1.73)
R_{merge}	0.129 (0.517)	0.034 (0.693)
$I / \sigma I$	55.7 (20.2)	18.4 (1.9)
Completeness (%)	99.8 (99.9)	98.1 (99.6)
Redundancy	92.7 (89.8)	4.5 (4.5)
Refinement		
Resolution (Å)		56.04 - 1.73
No. reflections		48843 (4935)
R_{work} / R_{free}		0.1675 / 0.1956
No. atoms		3619
Protein		3403
Ligand/ion		9
Water		207
B-factors		41.04
Protein		40.82
Ligand/ion		51.25
Water		44.25
R.m.s. deviations		
Bond lengths (Å)		0.014
Bond angles (°)		1.10
Molprobity clash score		0.89

*Values in parentheses are for highest-resolution shell.

Table 3-S4 | Calculated masses for the BotA^{PC} and BotA^{PCCD} (mutant) peptides and [M+H]$^{4+}$, [M+H]$^{5+}$ and [M+H]$^{6+}$ions, which were used for the peak area calculation.

peptide	Formula	M$_{mono.}$	[M+H]$^{4+}$	[M+H]$^{5+}$	[M+H]$^{6+}$
BotA^{PC} wt	$C_{202}H_{294}N_{48}O_{73}S_3$	4655.990	1165.006	932.206	777.006
BotA^{PCC} wt	$C_{202}H_{292}N_{48}O_{72}S_3$	4637.983	1160.503	928.604	774.004
BotA^{PC} P2A	$C_{200}H_{292}N_{48}O_{73}S_3$	4629.977	1158.502	927.003	772.670
BotA^{PCCD} P2A	$C_{200}H_{290}N_{48}O_{72}S_3$	4611.967	1156.999	925.801	771.668
BotA^{PC} V3L	$C_{203}H_{296}N_{48}O_{73}S_3$	4670.009	1168.509	935.009	779.342
BotA^{PCCD} V3L	$C_{203}H_{294}N_{48}O_{72}S_3$	4651.998	1164.007	931.407	776.340
BotA^{PC} V4L	$C_{203}H_{296}N_{48}O_{73}S_3$	4670.009	1168.509	935.009	779.342
BotA^{PCCD} V4L	$C_{203}H_{294}N_{48}O_{72}S_3$	4651.998	1164.007	931.407	776.340
BotA^{PC} V4A	$C_{200}H_{290}N_{48}O_{73}S_3$	4627.962	1157.998	926.600	772.334
BotA^{PCCD} V4A	$C_{200}H_{288}N_{48}O_{72}S_3$	4609.951	1153.495	922.998	769.332
BotA^{PC} V5A	$C_{200}H_{290}N_{48}O_{73}S_3$	4627.962	1157.998	926.600	772.334
BotA^{PCCD} V5A	$C_{200}H_{288}N_{48}O_{72}S_3$	4609.951	1153.495	922.998	769.332
BotA^{PC} V5E	$C_{202}H_{292}N_{48}O_{75}S_3$	4685.967	1172.499	938.201	782.002
BotA^{PCCD} V5E	$C_{202}H_{290}N_{48}O_{74}S_3$	4667.957	1167.996	934.599	779.000
BotA^{PC} F6W	$C_{204}H_{295}N_{49}O_{73}S_3$	4695.004	1174.758	940.008	783.508
BotA^{PCCD} F6W	$C_{204}H_{293}N_{49}O_{72}S_3$	4676.993	1170.256	936.406	780.506
BotA^{PC} F6A	$C_{196}H_{290}N_{48}O_{73}S_3$	4579.962	1145.998	917.000	764.334
BotA^{PCCD} F6A	$C_{196}H_{288}N_{48}O_{72}S_3$	4561.951	1141.495	913.398	761.332
BotA^{PC} D7A	$C_{201}H_{294}N_{48}O_{71}S_3$	4612.003	1154.01	923.408	769.674
BotA^{PCCD} D7A	$C_{201}H_{292}N_{48}O_{70}S_3$	4593.993	1149.505	919.806	766.673
BotA^{PC} D7N	$C_{202}H_{295}N_{49}O_{72}S_3$	4655.009	1164.760	932.009	776.842
BotA^{PCCD} D7N	$C_{202}H_{293}N_{49}O_{71}S_3$	4636.999	1160.257	928.407	773.840
BotA^{PC} D7V	$C_{203}H_{298}N_{48}O_{71}S_3$	4640.035	1161.016	929.014	774.346
BotA^{PCCD} D7V	$C_{203}H_{296}N_{48}O_{70}S_3$	4622.024	1156.513	925.412	771.345
BotA^{PC} M9A	$C_{200}H_{290}N_{48}O_{73}S_2$	4595.990	1150.005	920.205	767.006
BotA^{PCCD} M9A	$C_{200}H_{288}N_{48}O_{72}S_2$	4577.979	1145.502	916.603	764.004
BotA^{PC} M9F	$C_{206}H_{294}N_{48}O_{73}S_2$	4672.021	1169.013	935.411	779.677
BotA^{PCCD} M9F	$C_{206}H_{292}N_{48}O_{72}S_2$	4654.010	1164.510	931.809	776.676

Table 3-S5 | Nucleotide sequences used for the site-directed mutagenesis (SDM) of PurAH (TS: top strand; BS: bottom strand).

Primer name	Primer sequence (5' to 3')
PurAH Y185F TS	GGCATCAAGGTTTTCGACGACGTTGAG
PurAH Y185F BS	CTCAACGTCGTCGAAAACCTTGATGCC
PurAH D348N TS	CTGGTCGGTACCAACTGTGGCAACCCG
PurAH D348N BS	CGGGTTGCCACAGTTGGTACCGACCAG

Material & Methods

Cloning, protein expression and purification

<u>Cloning, Expression and Purification of PurCD and IpoC</u>

PurCD and IpoC were purified as described previously.[8]

<u>Cloning of PurAH wt and mutants</u>

Full-length, codon-optimized *purAH* (Accession No. WP_019887083.1, Eurofins Genomics) was cloned into pET-28b (Novagen). Presence of the insert was confirmed by restriction digestion and DNA sequencing before being transformed into Rosetta™ (DE3). Mutations that lead to the exchange of one amino acid in the PurAH sequence were introduced with the site-directed mutagenesis (SDM) primers shown in Table 3-S5 using the reagents and protocol of the Quick Change II kit (Agilent Technologies). The introduction of the mutations was verified by DNA sequencing.

<u>Cloning, Expression and Purification BotA and BotA point mutants</u>

BotA, BotA[P] and BotA[P] variants were cloned, expressed and purified as described previously.[8] BotA[P] variants with one altered amino acid were generated by using mutation carrying primers. Peptide concentrations were determined at a Spectrophotometer (NanoDrop2000, Thermo Fisher Scientific) using the extinction coefficient calculated from the amino acid sequence by the ExPASY ProtParam Server[39] (http://web.expasy.org/protparam/).

<u>Protein expression and purification of PurAH wt and mutants</u>

A single colony was picked into LB liquid medium containing the appropriate antibiotics (50 µg/mL Kanamycin and / or 34 µg/mL Chloramphenicol) to make an overnight culture. The overnight culture was inoculated 1 to 100 into fresh LB medium supplemented with antibiotics and was grown at 37 °C until the optical density (OD_{600}) reached 0.6; protein expression was then induced by the addition of 1 mM IPTG and the cells grown at 16 °C overnight. Cells were harvested by centrifuging the culture at 6,000x g at 4 °C for 15 min and the cell pellets frozen at -80 °C.

The cell pellets were resuspended in lysis buffer (20 mM Tris pH 8.0, 200 mM NaCl, 20 mM imidazole, 10% glycerol (w/v) and 3 mM DTT) supplemented with 0.4 mg DNAse per gram of wet cell pellet and cOmplete EDTA-free protease-inhibitor tablets (Roche). The cells were lysed by sonication and cell debris removed by centrifugation (40,000x g, 4 °C, 20 min). The

supernatant was loaded onto a pre-equilibrated (lysis buffer) 5 mL Histrap HP column (GE Healthcare) and was washed with 20 column volumes of lysis buffer. The protein was eluted from the column with 250 mM imidazole and loaded onto a gel filtration column (HiLoad 16/600 Superdex 200 pg, GE Healthcare) pre-equilibrated in 50 mM Tris-Cl, pH 8.5, 200 mM NaCl, 10% (w/v) glycerol and 1mM TCEP. The fractions of the highest purity were pooled together and concentrated to 5-8 mg mL^{-1}.

Biochemistry

Heterocyclization reaction by IpoC

For heterocyclisation reactions by IpoC, 50 µM substrate (BotA^P or BotA^P variants) was incubated with 5 µM IpoC and 5 mM ATP/MgCl$_2$ in reaction buffer (50 mM Tris pH 8.5, 200 mM NaCl, 10% glycerol) for 12 h at 37 °C.

For further use, heterocyclized BotA^P (BotA^{PC}) was immediately purified by size-exclusion chromatography on a pre-equilibrated Superdex S30 column (GE Healthcare) using reaction buffer and concentrated using 5 kDa cutoff filters (Merk MilliPore).

Reactions of BotA^P variants with IpoC were set up and analyzed in triplicates. Turnover was analyzed by LC-MS. For LC-MS analyses, reactions were stopped and enzymes were precipitated by the addition of ACN.

Macrocyclization reaction by PurCD

For macrocyclisation reactions by PurCD, 50 µM substrate (BotA^{PC}) was incubated with 5 µM PurCD and 10 mM ATP/MgCl$_2$ in reaction buffer (50 mM Tris pH 8.5, 200 mM NaCl, 10% glycerol) for 2 h at room temperature. Macroamidine formation was determined by LC-MS analysis.

Product (BotA^{PCCD}) was immediately purified for further experiments by size-exclusion chromatography on a pre-equilibrated Superdex S30 column (GE Healthcare) using reaction buffer. BotA^{PCCD} was concentrated using 5 kDa cutoff filters (Merck Millipore). As we observe macroamidine re-opening by PurCD, the purification product is a mixture of BotA^{PC} and BotA^{PCCD}. Thus, this mixture was further purified by HPLC to obtain pure BotA^{PCCD}. Purification was performed on a Dionex Ultimate 3000 RSLC system using a Xbridge C18, 4.6 x 100 mm, 5 µm dp column equipped with a C18 precolumn (Waters). Aliquots of 15 µL were separated by a gradient from (A) 10 mM NH$_4$HCO$_3$ in H$_2$O to (B) ACN at a flow rate of 1 mL/min and 45 °C. The gradient was initiated by a 1 min isocratic step at 5% B, followed by an increase to 19.7% B in 2 min, to 21.0% B in 24 min and to 95% B in 0.5 min. After a 2 min

step at 95% B the system was re-equilibrated to the initial conditions (5% B). BotA^{PCCD} containing fractions (confirmed by LC-MS) were pooled and ACN evaporated on a rotary evaporator. Finally, the residue was injected on a Superdex S30 column (GE Healthcare) using reaction buffer.

The ability of PurCD to use different IpoC-cyclysed BotA^P variants as substrate was tested. 50 µM BotA^P variant or wt (as control) was first incubated with 5 µM IpoC and 10 mM ATP/MgCl$_2$ in reaction buffer for 12 h at 37 °C. 5 µM PurCD was added and samples were incubated for further 2 h at 37 °C. The addition of ACN was used to stop the reaction and precipitate the enzymes. The conversion was analyzed by LC-MS. Reactions were set up and analyzed in triplicates.

To test if macroamidine re-opening by PurCD occurs after the follower is cleaved, 50 µM purified BotA^{PCCD} or BotA^{PCCDAH} was incubated with 10 µM PurCD, 10 mM ATP/MgCl$_2$ in macrocyclisation buffer (50 mM Tris pH 8.5, 200 mM NaCl, 10% glycerol) at room temperature for 16 h. Controls without the addition of enzymes were also set up. Reactions were stopped by the addition of ACN and analyzed my LC-MS.

Follower cleavage by PurAH

To test possible substrates for PurAH, 50 µM BotA peptide (BotA, BotA^P, BotA^{PC} or BotA^{PCCD}) was incubated with 5 µM PurAH and 100 µM CoCl$_2$ in reaction buffer at 37 °C for 16 h. The reactions were set up in triplicates. After the 16 h incubation time, reactions were stopped and enzymes were precipitated by the addition of ACN and analyzed by LC-MS.

The activity of EDTA-treated PurAH was tested in the presence of no metal ions and using different metal ions. 50 µM BotA^{PCCD} was incubated with 0.2 µM PurAH and 100 µM ZnCl$_2$, CoCl$_2$, CdCl$_2$, FeCl$_2$, MgCl$_2$, MnCl$_2$, NiCl$_2$ or no metal ions, respectively, at 37 °C. Reactions were set up in triplicates and stopped after an incubation time of 45 min by the addition of ACN. Turnover was analyzed by LC-MS.

The ability of PurAH to cleave off the follower of BotA^P variants was tested. 50 µM BotA^P variant or wt (as control) were first incubated with 5 µM IpoC and 10 mM ATP/MgCl$_2$ in reaction buffer for 12 h at 37 °C. 5 µM PurCD and 2.5 µM PurAH were added and samples were incubated for further 2 h at 37 °C. The reactions were stopped and enzymes were precipitated by the addition of ACN and the conversion was analyzed by LC-MS. Reactions were set up and analyzed in triplicates.

To determine the initial conversion rate of BotA^{PC} and BotA^{PCCD} by PurCD, 50 µM substrate was incubated with 0.2 µM PurAH and 100 µM CoCl$_2$. Samples were prepared without addition

of enzyme and pre-incubated at 37 °C. After addition of enzyme, samples were incubated in the LC sample holder at 37 °C and 1 µl aliquots were automatically taken and analyzed every 13 min by the LC-MS system in the time range from 1 to 105 min. Reactions were set up and analyzed in triplicates.

The turnover of PurAH wt, PurAH Y185F and PurAH D348N was compared. 50 µM BotA[P] wt was first incubated with 5 µM IpoC and 10 mM ATP/MgCl$_2$ in reaction buffer for 12 h at 37 °C. 5 µM PurCD was added and samples were incubated for further 1.5 h at 37 °C. The reaction solution was split up and 2.5 µM PurAH (wt or mutant) was added. After 0, 30, 60, 120, 240 and 480 min, the reactions were stopped and enzymes were precipitated by the addition of two volumes of ACN.

To produce and purify BotA[PCCDAH], 50 µM BotA[PC] was incubated with 5 µM PurCD, 0.2 µM PurAH, 10 mM ATP/MgCl$_2$, 100 µM CoCl$_2$ in reaction buffer at 37 °C overnight. The reaction solution was applied on a Chromabond C18 Hydra column (6 mL, 2000 mg, Macherey-Nagel), pre-equilibrated with dest. H$_2$O. After washing with 2 CV dest. H$_2$O, bound peptides (**2** and Follower peptide) were eluted separately by increasing the ACN concentration in 5% steps. Fractions containing pure BotA[PCCDAH] (analyzed by LC-MS) were combined and all solvent evaporated on a rotary evaporator. Finally, **2** was dissolved in reaction buffer.

Mass spectrometry

<u>LC-ESI-MS analysis</u>

All measurements to analyze the mass of processed or unprocessed BotA peptides were performed on a Dionex Ultimate 3000 RSLC system using a Xbridge C18, 100 x 2.1 mm, 3.5 µm dp column equipped with a C18 precolumn (Waters). Samples of 1 µL were separated by a gradient from (A) 10 mM NH$_4$HCO$_3$ to (B) ACN at a flow rate of 600 µL/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step at 5% B, followed by an increase to 15% B in 1 min, to 27% B in 8.5 min and to 95% B in 0.5 min. After a 2 min step at 95% B the system was re-equilibrated to the initial conditions (5% B). UV spectra were recorded by a DAD in the range from 200 to 600 nm.

For MS measurements on maXis-2 UHR-TOF mass spectrometer (Bruker Daltonics), the LC flow was split 1:8 before entering the mass spectrometer using the Apollo ESI source. In the source region, the temperature was set to 250 °C, the capillary voltage was 4000 V, the dry-gas flow was 10.0 L/min and the nebulizer was set to 30 psi. After the generated ions passed the quadrupole with a low cut-off at 150 m/z they were trapped in the collision cell for 100 µs and then transferred within 10 µs through the hexapole into the ICR cell. Data were recorded in the

mass range from 250 to 2500 m/z. Peaks in the MS-spectra are labelled with the observed monoisotopic masses.

For quantification of the time course experiments, the consumption of the substrates (BotA[PC] and BotA[PCCD], respectively) was quantified. The peak areas of the $[M+H]^{4+}$, $[M+H]^{5+}$ and $[M+H]^{6+}$ ions (which were the predominant ions) from BotA[PC] or BotAP[CCD] (see Table 3-S4, EIC width ± 5 ppm) were integrated and concentrations were calculated using a calibration curve obtained from serial dilutions of the respective purified BotA (BotA[PC] or BotA[PCCD]) using the TASQ 1.1 software (Bruker Daltonics). The relationship between the peak areas and the peptide concentrations are mostly linear in the tested concentration range (2-75 μM) and no saturation of the detector was not observed (as an example the calibration curve for BotA[PCCD] is shown in Figure 3-S22).

For the comparison of the activity of EDTA-treated PurAH in presence of different metal ions, the area of the reaction product **2** ($[M+H]^+$=799.3807 ± 0.05 Da) was integrated and compared. For **2,** a linear relationship of peak areas and concentrations were tested (concentration range 2.6-100 μM) and verified (Figure 3-S23), which allows a comparison of the formation of **2** by comparison of the peak areas.

To compare the turnover of the BotA[PCCD] wt by PurAH wt and mutants, the peak areas of the reaction product **2** ($[M+H]^+$=799.3807 ± 0.05 Da) were integrated and compared.

For estimation (relative differences) of the turnover of the BotA substrate mutants by PurAH wt (categorized as "good" (turnover comparable to wild-type (wt) substrate), "fair" (> 50% turnover compared to wt substrate), (> 50% turnover compared to wt substrate) "poor" , and "no turnover/trace"), the sum of substrate peak areas of the $[M+H]^{4+}$, $[M+H]^{5+}$ and $[M+H]^{6+}$ ions from the respective BotA[PC] and BotA[PCCD] mutants (see Table 3-S3) before and after addition of PurAH wt were compared. A linear relationship between the peak areas and the peptide substrate concentrations (in the assay concentrations range) was confirmed: Dilutions of the enzyme assay solutions (wt and mutant peptides) and dilutions of purified BotA[PC] wt and BotA[PCCD] wt show a linear relationship between the dilution factor and the peak area. We thus conclude that the inference of relative differences is feasible.

MS[2] Fragmentation

MS[2] fragmentation experiments were executed by direct infusion MS using a solariX XR (7T) FT-ICR mass spectrometer (Bruker Daltonics) equipped with an Apollo ESI source. The source region temperature was set to 220 °C, the capillary voltage was 4500 V, the dry gas flow was 4.0 L min[-1] and the nebulizer was set to 1.0 bar. The collision energy was ramped from 10 to

22 V until sufficient dissociation of the target molecule was observed. MS[2] spectra were obtained by collision-induced dissociation (CID) in the collision cell.

Crystallization and structure determination

Crystals of PurAH were obtained at 18 °C in 3.0 - 4.2 M sodium formate. The crystals were cryoprotected in mother liquor supplemented with 35% glycerol and flash cooled in liquid nitrogen. Data was collected at ESRF (Beamlines: ID23-1 and ID30A-3). To solve the PurAH structure, a single-wavelength anomalous dispersion (SAD) data set and a high-resolution dataset were collected. Data were processed using XDS[40] and POINTLESS[41], AIMLESS[42] and ctruncate[43] implemented in ccp4[44]. The structure was solved using the SAD dataset with the Auto-Rickshaw[45] pipeline's MR-SAD phasing method[46-53]. The model was manually rebuilt in COOT[54] and refined using PHENIX[46]. The structure was validated using *MolProbity*[55], and all images presented were created using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). Interaction diagrams were created using PoseView [56, 57] and plotting was performed with Matplotlib v2.1.0.[58]

3.4 References

1.   Nakamura, S., Isolation and characterization of bottromycins A and B. *J. Antibiotics, Ser. A* **1965,** *18*, 47-52.

2.   Waisvisz, J. M., Bottromycin. I. A new sulfurcontaining antibiotic. *J. Am. Chem. Soc.* **1957,** *79*, 4520-4521.

3.   Sowa, S.; Masumi, N.; Inouye, Y.; Nakamura, S.; Takesue, Y.; Yokoyama, T., Susceptibility of methicillin-resistant Staphylococcus aureus clinical isolates to various antimicrobial agents. *Hiroshima J Med Sci* **1991,** *40* (4), 137-44.

4.   Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Omura, S., Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew Chem Int Ed Engl* **2009,** *48* (5), 914-7.

5.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2. Release of aminoacyl- or peptidyl-tRNA from ribosomes. *J Biol Chem* **1976,** *251* (8), 2299-306.

6.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect on peptide bond formation. *FEBS Lett* **1981,** *123* (2), 173-6.

7.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect of bottromycin A2 on polysomes. *FEBS Lett* **1983,** *153* (1), 53-9.

8.   Franz, L.; Adam, S.; Santos-Aberturas, J.; Truman, A. W.; Koehnke, J., Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme. *J Am Chem Soc* **2017,** *139* (50), 18158-18161.

9.   Schwalen, C. J.; Hudson, G. A.; Kosol, S.; Mahanta, N.; Challis, G. L.; Mitchell, D. A., In Vitro Biosynthetic Studies of Bottromycin Expand the Enzymatic Capabilities of the YcaO Superfamily. *J Am Chem Soc* **2017,** *139* (50), 18154-18157.

10. Horbal, L.; Marques, F.; Nadmid, S.; Mendes, M. V.; Luzhetskyy, A., Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metab Eng* **2018,** *49*, 299-315.

11. Eyles, T. H.; Vior, N. M.; Truman, A. W., Rapid and Robust Yeast-Mediated Pathway Refactoring Generates Multiple New Bottromycin-Related Metabolites. *ACS Synth Biol* **2018,** *7* (5), 1211-1218.

12. Yamada, T.; Yagita, M.; Kobayashi, Y.; Sennari, G.; Shimamura, H.; Matsui, H.; Horimatsu, Y.; Hanaki, H.; Hirose, T.; S, O. M.; Sunazuka, T., Synthesis and Evaluation of Antibacterial Activity of Bottromycins. *J Org Chem* **2018,** *83* (13), 7135-7149.

13. Crone, W. J.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angew Chem Int Ed Engl* **2016,** *55* (33), 9639-43.

14. Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Muller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *Chembiochem* **2016,** *17* (23), 2286-2292.

15. Jimenez-Morales, D.; Adamian, L.; Shi, D.; Liang, J., Lysine carboxylation: unveiling a spontaneous post-translational modification. *Acta Crystallogr D Biol Crystallogr* **2014,** *70* (Pt 1), 48-57.

16. Concha, N. O.; Rasmussen, B. A.; Bush, K.; Herzberg, O., Crystal structure of the wide-spectrum binuclear zinc beta-lactamase from Bacteroides fragilis. *Structure* **1996,** *4* (7), 823-36.

17. Concha, N. O.; Janson, C. A.; Rowling, P.; Pearson, S.; Cheever, C. A.; Clarke, B. P.; Lewis, C.; Galleni, M.; Frere, J. M.; Payne, D. J.; Bateson, J. H.; Abdel-Meguid, S. S., Crystal structure of the IMP-1 metallo beta-lactamase from Pseudomonas aeruginosa and its complex with a mercaptocarboxylate inhibitor: binding determinants of a potent, broad-spectrum inhibitor. *Biochemistry* **2000,** *39* (15), 4288-98.

18. Ullah, J. H.; Walsh, T. R.; Taylor, I. A.; Emery, D. C.; Verma, C. S.; Gamblin, S. J.; Spencer, J., The crystal structure of the L1 metallo-beta-lactamase from Stenotrophomonas maltophilia at 1.7 A resolution. *J Mol Biol* **1998,** *284* (1), 125-36.

19. Bounaga, S.; Laws, A. P.; Galleni, M.; Page, M. I., The mechanism of catalysis and the inhibition of the Bacillus cereus zinc-dependent beta-lactamase. *Biochem J* **1998,** *331 (Pt 3),* 703-11.

20. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N., ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **2016,** *44* (W1), W344-50.

21. Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **2010,** *38* (Web Server issue), W529-33.

22. Landau, M.; Mayrose, I.; Rosenberg, Y.; Glaser, F.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **2005,** *33* (Web Server issue), W299-302.

23. Glaser, F.; Pupko, T.; Paz, I.; Bell, R. E.; Bechor-Shental, D.; Martz, E.; Ben-Tal, N., ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* **2003,** *19* (1), 163-4.

24. Holm, L.; Laakso, L. M., Dali server update. *Nucleic Acids Res* **2016,** *44* (W1), W351-5.

25. Crone, W. J. K.; Leeper, F. J.; Truman, A. W., Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chemical Science* **2012,** *3* (12), 3516-3521.

26. Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chemical Science* **2012,** *3* (12), 3522-3525.

27.   Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Muller, R., Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem Biol* **2012,** *19* (10), 1278-87.

28.   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., Structure and biosynthesis of the antibiotic bottromycin D. *Org Lett* **2012,** *14* (19), 5050-3.

29.   Fast, W.; Tipton, P. A., The enzymes of bacterial census and censorship. *Trends Biochem Sci* **2012,** *37* (1), 7-14.

30.   Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A., Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proc Natl Acad Sci U S A* **2018,** *115* (12), 3030-3035.

31.   Dunbar, K. L.; Melby, J. O.; Mitchell, D. A., YcaO domains use ATP to activate amide backbones during peptide cyclodehydrations. *Nat Chem Biol* **2012,** *8* (6), 569-75.

32.   Dunbar, K. L.; Chekan, J. R.; Cox, C. L.; Burkhart, B. J.; Nair, S. K.; Mitchell, D. A., Discovery of a new ATP-binding motif involved in peptidic azoline biosynthesis. *Nat Chem Biol* **2014,** *10* (10), 823-9.

33.   Melby, J. O.; Li, X.; Mitchell, D. A., Orchestration of enzymatic processing by thiazole/oxazole-modified microcin dehydrogenases. *Biochemistry* **2014,** *53* (2), 413-22.

34.   Li, Y. M.; Milne, J. C.; Madison, L. L.; Kolter, R.; Walsh, C. T., From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* **1996,** *274* (5290), 1188-93.

35.   Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **1997,** *25* (17), 3389-402.

36.   Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L., Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim Biophys Acta* **2015,** *1854* (8), 1019-37.

37.   Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003,** *13* (11), 2498-504.

38.   Durrant, J. D.; Votapka, L.; Sorensen, J.; Amaro, R. E., POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics. *J Chem Theory Comput* **2014,** *10* (11), 5047-5056.

39.   Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F., Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **1999,** *112*, 531-52.

40.   Kabsch, W., Xds. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 125-32.

41.   Evans, P. R., An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 282-92.

42.   Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **2013,** *69* (Pt 7), 1204-14.

43.   French, S.; Wilson, K., On the treatment of negative intensity observations. *Acta Crystallographica Section A* **1978,** *34* (4), 517-525.

44.   Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 235-42.

45.   Panjikar, S.; Parthasarathy, V.; Lamzin, V. S.; Weiss, M. S.; Tucker, P. A., On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr D Biol Crystallogr* **2009,** *65* (Pt 10), 1089-97.

46.   Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

47.   Bricogne, G.; Vonrhein, C.; Flensburg, C.; Schiltz, M.; Paciorek, W., Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallographica Section D* **2003,** *59* (11), 2023-2030.

48.   Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L., Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **1998,** *54* (Pt 5), 905-21.

49.   Cowtan, K., The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D* **2006,** *62* (9), 1002-1011.

50.   McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J Appl Crystallogr* **2007,** *40* (Pt 4), 658-674.

51.   Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 355-67.

52.   Otwinowski, Z., Maximum likelihood refinement of heavy atom parameters. *Isomorphons Replacement and Anomalous Scattering* **1991**, 80-86.

53. Sheldrick, G. M., Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 479-85.

54. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.

55. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 1), 12-21.

56. Fricker, P. C.; Gastreich, M.; Rarey, M., Automated Drawing of Structural Molecular Formulas under Constraints. *Journal of Chemical Information and Computer Sciences* **2004,** *44* (3), 1065-1078.

57. Stierand, K.; Rarey, M.; Maaß, P. C., Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. *Bioinformatics* **2006,** *22* (14), 1710-1716.

58. Hunter, J. D., Matplotlib: A 2D graphics environment. *Computing in science & engineering* **2007,** *9* (3), 90-95.

Chapter 4

The bottromycin epimerase BotH defines a group of atypical α/β-hydrolase-fold enzymes

Asfandyar Sikandar,[+] **Laura Franz**,[+] Sebastian Adam, Javier Santos-Aberturas, Liliya Horbal, Andriy Luzhetskyy, Andrew W. Truman, Olga V. Kalinina, and Jesko Koehnke

[+] shared authorship

Author contributions:

L.F. established BotH activity, carried out the biochemical experiments, produced BotH substrates, carried out Marfey's analysis and performed the mass spectrometry. A.S. and J.K. established the production and purification of BotH and performed MST experiments. A.S. designed and performed crystallization experiments, determined the reported crystal structures, produced and purified bottromycin A_2 and performed pull-down experiments. S.A. established the purification of the BotH substrate. J.S.-A. and A.W.T. aided bioinformatic analyses. L.H. produced, purified and analyzed bottromycin A_2 and derivatives under the guidance of A.L. O.V.K. designed and performed the bioinformatic analyses and wrote the bioinformatics section. J.K. analyzed and visualized the crystal structures for publication and wrote the paper with contributions from all authors. The full program was carried out under the guidance and direction of J.K.

4 The bottromycin epimerase BotH defines a group of atypical α/β-hydrolase-fold enzymes

4.1 Abstract

D-amino acids endow peptides with diverse, desirable properties, but the post-translational and site-specific epimerization of L-amino acids into their D-counterparts is rare and very challenging. Bottromycins are ribosomally synthesized and post-translationally modified peptides that have overcome this challenge and feature a D-aspartate (D-Asp), which was proposed to arise spontaneously during biosynthesis. We have identified the highly unusual alpha/beta-hydrolase (ABH) fold enzyme BotH as a novel peptide epimerase responsible for the post-translational epimerization of L-Asp to D-Asp during bottromycin biosynthesis. The biochemical characterization of BotH combined with the structures of BotH and the BotH-substrate complex allowed us to propose a mechanism for this reaction. Bioinformatic analyses of BotH homologs show that similar ABH enzymes are found in diverse biosynthetic gene clusters. This places BotH as the founding member of a new group of atypical ABH enzymes that may be able to epimerize non-Asp stereocenters across different families of secondary metabolites.
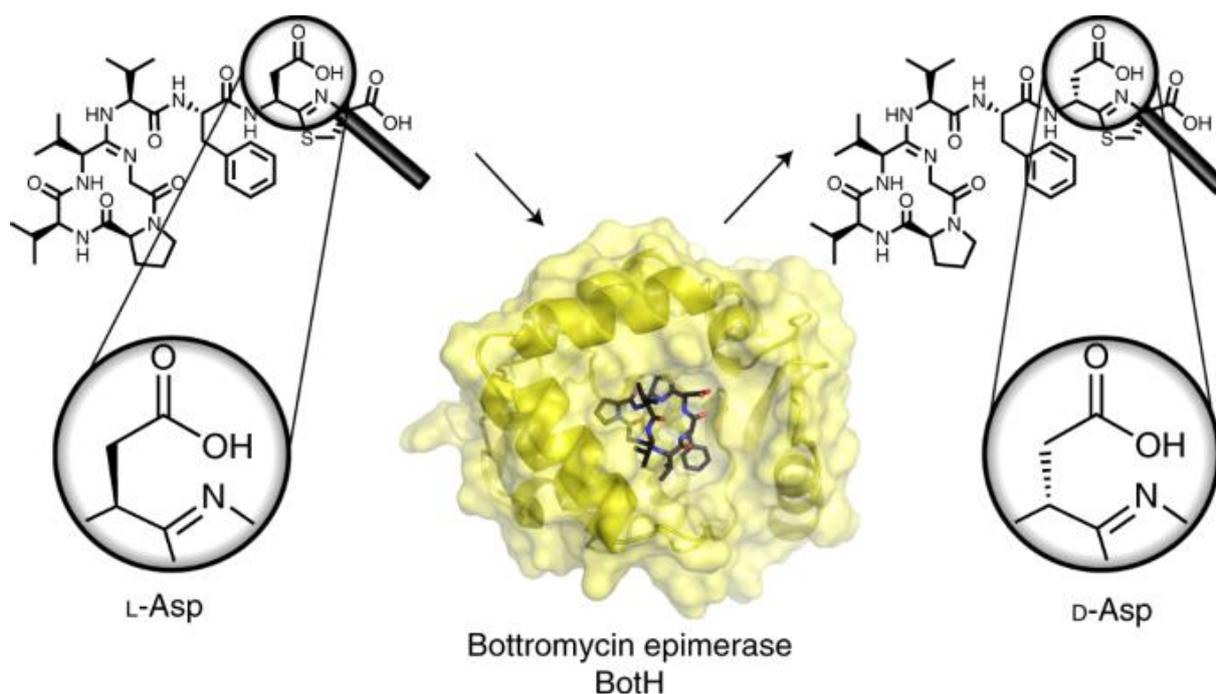

Figure 4-1 Graphical abstract.

4.2 Main Text

Introduction

Enzymes belonging to the superfamily of alpha/beta-hydrolase-fold proteins (ABHs) are found in all domains of life[1-3]. Their catalytic roles are highly diverse and they participate in primary and secondary metabolism, where they are usually responsible for the hydrolysis of (thio)ester and peptide bonds[4]. In addition, ABHs have also been reported to function as dehalogenases, epoxide hydrolases, dioxygenases, decarboxylases and haloperoxidases[3, 4]. Despite low overall sequence conservation, ABH enzymes share a remarkably conserved core fold[2, 3] that has a V-shaped lid domain above the active-site as a frequent addition[5, 6]. A majority of ABH family members possesses a canonical Ser, His, Asp catalytic triad at the active-site, but other catalytic residues have also been observed[3, 4]. In addition to catalytic roles, a ABH family members have also been reported to fulfil several other important functions, including small-molecule receptors that are involved in signal transduction, cell-cell interaction and channel regulation[5-8]. The ABH fold is thus a prime example for the reappropriation of a conserved core fold during evolution to fulfill a myriad of functions.

An aberrant ABH protein, BotH, is encoded in the biosynthetic gene cluster (BGC) for bottromycins[9-12]. This natural product family was first isolated from the terrestrial bacterium *Streptomyces bottropensis* and originally described as peptidic antibiotics with potent activity against Gram-positive bacteria[13, 14]. Bottromycins are active against problematic human pathogens, such as Methicillin-resistant *Staphylococcus aureus*[15, 16], address a novel target (the A-site of the prokaryotic ribosome)[17-19] and belong to the fast growing superfamily of ribosomally synthesized and post-translationally modified peptides (RiPPs)[9-12]. As is typical for RiPPs, their biosynthesis begins with the expression of a small structural gene to yield the precursor peptide (PP)[20]. Uniquely amongst bacterial RiPPs, the bottromycin PP contains an N-terminal core peptide (the eventual natural product) and a C-terminal follower peptide[9-12], which is important for substrate recognition by several of the biosynthetic enzymes[21, 22]. The order of biosynthetic steps (and responsible enzymes) has since been proposed based on an untargeted metabolomics approach using mass spectral networking[23]. Subsequent *in vitro* work has largely corroborated the metabolomics data[21, 22, 24, 25]. In the first phase of bottromycin biosynthesis, the N-terminal methionine is removed, proline, valine and phenylalanine residues of the core peptide are C-methylated and a cysteine-derived thiazoline is installed in no particular order[23]. The hallmark macroamidine linkage is then formed in this intermediate[21, 22], which is followed by proteolytic removal of the follower peptide to yield **1** (Figure 4-2a)[25]. To complete the biosynthesis of bottromycin A₂ (**2**, Figure 4-2a), **1** undergoes epimerization,

oxidative decarboxylation and O-methylation. While the latter two modifications have been attributed to specific enzymes, epimerization of the L-Asp of **1** was observed to progress spontaneously, albeit very slowly[23].
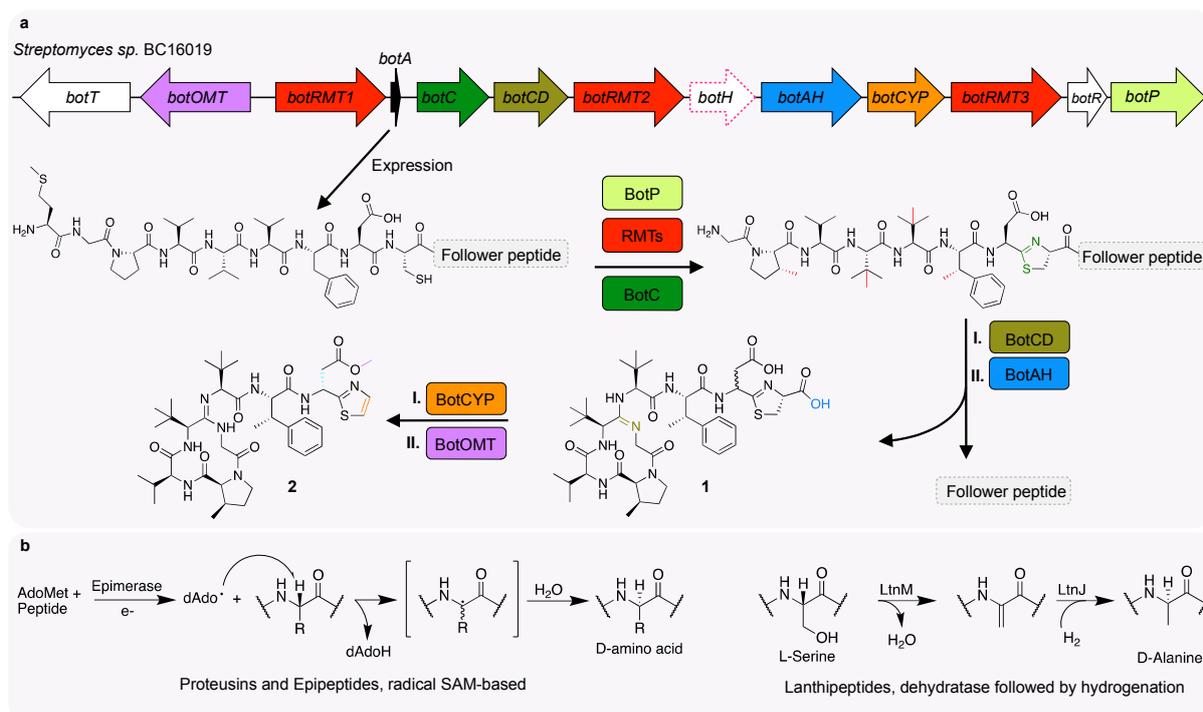
Figure 4-2 | Bottromycin biosynthesis and epimerization in RiPPs. **a** Bottromycin BGC found in *S.* sp. BC16019. After expression of the precursor peptide BotA, its N-terminal methionine is removed by BotP, three radical methyl transferases (RMT) perform four C-methylations and BotC installs a cysteine-derive thiazoline. These initial steps appear to follow no particular order. Next, BotCD catalyzes formation of the macroamidine, after which BotAH removes the follower peptide to yield **1**. Oxidative decarboxylation (BotCYP) and O-methylation complete the biosynthesis of bottromycin A$_2$ (**2**). Epimerization was proposed to occur spontaneously. Genes, enzymes and modifications have matching colors. White arrows represent genes for which no or a regulatory function have been proposed. The unusual ABH protein encoded in the pathway, BotH, is highlighted in pink. **b** Epimerization in RiPP biosynthesis reported to date. In proteusins, including polytheonamides, and epipeptides a radical SAM enzyme epimerizes a range of amino acids. In lanthipeptide biosynthesis, L-Ser can be converted to D-Ala in a two-step process.

Amino acid epimerization in non-ribosomal peptide synthesis is usually catalyzed by epimerization domains embedded within the assembly line that function on carrier protein-bound aminoacyl substrates[26]. Due to their ribosomal origin, RiPPs must undergo post-translational epimerization after the PP has been expressed as an all L-amino acid peptide. To date, only two enzymatic mechanisms for this process have been described in RiPPs, involving either a radical-SAM enzyme[27-30] or a two-step dehydration-hydrogenation process to generate D-alanine from L-serine (Figure 4-2b)[31, 32].

Here, we report the identification of BotH, an unusual ABH enzyme from the bottromycin BGC, as the epimerase of the biosynthetic pathway. This is the first reported instance of an ABH enzyme catalyzing peptide epimerization and thus expands the catalytic scope of this vast enzyme family. Biochemical data together with the structure of the BotH–substrate complex allowed us to propose a mechanism for this reaction. Interestingly, BotH is also able to bind bottromycins with high affinity, which hints at additional function(s) in the biosynthetic process. We show that all canonical ABH active-site residues required for hydrolase activity are absent in BotH, and bioinformatic analyses indicate that BotH homologs with comparable non-hydrolytic residues are widespread amongst BGCs and may catalyze similar biosynthetic steps.

Results

BotH is an unusual member of the ABH superfamily

In a search for an enzyme that may catalyze the epimerization in the bottromycin pathway, we noticed that the gene encoding for BotH had been annotated as an ABH, but that the predicted active-site residues (Ser/His/Asp) were not present and there were no prior experimental data on the role of this protein. We therefore expressed, purified and crystallized BotH (see Methods for details). A high-resolution (1.18 Å) native BotH dataset was collected from a crystal belonging to space group I222, which was phased using seleno-methionine BotH data (data collection and refinement statistics for all structures can be found in Table 4-S1). The crystals contained one protomer in the asymmetric unit and the electron density for residues 10-262 was continuous in the refined model. BotH is comprised of the prototypical ABH core structure (Figure 4-S1) and the putative active site was covered by a V-shaped loop consisting of four α-helices (Figure 4-S1). A search for similar structures using the DALI server[33] revealed 3-oxoadipate-enol-lactonase (PDB 2xua)[34] as the closest structural homolog (Figure 4-S1). A comparison of the two proteins revealed that in BotH, the active-site Ser has been mutated to a Phe, which is part of a Phe-Phe motif that spans a large, hydrophobic plane at the active site (Figure 4-S1). The remaining residues of the catalytic triad are either mutated (His to Ile) or missing (Asp) (Figure 4-S1). In spite of these mutations, the sizeable cavity found in this region of the structure appeared to be large enough for binding **1**.

BotH catalyzes the epimerization of 3b and 3a

To confirm this hypothesis, a des-methyl analog of **1**, **3** (des-methyl Pro2, Val4, Val5 and Phe6), was enzymatically produced as reported previously (Figure 4-3a and Figure 4-S2)[25]. Careful analysis revealed **3** to exist as an epimeric mixture of **3a** (D-Asp) and **3b** (L-Asp) (Figure 4-3a, Figure 4-S3 and Figure 4-S5). When 20 µM **3a/b** was incubated with 5 µM BotH and the reactions were analyzed by high-resolution electrospray ionization liquid chromatography−mass spectrometry (HR-ESI-LCMS), we observed a change in **3a** : **3b** ratios (Figure 4-3b), which was BotH-concentration dependent within the 2 h time-scale of the experiment (Figure 4-3c): Increasing the BotH concentration resulted in a shift to **3a** (D-Asp), which is the required epimer to proceed with biosynthesis (see chapter 5).
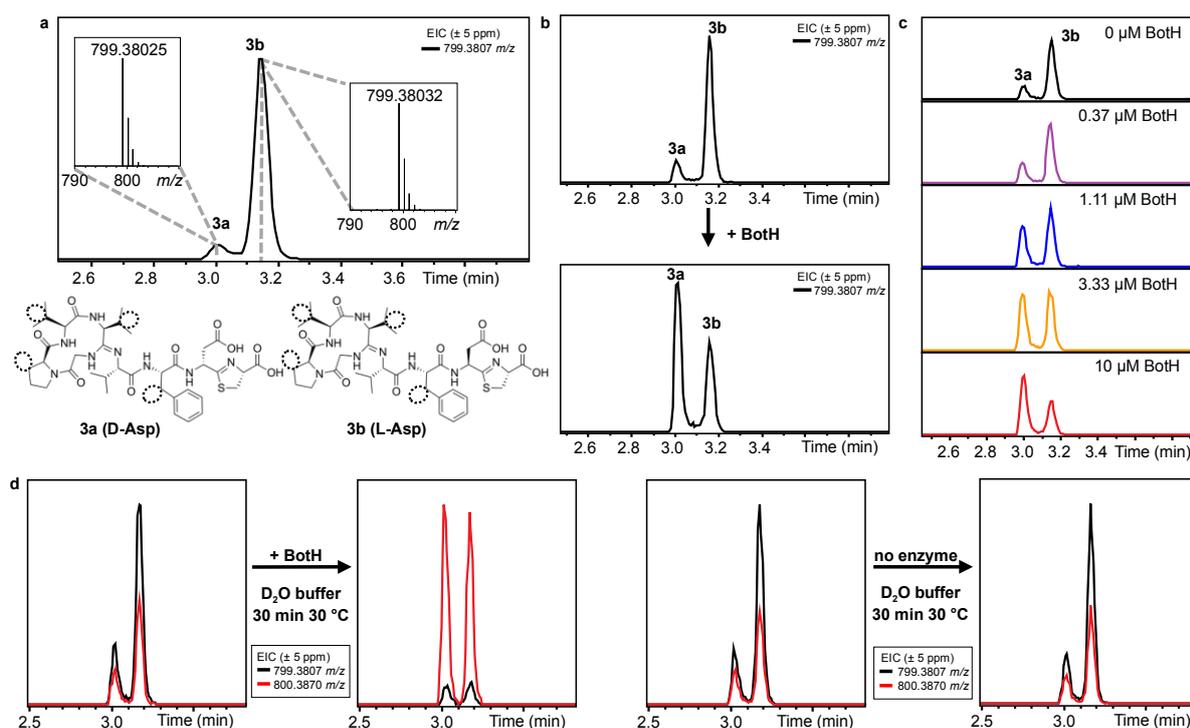
Figure 4-3 | BotH acts as an epimerase. **a** Spontaneous epimerization of the Asp in position 7 after proteolytic removal of the follower peptide. Marfey's reagent was used to assign the stereochemistry of the Asp for both peaks (Figure 4-S4). Extracted ion chromatograms (EIC) of **3** ([M+H]$^+_{calc.}$=799.3807; ± 5 ppm) and mass spectra at 3.00 and 3.15 min are shown. Missing methyl groups are highlighted by dashed ovals. **b** Incubation of **3a/b** with BotH results in a change of **3a** : **3b** ratios with **3a** (D-Asp) now the more abundant species. EICs of **3** ([M+H]$^+_{calc.}$=799.3807; ± 5 ppm) are shown. **c** Incubating 20 µM **3a/b** with increasing concentrations of BotH led to a shift of the equilibrium towards **3a**. Increasing BotH concentrations beyond 10 µM did not lead to a further shift of epimer ratios. EICs of **3** ([M+H]$^+_{calc.}$=799.3807; ± 5 ppm) are shown. **d** Deuteron incorporation into **3a/b** by BotH in D$_2$O buffer. EICs for **3** ([M+H]$^+_{calc.}$=799.3807; ± 5 ppm) (black) and deuterium incorporated **3** ([M+H]$^+_{calc.}$=800.3870; ± 5 ppm) (red), as well as mass spectra at the EIC maxima are shown. Representative experiments were repeated independently three times with similar results.

To probe if only **3b** or both epimers were substrates, we incubated **3a/b** with BotH in D$_2$O. This resulted in rapid (< 60 s) deuteron incorporation at the Asp7 position in both peaks, while very little deuteron incorporation was observed in the absence of BotH even after 24 hours (Figure 4-3d, Figure 4-S3 and Figure 4-S6). In fact, BotH concentrations low enough to leave epimer ratios unchanged still resulted in accelerated deuteron incorporation (Figure 4-S6). Repeating this experiment using the deuterated sample in H$_2$O showed an equally rapid exchange with solvent protons back to **3a/b** (Figure 4-S6). These data implied that BotH accepts both, **3a** and **3b**, as substrates in a reversible reaction, while favoring D-Asp (**3a**) as the product.

Structure of the BotH-3a complex

To better understand the mechanism of this intriguing enzyme, the high-resolution crystal structure of BotH in complex with its substrate **3a/b** was determined to 1.25 Å resolution. The overall structure of the complex was virtually unchanged when compared to the apo structure (C$_\alpha$ rmsd of 0.12 Å) (Figure 4-4a and Figure 4-S7). The substrate is curled into the active site in a way that places the thiazoline underneath the four amino acid macrocycle and the substrate engages in extensive hydrophobic interactions as well as inter- and intramolecular hydrogen bonds (Figure 4-4b and Figure 4-S7).
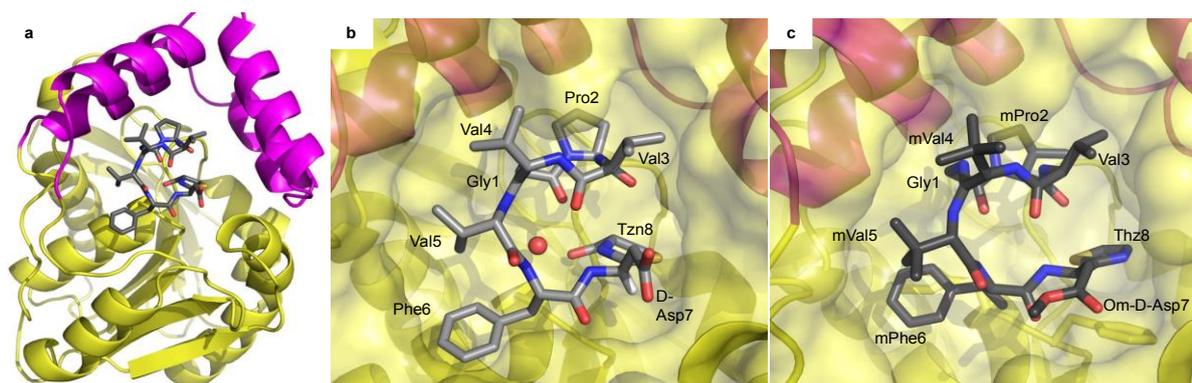
Figure 4-4 | BotH-**3a** and -bottromycin A$_2$ complexe structures. a Cartoon representation of the BotH-**3a** complex structure highlighting the V-shaped loop (magenta) positioned above the active site. The bound substrate is shown as sticks (grey (carbon), blue (nitrogen), red (oxygen) and yellow (sulfur). **b** The Close-up of **3a** (light gray sticks, labeled. Tzn = Thiazoline) and bottromycin A$_2$ (dark gray stick. mPhe bound in the BotH active site (cartoon, yellow/magenta with semi-transparent surface representation). The ordered water molecule trapped between substrate and protein is shown as a red sphere. The Asp7 C$_\alpha$ hydrogen is shown as a white stick for clarity. **c** Same as **b**, but bottromycin A$_2$ (dark gray sticks, labeled) bound to the active site of BotH. m = methylated residue, Om = O-methylated residue, Thz = Thiazole.

Of particular note are hydrogen bonds of the thiazoline carboxy group with the backbone NH of Val41 and Phe110. These two BotH residues are in the position of the oxyanion hole found in many hydrolases. The carboxy group of the substrate's Asp7 is involved in a hydrogen bonding network with ordered water molecules that ultimately link this side-chain to BotH residues (Figure 4-S7). An intramolecular hydrogen bond links the substrate's carbonyl of Val3 with the backbone NH of Asp7. The best fit to the electron density is achieved by choosing the amidine resonance structure that places the double-bond inside the macrocycle and a D-Asp in position 7, which indicates that the ligand observed in the complex structure is **3a** (Figure 4-S7). Since the epimerization mechanism would, at least formally, involve proton abstraction and addition, we scanned the complex structure for potential catalytic residues within 4 Å of the Cα hydrogen of Asp7, the site of catalysis, but could not identify any. This left two possibilities: bulk solvent or the side chain of substrate Asp7; its carboxy group is within 2.2 Å of the Cα hydrogen of interest. We would like to note that the carboxy group of Asp7 is surrounded closely (O – O distances 2.5 – 3.1 Å) by four ordered water molecules (Figure 4-S8), which appear well-positioned to facilitate the exchange of the abstracted proton/deuteron with bulk solvent during catalysis (Figure 4-S8).

To probe the importance of the side chain identity in position seven for epimerization, we first tested the conservative mutant substrates Asp7Ala and Asp7Asn. While both substrates stabilized BotH in thermal shift assays comparable to **3a**/**b**, we observed no epimerization (Figure 4-S9). The extension of the Asp7 side-chain by an additional methylene group (Asp7Glu), results in severely reduced turnover by BotH, but this substrate epimerizes (Figure 4-S9). These data imply that a side-chain carboxy group in position 7 of the substrate is essential for catalysis and that the appropriate distance of this carboxy group relative to the residue's Cα proton has a significant impact on turnover, which suggests that the BotH reaction may be an example of substrate-assisted catalysis.

BotH has relaxed substrate specificity

We had previously reported that the enzymes used to generate **3a**/**b** possess relaxed substrate specificities for core-peptide residues 2 – 7 (Figure 4-S12)[25]. Since substrate position 7 was restricted to Asp or Glu for epimerization, we wondered if BotH tolerates mutations in the remaining positions. To this end, an additional series of 13 BotA core-peptide mutants were used to produce the variant BotH substrates enzymatically. These substrates were then incubated with or without BotH and analyzed by LC-MS (Figure 4-5a, Figure 4-S13 and Figure 4-S14). All but two mutations in positions two to six were processed by BotH. The positions of

Val5 and Phe6 are intimately connected and it appeared that the orientation of the Phe6 side-chain is critical for epimerization. In agreement with this hypothesis, Phe6Ala could not be epimerized by BotH while Phe6Tyr and Phe6Trp were substrates. The side-chain of position 5 (Val) is engaged in hydrophobic interactions with the side-chains of Val4 and Phe6. Accordingly, Val5Thr is not a BotH substrate, while Val5Ala, Val5Leu and Val5Glu can be epimerized by the enzyme. The ability of BotH to process Val5Glu may appear surprising, but the flexibility of the Glu side chain should allow the C_β methylene of Glu5 to engage in hydrophobic interactions with Val4 and Phe6, while the C_γ methylene and terminal carboxy group point towards bulk solvent. Our data demonstrate that BotH is able to process a variety of substrates, which will be invaluable for the production of bottromycin derivatives.
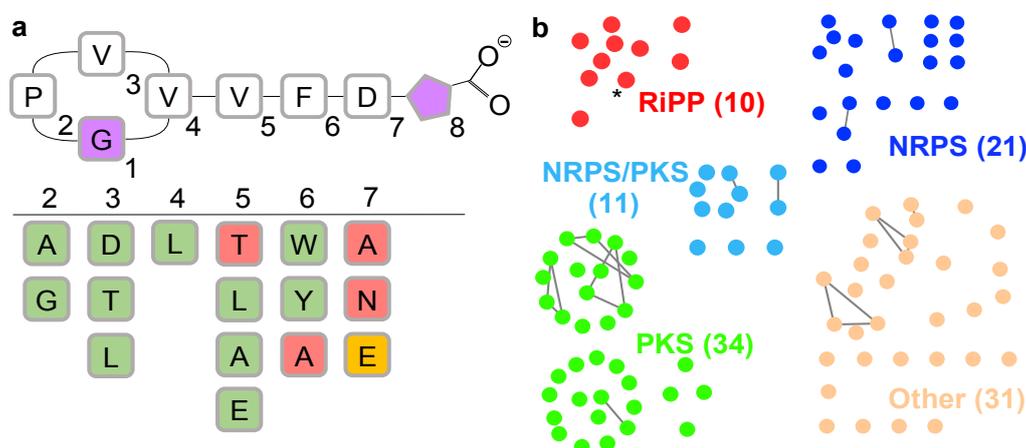
Figure 4-5 | BotH substrate specificity and distribution of BotH-like proteins in BGCs. a Summary of BotA point mutants tested as BotH substrates. Purple positions cannot be varied (pentagon represents thiazoline). Accepted mutations are highlighted in green, orange indicates a poor substrate and red mutations cannot be processed by BotH. The accompanying HR-LCMS data can be found in Figure 4-S13, 14. **b** Sequence similarity network of ABHs homologous to BotH (marked with an asterix). Of the 107 genes, 76 could be assigned to BGCs representing the three large bacterial natural product superfamilies (darker edges represent higher sequence similarity, NRPS = non-ribosomal peptide synthetase; PKS = polyketide synthase).

Bottromycins act as orthosteric inhibitors of BotH

Since **3a**/**b** are close structural homologs of bottromycin A$_2$ (**2**), it seemed feasible that BotH may be able to bind to bottromycins. We thus performed microscale thermophoresis (MST) experiments using heterologously produced **2** and three closely related variants (**4 - 6**, Figure 4-S10). BotH was able to bind all four bottromycins with K_D values in the high nM to low μM range but unable to epimerize oxidatively decarboxylated **3a** (Figure 4-S10). To understand the mode of binding, we determined the crystal structures of three bottromycin-BotH complexes to

1.40 (**2**), 1.70 (**5**), and 1.48 (**6**) Å resolution, respectively (Figure 4-4c, Figure 4-S11 and Figure 4-S17). These also represent the first crystal structures of any bottromycin. As observed for the complex with **3a**, the overall structural change in BotH due to ligand binding was minimal (C_α RMSD < 0.2 Å) and the bottromycins bound in a similar manner as **3a** (Figure 4-4b and 3c). Despite the high resolution, it is not obvious which way the thiazole is flipped, since the loss of the carboxy group due to thiazoline oxidation allows a fit of both rotamers without inducing a clash with BotH (Figure 4-S18). The bottromycins themselves are oriented in a twisted fashion that results in several tight, intramolecular hydrogen bonds (Figure 4-S17). Compared to the published NMR structure[35], bottromycin A_2 experiences significant strain as a result of binding to BotH, as it is forced to adopt a horseshoe shape with Val5 and Phe6 at its apex and the C-terminal thiazole stacked parallel under the macrocycle (Figure 4-S15). As expected, bottromycin A_2 acts as an orthosteric inhibitor of epimerization (Figure 4-S19), which suggests that BotH may be involved in a biosynthetic feedback mechanism to prevent self-poisoning of the producing strain (Figure 4-6b). In this model, an increase in the intracellular bottromycin concentrations offers a direct and faster means to reduce bottromycin production than altered gene expression since epimerization of Asp7 is highly important for the activity of the succeeding enzyme BotCYP (see chapter 5). It is of course also possible that BotH sequesters mature bottromycin to aid self-immunity.

It is still unclear why bottromycin A_2 contains a D-Asp, as studies with synthetic derivatives have shown that both epimers display the same bioactivity[36]. Selectivity of the bottromycin exporter BotT appears possible, despite a very small fraction of bottromycin A_2 in culture supernatant appearing to be the L-Asp epimer (Figure 4-S16). The main benefit of epimerization may be providing increased resistance to proteolytic degradation of bottromycin, since D-amino acid containing peptides have longer half-lives[37, 38]. It is of course also possible that the bottromycin target of organisms in a native setting requires a D-Asp for optimal target binding. Interestingly, the structure elucidation of novel bottromycin analogs via crystallization of their complex with BotH appears to require less compound and be more straightforward than NMR, due to the robust crystallization condition, high affinity and very well-diffracting crystals.

The ability of BotH to bind both the substrate and the mature natural product with high affinity may serve as a cautionary tale for attempting to identify the binding partners of "non-catalytic" ABHs. When using BotH to capture its ligand/substrate from either supernatant or lysate of a bottromycin producing strain, we were only able to detect bottromycin A_2, but not the actual substrate **1** (Figure 4-S21). Since biosynthetic pathway products tend to be present at higher concentrations than pathway intermediates, careful analysis of the biosynthetic pathway

supplying the ligand is required to exclude additional, non-canonical catalytic function(s) of the ABH under investigation.

**Evolution of BotH and its spread amongst bacterial BGCs**

ABHs comprise one of the largest protein families (~500,000 protein matches in InterPro[39], ~400,000 of them in Bacteria), of which 192,602 belong to Pfam[40] family PF00561 (Abhydrolase_1). The Ser-His-Asp catalytic triad is surprisingly poorly conserved: in 101,123 bacterial proteins from the PF00561 family, at least one of the catalytic residues is mutated or missing. In most cases (73,116 sequences), His is missing; and in the majority of these proteins (64,852 sequences) Asp is missing as well. Ser is mutated in only 42,810 sequences, most frequently to an aspartic acid, and completely missing in 1,283 sequences. Using these data, we identified 1,530 proteins that are unlikely to have hydrolase activity (see Methods section for details), including BotH. Taxonomic analysis revealed that these non-functional hydrolases are widespread among both Gram-positive and Gram-negative bacteria.

From these sequences, we then selected those that were in or near (less than 1 kb away) BGCs, which resulted in 107 proteins that were used to build a sequence–similarity network (Figure 4-5b and Figure 4-S22). In this network, all BotH homologs from bottromycin BGCs cluster together. Among the identified BGCs, two other major natural product superfamilies can be identified in addition to RiPPs: NRPS (non-ribosomal peptide synthetase) and PKS (polyketide synthase) clusters.

Discussion

We have identified the enzyme BotH as the epimerase of the bottromycin BGC, which is selective for Asp and Glu, but promiscuous with regards to mutations at other positions of the substrate. Based on the biochemical and structural data, we propose the following mechanism for epimerization in bottromycin biosynthesis (Figure 4-6a): Cleavage of the follower peptide converts the 2-thiazoline residue into a 2-thiazoline-4-carboxy moiety at the carboxy terminus of **3b**, which is bound by BotH. Interestingly, the two BotH residues that form hydrogen bonds with the thiazoline carboxy group are in the position of the canonical oxyanion hole found in most ABHs. By binding, **3b** traps an ordered water molecule within hydrogen-bonding distance of the thiazoline's nitrogen and carboxy group. The side-chain of Asp7 is positioned such that it may serve as a base to abstract the C_α proton from itself, which triggers enamine formation and leads to proton transfer from the ordered water molecule to the thiazoline nitrogen. The

resulting hydroxide is still hydrogen bonded to the thiazoline's carboxy group and the thiazoline nitrogen proton. Abstraction of this proton by the hydroxide triggers reprotonation of the enamine by the side chain of Asp7 and results in epimerization. The intermediate appears to be long-lived enough to allow exchange of the proton abstracted by Asp7 with bulk solvent (Figure 4-S17). The inability of BotH to epimerize substrate with an Asp7Asn mutation suggests an essential role of the carboxylic acid side chain and thus substrate-assisted catalysis. As demonstrated by our hydrogen-deuterium-hydrogen exchange experiments, this reaction is fully reversible. It appears that lowering of the energetic barrier for epimerization by BotH is sufficient to supply the succeeding enzyme with sufficient substrate for complete turnover, even without changing the **3a** : **3b** ratios (see chapter 5). This situation is reminiscent of non-ribosomal peptide synthetases that can contain epimerase domains, which produce a mixture of D- and L-epimers of a particular amino acid. The succeeding enzyme (condensation domain) then provides the stereochemical resolution of the pathway through selective incorporation of the D-amino acid[41].
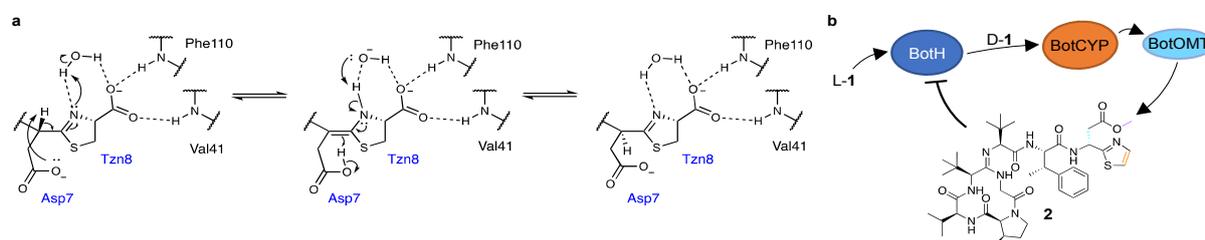
Figure 4-6 | Role(s) of BotH in bottromycin biosynthesis. a Proposed mechanism for the epimerization of **3b** to yield **3a**. BotH residues are labeled in black, substrate residues in blue. Hydrogen bonds are shown as dashed lines. Ordered water molecules surrounding the Asp7 carboxy group are not shown for clarity (see Figure 4-S8 for details). **b** Proposed role of BotH in self-resistance. A rise of intracellular bottromycin concentrations leads to an inhibition of BotH epimerase activity, which may in turn prevent self-poisoning of the producing strain and act as an intracellular buffer to store bottromycins.

The presence of BotH-like ABH protein encoding genes in over 100 BGCs encoding for the biosynthesis of diverse secondary metabolites places BotH as the founding member of a new subfamily of non-classical ABH enzymes. It will be fascinating to explore the functions of selected homologs, which may be able to epimerize non-Asp stereocenters across families of secondary metabolites via the unprecedented mechanism we have identified for bottromycin.

4.3 Supporting Information
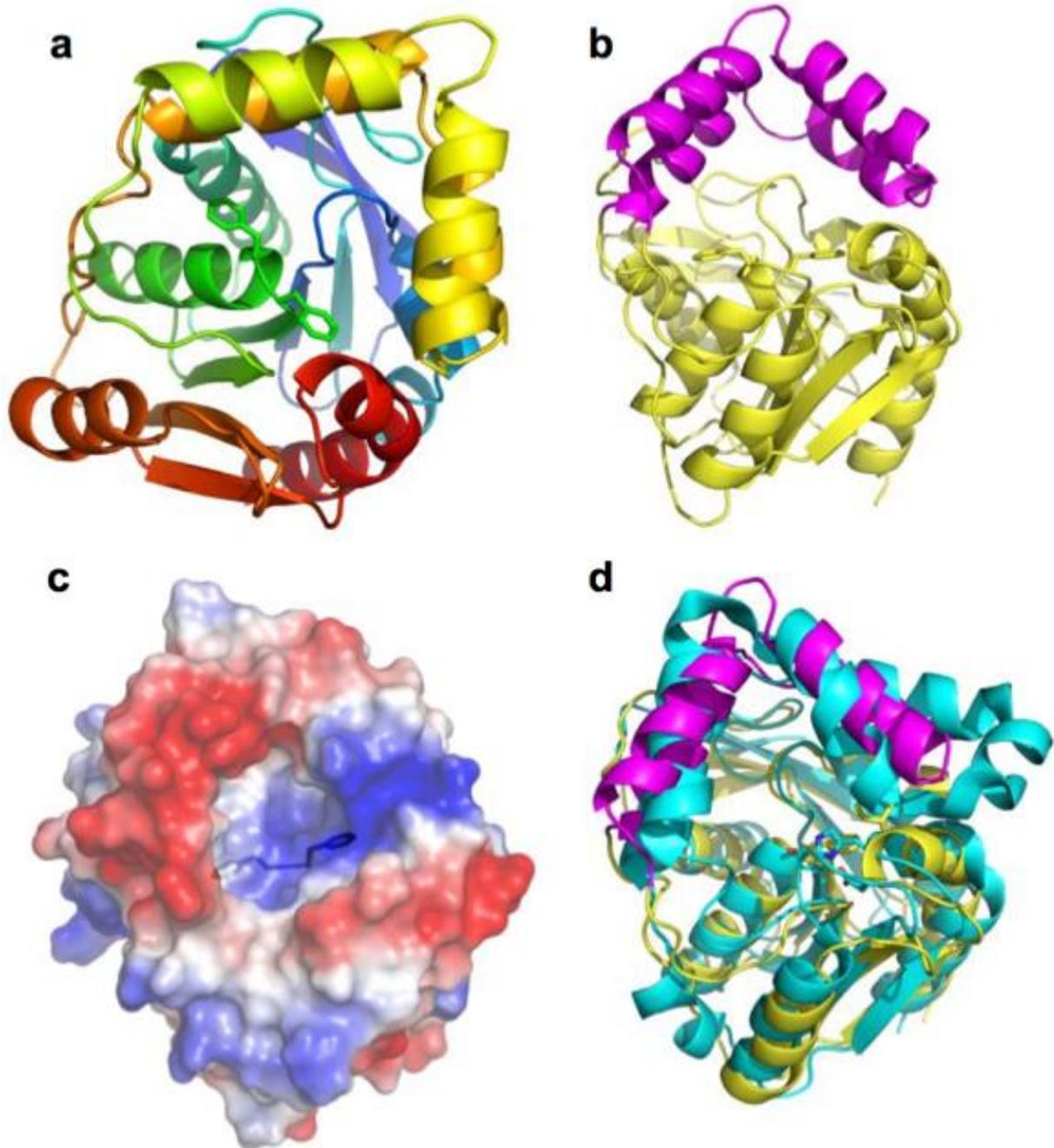
Figure 4-S1 | Crystal structure of BotH. **a** Cartoon representation of apo-BotH. **b** Cartoon representation of apo-BotH highlighting the V-shaped loop (magenta) positioned above the active site. **c** Electrostatic surface potential of BotH. The Phe-Phe motif at the active-site is shown as blue sticks. **d** Superposition of apo-BotH (yellow/magenta) with its closest structural homolog (PDB ID 2xua, cyan) gives a C_α-rmsd of 2.8 Å over 288 residues.
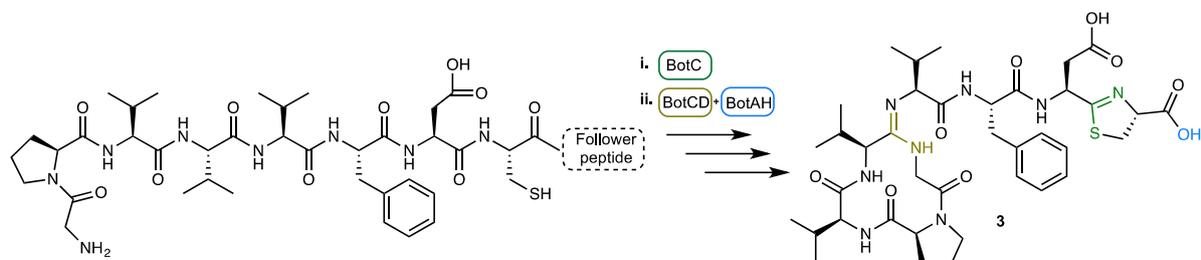
Figure 4-S2 | Biosynthesis of compound **3**. Production of **3** can be accomplished in a two-stage, one-pot reaction with quantitative yields.
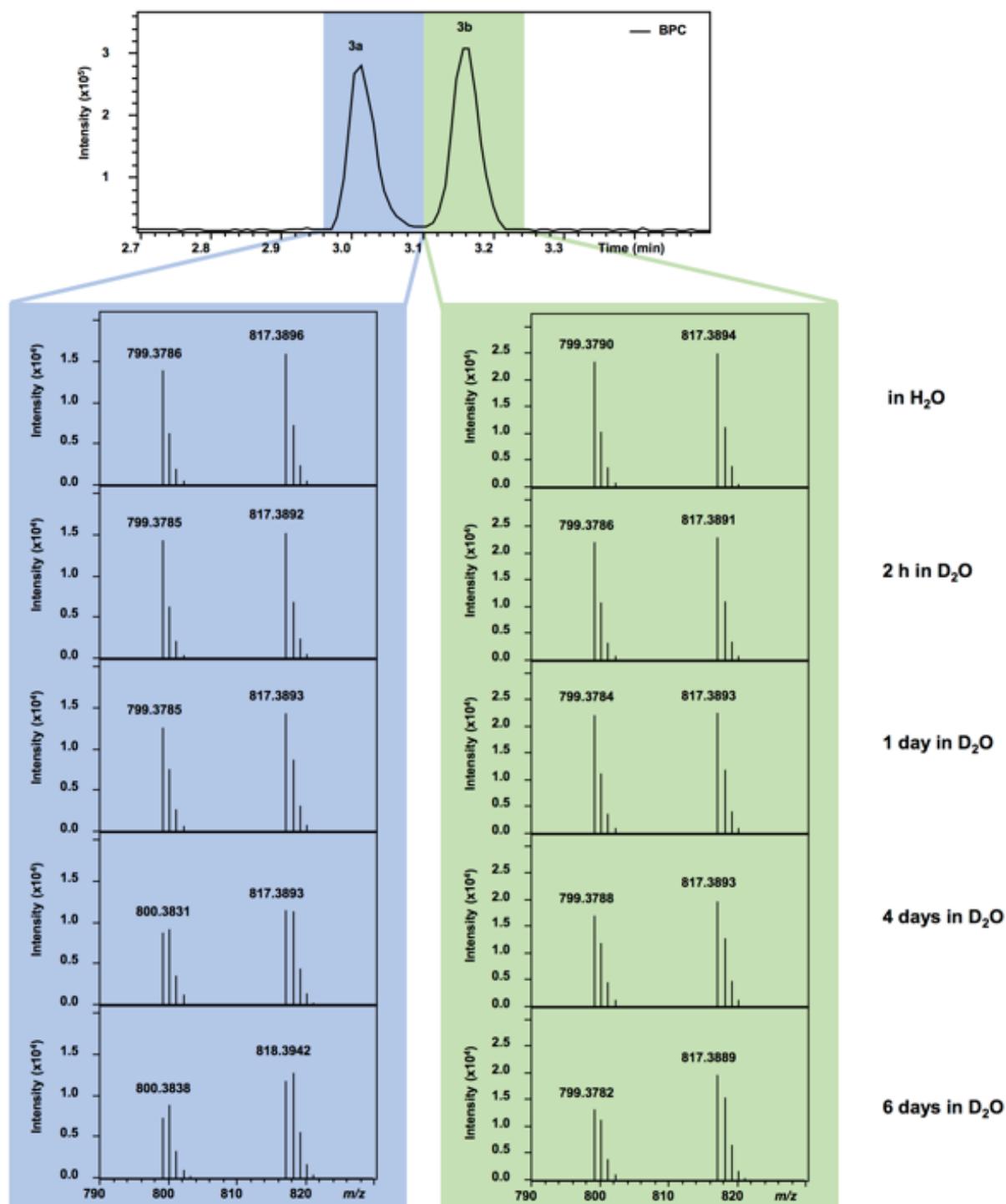
Figure 4-S3 | Spontaneous (non-enzymatic) epimerization of the **3a/b** Asp Cα. A roughly racemic mixture of **3a/b** was dissolved in D₂O to analyze epimerization. Samples were taken after 2 h, 1 day, 4 days and 6 days. In the acidic (0.1% FA) LC conditions the thaizoline partly (about 50%) re-opens, which leads to the addition of H₂O (+ 18.015 Da). Shown mass spectra are averaged from 2.96 - 3.10 min and 3.10 - 3.24 min respectively, as the retention times of **3a/b** and the respective compound with re-opened thiazoline slightly differ. Shown is the base peak chromatogram (BPC) in H₂O. No changes in the shape of the BPCs were observed at the different time points. Representative experiments were repeated independently three times with similar results. BPC = Base Peak Chromatogram.
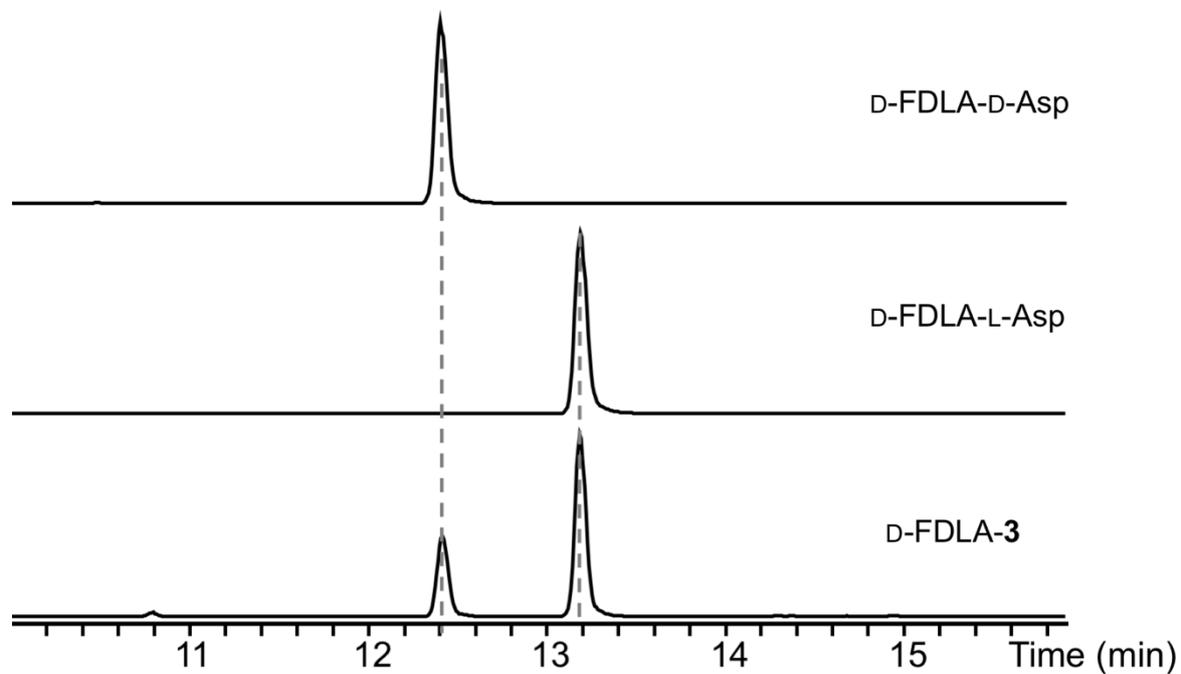
Figure 4-S4 | Marfey's method to confirm the stereochemistry of the Asp residue in **3a/b**. The analysis confirmed the presence of D-Asp (**3a**) and L-Asp (**3b**) and allowed us to assign **3a** to the less abundant peak. Shown are the extracted ion chromatograms (EICs ± 5 ppm) for the FDLA-Asp derivate (428.1412 *m/z*). The measurements were repeated independently three times with similar results.
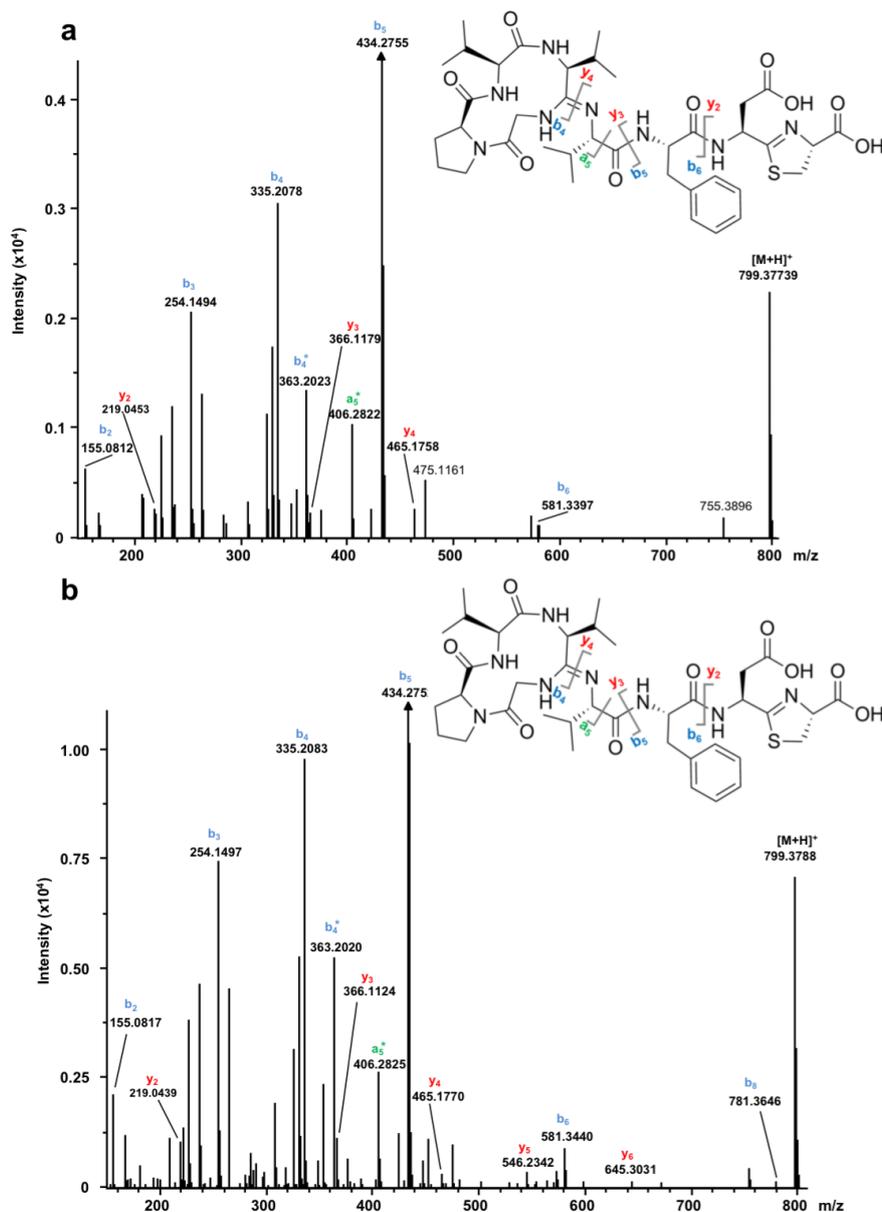
Figure 4-S5 | Structural determination of compound 3. a MS^2 fragmentation spectra of 3a (D-Asp). **b** MS^2 fragmentation spectra of **3b** (L-Asp). The associated peak lists and the structures of fragments b$_2$ and b$_3$ can be found in Table 4-S2. The experiments shown in this figure were repeated three times independently with similar results.
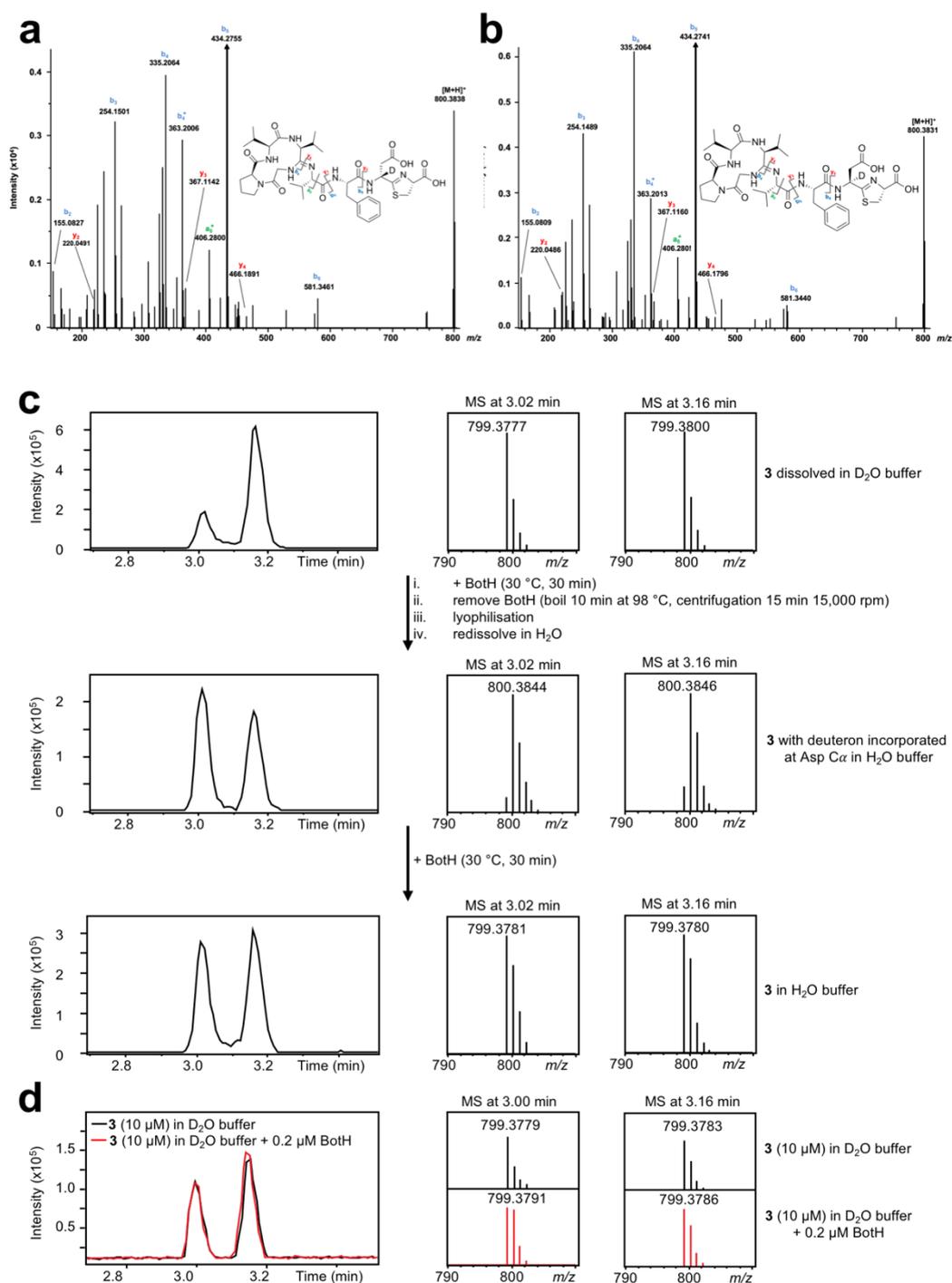
Figure 4-S6 | Analysis of deuterium exchange at **3** Asp Cα. **a** MS^2-fragmentation of deuterated **3a**. The associated peak lists and the structures of fragments b_2 and b_3 can be found in Table 4-S3. **b** MS^2-fragmentation of deuterated **3b**. The associated peak lists can be found in Table 4-S3. **c** Performing a BotH reaction in D$_2$O buffer leads to deuteron incorporation at the Asp Cα. In reverse, performing a BotH reaction with deuterated **3a/b** in H$_2$O buffer leads to the exchange of the deuteron with solvent protons. Shown are the base peak chromatograms (BPC). **d** At low BotH concentrations no change of the **3a** : **3b** ratio was observed within the time-frame of the experiment (7 min), but when the reactions are performed in D$_2$O buffer, a deuteron is incorporated. Shown are the BPCs. Representative experiments were repeated independently seven times with similar results.
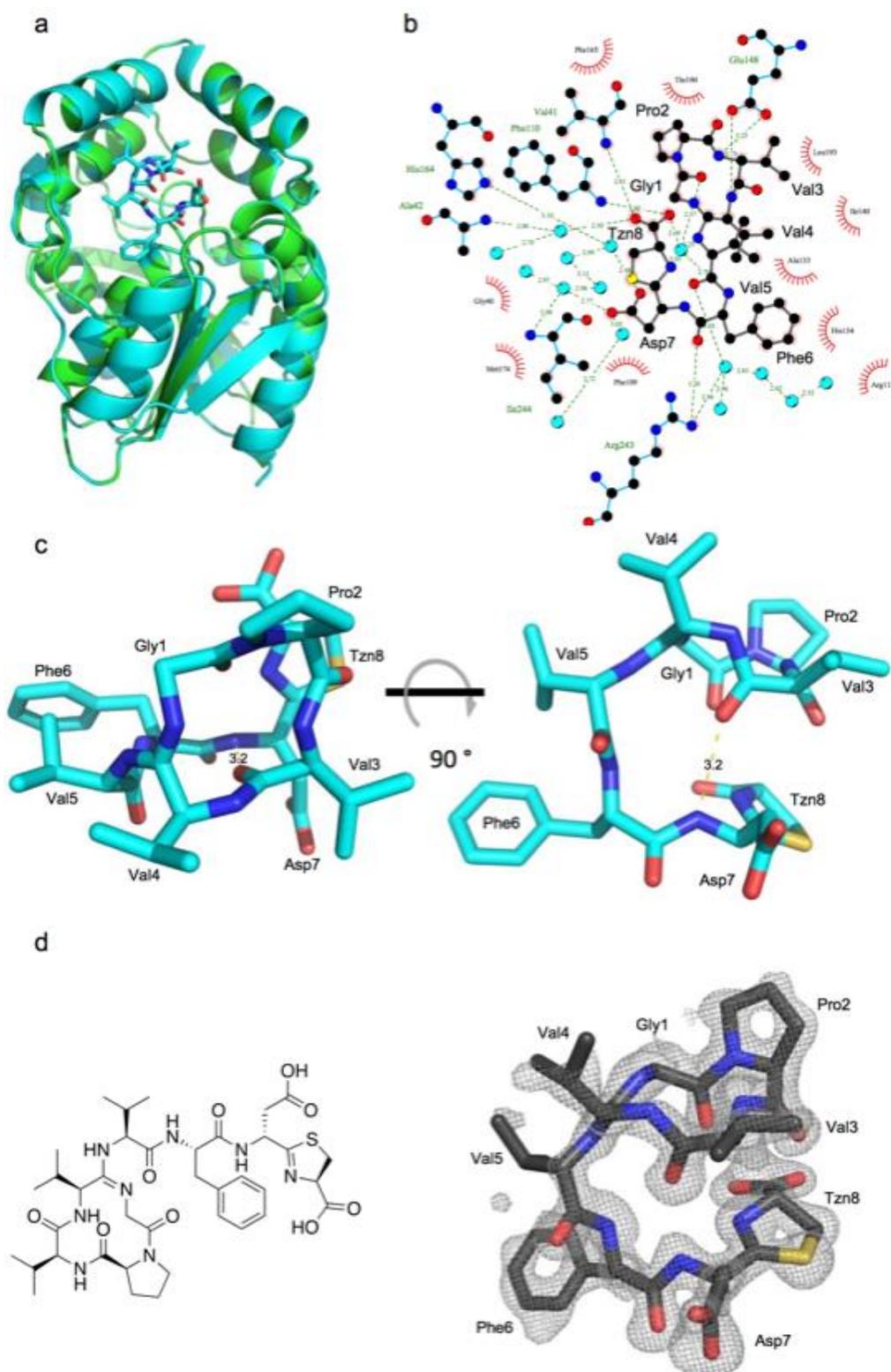
Figure 4-S7 | Crystal structure of BotH in complex with 3. a Superposition of the BotH apo structure (green) with the BotH-**3a** complex structure (cyan). **3a** is shown as cyan sticks. **b** LigPlus diagram of the interactions between **3a** and BotH. **3a** is shown with bonds in grey, BotH with bonds in cyan. Water molecules are depicted as cyan spheres, intermolecular hydrogen bonds shown as dashed lines with distances in Å and hydrophobic contacts are shown as red spoked arcs. **c 3a** as observed in the complex structure shown as sticks. The intramolecular hydrogen bond is shown as a dashed line with the distance given in Å. **d 3a** as observed in the complex crystal structure (left) and Polder map of **3a** contoured at 3 σ (right).
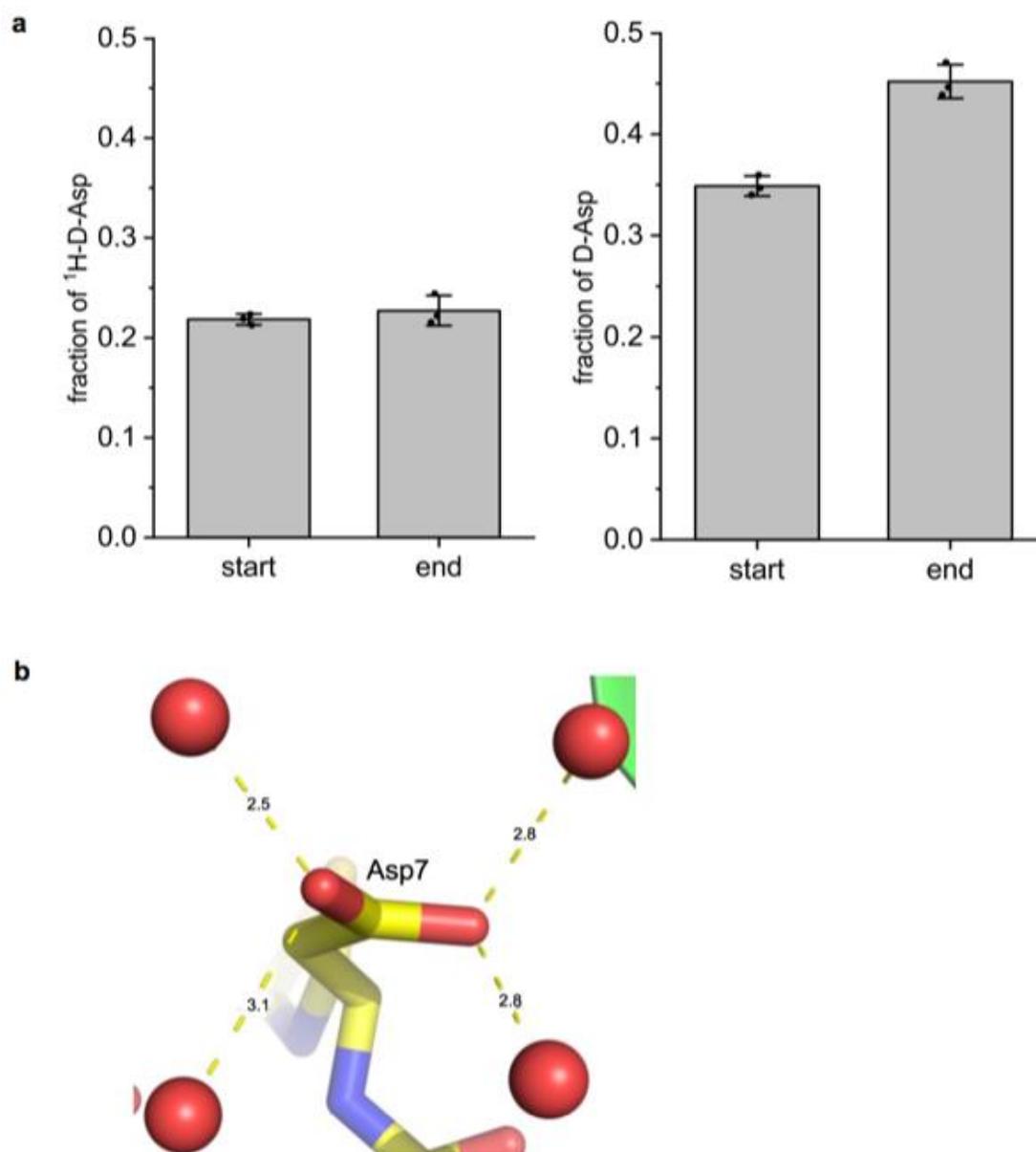
Figure 4-S8 | BotH single-turn over reaction and the role of water in epimerization. **a** Epimerization under single-turnover conditions in D_2O. While the ratio of D-Asp increases significantly (right), the ratio of ^{1}H-D-Asp remains unchanged within experimental error. These data indicate that the majority of protons abstracted by the carboxy group of Asp7 during enamine formation are exchanged with bulk solvent before reprotonation of the enamine intermediate, resulting in deuteron incorporation. Experiments were carried out in triplicate, shown are means ± SD (n=3) and black dots indicate results for individual measurements. Differences in the fraction of D-Asp at the start and end of the experiment were calculated to be extremely significant (p-value = 0.0008) using an unpaired two tailed t-test. **b** Four ordered water molecules surround the carboxy-group of Asp7 and may facilitate proton/deuteron exchange with bulk solvent during the epimerization reaction. O – O distances are given in Å.
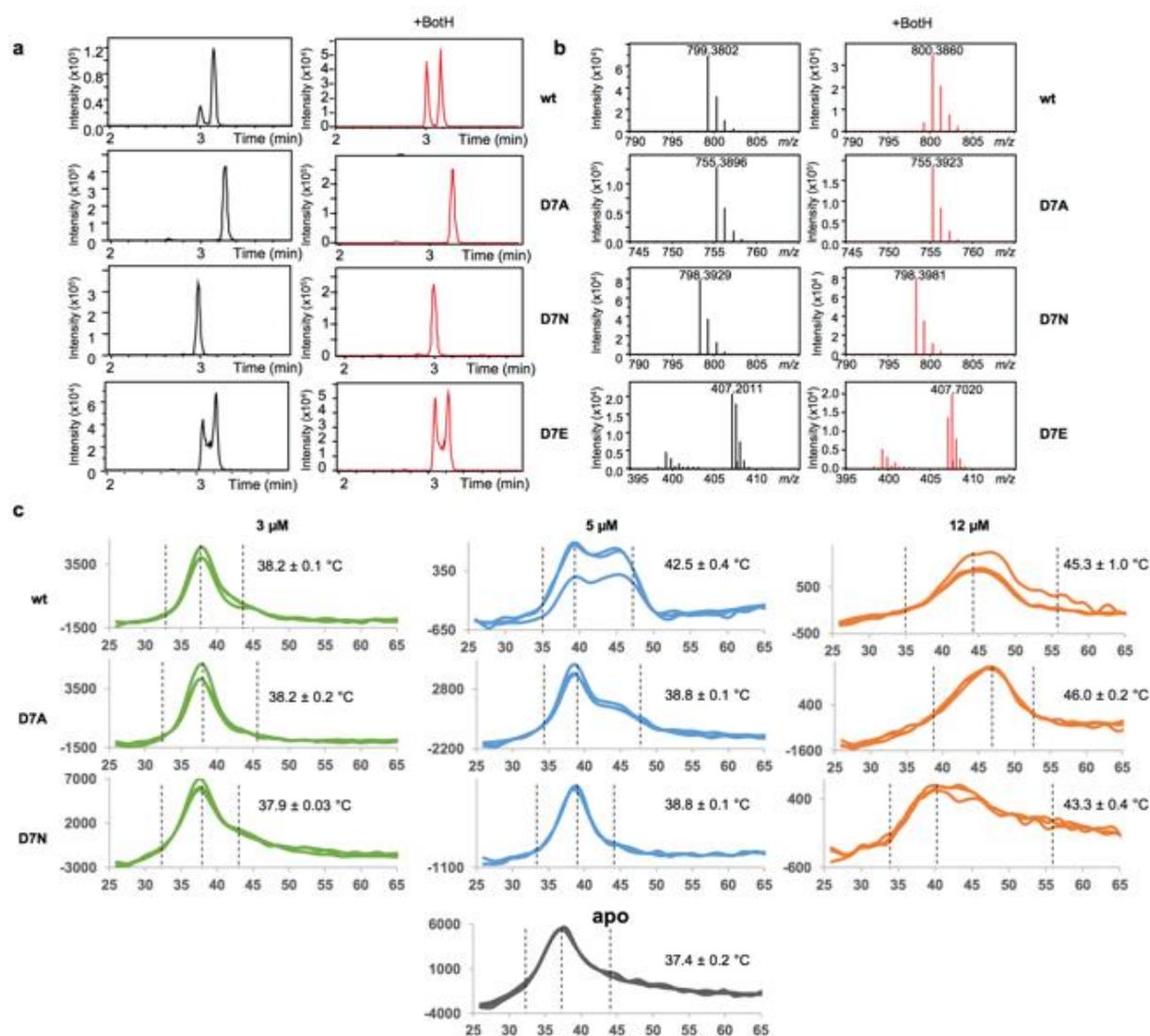
Figure 4-S9 | Epimerization of Asp7 mutants and their stabilization effect on BotH. **a** and **b** Epimerization of Asp7 mutants. **a** Graphs show extracted ion chromatograms (EICs, calculated mass see Table 4-S4 ± 5 ppm) of possible BotH substrates with (red) and without (black) addition of BotH in H_2O. No change of the retention time nor an additional peak could be observed in the respective EICs after addition of BotH for mutants Asp7Ala and Asp7Asn, while **3a/b** (wt) and Asp7Glu could be epimerized by BotH. **b** Mass spectra corresponding to **a** showing the incorporation of a deuteron for wt and Asp7Glu by BotH in D_2O buffer but not for Asp7Ala and Asp7Asn. Representative experiments were repeated independently three times with similar results. **c** Thermal shift assays of BotH incubated with different concentrations of wt, Asp7Ala and Asp7Asn substrates. As can be seen, all three substrates lead to a shift in melting temperature at comparable concentrations indicating that all three substrates bind to BotH. The double-peak at 5 μM **3a/b** concentration may indicate distinct melting temperatures for **3a**- and **3b**-complexes. Melting temperatures displayed within the graphs are means ± SD (n=3).
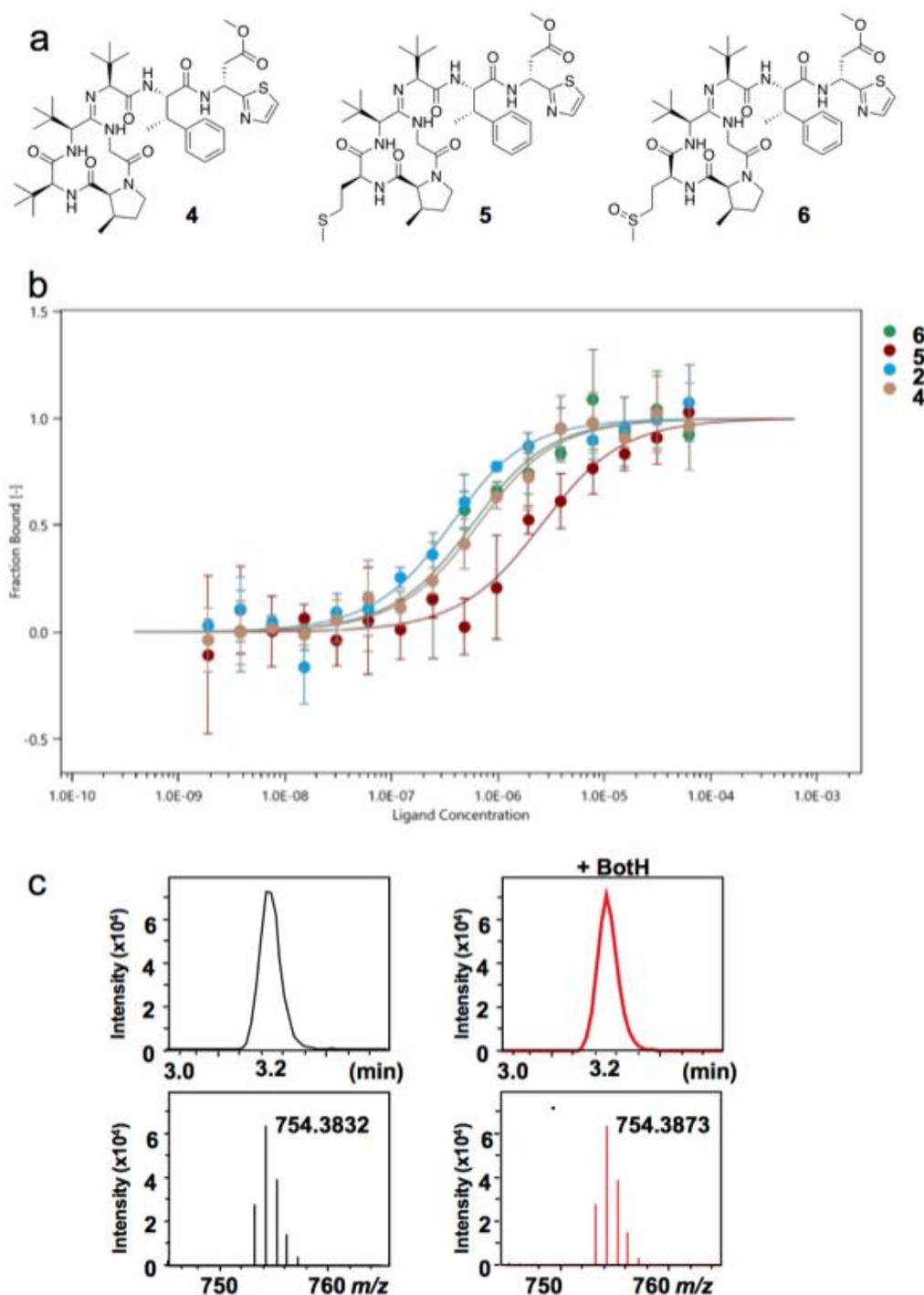
Figure 4-S10 | Binding affinity of BotH-bottromycin derivatives and effect of decarboxylation of **3a** on epimerization. **a** Chemical structure of the three bottromycin derivatives used in this study. **b** MST measurements to determine the affinities of bottromycin A_2 (**2**) and the three derivatives shown in a for BotH. The affinity of BotH for **2** is 232 ± 81 nM. The additional methyl group at the C_β position of Val3 of (**4**)[42] reduced the affinity to a K_D of 472 ± 62 nM, while the Val3Met mutation of **5**[42] reduced the affinity by an order of magnitude to a K_D of 3.2 ± 1.9 μM. Oxidation of the methionine sulfur of **5** (**6**) resulted in a K_D of 459 ± 178 nM. A MS^2 spectrum for **6** can be found in Figure 4-S20. Each curve represents three independent samples, data points represent the mean and the error bars represent standard deviations. **c** Oxidatively decarboxylated **3a**, with deuteron incorporated at the Asp $C\alpha$ ([M+H]$^+$calc.$_{mono.}$: 754.3815 Da), was incubated in H_2O buffer with (red) and without (black) BotH. BotH is unable to epimerize the decarboxylated **3a** as no change of the isotope pattern nor the retention time is observed. This experiments were performed in triplicate with similar results.
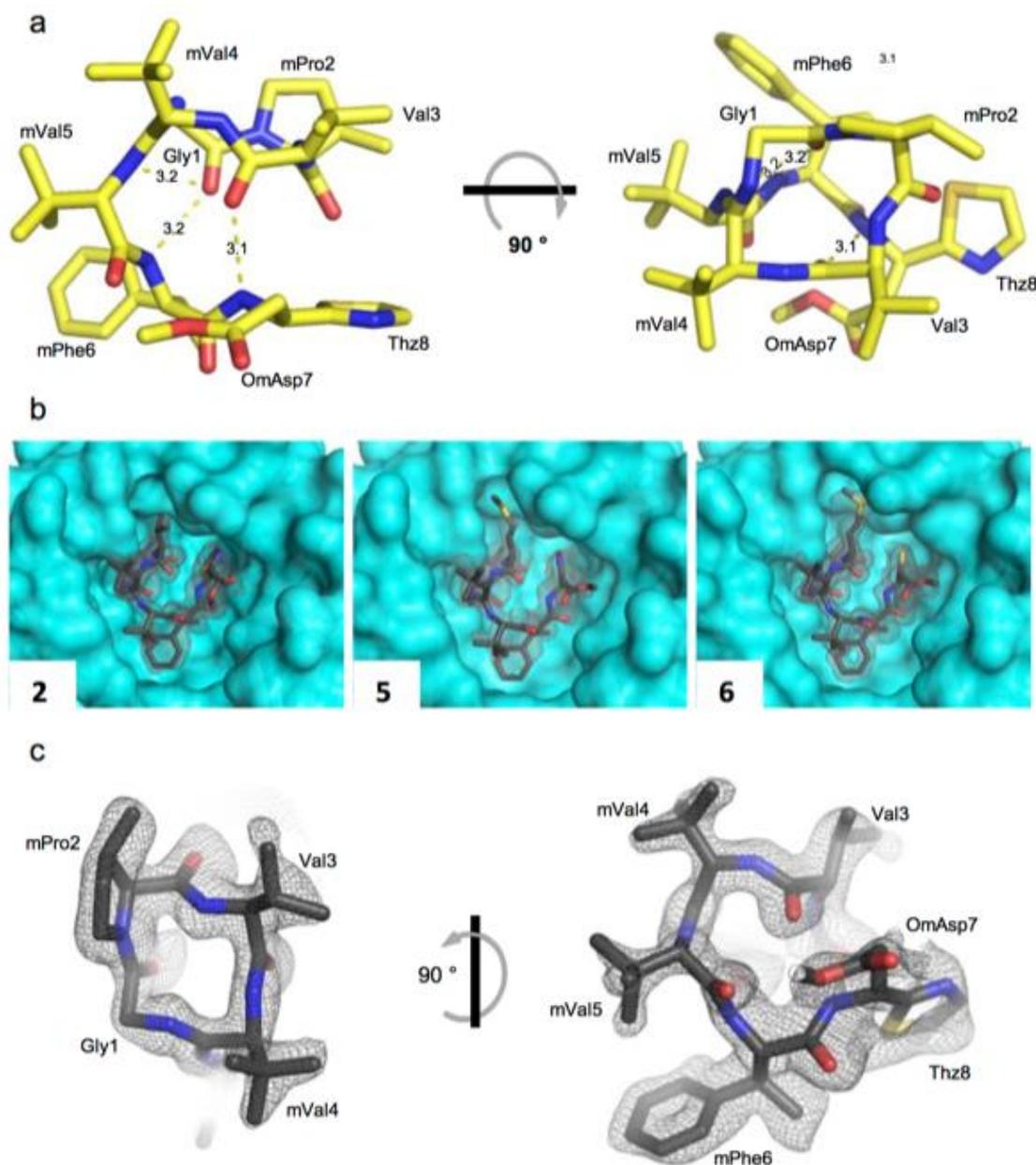
Figure 4-S11 | Crystal structure of BotH in complex with bottromycin A$_2$ and its derivatives. a Bottromycin A$_2$ (**2**) as observed in the active-site of BotH. Intramolecular hydrogen bonds are shown as dashed lines and distances given in Å. **b** Polder maps (grey isomesh) for bottromycin A$_2$ (**2**) and bottromycin derivatives **5** and **6** bound at the BotH active-site. BotH is shown as a cyan surface representation, the ligands as sticks. The polder maps were contoured at 3 (**2**), 2.5 (**5**) and 3 (**6**) σ. **c** Close-up polder map of bottromycin A$_2$ from **b** shown as a grey isomesh contoured at 3 σ. The structures also provide insights into the varying affinities of the bottromycin derivatives: Based on the BotH-**2** complex structure, the additional methyl group of **4** clashes with BotH Glu148, which is involved in hydrogen bonding to the compound and BotH residue Tyr160, which results in a 2-fold weaker affinity. The BotH-**5** complex shows that the hydrophobic side-chain of the methionine is forced into a very narrow, highly polar opening lined by Glu148 and His164, which may explain the marked loss in affinity. In the BotH-**6** complex structure, the oxidation of the methionine sulfur triggers a rotation of the His164 side-chain, which allows the protein to easily accommodate the bulky ligand side-chain and enables the formation of a new hydrogen bond between the substrate (methionine sulfoxide oxygen) and the side-chain of BotH Arg168. This rationalizes the increase in affinity when compared to unoxidized **5**.

G₁	P₂	V₃	V₄	V₅	F₆	D₇	C₈
	A	L	A	T	W	A	
	G		L	L	A	N	
				A	Y	E	
				E			

Figure 4-S12 | BotA point mutations that are compatible with the enzymes used to generate the BotH substrate (IpoC/PurCD/PurAH).
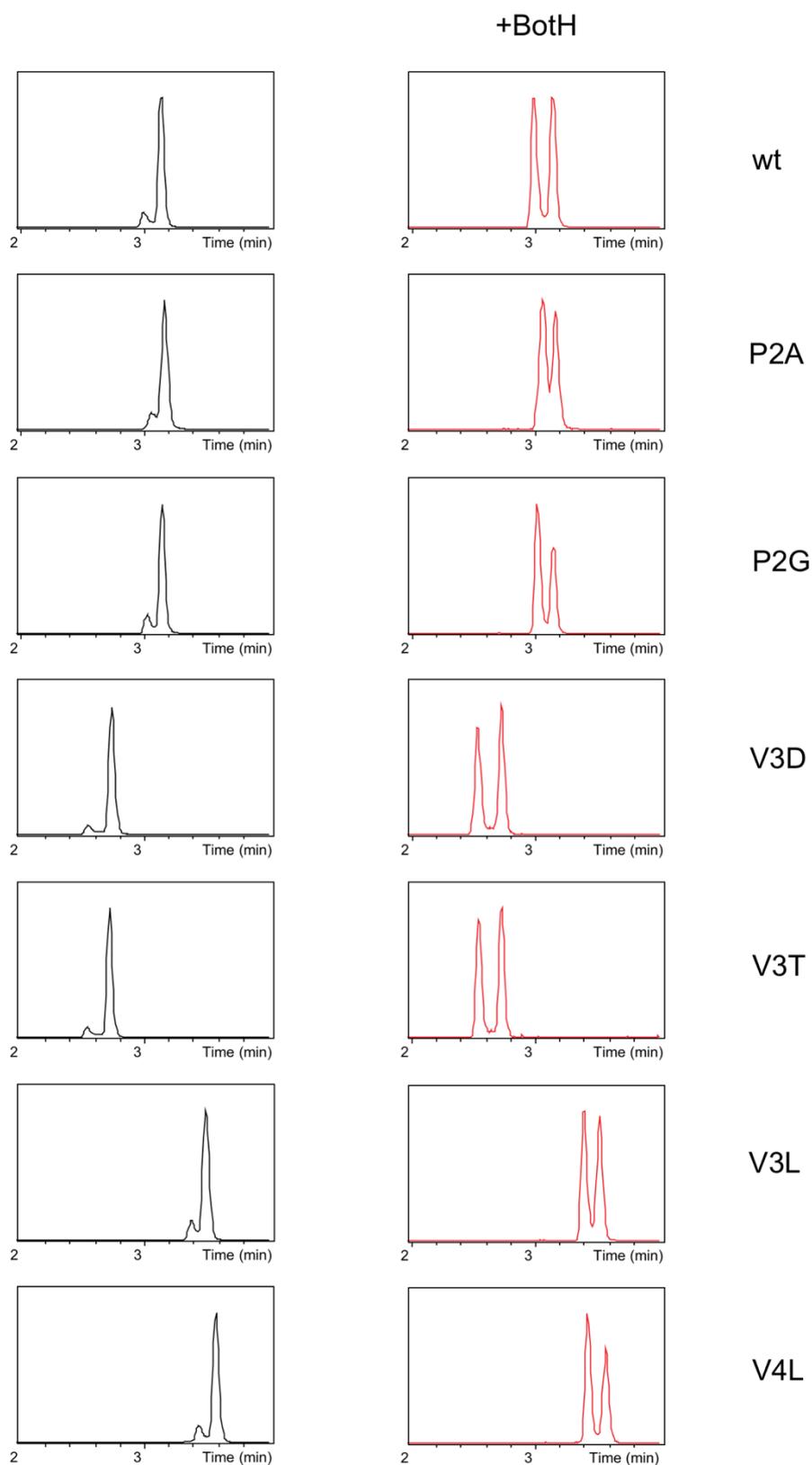
Figure 4-S13 | The effects of point mutations in the BotA core peptide on BotH activity. Each mutant was analyzed by HRLC-MS in the absence (left column, black) and presence (right column, red) of BotH. Shown are EICs for the monoisotopic [M+H]$^+$ mass ± 5 ppm. A complete list of masses is given in Table 4-S4. As can be seen, all mutations in positions 2–4 are accepted by the enzyme. The experiments shown in this figure were repeated independently at least three times with similar results.
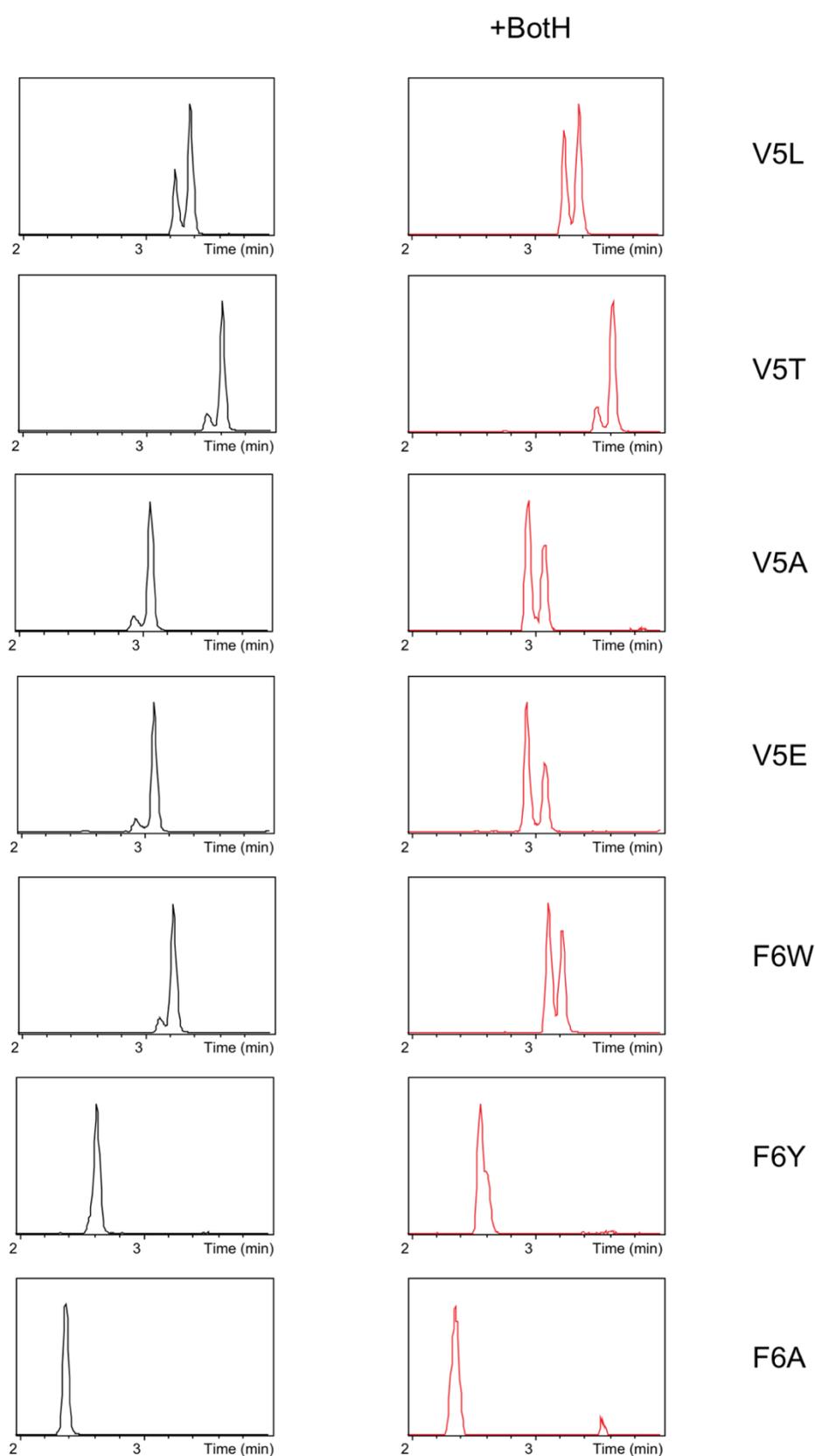
Figure 4-S14 | The effects of point mutations in the BotA core peptide on BotH activity. Each mutant was analyzed by HRLC-MS in the absence (left column, black) and presence (right column, red) of BotH. Shown are EICs for the monoisotopic [M+H]$^+$ mass ± 5 ppm. A complete list of masses is given in Table 4-S4. Mutations V5T and F6A are not accepted by the enzyme. The experiments shown in this figure were repeated independently at least three times with similar results.
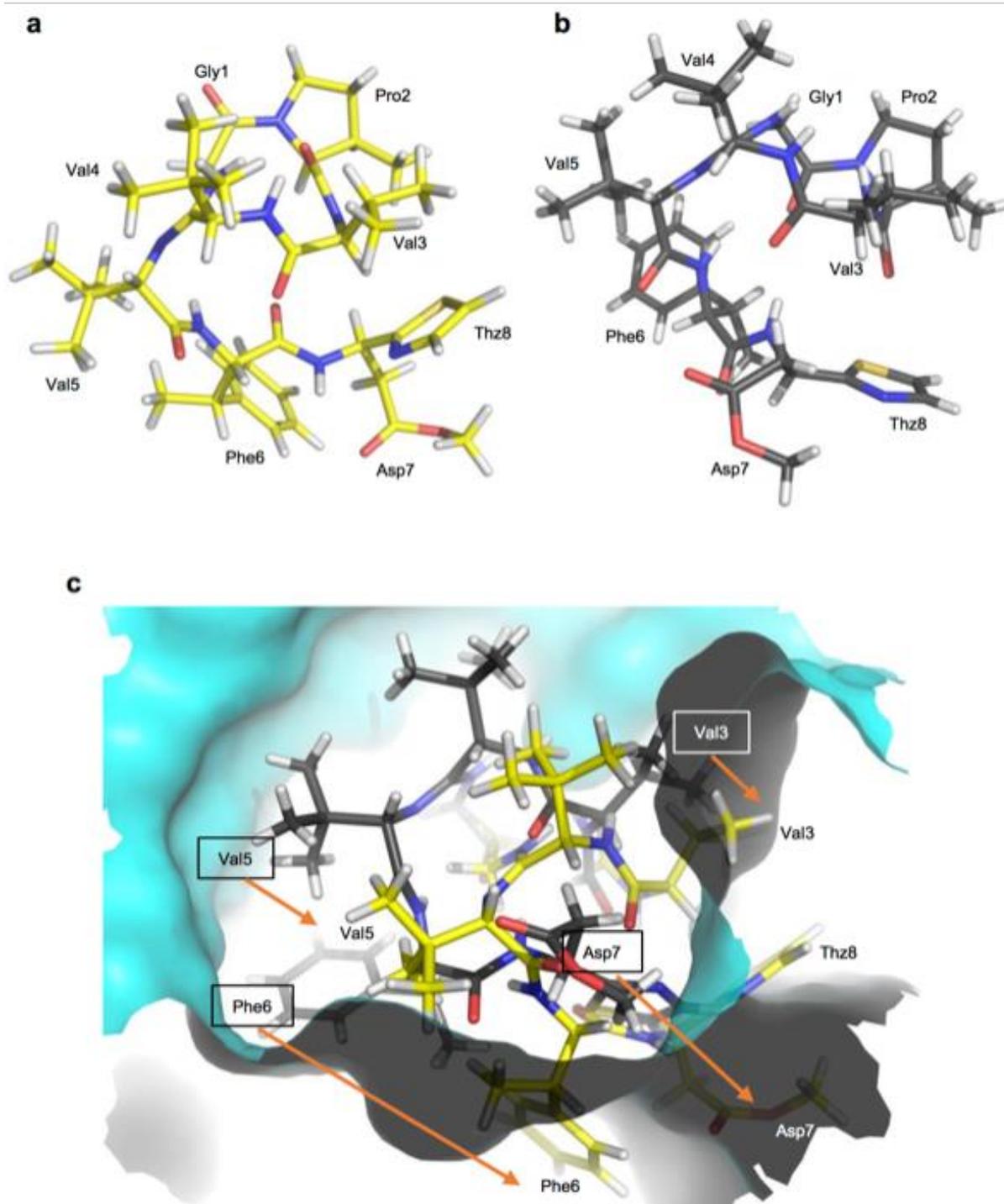
Figure 4-S15 | Comparison of bottromycin A_2 solution structure to bottromycin A_2 observed in BotH complex structure. **a** Solution NMR structure of bottromycin A_2 (Coordinates available upon request from Prof. Hiroaki Gouda[35]) shown as yellow sticks. **b** Crystal structure of bottromycin A_2 bound to BotH shown as grey sticks. The same orientation and magnification is used for **a** and **b**. Residues are labeled, Thz8 = Thiazole in position 8. **c** Superposition of the solution NMR structure of bottromycin A_2 and the bottromycin A_2 bound to BotH. Same color scheme as **a** and **b**, Pro2 was used as the reference residue. As can be seen, binding to BotH causes a significant conformational change and introduces strain. The more relaxed solution state (yellow) clashes with the protein in positions Val3, Phe6 and Thz8. Boxed labels belong to the complex crystal structure, orange arrows indicate movement required to reach unbound state.
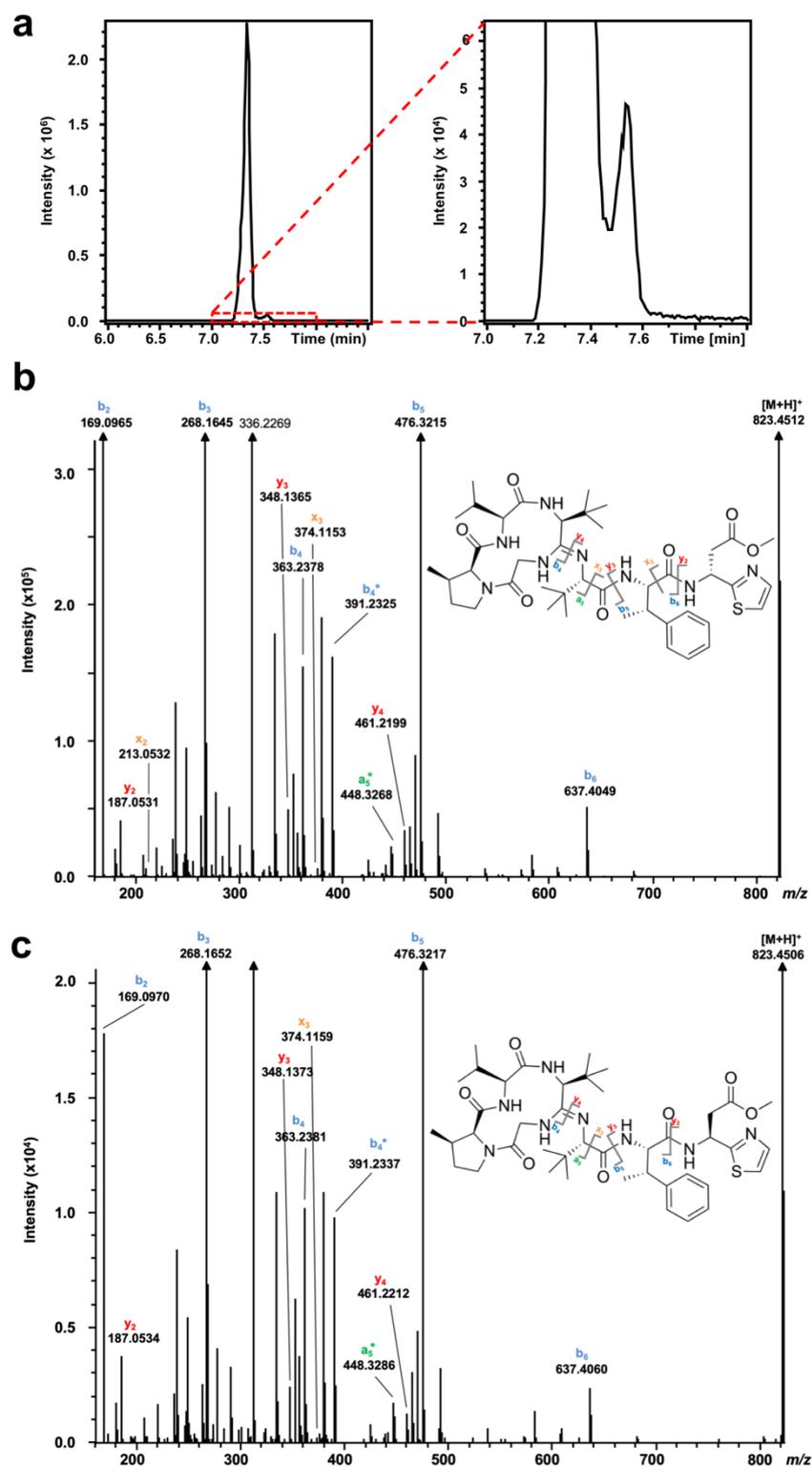
Figure 4-S16 | Identification of D/L-Asp bottromycin A₂. a EIC of bottromycin A₂ (**2**) crude extract (823.4540 ±1 ppm) **b** MS² fragmentation of bottromycin A₂ (**2**) main peak (O-Met-D-Asp-containing compound; retention time 7.3 min). The associated peak lists and the structures of fragments b₂ and b₃ can be found in Table 4-S5. **c** MS² fragmentation of minor bottromycin peak (most probably O-Met-L-Asp-containing compound; retention time 7.55 min). The associated peak lists can be found in Table 4-S5. The experiments shown in this figure were repeated three times independently with similar results.
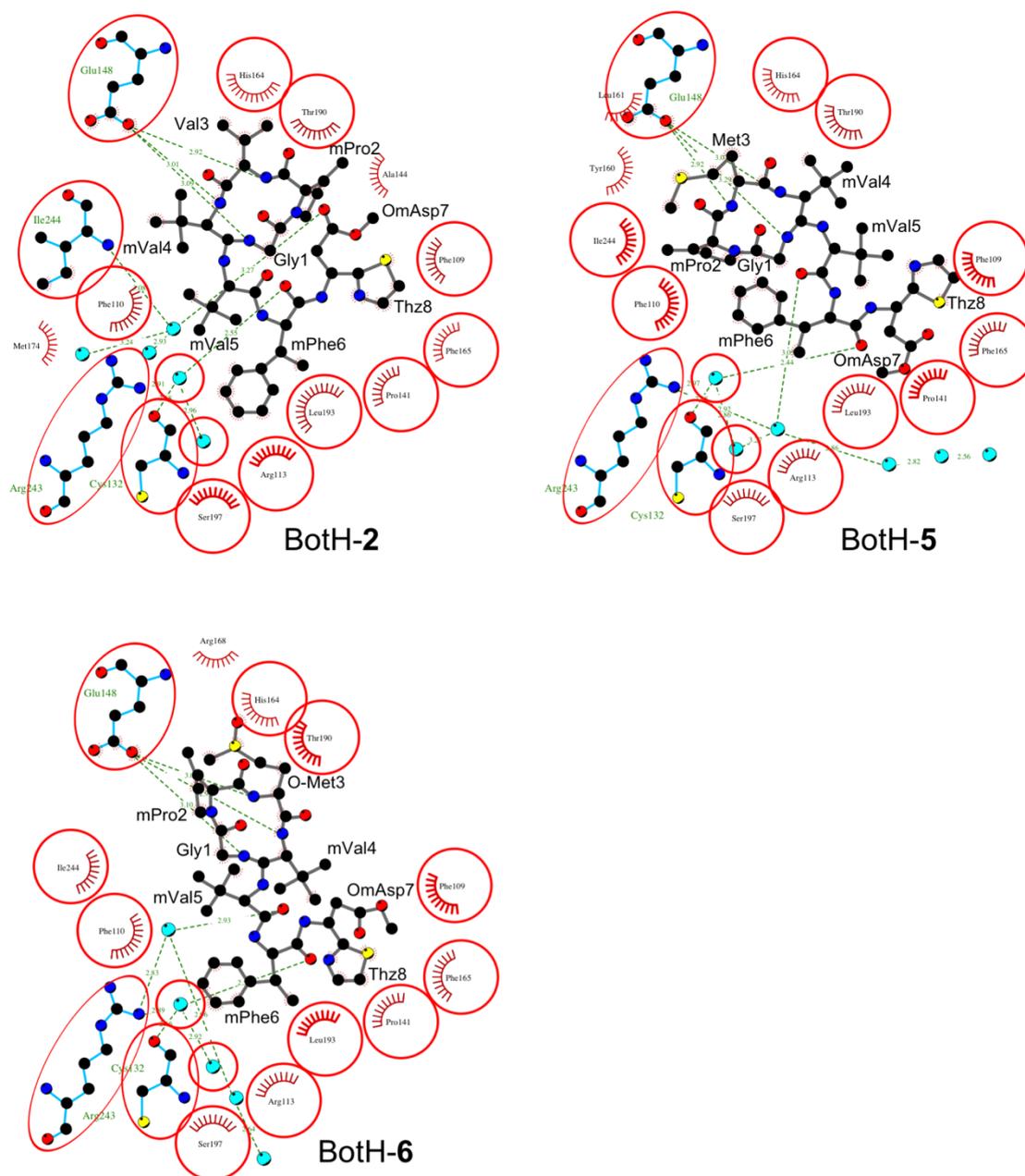
Figure 4-S17 | LigPlus diagrams of the three BotH complex structures with bottromycin A₂ (**2**) and two analogs (**5** and **6**). Ligands are shown with bonds in grey, BotH with bonds in cyan. Water molecules are depicted as cyan spheres, intermolecular hydrogen bonds are shown as dashed lines with distances in Å and hydrophobic contacts are shown as red spoked arcs. Residues and water molecules engaged in the same interactions in all three structures are circled in red.
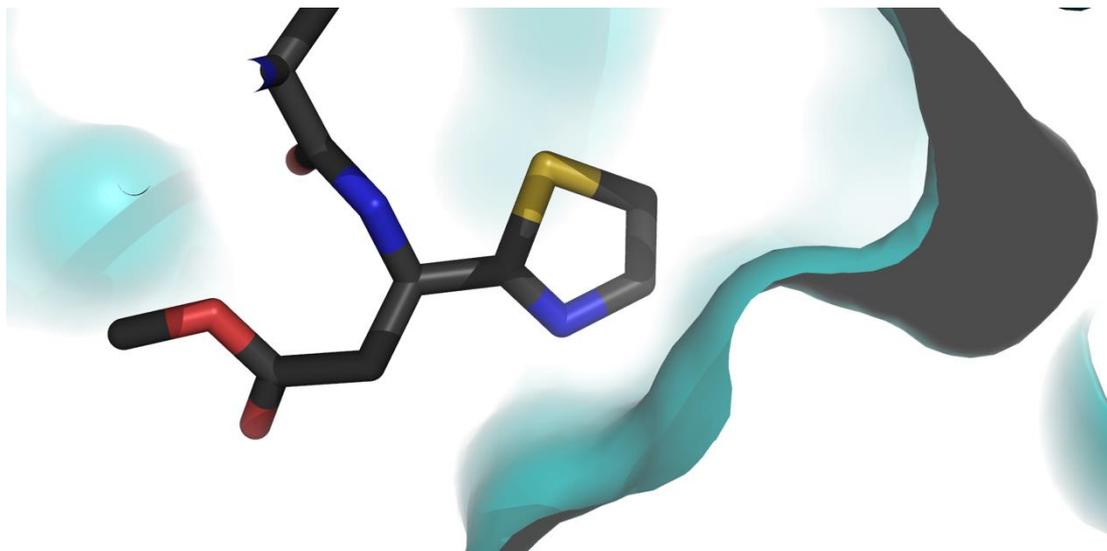
Figure 4-S18 | Close-up of the thiazole found in the BotH-bottromycin A_2 complex structure. Bottromycin A_2 is shown as sticks, BotH as a cyan surface representation. As can be seen, flipping the thiazole 180° such that sulfur and nitrogen swap positions does not lead to a clash and is thus also an allowed orientation.
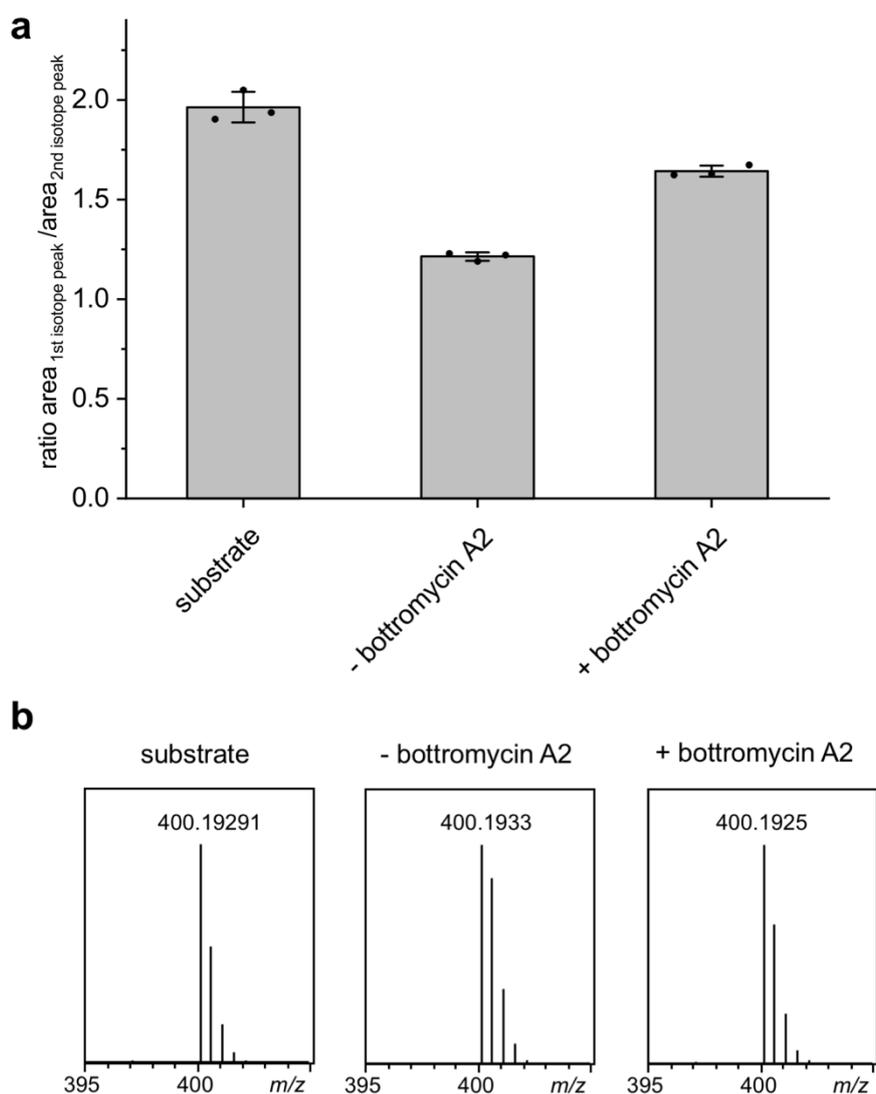
Figure 4-S19 | Orthosteric inhibition of BotH epimerization by bottromycin A_2. The effect of bottromycin A_2 on epimerization. The substrate **3a/b** (10 μM in D_2O buffer) was incubated with BotH (0.2 μM) with and without the presence of bottromycin A_2 (50 μM) for 7 min at 30 °C. BotH activity was measured by the incorporation of a deuteron, which leads to an increase of the 2nd isotope peak. In the presence of bottromycin A_2, the incorporation of a deuteron in **3a** by BotH is significantly lower. **a** Ratios of the area of the 1st to the 2nd isotope peak of $[3a+2H]^{2+}$. Shown are means ± SD (n=3). Differences were calculated to be extremely significant (p-value < 0.0001) using an unpaired two tailed t-test. Black dots represent individual data points. **b** Mass spectra (3a, 3.02 min) of the substrate (**3a/b** in D_2O buffer), the substrate after 7 min incubation with BotH and the substrate after 7 min incubation with BotH in presence of bottromycin A_2. The experiments were repeated three times independently with similar results.
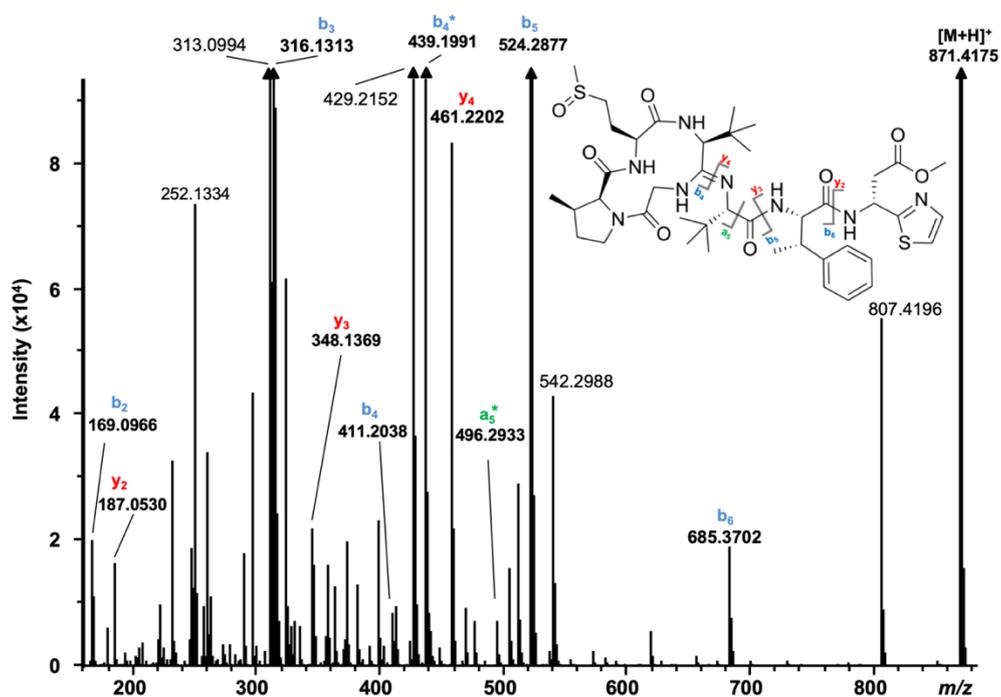
Figure 4-S20 | MS² fragmentation spectrum of the bottromycin analog 6. The associated peak lists and the structures of fragments b_2 and b_3 can be found in Table 4-S6. The experiments shown in this figure were repeated three times independently with similar results.
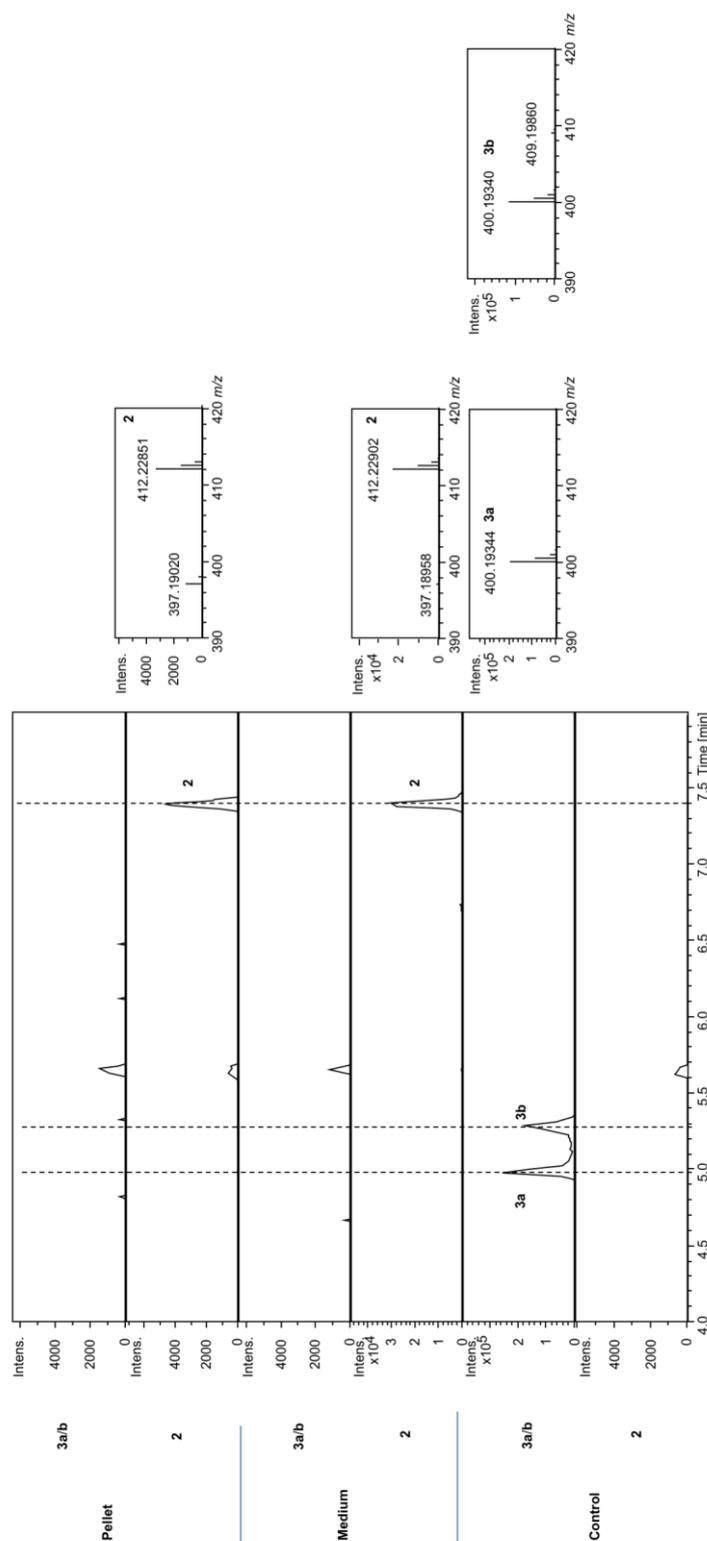
Figure 4-S21 | Pull-down data for **2** and **3a/b** with BotH. His$_6$-tagged BotH was incubated with equal volumes of either lysed cell pellet supernatant (top) or spent culture medium (middle) of the heterologous bottromycin producer *S.* sp. DG2-kmP41hyg[42]. Purified **3a/b** was used as a control (bottom). The BotH was then removed from the mixture using magnetic Ni^{2+}-NTA beads, washed thoroughly and eluted from the beads. Denaturing the enzyme with ACN liberated bound ligands. Only bottromycin A$_2$ (**2**) and traces of **4** (data not shown), but not the BotH substrate **3a/b** could be found in the experimental samples. Shown are the EICs (right) and mass spectra for labeled peaks (right) for either **3a/b** or **2** ± 5 ppm using the [M + 2H]$^{2+}$ masses. The experiments were repeated three times independently with similar results.
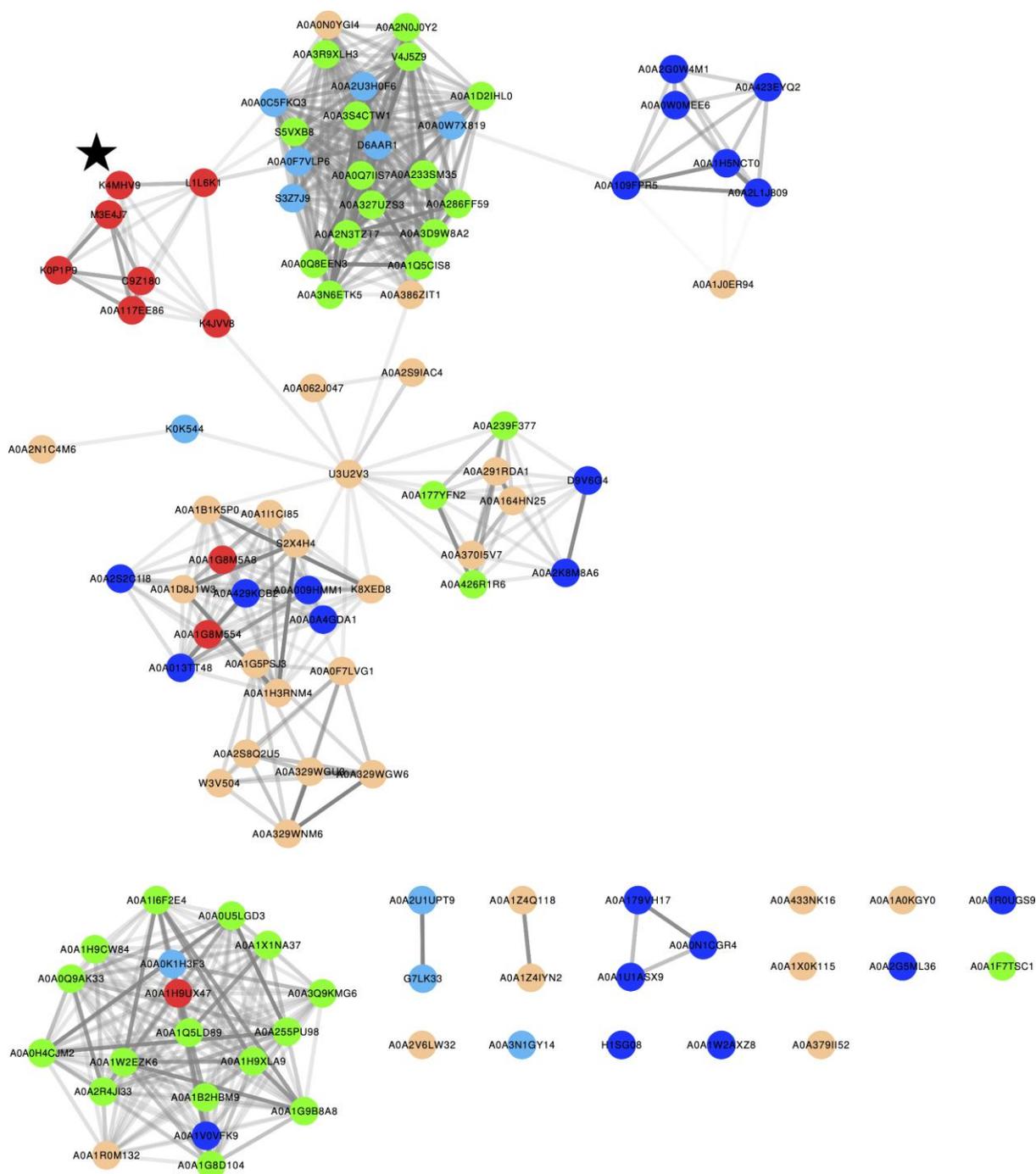
Figure 4-S22 | Sequence similarity network (SSN) of BotH-like proteins. SSN for BotH-like proteins lying in or near (closer than 1000 nt) biosynthetic gene clusters (BGCs) is depicted. Nodes represent individual proteins (Accession number given) and colored according to the cluster type (red: RiPP, blue: NRPS, green: PKS, sky: NRPS/PKS, cantaloupe: other), edges represent similarity relationships and are colored according to sequence identity (lighter: less identical, darker: more identical). Edges corresponding to sequence identity below 15% are omitted. BotH is indicated with a star (colored black).

Table 4-S1 | Data collection and refinement statistics.

	BotH_SeMet	BotH_apo	BotH-3a	BotH-2	BotH-5	BotH-6
PDB ID	6T6H	6T6H	6T6X	6T6Y	6T6Z	6T70
Data collection						
Space group	I222	I222	I222	I222	I222	I222
Cell dimensions						
a, b, c (Å)	66.7, 80.1, 88.7	66.7, 80.1, 88.7	66.7, 80.1, 89.5	67.0, 80.1, 89.4	66.7, 79.5, 89.0	66.5, 79.4, 88.4
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength (Å)	0.97941					
Resolution (Å)	1.70 (1.73-1.70) *		1.25 (1.27-1.25)	1.40 (1.42-1.40)	1.70 (1.73-1.70)	1.58 (1.61-1.58)
R_{sym} or R_{merge}	7.2 (50.0)	3.5 (63.4)	7.8 (72.5)	6.9 (80.0)	10.4 (77.8)	6.0 (71.2)
R_{pim}	1.4 (9.6)	1.6 (31.2)	4.1 (40.0)	2.9 (33.1)	4.9 (36.4)	3.2 (38.1)
CC(1/2)	1.000 (0.973)	1.000 (0.789)	0.995 (0.774)	0.998 (0.801)	0.997 (0.790)	0.999 (0.773)
$I / \sigma I$	47.6 (8.7)	21.4 (2.5)	11.4 (4.2)	14.5 (2.6)	10.5 (2.4)	13.5 (2.0)
Completeness (%)	100 (100)	98.7 (93.5)	99.0 (100)	99.4 (99.5)	99.2 (99.0)	99.8 (99.9)
Redundancy	26.7 (27.8)	6.4 (5.7)	4.3 (4.1)	6.6 (6.7)	5.3 (5.4)	4.3 (4.4)
Refinement						
Resolution (Å)		44.35-1.18	44.77-1.25	44.71-1.40	44.52-1.70	39.72-1.58
No. reflections		77,123	65,761	47,302	26,141	32,386
R_{work} / R_{free}		0.157 / 0.168	0.166 / 0.175	0.159 / 0.182	0.169 / 0.207	0.166 / 0.200
No. atoms		2,267	2,296	2,222	2,244	2,216
Protein		1,929	1,934	1,939	1,939	1,936
Ligand/ion		21	56	58	59	79
Water		317	306	225	246	201
B-factors		20.15	14.84	25.07	22.22	27.14
Protein		17.89	12.60	23.43	20.66	25.68
Ligand/ion		69.91	21.87	39.49	35.40	44.70
Water		30.62	27.69	35.52	31.38	34.24
R.m.s. deviations						
Bond lengths (Å)		0.011	0.008	0.016	0.005	0.012
Bond angles (°)		1.14	1.23	1.43	0.90	1.25
Molprobity Score		0.90	1.15	1.12	1.07	1.15

*1 crystal per structure. *Values in parentheses are for highest-resolution shell.

Table 4-S2 | MS^2 fragmentation of **3a** (**a**) and **3b** (**b**). Calculated and observed b- and y- ions. To produce **3**, bottromycin precursor peptide without the N-terminal methionine (BotA^P) was enzymatically processed by IpoC, PurCD and PurAH, to introduce the thiazoline, the macroamidine and cleave off the follower peptide, respectively.

a

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1	-	-	8	799.3793	799.3807
P	2	155.0812	155.0821	7	-	-
V	3	254.1494	254.1605	6	-	645.3070
V[a]	4	335.2078	335.2083	5	-	546.2386
V	5	434.2755	434.2767	4	465.1758	465.1808
F	6	581.3397	581.3451	3	366.1180	366.1124
D[b]	7	-	-	2	219.0453	219.0439
C[b]	8	-	781.3707	1	-	-

    c. Gly1 macrocyclisation with Val4 carbonyl
    d. Cys8 heterocyclisation with Asp7 carbonyl

Additionally, fragment a_5 (obs. 406.2822; calc. 406.2818) and b_4* (obs. 363.2023; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been observed previously for bottromycins and macrocyclic peptides, were detected.

b

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1	-	-	8	799.3788	799.3807
P	2	155.0817	155.0821	7	-	-
V	3	254.1497	254.1605	6	645.3031	645.3070
V[a]	4	335.2076	335.2083	5	546.2343	546.2386
V	5	434.2752	434.2767	4	465.1770	465.1808
F	6	581.3440	581.3451	3	366.1098	366.1124
D[b]	7		-	2	219.0422	219.0439
C[b]	8	781.3646	781.3707	1	-	-

    a. Gly1 macrocyclisation with Val4 carbonyl
    b. Cys8 heterocyclisation with Asp7 carbonyl

Additionally, fragment a_5 (obs. 406.2825; calc. 406.2818) and b_4* (obs. 363.2020; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been observed previously for bottromycins and macrocyclic peptides, were detected.

fragment b$_2$ fragment b$_3$
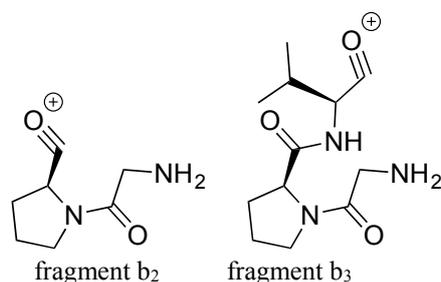
Table 4-S3 | MS^2 fragmentation of Asp Cα deuterated **3a** (**a**) and **3b** (**b**). Calculated and observed b- and y-ions. To produce **3,** bottromycin precursor peptide without the N-terminal methionine (BotA^P) was enzymatically processed by IpoC, PurCD and PurAH, to introduce the thiazoline, the macroamidine and cleave off the follower peptide, respectively.

a

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y	x_n	Obs. x	Calc. x
G^a	1	-	-	8	800.3838	800.3870	8	-	-
P	2	155.0827	155.0821	7	-	-	7	-	-
V	3	254.1501	254.1605	6	-	646.3133	6	-	674.3082
V^a	4	335.2064	335.2083	5	-	547.2449	5	575.2318	575.2398
V	5	434.2767	434.2767	4	466.1891	466.1871	4	476.1676	476.1714
F	6	581.3461	581.3451	3	367.1142	367.1186	3	-	393.0979
D^{b,c}	7	-	-	2	220.0491	220.0502	2	-	246.0995
C^b	8	-	782.3770	1	-	-	1	-	-

  a. Gly1 macrocyclisation with Val4 carbonyl
  b. Cys8 heterocyclisation with Asp7 carbonyl
  c. Asp7 Cα-positon labeled with deuteron

Fragments a$_5$ (obs. 406.2800; calc. 406.2818) and b$_4$* (obs. 363.2006; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been previously observed for bottromycins and macrocyclic peptides, were detected. Additionally, fragment z$_5$ (obs. 530.2191; calc. 530.2184) was detected.

b

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y	x_n	Obs. x	Calc. x
G^a	1	-	-	8	800.3831	800.3870	8	-	-
P	2	155.0809	155.0821	7	-	-	7	-	-
V	3	254.1489	254.1605	6	-	646.3133	6	-	674.3082
V^a	4	335.2064	335.2083	5	547.2449	547.2449	5	575.2329	575.2398
V	5	434.2741	434.2767	4	466.1796	466.1871	4	476.1685	476.1714
F	6	581.3440	581.3451	3	367.1160	367.1186	3	-	393.0979
D^{b,c}	7	-	-	2	220.0486	220.0502	2	-	246.0995
C^b	8	-	782.3770	1	-	-	1	-	-

  a. Gly1 macrocyclisation with Val4 carbonyl
  b. Cys8 heterocyclisation with Asp7 carbonyl
  c. Asp7 Cα-positon labeled with deuteron

Fragments a$_5$ (obs. 406.2801; calc. 406.2818) and b$_4$* (obs. 363.2013; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been previously observed for bottromycins and macrocyclic peptides, were detected. Additionally, fragment z$_5$ (obs. 530.2181; calc. 530.2184) was detected.

fragment b$_2$ fragment b$_3$
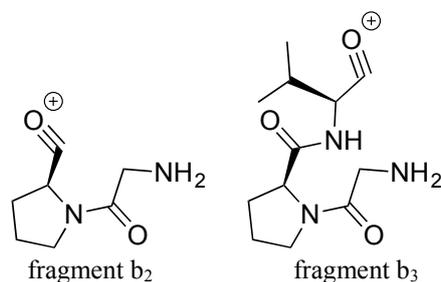
Table 4-S4 | Sum formulas and calculated masses for the IpoC, PurCD and PurAH treated BotA^P variants.

BotA^P variant	sum formula	monoiso. neutral mass	[M+H]$^+$
P2A	C$_{36}$H$_{52}$N$_8$O$_9$S	772.3578	773.3651
P2G	C$_{35}$H$_{50}$N$_8$O$_9$S	758.3421	759.3494
V3D	C$_{37}$H$_{50}$N$_8$O$_{11}$S	814.3320	815.3393
V3T	C$_{37}$H$_{52}$N$_8$O$_{10}$S	800.3527	801.3600
V3L	C$_{39}$H$_{56}$N$_8$O$_9$S	812.3891	813.3964
V4L	C$_{39}$H$_{56}$N$_8$O$_9$S	812.3891	813.3964
V5T	C$_{37}$H$_{52}$N$_8$O$_{10}$S	800.3527	801.3600
V5L	C$_{39}$H$_{56}$N$_8$O$_9$S	812.3891	813.3964
V5A	C$_{36}$H$_{50}$N$_8$O$_9$S	770.3421	771.3494
V5E	C$_{38}$H$_{52}$N$_8$O$_{11}$S	828.3476	829.3549
F6W	C$_{40}$H$_{55}$N$_9$O$_9$S	837.3843	838.3916
F6Y	C$_{38}$H$_{54}$N$_8$O$_{10}$S	814.3684	815.3756
F6A	C$_{32}$H$_{50}$N$_8$O$_9$S	722.3421	723.3494
D7A	C$_{37}$H$_{54}$N$_8$O$_7$S	754.3836	755.3909
D7N	C$_{38}$H$_{55}$N$_9$O$_8$S	797.3894	798.3967
D7E*	C$_{39}$H$_{56}$N$_8$O$_8$S	812.3891	[M+2H]$^{2+}$407.2018
wt	C$_{38}$H$_{54}$N$_8$O$_9$S	798.3734	799.3807

*the turnover by the enzymes IpoC, PurCD and PurAH is quite incomplete. Because of the low concentration of the Asp7Glu substrate the doubly charged ion was used for analysis.

Table 4-S5 | MS^2 fragmentation of the bottromycin A$_2$ main (**a**) and minor (**b**). Calculated and observed b-, x- and y- ions. For bottromycin production, the *S.* sp. DG2-kmP41hyg strain[42] was cultivated in TSB seed medium at 28 °C. After 2-3 days the seed culture was used to inoculated SG production medium (Glucose: 20 g; Peptone: 10 g and CaCO$_3$: 2 g per liter) and left to grow for 7 days at 28 °C. To the pooled cultures an equal volume of ethyl acetate was added and mixed for 1 hour. The organic layer was dried and resuspended in methanol before being loaded onto a sephadex LH-20 1-meter long column. The fractions containing bottromycin were pooled together and purified further by RP-HPLC.

a

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y	x_n	Obs. x	Calc. x
G^a	1	-	-	8	823.4512	823.4540	8	-	-
P	2	169.0965	169.0977	7	-	-	7	-	-
V	3	268.1645	268.1661	6	-	-	6	-	-
V^a	4	363.2378	363.2396	5	-	-	5	-	-
V	5	4763215	476.3237	4	461.2199	461.2223	4	-	-
F	6	637.4049	637.4077	3	348.1365	348.1382	3	374.1153	374.1175
D^b	7	-	-	2	187.0531	187.0541	2	213.0532	213.0334
C^b	8	-	-	1	-	-	1	-	-

    a.   Gly1 macrocyclisation with Val4 carbonyl
   b.   Cys8 heterocyclisation with Asp7 carbonyl

Additionally, fragments a$_5$ (obs. 448.3268; calc. 448.3288) and b$_4$* (obs. 391.2325; calc. 391.2345), which are characteristically associated with the bottromycin macrocycle and have been previously observed for bottromycins and macrocyclic peptides, were detected.

b

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y	x_n	Obs. x	Calc. x
G^a	1	-	-	8	823.4506	823.4540	8	-	-
P	2	169.0970	169.0977	7	-	-	7	-	-
V	3	268.1652	268.1661	6	-	-	6	-	-
V^a	4	363.2381	363.2396	5	-	-	5	-	-
V	5	476.3217	476.3237	4	461.2212	461.2223	4	-	-
F	6	637.4060	637.4077	3	348.1373	348.1382	3	374.1159	374.1175
D^b	7	-	-	2	187.0534	187.0541	2	-	213.0334
C^b	8	-	-	1	-	-	1	-	-

    a.   Gly1 macrocyclisation with Val4 carbonyl
    b.   Cys8 heterocyclisation with Asp7 carbonyl

Additionally, fragments a$_5$ (obs. 448.3286; calc. 448.3288) and b$_4$* (obs. 391.2337; calc. 391.2345), which are characteristically associated with the bottromycin macrocycle and have been previously observed for bottromycins and macrocyclic peptides, were detected.

fragment b$_2$ fragment b$_3$
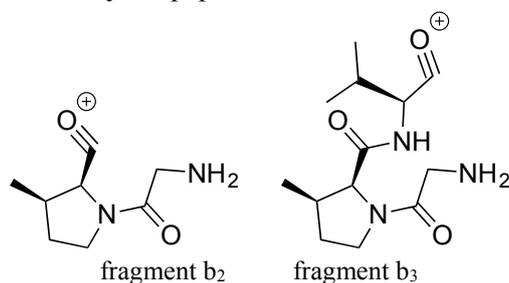
Table 4-S6 | MS^2 fragmentation of the bottromycin analog **6**. Calculated and observed b-, x- and y- ions.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1	-	-	8	871.4175	871.4210
P	2	169.0966	169.0977	7	-	-
M-O	3	316.1313	316.1331	6	-	-
V[a]	4	411.2038	411.2066	5	-	-
V	5	524.2877	524.2907	4	461.2202	461.2223
F	6	685.3702	685.3747	3	348.1369	348.1382
D[b]	7	-	-	2	187.0530	187.0541
C[b]	8	-	-	1	-	-

    a.   Gly1 macrocyclisation with Val4 carbonyl
    b.   Cys8 heterocyclisation with Asp7 carbonyl

Additionally, fragments a_5 (obs. 496.2933; calc. 496.2958) and b_4* (obs. 439.1991; calc. 439.2015), which are characteristically associated with the bottromycin macrocycle and have been previously observed for bottromycins and macrocyclic peptides, were detected.
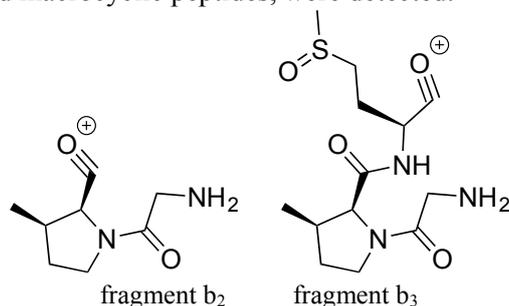

fragment b_2 fragment b_3

Methods

Protein expression and purification

The BotH coding sequence was amplified from genomic DNA isolated from *Streptomyces* sp. BC16019 and cloned into the pET-28b plasmid (Novagen). The resulting construct was verified by enzymatic restriction digestion and DNA sequencing before being transformed into *E. coli* BL21(DE3) competent cells.

A single colony was selected and grown in LB liquid medium supplemented with kanamycin (50 μg / mL) to make an overnight culture. This culture was used at a dilution of 1 to 100 to inoculate fresh LB medium containing the appropriate antibiotic, and the culture was grown at 37 °C, 180 rpm. Upon reaching an optical density (OD_{600}) of 0.6, the cultures were transferred to a precooled shaker at 16 °C, and protein expression was induced by addition of 1 mM IPTG. The cells were grown at 16 °C and 180 rpm over night before being harvested by centrifugation. Cell pellets were stored at -80 °C until further use.

The cell pellets were resuspended in lysis buffer (20 mM Tris pH 8.0, 500 mM NaCl, 20 mM Imidazole, and 3 mM DTT) supplemented with 0.4 mg DNAse per gram of wet cell pellet and cOmplete EDTA-free protease inhibitor tablets (Roche). The cell suspension was lysed via passage through a cell disrupter (Constant Systems) at 30,000 psi, and cell debris was removed by centrifugation (40,000 x g, 4 °C, 15 min). The supernatant was collected and directly loaded onto to a 5 mL Histrap HP column (GE healthcare) preequilibrated with lysis buffer. The column was washed extensively with lysis buffer (30 CV) before the protein was eluted with lysis buffer supplemented with 250 mM imidazole. Fractions containing BotH were directly loaded onto a gel filtration column (HiLoad 16/600 Superdex 200 pg, GE healthcare) preequilibrated in gel filtration buffer A (20 mM HEPES pH 7.4, 200 mM NaCl, 1 mM TCEP). The fractions of the highest purity as judged by SDS-PAGE were pooled and concentrated to 5 mg / mL.

Crystallization and structure determination

For crystallization, BotH was treated with thermolysin (1 : 100) on ice for 2 h, after which the protein was passed over a HiLoad 16/600 Superdex 200 pg gel filtration column as described above and concentrated to 5 mg / mL. Crystals of apo BotH and BotH in complex with ligands were obtained at 18 °C in 1.2 - 1.8 M ammonium sulfate and 0.1 M Tris-HCl using the hanging drop vapor diffusion method. For the complex crystallization, the thermolysin-treated protein was incubated with excess ligand (1 - 2 mM) on ice overnight before setting up crystallization drops. Crystals appeared after a few days and were allowed to grow for an additional week. The

crystals were cryoprotected in mother liquor supplemented with 30% glycerol and 0.5 mM ligand. Data was collected at ESRF (Beamlines ID23-1 and ID23-2), DESY (Beamline P11) and SLS (Beamline X06DA). To solve the apo structure, a single wavelength anomalous dispersion (SAD) data set was collected at the Se K absorption edge. Data were processed using Xia2[43], the structure was solved using PHENIX *AutoSol*[44] and the initial model used to obtain a molecular replacement solution for the high-resolution native data set using Phaser[45]. The solution was manually rebuilt in COOT[46] and refined using PHENIX Refine[44]. This apo structure was then used as a search model for structure determination of the complex crystal structures by molecular replacement (Phaser). Data for all complex crystal structures were processed using XDS[47] and POINTLESS[48], AIMLESS[49] and Ctruncate[50] implemented in ccp4[51]. All structures were validated using MolProbity. The images presented were created using PyMOL (Schrödinger, LLC) and LigPlot+[52].

Enzymatic reaction of BotH with **3**

To investigate the effect of BotH on **3a/b**, 20 μM **3a/b** were incubated with and without the addition of 5 μM BotH in GF buffer for 30 min. Reactions were stopped by the addition of 2 volumes ACN and were frozen at -80 °C until analysis. Reactions were set up and analyzed in at least triplicates.

To test the effect of different BotH concentrations on the **3a** : **3b** ratio, 20 μM **3a/b** were incubated with 0, 0.37, 1.11, 3.33 or 10 μM BotH at 30 °C for 2 h in GF buffer. Reactions were stopped by the addition of 2 volumes ACN and were frozen at -80 °C until analysis. Higher concentrations than 10 μM or longer incubation times were also tested, but did not lead to a further shift of epimer ratios. Reactions were set up and analyzed in triplicates.

To produce a roughly racemic mixture of **3a** : **3b**, 20 μM **3a/b** were incubated with 4 μM BotH for 30 min at 30 °C in GF buffer. BotH was denatured at 98 °C, 10 min, pelleted by centrifugation (15 min, 15,000 rpm) and the supernatant was lyophilized to remove all solvent.

Production of seleno-methionine (SeMet) BotH

L-Selenomethionine-labeled (SeMet) protein was expressed in *E. coli* BL21 cells grown in minimal medium supplemented with glucose-free nutrient mix (Molecular Dimensions), kanamycin (50 μg / mL) and glycerol (5% w / v). After 15 min growth at 37 °C and 180 rpm, 60 mg / mL selonomethionine was added. Upon reaching an OD_{600} of 0.6, amino acids (lysine, phenylalanine, threonine (100 mg / L each) and isoleucine and valine (50 mg / L each)) were added and the culture was grown for an additional 30 min at 37 °C, 180 rpm. Afterwards the

cultures were transferred to 16 °C, 180 rpm, and protein expression was induced by addition of 1 mM IPTG. The cells were harvested the next day and protein purified as described above.

Bottromycin extraction and purification

For bottromycin production, the *S. sp.* DG2-kmP41hyg strain[42] was cultivated in TSB seed medium at 28 °C. After 2 - 3 days the seed culture was used to inoculated SG production medium (Glucose: 20 g; Peptone: 10 g and $CaCO_3$: 2 g per liter) and left to grow for 7 days at 28 °C. To the pooled cultures an equal volume of ethyl acetate was added and mixed for 1 hour. The organic layer was dried and resuspended in methanol before being loaded onto a sephadex LH-20 1-meter long column. The fractions containing bottromycin were pooled together and purified further by RP-HPLC (XBridge® Peptide BEH^{TM} CSH C_{18} OBD Prep Column, 130 Å, 5 μm, 10 mm x 250 mm, 1 / pkg) using a linear gradient from 95% A (H_2O, 0.1% formic acid) to 95% B (acetonitrile, 0.1 % formic acid) over 40 min. The fractions containing bottromycin of the highest purity as judged by LC-MS were pooled and dried on a rotary evaporator. The resulting white amorphous solid was resuspend in methanol and dried under nitrogen to yield pure bottromycin A_2.

In vitro production and purification of **3**

To produce **3,** bottromycin precursor peptide without the N-terminal methionine (BotA^P) was enzymatically processed by IpoC, PurCD and PurAH, to introduce the thiazoline, the macroamidine and cleave off the follower peptide, respectively (see Figure S2).

BotA^P, IpoC, PurCD and PurAH were expressed and purified as described previously [21, 25]. Large scale IpoC reactions were carried out on a 9 mL scale in GF buffer (150 mM NaCl, 10 mM HEPES, 0.5 mM TCEP, pH 7.4) using the reaction setup 50 μM BotA^P, 5 μM IpoC, 5 mM ATP as well as 5 mM $MgCl_2$ and incubating the reaction mixture for 16 h at 37 °C. The reaction mixture was filtered through a 0.22 μm filter and loaded onto a Superdex 30 16/60 size exclusion chromatography column pre-equilibrated in GF buffer. Heterocyclized-peptide-containing fractions were pooled, analyzed by MS and concentrated using a 5 kDa cutoff filter. Next, large scale reactions with PurCD and PurAH were carried out on a 9 mL scale by incubating 50 μM BotA^{PC}, 5 μM PurCD, 1 μM PurAH, 10 mM ATP, 10 mM $MgCl_2$ and 100 μM $CoCl_2$ for 12 h at 37 °C in reaction buffer (200 mM NaCl, 50 mM Tris, 10% glycerol, pH 8.5). The reaction process was monitored by LC-MS. The analysis under acidic LC conditions, which leads to partial (about 50%) opening of the thiazoline ring, revealed the

presence of two peaks (**3a/b**) with identical mass and fragmentation patterns, with a ratio of **3a** : **3b** of ~ 1 : 9.

After completion of the reaction, the mixture was filtered through a 0.22 μm filter and applied to a Superdex 30 16/60 size exclusion chromatography column pre-equilibrated in ultrapure water. Every fraction was analyzed by mass spectrometry and the pure **3a/b**-containing fractions were pooled.

For further purification, **3a/b** was purified in H_2O and ACN containing 0.1% ammonium acetate pH 8.5. Separation was carried out on a Kinetex XB-C18 2.6 μm, 4.6 x 100 mm column (Phenomenex) at 60 °C using a 8 min gradient from 0 to 80% ACN. Pure fractions were lyophilized and the remaining **3a/b** powder was weighed using a precision scale. The yield of purified product was approximately 70%. A large-scale purification was carried out by Peptide Protein Research Ltd.

Microscale Thermophoresis (MST)

Microscale thermophoresis experiments were carried out on a Monolith NT.115 (NanoTemper) in MST buffer (PBS with 0.05% Tween 20) using BotH labelled with the RED-tris-NTA fluorescent dye NT-647 following manufacture's protocol. The concentration of His-labelled BotH was 200 nM. Ligands were diluted in MST buffer to a starting concentration of 62.5 μM and then used in serial dilutions. Instrument settings were 40% excitation power and 40% MST power. Data fitting and evaluation was performed using *MO.Affinity* analysis software (Nanotemper).

Thermal Shift Assays

Ligands used for thermal shift assays were prepared as described for the in vitro production and purification of **3** (see above) with the following modification: After incubation with PurCD/AH, enzymes were removed by passing the reactions through a 3 kDa cut-off filter and used for thermal shift assays. Protein melting temperatures were determined by monitoring protein unfolding using SYPRO orange as a fluorescence probe. BotH was diluted to 5 μM in buffer containing 5x SYPRO orange (Sigma). For the determination of stabilizing effects of the substrates on BotH, buffer containing 150 mM NaCl and 10 mM HEPES (pH 7.4) was used. Final substrate concentrations of 3.0, 5.0 and 12 μM were used. Samples (20 μL/well) were analyzed in 96 well plates (Microamp[TM] Reaction Plates) that were sealed (MicroAmp[TM] Adhesive Film). Measurements were carried out with a realtime PCR machine (StepOnePlus[TM] Real-Time PCR System) using manufacture's guidelines and a temperature gradient from 25 to

95 °C with stepwise increments of 1 °C and 1 min hold. After each temperature step the fluorescence intensity was measured. The melting temperatures were obtained using the derivative method (Protein Thermal Shift^{TM} Software v1.4). All conditions were tested in triplicates and mean values were calculated for the graphic presentation.

Epimerization under single-turnover conditions

To conduct an epimerization reaction under single-turnover conditions, 5 μM **3a/b** was added to 20 μM BotH in D_2O and the reaction quenched after 2 seconds by adding 2 reaction volumes of ACN and flash-freezing the samples immediately. Selection of an appropriate time-point was critical because BotH catalyzes the forward and backward reactions, and the substrate exists as an epimeric mixture. At the start of the reaction, the substrate consisted of approximately 35% D-Asp and 65% L-Asp, while the end-point presented as a roughly equimolar mixture of epimers. This difference is not the actual end-point of the reaction with excess BotH, but was chosen because it resulted in a shift of epimer ratios that was clearly distinguishable from experimental error while minimizing the amount of D-Asp to L-Asp conversion. We then used these data to calculate the proportion of hydrogen-containing D-Asp and deuteron-containing D-Asp in the total Asp-pool at the start and finish of the experiment. The data are presented in Figure 4-S8.

Marfey derivatization

For stereochemical assignment of the aspartate α-carbon in **3**, Marfey derivatization was performed. For the assay, 300 μg **3** (ratio of **3a** : **3b** of about 1 : 9) was used. Solvent was evaporated at 110 °C and the compound was hydrolyzed by addition of 100 μl 6 N HCl and incubation at 110 °C for 35 min in a closed vial with nitrogen. The vial was opened and incubated for further 15 min at 110 °C to dry the contents. Residues were dissolved in 110 μl H_2O.

For derivatization of the amino acids in the hydrolyzed compound **3** or the amino acid standards (2 mg / mL D- or L- aspartic acid), 50 μL solution were mixed with 20 μl 1 N $NaHCO_3$ and 20 μl of 1% D-FDLA (*Nα*-(2,4-dinitro-5-fluorophenyl)-D-leucinylamide) in acetone. The mixture was incubated for 1.5 h at 40 °C, 700 rpm. To stop the reaction, 10 μl of 2 N HCl and 300 μl ACN were added. The derivatized amino acids were separated by RP-HPLC-MS and the stereochemistry of the aspartate in the compounds **3a** and **3b** were assigned by comparing retentions times of the amino acid standards.

Spontaneous epimerization of 3

To determine the propensity for spontaneous (non-enzymatic) epimerization of the Asp Cα in **3a/b**, a lyophilized, racemic mixture of **3a** : **3b** was dissolved in D_2O. Samples were taken after 2 h, 1 day, 4 days and 6 days and analyzed by LC-MS. Three independent experiment were performed.

Pull-down of bottromycin A_2 and 3a/b using BotH

S. sp. DG2-KMp41hyg strain was cultivated as described above. 50 mL cultures were centrifuged for 10 min at room temperature to separate cells from the medium. The cell pellet was subsequently washed twice with PBS buffer before being resuspend in fresh 50 mL PBS buffer. The cell suspension was then lysed by sonication, centrifuged and the supernatant was decanted. His_6-BotH (75 µg) was subsequently incubated with medium (1 mL), supernatant (1 mL) and **3a/b** (50 µg dissolved PBS) for 30 min before being purified using a KingFisher[TM]mL Purification System. The protein was precipitated using ice cold ACN to liberate bound compounds and the supernatant was analyzed by HRLC-MS.

Incorporation of a deuteron at Asp Cα of 3

Deuteron incorporation at the Asp Cα was achieved by performing a BotH reaction in D_2O GF buffer. Lyophilized **3a/b** was dissolved in D_2O GF buffer and incubated for 30 min at 30 °C with (or without as control) 5 µM BotH. The reactions were stopped by addition of 2 volumes ACN and stored at -80 °C until analysis. Reactions were set up in triplicates and analyzed by LC-MS and MS-fragmentation.

Back exchange of deuterated **3a/b** with solvent protons was shown by performing a BotH reaction using deuterated **3a/b** in H_2O GF buffer: A reaction of **3a/b** with 5 µM BotH in D_2O GF was stopped after 30 min at 30 °C by denaturation of BotH at 98 °C for 10 min. Denatured BotH was pelleted by centrifugation (15 min, 15,000 rpm). The supernatant was lyophilized to remove the D_2O and re-dissolved in the same volume H_2O. Samples were split up and incubated 30 min at 30 °C with or without the addition of 5 µM BotH. Reactions were stopped by the addition of 2 volumes ACN and analyzed by LC-MS.

Reaction of BotH with derivatives of 3

To generate derivatives of **3,** bottromycin precursor peptides (BotA[P]) with singe amino acid changes in the core peptide were cloned, expressed and purified as described previously[21, 25]. Shortly, *botA* was mutated by using mutation carrying primers and cloned into a

pHisSUMOTEV vector, which was a gift from Dr. Huanting Liu, St. Andrews University[53]. To produce the native N-terminus of the bottromycin precursor peptide BotA[P] (after cleavage of the N-terminal methionine by BotP) (GPVVV….) without cloning artefacts, primers which introduce a lysine residue were used (...GAMAGK<u>GPVVV...</u>). Peptides were expressed in *E. coli* Lemo21(DE3) cells carrying the respective expression vector and purified by Ni^{2+}-NTA-chromatography and, after cleavage with Trypsin, by gel filtration[21, 25].

To produce the **3** derivatives, 50 µM BotA[P] derivative was incubated with 5 µM IpoC, 10 mM ATP as well as 5 mM $MgCl_2$ for 12 h at 37 °C in reaction buffer (200 mM NaCl, 50 mM Tris, 10 % glycerol, pH 8.5). Then, 5 µM PurCD, 1 µM PurAH and 100 µM $CoCl_2$ were added and incubated for a further 4 h at 37 °C. Each reaction was then divided and incubated with or without the addition of 5 µM BotH for 2 h at 30 °C. For the BotA[P] Asp7-mutants, deuteron incorporation by BotH was tested additionally: After incubation with IpoC, PurCD and PurAH, reactions were lyophilized and redissolved in D_2O. Afterwards the samples were divided and incubated with and without the addition of BotH (as described above).

Reactions were stopped by the addition of 2 volumes ACN and frozen at -80 °C until analysis. Reactions were set up and analyzed in triplicates.

Orthosteric inhibition of BotH epimerization by bottromycin A2

0.2 µM BotH in D_2O GF buffer was pre-incubated with and without 50 µM bottromycin A2 for 30 min at 30 °C. **3a/b** in D_2O GF buffer was added (final concentration 10 µM) and the solution was incubated for 7 min at 30 °C. The reactions were stopped by the addition of 2 volumes ACN and the samples were frozen at -80 °C until analysis by LC-MS. As a control for non-enzymatic deuteron incorporation, samples without enzyme were set up. Reactions were performed in triplicates and analyzed by LC-MS. For analysis of the LC-MS data, the EICs (± 5 ppm) for the 1st and the 2nd isotope peak of the doubly charged ion of **3** (1st isotope peak $[M+2H]^{2+} = 400.1940$; 2nd isotope peak $[M+2H]^{2+} = 400.6955$) were generated. The respective areas for **3a** (2.96 - 3.10 min) were calculated using the Bruker Compass DataAnalysis 4.2. software. The ratio of the area of the 1st isotope peak to the 2nd isotope peak was used to quantify the deuteron incorporation. The significance (*p*-value) of the ratio differences was calculated using an unpaired two-tailed t-test.

In vitro production of oxidatively decarboxylated **3a**, with deuteron incorporation at the Asp Cα

To investigate if BotH can epimerize decarboxylated **3a,** we produced decarboxylated **3a** with deuteron incorporation at the Asp Cα. 20 µM BotCYP substrate analogue **3a/b** with deuteron incorporated (see above) was incubated with 5 µM BotCYP, 5 µM BmCPR, 50 µM Fdx2, 1 µM BotH and 2.5 mM NADPH at 30 °C for 2 h in H_2O GF buffer. Enzymes were denatured at 98 °C for 10 min and pelleted by centrifugation (15.000 x g, 20 min, 4 °C). The supernatant was incubated with and without 5 µM BotH for 30 min at 30 °C. Reactions were stopped and enzymes were precipitated by addition of two volumes of ACN. The samples were frozen at -80 °C and precipitated enzymes were pelleted by centrifugation at 15.000 x g for 20 min. The turnover was analyzed by LC-MS measurements. Reactions were set up and analyzed in triplicates.

LC-MS and MS[2] analysis

Measurements were performed on a Dionex Ultimate 3000 RSLC system (Thermo Fisher Scientific) using a flow rate of 600 µL min^{-1} and column oven temperature of 45 °C. Standard measurements were performed using a BEH C18, 50 x 2.1 mm, 1.7 µm dp column equipped with a C18 precolumn (Waters). Samples were separated by a gradient from (A) H_2O + 0.1% formic acid to (B) ACN + 0.1% formic acid. The linear gradient was initiated by a 1 min isocratic step at 5% B, followed by an increase to 95% B in 9 min to end up with a 1.5 min plateau step at 95% B before re-equilibration under the initial conditions. UV spectra were recorded by a DAD in the range from 200 to 600 nm.

For Marfey's analysis, samples were separated on a BEH C18, 100 x 2.1 mm, 1.7 µm dp column equipped with a C18 precolumn (Waters). A multistep gradient using (A) H_2O + 0.1% formic acid and (B) ACN + 0.1% formic acid was applied for sample preparation. The gradient was initiated by a 1 min isocratic step at 5% B, an increase to 10% B in 1 min, an increase to 35% B in 14 min, followed by an increase to 55% B in 7 min and an increase to 80% B in 3 min to end up with a 1 min plateau step at 80% B before re-equilibration to the initial conditions.

For MS measurements on maXis-4 hr-qToF mass spectrometer (Bruker Daltonics), the LC flow was split 1:8 before entering the mass spectrometer using the Apollo II ESI source. In the source region, the temperature was set to 200 °C, the capillary voltage was 4000 V, the dry-gas flow was 5.0 L/min and the nebulizer was set to 1 bar. Ion transfer settings were set to Funnel 1 RF 350 Vpp and Multipole RF 400 Vpp, quadrupole settings were set to an ion energy of 5.0 eV and a low mass cut of 120 *m/z*. The collision cell was set to an energy of 5.0 eV and the pulse

storage time was 5 μs. Data were recorded in centroid mode ranging from 150 to 2500 *m/z* at a 2 Hz scan rate. Calibration of the maXis4G qTOF spectrometer was achieved with sodium formate clusters before every injection to avoid mass drifts. All MS analyses were acquired in the presence of the lock masses $C_{12}H_{19}F_{12}N_3O_6P_3$, $C_{18}H_{19}O_6N_3P_3F_2$ and $C_{24}H_{19}F_{36}N_3O_6P_3$ which generate the [M+H]$^+$ Ions of 622.028960, 922.009798 and 1221.990638.

LC-MS^2 fragmentation spectra were recorded using a scheduled precursor list (SPL). Separation was achieved using a BEH C18, 100 x 2.1 mm, 1.7 μm dp column equipped with a C18 precolumn (Waters) and a linear gradient from (A) H_2O + 0.1% formic acid to (B) ACN + 0.1% formic acid. The gradient was initiated by a 1 min isocratic step at 5% B, followed by an increase to 95% B in 18 min to end up with a 2 min plateau step at 95% ACN before re-equilibration to the initial conditions. SPL entries and parameters were set to fragment only the ions of interest. SPL tolerance parameters for precursor ion selection were 0.17 min and 0.05 *m/z*. The CID energy was ramped from 35 eV for 500 *m/z* to 45 eV for 1000 *m/z*.

Data were displayed and analyzed using the Bruker Compass DataAnalysis software (Version 4.2). Shown MS spectra are in general single spectra (taken at the EIC maximum), except for Figure 4-S3 and Figure 4-S9 where spectra are averaged. Signals in the MS-spectra are labelled with the observed monoisotopic mass. Extracted-ion chromatograms were generated using the calculated monoisotopic mass with a range of 5 ppm. All EICs and MS-spectra where no specific intensity is stated are scaled to a relative intensity of 1.

Bioinformatics analysis

We downloaded bacterial sequences of the InterPro[54] family IPR000073 and aligned them with HMMer (http://hmmer.org/)[55] using the HMM for the Pfam[40] family PF00561 (Abhydrolase_1), retaining only those from this Pfam family. Then sequences containing the following mutations at the catalytic site positions were selected with a custom Perl script: Ser mutated to either Val, Ile, Leu, Met, Ala, Phe, Tyr, or Trp; Asp mutated to anything except Asn or Glu or missing; His mutated to any amino acid or missing.

The corresponding genomes were identified by parsing UniProt[56] data files and retrieved from NCBI using the batch download option. Biosynthetic gene clusters (BGCs) were predicted using antiSMASH[57]. All BotH homologs whose genomic coordinates fall within the predicted boundaries of BGCs or lie outside them separated by no more than 1000 nt were considered to be associated with these BGCs.

The distance matrix between all proteins that are associated with BGCs was calculated using the R package seqinr[58]. The sequence similarity networks were visualized in Cytoscape[59], all edges connecting pairs of proteins that shared less than 15% sequence identity were omitted.

4.4 References

1. Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; et al., The alpha/beta hydrolase fold. *Protein Eng* **1992,** *5* (3), 197-211.

2. Nardini, M.; Dijkstra, B. W., Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* **1999,** *9* (6), 732-7.

3. Mindrebo, J. T.; Nartey, C. M.; Seto, Y.; Burkart, M. D.; Noel, J. P., Unveiling the functional diversity of the alpha/beta hydrolase superfamily in the plant kingdom. *Curr Opin Struct Biol* **2016,** *41*, 233-246.

4. Rauwerdink, A.; Kazlauskas, R. J., How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: the Serine-Histidine-Aspartate Catalytic Triad of alpha/beta-Hydrolase Fold Enzymes. *ACS Catal* **2015,** *5* (10), 6153-6176.

5. Hamiaux, C.; Drummond, R. S.; Janssen, B. J.; Ledger, S. E.; Cooney, J. M.; Newcomb, R. D.; Snowden, K. C., DAD2 is an alpha/beta hydrolase likely to be involved in the perception of the plant branching hormone, strigolactone. *Curr Biol* **2012,** *22* (21), 2032-6.

6. Guo, Y.; Zheng, Z.; La Clair, J. J.; Chory, J.; Noel, J. P., Smoke-derived karrikin perception by the alpha/beta-hydrolase KAI2 from Arabidopsis. *Proc Natl Acad Sci U S A* **2013,** *110* (20), 8284-9.

7. Shimada, A.; Ueguchi-Tanaka, M.; Nakatsu, T.; Nakajima, M.; Naoe, Y.; Ohmiya, H.; Kato, H.; Matsuoka, M., Structural basis for gibberellin recognition by its receptor GID1. *Nature* **2008,** *456* (7221), 520-3.

8. Marchot, P.; Chatonnet, A., Enzymatic activity and protein interactions in alpha/beta hydrolase fold proteins: moonlighting versus promiscuity. *Protein Pept Lett* **2012,** *19* (2), 132-43.

9. Crone, W. J. K.; Leeper, F. J.; Truman, A. W., Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chemical Science* **2012,** *3* (12), 3516-3521.

10. Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chemical Science* **2012,** *3* (12), 3522-3525.

11. 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., Structure and biosynthesis of the antibiotic bottromycin D. *Org Lett* **2012,** *14* (19), 5050-3.

12. Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Muller, R., Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem Biol* **2012,** *19* (10), 1278-87.

13. Waisvisz, J. M., Bottromycin. I. A new sulfurcontaining antibiotic. *J. Am. Chem. Soc.* **1957,** *79*, 4520-4521.

14. Nakamura, S., Isolation and characterization of bottromycins A and B. *J. Antibiotics, Ser. A* **1965,** *18*, 47-52.

15. Sowa, S.; Masumi, N.; Inouye, Y.; Nakamura, S.; Takesue, Y.; Yokoyama, T., Susceptibility of methicillin-resistant Staphylococcus aureus clinical isolates to various antimicrobial agents. *Hiroshima J Med Sci* **1991,** *40* (4), 137-44.

16. Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Omura, S., Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew Chem Int Ed Engl* **2009,** *48* (5), 914-7.

17. Otaka, T.; Kaji, A., Mode of action of bottromycin A2. Release of aminoacyl- or peptidyl-tRNA from ribosomes. *J Biol Chem* **1976,** *251* (8), 2299-306.

18. Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect on peptide bond formation. *FEBS Lett* **1981,** *123* (2), 173-6.

19. Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect of bottromycin A2 on polysomes. *FEBS Lett* **1983,** *153* (1), 53-9.

20. 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.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* **2013,** *30* (1), 108-60.

21. Franz, L.; Adam, S.; Santos-Aberturas, J.; Truman, A. W.; Koehnke, J., Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme. *J Am Chem Soc* **2017,** *139* (50), 18158-18161.

22. Schwalen, C. J.; Hudson, G. A.; Kosol, S.; Mahanta, N.; Challis, G. L.; Mitchell, D. A., In Vitro Biosynthetic Studies of Bottromycin Expand the Enzymatic Capabilities of the YcaO Superfamily. *J Am Chem Soc* **2017,** *139* (50), 18154-18157.

23. Crone, W. J.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angew Chem Int Ed Engl* **2016,** *55* (33), 9639-43.

24. Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Muller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *Chembiochem* **2016,** *17* (23), 2286-2292.

25.   Sikandar, A.; Franz, L.; Melse, O.; Antes, I.; Koehnke, J., Thiazoline-Specific Amidohydrolase PurAH Is the Gatekeeper of Bottromycin Biosynthesis. *J Am Chem Soc* **2019,** *141* (25), 9748-9752.

26.   Hur, G. H.; Vickery, C. R.; Burkart, M. D., Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat Prod Rep* **2012,** *29* (10), 1074-98.

27.   Freeman, M. F.; Gurgui, C.; Helf, M. J.; Morinaka, B. I.; Uria, A. R.; Oldham, N. J.; Sahl, H. G.; Matsunaga, S.; Piel, J., Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* **2012,** *338* (6105), 387-90.

28.   Morinaka, B. I.; Vagstad, A. L.; Helf, M. J.; Gugger, M.; Kegler, C.; Freeman, M. F.; Bode, H. B.; Piel, J., Radical S-adenosyl methionine epimerases: regioselective introduction of diverse D-amino acid patterns into peptide natural products. *Angew Chem Int Ed Engl* **2014,** *53* (32), 8503-7.

29.   Parent, A.; Benjdia, A.; Guillot, A.; Kubiak, X.; Balty, C.; Lefranc, B.; Leprince, J.; Berteau, O., Mechanistic Investigations of PoyD, a Radical S-Adenosyl-l-methionine Enzyme Catalyzing Iterative and Directional Epimerizations in Polytheonamide A Biosynthesis. *J Am Chem Soc* **2018,** *140* (7), 2469-2477.

30.   Benjdia, A.; Guillot, A.; Ruffie, P.; Leprince, J.; Berteau, O., Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in Bacillus subtilis. *Nat Chem* **2017,** *9* (7), 698-707.

31.   Cotter, P. D.; O'Connor, P. M.; Draper, L. A.; Lawton, E. M.; Deegan, L. H.; Hill, C.; Ross, R. P., Posttranslational conversion of L-serines to D-alanines is vital for optimal production and activity of the lantibiotic lacticin 3147. *Proc Natl Acad Sci U S A* **2005,** *102* (51), 18584-9.

32.   Yang, X.; van der Donk, W. A., Post-translational Introduction of D-Alanine into Ribosomally Synthesized Peptides by the Dehydroalanine Reductase NpnJ. *J Am Chem Soc* **2015,** *137* (39), 12426-9.

33.   Holm, L., Benchmarking Fold Detection by DaliLite v.5. *Bioinformatics* **2019**.

34.   Bains, J.; Kaufman, L.; Farnell, B.; Boulanger, M. J., A product analog bound form of 3-oxoadipate-enol-lactonase (PcaD) reveals a multifunctional role for the divergent cap domain. *J Mol Biol* **2011,** *406* (5), 649-58.

35.   Gouda, H.; Kobayashi, Y.; Yamada, T.; Ideguchi, T.; Sugawara, A.; Hirose, T.; Omura, S.; Sunazuka, T.; Hirono, S., Three-dimensional solution structure of bottromycin A2: a potent antibiotic active against methicillin-resistant Staphylococcus aureus and vancomycin-resistant Enterococci. *Chem Pharm Bull (Tokyo)* **2012,** *60* (2), 169-71.

36.   Yamada, T.; Yagita, M.; Kobayashi, Y.; Sennari, G.; Shimamura, H.; Matsui, H.; Horimatsu, Y.; Hanaki, H.; Hirose, T.; S, O. M.; Sunazuka, T., Synthesis and Evaluation of Antibacterial Activity of Bottromycins. *J Org Chem* **2018,** *83* (13), 7135-7149.

37.   Jungheim, L. N.; Shepherd, T. A.; Baxter, A. J.; Burgess, J.; Hatch, S. D.; Lubbehusen, P.; Wiskerchen, M.; Muesing, M. A., Potent human immunodeficiency

virus type 1 protease inhibitors that utilize noncoded D-amino acids as P2/P3 ligands. *J Med Chem* **1996,** *39* (1), 96-108.

38.   Rink, R.; Arkema-Meter, A.; Baudoin, I.; Post, E.; Kuipers, A.; Nelemans, S. A.; Akanbi, M. H.; Moll, G. N., To protect peptide pharmaceuticals against peptidases. *J Pharmacol Toxicol Methods* **2010,** *61* (2), 210-8.

39.   Hunter, S.; Apweiler, R.; Attwood, T. K.; Bairoch, A.; Bateman, A.; Binns, D.; Bork, P.; Das, U.; Daugherty, L.; Duquenne, L.; Finn, R. D.; Gough, J.; Haft, D.; Hulo, N.; Kahn, D.; Kelly, E.; Laugraud, A.; Letunic, I.; Lonsdale, D.; Lopez, R.; Madera, M.; Maslen, J.; McAnulla, C.; McDowall, J.; Mistry, J.; Mitchell, A.; Mulder, N.; Natale, D.; Orengo, C.; Quinn, A. F.; Selengut, J. D.; Sigrist, C. J.; Thimma, M.; Thomas, P. D.; Valentin, F.; Wilson, D.; Wu, C. H.; Yeats, C., InterPro: the integrative protein signature database. *Nucleic Acids Res* **2009,** *37* (Database issue), D211-5.

40.   El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D., The Pfam protein families database in 2019. *Nucleic Acids Res* **2019,** *47* (D1), D427-D432.

41.   Stachelhaus, T.; Walsh, C. T., Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. *Biochemistry* **2000,** *39* (19), 5775-87.

42.   Horbal, L.; Marques, F.; Nadmid, S.; Mendes, M. V.; Luzhetskyy, A., Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metab Eng* **2018,** *49*, 299-315.

43.   Winter, G., xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* **2010,** *43* (1), 186-190.

44.   Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

45.   McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J Appl Crystallogr* **2007,** *40* (Pt 4), 658-674.

46.   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.

47.   Kabsch, W., Xds. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 125-32.

48.   Evans, P. R., An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 282-92.

49.   Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **2013,** *69* (Pt 7), 1204-14.

50. French, S.; Wilson, K., On the treatment of negative intensity observations. *Acta Crystallographica Section A* **1978,** *34* (4), 517-525.

51. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 235-42.

52. Laskowski, R. A.; Swindells, M. B., LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **2011,** *51* (10), 2778-86.

53. Liu, H.; Naismith, J. H., A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr Purif* **2009,** *63* (2), 102-11.

54. Mitchell, A. L.; Attwood, T. K.; Babbitt, P. C.; Blum, M.; Bork, P.; Bridge, A.; Brown, S. D.; Chang, H. Y.; El-Gebali, S.; Fraser, M. I.; Gough, J.; Haft, D. R.; Huang, H.; Letunic, I.; Lopez, R.; Luciani, A.; Madeira, F.; Marchler-Bauer, A.; Mi, H.; Natale, D. A.; Necci, M.; Nuka, G.; Orengo, C.; Pandurangan, A. P.; Paysan-Lafosse, T.; Pesseat, S.; Potter, S. C.; Qureshi, M. A.; Rawlings, N. D.; Redaschi, N.; Richardson, L. J.; Rivoire, C.; Salazar, G. A.; Sangrador-Vegas, A.; Sigrist, C. J. A.; Sillitoe, I.; Sutton, G. G.; Thanki, N.; Thomas, P. D.; Tosatto, S. C. E.; Yong, S. Y.; Finn, R. D., InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* **2019,** *47* (D1), D351-D360.

55. Eddy, S. R., Profile hidden Markov models. *Bioinformatics* **1998,** *14* (9), 755-63.

56. UniProt, C., UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **2019,** *47* (D1), D506-D515.

57. Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H., antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* **2017,** *45* (W1), W36-W41.

58. Charif, D.; Lobry, J. R., SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. In *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*, Bastolla, U.; Porto, M.; Roman, H. E.; Vendruscolo, M., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2007; pp 207-232.

59. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003,** *13* (11), 2498-504.

Chapter 5

Characterization of the Stereospecific P450 Enzyme BotCYP Enables the In Vitro Biosynthesis of the Bottromycin Core Scaffold

Sebastian Adam,[+] **Laura Franz**,[+] Mohammed Milhim, Rita Bernhardt, Olga V. Kalinina, and Jesko Koehnke

[+] shared authorship

Author contributions:

Laura Franz contributed to the publication by performing and evaluation the *in vitro* biochemical experiments and mass spectrometry measurements, demonstrating that BotCYP enables stereospecific resolution of the pathway. Laura Franz performed enzyme expression and purification of the redox partners and the biosynthetic enzymes preceding CYP of the bottromycin pathway. She performed cloning of the SalCYP enzyme point mutants and designed figures. Sebastian Adam established the expression and purification protocols of the CYP enzymes, crystallized the enzyme SalCYP and solved the x-ray structure. He performed binding affinity experiments and designed figures. Laura Franz and Sebastian Adam produced and purified the enzyme substrate. Mohammed Milhim, supervised by Rita Bernhardt, performed heme reduction experiments, provided plasmids and protocols for the purification of redox partners, and established an optimized BotCYP construct and BotCYP expression. Olga V. Kalinina performed the large-scale, bioinformatic analysis of P450 enzymes associated with RiPP biosynthetic gene clusters.

Jesko Köhnke wrote the paper with contributions from Laura Franz, Sebastian Adam and Olga V. Kalinina. The full program was carried out under the guidance and direction of Jesko Köhnke. All authors provided critical feedback and discussed the results of the manuscript.

5 Characterization of the Stereospecific P450 Enzyme BotCYP Enables the In Vitro Biosynthesis of the Bottromycin Core Scaffold

5.1 Abstract

Bottromycins are ribosomally synthesized and post-translationally modified peptide natural product anti-biotics that are effective against high-priority human pathogens such as methicillin-resistant Staphylococcus aureus. The total synthesis of bottromycins involves at least 17 steps with a poor overall yield. Here, we report the characterization of the cytochrome P450 enzyme BotCYP from a bottromycin biosynthetic gene cluster. We determined the structure of a close BotCYP homolog, and used our data to conduct the first large-scale survey of P450 enzymes associated with RiPP biosynthetic gene clusters. We demonstrate that BotCYP converts a C-terminal thiazoline to a thiazole via an oxidative decarboxylation reaction and provides stereochemical resolution for the pathway. Our data enable the 2-pot in vitro production of the bottromycin core scaffold and may allow the rapid generation of bottromycin analogues for compound development.
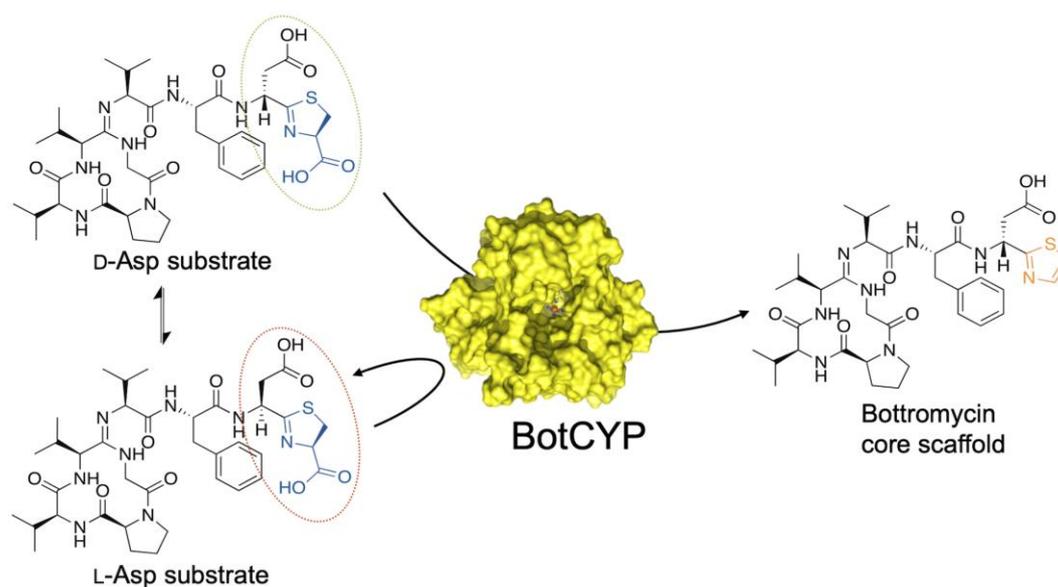
Figure 5-1 | Graphical abstract.

5.2 Main Text

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a rapidly growing natural product superfamily with interesting and diverse bioactivities.[1-3] Bottromycins (**1**, Figure 5-2A) are RiPPs with potent activity against Gram-positive bacteria, including the problematic human pathogens MRSA and VRE.[4, 5] The bioactivity triggered several attempts at total synthesis, which was finally achieved by an inspired route requiring at least 17 steps.[5] *In vitro* and *in vivo*[6] analyses of bottromycin biosynthesis have largely established the order of steps: Removal of the N-terminal methionine[7] is followed by thiazoline and macroamidine formation[8, 9] and removal of the follower peptide to yield **2a** (Figure 5-2B).[10] This intermediate is then converted to **2b** by the epimerase BotH, which results in a mixture of the two epimers (Figure 5-2B).[11] The D-Asp configuration of **2b**, tentatively assigned by Marfey´s analysis of the 2 epimers and protein crystallography,[11] is also found in the final natural product **1**.[5, 12]

Five-membered heterocycles that are enzymatically derived from serine, threonine or cysteine residues are frequently found in RiPPs.[1, 3] Their oxidation state, azoline or azole, can have a profound effect on RiPP bioactivity,[13, 14] and all enzymes linked with azoline oxidation in RiPP biosynthesis thus far are flavin-dependent.[15-17] In contrast, it has been proposed that the oxidative decarboxylation reaction of the thiazoline in bottromycin biosynthesis is catalyzed by a P450 enzyme (BotCYP) found in the bottromycin biosynthetic gene cluster (BGC, Figure 5-S1).[6] To investigate the role of this enzyme *in vitro*, we used heterologously expressed and purified BotCYP in spectral analyses to confirm an intact heme-containing protein. Incubation of BotCYP with sodium dithionite led to the observation of the typical Soret band at 448 nm in the spectrum with bound carbon monoxide, indicating the chemical reduction of the heme iron (Figure 5-S2).[18]

Notably, a typical reductase-ferredoxin pair required as a cofactor for cytochrome P450 enzymes is not found in the Bottromycin BGC. Enzymatic reduction of the heme iron was thus attempted via several other P450 reductase-ferredoxin pairs and the highest peak at 448 nm observed using the P450 reductase-ferredoxin pair BmCPR-Fdx2 from *Bacillus megaterium* (Figure 5-S2).[19]
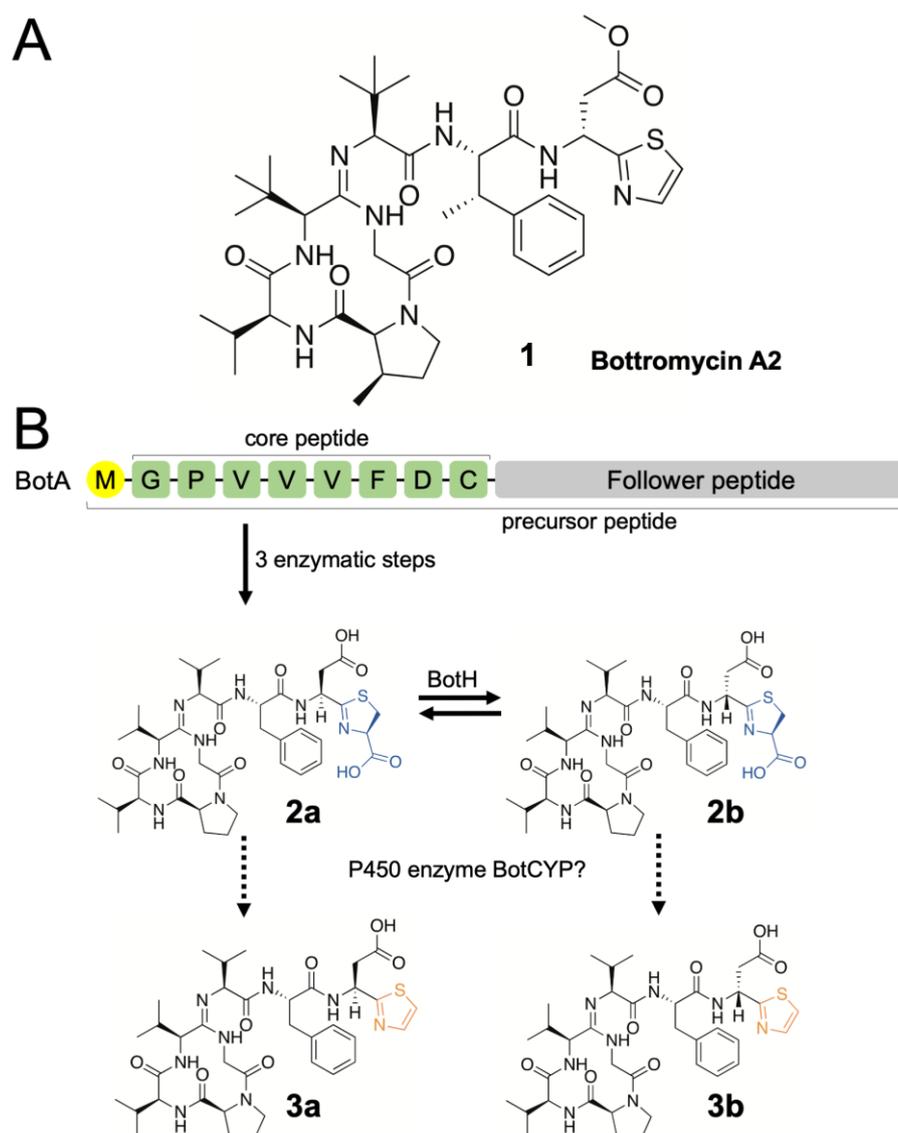
Figure 5-2 | A Structure of bottromycin A$_2$. **B** *In vitro* biosynthesis of **2a/b**, with a proposed role for the P450 enzyme BotCYP.

We then produced the putative substrate mixture **2a**/**b** biochemically *in vitro* using three enzymes (Details can be found in the section 5.3).[7, 8, 10] **2a/b** was incubated with an optimized BotCYP : BmCPR : Fdx2 ratio and a molar excess of NADPH in the absence and presence of BotCYP at 30 °C over night. The reaction mixture was subsequently analyzed by liquid chromatography-high resolution mass spectrometry (LC-HRMS, Figure 5-3A and B). In the absence of BotCYP, we observed no effect on **2a/b** ([M+H]$^+$calc.$_{mono.}$: 799.3807 Da; [M+H]$^+$obs.$_{mono.}$: r.t. 3.18 min 799.3801 Da, error -0.75 ppm, r.t. 3.02 min 799.3796 Da, error -1.38 ppm). In the presence of BotCYP, however, one major and one minor peak appeared, which had identical masses that corresponded to the decarboxylated, oxidized reaction product **3a/b** ([M+H]$^+$calc.$_{mono.}$: 753.3752 Da, [M+H]$^+$obs.$_{mono.}$: r.t. 3.41 min 753.3743 Da,

error -1.19 ppm, r.t. 3.22 min 753.3743 Da, error -1.19 ppm)) (Figure 5-3B). Analysis of both peaks by tandem mass spectrometry (MS^2) strongly supports the proposed structures (Figure 5-3C, Table 5-S1 and Table 5-S2).
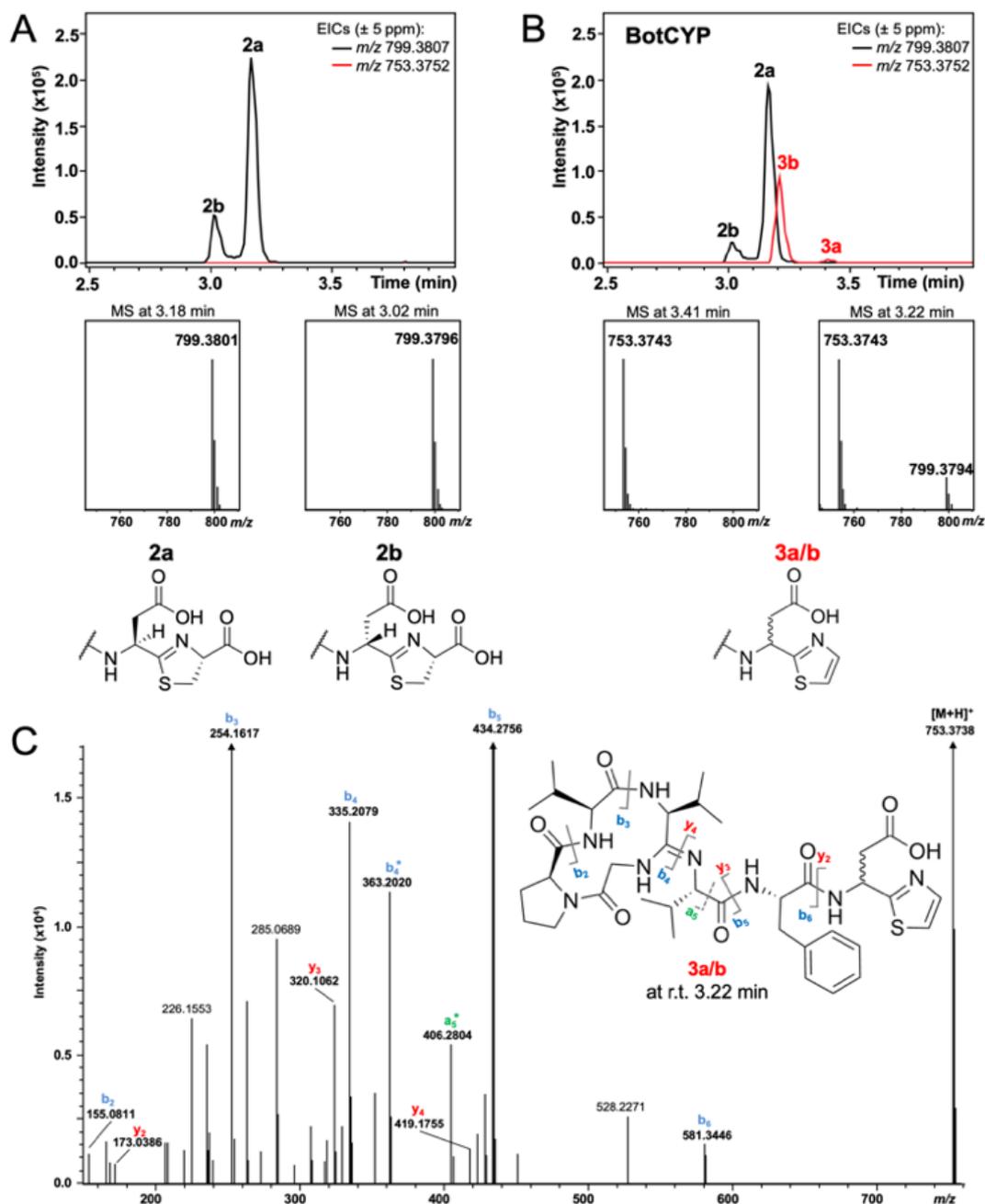
Figure 5-3 | A Incubation of putative BotCYP substrates **2a/b** with all co-factors results in no detectable change in the LC-HRMS spectrum. **B** Same as A, but with BotCYP added. The new species **3a/b** showed a loss of 46 Da, which is in agreement with the expected product. **C** MS^2 spectrum of **3a/b**. a, b and y ions are indicated. The peak list and structures of the fragments b_2 and b_3 can be found in Table 5-S1 and Table 5-S2.

These data establish BotCYP as sufficient for the oxidative decarboxylation of the C-terminal thiazoline to a thiazole. We estimate the yield of this reaction to be very poor (< 25%). The presence of a very small second product peak indicated that the enzyme may be stereospecific and selects one of the two available substrate epimers. Using an enzymatically produced, roughly equimolar mixture of **2a** and **2b** in time-course experiments allowed us to demonstrate that the D-Asp containing **2b** peak is preferentially consumed by BotCYP (Figure 5-4A). To better understand BotCYP, we attempted to determine its crystal structure, but extensive screening did not yield crystals. The close homolog SalCYP from *Salinispora tropica* (Figure 5-S1) was also able to perform the oxidative decarboxylation reaction on similar time-scales and with comparable yields using identical redox partners (Figure 5-4B, Figure 5-S3 and Figure 5-S4). The high-resolution (1.85 Å) structure of SalCYP was determined using single-wavelength anomalous dispersion from iron. All data collection and refinement statistics can be found in Table 5-S3. The overall structure of SalCYP reveals a heme b bound at the reactive centre, which is freely accessible via a solvent channel that leads to a large active site cavity suitable for binding very bulky substrates (Figure 5-S5). The heme b is coordinated by Cys354, forms two salt bridges between the two heme carboxylic acid moieties and Arg104 and Arg298 side chains of SalCYP as well as extensive hydrophobic interactions (Figure 5-S6). Surprisingly, we found the N-terminus (residues 1 – 20) to be ordered and extended away from the protein, with the amino group of the N-terminus coordinating the heme iron of a symmetry mate (Figure 5-S7). To allow an unbiased analysis of the binding pocket, we engineered a new crystal contact. This mutant protein, SalCYP^T, retained enzymatic activity (Figure 5-S8), readily crystallized in space group P4$_1$ and the structure was determined to 1.50 Å resolution. The active site was now unoccupied by protein, but retained its shape and solvent channel at the entrance to the active site cavity.

A search for structural homologs using the DALI server[20] revealed the camphor hydroxylase CYP101D2[21] as the closest structural homolog (PDB ID 4dxy) with a C$_\alpha$ rmsd of 2.4 Å over 355 residues (Figure 5-S9). The structural homology of SalCYP to the well-characterized OleTJE (PDB ID 5m0n),[22] which catalyzes the oxidative decarboxylation of fatty acids, was calculated as a C$_\alpha$ rmsd of 2.9 Å over 310 residues (Figure 5-S9).The most striking difference between SalCYP and virtually all close structural homologs was the wide and deep active-site cleft found in SalCYP (Figure 5-S10). The exception was the structure of TbtJ1 (PDB ID 5vws), which is the only other P450 enzyme structure from a RiPP pathway.[23] TbtJ1 catalyzes the hydroxylation of thiomuracin[24, 25] and possesses a similar active site topology. The residues found at the SalCYP active site are very well conserved (Figure 5-4C), which may reflect the

need for stereocontrol for this step of bottromycin biosynthesis. Taken together, these structures thus indicate that RiPP P450 enzymes have evolved to perform selective reactions within complex linear and cyclic RiPP core peptides (Figure 5-S10).
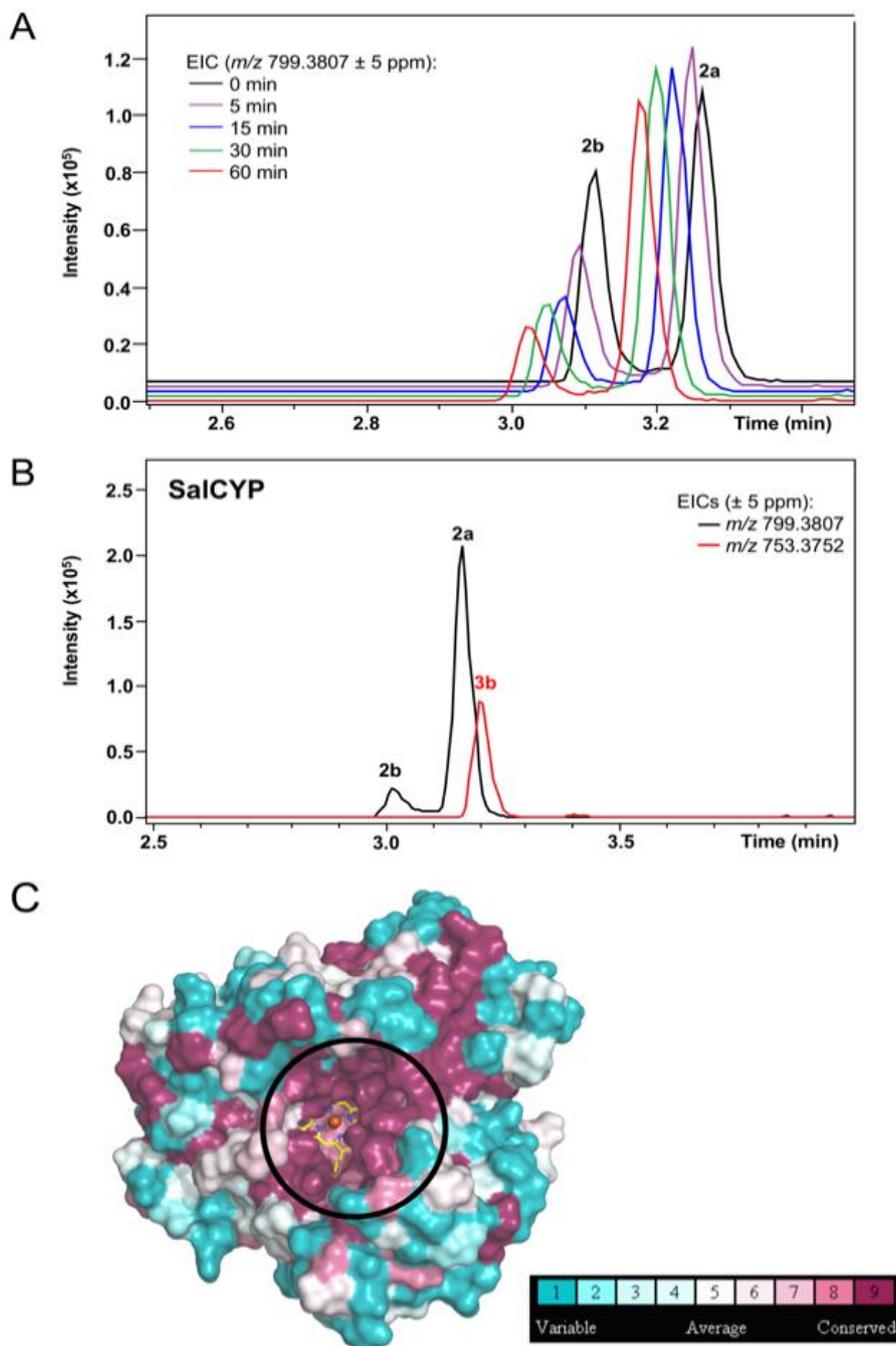
Figure 5-4 | A LC-HRMS data show that BotCYP preferentially consumes **2b**. **B** Analysis of a SalCYP reaction with **2a/b**. One product peak is observed when SalCYP is used. **C** ConSurf map showing the conservation of residues around the SalCYP active site (circle) for SalCYP homologues from other bottromycin biosynthetic gene clusters (Figure 5-5B).
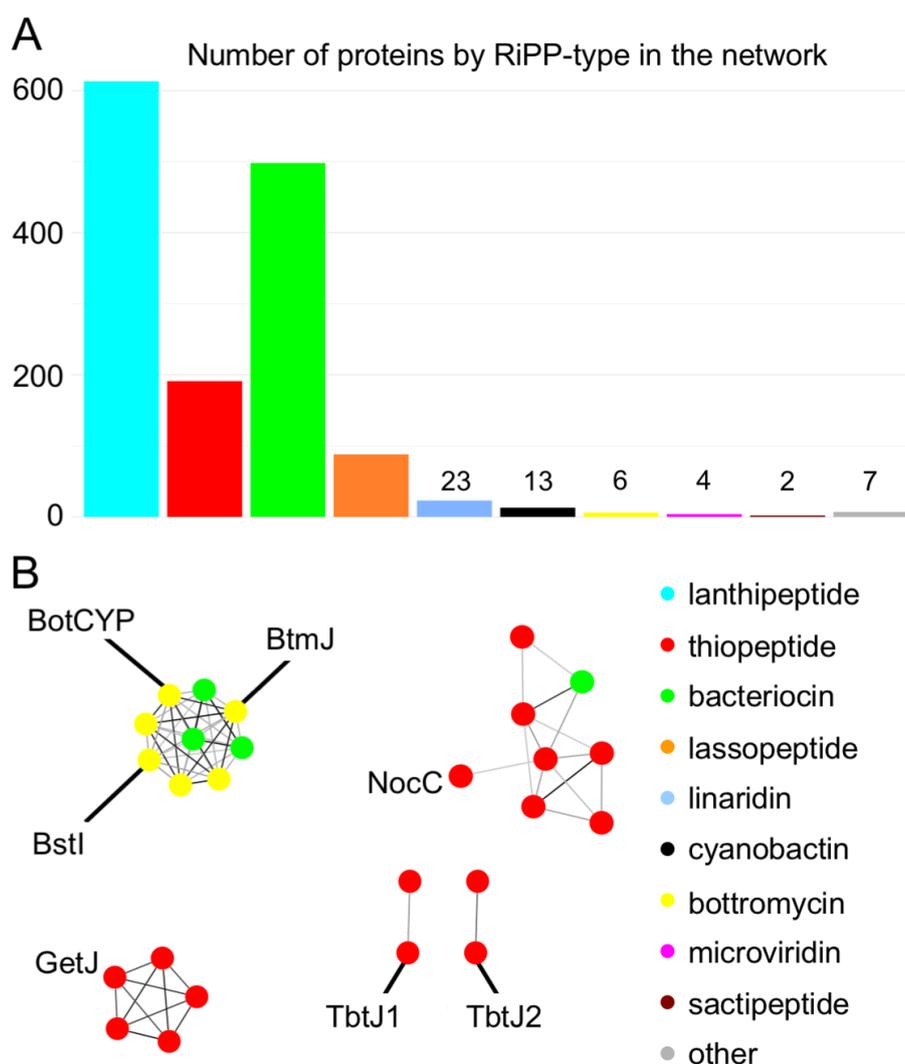
Figure 5-5 | A Histogram displaying the number of P450 enzyme containing predicted RiPP BGCs, grouped and colored by RiPP-type (See Figure 5-S11). **B** Clusters with characterized enzymes are annotated. MibO and NocV (see main text) are singlets and not displayed. Manual curation of the BotCYP-containing cluster identifies the three bacteriocin nodes as bottromycin BGCs.

We computed a sequence similarity network for BotCYP to explore the distribution of P450-like enzymes in RiPP BGCs (Figure 5-5 and Figure 5-S11). Although P450-like enzymes are widespread in bacteria (109,039 members of the Pfam family PF00067), and more than 15,000 of them are located in BGCs, only ~1.5% (1,786) were found in a predicted RiPP cluster. Sequence diversity of these proteins is very large (mean sequence identity 28%), and they are found in diverse Proteobacteria, as well as Actinobacteria, Cyanobacteria, Chloroflexi, and Bacteroidetes. BotCYP is a part of a small cluster that contains all P450 enzymes from putative bottromycin biosynthetic gene clusters, which echoes the findings for the amidohydrolase PurAH.[10] We would like to note that only a handful of RiPP P450 enzymes within this network have an assigned biosynthetic function to date: MibO (microbisporicin), TbtJ1 and TbtJ2 (thiomuracin), TsrR (thiostrepton) and NocV (nocathiacin) all catalyze the hydroxylation of

amino acids.[24-27] GetJ from GE37468 biosynthesis catalyzes the conversion of isoleucine to δ-hydroxy-proline and TsrP, also involved in thiostrepton biosynthesis, catalyzes an epoxidation that triggers additional changes.[26, 28] None of these proteins catalyze an oxidative decarboxylation reaction, nor are they involved in heterocycle biochemistry. BotCYP thus expands the catalytic scope of P450 enzymes involved in RiPP biosynthesis, and the map of the RiPP P450 landscape we provide here may serve as a very valuable starting point to explore the functionalities of P450 enzymes in RiPP biosynthesis and discover novel transformations.

The spontaneous exchange between **2a** and **2b** is slow,[6, 11] and **2b** (D-Asp), the much less abundant epimer, was the preferred substrate for BotCYP. This led us to speculate that the poor yield of the BotCYP and SalCYP reactions may be the result of depleting **2b**. We thus added the recently characterized epimerase BotH,[11] which catalyzes the rapid conversion of **2a** to **2b**, to the reaction mixture. Analysis of reactions containing BotH and BotCYP by LC-HRMS demonstrated that the oxidative decarboxylation reaction now nearly goes to completion (Figure 5-6A). Since **2b** contains D-Asp and was preferentially consumed by BotCYP, it may be intuitive to assume that the reaction product, **3b**, also harbors the D-Asp found in the natural product. To probe the stereochemistry at the Asp C_α-position of **3b**, we enzymatically produced substrate selectively labeled with a deuteron at this position (**4a/b**, $[M+H]^+$calc._mono._: 800.3870 Da; r.t. 3.15 min (**4a**) $[M+H]^+$obs._mono._: 800.3851 Da, error -2.37 ppm and r.t. 3.00 min (**4b**) $[M+H]^+$obs._mono._: 800.3852 Da, error -2.24 ppm) (Figure 5-S12).

We then used **4a/b** as a substrate in a BotCYP reaction in H_2O and the resulting peak had the same retention time and fragmentation pattern as **3b** (Figure 5-6B and Figure 5-S13). Crucially, the product peak (**5b**) retained the mass shift of + 1 Da ($[M+H]^+$calc._mono._: 754.3815 Da, $[M+H]^+$obs._mono._: 754.3801 Da, error -1.86 ppm), indicating that the configuration of the Asp C_α-position is retained during the oxidative decarboxylation reaction. Presence of the deuteron at the Asp-position was confirmed by MS^2 (Figure 5-S13). Our data thus strongly imply that **3b** represents the bottromycin core scaffold (des-methyl **1**) harboring the D-Asp moiety found in bottromycins. Incubation of **5b** with BotH in H_2O does no longer lead to an exchange of the deuteron at the Asp C_α-position, indicating that heterocycle oxidation to a thiazole locks the configuration at the Asp position.[11]
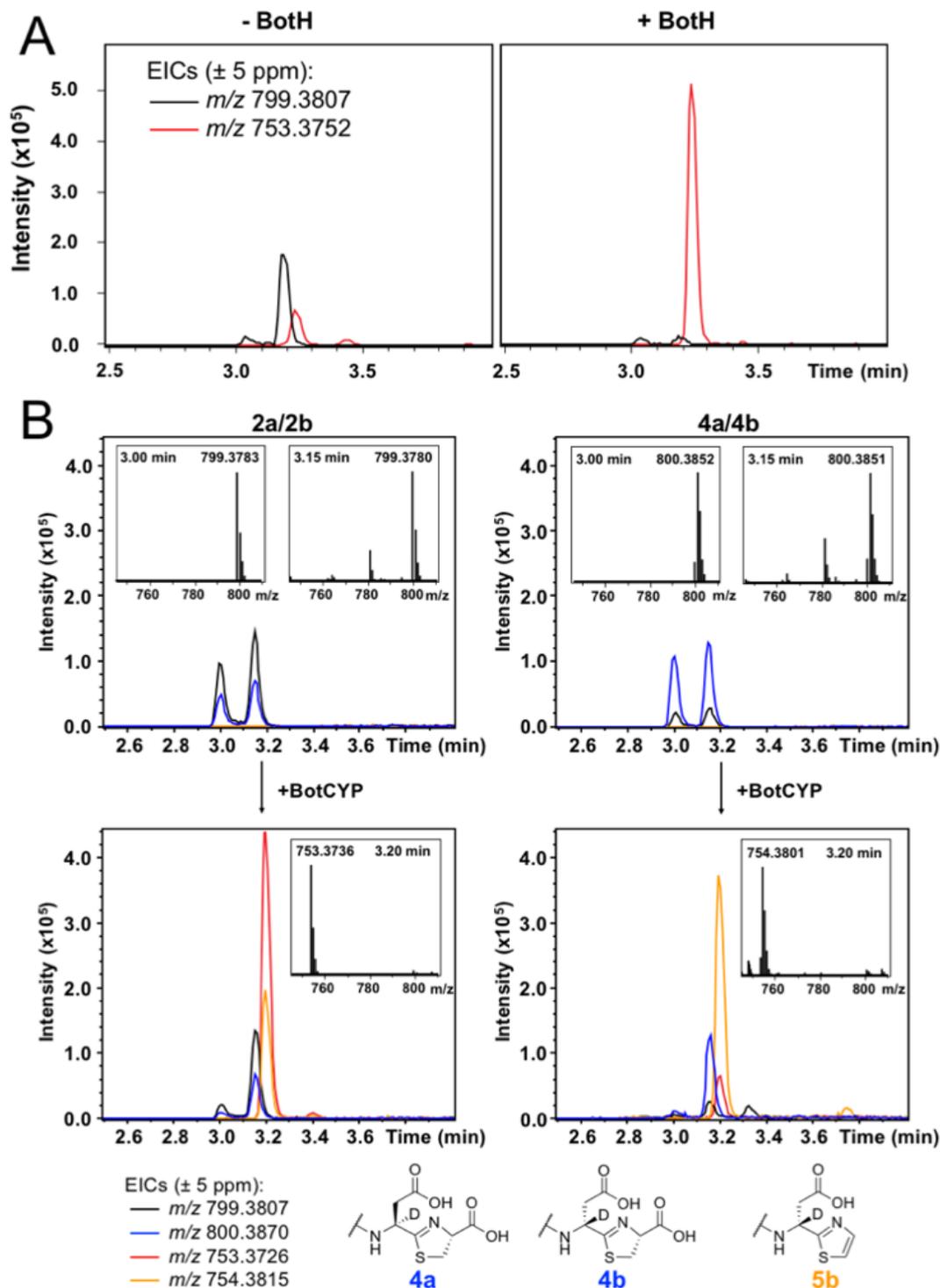
Figure 5-6 | A Comparison of the BotCYP reaction with (right) and without (left) BotH. Addition of BotH enables virtually complete turnover. **B** BotCYP reaction using **4a/b** as a substrate. The product **5b** has the same retention time as **3b**, but retains its mass shift of +1 Da.

We have demonstrated that the P450 enzyme found in the bottromycin BGC is sufficient for the oxidative decarboxylation of the C-terminal thiazoline found in the bottromycin biosynthetic intermediate **2a/b**. The stereospecificity of BotCYP (and SalCYP) provides stereochemical resolution for the pathway, since the epimerase BotH produces a product mixture (D- and L-Asp).[11] Our data also rationalize the results of a previous study attempting to

produce bottromycin derivatives at the Asp-position (Asp to Thr, Ala, or Asn).[4] Mutations in this position appear incompatible with BotH epimerase activity,[11] resulting in very little, if any, D-amino acid substrate for BotCYP and thus no bottromycin derivative production. To function efficiently, BotCYP needs to cooperate with the epimerase BotH,[11] but we were unable to detect complex formation between the two proteins (data not shown). This mirrors the results for the bottromycin biosynthetic enzymes PurCD (macrocyclase) and PurAH (amidohydrolase), which are both required for efficient macroamidine formation, but do not form a complex.[10] It is still unclear if the bottromycin biosynthetic machinery is organized within the producing strains, for example by co-localizing enzymes without physical complex formation. The reconstitution of BotCYP also opens a facile, 2-pot *in vitro* biosynthetic route to the bottromycin core scaffold. Consequently, it might be a valuable alternative to synthetic approaches of bottromycin production, which currently still involve steps with only a moderate yield for the final macrocyclization reaction.[5, 29] Our data may thus enable the rapid generation of bottromycin analogues for antibiotic activity testing.

5.3 Supporting Information

A

S. sp. BC16019

botT — botOMT — botRMT1 — botA / botC — botCD — botRMT2 — botH — botAH — botCYP — botRMT3 — botR / botP

S. tropica

Integrase — salRMT2 / salH — salAH — salCYP — salRMT3 — salR / salP
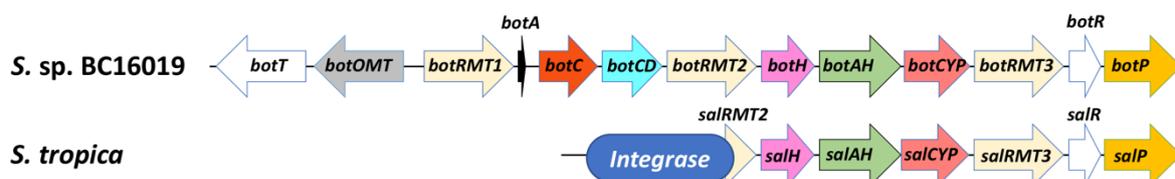
B

```
CLUSTAL O(1.2.4) multiple sequence alignment


BotCYP      MQADREPDRTGREAKRPAEPDHPPQQATCPVDFDFFAAPQTYRDAAAEHAGEHGAFYSD-      59
SalCYP      MQADKCPV----------TGATATTARGDSVEFDFFAAPQAYRRVAAEHARD-GAFHSSR      49
            ****: *                    *:********:** .***** : ***:*.

BotCYP      -RGFWVLTTYDGIVDAFKDEGTFTVGRVSAAEGAEEERWIPLTVEGREHTAWRQRLGAWF      118
SalCYP      GDSFWVLSTYEGICAAFRDEDTFSVSRVSAADGAEDERWIPLTIQGRTHNEWRRRLAAWF      109
             .****:**:**   **:**.**:*.*****:***:********:.** *. **:**.***

BotCYP      TPQRVRELTPSMRAGARRRIEGFLEKGEVSFNEDFARPYVLENLMTAVGWPPDGFDLLIA      178
SalCYP      TPQRARDLTPAIRANARRRISAFVDRGEVSFSDEFARPYVLENLMLAVGWPLADLDHLLA      169
            ****.*:***::**.*****..*::*****.:: *********** *****   .:* *:*

BotCYP      INRAMIDSRSAPDPRAAAYGELGLPALERFAREHIARRRAEPAD-DLTTASFGWEIDGAE      237
SalCYP      INVAMIRSREAPDPRQAFNAETAFPALQEYVRRHVARRRAEPVEGDLTSATFDWEIDGTP      229
            ** *** **.***** *   .*  .:***::..*.*:******.: ***:*:*.*****:

BotCYP      VTDDDRASLLCTLFLAGIDSTVNHLANAVQHLAHHEEDRRRFLAGPEVRPPAVEEFLRAN      297
SalCYP      VSDADRESLLTVLFLAGVDSTVNHMANGIQHLAHHPGDRHRFLRDPEVRPAAVEEFLRVN      289
            *:* ** *** .*****:******:** :****** **:*** .***** *******.*

BotCYP      SCMYPGRQAATGGAGGVADRGDTVLLPLALANHDPEVFPEPGRIDFDRTRNPHIAFGTGP      357
SalCYP      SCMYPGRLATREGAGGVASQGDTVLLPLALANYDPAVFPEPERVDFDREQNPHIAFGTGH      349
            ******* *:  ******.:*************:** ***** *:**** :*********

BotCYP      HQCLGAAFARAQILVALEEWHALVPDYAPSPEQRTAEPPFLRNSYDLRLVW      408
SalCYP      HQCLGAAYARAQILTAWEEWHELIPDYRLP--DPTVEPPFLRNVYDLRIVW      398
            *******:******.* **** *:***    : *.******* ****:**
```

Figure 5-S1 | A Comparison of the bottromycin biosynthetic gene cluster from *Streptomyces* sp. BC16019 and the truncated gene cluster found in *Salinispora tropica*. An integrase in the 5' region of *salRMT2* resulted in the deletion of the first half of the cluster, including the precursor peptide. **B** Pairwise sequence alignment of SalCYP (RefSeq WP_051425703) with BotCYP (UniProt ID K4MJU2). The sequence identity is 70%, not accounting for the non-conserved N-terminus (alignment starting -VDFDFF). The alignment was generated with *Clustal Omega*.[30]
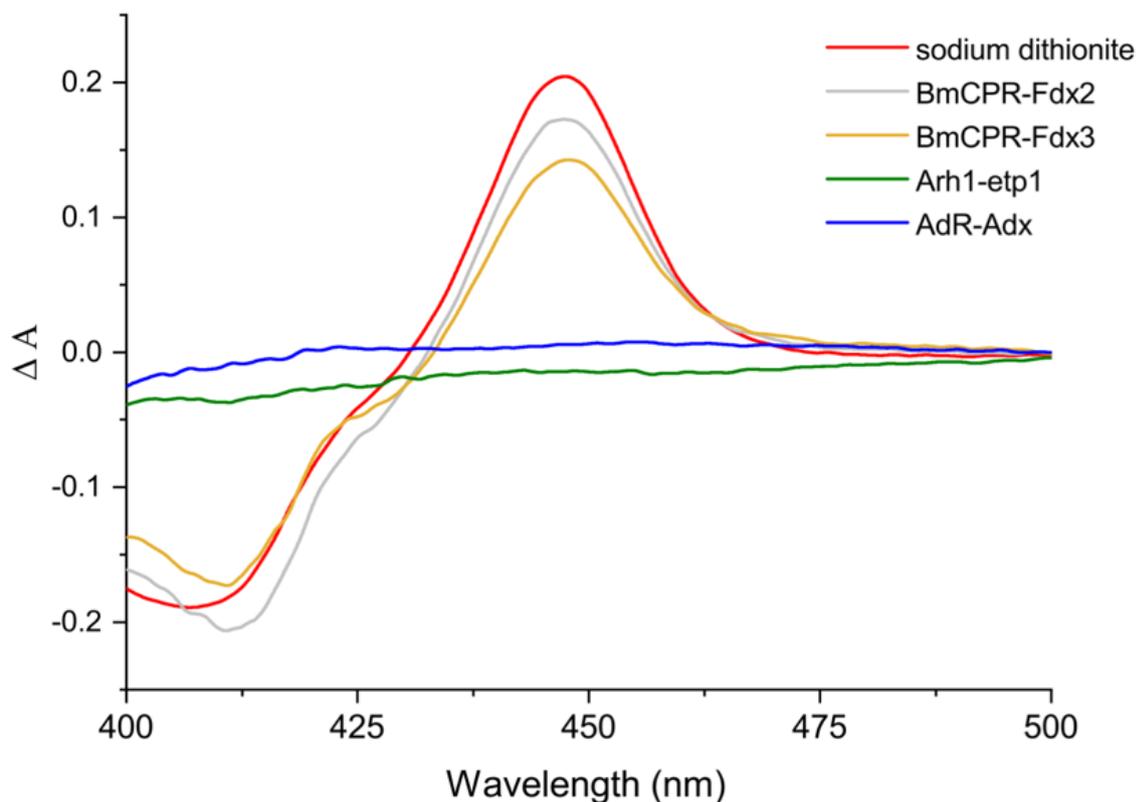
Figure 5-S2 | Identification of suitable redox partners for BotCYP. The dithionite reduced CO-difference spectrum of BotCYP (red line) shows the typical peak maximum at 448 nm for the Fe^{II}-CO complex. This spectrum was compared with the NADPH reduced BmCPR-Fdx2 (grey line), BmCPR-Fdx3 (yellow line), Arh1-Etp1^{fd} (green line) and AdR-Adx$_{4-108}$ (blue line) CO-complexed spectra. Details can be found in the methods section. The NADPH (1 mM) reduced CO-difference spectra were recorded in a 200 µl mixture of CYP/ferredoxin/reductase with a 2:40:6 molar ratio in KPP buffer (10 mM K_2HPO_4 and KH_2PO_4). These results showed that the redox partners BmCPR and Fdx2/3 were able to transfer, at least, the first electron to BotCYP. The BmCPR-Fdx2 redox system was selected for further investigation with BotCYP.
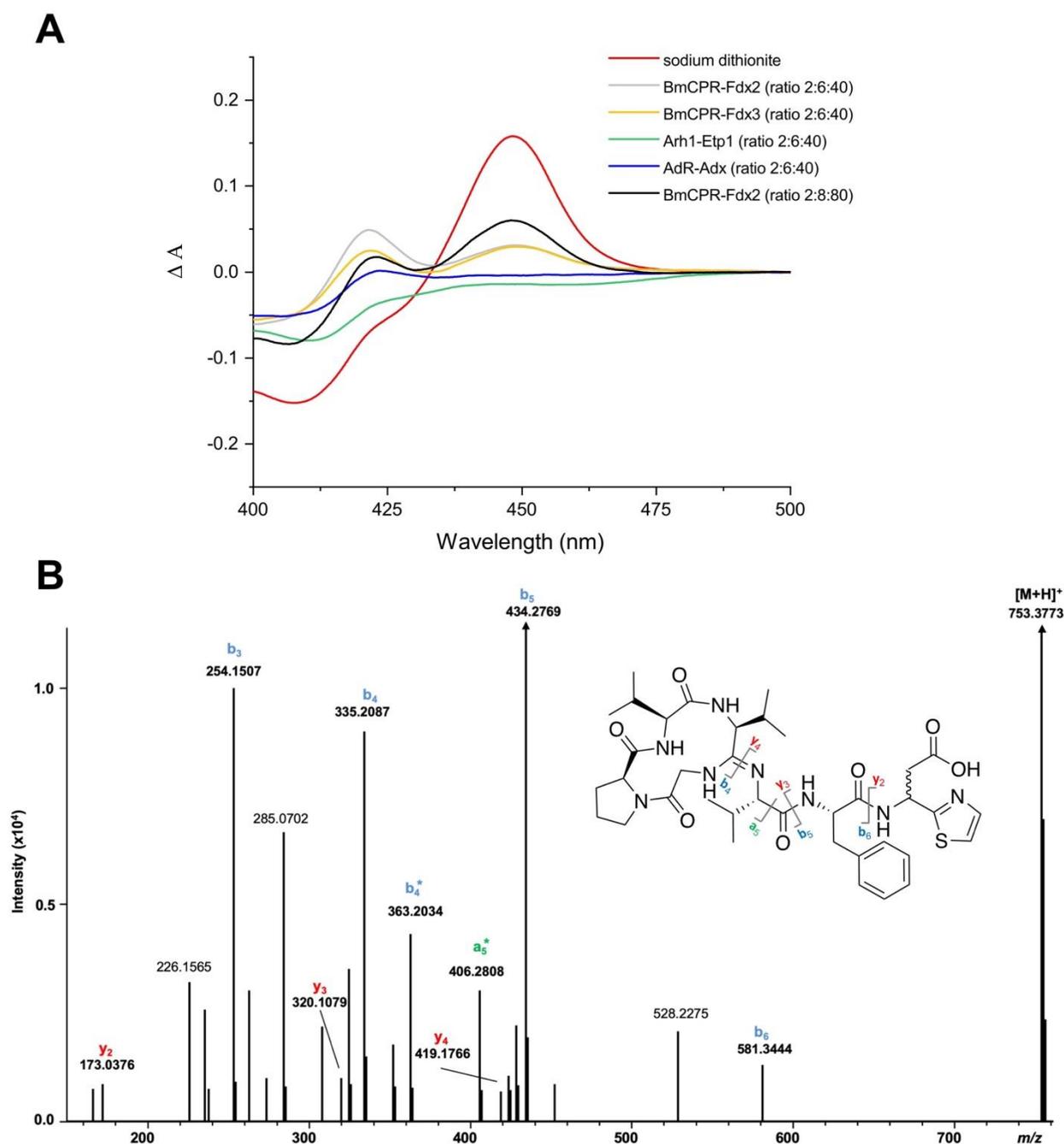
Figure 5-S3 | A Identification of suitable redox partners for SalCYP. The dithionite reduced CO-difference spectrum of SalCYP (red line) shows the typical peak maximum at 448 nm for the Fe^{II}-CO complex. This spectrum was compared with the NADPH reduced BmCPR-Fdx2 (grey and black line), BmCPR-Fdx3 (yellow line), Arh1-Etp1^{fd} (green line) and AdR-Adx$_{4-108}$ (blue line) CO-complexed spectra. Details can be found in the methods section. The NADPH (1 mM) reduced CO-difference spectra were recorded in a 200 µl mixture of CYP/ferredoxin/reductase with a 2:40:6 (and 2:80:8 for BmCPR-Fdx2) molar ratio in KPP buffer (10 mM K_2HPO_4 and KH_2PO_4). These results showed that the redox partners, BmCPR and Fdx2/3, were able to transfer, at least, the first electron to SalCYP. The BmCPR-Fdx2 redox system was selected for further investigation with SalCYP. **B** MS^2 spectrum of the SalCYP product. a, b and y ions are indicated.
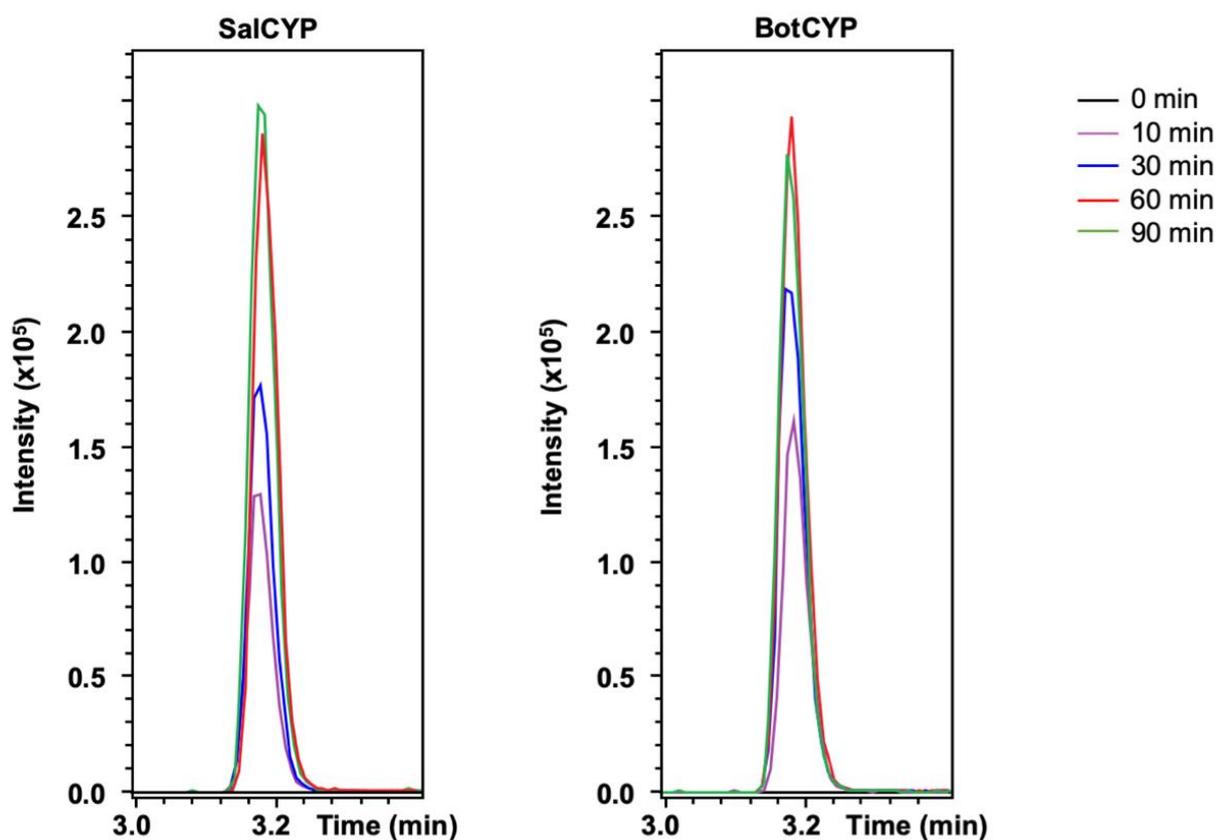
Figure 5-S4 | Comparison of BotCYP and SalCYP reactions with samples taken at different time points. Both proteins are able to process the **2a**/**b** substrate at comparable speed.
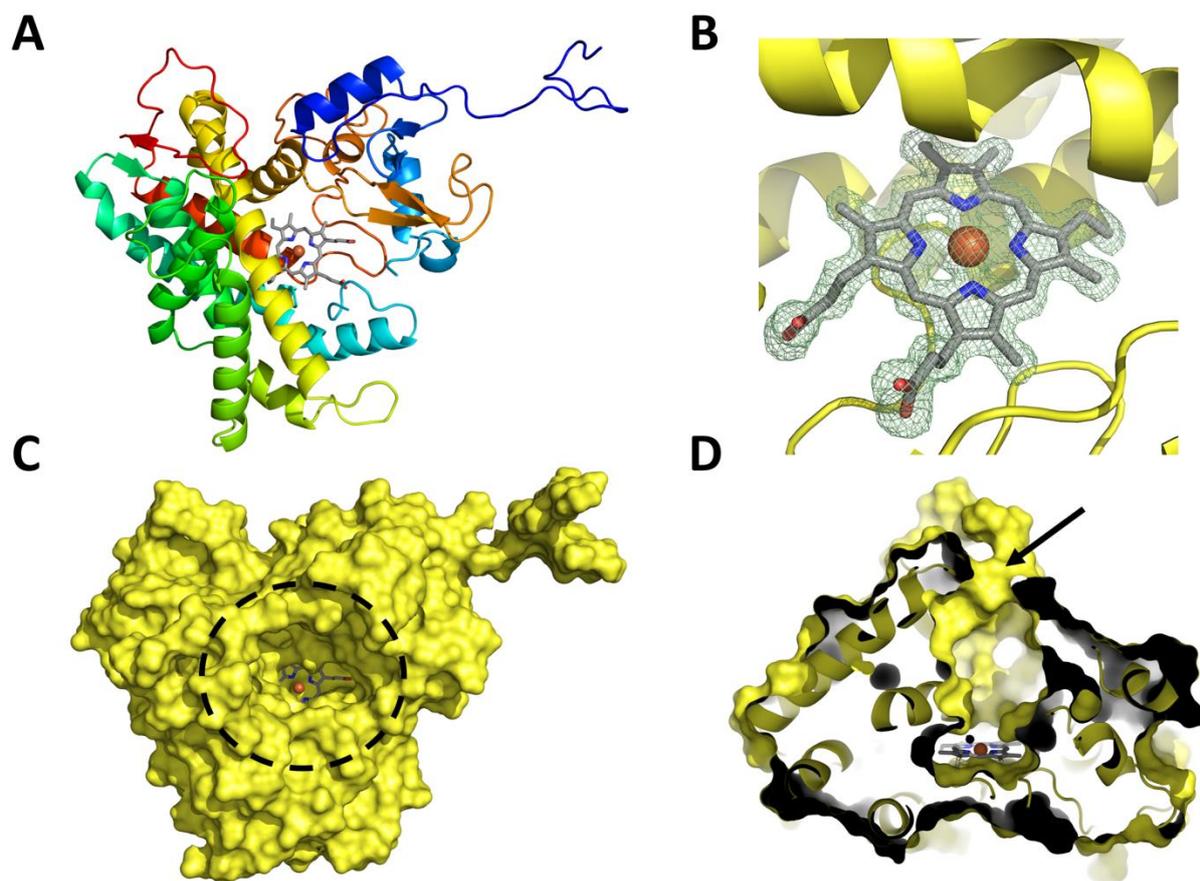
Figure 5-S5 | A Rainbow cartoon representation of SalCYP, blue is representing the N- whereas red is representing the C-terminus. We observe the N-terminus extending away from the rest of the protein. **B** Electron difference map for the heme b found at the active site. The electron difference map (F_o-F_c) was calculated using a model without the ligand and contoured at 3.0 σ. Heme molecules are shown as sticks, iron ions as orange spheres. **C** Surface representation of SalCYP, highlighting accessibility of the heme b in the active site (circle) for potential substrates. **D** Clipped surface of SalCYP, highlighting the entrance to the active site cleft (arrow) leading to the heme b.
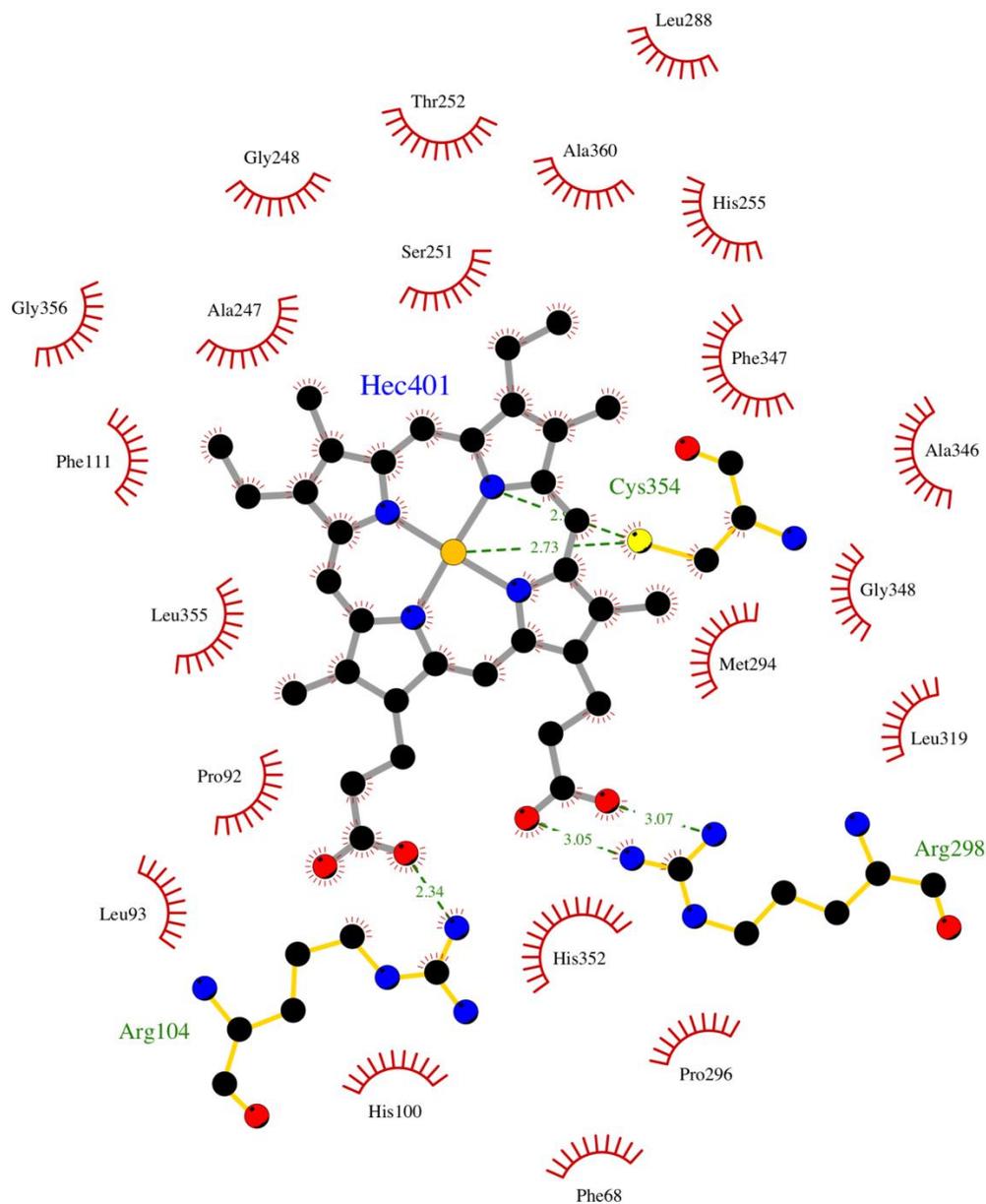
Figure 5-S6 | LigPlot[+31] diagram for the interactions between the heme b and SalCYP. Ligand bonds are grey, protein bonds yellow. Hydrogen bonds and salt bridges are indicated by green, dashed lines and distances given. Hydrophobic interactions are represented as red spoked arcs.
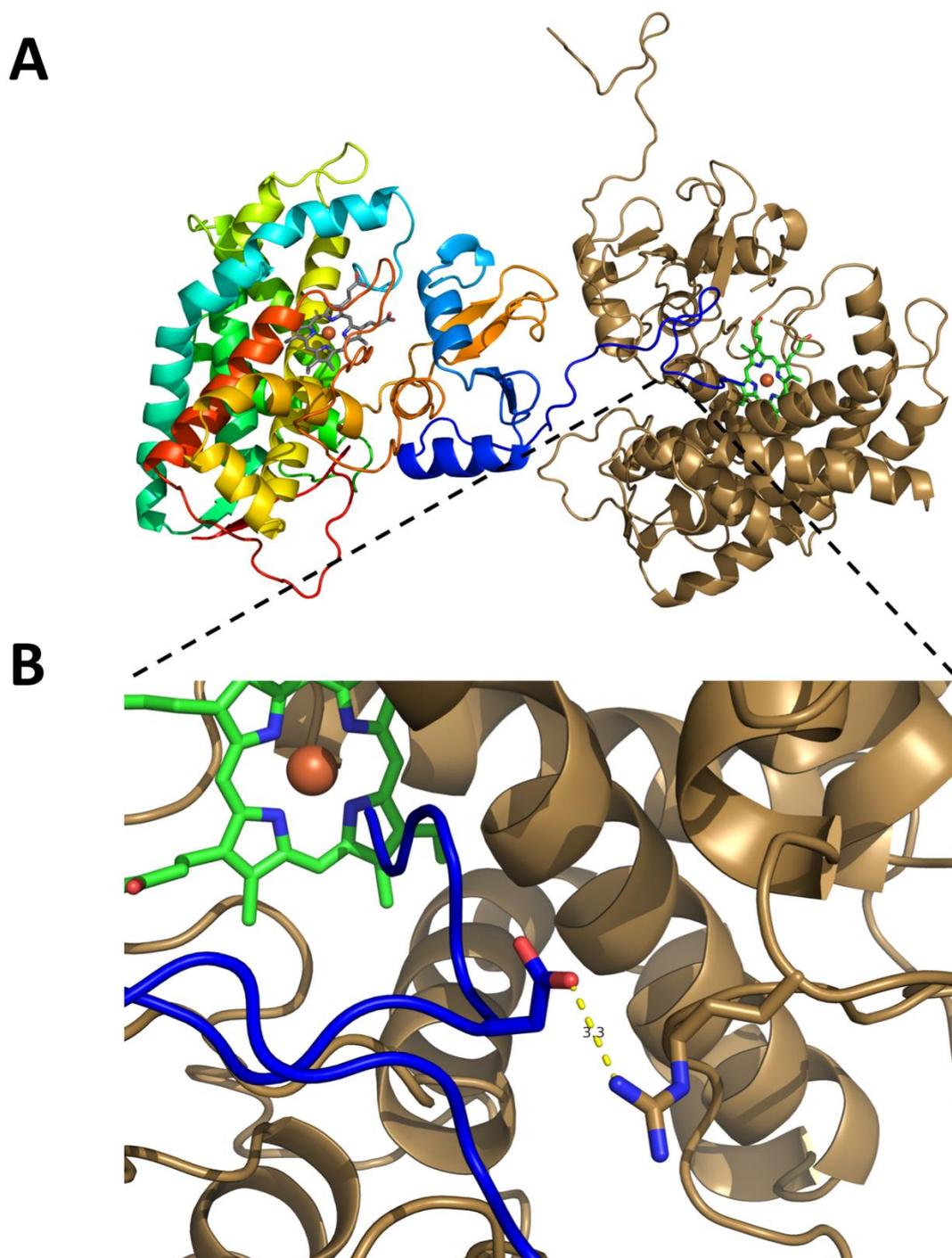
Figure 5-S7 | A Interaction of the SalCYP N-terminus (blue) with the heme iron of a symmetry mate (brown cartoon). Heme molecules are shown as sticks, iron ions as orange spheres. **B** Formation of a salt bridge between Asp4 and Arg178 of each monomer. It is a crucial interaction for the coordination of the N-terminus into the active site of the symmetry mate and possibly required for crystallization. The heme molecules are shown as sticks, the iron ion as an orange sphere.
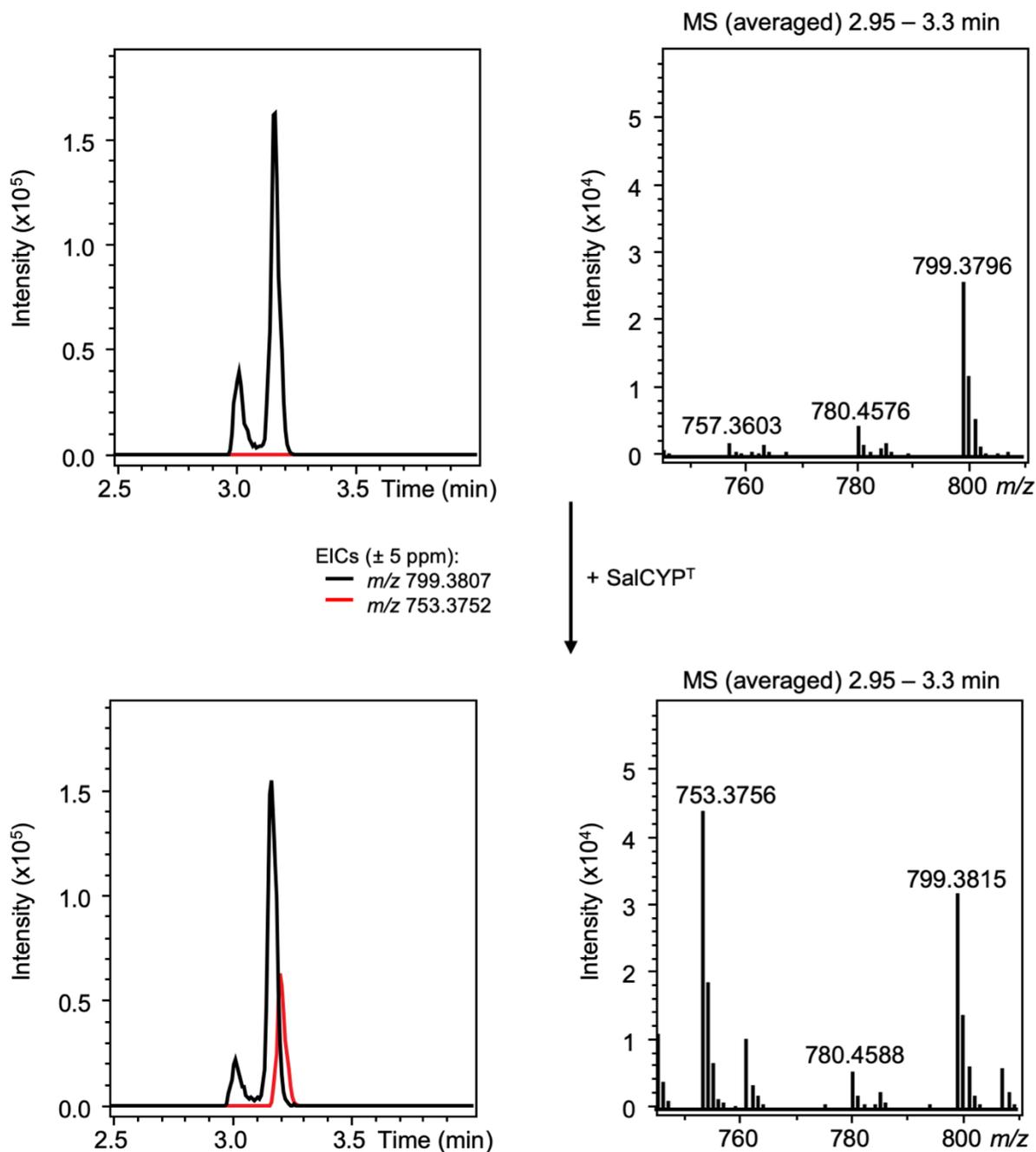
Figure 5-S8 | LC-HRMS analysis of a SalCYP^T reaction with **2a/b**. As can be seen, the truncated protein is able to turn over substrate.

211
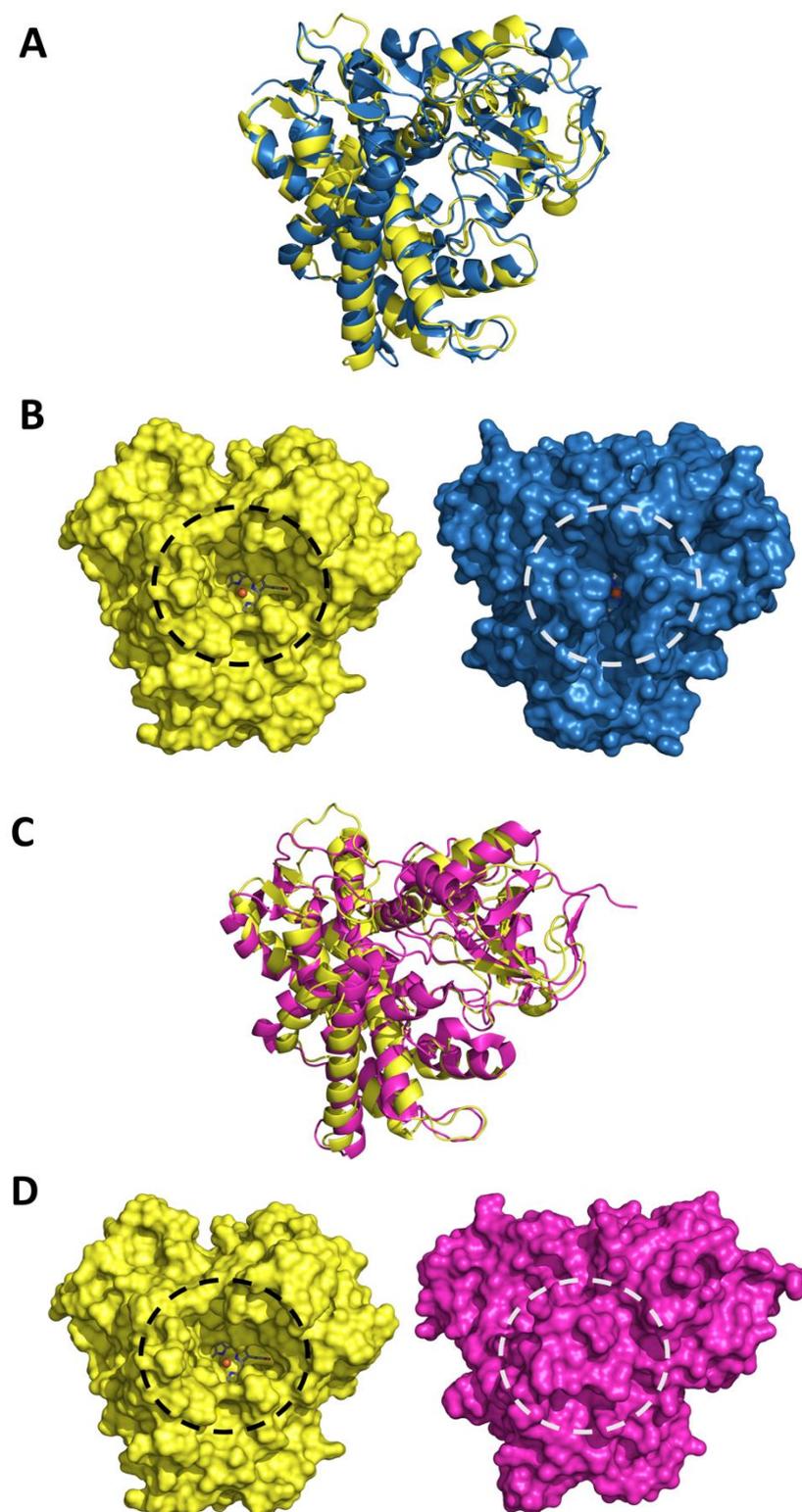
Figure 5-S9 | Comparison of the active sites of the closest structural homologue CYP101D2 (blue, PDB ID 4dxy) and the well characterized homologue OleTJE catalyzing an oxidative decarboxylation (magenta, PDB ID 5m0n) with SalCYP^T (yellow). **A** Superposition of cartoon representations of SalCYP^T and CYP101D2 (C$_\alpha$ rmsd 1.5 Å over 275 atoms). **B** Surface comparison of SalCYP^T and CYP101D2. **C** Superposition of cartoon representations of SalCYP^T and OleTJE (C$_\alpha$ rmsd 2.9 Å over 310 atoms). **D** Surface comparison of SalCYP^T and OleTJE. Compared to SalCYP, the active site heme of OleTJE is inaccessible for larger substrates.
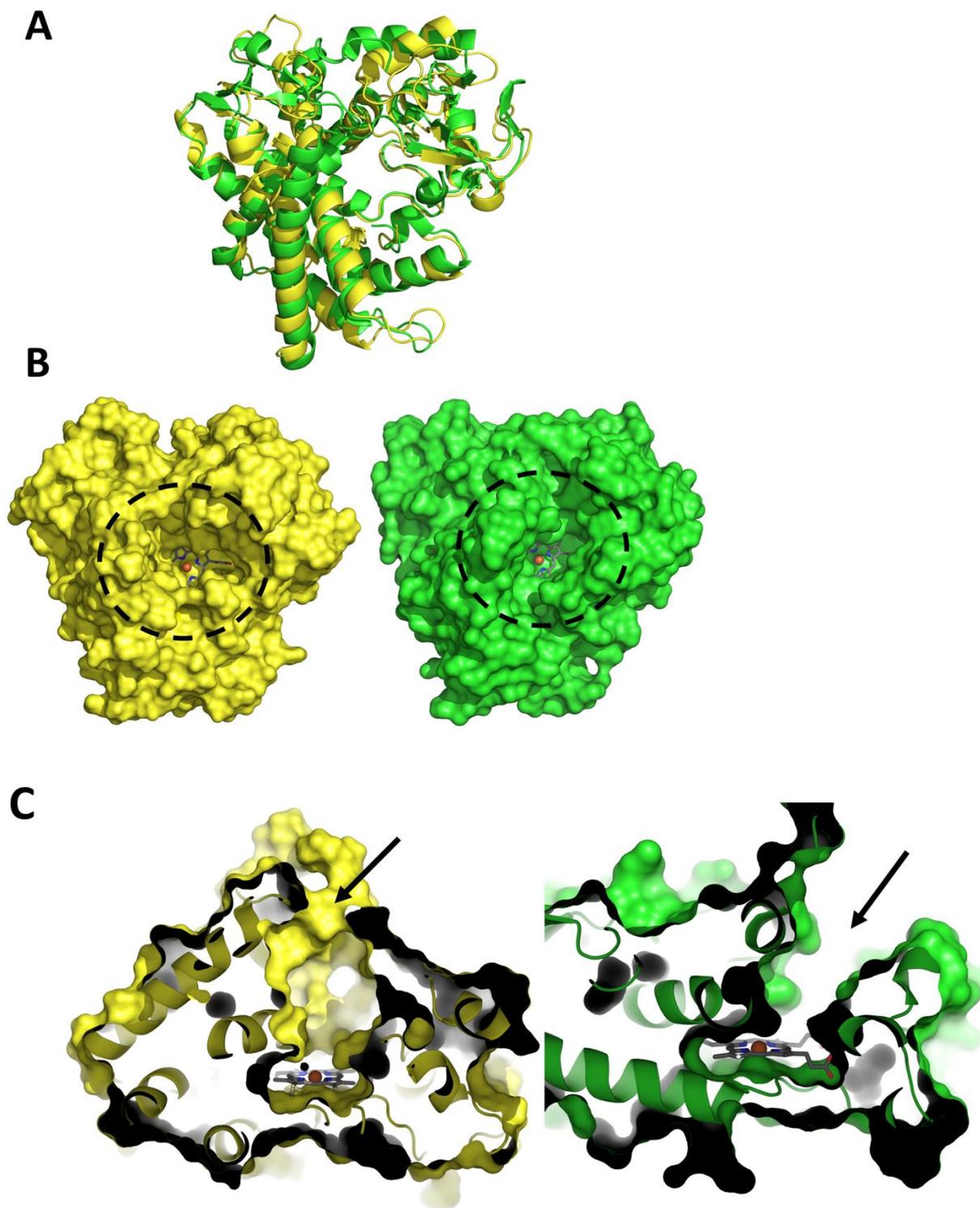
Figure 5-S10 | Comparison of the accessibility of the active sites of SalCYP^T (yellow) and thiomuracin hydroxylase TbtJ1 (PDB ID 5vws, green). **A** Superposition of the cartoon representations of both proteins. The structural homology was calculated as a C$_\alpha$ rmsd of 1.6 Å over 285 atoms. **B** Surface representation of both proteins. Heme molecules are shown as grey sticks, while iron ions are shown as orange spheres. The active site of TbtJ1 is also open, although less so than SalCYP. Both proteins accept much larger substrates than most small molecule-modifying Cytochrome P450 enzymes. **C** Clipped surfaces of SalCYP and TbtJ1, highlighting the entrance to the active site cleft (arrows) leading to the heme.

Figure 5-S11 | Sequence similarity network generated from all bacterial proteins of the PF00067 Pfam family that were predicted to lie in or near to a RiPP biosynthetic gene cluster. Only edges corresponding to sequence identity >50% are shown, singleton nodes are omitted. Predicted cluster specificity is shown in color. Clusters with characterized enzymes are annotated. MibO and NocV (see main text) are singlets and not displayed. The associated Cytoscape files and a full list of identified homologs can be found in separate supplementary files.
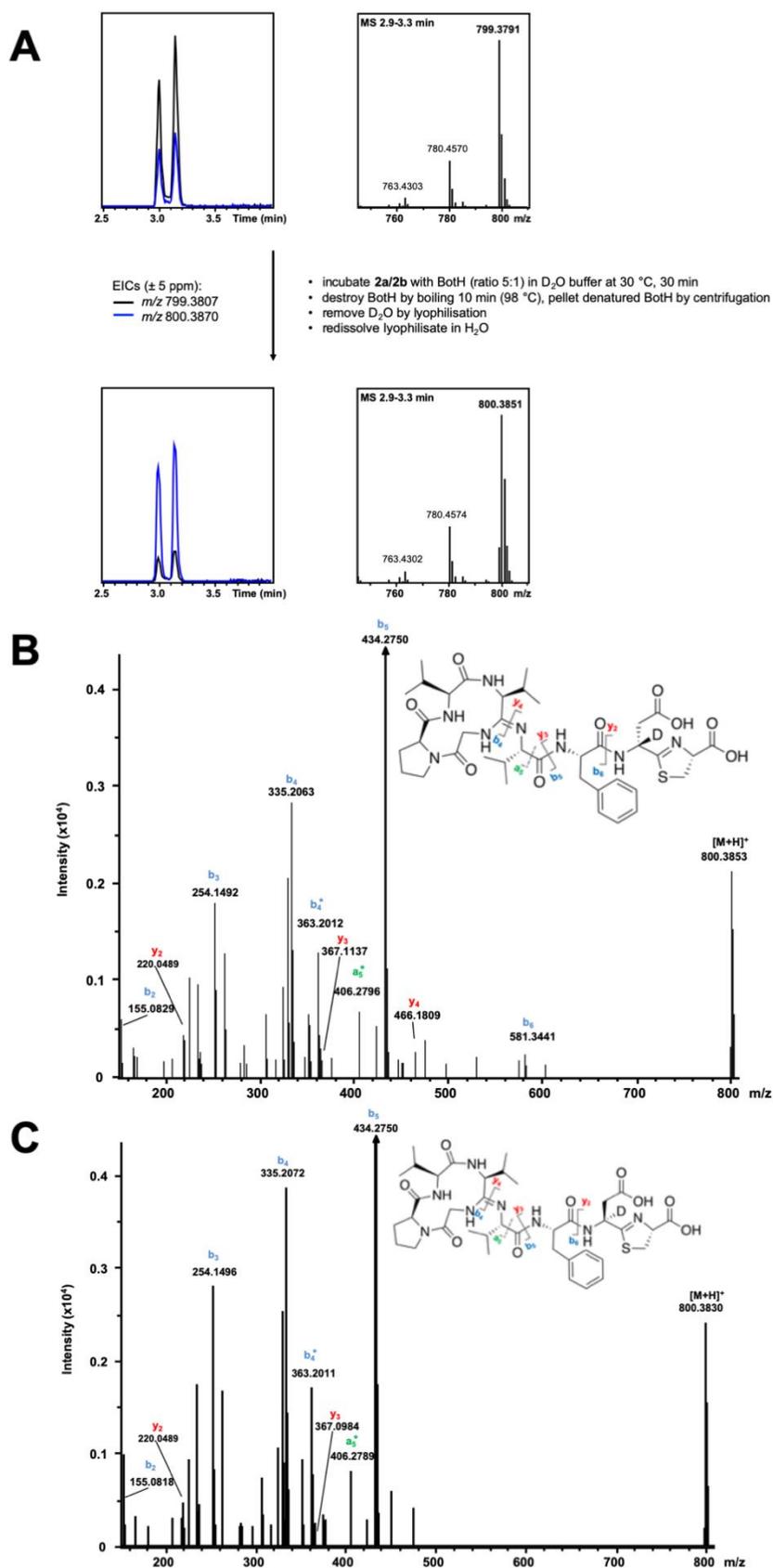
Figure 5-S12 | A Production of substrates **4a/4b**, which contain a deuteron at the α-Carbon of the aspartate residue. Substrates **4a/4b** show the same retention times that were observed for **2a/2b** and a mass shift of +1 Da. **B** MS² spectrum of **4b**. a, b and y ions are indicated. The peak list can be found in Table 5-S4. **C** MS² spectrum of **4a**. a, b and y ions are indicated. The peak list can be found in Table 5-S5.
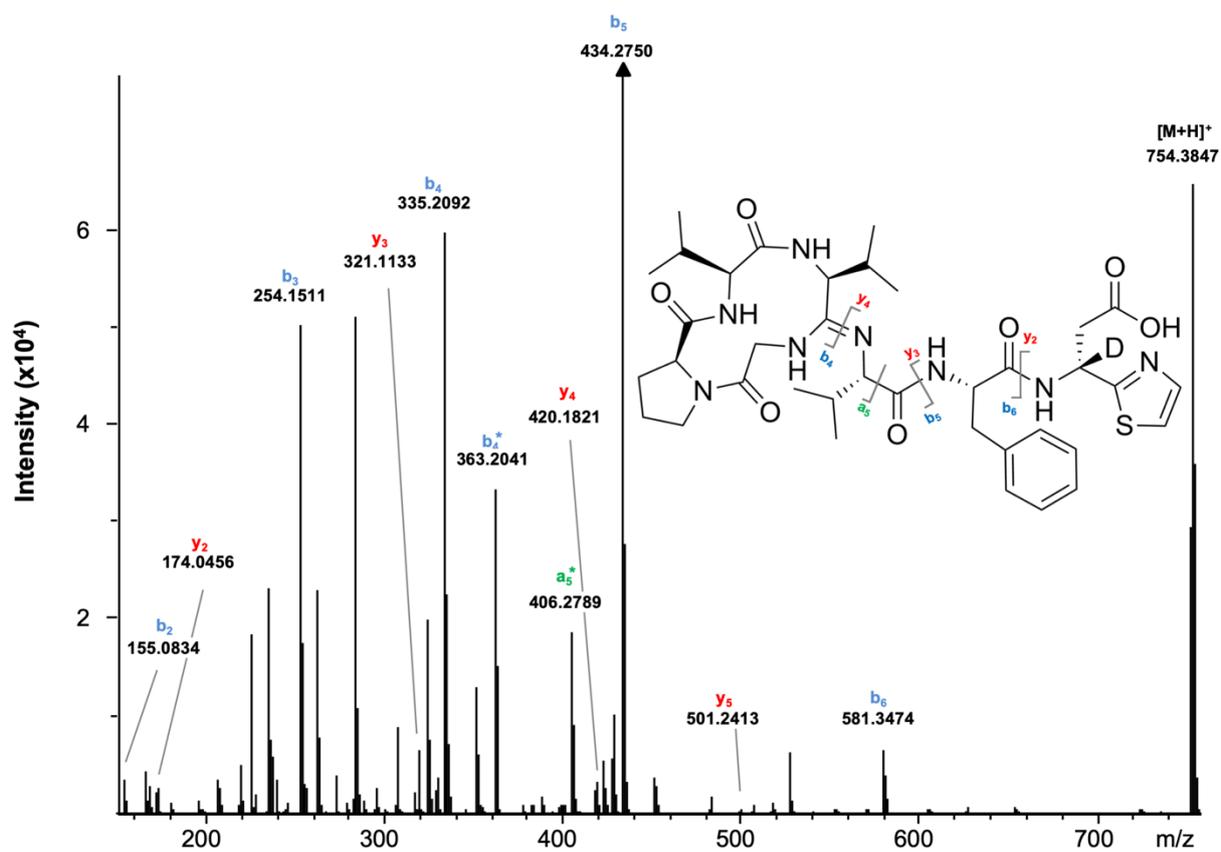
Figure 5-S13 | MS^2 spectrum of **5b**. a, b and y ions are indicated. The observed mass shift of +1 Da during the production of **4a/4b** is retained after the oxidative decarboxylation has occurred. The peak list can be found in Table 5-S6.

216

Table 5-S1 | MS^2 fragmentation data for **3a** from a BotCYP reaction.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1			8	753.3743	753.3752
P	2	-	155.0821	7	-	696.3543
V	3	254.1507	254.1605	6	-	599.3016
V[a]	4	335.2192	335.2083	5	-	500.2332
V	5	434.2765	434.2767	4	419.1606	419.1753
F	6	581.3420	581.3451	3	320.1040	320.1069
D[b]	7			2	173.0392	173.0385
C[b]	8			1		

[a] Gly1 macrocyclisation with Val4 carbonyl
[b] Cys8 heterocyclisation with Asp7 carbonyl and decarboxylation

Additionally, fragment a_5 (obs. 406.2808; calc. 406.2818) and b_4* (obs. 363.2005; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been observed previously for bottromycins and macrocyclic peptides, were detected.
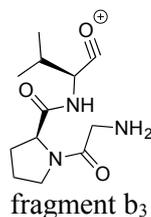
fragment b$_3$

Table 5-S2 | MS^2 fragmentation data for **3b** from a BotCYP reaction, corresponding to Figure 5-3.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1			8	753.3746	753.3752
P	2	155.0811	155.0821	7	-	696.3543
V	3	254.1617	254.1605	6	-	599.3016
V[a]	4	335.2079	335.2083	5	-	500.2332
V	5	434.2756	434.2767	4	419.1755	419.1753
F	6	581.3446	581.3451	3	320.1062	320.1069
D[b]	7			2	173.0386	173.0385
C[b]	8			1		

[a] Gly1 macrocyclisation with Val4 carbonyl
[b] Cys8 heterocyclisation with Asp7 carbonyl and decarboxylation

Additionally, fragments a_5 (obs. 406.2804; calc. 406.2818) and b_4* (obs. 363.2020; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been previously observed for bottromycins and macrocyclic peptides, were detected.

fragment b$_3$ fragment b$_2$
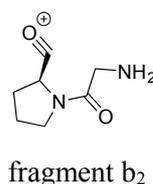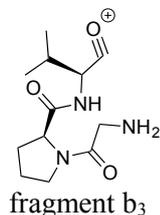
Table 5-S3 | Data collection and refinement statistics.

	SalCYP	**SalCYP^T**
PDB code	7aba	7abb
Data collection		
Space group	P 21	P 41
Cell dimensions		
a, b, c (Å)	50.45 97.40 93.28	96.53 96.53 49.99
α, β, γ (°)	90.00 104.57 90.00	90.00 90.00 90.00
Resolution (Å)	48.82 - 1.85	68.26 - 1.50
	(1.92 - 1.85)	(1.58 - 1.50)
R_{merge}	0.158 (0.699)	0.072 (0.889)
$I / \sigma I$	17.9 (4.7)	15.1 (2.1)
Completeness (%)	99.61 (96.61)	99.89 (99.90)
Redundancy	28.7 (23.2)	9.8 (9.9)
Refinement		
Resolution (Å)	48.82 - 1.85	48.26 - 1.50
No. reflections	74132 (7163)	73831 (7347)
R_{work} / R_{free}	0.1415/0.1757	0.1634/0.1812
No. atoms	7387	3245
Protein	6214	2896
Ligand/ion	86	43
Water	1087	306
B-factors	19.62	29.90
Protein	18.18	29.10
Ligand/ion	8.43	19.21
Water	28.69	39.00
R.m.s. deviations		
Bond lengths (Å)	0.005	0.008
Bond angles (°)	0.78	0.97

*Values in parentheses are for the highest-resolution shell.

Table 5-S4 | MS^2 fragmentation data for **4b**.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1	-	-	8	800.3853	800.3870
P	2	155.0829	155.0821	7	-	-
V	3	254.1492	254.1605	6	-	646.3133
V[a]	4	335.2063	335.2083	5	-	547.2449
V	5	434.2750	434.2767	4	466.1809	466.1871
F	6	581.3341	581.3451	3	367.1137	367.1186
D[b,c]	7	-	-	2	220.0489	220.0502
C[b]	8	-	782.3770	1	-	-

[a] Gly1 macrocyclisation with Val4 carbonyl
[b] Cys8 heterocyclisation with Asp7 carbonyl
[c] Asp7 C$_\alpha$-positon labeled with deuteron

Additionally, fragment a$_5$ (obs. 406.2796; calc. 406.2818) and b$_4$* (obs. 363.2012; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been observed previously for bottromycins and macrocyclic peptides, were detected.

Table 5-S5 | MS^2 fragmentation data for **4a**.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1	-	-	8	800.3830	800.3870
P	2	155.0818	155.0821	7	-	-
V	3	254.1496	254.1605	6	-	646.3133
V[a]	4	335.2072	335.2083	5	-	547.2449
V	5	434.2750	434.2767	4	-	466.1871
F	6	-	581.3451	3	367.0984	367.1186
D[b,c]	7	-	-	2	220.0489	220.0502
C[b]	8	-	782.3770	1	-	-

[a] Gly1 macrocyclisation with Val4 carbonyl
[b] Cys8 heterocyclisation with Asp7 carbonyl
[c] Asp7 C$_\alpha$-positon labeled with deuteron

Additionally, fragment a$_5$ (obs. 406.2789; calc. 406.2818) and b$_4$* (obs. 363.2011; calc. 363.2032), which are characteristically associated with the bottromycin macrocycle and have been observed previously for bottromycins and macrocyclic peptides, were detected.

Table 5-S6 | MS^2 fragmentation data for **5b**.

Seq.	b_n	Obs. b	Calc. b	y_n	Obs. y	Calc. y
G[a]	1	-	-	8	754.3847	754.3821
P	2	155.0834	155.0821	7	-	697.3606
V	3	254.1511	254.1605	6	-	600.3078
V[a]	4	335.2092	335.2083	5	501.2413	501.2394
V	5	434.2785	434.2767	4	420.1821	420.1816
F	6	581.3474	581.3451	3	321.1133	321.1132
D[b,c]	7	-	-	2	174.0456	174.0448
C[b]	8	-	-	1		

[a] Gly1 macrocyclisation with Val4 carbonyl
[b] Cys8 heterocyclisation with Asp7 carbonyl and decarboxylation
[c] Asp7 C$_\alpha$-position labeled with deuteron

Additionally, fragment a$_5$ (obs. 406.2832; calc. 406.2818) and b$_4$* (obs. 363.2041; calc. 363.2032), which is characteristically associated with the bottromycin macrocycle and has been previously observed for bottromycins and macrocyclic peptides, was detected.

Methods

<u>Cloning of the BotCYP, SalCYP and SalCYP$^{\text{T}}$ expression constructs</u>

Codon-optimized DNA for SalCYP expression in *E. coli* with a 5' sequence encoding for a TEV protease site (ENLYFQGG) and *NdeI* and *HindIII* restriction sites was purchased from Eurofins Genomics. After digestion with *NdeI* and *HindIII*, the DNA was ligated into pET28b+-vector treated with the same restriction enzymes with T4 DNA Ligase. The Thr14Asp mutation for SalCYP$^{\text{T}}$ was introduced by PCR using a primer carrying the same N-terminal mutation within its DNA sequence. The BotCYP construct was amplified by PCR using a cosmid containing the bottromycin biosynthetic gene cluster from *Streptomyces* sp. BC16019. The amplification products of both BotCYP and SalCYP$^{\text{T}}$ were extracted from an agarose gel, treated with *NdeI* and *HindIII* and cloned into a pET28b+-vector in the same way as the full-length construct. Clones were verified by DNA sequencing. Selected SalCYP plasmids were transformed into *E. coli* Lemo21(DE3) cells using a standard heat shock procedure and plated onto Luria-Bertani (LB) agar plates using a double selection of Kanamycin and Chloramphenicol at 50 and 34 µg / mL, respectively. pET28b+-BotCYP was transformed into *E. coli* C43(DE3) cells together with pGro7 (Chaperone plasmid set, Takara Bio Inc.) using a standard heat shock procedure. A selection for both plasmids was carried out by plating the cells onto a Luria-Bertani (LB) agar plate containing Kanamycin and Chloramphenicol at 50 and 34 µg / mL, respectively.

<u>Expression of BotCYP, SalCYP, SalCYP$^{\text{T242C}}$, SalCYP$^{\text{V243T}}$, SalCYP$^{\text{T242C,V243T}}$ and SalCYP$^{\text{T}}$</u>

SalCYP constructs were expressed in *E. coli* Lemo21(DE3), BotCYP was expressed in *E. coli* C43(DE3). Single colonies of transformants were used to inoculate starter cultures of Luria-Bertani(LB) medium with Kanamycin and Chloramphenicol at 50 and 34 µg / mL, respectively. Large scale expression of all constructs was carried out in modified Terrific Broth (TB) medium with the appropriate antibiotic concentrations added. Modified Terrific Broth was obtained by adding 1 mM MgSO$_4$ and NPS buffer[32] to a final concentration of 1x after autoclaving instead of the usual potassium phosphate buffer system. The cultures were inoculated 1 : 100 with starter culture and grown at 37 °C and 180 rpm until an OD$_{600}$ of 0.8 was reached, at which point the temperature was reduced to 25 °C and 0.5 mM 5-aminolevulinic acid as well as 0.25 mM FeCl$_3$ were added. For BotCYP, expression of the GroEL/GroES chaperone system was additionally induced by the addition of 4 mg / mL *L*-Arabinose to the culture. After growing the cultures for another 30 min at 25 °C, 0.25 mM IPTG was added and the flasks were

incubated for another 40 h. After 40 h, the cells were pelleted by centrifugation for 10 min at 5000 rpm and 4 °C and cell pellets were stored at - 80 °C until further use.

Purification of BotCYP, SalCYP and SalCYP^T

The protein purification of all constructs was carried out in the same lysis buffer (500 mM NaCl, 20 mM Bis-Tris pH 6.8, 20 mM Imidazole, 3 mM β-mercaptoethanol). For every 25 g of wet cell pellet, 100 mL of lysis buffer was added. The lysis buffer was supplemented with cOmplete EDTA-free protease inhibitor tablets (Roche) and DNAse (0.4 mg / g wet cell pellet, Sigma). The cell suspension was lysed via passage through a cell disruptor (30 k psi, Constant Systems) and the cell debris was removed by centrifugation (19,000 rpm, 4 °C, 15 min). The supernatant was filtered through a 0.45 μm filter and applied to a His-Trap HP 5 mL column pre-equilibrated in lysis buffer at a flow rate of 5 mL / min. After application of the lysate and an extensive column wash (30 CV lysis buffer), 3 mg of purified TEV-protease, diluted in 5 mL of lysis buffer, was loaded onto the column. TEV cleavage was carried out on the column for 2 h at room temperature. Cleaved CYP proteins were eluted off the column by washing the column with another 20 mL of lysis buffer while collecting 2 mL elution fractions. Eluted fractions with the most color were pooled and loaded onto a Superdex 200 16/600 pre-equilibrated in the respective gel filtration buffers (SalCYP, BotCYP and SalCYP point mutants: 150 mM NaCl, 10 mM HEPES pH 7.4, 0.5 mM TCEP; SalCYP^T: 10 mM HEPES pH 7.4). Size exclusion chromatography was carried out at 1 mL / min and the fractions with the most color were again pooled and concentrated using a 30 kDa cutoff filter (Amicon). The protein concentration was determined by spectrophotometry at A$_{280}$ using the extinction coefficients calculated from the primary amino acid sequence. SalCYP was diluted to 130 μM whereas SalCYP^T was diluted to 450 μM in their respective gel filtration buffers for crystallography. Protein purity and integrity was confirmed by SDS-PAGE as well as intact protein mass spectrometry (data not shown).

X-ray crystallography

Initial screening experiments for both proteins were carried out at their aforementioned concentrations in sitting-drop SwissSCI plates with commercially available crystallization screens (Qiagen) using a Gryphon crystallization robot. A needle cluster for SalCYP was observed after 3 days at 277 K in a well solution of 20% PEG 3350 and 200 mM calcium acetate. Initial crystals were collected, crushed and used as seeds at different concentrations for crystal optimization. Single crystals were observed with an added seed concentration of 0.75% and cryoprotected in mother liquor supplemented with a final concentration of 32% glycerol

prior to freezing in a cryo loop. A high redundancy dataset for Fe-SAD was collected at the ESRF Beamline ID-29 at 100 K and a wavelength of 1.738 Å. Data were processed using XDS[33] and POINTLESS[34], AIMLESS[35] and ctruncate[36] implemented in ccp4[37]. The structure was solved using Phenix.AutoSol,[38] followed by several rounds of manual rebuilding in COOT[39] and refinement in Phenix.refine[38]. Crystals of SalCYP[T] appeared after 5 days at 277 K in a well solution containing 20% PEG 8000, 16.6% glycerol, 200 mM calcium acetate and 80 mM sodium cacodylate pH 6.5. As for SalCYP, microseeding improved the quality as well as the nucleation probability of SalCYP[T] and was used in the same fashion to obtain single crystals. Single crystals were observed with an added seed concentration of 1% and cryoprotected in mother liquor supplemented with a final concentration of 32% glycerol prior to freezing in a cryo loop. A dataset of SalCYP[T] was collected at ESRF Beamline MASSIF-3 at 100 K and a wavelength of 0.961 Å. Data were processed as for SalCYP and molecular replacement in Phaser[40] was carried out using SalCYP as a search model, followed by several rounds of manual rebuilding in COOT[39] and refinement in Phenix.refine[38]. The structures were validated using *MolProbity[41]*, and all images presented were created using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.). Interaction diagrams were created using LigPlot[+31].

In vitro production of **2a/b**

BotA[P], IpoC, PurCD and PurAH were expressed and purified as described previously.[8, 10] Large scale IpoC reactions were carried out on a 9 mL scale in GF buffer (150 mM NaCl, 10 mM HEPES, 0.5 mM TCEP, pH 7.4) using the reaction setup 50 μM BotA[P], 5 μM IpoC, 5 mM ATP as well as 5 mM $MgCl_2$ and incubating the reaction mixture for 16 h at 37 °C. The reaction mixture was filtered through a 0.22 μm filter and loaded onto a Superdex 30 16/60 size exclusion chromatography column pre-equilibrated in GF buffer. Heterocyclized-peptide-containing fractions were pooled, analyzed by mass spectrometry and concentrated using a 5 kDa cutoff filter. The coupled large scale reactions of PurCD and PurAH were carried out on a 9 mL scale by incubating 50 μM BotA[PC], 5 μM PurCD, 0.25 μM PurAH, 10 mM ATP, 10 mM $MgCl_2$ and 100 μM $CoCl_2$ for 12 h at 37 °C in PurCD rxn buffer (200 mM NaCl, 50 mM CHES, 10% glycerol, pH 9.5). The reaction mixture was filtered through a 0.22 μm filter and applied to a Superdex 30 16/60 size exclusion chromatography column pre-equilibrated in ultrapure water. Every chromatography fraction was analyzed by mass spectrometry and the pure **2a/b**-containing fractions were pooled and lyophilized. Powder of **2a/b** was dissolved in ultrapure water and pipetted into a pre-weighed HPLC tube. All solvent was removed by vacuum

centrifugation and remaining powder of **2a/b** was weighed using a precision scale. A concentration of 10 mM for **2a/b** was adjusted in ultrapure water, subsequently analyzed by mass spectrometry and kept on ice until further use.

Redox partner purification

BmCPR and Fdx2 were purified as reported previously.[42, 43] The preparation of the fission yeast proteins Arh1 and Etp1$^{\text{fd}}$ was carried out as described before.[44, 45] Recombinant bovine AdR and the Adx4-108 (truncated form of Adx comprising amino acids 4–108) were purified as mentioned elsewhere.[46, 47]

The concentration of recombinant P450 was estimated using the CO-difference spectral assay as described previously with $\varepsilon_{450-490} = 91$ mM^{-1} cm^{-1}.[48] The concentration of BmCPR was quantified by measuring the flavin absorbance at 456 nm with $\varepsilon_{456} = 21$ mM^{-1} cm^{-1} for the oxidized enzyme.[43] The concentrations of the AdR and Arh1 were measured using the extinction coefficient $\varepsilon_{450} = 11.3$ mM^{-1} cm^{-1}.[45, 49] The concentrations of Fdx2 and Etp1$^{\text{fd}}$ were measured using the extinction coefficient $\varepsilon_{390} = 6.671$ mM^{-1} cm^{-1} and $\varepsilon_{414} = 9.8$ mM^{-1} cm^{-1}, respectively.[42, 50]

Investigation of electron transfer partners

The functional interaction of the electron transfer partners for a particular P450 was determined by observing the NADPH reduced CO-complex peak at 450 nm when P450 was coupled with the different redox partners in the absence of substrate. For this, SalCYP or BotCYP was mixed with ferredoxins and ferredoxin reductases with ratios of SalCYP : Fdx2/Etp1$^{\text{fd}}$/Adx : BmCPR/Arh1/AdR of 1:20:3 µM (and 1:40:4 for BmCPR-Fdx2) in 50 mM HEPES buffer pH 7.4 and NADPH was added to a final concentration of 1 mM. The spectra of NADPH-reduced samples were recorded after bubbling the sample with carbon monoxide (CO) gas.

LC-MS and MS^2 analysis

Measurements were performed on a Dionex Ultimate 3000 RSLC system (Thermo Fisher Scientific) using a BEH C18, 50 x 2.1 mm, 1.7 µm dp column equipped with a C18 precolumn (Waters). Samples were separated by a gradient from (A) H$_2$O + 0.1% formic acid to (B) ACN + 0.1% formic acid at a flow rate of 600 µL min^{-1} and 45 °C. The linear gradient was initiated by a 1 min isocratic step at 5% B, followed by an increase to 95% B in 9 min to end up with a 1.5 min plateau step at 95% ACN before re-equilibration under the initial conditions. UV spectra were recorded by a DAD in the range from 200 to 600 nm.

For MS measurements on maXis-4 hr-qToF mass spectrometer (Bruker Daltonics), the LC flow was split 1:8 before entering the mass spectrometer using the Apollo II ESI source. In the source region, the temperature was set to 200 °C, the capillary voltage was 4000 V, the dry-gas flow was 5.0 L / min and the nebulizer was set to 14.5 psi. Ion transfer setting were set to Funnel 1 RF 350 Vpp and Multipole RF 400 Vpp, quadrupole settings were set to an ion energy of 5.0 eV and a low mass cut of 120 m/z. The collision cell was set to an energy of 5.0 eV and the pulse storage time was 5 µs. Data were recorded in the mass range from 150 to 2500 m/z. Calibration of the maXis4G qTOF spectrometer was achieved with sodium formate clusters before every injection to avoid mass drifts. All MS analyses were acquired in the presence of the lock masses $C_{12}H_{19}F_{12}N_3O_6P_3$, $C_{18}H_{19}O_6N_3P_3F_2$ and $C_{24}H_{19}F_{36}N_3O_6P_3$ which generate the $[M+H]^+$ Ions of 622.028960, 922.009798 and 1221.990638. LC-MS^2 fragmentation was performed using an automatic precursor selection mode.

Signals in the MS-spectra are labelled with the observed monoisotopic mass. Extracted-ion chromatograms were generated with the Bruker Compass DataAnalysis software (Version 4.2) using the calculated monoisotopic mass with a range of 5 ppm. EICs shown in one figure always have the same time and intensity range, except for Figure 5-S12. MS-spectra in one figure are scaled to a relative intensity of 1.

In vitro reactions of BotCYP and SalCYP

To investigate the role of the P450 enzymes *in vitro*, 20 µM CYP substrate analogue **2a/b** was incubated with 5 µM BotCYP or SalCYP, respectively, 5 µM BmCPR, 50 µM Fdx2 and 2.5 mM NADPH at 30 °C overnight in reaction buffer (150 mM NaCl, 10 mM HEPES, pH 7.4). Reactions without the P450 enzyme were used as controls. Reactions were stopped and enzymes were precipitated by the addition of two volumes of ACN. The samples were frozen at -80 °C and precipitated enzymes were pelleted by centrifugation at 15.000 x g for 20 min. The turnover was analyzed by LC-MS measurements. Reactions were set up and analyzed in triplicates.

To investigate the effect of the addition of BotH to a CYP reaction, reactions were set up as described above with and without the addition of 1 µM BotH. Reactions were stopped after 120 min incubation at 30 °C.

To test if the absolute configuration at the Asp Cα-position was retained during the oxidative decarboxylation (Figure 5-6B), **4a/b** was produced by incubating 30 µM **2a/b** with 6 µM BotH in D$_2$O reaction buffer at 30 °C for 30 min. Afterwards, BotH was denatured by incubating the solution at 98 °C for 10 min. Denatured BotH was removed by centrifugation (15.000 x g, 20

min, 4 °C). D_2O was removed by lyophilization. The dried compound was then dissolved in H_2O and reactions were performed as described above. In parallel, the same reaction was performed in H_2O.

To investigate the substrate preference of BotCYP (Figure 5-4A), reactions were set up as described above, but were stopped after 5, 15, 30 and 60 min by the addition of ACN. For this experiment, a BotH treated **2a/b** solution was used. Upon treatment with BotH, an approximate ratio of **2a/b** of 1:1 was achieved. BotH was denatured and removed from the solution before the experiments, as described above.

Bioinformatics analysis

To study taxonomic distribution of BotCYP-like enzymes and their association with biosynthetic gene clusters, we first selected all bacterial member of the Pfam[51] family PF00067 (Cytochrome P450), to which BotCYP belongs. Then we predicted all biosynthetic gene clusters in the corresponding genomes with antiSMASH[52] and analyzed their genomic coordinates with respect to the location of cytochrome P450 proteins and calculated their sequence identity matrix with custom scripts. The sequence similarity network was visualized using Cytoscape[53]. The distance matrix between selected proteins was calculated with the R package seqinr[54], and the sequence similarity network was visualized using Cytoscape[53]. All edges representing sequence identity less than 50% and the corresponding singleton nodes were removed for clarity.

5.4 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.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* **2013,** *30* (1), 108-60.

2.   Ortega, M. A.; van der Donk, W. A., New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem Biol* **2016,** *23* (1), 31-44.

3.   Montalban-Lopez, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martinez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H. G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Sussmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J. K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A., New developments in RiPP discovery, enzymology and engineering. *Nat Prod Rep* **2020**.

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., Structure and biosynthesis of the antibiotic bottromycin D. *Org Lett* **2012,** *14* (19), 5050-3.

5.   Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Omura, S., Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew Chem Int Ed Engl* **2009,** *48* (5), 914-7.

6.   Crone, W. J.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angew Chem Int Ed Engl* **2016,** *55* (33), 9639-43.

7.   Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Muller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *Chembiochem* **2016,** *17* (23), 2286-2292.

8.   Franz, L.; Adam, S.; Santos-Aberturas, J.; Truman, A. W.; Koehnke, J., Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme. *J Am Chem Soc* **2017,** *139* (50), 18158-18161.

9. Schwalen, C. J.; Hudson, G. A.; Kosol, S.; Mahanta, N.; Challis, G. L.; Mitchell, D. A., In Vitro Biosynthetic Studies of Bottromycin Expand the Enzymatic Capabilities of the YcaO Superfamily. *J Am Chem Soc* **2017,** *139* (50), 18154-18157.

10. Sikandar, A.; Franz, L.; Melse, O.; Antes, I.; Koehnke, J., Thiazoline-Specific Amidohydrolase PurAH Is the Gatekeeper of Bottromycin Biosynthesis. *J Am Chem Soc* **2019,** *141* (25), 9748-9752.

11. Sikandar, A.; Franz, L.; Adam, S.; Santos-Aberturas, J.; Horbal, L.; Luzhetskyy, A.; Truman, A. W.; Kalinina, O. V.; Koehnke, J., The bottromycin epimerase BotH defines a group of atypical alpha/beta-hydrolase-fold enzymes. *Nat Chem Biol* **2020,** *16* (9), 1013-1018.

12. Gouda, H.; Kobayashi, Y.; Yamada, T.; Ideguchi, T.; Sugawara, A.; Hirose, T.; Omura, S.; Sunazuka, T.; Hirono, S., Three-dimensional solution structure of bottromycin A2: a potent antibiotic active against methicillin-resistant Staphylococcus aureus and vancomycin-resistant Enterococci. *Chem Pharm Bull (Tokyo)* **2012,** *60* (2), 169-71.

13. Gao, S.; Ge, Y.; Bent, A. F.; Schwarz-Linek, U.; Naismith, J. H., Oxidation of the Cyanobactin Precursor Peptide Is Independent of the Leader Peptide and Operates in a Defined Order. *Biochemistry* **2018,** *57* (41), 5996-6002.

14. Houssen, W. E.; Jaspars, M., Azole-based cyclic peptides from the sea squirt Lissoclinum patella: old scaffolds, new avenues. *Chembiochem* **2010,** *11* (13), 1803-15.

15. Bent, A. F.; Mann, G.; Houssen, W. E.; Mykhaylyk, V.; Duman, R.; Thomas, L.; Jaspars, M.; Wagner, A.; Naismith, J. H., Structure of the cyanobactin oxidase ThcOx from Cyanothece sp. PCC 7425, the first structure to be solved at Diamond Light Source beamline I23 by means of S-SAD. *Acta Crystallogr D Struct Biol* **2016,** *72* (Pt 11), 1174-1180.

16. Li, Y. M.; Milne, J. C.; Madison, L. L.; Kolter, R.; Walsh, C. T., From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* **1996,** *274* (5290), 1188-93.

17. Melby, J. O.; Li, X.; Mitchell, D. A., Orchestration of enzymatic processing by thiazole/oxazole-modified microcin dehydrogenases. *Biochemistry* **2014,** *53* (2), 413-22.

18. Guengerich, F. P.; Martin, M. V.; Sohl, C. D.; Cheng, Q., Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. *Nat Protoc* **2009,** *4* (9), 1245-51.

19. Milhim, M.; Gerber, A.; Neunzig, J.; Hannemann, F.; Bernhardt, R., A Novel NADPH-dependent flavoprotein reductase from Bacillus megaterium acts as an efficient cytochrome P450 reductase. *J Biotechnol* **2016,** *231*, 83-94.

20. Holm, L.; Laakso, L. M., Dali server update. *Nucleic Acids Res* **2016,** *44* (W1), W351-5.

21. Bell, S. G.; Yang, W.; Dale, A.; Zhou, W.; Wong, L. L., Improving the affinity and activity of CYP101D2 for hydrophobic substrates. *Appl Microbiol Biotechnol* **2013,** *97* (9), 3979-90.

22. Wang, J. B.; Lonsdale, R.; Reetz, M. T., Exploring substrate scope and stereoselectivity of P450 peroxygenase OleTJE in olefin-forming oxidative decarboxylation. *Chem Commun (Camb)* **2016,** *52* (52), 8131-3.

23. Gober, J. G.; Ghodge, S. V.; Bogart, J. W.; Wever, W. J.; Watkins, R. R.; Brustad, E. M.; Bowers, A. A., P450-Mediated Non-natural Cyclopropanation of Dehydroalanine-Containing Thiopeptides. *ACS chemical biology* **2017,** *12* (7), 1726-1731.

24. Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A., In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *Journal of the American Chemical Society* **2015,** *137* (51), 16012-16015.

25. Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A., Biosynthetic Timing and Substrate Specificity for the Thiopeptide Thiomuracin. *Journal of the American Chemical Society* **2016,** *138* (48), 15511-15514.

26. Zheng, Q.; Wang, S.; Liao, R.; Liu, W., Precursor-Directed Mutational Biosynthesis Facilitates the Functional Assignment of Two Cytochromes P450 in Thiostrepton Biosynthesis. *ACS Chem Biol* **2016,** *11* (10), 2673-2678.

27. Bai, X.; Guo, H.; Chen, D.; Yang, Q.; Tao, J.; Liu, W., Isolation and structure determination of two new nosiheptide-type compounds provide insights into the function of the cytochrome P450 oxygenase NocV in nocathiacin biosynthesis. *Organic Chemistry Frontiers* **2020,** *7* (3), 584-589.

28. Young, T. S.; Walsh, C. T., Identification of the thiazolyl peptide GE37468 gene cluster from *Streptomyces* ATCC 55365 and heterologous expression in *Streptomyces lividans*. *Proceedings of the National Academy of Sciences* **2011,** *108* (32), 13053-13058.

29. Yamada, T.; Yagita, M.; Kobayashi, Y.; Sennari, G.; Shimamura, H.; Matsui, H.; Horimatsu, Y.; Hanaki, H.; Hirose, T.; S, O. M.; Sunazuka, T., Synthesis and Evaluation of Antibacterial Activity of Bottromycins. *J Org Chem* **2018,** *83* (13), 7135-7149.

30. Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R., The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research* **2019,** *47* (W1), W636-W641.

31. Laskowski, R. A.; Swindells, M. B., LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *Journal of Chemical Information and Modeling* **2011,** *51* (10), 2778-2786.

32. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **2005,** *41* (1), 207-234.

33. Kabsch, W., XDS. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 125-32.

34. Evans, P. R., An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **2011,** *67* (Pt 4), 282-92.

35. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **2013,** *69* (Pt 7), 1204-14.

36. French, S.; Wilson, K., On the treatment of negative intensity observations. *Acta Crystallographica Section A* **1978,** *34* (4), 517-525.

37. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **2011,** *67* (Pt 4), 235-242.

38. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.

39. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallographica Section D* **2010,** *66* (4), 486-501.

40. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of Applied Crystallography* **2007,** *40* (4), 658-674.

41. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 1), 12-21.

42. Brill, E.; Hannemann, F.; Zapp, J.; Brüning, G.; Jauch, J.; Bernhardt, R., A new cytochrome P450 system from Bacillus megaterium DSM319 for the hydroxylation of 11-keto-β-boswellic acid (KBA). *Appl Microbiol Biotechnol* **2014,** *98* (4), 1701-17.

43. Milhim, M.; Gerber, A.; Neunzig, J.; Hannemann, F.; Bernhardt, R., A Novel NADPH-dependent flavoprotein reductase from Bacillus megaterium acts as an efficient cytochrome P450 reductase. *J Biotechnol* **2016,** *231*, 83-94.

44. Bureik, M.; Schiffler, B.; Hiraoka, Y.; Vogel, F.; Bernhardt, R., Functional expression of human mitochondrial CYP11B2 in fission yeast and identification of a new internal electron transfer protein, etp1. *Biochemistry* **2002,** *41* (7), 2311-21.

45. Ewen, K. M.; Schiffler, B.; Uhlmann-Schiffler, H.; Bernhardt, R.; Hannemann, F., The endogenous adrenodoxin reductase-like flavoprotein arh1 supports heterologous cytochrome P450-dependent substrate conversions in Schizosaccharomyces pombe. *FEMS Yeast Res* **2008,** *8* (3), 432-41.

46. Uhlmann, H.; Beckert, V.; Schwarz, D.; Bernhardt, R., Expression of bovine adrenodoxin in E. coli and site-directed mutagenesis of /2 Fe-2S/ cluster ligands. *Biochem Biophys Res Commun* **1992,** *188* (3), 1131-8.

47.   Sagara, Y.; Wada, A.; Takata, Y.; Waterman, M. R.; Sekimizu, K.; Horiuchi, T., Direct expression of adrenodoxin reductase in Escherichia coli and the functional characterization. *Biol Pharm Bull* **1993,** *16* (7), 627-30.

48.   Omura, T.; Sato, R., The Carbon Monoxide-binding Pigment of Liver Microsomes: II. Solubilization, Purification, and Properties. *J Biol Chem* **1964,** *239*, 2379-85.

49.   Hiwatashi, A.; Ichikawa, Y.; Maruya, N.; Yamano, T.; Aki, K., Properties of crystalline reduced nicotinamide adenine dinucleotide phosphate-adrenodoxin reductase from bovine adrenocortical mitochonria. I. Physicochemical properties of holo- and apo-NADPH-adrenodoxin reductase and interaction between non-heme iron proteins and the reductase. *Biochemistry* **1976,** *15* (14), 3082-90.

50.   Schiffler, B.; Bureik, M.; Reinle, W.; Müller, E.-C.; Hannemann, F.; Bernhardt, R., The adrenodoxin-like ferredoxin of Schizosaccharomyces pombe mitochondria. *J Inorg Biochem* **2004,** *98* (7), 1229-1237.

51.   El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D., The Pfam protein families database in 2019. *Nucleic acids research* **2019,** *47* (D1), D427-D432.

52.   Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H., antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic acids research* **2017,** *45* (W1), W36-W41.

53.   Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003,** *13* (11), 2498-2504.

54.   Charif, D.; Lobry, J. R., SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. In *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*, Bastolla, U.; Porto, M.; Roman, H. E.; Vendruscolo, M., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2007; pp 207-232.

Chapter 6

Leader peptide exchange to produce hybrid, new-to-nature ribosomal natural products

Laura Franz, and Jesko Koehnke

Manuscript in revision

6 Leader peptide exchange to produce hybrid, new-to-nature ribosomal natural products

6.1 Abstract

Ribosomal natural products contain exquisite post-translational peptide modifications that are installed by a range of pathway-specific enzymes. We present proof of principle for a sortase A-based approach that enables peptide modification by enzymes from unrelated pathways. This approach allowed the one-pot synthesis of a new-to-nature, hybrid ribosomal natural product.
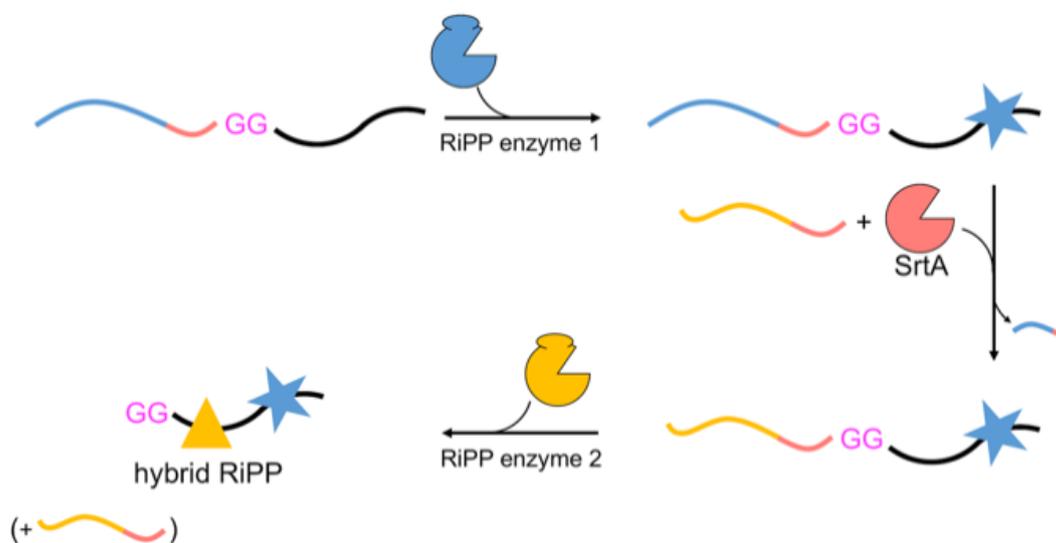
Figure 6-1 | Graphical abstract.

6.2 Main Text

Natural products and their derivatives have been an invaluable, rich source for drugs.[1, 2] In recent years, the fast-expanding superfamily of ribosomally synthesized and post-translationally modified peptide (RiPP) natural products has been recognised as a major source for novel compounds with potent biological (e.g. antibiotics) activities.[3, 4] RiPP biosynthesis begins with the production of a short, ribosomally synthesized precursor peptide (PP), usually comprised of a N-terminal leader peptide (LP) and a C-terminal core peptide (CP) (Figure 6-2).[3, 4] While the LP is essential for binding (and activation) of biosynthetic enzymes and typically discarded during biosynthesis, the CP is transformed into the mature RiPP. Primary RiPP enzymes require specific parts of the LP, or recognition sequences (RSs), for activity and introduce characteristic, RiPP class-defining post-translational modifications (PTMs) in the CP. The spatial separation of substrate recognition (LP) and the site of catalysis (CP) permits primary RiPP enzymes to be very promiscuous with respect to the CP sequence.[3-11] Secondary RiPP enzymes do not depend on the LP and are responsible for specific tailoring steps.
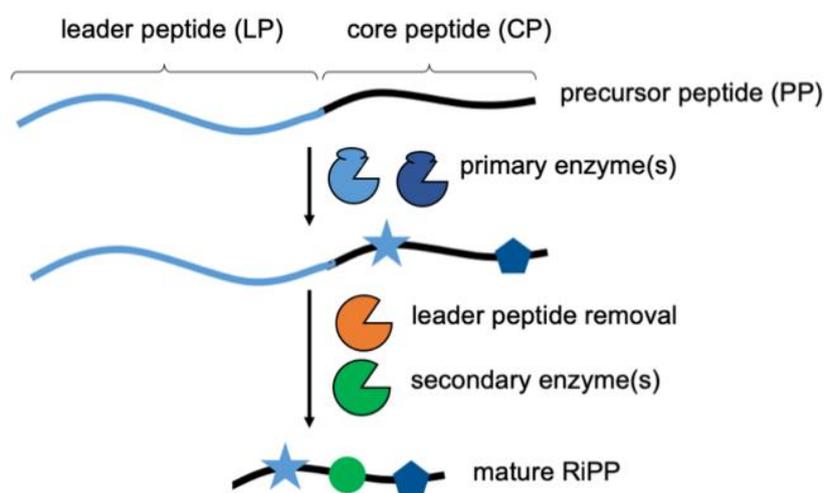
Figure 6-2 | Overview of RiPP biosynthesis.

The PTMs introduced by primary enzymes are incredibly diverse and expand the chemical space accessible to RiPPs far beyond the 20 canonical amino acids that comprise the PP (ribosomal expression).[3, 4] They include, but are not limited to, azol(in)e heterocycles, lanthionine-, sactionine-, C-C and ester-crosslinks, thioamidation, and N-methylation of the CP backbone.[3, 4] These PTMs endow RiPPs with several desirable properties, including potent bioactivities. The ability to combine the modifications from different pathways could facilitate the generation of RiPP-inspired natural product

libraries for drug discovery. It would require the attachment of different LPs to a given CP.

To accomplish this feat the initial focus had been on establishing conditions for LP-independent peptide processing. This was attempted by primary RiPP enzyme engineering, or to supply the LP *in trans*.[12-17] Recently the improved understanding of RiPP enzymology was used in an inspired approach to engineer new to nature, hybrid RiPPs.[18, 19] By packaging two RSs from different RiPP families into a single, chimeric LP it was possible to generate unprecedented molecules (Figure 6-S1). Thus far chimeric LPs have been limited to two RSs and the approach requires a thorough understanding of the substrate recognition requirements of each enzyme. We sought a less involved method, that would allow us to simply swap LPs between biosynthetic steps. This approach would allow iterative processing and not require a detailed understanding of the substrate recognition sequences required by the enzymes used, making it potentially plug-and-play. Our leader peptide exchange (LPX) technique is based on sortase A (SrtA)-mediated transpeptidation (Figure 6-3).
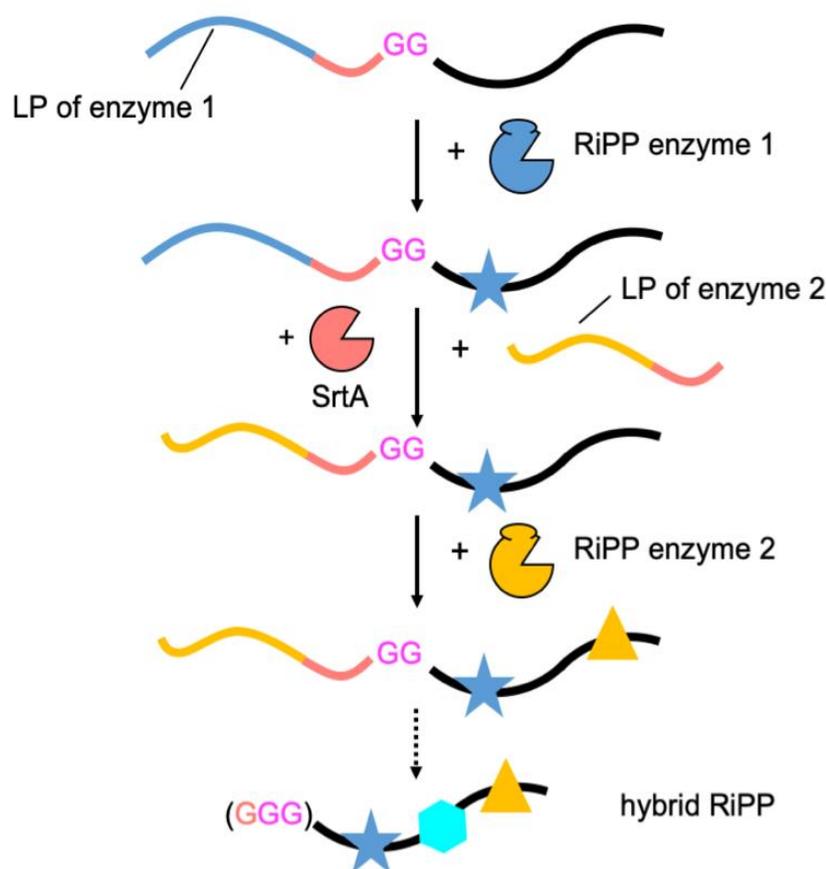
Figure 6-3 | The leader peptide exchange (LPX) strategy enables Sortase A (SrtA, red)-mediated LPX and thus iterative core peptide processing by biosynthetic enzymes of different RiPP classes. A final step using SrtA or another protease liberates the hybrid RiPP. SrtA recognition sequence is shown in red, di-Gly linker in magenta. Symbols (e.g. stars) represent PTMs.

To demonstrate the utility and feasibility of our idea we decided to generate a new-to-nature RiPP by combining enzymes from the unrelated cyanobactin and microviridin pathways (Figure 6-S2 and Figure 6-S3). The enzyme LynD is found in a cyanobactin pathway discovered in *Lyngbya* sp. and converts cysteine residues into thiazoline heterocycles during aesturamide biosynthesis (Figure 6-S2).[13, 20] We chose to combine this heterocyclase with MdnC of the microviridin J (derived from PP MdnA) pathway found in *Microcystis aeruginosa*.[21, 22] This enzyme catalyses the formation of two ω-ester bonds from Ser/Thr and Asp/Glu side chains, which leads to the formation of macrocycles (Figure 6-S3). An engineered variant of SrtA, SrtA$_{7m}$,[23] was chosen to facilitate LPX. We designed two synthetic peptides: Peptide **1** was comprised of a truncated cyanobactin LP suitable for LynD, followed by the SrtA recognition motif (LPXTG), a di-glycine and the MdnA core peptide harbouring two point mutations (Ser2 and Pro7 to Cys) (Figure 6-S4). Peptide **2** consisted of a truncated MdnA LP containing the MdnC RS and a C-terminal SrtA recognition motif (Figure 6-S4). To probe SrtA-mediated LPX, we incubated peptide **1** (All calculated and observed masses including mass errors can be found in Table 6-S1) and peptide **2** with SrtA$_{7m}$ (2 h, 37 °C) and analysed the reaction products by liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS). We readily detected the desired reaction product, peptide **3** (Figure 6-S5). Since the SrtA$_{7m}$ reaction is an equilibrium, we also detected the other expected masses (Figure 6-S5). Several techniques have been developed to pull the SrtA$_{7m}$ equilibrium into the desired direction and may be employed to optimize this system in the future.[24, 25]

With LPX successfully established, we next tested the tolerance of LynD and MdnC for the presence of the SrtA recognition sequence that we had introduced. Incubation of peptide **1** with LynD in the presence of ATP/Mg^{2+} led to an observed loss of 36 Da, which is in agreement with the desired two cyclodehydration reactions required to convert the two CP Cys residues into thiazolines (Figure 6-S6). Treatment of this reaction product with iodoacetamide, which alkylates free Cys residues, did not result in a mass shift. This strongly implies the conversion of the two Cys residues to thiazolines in the LynD-treated peptide (**4**) (Figure 6-S6). Incubation of peptide **3** with MdnC in the presence of ATP/Mg^{2+} also resulted in the expected loss of 36 Da. This is in agreement with the formation of two ω-ester bonds in peptide **3** (Figure 6-S7). These data confirm the compatibility of both enzymes with the SrtA recognition sequence and highlights the malleability of RiPP systems.

With all components of our model system tested, we combined them in one pot. After incubation of peptide **1** with LynD and ATP/Mg^{2+}, SrtA$_{7m}$ was added and we rapidly observed the liberation of the core peptide with two thiazolines (**6**) (Figure 6-4 III.). After addition of peptide **2**, the SrtA-mediated ligation product with **6** was observed (**7**, Figure 6-4 IV.). Subsequent addition of MdnC led to another loss of 36 Da, indicating that MdnC tolerates the two non-natural thiazolines in the core peptide and introduces two ω-esters into the peptide (**8**, Figure 6-4 V.). Since SrtA$_{7m}$ was not removed, the MdnA LP was cleaved by SrtA$_{7m}$ and we obtained a new to nature, hybrid RiPP **9** (Figure 6-4 V.).
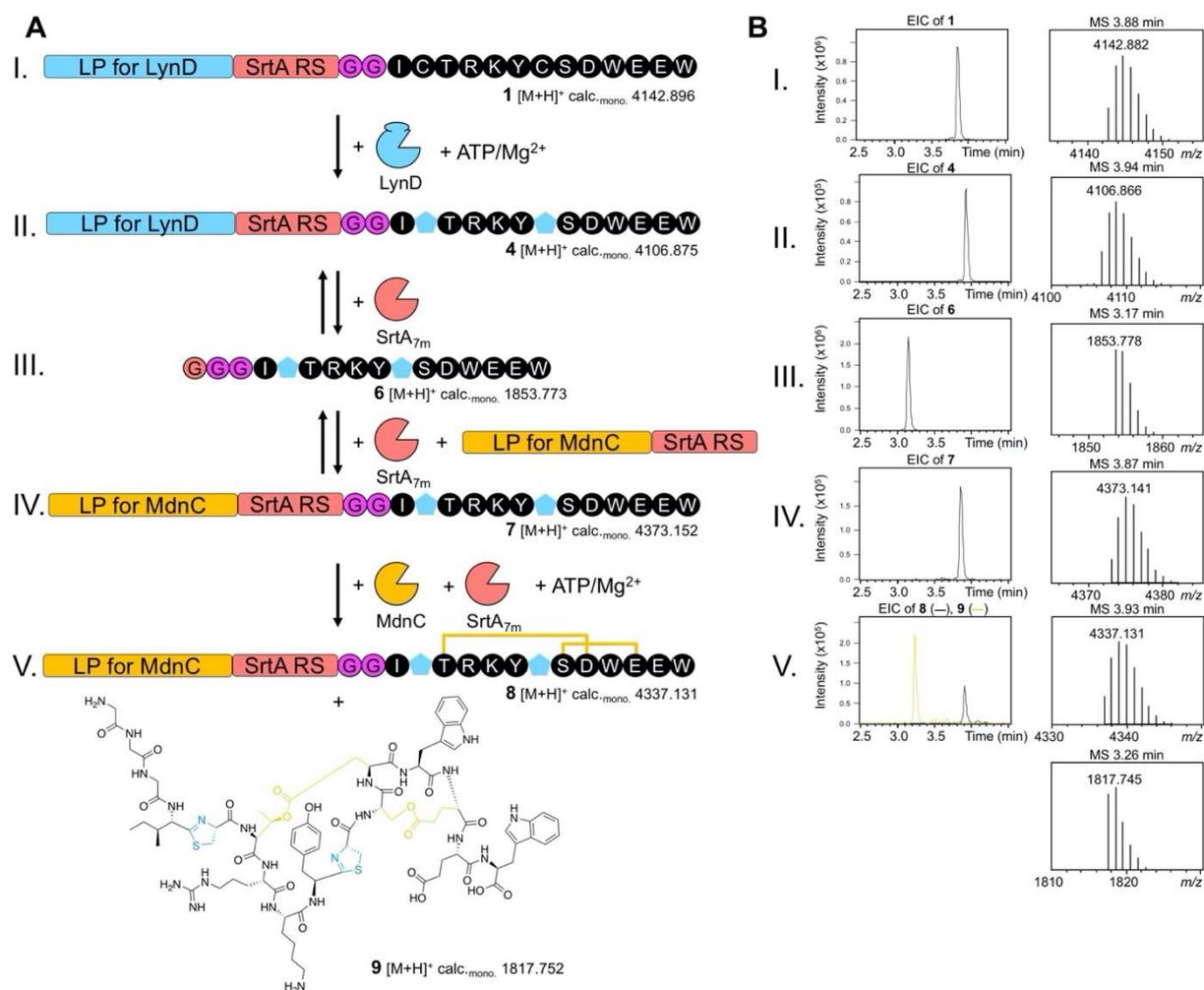
Figure 6-4 | Combination of the cyanobactin heterocyclase LynD and the microviridn J ATP-grasp ligase MdnC using the LPX strategy to produce a heterocycle-containing graspetide (**9**). **A** Schematic of the LPX model system. After introduction of the thiazolines (blue pentagons) in **1** by LynD, the LynD LP is exchanged for the MdnA LP (**2**) using SrtA-mediated transpeptidation yielding **7**. LPX allows the modification of the core peptide (black) by MdnC that introduces two ω-ester cross-links (**8**). SrtA finally cleaves the MdnA LP yielding the new-to-nature hybrid RiPP **9**. **B** Extracted ion chromatograms (EICs, ±5 ppm) and mass spectra for the desired products of each step shown in A. Shown are the masses of the singly-charged monoisotopic ions [M+H]$^+$. An extensive presentation of all products and secondary products (SrtA reactions are reversible) in all steps (I.-V.) is given in the Figure 6-S8.

Reactions without the addition of SrtA stalled after the introduction of the two thiazolines and did not result in ω-ester formation, or the production of **9**. To the best of our knowledge **9** represents the first heterocycle-containing graspetide, the newly coined name for microviridins and related compounds.[4]

We have demonstrated that the SrtA-based LPX technique is compatible with RiPP enzymes. Since our approach dispenses with the need for a detailed understanding of enzyme-substrate recognition for a particular enzyme / pathway, it has the potential to greatly expedite the creation of hybrid RiPPs. It is going to be important to consider the combination of enzymes used when designing these experiments. In the case presented here we used LynD before MdnC, because it was reasonable to assume that the topology introduced by MdnC would preclude processing of the cysteine residues by LynD. In addition, we selected LynD because it is specific to cysteine residues and leaves Ser/Thr available for MdnC. The use of other heterocyclases that are able to convert Ser/Thr residues to oxazolines would preclude MdnC from forming ω-esters. Depending on the timing of addition this could also lead to competition for reaction sites that would reduce yields and complicate analysis. It could of course also be viewed as an opportunity to generate a set of products from a single reaction, which may yield unexpected, interesting candidates for a given activity screen. The protection of side-chains may offer a solution to this issue, but would then require the purification of intermediates prior to addition of the next RiPP enzyme. In this context it is going to be interesting to trial this approach with RiPPs that depend on a particular sequence N-terminal of the core peptide for proper processing, such as thiopeptides. Here it may be necessary to include the native protease site as a spacer between the SrtA recognition sequence and the core peptide. It may be of particular interest to use the LPX approach in combination with self-sacrificing RiPP enzymes, such as the backbone N-methylating enzyme OphA, which cannot be combined with other leader-dependent RiPP enzymes through hybrid leader peptides.

We focussed our efforts on proof-of-concept and thus did not attempt to maximise reaction yields. The two main aspects that would improve yields are 1. the use of optimized SrtA substrates, which have been specifically developed to drive the ligation reaction in the desired reaction and 2. the inactivation or removal of SrtA prior to addition of the second enzyme (compare Figure 6-S8). A large variety of tools to accomplish the former task task have been developed and established for SrtA and will

be used in future studies. The latter point arose because the active $SrtA_{7m}$ in the reaction led to a competition with MdnC.

Truncated leader peptides were used here to reduce synthesis costs, but full-length leader peptides are of course compatible. Intriguingly, the peptides used as starting material in this study can of course be produced ribosomally, for example through expression in *Escherichia coli*. As a consequence the entire system could be moved into an *in vivo* setting. While the yields of *in vivo* ligation reactions using SrtA are not ideal, it would allow the leverage of genetics to create large libraries based on core peptide randomization. In targeted bioactivity screens the *in vivo* efficiency of SrtA may well prove sufficient to discover novel, exciting molecules.

6.3 Supporting Information
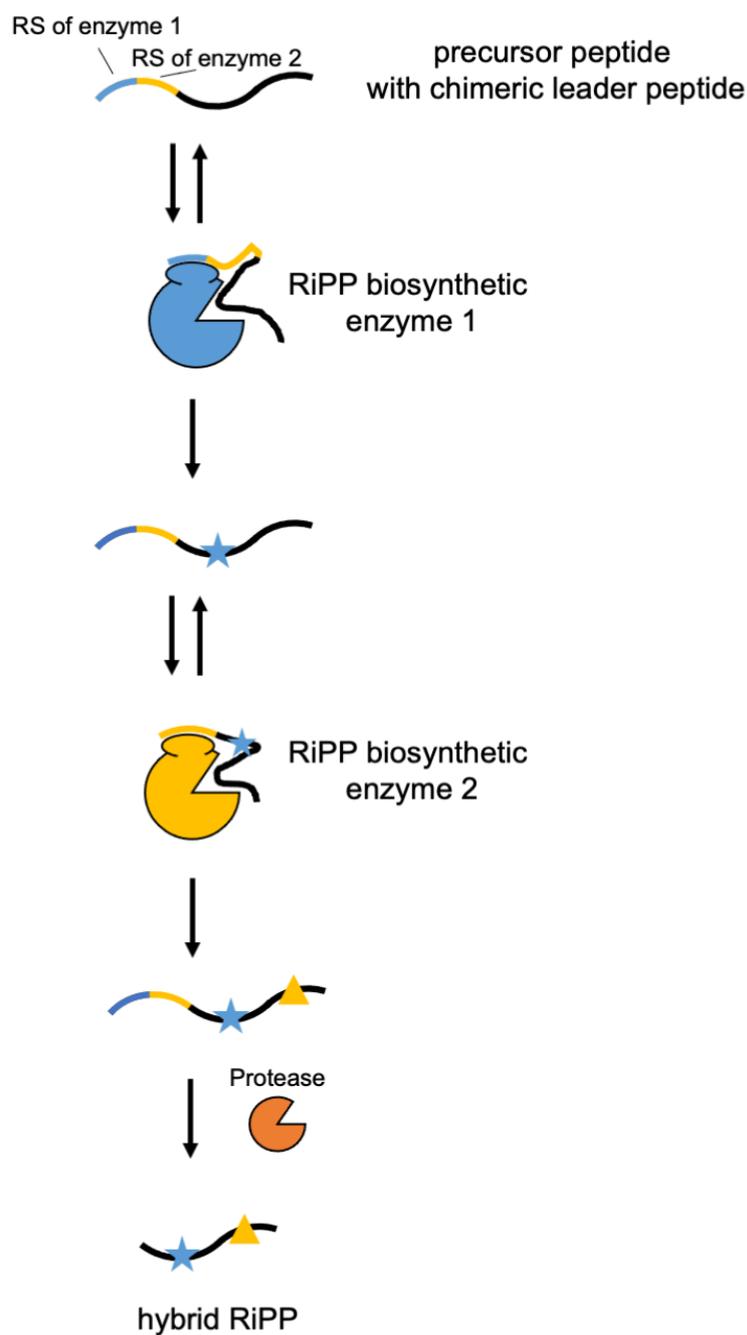
Figure 6-S1 | Chimeric leader peptide strategy as introduced by Mitchell and co-workers.[18] The combination of recognition sequences (RS) of primary enzymes from unrelated RiPP pathways allows the generation of hybrid RiPPs. Thus far this strategy has been limited to the combination of two RSs (as shown in the Scheme) and requires an understanding of the substrate recognition requirements of each enzyme. Figure adapted from Burkhart *et al.*, 2017.[18]
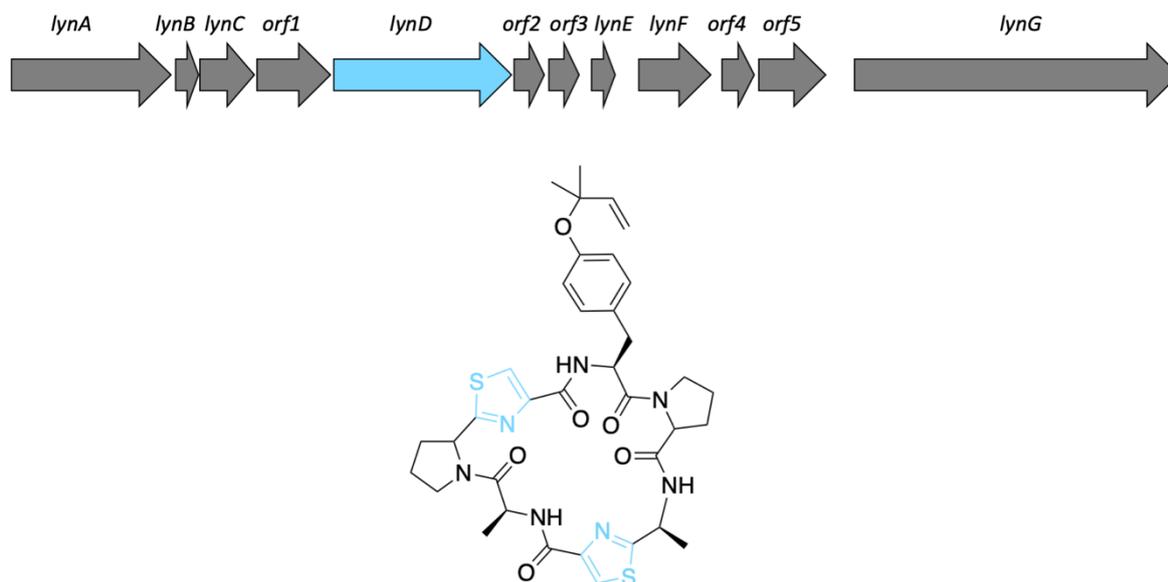
Figure 6-S2 | Aesturamide structure (bottom) and the related cyanobactin biosynthetic gene cluster (BGC) (top) from *Lyngbya* sp..[20] In the BGC the LynD-encoding gene is highlighted in light blue. LynD installs the cysteine-derived thiazolines, which are subsequently oxidised to thiazoles (light blue) by a different enzyme.
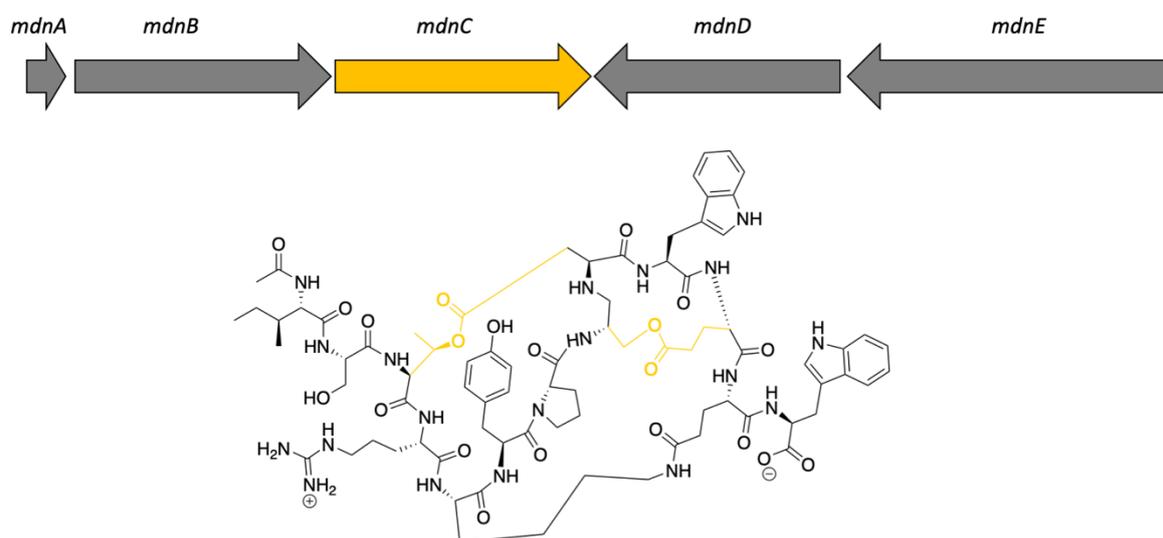
Figure 6-S3 | The ATP grasp ligase MdnC (yellow) from the microviridin BGC (top) catalyses the formation of ω-ester bonds (yellow) from Ser/Thr and Asp/Glu side chains during the biosynthesis of microviridin J (bottom).[26, 27]

```
PatE1        MNKKNILPQQGQPVIRLTAGQLSSQLAELSEEALGDAGLEASVTACITFCAYDGVEPSITVCISVCAYDGE
Peptide 1                       SQLAELSEEALGDAGLEASLPATGGGICTRKYCSDWEEW
                                        LynD RS

MdnA         MAYPNDQQGKALPFFARFLSVSKEESSIKSPSPDHEISTRKYPSDWEEW
Peptide 2               ALPFFARFLSVSKEESSIKLPATGG
Peptide 3               ALPFFARFLSVSKEESSIKLPATGGGICTRKYCSDWEEW
                                 MdnC RS
```

Figure 6-S4 | Peptide sequences of full-length precursor peptides (PPs) PatE1 and MdnA and the peptides **1-3**. In the PP sequences, the leader peptides are shown in light blue resp. yellow, and core peptide (CP) regions are highlighted in bold. The respective recognition sequences (RS) of LynD and MdnC are underlined. Previous studies demonstrated that the PP PatE1 from related patellamide BGC is suitable for LynD.[13] We designed peptide **1** by combining a truncated PatE1 leader peptide (LP), the sortase A recognition motif (LPATG), a di-glycine and the MdnA CP sequence harbouring two point mutations (Ser2 and Pro7 to Cys)[26, 27]. Peptide **2** was designed to contain a truncated MdnA LP and a C-terminal SrtA recognition motif. Peptide **3** can be obtained by sortase A-mediated transpeptidation from peptides **1** and **2**.
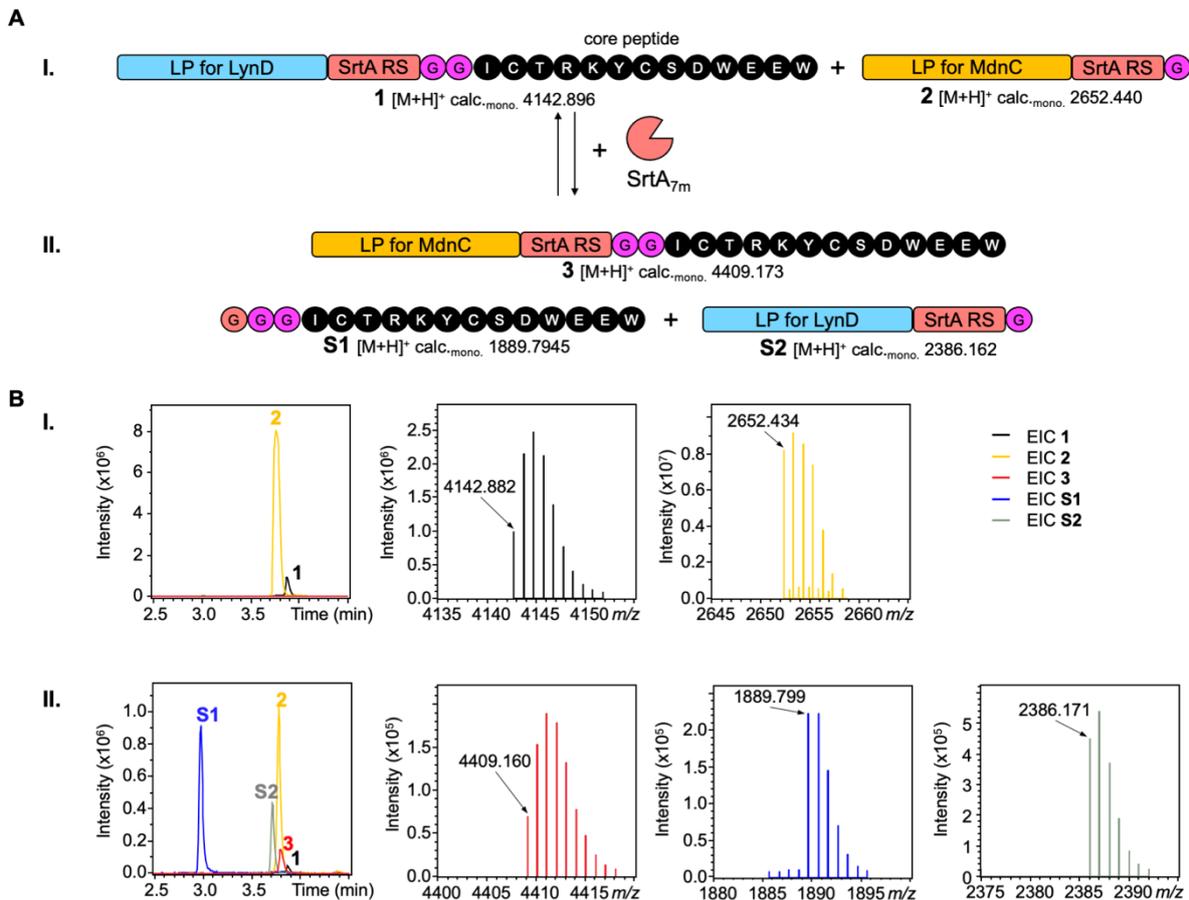
Figure 6-S5 | SrtA$_{7m}$-catalysed leader peptide (LP) exchange. **A** SrtA$_{7m}$ cleaves the threonine-glycine bond in the SrtA recognition motif (LPAT|G) in peptides **1** and **2**, releasing peptide **S1** (from peptide **1**) or di-glycine (from peptide **2**), and forming acyl-enzyme intermediates with the threonine of the peptides ([LP]-LPAT-SrtA). The intermediates are then attacked by **S1** or di-glycine, generating peptide **3** and **S2** (or the peptides **1** and **2** are again produced). In the experiment, peptides **1** and **2** were incubated with (II.) and without (I.) the addition of SrtA$_{7m}$ (2 h, 37 °C). The reactions were analysed by LC-ESI-MS (**B**). Deconvoluted mass spectra for the peptides at the EIC maxima are shown. The masses of the singly-charged monoisotopic ions [M+H]$^+$ are stated in the spectra.
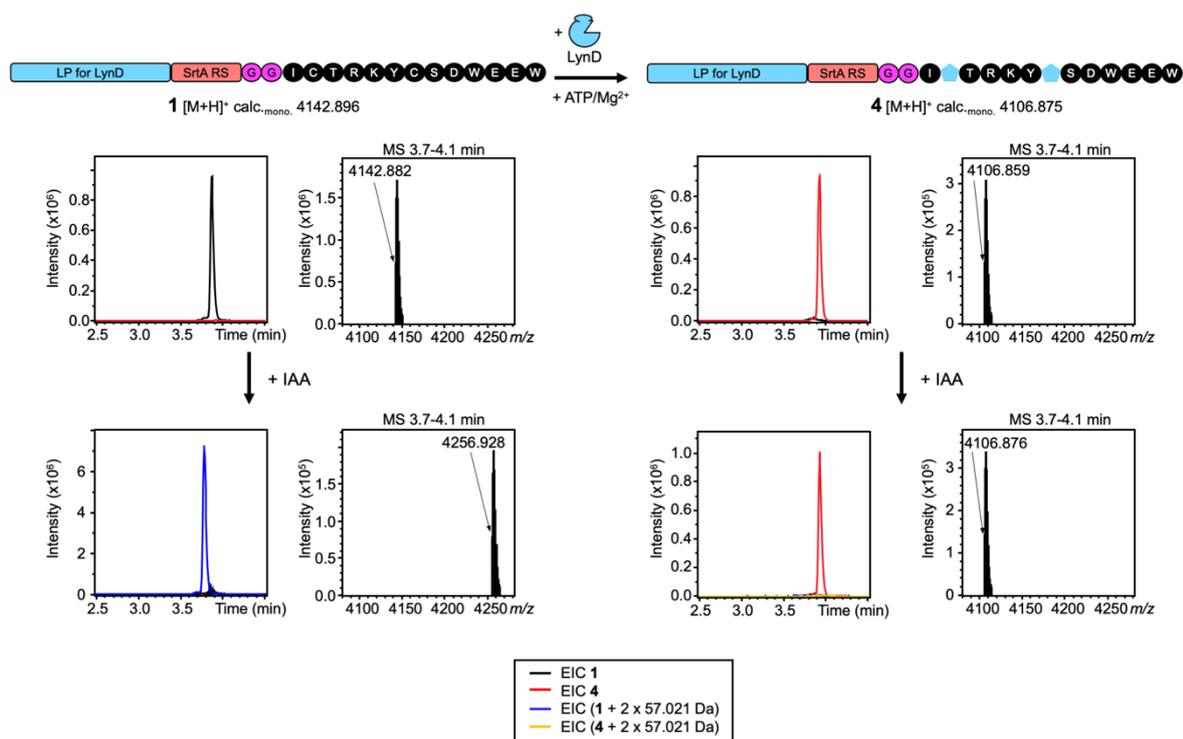
Figure 6-S6 | LynD converts the Cys residues in peptide **1** into thiazolines (light blue pentagons). Peptide **1** was incubated in presence of ATP/Mg^{2+} with and without (negative control) LynD. Analysis of the reactions by LC-ESI-MS showed a loss of 36 Da in the LynD-treated sample, which is in agreement with the formation of two thiazolines. Treatment of the reaction products with iodoacetamide, which alkylates free Cys residues, led to a mass shift in the negative control sample, but not in the LynD-treated sample. This strongly implies that the Cys residues are heterocyclized by LynD as expected. EICs were generated for the possible peptide masses. In the deconvoluted mass spectra the masses of the singly-charged monoisotopic ions [M+H]$^{+}$ are stated.
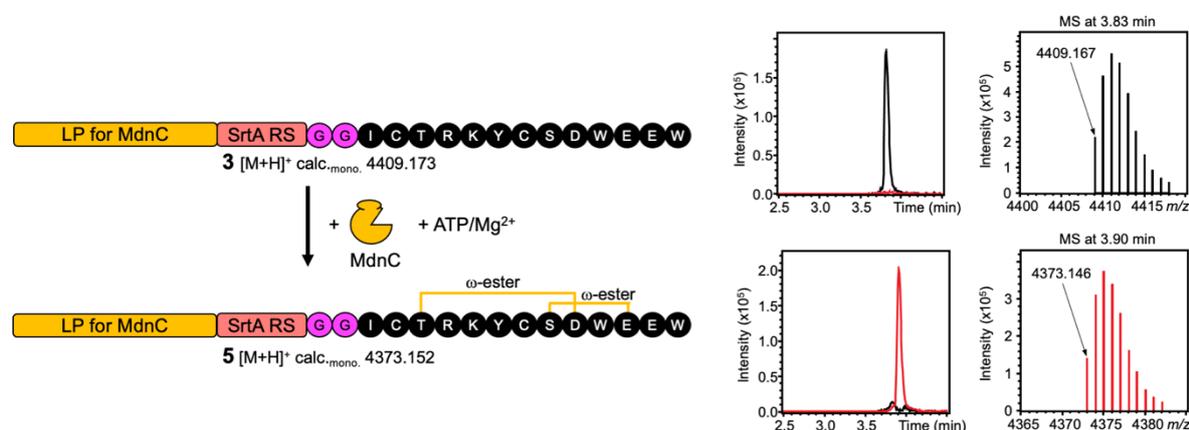
Figure 6-S7 | MdnC introduces two ω-ester bonds into peptide **3**. The incubation of MdnC with peptide **3** in presence of ATP/Mg^{2+} results in the loss of 36 Da, as analysed by LC-ESI-MS. EICs for the expected masses of **1** (black) and **3** (red) were generated. Deconvoluted mass spectra at the EIC maxima state the masses of the singly-charged monoisotopic ions [M+H]$^{+}$.
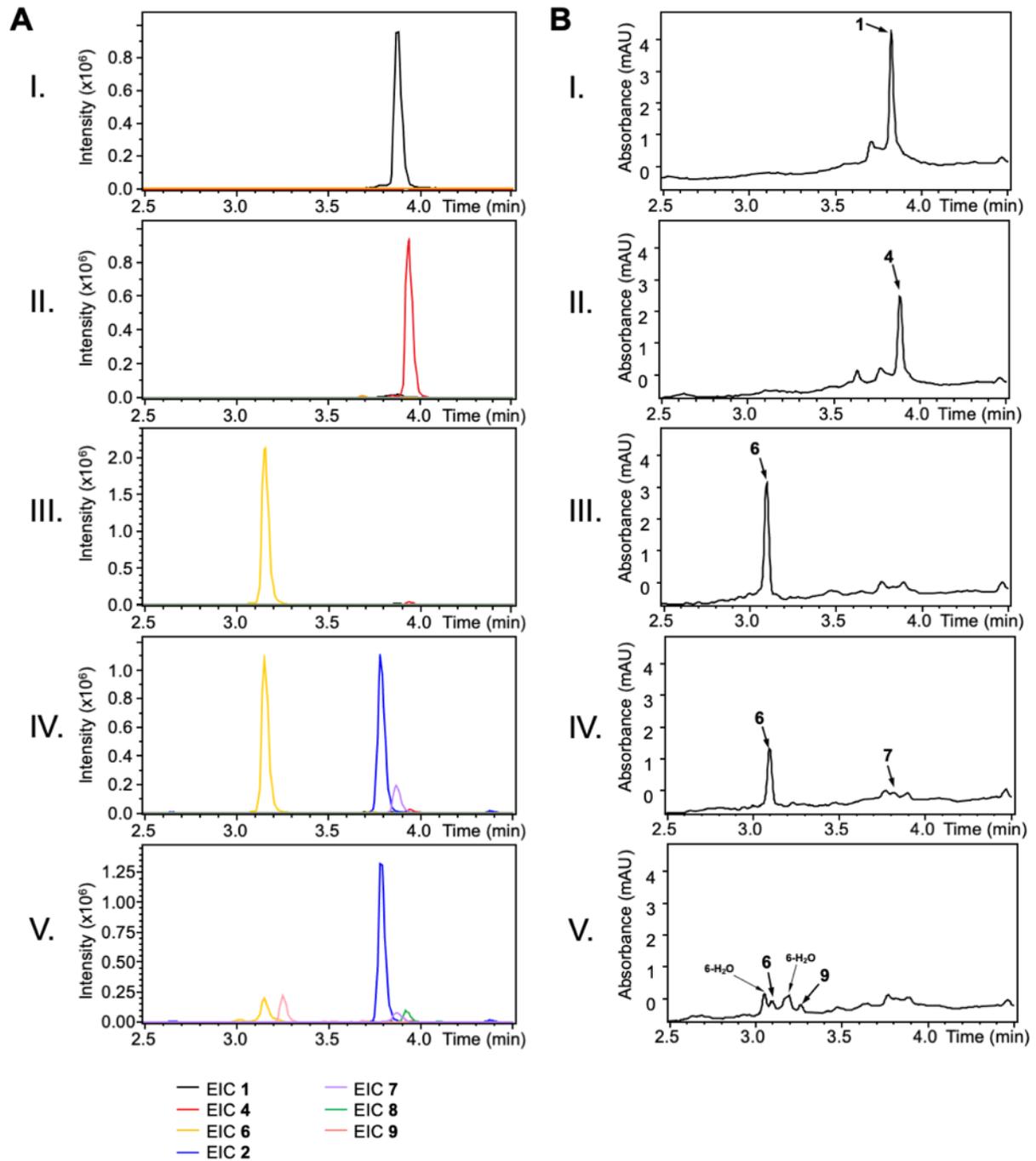
Figure 6-S8 | Combination of the cyanobactin heterocyclase LynD and the microviridn J ATP-grasp ligase MdnC using the ILP strategy to produce a heterocycle-containing graspetide (**9**). **A** EICs for peptides **1-2, 4, 6-9** for the steps I.-V. (see Figure 6-4). **B** UV chromatograms ($\lambda=280$ nm) for the steps I.-V.

fragment	sum formula	theo. mass	obs. mass	error (ppm)
z_1	$C_{11}H_{10}NO_2^+$	188.0712	188.0710	-1.06
y_1	$C_{11}H_{13}N_2O_2^+$	205.0977	205.0973	-1.95
y_4	$C_{32}H_{37}N_6O_9^+$	649.2622	649.2617	-0.77
a_{15}	$C_{70}H_{97}N_{20}O_{19}S_2^+$	1585.6680	1585.6643	-2.33
b_{15}	$C_{71}H_{97}N_{20}O_{20}S_2^+$	1613.6629	1613.6582	-2.91
y_{13}	$C_{76}H_{100}N_{19}O_{19}S_2^+$	1646.6884	1646.6829	-3.34
y_{14}	$C_{78}H_{103}N_{20}O_{20}S_2^+$	1703.7099	1703.7056	-2.52
y_{15}	$C_{80}H_{106}N_{21}O_{21}S_2^+$	1760.7314	1760.7284	-1.70
a_{16}	$C_{81}H_{107}N_{22}O_{20}S_2^+$	1771.7473	1771.7383	-5.07
b_{16}	$C_{82}H_{107}N_{22}O_{21}S_2^+$	1799.7423	1799.7367	-3.11
y_{16}	$C_{82}H_{109}N_{22}O_{22}S_2^+$	1817.7528	1817.7461	-3.69

Figure 6-S9 | MS fragmentation of **9**. The detected fragments of the a, b and y series are shown in the schematic of **9**, in the MS spectrum and in the table**.**
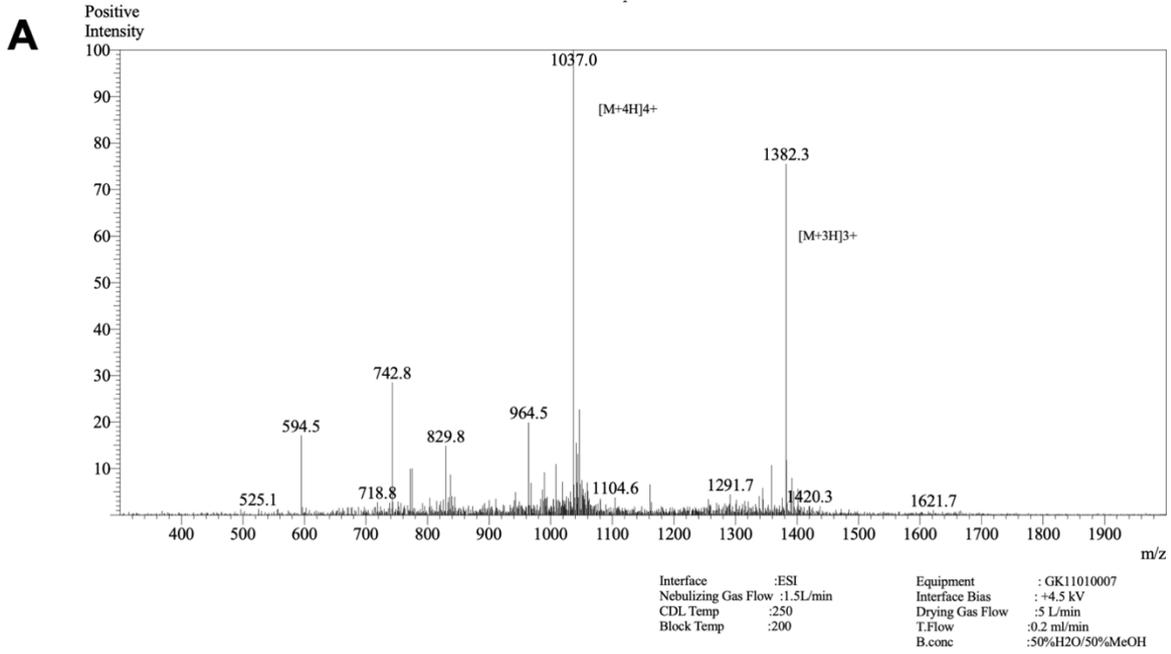
A

Positive
Intensity

1037.0

[M+4H]4+

1382.3

[M+3H]3+

742.8

594.5

964.5

829.8

525.1   718.8   1104.6   1291.7   1420.3   1621.7

m/z

Interface   :ESI   Equipment   : GK11010007
Nebulizing Gas Flow :1.5L/min Interface Bias : +4.5 kV
CDL Temp   :250   Drying Gas Flow :5 L/min
Block Temp   :200   T.Flow   :0.2 ml/min
                               B.conc   :50%H2O/50%MeOH

B

Pump A : 0.065% trifluoroacetic in 100% water (v/v)
Pump B : 0.05% trifluoroacetic in 100% acetonitrile (v/v)
Total Flow:1 ml/min
Wavelength:220 nm
<<LC Time Program>>

Time	Module	Command	Value
0.01	Pumps	Pump A B.Conc	15
25.00	Pumps	Pump A B.Conc	75
25.01	Pumps	Pump A B.Conc	95
31.00	Pumps	Pump A B.Conc	95
31.01	Pumps	Pump A B.Conc	15
40.00	Pumps	Pump A B.Conc	15
40.01	Controller	Stop	

<<Column Performance>>
<Detector A>
Column : Inertsil ODS-3 4.6 x 250 mm
Equipment:ZJ19010324

<Chromatogram>

mV

1 Detector A Channel 1 220nm

<Peak Table>

Detector A Channel 1 220nm
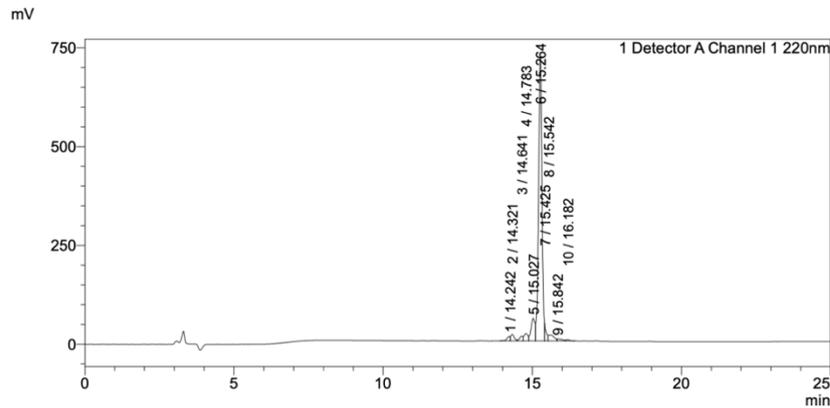
Peak#	Ret. Time	Area	Height	Area%
1	14.242	91565	11057	1.271
2	14.321	131846	16085	1.829
3	14.641	95453	12011	1.325
4	14.783	167047	18643	2.318
5	15.027	535779	56791	7.434
6	15.264	5746482	718456	79.738
7	15.425	155882	36828	2.163
8	15.542	198961	14451	2.761
9	15.842	59251	4816	0.822
10	16.182	24465	2948	0.339
Total		7206732	892085	100.000

Figure 6-S10 | Analytic certificate for peptide **1**. The analytic certificate was created and provided by the vendor GenScript.
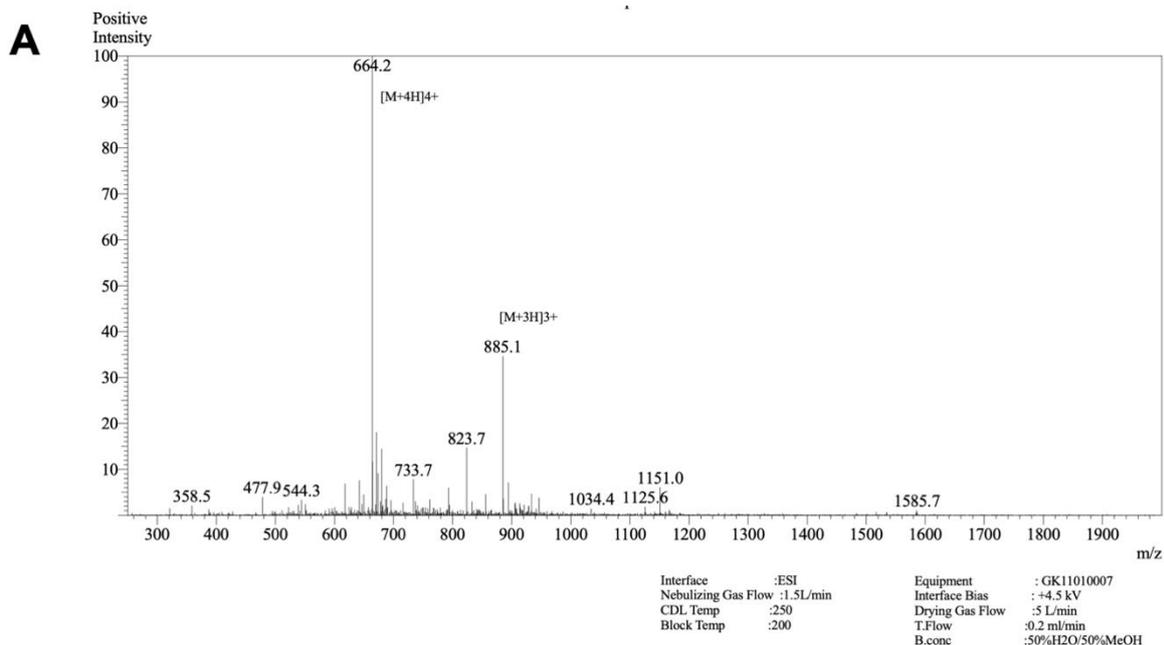
A

Positive
Intensity

664.2
[M+4H]4+

[M+3H]3+
885.1

823.7

733.7

358.5   477.9 544.3   1034.4 1125.6   1151.0   1585.7

m/z

Interface	:ESI	Equipment	: GK11010007
Nebulizing Gas Flow	:1.5L/min	Interface Bias	: +4.5 kV
CDL Temp	:250	Drying Gas Flow	:5 L/min
Block Temp	:200	T.Flow	:0.2 ml/min
		B.conc	:50%H2O/50%MeOH

B

Pump A : 0.065% trifluoroacetic in 100% water (v/v)
Pump B : 0.05% trifluoroacetic in 100% acetonitrile (v/v)
Total Flow:1 ml/min
Wavelength:220 nm

Time	Unit	Command	Value	Comment
0.01	Pumps	Pump A B.Conc	5	
25.00	Pumps	Pump A B.Conc	65	
25.01	Pumps	Pump A B.Conc	95	
27.00	Pumps	Pump A B.Conc	95	
27.01	Pumps	Pump A B.Conc	5	
35.00	Pumps	Pump A B.Conc	5	
35.01	Controller	Stop		

<<Column Performance>>
<Detector A>
Column : Inertsil ODS-3 4.6 x 250 mm
Equipment: GK1101009

Chromatogram

1 Det.A Ch1 / 220nm
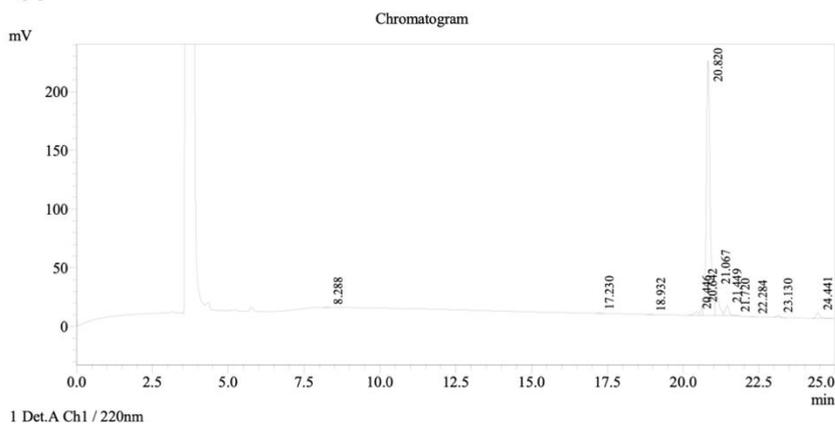
Peak Table

Detector A Ch1 220nm

Peak#	Ret. Time	Area	Height	Area %
1	8.288	1123	140	0.049
2	17.230	3518	546	0.155
3	18.932	1442	249	0.063
4	20.446	32872	2864	1.445
5	20.642	39320	8608	1.729
6	20.820	1859997	217078	81.773
7	21.067	206576	24332	9.082
8	21.449	66638	8364	2.930
9	21.720	3058	477	0.134
10	22.284	1054	159	0.046
11	23.130	14257	1884	0.627
12	24.441	44743	4754	1.967
Total		2274600	269455	100.000

Figure 6-S11 | Analytic certificate for peptide **2**. The analytic certificate was created and provided by the vendor GenScript.
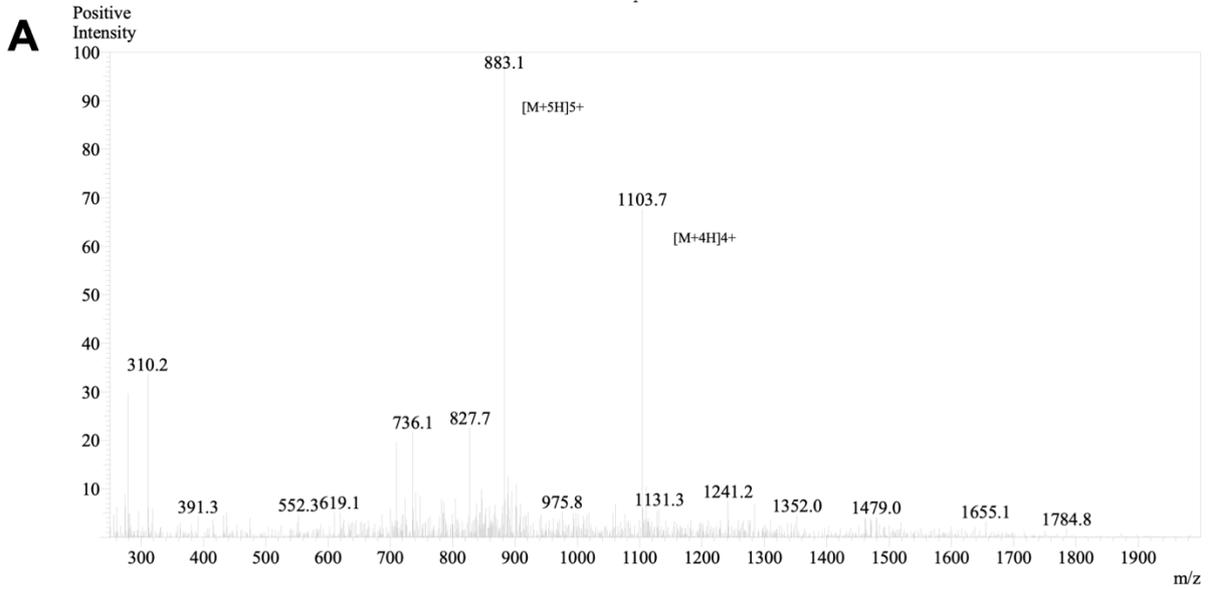
248

A

Positive Intensity

Mass spectrum showing peaks:

883.1 [M+5H]5+
1103.7 [M+4H]4+
310.2
736.1 827.7
391.3 552.3 619.1 975.8 1131.3 1241.2 1352.0 1479.0 1655.1 1784.8

m/z

Interface	:ESI	Equipment	: GK11010007
Nebulizing Gas Flow	:1.5L/min	Interface Bias	: +4.5 kV
CDL Temp	:250	Drying Gas Flow	:5 L/min
Block Temp	:200	T.Flow	:0.2 ml/min
		B.conc	:50%H2O/50%MeOH

B

Pump A : 0.065% trifluoroacetic in 100% water (v/v)
Pump B : 0.05% trifluoroacetic in 100% acetonitrile (v/v)
Total Flow:1 ml/min
Wavelength:220 nm
<<LC Time Program>>

Time	Module	Command	Value
0.01	Pumps	Pump B Conc.	5
25.00	Pumps	Pump B Conc.	65
25.01	Pumps	Pump B Conc.	95
27.00	Pumps	Pump B Conc.	95
27.01	Pumps	Pump B Conc.	5
32.00	Pumps	Pump B Conc.	5
32.01	Controller	Stop	

<<Column Performance>>
<Detector A>
Column :Inertsil ODS-3 4.6 x 250 mm
Equipment:GK12010012

<Chromatogram>

<Peak Table>

Detector A Channel 1 220nm
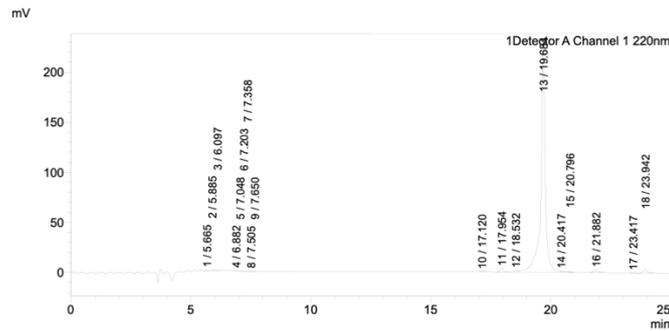
Peak#	Ret. Time	Area	Height	Area%
1	5.665	6944	966	0.225
2	5.885	4712	695	0.153
3	6.097	2683	415	0.087
4	6.882	1848	394	0.060
5	7.048	2009	447	0.065
6	7.203	2088	440	0.068
7	7.358	2091	490	0.068
8	7.505	2141	469	0.069
9	7.650	1824	427	0.059
10	17.120	3210	327	0.104
11	17.954	35716	3696	1.157
12	18.532	23614	1654	0.765
13	19.684	2906153	223546	94.131
14	20.417	1034	215	0.033
15	20.796	2541	354	0.082

Peak#	Ret. Time	Area	Height	Area%
16	21.882	36915	2912	1.196
17	23.417	3505	416	0.114
18	23.942	48311	4008	1.565
Total		3087338	241870	100.000

Figure 6-S12 | Analytic certificate for peptide **3**. The analytic certificate was created and provided by the vendor GenScript.

249

Table 6-S1 | Calculated and observed peptide masses from this study. The masses for the singly-charged monoisotopic ions [M+H]$^+$ are stated. The observed singly-charged masses were obtained by deconvolution using the Bruker Compass DataAnalysis software (Version 4.4).

peptide	corresponding figure	[M+H]$^+$calc.$_{mono.}$ (Da)	[M+H]$^+$obs.$_{mono.}$ (Da)	error (ppm)
1	Fig. 6-S5	4142.896	4142.882	-3.38
2	Fig. 6-S5	2652.440	2652.434	-2.26
3	Fig. 6-S5	4409.173	4409.160	-2.94
4	Fig. 6-S6	4106.875	4106.859	-3.89
5	Fig. 6-S7	4373.152	4373.146	-1.37
1	Fig. 6-4+6-S8	4142.896	4142.882	-3.38
4	Fig. 6-4+6-S8	4106.875	4106.866	-2.19
6	Fig. 6-4+6-S8	1853.773	1853.778	+2.70
7	Fig. 6-4+6-S8	4373.152	4373.141	-2.52
8	Fig. 6-4+6-S8	4337.131	4337.131	-0.05
9	Fig. 6-4+6-S8	1817.752	1817.745	-3.85

Materials & Methods

<u>Protein expression and purification</u>

Synthetically produced, codon-optimized genes coding for SrtA$_{7m}$[23] and MdnC (PDB 5IG9) were obtained from Eurofins Genomics. *srtA$_{7m}$* was cloned into the pHisTEV plasmid[28], *mdnC* was cloned into the pBMS plasmid, both were a gift from Dr. Huanting Liu (University of St Andrews, UK). The synthetic gene for LynD was a gift from Prof. James Naismith (University of Oxford, UK). The resulting constructs were verified by DNA sequencing using T7 promotor and terminator (LGC genomics) and transformed into chemically competent *E. coli* Lemo21(DE3) cells (New England Biolabs) using a standard heat shock procedure. A single colony was inoculated into LB medium supplemented with the appropriate antibiotics and incubated overnight at 37 °C, 200 rpm for an overnight culture.

For expression of SrtA$_{7m}$, the overnight culture was diluted 1:100 into fresh LB medium supplemented with the appropriate antibiotics. The resulting cultures were grown at 37 °C, 180 rpm until an OD$_{600}$ of 0.8, at which point the cultures were transferred to a pre-cooled 20 °C shaker (180 rpm) and protein expression was induced by the addition of 0.5 mM IPTG. Cultures were harvested after 16 h incubation by centrifugation. The cell pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 20 mM Imidazole, 3 mM beta-ME, 10% glycerol) supplemented with cOmplete EDTA-free protease inhibitor tablets (Roche) and DNase (0.4 mg/g wet cells, Sigma). The cell suspension was lysed via passage through a cell disrupter (30 kpsi, Constant Systems), and cell debris was removed by centrifugation (40,000 x g, 4 °C, 15 min). The supernatant was loaded onto a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated in lysis buffer. After an extensive column wash (30 CV), the His$_6$-tagged protein was eluted using lysis buffer supplemented with 250 mM imidazole. The imidazole was removed by passing the eluate through a 16/10 desalting column (GE Healthcare), equilibrated in lysis buffer. To cleave the His$_6$-tag, the protein was incubated with TEV protease (ratio 1:10 TEV : protein, 4 °C, 16 h). After the incubation, the solution was applied to a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated in lysis buffer. The flow through was concentrated and loaded onto a gel filtration column (HiLoad 16/600 Superdex 200 pg, GE healthcare) pre-equilibrated in gel filtration buffer (20 mM HEPES, 200 mM NaCl, 10% glycerol, 1 mM TCEP, pH 7.4). Fractions of highest purity were determined by SDS-PAGE, pooled and concentrated to 500 μM.

MdnC and LynD were expressed and purified similarly using the expression conditions and buffers as previously described.[13, 27]

In vitro experiments

Peptides used in this study were purchased from GenScript. The peptide purity was; peptide **1** 79.7%, peptide **2** 81.7%, peptide **3** 94.1%. The vendors analytic certificates are shown in figures S9-S11.

To probe the SrtA$_{7m}$ catalysed exchange of the LP for LynD with the LP for MdnC, 100 µM peptide **1** and 100 µM peptide **2** were incubated with and without the addition of 5 µM SrtA$_{7m}$ in reaction buffer 1 (20 mM HEPES, 200 mM NaCl, 2.5 mM TCEP, pH 7.4) for 2 h at 37 °C.

To test the ability of LynD to convert the two Cys residues in the peptide **1** into thiazolines, 100 µM peptide **1** was incubated with and without the addition of LynD in reaction buffer 2 (5 mM ATP, 5 mM MgCl$_2$, 20 mM HEPES, 200 mM NaCl, 2.5 mM TCEP, pH 7.4) for 2 h at 37 °C.

The ability of MdnC to introduce two ω-ester bonds into peptide **3** was tested by incubating 50 µM peptide **3** with 5 µM MdnC in reaction buffer 2 for 8 h at 37°C.

For the combination of all components, 100 µM peptide **1** was incubated with 5 µM LynD in reaction buffer 2 for 2 h at 37 °C. Next, 100 µM peptide **2**, 5 µM SrtA$_{7m}$ and 5 µM MdnC were added and incubated for 16 h at 37 °C. Negative controls were set up analogously without the addition of SrtA$_{7m}$.

For iodoacetamide alkylation, 20 mM iodoacetamide was added to the reaction solutions and incubated for 1 h at 20 °C in the dark.

All reactions were stopped by the addition of 2 volumes ACN and frozen at -80 °C until analysis by LC-ESI-MS. All reactions were run at least three times in independent experiments with similar results.

LC-MS and MS^2 analysis

All measurements were performed on a Dionex Ultimate 3000 RSLC system (Thermo Fisher Scientific) using a flow rate of 600 µL min^{-1} and a column oven temperature of 45 °C. Samples were separated by a gradient from (A) H$_2$O + 0.1% formic acid to (B) ACN + 0.1% formic acid using a BEH C18, 50 x 2.1 mm, 1.7 µm dp column equipped with a C18 precolumn (Waters). The linear gradient was initiated by a 1 min isocratic step at 5% B, followed by an increase to 95% B in 9 min to end up with a 1.5 min plateau step at 95% B before re-equilibration to the initial conditions.

For MS measurements on a maXis-4G hr-qToF mass spectrometer (Bruker Daltonics), the LC flow was split 1:8 before entering the mass spectrometer using the Apollo II ESI source. In the source region, the temperature was set to 200 °C, the capillary voltage was 4000 V, the dry-gas flow was 5.0 L / min and the nebulizer was set to 1 bar. Ion transfer settings were set to Funnel

1 RF 350 Vpp and Multipole RF 400 Vpp, quadrupole settings were set to an ion energy of 5.0 eV and a low mass cut of 120 m/z. The collision cell was set to an energy of 5.0 eV and the pulse storage time was 5 µs. Data were recorded in centroid mode ranging from 150 to 2500 m/z at a 2 Hz scan rate. Calibration of the maXis4G qTOF spectrometer was achieved with sodium formate clusters before every injection to avoid mass drifts. All MS analyses were acquired in the presence of the lock masses $C_{12}H_{19}F_{12}N_3O_6P_3$, $C_{18}H_{19}O_6N_3P_3F_2$ and $C_{24}H_{19}F_{36}N_3O_6P_3$ which generate the $[M+H]^+$ ions of 622.028960, 922.009798 and 1221.990638.

LC-MS^2 fragmentation spectra were recorded using a scheduled precursor list (SPL). The SPL entries and parameters were set to fragment the $[M+6H]^{6+}$ ion of **9**. SPL tolerance parameters for precursor ion selection were 0.15 min and 0.05 m/z. The CID energy was ramped from 35 eV for 500 m/z to 45 eV for 1000 m/z.

Data were displayed and analyzed using the Bruker Compass DataAnalysis software (Version 4.4). Shown MS spectra are deconvoluted and are single spectra (taken at the EIC maximum), except for Figure 6-S6, where spectra are averaged. The signals in the MS-spectra are labelled with the observed monoisotopic mass. Extracted-ion chromatograms were generated using the calculated monoisotopic masses of the ions with the charge states 1-6 and a mass range of 0.01 Da.

6.4 References

1. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod* **2020,** *83* (3), 770-803.

2. Katz, L.; Baltz, R. H., Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol* **2016,** *43* (2-3), 155-76.

3. 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.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports* **2013,** *30* (1), 108-160.

4. Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A., New developments in RiPP discovery, enzymology and engineering. *Natural Product Reports* **2020.**

5. Oman, T. J.; van der Donk, W. A., Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nature Chemical Biology* **2010,** *6* (1), 9-18.

6. Schmitt, S.; Montalbán-López, M.; Peterhoff, D.; Deng, J.; Wagner, R.; Held, M.; Kuipers, O. P.; Panke, S., Analysis of modular bioengineered antimicrobial lanthipeptides at nanoliter scale. *Nature Chemical Biology* **2019,** *15* (5), 437-443.

7. Hetrick, K. J.; Walker, M. C.; van der Donk, W. A., Development and Application of Yeast and Phage Display of Diverse Lanthipeptides. *ACS Central Science* **2018,** *4* (4), 458-467.

8. Urban, J. H.; Moosmeier, M. A.; Aumüller, T.; Thein, M.; Bosma, T.; Rink, R.; Groth, K.; Zulley, M.; Siegers, K.; Tissot, K.; Moll, G. N.; Prassler, J., Phage display and selection of lanthipeptides on the carboxy-terminus of the gene-3 minor coat protein. *Nature Communications* **2017,** *8* (1), 1500.

9. Oueis, E.; Jaspars, M.; Westwood, N. J.; Naismith, J. H., Enzymatic Macrocyclization of 1,2,3-Triazole Peptide Mimetics. *Angewandte Chemie International Edition* **2016,** *55* (19), 5842-5845.

10. Vinogradov, A. A.; Shimomura, M.; Goto, Y.; Ozaki, T.; Asamizu, S.; Sugai, Y.; Suga, H.; Onaka, H., Minimal lactazole scaffold for in vitro thiopeptide bioengineering. *Nature Communications* **2020,** *11* (1), 2272.

11. Pan, S. J.; Link, A. J., Sequence diversity in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. *J Am Chem Soc* **2011,** *133* (13), 5016-23.

12. Reyna-González, E.; Schmid, B.; Petras, D.; Süssmuth, R. D.; Dittmann, E., Leader Peptide-Free In Vitro Reconstitution of Microviridin Biosynthesis Enables Design of Synthetic Protease-Targeted Libraries. *Angew Chem Int Ed Engl* **2016,** *55* (32), 9398-401.

13. Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W.; Jaspars, M.; Naismith, J. H., Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat Chem Biol* **2015,** *11* (8), 558-563.

14. Gu, W.; Sardar, D.; Pierce, E.; Schmidt, E. W., Roads to Rome: Role of Multiple Cassettes in Cyanobactin RiPP Biosynthesis. *J Am Chem Soc* **2018,** *140* (47), 16213-16221.

15. Oman, T. J.; Knerr, P. J.; Bindman, N. A.; Velásquez, J. E.; van der Donk, W. A., An engineered lantibiotic synthetase that does not require a leader peptide on its substrate. *J Am Chem Soc* **2012,** *134* (16), 6952-5.

16. Goto, Y.; Ito, Y.; Kato, Y.; Tsunoda, S.; Suga, H., One-pot synthesis of azoline-containing peptides in a cell-free translation system integrated with a posttranslational cyclodehydratase. *Chem Biol* **2014,** *21* (6), 766-74.

17. Khusainov, R.; Kuipers, O. P., When the leader gets loose: in vivo biosynthesis of a leaderless prenisin is stimulated by a trans-acting leader peptide. *Chembiochem* **2012,** *13* (16), 2433-8.

18. Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A., Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Central Science* **2017,** *3* (6), 629-638.

19. Fleming, S. R.; Bartges, T. E.; Vinogradov, A. A.; Kirkpatrick, C. L.; Goto, Y.; Suga, H.; Hicks, L. M.; Bowers, A. A., Flexizyme-Enabled Benchtop Biosynthesis of Thiopeptides. *Journal of the American Chemical Society* **2019,** *141* (2), 758-762.

20. Martins, J.; Vasconcelos, V., Cyanobactins from Cyanobacteria: Current Genetic and Chemical State of Knowledge. *Mar Drugs* **2015,** *13* (11), 6910-6946.

21. Ziemert, N.; Ishida, K.; Weiz, A.; Hertweck, C.; Dittmann, E., Exploiting the Natural Diversity of Microviridin Gene Clusters for Discovery of Novel Tricyclic Depsipeptides. *Applied and Environmental Microbiology* **2010,** *76* (11), 3568-3574.

22. Li, K.; Condurso, H. L.; Li, G.; Ding, Y.; Bruner, S. D., Structural basis for precursor protein–directed ribosomal peptide macrocyclization. *Nature Chemical Biology* **2016,** *12* (11), 973-979.

23.   Hirakawa, H.; Ishikawa, S.; Nagamune, T., Ca2+-independent sortase-A exhibits high selective protein ligation activity in the cytoplasm of Escherichia coli. *Biotechnology Journal* **2015,** *10* (9), 1487-1492.

24.   Liu, F.; Luo, E. Y.; Flora, D. B.; Mezo, A. R., Irreversible Sortase A-Mediated Ligation Driven by Diketopiperazine Formation. *The Journal of Organic Chemistry* **2014,** *79* (2), 487-492.

25.   Antos, J. M.; Chew, G.-L.; Guimaraes, C. P.; Yoder, N. C.; Grotenbreg, G. M.; Popp, M. W.-L.; Ploegh, H. L., Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity. *J Am Chem Soc* **2009,** *131* (31), 10800-10801.

26.   Ziemert, N.; Ishida, K.; Liaimer, A.; Hertweck, C.; Dittmann, E., Ribosomal Synthesis of Tricyclic Depsipeptides in Bloom-Forming Cyanobacteria. *Angewandte Chemie International Edition* **2008,** *47* (40), 7756-7759.

27.   Li, K.; Condurso, H. L.; Li, G.; Ding, Y.; Bruner, S. D., Structural basis for precursor protein-directed ribosomal peptide macrocyclization. *Nature chemical biology* **2016,** *12* (11), 973-979.

28.   Liu, H.; Naismith, J. H., A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr Purif* **2009,** *63* (2), 102-11.

7 Discussion and Outlook

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a major class of natural products with a wide array of biological activities (e.g. anti-bacterial, -viral, -cancer).[1] Although RiPPs are of ribosomal origin, they are structurally and chemically highly diversified through post-translational modifications by biosynthetic enzymes.[2] In the recent years the superfamily of RiPP natural products has been expanding rapidly and attracted interest from basic and applied researchers due to the biological activities of RiPPs, the malleability of their biosynthetic pathways, and the use of RiPP biosynthetic enzymes for chemoenzymatic applications.

In this thesis, the biosynthesis of the RiPP bottromycins was elucidated using an *in vitro* approach. Bottromycins have a unique scaffold that contains a four-amino acid macrocycle formed via an amidine linkage, several β-methylated amino acids, an O-methylated D-aspartate and a C-terminal thiazole.[3] These structural features endow them with potent activity against clinically relevant human pathogens such as MRSA and VRE and their mechanism of action differs from all other antibiotics currently in clinical use.[4-7] The BGC of these unique molecules was discovered in 2012, which revealed their ribosomal origin.[8-11] In 2017, when this PhD project was initiated, bottromycin biosynthesis had mainly been proposed based on gene deletion and untargeted metabolomics approaches, which led to the tentative assignment of enzyme functions for most of the biosynthetic enzymes.[12] However, the metabolomics data was insufficient to assign the functions of all the enzymes encoded in the BGC. Studying the bottromycin biosynthesis *in vitro* enabled the individual characterization of the enzyme functions, kinetics, the elucidation of the sequence of the biosynthetic steps based on the enzyme substrate preferences and - by pairing biochemical with structural data - the understanding of enzyme mechanisms. This process held a few surprises in store and led to the partial revision of the proposed biosynthetic pathway. The *in vitro* characterization and functional reconstitution of the YcaO domain enzymes and the α/β-hydrolase encoded in the bottromycin BGC led to the expansion of the catalytic repertoire of these two enzyme families, namely by amidine formation (YcaO domain enzymes) and epimerization (α/β-hydrolases). A detailed analysis of macroamidine formation led to the realization that amidohydrolases can act as YcaO accessory proteins. The function of the bottromycin cytochrome P450 (CYP) enzyme was confirmed by the *in vitro* reconstitution of the reaction: it catalyzes the oxidative decarboxylation of the thiazoline. The bottromycin CYP enzyme is the first discovered CYP enzyme in the context of RiPPs that catalyzes an oxidative carboxylation reaction. As a result

of the work conducted during this thesis, the *in vitro* biosynthesis of the bottromycin core scaffold was achieved and will now enable the rapid generation bottromycin analogs for compound development. Furthermore, the results of this work increased our understanding how the bottromycin pathways can be manipulated (e.g. for enhanced production yields *in vivo*). The insights into bottromycin biosynthesis gained in this work are also of significance in a more general way; they contribute to our understanding RiPP biosynthesis, broaden the knowledge and capabilities of the studied enzyme classes and aid the mining for and prediction of RiPP natural products.

In addition to the elucidation of bottromycin biosynthesis, a novel strategy to generate new-to-nature ribosomal natural products by combination of RiPP enzymes from unrelated RiPP pathways through sortase A-mediated leader peptide exchange was accomplished. This approach has the potential to unlock the immense combinatorial power of RiPP pathways to produce novel peptide scaffolds for drug discovery.

7.1 Elucidation of bottromycin biosynthesis

Over the course of this thesis, five bottromycin biosynthetic enzymes, which are sufficient to produce the bottromycin core scaffold, were characterized: The two YcaO domain enzymes BotC and BotCD, the amidohydrolase PurAH, the α/β-hydrolase BotH and the cytochrome P450 enzyme BotCYP. The results led to a revised biosynthetic pathway (Figure 7-1).

The first enzymes I investigated were the two YcaO domain enzymes present in the bottromycin BGC (chapter 2). The function of YcaO enzymes in RiPP biosynthesis had been studied quite extensively, and they have been shown to convert Ser, Thr and Cys residues into the corresponding azolines by catalyzing the nucleophilic attack of the amino acid side chains to the preceding carbonyl carbon. YcaO domain enzymes are found in the BGCs of several RiPP classes, e.g. linear azole-containing peptides (LAPs), cyanobactins, and thiopeptides, and always catalyze ATP-dependent backbone modifications. Previously well-characterized azoline-forming YcaO enzymes have been shown to require RRE-containing Ocin-ThiF- or E1-like partner proteins, expressed in fusion or separately, to recruit the substrate peptides and/or activate the enzymes.[2, 13-15]

The YcaO domain enzymes of the bottromycin BGC lack an Ocin-ThiF-/E1-like domain and are hence referred to as 'standalone' YcaO domain enzymes. Based on the known function of characterized YcaO enzymes, it was proposed that one of the two bottromycin YcaO domain enzymes catalyzes thiazoline formation, but the function of the second YcaO domain enzyme was enigmatic. Gene deletion experiments coupled with mass spectrometry-based molecular networking demonstrated that the YcaO domain enzyme BotC (the *S.* sp. BC16019 BGC nomenclature is used throughout the discussion) installs the thiazoline, which is later on oxidized to thiazole, into the bottromycin molecule.[12] Macroamidine formation was abolished after deletion of the gene encoding for the second YcaO domain enzyme, BotCD. However, no macroamidine containing intermediates were detected after the deletion of the gene encoding for the putative amidohydrolase BotAH either.[12] Hence, it was proposed that BotCD and BotAH were both involved in macroamidine formation.

The *in vitro* characterization of the BotC homolog IpoC and the BotCD homolog PurCD (chapter 2), demonstrated that former catalyzes the thiazoline formation, while the latter is sufficient to catalyze the macroamidine formation in an ATP-dependent manner even without the presence of BotAH. Both YcaO domain enzymes bind to the follower peptide and are thus responsible for both, substrate binding and catalysis.
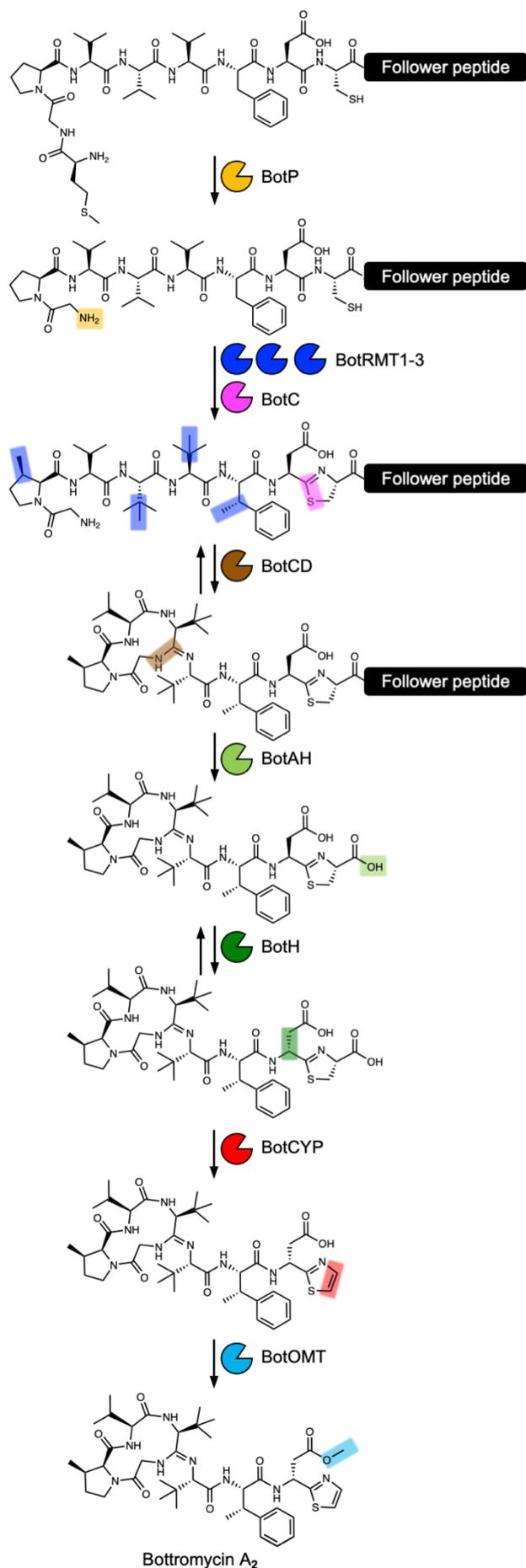
Figure 7-1 | Bottromycin biosynthetic pathway elaborated in this thesis.

Our results expanded the catalytic repertoire of YcaO enzymes to amidine formation and described the first biochemical studies of standalone YcaO enzymes. It has since been reported that YcaO enzymes also catalyze thioamide formation (TfuA-dependent and -independent), e.g. in thioviridamides, and are potentially responsible for amide backbone modifications in bacterial ribosomal proteins.[16-25] A RODEO[26] (a tool for RiPP BGC analysis) survey of protein families co-occurring with the about 13,000 YcaOs deposited in GenBank revealed that more than 60% of all YcaOs do not co-occur with known YcaO partner proteins (Ocin-ThiF-/E1- or TfuA-associated) and are classified as standalone YcaOs.[27] Further analyses from this study showed a higher similarity of non-canonical YcaOs to the amidine-forming than to the thiazoline-forming bottromycin YcaO, suggesting that amidine-forming YcaOs are also present in other uncharacterized RiPP BGGs.[27]

Indeed, following to our publication, two additional YcaO enzymes catalyzing amidine formation were identified. The bifunctional YcaO KlpD from klebsazolicin biosynthesis catalyzes the formation of azolines and the N-terminal amidine ring found in the klebsazolicin structure in complex with an E1-like protein and a dehydrogenase (KlpBCD) (Figure 7-2).[28] While in bottromycin biosynthesis BotCD catalyzes the nucleophilic attack from Gly1 onto the carbonyl carbon of Val4, KlpD catalyzes the attack of the amino group of Ser1 onto the carbonyl carbon of Gln2, forming a much smaller ring (6-membered) than observed in bottromycins (12-membered ring). The second amidine-forming YcaO was recently found during a genome mining approach identifying standalone actinobacterial YcaO enzymes associated with RiPP BGCs. This led to the discovery of the novel RiPP streptamidine, which contains a 6-membered amidine ring (Figure 7-2).[29] The streptamidine YcaO AmiD catalyzes the formation of an amidine ring formed between the amino group of His1 and the carbonyl carbon of Leu2. The genome mining approach describing stepamidine formation identified over 230 further uncharacterized novel actinobacterial RiPP BGCs associated with standalone YcaOs.[29] These recent data support the idea that amidine-containing RiPPs may be more widespread than previously anticipated, and that the bottromycin amidine-forming YcaO, whose function we determined here, was probably the founding member of a larger, amidine-forming YcaO family.
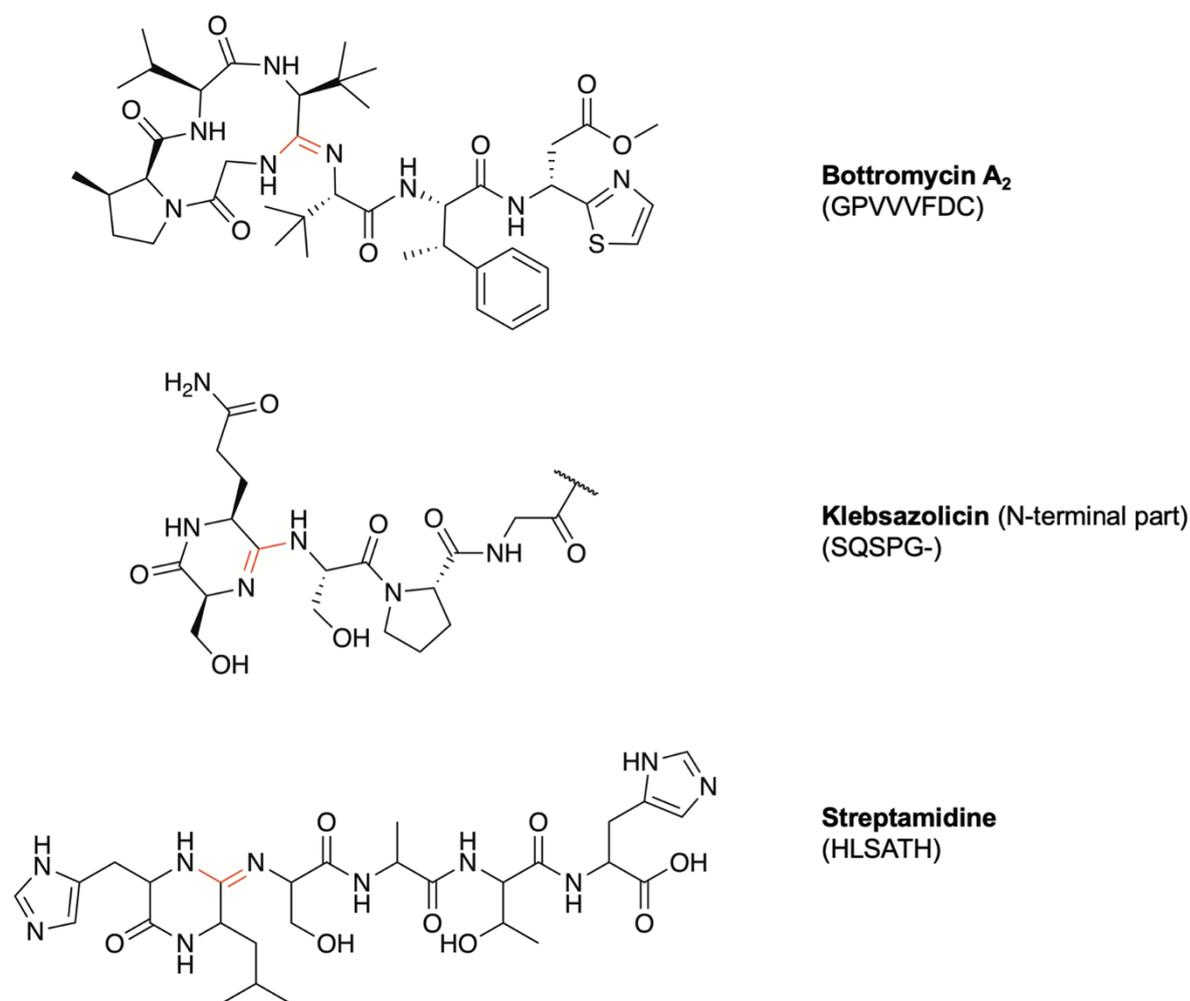
Figure 7-2 | Structures of currently known amidine-containing RiPPs. The respective core peptide sequences are shown in the brackets.

As mentioned above, the bottromycin YcaOs do not require a partner protein and are thus 'RRE-independent', but they were shown to bind to the follower peptide. The minimal length of the follower peptide for enzyme stabilization and efficient turnover is consistent with the sequence conservation of the follower peptide in homologous Streptomyces bottromycin BGCs. In addition to our study, another group characterized the bottromycin YcaO enzymes's follower binding and suggested differences in follower binding between the thiazoline- and macroamidine-forming bottromycin YcaO.[27] To understand the follower binding of standalone YcaO enzymes, the determination of YcaO-follower peptide complex structures are of high interest, but remain elusive despite of significant efforts. Structural information on YcaO enzymes (apo-structures) and complex structures of RiPP enzymes with the respective follower peptides are scarce. A complex structure of a YcaO in complex with its precursor peptide was also only achieved for one enzyme, the cyanobactin heterocyclase LynD (and McbD within the

microcin B17 synthetase complex), which emphasizes the difficulty to obtain YcaO-peptide complex structures.[30, 31]

Thiazoline formation catalyzed by BotC was determined to proceed very slowly *in vitro*. Since bottromycin production is severely diminished by deletion of the RMTs,[8] we speculated that the missing methylations on Pro, Vals and Phe are responsible for the slow turnover. Nevertheless, using a synthesized precursor peptide β-methylated at Val4, Val5 and Phe6 (methylation only at Pro2 missing) and comparison of the turnover of the methylated and non-methylated precursor peptide, did not show a change in reaction speed (unpublished data). Unfortunately, the correct MePro isomer was not commercially available at the time, and it can thus not be excluded that methylation at Pro2 is key to accelerate heterocyclization by IpoC. Bottromycins that only differ in their proline methylation pattern, bottromycins A_2, B_2 and C_2, are observed naturally, with bottromycin A_2 being the major product. While bottromycin A_2 contains MePro, bottromycin B_2 contains Pro and bottromycin C_2 contains Me_2Pro at the second position.[32] Bottromycin A_2 and C_2 are more potent than bottromycin B_2. Complex formation between BotC and BotCD *in vitro* was not observed (unpublished data), but an interaction with another biosynthetic enzyme (e.g. BotRMT1-3) for efficient turnover cannot be excluded. If slow cysteine to thiazoline processing is not a result of the *in vitro* environment, it may simply *be* slow to ensure full processing (esp. β-methylation) by the preceding enzymes before further modifications are introduced.

The amidine-forming YcaO, BotCD, rapidly converts precursor peptides that contain a thiazoline. Without the thiazoline turnover was minimal, which suggests that BotCD acts after BotC (Figure 7-1). Surprisingly, macroamidine formation was reversible, with a very small equilibrium constant (< 1).

Next, we studied a very close homolog of the putative amidohydrolase BotAH, PurAH, which had been suggested to be directly involved in macroamidine formation.[33] However, we had demonstrated that BotCD alone was sufficient to catalyze the macroamidine formation. The *in vitro* study of PurAH (chapter 3) revealed that the enzyme is responsible for cleavage of the follower peptide and highly selective for heterocyclized and macrocyclized substrate. This specificity ensures the modification by the preceding enzymes, that are necessary for biological activity, are installed before follower cleavage. The macroamidine cycle cannot be re-opened by BotCD after the follower has been removed. While a BotCD reaction will always show limited turnover, even under optimized conditions, addition of PurAH results in complete turnover. These results explain the disparity between the *in vivo* and *in vitro* data and the

accumulation of the thiazoline- rather than the macroamidine-containing intermediate after deletion of *botAH*.

The highest PurAH activity of this metallo-dependent enzyme was observed in the presence of Co^{2+} ions. This mirrors the results for the aminopeptidase BotP, which removes the N-terminal methionine of the precursor peptide and also requires Co^{2+} for highest activity. The RMTs are also predicted to possess cobalamin-binding domains.[8, 34] In light of these data it may not be surprising that the addition of Co^{2+} to production medium resulted in an increase of bottromycin production.[9] The Co^{2+}-dependence of gene regulation or expression levels for the bottromycin BGC was recently ruled out based on RT-qPCR analyses.[35] However, the same study showed that an increase in the Co^{2+} concentration led to an increase in mycelial growth of the bottromycin producing strain up to a certain concentration, which may provide an alternative explanation for the observed increase in bottromycin production.[35]

Follower cleavage by the amidohydrolase BotAH generates a C-terminal 2-thiazoline-4-carboxylic acid moiety (Figure 7-1). Azoline heterocycles are usually oxidized to the chemically more stable aromatic azoles in natural products. Azoline heterocycles that are not oxidized in the final natural product are often protected by steric and electronic effects from nucleophilic attack.[36] In RiPPs, the oxidation of thiazolines/oxazolines is usually catalyzed by FMN-dependent dehydrogenases.[2, 37] The bottromycin BGC however lacks a gene encoding for a dehydrogenase.[8-11] Gene deletion experiments demonstrated that the bottromycin cytochrome P450 enzyme BotCYP catalyzes the oxidative decarboxylation of the thiazoline into the thiazole that is found in the natural product.[33] A similar reaction of a CYP enzyme is known from the biosynthesis of the plant alkaloid camalexin, which catalyzes the oxidative decarboxylation reaction of the thiazoline-4-carboxylic acid moiety in the last step of camalexin biosynthesis.[37, 38]

RiPPs with a C-terminal azole heterocycle are very rare. The linear cyanobactins aeruginosamides and viridisamides contain a C-terminal thiazole-4-carboxylic acid methyl ester and are oxidized without decarboxylation.[39, 40] To the best of our knowledge, no other RiPP with a C-terminal decarboxylated thiazole has been discovered. BotCYP therefore expands the catalytic scope of RiPP CYP enzymes and is also one of the very few RiPP CYP enzymes that has been characterized. Other characterized RiPP CYP enzymes are not involved in heterocycle oxidation but catalyze the hydroxylation of amino acids, conversion of isoleucine to β-methyl-δ-hydroxy-proline, epoxidation, or - as recently reported - the formation of unusual biaryl (Tyr-His and Tyr-Tyr) and aryl-oxygen-aryl (Tyr-Tyr) linkages.[41-45]

Although the function of the bottromycin CYP enzyme had already been demonstrated *in vivo*, we aimed to reconstitute the function of BotCYP *in vitro* to characterize the enzyme in detail (chapter 5). A reductase-ferredoxin pair is not encoded in or near the bottromycin BGC. Since bottromycins can be expressed in heterologous hosts (e.g. *S. lividans*),[46] we aimed to use *S. lividans* lysate to reconstitute a functional redox system. This experiment did not yield the desired functional redox system, but we were able to find a P450 reductase-ferredoxin pair from *Bacillus megaterium* that was suitable for the bottromycin CYP enzyme.

In vitro reactions, using the pathway intermediate obtained from BotAH reactions as a substrate, yielded the expected intermediate with a C-terminal, decarboxylated thiazole. However, it was not possible to optimize the reaction towards complete turnover. Further analysis of the enzymatic reaction demonstrated that the CYP enzyme is specific for D-Asp containing intermediate. D-Asp is also found in the natural product itself despite the ribosomal origin of the substrate peptide. After follower cleavage by BotAH, the L-Asp spontaneously epimerizes slowly to D-Asp, albeit in small proportion and the epimerization was determined to be very slow (see Figure 4-S3).

Although D-configured amino acids are generally a hallmark of NRPs, several RiPPs have also been found to contain D-configured amino acids. Due to the ribosomal synthesis of the substrate peptide, which only allows the incorporation of L-configured amino acids, epimerization to D-amino acids has to occur post-translationally, either enzymatically or non-enzymatically. In polytheonamide and epipeptide biosynthesis, radical *S*-adenosyl-L-methionine (SAM) enzymes catalyze the epimerization of L- to D-amino acids.[47-49] In lanthipeptides, e.g. lactosin S, D-configured amino acids (D-Ala and D-butyric acid) are observed.[50, 51] They are produced by hydrogenation of dehydroalanine (Dha) or dehydrobutyrine (Dhb) moieties. Dha and Dhb themselves are formed by Ser and Thr dehydration (reductase).[52-54] However, these two mechanisms were not feasible in the bottromycin biosynthesis. Similarly to cyanobactins, it was proposed that the epimerization in bottromycin biosynthesis happens non-enzymatically. In cyanobactins, e.g. patellamide D and trunkamide A, epimerization of amino acids N-terminal to thiazole heterocycles is observed.[55-57] It has been proposed that epimerization occurs spontaneously (non-enzymatically), after thiazoline and macrocycle formation, and before thiazoline oxidation.[57-59] The proton of a C_α N-terminal to a thiazoline is more acidic due to the neighboring imine bond.[55] An increase in conformational stability of the molecules containing the D-configured amino acid, as shown for trunkamide A,[60] favors the D-configuration and may explain the gradually transformation into the D-configured molecule.[58]

NMR-based bottromycin A_2 structure determination in $CDCl_3$ (see Figure 8-2) showed that the C-terminal tail of bottromycins folds back onto the N-terminal macrocycle. This may represent the lowest-energy, native state, and was thought to be inaccessible with a L-configured Asp-OMe.[61] However, chemical synthesis of a bottromycin A_2 analog containing L-Asp-OMe instead of the D-Asp-OMe revealed that the non-natural epimer did not lose biological activity.[62] Another possibility is that epimerization increases resistance to proteolytic degradation,[63, 64] but the function of epimerization in bottromycin biosynthesis is still unclear. In chapter 4, we demonstrated that L-Asp epimerization in bottromycin biosynthesis is enzymatic. BotH, an unusual α/β-hydrolase-fold enzyme, catalyzes this reaction and provides the D-configured substrate for BotCYP (Figure 7-1). This expands the catalytic scope of α/β-hydrolase-fold enzymes in general and known epimerization mechanisms in RiPP biosynthesis in particular. Recently, a further enzyme responsible for amino acid epimerization in RiPPs has been discovered. MslH, which belongs to the metallo-dependent phosphatase family, catalyzes cofactor-independent epimerization of the C-terminal Trp in the biosynthesis of the lasso peptide MS-271.[65] Currently, the enzymatic mechanism behind this reaction is not understood. Further biosynthetic gene clusters (BGCs) of siamycin-like lasso peptides contain MslH homologs.[66] MslH has 52% identity to PgsA, which has recently been shown to catalyze the epimerization of L-Glu to D-Glu in poly-γ-glutamic acid (PGA) biosynthesis.[67] Sequence analysis and structure determination of the BotH-apo and -complex structures allowed us to propose its enzymatic mechanism (chapter 4) and understand the enzyme flexibility regarding amino acid changes in the core peptide (Figure 7-3).

We further propose that BotH may be involved in a biosynthetic feedback mechanism and aid self-immunity. BotH does not only bind its substrate but also bottromycin A_2 with high affinity. However, after oxidative decarboxylation of the thiazoline, BotH can no longer catalyze epimerization. Bottromycin A_2 acts as orthosteric inhibitor of epimerization, which may serve as a negative feedback mechanism to prevent self-poisoning. To study the self-immunity effect of BotH *in vivo*, the sensitivity of *S. coelicolur* M145 strains, expressing BtmH (BotH homolog), BtmA (BotT homolog, transporter) or BtmH and BmtA, to bottromycin A_2 were determined in a zone of inhibition assay (unpublished data, J. Santos-Aberturas, A.W. Truman). While expression of BtmH in *S. coelicolur* M145 had no significant effect on the resistance to bottromycin A_2, expression of BtmH and BtmA, led almost to complete resistance within the tested concentration range. The effect of BtmH and BtmA expression in combination was enhanced in comparison to the effect of BtmA alone. As mentioned in section 1.2.3, overexpression of the transporter BotT led to increased bottromycin production.[8] These results

suggest that BotH overexpression may increase bottromycin production levels, which are - despite several pathways engineering approaches - very low (< 1 mg/L).[8, 46, 68] As an effect of BotH on self-immunity was only observed when BotT was also expressed, it will be interesting to study the interaction of BotH and BotT.

In summary, the *in vitro* characterization of the bottromycin biosynthetic enzymes in this thesis contributed to the understanding of the bottromycin biosynthesis significantly and led to a revised bottromycin biosynthetic pathway (Figure 7-1). In addition, the *in vitro* reconstitution of the complete enzymatic cascade paved the way for the generation of new bottromycin analogs *in vitro* (section 7.2), but also facilitated new approaches for gene cluster engineering *in vivo*. Furthermore, novel enzymatic capabilities of two enzyme classes in general (YcaO enzymes and α/β-hydrolases) and one enzyme class in the context of RiPP (cytochrome P450 enzymes) were discovered and validated.

7.2 Generation of bottromycin analogss and pathway engineering

Even though bottromycins have shown some potent antimicrobial activities, their development towards a potential use as a drug was hampered as a result of their poor blood plasma stability due to methyl ester hydrolysis.[7, 69] Furthermore, decades of structure revisions until the final elucidation of the correct bottromycin structure impeded the development of a total synthesis route, which is still very difficult and inefficient.[70] In addition, *in vivo* engineering approaches suffer from very low production yields.[8, 46, 68] The insights into bottromycin biosynthesis acquired in this thesis may present a turning point and help to overcome these issues. They may allow the production of biologically active bottromycin analogs *in vitro*, and also contribute to pathway engineering and the production of analogs *in vivo*.

As described in section 1.3.2, the biosynthetic logic of RiPPs allows the easy generation of analogs by amino acid substitution(s) in the core peptide. To produce a RiPP analog with an amino acid substitution, all biosynthetic enzymes need to tolerate this change in the respective substrate. RiPP enzymes are often promiscuous towards changes in the core peptide due to the physical separation of substrate binding and catalysis and core peptides variations are often observed between homologous BGCs. The variation of amino acids is usually observed for amino acids that are not post-translationally modified. The bottromycin core peptide sequence is highly conserved in homologous bottromycin BGCs from several *Streptomyces* species. The only variation is observed for Val3. The precursor peptide genes in *S.* sp. WMMB272 and *S. sulphurous* DSM 40104 encode for an Ala at position 3, and the gene in *S. orinoci* encodes

for a Leu residue at this position. Val3 is the only amino acid in bottromycin A_2 that is not post-translationally modified. Thus, we were skeptically if bottromycin biosynthetic enzymes are promiscuous. Reconstitution of the biosynthetic enzymes *in vitro* allowed the analysis of each enzyme individually. Structural data of the enzymes with the respective pathway intermediate may serve as starting point for rational enzyme engineering to modify the substrate specificity of the enzymes.

25 precursor peptide variants were tested *in vitro* for their suitability for the enzymes IpoC, PurCD, PurAH and BotH. 13 could be processed by all four enzymes. The results of this investigation are summarized in Figure 7-3. From the 25 precursor peptides that were tested *in vitro*, 11 were also tested in *in vivo* studies.[71-73] The results of the *in vivo* mutagenesis studies, which are summarized in Figure 8-14, and our *in vitro* studies mostly agree. Only for the mutations Val3Leu, Val4Leu, and Val5Ala, which were accepted by the four tested enzymes studied *in vitro*, no bottromycin production was observed *in vivo*. While disparities in substrate tolerance *in vivo* and *in vitro* are not unusual, the substrate tolerances tested *in vitro* were only determined for four enzymes and the remaining biosynthetic enzymes may have halted bottromycin biosynthesis *in vivo*.
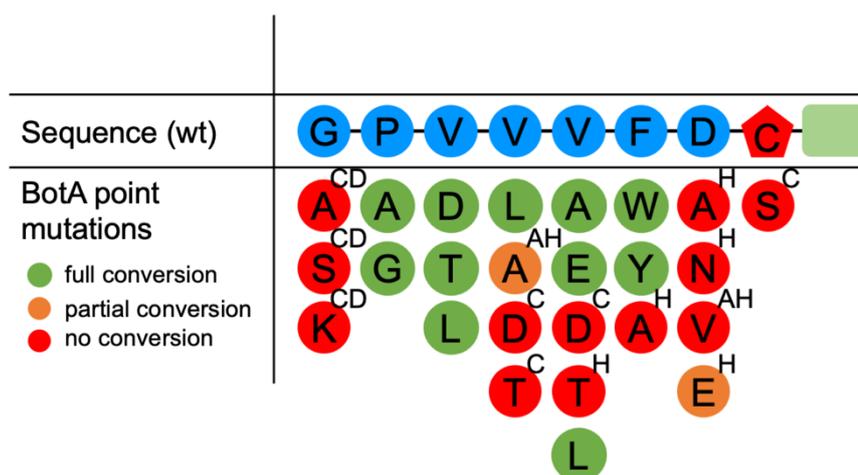
Figure 7-3 | Promiscuity of the enzymes IpoC, PurCD, PurAH and BotH towards changes in the core peptide sequence, tested in the course of this thesis. The enzyme that halted turnover is stated by the letter in superscript.

The two bottromycin YcaO enzymes were unable to catalyze cyclization reactions using a nucleophile different from Gly1 (macroamidine formation) and Cys8 (thiazoline formation), but changes in the macroamidine cycle and other positions were accepted. Unfortunately, there are no complex crystal structures available for these enzymes to understand the tolerance or intolerance to amino acids changes and to serve as a starting point for enzyme engineering to possibly expand the substrate tolerance of these enzymes.

Azoline-forming YcaOs often introduce several heterocycles into a RiPP. These YcaOs are often tolerant towards changes of amino acids neighboring the Cys/Ser/Thr residues they modify. This is to be expected since the substrate peptides are not simple repeats and thus contain different amino acids immediately adjoining the modified residue.[14] Some YcaOs are capable of using Cys, Ser and Thr side chains for heterocyclization reactions, thus introducing thiazoline as well as (methyl)oxazolines in a RiPP, as observed for PatD (patellamide biosynthesis).[59, 74] Furthermore, alterations of the relative and/or absolute positions of the Cys/Ser/Thr residues within the core peptide are accepted by certain YcaO enzymes. However, YcaOs prefer Cys over Ser/Thr residues, likely because of the higher nucleophilicity of Cys to Ser/Thr side chains.[14] One azoline forming YcaO, PatD, has even been demonstrated to form non-natural heterocycles, e.g. using homocysteine or 2,4-diaminobutanoic acid as a nucleophile to form 6-membered rings.[75, 76] However, other azoline forming YcaOs are not tolerant of changes of the nucleophile and are regioselective. This is nicely illustrated in nosiheptide biosynthesis, where 5 of the 6 core peptide Cys residues are heterocyclized. The Cys residue that is needed to form a macrocycle via its side-chain is skipped by the YcaO.[14]

Bottromycins however are one of the few reported cases, where the YcaO (BotC) is introducing only one heterocycle into the precursor peptide. BotC catalyzes the thiazoline formation in bottromycins using the Cys side chain as nucleophile. It is not able to catalyze oxazoline formation using Ser or Thr side chains as nucleophiles, and Cys residues at different positions in the core peptide are not cyclized (unpublished data). Irrespective of its high selectivity for the substrate residue, BotC accepts several changes of neighboring amino acids (Figure 7-3), which enables the production of bottromycin analogs. Unfortunately, *trans* activation of BotC by supplying the follower as a separate peptide, analogously to the lacticin 481 synthetase LctM,[77] was not successful (unpublished data) and thus such an approach to minimize the peptide synthesis cost (see section 1.3.2) is not applicable for bottromycin biosynthesis *in vitro*. IpoC showed poor stability and was very temperature sensitive. In addition, it does not tolerate all amino acid changes in the core peptide. Thus, it may be interesting to substitute the bottromycin azoline-forming YcaO *in vitro* (and *in vivo*) with a more promiscuous and more stable YcaO, e.g. the *cis*-activated LynD (LynD-leader fusion) construct. With that approach, the generation of bottromycin analogs would not be impeded by the substrate tolerance of IpoC. Oxazole-containing bottromycins may also be produced this way and problems with the poor stability of IpoC would be avoided.

The macroamidine-forming PurCD did not even accept minor amino acid changes (e.g. Gly1Ala), although the amino acid side chain is not directly involved in the amidine formation.

However, the unique flexibility of Gly compared to all other amino acids that contain a side chain may explain these results. The other currently identified amidine-forming YcaOs KlpD and AmiD use the amino group of Ser and His respectively, for amidine formation. No studies reporting on the ability of KlpD and AmiD to use other amino acids for amidine formation have been published. KlpD, which also introduces the azoline heterocycles in Klebsazolicin, may be more promiscuous than AmiD due to its bifunctional role. BotCD was not able to process synthetic precursor peptides in which the N-terminal glycine was replaced by either glycolic or thioglycolic acid (unpublished data). The hydroxyl group of glycolic acid is less nucleophilic than the glycine amino group, but the thiol group of thioglycolic acid is more nucleophilic. Hence, in this case the inability to use these substates cannot be explained by the nucleophilicity but with the substrate specificity.

BotCD is one of the most interesting enzymes of the pathway and may also be of interest for chemoenzymatic applications. Although PurCD accepts several changes in the core peptide, it is specific for azoline-containing intermediates, follower-dependent and due to its reverse reaction needs to be coupled with a macroamidine-specific follower-protease (BotAH). The azoline heterocycle could also not be mimicked by proline (unpublished data). These aspects make PurCD, without further enzyme engineering, unattractive for the generation of macroamidine peptide libraries exceeding the context of bottromycin. The macroamidine-forming YcaO is also specific to only form 12-membered macroamidine cycles. Different ring sizes as observed for other RiPP macrocycles (PatG$_{mac}$ or PCY1) are not possible.[78-80] Determination of the crystal structure with its substrate will be important to understand the macroamidine formation and enable enzyme engineering to create a more promiscuous enzyme.

For PurAH, a BotAH homolog responsible for cleavage of the follower peptide, a crystal structure was determined (chapter 3). PurAH contains a huge, highly conserved binding pocket and was determined to be specific for the heterocyclized and macrocyclized pathway intermediate. Modeling of the modified core peptide into the active site (unpublished data, O. Melse and I. Antes), confirmed a binding pocket suitable for binding the macrocyclized peptide, while the linear peptide did not fit into the binding pocket. The macrocycle, Val5 and Phe6 form hydrophobic interactions with PurAH. Enzyme and substrate mutations agree with our proposed mechanism and turnover of the peptide mutants Val4Ala, Val5Glu and Phe6Ala was reduced as the substrate-enzyme interactions are impaired for those mutants. However, almost all tested peptide mutants can be processed, albeit (much) slower (chapter 3).

BotH is the only enzyme in this thesis for which a complex crystal structure was obtained. Combining the structural information with biochemical results, we proposed an enzymatic mechanism for epimerization of Asp7 (chapter 4). The substrate-assisted mechanism of BotH is inconvenient for the development of bottromycin analogs by biosynthesis (*in vivo* and *in vitro*) rather than total synthesis, as a D-configured intermediate is necessary for processing with BotCYP. However, it was not tested if BotCYP would accept other D-configured amino acids besides D-Asp. Preliminary results using a synthetic peptide with D-Ala at position 7, indicate that D-configured amino acids are not accepted by the azoline-forming YcaO. For the further development of bottromycins, the generation of analogs altered at position 7 (Asp-OMe) are of high interest to increase *in vivo* stability. The lability of the methyl ester in physiological condition leads to the inactivity of bottromycins *in vivo*, while the *in vitro* biological activities of this compound were promising (see chapter 8). As BotH only accepts Asp or Glu at this position and enzymes that must act before BotH (e.g. BotC) probably do not accept D-configured amino acids, the variation of Asp7 is not feasible using the biosynthetic path. However, the Asp7-OMe may be derivatized chemically after biosynthesis, to produce analogs that are both highly active and stable in blood plasma.[7, 81]

Testing the substrate promiscuity of the enzymes *in vitro* demonstrated that the bottromycin forming enzymes are promiscuous. However, the positions Gly1, Cys8 and Asp7 cannot be modified, except for replacement of Asp7 with Glu. Bottromycin analogs with amino acids altered at positions 2 to 6 can be produced *in vivo* and *in vitro*, but to obtain bottromycin analogs that are stable in plasma a semi-synthetic approach is necessary. As bottromycin total synthesis remains highly ineffective, such approaches for bottromycin compound development are more attractive than complete chemical synthesis. Currently, the bottleneck of semi-synthetic approaches remains the very low production levels *in vivo*. With the *in vitro* reconstitution of the biosynthetic steps leading to the bottromycin core structure presented in this thesis, *in vitro* production of active and stable bottromycins for subsequent semi-synthetic modification came into reach. All biosynthetic enzymes except the RMTs and the O-methyltransferase (OMT) were reconstituted *in vitro*. The RMTs are not necessarily required, as methylated amino acids can be used in chemical peptide synthesis instead of enzymatic β-methylation. The same holds true for the OMT, as the Asp7 carboxy group can be esterified or otherwise modified (see section 8.5) synthetically to yield active bottromycins. Thus, it is possible to produce active bottromycins with the *in vitro* pathway presented in this thesis by using a synthetic precursor peptide, with β-methylated amino acids and post-biosynthetic esterification of Asp7.

It may also be feasible, to use a semi-synthetic approach that only utilizes the two bottromycin YcaO enzymes. In a bottromycin SAR study from Yamada *et al.*, the importance of the thia-β-Ala-OMe was determined.[62] For this study they established a synthesis starting with a bottromycin fragment that represents a bottromycin in which the peptide bond between Phe6-Asp7 was hydrolyzed. They synthesized bottromycin and the analogs by condensation of this fragment with amines. Enzymatic macrocyclization of a peptide with a hydrolysable ester bond instead of an amide bond between Phe6 and Asp7 and subsequent chemical hydrolysis would produce the fragment used by Yamada *et al.*. Using this proposed semi-synthetic strategy, the elaborate and inefficient chemical synthesis of the macroamidine can be evaded, while only two enzymes and their limitations in the substrate tolerance have to be considered.

However, these semi-synthetic *in vitro* approaches require peptides with un-natural amino acids (e.g. β-methylations) that cannot be produced *in vivo*. For the generation of bottromycin peptide libraries it would be of interest to reduce the effort for peptide synthesis. Generally, the five C-terminal amino acids of the bottromycin follower peptide can be truncated since they are not required for enzyme binding (see chapter 2). Furthermore, there are two feasible strategies to reduce the synthesis effort; Firstly, instead of synthesizing a library of full length bottromycin precursor peptides, it is possible to just produce a library of GPVVVFD variants synthetically and then obtain full length bottromycin precursor peptides by native chemical ligation (NCL) with the follower peptide that has N-terminal Cys. The feasibility of this approach was tested during this PhD project.

The second approach to produce bottromycin analogs, uses the flexible *in vitro* translation (FIT) system established from the Suga group (see section 1.3.2). This approach would allow the DNA-templated synthesis of precursor peptides (with unnatural amino acids) *in vitro* and subsequent production of bottromycin analogs using purified bottromycin biosynthetic enzymes.

In summary, the determination of the enzyme substrate tolerances demonstrated the feasibility to produce bottromycin analogs *in vivo* and *in vitro*, as changes at several core peptide positions were accepted. However, combining structural and biochemical data also led to an understanding of why some amino acid substitutions are not accepted, e.g. as these amino acids are important for binding or are involved in the enzymatic mechanism. The positions Gly1, Asp7 and Cys8 were determined to be not alterable. The *in vitro* reconstitution of the biosynthetic steps leading to the production of the bottromycin core scaffold enables the

production of active bottromycin (analogs) using synthetic precursor peptides and post-biosynthetic derivatization of Asp7.

7.3 New-to-nature RiPPs by sortase-mediated leader peptide exchange

The aim of combinatorial biosynthesis is to produce new-to-nature natural product (NP)-like molecules with desired properties and biological activities. This is achieved by combining enzymes of different biosynthetic pathways (section 1.3). RiPPs are a structurally very diverse class of natural products. They owe their structural diversity to the manifold post-translational modifications (PTMs) introduced by the RiPP biosynthetic enzymes. Primary RiPP biosynthetic enzymes bind to specific recognition sequences (RS) in the leader or follower peptide (section 1.2.2), but they introduce the PTM in the core peptide sequence. Recently an approach was introduced that allows the modification of precursor peptides by two RiPP enzymes from unrelated RiPP pathways using chimeric leader peptides, which are formed by combination of RSs of the respective enzymes (section 1.3.2).[82] In chapter 6 we presented a novel strategy to produce new-to-nature ribosomal peptides, utilizing sortase A-mediated transpeptidation for leader peptide exchange (LPX), which enables the modification of one core peptide by RiPP enzymes from unrelated pathways. Compared to the chimeric leader peptide strategy, this strategy does not require an understanding of enzyme binding to RSs as the whole leader peptide sequence can be exchanged. The LPX strategy facilitates the generation of RiPP-inspired natural product libraries for drug discovery in a plug-and-play system. A further advantage of the LPX strategy is the unlimited number of exchanges and enzymes that could theoretically be used, as the N-terminus of the hybrid core peptide, di-glycin, which is required for the sortase-mediated ligation, is recovered after sortase-mediated cleavage. The sortase approach does not require any chemical modification of the peptides or synthetic linkers, thus the LPX strategy can also be applied *in vivo*. However, it has to be mentioned that the sortase reaction is reversible and thus cannot provide full turnover when set up as a one pot reaction. The one pot LPX strategy is a promising method to produce novel RiPP-like molecules in high throughput to explore the combinatorial potential of RiPP pathways. In a simple assay looking for initial 'hits' a mixture of products may even be desirable. Compound production may be achieved by modification of the LPX method. Efficient turnover and production of a single product can be achieved by addition of e.g. primary amines or hydrazines, and purification of the core peptide intermediates after each round of modification.[83]

In its current form the LPX method has limited applicability for RiPP modifications that include N-terminal cyclizations, e.g. in thiopeptides, as this strategy requires a N-terminal di-glycine for sortase-mediated transpeptidation. However, this problem can be solved by introduction of a pathway specific protease site or a lysine residue (for trypsin cleavage) between the di-glycine motif and the core peptide to generate the N-terminus required for cyclisation.

The leader peptide exchange (LPX) strategy is a promising approach to produce hybrid RiPPs in a plug-and-play system. The method is broadly applicable and does not require a detailed understanding of the enzyme substrate recognition. This method may unlock the RiPP biosynthetic potential to develop novel RiPP-like drugs.

7.4 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.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* **2013,** *30* (1), 108-60.

2.   Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A., New developments in RiPP discovery, enzymology and engineering. *Natural Product Reports* **2020**.

3.   Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Ōmura, S., Structure Determination and Total Synthesis of Bottromycin A2: A Potent Antibiotic against MRSA and VRE. *Angewandte Chemie International Edition* **2009,** *48* (5), 914-917.

4.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2. Release of aminoacyl or peptidyl tRNA from ribosomes. *The Journal of biological chemistry* **1976,** *251*, 2299-306.

5.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2: Effect of bottromycin A2 on polysomes. *FEBS Letters* **1983,** *153* (1), 53-59.

6.   Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect on peptide bond formation. *FEBS Letters* **1981,** *123* (2), 173-176.

7.   Kobayashi, Y.; Ichioka, M.; Hirose, T.; Nagai, K.; Matsumoto, A.; Matsui, H.; Hanaki, H.; Masuma, R.; Takahashi, Y.; Omura, S.; Sunazuka, T., Bottromycin derivatives: efficient chemical modifications of the ester moiety and evaluation of anti-MRSA and anti-VRE activities. *Bioorganic & Medicinal Chemistry Letters* **2010,** *20* (20), 6116-6120.

8.   Huo, L.; Rachid, S.; Stadler, M.; Wenzel, Silke C.; Müller, R., Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis. *Chemistry & Biology* **2012,** *19* (10), 1278-1287.

9.   Crone, W. J. K.; Leeper, F. J.; Truman, A. W., Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chemical Science* **2012,** *3* (12), 3516-3521.

10.   Hou, Y.; Tianero, M. D. B.; 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., Structure and Biosynthesis of the Antibiotic Bottromycin D. *Organic Letters* **2012,** *14* (19), 5050-5053.

11.   Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chemical Science* **2012,** *3* (12), 3522-3525.

12.   Crone, W. J.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angew Chem Int Ed Engl* **2016,** *55* (33), 9639-43.

13.   Dunbar, K. L.; Tietz, J. I.; Cox, C. L.; Burkhart, B. J.; Mitchell, D. A., Identification of an Auxiliary Leader Peptide-Binding Protein Required for Azoline Formation in Ribosomal Natural Products. *J Am Chem Soc* **2015,** *137* (24), 7672-7.

14.   Burkhart, B. J.; Schwalen, C. J.; Mann, G.; Naismith, J. H.; Mitchell, D. A., YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function. *Chem Rev* **2017,** *117* (8), 5389-5456.

15.   Dunbar, K. L.; Chekan, J. R.; Cox, C. L.; Burkhart, B. J.; Nair, S. K.; Mitchell, D. A., Discovery of a new ATP-binding motif involved in peptidic azoline biosynthesis. *Nat Chem Biol* **2014,** *10* (10), 823-9.

16.   Nayak, D. D.; Mahanta, N.; Mitchell, D. A.; Metcalf, W. W., Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea. *eLife* **2017,** *6*, e29218.

17.   Santos-Aberturas, J.; Chandra, G.; Frattaruolo, L.; Lacret, R.; Pham, T. H.; Vior, N. M.; Eyles, T. H.; Truman, A. W., Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. *Nucleic Acids Res* **2019,** *47* (9), 4624-4637.

18.   Frattaruolo, L.; Lacret, R.; Cappello, A. R.; Truman, A. W., A Genomics-Based Approach Identifies a Thioviridamide-Like Compound with Selective Anticancer Activity. *ACS Chemical Biology* **2017,** *12* (11), 2815-2822.

19.   Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A., Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proceedings of the National Academy of Sciences* **2018,** *115* (12), 3030-3035.

20.   Dong, S.-H.; Liu, A.; Mahanta, N.; Mitchell, D. A.; Nair, S. K., Mechanistic Basis for Ribosomal Peptide Backbone Modifications. *ACS Cent Sci* **2019,** *5* (5), 842-851.

21.   Liu, J.; Lin, Z.; Li, Y.; Zheng, Q.; Chen, D.; Liu, W., Insights into the thioamidation of thiopeptins to enhance the understanding of the biosynthetic logic of thioamide-containing thiopeptides. *Organic & Biomolecular Chemistry* **2019,** *17* (15), 3727-3731.

22.   Mahanta, N.; Szantai-Kis, D. M.; Petersson, E. J.; Mitchell, D. A., Biosynthesis and Chemical Applications of Thioamides. *ACS chemical biology* **2019,** *14* (2), 142-163.

23.   Eyles, T.; Vior, N.; Lacret, R.; Truman, A., Understanding Thioamitide Biosynthesis Using Pathway Engineering and Untargeted Metabolomics. bioRxiv: 2020.

24.   Watson, Z. L.; Ward, F. R.; Méheust, R.; Ad, O.; Schepartz, A.; Banfield, J. F.; Cate, J. H. D., Structure of the bacterial ribosome at 2 Å resolution. *eLife* **2020,** *9*, e60482.

25.   Strader, M. B.; Costantino, N.; Elkins, C. A.; Chen, C. Y.; Patel, I.; Makusky, A. J.; Choy, J. S.; Court, D. L.; Markey, S. P.; Kowalak, J. A., A proteomic and transcriptomic approach reveals new insight into beta-methylthiolation of Escherichia coli ribosomal protein S12. *Mol Cell Proteomics* **2011,** *10* (3), M110.005199.

26.   Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H. C.; Zakai, U. I.; Mitchell, D. A., A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol* **2017,** *13* (5), 470-478.

27.   Schwalen, C. J.; Hudson, G. A.; Kosol, S.; Mahanta, N.; Challis, G. L.; Mitchell, D. A., In Vitro Biosynthetic Studies of Bottromycin Expand the Enzymatic Capabilities of the YcaO Superfamily. *Journal of the American Chemical Society* **2017,** *139* (50), 18154-18157.

28.   Travin, D. Y.; Metelev, M.; Serebryakova, M.; Komarova, E. S.; Osterman, I. A.; Ghilarov, D.; Severinov, K., Biosynthesis of Translation Inhibitor Klebsazolicin Proceeds through Heterocyclization and N-Terminal Amidine Formation Catalyzed by a Single YcaO Enzyme. *Journal of the American Chemical Society* **2018,** *140* (16), 5625-5633.

29.   Russell, A.; Vior, N.; Hems, E.; Lacret, R.; Truman, A., Discovery and characterisation of an amidine-containing ribosomally-synthesised peptide that is widely distributed in nature. bioRxiv: 2020.

30.   Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W.; Jaspars, M.; Naismith, J. H., Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat Chem Biol* **2015,** *11* (8), 558-563.

31.   Ghilarov, D.; Stevenson, C.; Travin, D.; Piskunova, J.; Serebryakova, M.; Maxwell, A.; Lawson, D.; Severinov, K., Architecture of Microcin B17 Synthetase: An Octameric Protein Complex Converting a Ribosomally Synthesized Peptide into a DNA Gyrase Poison. *Molecular Cell* **2019,** *73*.

32.   Nakamura, S.; Yajima, T.; Lin, Y.; Umezawa, H., Isolation and characterization of bottromycins A2, B2, C2. *J Antibiot (Tokyo)* **1967,** *20* (1), 1-5.

33.   Crone, W. J. K.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angewandte Chemie International Edition* **2016,** *55* (33), 9639-9643.

34.   Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Müller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *ChemBioChem* **2016,** *17* (23), 2286-2292.

35. Vior, N. M.; Cea-Torrescassana, E.; Eyles, T. H.; Chandra, G.; Truman, A. W., Regulation of Bottromycin Biosynthesis Involves an Internal Transcriptional Start Site and a Cluster-Situated Modulator. *Frontiers in Microbiology* **2020,** *11* (495).

36. Dunbar, K. L.; Mitchell, D. A., Insights into the mechanism of peptide cyclodehydrations achieved through the chemoenzymatic generation of amide derivatives. *Journal of the American Chemical Society* **2013,** *135* (23), 8692-8701.

37. Böttcher, C.; Westphal, L.; Schmotz, C.; Prade, E.; Scheel, D.; Glawischnig, E., The Multifunctional Enzyme CYP71B15 (PHYTOALEXIN DEFICIENT3) Converts Cysteine-Indole-3-Acetonitrile to Camalexin in the Indole-3-Acetonitrile Metabolic Network of *Arabidopsis thaliana*. *The Plant Cell* **2009,** *21* (6), 1830-1845.

38. Schuhegger, R.; Nafisi, M.; Mansourova, M.; Petersen, B. L.; Olsen, C. E.; Svatoš, A.; Halkier, B. A.; Glawischnig, E., CYP71B15 (PAD3) Catalyzes the Final Step in Camalexin Biosynthesis. *Plant Physiology* **2006,** *141* (4), 1248-1254.

39. Martins, J.; Vasconcelos, V., Cyanobactins from Cyanobacteria: Current Genetic and Chemical State of Knowledge. *Mar Drugs* **2015,** *13* (11), 6910-6946.

40. Sardar, D.; Pierce, E.; McIntosh, J. A.; Schmidt, E. W., Recognition Sequences and Substrate Evolution in Cyanobactin Biosynthesis. *ACS Synthetic Biology* **2015,** *4* (2), 167-176.

41. Young, T. S.; Walsh, C. T., Identification of the thiazolyl peptide GE37468 gene cluster from *Streptomyces* ATCC 55365 and heterologous expression in *Streptomyces lividans*. *Proceedings of the National Academy of Sciences* **2011,** *108* (32), 13053-13058.

42. Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A., In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *Journal of the American Chemical Society* **2015,** *137* (51), 16012-16015.

43. Zheng, Q.; Wang, S.; Liao, R.; Liu, W., Precursor-Directed Mutational Biosynthesis Facilitates the Functional Assignment of Two Cytochromes P450 in Thiostrepton Biosynthesis. *ACS Chem Biol* **2016,** *11* (10), 2673-2678.

44. Hug, J. J.; Dastbaz, J.; Adam, S.; Revermann, O.; Koehnke, J.; Krug, D.; Müller, R., Biosynthesis of Cittilins, Unusual Ribosomally Synthesized and Post-translationally Modified Peptides from Myxococcus xanthus. *ACS Chemical Biology* **2020,** *15* (8), 2221-2231.

45. Zdouc, M. M.; Alanjary, M. M.; Zarazúa, G. S.; Maffioli, S. I.; Crüsemann, M.; Medema, M. H.; Donadio, S.; Sosio, M., A biaryl-linked tripeptide from Planomonospora reveals a widespread class of minimal RiPP gene clusters. *Cell Chemical Biology* **2020**.

46. Horbal, L.; Marques, F.; Nadmid, S.; Mendes, M. V.; Luzhetskyy, A., Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metabolic Engineering* **2018,** *49*, 299-315.

47. Morinaka, B. I.; Vagstad, A. L.; Helf, M. J.; Gugger, M.; Kegler, C.; Freeman, M. F.; Bode, H. B.; Piel, J., Radical S-Adenosyl Methionine Epimerases: Regioselective

Introduction of Diverse D-Amino Acid Patterns into Peptide Natural Products. *Angewandte Chemie International Edition* **2014,** *53* (32), 8503-8507.

48. Freeman, M. F.; Gurgui, C.; Helf, M. J.; Morinaka, B. I.; Uria, A. R.; Oldham, N. J.; Sahl, H.-G.; Matsunaga, S.; Piel, J., Metagenome Mining Reveals Polytheonamides as Posttranslationally Modified Ribosomal Peptides. *Science* **2012,** *338* (6105), 387-390.

49. Benjdia, A.; Guillot, A.; Ruffié, P.; Leprince, J.; Berteau, O., Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in Bacillus subtilis. *Nature Chemistry* **2017,** *9* (7), 698-707.

50. Skaugen, M.; Nissen-Meyer, J.; Jung, G.; Stevanovic, S.; Sletten, K.; Inger, C.; Abildgaard, M.; Nes, I. F., In vivo conversion of L-serine to D-alanine in a ribosomally synthesized polypeptide. *J Biol Chem* **1994,** *269* (44), 27183-5.

51. Lohans, C. T.; Li, J. L.; Vederas, J. C., Structure and biosynthesis of carnolysin, a homologue of enterococcal cytolysin with D-amino acids. *J Am Chem Soc* **2014,** *136* (38), 13150-3.

52. Yang, X.; van der Donk, W. A., Post-translational Introduction of d-Alanine into Ribosomally Synthesized Peptides by the Dehydroalanine Reductase NpnJ. *Journal of the American Chemical Society* **2015,** *137* (39), 12426-12429.

53. Huo, L.; van der Donk, W. A., Discovery and Characterization of Bicereucin, an Unusual d-Amino Acid-Containing Mixed Two-Component Lantibiotic. *Journal of the American Chemical Society* **2016,** *138* (16), 5254-5257.

54. Xu, M.; Zhang, F.; Cheng, Z.; Bashiri, G.; Wang, J.; Hong, J.; Wang, Y.; Xu, L.; Chen, X.; Huang, S.-X.; Lin, S.; Deng, Z.; Tao, M., Functional Genome Mining Reveals a Class V Lanthipeptide Containing a d-Amino Acid Introduced by an F420H2-Dependent Reductase. *Angewandte Chemie International Edition* **2020,** *59* (41), 18029-18035.

55. Milne, B. F.; Long, P. F.; Starcevic, A.; Hranueli, D.; Jaspars, M., Spontaneity in the patellamide biosynthetic pathway. *Org Biomol Chem* **2006,** *4* (4), 631-8.

56. Schmitz, F. J.; Ksebati, M. B.; Chang, J. S.; Wang, J. L.; Hossain, M. B.; Van der Helm, D.; Engel, M. H.; Serban, A.; Silfer, J. A., Cyclic peptides from the ascidian Lissoclinum patella: conformational analysis of patellamide D by x-ray analysis and molecular modeling. *The Journal of Organic Chemistry* **1989,** *54* (14), 3463-3472.

57. Wipf, P.; Uto, Y., Total Synthesis and Revision of Stereochemistry of the Marine Metabolite Trunkamide A. *The Journal of Organic Chemistry* **2000,** *65* (4), 1037-1049.

58. Gu, W.; Dong, S.-H.; Sarkar, S.; Nair, S. K.; Schmidt, E. W., The Biochemistry and Structural Biology of Cyanobactin Pathways: Enabling Combinatorial Biosynthesis. *Methods Enzymol* **2018,** *604*, 113-163.

59. Houssen, W. E.; Bent, A. F.; McEwan, A. R.; Pieiller, N.; Tabudravu, J.; Koehnke, J.; Mann, G.; Adaba, R. I.; Thomas, L.; Hawas, U. W.; Liu, H.; Schwarz-Linek, U.; Smith, M. C.; Naismith, J. H.; Jaspars, M., An efficient method for the in vitro production of azol(in)e-based cyclic peptides. *Angew Chem Int Ed Engl* **2014,** *53* (51), 14171-4.

60. Salvatella, X.; Caba, J. M.; Albericio, F.; Giralt, E., Solution Structure of the Antitumor Candidate Trunkamide A by 2D NMR and Restrained Simulated Annealing Methods. *The Journal of Organic Chemistry* **2003,** *68* (2), 211-215.

61. Gouda, H.; Kobayashi, Y.; Yamada, T.; Ideguchi, T.; Sugawara, A.; Hirose, T.; Omura, S.; Sunazuka, T.; Hirono, S., Three-dimensional solution structure of bottromycin A2: a potent antibiotic active against methicillin-resistant Staphylococcus aureus and vancomycin-resistant Enterococci. *Chem Pharm Bull (Tokyo)* **2012,** *60* (2), 169-71.

62. Yamada, T.; Yagita, M.; Kobayashi, Y.; Sennari, G.; Shimamura, H.; Matsui, H.; Horimatsu, Y.; Hanaki, H.; Hirose, T.; S, O. M.; Sunazuka, T., Synthesis and Evaluation of Antibacterial Activity of Bottromycins. *J Org Chem* **2018,** *83* (13), 7135-7149.

63. Jungheim, L. N.; Shepherd, T. A.; Baxter, A. J.; Burgess, J.; Hatch, S. D.; Lubbehusen, P.; Wiskerchen, M.; Muesing, M. A., Potent human immunodeficiency virus type 1 protease inhibitors that utilize noncoded D-amino acids as P2/P3 ligands. *J Med Chem* **1996,** *39* (1), 96-108.

64. Rink, R.; Arkema-Meter, A.; Baudoin, I.; Post, E.; Kuipers, A.; Nelemans, S. A.; Akanbi, M. H.; Moll, G. N., To protect peptide pharmaceuticals against peptidases. *J Pharmacol Toxicol Methods* **2010,** *61* (2), 210-8.

65. Feng, Z.; Ogasawara, Y.; Nomura, S.; Dairi, T., Biosynthetic Gene Cluster of a d-Tryptophan-Containing Lasso Peptide, MS-271. *Chembiochem* **2018,** *19* (19), 2045-2048.

66. Sánchez-Hidalgo, M.; Martín, J.; Genilloud, O., Identification and Heterologous Expression of the Biosynthetic Gene Cluster Encoding the Lasso Peptide Humidimycin, a Caspofungin Activity Potentiator. *Antibiotics (Basel)* **2020,** *9* (2).

67. Ogasawara, Y.; Shigematsu, M.; Sato, S.; Kato, H.; Dairi, T., Involvement of Peptide Epimerization in Poly-γ-glutamic Acid Biosynthesis. *Organic Letters* **2019,** *21* (11), 3972-3975.

68. Eyles, T. H.; Vior, N. M.; Truman, A. W., Rapid and Robust Yeast-Mediated Pathway Refactoring Generates Multiple New Bottromycin-Related Metabolites. *ACS Synth Biol* **2018,** *7* (5), 1211-1218.

69. Waisvisz, J. M.; van der Hoeven, M. G., The Chemistry and Partial Structure of Bottromycin. IV. *Journal of the American Chemical Society* **1958,** *80*, 383-385.

70. Kazmaier, U., The Long, Long Way to Bottromycin. *Israel Journal of Chemistry n/a* (n/a).

71. Horbal, L.; Marques, F.; Nadmid, S.; Mendes, M. V.; Luzhetskyy, A., Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metab Eng* **2018,** *49*, 299-315.

72. 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., Structure and biosynthesis of the antibiotic bottromycin D. *Org Lett* **2012,** *14* (19), 5050-3.

73. Crone, W. J. K. Investigations into the biosynthesis of bottromycin University of Cambridge Cambridge, 2015.

74. McIntosh, J. A.; Donia, M. S.; Schmidt, E. W., Insights into Heterocyclization from Two Highly Similar Enzymes. *Journal of the American Chemical Society* **2010,** *132* (12), 4089-4091.

75. Goto, Y.; Suga, H., In Vitro Biosynthesis of Peptides Containing Exotic Azoline Analogues. *ChemBioChem* **2020,** *21* (1-2), 84-87.

76. 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., An enzymatic route to selenazolines. *Chembiochem* **2013,** *14* (5), 564-7.

77. Levengood, M. R.; Patton, G. C.; van der Donk, W. A., The leader peptide is not required for post-translational modification by lacticin 481 synthetase. *Journal of the American Chemical Society* **2007,** *129* (34), 10314-10315.

78. Ludewig, H.; Czekster, C.; Oueis, E.; Munday, E.; Arshad, M.; Synowsky, S.; Bent, A.; Naismith, J., Characterization of the Fast and Promiscuous Macrocyclase from Plant PCY1 Enables the Use of Simple Substrates. *ACS Chemical Biology* **2018,** *13*.

79. Oueis, E.; Stevenson, H.; Jaspars, M.; Westwood, N.; Naismith, J., Bypassing the proline/thiazoline requirement of the macrocyclase PatG. *Chemical Communications* **2017,** *53*.

80. Oueis, E.; Nardone, B.; Jaspars, M.; Westwood, N.; Naismith, J., Synthesis of Hybrid Cyclopeptides through Enzymatic Macrocyclization. *ChemistryOpen* **2017,** *6*, 11 - 14.

81. Lerchen, H.-G.; Schiffer, G.; Broetz-Oesterhelt, H.; Mayer-Bartschmid, A.; Eckermann, S.; Freiberg, C.; Endermann, R.; Schuhmacher, J.; Meier, H.; Svenstrup, N.; Seip, S.; Gehling, M.; Haebich, D. Fermentative and chemical synthesis of bottromycin derivative analogs for use in treatment or prevention of bacterial infections in humans or animals. WO2006103010A1, 2006.

82. Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A., Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Cent Sci* **2017,** *3* (6), 629-638.

83. Dai, X.; Böker, A.; Glebe, U., Broadening the scope of sortagging. *RSC Advances* **2019,** *9* (9), 4700-4721.

Appendix

Chapter 8

Bottromycins - Biosynthesis, Synthesis and Activity

**Laura Franz**, Uli Kazmaier, Andrew W. Truman, and Jesko Koehnke

8 Bottromycins-Biosynthesis, Synthesis and Activity

8.1 Abstract

Bottromycins are a class of macrocyclic peptide natural products that are produced by several *Streptomyces* species and possess promising antibacterial activity against clinically relevant multidrug-resistant pathogens. They belong to the ribosomally synthesised and post-translationally modified peptide (RiPP) superfamily of natural products. The structure contains a unique four-amino acid macrocycle formed via a rare amidine linkage, *C*-methylation and a D-amino acid. This review covers all aspects of bottromycin research with a focus on recent years (2009-2020), in which major advances in total synthesis and understanding of bottromycin biosynthesis were achieved.
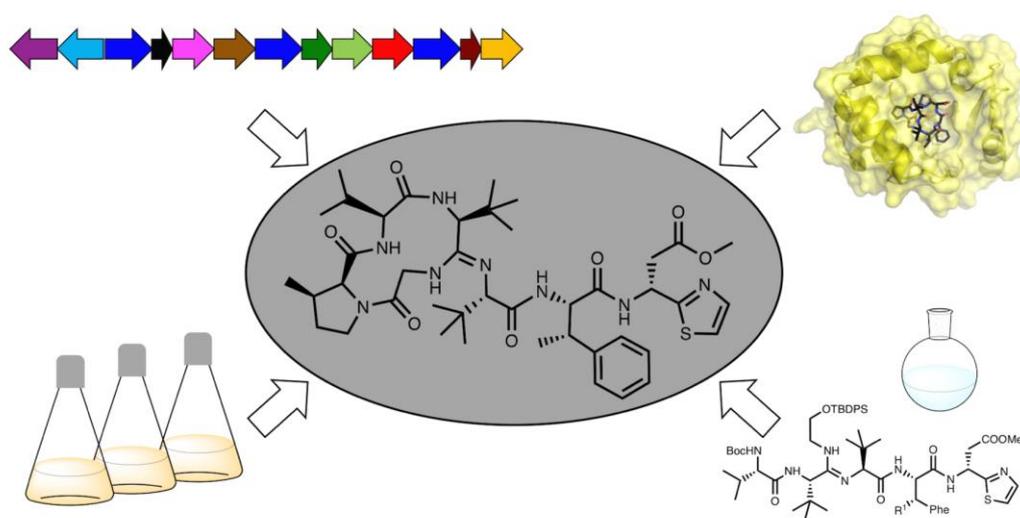
Figure 8-1 | Graphical abstract.
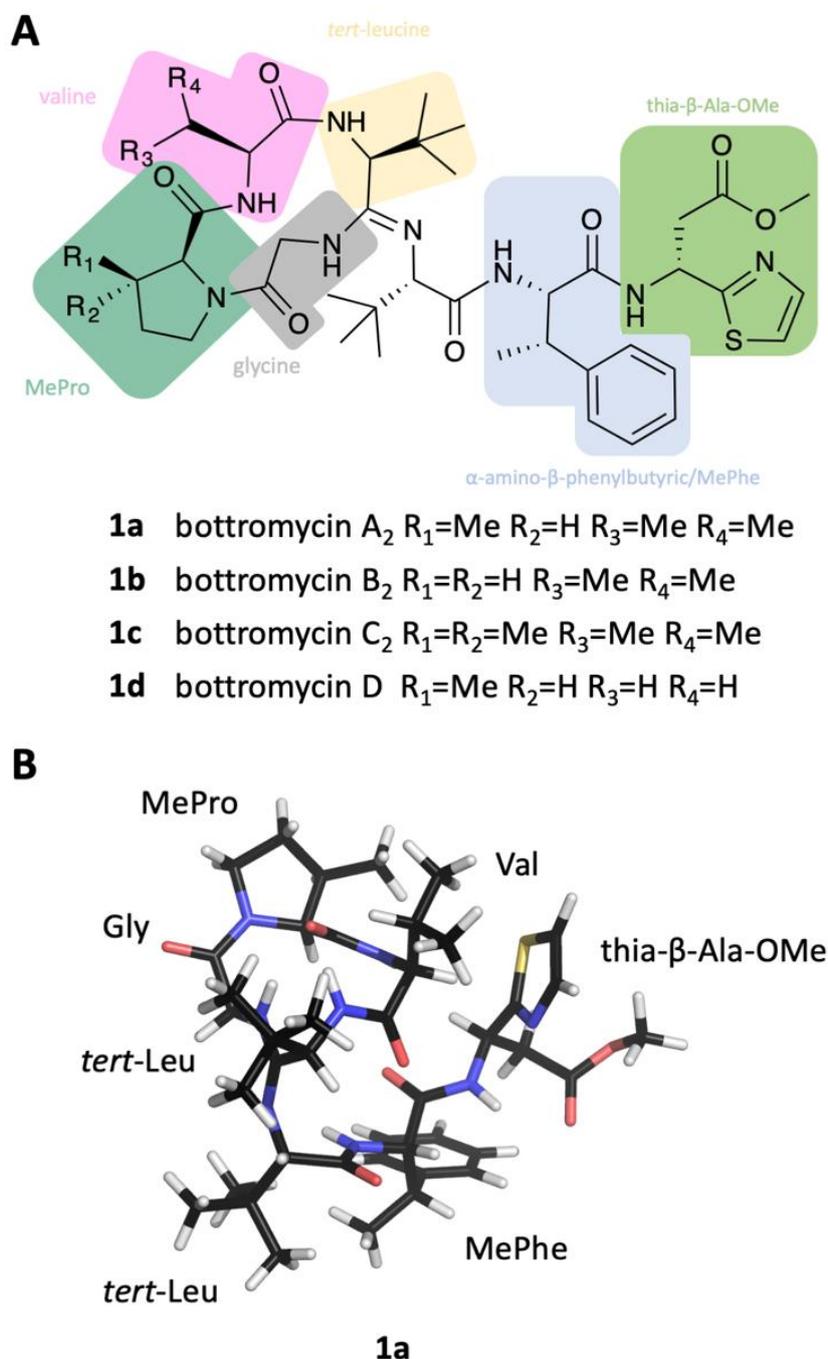
8.2 Introduction

Bottromycins were originally isolated from the fermentation broth of *Streptomyces bottropensis* in the late 1950s and described as peptidic natural products with antibacterial activity against Gram-positive pathogens.[1, 2] A series of mechanism of action (MoA) studies revealed that bottromycins inhibit protein synthesis by binding to the aminoacyl-tRNA binding site (A-site) of the prokaryotic 50S ribosome.[3-5] The bottromycin target site is currently not addressed by any antibiotic used in the clinic, which makes cross-resistance unlikely. As a result, bottromycins are effective against the problematic human pathogens methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE).

The chemical structure of bottromycins proved very difficult to elucidate and consequently underwent several revisions that ultimately led to the assignment of **1**. It showed that bottromycins are highly modified heptapeptides that are comprised of an *N*-terminal, four-amino acid macrocycle formed via a unique amidine linkage, several *C*-methylated residues, D-aspartate and a *C*-terminal thiazole (Figure 8-2A). This structure was confirmed by total synthesis in 2009.[6] Shortly after the successful total synthesis, several groups reported the discovery of the bottromycin biosynthetic gene cluster (BGC), which revealed that they are unusual ribosomally synthesized and post-translationally modified peptides (RiPPs).[7-10]

The emerging RiPP superfamily encompasses highly diverse molecules with interesting bioactivities.[11] The unifying feature is their biosynthetic logic: a short structural gene is expressed and yields the precursor peptide, which consists of one or more core peptide(s) (the eventual natural product(s)) and an *N*-terminal leader or *C*-terminal follower peptide that is important for recognition of the precursor peptide by parts of the biosynthetic machinery. The post-translational modifications introduced in the core peptide have been reviewed extensively elsewhere and expand the chemical and structural features far beyond the 20 canonical amino acids.[11, 12] These modifications include, but are not limited to, heterocyclisation of Ser/Thr and Cys residues to oxazolines and thiazolines, oxidation of these heterocycles to the corresponding azoles, epimerisation of amino acids to give D-stereocentres, methylation, Ser/Thr/Tyr prenylation, dehydration, hydroxylation, macrocycle formation and the formation of new C–C bonds through different chemistries. In fact, in some RiPPs such as pyrroloquinoline quinone, the final product does not contain any peptide bonds.[12] Bottromycins are the only RiPP of bacterial origin that utilises a follower peptide rather than the canonical leader peptide for biosynthesis.[7-10]

This review aims to cover all aspects of bottromycin research with a focus on the recent years (2009-2020). We will place particular emphasis on the total synthesis of bottromycins, studies

conducted to investigate the biosynthesis and produce derivatives *in vivo* and very recent progress on the enzymology of individual steps involved in bottromycin biosynthesis.

1a bottromycin A$_2$ R$_1$=Me R$_2$=H R$_3$=Me R$_4$=Me

1b bottromycin B$_2$ R$_1$=R$_2$=H R$_3$=Me R$_4$=Me

1c bottromycin C$_2$ R$_1$=R$_2$=Me R$_3$=Me R$_4$=Me

1d bottromycin D R$_1$=Me R$_2$=H R$_3$=H R$_4$=H

Figure 8-2 | Structures of bottromycins. A Structural formular of bottromycin A$_2$, B$_2$, C$_2$ and D. B NMR structure of bottromycin A$_2$ in CDCl$_3$. Coloured fragments in A represent the identified ninhydrin positive compounds after hydrolysis of bottromycin during the structure elucidation process of bottromycin.

8.3 Discovery, structure elucidation and activity

In 1955, a new *Streptomyces* species, *Streptomyces bottropensis*, isolated from a soil sample taken in the German town Bottrop, was reported.[13] Its fermentation broth displayed antibiotic activity and led to the discovery of a new antibiotic, named bottromycin.[1, 13]

Waisvisz and coworkers were the first group to investigate the bottromycin structure and reported that although bottromycin itself gave a negative ninhydrin test, hydrolysis of bottromycins gave 7 ninhydrin positive compounds. Four remained unspecified, but two that could be identified as glycine and valine.[1, 14-16] Acetylation of bottromycin yielded two degradation products, which were further analysed. This resulted in the identification of two ninhydrin-positive compounds, α-amino-β-phenylbutyric (MePhe) acid and β-(2-thiazole)-β-alanine methyl ester (thia-β-Ala-OMe)[15] as well as the previously identified glycine and valine.[16] Thus, only two of the unknown hydrolysis products remained obscure. As part of these studies, a methyl ester moiety was identified and further analysed: mild alkaline hydrolysis yielded a biologically inactive bottromycin[14], while re-esterification in methanolic HCl solution yielded a biologically active compound identical with bottromycin. Esterification with ethanol or *n*-butyl alcohol produced less active bottromycin derivatives.[14]

In 1965, Nakamura and colleagues reported the isolation of bottromycin from *Streptomyces* No. 3668-L2.[17] They were able to identify the two remaining unknown ninhydrin-positive substances from acid hydrolysis as L-β,β-dimethyl-alpha-aminobutyric acid (*tert*-leucine, *t*-Leu) and L-*cis*-3-methylproline (MePro).[18] Analysis of the isolated antibiotic by thin-layer chromatography revealed it to contain a major component, which was identical to the previously studied bottromycin and designated as bottromycin A, and two minor components, designated as bottromycin B and C.[18, 19] Bottromycin B and C are almost identical to bottromycin A, but contain L-proline (bottromycin B) and L-3,3-dimethylproline (Me$_2$Pro) (bottromycin C) instead of MePro. Bottromycins B and C are biologically active, but bottromycin B displayed 3-4 times less potency than bottromycin A and C (see Table 8-1 and Table 8-5).[19] Nakamura and colleagues also reported the recovery of pivalic acid after hydrolysis of bottromycin beside the 6 previously described compounds.[18] Different tests (i.e. van Slyke test) also suggested the existence of an amidine group in bottromycin in the tetrapeptide moiety.[18] It was concluded that the *N*-terminus of bottromycins must not be free, because they are negative in ninhydrin reactions, Edman degradation and Sanger decomposition.[18] The structure (**2**) that was proposed based on these data harboured pivalic acid at the *N*-terminus of the tetrapeptide (*t*-Leu, Val, MePro, Gly) (Figure 8-3). A subsequent revision of the structure postulated 1-Δ^1-caproic acid instead of pivalic acid at the *N*-terminus.[20]

Bottromycin with pivalic acid was designated as bottromycin A_1, and bottromycin containing $1-\Delta^1$-caproic acid was designated as bottromycin A_2, but synthetic attempts by Yamada *et al.* indicated that the proposed structure was incorrect.[21]

R = H, Me

2
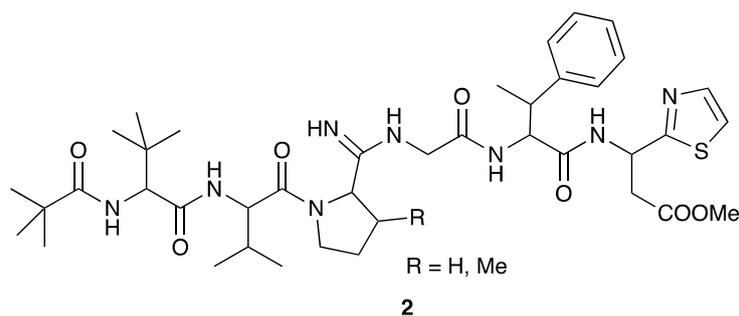
Figure 8-3 | Structure of bottromycin (**2**) according to Nakamura *et al.*[18]

Ten years later, Takita and colleagues proposed a new structure for bottromycin, based on mass spectrometry and ^{1}H NMR data.[22] This cyclic structure of the tetrapeptide was revised by Shipper in 1983,[23] who demonstrated that an unusual amidine moiety links the cyclic tetrapeptide and the linear chain and is formed by condensation of *N*-terminal amino group and the backbone amide carbonyl. In spite of these iterative revisions, and the fact that Bottromycin A_1 and A_2 were actually identical, the designation bottromycin A_2 (as well as B_2 and C_2) for bottromycin A, B and C was retained. It took until 2009 before the correct structure of bottromycin A_2 (**1a**) (see Figure 8-2A) was determined by Shimamura *et al.* through the total synthesis of bottromycin A_2, demonstrating the D-configuration of the thia-β-Ala-OMe.[6] In 2012, Gouda *et al.* determined the three-dimensional structure of bottromycin A_2 in $CDCl_3$ based on NMR data.[24] In this structure (Figure 8-2B) the *C*-terminal residues fold back on the macrocycle made by the four *N*-terminal amino acids. Hence, the MePro and the thia-β-Ala-OMe, which are essential for activity, are on one side of the three-dimensional structure, which suggests an involvement of this region in target engagement.

Biological data for bottromycins A_2–C_2 were reported by Nakamura *et al.*[19] They determined the minimal growth inhibitory concentrations (MIC) towards a wide range of bacterial strains (Table 8-1). In addition, bottromycin A_2 also showed strong inhibition against mycoplasma (0.001-0.01 μg/ml),[2, 25] the multidrug-resistant human pathogens MRSA (1 μg/ml) and VRE (0.5 μg/ml)[2] and *Xanthomonas oryzae pv. oryzae* KACC 10331, a pathovar that causes rice bacterial blight,[26] which makes bottromycins potentially interesting for agrochemical use.

Although highly active *in vitro*, the bottromycins showed no convincing *in vivo* efficiency because of their instability in oral and parenteral administration,[27] which is mainly the result of the lability of the methyl ester under physiological conditions.[2] Synthetic approaches to change the methyl ester moiety could increase plasma stability without decreasing the activity (see

Table 8-2 - Table 8-5). *In vivo* studies using bottromycin derivatised at the methyl ester moiety displayed *in vivo* activity against staphylococcal and streptococcal infection in animals,[27] and *Mycoplasma gallispetium* (pleuropneumonia-like organisms) in chicken using subcutaneous administration.[25, 28, 29]

Table 8-1 | MIC-values for bottromycins A_2-C_2. MIC values from Nakamura *et al.*[19]

Strain	MIC (μg/ml)		
	A_2	B_2	C_2
Staphylococcus aureus (Smith)	0.2	0.8	0.1
Staphylococcus aureus (209 P)	0.1	0.8	0.1
Staphylococcus aureus (BR4)	0.4	0.8	0.4
Staphylococcus aureus (R1)	0.4	1.5	0.4
Staphylococcus aureus (R5)	0.4	1.5	0.4
Staphylococcus aureus (R6)	0.4	1.5	0.2
Micrococcus flavus	0.4	1.5	1.5
Bacillus subtilis (PCI 219)	0.06	0.2	0.06
Bacillus cereus (IAM 1729)	0.4	0.8	0.25
Cornybacterium xerosis	0.06	0.2	0.06
Mycobacterium phlei	0.1	1.5	0.1
Mycobacterium 607	25	(25)	(12)
Shigella dysenteria	12	25	25
Shigella sonnei	25	>100	50
Salmonella typhosa	50	>100	>100
Salmonella paratyphi	25	6	6
Escherichia coli B	3	25	6
Escherichia coli K12	25	100	50
Klebsiella pneumoniae 602	50	>100	>100
Pseudomonas aeroginosa A3	50	>100	100
Sarcina lutea	0.4		0.8
Proteus vulgaris OX-19	12	100	50

S. aureus BR4, R1, R5 and R6 are clinical isolates resistant to antibiotics: BR4 is erythromycin-carbamycin resistant, R1 and R6 are penicillin-tetracycline resistant.

8.4 Mechanism of action

In MoA studies, it was demonstrated that bottromycin A_2 inhibits protein synthesis,[30] which is a common target of antibiotics (e.g. aminoglycosides and macrolides). How bottromycin inhibits protein synthesis was investigated by different groups, who came to divergent conclusions.

Tanaka and co-workers were the first group that studied the MoA of bottromycin and demonstrated the inhibition of protein biosynthesis *in vivo* and *in vitro*[30]. The inhibition was reported to be highly dependent on the base composition.[30] They showed that bottromycin inhibits neither aminoacyl-tRNA synthesis nor its binding to the ribosome.[30] The puromycin reaction, which is regarded as an analogous reaction to the peptide bond formation, was not significantly affected by bottromycin A_2 in the absence of guanosine triphosphate (GTP) and G factor. This reaction does not require GTP or G factor. Addition of GTP and G factor stimulate translocation of the peptidyl-tRNA from the A- to the P-site. In the presence of GTP and G factor the puromycin reaction was inhibited by bottromycin A_2.[31, 32] They concluded from these results that bottromycin A_2 interferes with the translocation of peptidyl-tRNA and movement of mRNA on the ribosomes.[32, 33] In a cell-free system, it was determined on which subunit of the ribosome the antibiotic acts. Examining the inhibitory effect of bottromycin A_2 in a protein synthesising system containing excess of either 30S or 50S ribosomal subunit, the excess of 50S over 30S subunit decreased the inhibitory effect by bottromycin. This effect could not be observed using an excess of 30S over 50S subunit. From these results it was concluded that bottromycin interacts with the 50S subunit of the ribosome.[34]

Pestka and Brot also examined the effect of bottromycin on several steps of the protein synthesis.[35] They also determined an effect of bottromycin on the translocation process, using an oligophenylalanine formation assay. An inhibitory effect was observed in the absence and presence of G protein and GTP. In contrast to Tanaka *et al.*, they also observed an effect of bottromycin on peptide bond synthesis using an acetyl-phenylalanyl-puromycin formation assay.[35] As the degree of inhibition on oligophenylalanine synthesis and the puromycin reaction were comparable, they suggested that the inhibition of the peptide bond formation may be the primary action of bottromycin A_2.[35]

The latest studies examining the MoA of bottromycins were carried out in the early 1980s by Otaka and coworkers.[3-5] They reported that bottromycin interferes with the interaction of aminoacyl- or peptidyl-tRNA with the A (aminoacyl) site of ribosomes[3, 5] and proposed the hypothesis that bottromycin binds to (or close to) the A site of the ribosome and lowers the affinity of aminoacyl-, peptidyl-tRNA or puromycin.[4, 5]

The proposed MoA of bottromycin from Otaka and coworkers is similar to the mechanism of tetracyclines. Tetracyclines block the binding of aminoacyl-tRNAs to the A site of the ribosome,[36] but only bottromycins are able to release bound tRNA from the A site. While tetracyclines bind to the 30S ribosomal subunit,[36, 37] bottromycins are reported to bind the 50S subunit of the ribosome.[34] The different MoAs and binding sites between tetracycline and bottromycin are also supported by the observation that no cross-resistance to the tetracycline-resistant strains *S. aureus* R1 and R6 (Table 8-1) is observed. Other antibiotics that act at the A-site can also have different functions, such as negamycin, which inhibits translocation and stimulates miscoding.[37-39] In the past decade, multiple structures of the 70S ribosome or its subunits in complex with antibiotics have been determined, which provided insights into their mechanism of action.[39, 40] Unfortunately, no ribosome-bottromycin complex structure has been published yet, so our understanding of the MoA of bottromycin remains limited.

8.5 Synthetic approaches and total syntheses towards bottromycins

Synthesis of the unusual amino acids.[41]

Based on the early structure proposals, the first syntheses of the unusual amino acid building blocks were reported from the mid 1970's on. A first stereospecific synthesis of (2*S*,3*R*)-3-methylproline (MePro) was reported by Titouani *et al.* in 1980 using a Hofmann-Löffler-Freitag reaction.[42] Herdeis *et al.* reported the syntheses of both, the (2*S*,3*S*)- and the (2*S*,3*R*)-isomer of MePro from a pyroglutaminol derivate,[43] while Karoyan and Chassaing used a 5-*exo* trig cyclisation approach to generated the five-membered ring.[44] Kamenecka *et al.* developed a protocol starting from commercially available 3-hydroxy-(*S*)-proline using Stille cross coupling chemistry,[45] and a stereoselective cuprate addition was the key step in the synthesis by Flamant-Robin *et al.*[46]

3,3-Dimethyl-(2*S*)-proline (Me$_2$Pro) is one of the unusual amino acids found in bottromycin C$_2$. A first enantioselective synthesis was reported by Sharma and Lubell in 1996.[47] A regioselective enolisation of a 4-oxo-proline derivative followed by alkylation with different alkyl halides allowed the synthesis of a variety of proline derivatives. Two approaches towards racemic Me$_2$Pro were described by Medina[48] and Bott *et al.*[49]

Most investigations focused on the synthesis of (2*S*,3*S*)-3-methylphenylalanine (MePhe) because it also appears in some other natural products, such as mannopeptimycin[50] and the isoleucyl-tRNA-synthetase inhibitor SB-203208.[51] In connection with one of the first synthetic studies towards bottromycins, Kataoka *et al.* described the synthesis and optical resolution of MePhe via condensation of racemic 1-bromo-1-phenyl-ethane with acetaminomalonate.[52] Many attempts have been undertaken to separate the stereoisomers more easily using modern chromatographic techniques.[53-61] Ogawa *et al.* reported an enzymatic approach to MePhe,[62] while Tsuchihashi *et al.* used the Michael addition of malonate onto a chiral vinyl sulfoxide as a key step.[63] Dharanipragada *et al.*[64, 65] and Fioravanti *et al.*[66] described the asymmetric syntheses of MePhe using auxiliary-controlled enolate chemistry, while the groups of Pericas and Rieva developed a protocol using a Sharpless epoxidation as a stereo-controlling step.[67] O´Donnell *et al.* reported an acyclic stereoselective boron alkylation as a key step using a chiral boron reagent in the presence of cinchona alkaloids,[68] while the group of Turner developed a chemo-enzymatic route towards enantiomerically pure MePhe derivatives, based on an oxidation-reduction sequence.[69] Doi *et al.* described a phase transfer-catalysed alkylation of a glycinate Schiff base with 1-bromo-1-phenylethane under the influence of chiral quaternary ammonium bromide and 18-crown-6.[70] And finally, Zhang. *et al.* reported a palladium-

catalysed C-H functionalisation of C(sp^3)-H bonds using 8-aminoquinoline (AQ) as a directing group, giving access to fully protected MePhe derivatives.

The unusual *C*-terminal thiazolyl amino acid thia-β-Ala was the last one whose configuration was determined. It required the total synthesis of the bottromycins to establish it definitively. The problem arose from the structure elucidation of bottromycin. By hydrolysis of the natural product with conc. HCl, Waisvisz *et al.* obtained a "sulfur-containing amino acid", which unfortunately showed no optical activity.[15] Umezawa´s group subsequently obtained an optically active amino acid ([α]^{18}$_D$: +9) by hydrolysing the antibiotic with acetic anhydride.[71] To determine the structure of the *C*-terminal amino acid Waisvisz prepared racemic thia-β-Ala by addition of hydroxylamine towards β-2-thiazolacrylic acid, unfortunately only with moderate yield.[72]

Seto *et al.* tried to obtain optically active (*S*)-thia-β-Ala starting from (*S*)-aspartic acid.[72] Thiazole formation was performed by condensation of the corresponding protected aspartic acid thioamide with bromoacetaldehyde, but unfortunately, these derivatives were also optically inactive. Obviously, complete epimerisation occurred in the thiazole formation step. The racemic amino acid, however, could be resolved into its enantiomers by treating the Phth-derivative with brucine.[46] After cleavage of the Phth-protecting group, the (+)-amino acid, the constituent of bottromycin, was isolated in pure form. It should be mentioned that the thia-β-Ala derivatives prepared also lost their optical activity after heating under reflux in 6 N HCl for 8 h, while the same compounds were stable at room temperature or under slightly basic conditions, illustrating the configurational lability of these compounds. The only enantioselective synthesis of enantiomerically pure (*S*)-and (*R*)-thia-β-Ala so far was reported by the groups of Sunazuka and Ōmura,[6] taking advantage of the chiral sulfinamide chemistry developed by Davis and Ellman,[73, 74] which allowed the synthesis of both enantiomers in a highly stereoselective fashion.

Synthetic studies towards Bottromycins

The total synthesis of bottromycins was adversely affected by the long-standing erroneous structure proposals for the compounds, and so far only one total synthesis exists.[6, 75] It confirmed that the *C*-terminal thia-β-Ala is (*R*)-configured and not (*S*)- as originally reported. While the wrong structural assumptions meant that early synthetic work could only ever be unsuccessful, significant efforts were directed towards the synthesis of the partial structure of this rather unique peptide. The first investigations were already reported by Yamada *et al.* in

1997.[76] Their synthetic route was based on the linear hexapeptide **2** proposed by Nakamura *et al.* (Figure 8-3).[19, 71, 77, 78]

Yamada *et al.* focused on the synthesis and properties of the central amidine unit. Several amidines were prepared by condensation of Cbz-protected amino acid imido esters with amino acid esters (Scheme 8-1).[21] The desired amidine **3** could be obtained without problems, but it was impossible to extend the dipeptides at the *C*-terminus. On activation, or even on standing under basic conditions cyclisation occurred to the corresponding imidazolone **4**. Therefore, the authors decided to form the amidine unit of **7** by coupling two model tripeptide fragments, the tripeptide imido ester **5** and tripeptide **6**.[21]
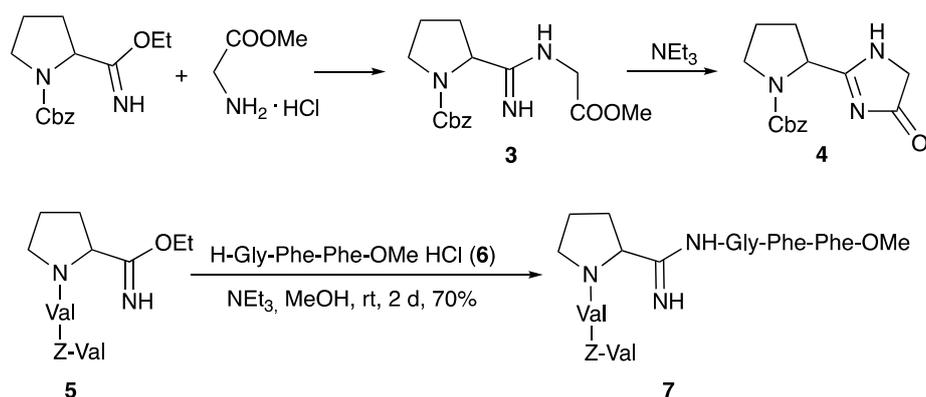

Scheme 8-1 | Formation of amidine **7.**

Interestingly, the pK_a of all synthesised amidines (pK_a ~ 9,3) were around 1 pK_a higher than in the natural product (~ 8.2), a first indication that the structure proposal might not be correct. The antimicrobial activities of these amidines were examined, but no activity was observed. Based on the revised structures by Schipper[23] and Kaneda,[79] who proposed a cyclic tetrapeptide with a tripeptide chain connected via an unusual amidine moiety, Kazmaier *et al.* focused on the synthesis of the corresponding peptide ring and the highly substituted amidine.[80] A key step of their approach was an Ugi reaction using a protected thioamino acid and NH_3 as the amine compound (Scheme 8-2). Although Ugi reactions with NH_3 are often non-specific and yield a range of side products, good results were obtained with sterically demanding aldehydes.[81, 82] With thiocarboxylic acids this approach allowed the synthesis of endothiopeptides.[83-85] With isocyanoacetate the linear tripeptide **8** was obtained, which could be extended to the desired tetrapeptide **9** under standard conditions. Attempts to cyclise **9** or to connect the side chain via peptide coupling failed, because the thioamide underwent cyclisation to the thiazolinone **10**, comparable to the imidazolone formation reported by Yamada (Scheme 8-1).[21]
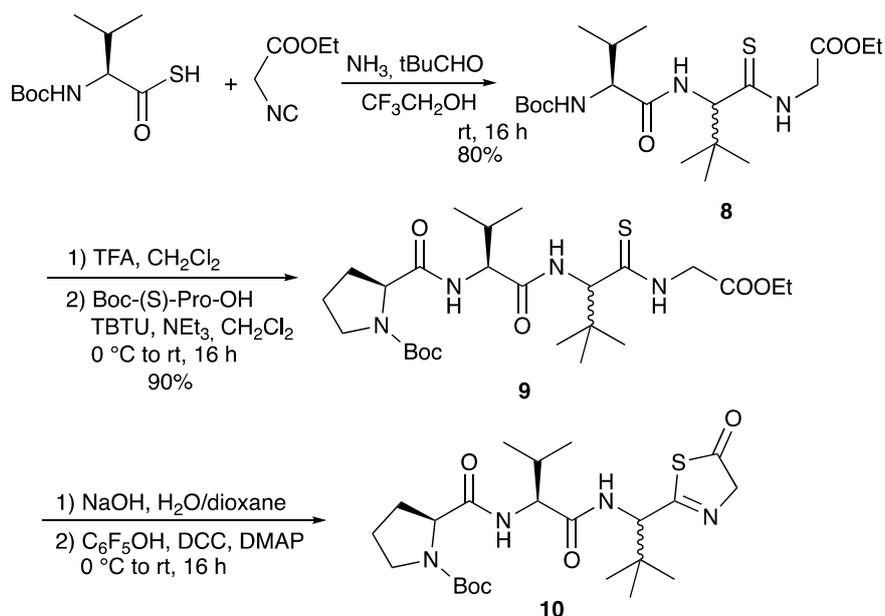
Scheme 8-2 | Synthesis of endothiopeptides via Ugi reaction.

To figure out if amidine formation is possible between sterically demanding amino acids, thiopeptide **11** was synthesised in an analogous fashion (Scheme 8-3). Attempts to couple **11** directly with amines failed, so the thioamide was converted into the corresponding thioimidoester **12**. In the presence of $Hg(OOCCF_3)_2$ **12** could be coupled with valine methyl ester to obtain amidine **13** in good yield.
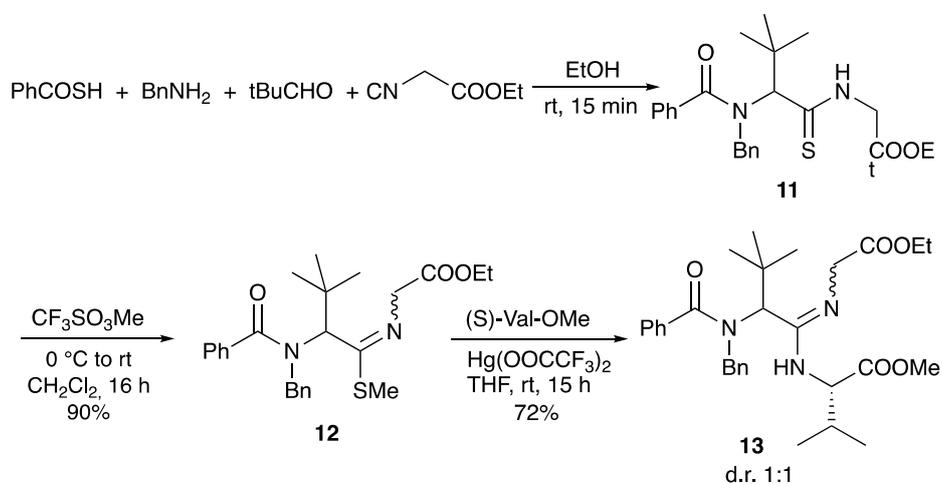
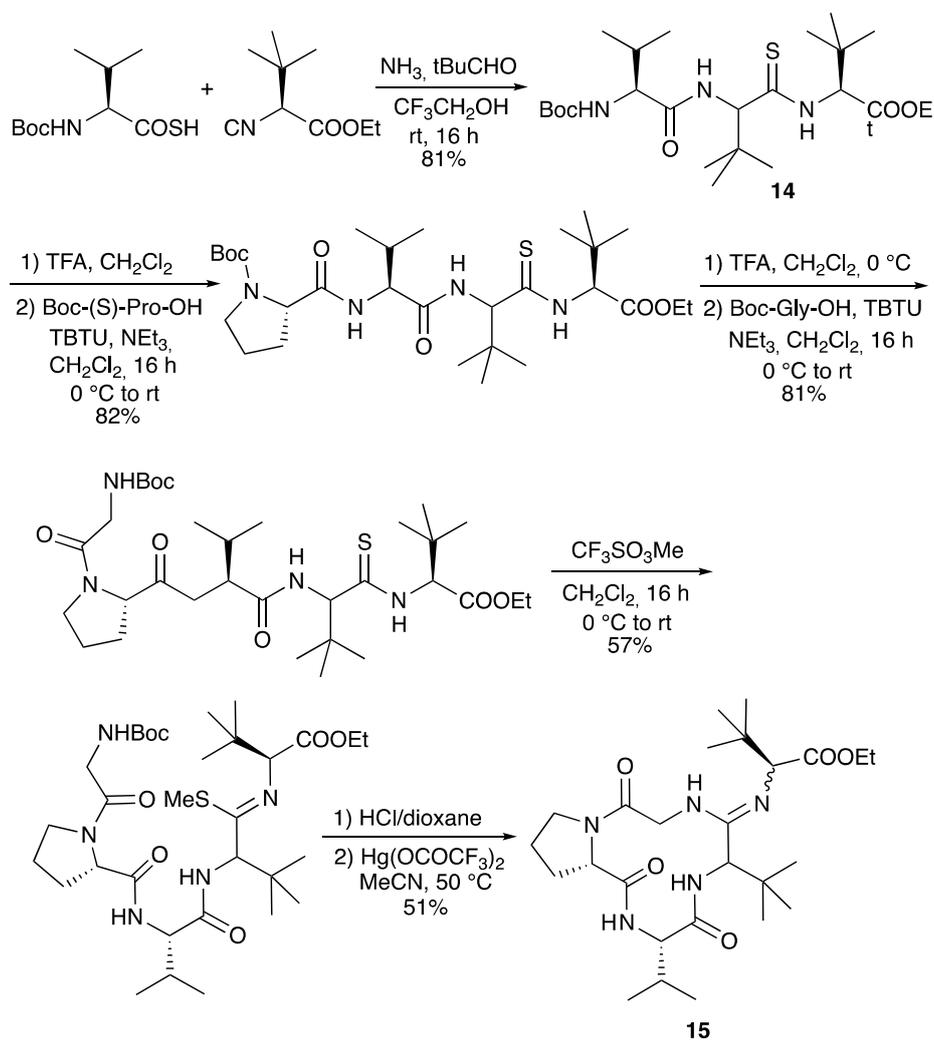**Scheme 8-3 |** Synthesis of amidine **13** via Ugi reaction.

The diastereomers formed could be separated by flash chromatography, but unfortunately this protocol could not be carried out with endothiopeptide **8**. This resulted in a change in the strategy, replacing the intermolecular amidine formation with an intramolecular one by using

the isocyanide of tLeu-OMe (Scheme 8-4). The endothiopeptide **14** was obtained in high yield and could be extended on the N-terminus. S-Methylation and cyclisation in the presence of Hg(OOCCF$_3$)$_2$ gave access to cyclic amidine **15**.
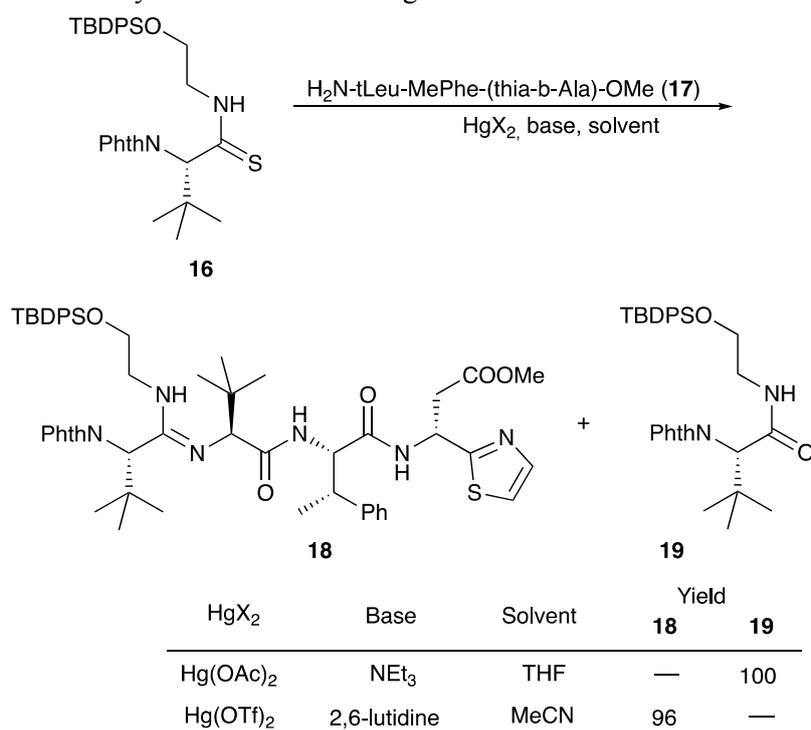
Amidine formation as the key step was also investigated in detail by Ōmura and Sunazuka *et al.* during their synthesis of bottromycin A$_2$ (**1a**) and B$_2$ (**1b**) (Scheme 8-5).[75] They investigated the reaction of thioamide **16** with the tripeptide side chain **17**. No reaction was observed in THF using NEt$_3$ as a base, while in the presence of Hg(OAc)$_2$ the desired amidine **18** was not obtained, and instead the amide **19** was produced. Better results were obtained using HgCl$_2$ and Hg(OTf)$_2$ as Lewis acids. Finally, 2,6-lutidine in acetonitrile was the method of choice to yield **19**.

The same groups also performed degradation studies of bottromycin obtained by fermentation (Scheme 8-6).[86] They subjected **1a** to pyrolysis in MeOH in a sealed tube at 130 °C, resulting in cleavage of the tripeptide side chain. Besides dipeptide **20**, cyclic product **21** was also obtained as a diastereomeric mixture. Obviously, the epimerisation of the tLeu in the side chain occurred via the enol-form of **21**. This could explain why the tLeu obtained by total hydrolysis of the bottromycins has a lower optical rotation than the synthetic enantiopure amino acid. Reduction under mild conditions converted the natural product into alcohol **22**, which could be used to investigate cyclisation conditions.
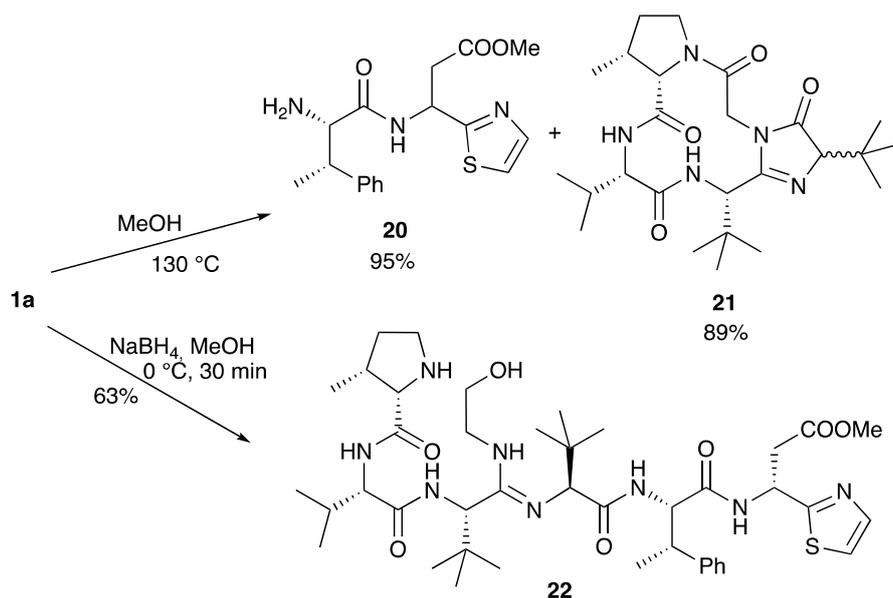
Dipeptide **20** was also used to determine the configuration of the thia-β-Ala, an amino acid that is rather configurational labile.[72] Both enantiomers of thia-β-Ala-OMe were synthesised via the sulfinamide protocol and subsequently coupled with azido-MePhe **23** (Scheme 8-7). Reduction of the azido functionality of **24** provided the two diastereomeric dipeptides **20**. Comparison of their ^{1}H NMR spectra with the spectrum of **20** obtained via pyrolysis clearly indicated that the (R)-isomer is incorporated into the bottromycins and that the original structure proposal (S) was incorrect. Coupling of **20** with Boc-(S)-tLeu and subsequent Boc-cleavage provided tripeptide **17**, which was also used in the amidine formation experiments (Scheme 8-5).
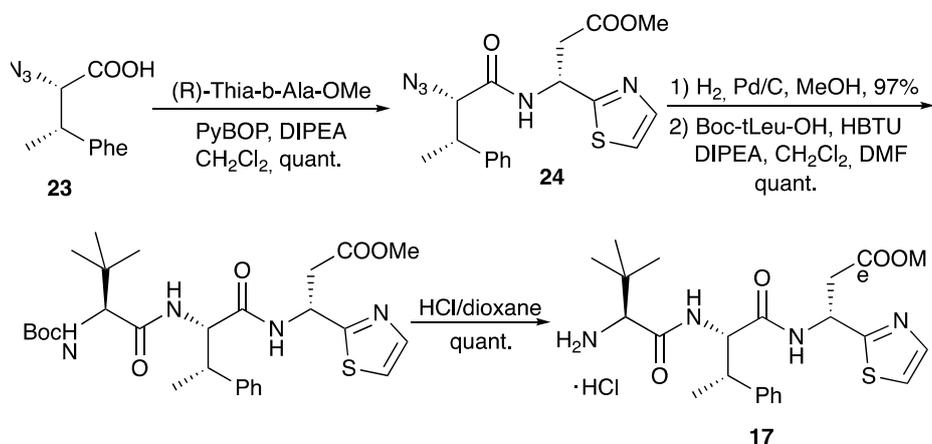
Scheme 8-4 | Synthesis of cyclic amidine **15** via Ugi reaction.

Scheme 8-5 | Synthesis of linear amidine **18**.

HgX$_2$	Base	Solvent	Yield	
			18	**19**
Hg(OAc)$_2$	NEt$_3$	THF	—	100
Hg(OTf)$_2$	2,6-lutidine	MeCN	96	—

**Scheme 8-6** | Degradation of bottromycin A$_2$ (**1a**).

Scheme 8-7 | Synthesis of tripeptide side chain **17**.

Total synthesis of bottromycin and analogs

Based on their own synthetic studies and with all building blocks in hand, Sunazuka and Ōmura *et al.* developed the first and so far only complete total synthesis of bottromycin (Scheme 8-8).[6, 75] To extend the peptide chain, the Phth-group of amidine **18** was removed and the free amine was coupled with Boc-(*S*)-Val. Further elongation gave rise to hexapeptide **25,** which was subjected to desilylation and oxidation. These last two steps had to be carried out on the stage of the hexapeptide, as attempts to oxidise tetrapeptide **18** resulted in the formation of a diketopiperazine. The oxidation was a very critical step due to the nucleophilicity of the internal amidine. Thus all oxidation methods proceeding via an aldehyde failed, because this aldehyde intermediate was trapped by the amidine forming an imidazole. Only Jones oxidation was successful, providing an acceptable yield of the desired carboxylic acid **26**. The amidine was

also problematic in the final macrocyclisation step, and the best results were obtained using EDCI/DIPEA in CH$_2$Cl$_2$, although the yield of **1a** was only moderate.
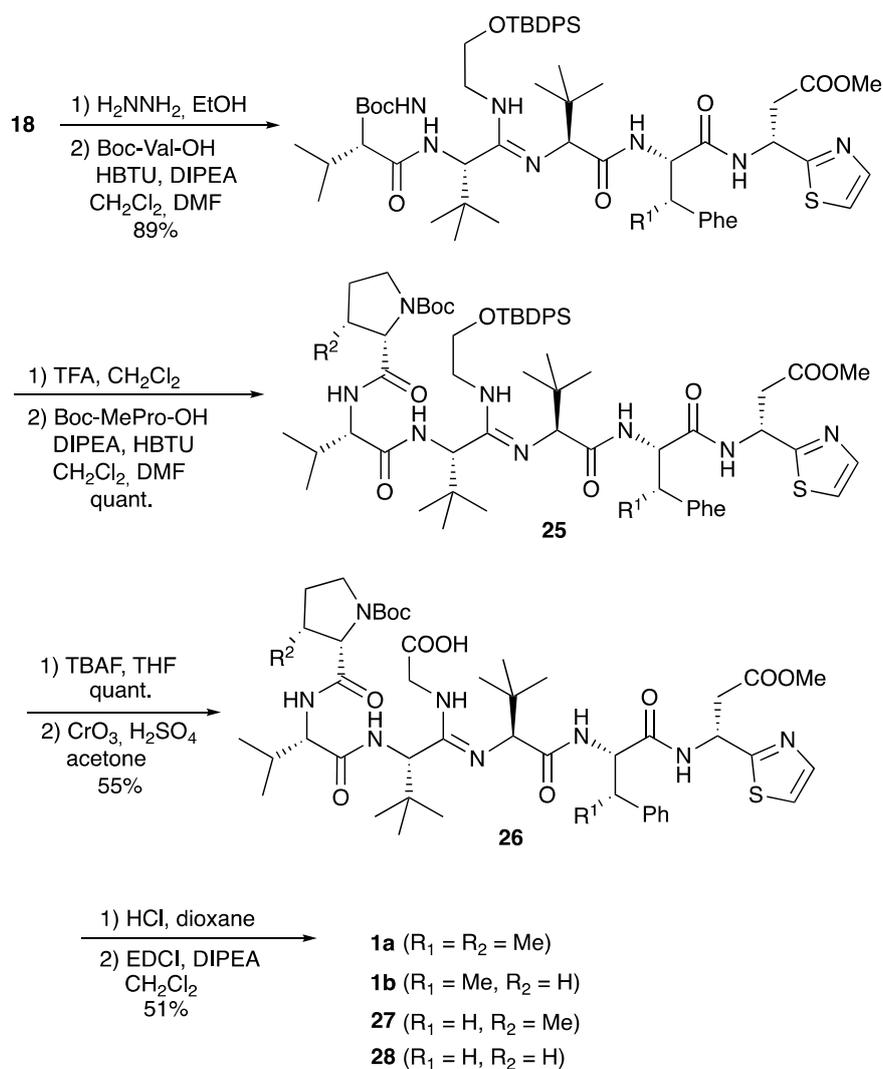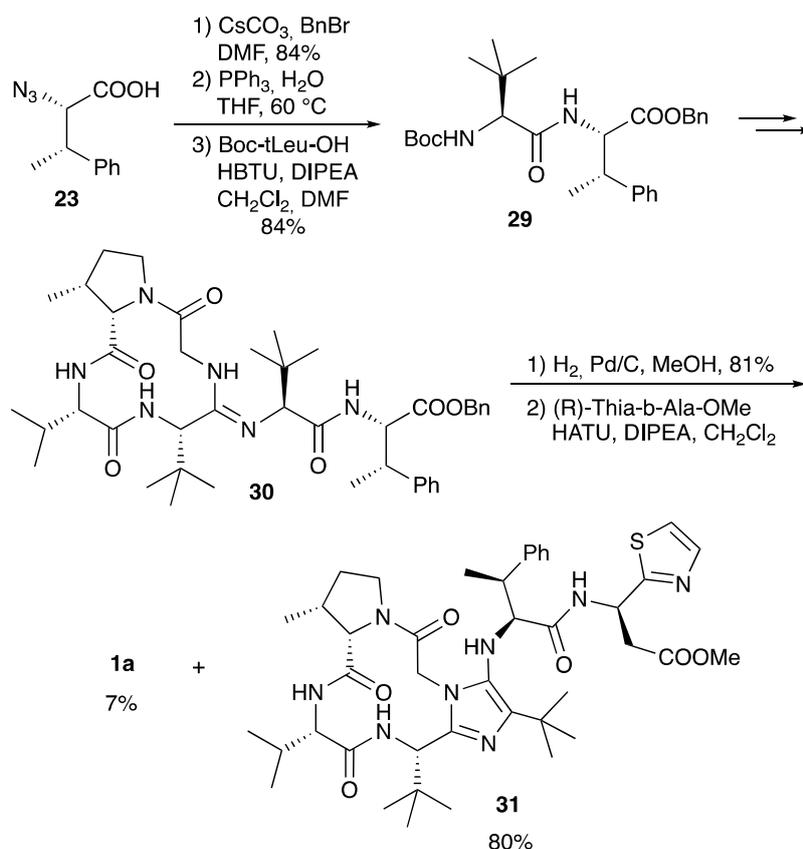

Scheme 8-8 | Total synthesis of bottromycin A$_2$ (**1a**).

This protocol was also used to generate derivatives missing some β-methyl groups, such as bottromycin B$_2$ (**1b**) (Pro instead of MePro), or derivatives where β-MePhe was replaced by Phe [Phe-BotA$_2$ (**27**), PheBotB$_2$ (**28**)]. Their NMR spectra were rather complicated (existence of conformers), which suggests that the methyl group of the β-MePhe is important for the three-dimensional structure of the bottromycins.

Since it is known that the methyl ester of the thia-β-Ala has an effect on the biological activity of bottromycins *in vitro* and *in vivo*[2] Sunazuka and Ōmura considered the synthesis of a bottromycin derivative missing the *C*-terminal amino acid, so that this position can be varied in the last step by coupling a wide range of amines to the "shortened" hexapeptide. Although this is a highly interesting approach, it was not as trivial as hoped. Azido-MePhe **23** was converted

into the corresponding benzyl ester and, after reduction of the azide, coupled to Boc-(*S*)-*t*Leu (Scheme 8-9). The dipeptide **29** was incorporated into bottromycin derivative **30** according to Scheme 8-9. The benzyl ester could be cleaved easily to the carboxylic acid, the key intermediate for the synthesis of analogs. To validate the concept, the acid was coupled with (*R*)-thia-β-Ala-OMe to the original natural product **1a**. The reaction proceeded smoothly, but **1a** was only a side product. The main product was derivative **31** containing an imidazole on the tetrapeptide ring.

Scheme 8-9 | Variable synthesis of **1a**.

So far, HATU as the coupling reagent gave the best yields for bottromycin A$_2$ analogs (**32**) and was used to generate a range of amides (Scheme 8-10), but the corresponding imidazole was the main product in all cases.
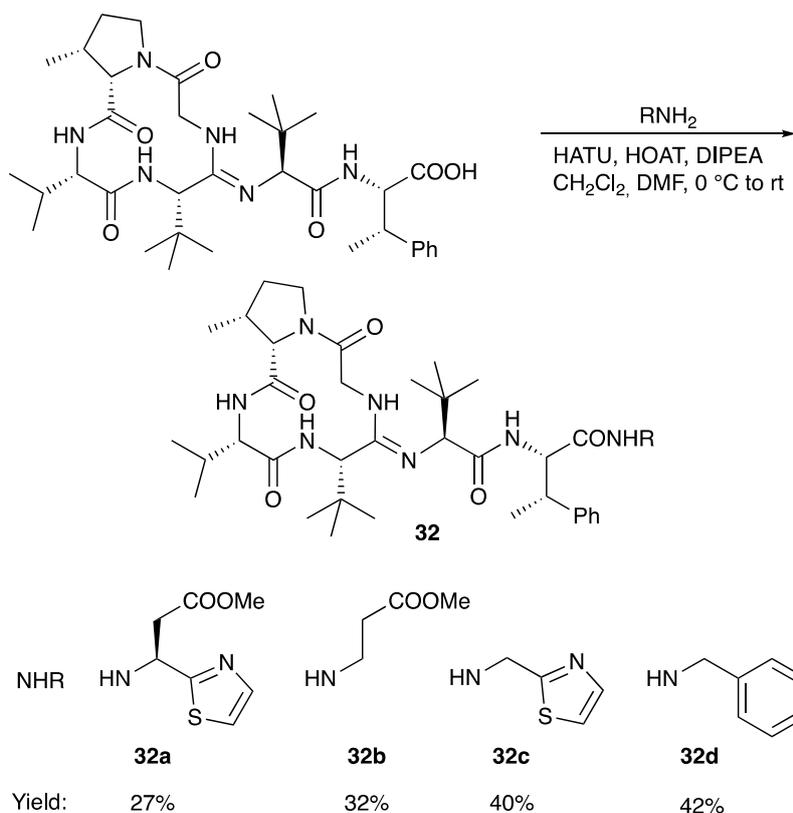
Scheme 8-10 | Synthesis of bottromycin A$_2$-analogs **32**.

Further bottromycin derivatives were obtained by saponification of the natural product **1a** at the *C*-terminus and coupling the free acid **33** with suitable nucleophiles (Scheme 8-11). Researchers at AiCuris used this approach for the synthesis of "Weinreb-amide"-type *N*-alkyl-*N*-alkoxyamides **34** by reaction with linear or cyclic *N,O*-dialkylhydroxylamines (Table 8-2).[87]
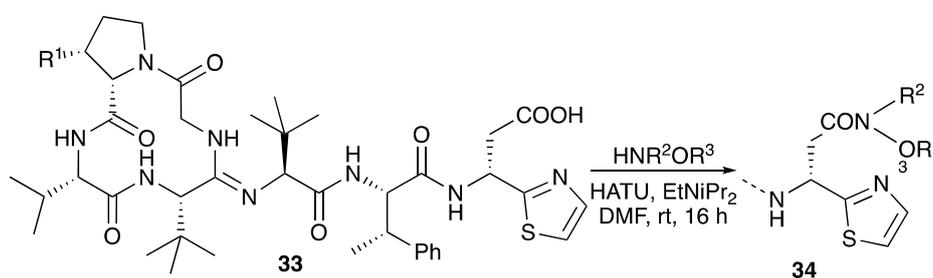
Scheme 8-11 | Synthesis of bottromycin A$_2$-analogs **34**.

301

Table 8-2 | Synthesis of bottromycin-analogs **34**.

34	R^1	R^2	R^3	Yield (%)
34a	Me	Me	Me	70
34b	Me	-(CH$_2$)$_3$-		40
34c	Me	-(CH$_2$)$_4$-		35
34d	H	-(CH$_2$)$_4$-		14
34e	Me	Et	Et	50
34f	Me	Me	Et	47
34g	Me	-CH$_2$-CHOH-CH$_2$-		63
34h	Me	-(CH$_2$)$_3$-CH(COOEt)-		52
34i	Me	-(CH$_2$)$_3$-CH(COOEt)-		84
34k	Me	⬡		90
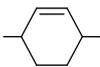
The groups of Ōmura and Sunazuka synthesised a range of different derivatives via the corresponding hydrazide **35** as a common intermediate (Scheme 8-12).[2] The hydrazide was easily obtained by heating the solution of bottromycin A$_2$ with hydrazine. Nitrosation gave rise to acyl azide **36** as an active intermediate, which could be coupled with a range of amines to the corresponding amides **37**. Application of mono Boc-protected piperazine allowed further modification by replacing the Boc-protecting group (**38**). On the other hand, heating the acyl azide to 60 °C resulted in a Curtius rearrangement that gave rise to an isocyanante **39**, which on treatment with amines provided ureas **40**. Reacting **36** with thiols gave rise to thioesters such as **41** which could be subjected to palladium-catalysed cross coupling reactions with organozinc reagents generating ketones **42.**

Biological activities and structure-activity relationship (SAR) studies of bottromycins and derivatives

Based on the importance of the methyl ester for bioactivity and its lability *in vivo* a range of amide derivatives **37** were prepared and their *in vitro* and *in vivo* acitivity towards *S. aureus* (Table 8-3) was compared.[29] The tests were either carried out *in vitro* using a rapid tube dilution test, or in an *in vivo* mouse model, where compounds were administered intraperitoneally once. All compounds were active, but the primary and secondary amides were less active than the esters in the *in vitro* test, but more active *in vivo*. Aromatic amides (**37k-m**) and those with a basic substituent (**37n**) were less active to almost inactive *in vivo*, and tertiary amides (**37o,p**)

302

were also significantly less active, while hydrazides (**35**, **37q**) and hydroxamates (**37r**) were relatively active *in vivo*.

In a patent, researchers at AiCuris described the synthesis of *N,O*-dialkylated bottromycin hydroxamates **34a-k** (Table 8-2) and their biological evaluation (Table 8-4).[87]
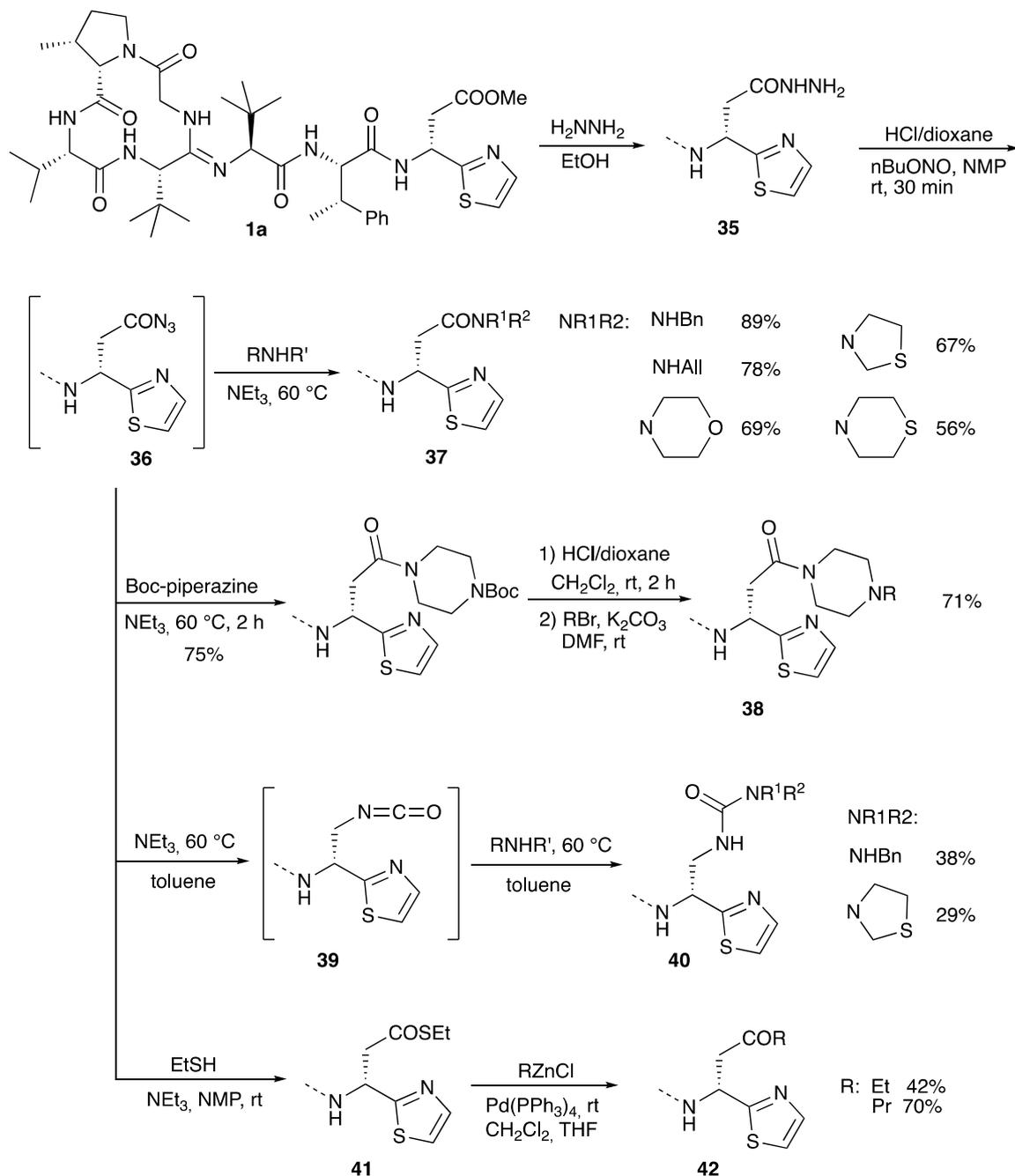
Scheme 8-12 | Synthesis of bottromycin A$_2$-analogs via hydrazine **35**.

Table 8-3 | Biological activities of bottromycins and *C*-terminal amide derivatives (–NR^{1}R^2).

Compound	R^1	R^2	*In vitro* MIC (μg/ml)	*In vivo* ED$_{50}$ (μg/dose)	Ratio *in vivo*: *in vitro*
1a			0.01	50	5000
1b			0.04	200	5000
37a	H	H	0.10	25	250
37b	Me	H	0.05	10	200
37c	Et	H	0.05	15	300
37d	*n*Pr	H	0.5	10	20
37e	*i*Pr	H	1.0	10	10
37f	*t*Bu	H	0.25	10	40
37g	Bn	H	0.5	18	36
37h	CH$_2$CHOHCH$_2$OH	H	1.0	25	25
37i	CH$_2$CH$_2$OH	H	0.5	18	36
37k	Ph	H		60	
37l	*p*-F-C$_6$H$_4$			>100	
37m	α-Naphthyl	H		>100	
37n	NH(CH$_2$)$_2$NEt$_3$	H		>100	
37o	*i*Pr	*i*Pr		82	
37p	-CH$_2$CH$_2$CH$_2$CH$_2$-			95	
37q	NMe$_2$	H		35	
37r	Me	OH		28	
35	NH$_2$	H		46	

Table 8-4 | Biological activities of bottromycin hydroxamate derivatives **34**.

34	MIC (μM)			
	S. aureus 133	*S. pneumoniae* G9a	*E. faecium* BM4147	*E. faecalis* ICB27159
34a	0.78	<0.05	0.39	0.78
34c	0.39	<0.05	0.1	0.39
34d	3.13	<0.05	0.39	1.56

By far the most detailed SAR studies were reported by Ōmura and Sunazuka, who also investigated the desmethyl derivatives **27** and **28**, which were obtained by total synthesis (Scheme 8-8). A wide range of different derivatives such as amides **32** (Scheme 8-12), **37** and **38**, hydrazide **35**, ureas **40**, thioester **41** and ketones **42** were prepared from bottromycins obtained by fermentation (Scheme 8-12). Their activity was tested towards a panel of Gram-positive strains, using vancomycin (VCM) and linezolid (LZD) as references (Table 8-5).[2, 75] The results of the SAR studies are summarised schematically in Figure 8-4.
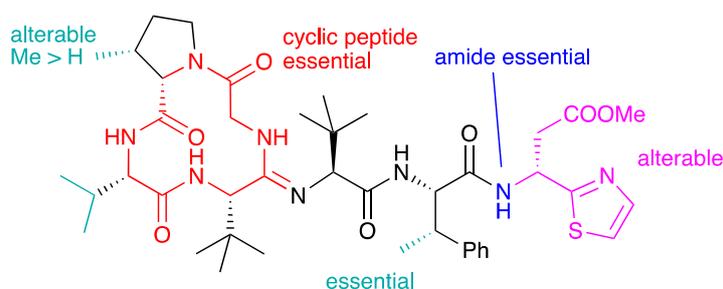
Figure 8-4 | Summary of SAR for bottromycin derivatives.

The unusual methylation pattern (cyan) has a significant effect on the bioactivity towards *S. aureus*. Bottromycin D (**1d**), where the valine is replaced by an alanine, and bottromycin B_2 (**1b**), which does not have the methyl group at the proline, were less active than bottromycin A_2 (**1a**) (Table 8-2 Table 8-5).[19] Bottromycin C_2 (**1c**), the analogue dimethylated on proline, was roughly as active as bottromycin A_2. The β-methyl group on the Phe seems to be essential and its removal (**27**, **28**) causes a dramatic drop in activity (Table 8-5). It appears that this methyl group influences the conformation of the side chain and controls the three-dimensional structure of the whole molecule, an assumption which is supported by ^{1}H NMR.[6, 75] Bicyclic derivatives such as **31** and linear peptides do not show significant activity, probably due to an undesired three dimensional conformation, which clearly indicates that the cyclic peptide ring (red) is essential.[75] No activity was observed for derivatives with either a COOH-group at the *C*-terminus, such as **33**, or if the thia-β-Ala is missing completely. This might be caused by a drop in the hydrophobicity. Interestingly, incorporating the opposite (*S*)-isomer of thia-β-Ala (**32a**) had no significant effect on activity (2 μg/mL). The thia-β-Ala (purple) was not essential at all for activity - derivatives missing the acetate side chain (**32c**) or the thiazole unit (**32b**) were only slightly less active.

Surprisingly benzyl amide **32d** is almost as active as **1a**, while the corresponding benzyl ester **30** is not very effective. The data showed that the amide functionality (blue) is necessary for good activities. Benzyl amide **32d** is more active than the dethiazolyl analogue **32b**, which indicates that an (hetero)aromatic substituent at the *C*-terminus has a positive effect on activity.

The moderate *in vivo* activity of the methyl ester in the natural products probably results from its low hydrolytic stability under physiological conditions and its cleavage towards the almost inactive carboxylic acid **33**. Although significantly less active *in vitro*, better *in vivo* stabilities were observed for secondary aliphatic amides (Table 3, **37a-i**), while aromatic (**37j-m**) and tertiary amides (**37o,p**) as well as those with basic side chains (**37n**) were almost inactive.[29] Piperazino derivatives **38** and ureas **40** exhibited 4- to 32-fold weaker activity *in vitro*, but better stability.[2] Thioesters such as **41** were significantly more active than **1a**, but due to their great reactivity completely unstable in mouse plasma. Ketones **42**, which cannot undergo hydrolysis, are perfectly stable and showed activities comparable to **1a** and vancomycin, but importantly were also active against vancomycin-resistant strains. Subsequent biological evaluation using MRSA-infected mice showed that propyl ketone **42b** might be a good candidate for drug development. 100 mg/kg given to mice orally resulted in survival for at least five days after administration, while all non-treated animals died in the same time frame. Hydroxamates **34** (Table 8-4) might also be suitable for this purpose.[87]

Table 8-5 | Biological activities of bottromycin derivatives towards Gram-positive strains.

Comp.	R^1	R^2	MIC (µg/ml)						Rates of residual anti-MRSA activity (%)
			S. aureus FDA209P[a]	*S. aureus* Smith[a]	MRSA HH-1[b]	MRSA 92-1191[b]	VRE NCTC12201[c]	VRE NCTC12203[c]	
1a			1	1	1	2	1	0.5	0
1b			4	4	4	4			
27			32	>32	>32	32			
28			>32	>32	>32	>32			
30			>32	>32	>32	32			
31			>32	>32	>32	>32			
32a			2	2	2	2			
32b			8	8	8	8			
32c			4	4	4	2			
32d			2	4	2	2			
33	Me	H	64	64	64	128	128	32	--
35			16	16	16	32	8	4	86
37g	Bn	H	8	8	8	8	8	2	71
37s	CH₂CCH	H	8	8	8	16	4	2	100
37t	-(CH₂)₂O(CH₂)₂-		16	8	16	32	16	4	100
37u	-(CH₂)₂SCH₂-		4	4	8	8	8	2	100
37v	-(CH₂)₂S(CH₂)₂-		8	4	8	8	4	4	100
38a	Boc	H	8	4	8	8	8	4	42
38b	H		64	32	64	128	32	32	--
38c	CH₂CCH		16	16	16	32	16	16	67
38d	Bn		4	4	4	4	4	4	84
40a	-(CH₂)₂SCH₂-		4	4	4	4	4	2	100
40b	Bn		8	16	16	16	8	4	100
41			<0.25	0.5	<0.25	0.5	<0.25	<0.25	0
42a	Et		1	1	2	2	2	1	100
42b	*n*Pr		1	1	1	2	2	0.5	100
VCM[d]			1	1	0.5	1	>128	>128	--
LZD[e]			2	2	2	2	2	2	--

[a] *S. aureus* FDA209P and Smith: susceptible strains. [b] *MRSA* HH-1 and 92-1191: MRSA strains isolated from clinical patients. [c] Vancomycin resistant *Enterococcus faecalis* NCTC12201 and NCTC12203: encoded by van A gene. [d] Vancomycin. [e] Linezolid

8.6 Biosynthetic gene cluster and biosynthesis

The biosynthetic origin of the bottromycins was unknown for decades following their discovery. They could feasibly be synthesised via either ribosomal or non-ribosomal pathways, as all residues could theoretically be produced from proteinogenic amino acids. Early studies of bottromycin biosynthesis involved elegant isotope labelling experiments from Arigoni and colleagues, where it was shown that the methyl groups at the β-positions of proline, two valines and phenylalanine, along with the thia-β-Ala-OMe, were derived from methionine.[88] They also showed that β-methylated valine and phenylalanine derive from L-valine and L-phenylalanine. Feeding *S. bottropensis* with methionine featuring an isotopically labelled chiral methyl group ([methyl-(^{2}H,^{3}H)]-(2*S*,methyl-*R*)-methionine) showed that the β-methylation occurs with a retention of configuration. This was consistent with double inversion of configuration via a radical SAM mechanism.[89] Additional experiments using isotopically labelled amino acids were also consistent with a radical SAM mechanism.[88]

In 2012, four independent research teams identified bottromycin biosynthetic gene clusters (BGCs) in four different *Streptomyces* species: the known producers *S. bottropenesis*[7] and *Streptomyces* sp. BC16019[10], the plant pathogen *Streptomyces scabies*[9] and the marine ascidian-derived *Streptomyces* sp. WMMB272[8] (Figure 8-5A). These reports corroborated the earlier feeding studies by showing that bottromycins are ribosomally synthesised and post-translationally modified peptides (RiPPs)[12] and that the BGCs encode three radical SAM methyltransferases. In each study, the BGC was identified by BLAST searches for genes that could encode a putative bottromycin core peptide, GPVVVFDC (or GPAVVFDC for bottromycin D in *S.* sp. WMMB272) (Figure 8-5B). RiPPs originate from a larger ribosomally synthesised precursor peptide that usually consists of a leader peptide and a core peptide that is post-translationally modified by tailoring enzymes. However, the discovery of the bottromycin BGC provided the first (and still only) example of a bacterial RiPP that derives from an *N*-terminal core peptide that has no leader peptide and is attached to a "follower" peptide (Figure 8-5C).

The genetic organisation of these BGCs is effectively identical, and while there are significant differences in protein sequence identity between each BGC, *S.* sp. BC16019 nomenclature will be used here onwards for clarity. The bottromycin BGC encodes 13 proteins (Figure 8-5A): one precursor peptide (BotA), two YcaO-domain proteins (BotC and BotCD), three radical SAM methyltransferases (BotRMT1-3), three putative hydrolases (BotH, BotAH, BotP), one cytochrome P450 (BotCYP), one *O*-methyltransferase (BotOMT), one putative regulatory protein (BotR) and one major facilitator superfamily transporter (BotT). These initial studies

revealed a number of key details relating to bottromycin biosynthesis. Gene inactivation experiments in *S. bottropenesis*[7], *S. scabies*[9] and *S.* sp. BC16019[10] confirmed the identity of the BGC, which was further validated by heterologous expression of the *S.* sp. BC16019 BGC. The identity of the BGC in *S.* sp. WMMB272 was demonstrated by the production of bottromycin A_2 upon expression of a mutant precursor peptide gene that encoded the bottromycin A_2 core peptide instead of the natural bottromycin D core peptide.[8]
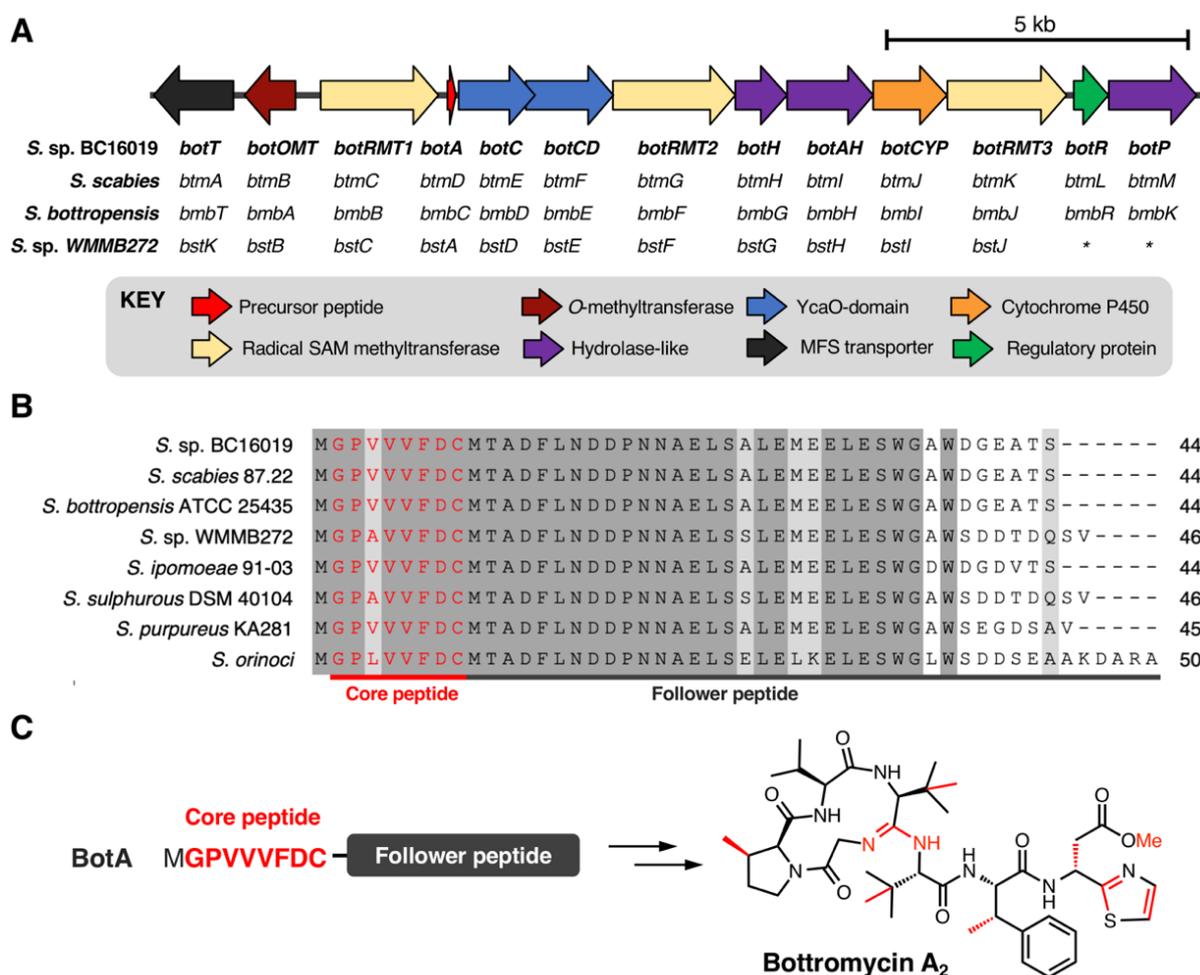

Figure 8-5 | Identification of the bottromycin BGC. **A** Organisation of the BGC. Gene nomenclature from each strain is shown (* = genes not identified in the *S.* sp. WMMB272 study as this BGC was at the end of a contig). **B** Sequence alignment of precursor peptides from every bottromycin BGC identified bioinformatically. Identical residues are shown with a dark grey background, similar residues are shown with a light grey background (Risler matrix score >0.7) and the core peptide is highlighted in red text. **C** Schematic showing the conversion of the BotA precursor peptide into bottromycin A_2, where all post-translational modifications are coloured red in the final product.

Notably, gene deletions in the *S. scabies* BGC[9] and insertional inactivation of genes in the *S.* sp. BC16019 BGC[10] demonstrated the roles of the radical SAM methyltransferases BotRMT1-3 via the production of differentially methylated bottromycin derivatives by each mutant, thereby validating the earlier isotopic labelling studies.[88] BotRMT1 catalyses radical *C*-methylation of Phe6, BotRMT2 catalyses radical *C*-methylation of both Val4 and Val5, and

BotRMT3 catalyses radical *C*-methylation of Pro2. At the time, this represented one of the first examples of radical β-methylation of amino acid residues, along with the polytheonamides, 49-amino acid RiPPs produced by '*Candidatus* Entotheonella factor', a member of a marine sponge microbiome.[90] Multiple non-ribosomal peptides contain β-methylated amino acids, but these are generated via conventional methylation of a precursor keto acid.[91]

Inactivation of *botOMT* in the *S.* sp. BC16019 BGC confirmed its role in *O*-methylation of Asp7[10]. However, little else was known about the biosynthetic steps required to convert BotA into mature bottromycin, although plausible routes were initially proposed based on the predicted catalytic roles of *bot* proteins. In-frame gene deletions in the *S. scabies* BGC had demonstrated the essentiality of numerous putative biosynthetic genes[9], including *botC* and *botCD* (encoding the two YcaO-domain proteins), and *botCYP* (encoding a P450), but no bottromycin related metabolites could be initially identified from these mutants. The challenge with identifying molecules related to RiPPs following gene deletions is that a pathway may "stall" if a key step is disrupted, with the core peptide still attached to the leader/follower peptide. This therefore is likely to undergo further degradation into a very short modified peptide that may be distantly related to the final product and therefore difficult to detect.
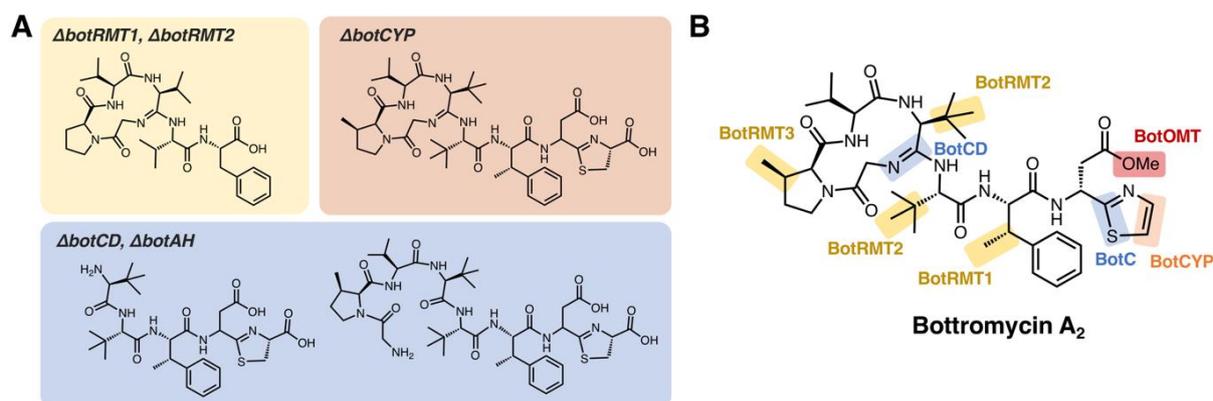
Figure 8-6 | Key metabolites identified by mutational analysis of *bot* biosynthetic genes. **A** Structures of metabolites based on tandem MS data from Crone *et al*. **B** Proposed roles of biosynthetic proteins based on the studies of Crone *et al.*[9] and Huo *et al.*[10]

To improve the detection of bottromycin-related metabolites from pathway mutants, Truman and co-workers used mass spectrometry-based molecular networking[92] and untargeted metabolomics to study in-frame deletions of orthologues of *botA*, *botC*, *botCD*, *botAH, botRMT1, botRMT2* and *botCYP* in *S. scabies*[93]. This analysis identified a series of bottromycin-related molecules (intermediates or shunt metabolites) associated with each mutant strain, which were then used to propose a feasible pathway based on where the pathway stalled (Figure 8-6). This indicated that radical methylation by BotRMT1 and BotRMT2 were

early steps in the pathway in *S. scabies*, as was heterocyclisation of Cys8 by the standalone YcaO-domain protein BotC. Additionally, this study showed that the M17-family leucine aminopeptidase BotP removes methionine from the *N*-terminus of BotA.

Following removal of the *N*-terminal methionine by BotP, the unique macrocyclic amidine of bottromycin was proposed to be formed by BotCD (YcaO-domain protein) and BotAH (hydrolase), based on the production of linear bottromycin-related peptides by each mutant (Figure 8-6). Deletion of *botCYP* led to the accumulation of *O*-desmethyl bottromycins A_2 and B_2 carboxylated at their *C*-termini. This was consistent with BotCYP catalysing late-stage oxidative decarboxylation of the thiazoline moiety to generate a terminal thiazole. Each compound mass appeared as twin peaks via liquid chromatography - mass spectrometry (LC-MS), suggesting a mixture of aspartate epimers, which was supported by deuterium labelling and therefore provided a potential route to D-aspartate in mature bottromycin. Epimerisation was shown to happen spontaneously, but it was not clear whether other proteins were involved in accelerating this key step. No mutant strains produced metabolites that are *O*-methylated on Asp7. This suggested that *O*-methylation is the final biosynthetic step, and it was shown that purified BotOMT could methylate *O*-desmethyl bottromycin A_2. This is in agreement with the earlier gene inactivation work by Huo *et al.*.[10]
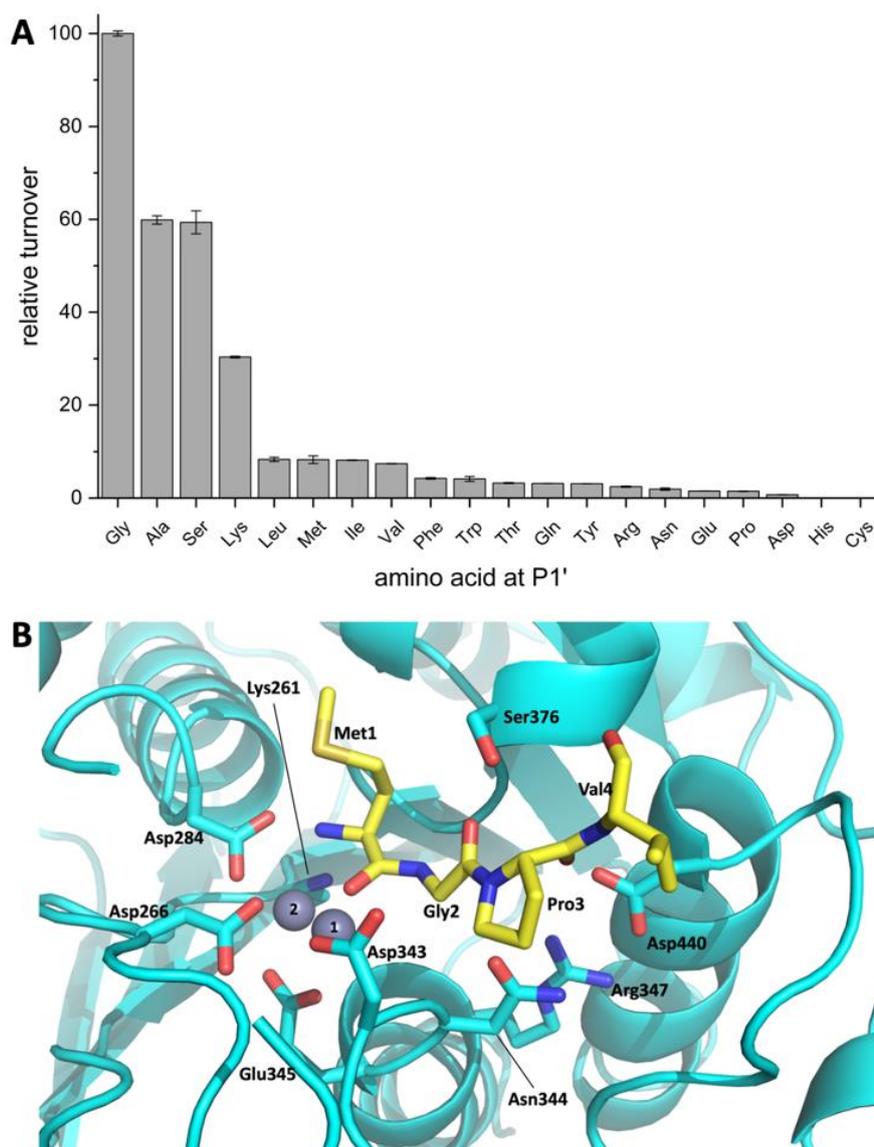
8.7 Biosynthetic enzymes

The following paragraphs will be dedicated to the enzymes involved in the biosynthesis of the bottromycin core scaffold (desmethyl bottromycin): BotP, BotC, BotCD, BotAH, BotH, and BotCYP. The activity of these enzymes has been reconstituted *in vitro*, and structural information is available for several of them. Although the enzymes were studied using enzyme homologues from different BGC containing *Streptomyces* species (see Figure 8-5B), the *S.* sp. BC16019 nomenclature will be used for clarity.

Aminopeptidase BotP

The aminopeptidase BotP belongs to the family of hexameric M17 leucine aminopeptidases (LAP). LAPs are metallo-exopeptidases found in all kingdoms of life and cleave *N*-terminal residues from proteins and peptides. They do not only hydrolyse *N*-terminal leucine residues, but often are promiscuous. The function of LAPs is diverse and goes beyond the function of recycling amino acids and includes the processing of bioactive peptides and peptides for major histocompatibility complex (MHC) class I antigen presentation, gene regulation and vesicular trafficking.[94, 95]

Usually, *N*-terminal methionine is hydrolysed by endogenous aminopeptidases, but these do not function efficiently with the MGP sequence found at the *N*-terminus of BotA.[96] The aminopeptidase BotP was predicted[9, 10] and confirmed[93, 97] to remove the *N*-terminal methionine from the precursor peptide BotA (**43**), which generates the free glycine amino group (**44**) (Figure 8-8) that is necessary for the cyclisation onto an internal amide carbonyl to generate the unique amidine macrocycle found in bottromycins.

Koehnke and coworkers determined the crystal structure of BotP (Figure 8-7B) and assessed the substrate promiscuity of BotP using pentapeptide mimics of BotA (Figure 8-7A).[97] BotP showed a hexameric structure typical for M17 LAPs and the activity of recombinant BotP, isolated from *E. coli*, could be reconstituted in presence of Co^{2+} (or Mn^{2+}) ions. RiPP enzymes catalysing the initial biosynthetic steps often bind to the follower (or for other RiPPs leader) peptide to aid substrate recognition and enzyme activity.[98] For BotP modelling suggests that only the first 3-4 amino acids contribute to substrate binding, but an *in vitro* assay using pentapeptides showed that these truncated substrates were processed slower than full-length BotA. The reasons for this discrepancy remain to be determined. BotP tolerates several amino acid changes in P1-P3', but processing is reduced drastically.

Figure 8-7 | A BotP substrate promiscuity using pentapeptide mimics of BotA with amino acid changes in P1`position. **B** Model of BotP-Mn^{2+} (cyan) with the peptide MGPV (yellow).

YcaO domain enzymes BotC and BotCD

The bottromycin biosynthetic gene cluster contains two genes encoding for YcaO-domain enzymes, BotC and BotCD[7-10]. YcaO proteins can be found in bacteria and archaea. Their function is best characterised in RiPPs such as linear azol(in)e-containing peptides (LAPs), thiopeptides and cyanobactins, where the YcaO enzymes catalyse the ATP dependent cyclodehydration reaction of Cys, Ser or Thr side chains onto the preceding backbone amide carbonyl to form azoline heterocycles.[99-101] They can however also be used to generate non-natural heterocycles.[102, 103]

From an untargeted metabolomic approach it was predicted that BotC catalyses the heterocyclisation of Cys to thiazoline and BotCD, together with BotAH, catalyses the formation

of the unique macroamidine linkage.[93] The function of the two bottromycin YcaO enzymes were independently studied in *in vitro* approaches by the Mitchell and Koehnke groups.[104, 105] It was demonstrated that BotC catalyses the heterocyclisation reaction that converts the core peptide's Cys residue to a thiazoline. The second YcaO enzyme, BotCD, was sufficient to catalyse macroamidine formation. Both proteins bind to the follower peptide, but with low affinity.[104] BotC and BotCD were quite tolerant to changes in the core peptide sequence but recalcitrant to changes of the nucleophile.[104, 105] BotC was unable to utilise Ser or Thr instead of Cys to generate oxazolines.[104] While the turnover of all bottromycin biosynthetic enzymes has been shown to be relatively fast, heterocyclisation by BotC was shown to be slow and could be the rate limiting step of the pathway, since BotCD strongly prefers a heterocyclised substrate for macrocyclisation.[105] In contrast to all other YcaO enzymes studied to date, the BotCD reaction was shown to be reversible: The enzyme catalysed amidine formation and ring opening, both in an ATP and Mg^{2+}-dependent fashion.[105] Thus BotCD expanded the catalytic scope of YcaO enzymes in RiPP pathways to amidine formation, but also raised questions as to possible partner proteins for BotCD, which may prevent ring opening.

Based on the biochemical data it was proposed that macroamidine formation proceeds analogously to heterocyclization: After nucleophilic attack of the *N*-terminal amino group onto the amide carbon a hemiorthoamide intermediate is formed, which is then ATP-dependent *O*-phosphorylated, followed by subsequent phosphate elimination to form the macroamidine (Figure 8-8).
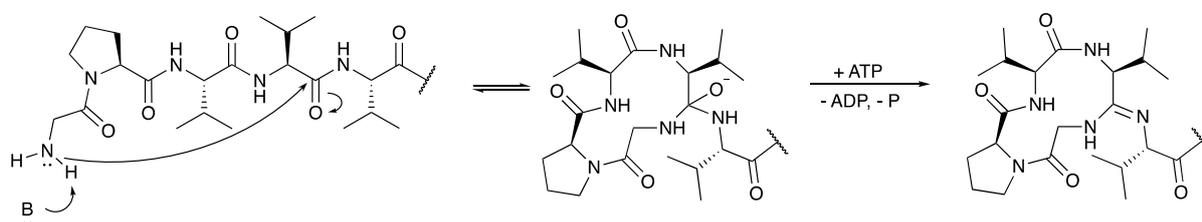
Figure 8-8 | Proposed mechanisms for thiazoline and macroamidine formation by the YcaO domain enzymes BotC and BotCD.
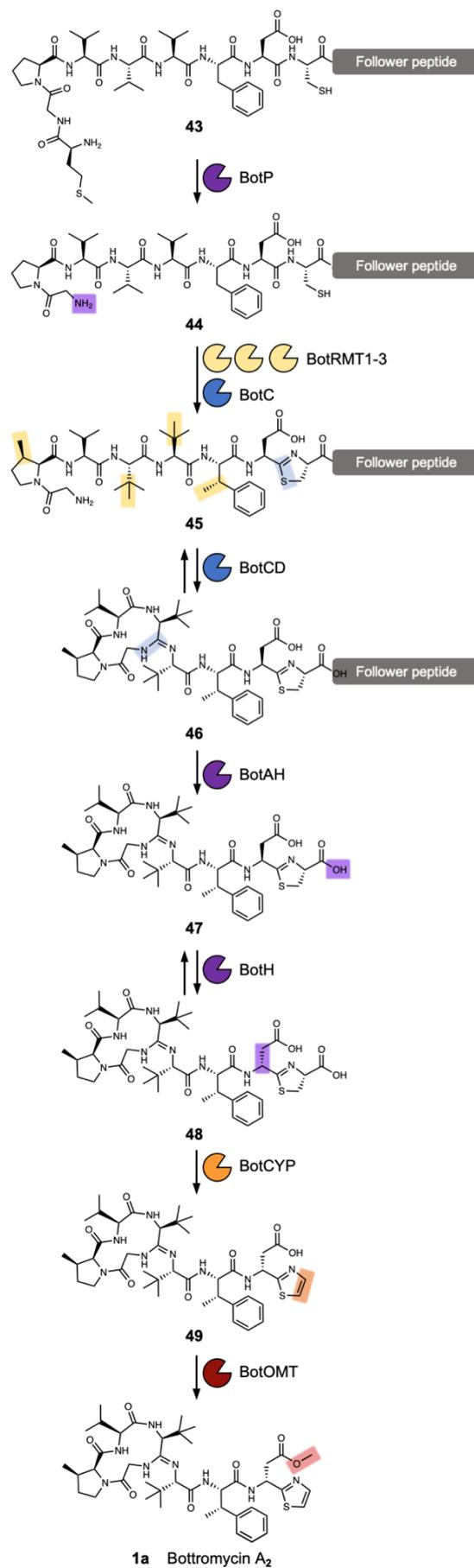
Figure 8-9 | Biosynthetic pathway for bottromycin A$_2$.

Amidohydrolase BotAH

Analysis of the metabolomic network from *S. scabies* Δ*btmI* (*botAH* orthologue) and Δ*btmF* (*botCD* orthologue) revealed two new bottromycin related molecules, which could not be found in *S. scabies* WT and that were not macrocyclised. From this observation it was proposed that BotCD and BotAH are involved in macrocyclisation.[93] However, *in vitro* data showed that BotCD is sufficient to catalyse macrocyclisation. Pfam-analysis classifies the protein encoded by the gene *botAH* as a putative metallo-dependent amidohydrolase. From its putative function it was assumed that either BotAH, or the putative α/β hydrolase BotH, catalyses the liberation of the modified core from the follower peptide for the final biosynthetic steps. The crystal structure of the BotAH homologue PurAH confirmed the Pfam prediction. When the role of BotAH was probed *in vitro* it was discovered that it cleaved the follower peptide off the hetero- and macrocyclised core peptide, and a mechanism was proposed (Figure 8-10).[106] As observed for BotP, the highest activity of BotAH was observed after addition of Co^{2+}. This may explain increasing bottromycin production levels in *S. scabies* by supplementation of production medium with Co^{2+}.[107]
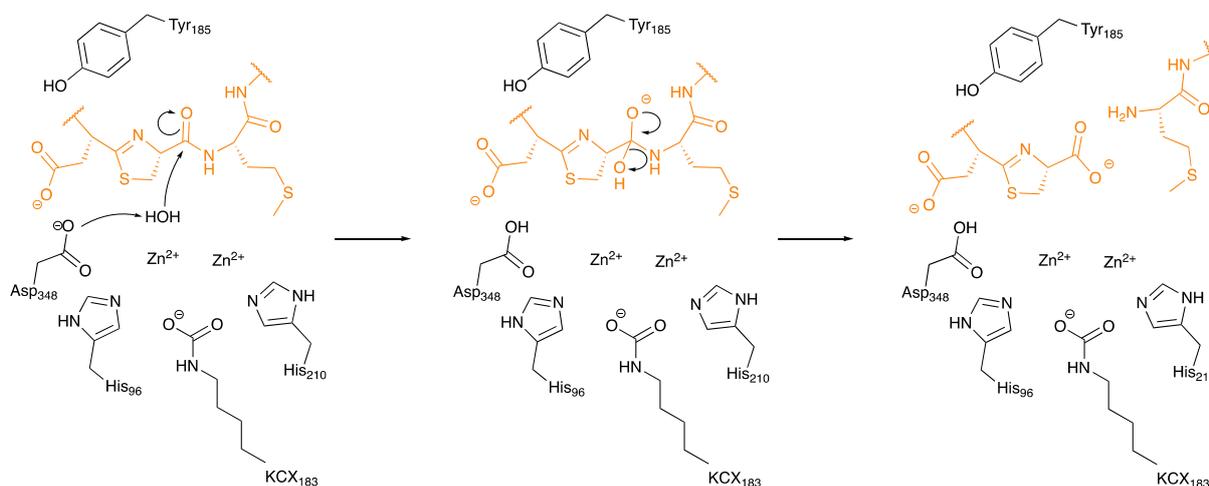
Figure 8-10 | Proposed mechanism of the follower peptide cleavage by the amidohydrolase AH. BotAH active side residues are shown in back, the peptide substrate is shown in orange.

The biosynthetic role of BotAH raised the question of whether it could aid macrocyclisation by influencing the equilibrium of the BotCD reaction. After all, the activity of BotCD was dependent on the follower peptide (attached to the modified core peptide). Indeed, when added to macrocyclisation reactions, BotAH's activity precluded ring opening and pulled the macroamidine formation equilibrium to the side of macrocyclised product.[106] As a result, BotAH can be viewed as a YcaO accessory protein, the first hydrolase for which such a function

has been reported. These observations also rationalise why the amidohydrolase is essential for macroamidine formation *in vivo*.

Atypical α/β hydrolase BotH

The function of the α/β hydrolase (ABH) BotH was not studied *in vivo* by gene deletion experiments and remained unclear until very recently. Enzymes of the α/β hydrolase family catalyse the hydrolysis of (thio)ester and peptide bonds, but some members have diverse functions, such as dehalogenases, dioxygenases and decarboxylases.[108, 109]

BotH is annotated as an ABH, and its crystal structure revealed a typical ABH fold, but the canonical Ser-His-Asp catalytic triad residues, which the majority of the ABH family members possess, is mutated (Ser to Phe, His to Ile) or missing (Asp).[110] Accordingly, no hydrolytic activity was detected for BotH. Instead, the follower-cleaved intermediate **47** binds to the large cavity of the designated BotH active site and *in vitro* biochemical assays determined BotH to function, unexpectedly, as the epimerase of the Asp residue found in bottromycins (**1**).[110] Deuteron labelling experiments revealed, that the enzyme catalyses the rapid epimerisation of L-Asp to D-Asp, and its back-reaction (Figure 8-9). The action of BotH leads to a mixture of **47** and **48**, but provides a much greater abundance of the D-Asp containing intermediate **48** than the spontaneous epimerization, which proceeds a glacial speeds and favours **47**. In the complex crystal structure of BotH with **48**, no potential catalytic residues of BotH were in a reasonable distance of the substrate Asp Cα proton. In addition, mutation of the core peptide Asp residue to Ala or Asn prevented epimerisation, but not binding. Substitution of Asp with Glu still allowed epimerization, which led the suggestion of substrate-assisted catalysis. BotH is the first reported ABH to catalyse peptide epimerisation.

In other RiPP pathways, epimerisation usually involves radical SAM enzymes or a two-step dehydration–hydrogenation process to generate D-alanine from L-serine.[111-114] BotH is thus the founding member of a group of atypical ABH enzymes that may be able to epimerise amino acids post-translationally and but also other secondary metabolites. Further, BotH binds the pathway product bottromycin A_2, but is not able to epimerise it. The complex crystal structure of BotH and bottromycin A_2 was the first crystal structure of any bottromycin (Figure 8-11). The resulting orthosteric inhibition of BotH by bottromycin A_2 may results in a biosynthetic feedback mechanism to prevent self-poisoning.[110]
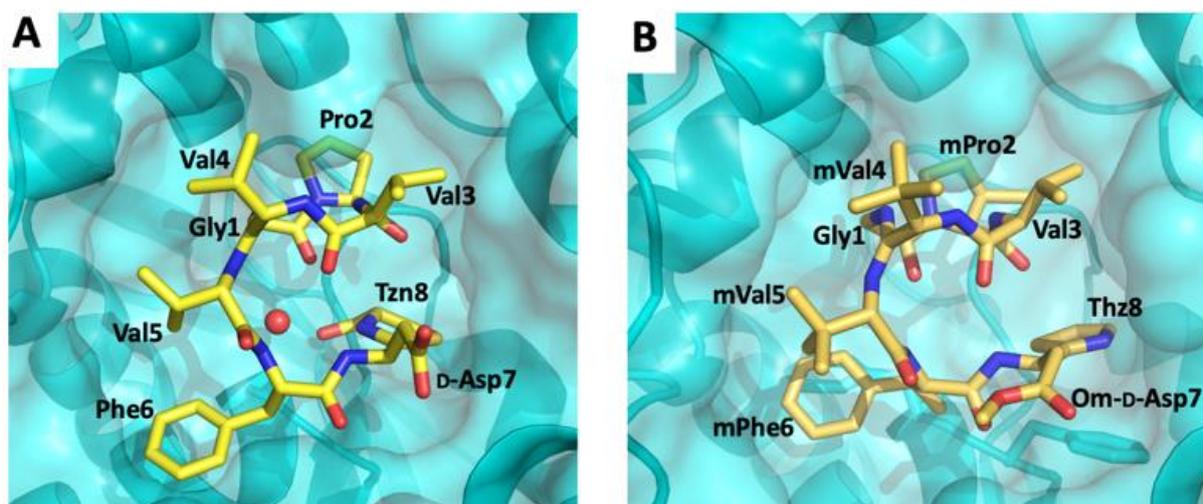
Figure 8-11 | BotH-48 **A** and BotH-bottromycin A$_2$ **B** complex structures.

Cytochrome P450 enzyme BotCYP

Thiazoles (and oxazoles) as heterocycles derived from cysteine (or serine/threonine) are a common motif in RiPPs and are associated with linear azol(in)e-containing peptides (LAPs), cyanobactins, thiopeptides, and bottromycins. Thiazoline oxidation is more frequently observed than oxazoline oxidation and usually involves an FMN-dependent dehydrogenase.[11, 12] In bottromycin biosynthesis the oxidative carboxylation of the *C*-terminal thiazoline is catalysed by the P450 enzyme found in the gene cluster.[93] P450 enzymes are not very common in RiPP pathways, although a recent large-scale survey of the RiPP landscape identified ~ 1800 P450 genes associated with putative RiPP BGCs.[115] The handful of P450 enzymes that had been characterized from RiPP biosynthetic pathways predominantly catalysed amino acids hydroxylation,[116-119] but other functions had also been reported.[118, 120, 121] The P450-catalyzed oxidative decarboxylation of a heterocycle in RiPP biosynthesis had not been investigated *in vitro* but recent work by the Koehnke group demonstrated that BotCYP selectively acts on D-Asp containing intermediate **48** (Figure 8-12).[115] Given the substrate-assisted mechanism of BotH, this observation rationalizes why all attempts to produce bottromycin analogs with residues other than Asp/Glu were unsuccessful (Figure 8-14). Without a D-amino acid in core peptide position 7 the pathway stalls at the point of oxidative decarboxylation. Unfortunately the crystal structure of a close BotCYP homologue did not provide an answer as to how this enzyme may provide stereochemical resolution for the pathway. The reconstitution of this step has nevertheless enabled the production of the bottromycin core scaffold, which may aid the production of bottromycin analogues for bioactivity testing in the future.
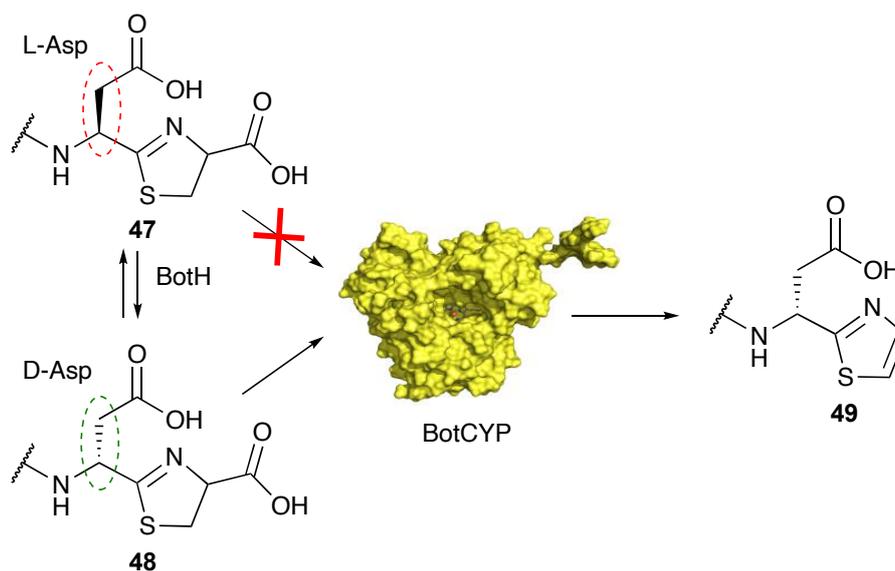
Figure 8-12 | BotCYP catalyses the oxidative decarboxylation reaction of the thiazoline in **48** to the thiazole found in the bottromycin core structure. The cytochrome P450 enzyme is stereoselective for the D-Asp containing intermediate and thus provides stereochemical resolution for the pathway.

8.8 Heterologous production of bottromycins and derivatives

Early studies on bottromycin benefitted from large-scale fermentation to aid with product isolation and characterisation, such as a 3,000-litre fermentation used by Waisvisz and colleagues in 1957 to first characterise bottromycin.[1] However, more recently, numerous groups have reported very low bottromycin yields with laboratory-scale fermentations,[10, 93, 122] which has presented challenges for understanding bottromycin biosynthesis and engineering the pathway to generate derivatives.

Regulation of the bottromycin biosynthetic gene cluster

To identify the transcriptional organisation of the bottromycin BGC, Vior *et al.*[107] used 5′-tag-RNA-seq in *S. scabies*. This revealed transcriptional start sites preceding *botRMT1* and *botOMT* (Figure 8-13A), as well as an internal transcriptional start site within *botRMT1*, which precedes the precursor peptide gene, *botA*. This is hypothesised to increase expression of the precursor peptide in relation to the tailoring enzymes. RT-PCR (reverse transcription - polymerase chain reaction) and 5'-RACE (rapid amplification of cDNA ends) experiments were consistent with the 5′-tag-RNA-seq analysis, although inferred a possible additional transcriptional start site before *botT* that was not detected by 5′-tag-RNA-seq. This study showed that the bottromycin BGC does not encode a pathway-specific master regulator, and instead encodes a regulatory protein, BotR, that specifically modulates the expression of *botA* but not other genes in the BGC. The precise mechanism of BotR-mediated modulation remains enigmatic, especially as no strong promoter activity could be detected from the internal transcriptional start site that precedes *botA*.

Heterologous expression of the *bot* BGC

Huo *et al.* were the first to express a bottromycin BGC in a heterologous host.[10] Here, a pOJ436-based[123] cosmid library from *Streptomyces* sp. BC16019 was screened for the intact *bot* BGC. The resulting DG2 construct was then expressed in *Streptomyces albus* J1074 and *Streptomyces coelicolor* A3(2) to successfully produce bottromycin A_2, although estimated production levels (1-4 µg/L) were much lower than from the native producer. This large construct was simplified by removing a 16 kb fragment of non-*bot* genes from the 5'-end of the BGC. This was replaced with a kanamycin resistance gene via Red/ET recombineering[124] to generate cosmid DG2-kan (Figure 8-13B).

To increase production levels from DG2-kan, an approach pioneered by Ochi and colleagues[125, 126] was employed. This strategy involved obtaining rifampicin-resistant isolates that contain

mutations in the *rpoB* gene, which encodes the RNA polymerase β-subunit. These *rpoB* mutations can enhance levels of antibiotic production in *Streptomyces* without affecting growth. Accordingly, bottromycin production was increased by about 10-fold in rifampicin-resistant mutants of *S. coelicolor*-DG2-kan. Huo *et al.* hypothesised that a limiting factor preventing higher yields was a lack of bottromycin resistance in *S. coelicolor*. The most likely self-immunity gene in the cluster is *botT*, which encodes a major facilitator superfamily (MFS) transporter.[127] Therefore, to increase *botT* expression, the region preceding the *botT* gene was replaced with the strong P*ermE** promoter using Red/ET recombineering. This further increased bottromycin production levels two-fold compared to the rifampicin-resistant mutants. Given that 11 *bot* biosynthetic genes (*botRMT1-botP*) are present on a polycistronic operon,[107] it should be possible to engineer pathway regulation by simply changing the promoter preceding *botRMT1* and optionally modifying the promoter(s) for *botOMT* and *botT*. Truman and colleagues used transformation-associated recombination (TAR) cloning[128, 129] in *Saccharomyces cerevisiae* (yeast) to directly capture the bottromycin BGC from the genomic DNA of *S. scabies* DSM 41658.[130] This used the yeast/*E. coli* shuttle vector pCAP01,[128] which can also integrate into actinobacterial genomes via the φC31 attachment site. The resulting vector, pCAPbtm, was introduced into *S. coelicolor* M1146,[131] but bottromycin was produced in negligible amounts. Therefore, it was hypothesised that the bottromycin BGC could be efficiently engineered to improve productivity by use of homologous recombination in yeast. This strategy involved introducing double-strand breaks in pCAPbtm via restriction sites naturally found in the bottromycin BGC. This fragmented BGC was then repaired using a combination of double- and single-stranded DNA fragments to introduce new genetic features in a marker-free way (Figure 8-13C).[130]

This approach was used to generate a series of modified pCAPbtm-derived plasmids that contained a variety of strong promoters (PSF14, P*hrdB*, P*aac3*, P*ermE**) in front of *botRMT1* and *botT*, as well as rearranged the *botA* and *botRMT1* genes. In an effort to limit the potential toxicity of bottromycin overproduction, the *botOMT* gene was removed, as previous work had shown that BotOMT-catalysed *O*-methylation is important for bottromycin activity.[2] To fully understand the metabolic consequences of engineering regulation, the total productivity of the pathway was assessed using LC-MS-based metabolomics to detect multiple peptides derived from the bottromycin pathway that likely resulted from incomplete biosynthesis and subsequent hydrolysis of modified BotA. One surprising challenge was that BotRMT1 was inactive in all conditions tested, which resulted in all molecules lacking a β-methyl group on phenylalanine. Sequencing revealed no mutations to the gene cluster in any of the constructed vectors, so to

further control expression levels, an inducible theophylline-dependent riboswitch[132] was incorporated upstream of *botRMT1*. While this did not lead to active BotRMT1, theophylline induction did lead to the most productive heterologous expression system tested in this study, which was 120 times more productive than heterologous expression of the wild type BGC, as well as producing higher levels of bottromycin-related metabolites than the native *S. scabies* producer (Figure 8-13D).

The DG2-kan cosmid generated by Huo *et al.*[10] was used as the basis for BGC engineering by Luzhetskyy and colleagues.[122] DG2-kan contains the *bot* BGC from *Streptomyces* sp. BC16019, and this cosmid was expressed in *Streptomyces lividans* TK24 to yield 0.23 mg/L bottromycin A_2. To improve pathway productivity, the BGC was initially engineered to replace the native promoters of *botOMT* and *botRMT1* with strong promoters from a promoter library previously generated by the Luzhetskyy group.[133] The selection of strong promoters initially led to *S. lividans* growth problems, potentially caused by toxicity of the pathway to the host. This was overcome by selecting for bottromycin-resistant *S. lividans* mutants, which led to bottromycin A_2 production levels of up to 3-fold higher than a control strain.

As it can be difficult to predict the precise relationship between promoter strength and pathway productivity, a 'random rational strategy' was employed to further boost yields from DG2-kan. Here, a library of *bot* BGCs was generated that featured random synthetic promoters inserted between *botOMT* and *botRMT1*. These random promoters were created based on the consensus -35 and -10 sequences of the *ermE*p1[134] promoter (Figure 8-13B). Degenerate primers were then used to randomise the sequences upstream, between and downstream of these consensus sequences. These were then introduced into the BGC using Red/ET cloning, and the mutated cosmids were conjugated into *S. lividans* TK24. Screening of 100 randomly selected strains harbouring mutated BGCs (DG2-KmRandom, Figure 8-13B) revealed that 10% produced 5-50 fold more bottromycin A_2 than a control strain harbouring unmodified DG2-kan. Quantitative RT-PCR (RT-qPCR) revealed that the transcription from both promoters had increased in the best producer, but that the strength for each promoter was very different (1:59 ratio), which emphasises the benefit of screening a promoter library. Pathway productivity was further increased by introducing this mutated BGC into the native producer, *Streptomyces* sp. BC16019, which therefore contained one copy each of the wild type and mutated *bot* BGCs. This led to a 37-fold increase in bottromycin production in relation to wild type *Streptomyces* sp. BC16019.
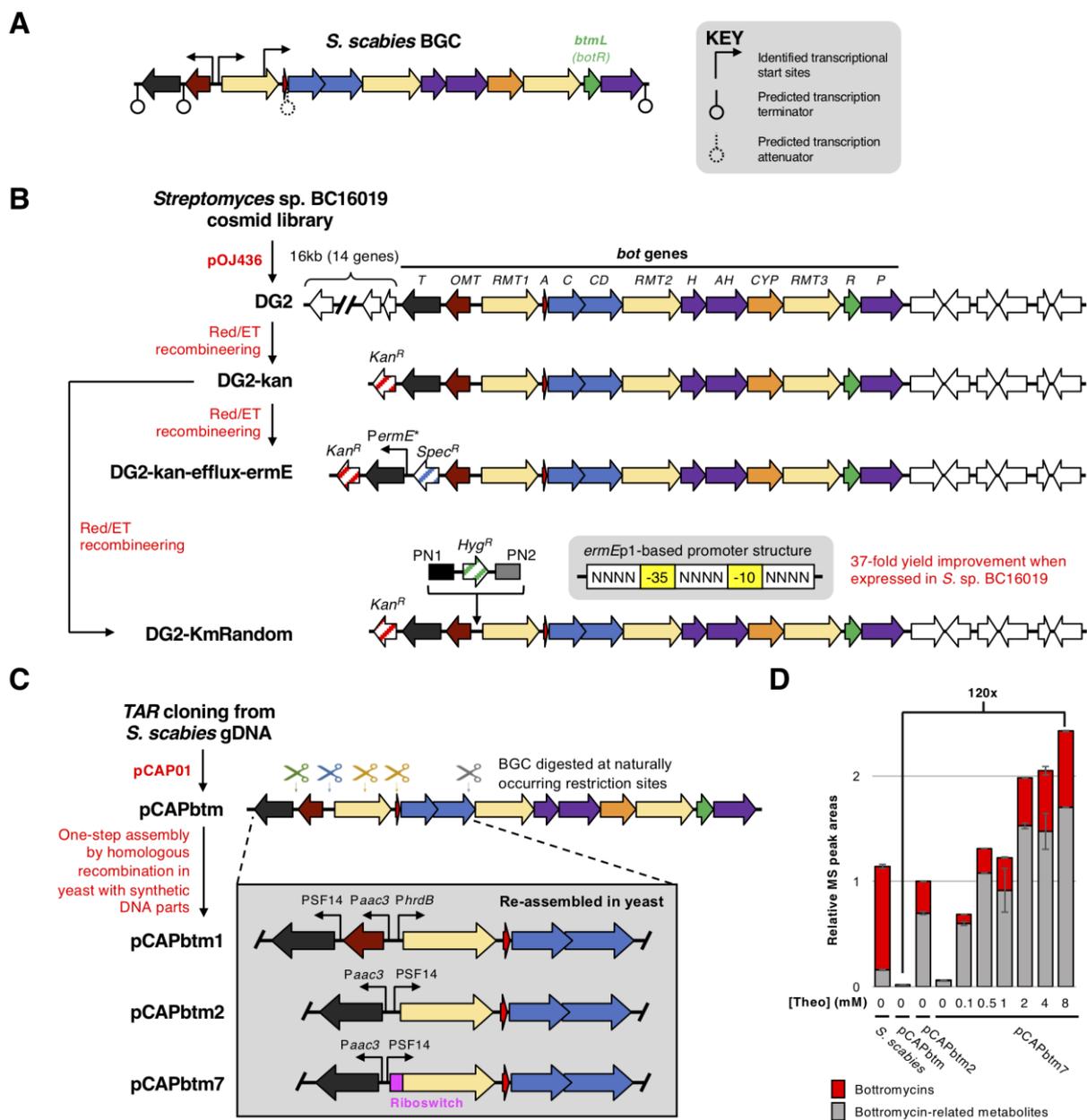
Figure 8-13 | Overview of bottromycin BGC regulation and methods used genetically engineer the BGC. **A** Natural transcriptional start sites and predicted terminators determined by Vior et al.[107] The gene encoding the pathway-situated transcriptional modulator BtmL (BotR) is highlighted. **B** Overview of cloning and engineering the *Streptomyces* sp. BC16019 BGC by Huo et al.[10] and Horbal et al.[122] **C** Overview of cloning and engineering the *Streptomyces scabies* BGC by Eyles et al.[130] showing selected examples of engineered BGCs. **D** Bottromycin production chart adapted from Eyles et al.[130] [Theo] = theophylline induction concentration.

There is a lack of natural diversity within the core peptides of known bottromycins (Figure 8-5B), which is in contrast to most other RiPP families12. RiPPs are uniquely suited to pathway engineering to generate derivatives, as mutations to the core peptide lead to predictable changes to the final chemical structure. The lack of diversity amongst natural bottromycins means that there is a lot of chemical space to explore for bottromycin-like molecules, which could be important in the context of antibiotic discovery. Therefore, Luzhetskyy and colleagues

323

generated DG2-kan cosmids with mutated *botA* genes,[122] and then introduced these into *Streptomyces* sp. BC16019 Δ*botA*, which is unable to make the wild type BotA precursor peptide. 12 mutants were generated, although most led to no detectable bottromycin-like metabolites, indicating a lack of pathway tolerance to these unnatural substrates, which is unlike many other RiPP pathways.[135] Based on these data, the position most tolerant of modifications was Val3 of the core peptide (Figure 8-14). Peptides with isoleucine and methionine residues at this position were successfully converted into bottromycin derivatives containing all expected post-translational modifications. Figure 8-14 shows methionine-containing bottromycin M. Position 3 of the core peptide is effectively the only amino acid that is not post-translationally modified during biosynthesis and is the only position where natural variation has been observed, as bottromycin D from *Streptomyces* sp. WMMB272 features an alanine at this position.[8] In this strain, an Ala3Val mutant was tolerated and therefore generated bottromycin A$_2$. In contrast, no products were detected from three different mutants of Asp7 in this strain (Figure 8-14).
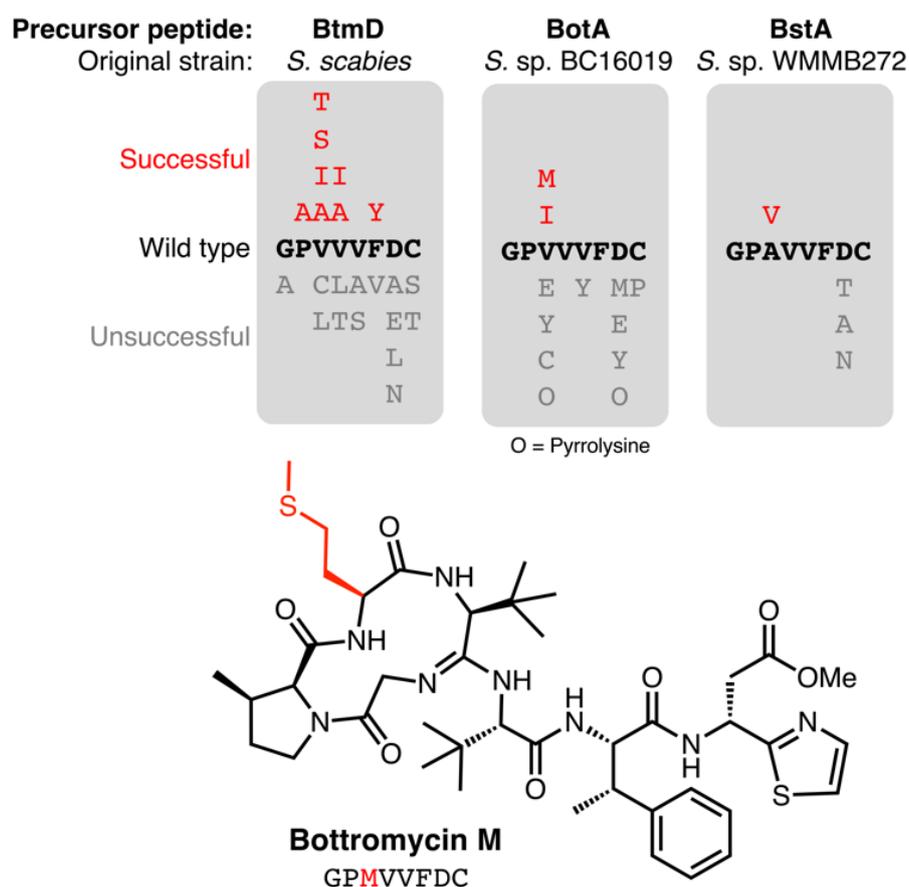
Figure 8-14 | Mutations made to precursor peptide BotA and homologues in other bottromycin BGCs.[2, 122, 136] The characterised structure of bottromycin M is shown.

Comparable mutation results were obtained by Crone *et al.*[136] Here, 22 *botA* mutants were generated using the precursor peptide complementation strategy previously reported for *S. scabies* Δ*btmD*[9] (the *botA* orthologue in this strain). As with *Streptomyces* sp. BC16019, Val3 of the core peptide was most tolerant of mutations, where mutations to Ala, Ile, Ser and Thr residues all led to bottromycin analogues with expected masses and tandem MS fragmentation patterns (Figure 8-14). As with the other mutant studies, no bottromycin analogues were produced by mutations to Asp7, which was targeted due to the reported importance of this residue for activity.

8.9 Conclusions

Over 60 years after their original discovery bottromycins continue to capture the attention of research groups and provide riddles to be solved. Given the developing antibiotic crisis and the effectiveness of bottromycins against (multi) drug-resistant pathogens, the recent advances in bottromycin research should be capitalized on: Through chemical synthesis and a combination of biochemistry/biotechnology new chemical space has become available to explore a more comprehensive SAR and address the *in vivo* instability of the compounds. In addition, the actual target within the A-site of the prokaryotic 50S ribosome has remained elusive and the MICs of some derivatives suggest an alternative, unknown target. We therefore think that despite its long history, the story of bottromycins may be just beginning.

8.10 References

1. Waisvisz, J. M.; van der Hoeven, M. G.; van Peppen, J.; Zwennis, W. C. M., Bottromycin. I. A New Sulfur-containing Antibiotic. *Journal of the American Chemical Society* **1957,** *79* (16), 4520-4521.

2. Kobayashi, Y.; Ichioka, M.; Hirose, T.; Nagai, K.; Matsumoto, A.; Matsui, H.; Hanaki, H.; Masuma, R.; Takahashi, Y.; Omura, S.; Sunazuka, T., Bottromycin derivatives: Efficient chemical modifications of the ester moiety and evaluation of anti-MRSA and anti-VRE activities. *Bioorg. Med. Chem. Lett.* **2010,** *20* (20), 6116-6120.

3. Otaka, T.; Kaji, A., Mode of action of bottromycin A2. Release of aminoacyl or peptidyl tRNA from ribosomes. *The Journal of biological chemistry* **1976,** *251*, 2299-306.

4. Otaka, T.; Kaji, A., Mode of action of bottromycin A2: Effect of bottromycin A2 on polysomes. *FEBS Letters* **1983,** *153* (1), 53-59.

5. Otaka, T.; Kaji, A., Mode of action of bottromycin A2: effect on peptide bond formation. *FEBS Letters* **1981,** *123* (2), 173-176.

6. Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Omura, S., Structure determination and total synthesis of bottromycin A2:a potent antibiotic against MRSA and VRE. *Angew. Chem., Int. Ed.* **2009,** *48* (5), 914-917.

7. Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chemical Science* **2012,** *3* (12), 3522-3525.

8. Hou, Y.; Tianero, M. D. B.; 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., Structure and Biosynthesis of the Antibiotic Bottromycin D. *Organic Letters* **2012,** *14* (19), 5050-5053.

9. Crone, W. J. K.; Leeper, F. J.; Truman, A. W., Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chemical Science* **2012,** *3* (12), 3516-3521.

10. Huo, L.; Rachid, S.; Stadler, M.; Wenzel, Silke C.; Müller, R., Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis. *Chemistry & Biology* **2012,** *19* (10), 1278-1287.

11. Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A., New developments in RiPP discovery, enzymology and engineering. *Natural Product Reports* **2020**.

12.   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.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports* **2013,** *30* (1), 108-160.

13.   J.N. Elgersma, I. N. R. e. I. J. H. v. B. t. D., Haag Werkwijze voor de bereiding van een antibiotische stof met behulp van actinomyceten van het geslacht Streptomyes 79749, 1955.

14.   Waisvisz, J. M.; van der Hoeven, M. G., The Chemistry and Partial Structure of Bottromycin. IV. *Journal of the American Chemical Society* **1958,** *80*, 383-385.

15.   Waisvisz, J. M.; Vanderhoeven, M. G.; Nijenhuis, B. T., The Structure of the Sulfur-containing Moiety of Bottromycin. *Journal of the American Chemical Society* **1957,** *79* (16), 4524-4527.

16.   Waisvisz, J. M. v. d. H., M. G.; Hölscher, J. F.; te Nijenhuis, B., Bottromycin. II. Preliminary Degradation Studies. *Journal of the American Chemical Society* **1957,** *79*, 4522-4524.

17.   Nakamura, S.; Chikaike, T.; Karasawa, K.; Tanaka, N.; Yonehara, H.; Umezawa, H., Isolation and characterization of bottromycins A and B. *The Journal of antibiotics, Ser. A* **1965,** *18*, 47-52.

18.   Nakamura, S.; Chikaike, T.; Yonehara, H.; Umezawa, H., Structures of Bottromycins A and B. *The Journal of antibiotics, Ser. A* **1965,** (18), 60-61.

19.   Nakamura, S.; Yajima, T.; Lin, Y.; Umezawa, H., Isolation and characterization of bottromycins A2, B2, C2. *J Antibiot (Tokyo)* **1967,** *20* (1), 1-5.

20.   Nakamura, S.; Tanaka, N.; Umezawa, H., Bottromycin A1, A2 and their structures. *J Antibiot (Tokyo)* **1966,** *19* (1), 10-12.

21.   Yamada, T.; Miyazawa, T.; Kuwata, S.; Watanabe, H., Studies of unusual amino acids and their peptides. VIII. The syntheses of an iminohexapeptide as a model of bottromycin and its related iminopeptides. *Bull. Chem. Soc. Jpn.* **1977,** *50* (7), 1827-30.

22.   Takahashi, Y.; Naganawa, H.; Takita, T.; Umezawa, H.; Nakamura, S., The revised structure of bottromycin A2. *J Antibiot (Tokyo)* **1976,** *29* (10), 1120-1123.

23.   Schipper, D., The revised structure of bottromycin A2 *Journal of Antibiotics* **1983,** *36* (8), 1076-1077.

24.   Gouda, H.; Kobayashi, Y.; Yamada, T.; Ideguchi, T.; Sugawara, A.; Hirose, T.; mura, S.; Sunazuka, T.; Hirono, S., Three-Dimensional Solution Structure of Bottromycin A2: A Potent Antibiotic Active against Methicillin-Resistant *Staphylococcus aureus* and Vancomycin-Resistant *Enterococci*. *Chemical and Pharmaceutical Bulletin* **2012,** *60* (2), 169-171.

25.   Tanaka, N.; Nishimura, T.; Nakamura, S.; Umezawa, H., Activity of bottromycin against Mycoplasma gallisepticum. *J Antibiot (Tokyo)* **1968,** *21* (1), 75-76.

26.   Park, S.-B.; Lee, I.; Suh, J.-W.; Kim, J.-G.; Lee, C., Screening and identification of antimicrobial compounds from Streptomyces bottropensis suppressing rice bacterial blight. *Journal of microbiology and biotechnology* **2011,** *21*, 1236-42.

27.   Tanaka, N.; Nishimura, T.; Nakamura, S.; Umezawa, H., Biological studies on bottromycin A and its hydrazide. *J. Antibiot., Ser. A* **1966,** *19* (4), 149-54.

28.   Wolf, F. J.; Miller, W. J. Amides of methobottromycin. 1972.

29.   Miller, W. J.; Chaiet, L.; Rasmussen, G.; Christensen, B.; Hannah, J.; Miller, A. K.; Wolf, F. J.; Wolf, F. J., Bottromycin. Separation of biologically active compounds and preparation and testing of amide derivatives. *J. Med. Chem.* **1968,** *11*, 746-9.

30.   Tanaka, N.; Sashikata, K.; Yamaguchi, H.; Umezawa, H., Inhibition of Protein Synthesis by Bottromycin A2 and Its Hydrazide. *The Journal of Biochemistry* **1966,** *60* (4), 405-410.

31.   Lin, Y.-C.; Tanaka, N., Mechanism of Action of Bottromycin in Polypeptide Biosynthesis. *The Journal of Biochemistry* **1968,** *63* (1), 1-7.

32.   Lin, Y.-C.; Kinoshita, T.; Tanaka, N., Mechanism of protein synthesis inhibition by bottromycin A2: studies with puromycin. . *J Antibiot (Tokyo)* **1968,** *21* (8), 471-476.

33.   Tanaka, N.; Lin, Y.-C.; Okuyama, A., Studies on translocation of F-met-tRNA and peptidyl-tRNA with antibiotics. *Biochemical and Biophysical Research Communications* **1971,** *44* (2), 477-483.

34.   Kinoshita, T.; Tanaka, N., On the site of action of bottromycin A2. . *J Antibiot (Tokyo)* **1970,** *23* (6), 311-312.

35.   Pestka, S.; Brot, N., Studies on the formation of transfer ribonucleic acid-ribosome complexes. IV. Effect of antibiotics on steps of bacterial protein synthesis: some new ribosomal inhibitors of translocation. *Journal of Biological Chemistry* **1971,** *246* (24), 7715-7722.

36.   Brodersen, D. E.; Clemons, W. M.; Carter, A. P.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., The Structural Basis for the Action of the Antibiotics Tetracycline, Pactamycin, and Hygromycin B on the 30S Ribosomal Subunit. *Cell* **2000,** *103* (7), 1143-1154.

37.   Polikanov, Y. S.; Szal, T.; Jiang, F.; Gupta, P.; Matsuda, R.; Shiozuka, M.; Steitz, T. A.; Vázquez-Laslop, N.; Mankin, A. S., Negamycin interferes with decoding and translocation by simultaneous interaction with rRNA and tRNA. *Mol Cell* **2014,** *56* (4), 541-50.

38.   Olivier, N. B.; Altman, R. B.; Noeske, J.; Basarab, G. S.; Code, E.; Ferguson, A. D.; Gao, N.; Huang, J.; Juette, M. F.; Livchak, S.; Miller, M. D.; Prince, D. B.; Cate, J. H. D.; Buurman, E. T.; Blanchard, S. C., Negamycin induces translational stalling and miscoding by binding to the small subunit head domain of the *Escherichia coli* ribosome. *Proceedings of the National Academy of Sciences* **2014,** *111* (46), 16274-16279.

39.   Polikanov, Y. S.; Aleksashin, N. A.; Beckert, B.; Wilson, D. N., The Mechanisms of Action of Ribosome-Targeting Peptide Antibiotics. *Frontiers in Molecular Biosciences* **2018,** *5* (48).

40.   Lin, J.; Zhou, D.; Steitz, T. A.; Polikanov, Y. S.; Gagnon, M. G., Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annual Review of Biochemistry* **2018,** *87* (1), 451-478.

41.   Kazmaier, U., The Long, Long Way to Bottromycin. *Israel Journal of Chemistry n/a* (n/a).

42.   Titouani, S. L.; Lavergne, J. P.; Viallefont, P.; Jacquier, R., New syntheses of L-amino acids. I. Stereospecific synthesis of L-proline, and of the cis- and trans-3- and 4-methyl L-prolines. *Tetrahedron* **1980,** *36* (20-21), 2961-5.

43.   Herdeis, C.; Hubmann, H. P.; Lotter, H., Synthesis of homochiral 3-substituted glutamic acids and prolines from pyroglutamic acid. *Tetrahedron: Asymmetry* **1994,** *5* (3), 351-4.

44.   Karoyan, P.; Chassaing, G., New strategy for the synthesis of 3-substituted prolines. *Tetrahedron Lett.* **1997,** *38* (1), 85-88.

45.   Kamenecka, T. M.; Park, Y.-J.; Lin, L. S.; Lanza, T.; Hagmann, W. K., Enantioselective approach to 3-substituted prolines. *Tetrahedron Lett.* **2001,** *42* (49), 8571-8573.

46.   Flamant-Robin, C.; Wang, Q.; Chiaroni, A.; Sasaki, N. A., An efficient method for the stereoselective synthesis of cis-3-substituted prolines: conformationally constrained α-amino acids. *Tetrahedron* **2002,** *58* (52), 10475-10484.

47.   Sharma, R.; Lubell, W. D., Regioselective enolization and alkylation of 4-oxo-N-(9-phenylfluoren-9-yl)proline: Synthesis of enantiopure proline-valine and hydroxyproline-valine chimeras. *Journal of Organic Chemistry* **1996,** *61* (1), 202-209.

48.   Medina, J. R.; Blackledge, C. W.; Erhard, K. F.; Axten, J. M.; Miller, W. H., Benzyl 2-cyano-3,3-dimethyl-1-pyrrolidinecarboxylate, a versatile intermediate for the synthesis of 3,3-dimethylproline derivatives. *Journal of Organic Chemistry* **2008,** *73* (10), 3946-3949.

49.   Bott, T. M.; Vanecko, J. A.; West, F. G., One-Carbon Ring Expansion of Azetidines via Ammonium Ylide 1,2-Shifts: A Simple Route to Substituted Pyrrolidines. *Journal of Organic Chemistry* **2009,** *74* (7), 2832-2836.

50.   Fuse, S.; Koinuma, H.; Kimbara, A.; Izumikawa, M.; Mifune, Y.; He, H.; Shin-ya, K.; Takahashi, T.; Doi, T., Total Synthesis and Stereochemistry Revision of Mannopeptimycin Aglycone. *J. Am. Chem. Soc.* **2014,** *136* (34), 12011-12017.

51. Houge-Frydrych, C. S. V.; Gilpin, M. L.; Skett, P. W.; Tyler, J. W., SB-203207 and SB-203208, two novel isoleucyl tRNA synthetase inhibitors from a Streptomyces sp. II. Structure determination. *J. Antibiot.* **2000,** *53* (4), 364-372.

52. Kataoka, Y.; Seto, Y.; Yamamoto, M.; Yamada, T.; Kuwata, S.; Watanabe, H., Studies of Unusual Amino Acids and Their Peptides. VI. The Syntheses and the Optical Resolutions of β-Methylphenylalanine and Its Dipeptide Present in Bottromycin *Bulletin of the Chemical Society of Japan* **1976,** *49* (4), 1081-1084.

53. Alias, M.; Lopez, M. P.; Cativiela, C., An efficient and stereodivergent synthesis of threo- and erythro-beta-methylphenylalanine. Resolution of each racemic pair by semipreparative HPLC. *Tetrahedron* **2004,** *60* (4), 885-891.

54. Grobuschek, N.; Schmid, M. G.; Tuscher, C.; Ivanova, M.; Gubitz, G., Chiral separation of beta-methyl-amino acids by ligand exchange using capillary electrophoresis and HPLC. *Journal of Pharmaceutical and Biomedical Analysis* **2002,** *27* (3-4), 599-605.

55. Nogle, L. M.; Mann, C. W.; Watts, W. L.; Zhang, Y. R., Preparative separation and identification of derivatized beta-methylphenylalanine enantiomers by chiral SFC, HPLC and NMR for development of new peptide ligand mimetics in drug discovery. *Journal of Pharmaceutical and Biomedical Analysis* **2006,** *40* (4), 901-909.

56. Peter, A.; Toth, G.; Torok, G.; Tourwe, D., Separation of enantiomeric beta-methyl amino acids and of beta-methyl amino acid containing peptides. *Journal of Chromatography A* **1996,** *728* (1-2), 455-465.

57. Torok, G.; Peter, A.; Armstrong, D. W.; Tourwe, D.; Toth, G.; Sapi, J., Direct chiral separation of unnatural amino acids by high-performance liquid chromatography on a ristocetin A-bonded stationary phase. *Chirality* **2001,** *13* (10), 648-656.

58. Vekes, E.; Torok, G.; Peter, A.; Sapi, J.; Tourwe, D., Indirect high-performance liquid chromatographic separation of stereoisomers of beta-alkyl-substituted amino acids by the application of (S)-N-(4-nitrophenoxycarbonyl)phenylalanine methoxyethyl ester as chiral derivatizing agent. *Journal of Chromatography A* **2002,** *949* (1-2), 125-139.

59. Ilisz, I.; Fodor, G.; Ivanyi, R.; Szente, L.; Toth, G.; Peter, A., Enantioseparation of β-methyl-substituted amino acids with cyclodextrins by capillary zone electrophoresis. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2008,** *875* (1), 273-279.

60. Jiang, C.; Armstrong, D. W.; Lantz, A. W.; Peter, A.; Toth, G., Enantiomeric separation of synthetic amino acids using capillary zone electrophoresis. *Journal of Liquid Chromatography & Related Technologies* **2007,** *30* (9-12), 1421-1436.

61. Torok, R.; Berkecz, R.; Peter, A., Enantioseparation of phenylalanine analogs on a quinine-based anion-exchanger chiral stationary phase: Structure and temperature effects. *Journal of Separation Science* **2006,** *29* (16), 2523-2532.

62. Ogawa, J.; Ryono, A.; Xie, S. X.; Vohra, R. M.; Indrati, R.; Akamatsu, M.; Miyagawa, H.; Ueno, T.; Shimizu, S., Separative preparation of the four stereoisomers of β-methylphenylalanine with N-carbamoyl amino acid amidohydrolases. *J. Mol. Catal. B: Enzym.* **2001,** *12* (1-6), 71-75.

63. Tsuchihashi, G.; Mitamura, S.; Ogura, K., Synthesis of (S,S)-2-Amino-3-phenylbutyric Acid. *Bulletin of the Chemical Society of Japan* **1979,** *52* (7), 2167-2168.

64. Dharanipragada, R.; Nicolas, E.; Toth, G.; Hruby, V. J., Asymmetric synthesis of unusual amino acids: Synthesis of optically pure isomers of β-methylphenylalanine. *Tetrahedron Letters* **1989,** *30* (49), 6841-6844.

65. Dharanipragada, R.; Vanhulle, K.; Bannister, A.; Bear, S.; Kennedy, L.; Hruby, V. J., Asymmetric synthesis of unusual amino acids: An efficient synthesis of optically pure isomers of β-methylphenylalanine. *Tetrahedron* **1992,** *48* (23), 4733-4748.

66. Oppolzer, W.; Moretti, R., Enantioselective Syntheses of α-Amino Acids from 10-(Aminosulfonyl)-2-bornyl Esters and Di(tert-butyl) Azodicarboxylate. *Tetrahedron* **1988,** *44* (17), 5541-5552.

67. Medina, E.; Moyano, A.; Pericas, M. A.; Riera, A., Synthesis of N-Boc-β-Aryl Alanines and of N-Boc-β-Methyl-β-aryl Alanines by Regioselective Ring-Opening of Enantiomerically Pure N-Boc-Aziridines. *J. Org. Chem.* **1998,** *63* (23), 8574-8578.

68. O'Donnell, M. J.; Cooper, J. T.; Mader, M. M., Acyclic Stereoselective Boron Alkylation Reactions for the Asymmetric Synthesis of β-Substituted α-Amino Acid Derivatives. *J. Am. Chem. Soc.* **2003,** *125* (9), 2370-2371.

69. Enright, A.; Alexandre, F.-R.; Roff, G.; Fotheringham, I. G.; Dawson, M. J.; Turner, N. J., Stereoinversion of β- and γ-substituted α-amino acids using a chemo-enzymatic oxidation-reduction procedure. *Chem. Commun. (Cambridge, U. K.)* **2003,** (20), 2636-2637.

70. Ooi, T.; Kato, D.; Inamura, K.; Ohmatsu, K.; Maruoka, K., Practical Stereoselective Synthesis of β-Branched α-Amino Acids through Efficient Kinetic Resolution in the Phase-Transfer-Catalyzed Asymmetric Alkylations. *Org. Lett.* **2007,** *9* (20), 3945-3948.

71. Nakamura, S.; Chikaike, T.; Yonehara, H.; Umezawa, H., Isolation, Characterization and Structural Elucidation of New Amino Acids from Bottromycin A. *Chemical & Pharmaceutical Bulletin* **1965,** *13* (5), 599-602.

72. Seto, Y.; Torii, K.; Bori, K.; Inabata, K.; Kuwata, S.; Watanabe, H., Studies of Unusual Amino Acids and Their Peptides. VI. The Syntheses and the Optical Resolutions of β-Methylphenylalanine and Its Dipeptide Present in Bottromycin *Bulletin of the Chemical Society of Japan* **1974,** *47* (1), 151-155.

73. Davis, F. A.; Reddy, R. E.; Szewczyk, J. M., Asymmetric Synthesis of (R)-(+)-.beta.-Phenylalanine from (S)-(+)-Benzylidene-p-toluenesulfinamide. Regeneration of the Sulfinimine Precursor. *Journal of Organic Chemistry* **1995,** *60* (21), 7037-7039.

74. Tang, T. P.; Ellman, J. A., The tert-butanesulfinyl group: An ideal chiral directing group and boc-surrogate for the asymmetric synthesis and applications of beta-amino acids. *Journal of Organic Chemistry* **1999,** *64* (1), 12-13.

75. Yamada, T.; Yagita, M.; Kobayashi, Y.; Sennari, G.; Shimamura, H.; Matsui, H.; Horimatsu, Y.; Hanaki, H.; Hirose, T.; Omura, S.; Sunazuka, T., Synthesis and Evaluation of Antibacterial Activity of Bottromycins. *J. Org. Chem.* **2018,** *83* (13), 7135-7149.

76.   Yamada, T.; Suegane, K.; Kuwata, S.; Watanabe, H., Studies of unusual amino acids and their peptides. VII. The syntheses and the reactions of iminopeptides. *Bull. Chem. Soc. Jpn.* **1977,** *50* (5), 1088-93.

77.   Nakamura, S.; Chikaike, T.; Yonehara, H.; Umezawa, H., STRUCTURES OF BOTTROMYCINS A AND B. *J Antibiot (Tokyo)* **1965,** *18*, 60-1.

78.   Nakamura, S.; Umezawa, H., The structure of bottromycin A2, a new component of bottromycins *Chemical & Pharmaceutical Bulletin* **1966,** *14* (9), 981-986.

79.   Kaneda, M., Studies on bottromycins. I. 1H and 13C NMR assignments of bottromycin A2, the main component of the complex *Journal of Antibiotics* **1992,** *45* (5), 792-796.

80.   Ackermann, S.; Lerchen, H.-G.; Haebich, D.; Ullrich, A.; Kazmaier, U., Synthetic studies towards bottromycin. *Beilstein J. Org. Chem.* **2012,** *8*, 1652-1656, No. 189.

81.   Kazmaier, U.; Hebach, C., Peptide syntheses via Ugi reactions with ammonia. *Synlett* **2003,** (11), 1591-1594.

82.   Pick, R.; Bauer, M.; Kazmaier, U.; Hebach, C., Ammonia in Ugi reactions - Four-component versus six-component couplings. *Synlett* **2005,** (5), 757-760.

83.   Bauer, M.; Maurer, F.; Hoffmann, S. M.; Kazmaier, U., Hydroxyalkyl Thiazolines, a New Class of Highly Efficient Ligands for Carbonyl Additions. *Synlett* **2008,** (20), 3203-3207.

84.   Kazmaier, U.; Ackermann, S., A straightforward approach towards thiazoles and endothiopeptides via Ugi reaction. *Organic & Biomolecular Chemistry* **2005,** *3* (17), 3184-3187.

85.   Kazmaier, U.; Persch, A., A straightforward approach towards 5-substituted thiazolylpeptides via the thio-Ugi-reaction. *Organic & Biomolecular Chemistry* **2010,** *8* (23), 5442-5447.

86.   Hata, F.; Matsumae, A.; Abe, K.; Sano, Y.; Otani, M.; Omura, S. Fermentative preparation of bottromycin. JP47010036B, 1972.

87.   Lerchen, H.-G.; Schiffer, G.; Broetz-Oesterhelt, H.; Mayer-Bartschmid, A.; Eckermann, S.; Freiberg, C.; Endermann, R.; Schuhmacher, J.; Meier, H.; Svenstrup, N.; Seip, S.; Gehling, M.; Haebich, D. Fermentative and chemical synthesis of bottromycin derivative analogs for use in treatment or prevention of bacterial infections in humans or animals. WO2006103010A1, 2006.

88.   Kellenberger, J. On the Stereochemical Course of the C-methylation Steps in Bottromycin Biosynthesis. Ph.D. thesis, ETH Switzerland, 1997.

89.   Mosimann, H.; Kräutler, B., Methylcorrinoids Methylate Radicals-Their Second Biological Mode of Action? *Angewandte Chemie (International ed. in English)* **2000,** *39* (2), 393-395.

90.   Freeman, M. F.; Gurgui, C.; Helf, M. J.; Morinaka, B. I.; Uria, A. R.; Oldham, N. J.; Sahl, H.-G.; Matsunaga, S.; Piel, J., Metagenome Mining Reveals Polytheonamides as Posttranslationally Modified Ribosomal Peptides. *Science* **2012,** *338* (6105), 387-390.

91.   Mahlert, C.; Kopp, F.; Thirlway, J.; Micklefield, J.; Marahiel, M. A., Stereospecific Enzymatic Transformation of α-Ketoglutarate to (2S,3R)-3-Methyl Glutamate during Acidic Lipopeptide Biosynthesis. *Journal of the American Chemical Society* **2007,** *129* (39), 12011-12018.

92.   Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* **2016,** *34* (8), 828-837.

93.   Crone, W. J. K.; Vior, N. M.; Santos-Aberturas, J.; Schmitz, L. G.; Leeper, F. J.; Truman, A. W., Dissecting Bottromycin Biosynthesis Using Comparative Untargeted Metabolomics. *Angewandte Chemie International Edition* **2016,** *55* (33), 9639-9643.

94.   Matsui, M.; Fowler, J. H.; Walling, L. L., Leucine aminopeptidases: diversity in structure and function. *Biol Chem* **2006,** *387* (12), 1535-44.

95.   Dorus, S.; Wilkin, E. C.; Karr, T. L., Expansion and functional diversification of a leucyl aminopeptidase family that encodes the major protein constituents of Drosophila sperm. *BMC Genomics* **2011,** *12*, 177-177.

96.   Frottin, F.; Martinez, A.; Peynot, P.; Mitra, S.; Holz, R. C.; Giglione, C.; Meinnel, T., The proteomics of N-terminal methionine cleavage. *Molecular & cellular proteomics : MCP* **2006,** *5* (12), 2336-49.

97.   Mann, G.; Huo, L.; Adam, S.; Nardone, B.; Vendome, J.; Westwood, N. J.; Müller, R.; Koehnke, J., Structure and Substrate Recognition of the Bottromycin Maturation Enzyme BotP. *ChemBioChem* **2016,** *17* (23), 2286-2292.

98.   Oman, T. J.; van der Donk, W. A., Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat Chem Biol* **2010,** *6* (1), 9-18.

99. Lee, S. W.; Mitchell, D. A.; Markley, A. L.; Hensler, M. E.; Gonzalez, D.; Wohlrab, A.; Dorrestein, P. C.; Nizet, V.; Dixon, J. E., Discovery of a widely distributed toxin biosynthetic gene cluster. *Proc Natl Acad Sci U S A* **2008,** *105* (15), 5879-84.

100. Wieland Brown, L. C.; Acker, M. G.; Clardy, J.; Walsh, C. T.; Fischbach, M. A., Thirteen posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin. *Proc Natl Acad Sci U S A* **2009,** *106* (8), 2549-53.

101. Schmidt, E. W.; Nelson, J. T.; Rasko, D. A.; Sudek, S.; Eisen, J. A.; Haygood, M. G.; Ravel, J., Patellamide A and C biosynthesis by a microcin-like pathway in Prochloron didemni, the cyanobacterial symbiont of Lissoclinum patella. *Proc Natl Acad Sci U S A* **2005,** *102* (20), 7315-20.

102. Goto, Y.; Suga, H., In Vitro Biosynthesis of Peptides Containing Exotic Azoline Analogues. *ChemBioChem* **2020,** *21* (1-2), 84-87.

103. 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. M.; Jaspars, M.; Naismith, J. H., An Enzymatic Route to Selenazolines. *ChemBioChem* **2013,** *14* (5), 564-567.

104. Schwalen, C. J.; Hudson, G. A.; Kosol, S.; Mahanta, N.; Challis, G. L.; Mitchell, D. A., In Vitro Biosynthetic Studies of Bottromycin Expand the Enzymatic Capabilities of the YcaO Superfamily. *Journal of the American Chemical Society* **2017,** *139* (50), 18154-18157.

105. Franz, L.; Adam, S.; Santos-Aberturas, J.; Truman, A. W.; Koehnke, J., Macroamidine Formation in Bottromycins Is Catalyzed by a Divergent YcaO Enzyme. *J Am Chem Soc* **2017,** *139* (50), 18158-18161.

106. Sikandar, A.; Franz, L.; Melse, O.; Antes, I.; Koehnke, J., Thiazoline-Specific Amidohydrolase PurAH Is the Gatekeeper of Bottromycin Biosynthesis. *Journal of the American Chemical Society* **2019,** *141* (25), 9748-9752.

107. Vior, N. M.; Cea-Torrescassana, E.; Eyles, T. H.; Chandra, G.; Truman, A. W., Regulation of Bottromycin Biosynthesis Involves an Internal Transcriptional Start Site and a Cluster-Situated Modulator. *Frontiers in Microbiology* **2020,** *11* (495).

108. Mindrebo, J. T.; Nartey, C. M.; Seto, Y.; Burkart, M. D.; Noel, J. P., Unveiling the functional diversity of the alpha/beta hydrolase superfamily in the plant kingdom. *Curr Opin Struct Biol* **2016,** *41*, 233-246.

109. Rauwerdink, A.; Kazlauskas, R. J., How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: the Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. *ACS Catal* **2015,** *5* (10), 6153-6176.

110. Sikandar, A.; Franz, L.; Adam, S.; Santos-Aberturas, J.; Horbal, L.; Luzhetskyy, A.; Truman, A. W.; Kalinina, O. V.; Koehnke, J., The bottromycin epimerase BotH defines a group of atypical α/β-hydrolase-fold enzymes. *Nature Chemical Biology* **2020**.

111. Morinaka, B. I.; Vagstad, A. L.; Helf, M. J.; Gugger, M.; Kegler, C.; Freeman, M. F.; Bode, H. B.; Piel, J., Radical S-Adenosyl Methionine Epimerases: Regioselective Introduction of Diverse D-Amino Acid Patterns into Peptide Natural Products. *Angewandte Chemie International Edition* **2014,** *53* (32), 8503-8507.

112.   Benjdia, A.; Guillot, A.; Ruffié, P.; Leprince, J.; Berteau, O., Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in Bacillus subtilis. *Nature Chemistry* **2017,** *9* (7), 698-707.

113.   Yang, X.; van der Donk, W. A., Post-translational Introduction of d-Alanine into Ribosomally Synthesized Peptides by the Dehydroalanine Reductase NpnJ. *Journal of the American Chemical Society* **2015,** *137* (39), 12426-12429.

114.   Huo, L.; van der Donk, W. A., Discovery and Characterization of Bicereucin, an Unusual d-Amino Acid-Containing Mixed Two-Component Lantibiotic. *Journal of the American Chemical Society* **2016,** *138* (16), 5254-5257.

115.   Adam, S.; Franz, L.; Milhim, M.; Bernhardt, R.; Kalinina, O. V.; Koehnke, J., Characterization of the Stereoselective P450 Enzyme BotCYP Enables the In Vitro Biosynthesis of the Bottromycin Core Scaffold. *Journal of the American Chemical Society* **2020,** *142* (49), 20560-20565.

116.   Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A., In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *Journal of the American Chemical Society* **2015,** *137* (51), 16012-16015.

117.   Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A., Biosynthetic Timing and Substrate Specificity for the Thiopeptide Thiomuracin. *Journal of the American Chemical Society* **2016,** *138* (48), 15511-15514.

118.   Zheng, Q.; Wang, Q.; Wang, S.; Wu, J.; Gao, Q.; Liu, W., Thiopeptide Antibiotics Exhibit a Dual Mode of Action against Intracellular Pathogens by Affecting Both Host and Microbe. *Chem Biol* **2015,** *22* (8), 1002-7.

119.   Bai, X.; Guo, H.; Chen, D.; Yang, Q.; Tao, J.; Liu, W., Isolation and structure determination of two new nosiheptide-type compounds provide insights into the function of the cytochrome P450 oxygenase NocV in nocathiacin biosynthesis. *Organic Chemistry Frontiers* **2020,** *7* (3), 584-589.

120.   Young, T. S.; Walsh, C. T., Identification of the thiazolyl peptide GE37468 gene cluster from *Streptomyces* ATCC 55365 and heterologous expression in *Streptomyces lividans*. *Proceedings of the National Academy of Sciences* **2011,** *108* (32), 13053-13058.

121.   Hug, J. J.; Dastbaz, J.; Adam, S.; Revermann, O.; Koehnke, J.; Krug, D.; Müller, R., Biosynthesis of Cittilins, Unusual Ribosomally Synthesized and Post-translationally Modified Peptides from Myxococcus xanthus. *ACS Chemical Biology* **2020,** *15* (8), 2221-2231.

122.   Horbal, L.; Marques, F.; Nadmid, S.; Mendes, M. V.; Luzhetskyy, A., Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metab Eng* **2018,** *49*, 299-315.

123.   Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Nagaraja Rao, R.; Schoner, B. E., Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* **1992,** *116* (1), 43-49.

124.   Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F., PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of

the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences* **2003,** *100* (4), 1541-1546.

125. Hu, H.; Ochi, K., Novel Approach for Improving the Productivity of Antibiotic-Producing Strains by Inducing Combined Resistant Mutations. *Applied and Environmental Microbiology* **2001,** *67* (4), 1885-1892.

126. Hu, H.; Zhang, Q.; Ochi, K., Activation of Antibiotic Biosynthesis by Specified Mutations in the *rpoB* Gene (Encoding the RNA Polymerase β Subunit) of *Streptomyces lividans*. *Journal of Bacteriology* **2002,** *184* (14), 3984-3991.

127. Quistgaard, E. M.; Löw, C.; Guettou, F.; Nordlund, P., Understanding transport by the major facilitator superfamily (MFS): structures pave the way. *Nature Reviews Molecular Cell Biology* **2016,** *17* (2), 123-132.

128. Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S., Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proceedings of the National Academy of Sciences* **2014,** *111* (5), 1957-1962.

129. Noskov, V.; Kouprina, N.; Leem, S.-H.; Koriabine, M.; Barrett, J. C.; Larionov, V., A genetic system for direct selection of gene-positive clones during recombinational cloning in yeast. *Nucleic Acids Research* **2002,** *30* (2), e8-e8.

130. Eyles, T. H.; Vior, N. M.; Truman, A. W., Rapid and Robust Yeast-Mediated Pathway Refactoring Generates Multiple New Bottromycin-Related Metabolites. *ACS Synthetic Biology* **2018,** *7* (5), 1211-1218.

131. Gomez-Escribano, J. P.; Bibb, M. J., Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. *Microbial Biotechnology* **2011,** *4* (2), 207-215.

132. Rudolph, M. M.; Vockenhuber, M.-P.; Suess, B., Synthetic riboswitches for the conditional control of gene expression in Streptomyces coelicolor. *Microbiology* **2013,** *159* (Pt_7), 1416-1422.

133. Siegl, T.; Tokovenko, B.; Myronovskyi, M.; Luzhetskyy, A., Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metabolic Engineering* **2013,** *19*, 98-106.

134. Strohl, W. R., Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Research* **1992,** *20* (5), 961-974.

135. Hudson, G. A.; Mitchell, D. A., RiPP antibiotics: biosynthesis and engineering potential. *Curr Opin Microbiol* **2018,** *45*, 61-69.

136. Crone, W. J. K. Investigations into the biosynthesis of bottromycin University of Cambridge Cambridge, 2015.