Published in final edited form as:

Chembiochem. 2010 July 5; 11(10): 1413–1421. doi:10.1002/cbic.201000196.

Marine Molecular Machines: Heterocyclization in Cyanobactin Biosynthesis

John A. McIntosh and Prof. Eric W. Schmidt^[a]

Eric W. Schmidt: ews1@utah.edu

^[a]J.A. McIntosh, and E.W. Schmidt, Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, UT USA 84112, Fax: (+1) 801-585-9119

Abstract

Natural products that contain amino acid-derived (Cys, Ser, Thr) heterocycles are ubiquitous in nature, yet key aspects of their biosynthesis remain undefined. Cyanobactins are heterocyclic ribosomal peptide natural products from cyanobacteria, including symbiotic bacteria living with marine ascidians. In contrast to other ribosomal peptide heterocyclases that have been studied, the cyanobactin heterocyclase is a single protein that does not require an oxidase enzyme. Using this simplifying condition, we provide new evidence to support the hypothesis that these enzymes are molecular machines, using ATP in a product binding or orientation cycle. Further, we show that both protease inhibitors and ATP analogues inhibit heterocyclization and define the order of biochemical steps in the cyanobactin biosynthetic pathway. The cyanobactin pathway enzymes, PatD and TruD, are thiazoline and oxazoline synthetases.

Keywords

patellamide; trunkamide; microcin; thiazole; thiazoline

“A heterocycle sounds like a wonderful thing to ride, especially with someone you love.”^[1]

— George Carlin

Introduction

The heterocyclic thiazole, thiazoline, oxazole, and oxazoline motifs are commonly found in natural products possessing diverse biological activities. These Cys-, Ser-, and Thr-derived heterocycles are present in approved drugs, as well as in drug leads and in toxins produced by human pathogens.^[2-5] For example, in the ribosomal peptide group microcin B17 is a DNA gyrase inhibitor produced by *Escherichia coli*,^[6] thiostrepton and relatives are potent antibiotics,^[7] and cyanobactins are ubiquitous bioactive peptides from cyanobacteria (Figure 1).^[8] More recently, heterocyclic relatives of the cyanobactin / microcin B17 group have been shown to occur commonly in diverse bacteria.^[5, 9] In the nonribosomal peptide group, an epothilone derivative is an FDA-approved anti-cancer agent, bacitracin is used as an antibiotic, and certain siderophores are key virulence factors (Figure 1).^[2, 3] In addition, the heterocyclic motif itself is potentially bioactive depending upon context: thiazol(in)e and oxazol(in)e can interact with nucleic acid, protein, or metal ligands.^[3] Thus, routes to their enzymatic synthesis are of keen interest.

Correspondence to: Eric W. Schmidt, ews1@utah.edu.

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

Despite extensive and groundbreaking studies of heterocyclization enzymes in both ribosomal and nonribosomal peptide natural products,^[10, 11] their chemical mechanism(s) remain unknown.^[12] Previous biochemical studies of biosynthesis in heterocycle-containing ribosomal peptides involve a complex of three proteins whose activities are intertwined and difficult to separate. In particular, ATPase, oxidase, and Zn-binding domains are invariant requirements for heterocycle biosynthesis in the cases of microcin B17 and streptolysin S.^[5, 13] The enzymes modify microcin and streptolysin precursor peptides in multiple positions by synthesizing oxidized heterocycles, thiazoles and oxazoles, and not their unoxidized presumed precursors, thiazolines, and oxazolines. Of the apparent enzymatic activities (oxidase, ATPase, and heterocyclase), none has been observed in isolation from the others, and while ATPase and oxidase activities are readily attributable to specific proteins, their respective roles in enzymatic heterocyclization are unclear. An oxidase is required, explaining the lack of observed thiazoline or oxazoline either in vitro or in vivo. It has been proposed that ATP consumption fuels a molecular machine driving heterocyclization, although a direct effect on catalysis has not been ruled out. These conserved protein domains also appear to be present in thiopeptide biosynthesis,^[7] though to our knowledge, they have not yet been studied in vitro. Heterocyclization has also been studied in nonribosomal peptide systems,^[10, 14] where the enzymes and resulting chemical mechanisms (excepting the oxidase) are non-homologous to the ribosomal peptide group.

Here we present results regarding the biosynthesis of heterocycles among the cyanobactins, which are a group of circular, heterocyclic, ribosomally derived peptides from cyanobacteria, including the marine animal symbionts, *Prochloron spp.*^[8] Two pathways to *Prochloron*-derived cyanobactins exist: the patellamide (*pat*) pathway, whose members contain heterocycles derived from Cys, Ser, and Thr, and the trunkamide pathway (*tru*), whose members contain heterocycles derived only from Cys. Heterocycle oxidation is somewhat variable, with both oxidized (thiazole and oxazole) and unoxidized (thiazoline and oxazoline) heterocycles present in both families. By contrast, the products of the microcin pathway, and presumably the streptolysin pathways, are always oxidized, although the latter has not been completely defined.

The biosynthetic gene clusters of the patellamide and trunkamide pathways have been previously characterized, as has the biochemical basis of circularization and certain aspects of heterocyclization.^[12, 15] In brief, PatE and TruE are precursor peptides, each encoding two natural products on a single short peptide. Importantly, PatE and TruE both contain a leader peptide sequence and “enzyme recognition” sequences that flank the N- and C-termini of the product coding cassettes (Figure S1). PatD and TruD are heterocyclases that, in vivo, operate regioselectively to modify Cys, Ser, and Thr in the case of PatD, but only Cys in the case of TruD. PatA or TruA proteases cleave N-terminal recognition sequences of cassettes 1 and 2,^[15] though it was not clear prior to this study whether PatA/TruA acted prior to or after heterocyclization. Lastly, PatG or TruG proteases cleave the C-terminal recognition sequences in tandem with macrocyclization (Figure S2).^[15] All three of these enzyme groups are capable of modifying a diverse set of cassette mutants to yield libraries of natural products.^[16]

TruD, which in vivo heterocyclizes only Cys residues, is a didomain protein with an N-terminal region that bears distant similarity to heterocyclase proteins in the streptolysin and microcin systems. The C-terminal region is very distantly similar to the protein proposed to be responsible for peptide binding and ATPase activity in other heterocyclization systems. PatD is >99% identical to TruD in the N-terminal domain, but only 77% identical in the C-terminal domain (Figure 2). No oxidase domain is present in TruD or in the *tru* pathway, and unlike the microcin or streptolysin cases oxidation is not required to reconstitute heterocyclase activity. Thus, TruD is a single protein that fully reconstitutes

heterocyclization activity in a regioselective manner. The relative simplicity of this system in comparison to other studied heterocyclases enabled us to obtain new insights into the function of this important enzyme family. Here, we present evidence that strongly supports the molecular machine hypothesis regarding the role of ATP, determine the order of enzymatic steps en route to cyanobactin synthesis, and describe inhibitors of heterocyclization.

Results and Discussion

Metal and Cofactor Requirements of PatD and TruD

The enzymes PatD and TruD, and substrates TruE2 and TruE4 were cloned as described elsewhere.^[17] Two point mutations were found in *patD* in these studies, but they lacked any apparent functional consequence. In addition, *patD* was used as a template to clone *truD*, ensuring that both PatD and TruD are 100% identical in their N-terminal catalytic domains.

As previously described, when incubated with substrates, TruD and PatD are fully competent heterocyclase enzymes.^[18] A robust SDS-PAGE assay was developed in which heterocyclization can clearly be tracked by mobility shift (Figure S3), as described elsewhere.^[18] In general, TruD products migrate more rapidly by SDS-PAGE analysis than unreacted precursor, while PatD products migrate more rapidly still. This assay allowed enzyme requirements and timing to be rigorously defined.

All enzyme reactions described in this work, including determination of enzyme requirements, substrate and product measurements, and kinetic analysis, were performed at minimum in triplicate in independent runs. ATP, MgCl₂, and DTT were found to be necessary for the heterocyclization reaction with both PatD and TruD. The minimum Mg²⁺ concentration that could support catalysis was 1 mM, which is somewhat lower than the reported Mg²⁺ requirements of microcin B17 and streptolysin S synthetases (2-20 mM Mg).^[5, 11] Ultrapure MgCl₂ was used in some experiments and was found to support catalysis. Additionally, other additives (Tris, ATP, and DTT) could be passed through Chelex resin without inhibiting catalysis, indicating that traces of other metals were not required for catalysis.

Like McbB from the microcin B17 pathway,^[19] PatD and TruD are both strongly associated with Zn even after extensive dialysis, as indicated by ICP-OES experiments. The apparent binding stoichiometry of the enzyme-metal complex was found to be roughly 1 mol of Zn per mol of enzyme. The role of this Zn is unknown but has been proposed to be structural in the microcin B17 context.^[19] No other strong associations of metals with enzyme were found using ICP-OES. Taken together, this indicates that Zn and Mg are the sole metal cofactors required to catalyze reaction.

The fate of ATP was probed by HPLC analysis. During the course of the reaction, ATP was shown to be hydrolyzed to ADP, as was found for microcin B17 synthetase (Figure S4).^[20] Overall, these requirements are quite similar to those defined in microcin B17 biosynthesis, indicating that the enzymes could function in a similar manner despite their nearly complete lack of protein sequence similarity. The major differences between microcin B17 biosynthetic enzymes and PatD/TruD are (1) the lack of a requirement for an oxidase domain in our system, and (2) the presence of putative heterocyclase and peptide-binding domains on a single polypeptide here, whereas in the microcin B17 system these domains exist as separate polypeptides.

Order of Heterocyclization Events in TruE2

Enzymatic reactions with TruE2 and TruD or PatD, reactions generally went to completion within 2 hours. Time-course experiments from 15 min to 24 hours revealed that Cys reactivity was fast, with TruD-TruE2 reactions being complete within 1 hour. Catalysis of the third dehydration event (Thr cyclization) by PatD was slower, requiring up to 2 hours for complete modification (Figure S5). No further reaction was observed with extended incubation periods, even with batch addition of further enzyme.

In time-course experiments with TruD, TruE2 was observed to proceed directly to the doubly dehydrated product. Only by supplying inadequate amounts of ATP could we observe an appreciable accumulation of singly-dehydrated product, and even under those conditions the amount of doubly-dehydrated product was still greater (Figure S6). These experiments are most consistent with the idea that the substrate can dissociate from enzyme between heterocyclizations, but that the singly-heterocyclized product is a much faster reactant than the unmodified precursor peptide.

By contrast, heterocyclization of TruE2 to afford a third heterocycle (oxazoline) by PatD was slower, and intermediates could be captured and readily observed both by SDS-PAGE and mass spectrometry. Quite clearly then, the precursor peptide leaves the enzyme after the second heterocyclization and before the third heterocyclization. One of the most convincing pieces of evidence was found upon incubation of PatD with TruE2. A sample taken before the PatD reaction was complete showed three bands (listed in order of increasing mobility): (1) unmodified TruE2 (2) a band consistent with two thiazolines, and (3) the fully modified TruE2.

Relative rates have also been determined in microcin B17 biosynthesis, and the results observed in that system are quite similar to what we present here: in both cases, the reaction to form oxygen-containing heterocycles is slower, as would be expected from the reduced nucleophilicity of hydroxyl in comparison to thiol.^[20]

Timing of Biosynthetic Steps in Cyanobactin Synthesis

Previous work has shown that PatA and PatG enzymes, which cleave and circularize PatE, accept broadly different substrates.^[15, 16] A mystery has been how these enzymes produce only the natural products and not other derivatives. One idea, consistent with work on microcin B17 and previous cyanobactin coexpression experiments, is that the enzymes could form a complex that would sequester substrates. In numerous conditions, however, we could not observe any requirement for complex formation. For example, PatA, PatG, and PatD are all competent catalysts without the addition of other enzymes, and addition of multiple enzymes does not appreciably increase reaction rate of single steps. Pull-down experiments using various Pat proteins as bait in different conditions were unsuccessful, as were experiments involving co-expression. Thus, there is no evidence that protein complexes are required for cyanobactin biosynthesis.

Nonetheless, when PatA is co-incubated with TruD or PatD and TruE2 or other precursor peptides, we do not detect any PatA cleavage fragments that lack heterocycle modifications. This observation was surprising given that PatA is capable of cleaving unmodified precursor peptides in the absence of PatD or TruD.^[15] To further probe this issue, PatA was added to the reaction mixture and allowed to react prior to the addition of PatD or TruD, and vice versa. Under no condition could we observe predicted dehydration products if PatA was allowed to react first, cleaving the leader sequence, whereas the expected products were obtained if PatD or TruD were added first or if the A / D combinations were co-incubated. Therefore, it appears that the leader sequence, which PatA cleaves, is required in cis for PatD/TruD modification, as observed in the case of microcin B17.^[21, 22] In other words, the

products of PatD/TruD are substrates for PatA, but the reverse is not true. Thus, the fidelity of cyanobactin synthesis is most likely encoded at the level of substrate recognition, and not at the level of protein complexes. Through these experiments, the order of catalytic events in cyanobactin synthesis was shown to be heterocyclization, followed by N-C circularization. The relative preference of PatD or TruD for heterocyclizing certain positions in competition with the rate of PatA leader-sequence cleavage dictates the structures of the natural products. The precise timing of *O*-prenylation (as observed in the *tru* pathway) and heterocycle oxidation (as obtains in the *pat* pathways) remains unclear, but oxidation must take place after heterocyclization, and prenylation also seems likely to occur after heterocyclization.

In comparison, these additional protease events are not tied to the microcin B17 gene cluster,^[23] nor are microcin B17 or streptolysin S further modified beyond heterocycle synthesis except for (as-yet undefined in the case of streptolysin S) leader sequence cleavage.

Kinetic Analysis of TruD

TruD was used for all rate experiments because it is more readily purified compared with PatD, and its reactivity (Cys-only in natural precursor peptides) is substantially simpler than that of PatD. Rates of reaction were analyzed using varying ATP, TruE2, and TruD concentrations. With highly purified TruD, the background hydrolysis of ATP due to enzyme in absence of substrate was virtually nonexistent. The background with TruE2 alone is quite small, but is measurable. By contrast, when substrate and enzyme are coincubated, ATP hydrolysis occurs at a robust pace. Reactions proceeded at an essentially linear rate for the conditions attempted in the first 40 min. Therefore, reactions were sampled at 0, 20, and 40 min and analyzed by SDS-PAGE for TruE2 turnover and HPLC for ADP formation. Results of this analysis were fit using the Solver function in Excel to estimate kinetic constants.

Using variable TruE2 concentrations, from 2.5 to 7.4 μM , TruE2 K_m was estimated as $\sim 1 \mu\text{M}$. The apparent k_{cat} for ATP hydrolysis under these conditions was 2.6 min^{-1} (Figures S7 and S8). This study was not designed to specifically measure the k_{cat} of the enzyme for TruE2, which is a complex problem. For example, the clear dissociation of the enzyme-substrate complex — especially with oxazoline synthesis where the rate of reaction differs — indicates that different intermediates have different k_{cat} values, and possibly even different K_m values. However, it was possible to estimate a turnover number using saturating substrate and enzyme conditions. Using standard reaction conditions with 7.4 μM TruE2 and 140 nM TruD, the reaction proceeded to $\sim 90\%$ completion at $t=60 \text{ min}$ (Figure 3). Based upon this experiment, a TruE2 turnover number of $\sim 0.8 \text{ enz}^{-1}\text{min}^{-1}$, or a heterocycle turnover number of $\sim 1.6 \text{ enz}^{-1}\text{min}^{-1}$, could be calculated. This is not a true k_{cat} but gives an estimate of the speed of the enzyme under the reaction conditions.

Using variable ATP from 200 to 800 μM , the K_m for ATP was estimated as 300 μM . The calculated k_{cat} value was 2.4 min^{-1} , which is in excellent agreement for that calculated using variable TruE2 concentration. These rates also scaled precisely as enzyme concentration was doubled. At 104 nM TruD, TruE2 reactions were 1.5 times faster than at 69 nM TruD; at 140 nM TruD, reactions were 2.1 times faster than 69 nM TruD.

The kinetic constants described above are strikingly similar to those reported for microcin B17 synthetase,^[11, 20] despite the greatly different experimental conditions and the absence of both a protein complex and an oxidase here. In particular, the relative K_m (1 μM) measured here using kinetic methods is similar to that reported for microcin B17 (2.3 μM),^[11] but much higher than that reported for the streptolysin S leader peptide (6.7 nM),

which was measured using surface plasmon resonance.^[24] Overall, these studies indicate that the enzymes function similarly, and that results reported here likely are applicable to the distant protein relatives previously studied.

Inhibition of Heterocyclization

In the course of our studies, two types of inhibitors were shown to slow heterocyclization: protease inhibitors AEBSF and PMSF; and a non-hydrolyzable ATP analogue, β,γ -methylene ATP. As found in the case of microcin B17,^[20] addition of the aforementioned non-hydrolyzable ATP analogue to reaction mixtures inhibited heterocycle formation (Figure S11). By contrast, the use of irreversibly acting protease inhibitors could not be anticipated, and this was discovered in a serendipitous manner. In studies of possible AEBSF inhibition of PatA protease, it was observed that the protease cleavage was not blocked, but strikingly heterocyclization was inhibited. This led to further studies demonstrating that AEBSF is a covalent inhibitor of heterocyclization, which are described below.

An inhibition curve using 1-10 mM AEBSF and monitoring TruE2 by SDS-PAGE showed that inhibition was complete at 10 mM for both PatD and TruD (Figure S12). Inhibition was monitored with an appropriate set of controls, ensuring specific inhibition and not an effect of solvent or conditions. Since in principle AEBSF could act competitively or allosterically instead of irreversibly, reactions were also incubated with PMSF. At a concentration of 1 mM, the reaction was strongly inhibited by PMSF. These data implicate direct nucleophilic displacement of fluoride in AEBSF and PMSF as the mechanism of inhibition. Moreover, a kinetic analysis was performed in which increasing concentrations of AEBSF were applied to enzyme reactions with TruD, measuring ADP production from ATP (Figure S12). The kinetic profile of reactions containing AEBSF strongly suggest that AEBSF does indeed act as an irreversible inhibitor; that is, ATP hydrolysis is not initially prevented by AEBSF, but as time goes on, virtually all ATP hydrolysis is halted.

Role of ATP in Heterocyclization

Upon beginning our study of this reaction, there were two plausible, extant hypotheses regarding the role of ATP in heterocyclization of ribosomal peptides. First is the idea that ATP could be a necessary cofactor in activating the electrophilic carbonyl carbon for attack by sulfur or oxygen (Scheme 1B).^[11] An alternative hypothesis holds that ATP is used by the heterocyclase enzyme in the manner of a molecular motor or G-protein.^[13] One of the main lines of evidence in favor of the molecular machine hypothesis was that ATP was used in super-stoichiometric amounts by microcin B17 synthetase; ATP was not used unless all enzyme components and the substrate were present, but upon incubation with substrate “excess” ATP was consumed. In addition, ATP hydrolysis could be uncoupled from heterocyclization when a large excess of substrate was employed.^[20] We therefore performed several experiments to illuminate the role of ATP in the PatD/TruD mechanisms, using TruD and TruE2.

Virtually no ATP was consumed in the absence of TruE2 substrate. In addition, a negligible ATP background hydrolysis was detected without enzyme. Upon addition of TruE2, a rapid burst in ATP hydrolysis could be observed. In experiments to determine the “minimum” amount of ATP leading to complete TruE2 modification, it was found that addition of 40 μ M ATP led to complete modification of 9 μ M TruE2 (or 18 μ M of heterocycle formed). Although this ATP concentration is well below the K_m , this experiment indicated fewer than 2.5 ATP hydrolysis events are required per heterocycle formed (Figure S9).

Although no ATP is consumed when the TruE2 substrate is absent, we wondered whether ATP would be used when only fully-modified TruE2 substrate is present. In a 6h time-course experiment with 140 nM TruD and 7.4 μ M TruE2, heterocycle formation is essentially complete after 1 h; during this initial hour the rate of ADP production is 0.24 μ M/min. During the following five hours, the rate of ADP production is significantly reduced, and remains linear with a rate of 0.06 μ M/min (Figures 3 and S10). When the background rate of hydrolysis due to enzyme only, substrate only and buffer only is accounted for, the 'excess' rate of hydrolysis after heterocycle formation is complete remains 0.05 μ M/min, a value roughly 20% of the initial fast rate of ATP hydrolysis.

These results strongly support the idea that TruD uses ATP in the manner of a molecular machine, since TruD continues to hydrolyze ATP even after all of the substrate for chemical reaction has been consumed. These results are superficially different than those obtained for the microcin B17 synthetase, in which 5-fold more ATP was used than heterocycles-formed (whereas TruD uses closer to a stoichiometric amount of ATP per heterocycle-formed). However, the microcin enzymes and substrates are quite different, and stoichiometry experiments in that system involved a precursor peptide that was processed to a tandem heterocyclic system (McbA 1-46 where GSC becomes G-Oxazole-Thiazole). The stoichiometry of ATP-used to heterocycles formed was not calculated explicitly for a single-turnover substrate in that system (McbA 1-47 where GSC was mutated to GGC, which is processed to GG-Thiazole). However, other experiments showed that the rate of heterocyclization was essentially unchanged compared to wild-type McbA 1-46, but it induced a 7-fold lower level of ATP consumption.^[20] Consequently, our results, which show nearly stoichiometric ATP consumption are actually quite consistent with previously reported results, given that there might be something distinctly different about the synthesis of a bisheterocyclic system as found in wild-type McbA 1-46.

We considered several further mechanistic hypotheses consistent with the hydrolysis of ATP in a molecular machine: (1) ATP binding and hydrolysis allows dissociation of a tightly-bound enzyme-substrate complex formed after heterocyclization; (2) ATP binding places the enzyme in an active conformation, in which heterocyclization can proceed readily; (3) ATP is used to autophosphorylate the enzyme, thus placing the enzyme into an active conformation. To test hypothesis (1) we conducted an extensive incubation of TruD with the single-cassette precursor TruE4 in the presence and absence of ATP. We used a very large amount of enzyme, such that a single-turnover would be apparent in the resulting mass spectrum. No turnover whatsoever was observable in the absence of ATP, even after extended incubation, while the reaction proceeded readily with ATP (Figure S11), a finding that does not support hypothesis (1). To test hypothesis (2) the non-hydrolyzable ATP analogue, β,γ -methylene ATP was added in increasing amounts to the TruD-TruE2 modification reaction; in addition, we attempted to observe TruE4 modification when the non-hydrolyzable analogue was substituted for ATP in the reaction. We found that the non-hydrolyzable ATP analogue strongly inhibited the TruE2 modification reaction, and that no modification to TruE4 could be observed when non-hydrolyzable ATP was employed (Figure S11), arguing for rejection of hypothesis (2). Finally, hypothesis (3) was tested by incubating enzyme, both alone and with substrate with radiolabeled 32 P ATP. Reactions were then analyzed by SDS-PAGE and autoradiography. However, we were unable to trap any autophosphorylated enzyme intermediate under any condition attempted. Consequently, it seems that the molecular machine might function via a more complicated mechanism than any proposed above.

Sequence Analysis and Comparison

PatD and TruD are didomain proteins that are nearly identical in their N-terminal domains (in the constructs used here, they are 100% identical through residue 323). These N-terminal

domains have elsewhere been proposed to directly catalyze heterocyclization.^[13] In addition, these enzymes share low similarity with related heterocyclization enzymes from other families, such as goadsporin, streptolysin, thiostrepton and relatives, and microcins. The percent identity to microcins is low enough that the sequences are essentially unalignable, but they can be transitively related as the streptolysin enzymes are related to both microcin and cyanobactin enzymes.^[5] Although there are highly conserved residues within the cyanobactin genes that could be involved in catalysis, there are no universally conserved sequence features that are shared in common amongst the different heterocyclizing proteins.

By contrast to the N-terminal domains, the C-terminal domains of PatD and TruD are only 77% identical to each other (Figure 2). These domains share homology with “YcaO” domains, which have no known function. In the context of heterocyclization, indirect experimental evidence indicates that they are likely involved largely in substrate binding and ATP hydrolysis, and not directly in catalysis of heterocycle formation.^[13, 20]

Interestingly, the N- and C-terminal domains of PatD and TruD share short conserved sequence features with MccB, an enzyme involved in adenylation of a peptide intermediate in microcin C7 biosynthesis (though it must be emphasized that microcin C7 does *not* contain any heterocycles and the biosynthetic purpose of adenylation is distinct from any reaction described in this work). A recent crystal structure of MccB shows that these residues line a substrate recognition pocket that strongly associates with the C-terminal residues of microcin C7.^[25] These microcin C7 residues are very similar to residues found in the C-terminal regions of PatE / TruE. Within MccB, the region of homology ends just before the ATP binding site and the reaction center. This alignment suggests that the homologous residues in PatD / TruD may bind to PatE / TruE in a similar way, but it does not inform on the role of ATP in heterocyclization. Speculatively, PatE / TruE may be held in place in the region adjacent to ATP binding, and hydrolysis of ATP leads to substrate release. ATP binding and hydrolysis could be either covalent (for example, to the C-terminus of the substrate), or non-covalent.

The conserved PatE / TruE leader sequence contains a short region that is predicted to be helical by multiple methods.^[26, 27] This region is of about the same length as an experimentally determined helical region (in trifluoroethanol) of the microcin B17 precursor peptide. This microcin helix was shown to be critical in interaction with the peptide binding protein,^[28] and it may serve the same function here. Indeed, without the leader sequence PatD and TruD cannot synthesize heterocycles, as shown in this study. The same precursor peptide region has been studied through extensive mutagenesis in the streptolysin S group.^[24] It is clear from these studies and others in the ribosomal peptide field that the leader sequence is critical to Cys/Ser/Thr modification in the lantibiotic- and microcin-like peptides.^[29]

Mechanistic Hypothesis

Our results allow us to rule out a number of alternative hypotheses regarding the mechanism of ribosomal peptide heterocyclases. The apparent requirement for the oxidase in heterocyclization of microcins and streptolysin led to the proposal that oxidation could occur prior to formation of Ser- or Cys-derived heterocycles (Scheme 1A), meaning that the relevant nucleophiles in the heterocyclization reaction would be enehydroxyl and enethiol.^[30] We have shown here that the flavin-containing oxidase is not required for heterocyclization, thus demonstrating that this mechanism cannot be correct. Kinetic analysis strongly supports the hypothesis that ATP drives a molecular machine that promotes heterocyclization, as proposed for microcin B17 synthesis (Scheme 1D); by

contrast, a mechanism wherein ATP is used to activate the adjacent carbonyl for attack by sulfur or oxygen (Scheme 1B) is rendered less plausible by the results presented here.

Recently, it was shown that intein chemistry can yield heterocyclic thiazolines as side products (Scheme 1C).^[31] Thus, it seems possible that heterocyclization chemistry could be quite similar to that required for intein splicing, as well as the chemical strategy employed by many autoproteolyzing enzymes.^[32] On this account, the peptide merely needs to be held in the right conformation, with appropriate nearby acids and bases to accelerate nucleophilic attack on the carbonyl adjacent to Cys or Ser / Thr.

Given intein chemistry (which proceeds without added energy, i.e. ATP), it seems mysterious why ATP would be required by heterocyclase enzymes. One possible explanation for this discrepancy is found in the fact that inteins are, in effect, single-turnover enzymes. In the case of multiple turnover heterocyclases, the enzyme-substrate complex may be so tightly bound that ATP is needed for product release, or the enzyme-substrate complex may be unable to form without the use of ATP. In addition, this system is strongly reminiscent of AAA-proteases, which require ATP for the detection of misfolded proteins and then hydrolyze these aberrant proteins via separate protease domains.^[33] It is noteworthy that there is sequence homology, both in substrate and enzyme, between cyanobactin biosynthetic proteins and microcin C7 proteins, which do not catalyze heterocyclization but instead adenylation. This similarity suggests a possible evolutionary relationship connecting the microcin-group biosynthetic pathways.

Conclusion

We have characterized two heterocyclization enzymes from cyanobactin pathways, PatD and TruD. In both cases, single enzymes were sufficient to recapitulate heterocyclization activity, making this the first single-protein reconstitution of heterocyclization activity and the first thiazoline / oxazoline synthetases to be characterized in ribosomal systems. Previous work in the microcin B17 and streptolysin S systems had shown that three protein domains are required for the biosynthesis of ribosomal peptide-derived heterocycles: an ATPase, a Zn-binding domain, and a flavin-containing oxidase.^[5, 13] By contrast, in the PatD / TruD group, oxidation is not a required component of catalysis, and the heterocyclization activity exists within a single polypeptide. These differences substantially simplified our approaches to gain mechanistic insights into the function and mechanism of heterocyclases in nature, allowing us to provide confirmatory evidence that ATP is used to drive a molecular heterocyclization machine. These results also define the order of steps in cyanobactin biosynthetic pathways, which lead to diverse and ubiquitous cyanobacterial natural products.

Experimental Section

General methods

Isopropyl β -D-1-thiogalactopyranoside (IPTG), dithiothreitol (DTT), leupeptin, pepstatin, 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), and phenylmethanesulfonyl fluoride (PMSF) were purchased from ISC Bioexpress. Metal-free nitric acid (Optima) was purchased from Fisher Scientific. Ultra-pure MgCl_2 , β , γ -adenosine triphosphate (β , γ -ATP) were purchased from Sigma-Aldrich. γ - ^{32}P -labeled ATP was purchased from Perkin-Elmer. Ni-NTA resin was purchased from Qiagen. ZipTip C18 pipette tips were purchased from Millipore. Krypton fluorescent protein stain was purchased from Pierce. All expression vectors were purchased from Novagen. *Escherichia coli* strain DH5 α was used for all cloning steps, while *E. coli* strain BL21(DE3)Star was used for all protein expressions.

Gene cloning and expression

Genes were obtained from ascidian symbiont metagenomes and cloned as previously described.^[16, 17] Enzymes were expressed as previously described, except that for kinetic analysis additional enzyme purification was performed. To wit, nickel-purified TruD was loaded onto a HiPrep 16/10 Q FF column (GE Healthcare), and run on an AKTA purifier FPLC system. The column was washed with 0.5 column volumes of buffer A, a linear gradient from 100% buffer A-0% buffer B to 20% buffer A-80% buffer B over 30 column volumes was run, followed by a 5 column volume wash at 0% buffer A-100% buffer B. Buffer A consisted of NaCl (0.1 M) buffered to pH 8.0 with HEPES (25 mM), while buffer B consisted of NaCl (1 M) buffered to pH 8.0 with HEPES (25 mM). TruE2 used for kinetic analysis was purified as previously described,^[17] except that the protein was stocked in a solution containing NaCl (350 mM), DTT (10 mM), sucrose (1% w/v) buffered to pH 8.0 with HEPES (20 mM).

Enzyme reactions

All enzyme reactions used in this study were performed at least in triplicate in independent runs. Reaction mixtures were incubated at 34°C in an MJ research minicycler for varying amounts of time. Enzyme, precursor peptide, and ATP concentrations varied, as described below. The following additives were present in standard reactions but were varied in early optimization experiments: tris pH 8.0 (40 mM), DTT (8 mM), and MgCl₂ (4 mM). Enzyme reactions generally contained the optimized additive mixture and PatD or TruD (0.6 μM), TruE2 (8 μM), or TruE4 (12 μM), and ATP (0.8 mM). Reactions were run using varying times, from 15 min to 27 h. To confirm modification, 10 μL of the reaction mixture was removed and analyzed by SDS-PAGE. At minimum, at least 3 separate experiments were performed for each enzyme-substrate concentration.

Sulfonyl fluoride inhibition

AEBSF and PMSF were used with standard concentrations of reagents, including TruE2 and TruD or PatD. Concentrations of sulfonyl fluorides were tested at 1 mM and 10 mM and compared to controls containing equivalent amounts of vehicle only (methanol for PMSF, water for AEBSF). PMSF was tested only against TruD. After 1.5 h reactions, the mixtures were analyzed by SDS-PAGE.

ATP conversion

To ascertain whether PatD and TruD hydrolyze ATP to ADP, or to AMP, PatD or TruD (0.6 μM) were incubated with TruE2 (2 μM), and ATP (100 μM) for 30 min. Reactions were then quenched by adding 10 μL of a saturated solution of urea and analyzed by HPLC as described below.

ATP stoichiometry

All reactions were performed in triplicate, using standard reaction conditions with TruE2 and PatD or TruD. No enzyme and no substrate controls were performed. Further controls contained 10 mM AEBSF and were performed with enzyme and substrate or without substrate. In one round, reactions were run with 800 μM ATP for 0.5 h, then quenched with 8 M urea.

Non-hydrolyzable ATP inhibition

Reactions were performed according to standard conditions except that two sets of reactions one containing ATP (800 μM) and the other ATP (200 μM) were used. β,γ-methylene-ATP was added to each set (800 μM or 200 μM ATP) at the following concentrations: 0, 100,

200, 400, 800, and 1600 μM . Reactions were allowed to proceed for 0.5h, and then analyzed by SDS-PAGE.

^{32}P -ATP labeling experiments

γ - ^{32}P -ATP (1 μCi) was doped into reactions containing cold ATP (40 μM). Standard reaction conditions were used with TruD and TruE2. No substrate, no enzyme, and AEBSF-inhibited controls were performed alongside these reactions. Reactions were analyzed both by adsorption to Nytran paper followed by scintillation counting, as well as by SDS-PAGE followed by autoradiography.

SDS-PAGE assays

18% acrylamide gels were used for all assays. Prior to electrophoresis, samples were brought up in 1X SDS sample buffer diluted from 6X SDS sample buffer: tris pH 6.8 (7mL, 0.5 M) glycerol (3 mL), SDS (1 g), DTT (0.93 g), bromophenol blue (1.2 mg), H_2O (up to 10 mL), and then boiled for 3 min.^[34] After electrophoresis, gels were placed in boiled fixing solution consisting of H_2O (53% v/v), ethanol (40% v/v), acetic acid (7% v/v), incubated 10-20 min with gentle rocking, and then placed in boiled stain solution consisting of: Coomassie R250 (0.02% w/v) in H_2O (85% v/v), acetic acid (10% v/v), ethanol (5% v/v). Gels were then destained in a solution consisting of H_2O (85% v/v), acetic acid (10% v/v), ethanol (5% v/v) for several h, and then photographed. For TruE2 time course experiments, gels were stained using Krypton fluorescent protein stain (Pierce) according to the manufacturer's instructions, and imaged using a Typhoon fluorescence reader (GE).

HPLC analysis

The HPLC method for all ATP-usage and stoichiometry experiments employed a Vydac 302IC4.6 ion exchange column, and a Hitachi LaChrom Elite HPLC system. A linear gradient proceeding from 100% buffer A, which consisted of formic acid (45 mM), adjusted to pH 4.5 using NaOH, to 100% buffer B, which consisted of NaH_2PO_4 (0.5 M), adjusted to pH 2.5 using formic acid) over 12 min was used to effect separation of ATP, ADP, and AMP. The elution profiles of the experimental runs were compared to those of authentic ATP, ADP, and AMP standards. Peaks were quantified by comparison with a calibration curve constructed by injecting known quantities of AMP.

Metal requirements

To assess whether or not the enzymes bound Zn as predicted, PatD and TruD were dialyzed over two days with stirring at 4°C against a solution (2 L) containing glycerol (5% v/v), NaCl (500 mM), Sepharose chelating resin (10 g) buffered to pH 7.8 using HEPES (25 mM). The purified, dialyzed enzyme samples were then digested using metal-free nitric acid, heated to 95°C, and read on a Perkin-Elmer Optima 3100 XL ICP-OES instrument.

The potential requirement for other metals (aside from Mg and Zn) was tested by using ultrapure MgCl_2 (Aldrich 255777) with TruD (1.6 μM), TruE2 (2 μM), using Chelex-treated ATP and DTT in standard concentrations without tris buffer. ATP and DTT were passed through Chelex 100 (Biorad) resin (700 μL) prior to addition to reaction. An additional set of controls were performed without using Chelex treatment.

Enzyme kinetics

Reactions were performed in the same manner as the standard conditions described above, except that enzyme, substrate, ATP, and AEBSF concentrations were varied. In the experiments where [ADP] was measured at 0, 20, and 40 minutes, experiments were performed with variable TruD, TruE2, ATP, and AEBSF. Experiments varying TruD were

performed as follows: TruD was added at variable concentrations (140, 104, and 69 nM) while holding constant TruE2 (7.4 μ M), ATP (800 μ M), and AEBSF (0 μ M). Experiments varying TruE2 were performed as follows: TruE2 was added at variable concentrations (7.4, 5, and 2.5 μ M) while holding constant TruD (104 nM), ATP (800 μ M), and AEBSF (0 μ M). Experiments varying ATP were performed as follows: ATP was added at variable concentrations (800, 400, and 200 μ M) while holding constant TruD (104 nM), TruE2 (7.4 μ M), and AEBSF (0 μ M). Experiments varying AEBSF were performed as follows: AEBSF added at variable concentrations (10, 5, and 1 mM) while holding constant TruD (104 nM), TruE2 (7.4 μ M), and ATP (800 μ M). Controls were run that lacked TruD while holding constant TruE2 (7.4 μ M), ATP (800 μ M), and AEBSF (0 μ M). Additionally, controls were run that lacked TruE2 while holding constant TruD (104 nM), ATP (800 μ M), and AEBSF (0 μ M). In experiments measuring [ADP] at 0, 60, 120, 240, and 360 minutes, the reactions contained TruD (140 nM), TruE2 (7.4 μ M), ATP (800 μ M). Controls were run that lacked either enzyme, substrate, or both. All reactions were performed in triplicate. After removal of aliquots at each time point, the reactions were quenched by addition of an equal volume of 8M urea, and then frozen at -80°C until analysis. Nucleotide content was assessed by HPLC as described above. Additionally, selected reactions were analyzed by SDS-PAGE to ensure that the heterocyclization reaction was proceeding at the expected pace. Further, to analyze the stoichiometry of ATP-hydrolyzed to heterocycles-formed, samples of the 0, 60, 120, 240, and 360 minute time points described above were analyzed by SDS-PAGE, stained with Krypton fluorescent protein stain according to the manufacturer's instructions, and then imaged on a Typhoon fluorescence reader. The resulting images were analyzed for band densitometry using the program ImageJ.

Protein Quantitation

TruE2 used for kinetic analysis was quantitated via amino acid analysis. The protein was dialyzed extensively against a solution containing NaCl (0.35 M), DTT (10mM), sucrose (1% w/v), and then subjected to amino acid analysis. The concentration was calculated based on the nanomoles recovered of Gly, Ala, Leu, Tyr, Phe, Lys, His, and Arg, which were averaged together between two separate runs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by NIH GM071425 and by a Willard Eccles Fellowship and an ACS Medicinal Chemistry Fellowship funded by Sanofi-Aventis to J.A.M. We thank Pam Smith and Dennis Winge for ICP-OES assistance and Darrell Davis for help conducting the ATP-usage experiments. We also thank Archana Yerra for technical assistance, Adele Flail for graphical assistance, and Janet Lindsley for helpful discussions. Lastly we thank an anonymous reviewer for suggesting several additional experiments in a previous revision of this paper.

References

1. Schmidt EW. The UCSD Guardian. 1991; 74
2. Lee FY, Borzilleri R, Fairchild CR, Kim SH, Long BH, Reventos-Suarez C, Vite GD, Rose WC, Kramer RA, Kelly WL, Hillson NJ, Walsh CT. Clin Cancer Res. 2001; 7:1429. [PubMed: 11350914]
3. Roy RS, Gehring AM, Milne JC, Belshaw PJ, Walsh CT. Nat Prod Rep. 1999; 16:249. [PubMed: 10331285]
4. Anderson B, Hodgkin D, Viswamitra MA. Nature. 1970; 225:233. [PubMed: 5409975]
5. Lee SW, Mitchell DA, Markley AL, Hensler ME, Gonzalez D, Wohlrab A, Dorrestein PC, Nizet V, Dixon JE. Proc Natl Acad Sci U S A. 2008; 105:5879. [PubMed: 18375757]

6. Yorgey P, Lee J, Kordel J, Vivas E, Warner P, Jebaratnam D, Kolter R. *Proc Natl Acad Sci U S A*. 1994; 91:4519. [PubMed: 8183941]
7. Li C, Kelly WL. *Nat Prod Rep*. 2010; 27:153. [PubMed: 20111801]
8. Donia MS, Schmidt EW. *Comprehensive Natural Products Chemistry II*. 2008 submitted.
9. Donia MS, Ravel J, Schmidt EW. *Nat Chem Biol*. 2008; 4:341. [PubMed: 18425112]
10. Kelly WL, Hillson NJ, Walsh CT. *Biochemistry*. 2005; 44:13385. [PubMed: 16201763]
11. Li YM, Milne JC, Madison LL, Kolter R, Walsh CT. *Science*. 1996; 274:1188. [PubMed: 8895467]
12. McIntosh JA, Donia MS, Schmidt EW. *Nat Prod Rep*. 2009; 26:537. [PubMed: 19642421]
13. Milne JC, Roy RS, Eliot AC, Kelleher NL, Wokhlu A, Nickels B, Walsh CT. *Biochemistry*. 1999; 38:4768. [PubMed: 10200165]
14. Schneider TL, Shen B, Walsh CT. *Biochemistry*. 2003; 42:9722. [PubMed: 12911314]
15. Lee J, McIntosh JA, Hathaway BJ, Schmidt EW. *J Am Chem Soc*. 2009; 131:2122. [PubMed: 19166292]
16. Donia MS, Hathaway BJ, Sudek S, Haygood MG, Rosovitz MJ, Ravel J, Schmidt EW. *Nat Chem Biol*. 2006; 2:729. [PubMed: 17086177]
17. McIntosh JA, Donia MS, Schmidt EW. *J Am Chem Soc*. Article ASAP.
18. McIntosh JA, Donia MS, Schmidt EW. *J Am Chem Soc*. 2010; 132:4089. [PubMed: 20210311]
19. Zamble DB, McClure CP, Penner-Hahn JE, Walsh CT. *Biochemistry*. 2000; 39:16190. [PubMed: 11123948]
20. Milne JC, Eliot AC, Kelleher NL, Walsh CT. *Biochemistry*. 1998; 37:13250. [PubMed: 9748332]
21. Madison LL, Vivas EI, Li YM, Walsh CT, Kolter R. *Mol Microbiol*. 1997; 23:161. [PubMed: 9004229]
22. Sinha Roy R, Belshaw PJ, Walsh CT. *Biochemistry*. 1998; 37:4125. [PubMed: 9521734]
23. Allali N, Afif H, Couturier M, Van Melder L. *J Bacteriol*. 2002; 184:3224. [PubMed: 12029038]
24. Mitchell DA, Lee SW, Pence MA, Markley AL, Limm JD, Nizet V, Dixon JE. *J Biol Chem*. 2009; 284:13004. [PubMed: 19286651]
25. Regni CA, Roush RF, Miller DJ, Nourse A, Walsh CT, Schulman BA. *Embo J*. 2009; 28:1953. [PubMed: 19494832]
26. Cole C, Barber JD, Barton GJ. *Nucleic Acids Res*. 2008; 36:W197. [PubMed: 18463136]
27. Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, J Dt. *Nucleic Acids Res*. 2005; 33:W36. [PubMed: 15980489]
28. Roy RS, Kim S, Baleja JD, Walsh CT. *Chem Biol*. 1998; 5:217. [PubMed: 9545435]
29. Oman TJ, van der Donk WA. *Nat Chem Biol*. 2010; 6:9. [PubMed: 20016494]
30. Kupke T, Gotz F. *J Biol Chem*. 1997; 272:4759. [PubMed: 9030529]
31. Ludwig C, Schwarzer D, Mootz HD. *J Biol Chem*. 2008; 283:25264. [PubMed: 18625708]
32. Walsh, CT. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*. 1. Roberts and Company; Greenwood Village, Colorado: 2006.
33. Striebel F, Kress W, Weber-Ban E. *Curr Opin Struct Biol*. 2009; 19:209. [PubMed: 19362814]
34. Gallagher S. *Curr Protoc Mol Biol*. 1998; 6:11.

Figure 1. Heterocyclic natural products. Shown are patellamide C (1), epothilone B (2), thiostrepton (3), and microcin B17 (4).

Figure 2.

A) Shown are alignments between PatD and TruD. Darker regions indicate regions of higher identity. B) Sequence of TruE2 precursor peptide is shown, with naturally heterocyclized residues highlighted in red. C) A zoomed-in view of the C-terminal cassette in TruE2. In vitro, PatD modifies one Thr and one Cys in this cassette, while TruD modifies one Cys both in vitro and in vivo. In nature, in combination with other biosynthetic enzymes the TruD product shown is converted to the prenylated, heterocyclic natural product patellin 6.

Figure 3. Stoichiometry of heterocycle formation. A) rates of ADP formation and thiazoline synthesis are overlaid; corrected slope denotes the rate of ATP hydrolysis when corrected for the background hydrolysis. B) %-completion of the heterocyclization reaction as determined by SDS-PAGE gel densitometry. C) SDS-PAGE gel used to derive [thiazoline] and %-completion.

Scheme 1.

Shown above are mechanistic possibilities for heterocyclization. A) oxidation preceding heterocycle formation B) activation of the adjacent carbonyl oxygen, perhaps with phosphate from ATP C) intein mechanism for thiazoline formation as a side-product D) molecular machine mechanism for heterocyclase enzymes.