Enzyme-Primed Native Chemical Ligation Produces Autoinducing Cyclopeptides in Clostridia

Evelyn M. Molloy, Maria Dell, Veit G. Hänsch, Kyle L. Dunbar, Romy Feldmann, Ansgar Oberheide, Lydia Seyfarth, Jana Kumpfmüller, Therese Horch, Hans-Dieter Arndt and Christian Hertweck*

Abstract: Clostridia coordinate many important processes such as toxin production, infection, and survival by density-dependent communication (quorum sensing) using autoinducing peptides (AIPs). Although clostridial AIPs have been proposed to be (thio)lactone-containing peptides, their true structures remain elusive. Here, we report the genome-guided discovery of an AIP that controls endospore formation in Ruminoclostridium cellulolyticum. Through a combination of chemical synthesis and chemical complementation assays with a mutant strain, we reveal that the genuine chemical mediator is a homodetic cyclopeptide (cAIP). Kinetic analyses indicate that the mature cAIP is produced via a cryptic thiolactone intermediate that undergoes a rapid S=N acyl shift, in a manner similar to intramolecular native chemical ligation (NCL). Finally, by implementing a chemical probe in a targeted screen, we show that this novel enzyme-primed, intramolecular NCL is a widespread feature of clostridial AIP biosynthesis.

Introduction

The Clostridia are a heterogeneous class of anaerobic bacteria that play major roles in human and animal health, industrial biotechnology, and the environment.\cite{1} There is growing evidence that Clostridia coordinate a broad range of crucial group behaviors that affect their physiology and pathogenicity by sensing the accumulation of secreted peptidic signals that regulate gene expression.\cite{2} Since the concentration of these chemical mediators correlates with cell density (quorum sensing, QS), gene expression can be synchronized with specific environmental conditions and growth phases.\cite{3} An impressive number of genetic studies have revealed that the accessory gene regulator (Agr) system is central to QS in diverse Clostridia\cite{4–m} For example, Agr-based QS is linked to the virulence potential of the infamous human pathogens Clostridium difficile\cite{2a–l} and Clostridium perfringens\cite{2n} and regulates toxin production in C. difficile\cite{2a–b} C. perfringens\cite{2a–i} and Clostridium botulinum.\cite{2l} Furthermore, Agr-mediated QS is implicated in the regulation of clostridial sporulation\cite{2a–i} as well as granulose production in the industrially important solvent producer Clostridium acetobutylicum.\cite{2a}

The QS signals derived from the Agr system are called autoinducing peptides (AIPs) and belong to the ribosomally synthesized and post-translationally modified peptide (RiPP) class of natural products.\cite{4} The mechanistic details of AIP biosynthesis and signaling were deciphered in pathogenic Staphylococcus spp.,\cite{5} with the prototypical Agr system consisting of four components: a gene-encoded precursor peptide (AgrD), an integral membrane macrocyclase (AgrB), and a sensor kinase (AgrC) and response regulator (AgrA) pair that transduce the signal (Figure 1A).\cite{6} AgrD is comprised of three regions: an N-terminal leader peptide that aids in membrane localization,\cite{7} a core peptide that encodes the QS signal,\cite{8} and a C-terminal follower peptide\cite{9} that is recognized by AgrB. The biosynthesis of the mature QS signal from this unmodified precursor peptide occurs over two steps (Figure 1B). First, the AgrB transmembrane cysteine protease\cite{10–12} removes the follower peptide and catalyzes intramolecular thiolactone bond formation between a conserved cysteine in the core peptide and the liberated C-terminus.\cite{13} The leader peptide is then removed by an extracellular peptidase,\cite{14} leaving the thiolactone macrocycle with an N-terminal extension (NTE) of varying length (Figure 1A and Figure 1B).\cite{15–17} In rare cases, cyclization involves a serine residue, resulting in the formation of a lactone macrocycle.\cite{18–20} Akin to other QS systems, AIP production and sensing are mutually enhancing, leading to a positive-feedback autoinduction circuit.\cite{21}

By analogy to their staphylococcal counterparts, it has been assumed that clostridial AIPs generally possess thiolactone moieties.\cite{2a–h,k,l} Considering the prevalence\cite{2a–m} and essential roles of Agr signaling circuits in Clostridia,\cite{2a–m} this assumption should be substantiated. Until now, the chemical...
Genomics-guided discovery of an AIP from a cellulytic, antibiotic-producing anaerobe

\textit{Ruminiclostridium cellulolyticum} DSM 5812 has been extensively studied for its ability to degrade cellulose\cite{16} and to produce the DNA-gyrase inhibiting antibiotic clothioamide,\cite{17} but there has been no report on AIP-mediated QS in this genetically tractable strain.\cite{18} Indeed, AIP-dependent signaling pathways have not been studied in any \textit{Ruminiclostridium} species to date. We therefore chose \textit{R. cellulolyticum} as a starting point for the characterization of clostridial AIPs. To gain insight into the potential of \textit{R. cellulolyticum} to produce AIPs, we mined its genome for genes encoding members of the AgrB protein family (PF04647). We identified four homologs of the AgrB peptidase gene (Table S1). Next, we scanned the genome region surrounding each candidate agrB for genes tentatively encoding a AIP precursor peptide (AgrD) and detected a candidate agrD in all four loci (Table S1 and Figure S1). In theory, the four \textit{agrBD} gene pairs could enable \textit{R. cellulolyticum} to produce four different AIPs (\textit{Rc-AIP1-4}) (Figure S1). This finding is unusual since previous bioinformatic analyses detected just one or two AIP gene clusters in members of the Clostridia.\cite{2b,2l,15}

The core peptide sequences of each possible \textit{R. cellulolyticum} AIP were predicted by aligning \textit{Rc-AgrD1-4} to AgrD homologues for which the corresponding products have been structurally characterized (Figure S2). Based on previous synthetic studies involving complementation of mutant phenotypes in various Clostridia\cite{2b,19} and a preliminary MS-based detection of an AIP from \textit{Hungateiclostridium thermocellum},\cite{20} we predicted that the AIPs from \textit{R. cellulolyticum} would not have NTEs, i.e. they should present a free N-terminus at the ultimate cysteine residue.

In order to assess the production of AIPs by \textit{R. cellulolyticum}, the strain was cultivated, the culture was extracted with ethyl acetate, and the extract was analyzed by HPLC-HRMS. The metabolite profile was searched for the molecular ion masses calculated for each macrocyclic AIP (\textit{Rc-AIP1-4}) without an NTE. No signals were detected that could account for the \textit{Rc-AIP2-4} predictions, regardless of NTE length (0 to 4 residues). However, we detected a metabolite with an exact mass consistent with the prediction for mature \textit{Rc-AIP1} devoid of an NTE (calc. m/z 873.3389 [M+H]+; found m/z 873.3398 [M+H]+) (Figure 2A and Figure 2B). We also checked the metabolite profile for the exact masses of \textit{Rc-AIP1} bearing NTEs of up to four residues; however, no

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Figure1.png}
\caption{The prototypical Agr system of \textit{Staphylococcus aureus} A) Representation of an \textit{S. aureus} AIP biosynthetic gene cluster showing genes encoding the integral membrane macrolactase (AgrB), the precursor peptide (AgrD), the cognate two-component transduction system (AgrC-AgrA), as well as the distal gene encoding the membrane protease (SspB). Also shown is an example of a mature staphylococcal thiolactone-containing AIP. B) Overview of the biosynthetic pathway to staphylococcal thiolactone-containing AIPs. The AIP is derived from an internal fragment of the ribosomally synthesized precursor peptide AgrD. First, AgrB catalyzes C-terminal proteolysis of the follower peptide (FP) and subsequent macrocyclization, followed by secretion. Then, the N-terminal leader peptide (LP) is removed by SspB, leaving an N-terminal extension (NTE) of varying length. The mature AIP, which varies among strains, consists of a five-membered thiolactone ring with a 2–4 residue NTE. In response to binding of the AIP, the sensor kinase AgrC is autophosphorylated and the signal is transduced by the response regulator AgrA leading to global alterations in gene expression.}
\end{figure}

structures of clostridial AIPs have been inferred from indirect experimental evidence such as signal degradation,\cite{21} and by chemical complementation of mutant phenotypes with synthetic AIPs designed from bioinformatics-based structure predictions.\cite{2b,2j} Here, we report the first complete structural assignment of a clostridial AIP and provide compelling evidence that the previously suggested structures are transient precursors of the true signaling molecules. Our findings not only prompt a re-examination of some previously reported AIP structures, but also reveal a novel pathway to cyclic RiPPs by natural native chemical ligation (NCL).
corresponding signals were found (Figure S3). HRMS\textsuperscript{2}-based fragmentation of the ion corresponding to the \textit{Rc}-AIP1 candidate resulted in a pattern in which all observed fragment masses were in agreement with a macrocyclic peptide of the predicted amino acid sequence (CWFWSY) (Figure S4). Notably, \textit{Rc}-AIP1 is only the second example of an active clostridial AIP detected to date and, in both cases, the signaling molecules lack an NTE.\textsuperscript{[19]}

\textbf{Rc-AIPs is a homodetic cyclopeptide derived from a thiolactone-containing intermediate}

The possibility of spontaneous \(S\rightarrow N\) acyl migrations occurring in AIPs with thiolactones formed between the \(C\)- and \(N\)-termini was only recently posited.\textsuperscript{[20]} The identification of \textit{Rc}-AIP1, in which an NTE is absent, provided us with an opportunity to investigate this prospect. We attempted to isolate \textit{Rc}-AIP1 from \textit{R. cellulolyticum}, but this proved challenging due to low production titers. As an alternative approach, we synthesized a reference compound for comparison to \textit{Rc}-AIP1. Specifically, we employed Fmoc-based solid-
phase peptide synthesis (SPPS, for more experimental details see SI) to prepare the linear βBu-protected peptide, WFWSY. The N-terminal amino acid (Cys) was coupled to the linear protected peptide as a S-Trt- and N-Boc-protected version by using ethyl cyanohydroxyminoacetate (Oxyma), (1-cyano-2-ethoxy-2-oxoethylideneaminoxy)-dimethylamino-morpholino-carbenium hexafluorophosphate (COMU), and diisopropylpyrrolidylamine (DIPA) as coupling reagents. The Trt group was removed simultaneously as the Boc- and βBu-protected peptide was liberated from the resin by treatment with hexafluoroisopropanol (HFIP) and purified by preparative HPLC. The obtained thiol peptide acid 1a was cyclized with 1-hydroxybenzotriazole hydrate (HOBt) and N-(3-dimethylaminopropyl)-N'-ethycarbodiimide hydrochloride (EDCI) as coupling reagents, and DIPA and 4-dimethylamino pyridine (DMAP) as bases in DMF,[20] to yield the corresponding thiolactone 1b (Figure 2C and Figure S5). After removal of the βBu-based protecting groups with trifluoroacetic acid (TFA), the labile thiolactone 1, which could be handled under strictly acidic conditions only, was purified by preparative HPLC (3.4% overall yield starting from SPPS). The characteristic $^{13}$C-NMR signal at ≈ 200 ppm confirmed the successful formation of the thiolactone-containing peptide 1 (Figure 2D and Figure S6).

Immediately after deprotection, we subjected the synthetic thiolactone 1 to HPLC-HRMS analysis in order to compare its retention time, accurate mass, and fragmentation patterns with those of Rc-AIP1 from R. cellulolyticum culture extracts. We observed a divergence in the retention times of 1 (peak 1) and Rc-AIP1 (Figure 2B and Figure S7). However, we noted that a minor congener (peak 2) in the reference chromatogram had the same retention time as Rc-AIP1. We also found that both compounds show the same m/z values and very similar HRMS$^2$ fragmentation patterns, yet differ in their stabilities to the collision energies used (Figure S8). These observations led us to conclude that 1 is unstable and undergoes an S→N acyl shift to form the corresponding homodetic cyclopeptide 2 with a more stable amide bond (Figure 2C).

To scrutinize this conjecture, we used HPLC-HRMS to monitor the stability of thiolactone 1 in MeOH. We noted that 1 was indeed unstable, since the corresponding peak decreased upon formation of a new peak that came to predominate in the HPLC-HRMS profile (Figure S9). In order to obtain a synthetic reference, we incubated 1 in 50 mM 3-(N-morpholino)propanesulfonic acid (MOPS) at pH 7 until peak 2 became the major signal in the chromatogram. The corresponding compound 2 was purified by preparative HPLC and its lactam structure was confirmed by NMR analyses (Figure 2D and Figure S10). Notably, the retention time of 2 proved to be identical to that of the natural product (Rc-AIP1) (Figure 2B and Figure S7). Consequently, Rc-AIP1 must exclusively possess amide bonds and hence deviates from the hitherto assumed (thio)lactone-containing architecture of AIPs. In light of this difference, we designate Rc-AIP1 as a „cAIP“ to emphasize that it is a homodetic cyclopeptide.

To exclude the possibility that lactam 2 arose in the R. cellulolyticum metabolite profile as an artifact of the extraction procedure, we freshly prepared thiolactone 1 and, immediately after deprotection, determined its rearrangement kinetics under conditions mimicking those of the cultivation (at 37°C in 50 mM MOPS, pH 7.1). As the conversion of 1 to 2 results in the formation of a thiol, the transformation was monitored by using a spectrophotometer and a thiol detection assay based on Ellman’s reagent (Figure 2E).[21] We determined an average half-life of 5.4 ± 0.6 min for 1 (Figure 2E and Figure S11), indicating that the NTE-free thiolactone-containing AIP would not persist over the estimated generation time (6 h) of R. cellulolyticum determined under similar cultivation conditions.[22] We propose that by the time the threshold concentration needed to trigger a QS-dependent response would be reached, the spontaneously formed 2 would dominate over 1. We therefore propose that 2 is the native structure of Rc-AIP1.

**Rc-AIP1 controls endospore formation in Ruminiclostridium cellulolyticum**

Our elucidation of the structure of Rc-AIP1 revealed an unprecedented structural deviation for an AIP family member, but the question remained as to whether the homodetic cyclopeptide 2 acts as a competent signal. If so, it should be sensed by R. cellulolyticum and elicit a response. However, no physiological role has been reported for an R. cellulolyticum AIP. To this end, we created a null mutant by using CRISPR-nCas9 genome editing[20] to incorporate a nonsense mutation in Rc-agrD1 (Ccel_2126), the gene encoding Rc-AIP1 (Figure 3A and Figure S12). As expected, no signal for Rc-AIP1 could be detected in the metabolite profile of R. cellulolyticum ΔagrD1 (Figure 3B). This biosynthetic deficiency could be complemented by re-introducing an intact copy of Rc-agrD1 on a plasmid (Figure 3B).

Since Agr-based QS has been implicated in the initiation of sporulation in Clostridia,[21,22] we assessed the sporulation efficiency of R. cellulolyticum ΔagrD1 (Figure 3C). We found that the number of heat-resistant endospores formed by R. cellulolyticum ΔagrD1 was indeed reduced by greater than two orders of magnitude compared to the wild-type strain (Figure 3D). Exogenous supplementation of R. cellulolyticum ΔagrD1 with 2 could rescue the spore formation deficiency in a concentration-dependent manner (Figure 3D). Importantly, 2 restored wild-type levels of sporulation at 200 nM, a concentration considered physiologically relevant[23] and at least an order of magnitude lower than previously reported for clostridial Agr-dependent phenotypes.[24] Combined with the inherent instability of 1 over a biologically relevant timescale, these chemical complementation experiments corroborate that 2 is an authentic signaling molecule involved in the orchestration of endospore formation in R. cellulolyticum.
Figure 3. Control of endospore formation in Ruminiclostridium cellulolyticum by the cAIP Rc-AIP1 and evaluation of synthetic surrogates: A) Scheme depicting the incorporation of a nonsense mutation into Re-agrD1 (Ccel_2126) of R. cellulolyticum to generate the null mutant R. cellulolyticum ΔagrD1. The simultaneous introduction of a SpeI recognition site enabled the differentiation of successfully edited R. cellulolyticum from the wild type by restriction digest of colony PCR products. B) Extracted ion chromatogram (EIC) showing the presence or absence of Rc-AIP1 in the metabolite profiles of R. cellulolyticum wild type (WT), R. cellulolyticum ΔagrD1, R. cellulolyticum ΔagrD1 pCPthl-empty (vector control) and R. cellulolyticum ΔagrD1 pCPthl-agrD1 (complemented mutant). C) Overview of the endospore formation assay. SPPS, solid-phase peptide synthesis. D) Quantification of endospore formation (colony forming units per milliliter, CFU · mL⁻¹) by heat-treated cultures of R. cellulolyticum wild type, R. cellulolyticum ΔagrD1, and R. cellulolyticum ΔagrD1 supplemented with the listed concentrations of 2, 3, 4 and 5. The results depicted are derived from at least three independent experiments. The colony counts for R. cellulolyticum ΔagrD1 + DMSO (vehicle control) and R. cellulolyticum ΔagrD1 + 200 nM 3 were below the quantification limit (300 CFU · mL⁻¹) of the assay. For the other assay conditions, the displayed values are the mean colony counts (CFU · mL⁻¹) with error bars showing the standard deviation. All values were compared to that of the wild type (ns, not significant; *, p ≤ 0.05). DMSO, dimethyl sulfoxide. E) Synthetic route to compounds 3–5. The linear sequences for 4 and 5 were produced in the same manner as for 1 and 2 (see Figure 2C). The asterisk indicates the fBu-protection of a hydroxyl group.
Synthetic surrogates reveal the importance of a macrocycle to Rc-AIPs recognition

The rapid transformation of the thiolactone 1 into lactam 2 under cultivating conditions precludes a direct comparison of their potencies in the endospore formation assay. In order to gain preliminary insights into the structure–activity relationship of Rc-AIP1, we prepared three analogues: a linear derivative 3, an N-terminally acetylated derivative 4, and a derivative 5 containing a peptidomimetic thiazole moiety at the macrocyclization site (Figure 3E and Figure S5). N-acetylated analogue 4 features an increased ring size compared to the naturally occurring lactam 2, but preserves the thiolactone moiety since it cannot undergo an S→N acyl shift. Thiazole-containing analogue 5 has the same ring size as 2, but the risk of disulfide formation is eliminated. Moreover, it represents a stabilized analogue of the cAIP intermediate thiolactone 1 and is not prone to hydrolysis. Heterocyclic motifs are present in many natural products and their incorporation in cyclic peptides has led to molecules with improved properties.[24]

In brief, the linear peptide 3 was obtained by hydrolysis of the protected, SPPS-derived thiolactone 1b with aqueous NaOH, followed by TFA treatment to remove the protecting groups (0.8% overall yield starting from SPPS). We prepared the N-acetylated analogue 4 from the corresponding N-acetylated, fully-protected linear peptide 4a, which was assembled on solid support. Treatment of 4a with TFA and Et3SiH yielded the ω-mercaptoproxylic acid, which was transformed into the thiolactone by using benzotriazole-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and DIPEA.[25] Global cleavage of the protecting groups was performed under acidic conditions to give thiolactone 4 (8.4% overall yield starting from SPPS). The thiazole analogue 5 was synthesized from the fully protected linear peptide 5a, which was assembled by SPPS in analogy to 1a and 4a, but with (R)-2-azido-3-(tritylthio)propanoic acid replacing the protected cysteine.[25] Upon deprotection, the ω-mercaptoproxylic acid was cyclized to 5 by using PyBOP as a coupling reagent and DIPEA as a base. The heterocyclization was promoted by means of a Staudinger-aza-Wittig sequence[26] by using PPh3 in 2,6-lutidine. The resulting thiazoline was oxidized by using 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) and bromotrichloromethane. Subsequently the protecting groups were cleaved by TFA/Et3SiH/H2O (16% overall yield starting from SPPS).[26]

With these structural analogues at hand, we tested their potencies in the endospore formation bioassay with R. cel lulolyticum ΔagrD1. In stark contrast to the response to 2, supplementation with 200 nM of the linear peptide 3 was indistinguishable from the phenotype of the untreated mutant (Figure 3D). It is likely that the conformation needed for effective binding of Rc-AIP1 to its cognate receptor is favored by its cyclic nature compared to the conformationally more flexible linear peptide. Both macrocyclic analogues 4 and 5 were active in the endospore formation assay, albeit less so than 2. Specifically, both 4 and 5 were only capable of partially complementing the sporulation deficient phenotype of R. cel lulolyticum ΔagrD1 at 200 nM, a level at which 2 completely restored sporulation. As well as further supporting that a macrocyclic nature is important for recognition, the partial functionality of analogues 4 and 5 indicate that there is some variation tolerated in terms of linkage moiety and ring size of synthetic surrogates. However, the relative impact of these structural variations on biological function is difficult to distinguish. More elaborate structure-activity studies will be necessary to precisely evaluate the structural features that are crucial for Rc-AIP1 to adopt the conformation required for optimal interaction with its receptor.

Clostridial AIPs commonly undergo intramolecular S→N rearrangements

Prompted by the discovery that Rc-AIP1 is a potent QS signal, we wondered if the products of other clostridial Agr clusters might be homodetic cyclopeptides. We therefore mined the genomes of diverse Clostridia spanning four genera for genes encoding homologous cyclopeptides. We therefore identified the sequences of the candidate AIPs from the corresponding fragments (Figure S4 and Figure S14–18). Our detection of the AIP produced by C. acetobutylicum confirms the prediction that it is a 6-membered macrocyclic peptide without an NTE.[23]

Since the absence of an NTE is a prerequisite for the S→N acyl shift to occur, we presumed that the newly detected AIPs should be subject to NCL-like rearrangement to form cAIPs. However, such a connectivity difference would again be indistinguishable by HRMS[2]. Although it has been proposed that the AIPs of C. acetobutylicum[23] and C. perfringens[25] would not have NTEs based on chemical complementation experiments, the resulting possibility of rearrangement was not explored. Our detection of numerous NTE-free AIPs allowed us to investigate whether the S→N acyl shift observed for Rc-AIP1 is a more general phenomenon guided by the innate chemical reactivity of the α-amino thioester.

In order to circumvent the need for synthetic references to finalize the structures for all candidate AIPs, we developed a method that could distinguish between a thiolactone or lactam moiety (Figure S19). We employed N-benzylmaleimide to chemoselectively label the free thiols that would only be present in the lactam forms and exploited the fact that cAIPs would be resistant to basic hydrolysis (Figure 4A). The method was validated by way of synthetic standards (Figure S20–23) and by showing that Rc-AIP1 in a crude extract behaved as expected for a cAIP (Figure S24). When applied in a screen of the crude extracts containing candidate AIPs, a thiol adduct was detected in all (Figure 4B–E), indicating that the cysteine side chain of each AIP had indeed been liberated by an S→N acyl shift. In accordance with the results obtained from N-benzylmaleimide labeling, no signals corresponding to the linearized forms were detected upon NaOH
Figure 4. Identification of cAIPs in extracts of diverse Clostridia: A) General scheme depicting the reaction of a cAIP with \( \text{N-} \text{benzylmaleimide} \) resulting in the formation of a thiol adduct (red) and the absence of hydrolysis by \( \text{NaOH} \) (blue). B–E) After treatment with \( \text{N-} \text{benzylmaleimide} \) (+mal) and \( \text{NaOH} \), the AIPs detected in extracts of the named Clostridia behaved in a manner characteristic of homodetic peptides (cAIPs) (Figure 4A). Specifically, new species corresponding to the mal-adducts were detected (red), whereas no hydrolytic ring-openings were observed (blue). EICs are shown at either a 1.0 ppm window (panel B) or a 1.5 ppm window (panels C–E). †Further analysis by HRMS experiments (untreated: Figure S4 and Figure S14–18, + mal: Figure S25–28). *Unrelated peaks detected in the crude extract.
treatment of the extracts (Figure 4B–E). Hence, we conclude that, like Rc-AIP1, these NTE-free AIPs from Clostridia are homodetic cyclopeptides and should be classified as cAIPs.

**A biosynthetic route to cyclopeptides involving enzyme-primed native peptide ligation**

Two in vivo routes for the formation of the cAIPs may be conceived (Figure 5A). By analogy to the staphylococcal Agr system, AgrB should cleave the follower peptide of AgrD and catalyze the formation of a thiolactone-containing macrocycle, which is then exported. Concomitant proteolytic cleavage of the leader peptide would liberate a free N-terminus that can undergo a spontaneous $S\rightarrow N$ acyl shift (route (i)). Alternatively, the leader peptide may be removed before the core peptide is processed by AgrB. The resulting free amine would then liberate the AgrB-bound intermediate by directly forming the amide bond (route (ii)), mirroring other RiPP maturation routes (e.g. cyanobactins). Whereas there is no direct proof of a cryptic thiolactone intermediate, multiple lines of evidence support route (i) as the more probable pathway. Firstly, a cysteine residue at the cyclization position is indispensable for the formation of a cryptic thiolactone in route (i), whereas any amino acid would be suitable for route (ii). We and others noted that the cysteine residue is a conserved feature of clostridial AgrD peptides (Figure S13B).

Whereas there is no direct proof of a cryptic thiolactone intermediate, multiple lines of evidence support route (i) as the more probable pathway. Firstly, a cysteine residue at the cyclization position is indispensable for the formation of a cryptic thiolactone in route (i), whereas any amino acid would be suitable for route (ii). We and others noted that the cysteine residue is a conserved feature of clostridial AgrD peptides (Figure S13B).

Secondly, in characterized Agr systems the leader peptide is removed after the thiolactone-containing peptide is secreted from the cell. Route (ii) would be inconsistent with these data.

Taken together, these findings reveal a specialized biosynthetic pathway to cAIPs in Clostridia. By cleaving the follower peptide, AgrB generates a thioester-bound peptide and promotes an entropy-driven cyclization reaction to release the corresponding thiolactone. Liberation of the amine by cleavage of the NTE region sets the stage for a consecutive, maturing $S\rightarrow N$ acyl shift. This mechanism is strikingly similar to NCL strategies employed for the synthesis of (cyclo)peptides (Figure 5B). In particular, the AgrB-primed pathway mirrors intramolecular variants of NCL that permit the on-resin synthesis of cyclopeptides.

The reactivities and mechanisms are identical in both the synthetic and biosynthetic routes, yet the synthetic peptide is linked to a bead, whereas the cAIP precursor is tethered to an enzyme. Spontaneous $S\rightarrow N$ acyl migrations have been reported in protein splicing pathways as well as in a recently discovered compensatory pathway for glutathione biosynthesis in bacteria lacking the GshA glutathione ligase. By drawing on chemical logic, it had been suspected that the thiolactone formed from an unprotected N-terminal cysteine in a subset of AIPs could rearrange to yield an amide bond.

Notably, the propensity of thioesters to undergo exchange reactions and $S\rightarrow N$ acyl shifts was recently exploited in a discovery strategy for canonical thiolactone-possessing peptides.

**Figure 5.** Model for cAIP biosynthesis: A) Two possible routes for in vivo cAIP formation involving either (i) a native chemical ligation (NCL)-like $S\rightarrow N$ shift or (ii) the direct formation of the lactam. FP, follower peptide; LP, leader peptide. B) Comparison of of solid-phase peptide synthesis (SPPS)-based and enzyme-primed NCL strategies for cyclopeptide production. In chemical synthesis, direct attack of a free thiol leads to self-cleaving from the 3-amino-4-(methylamino)benzoic acid (MeDbz)-linked resin yielding the thiolactone. In cAIP biosynthesis, the same product is produced by the attack of a free thiol to the thioester. The intermediary thiolactones undergo intramolecular $S\rightarrow N$ acyl shifts to give the corresponding cyclopeptides.
AIPs produced in culture.[20, 32] Efforts to apply the NCL-based technique to identify the NTE-free AIP of the pathogen *Listeria monocytogenes* were unsuccessful.[20] On the basis of our findings, we propose, as did the authors, that the mature AIP presents as the lactam form, meaning it would be incompatible with chemoselective trapping by NCL. We speculate that the NTE-free AIP of the human-associated *Lactobacillus plantarum*,[34] an important anaerobe used for food fermentation, could also be a cAIP. The fate of the lactone in NTE-free AIPs cyclized through a serine residue, such as that of *H. thermocellum*, remains to be clarified.[19]

Finally, it should be highlighted that this previously overlooked natural NCL-like cyclopeptide formation adds to the known pathways for RiPP maturation. Previously, ATP-dependent and protease-dependent avenues have been reported for cyclopeptide biosynthesis in diverse RiPP families,[27, 35] with a maturation enzyme installing the lactam bond in each case. The mechanism for cAIP macrolactamization described herein markedly differs from these routes and represents a novelty in the RiPP biosynthetic toolbox.

**Conclusion**

Many genetic studies have indicated that Agr-derived peptides mediate important processes in Clostridia. However, the true nature of the implicated signaling molecules has remained uncertain. In this study, we provide the first rigorously determined structure of a clostridial AIP. On the basis of chemical synthesis and meticulous analysis, we ascertained that the signals detected in clostridial cultures do not correspond to thiolactones; they are homodetic intermediates. In the case of a model clostridial cAIP, *Rc*-AIP1 produced by *R. cellulosilyticum*, kinetic analyses exclude the possibility that the thiolactone precursor could play a significant biological role, which is also supported by chemical complementation experiments.

Until now, it was believed that cyclization through a thioester or ester was a conserved feature of AIPs.[4, 5] In expanding the chemical diversity of the AIP family to include cAIPs, this discovery highlights the importance of rigorous chemical analyses and synthesis, and serves as a cautionary tale regarding cursory bioinformatics-based assumptions about the structures of natural products. Moreover, the structures and pathways described herein are an unexpected divergence from the canonical AIP biosynthetic routes and prompt a revision of the clostridial AIPs and their formation. This novel route to cyclic RiPPs via cryptic thiolactone-

**Acknowledgements**

We thank A. Perner for MS measurements, H. Heinze for NMR measurements, and H. Büttner for providing culture extracts. Financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1127/2 ChemBioSys-239748522 (to H.D.A. and C.H.), Leibniz Award (to C.H.), and Cluster of Excellence „Balance of the Microverse“ (to R.F., H.D.A. and C.H.), the Alexander von Humboldt Foundation (postdoctoral fellowship to K.L.D.), and the Carl Zeiss Foundation (predoctoral fellowship to A.O.) is gratefully acknowledged. Open access funding enabled and organized by Projekt DEAL.

**Conflict of interest**

The authors declare no conflict of interest.

**Stichwörter**: Clostridia · cyclopeptide · native chemical ligation · quorum sensing · RiPP
