Natural Products

Nocathioamides, Uncovered by a Tunable Metabologenomic Approach, Define a Novel Class of Chimeric Lanthipeptides

Hamada Saad, Saefuddin Aziz, Matthias Gehringer, Markus Kramer, Jan Straetener,
*Anne Berscheid, Heike Brötz-Oesterhelt, and Harald Gross**

Abstract: The increasing number of available genomes, in combination with advanced genome mining techniques, unveiled a plethora of biosynthetic gene clusters (BGCs) coding for ribosomally synthesized and post-translationally modified peptides (RiPPs). The products of these BGCs often represent an enormous resource for new and bioactive compounds, but frequently, they cannot be readily isolated and remain cryptic. Here, we describe a tunable metabologenomic approach that recruits a synergism of bioinformatics in tandem with isotope- and NMR-guided platform to identify the product of an orphan RiPP gene cluster in the genomes of Nocardia terpenica IFM 0406 and 0706^T. The application of this tactic resulted in the discovery of nocathioamides family as a founder of a new class of chimeric lanthipeptides I.

Introduction

RiPPs outline a significantly expanding category of peptide-based natural products with multiple classes covering a vast chemical space. Such a broad scope of structural modifications translates into different biological activities, ranging from antibacterial, antifungal to antiviral potencies,[1] which can be therapeutically exploited. However, intriguing ecological and physiological roles are also reported for RiPPs.[2] Lanthipeptides, the most prevalent and exhaustively mined class of RiPPs, are structurally branded by their unifying installation of the thioether-bridged amino acids lanthionine (Lan) and methyllanthionine (MeLan) as one of

the frequent post-translational modifications (PTMs) of RiPPs.[3] The typical introduction of sulfur and oxygen heterocycles within the core peptide comprising thazol(in)e and (methyl)oxazol(in)e moieties has also been perceived in several RiPP variants, such as linear azoline scaffolds (LAPs), head-to-tail monocyclic skeletons (cyanobactins), and multi-cyclic frameworks with a nitrogen-containing heterocycle (pyritides).[4] Among the rare alterations of RiPPs, thioami-tides arise in which thioamide(s) can be exceptionally incorporated into the peptide backbone to leverage their structural diversification and bioactive potential.[5] With the explosion of affordable sequencing technologies in orchestration with improved bioinformatics algorithms for biosynthetic gene cluster (BGC) identification and defined motif annotation,[6] the quest has been paved towards the discovery of a hidden arsenal of novel RiPP-based molecules across various bacterial (meta)genomes. Nevertheless, connecting the RiPPs genetic locus of interest to the cognate metabolite-(s) is yet a challenging task[7] due to the production of complex metabolite mixtures, the silence of many BGCs under laboratory cultivation conditions or if the source organism is considered uncultivable. The latter two hurdles are nowadays overcome by the application of heterologous expression and synthetic biology, while analytical problems are foremost addressed by the integration of mass spectrom-etry (MS)-based metabolomics techniques.[8]

Bacteria from the genus *Nocardia* are foremost known in a medical context since they cause severe infections to

[*] M. Sc. H. Saad, M. Sc. S. Aziz, Prof. Dr. H. Gross
Department of Pharmaceutical Biology, Institute of Pharmaceutical Sciences, University of Tübingen
Auf der Morgenstelle 8, 72076 Tübingen (Germany)
E-mail: harald.gross@uni-tuebingen.de

M. Sc. H. Saad
Department of Phytochemistry and Plant Systematics, Division of Pharmaceutical Industries, National Research Centre
Dokki, Cairo (Egypt)

M. Sc. S. Aziz
Microbiology Department, Biology Faculty, Jenderal Soedirman University
Purwokerto (Indonesia)

Prof. Dr. M. Gehringer
Department of Pharmaceutical/Medicinal Chemistry Institute of Pharmaceutical Sciences, University of Tübingen
Auf der Morgenstelle 8, 72076 Tübingen (Germany)

Dr. M. Kramer
Institute of Organic Chemistry, University of Tübingen
Auf der Morgenstelle 18, 72076 Tübingen (Germany)

J. Straetener, Dr. A. Berscheid, Prof. Dr. H. Brötz-Oesterhelt
Department of Microbial Bioactive Compounds, Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen
Auf der Morgenstelle 28, 72076 Tübingen (Germany)

Prof. Dr. H. Brötz-Oesterhelt, Prof. Dr. H. Gross
Cluster of Excellence: EXC 2124: Controlling Microbes to Fight Infection, University of Tübingen, Tübingen (Germany)

humans and animals via nocardiosis. However, they have emerged over the last two decades also as talented producer strains of structurally diverse and highly bioactive natural products.[9] Beside nonribosomal peptides and polyketides, *Nocardia* spp. are also privileged producers of RiPPs, such as nocathiacins and nocardithiocin as thiopeptides[9] besides nocavionin[10] as lipolanthines. Within the course of our genome-driven investigations of *Nocardia* strains,[11] we noted that the highly similar genomes of *Nocardia terpenica* IFM 406 and 706[T] both contain, beside nocavionin,[10] a further RiPP-BGC which codes for a core ribosomal peptide co-localized with a unique combination of post-translational enzymes that have not been observed so far for lanthipeptides. Here, we report a genome-guided identification, isolation and characterization of three structurally novel RiPPs, designated as nocathioamides A–C, from these strains, which represent the first members of a new class of chimeric lanthipeptides.

Results and Discussion

Genome Mining of the Nocathioamide Biosynthetic Gene Cluster in Nocardia terpenica IFM 0406 and 0706

Mining the genomes of *N. terpenica* IFM 406 and 706, employing the bioinformatic tools antiSMASH 5.1[6a] and RODEO[6b, 12] readily revealed the presence of a putative lanthipeptide BGC (Table S1). The comparatively large *nta* gene cluster (17.1 kb) consists of 14 open reading frames (ORFs), annotated as *ntaA-ntaN* (Figure 1). NtaA-ntaJ are

transcribed in one direction, while *ntaK-ntaN* are translated in the opposite orientation. The gene *ntaA* encodes a 51-aa precursor peptide which showed no sequence similarity to any known RiPPs. The leader peptide possessed the conserved FDLN motif and towards its C-terminal end two proline residues (Figure 1), which typically occurs in prepeptides of class I lantibiotics and suggest, with P commonly located at position −2, two putative cleavage sites.[13] To complicate matters further, NtaA contains in addition one characteristic tandem alanine "AA" and two "PA" cleavage sites[14] which give rise to multiple possibilities for dividing the precursor peptide into an N-terminal leader peptide and a C-terminal core peptide. Application of the automated web-based tool RIPPMiner[15] corroborated these findings and predicted the formation of a 12 or 14mer. At this stage of the study, we assumed that the cleavage site is located within the "PAA-PAC" sequence (residues −4 to +2) of the NtaA precursor peptide, which culminates in 12 to 15mer. We additionally expected that the dehydratases NtaF (Protein family PF04738) and NtaG (PF14028) catalyze up to three dehydrations (on Thr4, Ser6 and Ser7), enabling thereby, in conjunction with the LanC-like enzyme NtaB, the formation of up to three (methyl)lanthionine bridges (with Cys2, Cys10 and Cys11). Moreover, the *nta* cluster contained two YcaO-encoding genes (*ntaJ* and *ntaL*) which would lead, depending on their type and partner protein, either to (methyl)azoline(s) formation or conversion of an amide(s) into a backbone thioamide(s).[12] The partner proteins were readily identified employing the RODEO[12] algorithm: the set of *ntaK/ntaL* was predicted to code for a pair of TfuA/YcaO proteins (PF07812/PF02624),[16] and *ntaI/ntaJ* to encode the tandem
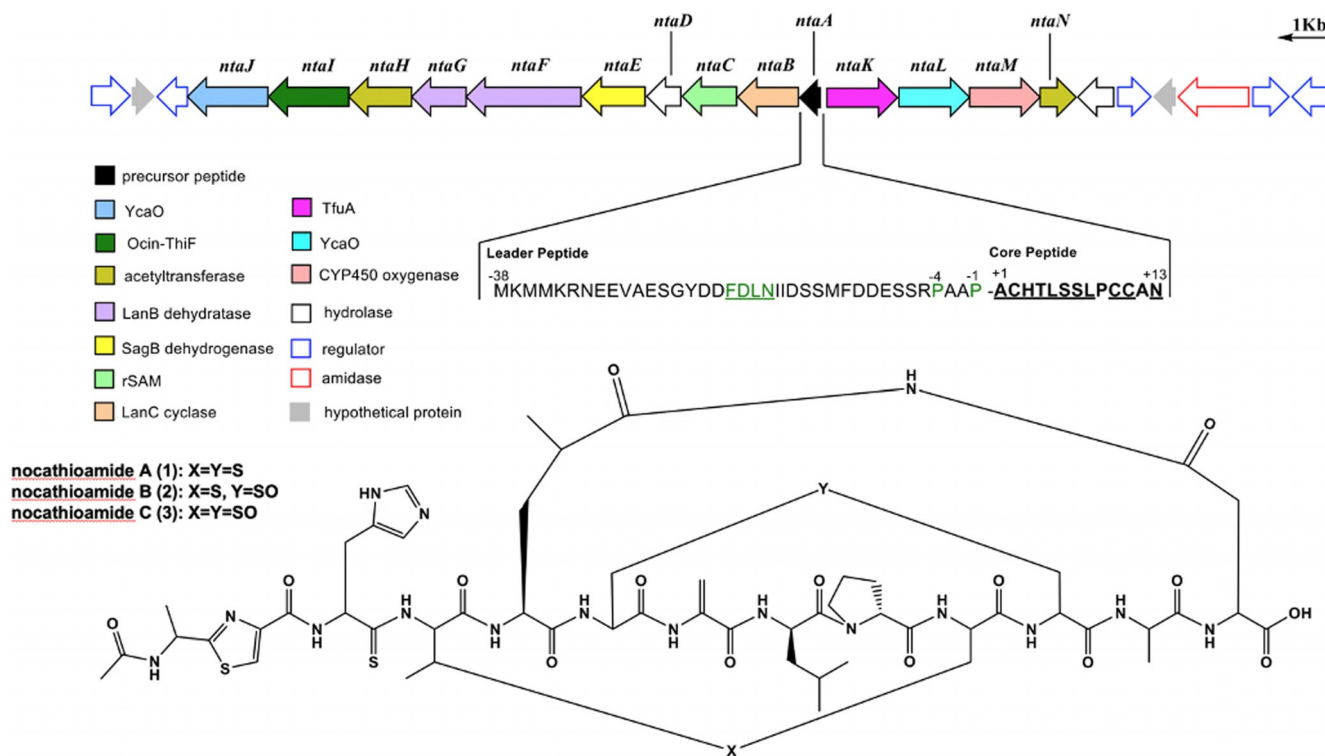

Figure 1. Top: Biosynthetic gene cluster coding for nocathioamides A–C. Bottom: Chemical structures of nocathioamides.

Ocin-ThiF/YcaO (TIGR03693/ PF02624). Thus, we hypothesized that at the resultant RiPP would be thioamidated by NtaK/NtaL and that at least one Ser/Thr/Cys residue is converted into a (methyl)azoline catalyzed by NtaI/NtaJ. Notably, the application of RRE-Finder implemented in antiSMASH 6.0[17] unveiled that the RiPP recognition element (RRE), which binds specifically to the precursor peptide (NtaA) guiding the post-translationally modifying enzymes to their substrates, is fused to NtaI (Table S1). Taking into account the presence of a SagB-dehydrogenase-encoding gene[18] (*ntaE*, PF00881), it is conceivable that (methyl)azoline would be oxidized to (methyl)azole during the maturation of the RiPP. Consequently, the presence of a (methyl)-azole ring system reduces, in turn, the maximum number of possible (methyl)lanthionine-bridges from three to two. The analysis of the remaining ORFs of the *nta* BGC revealed genes for two GNAT-N-acetyltransferases (*ntaH*, PF13527 and *ntaN*, PF00583), a CYP$_{450}$ oxygenase (*ntaM*, PF00067), a hydrolase (*ntaD*, PF01738) and a methyltransferase (*ntaC*, PF01135). Thus, we anticipated that the final peptide to be twice N-acetylated and harbors an additional hydroxyl or keto-group. Since methyltransferases catalyze a diverse spectrum of chemical reactions ranging from actual methyl transfer over epimerization to C—C crosslinking reactions,[1b, 19] no reliable prediction can be currently performed. In summary, the bioinformatic analysis suggested the production of a RiPP, formed by 12–15 amino acids featuring one to two (methyl)lanthionine bridges, thioamide bond(s), (methyl)azole ring(s) and possibly oxygenated in addition to acetylated amino acid residues. Due to the unprecedented peptide sequence and the predicted modifications of the resultant peptide, we embarked on the screening for the product(s) of the *nta* BGC.

Targeted Identification and Isolation of Nocathioamides using a Configured Metabologenomic Approach

Counting on the bioinformatically predicted structural features and the limitation of formulating only a very broad mass range of the mature peptide, a three-layered metabolomic workflow was designed to track down the resultant RiPP using OSMAC-mass spectrometry (OSMAC-MS), stable isotope labeling and ^{1}H-^{13}C HMBC NMR (Figure S1). A panel of extracts, prepared under different culture conditions (30 different culture recipes, while temperature and fermentation periods were kept constant), was screened by liquid chromatography-high resolution (LC-HR)-MS/MS with particular attention given to nitrogen-containing metabolites with a mass higher than 1000 Da. The pharma media-based extracts offered a reproducible candidate set of structurally related ions (*m/z* 1315, 1331 and 1347) with a peptidic nature inferred via their molecular formula predictions, and the intense formation of doubly charged species (Figures S2,-S2A). Initial attempts to interpret the peptides sequence using the MS/MS technique were not successful, which were attributed by then to the processing degree of the substrates under investigation.

To corroborate the candidate compounds, three feeding experiments, employing isotopically labeled (^{15}NH$_4$)$_2$SO$_4$, [^{2}H$_{10}$] L-leucine and [^{2}H$_7$] L- proline, respectively, were carried out on a small scale with strain IFM 0406. As postulated, the ^{15}N-labelling experiment led to uniformly ^{15}N-labeled peptides. In agreement with our hypothesis, each of the three candidate compounds showed in their corresponding LC/MS spectra a characteristic ^{15}N-isotope-envelope pattern that pointed to the presence of 12 to 13mer (Figures 2 A and S3,S4A). In order to prove that the targeted peptides are indeed the product of the *nta* gene cluster, [^{2}H$_{10}$] L-leucine was fed, which was expected to be incorporated twice and to produce a + 9/10 and a + 18/19/20 Da pattern. Strikingly, mass shifts of + 6, + 9, and + 15 Da were observed for all candidate compounds' pseudomolecular ions (Figures 2 A and S5,S6A), which readily hinted on the crude level prior to any isolation steps that one Leu residue must have undergone an unusual modification. Within the same context and driven by the numerous predicted possible cleavage sites, the labeled substrate [^{2}H$_7$] L-proline was supplemented into IFM 0406 cultures to validate more the features in question and the probable hydrolytic site within the prepeptide. Notably, the proline experiment exhibited an extra 7 Da as a clear indication of a single encoded proline residue within the peptides under investigation (Figures 2 A and S7,S8A), which on the one hand validated that we are tracking down the products of interest and on the other hand automatically shrank the cleavage position from the "PAAPAC" to the "AC" sequence (residues + 1 to + 2) which corroborates, in turn, the formation of a 12 to 13mer.

At this point and motivated by the suggested molecular formula which showed no fit with any known compound derived from a natural resource using SciFinder® (Figures S9–S11), IFM 0406 was fermented on a large scale (22L). Upon C$_{18}$-vacuum liquid chromatography (RP-VLC), we took advantage of the in silico predicted rare thioamide and common (methyl)azole moieties, which produce distinctive downfield resonances in NMR spectra.

Guided by LC-MS of RP-VLC fractions, a series of ^{1}H-^{13}C HMBC NMR experiments were complementarily executed on the (sub)fractions exhibiting masses > *m/z* 1300 (Figure S23). Cross correlations, typical for oxazole and/or thiazole entities ($\delta_{H/C}$ 8.18/148.32 and $\delta_{H/C}$ 8.18/174.91) were exclusively observed in the VLC-fraction, eluting with 70 % MeOH which was, beside the known secondary metabolite brasilicardin A, mainly enriched with the unknown peptides in pursuit (Figures 2 B and S23A). In contrast, pronounced downfield ^{1}H-^{13}C HMBC cross-peaks (δ_C > 200 ppm) were not detectable at all in any fractions, thusly crippling the envisioned thioamide-based NMR screening scheme. However, upon further purification and having this molecular family in hand, the thioamide motif was found to reveal weaker resonances under the applied standard NMR conditions. Using Sephadex LH-20 column chromatography, and followed by sequential rounds of RP-HPLC (Figures S13–S15), pure nocathioamides A (**1**), B (**2**) and traces of C (**3**) were obtained.
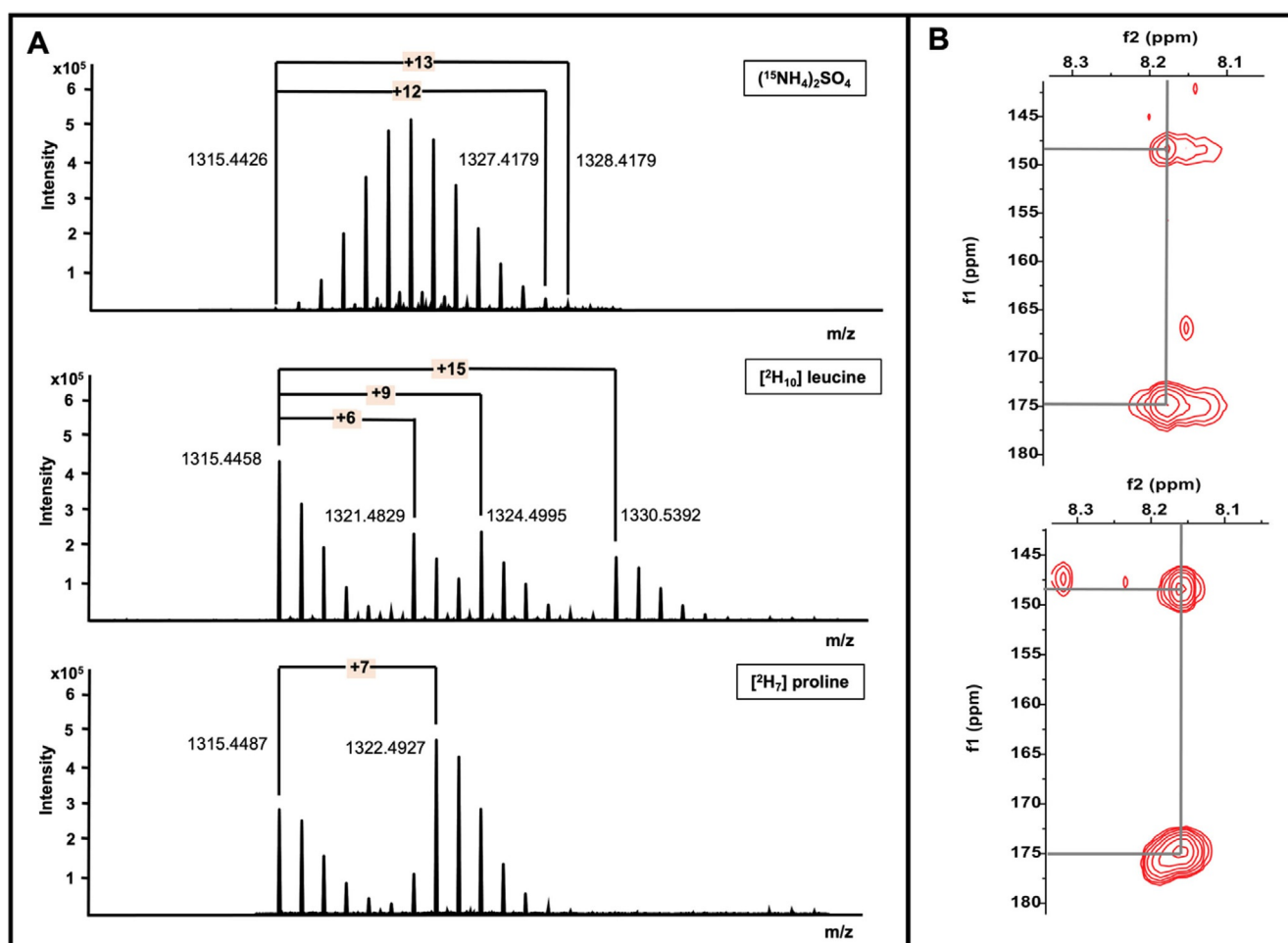
Figure 2. Details of the labeling experiments and decisive NMR key correlations, vital for the metabologenomic approach. A) HR-MS profiles of labeled nocathioamide A $[M+H]^+$; B) ^1H-^{13}C HMBC screening of azole(s) on (sub)fractions.

Structure Elucidation of Nocathioamides A–C

In combination with extensive NMR analyses, the suggested elemental composition of **1**, possessing m/z 1315 $[M+H]^+$ and 658 $[M+2H]^{2+}$, was deduced to be $C_{54}H_{74}N_{16}O_{15}S_4$ with 26 degrees of unsaturation (RDB) (Figure S9). The observation of a doubly charged species in the mass spectrum with a large number of exchangeable amide NH protons (δ_H 6.5–10.5 ppm) and carbonyl carbons (δ_C 160–185 ppm) from 1D NMR assured the peptidic nature of **1** (Figures S34,S35). The ^{13}C-NMR spectra in the differently applied solvents (d_4-, d_3-CH$_3$OH) showed the presence of two major conformers (ratio: 3:1) (Figures S25 and S49). Although a high degree of spectral overlap existed, 11 spin systems were assembled from the HSQC, COSY, TOCSY, and HSQC-TOCSY data. Briefly, among the deduced partial structures were the amino acids 2 × Ala, 1 × His, 1 × α-aminobutyric acid (Abu), 1 × Leu, 1 × 4-methylglutamate (4-Meglu), 3 × AlaS, 1 × Pro and 1 × Asn. ^1H-^{15}N and ^1H-^{13}C-based HMBC cross correlations between the side chain hydrogens with the remaining nitrogen and carbon atoms completed each amino acid moiety. Notably, one low-relaxing quaternary atom δ_C 207.99 ppm), indicative for a thiocarbonyl

functional group, showed cross-peaks with H$_2$-β (His) in the HMBC spectrum, which allowed the definition of thiocarbonyl-histidine (Figure 3 A). This fact was unequivocally apparent from a significant downfield shift of the nitrogen of Abu, involved in the thioamide bond formation (Figure 3 B). Moreover, HMBC resonances from a carbonyl group to H$_2$-β, H-γ and H$_3$-δ of a Leu residue proved it to be oxidized at the remaining C-δ in the form of 4-Meglu. Further interpretation of the ^1H-^{15}N- and ^1H-^{13}C HMBC data also revealed the presence of a thiazole (Thz) and a dehydroalanine (Dha) residues and acetylation of the N-terminal Ala moiety (Ala-1). The defined subunits were step-wise connected using a combination of foremost ^1H-^{13}C HMBC correlations from α-hydrogen resonances to the (thio)amide carbonyls of the adjacent amino acid and the observed H^{α,β}(i) → H^N(i + 1) cross-peaks in the ^1H-^1H NOESY spectra to give the primary linear sequence of **1** (Figure 4). The so far delineated structure accounted only for 23 out of the 26 RDBs, indicating that the final three degrees of unsaturation must arise from three cyclization events. Analysis of ^1H-^1H NOESY and ^1H-^{13}C HMBC spectra, recorded either in d_3- or d_4-MeOH and d_6-DMSO, proved that residues 4–10 (Abu–AlaS-10) and 6–11 (AlaS-6–AlaS-11) were connected via β-
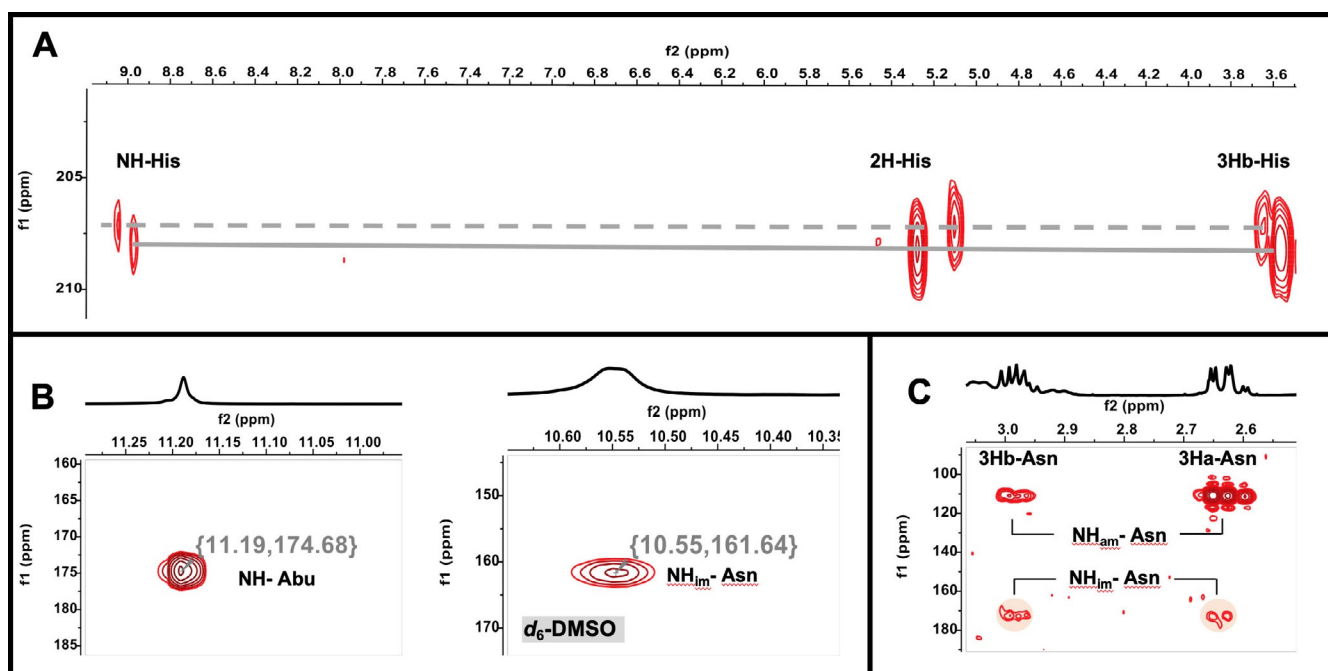
Figure 3. Key NMR correlations: A) ^{1}H-^{13}C HMBC NMR spectrum highlighting cross-peaks between C=S and 2 H, 3Hb and the amidic hydrogen of His; B) ^{1}H-^{15}N HSQC NMR spectra highlighting the chemical shifts of the NH of Abu connected to C=S and the macrocyclic imide NH of Asn; C) ^{1}H-^{15}N HMBC NMR spectrum showing the key cross-peaks between 3Ha, and 3Hb of Asn with the amidic and imidic NH of Asn.
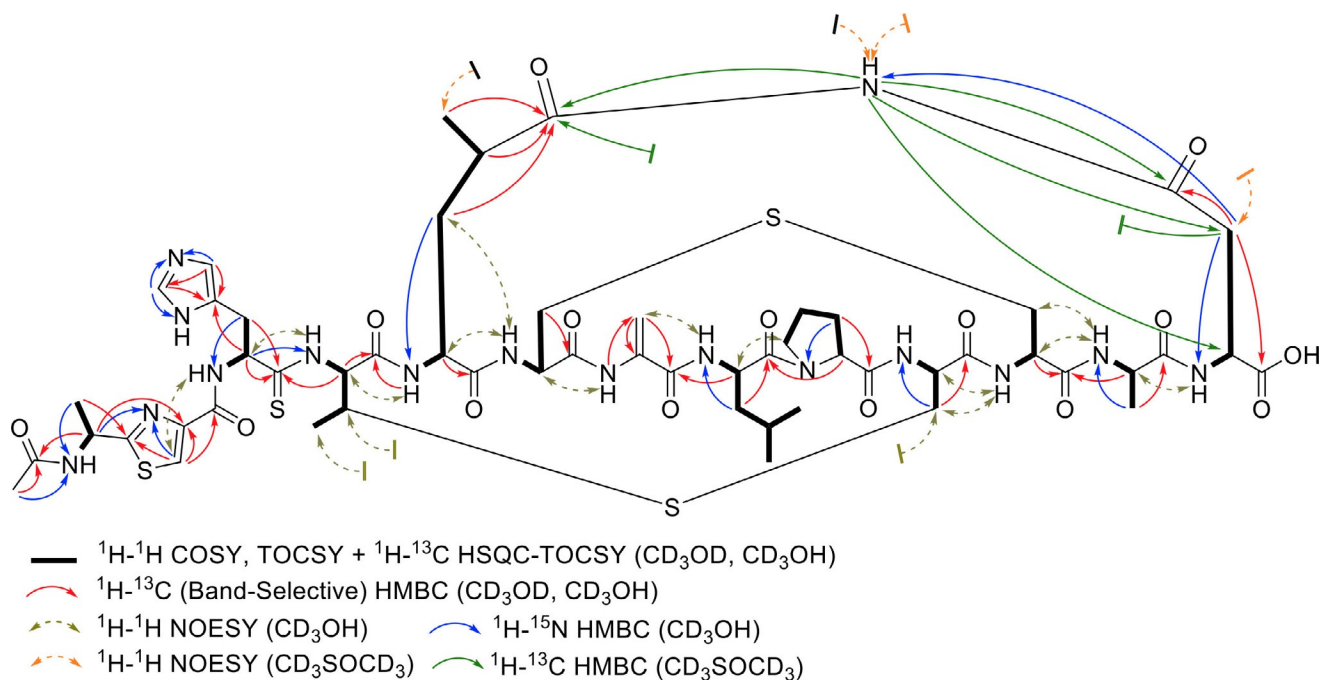

Figure 4. NMR key correlations of nocathioamide A (**1**).

thioether bonds while 5–13 (4-Meglu–Asn) were interlinked with an imide functional group (Figures 4, and S16–S21). The finding of the unprecedented imide fragment was essentially gleaned from ^{1}H-^{15}N HSQC spectrum of **1** (Figures 3 B, and S97), which provoked the reinspection of ^{1}H-^{15}N HMBC couplings, disclosing that H$_2$-β (Asn) were correlated to a nitrogen atom at 172.8 ppm, consistent with reported ^{15}N shift values for imides (Figures 3 C, and S114A).[20] Since literature about ^{15}N-NMR shift values is scarce, and the listed values strongly depend on the referencing method, we furthermore validated the ^{15}N shift readings by the measurement of the commercially available glutarimide, as an imide

model compound, whose imide nitrogen resonated at 172.3 ppm (Figure S100), thereby corroborating our conclusion and completing the 2D structure of nocathioamide A (**1**). Using the results from the initial labeling studies, we were also able to define the absolute configuration of three amino acids. Thus, the interpretation of the MS-shift patterns showed that both leucine residues are D-configured, while the proline unit possesses the L-configuration.[21]

Nocathioamide B (**2**) gave a prominent doubly charged $[M+2H]^{2+}$ peak by HRESIMS for $C_{54}H_{74}N_{16}O_{16}S_4$ (Figure S10), with an additional 16 mass units translated into an extra oxygen-atom relative to nocathioamide A (**1**). Furthermore, the NMR spectra of **2** were nearly identical with those of **1** (Figures S51, and S83), except that the α- and β-carbons of the lanthionine-bridge forming residues 6 and 11 (AlaSO-6, and AlaSO-11) showed a stark change in their chemical shifts where the α-carbons shifted upfield whilst the β-carbons resonated strongly downfield while the hydrogen values were merely not affected (Tables S3 and S6). Since such a diagnostic change in chemical shifts typically occurs upon the oxidation of the thioether functionality,[22] **2** was deduced to be the S-monooxidized product of **1**. Unlike **1** and **2**, nocathioamide C (**3**) was solely analyzed by HR-ESI-MS for $C_{54}H_{74}N_{16}O_{17}S_4$ (Figure S11), exhibiting 32 additional Da ($2 \times O$) in comparison with **1**. Thus, this difference could be readily accounted for two S-oxide groups instead of the typical thioether bridges present in **1**.

To expedite further the molecular family, the recruitment of a Global Natural Products Social (GNPS)-based molecular networking analysis seamlessly unveiled an additional suite of nocathioamide variants (Figure S124 and Table S8).[23]

Evaluation of the Biological Activity

In antibacterial, antifungal and cytotoxicity assays compounds **1** and **2** were both found to be inactive up to the highest concentration tested (MIC $> 32\ \mu g\,mL^{-1}$, IC_{50} $> 64\ \mu g\,mL^{-1}$, Table S9). Since **1** and **2** did not exhibit the classic biological activity of RiPPs, we assume that nocathioamides might possess an ecological or physiological function.[2] The identification of this function will be the subject of further studies.

Biosynthetic Considerations and Classification of Nocathioamides

The performed bioinformatic prediction of the final product was largely correct but was, owing to the novelty of the involved genes cocktail, not able to predict a) the exact cleavage site, and therefore the length of the peptide, b) the imide ring closure, c) the catalytic activity of the methyltransferase NtaC and d) that only one acetylation reaction will take place despite the presence of two acetyltransferase encoding genes (*ntaH* and *ntaN*). Regarding the latter issue, only one acetyltransferase is apparently contributing in the biosynthesis. We hypothesize that NtaH is the decisive enzyme for the N-acetylation of the Ala-1 residue due to

the analysis of the very recently sequenced genome of *Longimycelium tulufanense* CGMCC 4.5737[24] (NZ_BMMK01000000) which also contained the *nta* cluster (Table S1). Strikingly, all annotated *nta* genes exhibited a high identity and similarity score, except for NtaN which showed no significant similarity with all proteins of strain CGMCC 4.5737. Since the nocathioamides A–C (**1**–**3**) bear no C- or N-methylated groups, but instead carries two D-configured leucine residues, the methyltransferase, NtaC, was hypothesized to be involved in the amino acid epimerization process, which is well established in a RiPP biosynthetic context.[25]

From a PTM point of view, nocathioamides represent a hypermodified RiPPs. 11 out of its 13 residues underwent different and specific enzymatic modifications to deliver unique chemical alterations [(Me)Lan, thioamidation, azole formation, N-acetylation, (C and S) oxidation, dehydration, epimerisation, and macroimidation]. Most intriguing is the unusual oxidation of the δ-carbon atom of Leu5 to deliver 4-Meglu, possibly mediated by the action of the cytochrome P_{450} enzyme (CYP_{450}) NtaM, which prepares the molecule for a macrocyclic imide ring closure. To the best of our knowledge, such a regio-specific CYP_{450}-based oxidation reaction of Leu was formerly reported so far for only seven NRPS-derived natural products, such as echinocandins, nostopeptolide, polyoxypeptin B, griselimycin, perthamide C, monamycin D and ilamycin,[26] but never in a RiPP context (Figure S123). Moreover, the involved oxidative reactions never led to an imide macrocyclization, but rather to acyclic 4-methylglutamate or five- and six-membered heterocycles such as 4-methylproline or cyclic hemiaminal substructures, respectively. Imide moieties themselves exist in several secondary metabolites, however either as succinimide or glutarimide units and not as a key element for ring closures of macrocycles, highlighting the uniqueness of nocathioamides again.

Another feature of the nocathioamide family is the single and double sulfoxidation of Lan and MeLan bridges. Sulfoxidated variants are described still in only a few RiPPs[22,27] and the process is either enzymatically mediated by a clustered monooxygenase[22a] or occurs spontaneously under alkaline conditions.[22c]

For the classification of nocathioamides, according to the rules of the RiPP community,[1b] we consulted the *nta* BGC. Concerning the precursor peptide we found a hitherto unknown new AP cleavage site. Furthermore, we observed the presence of a split *lanB*-like dehydratase (*ntaF/ntaG*) beside a *lanC* like cyclase (*ntaB*), which clearly categorized the nocathioamides as a member of the class I lanthipeptides. Notably, the co-occurrence of a split *lanB* with *lanC* was singly reported for class I lantibiotics of the antifungal pinensins despite its frequent existence in LAPs and thiopeptides.[4,28] It is also noteworthy to mention the thioamidation, which exclusively occurs at the His residue. In contrast to the few so far discovered thioamide containing RiPPs, recent genome mining efforts commenced to leverage and expand this chemical space with new thioamidated candidates either as thioviridamide-similar entities[29] or as thiopeptides, exemplified by saalfelduracin and thiopeptin.[12] Interestingly, the associations of lanthionine synthetases with either F-dependent and/or TfuA-paired YcaOs were bioinformatically

projected before by the teams of Mitchell and van der Donk to chart novel multi-biosynthetic systems[30] which were followed up by their engineering efforts through combinatorial biosynthesis to mix and match new-to-nature hybrid RiPPs as thiazolinyl-lanthipeptides (I, II).[31] Along the same lines, the nocathioamide family elegantly symbolises the first-in-nature combinatorial tribrid RiPPs hovering over three different class-defining biosynthetic machineries of lanthi-peptides, LAPs and thioamitides, decorated with an additional unique PTM.

Conclusion

In this study, we demonstrated the utility of a newly optimized metabologenomic approach through its application to an orphan gene cluster found in the genomes of *N. terpenica* IFM 0406 and 0706[T]. This approach led to the isolation of three novel peptides, nocathioamides A–C (**1–3**), which would have been overlooked by regular LC/MS screening due to peak overlapping and/or the appearance as a minor peak (Figure S12). We achieved the isolation by the bioinformatic prediction of the resulting peptide and used this information, after a round of media optimization, to screen for specific mass shifts of some amino acid residues upon labeling besides particular NMR features after fractionation. The benefit of the isotopic labeling experiments was three-fold: They early connected the suspected chemotypes with the BGC of interest, narrowed down the cleavage site and allowed the determination of the absolute configuration of the involved amino acids. Such a developed hybrid MS/NMR tactic represents a valuable complement to the existing genome mining strategies, particularly for RiPPs lacking predictable bioactivity or mass ranges. Albeit the adopted approach applied in this study resulted in characterizing a thiazole-containing RiPP, it still could be modulated to target other rare functional groups of RiPPs. For example, the integration of a ^{1}H-^{15}N-HSQC NMR screening component can enable to delve more RiPPs featuring imide moieties or unusual nitrogen crosslinks. Furthermore, with a refocused ^{1}H-^{13}C-HMBC sequence that takes into account the low-relaxing quarternary atoms of thioamide carbonyl groups, it could be implemented to unearth further thioamidated RiPPs.

While the discovery of nocathioamides brings up several firsts into RiPPs, the *nta* BGC sets the stage for studying the interplay of the involved enzymes to elucidate the currently enigmatic biosynthetic events. This will possibly allow to expand the biobricks toolbox in synthetic biology towards engineering customized hybrid RiPPs exploiting the rare PTMs and the exceptional combinations of the different biogenesis machineries offered by nocathioamides chimeric gene cluster.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

[1] a) M. A. Ortega, W. A. van der Donk, *Cell Chem. Biol.* **2016**, *23*, 31–44; b) M. Montalbán-López, T. A. Scott, S. Ramesh, I. R. Rahman, A. J. van Heel, J. H. Viel, V. Bandarian, E. Dittmann, O. Genilloud, Y. Goto, M. J. Grande Burgos, C. Hill, S. Kim, J. Koehnke, J. A. Latham, A. J. Link, B. Martinez, S. K. Nair, Y. Nicolet, S. Rebuffat, H.-G. Sahl, D. Sareen, E. W. Schmidt, L. Schmitt, K. Severinov, R. D. Süssmuth, A. W. Truman, H. Wang, J.-K. Weng, G. P. van Wezel, Q. Zhang, J. Zhong, J. Piel, D. A. Mitchell, O. P. Kuipers, W. A. van der Donk, *Nat. Prod. Rep.* **2021**, *38*, 130–239.

[2] Y. Li, S. Rebuffat, *J. Biol. Chem.* **2020**, *295*, 34–54.

[3] L. M. Repka, J. R. Chekan, S. K. Nair, W. A. van der Donk, *Chem. Rev.* **2017**, *117*, 5457–5520.

[4] a) B. J. Burkhart, C. J. Schwalen, G. Mann, J. H. Naismith, D. A. Mitchell, *Chem. Rev.* **2017**, *117*, 5389–5456; b) G. A. Hudson, A. R. Hooper, A. J. DiCaprio, D. Sarlah, D. A. Mitchell, *Org. Lett.* **2021**, *23*, 253–256.

[5] N. Mahanta, D. M. Szantai-Kis, E. J. Petersson, D. A. Mitchell, *ACS Chem. Biol.* **2019**, *14*, 142–163.

[6] a) K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, *Nucleic Acids Res.* **2019**, *47*, W81–W87; b) J. I. Tietz, C. J. Schwalen, P. S. Patel, T. Maxson, P. M. Blair, H.-C. Tai, U. I. Zakai, D. A. Mitchell, *Nat. Chem. Biol.* **2017**, *13*, 470–478; c) J. Santos-Aberturas, G. Chandra, L. Frattaruolo, R. Lacret, T. H. Pham, N. M. Vior, T. H. Eyles, A. W. Truman, *Nucleic Acids Res.* **2019**, *47*, 4624–4637.

[7] Z. Zhong, B. He, J. Li, Y.-X. Li, *Synth. Syst. Biotechnol.* **2020**, *5*, 155–172.

[8] a) C. L. Cox, J. I. Tietz, K. Sokolowski, J. O. Melby, J. R. Doroghazi, D. A. Mitchell, *ACS Chem. Biol.* **2014**, *9*, 2014–2022; b) H. Mohimani, R. D. Kersten, W. T. Liu, M. Wang, S. O. Purvine, S. Wu, H. M. Brewer, L. Pasa-Tolic, N. Bandeira, B. S. Moore, P. A. Pevzner, P. C. Dorrestein, *ACS Chem. Biol.* **2014**, *9*, 1545–1551; c) L. Cao, A. Gurevich, K. L. Alexander, C. B. Naman, T. Leao, E. Glukhov, T. Luzzatto-Knaan, F. Vargas, R. Quinn, A. Bouslimani, L. F. Nothias, N. K. Singh, J. G. Sanders, R. A. S. Benitez, L. R. Thompson, M.-N. Hamid, J. T.

Morton, A. Mikheenko, A. Shlemov, A. Korobeynikov, I. Friedberg, R. Knight, K. Venkateswaran, W. H. Gerwick, L. Gerwick, P. C. Dorrestein, P. A. Pevzner, H. Mohimani, *Cell Syst.* **2019**, *9*, 1–9.

[9] D. Dhakal, V. Rayamajhi, R. Mishra, J. K. Sohng, *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 385–407.

[10] V. Wiebach, A. Mainz, M.-A. J. Siegert, N. A. Jungmann, G. Lesquame, S. Tirat, A. Dreux-Zigha, J. Aszodi, D. Le Beller, R. D. Süssmuth, *Nat. Chem. Biol.* **2018**, *14*, 652–654.

[11] a) A. Buchmann, M. Eitel, P. Koch, P. N. Schwarz, E. Stegmann, W. Wohlleben, M. Wolański, M. Krawiec, J. Zakrzewska-Czerwinska, C. Méndez, A. Botas, L. E. Núñez, F. Morís, J. Cortés, H. Gross, *Genome Announc.* **2016**, *4*, e01391; b) J. Chen, A. Frediansyah, D. Männle, J. Straetener, H. Brötz-Oesterhelt, N. Ziemert, L. Kaysser, H. Gross, *ChemBioChem* **2020**, *21*, 2205–2213; c) A. Buchmann, H. Gross, *Microbiol. Resour. Announc.* **2020**, *9*, e00689-20.

[12] C. J. Schwalen, G. A. Hudson, B. Kille, D. A. Mitchell, *J. Am. Chem. Soc.* **2018**, *140*, 9494–9501.

[13] A. Angelopoulou, A. K. Warda, P. M. O'Connor, S. R. Stockdale, A. N. Shkoporov, D. Field, L. A. Draper, C. Stanton, C. Hill, R. P. Ross, *Front. Microbiol.* **2020**, *11*, 788.

[14] S. M. Asaduzzaman, K. Sonomoto, *J. Biosci. Bioeng.* **2009**, *107*, 475–487.

[15] P. Agrawal, S. Khater, M. Gupta, N. Sain, D. Mohanty, *Nucleic Acids Res.* **2017**, *45*, W80–W88.

[16] a) N. Mahanta, A. Liu, S. Dong, A. K. Nair, D. A. Mitchell, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3030–3035; b) J. Liu, Z. Lin, Y. Li, Q. Zheng, D. Chen, W. Liu, *Org. Biomol. Chem.* **2019**, *17*, 3727–3731.

[17] A. M. Kloosterman, K. E. Shelton, G. P. van Wezel, M. H. Medema, D. A. Mitchell, *mSystems* **2020**, *5*, e00267-20.

[18] S. W. Lee, D. A. Mitchell, A. L. Markley, M. E. Hensler, D. Gonzalez, A. Wohlrab, P. C. Dorrestein, V. Nizet, J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5879–5884.

[19] a) N. Mahanta, G. A. Hudson, D. A. Mitchell, *Biochemistry* **2017**, *56*, 5229–5244; b) M. I. Gibson, C. C. Fornerisa, M. R. Seyedsayamdost in *Cyclic Peptides* (Eds.: J. Koehnke, J. Naismith, W. A. van der Donk), Royal Society of Chemistry, London, **2017**, pp. 141–163; c) J. Lu, Y. Li, Z. Bai, H. Lv, H. Wang, *Nat. Prod. Rep.* **2021**, https://doi.org/10.1039/D0NP00044B.

[20] G. C. Levy, R. L. Lichter in *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, Wiley, New York, **1979**.

[21] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chem. Eur. J.* **2012**, *18*, 2342–2348.

[22] a) S. Somma, W. Merati, F. Parenti, *Antimicrob. Agents Chemother.* **1977**, *11*, 396–401; b) L. Vertesy, W. Aretz, A. Bonnefoy, E. Ehlers, M. Kurz, A. Markus, M. Schiell, M. Vogel, J. Wink, H. Kogler, *J. Antibiot.* **1999**, *52*, 730–741; c) S. I. Maffioli, M. Iorio, M. Sosio, P. Monciardini, E. Gaspari, S. Donadio, *J. Nat. Prod.* **2014**, *77*, 79–84.

[23] M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderon, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya P, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodriguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. O. Palsson, K. Pogliano, R. G. Linington, M. Gutierrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein, N. Bandeira, *Nat. Biotechnol.* **2016**, *34*, 828–837.

[24] W. Linhuan, M. Juncai, *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 895–898.

[25] a) A. Bhushan, P. J. Egli, E. E. Peters, M. F. Freeman, J. Piel, *Nat. Chem.* **2019**, *11*, 931–939; b) M. Korneli, S. W. Fuchs, K. Felder, C. Ernst, L. V. Zinsli, J. Piel, *ACS Synth. Biol.* **2021**, *10*, 236–242.

[26] a) J. Ma, H. Huang, Y. Xie, Z. Liu, J. Zhao, C. Zhang, Y. Jia, Y. Zhang, H. Zhang, T. Zhang, J. Ju, *Nat. Commun.* **2017**, *8*, 391; b) M.-C. Tang, Y. Zou, K. Watanabe, C. T. Walsh, Y. Tang, *Chem. Rev.* **2017**, *117*, 5226–5333.

[27] M. Simone, P. Monciardini, E. Gaspari, S. Donadio, S. I. Maffiolo, *J. Antibiot.* **2013**, *66*, 73–78.

[28] K. I. Mohr, C. Volz, R. Jansen, V. Wray, J. Hoffmann, S. Bernecker, J. Wink, K. Gerth, M. Stadler, R. Müller, *Angew. Chem. Int. Ed.* **2015**, *54*, 11254–11258; *Angew. Chem.* **2015**, *127*, 11406–11410.

[29] a) L. Frattaruolo, R. Lacret, A. R. Cappello, A. W. Truman, *ACS Chem. Biol.* **2017**, *12*, 2815–2822; b) T. Kawahara, M. Izumikawa, I. Kozone, J. Hashimoto, N. Kagaya, H. Koiwai, M. Komatsu, M. Fujie, N. Sato, H. Ikeda, K. Shin-ya, *J. Nat. Prod.* **2018**, *81*, 264–269.

[30] a) Q. Zhang, J. R. Doroghazi, X. Zhao, M. C. Walker, W. A. van der Donk, *Appl. Environ. Microbiol.* **2015**, *81*, 4339–4350; b) M. C. Walker, S. M. Eslami, K. J. Hetrick, S. E. Ackenhusen, D. A. Mitchell, W. A. van der Donk, *BMC Genomics* **2020**, *21*, 387; c) C. Wu, W. A. van der Donk, *Curr. Opin. Biotechnol.* **2021**, *69*, 221–231.

[31] B. J. Burkhart, N. Kakkar, G. A. Hudson, W. A. van der Donk, D. A. Mitchell, *ACS Cent. Sci.* **2017**, *3*, 629–638.