Natural product biosynthesis in myxobacteria: Studies on enzymatic versatility and secondary metabolite diversity

Dissertation zur Erlangung des Grades des
Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes

von
Daniel Krug

Saarbrücken
2009

Tag des Kolloquiums:	02. Juni 2009
Dekan:	Prof. Dr. Stefan Diebels
Berichterstatter:	Prof. Dr. Rolf Müller
	Prof. Dr. Uli Kazmaier
	Prof. Dr. Andreas Bechthold

Vorveröffentlichungen der Dissertation

Ergebnisse dieser Arbeit wurden mit Genehmigung der Naturwissenschaftlich-Technischen Fakultät III, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht oder sind derzeit in Vorbereitung zur Veröffentlichung:

1. **Rachid S., Krug D., Weissman K.J., and Müller R. (2007).** Biosynthesis of (*R*)-β-tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chondromyces crocatus*. *J.Biol.Chem.* 282, 21810-21817.

2. **Krug D. and Müller R. (2009).** Discovery of additional members of the tyrosine aminomutase enzme family and the mutational analysis of CmdF. *ChemBioChem* 10 (4), 741-750.

3. **Rachid S., Krug D., Kochems I., Kunze B., Scharfe M., Blöcker H., Zabriski M., and Müller R. (2006).** Molecular and biochemical studies of chondramide formation - highly cytotoxic natural products from *Chondromyces crocatus* Cm c5. *Chem.Biol.* 14, 667-681.

4. **Krug, D., Dickschat, J. S., and Müller, R.** (**2009**). Imidacins: Structure and biosynthesis of novel secondary metabolites from *Stigmatella aurantiaca* incorporating an urocanate building block. *Manuscript to be submitted*.

5. **Berner M., Krug D., Bihlmaier C., Vente A., Müller R., and Bechthold A. (2006).** Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. *J.Bacteriol.* 188, 2666-2673.

6. **Krug D., Zurek G., Revermann O., Vos M., Velicer G.J., and Müller R. (2008).** Discovering the hidden secondary metabolome of *Myxococcus xanthus*: a study of intraspecific diversity. *Appl.Environ.Microbiol.* 74, 3058-3068.

7. **Krug D., Zurek G., Schneider B., Garcia R., and Müller R. (2008).** Efficient mining of myxobacterial metabolite profiles enabled by liquid-chromatography – electrospray ionization – time-of-flight mass spectrometry and compound-based principal component analysis. *Anal.Chim.Acta* 624, 97-106.

8. **Garcia R.O., Krug D., Müller R. (2009).** Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods. *In:* Hopwood, D. (Ed.), Methods in Enzymology, Vol. 458 (A).

Danksagung

An erster Stelle möchte ich Herrn Prof. Dr. Rolf Müller danken für die Aufnahme in seinen Arbeitskreis, die Überlassung der interessanten Themen sowie für unzählige Anregungen, das stets interessierte Zuhören und viele hilfreiche Diskussionen. Sein fortwährendes Engagement für den Aufbau einer außergewöhnlichen Forschungsumgebung während der letzten Jahre war beeindruckend. Ganz besonders dankbar bin ich auch für die erheblichen Freiräume bei der Bearbeitung der Themen und für die Chance, die Zielrichtung der eigenen Forschungsarbeiten mitzubestimmen. Bei Herrn Prof. Dr. Uli Kazmaier möchte ich mich herzlich für die Übernahme des Korreferates bedanken.

Ich danke Herrn Dr. Klaus Gerth vom HZI Braunschweig für seine Unterstützung und Hilfsbereitschaft zunächst während der kurzen Zeit in Braunschweig und danach „aus der Ferne“. Ein besonderes Dankeschön möchte ich weiterhin allen Co-Autoren der Veröffentlichungen, die Bestandteil der vorliegenden Arbeit sind, aussprechen. Im einzelnen möchte ich mich insbesondere bei Prof. Dr. Gregory J. Velicer für viele interessante Anregungen bedanken, sowie bei Dr. Kira J. Weissman für ihre unerschütterliche Hilfsbereitschaft in Angelegenheiten der Englischen Sprache. Darüber hinaus danke ich Dr. Gabriela Zurek und ihren Kollegen von der Firma Bruker für die gute Zusammenarbeit und ihre Unterstützung bei massenspektrometrischen Fragen.

Ole Revermann danke ich für die Aufnahme von NMR-Spektren. Sehr gefreut habe ich mich außerdem darüber, Eva Luxenburger als Mitstreiterin in Sachen HPLC- und MS-Technik zu haben. Allen jetzigen und vielen ehemaligen Mitgliedern der Arbeitsgruppe danke ich für die gute Zusammenarbeit und die freundliche Atmosphäre – beides unverzichtbare Voraussetzungen für Spaß an der Arbeit. Mein ganz besonderer Dank gilt außerdem Silke Wenzel, Tina Binz, Bettina Frank und Kathrin Buntin, für viele lustige Momente während und außerhalb des wissenschaftlichen Alltags.

Nicht zuletzt möchte ich meinen Eltern herzlich danken für ihre langjährige und vielfältige Unterstützung während des Studiums und der gesamten Doktorandenzeit.

Zusammenfassung

Diese Arbeit beschäftigt sich mit der Biosynthese von myxobakteriellen Naturstoffen, mit Schwerpunkt auf der funktionellen und biochemischen Charakterisierung von Enzymen, die ungewöhnliche Vorläufer für biosynthetische Stoffwechselwege zur Verfügung stellen. Es konnte gezeigt werden, dass (*R*)-β-Tyrosin – ein Bestandteil der Chondramide aus *Chondromyces crocatus* – durch die neuartige Tyrosin-Aminomutase (TAM) CmdF gebildet wird. Mittels gezielter Mutagenese wurden Möglichkeiten zur Beeinflussung der Stereoselektivität von CmdF aufgezeigt. Die Klonierung und Characterisierung weiterer TAM Enzyme offenbarte einen hohen Grad an evolutionärer Diversifizierung innerhalb der erweiterten TAM Familie. Außerdem wird die Strukturaufklärung und die Biosynthese neu entdeckter Metabolite aus *Stigmatella aurantiaca* beschrieben, welche Imidacine genannt wurden. Diese Verbindungen besitzen eine fettsäureartige Grundstruktur, zeichnen sich jedoch zusätzlich durch eine äußerst seltene Imidazolacrylgruppe und einen Cyclopropanring aus. Die Biosynthese der Imidacine ist abhängig von der Bereitstellung einer Urocansäure-Startereinheit durch eine Histidin-Ammoniumlyase (HAL), und stellt damit das bisher einzige Beispiel für die Beteiligung von HALs am bakteriellen Sekundärstoffwechsel dar.

Es wurde darüber hinaus ein „Metabolomics"-basierter Ansatz zur vertieften Analyse der Sekundärmetabolitenprofile von Myxobakterien entwickelt. Hochauflösende massenpektrometrische Messungen wurden mit einem statistischen Auswertungsverfahren (PCA, „Principal Component Analysis") kombiniert, um solche Substanzen zu identifizieren, die signifikante Unterschiede innerhalb einer Probenserie mit hoher Komplexität darstellen. Die Anwendung dieser Methode enthüllte die erstaunlich hohe Diversität des sekundären Metaboloms der Spezies *Myxococcus xanthus*.

Abstract

In this work, the biosynthesis of secondary metabolites in myxobacteria was investigated, with particular focus on the functional and biochemical characterization of enzymes furnishing biosynthetic pathways with unusual building blocks. A novel member of the tyrosine aminomutase (TAM) enzyme family, CmdF, was shown to supply (*R*)-β-tyrosine for incorporation into the chondramides by *Chondromyces crocatus*. Site-directed mutagenesis demonstrated the feasibility of modulating the stereoselectivity exhibited by CmdF through genetic engineering. The cloning and characterization of additional TAMs revealed a high degree of evolutionary diversification within the emerging TAM enzyme family. Furthermore, the structural elucidation and biosynthesis of newly discovered metabolites from *Stigmatella aurantiaca*, the imidacins, is reported. These compounds consist of a fatty-acid like scaffold, but additionally feature an unique imidazole acryl moiety and a rare cyclopropane ring. Imidacin formation depends on the supply of an unprecedented urocanic acid precursor by a histidine ammonia lyase (HAL), revealing for the first time the involvement of HAL in bacterial secondary metabolism.

In addition, a metabolomics-based approach for the mining of myxobacterial metabolite profiles was devised. High-resolution mass spectrometry, combined with statistical interpretation by principal component analysis (PCA), was employed to highlight compounds which constitute significant differences in a set of highly complex samples. This technique uncovered a strikingly high level of intraspecific diversity in the secondary metabolome of *Myxococcus xanthus*.

Contents

A. Introduction

B. Publications

C. Discussion

D. References

E. Appendix

A. Introduction

1. The role of natural products in drug discovery

Natural products from microbial sources exhibit a wide range of biological activities and are characterized by an extraordinary degree of structural diversity. A number of these so-called secondary metabolites are already of medical importance for the treatment of numerous diseases, or serve as leads for the development of novel drugs.[1] Most antibiotics are in fact secondary metabolites isolated from microorganisms or derivatives thereof, and medicines based on natural products account for a high percentage of annual pharmaceutical sales.[2,3] Macrolide antibiotics such as erythromycin, the vancomycin family of glycopeptides, and the lipopeptide daptomycin (Figure A-1, **1**-**3**) represent examples of valuable anti-infectives discovered from bacteria. These and various other compound classes are in clinical use today.[4-6] Besides antibacterial therapeutics, microbial secondary metabolites with potent antifungal, immunosuppressive or cytotoxic activity are known (Figure A-1, **4**-**6**).[2,7-9] Natural products from microbial sources are also used

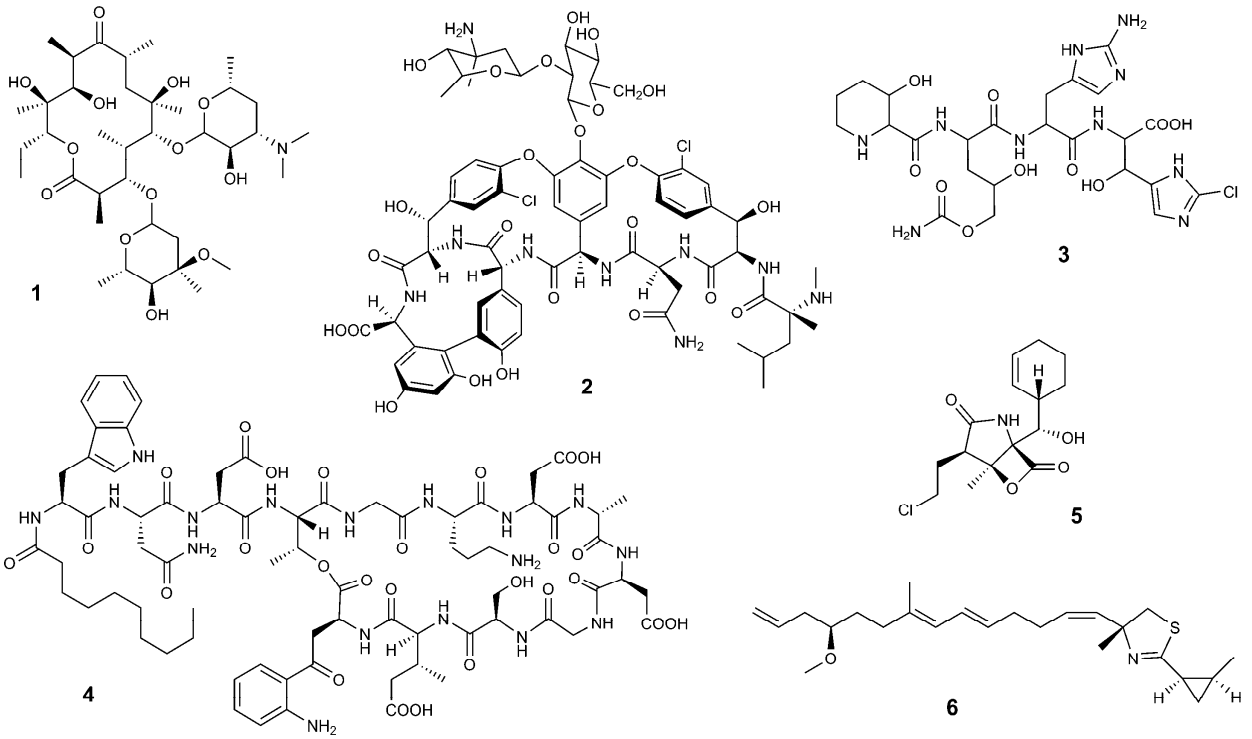


Figure A-1. Structural diversity of bioactive secondary metabolites from bacteria (activity and producing organism are given in parentheses). **1**, erythromycin (antibacterial, *Saccharopolyspora erythrea)*; **2**, vancomycin (antibacterial, *Amycolatopsis orientalis*); **3**, GE81112 (antibacterial, *Streptomyces* sp.); **4**, daptomycin (antibacterial, *Streptomyces roseosporus*); **5**, salinisporamide (cytotoxic, *Salinispora tropica*); **6**, curacin (cytotoxic, *Lyngba majuscula*).

in veterinary medicine, and some are employed in agriculture due to their insecticidal, herbicidal or plant growth-promoting activities.[2,10]

The emergence of clinically relevant pathogens resistant to an increasing number of antibiotics, and the appearance of novel infectious diseases, generate a continued need to discover and develop new biologically active lead structures. Natural products from bacteria are believed to yield an increased success rate when searching for drug candidates, in comparison with screens employing synthetic compound libraries. This perspective is substantiated by the notion that microbial natural products represent the outcome of an evolutionary optimization process involving selection for a specific biological activity.[11] Notably, the chemical complexity and structural diversity of bacterial secondary metabolites rivals that of natural products from other established sources, such as plants and fungi. However, the biosynthetic potential of certain orders of bacteria, for example the *Actinomycetales*, has already been extensively exploited for the discovery of natural products in the course of the last century.[12] The probability of finding a novel chemical structure produced by a species from such phylogenetic groups has therefore decreased, and the last several years have consequently seen increased efforts to tap alternative bacterial natural product resources.

Encouragingly, it is estimated that only a minor proportion (as low as 1%) of prokaryotic phyolgenetic diversity has been unearthed so far.[10] Thus, the identification of novel bacterial secondary metabolite producers is expected to lead to the discovery of many novel natural products with interesting biological activities in the future. This assumption seems to apply especially to bacterial isolates from habitats which have not been thoroughly investigated.[13] Evolutionary adaption to a specific ecological niche might be a driving force for the production of secondary metabolites which help to increase fitness of the producing organism. Moreover, novel producers could also be identified in bacterial taxa whose growth characteristics were formerly seen as an obstacle to laboratory cultivation, and by investigating bacteria with an unusual "lifestyle". Prokaryotic groups which have been established rather recently as viable sources of secondary metabolites (when compared e.g. to the streptomycetes) include the cyanobacteria,[14] certain bacterial endosymbionts,[15,16] the marine actinomycetes,[17] the pseudomonads and the myxobacteria.[18,19]

2. Myxobacteria – promising producers of bioactive secondary metabolites

Myxobacteria are soil-living, Gram-negative bacteria that exhibit some outstanding characteristics.[20] They can glide over solid surfaces as coordinated swarms and feed cooperatively by degrading biological macromolecules or even live prey microorganisms, which is accomplished,

at least in part, by secretion of lytic enzymes. Myxobacteria are also able to undergo multicellular development when they encounter starvation conditions. This process begins with the ordered aggregation of cells and culminates in the formation of fruiting bodies containing resting cells, the myxospores, which are resistant to environmental stress such as temperature changes, desiccation, and UV-radiation. When favorable nutritional and environmental conditions are restored, the spores germinate and retransform into viable cells. Myxobacterial fruiting bodies vary in shape and size, from simple cell clusters to sophisticated tree-like structures that emerge from the surface, and their morphology is routinely considered for taxonomic classification of myxobacterial isolates.[21]

Over the last three decades, the myxobacteria have made their way from exotic organisms which gained attention mainly due to their uncommon morphological characteristics and complex lifestyle, to an established source of natural products that exhibit diverse structures and biological activities.[22,23] More than 7500 myxobacterial strains have been isolated by workgroups at the German Research Center for Biotechnology (Braunschweig, Germany; recently renamed

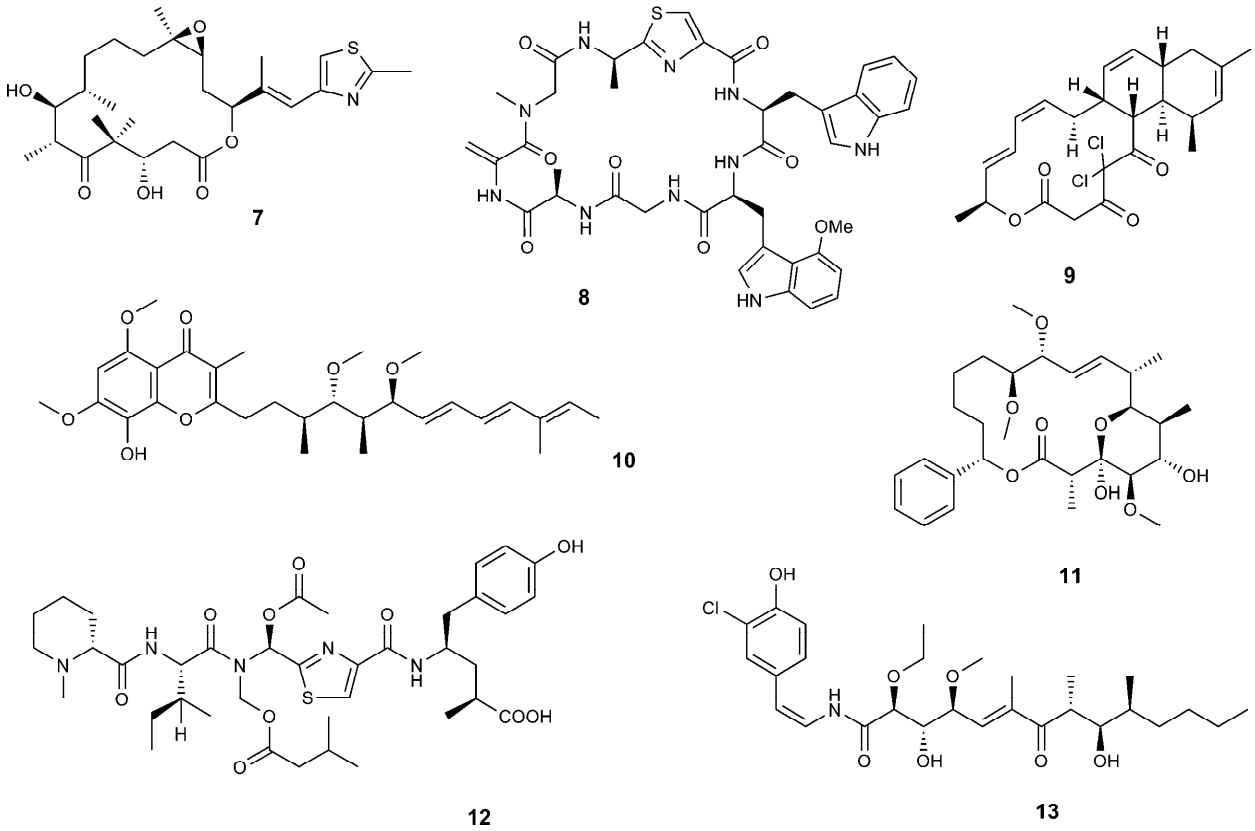


Figure A-2. Examples of myxobacterial secondary metabolites (activity and producing organism are given in parentheses). **7**, epothilon B (antitumor, *Sorangium cellulosum)*; **8**, argyrin A (antitumor, immunosupressive, *Archangium gephyra*); **9**, chlorotonil A (*Sorangium cellulosum*); **10**, stigmatellin A (antifungal, *Stigmatella aurantiaca*); **11**, soraphen A (antifungal, *Sorangium cellulosum*); **12**, tubulysin A (cytotoxic, *Angiococcus disciformis*); **13**, chondrochloren B (antifungal, *Chondromyces crocatus*).

to Helmholtz Center for Infection Research), and a high percentage of these have been revealed as proficient producers of secondary metabolites. To date, more than 100 basic structures and more than 500 derivatives thereof have been discovered from myxobacteria.[24] Many of these compounds display structural features that were novel at the time of their discovery. Likewise, many myxobacterial natural products have been found to exhibit biological activities that are otherwise rarely observed from compounds of bacterial origin, and a number of secondary metabolites from myxobacteria exhibit unique modes of action.[22,23] These findings make the discovery of natural products from myxobacteria especially appealing, because chances are increased to find compounds which may have the potential to serve as novel leads for drug development.

A few selected examples of myxobacterial secondary metabolites are shown in Figure A-2. Epothilone **7** and tubulysin **12** act on the cytoskeleton of eukaryotic cells with contrary effects: epothilone stabilizes microtubules, while tubulysin inhibits the polymerization of tubulin; either mechanism leads to an arrest of the cell cycle and finally induces apoptosis.[25,26] Epothilone exhibits activity against diverse tumor cell lines, and an epothilone lactame derivative was recently approved as an anti-cancer drug. Argyrin **8** is another myxobacterial compound with potent antitumor activity as this natural product was shown to exert its main effect through inhibition of the proteasome.[27] Stigmatellin **10** and soraphen **11** represent antifungal compounds with different modes of action.[28,29] Chondrochloren **13** displays antibiotic activity against a variety of microorganisms, and the biological activity of chlorotonil **9** is currently being evaluated.[30-32]

The striking diversity of myxobacterial natural product structures and biological activities has triggered considerable efforts to investigate the corresponding biosynthetic pathways.[23] Molecular and biochemical studies continue to contribute to an enhanced understanding of the complex biosynthetic machinery that underlies secondary metabolite production in myxobacteria. This knowledge provides the basis for attempts to improve production yields or manipulate biosynthetic pathways for the production of "unnatural" structural variants of natural products.[33] During the last ten years, a sizeable number of secondary metabolite biosynthesis gene clusters from various myxobacterial genera has been identified and characterized. The genetic information obtained thereby paves the way to use the enzymatic machineries encoded by secondary metabolite biosynthesis-related genes for combinatorial biosynthesis and for the genetic engineering of artificial biosynthetic pathways which yield novel compounds, sometimes even with predictable structures.[34-37] Furthermore, entire biosynthetic pathways may be transferred into heterologous hosts that offer biotechnological advantages, or which are more amenable to genetic manipulation than the original myxobacterial producers.[38]

Investigating natural product biosynthesis on the molecular level also provides insights into the mechanisms by which monomeric biosynthetic building blocks are made and sequestered from bacterial primary metabolism to enable secondary metabolite biosynthesis. Deciphering these "channeling" processes, as well as understanding in detail the modifications that occur on biosynthetic intermediates prior to their incorporation into a natural product, constitutes an important prerequisite for improving secondary metabolite production in the original producer strain or in a heterologous host. Moreover, single enzymes from complex pathways are sometimes found to be valuable biochemical tools, due to the intriguing transformations which they are able to catalyze.

With the advent of the "genomic era", the above research objectives have been greatly facilitated by the increasing availability of whole-genome data for bacterial secondary metabolite producers. The genomes of the myxobacteria *Myxococcus xanthus* DK622 and *Sorangium cellulosum* So ce56 were recently sequenced, revealing the latter as the largest bacterial genome known to date.[39,40] The strikingly high abundance of typical secondary metabolite biosynthesis-related gene clusters in the genomes of these strains suggested that their genetic potential for the production of secondary metabolites exceeds by far the number of previously reported compound classes from each strain. While this finding may be considered as quite encouraging with regard to the prospects of discovering novel natural products from myxobacteria, it highlights at the same time the need for improved analytical methods for the detection, identification and characterization of myxobacterial secondary metabolites.

3. Outline of this work

The primary objective of this thesis is to study aspects of natural product formation in myxobacteria, with particular focus on the identification and characterization of enzymes belonging to the ammonia-lyase/aminomutase enzyme family and their roles as suppliers of building blocks for secondary metabolite biosynthesis. Two compound classes were investigated: the chondramides produced by *Chondromyces crocatus*, and the imidacins, previosuly unknown metabolites from *Stigmatella aurantiaca* (Figure A-3).

The chondramides **14** are highly cytotoxic depsipeptides that exhibit several unusual structural features, including a rare β-tyrosine moiety.[41] Bioinformatic analysis of the chondramide biosynthetic gene cluster from *C. crocatus* Cm c5 revealed a hybrid polyketide synthase/nonribosomal peptide synthetase (PKS/NRPS) biosynthetic route and also highlighted the gene *cmdF*, encoding a putative tyrosine 2,3-aminomutase (TAM).[42] CmdF exhibits high

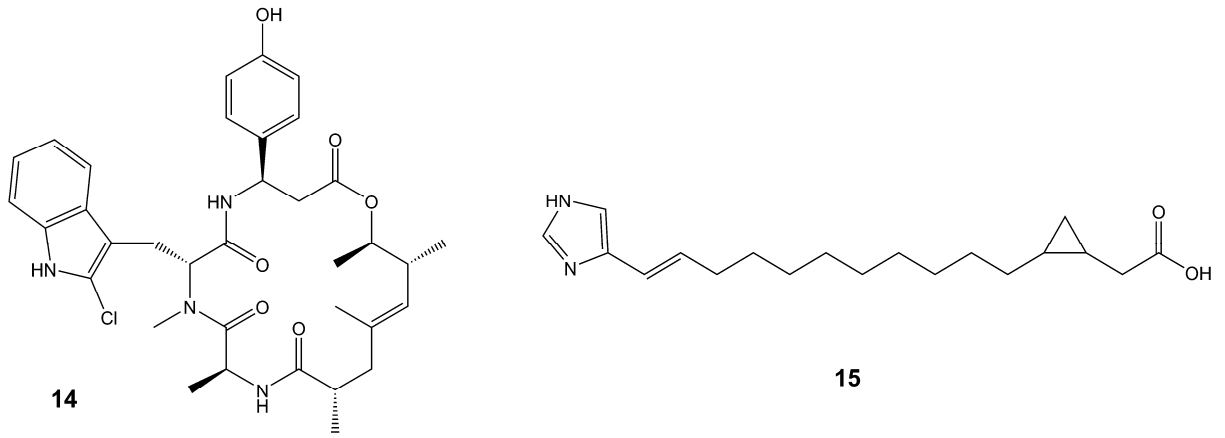


Figure A-3. Structure of the myxobacterial secondary metabolite chondramide C **14** and the structure of imidacin A_1 **15** proposed in this work.

similarity to the known tyrosine aminomutase SgcC4 from the enediyne biosynthetic pathway in *Streptomyces globisporus*.[43] However, it is shown here that the chondramides incorporate exclusively (*R*)-β-tyrosine, in contrast to the (*S*)-β-tyrosine-derived unit found in the enediynes. To explore the underlying biochemistry, CmdF was heterologously expressed in *E. coli* and biochemically characterized, revealing the protein as a novel tyrosine aminomutase with stereochemical preference for the production of (*R*)-β-tyrosine, consistent with its proposed function to provide the β-tyrosine building block during chondramide biosynthesis.[44] In order to establish a stronger basis for sequence comparison and phylogenetic analysis of tyrosine aminomutases, the hitherto small TAM enzyme family was further extended by the identification and biochemical characterization of candidate genes from *Myxococcus* and *Cupriavidus* species. The obtained data support a high degree of evolutionary diversification within the expanded TAM enzyme family, suggesting that family members may have been acquired by convergent evolution.[45] Aiming at identifying structural determinants of product specificity and stereoselectivity in TAM enzymes, mutational analysis of CmdF was carried out, yielding a mutant enzyme which produced an increased enantiomeric excess of (*R*)-β-tyrosine. This finding shows the potential to fine-tune the catalytic properties of tyrosine aminomutases by genetic engineering for their application as biocatalysts or in designed biosynthetic pathways.[45]

The imidacins **15** represent a novel compound class produced in minute amounts by *Stigmatella aurantiaca* Sg a15.[46] Structural elucidation of the imidacins revealed an unique fatty acid-like scaffold, consisting of an imidazole-acryl moiety linked via a linear alkyl chain of variable length to a carboxy group. In addition, the alkyl chain incorporates a rare cyclopropane moiety. The highly uncommon imidazol-acryl substructure could be shown to derive from

urocanic acid, representing a building block not previously reported for any natural product. Urocanate is supplied by the histidine ammonia-lyase enzyme HAL_{Sa}, which was demonstrated by targeted gene inactivation to be essential for imidacin biosynthesis. In addition, HAL_{Sa} was recombinantly produced in *E. coli* and biochemically characterized. The results of further gene inactivation experiments, in combination with feeding studies and the identification of additional candidate genes through screening of a transposon library, allowed to eventually identify the imidacin biosynthetic gene cluster and deduce a model for imidacin biosynthesis.[46] The data presented in this study reveal how a variety of enzymatic activities commonly associated with unrelated prokaryotic metabolic pathways cooperate in *S. aurantiaca* Sg a15 to produce the intriguing imidacin structure.

Another part of this thesis deals with the characterization of myxobacterial secondary metabolite profiles using liquid-chromatography-coupled high-resolution electrospray-ionization mass spectrometry (HR LC-ESI-MS). Based on the combined evaluation of mass accuracy, isotope pattern and retention time information, a highly sensitive method was devised for the targeted screening of complex natural product-containing extracts, allowing for compound identification with increased confidence. Furthermore, a statistical approach to explore high-resolution LC-MS datasets for the presence of compounds which give rise to significant differences in a set of samples was established. Principal component analysis (PCA) was employed for statistical interpretation; additionally, a new compound-finding algorithm was implemented for the preprocessing of raw LC-MS data. This data conditioning step was revealed as a crucial prerequisite for obtaining meaningful results from a PCA model.[47]

Motivated by the apparent discrepancy between the genetic potential of *Myxococcus xanthus* DK1622 for the production of secondary metabolites and the small number of natural products reported to date from this strain, the metabolite profiles of 98 *M. xanthus* strains isolated from locations worldwide were compared.[48] PCA was used to reveal distinct compounds exhibiting a non-ubiquitous distribution in the sample set. This screen revealed a strikingly high level of intraspecific diversity in the *M. xanthus* secondary metabolome. A significant number of potentially novel compounds was identified, suggesting that *M. xanthus* constitutes a viable source of as yet undiscovered natural products. Moreover, the obtained data help to correlate gene clusters of yet undefined biosynthetic function to secondary metabolites produced by these pathways. Due to its largely unbiased nature, the diversity-oriented approach employed in this work is also considered as useful for a wide range of metabolomics-based experiments, and shows promise to facilitate the efficient mining of the natural product inventory of various myxobacterial species in the future.[49]

4. The basic concepts of PKS- and NRPS-directed biosynthesis

Secondary metabolites from bacteria are usually the end products of complex multi-step biosynthetic pathways. Many bacterial natural products originate from biosynthetic machineries consisting of polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NPRSs), which act as enzymatic "assembly lines" by catalyzing the stepwise condensation of a starter unit with small monomeric building blocks. The building blocks and biosynthetic intermediates are tethered through thioester linkages to the terminal thiols of 4'-phosphopantetheine (4'-ppant) arms. These essential prosthetic groups are installed by phosphopantetheinyl transferase enzymes as covalent modifications onto conserved serine residues located in the carrier domains of PKS and NRPS proteins.[50] The biosynthetic logic of PKS and NRPS pathways thus follows a so-called "multiple carrier thiotemplate-model".[51] This section provides a brief overview of the mechanisms that form the basis of PKS- and NRPS-directed biosynthesis.

PKS biosynthetic systems are commonly classified with regard to the organization of their enzymatic functionalities and the basic structures of the compounds which they typically produce. PKSs of one particular architecture, termed "type I", occur widely in bacterial secondary metabolism and exhibit a multimodular organization where each module performs one cycle of chain extension (Figure A-4 **A**).[51,52] The minimum module comprises three functions: the acyltransferase (AT) domain selects an extender unit, usually the activated form of a short-chain dicarboxylic acid such as malonyl-CoA, and transfers it onto the 4'-ppant arm of an acyl carrier protein (ACP) domain within the same module. A ketosynthase domain (KS) subsequently decarboxylates the extender unit and performs a Claisen-like condensation reaction with the growing polyketide chain bound to the multienzyme. Additional domains are often integrated into a module and contribute to the structural diversity of PKS-derived metabolites: a ketoreductase (KR) domain reduces the β-keto group to an alcohol, a dehydratase domain (DH) then catalyzes the elimination of water to form a double bond, and finally an enoylreductase (ER) domain may reduce this double bond, resulting in a methylene moiety (Figure A-4 **A**).

Similar to the modular PKSs, NRPSs carry out the repeated condensation of amino acid monomers to yield non-ribosomally synthesized peptide products (Figure A-4 **B**).[51] The individual amino acid building blocks are selected and activated as aminoacyl adenylates by an adenylation (A) domain prior to their loading onto the integral peptidyl carrier protein (PCP) domain of the same module. Subsequently, amide bond formation with the growing peptide chain, tethered to the PCP of the preceeding module, is catalyzed by a condensation (C) domain.[53]

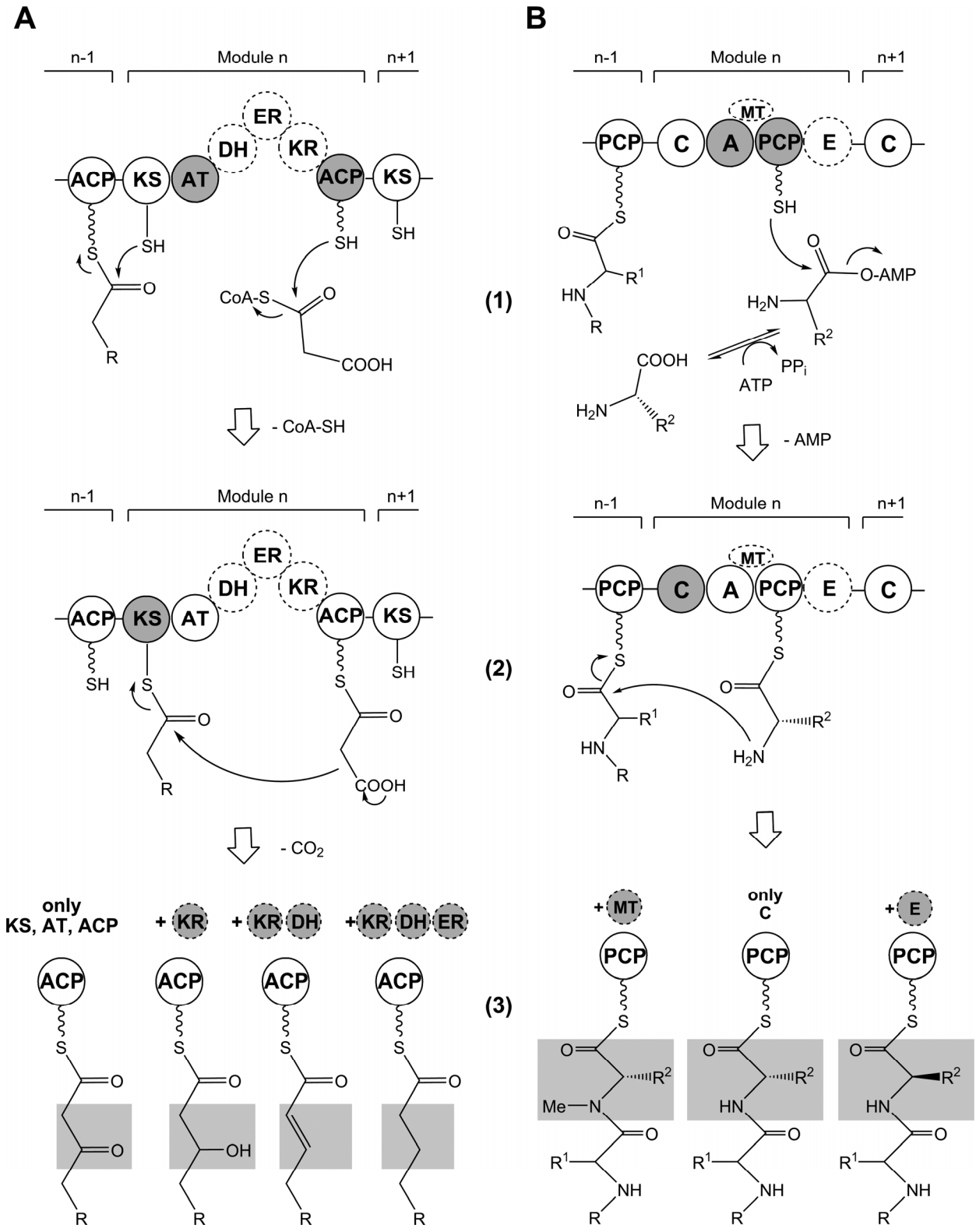


Figure A-4. The basic reactions catalyzed by PKS **(A)** and NRPS **(B)** biosynthetic systems. **(1)**, loading of monomeric building blocks; **(2)**, chain extension by claisen-like condensation; **(3)**, further modification of intermediates on the "assembly line" (grey boxes). The domains that direct each step are shown as filled circles, domains symbolized by dashed circles are optional. **ACP**, acyl carrier protein; **PCP**, peptidyl carrier protein; **KS**, ketosynthase; **AT**, acyl transferase; **KR**, ketoreductase; **DH**, dehydratase; **ER**, enoylreductase; **C**, condensation domain; **A**, adenylation domain; **MT**, methyltransferase domain; **E**, epimerase domain.

The PCP-bound peptidyl chain may be further modified, for example by *N*-methylation – usually performed by an optional methyltransferase (MT) domain integrated into the adenylation domain - or by inversion of the amino acid C_α stereocenter, carried out by an additional epimerase (E) domain (Figure A-4 **B**). Following the incorporation of cysteine, serine or threonine, the side chains of these amino acids may be intramolecularly cyclized and dehydrated to form thiazoline and oxazoline rings, respectively, by the action of heterocyclization domains.[51]

In contrast to the ribosomal synthesis of proteins, NRPS-directed biosynthesis is not restricted to the exclusive use of the standard set of 20 proteinogenic amino acids. Indeed, a sizeable number of A-domains have been characterized which activate and incorporate modified or highly unusual amino acids, or even non-aminated short-chain acids or aromatic carboxylic acids.[51] Furthermore, additional tailoring enzymes can modify individual building blocks or the nascent peptidyl chains during NRPS biosynthesis, while these are covalently attached to PCP domains. Examples include halogenation and the cross-linking of amino acid side chains.[54,55]

The products of PKS and NRPS biosynthetic pathways are most often released from the assembly line through the action of a thioesterase (TE) domain located downstream of the last module in the respective pathway. TE-catalyzed hydrolysis of the enzyme-bound thioester intermediate yields the free acid; alternatively, intramolecular nucleophilic attack onto the TE-bound acyl chain leads to macrocyclization.[55] The PKS- or NRPS-derived core structures may undergo further post-assembly line modifications by the action of dedicated enzymes, with the most frequently observed transformations being the introduction or removal of double bonds, hydroxylations, glycosylations, acylations, *O*-, *N*- or *C*-methylations, along with rather unusual modifications such as epoxidation.[51]

In textbook bacterial PKS and NRPS biosynthetic pathways, each module in the enzymatic assembly line is used only once, and the order of modules and arrangement of domains contained in these modules therefore correspond to the chemical structure of the polyketide or non-ribosomal peptide product – an observation that has been termed the co-linearity rule.[56] Bioinformatic analysis of many PKSs and NRPSs, in combination with the results of structural and biochemical studies, has enabled the development of guidelines for predicting the specificity of domains responsible for the loading of monomeric building blocks.[53,57,58] Thus, it is (to a certain extent) possible to devise a hypothetical structure for a compound produced by a PKS or NRPS pathway through *in-silico* analysis of the genes encoding the underlying biosynthetic enzymes. If the product structure is known, the deduced biosynthetic functions may be correlated to structural features and a biosynthetic model can be developed. However, in recent years the

molecular basis for the formation of many myxobacterial natural products has been investigated in detail, and these studies have frequently revealed unexpected genetic organizations of biosynthetic pathways and surprising biosynthetic features that challenge traditional textbook biosynthetic logic, such as the iterative use or the complete skipping of modules and the incorporation of unusual building blocks and starter units.[56,59]

The chondramides – products of a mixed PKS/NRPS pathway with unsual features.

The biosynthesis of the chondramides (Figure A-5; see also chapter B-3) provides an illustrative example for the aforementioned concepts of PKS- and NRPS-directed biosynthesis. The chondramides are highly cytotoxic depsipeptides which are biosynthesized via a hybrid PKS/NRPS route in the myxobacterium *Chondromyces crocatus*.[41,42,60] In fact, a significant number of myxobacterial secondary metabolites results from such "mixed" PKS and NRPS pathways.[59] The first two multienzymes encoded by the genes *cmdA* and *cmdB* in the chondra-

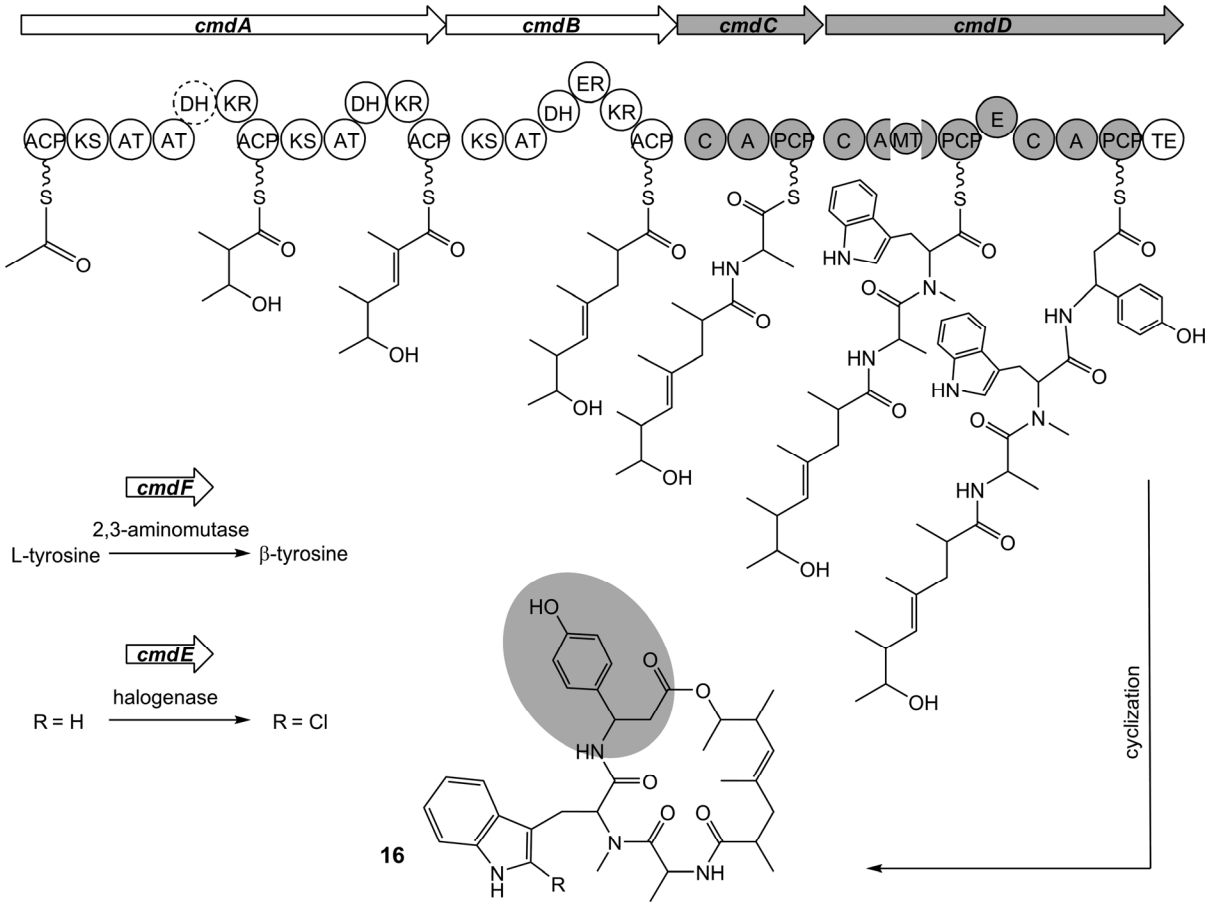


Figure A-5. Schematic representation of biosynthetic functions encoded by the hybrid PKS/NRPS chondramide gene cluster, and proposed assembly-line intermediates during biosynthesis of chondramides C (R = H) and D (R = Cl) **16**.

mide biosynthetic gene cluster contain typical PKS modules, which elongate an ACP-bound acetyl starter unit with three methylmalonyl-CoA extender units. The first extension step is followed by ketoreduction to yield a hydroxy group, in agreement with the presence of a KR domain and an inactive DH domain, while KR-catalyzed reduction and subsequent dehydration by a functional DH domain take place after the second chain extension cycle, resulting in the formation of a double bond between two extender units. Following incorporation of the third propionate unit, complete reduction leads to a methylene moiety – as anticipated with regard to the full complement of KR, DH and ER domains found in the last PKS module. Biosynthesis proceeds via the incorporation of alanine by the single-module NRPS CmdC, followed by condensation of tryptophane to the peptidyl chain, catalyzed by the first module of CmdD. The A-domain in this module contains an integrated MT domain, which likely accounts for *N*-methylation at the appropriate position. The presence of an additional epimerase domain suggests that the tryptophane building block in the final chondramide structure adopts the "unnatural" D-configuration. Furthermore, some chondramide derivatives exhibit a rare 2-chlorotryptophane substructure, where the chlorine is introduced by the dedicated halogenase CmdE, which is also encoded within the chondramide biosynthetic gene cluster.[42]

Another striking structural feature of the chondramides is the (*R*)-β-tyrosine moiety, which is incorporated by the second NRPS module of CmdD (Figure A-5). The biosynthetic origin of this unusual building block and its mode of incorporation into the chondramide scaffold are the subject of chapter B-1,[44] revealing the role of the tyrosine 2,3-aminomutase CmdF in the process. CmdF exhibits unprecedented stereochemical preference for the production of (*R*)-β-tyrosine, which is consistent with its function to provide the β-tyrosine building block for chondramide biosynthesis (see also the following chapter A-5 and chapter B-2).

The chondramide biosynthetic pathway from *Chondromyces crocatus* shows a high level of colinearity between genes in the cluster and the predicted biochemical steps required for chondramide biosynthesis.[42] In contrast, the biosynthesis of the imidacins by the myxobacterium *Stigmatella aurantiaca,* which is described in chapter B-4, only follows the rules of textbook biosynthetic logic to a low extent. The unique urocanate building block for imidacin biosynthesis is also supplied by a member of the ammonia-lyase enzyme family. Therefore, the following chapter provides an overview of the general characteristics and functions of ammonia-lyase and aminomutase enzymes.

5. Aromatic amino acid ammonia-lyase and aminomutase enzymes as suppliers of building blocks for natural product biosynthesis

Enzymes of the aromatic amino acid ammonia-lyase family catalyze the non-oxidative deamination of L-histidine, L-tyrosine and L-phenylalanine to yield the corresponding α,β-unsaturated acids and ammonia. The related aminomutase enyzmes catalyze the intramolecular 2,3-migration of the amino group (Figure A-6).[61] Histidine ammonia-lyase (HAL), tyrosine ammonia-lyase (TAL) and phenylalanine ammonia-lyase (PAL) play diverse roles in the metabolism of plants, microorganisms and animals, while aminomutase activity is markedly rare.[61]

PAL activity is found in most plants and many fungi, where the production of (*E*)-cinnamate constitutes the first step of phenylalanine degradation. Cinnamic acid **23**, in turn, is the precursor of many plant metabolites, including the flavonoids, the coumarins and lignin, the major constituent of wood. The PAL enzymes of certain plants are also able to convert tyrosine to coumaric acid due to their secondary TAL activity.[61] A phenylalanine aminomutase (PAM) has been shown to supply (*R*)-β-phenylalanine as a precursor for the biosynthesis of the important anti-cancer drug Taxol in *Taxus* species.[62]

HAL is present in prokaryotes and most higher animals. In bacteria, the HAL-catalyzed conversion of L-histidine **17** to urocanate **18** constitutes the first step of the general "histidine utilization" pathway (Figure A-7). Urocanate is further processed to imidazolonepropionate **25** by uro-

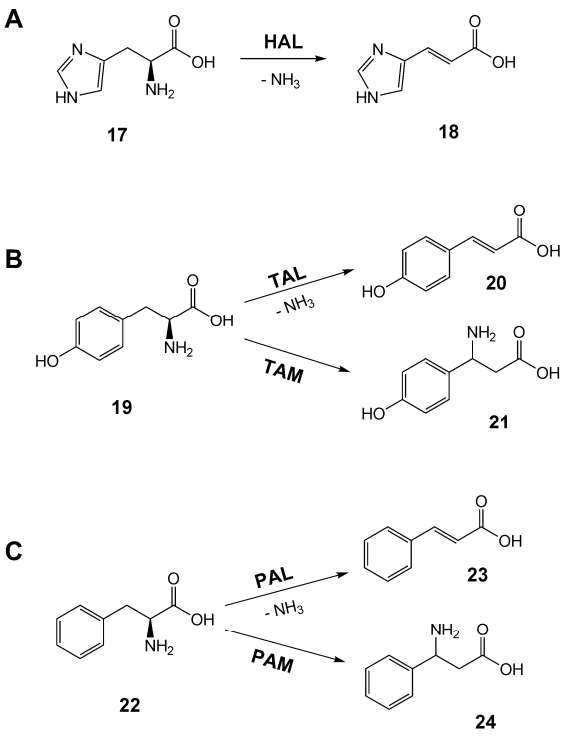


Figure A-6. Reactions catalyzed by members of the aromatic amino acid ammonia-lyase and aminomutase enzyme family. **(A)**, HAL converts L-histidine **17** to urocanic acid **18**; **(B)**, TAL converts L-tyrosine **19** to coumaric acid **20**, while TAM yields β-tyrosine **21**; **(C)**, PAL converts L-phenylalanine **22** to cinnamic acid **23**, and PAM produces β-phenylalanine **24**.

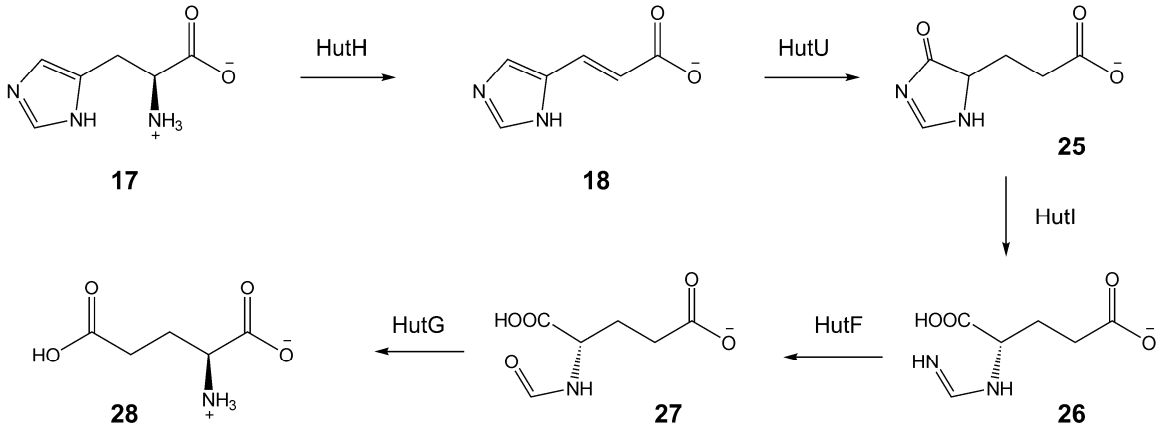


Figure A-7. Degradation of L-histidine **17** to L-glutamate **28** via the "histidine utiliziation" pathway in the bacterium *Pseudomonas putida*. The involved enzymes are: HutH, histidine ammonia-lyase (HAL); HutU, urocanase; HutI, imidazolonepropionate hydrolase; HutF, formiminoglutamate hydrolase; HutG, *N*-formylglutamate aminohydrolase.

canase, followed by enzymatic hydrolysis to formiminoglutamate **26**. Subsequent hydrolysis yields formylglutamate **27** and finally glutamate **28** (however, these two steps are combined into the elimination of formamide in some species). Genes encoding the required enzymes for this pathway are organized in the *hut* (histidine utilization) operon in many bacteria.[61]

In contrast to histidine ammonia-lyase, PAL- and TAL-type enzymes are rare in bacteria – an observation that apparently reflects the general absence of phenylpropanoid-type metabolites in prokaryotes. The first bacterial PAL enzyme, EncP, was reported from *Streptomyces maritimus*, where it is involved in secondary metabolite biosynthesis (Figure A-8 **A**): Cinnamic acid **23** produced by EncP is further transformed into cinnamoyl-CoA **29** by the CoA-Ligase EncH and is then degraded via a β-oxidative route to yield benzoyl-CoA **30**, which serves as starter unit for the biosynthesis of enterocin **31**.[63,64] The involvement of another PAL in the biosynthesis of isopropylstilbenes in *Photorhabdus luminescens* was recently demonstrated, and the corresponding enzyme, StlA, was biochemically characterized.[65] Very recently, genomic mining has revealed the presence of putative PAL-encoding genes in the cyanobacteria *Nostoc punctiforme* and *Anabaena variabilis*, and the proposed activity of these enzymes has been biochemically confirmed.[66] The PAL-encoding genes in these species seem to be associated with hypothetical secondary metabolite biosynthetic gene clusters, but the role of cinnamic acid in the respective pathways is to date not clear. Phenylalanine aminomutase activity in bacteria appears to be as rare as in plants: the only report so far is related to the production of (*S*)-β-Phe by the enzyme AdmH during biosynthesis of the antibiotic Andrimid in *Pantoea agglomerans*.[67]

Enzymes of the ammonia-lyase family which use primarily (or exclusively) L-tyrosine as substrate are exceedingly rare in nature. Until recently, dedicated TAL enzymes were only known from *Rhodobacter sphaeroides* and *Rhodobacter capsulatus*.[68,69] In these species, *p*-hydroxcinnamic acid (pHCA) **20** serves as the chromophoric group for the so-called "photo-active yellow protein" (PYP). The covalent attachment of pHCA constitutes a post-translational modification which is essential for the function of these light sensor proteins. The kinetic characterization of the two TALs from *Rhodobacter* has revealed that these enzymes additionally exhibit significant PAL activity.[68,69]

Studies on the molecular basis of enediyne antibiotics formation in *Streptomyces globisporus* have highlighted the L-tyrosine 2,3-aminomutase (TAM) SgcC4 as a novel member of the ammonia-lyase/aminomutase enzyme family (Figure A-8 **B**). SgcC4 converts L-tyrosine **19** prefer-

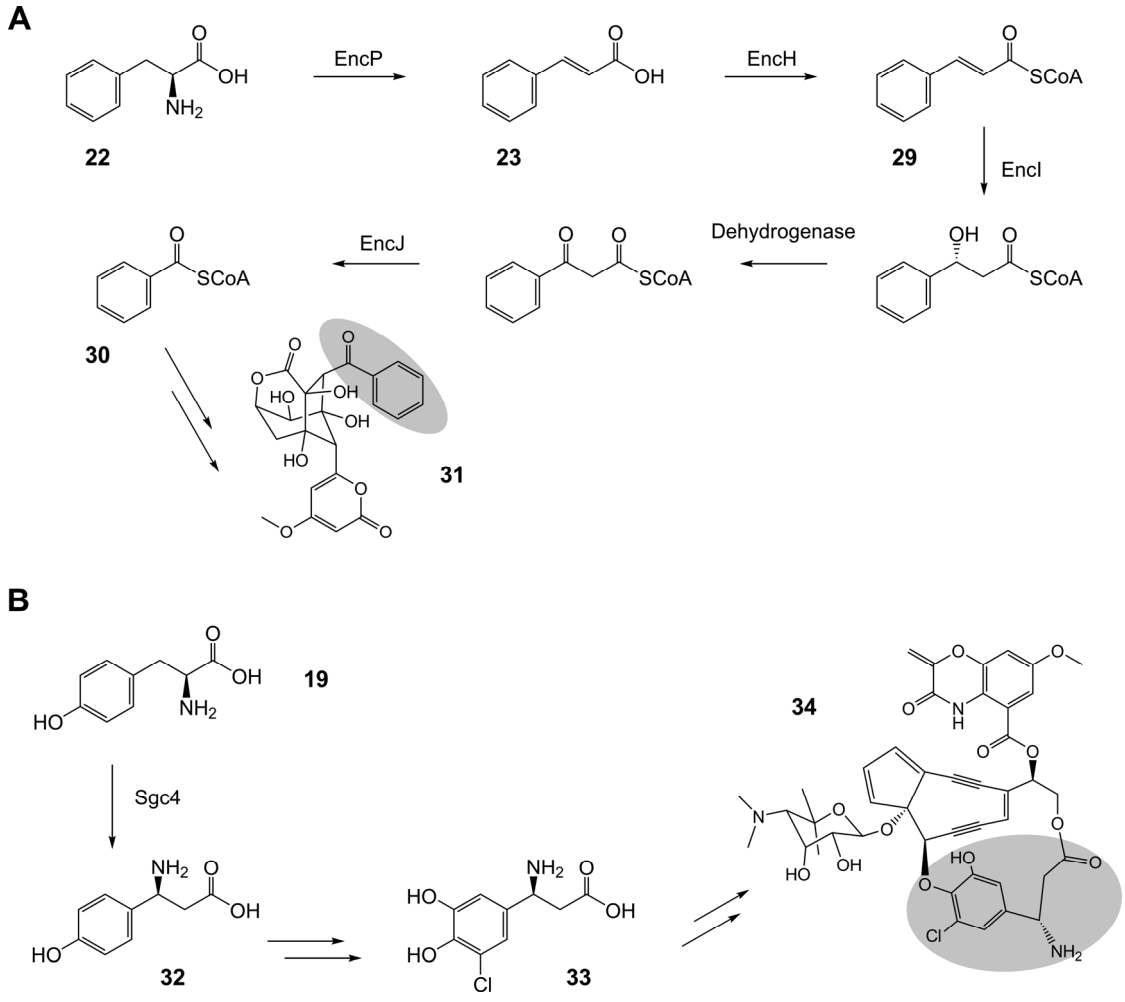


Figure A-8. (A) The PAL enzyme EncP converts L-Phe **22** to cinnamic acid **23**, which is further degraded to benzoyl-CoA **30**, a precursor for the biosynthesis of enterocin **31** by *Streptomyces maritimus*. **(B)** The tyrosine aminomutase SgcC4 from *Streptomyces globisporus* produces preferentially (*S*)-β-tyrosine **32**, which undergoes further transformation into 3-Cl-4,5-$(OH)_2$-β-Phe **33**, a biosynthetic building block for the endiyne antitumor antibiotic C-1027 **34**. The ammonia-lyase and aminomutase-derived moieties, respectively, are highlighted by grey background color.

entially to (*S*)-β-tyrosine **33**, which is further modified by hydroxylating and chlorinating enzymes during enediyne biosynthesis.[43,70,71] A highly similar TAM enzyme, MdpC4, has recently been shown to exert the analogous function to SgcC4 in the course of madurapeptin biosynthesis in the marine actinomycete *Actinomadura madurae*.[72]

Taken together, the following picture evolves from the literature on the occurence and function of ammonia-lyase/aminomutase enzymes in nature: PAL represents an ubiquitous enzyme in plants, where it constitutes an important branching point between primary and secondary metabolism. PALs are seldomly encountered in prokaryotes, but have been shown to supply cinnamic acid precursors for a small number of secondary metabolite biosynthesis-related pathways. As an exceptionally rare enzyme, the singular role of TAL appears to be the provision of coumaric acid as a chromophore for the photoactive yellow proteins in *Rhodobacter* species. HAL, in contrast, occurs widely in bacteria and has an established function for histidine utilization, catalyzing the first step of the general histidine degradation pathway. The rare aminomutase enzymes PAM and TAM have been found associated with natural product biosynthetic routes in all instances reported to date, where they fulfill specialized roles as suppliers of β-amino acid building blocks.

Discovery and characterization of new prokaryotic ammonia-lyase/aminomutase enzymes and elucidation of their roles in secondary metabolite biosynthesis.

In the course of this thesis, a number of ammonia-lyase enzymes were characterized which exhibit uncommon catalytic properties, and/or play an unusual role in the metabolism of the respective organism. These findings are summarized in the following.

A) The tyrosine 2,3-aminoumutase CmdF from the myxobacterium *C. crocatus* was shown to produce preferentially (*R*)-β-tyrosine, in contrast to the previously known tyrosine aminomutases, SgcC4 and MdpC4 from the enediyne biosynthetic pathways in *S. globisporus* und *A. madurae*, respectively (which predominantly yield (*S*)-β-tyrosine). Thus, CmdF is a novel member of the TAM enzyme family. Its product stereoselectivity matches the configuration of the β-tyrosine building block incorporated into the chondramides (Figure A-9**A**).[44] The detailed biochemical characterization and mutational analysis of CmdF as well as attempts to alter its catalytic properties by genetic engineering are reported in chapter B-2.[45] These studies were extended towards the discovery and characterization of additional members of

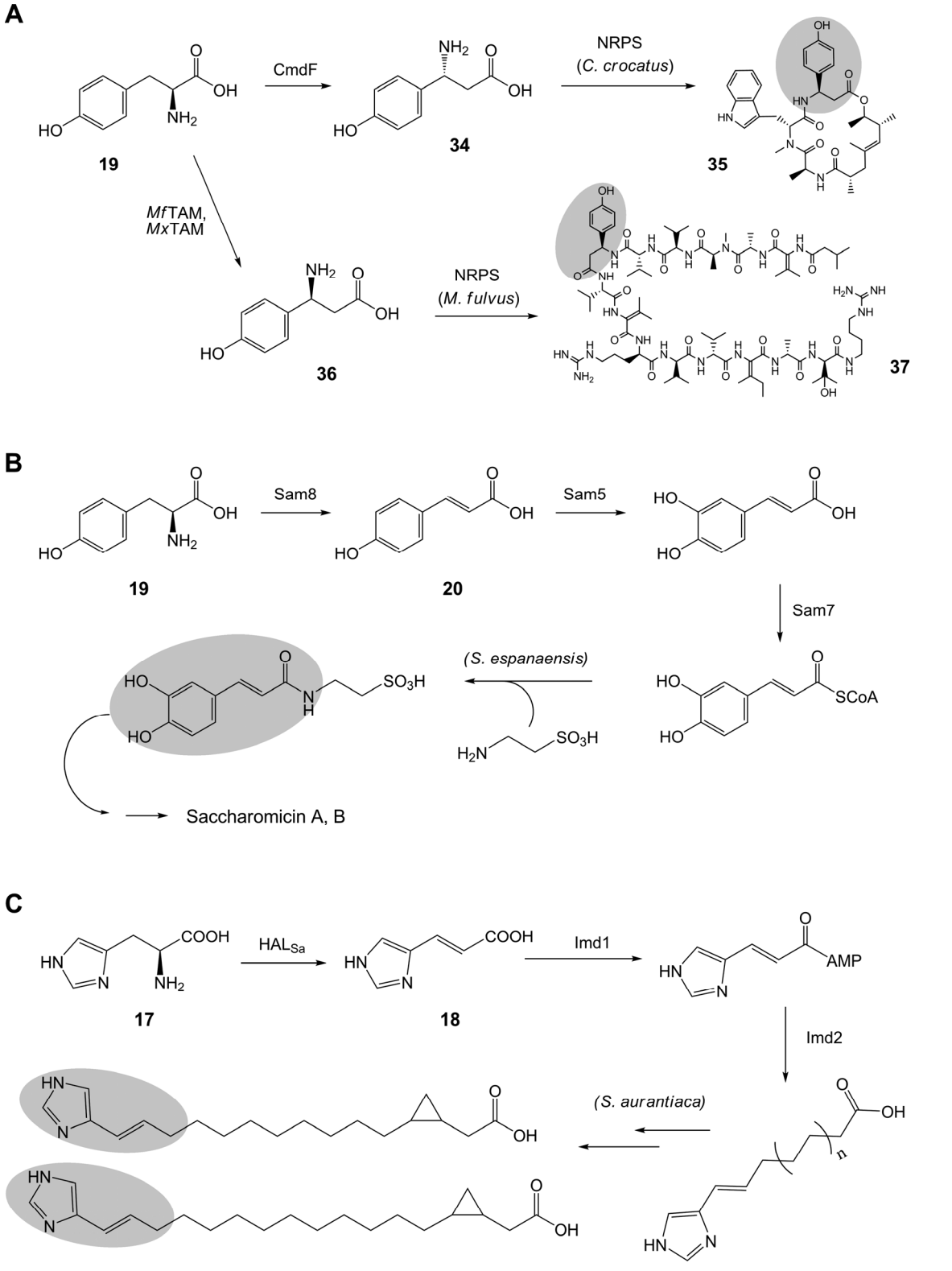


Figure A-9. Novel members of the ammonia-lyase/aminomutase enzyme family which were characterized in this work, and their roles as suppliers of building blocks (grey background) for natural product biosynthesis. **(A)** TAMs from *C. crocatus* (CmdF) and *Myxococcus* (*Mf*TAM, *Mx*TAM), providing (*R*)-β-Tyr **34** for chondramide **35** biosynthesis and (*S*)-β-Tyr **36** for myxovalargin **37** biosynthesis, respectively. **(B)** Tyrosine ammonia-lyase from *Saccharothrix espanaensis* (Sam8) and its role in Saccharomicin biosynthesis. **(C)** Histidine ammonia-lyase from *Stigmatella aurantiaca* (HAL_{Sa}) provides the urocanate starter unit **18** for the biosynthesis of imidacins.

the emerging tyrosine aminomutase family from other bacterial sources. Genes encoding enzymes with TAM activity were cloned from two myxovalargin-producing *Myxococcus* species and a strain of *Cupriavidus metallidurans*. The *Myxococcus* enzymes were heterologously expressed and could be shown to exhibit (*S*)-β-tyrosine stereoselectivity, consistent with their proposed function in supplying a (*S*)-β-tyrosine building block for myxovalargin biosynthesis (Figure A-9 **A**). TAM from *Cupriavidus* produced (*R*)-β-tyrosine, but the kinetic enzyme analysis revealed that this protein exhibits primarily TAL activity. Including the newly characterized proteins, the expanded tyrosine aminomutase enzyme family now comprises six representatives which differ markedly with respect to product stereospecificity and relative preference for aminomutase to ammonia-lyase activity. The increased availability of TAM sequences enabled – for the first time – a phylogenetic analysis of tyrosine aminomutases and led to the proposal of a new model for the evolution of this enzyme family (chapter B-2).

B) A putative ammonia-lyase from the actinomycete *Saccharothrix espanaensis,* Sam8, was also recombinantly produced and biochemically characterized (chapter B-5).[73] Sam8 clearly preferred L-tyrosine over L-phenylalanine as substrate, and thus constitutes one of the few authentic TALs from bacteria. Notably, the characterization of Sam8 represents the first report of a TAL from an actinomycete. Moreover, the suggested involvement of Sam8 in secondary metabolite biosynthesis in *S. espanaensis* is unprecedented in bacteria (Figure A-9 **B**). Coumaric acid, the product of Sam8-catalyzed tyrosine conversion, is apparently ring-hydroxylated to yield caffeic acid, which in turn is activated as caffeoyl-CoA prior to incorporation into the *N*-(dihydroxycinnamoyl) taurine aglycone of the saccharomicin antibiotics.

C) Targeted inactivation of a gene encoding a putative histidine ammonia-lyase in *Stigmatella aurantiaca* Sg a15 (HAL_{Sa}) led to the discovery of the imidacins, a novel class of myxobacterial secondary metabolites. Their biosynthesis depends on the supply of urocanic acid by HAL_{Sa} (Figure A-9 **C**), revealing for the first time a non-standard role for histidine ammonia-lyase in bacteria – namely, the involvement of HAL_{Sa} in myxobacterial secondary metabolism. This enzyme was heterologously expressed in *E. coli* and biochemically characterized (chapter B-4), and the pathway for the incorporation of urocanate into the newly identified metabolites was additionally delineated.[46] The incorporation of an urocanate building block has previously not been reported from any natural product.

On the structural and mechanistic basis for ammonia-lyase and aminomutase catalysis – implications for control of specificity and selectivity.

A number of biochemical, molecular biological and structural studies have addressed questions related to the catalytic properties of aromatic amino acid ammonia-lyase and aminomutase enzymes. The activity of this enzyme family has been shown to rely on the uncommon prosthetic group 4-methylidene-imidazole-5-one (MIO), which results from the autocatalytic rearrangement of a signature Ala-Ser-Gly motif present in the protein (Figure A-10 **A**).[74] The exact role of MIO in ammonia-lyase and aminomutase chemistry has been controversially debated. Results from a study on histidine ammonia-lyase from *Pseudomonas putida* suggested the attack of the highly electrophilic MIO on the substrate imidazole ring, leading to activation of the non-acidic β-protons. Following the abstraction of a β-proton by an enzymatic base, the reaction would proceed via the elimination of ammonia, rearomatization of the ring system to form the product (*E*)-urocanate, and regeneration of the MIO group (Figure A-10 **B**).[75] It has been pro-

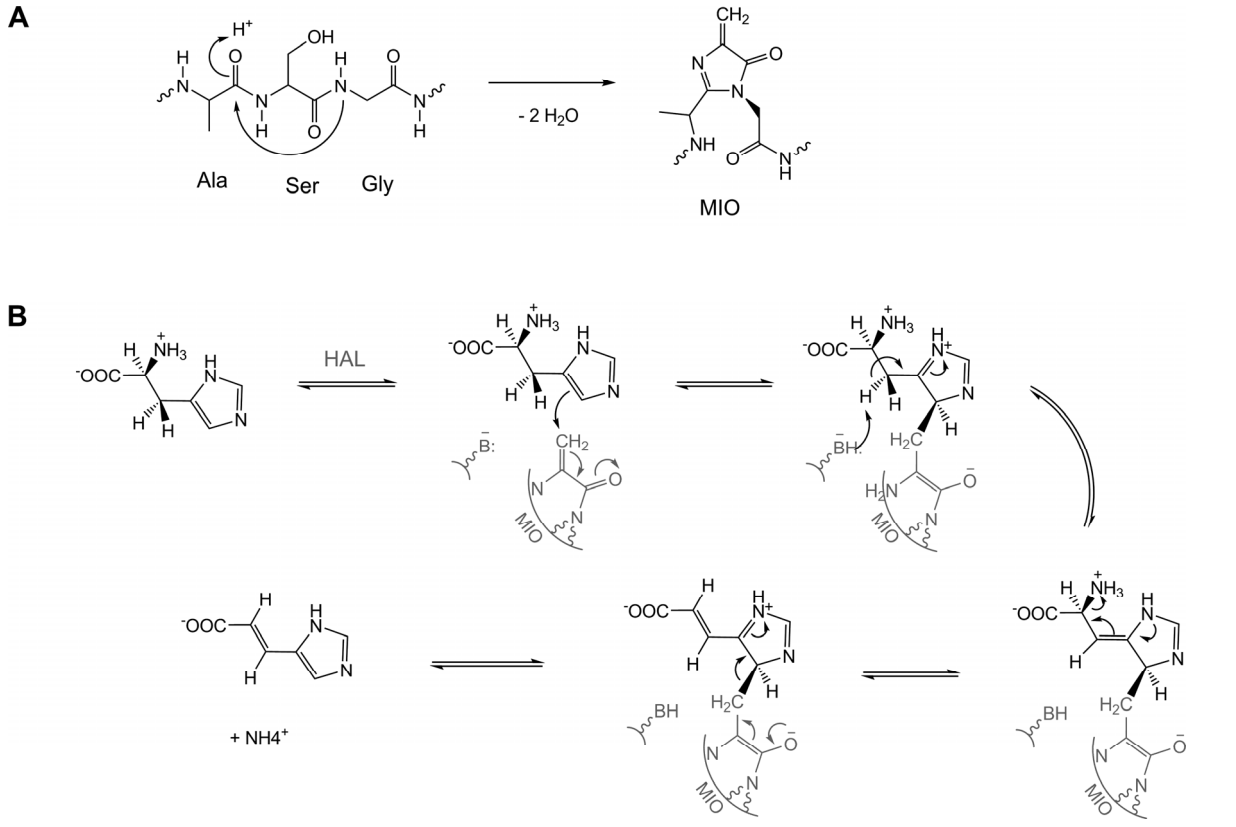


Figure A-10. **(A)** Self-modification of ammonia-lyase enzymes by condensation of residues from the internal tripeptidye Ala-Ser-Gly to form the MIO group. **(B)** Proposed catalytic mechanism for the conversion of L-histidine to urocanate by HAL through a Friedel-Crafts-like attack of the electrophile MIO onto the substrate imidazole ring, followed by abstraction of a β-proton and elimination of ammonia.

posed that this Friedel-Crafts-like mechanism could similiarly apply to PAL, TAL and the aminomutases.[61]

An alternative mechanistic route requires the attack of MIO on the substrate amine, generating a covalent adduct that promotes the α,β-elimination of ammonia from the amino acid. This mechanism is illustrated in Figure A-11 **A** for the TAM-catalyzed conversion of L-tyrosine to β-tyrosine, while Figure A-11 **B** depicts the Friedel-Crafts-type proposal for the same reaction.[76] Data from a recent crystallographic study on the tyrosine aminomutase SgcC4 from *S. globisporus* support the amino-adduct mechanism, because electron density continuous from the MIO group to the amino group of an substrate-analogous inhibitor has been detected in the active site of SgcC4 (Figure A-11 **C**).[77] Either way, the aminomutase reaction is thought to proceed via a *para*-hydroxycinnamate (pHCA) intermediate. Consistent with this assumption, it has been observed that SgcC4 exhibits significant TAL activity as a side reaction.[78] In the ammonia-lyase reaction pathway, the pHCA intermediate and ammonia are released, while retention of ammonia in the active site of aminomutases facilitates a subsequent 1,4-addition between the enzyme-

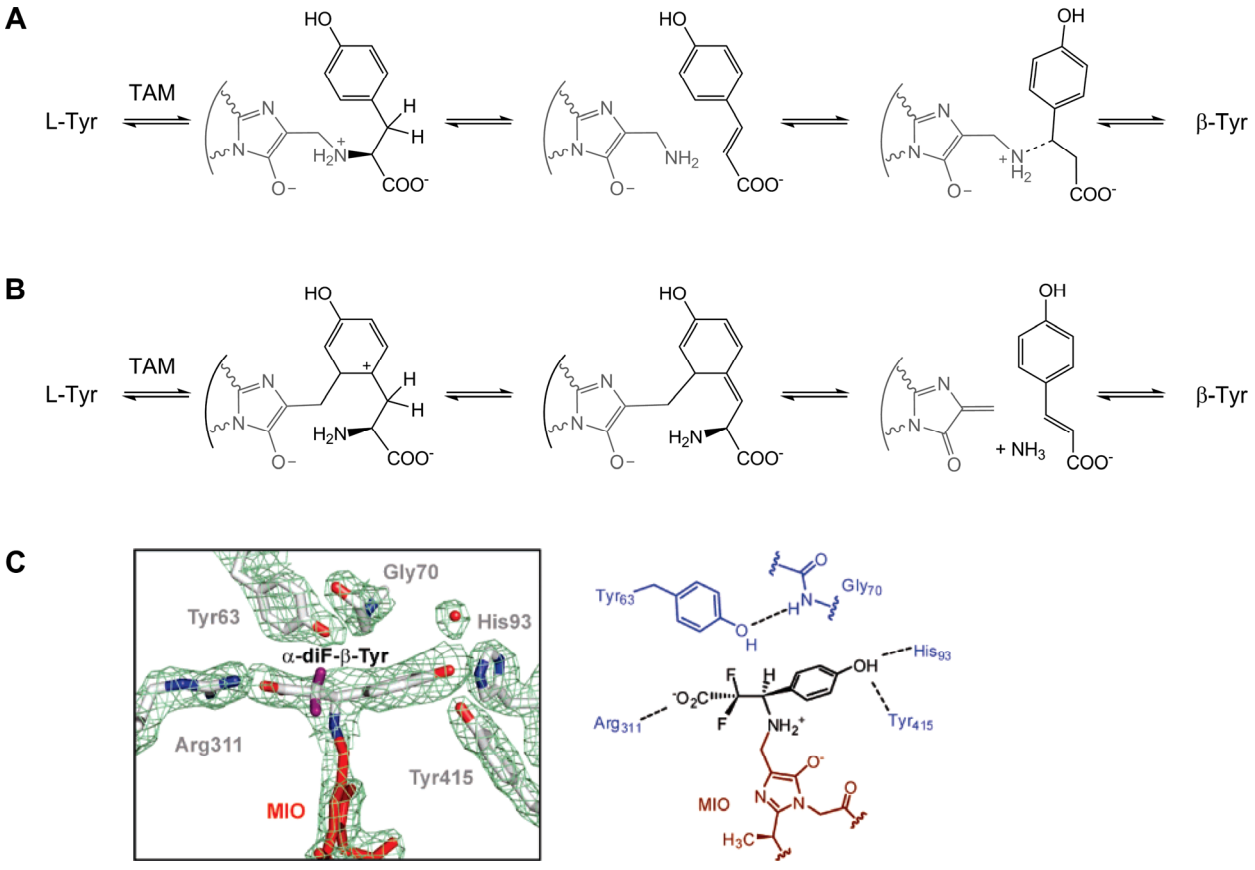


Figure A-11. **(A, B)** Two alternative mechanistic proposals for MIO-based tyrosine aminomutases. **(C)** Structure of SgcC4 bound to the substrate analogue α,α-difluoro-β-tyrosine (left), and interactions between enzyme and bound substrate analogue in the active-site region (right). Red, MIO group; black, substrate analogue; blue, amino acid residues from SgcC4. Adapted from [77] with permission.

bound intermediates, affording the β-amino acid as the end product. The orientation of the pHCA intermediate in the active-site pocket determines the facial selectivity of ammonia addition, and thus in principle, both enantiomers of the β-amino acid product may result.[77] Notably, the known tyrosine aminomutases SgcC4 and MdpC4 from the enediyne biosynthetic pathways have been demonstrated to produce preferentially (*S*)-β-tyrosine, while the newly reported CmdF from *C. crocatus* primarily yields (*R*)-β-tyrosine.[44] Moreover, two additional tyrosine aminomutases from *Myxococcus* species exhibit (*S*)-β-tyrosine stereoselectivity, but a new enzyme from *Cupriavidus* affords preferentially (*R*)-β-tyrosine.[45] These findings raise questions concerning the evolution of stereocontrol in the TAM enzyme family.

The availability of crystallographic data for HAL, PAL, TAL and TAM proteins has helped to shed light on the correlation between structural features and catalytic properties of these enzymes. Based on the combined results from a structural study and site-directed mutagenesis experiments, amino acid residues in the active site of HAL from *P. putida* with an implication in substrate binding and catalysis have been identified.[75] One highly conserved glutamic acid residue, Glu414, has been shown to be useful as a predicitve tool, allowing to distinguish HAL from both TAL and PAL through the *in-silico* analysis of a candidate protein sequence.[79] Very recently, key amino acids which control substrate selectivity have also been reported for TAL from *Rhodobacter*.[80] In particular, one conserved histidine residue, His89, has been found to direct substrate recognition. When this position in the amino acid sequence was changed to phenylalanine, the enzyme exhibited efficient PAL activity, identifying His89 as a selectivity switch.[66]

It has been suggested that these substrate selectivity principles could similarly apply to the aminomutases - a notion that is apparently supported by the recent crystal structure of the tyrosine aminomtase SgcC4.[77] However, it has to be pointed out that enzymes of the TAM family have not been extensively studied by molecular biological approaches prior to this work. Thus, the determinants of product specificity (i.e., what makes the enzymes aminomutases versus ammonia-lyases) and stereoselectivity remain to be elucidated. Such information constitutes an important prerequisite for future attempts to manipulate rationally the catalytic properties of aminomutases for various biotechnological applications. Therefore, the heterologous expression of new tyrosine aminomutases was carried out to enable their biochemical characterization. Furthermore, site-directed mutagenesis of the tyrosine aminomutase CmdF from *C. crocatus* was performed, aiming at the exploration of structure-function relationships for conserved sequence motifs found in the tyrosine aminomutase enzymes (chapter B-2).

6. A diversity-oriented approach for the analysis of myxobacterial secondary metabolite production patterns by high-resolution mass spectrometry

The myxobacteria as a group have been demonstrated to produce a wide variety of secondary metabolites, but the degree of metabolic diversity that can be found within a single species has not been systematically investigated. Strains belonging to certain myxobacterial species, e.g. *Sorangium cellulosum*, seem to exhibit low intraspecific overlap in their secondary metabolite profiles and yield at the same time a high number of natural products not known from other (myxobacterial) genera.[18] In contrast, strains of the species *Myxococcus xanthus* have been shown to date to produce only eight compound classes – five of which have been reported from other myxobacteria as well (Figure A-12, **38-42**),[81-85] while the three remaining structures are highly similar or even identical to compounds previously identified from members of the *Actinomycetale*s (Figure A-12, **43-45**).[86-88] These findings are apparently not in agreement with the genetic potential of *M. xanthus* for the production of secondary metabolites. Genomic mining

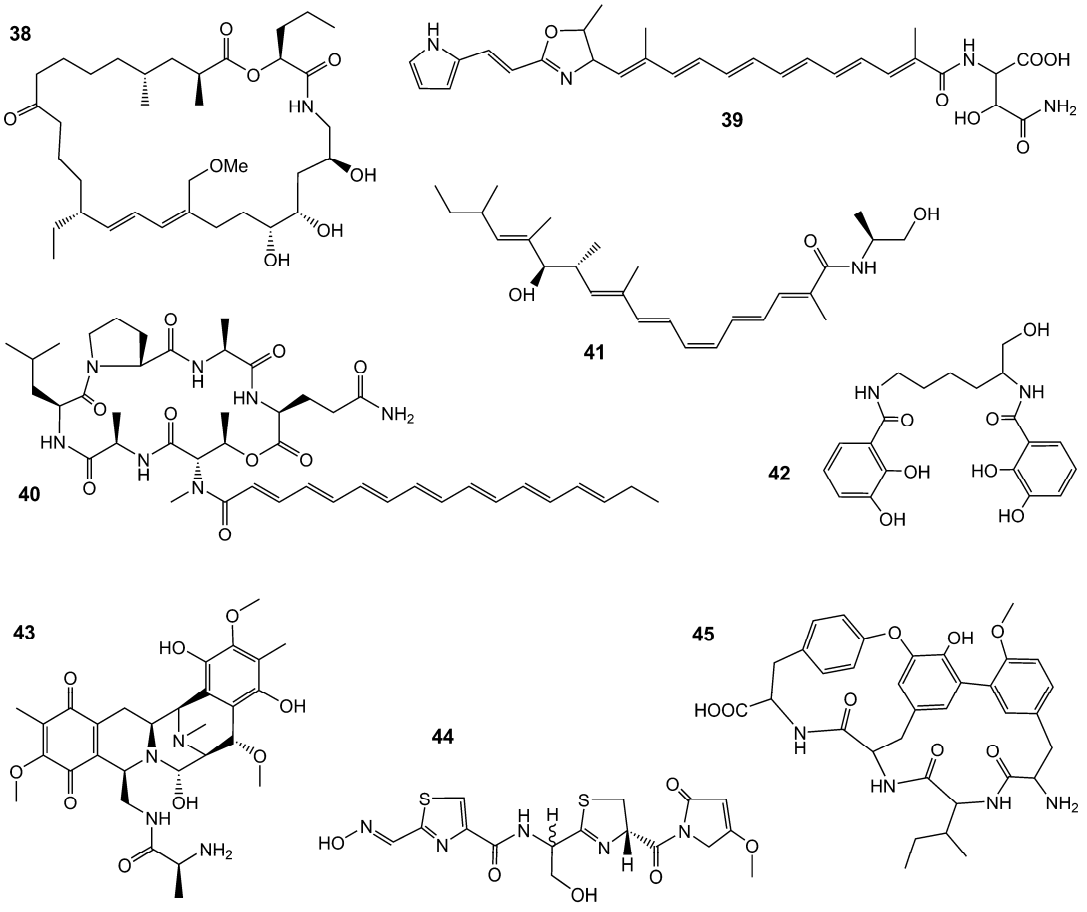


Figure A-12. Overview of natural products discovered to date from the species *Myxococcus xanthus*. **38**, myxovirescin; **39**, DKxanthene; **40**, myxochromide A; **41**, myxalamide; **42**, myxochelin; **43**, saframycin; **44**, althiomycin; **45**, cittilin.

has revealed at least 18 gene clusters encoding secondary metabolite assembly lines of the PKS-, NRPS- or mixed PKS/NRPS-type in the genome of the model strain *M. xanthus* DK1622, suggesting that this strain should produce a substantial number of secondary metabolites.[40] However, to date only five known *M. xanthus* metabolites have been identified in extracts from DK1622 cultivations (Figure A-12, **38-42**).

To evaluate further the secondary metabolite diversity in *M. xanthus*, a considerable number of *M. xanthus* strains isolated from locations worldwide was screened. This study aimed at determining whether the metabolic inventory of this species, as observable under laboratory conditions, is largely present in DK1622 or whether there is significant potential for novel compound discoveries in other strains. Liquid chromatography-coupled mass spectrometry (LC-MS) was considered the appropriate analytical technique for this purpose, as it can in principle capture a highly diverse range of chemical compounds and enables the detection of even minor constituents in a given sample. Thus, a screening platform consisting of fast chromatographic separation (UPLC), electrospray ioniaztion (ESI), and high-resolution (HR) time-of-flight mass spectrometry (ToF-MS) was implemented and is briefly described in the following.

Data from HR LC-MS measurements can be conveniently employed for compound identification in highly complex mixtures, based on the combined evaluation of retention time, accurate *m/z* value, and isotope pattern (Figure A-13). These properties also form the basis for a targeted

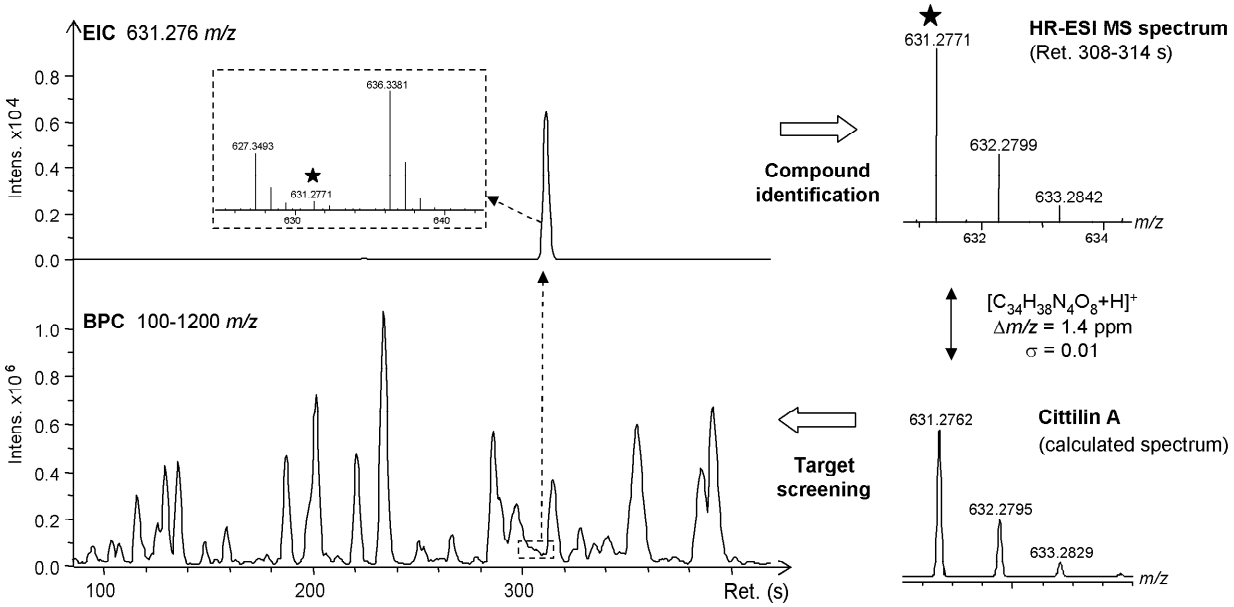


Figure A-13. Implementation of LC-coupled high-resolution mass spectrometry for compound identification and target screening, using both mass accuracy and isotope pattern information. This example illustrates the identification of cittilin A in extracts from a weak producer (note the different intensity scales of base peak chromatogram (BPC) and extracted ion chromatogram (EIC)). Adapted from [ref].

screening procedure, which is applied to HR LC-MS measurements in order to quickly gain an overview of known compounds present in a sample. This overview in turn greatly facilitates dereplication.[89] The required input for a high-resolution target screening database is the molecular formula plus the retention time determined under standardized chromatographic conditions. In a first step, accurate *m/z* values for pseudomolecular ions, $[M+H]^+$, are calculated from the molecular formulae and extracted ion chromatograms (EICs) are created in the dataset under investigation, using a low *m/z* tolerance (usually below 5 mDa). A chromatographic peak is subsequently detected by a peakfinder algorithm, and mass spectra found in the respective peak are averaged and compared to the theoretical spectrum, derived from the molecular formula (Figure A-13). Finally, the deviation of mass positions is evaluated and a match factor, the so-called "sigma value" (σ), is calculated based on the deviations of measured signal intensities from the theoretical isotope pattern. This approach, sometimes termed "SigmaFit", also enables the generation of molecular formulae for unknown compounds with increased confidence, in comparison to the evaluation of precise *m/z* values only.[90]

Data from high-resolution LC-MS measurements are also suitable for processing by an innovative "find-all-compounds" approach, which can effectively reduce the complexity of LC-

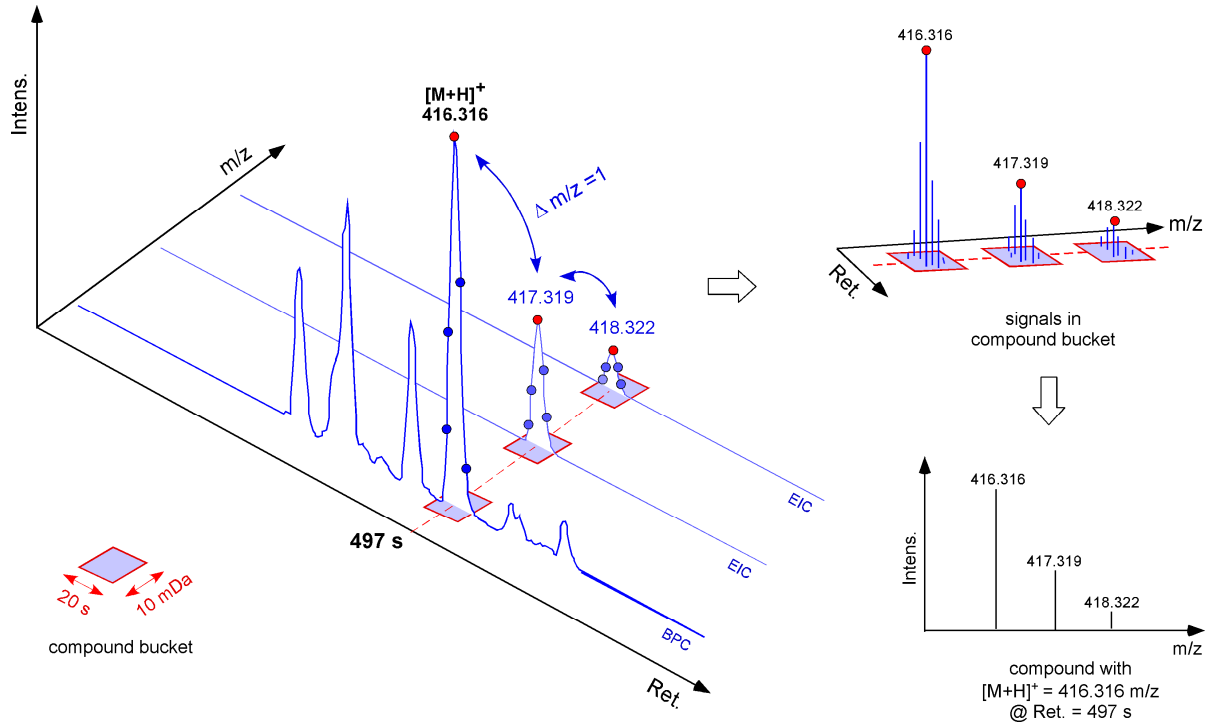


Figure A-14. Illustration of a pre-processing approach for high-resolution LC-MS data which implements a compound-finding algorithm. Isotope peaks are identified by correlation analysis, tolerances are applied in the time- and *m/z*-domain (here, Δ Ret. = 20 s and Δ m/z = 10 mDa), and all signals that originate from a parent ion are subsequently merged into one "compound bucket".

MS datasets prior to their interpretation by statistical methods. Briefly, the underlying idea is to differentiate all compounds with a chromatographic elution profile from randomly dispersed background noise (Figure A-14). All *m/z* signals that originate from one compound, including isotope peaks and multiply charged ions (if present), are merged into a "compound bucket" which is characterized by one noise-free mass spectrum, preserving the retention time as well as chromatographic peak area and intensity as additional information. It is expected that most of the putative compounds defined by this "compound-based bucketing" approach represent actual chemical species in the sample. Notably, the obtained list of reported compounds does also contain unknowns, which would not have been revealed by targeted screening (since their presence in a sample was not anticipated).

Using computational methods, the comparative analysis of untreated LC-MS data derived from different samples is usually unfeasible for a number of reasons, including the presence of a considerable degree of chemical noise and retention time variation among sample runs. Comparing the complement of compounds found by the above-described approach thus constitutes a practical alternative to the direct comparison of raw LC-MS data, when aiming at the discovery of significant differences between samples. Furthermore, the "compound buckets" may be directly used in the calculation of a statistical model for the classification of samples according to their metabolite profiles. One established method from the metabolomics field which was used in this work is "Principal Component Analysis" (PCA). PCA is an unsupervised pattern recognition technique which creates an overview of a multivariate dataset and thereby reveals groups of observations, trends and outliers.[91] In the context of LC-MS, PCA is applied as a statistical filter which highlights compounds responsible for significant variance in a set of samples.[92]

The combined use of target screening and compound-based PCA for the analysis of variation among myxobacterial production patterns, utilizing data from high-resolution ESI-TOF measurements, is described in more detail in chapter B-7.[47] These techniques were also employed to investigate the metabolite profiles from 98. *M. xanthus* strains that originate from locations worldwide (chapter B-6).[48] This study revealed a strikingly high level of intraspecific diversity in the *M. xanthus* secondary metabolome. A considerable number of non-ubiquitous candidate compounds was reported, greatly exceeding the small number of secondary metabolites previously known to derive from *M. xanthus*. Thus, *M. xanthus* should be regarded as a promising source of natural products. The implementation of high-resolution LC-MS-based methods within a general analytical framework for diversity-oriented mining of the myxobacterial secondary metabolite inventory is discussed in chapter B-8.[49]

B. Publications

Chapter B-1

Biosynthesis of (*R*)-β-tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chondromyces crocatus*

Rachid, S., Krug, D., Weissman, K. J., and Müller, R. (2007)

J. Biol. Chem. 282, 21810-21817

This article is available online at:
http://www.jbc.org/cgi/content/abstract/282/30/21810

Chapter B-2

Discovery of additional members of the tyrosine aminomutase enzyme family and the mutational analysis of CmdF

Krug, D. and Müller, R. (2009)

ChemBioChem 10, 4, 741-750

This article is available online at:
http://dx.doi.org/10.1002/cbic.200800748

Chapter B-3

Molecular and biochemical studies of chondramide formation - highly cytotoxic natural products from *Chondromyces crocatus* Cm c5

Rachid, S., Krug, D., Kochems, I., Kunze, B., Scharfe, M., Blöcker, H., Zabriskie, T.M., and Müller, R. (2006)

Chem. Biol. 14, 667-681

This article is available online at:
http://dx.doi.org/10.1016/j.chembiol.2006.06.002

Chapter B-4

Imidacins: Structure and biosynthesis of novel secondary metabolites from *Stigmatella aurantiaca* incorporating an urocanate building block

Krug, D., Dickschat, J. S., and Müller, R. (2009)

Manuscript to be submitted.

Imidacins: Structure and biosynthesis of novel secondary metabolites from *Stigmatella aurantiaca* incorporating an urocanate building block

Daniel Krug[1], Jeroen S. Dickschat[2], Rolf Müller[1]

[1] *Institute for Pharmaceutical Biotechnology, Universität des Saarlandes, Saarbrücken, Germany*
[2] *Institut für Organische Chemie, Technische Universität Braunschweig , Germany*

**Correspondence to: Rolf Müller, Pharmaceutical Biotechnology, Saarland University, P.O.Box 151150, 66041 Saarbrücken, Germany. Phone:* +49 681 302 70200*, Fax:* +49 681 302 70202

Running title: Incorporation of urocanate into novel metabolites from *S. aurantiaca*

The imidacins constitute a novel class of secondary metabolites produced in minute quantities by the myxobacterium *Stigmatella aurantiaca* Sg a15. They were discovered after inactivation of a gene encoding a histidine ammonium-lyase by comparative metabolite profiling. Structural elucidation of the imidacins revealed a unique fatty acid-like scaffold, consisting of an imidazole-acryl moiety linked via a linear alkyl chain of variable length to a carboxy group. In addition, some of these compounds exhibit a rare cyclopropane ring. The highly uncommon imidazole-acryl substructure could be shown to stem directly from urocanic acid, which in turn is derived from L-histidine. The histidine ammonium-lyase enzyme responsible for this transformation was cloned from Sg a15 and characterized biochemically. The results of targeted gene inactivation experiments and feeding studies, in combination with the identification of further candidate genes through screening of a transposon library, allowed to eventually identify the imidacin biosynthetic gene cluster and deduce a model for imidacin biosynthesis. The proposed biosynthetic route starts with the activation of urocanate, representing a building block not previously reported from any natural product. Extension by chain elongation starting from carrier protein bound urocanate is subsequently performed by an iterative type I polyketide synthase, and late biosynthetic steps include the introduction of a cyclopropane moiety by enzymatic methylenation of an unsaturated intermediate. Genes involved in imidacin biosynthesis were found distributed on at least three chromosomal loci in the *S. aurantiaca* Sg a15 genome. Thus, the data presented in this study uncover an intriguing interplay of diverse biosynthetic functionalities commonly associated with unrelated prokaryotic metabolic pathways.

Myxobacteria are soil-living organisms that exhibit a complex lifestyle, which includes coordinated swarming on surfaces, cooperative feeding on biological macromolecules or even live bacteria, and the development of multicellular fruiting bodies under starvation conditions.[1] Myxobacteria have also attracted attention as a rich source of natural products with diverse structures and biological activities.[2] The species *Stigmatella aurantiaca,* for example, is a well-established myxobacterial multiproducer of secondary metabolites. Two *S. aurantiaca* strains have been especially in the focus of research: strain Sg a15 produces the antibacterial myxalamids, the potent electron-transport inhibitor stigmatellin, the iron-chelating myxochelins as well as the alkaloid aurachin and a range of volatile compounds,[3-6] while the ACE-inhibitor myxochromid, the antibiotics myxothiazol and aurafuron, plus the DKxanthenes and dawenol have been reported from *S. aurantiaca* DW4/3-1.[7-11]

Most of these compounds originate from biosynthetic machinery consisting of polyketide synthases (PKS) and non-ribosomal peptide synthetases (NPRS), which act as enzymatic assembly lines by catalyzing the stepwise condensation of a starter unit with small monomeric building blocks, yielding more complex chemical structures. Polyketide synthases of a particular architecture, termed 'type I', exhibit a multimodular organization where each module performs one cycle of chain extension.[12] The minimum module contains three functionalities: the acyltransferase (AT) domain selects an extender

unit, usually the activated form of a short-chain dicarboxylic acid such as malonyl-CoA, and transfers it onto the acyl carrier protein (ACP) domain within the same module. The ketosynthase domain (KS) subsequently deacarboxylates the extender unit and performs a Claisen-like condensation with the growing polyketide chain bound to the multienzyme. Additional domains are often integrated into a module and contribute to the structural diversity of PKS-derived metabolites: a ketoreductase (KR) domain reduces the β-keto group to an alcohol, a dehydratase domain (DH) then catalyzes the elimination of water to form an α,β-double bond, and finally an enoylreductase (ER) domain may reduce this double bond, resulting in a methylene moiety. In textbook bacterial PKS biosynthetic pathways, each module in the enzymatic assembly line is used only once, and the order of modules and arrangement of domains contained in these modules therefore correspond to the chemical structure of the polyketide product – an observation that has been termed the 'colinearity rule'.[12] In the past years, the molecular basis for the formation of many myxobacterial natural products, including several secondary metabolites from *S. aurantiaca*, has been investigated in detail.[8,13-17] These studies have frequently revealed unexpected genetic organizations of biosynthetic pathways and surprising biosynthetic features that challenge traditional textbook biosynthetic logic, such as the iterative use or the complete skipping of modules and the incorporation of unusual building blocks and starter units.[18]

The starter units for PKS or NRPS biosynthesis are mostly recruited from primary metabolism, e.g. short-chain carboxylic acids or amino acids, and are usually activated as CoA-thioesters or (amino-)acyl adenylates prior to their loading onto the biosynthetic assembly line. However, rather uncommon building blocks may also be generated by dedicated enzymes and furnish specific pathways. One class of enzymes that has occasionally been described to fulfill this role are the aromatic amino acid ammonium-lyases.[19] These enzymes generally catalyze the non-oxidative deamination of L-tyrosine, L-phenylalanine or L-histidine to yield the corresponding α,β-unsaturated carboxylic acid. Phenylalanine ammonium-lyase (PAL), for example, has been shown to deliver cinnamic acid as a precursor for stilbene biosynthesis in *P. luminescens* and is involved in starter unit biogenesis in the enterocin biosynthetic pathway from *S. maritimus*.[20,21] Tyrosine ammonium-lyase (TAL) has been demonstrated to supply *p*-hydroxycinnamic acid for the formation of saccharomicin antibiotics in the actinomycete *S. espanaensis*.[22] In contrast to PAL and TAL which represent rare activities in bacteria, a gene encoding a histidine ammonium-lyase (HAL) can be found in the genomes of many species. HAL is commonly associated with prokaryotic primary metabolism, where it catalyzes the first step in the general histidine degradation pathway.[19]

In our current study we report, to the best of our knowledge for the first time, the involvement of HAL in secondary metabolite biosynthesis. Following the identification of urocanic acid, the product of HAL-catalyzed histidine degradation, as the starter unit for two novel natural products with intriguing structural features, we set out to elucidate their structures and additionally delineate the biosynthetic pathway for their formation in *S. aurantiaca* Sg a15.

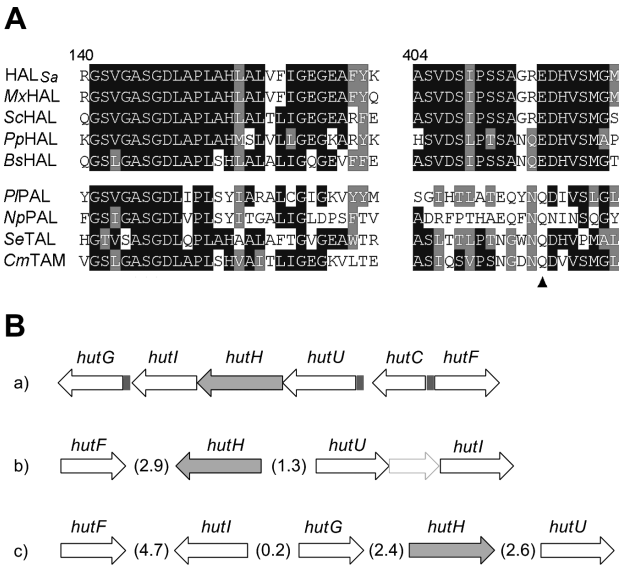


Figure 1. (A) Portion of an ClustalW alignment of prokaryotic ammonium-lyase protein sequences. Upper part, histidine ammonium-lyase enzymes (HAL): HAL_{Sa} (*S. aurantiaca* Sg a15), *Mx*HAL (*M. xanthus*), *Sc*HAL (*S. cellulosum*), *Pp*HAL (*P. putida*), *Bs*HAL (*B. subtilis*). Lower part, enzymes that use phenylalanine (PAL) or tyrosine (TAL, TAM) as substrates: *Pl*PAL (*P. luminescens*), *Np*PAL (*N. punctiforme*), *Se*TAL (*S. espanaensis*), *Cm*TAM (*C. crocatus*). The semi-conserved residue highlighted with a triangle is commonly used to distinguish HAL enzymes from ammonium-lyases which use alternative substrates. **(B)** Localization of genes encoding functions for the general histidine degradation pathway in the genomes of *P. putida* (a), *M. xanthus* (b) and *S. cellulosum* (c). Numbers in parentheses are distances in million basepairs. Grey rectangles adjacent to genes are promotor regions in *P. putida* (a). HAL is often referred to as HutH (grey arrows). See also Figure 2 and legend text.

Results

Biochemical characterization of the histidine ammonium-lyase HAL_{Sa}

In the course of previous investigations on the molecular basis of myxalamid formation in *S. aurantiaca* Sg a15 an open reading frame 1527 bp in size was identified, the deduced amino acid sequence of which exhibits similarity to many histidine ammonium-lyase (HAL) proteins from various bacterial sources.[14] Sequence comparison revealed the putative HAL from Sg a15 (HAL_{Sa}) as a homolog of *Mx*HAL from *Myxococcus xanthus* DK1622 (MXAN_3465, 86% identity, 93% similarity on the amino acid level), which was recently shown to be responsible for urocanic acid production in this myxobacterium.[23] HAL_{Sa} is also similar to biochemically characterized HAL enzymes from *Bacillus subtilis* (HutH, 46% / 64%) and *Pseudomonas putida* (HutH, 45% / 61%). Furthermore, a glutamic acid residue is found at position 415 in HAL_{Sa} which is strictly conserved in ammonium-lyase enzymes that accept histidine as substrate (Fig. 1A). This amino acid is commonly used as a predictive tool to distinguish HAL from members of the ammonium-lyase enzyme family that act on phenylalanine (PAL) or tyrosine (TAL).[20,22,24]

HAL enzymes catalyze the elimination of ammonia from L-histidine to yield *trans*-urocanic acid (Fig. 2).[19] To confirm that HAL_{Sa} can perform this function, the protein was overexpressed in *E. coli* BL21 as a C-terminal fusion with glutathione *S*-transferase, similar to a previously described procedure.[25] Purification of HAL_{Sa} (509 aa, calculated M_r, 56309 Da) was achieved by glutathione affinity chromatography, followed by on-column cleavage using PreScission protease (Fig. 3A). The purified enzyme was then incubated with L-tyrosine, L-phenylalanine and L-histidine in order to assay for ammonium-lyase activity. The biochemical characterization of recombinant HAL_{Sa} (Fig. 3B) revealed that it can efficiently convert L-histidine to urocanic acid, while no turnover with alternative substrates was observed. The determined k_{cat} (0.1 s^{-1}) and K_M (0.37 mM) values are significantly lower when compared to HutH from *P. putida* (86 s^{-1} and 3.9 mM, respectively),[26] but are found within the range of kinetic parameters reported for other enzymes of the ammonium-lyase family.[20] The highest activity, as judged by relative initial rates of product formation, was observed at pH ~9 (Fig. 3C), which is typical for ammonium-lyase-type enzymes. Thus, HAL_{Sa} indeed is a histidine ammonium-lyase, as anticipated on the basis of sequence similarity.

Inactivation of HAL_{Sa} reveals its unusual role in secondary metabolism of Sg a15

HAL is present in many bacterial species,[19] since it catalyzes the first step of the general degradation pathway from L-histidine to L-glutamate (Fig. 2). It is usually encoded within a "histidine utilization" (*hut*) operon, together with enzymes that carry out further chemical transformations on urocanate, the product of the HAL reaction (Fig. 1B). The HAL_{Sa}-encoding gene is apparently not part of such an operon in *S. aurantiaca* Sg a15, as judged by its position next to the myxalamide biosynthesis gene

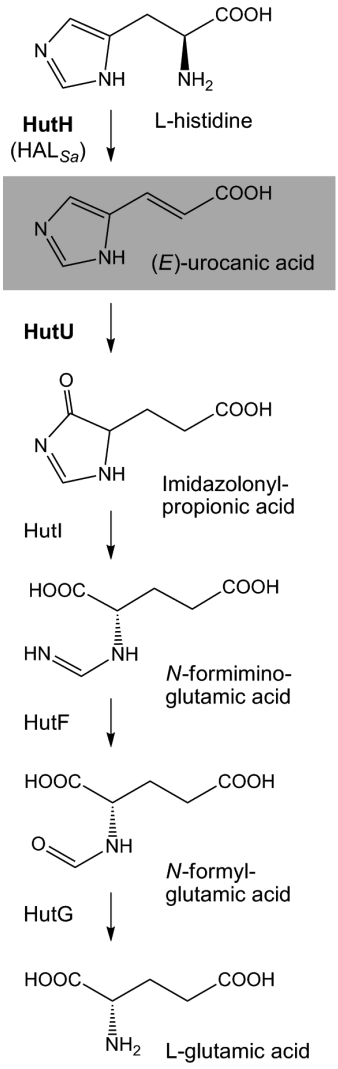


Figure 2. Chemical transformations during the degradation pathway from L-histidine via urocanic acid to L-glutamic acid. The involved enzymes are: HutH, histidine ammonium-lyase EC 4.3.1.3 (HAL); HutU, urocanase EC 4.2.1.49 ; HutI, EC 3.5.2.7; HutF, EC 3.5.3.8; HutG, EC 3.5.1.68. Urocanic acid constitutes the branching point into the biosynthesis of two novel metabolites in *S. aurantiaca* Sg a15.

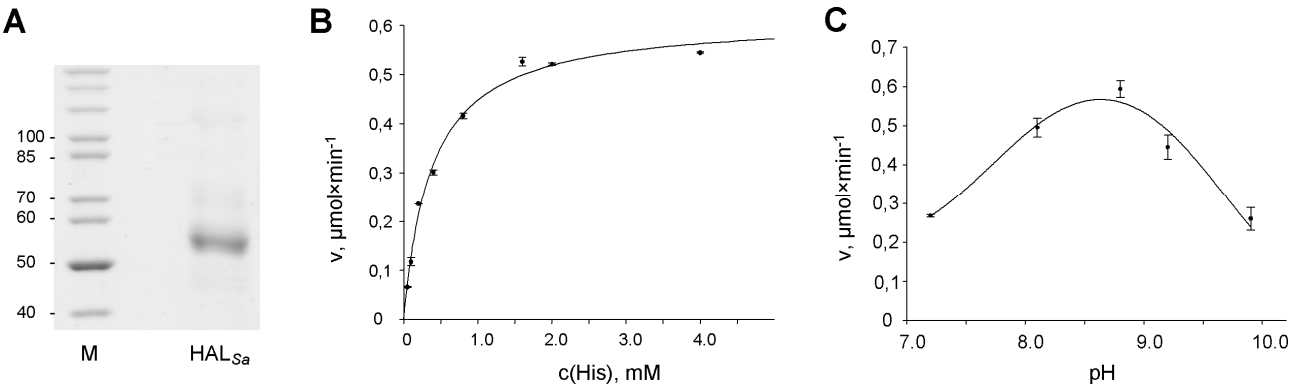


Figure 3. (A) SDS-PAGE to visualize the recombinant production in *E. coli* BL21 and purification of HAL_{Sa} (calculated M_r, ~56 kDa). **(B)** Non-linear regression of observed intial velocities for the conversion of L-histidine to urocanic acid by HAL_{Sa}, using substrate concentrations in the range from 50 µM to 4 mM. **(C)** pH-dependency of HAL_{Sa}-catalyzed urocanic acid formation.

cluster.[14] Furthermore, searching for typical histidine utilization genes in the genomes of two sequenced myxobacterial model strains, *M. xanthus* DK1622 and *S. cellulosum* So ce56, indicated that putative *hut* genes are distributed over distant locations on their chromosomes (Fig. 1B). Thus, it seems that a genetically disintegrated histidine utilization pathway is a common characteristic of various myxobacterial species. This observation prompted us to investigate the function of HAL_{Sa} by targeted gene inactivation, which was achieved by a previously established method of plasmid insertion mediated through single-crossover homologous recombination (see experimental procedures).[27] The resulting kanamycin-resistant mutant strain Sga15HAL$^-$ no longer produced urocanic acid, as confirmed by GC-MS analysis (supplemental Fig. S1). Comprehensive HPLC-DAD-MS analysis of extracts derived from liquid cultivations of wildtype Sg a15 and Sga15HAL$^-$ highlighted the absence of two signals with 319 *m/z* and 347 *m/z* in the Sga15HAL$^-$ extract (Fig. 4, traces A and B). On the basis of high-resolution ESI-MS measurements, the molecular formulae $C_{19}H_{30}N_2O_2$ (calculated *m/z* $[M+H]^+$, 319.2380 Da; measured, 319.2374 Da) and $C_{21}H_{34}N_2O_2$ (calculated *m/z* $[M+H]^+$, 347.2693 Da; measured, 347.2685 Da) could be assigned to these compounds (in the following referred to as 'c319' and 'c347' and named imidacin A_1 and A_2 later) with high confidence. These molecular compositions did not match any known myxobaterial natural product from our in-house compound database, and no plausible hits were encountered when searching multiple publicly accessible compound databases. The production of these two metabolites by the Sga15HAL$^-$ mutant was restored to wildtype level through feeding of (*E*)-urocanic acid to the liquid culture (Fig. 4, trace C). This finding suggests that their formation depends on the availability of either urocanic acid, or another downstream intermediate from the histidine degradation pathway (Fig. 2) as a potential bioynthetic precursor. Wildtype Sg a15 cultures were also supplemented with urocanic acid and with histidine, but neither addition did increase production yields of the two novel metabolites (Fig. 7C).

Inactivation of urocanase does not impair production of the novel metabolites

We next aimed at disrupting the gene *hutU*, encoding an urocanase (HutU) which catalyzes the next reaction in the histidine degradation pathway, following the initial conversion of histidine to urocanic acid. This inactivation experiment was expected to enable the discrimination between urocanic acid and its further degradation products as possible precursors for the newly identified metabolites (Fig. 2). Since genome sequence information is not available for *S. aurantiaca* Sg a15, degenerate primers were employed for the PCR-amplification of a 682-bp fragment from chromosomal DNA of that strain (see methods). This fragment was cloned into the pCR2.1Topo vector, and sequence analysis revealed convincing similarity to *hutU* genes from several organisms. The obtained plasmid pTOPO-hutU was used for insertional gene inactivation in Sg a15. HPLC-MS analysis of the resulting mutant strain Sga15HutU$^-$ showed unambiguously that the compounds with 319 *m/z* and 347 *m/z* were still produced (Fig. 4 trace D). Thus, it could be concluded that urocanic acid is recruited from the histidine degradation pathway for the biosynthesis of two

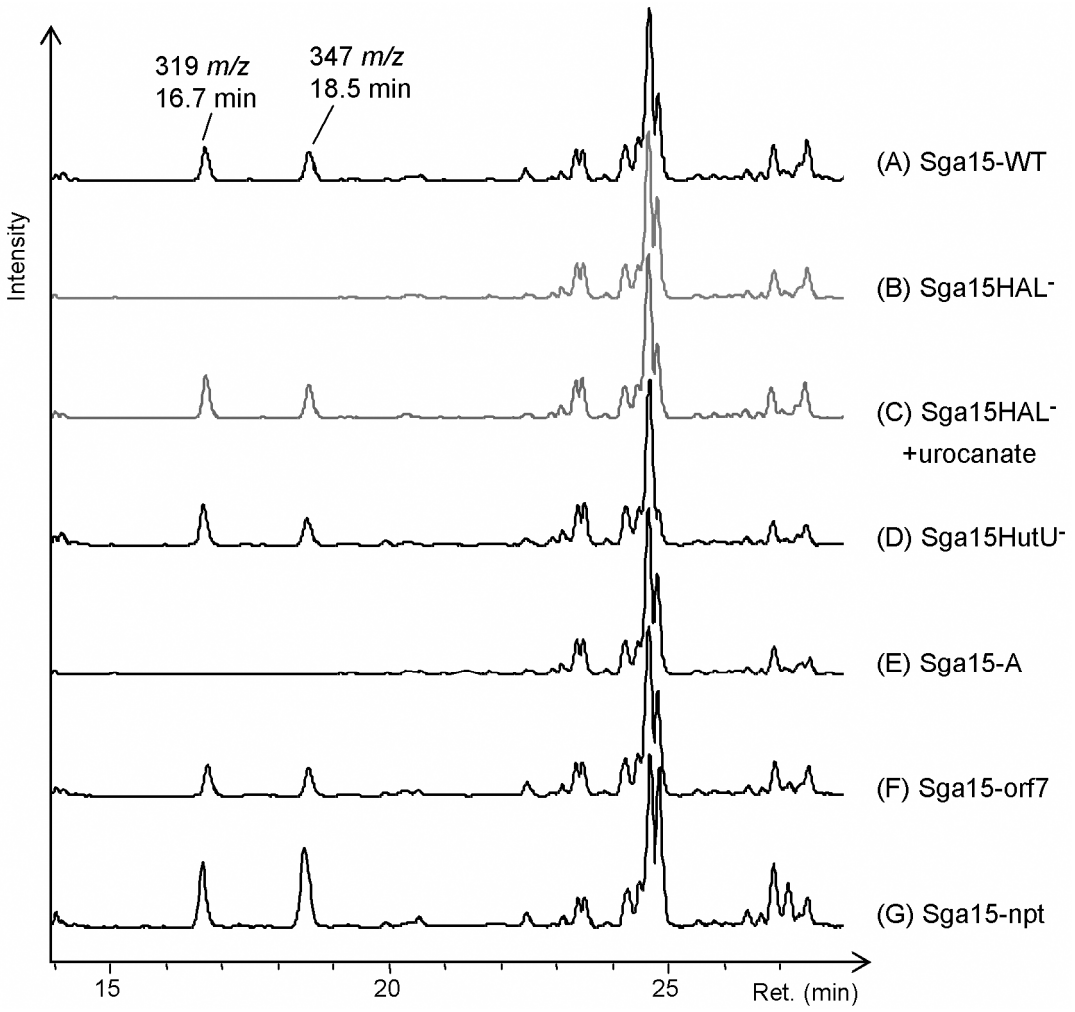


Figure 4. HPLC-MS analysis of extracts from cultivations of *S. aurantiaca* Sg a15 and selected mutants constructed in this study. **(A)** wildtype Sg a15, **(B)** HAL_{Sa} inactivation mutant Sga15HAL$^-$, **(C)** complementation of Sga15HAL$^-$ with urocanic acid, **(D)** urocanase knockout strain Sga15HutU$^-$, **(E)** strain Sga15-A with inactivated *imd1*, **(F)** strain Sga15-orf7 with inactivated *orf7*, **(G)** Sga15-npt mutant strain, harbouring the *npt* promotor sequence inserted upstream of *imd2*. See text for more details. Peaks at 16.7 min and 18.5 min represent imidacins A_1 and A_2 (compounds c319 and c347) with 319 *m/z* and 347 *m/z*, respectively. Basepeak chromatograms (BPC) in the mass range 300–400 *m/z* are shown.

novel metabolites by Sg a15. Production yields from the Sga15HutU$^-$ mutant were marginally increased in comparison to the cultivation of wildtype Sg a15, but did not rise further upon addition of histidine or urocanic acid to the culture (Fig. 7c).

Complementation of Sga15HAL$^-$ with $^{13}C_6$-urocanic acid

The outcome of the *hutU* inactivation experiment helped to define the branching point from the histidine degradation pathway, into an as yet uncharacterized biosynthetic route resulting in formation of c319 and c347, at the level of urocanic acid. Consequently, we set out to elaborate further on the fate of the proposed urocanate building block. We reasoned that feeding of ^{13}C-urocanic acid should be informative for determining whether the complete carbon scaffold, or a different molecule resulting from further transformation of urocanic acid offside the histidine degradation pathway, is actually incorporated into the novel metabolites. As uniformly ^{13}C-labeled urocanic acid was not commercially available, recombinant HAL_{Sa} was used for the *in-vitro* conversion of $^{13}C_6$-L-histidine to $^{13}C_6$-urocanic acid. The Sga15HAL$^-$ mutant was chosen for the subsequent feeding experiment, since it conveniently allowed for the incorporation of $^{13}C_6$-urocanate without intrinsic background that would result from the HAL-catalyzed conversion of non-labelled histidine in the wildtype strain. Extracts from a cultivation of Sga15HAL$^-$ in the presence of ~1 mM $^{13}C_6$-urocanic acid were analyzed by HPLC-

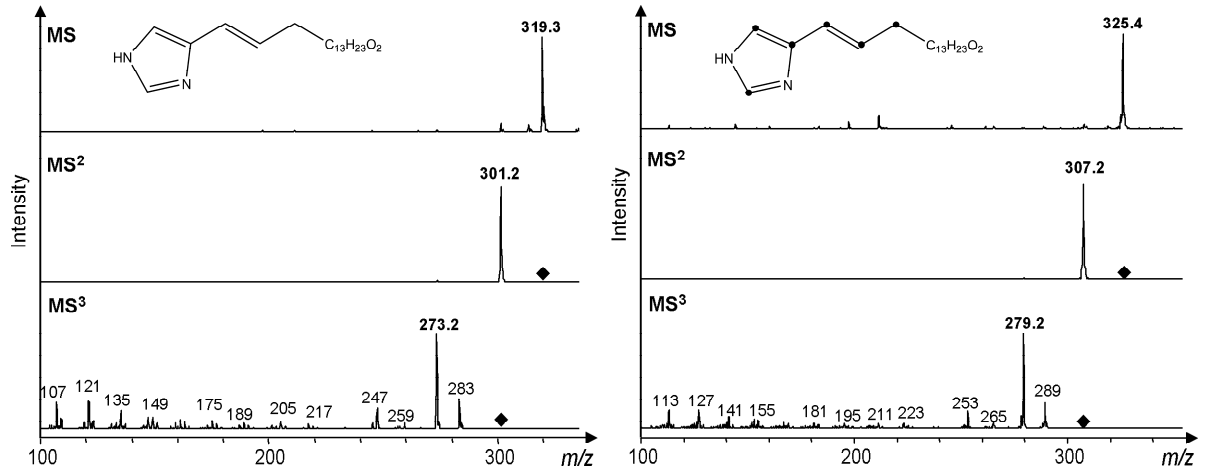


Figure 5. Incorporation of $^{13}C_6$-urocanic acid into imidacin A_1. Multiple stage tandem-MS analyses (MS^2, MS^3) of extracts form a control cultivation of strain Sga15HAL$^-$ (left side) and a culture suplemented with $^{13}C_6$-urocanic acid (right side) are shown. Incorporation of the complete carbon skeleton of urocanic acid is evident from the observed mass shifts of +6 Da.

coupled positive ESI-iontrap mass spectrometry. Signals with 325 *m/z* and 353 *m/z* were detected at the same retention times that were previously observed for the non-labelled compounds with 319 *m/z* and 347 *m/z*, and the corresponding ions were subjected to multiple stages of collisionally-induced dissociation (CID). This analysis revealed unambiguously that all carbon atoms from urocanic acid are incorporated into the novel metabolites, due to the observed mass shifts of +6 Da (Fig. 5). Furthermore, the fragmentation pattern permitted a preliminary assessment of putative compound structures. Taking into account the proposed elemental compositions, it could be estimated that the novel compounds are approximately linear molecules and consist of a terminal urocanic acid-derived substructure, linked to a carboxylate group via an alkyl chain bearing further modifications. Disappointingly, an initial isolation attempt from the supernatants of 10 L of *S. aurantiaca* Sg a15 batch cultures yielded only microgram quantities of these compounds.

Identification of biosynthetic candidate genes and inactivation experiments

Due to the limited availability of material, a comprehensive structure proposal for the novel metabolites could not be made at an early stage, and consequently genes and enzymes possibly involved in their biosynthesis could not be readily inferred. Therefore, a library of kanamycin-resistant Sg a15 mutants, harboring random chromosomal insertions of the *mariner*-based transposon plasmid pMycoMar,[16] was screened in order to identify genes with functions in the biosynthesis of the urocanate-derived metabolites. HPLC-MS analysis of supernatants from 1,920 mutants highlighted one clone, designated Sga15-29F10, which lacked formation of the target compounds with 319 *m/z* and 347 *m/z*. Production of these was not restored by feeding of urocanic acid to a Sga15-29F10 culture, indicating that a genetic locus different from HAL_{Sa} was affected by insertion of the transposable element. Plasmid recovery was carried out using *Sma*I-digested chromosomal DNA from strain Sga15-29F10 (see methods), and sequencing of plasmids p29F10 and p29F10Sma uncovered a 13 kbp stretch of sequence centered around the transposase insertion site (Fig. 6). The deduced amino acid sequences of genes located close to the termini of this region showed similarity to a partial KS domain and an α,β-hydrolase, respectively, and these genes were inactivated by plasmid insertion. The resulting mutant strains Sga15in2 and Sga15in3 were both deficient in production of the urocanate metabolites, as confirmed by HPLC-MS analysis (data not shown). PCR-screening of a pooled Sg a15 cosmid library (see methods) revealed two overlapping cosmids, Cos13F7 and Cos11D1, which contained the previously identified 13-kbp stretch of sequence, and further sequence information was acquired from these cosmids by primer walking.

The arrangement of predicted open reading frames in the sequenced region spanning a total of 46 kbp is depicted in Figure 6, and the results of the *in-silico* analysis are summarized in Table 1. The

Gene	Position (nucleotides)	Protein length (aa)	Proposed function, encoded domains, and/or similarity to:	Inactivation mutant strain
orf1	627–4039	1137	Serine/threonine protein kinase, STIAU_3097	Sga15-stpk
orf2	6194–4470	574	σ54 trancsriptional regulator, STIAU_2698	Sga15-sig54
imd2	8019–14699	2226	KS, AT, DH, ER, KR, T	Sga15in2
imd3	14700–18800	1366	KS, AT, DH, ACP	Sga15-29F10
imd4	18797–20212	471	Radical SAM superfamily, HemN	
imd5	20223–21137	304	α/β-hydrolase family enzyme, putative esterase	Sga15in3
orf3	21134–21979	284	Leucine-rich repeat domain protein	
imd1	22179–26231	1351	Adenylation domain, KR, ACP, KS	Sga15-A
orf4	32218–27748	1490	Hypothetical protein, YD repeat-containing protein	
orf5	33529–35100	523	Hypothetical protein, STIAU_4931	
orf6	36312–35086	408	Cyclomaltodextrin binding protein, STIAU_4932	
orf7	39567–36355	1070	CoA-Ligase, KS; STIAU_4934	Sga15-orf7
orf8	40030–40806	258	ABC exporter, permease protein, STIAU_4935	
orf9	40803–42128	441	ABC exporter, ATP-binding protein, STIAU_4936	
orf10	42125–42970	281	Glycosyltransferase superfamily, STIAU_4937	
orf11	42967–44009	480	Hypothetical protein, STIAU_4938	
orf12	44898–44458	146	Hypothetical protein, STIAU_4940	
orf13	46492–44900	530	Putative glycosyltransferase, STIAU_4941	

Table 1. Genes and proteins located in the sequenced region. The position in the DNA sequence and the length of deduced proteins and their proposed functions are given. The nucleotide sequence information determined in this study was deposited with EMBL (accession no. xxxxxxx). Proteins were designated as "hypothetical proteins" if no biochemically characterized representative with convincing similarity could be found by BLAST searching, using the deduced amino acid sequence as query. "STIAU" numbering relates to annotated genes in the *Stigmatella aurantiaca* DW4/3-1 draft genome sequence (GenBank accession no. AAMD00000000).

putative open reading frames *imd2*, *imd3*, *imd1* plus the three genes located in between the latter two are all transcribed in the same direction and constitute an apparent operon, flanked by intergenic gaps with a length of 1824 bp (upstream) and 1517 bp (downstream), respectively. These sequence segments exhibit an unusually low average GC-content – e.g., only 54 % are calculated for the upstream gap, and searching for proteins possibly encoded in this region provided evidence for the presence of a truncated transposase gene. The gene products of *imd1*, *imd2*, *imd3* and *orf7* harbor catalytic domains commonly associated with PKS-directed biosynthesis (Table 1 and Fig. 6). Inactivation of *imd1* – encoding a protein with predicted integral adenylation (A), enoylreductase (ER), acyl carrier-protein (ACP) and ketosynthase (KS) domains – by single-crossover plasmid insertion yielded mutant Sga15-A. This mutant did not produce the urocanate-derived metabolites, as revealed by HPLC-MS analysis (Fig. 4, trace E). In contrast, the inactivation of *orf7*, encoding a putative CoA-ligase (CL) and a ketosynthase (KS) domain, gave rise to a mutant with unchanged metabolite production pattern when compared to the wildtype (Fig. 4, trace F). Similarly, the HPLC-MS analysis of extracts from mutants Sga15-stpk and Sga15-sig54 revealed that production yields for the urocanate-derived metabolites were not significantly altered in these knockout strains (Table 1, Fig. 7C).

The deduced amino acid sequences of three putative genes interspersed in between *imd3* and *imd1* exhibit similarity to members of the SAM-binding enzyme superfamily (*imd4*), to proteins with a general α,β-hydrolase fold including enzymes with esterase activity (*imd5*), and to proteins of undefined function which possess leucine-rich repeat motifs as a common feature (*orf3*) (Fig. 6, Table 1). The intergenic distance between *orf3* and *imd1* may indicate the presence of a promoter in this region. Located downstream of *imd1*, but transcribed in the opposite direction, *orf4* encodes a hypothetical YD-repeat containing protein that is unlikely to be involved in any biosynthetic transformation.

When comparing the predicted gene products in the sequenced 46-kbp region from Sg a15 to the draft genome sequence of *S. aurantiaca* DW4/3-1,[28] it is evident that homologs of *orf5* to *orf13* plus *orf1* and *orf2* exist in that strain (Table 1). Furthermore, the similar arrangement in DW4/3-1 of genes corresponding to *orf5* to *orf13* from Sg a15 indicates local synteny between the genomes of both strains. Notably, genes with similarity to *imd1–imd3* could not be identified, consistent with the fact that strain DW4/3-1 does not produce the urocanate-derived metabolites (data not shown). The inactivation experiments described above confirm that the proteins encoded within this operon are essential for the biosynthesis of the novel urocanate-derived compounds in Sg a15. Thus, the biosynthetic genes are split into at least two chromosomal loci: one of these is the hal_{sa} gene, which is responsible for providing the urocanate precursor, and the second locus is the region encoding PKS-like biosynthetic machinery that was identified by the transposon mutagenesis approach.

Increased product yield via promoter insertion

Having identified a set of candidate genes for the biosynthesis of compounds c319 and c347, we aimed at using this information in order to genetically engineer *S. aurantiaca* Sg a15 for increased production of the novel metabolites. We hoped to thereby facilitate the isolation of sufficient material for further structural characterization because previous studies indicated the successful application of this strategy: Insertion of the *npt* promoter in front of the myxochromide biosynthetic gene cluster of *M. xanthus* DK1622 increased production approximately 50 fold (*Silke C. Wenzel and R.M., unpublished results*).

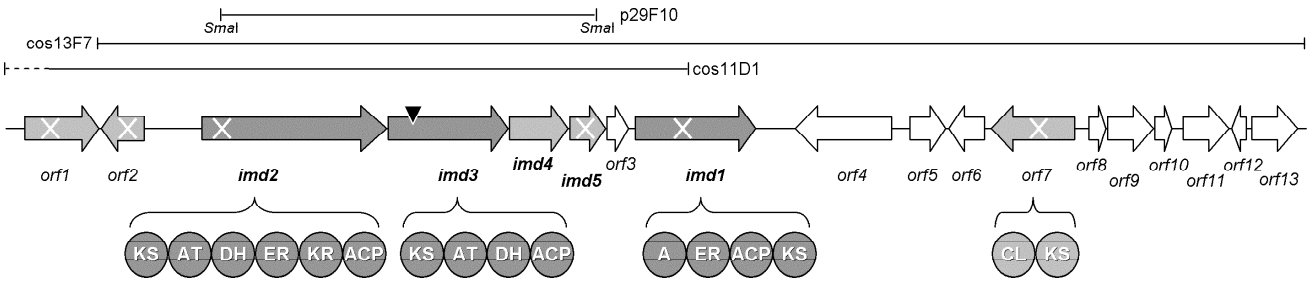


Figure 6. Putative open reading frames and proposed functions encoded in a 46-kbp sequence region covered by plasmid p29F10Sma and comsids Cos13F7 and Cos11D1. Genes marked with crosses have been inactivated by single-crossover plasmid insertion. The black triangle highlights the insertion site of the *mariner*-based transposon plasmid pMycoMar. Genes between *imd2* and *imd1* form an apparent operon. See results section for more information.

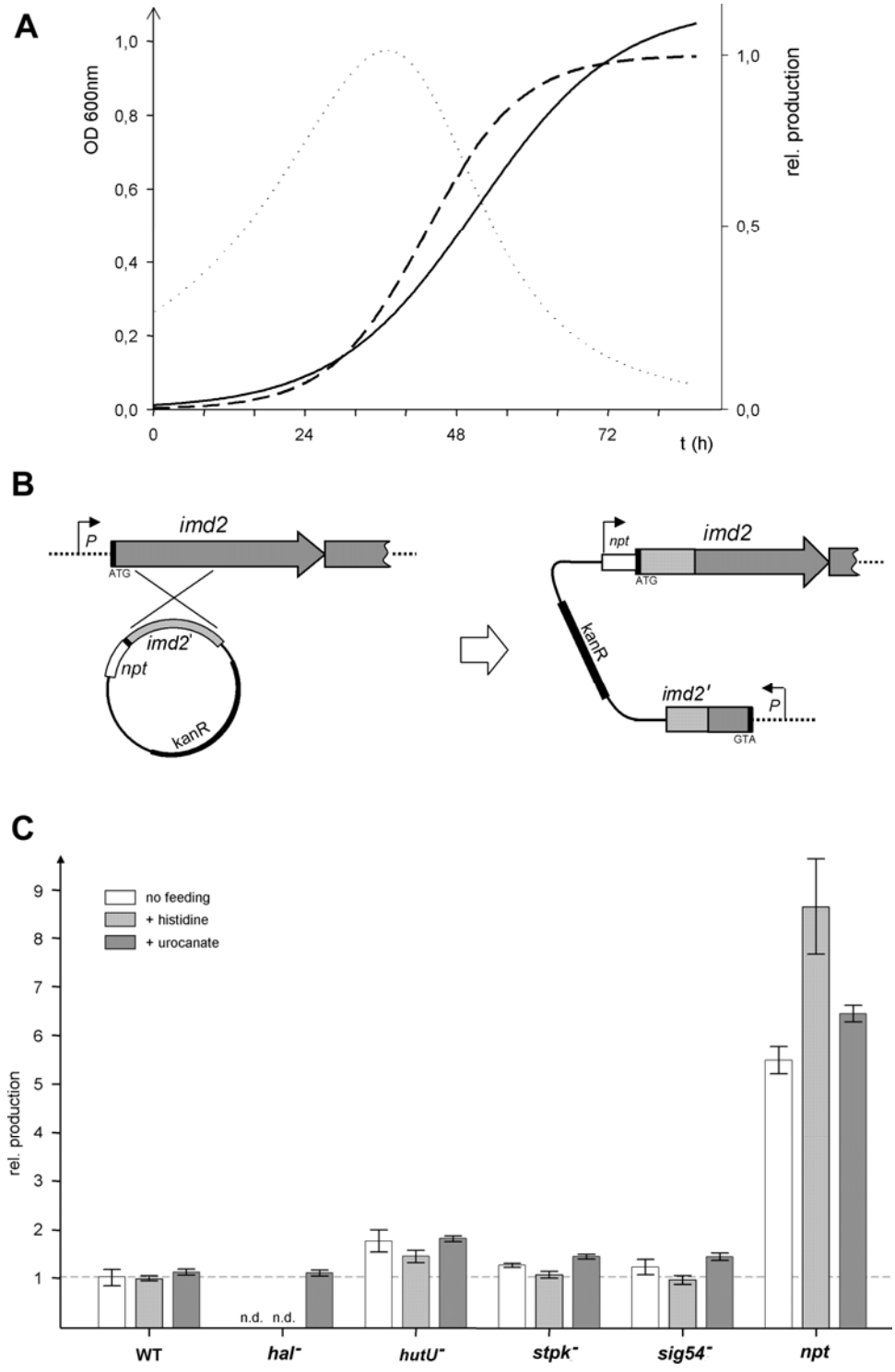


Figure 7. (A) Time course for the production of c319 and 347 by *S. aurantiaca* Sg a15. Dashed line, relative concentrations as determined by HPLC-MS analysis, summarized for both metabolites (right axis). Solid curve, bacterial growth followed by monitoring absorption at 600 nm (left axis). Results from sigmoidal curve fitting are shown. Dotted line, relative productivity expressed as ratio between the slope of product concentration and growth at every timepoint. **(B)** Insertion of the *npt* promotor sequence upstream of *imd2* via a single-crossover homologous recombination approach. **(C)** Relative roduction yields from cultivations of the *S. aurantica* Sg a15 wildtype (WT) and mutants Sga15HAL$^-$ (*hal*$^-$), Sga15HutU$^-$ (*hutU*$^-$), Sga15-stpk (*stpk*$^-$), Sga15-sig54 (*sig54*$^-$) and Sga15-npt (*npt*). White bars, without feeding; light grey bars, with feeding of histidine; dark grey bars, with feeding of urocanic acid. All values represent the mean of three independent cultivations, normalized to the average wildtype production level which was set as "1". Error bars were calculated from standard deviations (triplicate experiments).

The relative concentrations of compounds c319 and c347 were determined via LC-MS measurements (see methods) at various timepoints during growth of *S. aurantiaca* Sg a15 in liquid medium, in order to shed light on the correlation between their formation and growth of Sg a15. The results of this production kinetics experiment are shown in Fig. 7A: Maximum productivity was observed at the onset of the exponential growth phase (ca. 30 h after inoculation), and compound concentration did not increase further when the Sg a15 culture approached the stationary phase (at ~80 h and later). Thus, production of the urocanate metabolites seems to be modulated by a regulatory mechanism, which could possibly involve transcriptional control via a putative promotor region upstream of the *imd1–imd5* operon. Therefore, the mutant Sga15-npt was constructed, in which this operon was placed under the control of the constitutive *npt* promotor, utilizing a single-crossover plasmid insertion approach (Fig. 7B). The production yields of compounds c319 and c347 from Sga15-npt were approximately 5-folds higher when compared to the wildtype level, and increased further upon feeding of urocanate (ca. 6.5-fold) or histidine (up to 9-fold) (Fig. 4G and Fig. 7C). Growth characteristics and general phenotypic traits of the mutant strain Sga15-npt were largely unchanged in comparison with the wildtype.

Structural features of the urocanate-derived metabolites c319 and c347

Compound isolation and purification was carried out by extracting the supernatants from shaking flask cultivations of Sga15-npt with ethyl acetate, followed by size exclusion chromatography and successive rounds of reversed-phase liquid chromatography. The final yield from a total culture volume of 10 L was in the range of 0.2–0.3 mg. Compounds c319 and c347 were subjected to multiple stages of high-resolution tandem-MS analysis using static nano-electrospray ionization (nanoESI) and collisionally induced dissociation (CID) on a linear trap–orbitrap mass spectrometer. The obtained spectra (supplemental Fig. S2) suggest the initial two-step elimination of a terminal carboxylic acid group and the subsequent loss of multiple neutral fragments, with cleavages of carbon-carbon bonds occurring along an alkyl chain. This type of fragmentation pattern may be explained by a charge-remote fragmentation mechanism,[29] when the positive charge is located at the terminal imidazole moiety. This interpretation is supported by the mass shifts observed from incorporation of $^{13}C_6$-urocanic acid (Fig. 5) and by the calculation of high-confidence molecular formulae for fragments containing the imidazole-acryl unit (Fig. S3). The CID fragmentation analysis also provided evidence for a modification of the carbon chain near to the carboxy terminus that accounts for one double bond equivalent.

The proposed structure was refined by GC-MS analysis following derivatization of c319 and c347 with MSTFA (Fig. 8). The EI mass spectra indicate the presence of alkyl chains in both derivatives which differ in length by two methylene units, as concluded from the mass shift of 28 Da for a number of signals (marked with asterisks in Fig. 8). A distance of 40 Da between distinct fragments (marked with arrowheads in Fig. 8), likely resulting from C2–C3 and C4–C5 bond cleavage, respectively, indicates the presence of a $-C_3H_4-$ moiety between these positions, which is reasonably explained by a cyclopropane functionality. Indeed, a similar fragmentation behaviour has been observed previously in EI mass spectra of compounds with a cyclopropane group embedded in a comparable chemical environment, e.g. hexylcyclopropyl-pyrrolidinyl-octanone (Supplemental Figure S3). This structural detail was further substantiated by the analysis of ^{1}H-NMR data for compound c347, showing the expected signals for protons located at the disubstituted cyclopropane ring (supplemental Table S1). The NMR analysis also confirmed the presence of the imidazole acryl moiety and its connection to the alkyl chain. The structure of the saturated chain itself could not be determined in NMR experiments due to intense signal overlaps. Thus, compounds c319 and c347 were assigned the structures shown in Figure 9A based on the combined results of mass spectral- and NMR analysis and were named imidacins A_1 and A_2. These molecules may be regarded as formal analogs of fatty acids, featuring an urocanate starter unit and a cyclopropane group located in the unusual Δ^3-position.

Incorporation of (methyl-d_3)-L-methionine

A variety of microorganisms, including *E. coli* and *M. tuberculosis*, are known to produce membrane lipids that contain cyclopropyl fatty acids,[30] and it has been demonstrated that cyclopropane ring formation is accomplished through the action of cyclopropane synthases.[31,32] These SAM-dependent enzymes catalyze the methylenation ('cyclopropanation') of unsaturated fatty acid precursors via addition of a methionine-derived methyl group to the double bond, followed by rapid proton loss (Fig. 9B).[33] To probe the biosynthetic origin of the methylene group in the cyclopropane moiety of compounds c319 and c347, a culture of strain Sga15-npt

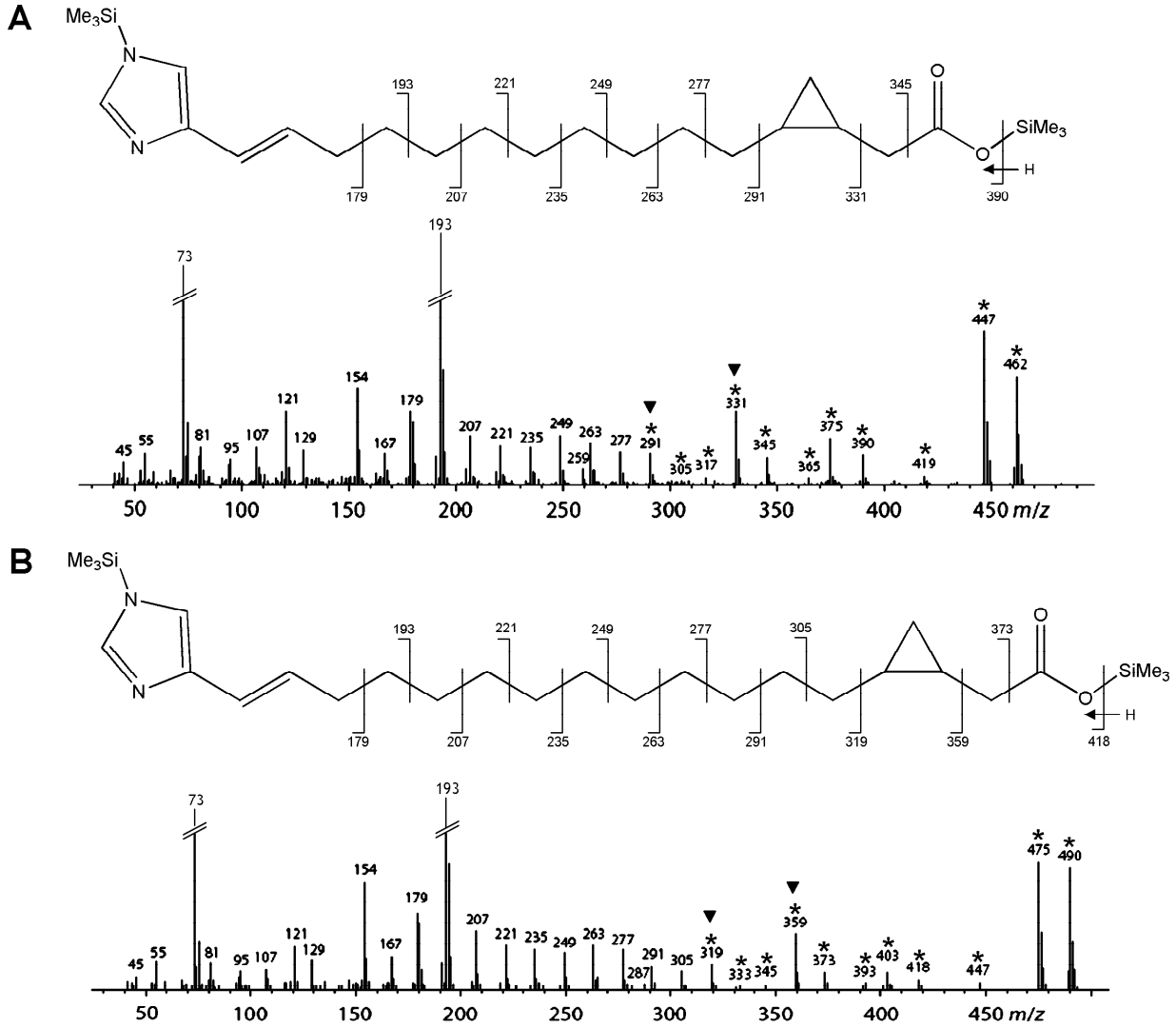


Figure 8. GC-MS analysis of compounds c319 (imidacin A_1) **(A)** and c347 (imidacin A_2) **(B)** following derivatization with MSTFA, and structural assignment of fragments. Signals marked with an asterisk experience a mass shift of 28 Da, when comparing both spectra. Black arrowheads highlight fragmentation flanking the cyclopropyl moiety.

was supplemented with (methyl-d_3)-L-methionine. The subsequent HPLC-MS analysis showed unambiguously that two deuterium labels were incorporated into the urocanate metabolites with high abundance (Fig. 10A). Thus, the C3' methylene unit is most likely derived from methionine, and a SAM-dependent cyclopropane synthase is probably responsible for ring formation. The logical substrate for such an enzyme would be a Δ^3-unsaturated precursor of the urocanate metabolites (Fig. 9B), and high-resolution LC-MS analysis revealed indeed the presence of a compound with matching elemental composition, plausible retention time and fragmentation pattern in extracts from Sg a15 (supplemental Fig. S4). Moreover, a signal corresponding to a putative precursor with a completely saturated alkyl chain was also detected, albeit with very low abundance. Taken together, these findings suggest that the late steps in the biosynthesis of compounds c319 and c347 include the introduction of a double bond in the Δ^3-position of a fatty acid-like precursor and subsequent SAM-dependent cyclopropanation (Fig. 9B). Interestingly, *imd4* encodes a protein with a conserved SAM binding site (Table 1 and supplemental Figure S5).

Discussion

In this article, the discovery of two novel secondary metabolites from *S. aurantiaca* Sg a15 is described and a biosynthetic route for their formation

is proposed which depends on the supply of urocanic acid as a starter unit by histidine ammonium-lyase. Related enzymes of the ammonium-lyase family that use phenylalanine or tyrosine as substrates occur rarely in bacteria, but have been reported previously to provide precursors for secondary metabolite biosynthetic pathways.[20,22,34] HAL, in contrast, is found in many prokaryotes, but has been investigated exclusively in the context of catabolic histidine utilization to date (Fig. 2). Notably, this study for the first time assigns a different role to HAL – specifically, to provide the urocanate building block for imidacin biosynthesis by *S. aurantiaca*. Both, the genomic localization of the gene encoding HAL_{Sa} and its signficantly different kinetic parameters in comparison with other biochemically characterized HAL enzymes,[19] might be interpreted as evolutionary adaption to this unusual function: Placement of the HAL_{Sa}-coding gene outside of a histidine utilization operon permits the independent regulation of its expression, while the low K_M value determined *in-vitro* for HAL_{Sa} enables it to retain catalytic efficiency even at low physiological histidine concentrations. Moreover, high turnover numbers as achieved for example by HAL from *P. putida* are not required to effectively furnish the imidacin biosynthetic pathway with an urocanate precursor. Any other physiological significance of HAL-mediated histidine degradation for the primary metabolism of *S. aurantiaca* is at present unclear. The non-essentiality of HAL may enable the identification of further urocanate-dependent metabolites from other strains by comparing their metabolite profiles to those of HAL-inactivated mutants.

To the best of our knowledge, the imidacins are the first natural products known to incorporate an urocanic acid moiety. The recruitment of carboxylic acids from bacterial primary metabolism for their use as starter units in secondary metabolite biosynthesis has been demonstrated for a number of pathways to proceed via the activation as acyl-adenylates by AMP-ligases, or as CoA-esters by CoA-ligase enzymes, respectively.[35] The activated building blocks are subsequently transferred onto the 4'-phosphopantheteine group of dedicated carrier proteins, where they are bound as thioesters. Alternatively, they can undergo further chemical transformations before being loaded onto the biosynthetic assembly line. Well-established examples (among many others) for these modes of starter unit sequestration include the activation of 2,3-dihydroxybenzoic acid by the adenylating enzymes during enterobactin biosynthesis (EntE) and myxochelin (MxcE) biosynthesis, respectively.[13,36] Another example is the formation of cinnamoyl-CoA by the CoA-ligase EncH as an early step in enterocin bioynthesis.[34] Two candidate proteins, which could accomplish the activation of urocanate in analogous ways, were found in the genomic region identified by transposon mutagenesis as essential for imidacin biosythesis: the first domain encoded by *imd1* shows similarity to EntE and MxcE, while one domain encoded by *orf7* exhibits similarity to CoA-ligases (Fig. 6). Targeted gene inactivation experiments revealed unambigously that the protein encoded by *orf7* has no function in imidacin biosynthesis; in contrast, disruption of *imd1* abolished imidacin production. Thus, the adenylation domain Imd1-A is likely to be responsible for the activation of urocanic acid as urocanyl-AMP, probably followed by transfer onto the integral carrier protein domain Imd1-ACP (Fig. 9B). Feeding of $^{13}C_6$-urocanate revealed that the complete urocanate backbone is preserved in the imidacin structure, arguing against the activation of a hypothetical truncated derivative of urocanate by Imd1-A. Furthermore, this finding rules out the possibility that activated urocanate undergoes further modifications associated with loss of carbons, for example via a β-oxidative route as it has been reported for cinnamate in the course of enterocin biosynthesis.[34]

An enoylreductase domain, Imd1-ER, is located in between the A domain and the ACP encoded by *imd1*. A similar domain arrangement is found in the loading module of the rapamycin biosynthetic gene cluster, where the enyolreductase has been shown to reduce the double bond of the dihydroxycyclohexene carboxylic acid starter unit.[37] However, the double bonds in the imidazole-acryl moiety of urocanate are apparently transferred into the final structure of the imidacins. Thus, Imd1-ER does presumably not process the urocanate-ACP intermediate, even though this domain should be regarded as functional as judged by the presence of intact signature motifs in its amino acid sequence.

Following selection and loading of the urocanic acid starter unit, we reason that the linear scaffold of the imidacins derives from multiple rounds of chain elongation according to PKS biosynthetic logic. The existence of two imidacin derivatives which differ in chain length by one $-C_2H_4-$ unit suggests that their biosynthesis involves an iterative mechanism for carbon chain elongation (Fig. 9B). A suitable candidate module is encoded by *imd2*, which has the full complement of KS, AT, DH, ER, KR and ACP domains to catalyze one step of chain extension including ketoreduction, dehydration and enoylreduction of the ACP-bound intermediate. The integral Imd2-AT domain exhibits the required substrate specificity for malonyl-CoA, according to the conserved

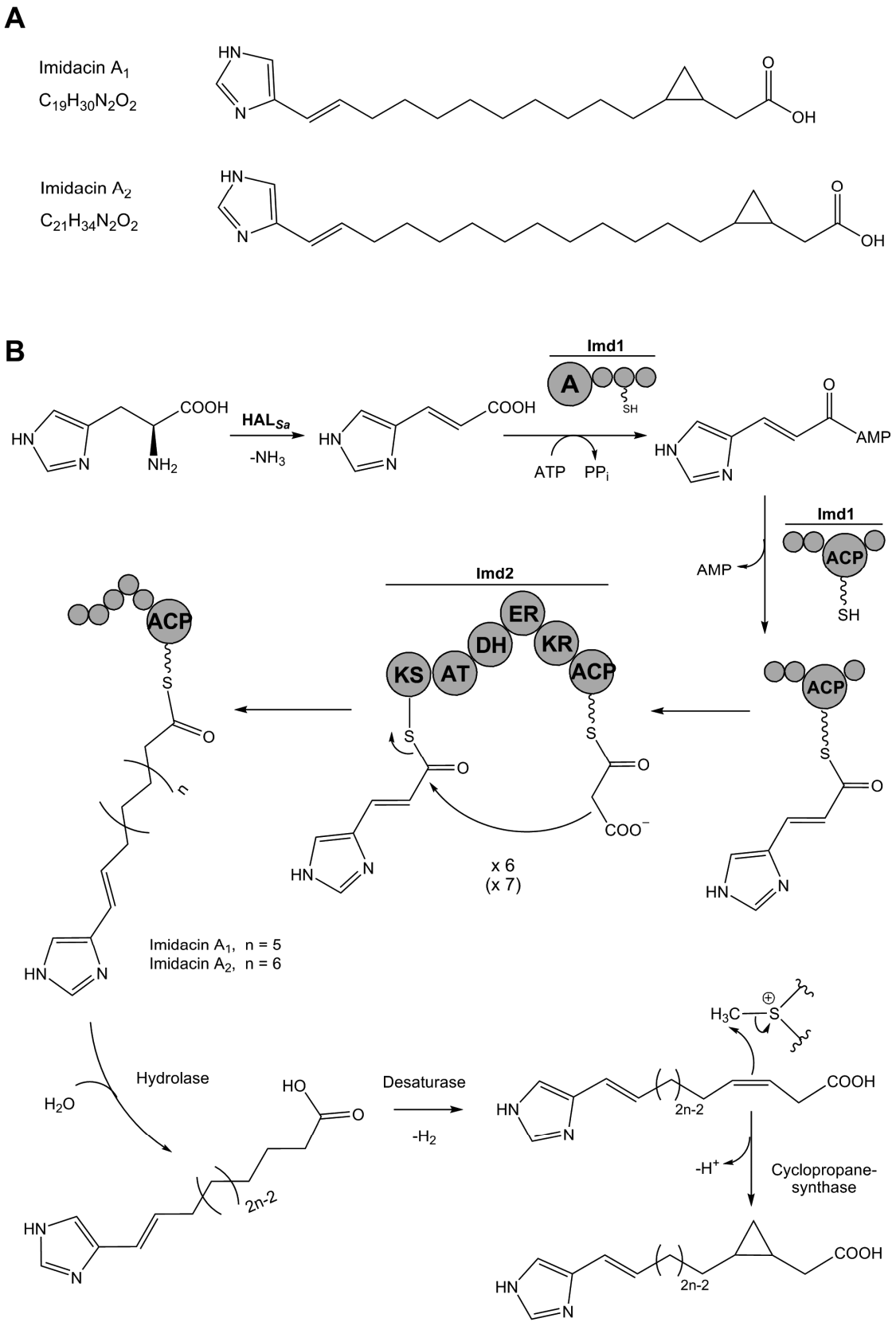


Figure 9. (A) Structures of the two novel metabolites imidacin A_1 and A_2. **(B)** The proposed biosynthetic route for the formation of imidacins by *S. aurantiaca* Sg a15, starting with the conversion of L-histidine to urocanate.

HAFH motif in its sequence.[38] Unusual but not unprecedented iterative use of module Imd2 would then result in formation of the linear, completely reduced carbon backbone of imidacins A_1 and A_2 after six and seven rounds of catalysis, respectively (Fig. 9B). The iterative use of modules in bacterial type I PKS has been reported previously from a growing number of myxobacterial and streptomycete megasynthetases.[18,39] For example, the polyene side chain of myxochromides is generated by only one PKS module (including a non-functional ER domain) which is capable of producing unsaturated polyketide chains of variable length.[7] Programmed iteration of a single module is believed to be involved in stigmatellin biosynthesis and was recently shown to lead to variable chain lengths during bioynthesis of the DKxanthenes.[11,40] Similarly, the natural products aureothin, avilamycin and borellidin from streptomycetes have been shown to originate, at least in part, from iterative type I PKS biochemistry.[41-43] However, to the best of our knowledge none of the PKS biosynthesis pathways investigated to date delivers a fully reduced and unbranched alkyl chain mimiking the outcome of fatty acid biosynthesis, by the repeated action of only one PKS module. The variable length of the PKS-derived scaffold in imidacins A_1 and A_2 may be explained by a lack of stringent chain length control. This hypothesis will have to be addressed in further studies as the mechanisms underlying chain length control in iterative PKS systems still await clarification.

An additional module harboring PKS biosynthesis-related domains (KS, AT, DH, ACP) is encoded by *imd3* (Fig. 6). All domains seem to be functional, as judged by the presence of conserved sequence motifs, and Imd3-AT is predicted to exhibit the same substrate specificity for malonyl-CoA as Imd2-AT. However, the absence of a ketoreductase domain renders the integral DH domain useless and thus argues against a role of this module in imidacin biosynthesis, since all incorporated building blocks are fully reduced in the final imidacin structure. Disruption of *imd3* by transposon insertion abolished imidacin production in mutant Sga15-29F10 (Fig. 6), but this finding might be explained by a polar effect on the downstream genes (see below).

The products of PKS-type biosynthetic pathways are usually released from the assembly line by a dedicated thioesterase (TE) domain located at the C-terminal end of the last module in the respective pathway. A typical TE domain is not encoded within the region that was sequenced here, but an enzyme with convincing similarity to proteins of the α/β-hydrolase family, including representatives that exhibit esterase activity, is encoded by the open reading frame *imd5,* located between *imd3* and *imd1* (Fig. 6). Targeted inactivation of this gene led to a mutant deficient in imidacin production. This result is unlikely to arise from polar effects, since *imd1* (which is preceeded by a non-coding region, sufficient in length to contain a promoter) is probably transcribed independently from *imd5*, and the hypothetical involvement of *orf3,* encoding a short 'leucine-rich repeat' protein, in imidacin biosynthesis is considered highly implausible. Thus, we propose that *imd5* encodes an enzyme that catalyzes hydrolytic release of the fully chain-elongated and completely saturated imidacin precursor from Imd2 as the free acid (Fig. 9B). Notably, two members of the α/β-hydrolase enzyme family, MonCII and NanCII, have recently been revealed as novel chain-terminating thioesterases in monensin and nanchangmycin biosynthesis, respectively.[44,45] The amino acid sequences of these proteins displayed low overall similarity to known thioesterases from a variety of PKS biosynthetic pathways, but share with them the typical α/β-hydrolase fold and an essential serine residue at the active site.[44] The same holds true for the putative α/β-hydrolase Imd5 (see also supporting information, Figure S5) and thus suggests Imd5 to function as a thioesterase.

The final imidacin structure exhibits a cyclopropane moiety likely resulting from the methylenation of a double-bond by an SAM-dependent cyclopropane synthase, as concluded from the incorporation of two deuterium labels after feeding of (methyl-d3)-L-methionine (Fig. 10). High-resolution LC-MS analysis of cell extracts provided evidence for the presence of both a fully saturated and a mono-unsaturated imidacin precursor, named imidacin C and B, respectively (supplemental Figure S4). We therefore propose that desaturation and cyclopropanation take place as post-PKS modifications with high efficiency (Fig. 9B). In the absence of a suitable candidate gene for a desaturase that could deliver the Δ^3-mono-unsaturated imidacin precursor for subsequent cyclopropanation, we assume that this function is carried out by an enzyme encoded elsewere in the Sg a15 genome. Intriguingly, the cloning and biochemical characterization of the first desaturase with Δ^3-regioselectivity has been reported very recently from a fungal source.[46] We cannot exclude, however, the possibility that the double bond is incorporated during PKS assembly. Nevertheless, this seems unlikely due to the unusual double bond geometry which is only found rarely in PKS products[17]. It may be speculated about the unorthodox formation of this double bond by an enoyl-

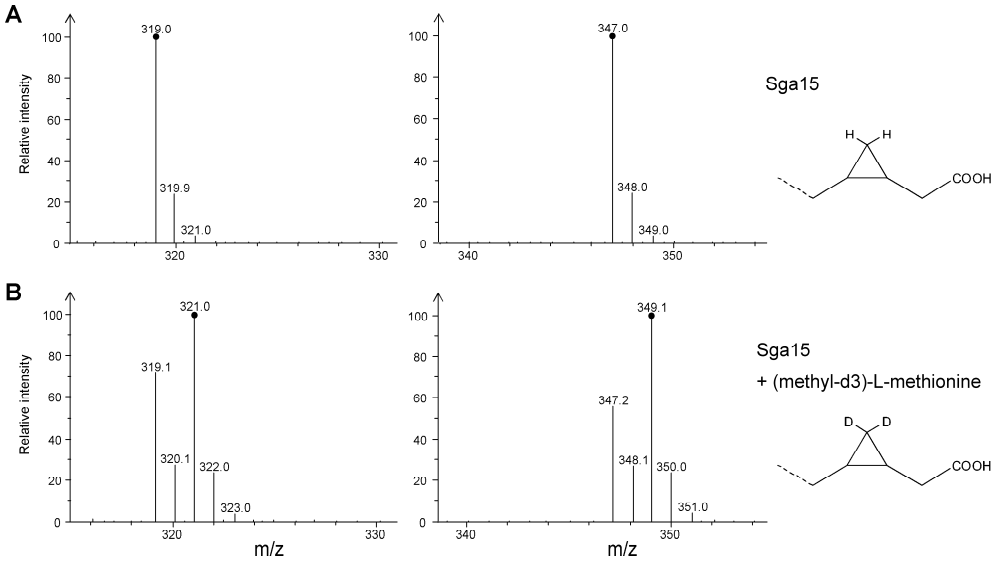


Figure 10. Mass spectra of imidacin A_1 (compound c319, left) and imidacin A_2 (compound c347, right) obtained from a culture of Sga15-npt **(A)** and a feeding experiment where the same strain was supplemented with (methyl-d_3)-L-methionine **(B)**, revealing the incorporation of two deuterium labels into the cyclopropane ring moiety.

reductase domain catalyzing the inverse reaction. The Imd1-ER domain (which lacks a role in the respective chain extension cycle) could theoretically introduce the Δ^3 double bond in this way. However, this would also require that Imd1-ER processes a chemical moiety in the imidacin structure which is located distantly to the part of the molecule assembled by Imd1.

Enzymes capable of introducing a cyclopropane ring into unsaturated fatty acid scaffolds via the addition of an SAM-derived methylene group have been described in *E. coli* and *M. tuberculosis*. However, these enzymes do not exhibit the regioselectivity that would be required for imidacin cyclopropanation: The *E. coli* cyclopropane synthase usually processes Δ^9-unsaturated fatty acids contained in cell wall phospholipids, while two distinct cyclopropane synthase enzymes in *M. tuberculosis* transform distantly positioned double bonds into cyclopropane moieties in the long-chain mycolic acids.[30,47] A candidate enzyme for imidacin cyclopropanation is encoded by *imd4*, as suggested by the presence of a conserved SAM-binding motif in the deduced amino acid sequence (supplemental Figure S5).[48] Although Imd4 is more similar to HemN-like proteins which belong to the 'radical SAM' superfamily rather than to known cyclopropane synthases,[49] it is most likely not a functional HemN-analog because a signature CxxxCxxC motif found to date in all known 'radical SAM' enzymes is absent (Figure S5).[48] Thus, Imd4 may instead constitute a novel type of cyclopropane synthase involved in the formation of imidacins A. However, we cannot exclude at present that the relevant enzyme is encoded outside the imidacin biosynthetic gene cluster, just as proposed above for the missing desaturase.

Genes involved in regulation of secondary metabolite production are sometimes encoded within or in the immediate vicinity of biosynthetic gene clusters. Two genes with conceiveable regulatory roles were identified upstream of the *imd* operon: *orf1* encodes a putative serine/threonine protein kinase (STPK) and *orf2* encodes a σ^{54} transcriptional activator (Fig. 6). However, targeted inactivation of *orf1* and *orf2* did not influence the production of imidacins by the respective mutant strains (Figure 7C), and we conclude therefore that the STPK and the σ^{54} transcriptional activator are not involved in the regulation of imidacin biosynthesis.

In summary, the genes that direct imidacin formation in Sg a15 are split across at least two, more likely three distant loci in the genome, and the organization of genes encoded within the *imd* operon is not colinear with the proposed biosynthesis. Furthermore, not all theoretically required genes could yet be identified, but on the other hand some of the encoded functions are apparently not required. It seems plausible that the *imd* operon was placed in the genome of strain Sg a15 as part of a sequence transposition event – this idea is supported by the unusually low GC content in the preceeding intergenic gap and the sequence remains of a putative transposase. The seemingly redundant or cryptic

genes (e.g. *imd3*, *orf3*) could thus represent evolutionary artifacts.

The assignment of some biological role for imidacins in the producing host Sg a15 must at present remain largely speculative. Cyclopropyl fatty acids, representing the closest structural analogs of imidacins, occur frequently as constituents of cell wall lipids in a variety of bacteria, where they are thought to modulate membrane fluidity in response to environmental influences.[30] Cyclopropyl-modified mycolic acids have also been implicated in pathogenicity of *M. tuberculosis*.[47] The imidacins could in principle be part of the *S. aurantiaca* cell envelope; however, tentative attempts to obtain free imidacins from crude cell lysates by saponification of hypothetical esters were unsuccessful (data not shown), suggesting that the imidacins are not major constituents of lipids in *S. aurantiaca*. In light of the extraordinarily low, growth phase-dependent production of imidacins, and taking into account the generally complex lifestyle of myxobacteria, an as yet unrecognized regulatory role or their use as strain-specific signal molecules should also be considered.

Besides the unclear biological significance of imidacin production for *S. aurantiaca* Sg a15, the question needs to be addressed in future studies whether genes without a predicted function for imidacin biosynthesis are actually dispensable. Moreover, the proposed roles of certain candidate genes and domains encoded within modules require confirmation by targeted inactivation and mutational analysis. Unfortunately, genetic manipulation of *S. aurantiaca* Sg a15 is limited by the availability of reliable selection markers and, despite serious efforts to establish better genetic systems, the lack of a method for introducing markerless mutations (*Julia Hovermann and R.M., unpublished results*). Thus, further investigations would clearly benefit from heterologous expression of the complete imidacin pathway, for which the stage has been set by cloning of relevant genes in this study. Experiments aimed at the identification of the missing desaturase gene and the biochemical characterization of the unusual cyclopropane synthase and hydrolase enzymes are currently under way.

Significance

The impressive structural diversity of microbial secondary metabolites stems from the virtually unlimited combinations of biosynthetic building blocks, the various ways of assembly into more complex biosynthetic intermediates, and the enzymatic transformations that decorate a core structure with additional features to finally result in each natural product's intriguing characteristics. The structure of the novel metabolites reported here, and the biosynthetic machinery that achieves their formation, are remarkable in several respects. First, an unprecedented urocanic acid starter unit is provided by a histidine ammonium-lyase. Urocanate is subsequently activated and elongated by an iterative PKS module to yield a completely reduced linear backbone. The resulting fatty acid-like scaffold is then further modified with a rare cyclopropane moiety, which presumably results from the successive action of a desaturase and methylenation by a cyclopropane synthase. The corresponding enzymes are commonly associated with fatty acid metabolism. In imidacin biosynthesis they seem to exhibit exceptional substrate specificity and regioselectivity not previously observed in bacteria.

In conclusion, a variety of enzymatic activities from unrelated biosynthetic pathways cooperate in *S. aurantiaca* Sg a15 to produce the unique imidacin structure. Genes encoding the required functions are distributed over multiple locations in the Sg a15 genome, rendering it highly unlikely that the biosynthetic route, let alone its product, would be predictable by bioinformatic methods using whole-genome data. Therefore, the discovery of the imidacins highlights the value of combined molecular biological and metabolic profiling approaches to mining the myxobacterial natural product repository – exemplified here by the identification of a novel compound class through comparison of metabolite profiles following inactivation of the general histidine utilization pathway. The biosynthetic capabilities that form the basis of such unexpected secondary metabolite diversity, in particular those that accomplish the activation of unusual monomeric building blocks or carry out rare chemical transformations, represent useful additions to the enzymatic 'toolbox' for the design and genetic engineering of biosynthetic pathways.

Experimental procedures

Chemicals, media, strains and fermentation conditions – All chemicals were purchased from Sigma-Aldrich unless indicated otherwise. *E. coli* strain DH10B was grown at 37 °C and strain BL21 was grown at 30 °C in Erlenmeyer flasks in liquid LB medium, supplemented with the following antibiotics whenever appropriate: ampicillin, 100 µg/mL; kanamycin sulfate, 50 µg/mL. *Stigmatella aurantiaca* Sg a15 and its decendants were grown in liquid tryptone medium (ref) at 30 °C and 200 rpm and maintained on TS agar plates,[1] containing kanamycin

when required for selection. The sequences of primers can be found in the supplemental material (Table S2).

Recombinant production, purification and activity assay of HAL_{Sa} – Heterologous expression of *Sa*HAL was achieved as C-terminal fusion to glutathione *S*-transferase (GST). The primer pair HALSga15start and HALSga15stop was employed for the PCR-amplification of a 1527 bp-fragment from genomic DNA of strain *S. aurantiaca* Sg a15 using Pfu polymerase (Promega), adding mutageneic *Bgl*II and *Sal*I restriciton sites. The fragment was digested accordingly and ligated to the *Bam*HI/*Sal*I-digested pGEX-6P-1 vector backbone (Invitrogen), using T4 DNA Ligase (MBI-Fermentas). Following transformation into *E. coli* DH10B, plasmid DNA was prepared using the GeneJet plasmid isolation kit (MBI-Fermentas) and the insert of the obtained plasmid pGEX-*Sa*HAL was sequence-verified (MWG Eurofins). *E. coli* BL21, transformed with the pGEX-*Sa*HAL expression plasmid, was grown overnight in 2 mL LB medium with ampicillin, and was inoculated at a 1:100 dilution into 50 mL LB medium containing ampicillin. When growth reached an OD_{600} of 1.0, IPTG was added to a final concentration of 0.2 mM, and cultivation was continued for 2.5 h at 30 °C. Cells were harvested by centrifugation, resuspended in PBS buffer and passed twice through a French pressure cell. The crude lysate was clarified by centrifugation and protein purification was carried out by affinity chromatography using GST-Sepharose. The GST tag was cleaved with PreScission Protease at 4 °C according to the manufacturer's recommendations and the purified protein was eluted in buffer containing 50 mM Tris (pH 7.0), 150 mM NaCl, 1 mM EDTA, and 1 mM dithiothreitol. Enzyme activity was assayed using 1–10 µg purified protein in 0.1 M Tris buffer (pH 8.8) at 30 °C, with substrate concentrations ranging from 50 µM–4 mM L-histidine, or up to 5 mM phenylalanine or 2 mM tyrosine. Protein concentrations were determined using the colorimetric assay described by Bradford with chemicals supplied by BioRad.[50]

Cloning of an hutU-*homologous fragment from S. aurantiaca Sg a15* – The degenerate primer pair HutU_QMTA, 5'-GCCTGATGATGTACGGCCARATGACNGC-3' and HutU_PGFV (5'-CGGATGTACGCCGGCACRAANCCNGG-3' (degenerate regions are underlined) was designed using the CODEHOP approach,[51] and used for the PCR-amplification of a 682-bp fragment from genomic DNA of *S. aurantiaca* Sg a15. The fragment was cloned into the pCR2.1Topo vector (Invitrogen) to yield the plasmid pTOPO-Sga15hutU, the insert sequence of which was determined by sequencing.

Targeted gene inactivation in S. aurantiaca *Sg a15* – In general, an internal fragment (500–800 bp in size) of the gene to be inactivated was PCR-amplified from genomic DNA using Taq polymerase and sequence-specific primers (listed in supplemental Table S2), and cloned into the pCR2.1Topo vector. The resulting inactivation plasmid was used for electroporation of *S. aurantiaca* Sg a15 as previously described.[27] Viable colonies were obtained on TS agar supplemented with kanamycin, and chromosomal DNA was pepared from liquid cultures using the Puregene Kit (Gentra). Integration of the plasmid into the correct genomic locus was routinely verified by PCR, using control primers (Table S2) which are located outside the amplified region, in combination with vector-specific primers pTOPOin and pTOPOout.

HPLC-MS analysis of S. aurantiaca *Sg a15 metabolite profiles* – Typically, a 50 mL shaking flask culture of *S. aurantiaca* Sg a15 was grown for 72–84 hours, clarified by centrifugation, and the supernatant was extracted with excess ethyl acetate. Following evaporation of the solvent until complete dryness, the solid residue was redissolved in 1 mL of methanol. A 5 µL aliquot was used for injection into an Agilent 1100 series HPLC-DAD system coupled to a Bruker HCTplus ESI-iontrap MS device. Chromatographic separation was performed on a Luna RP-C18 column (100 x 2 mm, 2.5 µm particle size; Phenomenex, Germany) equipped with a 4 x 2 mm C18 guard cartridge. The mobile phase system consisted of water (A) and acetonitrile (B), each containing 0.1 % formic acid. Linear gradients from 5 %B to 95 %B were routinely employed at a flow rate of 0.4 mL/min, and the column temperature was held at 32 °C. MS analysis was carried out in positive ionization mode. A tandem-MS method was used for quantitative measurements, which records the intensities of $[M\text{-}H_2O+H]^+$ fragments derived from the parent ions with $[M+H]^+ = 319$ *m/z* and $[M+H]^+ = 347$ *m/z*.

HPLC-MS based screening of a library of Sg a15 mutants and plasmid recovery – The construction of a library of *S. aurantiaca* mutants harbouring random insertional gene inactivations was reported previously.[16] Briefly, *S. aurantiaca* Sg a15 was transformed with the *mariner*-based transposon plasmid pMycoMar,[52] and viable clones were subsequently selected on TS agar containing kanamycin and transferred to 150 µL tryptone medium (with kanamycin added) in 96-well plates. After growth for 5 days at 30 °C and 700 rpm on a wellplate incubator (Infors), 50 µL aliquots from each well were added to 100 µL methanol, the suspension was clarified by centrifugation and 15 µL aliquots from the supernatant were subjected to HPLC-MS analysis. Separation was carried out on a RP-C18 column (Nucleodur, 125 x 2 mm, 3 µm particle size, Macherey & Nagel) equipped with a precolumn (8 x 3 mm, 5 µm partile size), with instrumentation and a mobile phase system essentially as described above. Detection of the target compounds with high sensitivity and selectivity was achieved by operating the ion-trap in manual tandem-MS mode and recording the intensity of $[M\text{-}H_2O+H]^+$ fragment ions derived from the precursors with $[M+H]^+ = 319$ *m/z* and $[M+H]^+ = 347$ *m/z*, respectively. For plasmid recovery from mutant Sga15-29F10, chromosomal DNA was digested with *Sma*I, re-ligated using T4 DNA ligase and used for the transformation of *E. coli* DH5αλpir. The obtained plasmid p29F10 was used to determine the cloned sequence flanking the transposable element, using primers K388 and K390[52]. Treatment of p29F10 with *Sma*I followed by ligation to the *Sma*I-digestested pBluescriptSK vector

backbone gave plasmid p29F10Sma, which was used for sequencing into the region in between the former chromosomally-encoded *Sma*I restriction sites, utilizing standard M13 primers.

PCR-screening of a Sg a15 cosmid library – The construction of a cosmid library in *E.coli* SURE using the Supercos Vector (Stratagene), containing 35-48 kbp fragments of the Sg a15 chromosome, was described previously.[27] Pools consisting of 96 cosmids were initially employed as template for the PCR-amplication of fragments from the genes *imd2* and *imd5* (Fig. 6), using primers listed in supplemental Table S2. Subsequently, additional rounds of PCR using pools of 32 cosmids and finally the single cosmids as templates were carried out to identify the overlapping cosmids cos13F7 and cos11D1. These cosmids were introduced into *E.coli* DH10B by electroporation, and cosmid DNA prepared with the GeneJet Kit (MBI Fermentas) was used for sequencing (MWG Eurofins).

Insertion of an upstream npt *promotor adjacent to orf1* – The *npt* (neomycin phosphotransferase) promotor sequence was PCR-amplified from the plasmid pCR2.1Topo (Invitrogen) using Phusion polymerase (Finnzymes) and primers Mch71 and Mch72,[53] adding a modified RBS sequence (*Silke C. Wenzel, unpublished*). The amplified fragment 157 bp in size was gel-purified and used in a subsequent overlap-extension PCR step, in combination with primers Mch71 and Sga15con4b and a 1073-bp fragment previously amplified from cosmid cos13F7 using the primer pair Sga15act3neu2 and Sga15con4b (supplemental Table S2). Following the addition of single A-overhangs to the resulting PCR product by Taq polymerase, the gel-purified fragment 1209 bp in size was cloned into the pCR2.1Topo vector to give plasmid pTOPO-Sga15act. The correct fusion of the *npt* sequence to the *orf1*-homologous region was verified by sequencing, and the plasmid was used for the transformation of *S. aurantiaca* Sg a15, yielding strain Sga15-npt.

Preparation of samples for GC-MS analysis – Extracts were separated by HPLC as described above, and fractions corresponding to the retention times of compounds with 319 *m/z* and 347 *m/z* were collected. Solvents were removed from the samples by evaporation to complete dryness, and the residue was dissolved in dimethylformamide or pyridine. An equal volume of MSTFA was added and the reaction was allowed to proceed for 1 h at 60 °C prior to injection into the 6890N GC-MS analyzer with MSD (Agilent Technologies, Germany).

Direct-infusion high-resolution ESI-MS measurements – Samples were seperated on a RP-C18 column (50 x 2 mm, BEH, 1.7 µm particle size; Waters, Germany) using an Accela ultra-high performance liquid chromatography system (Thermo-Fisher), and fractions were collected in 96-well plates (Abgene, Germany) with the Triversa NanoMate system (Advion, UK). The fractions were dried under nitrogen and a mixture of 50 % aqueous methanol, containing 0.1 % formic acid, was added. Static electrospray of sample aliquots was achieved using the automated chip-based nanoESI capability of the NanoMate device with parameters as follows: spray voltage, 1.7 kV; back pressure, 30 psi. Detection was performed on the LTQ-Orbitrap mass spectrometer (Thermo-Fisher) which was set to a resolution of 30,000 (FWHM at 400 *m/z*) with internal lock-mass calibration. Multiple stage tandem-MS studies were conducted in the linear trap using relative normalized CID energies in the range of 20–35%.

Acknowledgements

We would like to thank Eva Luxenburger for skilful technical assistance with HPLC-MS measurements and for her help in compound isolation. Research in the laboratory of RM was funded by the "Deutsche Forschungsgemeinschaft" (DFG) and the "Bundesministerium für Bildung und Forschung" (BMB+F).

References

1. **Shimkets L., Dworkin M., Reichenbach H. (2006).** The Myxobacteria. *In:* Dworkin, M. (Ed.), The Prokaryotes, Springer, Berlin, 31-115.

2. **Bode H.B., Müller R. (2007).** Secondary metabolism in myxobacteria. *In:* Whitworth, D. (Ed.), Myxobacteria: multicellularity and differentiation, ASM Press, Chicago, 259-282.

3. **Gerth K., Jansen R., Reifenstahl G., Höfle G., Irschik H., Kunze B., Reichenbach H., and Thierbach G. (1983).** The myxalamids, new antibiotics from *Myxococcus xanthus* (Myxobacterales). I. Production, physico-chemical and biological properties, and mechanism of action. *J.Antibiot.* 36, 1150-1156.

4. **Höfle G., Kunze B., Zorzin C., and Reichenbach H. (1984).** Antibiotika aus Gleitenden Bakterien, XXIII: Stigmatellin A und B - zwei neue Antibiotika aus *Stigmatella aurantiaca* (Myxobacterales). *Liebigs Ann.Chem.* 8, 1883-1904.

5. **Kunze B., Höfle G., and Reichenbach H. (1987).** The aurachins, new quinoline antibiotics from myxobacteria: production, physico-chemical and biological properties. *J.Antibiot.* 40, 258-265.

6. **Dickschat J.S., Bode H.B., Wenzel S.C., Müller R., and Schulz S. (2005).** Biosynthesis and identification of volatiles released by the myxobacterium *Stigmatella aurantiaca*. *ChemBioChem* 6, 2023-2033.

7. **Wenzel S.C., Kunze B., Höfle G., Silakowski B., Scharfe M., Blöcker H., and Müller R. (2005).**

Structure and biosynthesis of myxochromides S1-3 in *Stigmatella aurantiaca*: Evidence for an iterative bacterial type I polyketide synthase and for module skipping in nonribosomal peptide biosynthesis. *ChemBioChem* 6, 375-385.

8. **Silakowski B., Schairer H.U., Ehret H., Kunze B., Weinig S., Nordsiek G., Brandt P., Blöcker H., Höfle G., Beyer S., and Müller R. (1999).** New lessons for combinatorial biosynthesis from myxobacteria: the myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. *J.Biol.Chem.* 274, 37391-37399.

9. **Kunze B., Reichenbach H., Müller R., and Höfle G. (2005).** Aurafuron A and B, new bioactive polyketides from *Stigmatella aurantiaca* and *Archangium gephyra (myxobacteria)*. *J.Antibiot.* 58, 244-251.

10. **Söker U., Kunze B., Reichenbach H., and Höfle G. (2003).** Dawenol, a new polyene metabolite from the myxobacterium *Stigmatella aurantiaca*. *Z. Naturforsch. B* 58, 1024-1026.

11. **Meiser P., Weissman K.J., Bode H.B., Krug D., Dickschat J.S., Sandmann A., and Müller R. (2008).** DKxanthene biosynthesis -- understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism. *Chem. Biol.* 15, 771-781.

12. **Staunton J. and Weissman K.J. (2001).** Polyketide biosynthesis: a millennium review. *Nat.Prod.Rep.* 18, 380-416.

13. **Gaitatzis N., Kunze B., and Müller R. (2001).** In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proc.Natl.Acad.Sci.U.S.A.* 98, 11136-11141.

14. **Silakowski B., Nordsiek G., Kunze B., Blöcker H., and Müller R. (2001).** Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: The myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga15. *Chem. Biol.* 8, 59-69.

15. **Frank B., Wenzel S.C., Bode H.B., Scharfe M., Blöcker H., and Müller R. (2007).** From genetic diversity to metabolic unity: studies on the biosynthesis of aurafurones and aurafuron-like structures in myxobacteria and streptomycetes. *J. Mol. Biol.* 374, 24-38.

16. **Sandmann A., Dickschat J., Jenke-Kodama H., Kunze B., Dittmann E., and Müller R. (2007).** A type II polyketide synthase from the gram-negative bacterium *Stigmatella aurantiaca* is involved in aurachin alkaloid biosynthesis. *Angew.Chem.Int.Ed. Engl.* 46, 2712-2716.

17. **Weinig S., Mahmud T., and Müller R. (2003).** Markerless mutations in the myxothiazol biosynthetic gene cluster: A delicate megasynthetase with a superfluous nonribosomal peptide synthetase domain. *Chem.Biol.* 10, 953-960.

18. **Wenzel S.C. and Müller R. (2007).** Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat.Prod.Rep.* 24, 1211-1224.

19. **Poppe L. and Rétey J. (2005).** Friedel-Crafts-type mechanism for the enzymatic elimination of ammonia from histidine and phenylalanine. *Angew.Chem.Int.Ed.Engl.* 44, 3668-3688.

20. **Williams J.S., Thomas M., and Clarke D.J. (2005).** The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. *Microbiology* 151, 2543-2550.

21. **Xiang L. and Moore B.S. (2002).** Inactivation, complementation, and heterologous expression of encP, a novel bacterial phenylalanine ammonia-lyase gene. *J.Biol.Chem.* 277, 32505-32509.

22. **Berner M., Krug D., Bihlmaier C., Vente A., Müller R., and Bechthold A. (2006).** Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. *J.Bacteriol.* 188, 2666-2673.

23. **Krug D. and Müller R. (2008).** Discovery of additional members of the tyrosine aminomutase enzme family and mutational analysis of CmdF. *ChemBioChem*. in press.

24. **Calabrese J.C., Jordan D.B., Boodhoo A., Sariaslani S., and Vannelli T. (2004).** Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. *Biochemistry* 43, 11403-11416.

25. **Rachid S., Krug D., Weissman K.J., and Müller R. (2007).** Biosynthesis of (*R*)-beta-tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chondromyces crocatus*. *J.Biol.Chem.* 282, 21810-21817.

26. **Röther D., Poppe L., Viergutz S., Langer B., and Rétey J. (2001).** Characterization of the active site of histidine ammonia-lyase from Pseudomonas putida. *Eur.J.Biochem.* 268, 6011-6019.

27. **Beyer S., Kunze B., Silakowski B., and Müller R. (1999).** Metabolic diversity in myxobacteria: identi-

fication of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide-(poly)peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce90. *Biochim.Biophys.Acta* 1445, 185-195.

28. **Ronning C.M., Nierman W.C. (2007).** The genomes of *Myxococcus xanthus* and *Stigmatella aurantiaca*. *In:* Whitworth, D. (Ed.), Myxobacteria: multicellularity and differentiation, ASM Press, Chicago, 285-298.

29. **Denekamp C., Van den Heuvel H., Voinov V.G., Claeys M., Seto C., Grossert J.S., Waddell D.S., Curtis J.M., and Boyd R.K. (2000).** Charge-remote fragmentation characteristics of functionalized alkanes in high-energy collision-induced dissociation. *Rapid Commun.Mass Spectrom.* 14, 1035-1043.

30. **Grogan D.W. and Cronan J.E. (1997).** Cyclopropane ring formation in membrane lipids of bacteria. *Microbiol.Mol.Biol.Rev.* 61, 429-&.

31. **Courtois F., Guerard C., Thomas X., and Ploux O. (2004).** *Escherichia coli* cyclopropane fatty acid synthase. *Eur.J.Biochem.* 271, 4769-4778.

32. **Glickman M.S., Cahill S.M., and Jacobs W.R. (2001).** The *Mycobacterium tuberculosis* cmaA2 gene encodes a mycolic acid trans-cyclopropane synthetase. *J.Biol.Chem.* 276, 2228-2233.

33. **Stuart L.J., Buck J.P., Tremblay A.E., and Buist P.H. (2006).** Configurational analysis of cyclopropyl fatty acids isolated from *Escherichia coli*. *Organic Letters* 8, 79-81.

34. **Xiang L. and Moore B.S. (2003).** Characterization of Benzoyl Coenzyme A Biosynthesis Genes in the Enterocin-Producing Bacterium '*Streptomyces maritimus*'. *J.Bacteriol.* 185, 399-404.

35. **Moore B.S. and Hertweck C. (2002).** Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat.Prod.Rep.* 19, 70-99.

36. **Ehmann D.E., Shaw-Reid C.A., Losey H.C., and Walsh C.T. (2000).** The EntF and EntE adenylation domains of Escherichia coli enterobactin synthetase: sequestration and selectivity in acyl-AMP transfers to thiolation domain cosubstrates. *Proc.Natl.Acad.Sci.U.S.A.* 97, 2509-2514.

37. **Gregory M.A., Petkovic H., Lill R.E., Moss S.J., Wilkinson B., Gaisser S., Leadlay P.F., and Sheridan R.M. (2005).** Mutasynthesis of rapamycin analogues through the manipulation of a gene governing starter unit biosynthesis. *Angew.Chem.Int.Ed.Engl.* 44, 4757-4760.

38. **Yadav G., Gokhale R.S., and Mohanty D. (2003).** Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* 328, 335-363.

39. **Wilkinson B., Foster G., Rudd B.A., Taylor N.L., Blackaby A.P., Sidebottom P.J., Cooper D.J., Dawson M.J., Buss A.D., Gaisser S., Bohm I.U., Rowe C.J., Cortes J., Leadlay P.F., and Staunton J. (2000).** Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. *Chem.Biol.* 7, 111-117.

40. **Gaitatzis N., Silakowski B., Kunze B., Nordsiek G., Blöcker H., Höfle G., and Müller R. (2002).** The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J.Biol.Chem.* 277, 13082-13090.

41. **He J. and Hertweck C. (2005).** Functional analysis of the aureothin iterative type I polyketide synthase. *ChemBioChem* 6, 908-912.

42. **Gaisser S., Trefzer A., Stockert S., Kirschning A., and Bechthold A. (1997).** Cloning of an avilamycin biosynthetic gene cluster from Streptomyces viridochromogenes Tü57. *J.Bacteriol.* 179, 6271-6278.

43. **Olano C., Wilkinson B., Sanchez C., Moss S.J., Sheridan R., Math V., Weston A.J., Brana A.F., Martin C.J., Oliynyk M., Mendez C., Leadlay P.F., and Salas J.A. (2004).** Biosynthesis of the angiogenesis inhibitor borrelidin by Streptomyces parvulus Tu4055: Cluster analysis and assignment of functions. *Chem.Biol.* 11, 87-97.

44. **Harvey B.M., Hong H., Jones M.A., Hughes-Thomas Z.A., Goss R.M., Heathcote M.L., Bolanos-Garcia V.M., Kroutil W., Staunton J., Leadlay P.F., and Spencer J.B. (2006).** Evidence that a novel thioesterase is responsible for polyketide chain release during biosynthesis of the polyether ionophore monensin. *ChemBioChem* 7, 1435-1442.

45. **Liu T.G., Lin X., Zhou X.F., Deng Z.X., and Cane D.E. (2008).** Mechanism of thioesterase-catalyzed chain release in the biosynthesis of the polyether antibiotic nanchangmycin. *Chem. Biol.* 15, 449-458.

46. **Zauner S., Zahringer U., Lindner B., Warnecke D., and Sperling P. (2008).** Identification and functional characterization of the 2-hydroxy fatty N-acyl-delta-3(E)-desaturase from F*usarium graminearum*. *J.Biol.Chem.* in press.

47. **George K.M., Yuan Y., Sherman D.R., and Barry C.E. (1995).** The biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis* - Iden-

tification and functional analysis of Cmas 2. *J.Biol.Chem.* 270, 27292-27298.

48. **Sofia H.J., Chen G., Hetzler B.G., Reyes-Spindola J.F., and Miller N.E. (2001).** Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* 29, 1097-1106.

49. **Layer G., Verfurth K., Mahlitz E., and Jahn D. (2002).** Oxygen-independent coproporphyrinogen-III oxidase HemN from *Escherichia coli*. *J.Biol.Chem.* 277, 34136-34142.

50. **Bradford M.M. (1976).** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principleof protein-dye binding. *Anal. Biochemistry* 72, 248-254.

51. **Rose T.M., Henikoff J.G., and Henikoff S. (2003).** CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Res.* 31, 3763-3766.

52. **Sandmann A., Sasse F., and Müller R. (2004).** Identification and analysis of the core biosynthetic machinery of tubulysin, a potent cytotoxin with potential anticancer activity. *Chem.Biol.* 11, 1071-1079.

53. **Meiser P. and Müller R. (2008).** Two functionally redundant Sfp-type 4'-phosphopantetheinyl transferases differentially activate biosynthetic pathways in *Myxococcus xanthus*. *ChemBioChem* 9, 1549-1553.

Supplemental Material

Figure S1

Targeted gene inactivation in Sga15, genetic verification and phenotypic analysis.

Figure S2

High-resolution MS^3 CID spectra of imidacins A.

Figure S3

GC-MS library spectra (NIST) of hexylcyclopropyl-oxooctylpyrrolidine.

Figure S4

HPLC high-resolution ESI-MS analysis and proposed structures of imidacins in Sga15 extracts.

Figure S5

Annotated protein sequences of Imd4 and Imd5.

Table S1

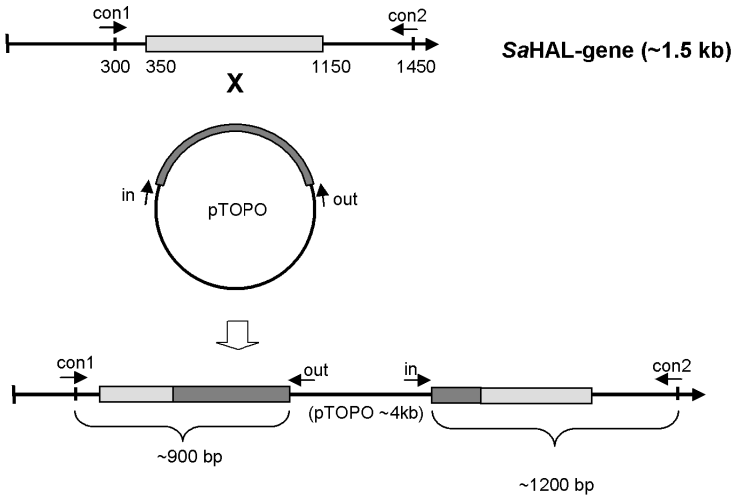
1H-NMR data for imidacin A_2 (c347).

Table S2

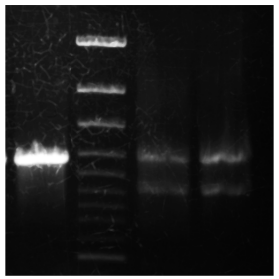
Oligonucleotide sequences.

Supplemental Figure S1

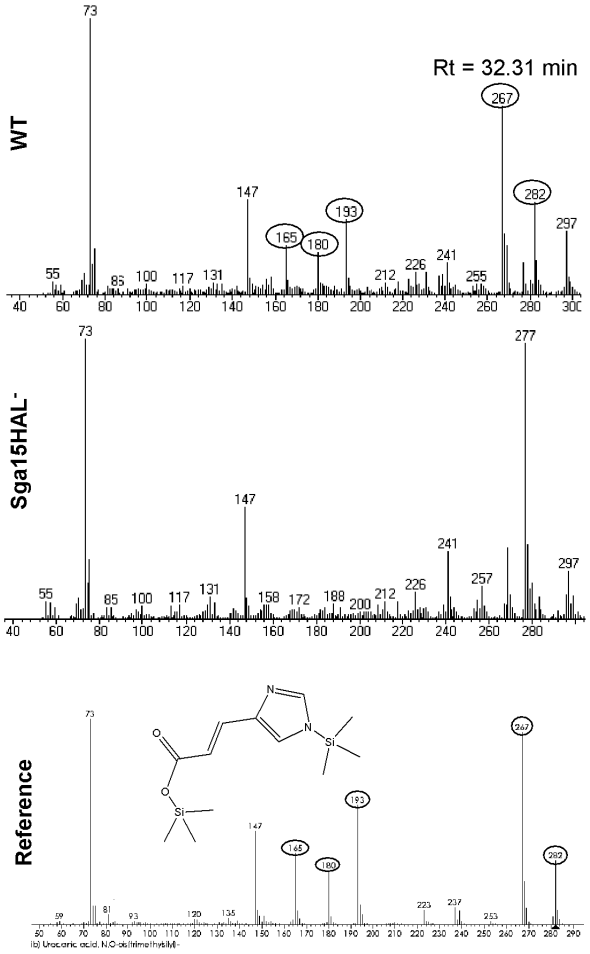
A Gene inactivation by plasmid insertion via homologous recombination. The pTOPO plasmid containing an internal ~800 bp fragment of the *Sa*HAL-coding gene was introduced by electroporation into Sg a15.



B Agarose gel picture: verification of correct integration by multiplex PCR using primers pTOPOin / pTOPOout and primers Sga15con1 / Sga15con2 with chromosomal DNA of two Sga15HAL⁻ clones.

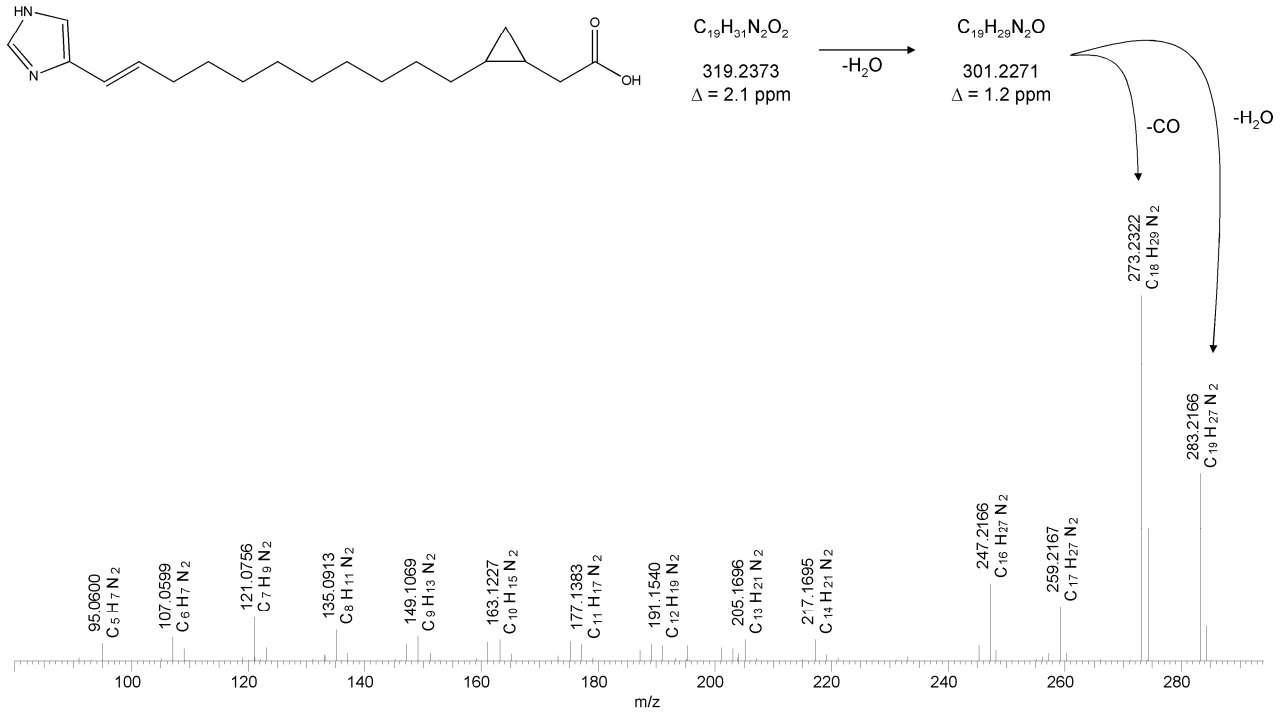


C GCMS analysis revealing the abscence of urocanic acid in supernatants of Sga15HAL⁻ mutant cultures.

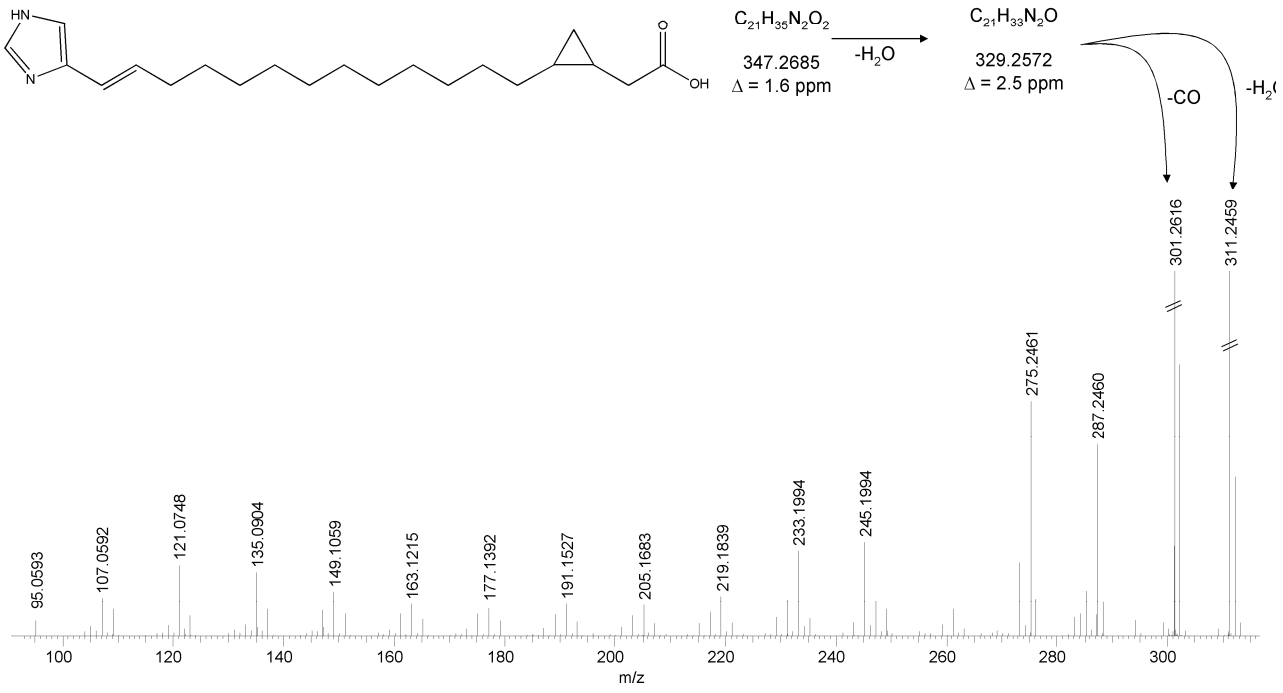


Supplemental Figure S2

A High-resolution MS^3 CID spectrum of imidacin A_1 (c319). All m/z values and formulae are given for $[M+H]^+$ ions.

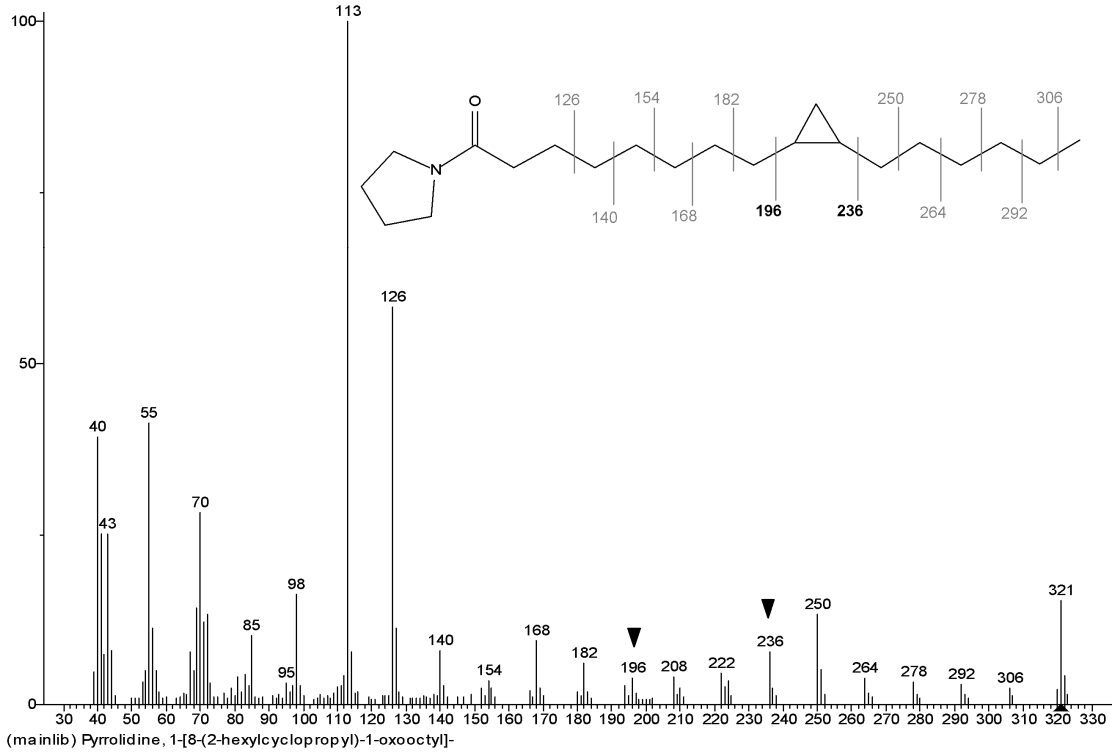


B High-resolution MS^3 CID spectrum of imidacin A_2 (c347). All m/z values and formulae are given for $[M+H]^+$ ions.

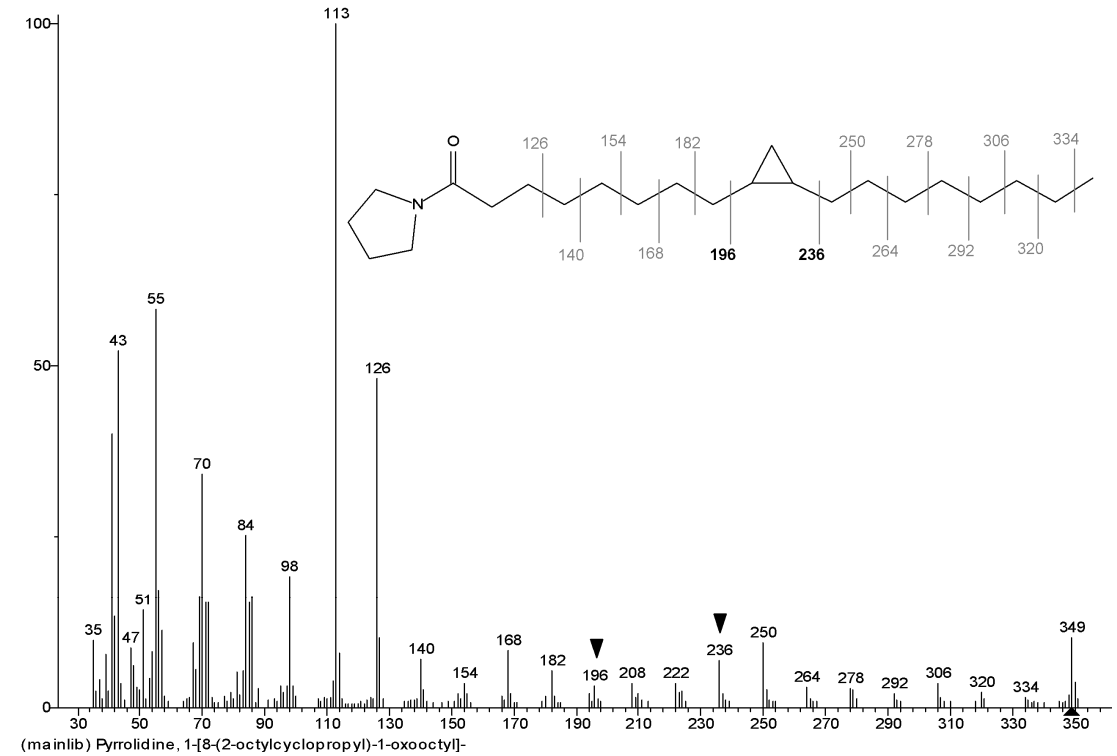


Supplemental Figure S3

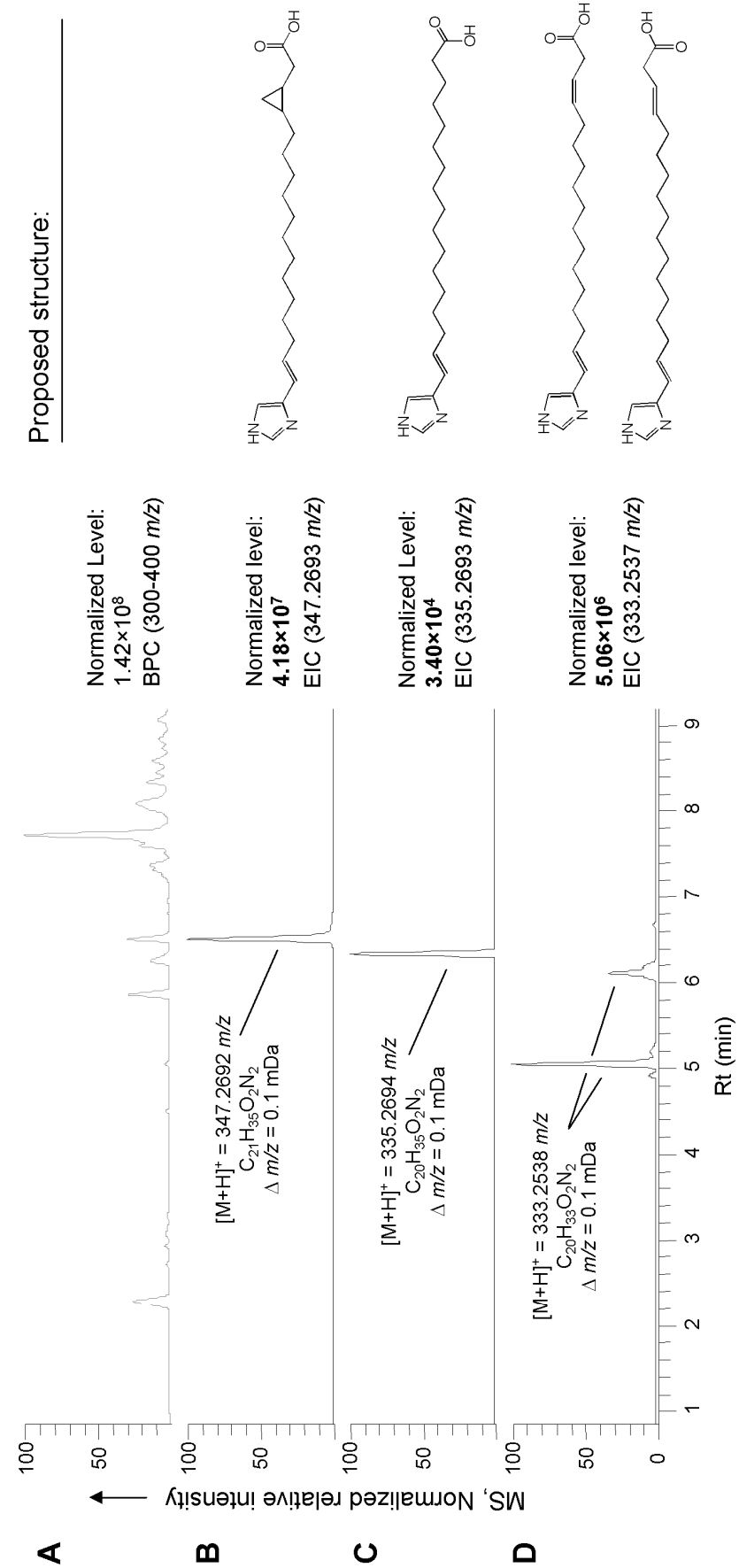
A



B



Supplemental Figure S4



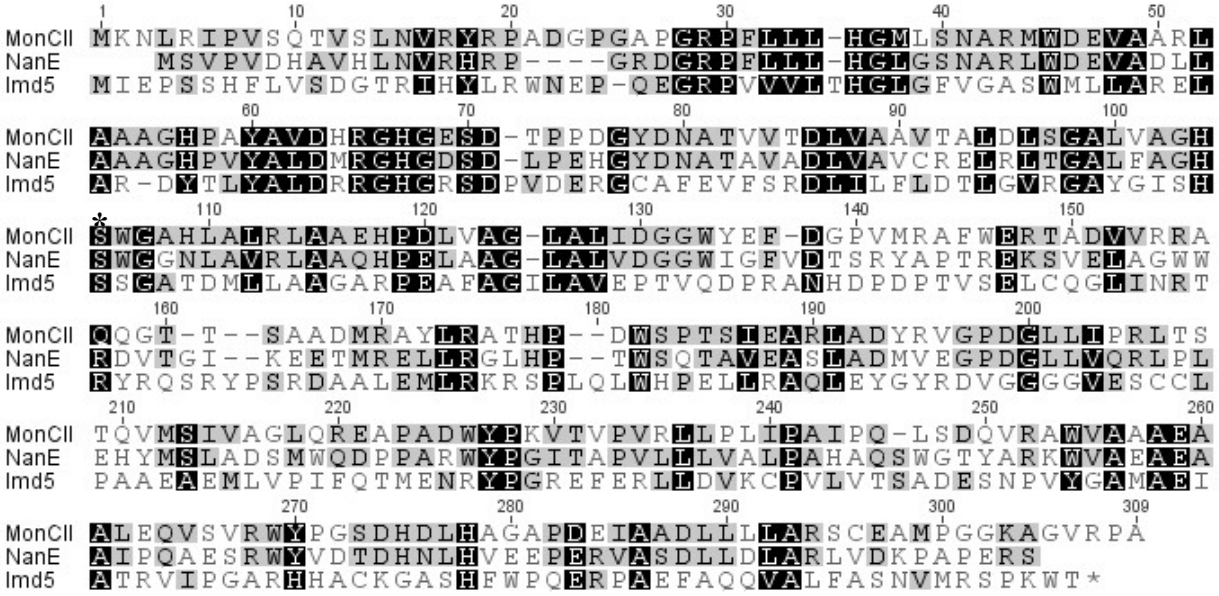
HPLC HR-ESI MS analysis of imidacin A_2 (trace **B**) and ist putative biosynthetic precursors imidacin B_2 (trace **D**) and imidacin C_2 (trace **C**). Two double bond isomers of imidacin B_2 may be present (trace **D**). A basepeak chromatogram (trace **A**) and extracted ion chromatograms (traces **B**, **C**, **D**) are shown.

Supplemental Figure S5

A) Protein sequence of Imd4. Residues on grey background: glycine-rich motif putatively involved in SAM binding; boxed residues: hypothetical artifacts of “radical SAM-like” cysteine motif.

MKPSVSRSARSSVEAARDFIERNLNVRQVNKVLHAFPSPRMWQETDIPVRERLEERKAAQESRE
LALYVAVPYCIRTDPDRCGYCLFPVEVYTGNNDLDTYFTYLRREGELYRGLFEKDRVVNIYFG
GGTSNLYRADKYPQLMDLVREHFTLPTDSSITLEGIPQLFTREKLQKIKESGMNRVSMGVQQMN
DELNALSGRKQTVKHTLQSIEWCQELGLPCNVDLIFGWPRQTVATMMKDLELLVSTGIHHITHY
ELNVGGPTDFALNRRHELPSVAENLEMYRVSRDFLKSRGYVQLTPYDWEKIDAPEQSLYEECDR
DFRAMDIFGWGYAGVSQFAGVGGRSAWTYRQHTSIKDYYKAIDEGRIPLEHGFTHQEVDYRLSQ
LFRNLQGLEVDLDAYQRAFGVELLDEYAGAWQALAERGFVEIEGRKLRLVGDGVFYTPMIQTLL
SRERIAELSSSLRASVRPAASVGV

B) Protein sequence alignment of Imd5 with MonCII and NanE. The asterisk above the sequences denotes the putative active-site serine residue.



Supplemental Table S1

^{1}H-NMR data for imidacin A_2 (c347).

A)

^{1}H NMR (MeOD, 400 MHz): δ = 0.65 – 0.70 (m, 1H, CH_2), 0.85 – 0.91 (m, 1H, CH_2), 1.04 – 1.09 (m, 1H, CH), 1.13 (dd, 1H, $^3J_{H,H}$ = 2.7, 6.2 Hz, CH), 1.25 – 1.42 (m, 18H, 9 x CH_2), 1.46 (quin, 2H, $^3J_{H,H}$ = 6.8 Hz, CH_2), 2.17 (q, 2H, $^3J_{H,H}$ = 7.1 Hz, CH_2), 6.15 (dt, 1H, $^3J_{H,H}$ = 6.9, 15.9 Hz, CH), 6.27 (d, 1H, $^3J_{H,H}$ = 16.0 Hz, CH), 6.94 (s, 1H, CH), 7.66 (s, 1H, CH), 8.45 (s, 1H, NH) ppm.

B)

H,H-COSY

Nr.	ppm	Multiplicity	Integral	$^3J_{H,H}$ / Hz	H,H-COSY[a]	Interpretation
1	0.65 – 0.70	m	1		3, 5	Cyclopropyl-CH_2[b]
2	0.85 – 0.91	m	1		5	Cyclopropyl-CH_2[b]
3	1.04 – 1.09	m	1		1, 5	Cyclopropyl-CH
4	1.13	dd	1	2.7, 6.2	5	Cyclopropyl-CH
5	1.25 – 1.42	m	18		1 – 4, 6	9 x CH_2, Alkyl chain
6	1.46	quin	2	6.8	5, 7	C*H*$_2$-CH_2CH=CH[c]
7	2.17	q	2	7.1	6, 8	CH_2-C*H*$_2$-CH=CH[c]
8	6.15	dt	1	6.9, 15.9	7, 9	CH_2CH_2-C*H*=CH[c]
9	6.27	d	1	16.0	8	CH_2CH_2CH=C*H*[c]
10	6.94	s	1			Imidazole-CH
11	7.66	s	1			Imidazole-CH
12	8.45	s	1			Imidazole-NH

[a] Number of signal according to column 1 in this table, for which a crosspeak can be seen in the H,H-COSY spectrum.

[b] ^{1}H-NMR signal represents one of the diastereotopic protons in the cyclopropyl ring CH_2-moiety.

[c] protons assigned to this ^{1}H-NMR-signal are printed in italics.

Supplemental Table S2

HALSga15start	5'-ATGTCCCAGATCTCCCCGCTTGAAC-3'
HALSga15stop	5'-CTACTGTCGTCGACCGCGTCGCGGTG-3'
HALSga15fwd1	5'-GTGCTGGCCAAGGGCTATTCG-3'
HALSga15rev1	5'-GCAATCATGAAGCCCGAGTTG-3'
Sga15con1	5'-GGTGCTGCTCTTGCTGCGTTG-3'
Sga15con2	5'-AGCTCCCGGTCCTTTTCCATG-3'
Sga15con3	5'-CCGAGGGCTGCGGCATGTTGGTC-3'
Sga15con4	5'-AGCGCCGCGCGGTGCTCAAAGTG-3'
Sga15inakt2f	5'-CACGGTGGCCTGTGGCAGCGCATGG-3'
Sga15inakt2r	5'-CTGCGGCCGCGGTGAAGCTGAGGTC-3'
Sga15inakt3f	5'-CTCGGCAGCGGGCAGGCAGCAAGAC-3'
Sga15inakt3r	5'-GGGCTTCGTGGGTGCCTCGTGGATG-3'
pTOPOin	5'-CCTCTAGATGCATGCTCGAGC-3'
pTOPOout	5'-TTGGTACCGAGCTCGGATCC-3'
Sga15con5	5'-GTAGCGGTTCTCCATGGTCTGGAAG-3'
Sga15con6	5'-TCGGTTGGAGTCTGAGGATCCAC-3'
cos13F7_19F2rev	5'-GTGATGTCCTGCAGTCGCTTG-3'
cos13F7_18F	5'-CCCGAACTTCGCCTATGAG-3'
cos11D1_7Rrev	5'- TCGCGCACCTGGGGATGATC-3'
cos13F7_6R	5'- GTGGCCCACCAATACGAAG-3'
Sga15act3neu2	5'-GATAAGGAGGTACAGATTATGCGAAGATCGACTG-3'
Sga15con4b	5'-CCCACTCGCAGGAGTTGTG-3'
HutU_QMTA	5'-GCCTGATGATGTACGGCCARATGACNGC-3'
HutU_PGFV	5'- CGGATGTACGCCGGCACRAANCCNGG-3'
K388	5'-TGGGAATCATTTGAAGGTTGG-3'
K390	5'-GGGTATCGCTCTTGAAGGGAAC-3'

Chapter B-5

Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*

Berner, M., Krug, D., Bihlmaier, C., Vente, A., Müller, R., and Bechthold, A. (2006)

J. Bacteriol. 188, 2666-2673

This article is available online at:

http://jb.asm.org/cgi/content/abstract/188/7/2666

Chapter B-6

Discovering the hidden secondary metabolome of *Myxococcus xanthus*: a study of intraspecific diversity

Krug, D., Zurek, G., Revermann, O., Vos, M., Velicer, G. J., and Müller, R. (2008)

Appl. Environ. Microbiol. 74, 3058-3068

This article is available online at:
http://aem.asm.org/cgi/content/abstract/74/10/3058

Chapter B-7

Efficient mining of myxobacterial metabolite profiles enabled by liquid-chromatography – electrospray ionization – time-of-flight mass spectrometry and compound-based principal component analysis

Krug, D., Zurek, G., Schneider, B., Garcia, R., and Müller, R. (2008)

Anal. Chim. Acta 624, 97-106

This article is available online at:
http://dx.doi.org/10.1016/j.aca.2008.06.036

Chapter B-8

Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods

Garcia, R. O., Krug, D., and Müller, R. (2009)

In: *Methods in Enzymology* (Editor: D. Hopwood), *in press*.

This article is available online at:

http://dx.doi.org/10.1016/S0076-6879(09)04803-4

C. Discussion

1. General scope of this work

This thesis deals with aspects of secondary metabolite biosynthesis in myxobacteria, with particular focus on the biosynthetic origin of unusual building blocks incorporated into myxobacterial natural products. Detailed molecular and biochemical studies were carried out, in order to delineate the involved metabolic pathways and to explore the underlying enzymology. Genetic engineering was used to establish structure-activity relationships for one selected enzyme from *Chondromyces crocatus*, demonstrating the feasibility to tailor its catalytic properties for potential biotechnological applications. Moreover, targeted gene inactivation and feeding studies led to the discovery and structure elucidation of a unique class of secondary metabolites from *Stigmatella aurantiaca*. Investigating the biosynthetic route for the formation of these compounds revealed a novel branching point from bacterial primary metabolism. Besides the sequestration of an unprecedented starter unit, this new pathway was also shown to include several intriguing enzymatic reactions. Both, the use of unusual building blocks and uncommon enzymatic transformations, contribute to the generally impressive structural diversity of myxobacterial secondary metabolites.

Furthermore, a mass spectrometric method was devised for the characterization of myxobacterial metabolite diversity. The metabolomics-based approach employed here, in combination with statistical evaluation, paves the way for an enhanced, comprehensive view of complex myxobacterial secondary metabolomes. The complexity of the species-wide secondary metabolite inventory of *Myxococcus xanthus* could be shown to match the previously established genetic capacity of this species for natural product biosynthesis. The implementation of advanced mass spectrometric techniques within an analytical framework for the mining of microbial metabolite profiles is described. These methods should facilitate the future discovery of novel natural products not only from myxobacteria.

2. Biosynthesis of β-tyrosine by tyrosine 2,3-aminomutase enzymes and its incorporation into natural products

Biosynthesis of (*R*)-β-tyrosine and its incorporation into the chondramides.

The chondramides produced by *C. crocatus* exhibit striking structural similarity to the jaspamides (jasplakinolides), a family of depsipeptides derived from the marine sponge *Jaspis spledens* (Figure C-1).[93] Based on the assumption that the corresponding biosynthetic pathways should be highly similar, it was suggested that the stereochemistry of the β-tyrosine moiety in the chondramides should match the (*R*)-configuration found in the jaspamides.[44] This expectation was confirmed in this work. Very recently, the configurational assignment of all stereocenters in the chondramide structure was accomplished by NMR analysis, and thus the presence of a (*R*)-β-tyrosine moiety has been independently validated.[94,95]

The biosynthesis of a (*R*)-β-Tyr building block and its mode of incorporation into a natural product have not been investigated previously, but the sequencing of the chondramide biosynthetic gene cluster afforded an opportunity to probe the biosynthetic origin of this unusual β-amino acid.[44] It was hypothesized that (*R*)-β-Tyr could be produced directly from L-tyrosine by CmdF, an enzyme encoded in the *cmd* gene cluster showing significant homology to the tyrosine aminomutase SgcC4 producing (S)-ß-Tyr in the *S. globisporus* enediyne pathway. CmdF was shown by targeted gene inactivation to be essential for chondramide biosynthesis. Complementation of the *cmdF*⁻ mutant could be achieved through feeding of (*R*)-β-Tyr, but not (*S*)-β-Tyr.

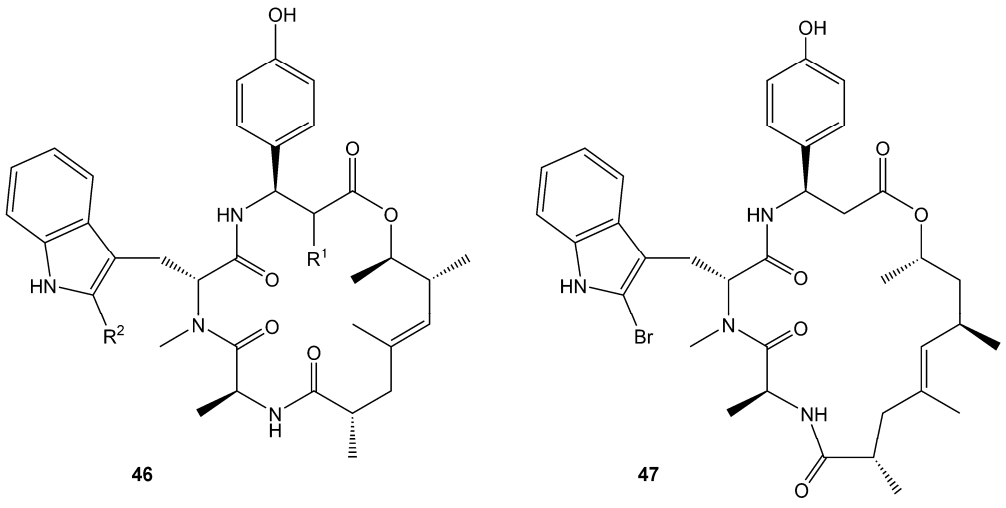


Figure C-1. Structures of chondramides **46** from the myxobacterium *C. crocatus* Cm c5 and jaspamide **47** from the marine sponge *Jaspis spledens*. Chondramide A, R^1= -OCH_3, R^2 = H; chondramide B, R^1 = -OCH_3, R^2 = Cl; chondramide C, R^1 = R^2 = H; chondramide D, R^1= H, R^2 = Cl.

On the contrary, the tyrosine aminomutase SgcC4 (the only characterized TAM enzyme at the onset of this study) has been shown to produce preferentially (*S*)-β-Tyr *in vitro*, which is found in the enediyne product C-1027. This difference prompted further investigation of the catalytic properties of CmdF, and therefore heterologous expression was carried out. The biochemical characterization of recombinant CmdF revealed this enzyme as a novel member of the TAM family, converting L-tyrosine predominantly to (*R*)-β-Tyr. CmdF thus exhibits stereoselectivity consistent with its function to supply the (*R*)-β-Tyr building block for chondramide formation in *C. crocatus*. However, CmdF also produced significant amounts of the opposite enantiomer *in vitro*, but (*S*)-β-Tyr is not incorporated into the chondramides at any detectable level (chapter B-1). The same stereochemical ambiguity was observed *in vivo* when CmdF was heterologously expressed in *M. xanthus*, and the newly identified aminomutase enzyme from *Cv. metallidurans,* CmTAL, exhibited similarly imperfect stereoselectivity (chapter B-2). Moreover, SgcC4 has been reported to produce minor amounts of (*R*)-β-Tyr, which is not incorporated into the enediynes.[78] Thus, these tyrosine aminomutases evidently fail to create an enantiopure intracellular β-amino acid pool, and therefore "gatekeeping" mechanisms must be in place to ensure stringent stereochemical control of β-tyrosine incorporation into the secondary metabolites from *C. crocatus* and *S. globisporus*.

In NRPS biosynthetic pathways, the adenylation (A) domains are responsible for the selection and activation of amino acid building blocks. The CmdD-A_7 domain responsible for β-tyrosine incorporation during chondramide biosynthesis (Figure C-2 **A**) was shown *in vitro* to activate primarily (*R*)-β-Tyr, a result that agrees well with the stereoselectivity exhibited by CmdF and the (*R*)-configuration of the β-tyrosine moiety in the chondramides.[44] Activity of CmdD-A_7 towards L-tyrosine and (*S*)-β-Tyr was at considerably lower levels. Similarly, the SgcC1 A domain which catalyzes the incorporation of β-tyrosine during enediyne biosynthesis in *S. globisporus* (Figure C-2 **B**) exhibits strong kinetic preference for the activation of (*S*)-β-Tyr. The observed activation of the opposite amino acid enantiomers is probably further substrate-limited *in vivo*. These data argue against an alternative biosynthetic model in which the A domains activate L-α-tyrosine, which would then be isomerized to the β-form while tethered to a PCP domain on the respective assembly line. Nevertheless, the finding that not even traces of the disfavoured enantiomers are found in chondramides and enediynes is astonishing. Thus, additional gatekeeping likely occurs in the later stages of the biosynthesis, preventing product formation to proceed with incorporation of the wrong enantiomer.

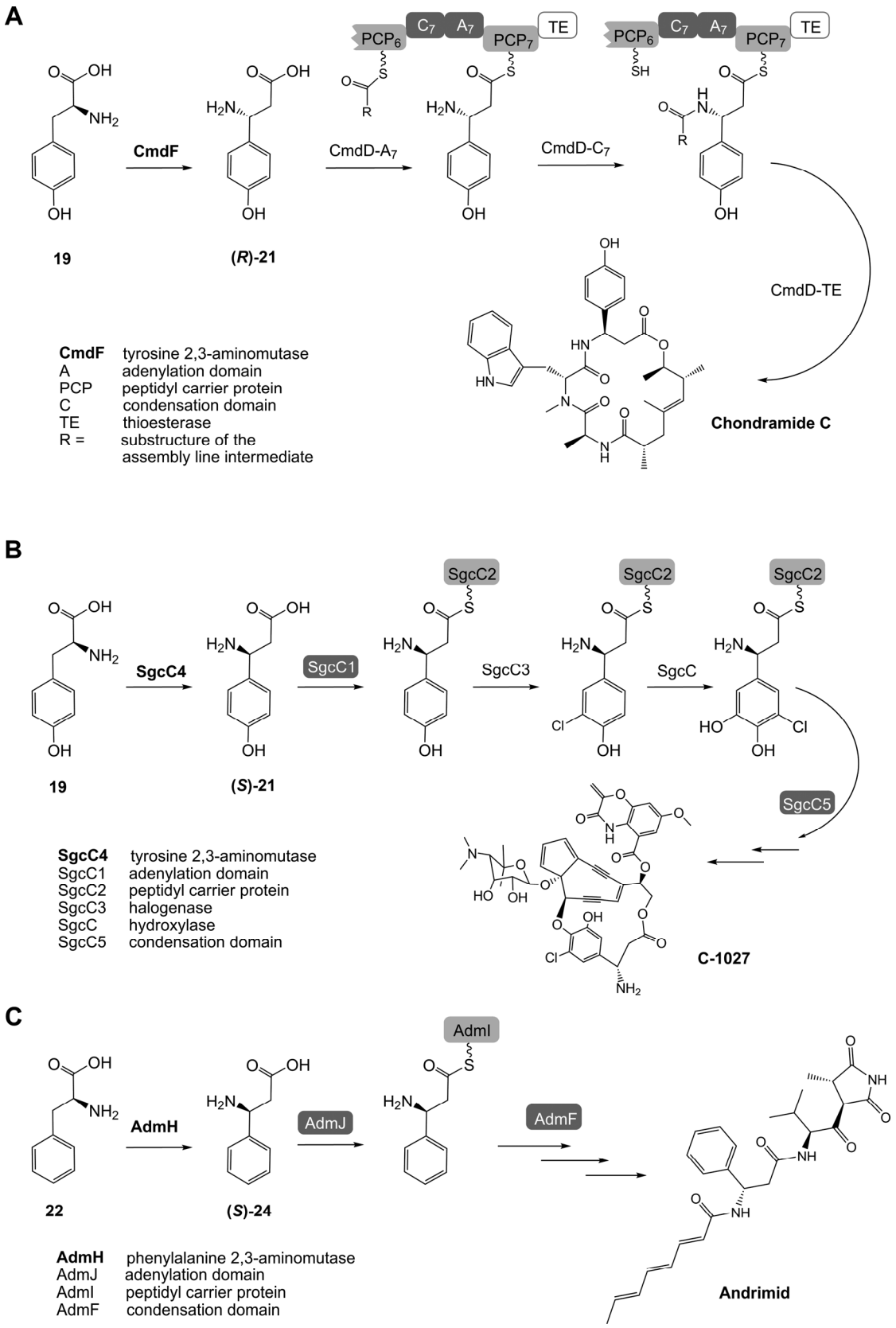


Figure C-2. Incorporation of β-amino acid building blocks during biosynthesis of chondramides (**A**), the enediyne compound C-1027 (**B**) and andrimid (**C**).

β-tyrosine incorporation is the last chain extension step in chondramide biosynthesis before the intermediate is released from the assembly line by lactonization. Besides the CmdD-A_7 domain, three additional functionalities which could exert stereochemical control are present in the terminal CmdD module (Figure C-2 **A**). Following activation as the β-aminoacyl adenylate by CmdD-A_7, the transfer of the β-tyrosine monomer onto the integral carrier domain (CmdD-PCP_7) may be subject to stereochemical discrimination, possibly due to spatial constraints. Should the "incorrect" β-amino acid enantiomer nevertheless become attached to PCP_7, then the condensation domain CmdD-C_7 could act as a gatekeeper, if it catalyzes the elongation of the peptidyl intermediate only with the correct enantiomer loaded on the downstream PCP_7. Several studies have revealed the ability of C domains to exert such stereochemical control.[96,97] Finally, the thioesterase (CmdD-TE) might represent a last instance of chiral selectivity, by performing cyclization only if an elongated intermediate is presented to it on PCP_7 exhibiting the appropriate configuration at the β-amino stereocenter. Further attempts to clarify the detailed mechanism of stereochemical control in the late stage of chondramide biosynthesis would greatly benefit from the heterologous expression of combinations of the underlying protein domains and, ideally, the PCP_6 to TE portion of CmdD (even though the purification of biosynthetic proteins of that size is challenging). Synthetic intermediates and activated β-amino acid precursors with differing chirality could then be enzymatically installed *in-vitro* on the recombinant partial CmdD module,[98] in order to probe its processing capability.

The above considerations could similarly apply to the biosynthesis of the (*S*)-β-Tyr-containing enediyne product C-1027, although the catalytic domains in the enediyne pathway are encoded on discrete proteins.[99] Additional enzymes are involved in hydroxylation and chlorination of (*S*)-β-Tyr while it is bound to the carrier protein SgcC2 (Figure C-2 **B**), but at least the halogenase SgcC3 has been shown to accept both β-Tyr enantiomers.[70,100] In contrast to the chondramide and enediyne biosynthetic routes, rigorous chiral gatekeeping seems to occur in the early stages of andrimid biosynthesis (Figure C-2 **C**).[67] In this pathway, an (*S*)-β-Phe building block is supplied by a phenylalanine ammonia lyase (AdmH), which does not produce detectable amounts of the opposite enantiomer. In addition, the A domain responsible for (*S*)-β-Phe selection (AdmJ) does not activate (*R*)-β-Phe to any appreciable degree. This two-tiered mechanism of stereochemical control apparently compensates the lack of stereselectivity in subsequent biosynthetic steps – indeed, the condensing enzyme AdmF involved in andrimid biosynthesis has been demonstrated to accept both β-phenylalanine enantiomers.[67]

β-amino acids are building blocks in a number of natural products, and their incorporation into biologically active peptides has been associated with increased stability toward proteolytic degradation.[53,101,102] Moreover, the β-amino acid moiety likely constitutes an important structural feature, reflecting the evolutionary adaption of a compound to a specific biological target. Therefore, β-amino acid-incorporating A domains represent interesting tools for the construction of artificial assembly lines, capable of producing "designed" secondary metabolites. However, problems (in particular a reduced yield) are sometimes encountered when domains are excised from one biosynthetic multienzyme and inserted into a non-cognate assembly line.[103] To circumvent such difficulties, the genetic re-engineering of adenylation domains to alter substrate specificity has been proposed. Analysis of the substrate binding pockets of A domains has resulted in the identification of eight residues which can be used to predict the activation specificity. The feasibility of manipulating this "non-ribosomal code" through single amino acid mutations, in order to achieve activation of different amino acids, has been demonstrated.[104]

It has been pointed out, however, that the current predictive model, which is based essentially on the crystal structure of the L-α-Phe activating domain PheA from the gramicidin pathway, is perhaps not applicable to β-amino acid-activating A domains. While the comparison of non-ribosomal codes (Table C-1) reveals slight similarity between certain L-α-Tyr and β-Tyr activating A domains, considerable divergence is observed between the β-amino acid activating domains from different pathways. Interpretation is further hindered by the hitherto small number of characterized A domains with β-amino acid specificity. Significantly, it may not even be correct to assume that the same conserved key residues highlighted by the existing model interact with the substrate in β-amino acid-specific A domains, given the fundamental structural difference between α- and β-amino acids. Thus, the crystal structure of an β-amino acid activating A domain is eagerly awaited and is expected to shed light on the molecular basis for selective acti-

Table C-1. Key residues which determine the substrate specificity of A-domains that activate L-Phe or L-Tyr, and residues found at the corresponding positions in β-amino acid activating A domains. The domains originate from the following pathways: PheA, gramicidin; tycC, tyrocidine; BacC, bacitracin; SgcC1, enediyne C-1027; CmdD-A_7, chondramide; AdmJ, andrimid. Numbering relates to PheA.

Domain	Substrate	Key residues									
		235	*236*	*239*	*278*	*299*	*301*	*322*	*330*	*331*	*517*
PheA	L-α-Phe	D	A	W	T	I	A	A	V	C	K
TycC	L-α-Tyr	D	A	S	T	V	A	A	V	C	K
BacC	L-α-Tyr	D	G	T	L	T	A	E	V	A	K
SgcC1	*(S)*-β-Tyr	D	P	A	Q	L	M	L	I	A	K
MdpC1	*(S)*-β-Tyr	D	P	C	Q	V	M	V	I	A	K
CmdD-A_7	***(R)*-β-Tyr**	**D**	**G**	**S**	**T**	**I**	**T**	**A**	**V**	**A**	**K**
AdmJ	*(S)*-β-Phe	D	M	L	S	A	K	A	M	C	K

vation of β-amino acids. The establishment of a non-ribosomal code for β-amino acid activating A domains is a crucial prerequisite for predicting reliably the incorporation of β-amino acids by biosynthetic routes revealed through genome mining. Furthermore, this knowledge could facilitate future attempts to rationally modify the substrate sepcificity of A domains, with the aim to create novel β-amino acid-incorporating NPRS assembly lines.

Discovery of novel aminomutases and mutational analysis of CmdF.

The production of β-Tyr and β-Phe via aminomutase-catalyzed conversion of the corresponding α-amino acids acids, followed by incorporation according to NRPS biosynthetic logic, appears to be a common strategy for the formation of natural products which contain aromatic β-amino acids.[67] Thus, the "function-guided" discovery of new aminomutase enzymes from strains producing such secondary metabolites should be anticipated. In the course of this work, additional tyrosine aminomutases were cloned from two myxovalargin-producing *Myxococcus* strains and from one strain of *Cupriavidus metallidurans*, thereby expanding the formerly small TAM enzyme family.[45] These new data were expected to provide a stronger basis for sequence comparison and to pave the way for attempts to manipulate the properties of CmdF by site-directed mutagenesis. Biochemical characterization revealed that the *Myxococcus* enzymes produced (*S*)-β-Tyr, consistent with their putative role in supplying an (*S*)-β-Tyr building block for myxovalargin biosynthesis in these strains. The *Cupriavidus* enzyme (CmTAL) was shown to form preferentially (*R*)-β-Tyr, but CmTAL was in fact more efficient as a TAL than a TAM. Notably, the secondary TAM activity exhibited by CmTAL distinguishes this enzyme from other known TALs. For example, the TAL from *Saccharothrix espanaensis*, Sam8, did not show any aminomutase activity.[73]

The preferential production of opposite β-Tyr enantiomers constitutes a striking difference between members of the TAM enzyme family. Very recently, a structural study on SgcC4 has established the mechanistic basis for aminomutase catalysis.[77] It has been proposed that stereochemical differentiation results from the attack of an MIO-amine adduct on opposite faces of a planar pHCA intermediate, and crystallographic data have confirmed that the active site of SgcC4 can accommodate different substrate orientations. On the basis of this model, the stereochemical outcome of the aminomutase reaction should be determined by amino acid residues in proximity to the active site, which help to orient the substrate appropriately. However, residues which could potentially play such roles are not obvious from the existing structural and biochemical data.

Despite the absence of structural data for the myxobacterial aminomutase CmdF (crystallisation attempts in the course of this study in collaboration with the group of Prof. Schulz, Freiburg, yielded diffraction data with insufficient resolution), efforts were made to modulate its catalytic properties by site-directed mutagenesis. These experiments were inspired by the recent report of a single active site residue which is able to act like a substrate selectivity switch, converting TAL from *Rhodobacter* into an efficient PAL.[66] A number of CmdF variants were generated, in which certain conserved motifs found in ammonia lyase or aminomutase enzymes were mutated in the corresponding positions in CmdF. It was expected that the mutant CmdF enzymes might exhibit altered product specificity (i.e. ammonia lyase versus aminomutase activity) or modified substrate selectivity, or display a stereoselectivity shift (chapter B-2). Unfortunately, these attempts were largely unsuccessful – in fact, most experiments only allowed the general conclusion that the aminomutase scaffold does not readily tolerate the introduction of conserved motifs from HAL, PAL, TAL or TAM protein sequences. However, the variation of one semiconserved amino acid, Glu399, was found to impact stereoselectivity: Replacement of this residue by lysine ($CmdF_{E399K}$) increased the enantiomeric excess of (*R*)-β-Tyr from 69 to 97% ee, while substitution with methionine ($CmdF_{E399M}$) promoted racemization (chapter B-2).

Prospects for the genetic manipulation of stereochemical control in TAMs.

In the following, a possible role of Glu399 in the aminomutase mechanism is illustrated with the aid of a homology model for CmdF, which was calculated using the published SgcC4 structure as a template (Figure C-3 **A**).[105] Besides the crystal structure of SgcC4,[76] structural data also exist for a number of ammonia lyase enzymes.[79,80,106] These crystallographic studies have pictured the common aminomutase/ammonia lyase scaffold as a largely α-helical bundle (Figure C-3 **A**), and the active form has been described as a homotetramer (a head-to-tail dimer of dimers; Figure C-3 **B**, **C**). Importantly, the MIO-containing active site has been shown to be framed by residues from three of the four subunits in the enzyme complex. Certain key residues within interacting distance to the amino acid substrate in the binding pocket have been assigned specific functions for aminomutase and/or ammonium lyase catalysis. His83 is the already mentioned TAL/PAL "substrate selectivity switch" which has been proposed to direct substrate recognition in TAL and TAM; structural data for SgcC4 suggest that this residue and Tyr403 (from a neighboring protein chain) form a bifurcated hydrogen bond with the tyrosine ring-hydroxyl (see also Fig. A-11).[77] Intriguingly, the residue Glu399, which was identified in this thesis as a "stereoselectivity modulator" in CmdF, is located on the Tyr403-neighboring helix turn (Figure

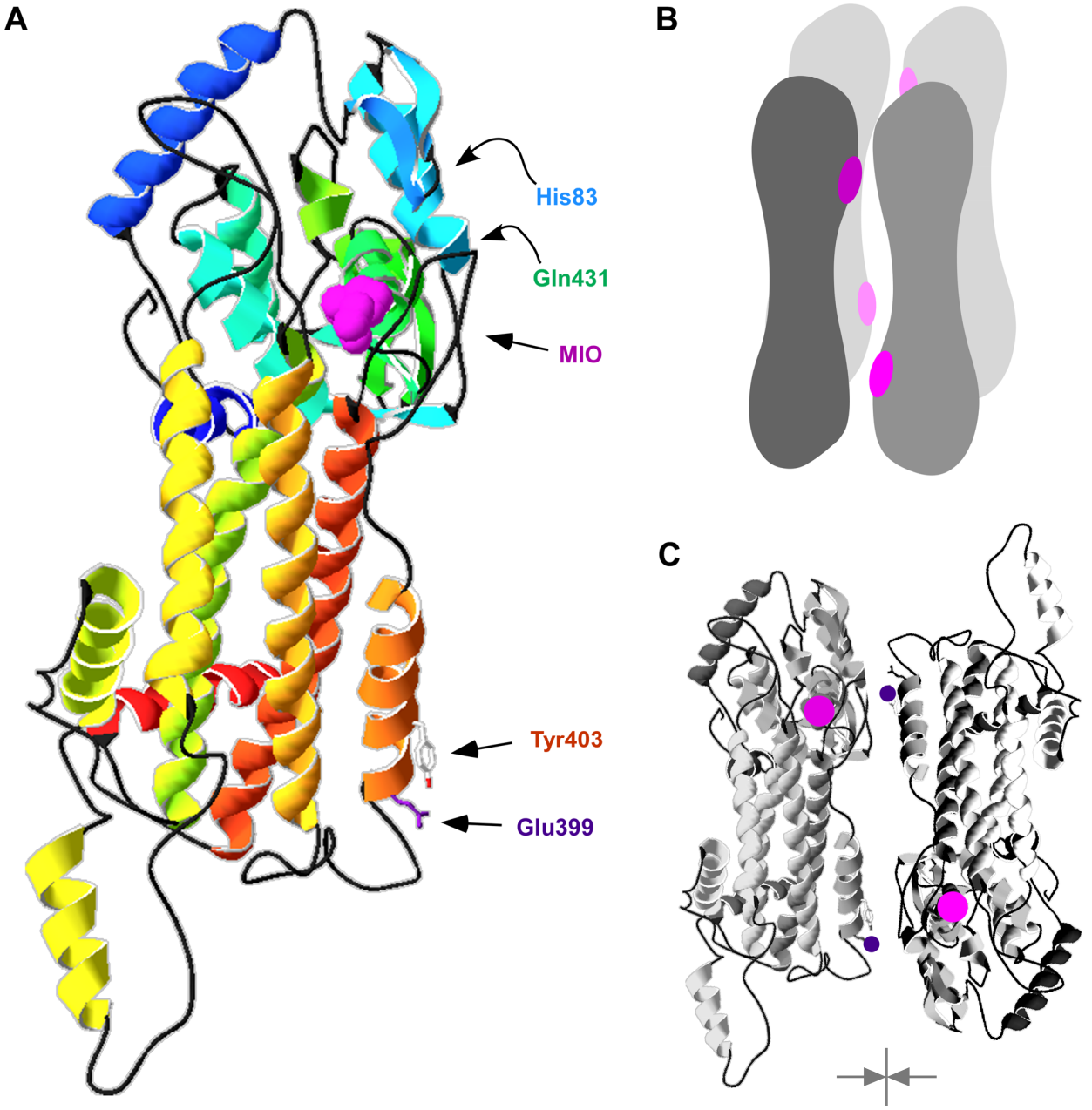


Figure C-3. **A**, homology model for CmdF, using the SgcC4 structure (PDB: 2OHY) as template for SWISS-MODEL. The active site MIO is drawn in the space-filling representation (pink). **B**, arrangement of subunits in the active-form tetramer; **C**, positioning of dimers.

C-3 **A**). Thus, replacing Glu399 with amino acids that exhibit significantly altered chemical properties and/or steric requirements of their side chains, is likely to have an impact on the electronic and spatial characteristics of the binding pocket – either directly, or indirectly through repositioning of the nearby Tyr403. This influence could induce a change in substrate orientation, which would then be reflected by a modified stereochemical course of the aminomutase reac-

tion. A crystallographic investigation of CmdF and its variants with altered stereoselectivity, bound to the inhibitor used in the SgcC4 study, should help to reveal further details and additional stereochemical determinants.

It must be pointed out that Glu399 is not strictly conserved in the aminomutases, and the amino acid at this position does not correlate with the stereoselectivity exhibited by different TAMs. Several independent evolutionary solutions might have been established in Nature, regarding the key residues that modulate the structural framework of the active site of TAM enzymes (see below). Indeed, a comparison of the amino acid sequences of all TAMs characterized to date did not reveal sequence motifs which would allow to predict the stereochemical preference by sequence analysis. Therefore, sequence-swapping or directed evolution techniques should be employed in future studies on the molecular basis for of stereoselectivity in the TAM enzyme family.

Nonetheless, the finding that a single residue modulates stereocontrol in CmdF is encouraging with respect to possible applications for TAMs, because it indicates the feasibility to tailor their catalytic properties and thereby increase their potential as highly stereoselective biocatalysts. For example, $CmdF_{E399K}$ could be used for the supply of an (*R*)-β-Tyr building block in an engineered NRPS pathway. Implementing stringent stereochemical control in such artificial biosynthetic routes could prove difficult, and an enantiopure pool of the β-amino acid precursor might thus be desirable. Another possible application could make use of the "reverse" aminomutase reaction to produce (*R*)-β-tyrosine with high enantiomeric excess, starting from coumaric acid, the usual product of TAL side activity exhibited by the TAMs. Ideally, ring-hydroxylated β-Tyr derivatives would also be accessible, using various aromatic acrylic acids as substrates.[107] However, this biotransformation would definitely require thorough optimization of the catalyst to achieve significantly higher turnover numbers, in order to become practically useful. The enzymatic synthesis of enantiopure α- and β-amino acids using recombinant phenylalanine aminomutase (PAM) from *Taxus chinensis* has been reported very recently.[108] In that study, it has been highlighted that the availability of a highly enantioselective aminomutase enzyme was a crucial prerequisite to obtain products with satisfactory chiral purity. Interestingly, turnover with ring-hydroxylated substrates could not be achieved using PAM – the potential of TAM to fill this gap remains to be explored.

The tyrosine 2,3-aminomutase enzyme family: evolutionary implications.

In this work, the TAM enzyme family was extended through the biochemical characterization of candidate enzymes from prokaryotic sources, and the increased availability of sequence information permitted for the first time a phylogenetic analysis. The TAM family now comprises six members, which share the same strict substrate specificity for L-tyrosine, but differ significantly with respect to product stereochemistry and relative preference for aminomutase versus ammonia lyase activity. An amino acid alignment was compiled which includes all bacterial PAL, TAL and TAM proteins characterized to date (chapter B-2), plus a selection of enzymes with confirmed HAL activity (albeit a much higher number of HAL homologs exist in publicly accessible databases). Figure C-4 illustrates the outcome of this comparison in the form of an unrooted similarity tree.

Two conclusions may be drawn from this picture. Firstly, the TAM enzymes are located in a separate node of the tree, and the branching pattern indicates greater similarity to HALs than to TALs and PALs. It has been speculated previously that tyrosine aminomutases may have emerged through the evolutionary conversion of a TAL towards tyrosine aminomutase activity,

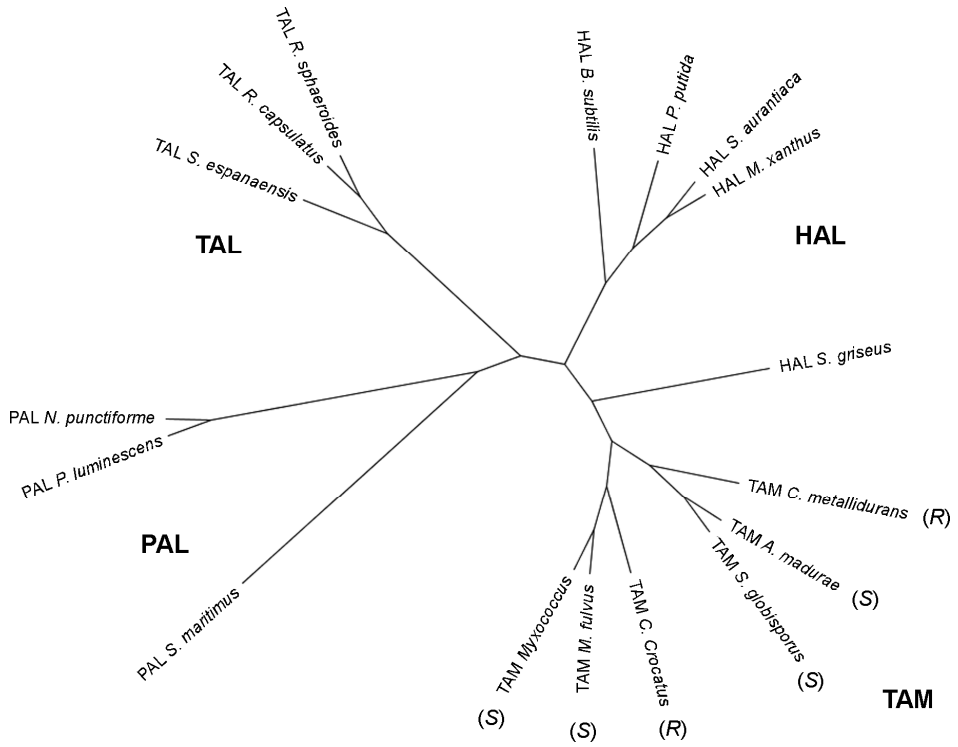


Figure C-4. Phylogenetic comparison of prokaryotic ammonium lyase and aminomutase protein sequences, represented by an unrooted tree diagram.

which would also offer a convenient explanation for the significant ammonia lyase side activity exhibited by aminomutases. Contrary to this expectation, the evidence presented here strongly suggests that TAMs do not constitute a side-branch of the TAL group, but may instead have evolved from a "HAL-like" ancestor. This proposal also appears plausible with regard to the widely spread occurence of HAL enzymes in bacteria, which are in fact present in most species, while prokaryotic TALs and PALs are extremely rare. Speculations have actually been raised about bacterial PALs and TALs being themselves derived from HAL, considering the immense phylogenetic distance to their PAL and TAL counterparts of plant origin.[109]

Secondly, no clustering according to stereochemical preference for the production of (*R*)- or (*S*)-β-tyrosine occurs within the TAM family. The observed grouping instead reflects the underlying phylogenetic relationships. Notably, the preferred TAM stereoselectivity matches the configuration of a chiral building block found in a secondary metabolite from the respective strain, in all instances where such a natural product has been identified to date. In conclusion, the catalytic properties of each individual TAM enzyme may have evolved in order to furnish a specific pathway with a β-tyrosine building block. Thus, convergent evolution might have taken place: characteristic traits of the TAMs could have evolved independently in distant phylogenetic clades. This notion implies that specific amino acid substitutions, which bring about this high degree of evolutionary diversification within the TAM family, are not necessarily conserved in proteins exhibiting the same phenotypic characteristics (i.e., the same stereochemical preference). The finding that the known TAM sequences are polyphyletic may therefore explain the above-mentioned difficulties to predict signature motifs correlating with the stereoselectivity of TAM enzymes from different sources.

Aminomutase genes are frequently encountered as integral parts or in the immediate vicinity of gene clusters encoding β-amino acid-incorporating pathways (e.g. chondramide, enediynes, and andrimid biosynthesis). This observation indicates the possibility to identify such biosynthetic routes by searching for genes potentially encoding aminomutases – an approach greatly facilitated by (but also practicable without) the availability of whole-genome data or metagenomic sequence information. In this regard, it is worth mentioning that the biosynthetic pathways for myxovalargin production in *Myxococcus* and jaspamide formation in *Jaspis* (or in a hypothetical bacterial symbiont of this sponge) have yet to be discovered.

3. Structure and biosynthesis of imidacins – novel secondary metabolites from *Stigmatella aurantiaca* Sg a15 featuring an unique urocanate-moiety

The biosynthesis of microbial secondary metabolites is frequently investigated by a "retro-biosynthetic" approach: based on a known structure, genes putatively involved in the underlying biosynthetic route are first predicted and then searched for in the genome of a producer strain, using molecular biological techniques.[33] The increasing availability of genomic sequence information has further simplified this procedure, facilitating in many cases the *in-silico* correlation of gene clusters to a given compound. Whole-genome data has also enabled the reverse "from genes to products" approach, termed gemome mining: bioinformatic methods are initially used to predict biosynthetic assembly lines as well as the approximate structures potentially produced by these pathways.[110] An attempt is subsequently made to identify the corresponding compounds in the metabolite inventory of a strain under investigation. Such approaches can only be successfully applied if the genetic information permits reliable forecasts about the monomers incorporated by a pathway as well as the type and order of all involved enzymatic transformations. Consequently, the above procedures are likely to fail when analyzing what is regarded as non-textbook biosynthetic routes – i.e., pathways exhibiting a non-contiguous genetic organization or encoding superfluous functions. Moreover, biosynthetic routes incorporating unprecedented building blocks or employing unfamiliar biochemical transformations cannot be adequately predicted. Importantly, such pathways seem to rather represent the rule then the exception and are detected with high frequency in bacterial genomes.

In this work, an alternative approach, best described as "from precursors to products", lead to the discovery of novel secondary metabolites from *S. aurantiaca* Sg a15 (chapter B-4). Following inactivation of the general histidine utilization pathway in Sg a15, comparative metabolite profiling highlighted the existence of previously unrecognized compounds, later named imidacins, the production of which depends on the activity of histidine ammonia lyase. Additional targeted gene inactivation experiments and the screening of a transposon mutant library, combined with feeding studies, enabled the identification of further genes involved in imidacin formation and allowed to eventually deduce a biosynthetic model (Figure C-5). The obtained genetic information was also used to engineer *S. aurantiaca* for increased imidacin production – this manipulation was actually a crucial factor for the structural elucidation of the imidacins, since they were produced only in minute quantities by the wildtype strain. Genes encoding functions for imidacin biosynthesis were found distributed over distant locations in the genome of

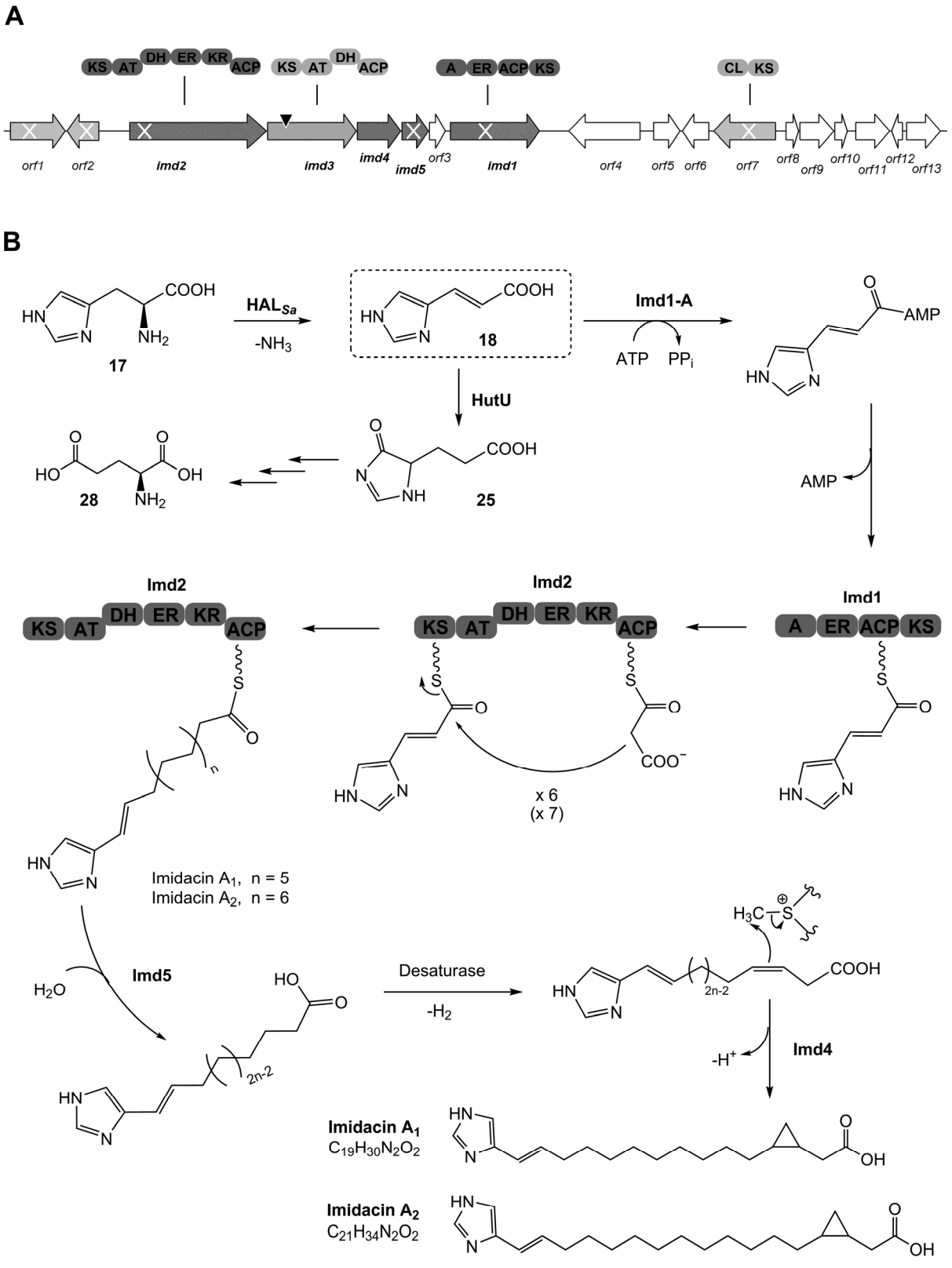


Figure C-5. **A**, Genes involved in imidacin biosynthesis (dark grey). White crosses denote targeted gene inactivations. The reverse black triangle highlight the transposon insertion site. CL, CoA ligase. **B**, The proposed biosynthetic route for imidacin formation, starting from a novel branching point in bacterial histidine metabolism. HAL_{Sa}, histidine ammonia-lyase; HutU, urocanase; Imd4, proposed cyclopropane synthase; Imd5, putative hydrolase. The actual double bond geometry of the precursor processed by Imd4 is not known.

S. aurantiaca. Imidacin formation may thus be seen as the result of an interplay between diverse biosynthetic functions which are otherwise associated with unrelated metabolic pathways. Several uncommon features of this intriguing biosynthetic route are pointed out in the following.

According to search results from publicly accessible databases, the imidacins represent the only natural products known to date to incorporate an imidazole-acryl moiety. This unique substructure was shown in this study to derive directly from urocanic acid, the product of HAL-catalyzed conversion of L-histidine. In fact, the identification of urocanate as the starter unit for imidacin biosynthesis revealed an unprecedented branching point from the prokaryotic HAL-dependent histidine degradation pathway into secondary metabolism in *S. aurantiaca* Sg a15 (Figure C-5, **B**). Future studies, using a combined gene inactivation and metabolite profiling approach as implemented here, should help to clarify whether more urocanate-derived metabolites exist in myxobacteria as well as in other bacterial groups.

For a number of secondary metabolite biosynthesis routes, the recruitment of monomeric precursors such as carboxylic acids from primary metabolism has been demonstrated to proceed via the activation as acyl-adenylates by AMP-ligases, or as CoA-esters by CoA-ligase enzymes.[111] Within the imidacin operon, *imd1* was found to encode a putative AMP-ligase domain (Figure C-5), and module Imd1 was subsequently shown by targeted inactivation to be essential for imidacin formation. The hypothetical involvement of a CoA-ligase encoded by a gene located in close proximity was disproven by targeted inactivation. Imidacin biosynthesis is thus proposed to start with the activation of urocanic acid as urocanyl-AMP by Imd1-A, probably followed by transfer onto the carrier protein domain Imd1-ACP. This mode of starter unit incorporation should be further substantiated through the direct analysis of the loading of urocanate onto an excised Imd1-ACP by recombinant Imd1-A *in-vitro*. High-resolution mass spectrometry could then be used to detect the resulting urocanyl-S-ACP. The mass spectrometric interrogation of assembly line intermediates bound as thioesters to stand-alone carrier proteins, or tethered to integral carrier domains on intact multi-domain enzymes, has been reported recently for a wide range of biosynthetic building blocks.[112] If the complete Imd1 module could be heterologously expressed and purified, it would be possible to additionally investigate the uncertain role of the apparently superfluous Imd1-ER domain by mutation of active-site residues.

Following activation of the urocanic acid starter unit, the linear scaffold of the imidacins is proposed to derive from several rounds of chain extension according to PKS biosynthetic logic (chapter B-4). Two imidacin derivatives with chain length differences corresponding to one C_2H_4- unit were found, suggesting that imidacin biosynthesis involves an iterative mechanism for carbon chain elongation lacking perfect chain length control. Only one module encoded

within the *imd* operon, Imd2, has the full complement of domains necessary to catalyze chain extension plus subsequent ketoreduction, dehydration and enoylreduction (Figure C-5). Additionally, *in silico* analysis indicated that the integral Imd2-AT domain likely exhibits substrate specificity for malonyl-CoA. Thus, iterative use of module Imd2 could afford the linear, completely reduced imidacin backbone after six or seven cycles of chain extension and β-keto processing. The iterative use of a bacterial multifunctional type I PKS module has been demonstrated first in the biosynthesis of orsellinic acid by *S. viridochromogenes*.[113] Subsequently, the itera-

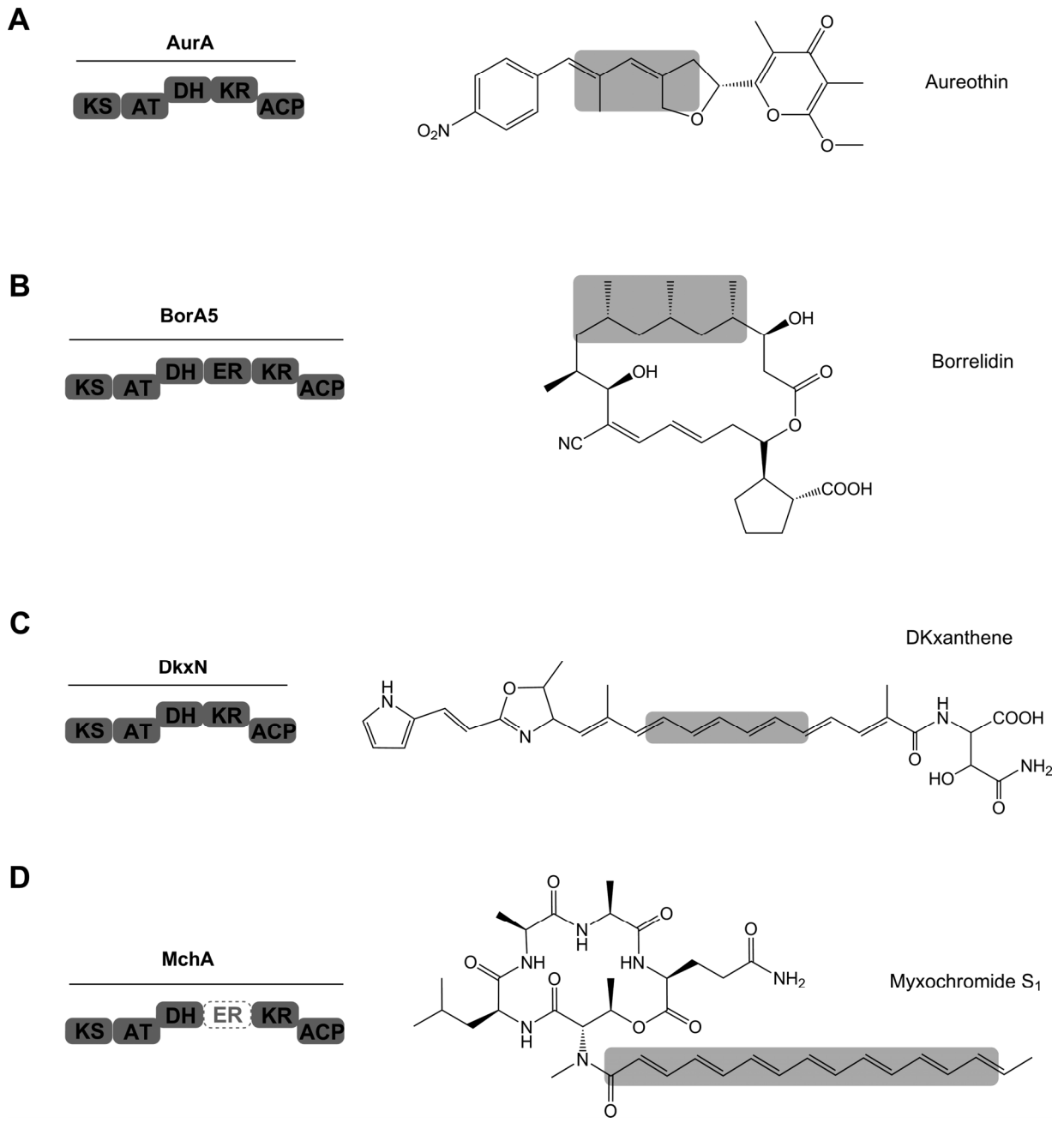


Figure C-6. Secondary metabolites derived in part from iterative type I PKS biosynthetic logic, and functional organization of the modules proposed to act iteratively in the respective pathways.

tive action of a type I PKS system, which usually carries out only one round of catalysis per module, has been observed as an aberrant function (termed "stuttering") in erythromycin biosynthesis.[114] In another early study, iterative type I PKS biochemistry has been shown to be involved in the biosynthesis of stigmatellin from *S. aurantiaca*.[115] Iterative use of type I PKS modules has since then been reported from a growing number of myxobacterial and streptomycete megasynthetases, and is increasingly recognized as a programmed event in type I PKS biosynthetic pathways.[59] Selected examples of secondary metabolites which originate in part from iterative type I PKS biosynthesis are shown in Figure C-6, along with the domain organization of the iteratively acting modules. Modules AurA and BorA5 incorporate two and three methylmalonyl-CoA monomers during biosynthesis of aureothin and borrelidin, repectively, but differ with respect to the degree of ketoreduction they perform.[116,117] The malonyl-CoA specific module DkxN has recently been proposed as one of two iterating modules involved in the biosynthesis of the variable-length polyene portion of the DKxanthenes from *M. xanthus*.[84] Intriguingly, the polyenic side chain of the myxochromides from *S. aurantiaca* is synthesized by one single module (MchA), representing the only PKS module in the hybrid PKS/NRPS *mch* gene cluster.[85] Although MchA contains an ER domain, this domain is presumably inactive due to several mutations in conserved motifs. Thus, iterative use of MchA for seven or eight rounds of malonyl-CoA incorporation has been proposed to give rise to the unsaturated C_{16} to C_{18} moiety found in the myxochromides S_1 – S_3.

In contrast to the aforementioned biosyntheses, the imidacin pathway currently represents the only investigated type I PKS biosynthetic route producing a fully reduced and unbranched alkyl chain, by the repeated use of one single PKS module. Imidacin formation thus evidently mimicks the outcome of fatty acid biosynthesis. The variable length of the imidacin scaffold may be explained by a lack of stringent chain length control. This assumption may be addressed in future studies, as the mechanisms underlying chain length control in iterative PKS systems still await an explanation.

The structures of imidacins A_1 and A_2 comprise a cyclopropane moiety (Figure C-5). Besides certain fatty acids from *E. coli*, *M. tuberculosis* and other prokaryotes,[118,119] only a small number of bacterial natural products have been shown to contain this structural feature. Examples include coronatine from *P. syringae*, curacin from the marine cyanobacterium *L. majuscula*, and ambruticin from *S. cellulosum* (Figure C-7).[120-122] Despite the rareness of cyclopropane rings in microbial secondary metabolites, the biosynthetic routes leading to their formation are amazingly diverse. For example, a cryptic chlorination strategy has been proposed for the formation of the aminocarboxycyclopropane moiety of coronamic acid, the precursor of coronatine

(Figure C-7 **A**).[123] Following enzymatic chlorination by CmaB, the PCP-bound γ-Cl intermediate undergoes nucleophilic attack by the α-carbon (supported by CmaC), to yield the cyclopropane ring.[120] Curacin biosynthesis is also believed to proceed via a cyclopropyl-ACP intermediate (Figure C-7 **B**), but details about the transformation of the proposed 3-methylcrotonyl-

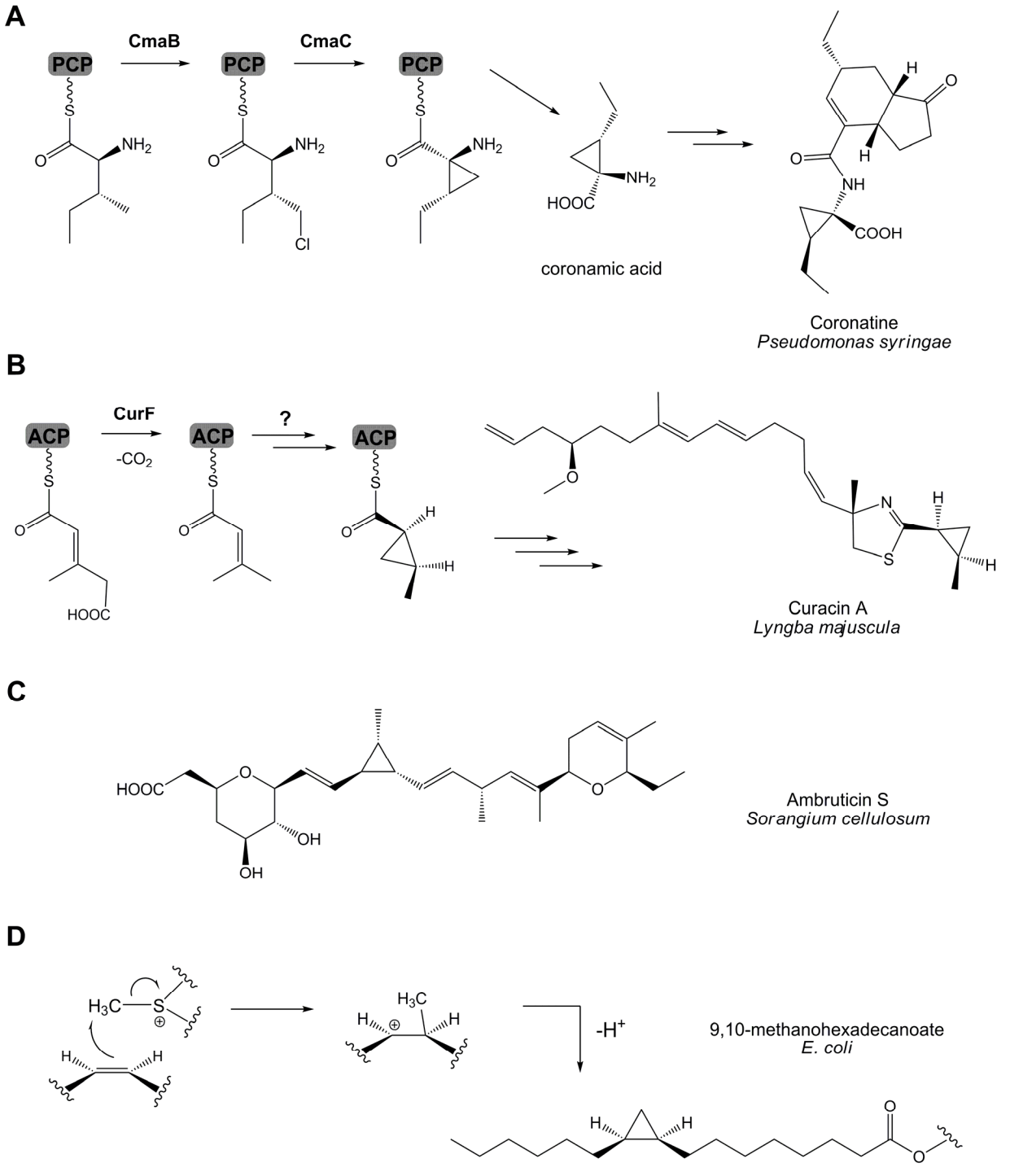


Figure C-7. A-D, Bacterial metabolites featuring a cyclopropane moiety, and proposed biosynthetic transformations involved in cycopropane ring formation (**A**, **B**, **D**).

ACP precursor have yet to be elucidated.[121] With regard to the biosynthesis of ambruticins (Figure C-7 **C**), it has been hypothesized that formation of the methylcyclopropane ring could be induced by a Favorskii rearrangement of an assembly line intermediate; however, nothing is known about the enzyme hypothetically involved in establishing the final cyclopropane stereochemistry.[122]

Notably, enzymes capable of introducing a cyclopropane moiety into unsaturated fatty acid scaffolds have been reported from *E. coli* and *M. tuberculosis*.[124,125] These *S*-adenosylmethionine (SAM)-dependent cyclopropane synthases have been shown to catalyze the methylenation ("cyclopropanation") of unsaturated fatty acid precursors via addition of an SAM-derived methyl group to a double bond, followed by rapid proton loss (Figure C-7 **D**). Due to the structural similarity between fatty acids and imidacins, we suggested that this mechanism for cyclopropanation could also apply to imidacin biosynthesis (chapter B-4). Therefore, an *S. aurantiaca* Sga15 culture was supplemented with (methyl-d3)-L-methionine to create an intracellular pool of (methyl-d3)-SAM. The observed incorporation of two deuterium labels into the imidacins provided strong evidence for the formation of the cyclopropane ring in imidacins via SAM-dependent cyclopropanation. The logical substrate for this transformation would be a Δ^3-unsaturated imidacin precursor, which in turn could derive from a molecule with a completely saturated alkyl chain. The mass spectrometric investigation of Sg a15 extracts indicated the existence of both putative precursors (chapter B-4, Figure S4). Thus, the late steps in imidacin biosynthesis most likely include the introduction of an unsual Δ^3-double bound into a fatty acid-like precursor and subsequent conversion of this double bond to yield a cyclopropane ring.

This hypothesis is further corroborated by the fact that a cyclopropane synthase candidate enzyme, Imd4, is encoded within the imidacin gene cluster. Imd4 has a conserved SAM binding site, but otherwise exhibits rather weak similarity to known cyclopropane synthases. However, this divergence might be attributable to the unusual substrate specificity required for imidacin conversion. Further characterization of the potentially novel cyclopropane synthase Imd4 should be performed *in-vitro*, using recombinantly produced Imd4 and the unsaturated presursor (obtained through synthesis) as substrate. A plausible candidate enzyme for the required Δ^3-desaturation step was not identified within the sequenced region, but may instead be encoded elsewere in the Sg a15 genome. A feeding study, using the completely saturated imidacin precursor and the HAL-deficient mutant strain generated in this work, should corroborate the presence of this Δ^3-desaturase function in Sg a15. Such an enzyme has actually not been reported from bacteria to date, but has been characterized very recently, for the first time, from the fungus

Fusarium graminearum.[126] Thus, the sequence-guided discovery of a gene encoding the hypothetical imidacin desaturase might be attempted, provided that whole-genome sequence information for *S. aurantiaca* Sg a15 becomes available. At present, the alternative possibility that the $\Delta^{3\text{-}}$double bond could be incorporated during PKS assembly cannot be excluded. Indeed, certain iterative PKS systems have been reported to exhibit a variable behaviour with respect to the oxidation state of the assembly line intermediate, depending on chain length.[52] However, due to the rather unusal positioning of the double bond a PKS-derived origin appears less likely.

Taken together, imidacin biosynthesis was shown to comprise several intriguing transformations. The pathway exhibits an unusually disintegrated genetic organization, as genes encoding functions for precursor supply and post-assembly line modification are located distantly from the core biosynthetic gene set. Moreover, the clustered region itself appears partially cryptic, due to the non-linear arrangement of modules and the presence of seemingly nonessential genes. Therefore, further gene inactivation experiments, using markerless in-frame deletions or point mutations, are needed to confirm that genes without a predicted function for imidacin biosynthesis are actually dispensable. However, such efforts are currently hindered by the inavailability of the appropriate genetic tools for *S. aurantiaca* Sg a15. Thus, heterologous expression of the imidacin pathway is envisaged. Genetic manipulation in the heterologous host could then help to determine the minimally required gene set and should also allow to confirm the proposed roles of certain candidate genes and domains encoded within the cluster.

Finally, further studies should evaluate the potential biological activity of imidacins. Additional investigations are also necessary to shed light on the significance of imidacin production for *Stigmatella aurantiaca*.

4. The impact of advanced mass spectrometry-based methods on the analysis of microbial natural product diversity

The last several years have seen an enormous increase in the amount of microbial genomic information added to public databases. In light of the current advances in sequencing technology, this trend is definitely expected to accelerate further. When analyzing the genomes of many bacterial species for the presence of typical secondary metabolite biosynthetic gene clusters, it has recently become a common finding that the apparent genetic potential for natural product biosynthesis exceeds by far the number of known compounds from a species. This observation even holds true for many well-established producers of secondary metabolites.[110] Thus, the discovery of novel metabolites and their subsequent correlation to biosynthetic gene clusters now represents the new bottleneck, rather than the process of obtaining the sequence information.

A strikingly high number of predicted secondary metabolite biosynthesis pathways were also revealed in the course of two recent myxobacterial genome-sequencing projects: the genomes of *Myxococcus xanthus* DK1622 and *Sorangium cellulosum* So ce56 contain at least 18 and 17 putative PKS, NRPS or hybrid gene loci, respectively.[39,40] In contrast, intense screening efforts have so far yielded merely a handful of secondary metabolites from each of these strains. Various reasons could account for this discrepancy. Biosynthetic pathways may theoretically be non-functional due to mutations, thus representing evolutionary artifacts. However, this assumption is unlikely for a high proportion of biosynthetic gene clusters in bacteria, since the elimination of large, "useless" sequence regions (i.e. defective genes which do not confer a fitness advantage) is expected to occur rapidly in terms of the evolutionary timescale. More likely, low productivity may be a reason for non-detectability. Both, biosynthetic routes and metabolic pathways for precursor supply, may be down-regulated in the laboratory environment for a number of reasons – including nutritional conditions, the generally "unnatural" cultivation parameters (liquid medium in shaking flask vs. soil; higher temperature, better aeration), and feedback inhibition. In addition, precursors required to biosynthesize unusual building blocks may not be available, or very specific inducers of secondary metabolite production from the natural habitat could be missing. Moreover, the produced metabolites could suffer from limited stability in the cultivation broth, and they may even be actively degraded by the producer upon prolonged cultivation. Some of these issues, namely feedback inhibition and the stability/degradation problem, are partly circumvented by the routine addition of adsorber resin to myxobacterial fermentations. It should be noted, however, that this widespread practice introduces a certain bias with respect

to the metabolite inventory observed from a strain, as it can disproportionally increase the yield of well-adsorbed compounds while potentially failing to collect others.

Another important point of concern relates to the analytical methods used in the past for the discovery of natural products from myxobacteria. Over the last decades, the screening of myxobacterial extracts has largely relied on bioassays and HPLC-coupled UV/Vis detection. Activity-guided screening is convenient because it can directly deliver candidate compounds for a specific biological activity – however, this approach is inherently biased by the choice of assays and targets. HPLC-UV/Vis-based screening, on the other hand, inevitably fails to highlight compounds without "eye-catching" light-absorption characteristics. In addition, it cannot identify minor constituents in complex mixtures. Significantly, it is highly desirable to reliably detect compounds regardless of their biological activity and in spite of their low abundance, when the aim is to correlate a strain's compound inventory to its genetic capacity for the production of secondary metabolites.

Modern liquid chromatography-coupled electrospray ionization mass spectrometry (LC-ESI-MS) offers a solution to this challenge, as it can in principle detect a highly diverse range of chemical compounds with high sensitivity. Depending on the MS platform used, different approaches may be chosen to screen bacterial metabolite profiles. Fragmentation patterns derived from data-dependent tandem-MS measurements have occassionally been employed for compound identification.[127] However, it should be pointed out that the self-limitation to compounds giving rise to sufficiently intense signals for automatic MS^2-fragmentation, represents a severe shortcoming of this strategy. Therefore, a different approach for the analysis of myxobacterial extracts was devised in this work (chapters B-6 and B-7). Using data from high-resolution LC-ESI-ToF measurements, compound identification was accomplished through the combined evaluation of accurate mass values and isotope patterns. In addition, retention times were incorporated as an orthogonal analytical parameter. Importantly, the implementation of high-resolution LC-MS also permitted the semi-quantitative treatment of all obtained data, without the need to define in advance the specific signals to monitor.[128] Moreover, a compound finding algorithm was applied to the high-resolution LC-MS data, thereby creating a "sample inventory" which contained both known substances and potentially novel compounds. This processing step in turn facilitated the subsequent statistical interpretation by Principal Component Analysis (PCA) (chapter B-7).

Because of the aforementioned possibilities, LC-coupled high-resolution MS techniques similar to the methods employed in this work are now increasingly used for a range of metabolic profiling applications.[128-130] In contrast, statistical methods for the mining of LC-MS data cur-

rently appear to be underutilized in screening projects aimed at the discovery of novel natural products from bacteria.

Discovering the hidden secondary metabolome of *Myxococcus xanthus.*

Myxococcus xanthus is the biologically best characterized myxobacterium to date, but in terms of known secondary metabolites this species apparently lags behind other established myxobacterial producers, such as *Sorangium cellulosum*. Nevertheless, genomic mining has revealed the considerable potential of *M. xanthus* for the production of secondary metabolites, as judged by the presence of at least 18 PKS, NRPS or hybrid gene clusters in the genome of the model strain DK1622.[131] To date, five secondary metabolite classes have been identified in extracts from DK1622, and the availability of whole-genome information has greatly facilitated their correlation to biosynthetic gene clusters. These known compounds are the myxochelins, the myxalamids, the myxovirescins, the myxochromides, and the DKxanthenes.[84,132-135] Encouragingly, a recent global gene expression study has provided evidence that many of the predicted natural product assembly lines in DK1622, which could so far not be correlated to metabolites, are actually active under vegetative growth conditions.[136] Consistent with this finding, a previous proteomic study also suggested that many gene clusters are expressed, since the corresponding peptides could be identified by mass spectrometric analysis.[137] Taken together, the combined data from these two studies provide evidence that most of the secondary metabolite gene clusters (i.e. 17 out of 18) in *M. xanthus* DK1622 are not silent in the laboratory environment. Nevertheless, only five natural products have been identified to date. Thus, it was anticipated that the application of improved LC-MS-based methods to *M. xanthus* extracts should lead to the

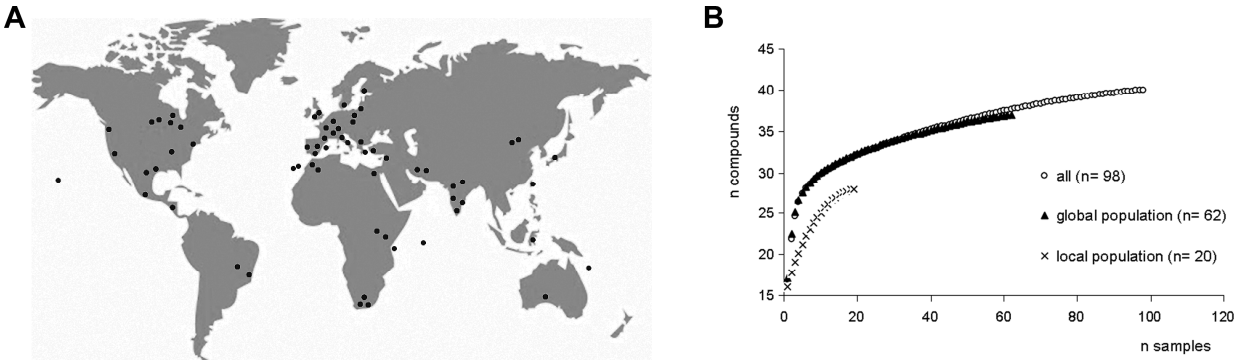


Figure C-8. **A**, Global distribution of sampling sites from which the 98 *M. xanthus* strains in this study originated. Multiple isolates from regional or local spatial ranges were sumamrized as single spots. **B**, Diversity accumulation curves, based on the distribution of non-ubiquituous compounds in 98 samples. Each data point represents the average number of discovered compounds that would result if *n* samples were taken (see also chapter B-6).

discovery of novel compounds. However, a comparative profiling study which involved knock-out mutants of all predicted biosynthetic gene clusters in *M. xanthus* DK1622, did not succeed to disclose further candidate metabolites (*Nina Cortina, DK and RM, unpublished results*).

A different approach to discover the hidden secondary metabolite inventory of *M. xanthus* was therefore taken. The metabolite profiles of 98 *M. xanthus* strains that originated from 78 locations worldwide were compared, and 20 centimeter-scale isolates from one location (Tübingen, Germany) were also included in this screen (Figure C-8, **A**). This study had in fact two objectives: firstly, to determine whether there is significant potential for novel compound discoveries in other strains besides DK1622 (e.g. if these other strains exhibit increased yield); and secondly, to compare the intraspecific diversity of *M. xanthus* metabolite profiles on global versus local scales.

Targeted screening revealed the presence of four known secondary metabolites (myxochromides, myxochelins, myxalamids, DKxanthenes) in all 98 *M. xanthus* extracts, while only three natural products (althiomycin, cittilin, myxovirescin) exhibited a non-ubiquitous pattern.[48] Thus, the known natural products seemed to contribute rather marginally to secondary metabolite diversity in *M. xanthus*. On the contrary, the statistical evaluation of all LC-MS datasets, implementing a compound-based PCA approach, revealed 37 additional non-ubiquitous compounds representing putative novel metabolites (chapter B-6).[48] The diversity-oriented LC-MS based screening approach presented here thus uncovered a strikingly high level of intraspecific diversity in the secondary metabolome of *M. xanthus*. Accordingly, the species *M. xanthus* should be regarded as a promising source of as yet unidentified natural products, in agreement with the previous findings from genomic, transcriptomic and proteomic studies with DK1622.

By applying basic population statistics to the matrix of observed *M. xanthus* "chemotypes", it was estimated that the 37 newly reported candidate compounds are likely to constitute ~80% of the total secondary metabolome diversity in *M. xanthus* (Figure C-8, **B**). The metabolite profiles of the 20 local isolates were more similar to each other than to most of the global isolates, but the centimeter-scale population nonetheless housed substantial metabolite diversity. This finding is consistent with the result of a recent study on the genetic population structure of *M. xanthus* on the centimeter scale, which concluded that *M. xanthus* cells are "surrounded in soil by a wide range of genetically distinct conspecifics".[138] Interestingly, the average degree of chemotype similarity among strain pairs isolated within the meter-scale range at the same location was not significantly different from that among global-scale strain pairs. Thus, much mixing of global chemotypes seems to occur at the meter scale. In conclusion, intense sampling at indi-

vidual sites may substantially increase the estimate of the total number of metabolites beyond that obtained from analysis of single samples from many sites (Figure C-8, **B**).

In this study, all observations (i.e. high-resolution *m/z* signals at specific retention times in an LC-MS dataset) constituting a distinct difference between the metabolite profiles of two or more strains were regarded as indicating putative compounds of interest, and proposed molecular formulae served as descriptors for the novel compounds. For a number of reasons, the structural elucidation of all these compounds and their correlation to a biosynthetic gene cluster may not be a straightforward undertaking (chapter B-6). However, the finding that several putative compounds occur with a very similar pattern, and additionally exhibit similar retention times plus molecular formulae is suggestive of "biosynthetically meaningful" differences. In addition, it indicates that at least some of these candidate compounds are likely to be secondary metabolites synthesized by PKS and/or NRPS biosynthetic routes. Intriguingly, the targeted analysis of extracts from *M. xanthus* DK1622, utilizing the list of 37 candidate metabolites highlighted by the species-wide screen, revealed this strain as a low-level producer of 10 of these compounds. Follow-up studies aimed at identifying the corresponding biosynthetic gene clusters are now beginning to correlate genes to compound masses (*Nina Cortina, DK and RM, personal communication*). Thus, the diversity-oriented screening approach devised here can help to discover the products assembled by biosynthetic pathways which could not previously be correlated to a specific secondary metabolite.

The analytical framework for metabolomics-based mining of the myxobacterial secondary metabolite inventory: basic considerations and outlook.

Metabolomics studies in principle aim at generating a quantitative molecular phenotype for an organism under investigation.[130] In the context of bacterial natural product research, primarily the secondary metabolome is of interest. It may be defined as the full range of small molecules released by a strain into the surrounding medium which are neither end products nor intermediates of primary metabolic pathways. The secondary metabolome thus represents a subset of the extracellular metabolome, and is commonly analyzed as an integral part of the latter. Consequently, sample complexity is generally high and is oftentimes increased further by the presence of abundant matrix compounds derived from cultivation broth, which may occasionally even dominate a sample in terms of signal strength. Thus, studies of the bacterial secondary metabolome require techniques with high resolving power and sensitivity.[139] In the last several years, this demand has been increasingly met by combining advanced liquid chromatography

techniques (i.e. UPLC using sub-2 µm particle size colums) with high-resolution mass spectrometric detection.[140] The utility of such an analytical platform was demonstrated in the *M. xanthus* study discussed in the preceeding section. In comparison to bioassay-guided approaches or UV/Vis screening, metabolomics-style experiments based on mass spectrometry are far less biased with respect to the properties of the detected compounds. However, a certain bias is always introduced by the sample preparation procedure and by the analytical setup (e.g. choice of stationary and mobile phase, ionization, mass range). To address the latter issue, a metabolomics experiment with the highest possible coverage of "chemical space" may require an individual sample to be analyzed using multiple instrumental configurations. With regard to the characterization of myxobacterial secondary metabolomes, the potential benefit of changing certain key attributes of the analytical system remains to be evaluated in the future. For example, negative electrospray ionization or APCI could be used in addition to the standard positive ESI. The implementation of nano-electrospray ionization (nanoESI) appears especially promising, as it affords reduced in-source fragmentation of labile compounds and less unwanted adduct formation along with superior sensitivity (DK, unpublished observations).

Metabolomic studies may adopt different strategies, commonly termed the "targeted" and the "comprehensive" approach. The former is often selective for certain compound classes, such as amino acids or lipids, and the analysis is usually limited to a number of known compounds. In the second approach, also called "quantitative metabolomics", the focus is on attempting to identify and quantify as many compounds in a given sample as possible.[130] The worldwide LC-MS survey of *M. xanthus* metabolite diversity which was performed in this work clearly falls in the second category (although "targeted screening" was also employed). In such comprehensive studies, the implementation of statistical tools for data evaluation is crucial. One statistical method which has been previously used in metabolomic applications is Principal Component Analysis (PCA). As explained earlier, PCA is a multivariate method for the classification of samples which acts like a statistical filter for LC-MS data, pointing towards compounds representing the most significant differences between samples. However, the essential challenge is to properly prepare LC-MS data, which always contain a considerable degree of "chemical noise", for statistical evaluation. Pre-processing of high-resolution data can be accomplished by applying a compound finder algorithm (a procedure also referred to as "chemical feature detection").[141] Statistical interpretation is then carried out on the resulting matrix, representing the distribution of all putative compounds (denoted by an exact mass/retention time pair plus an intensity value) in all samples. It was one objective of this work to evaluate the effectiveness of a new integrated compound-finding/PCA strategy for the characterization of myxobacterial me-

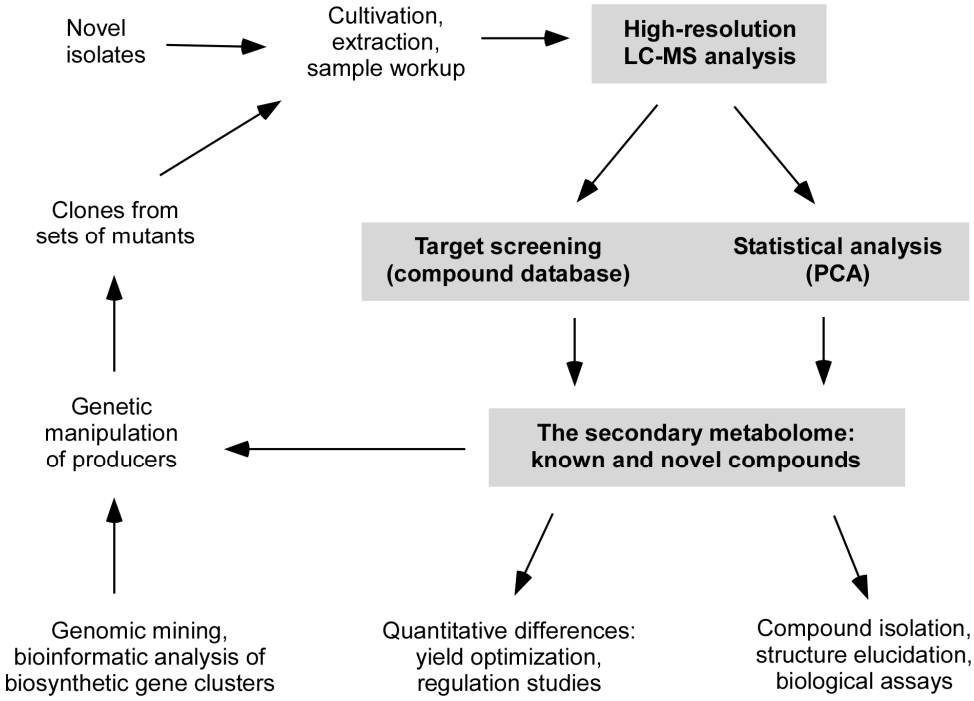


Figure C-9. The analytical framework for LC-MS based mining of myxobacterial secondary metabolomes.

tabolite profiles. The results clearly demonstrated that the "compound-based bucketing" approach devised here can indeed generate meaningful output using high-resolution LC-MS data derived from complex *M. xanthus* samples. Furthermore, the compound-based PCA strategy was shown to be superior to a more conventional "rectangular bucketing" approach (chapter B-7).

Besides the discovery of strain-specific metabolites, as carried out in the diversity-oriented screening approach with *M. xanthus*, the identification of compounds giving rise to distinct differences in a set of highly complex samples is an important concern in a number of typical applications related to the investigation of secondary metabolite biosynthesis in myxobacteria (Figure C-9). For example, the metabolite profiles from a library of bacterial clones, generated by targeted knockout of candidate biosynthetic genes could be analyzed for the absence of specific metabolites relative to the wildtype strain. Even without any information about biosynthetic genes, samples for such an analysis can be generated by random gene inactivation using transposon mutagenesis. Furthermore, quantitative variation in secondary metabolomes could be addressed through high-resolution LC-MS measurements with subsequent interpretation by PCA, providing the opportunity to assay changes in metabolite production pattern in response to different nutritional conditions – notably, without the need to define *a priori* the specific com-

pounds to monitor. Similarly, a metabolomics-based approach seems appropriate for investigating the molecular basis for regulation of secondary metabolite biosynthesis in myxobacteria in the future. Finally, PCA-based LC-MS data mining could also become useful for comparative metabolite profiling in the course of heterologous expression projects. It is an obvious prerequisite for all these applications that the strains under investigation exhibit reproducible growth and production characteristics (ideally even when using a miniaturized cultivation setup). These requirements are generally not a problem when working with *M. xanthus*. However, they impose a certain limitation on the applicability of PCA-based LC-MS experiments, considering the extremely slow and sometimes irregular growth characteristics experienced from a number of myxobacterial species, like *S. cellulosum* and *C. crocatus*.

It has to be pointed out that the result of PCA, owing to its essentially unsupervised nature, may in some cases not reflect the meaningful differences between samples, in relation to the underlying experimental design. This can particularly occur in sample sets where intense variation exists within highly abundant matrix compounds – the random variation may then supersede the subtle, but systematic differences pointing at the relevant metabolites. In such cases, it is necessary to manually curate a PCA model by exclusion of compounds that give rise to random variation, thereby adding a certain degree of supervision. If the respective experiment is based on groups of replicate samples, this refinement procedure might even be performed automatically by excluding compounds responsible for a large degree of group-internal variance (albeit it seems that this possibility is not currently realized in the available PCA software packages). Alternatively, it may also be feasible to directly query the comprehensive compound list generated from a sample set (the so-called “bucket table”), with the aim to filter out compounds exhibiting a defined occurrence pattern (e.g., creating a list of compounds present in all samples belonging to the wildtype group, but systematically absent from all samples in one “knockout” group).

Once a putative compound of interest has been highlighted by statistical data evaluation, high-resolution measurements allow the direct calculation of a molecular formula proposal, thus facilitating compound identification. The generation of molecular formulae for putative novel compounds with high confidence due to the evaluation of both exact mass and isotopic pattern has been exemplified throughout this work as an important key for de-replication and prioritization of candidate compounds for further characterization. In addition, tandem-MS measurements may be employed to confirm a compound identification. They can also serve to obtain initial information about the structure of unknown compounds (e.g. peptide-like fragmentation pattern), possibly in combination with feeding studies employing isotopically labeled precursors. The ability of novel mass spectrometers such as the linear trap-orbitrap or quadru-

pole/quadrupole-ToF devices to perform multi-stage fragmentation meaurements at high resolution, is especially advantageous in this respect.

Taken together, metabolomics-based approaches to mining the myxobacterial metabolite inventory, implementing advanced LC-coupled high-resolution mass spectrometry techniques in combination with improved statistical tools, are likely to contribute to the future discovery of novel natural products from myxobacterial species. Likewise, the in-depth investigation of myxobacterial secondary metabolomes looks set to enhance our understanding of various aspects of secondary metabolite production in myxobacteria.

D. List of citations

1. **Grabley S., Thiericke R. (1999).** The impact of natural products on drug discovery. *In:* Grabley, S. and Thiericke, R. (Eds.), Drug discovery from nature, Springer, Berlin, 3-37.

2. **Demain A.L. (1999).** Pharmaceutically active secondary metabolites of microorganisms. *Appl. Microbiol. Biotechnol.* 52, 455-463.

3. **Baker D.D., Chu M., Oza U., and Rajgarhia V. (2007).** The value of natural products to future pharmaceutical discovery. *Nat.Prod.Rep.* 24, 1225-1244.

4. **Hubbard B.K. and Walsh C.T. (2003).** Vancomycin assembly: nature's way. *Angew.Chem.Int.Ed.* 42, 730-765.

5. **Baltz R.H., Miao V., and Wrigley S.K. (2005).** Natural products to drugs: daptomycin and related lipopeptide antibiotics. *Nat.Prod.Rep.* 22, 717-741.

6. **Washington J.A. and Wilson W.R. (1985).** Erythromycin: a microbial and clinical perspective after 30 years of clinical use. *Mayo.Clin.Proc.* 60, 189-203.

7. **Williams P.G., Buchanan G.O., Feling R.H., Kauffman C.A., Jensen P.R., and Fenical W. (2005).** New cytotoxic salinosporamides from the marine actinomycete *Salinispora tropica*. *J.Org.Chem.* 70, 6196-6203.

8. **Chang Z., Sitachitta N., Rossi J.V., Roberts M.A., Flatt P., Jia J., Sherman D.H., and Gerwick W.H. (2004).** Biosynthetic pathway and gene cluster analysis of Curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *J.Nat.Prod.* 67, 1356-1367.

9. **Brandi L., Fabbretti A., La Teana A., Abbondi M., Losi D., Donadio S., and Gualerzi C.O. (2006).** Specific, efficient, and selective inhibition of prokaryotic translation initiation by a novel peptide antibiotic. *Proc.Natl.Acad.Sci.USA* 103, 39-44.

10. **Demain A.L. (2006).** From natural products discovery to commercialization: a success story. *J.Ind.Microbiol.Biotechnol.* 33, 486-495.

11. **Cordier C., Morton D., Murrison S., Nelson A., and O'Leary-Steele C. (2008).** Natural products as an inspiration in the diversity-oriented synthesis of bioactive compound libraries. *Nat.Prod.Rep.* 25, 719-737.

12. **Baltz R.H. (2006).** Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J.Ind.Microbiol.Biotechnol.* 33, 507-513.

13. **Busti E., Monciardini P., Cavaletti L., Bamonte R., Lazzarini A., Sosio M., and Donadio S. (2006).** Antibiotic-producing ability by representatives of a newly discovered lineage of actinomycetes. *Microbiology* 152, 675-683.

14. **Tan L.T. (2007).** Bioactive natural products from marine cyanobacteria for drug discovery. *Phytochemistry* 68, 954-979.

15. **Scherlach K., Partida-Martinez L.P., Dahse H.M., and Hertweck C. (2006).** Antimitotic rhizoxin derivatives from a cultured bacterial endosymbiont of the rice pathogenic fungus Rhizopus microsporus. *J.Am.Chem.Soc.* 128, 11529-11536.

16. **Li J., Hu K., and Webster J.M. (1998).** Antibiotics from Xenorhabdus spp. and Photorhabdus spp. (Enterobacterialceae). *Chem.Hetero.Comp.* 34, 1331-1339.

17. **Lam K.S. (2006).** Discovery of novel metabolites from marine actinomycetes. *Curr.Opin.Microbiol.* 9, 245-251.

18. **Reichenbach H., Höfle G. (1999).** Myxobacteria as producers of secondary metabolites. *In:* Grabley, S. and Thiericke, R. (Eds.), Drug Discovery from Nature, Springer, Berlin, 149-179.

19. **de Bruijn I., de Kock M.J.D., Yang M., de Waard P., van Beek T.A., and Raaijmakers J.M. (2007).** Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in Pseudomonas species. *Mol.Microbiol.* 63, 417-428.

20. **Whitworth, D. (2007).** Myxobacteria: multicellularity and differentiation. ASM Press, Chicago.

21. **Shimkets L., Dworkin M., Reichenbach H. (2006).** The Myxobacteria. *In:* Dworkin, M. (Ed.), The Prokaryotes, Springer, Berlin, 31-115.

22. **Weissman K.J. and Müller R. (2008).** A brief tour of myxobacterial secondary metabolism. *Bioorg.Med.Chem.* In press.

23. **Bode H.B. and Müller R. (2006).** Analysis of myxobacterial secondary metabolism goes molecular. *J.Ind.Microbiol.Biotechnol.* 33, 577-588.

24. **Gerth K., Pradella S., Perlova O., Beyer S., and Müller R. (2003).** Myxobacteria: proficient producers of novel natural products with various biological activities - past and future biotechnological aspects with the focus on the genus *Sorangium*. *J.Biotechnol.* 106, 233-253.

25. **Mulzer J. (2008).** The Epothilones - An Outstanding Family of Anti-Tumour Agents: From Soil to the Clinic. Springer, New York.

26. **Khalil M.W., Sasse F., Lunsdorf H., Elnakady Y.A., and Reichenbach H. (2006).** Mechanism of action of tubulysin, an antimitotic peptide from myxobacteria. *ChemBioChem* 7, 678-683.

27. **Nickeleit I., Zender S., Sasse F., Geffers R., Brandes G., Sörensen I., Steinmetz H., Kubicka S., Carlomagno T., Menche D., Gütgemann I., Buer J., Gossler A., Manns M.P., Kalesse M., Frank R., and Malek N.P. (2008).** Argyrin A reveals a critical role for the tumor suppressor protein $p27_{kip1}$ in mediating antitumor activities in response to proteasome inhibition. *Cancer Res.* 14, 23-35.

28. **Thierbach G., Kunze B., Reichenbach H., and Höfle G. (1984).** The mode of action of stigmatellin, a new inhibitor of the cytochorme b-c1 segment of the respiratory chain. *Biochim.Biophys.Acta* 765, 227-235.

29. **Vahlensieck H.F., Pridzun L., Reichenbach H., and Hinnen A. (1994).** Identification of the yeast ACC1 gene product (acetyl-CoA carboxylase) as the target of the polyketide fungicide soraphen A. *Curr.Genet.* 25, 95-100.

30. **Gerth K., Steinmetz H., Hofle G., and Jansen R. (2008).** Chlorotonil A, a macrolide with a unique gem-dichloro-1,3-dione functionality from *Sorangium cellulosum*, So ce1525. *Angew.Chem.Int.Ed.* 47, 600-602.

31. **Rahn N. and Kalesse M. (2008).** The total synthesis of chlorotonil A. *Angew.Chem.Int.Ed.* 47, 597-599.

32. **Jansen R., Kunze B., Reichenbach H., and Höfle G. (2003).** Chondrochloren A and B, new beta-amino styrenes from *Chondromyces crocatus* (myxobacteria). *Eur.J.Org.Chem.* 2684-2689.

33. **Bode H.B. and Müller R. (2005).** The impact of bacterial genomics on natural product research. *Angew.Chem.Int.Ed.* 44, 6828-6846.

34. **Menzella H.G. and Reeves C.D. (2007).** Combinatorial biosynthesis for drug development. *Curr.Opin.Microbiol.* 10, 238-245.

35. **Floss H.G. (2006).** Combinatorial biosynthesis - Potential and problems. *J.Biotechnol.* 124, 242-257.

36. **Weissman K.J. and Leadlay P.F. (2005).** Combinatorial biosynthesis of reduced polyketides. *Nat.Rev.Microbiol.* 3, 925-936.

37. **Shen B. (2004).** Accessing natural products by combinatorial biosynthesis. *Sci STKE* 2004, e14.

38. **Wenzel S.C. and Müller R. (2005).** Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways. *Curr.Opin.Biotechnol.* 16, 594-606.

39. **Schneiker S., Perlova O., Kaiser O., Gerth K., Alici A., Altmeyer M.O., Bartels D., Bekel T., Beyer S., Bode E., Bode H.B., Bolten C.J., Choudhuri J.V., Doss S., Elnakady Y.A., Frank B., Gaigalat L., Goesmann A., Groeger C., Gross F., Jelsbak L., Jelsbak L., Kalinowski J., Kegler C., Knauber T., Konietzny S., Kopp M., Krause L., Krug D., Linke B., Mahmud T., Martinez-Arias R., McHardy A.C., Merai M., Meyer F., Mormann S., Munoz-Dorado J., Perez J., Pradella S., Rachid S., Raddatz G., Rosenau F., Ruckert C., Sasse F., Scharfe M., Schuster S.C., Suen G., Treuner-Lange A., Velicer G.J., Vorholter F.J., Weissman K.J., Welch R.D., Wenzel S.C., Whitworth D.E., Wilhelm S., Wittmann C., Blöcker H., Pühler A., and Müller R. (2007).** Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat.Biotechnol.* 25, 1281-1289.

40. **Goldman B.S., Nierman W.C., Kaiser D., Slater S.C., Durkin A.S., Eisen J., Ronning C.M., Barbazuk W.B., Blanchard M., Field C., Halling C., Hinkle G., Iartchuk O., Kim H.S., Mackenzie C., Madupu R., Miller N., Shvartsbeyn A., Sullivan S.A., Vaudin M., Wiegand R., and Kaplan H.B. (2006).** Evolution of sensory complexity recorded in a myxobacterial genome. *Proc.Natl.Acad.Sci.USA* 103, 15200-15205.

41. **Kunze B., Jansen R., Reichenbach H., and Höfle G. (1996).** Chondramides A-D, new cytostatic and antifungal cyclodepsipeptides from *Chondromyces crocatus* (myxobacteria): Isolation and structure elucidation. *Liebigs Ann.* 285-290.

42. **Rachid S., Krug D., Kochems I., Kunze B., Scharfe M., Blöcker H., Zabriski M., and Müller R. (2006).** Molecular and biochemical studies of chondramide formation - highly cytotoxic natural products from *Chondromyces crocatus* Cm c5. *Chem.Biol.* 14, 667-681.

43. **Christenson S.D., Liu W., Toney M.D., and Shen B. (2003).** A novel 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis. *J.Am.Chem.Soc.* 125, 6062-6063.

44. **Rachid S., Krug D., Weissman K.J., and Müller R. (2007).** Biosynthesis of (*R*)-beta-tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chondromyces crocatus*. *J.Biol.Chem.* 282, 21810-21817.

45. **Krug D. and Müller R. (2009).** Discovery of additional members of the tyrosine aminomutase enzme family and the mutational analysis of CmdF. *ChemBioChem* 10, 741-750.

46. **Krug, D., Dickschat, J. S., and Müller, R.** (**2009**). Imidacins: Structure and biosynthesis of novel secondary metabolites from *Stigmatella aurantiaca* incorporating an urocanate building block. *(manuscript to be submitted)*.

47. **Krug D., Zurek G., Schneider B., Garcia R., and Müller R. (2008).** Efficient mining of myxobacterial metabolite profiles enabled by liquid-chromatography – electrospray ionization – time-of-flight mass spectrometry and compound-based principal component analysis. *Anal.Chim.Acta* 624, 97-106.

48. **Krug D., Zurek G., Revermann O., Vos M., Velicer G.J., and Müller R. (2008).** Discovering the hidden secondary metabolome of *Myxococcus xanthus*: a study of intraspecific diversity. *Appl.Environ.Microbiol.* 74, 3058-3068.

49. **Garcia R.O., Krug D., Müller R. (2009).** Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods. *In:* Methods in enzymology, Hopwood, D. (Ed.), Vol. 458(A)

50. **Lai J.R., Koglin A., and Walsh C.T. (2006).** Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. *Biochemistry* 45, 14869-14879.

51. **Fischbach M.A. and Walsh C.T. (2006).** Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem.Rev.* 106, 3468-3496.

52. **Staunton J. and Weissman K.J. (2001).** Polyketide biosynthesis: a millennium review. *Nat.Prod.Rep.* 18, 380-416.

53. **Schwarzer D., Finking R., and Marahiel M.A. (2003).** Nonribosomal peptides: From genes to products. *Nat.Prod.Rep.* 20, 275-287.

54. **Neumann C.S., Fujimori D.G., and Walsh C.T. (2008).** Halogenation strategies in natural product biosynthesis. *Chem. Biol.* 15, 99-109.

55. **Walsh C.T. (2007).** The chemical versatility of natural-product assembly lines. *Acc.Chem.Res.* 41, 4-10.

56. **Moss S.J., Martin C.J., and Wilkinson B. (2004).** Loss of co-linearity by modular polyketide synthases: a mechanism for the evolution of chemical diversity. *Nat.Prod.Rep.* 21, 575-593.

57. **Yadav G., Gokhale R.S., and Mohanty D. (2003).** Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* 328, 335-363.

58. **Finking R. and Marahiel M.A. (2004).** Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol.* 58, 453-488.

59. **Wenzel S.C. and Müller R. (2007).** Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat.Prod.Rep.* 24, 1211-1224.

60. **Sasse F., Kunze B., Gronewold T.M., and Reichenbach H. (1998).** The chondramides: cytostatic agents from myxobacteria acting on the actin cytoskeleton. *J. Nat. Cancer. Inst.* 90, 1559-1563.

61. **Poppe L. and Retey J. (2005).** Friedel-Crafts-type mechanism for the enzymatic elimination of ammonia from histidine and phenylalanine. *Angew.Chem.Int.Ed.* 44, 3668-3688.

62. **Walker K.D., Klettke K., Akiyama T., and Croteau R. (2004).** Cloning, heterologous expression, and characterization of a phenylalanine aminomutase involved in Taxol biosynthesis. *J.Biol.Chem.* 279, 53947-53954.

63. **Xiang L. and Moore B.S. (2002).** Inactivation, complementation, and heterologous expression of encP, a novel bacterial phenylalanine ammonia-lyase gene. *J.Biol.Chem.* 277, 32505-32509.

64. **Xiang L. and Moore B.S. (2003).** Characterization of Benzoyl Coenzyme A Biosynthesis Genes in the Enterocin-Producing Bacterium '*Streptomyces maritimus*'. *J.Bacteriol.* 185, 399-404.

65. **Williams J.S., Thomas M., and Clarke D.J. (2005).** The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. *Microbiology* 151, 2543-2550.

66. **Watts K.T., Mijts B.N., Lee P.C., Manning A.J., and Schmidt-Dannert C. (2006).** Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino Acid lyase family. *Chem. Biol.* 13, 1317-1326.

67. **Magarvey N.A., Fortin P.D., Thomas P.M., Kelleher N.L., and Walsh C.T. (2008).** Gatekeeping versus promiscuity in the early stages of the andrimid biosynthetic assembly line. *Acs Chemical Biology* 3, 542-554.

68. **Kyndt J.A., Meyer T.E., Cusanovich M.A., and Van Beeumen J.J. (2002).** Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein. *FEBS Lett.* 512, 240-244.

69. **Schroeder A.C., Kumaran S., Hicks L.M., Cahoon R.E., Halls C., Yu O., and Jez J.M. (2008).** Contributions of conserved serine and tyrosine residues to catalysis, ligand binding, and cofactor processing in the active site of tyrosine ammonia lyase. *Phytochemistry* 69, 1496-1506.

70. **Lin S., Van Lanen S.G., and Shen B. (2007).** Regiospecific chlorination of (*S*)-ß-Tyrosyl-*S*-Carrier protein catalyzed by SgcC3 in the biosynthesis of the enediyne antitumor antibiotic C-1027. *J.Am.Chem.Soc.* 129, 12432-12438.

71. **Van Lanen S.G., Dorrestein P.C., Christenson S.D., Liu W., Ju J., Kelleher N.L., and Shen B. (2005).** Biosynthesis of the beta-amino acid moiety of the enediyne antitumor antibiotic C-1027 featuring beta-amino acyl-S-carrier protein intermediates. *J.Am.Chem.Soc.* 127, 11594-11595.

72. **Van Lanen S.G., Oh T.J., Liu W., Wendt-Pienkowski E., and Shen B. (2007).** Characterization of the maduropeptin biosynthetic gene cluster from *Actinomadura madurae* ATCC 39144 supporting a unifying paradigm for enediyne biosynthesis. *J.Am.Chem.Soc.* 129, 13082-13094.

73. **Berner M., Krug D., Bihlmaier C., Vente A., Müller R., and Bechthold A. (2006).** Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. *J.Bacteriol.* 188, 2666-2673.

74. **Röther D., Merkel D., and Rétey J. (2000).** Spectroscopic evidence for a 4-methylidene imidazol-5-one in histidine and phenylalanine ammonia-lyases. *Angew.Chem.Int.Ed.* 39, 2462-2464.

75. **Röther D., Poppe L., Viergutz S., Langer B., and Retey J. (2001).** Characterization of the active site of histidine ammonia-lyase from Pseudomonas putida. *Eur.J.Biochem.* 268, 6011-6019.

76. **Christianson C.V., Montavon T.J., Van Lanen S.G., Shen B., and Bruner S.D. (2007).** The structure of L-tyrosine 2,3-aminomutase from the C-1027 enediyne antitumor antibiotic biosynthetic pathway. *Biochemistry* 46, 7205-7214.

77. **Christianson C.V., Montavon T.J., Festin G.M., Cooke H.A., Shen B., and Bruner S.D. (2007).** The mechanism of MIO-based aminomutases in beta-amino acid biosynthesis. *J.Am.Chem.Soc.* 129, 15744-15745.

78. **Christenson S.D., Wu W., Spies M.A., Shen B., and Toney M.D. (2003).** Kinetic analysis of the 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis. *Biochemistry* 42, 12708-12718.

79. **Calabrese J.C., Jordan D.B., Boodhoo A., Sariaslani S., and Vannelli T. (2004).** Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. *Biochemistry* 43, 11403-11416.

80. **Louie G.V., Bowman M.E., Moffitt M.C., Baiga T.J., Moore B.S., and Noel J.P. (2006).** Structural determinants and modulation of substrate specificity in phenylalanine-tyrosine ammonia-lyases. *Chem. Biol.* 13, 1327-1338.

81. **Gerth K., Irschik H., Reichenbach H., and Trowitsch W. (1982).** The myxovirescins, a family of antibiotics from *Myxococcus virescens* (Myxobacterales). *J.Antibiot.* 35, 1454-1459.

82. **Jansen R., Reifenstahl G., Gerth K., Reichenberg H., and Höfle G. (1983).** Myxalamide A,B,C und D, eine Gruppe homologer Antibiotika aus *Myxococcus xanthus* Mx x12 (Myxobacterales). *Liebigs Ann. Chem.* 1081-1095.

83. **Kunze B., Bedorf N., Kohl W., Höfle G., and Reichenbach H. (1989).** Myxochelin A, a new iron-chelating compound from *Angiococcus disciformis* (Myxobacterales). Production, isolation, physico-chemical and biological properties. *J.Antibiot.* 42, 14-17.

84. **Meiser P., Weissman K.J., Bode H.B., Krug D., Dickschat J.S., Sandmann A., and Müller R. (2008).** DKxanthene biosynthesis -- understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism. *Chem. Biol.* 15, 771-781.

85. **Wenzel S.C., Kunze B., Höfle G., Silakowski B., Scharfe M., Blöcker H., and Müller R. (2005).** Structure and biosynthesis of myxochromides S1-3 in *Stigmatella aurantiaca*: Evidence for an iterative bacterial type I polyketide synthase and for module skipping in nonribosomal peptide biosynthesis. *ChemBioChem* 6, 375-385.

86. **Irschik H., Trowitzsch-Kienast W., Gerth K., Höfle G., and Reichenbach H. (1988).** Saframycin Mx1, a new natural saframycin isolated from a myxobacterium. *J.Antibiot.* 41, 993-998.

87. **Kunze B., Reichenbach H., Augustiniak H., and Höfle G. (1982).** Isolation and identification of althiomycin from *Cystobacter fuscus* (myxobacterales). *J.Antibiot.* 35, 635-636.

88. **Trowitzsch-Kienast, W.** (**1993**). Cittilins: Bicyclic Isotrityrosines from *Myxococcus xanthus*. German Chemists' Society, 24th General meeting, Sep 5-11, 1993. 496-497. 1993. Conference Proceeding.

89. **Koehn F.E. (2008).** High impact technologies for natural products screening. *Prog.Drug Res.* 65, 175, 177-175, 210.

90. **Ojanperä S., Pelander A., Pelzing M., Krebs I., Vuori E., and Ojanperä I. (2006).** Isotopic pattern and accurate mass determination in urine drug screening by liquid chromatography/time-of-flight mass spectrometry. *Rapid Commun.Mass Spectrom.* 20, 1161-1167.

91. **Lavine B.K. (2000).** Clustering and classification of analytical data. *In:* Meyers, R. A. (Ed.), Encyclopedia of Analytical Chemistry: Applications, Theory, and Instrumentation, Wiley & Sons, New York, 1-20.

92. **Idborg-Bjorkman H., Edlund P.O., Kvalheim O.M., Schuppe-Koistinen I., and Jacobsson S.P. (2003).** Screening of biomarkers in rat urine using LC/electrospray ionization-MS and two-way data analysis. *Anal.Chem.* 75, 4784-4792.

93. **Zabriskie T.M., Klocke J.A., Ireland C.M., Marcus A.H., Molinski T.F., Faulkner D.J., Xu C., and Clardy J. (1986).** Jaspamide, a modified peptide from a Jaspis sponge, with insecticidal and antifungal activity. *J.Am.Chem.Soc.* 108, 3123-3124.

94. **Eggert U., Diestel R., Sasse F., Jansen R., Kunze B., and Kalesse M. (2008).** Chondramide C: Synthesis, configurational assignment, and structure-activity relationship studies. *Angew. Chem.Int.Ed.* 47, 6478-6482.

95. **Waldmann H., Hu T.S., Renner S., Menninger S., Tannert R., Oda T., and Arndt H.D. (2008).** Total Synthesis of Chondramide C and Its Binding Mode to F-Actin. *Angew.Chem.Int.Ed Engl.* 47, 6473-6477.

96. **Clugston S.L., Sieber S.A., Marahiel M.A., and Walsh C.T. (2003).** Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: The C-5 domain of tyrocidine synthetase is a C-D(L) catalyst. *Biochemistry* 42, 12095-12104.

97. **Ehmann D.E., Trauger J.W., Stachelhaus T., and Walsh C.T. (2000).** Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. *Chem.Biol.* 7, 765-772.

98. **Belshaw P.J., Walsh C.T., and Stachelhaus T. (1999).** Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* 284, 486-489.

99. **Liu W., Christenson S.D., Standage S., and Shen B. (2002).** Biosynthesis of the enediyne antitumor antibiotic C-1027. *Science* 297, 1170-1173.

100. **Van Lanen S.G., Lin S.J., Dorrestein P.C., Kelleher N.L., and Shen B. (2006).** Substrate specificity of the adenylation enzyme SgcC1 involved in the biosynthesis of the enediyne antitumor antibiotic C-1027. *J.Biol.Chem.* 281, 29633-29640.

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

102. **Steer D.L., Lew R.A., Perlmutter P., Smith A.I., and Aguilar M.I. (2002).** Beta-amino acids: versatile peptidomimetics. *Curr.Med.Chem.* 9, 811-822.

103. **Fischbach M.A., Lai J.R., Roche E.D., Walsh C.T., and Liu D.R. (2007).** Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. *Proc.Natl.Acad.Sci.U.S.A* 104, 11951-11956.

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

105. **Bordoli L., Kiefer F., Arnold K., Benkert P., Battey J., and Schwede T. (2009).** Protein structure homology modeling using SWISS-MODEL workspace. *Nat.Protoc.* 4, 1-13.

106. **Schwede T.F., Retey J., and Schulz G.E. (1999).** Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. *Biochemistry* 38, 5355-5361.

107. **Gloge A., Zon J., Kovari A., Poppe L., and Retey J. (2000).** Phenylalanine ammonia-lyase: The use of its broad substrate specificity for mechanistic investigations and biocatalysis - Synthesis of L-arylalanines. *Chem.Eur.J.* 6, 3386-3390.

108. **Wu B., Szymanski W., Wietzes P., de Wildeman S., Poelarends G.J., Feringa B.L., and Janssen D.B. (2009).** Enzymatic synthesis of enantiopure alpha- and beta-amino acids by phenylalanine aminomutase-catalysed amination of cinnamic acid derivatives. *ChemBioChem* DOI: 10.1002/cbic.200800586.

109. **Ritter H. and Schulz G.E. (2004).** Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. *Plant Cell* 16, 3426-3436.

110. **Zerikly M. and Challis G.L. (2009).** Strategies for the discovery of new natural products by genome mining. *ChemBioChem* DOI: 10.1002/cbic.200800389.

111. **Moore B.S. and Hertweck C. (2002).** Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat.Prod.Rep.* 19, 70-99.

112. **Dorrestein P.C., Bumpus S.B., Calderone C.T., Garneau-Tsodikova S., Aron Z.D., Straight P.D., Kolter R., Walsh C.T., and Kelleher N.L. (2006).** Facile Detection of Acyl and Peptidyl Intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions during Tandem Mass Spectrometry. *Biochemistry* 45, 12756-12766.

113. **Gaisser S., Trefzer A., Stockert S., Kirschning A., and Bechthold A. (1997).** Cloning of an avilamycin biosynthetic gene cluster from Streptomyces viridochromogenes Tü57. *J.Bacteriol.* 179, 6271-6278.

114. **Wilkinson B., Foster G., Rudd B.A., Taylor N.L., Blackaby A.P., Sidebottom P.J., Cooper D.J., Dawson M.J., Buss A.D., Gaisser S., Bohm I.U., Rowe C.J., Cortes J., Leadlay P.F., and Staunton J. (2000).** Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. *Chem.Biol.* 7, 111-117.

115. **Gaitatzis N., Silakowski B., Kunze B., Nordsiek G., Blöcker H., Höfle G., and Müller R. (2002).** The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J.Biol.Chem.* 277, 13082-13090.

116. **He J. and Hertweck C. (2005).** Functional analysis of the aureothin iterative type I polyketide synthase. *ChemBioChem* 6, 908-912.

117. **Olano C., Wilkinson B., Sanchez C., Moss S.J., Sheridan R., Math V., Weston A.J., Brana A.F., Martin C.J., Oliynyk M., Mendez C., Leadlay P.F., and Salas J.A. (2004).** Biosynthesis of the angiogenesis inhibitor borrelidin by Streptomyces parvulus Tu4055: Cluster analysis and assignment of functions. *Chem.Biol.* 11, 87-97.

118. **Stuart L.J., Buck J.P., Tremblay A.E., and Buist P.H. (2006).** Configurational analysis of cyclopropyl fatty acids isolated from *Escherichia coli*. *Org. Lett.* 8, 79-81.

119. **George K.M., Yuan Y., Sherman D.R., and Barry C.E. (1995).** The biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis* - Identification and functional analysis of Cmas 2. *J.Biol.Chem.* 270, 27292-27298.

120. **Kelly W.L., Boyne M.T., Yeh E., Vosburg D.A., Galonic D.P., Kelleher N.L., and Walsh C.T. (2007).** Characterization of the aminocarboxycyclopropane-forming enzyme CmaC. *Biochemistry* 46, 359-368.

121. **Gu L., Jia J., Liu H., Hakansson K., Gerwick W.H., and Sherman D.H. (2006).** Metabolic coupling of dehydration and decarboxylation in the curacin A pathway: Functional identification of a mechanistically diverse enzyme pair. *J.Am.Chem.Soc.* 128, 9014-9015.

122. **Julien B., Tian Z.Q., Reid R., and Reeves C.D. (2006).** Analysis of the ambruticin and jerangolid gene clusters of *Sorangium cellulosum* reveals unusual mechanisms of polyketide biosynthesis. *Chem.Biol.* 13, 1277-1286.

123. **Vaillancourt F.H., Yeh E., Vosburg D.A., O'Connor S.E., and Walsh C.T. (2005).** Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis. *Nature* 436, 1191-1194.

124. **Courtois F., Guerard C., Thomas X., and Ploux O. (2004).** *Escherichia coli* cyclopropane fatty acid synthase. *Eur.J.Biochem.* 271, 4769-4778.

125. **Glickman M.S., Cahill S.M., and Jacobs W.R. (2001).** The *Mycobacterium tuberculosis* cmaA2 gene encodes a mycolic acid trans-cyclopropane synthetase. *J.Biol.Chem.* 276, 2228-2233.

126. **Zäuner S., Zahringer U., Lindner B., Warnecke D., and Sperling P. (2008).** Identification and functional characterization of the 2-hydroxy fatty N-acyl-delta-3(E)-desaturase from *Fusarium graminearum*. *J.Biol.Chem.* 283, 36734-36742.

127. **Kim J., Choi J.N., Kim P., Sok D.E., Nam S.W., and Lee C.H. (2009).** LC-MS/MS Profiling-Based Secondary Metabolite Screening of *Myxococcus xanthus*. *J.Microbiol.Biotechnol.* 19, 51-54.

128. **Zhang N.R., Yu S., Tiller P., Yeh S., Mahan E., and Emary W.B. (2009).** Quantitation of small molecules using high-resolution accurate mass spectrometers - a different approach for analysis of biological samples. *Rapid Commun.Mass Spectrom.* 23, 1085-1094.

129. **Dunn W.B., Broadhurst D., Brown M., Baker P.N., Redman C.W.G., Kenny L.C., and Kell D.B. (2008).** Metabolic profiling of serum using Ultra Performance Liquid Chromatography and the LTQ-Orbitrap mass spectrometry system. *J.Chromatogr.B* 871, 288-298.

130. **Wishart D.S. (2008).** Applications of metabolomics in drug discovery and development. *Drugs R.D.* 9, 307-322.

131. **Bode H.B., Müller R. (2007).** Secondary metabolism in myxobacteria. *In:* Whitworth, D. (Ed.), Myxobacteria: multicellularity and differentiation, ASM Press, Chicago, 259-282.

132. **Wenzel S.C., Meiser P., Binz T., Mahmud T., and Müller R. (2006).** Nonribosomal peptide biosynthesis: Point mutations and module skipping lead to chemical diversity. *Angew.Chem.Int.Ed.* 45, 2296-2301.

133. **Bode H.B., Meiser P., Klefisch T., Socorro D.J.Cortina N., Krug D., Göhring A., Schwär G., Mahmud T., Elnakady Y.A., and Müller R. (2007).** Mutasynthesis-derived myxalamids and origin of the isobutyryl-CoA starter unit of myxalamid B. *ChemBioChem* 8, 2139-2144.

134. **Meiser P., Bode H.B., and Müller R. (2006).** The unique DKxanthene secondary metabolite family from the myxobacterium *Myxococcus xanthus* is required for developmental sporulation. *Proc.Natl.Acad.Sci.USA* 103, 19128-19133.

135. **Simunovic V., Zapp J., Rachid S., Krug D., Meiser P., and Müller R. (2006).** Myxovirescin biosynthesis is directed by hybrid polyketide synthases/ nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl CoA synthases and trans-acting acyltransferases. *ChemBioChem* 7, 1206-1220.

136. **Bode H.B., Ring M.W., Schwär G., Altmeyer M.O., Kegler C., Jose I.R., Singer M., and Müller R. (2009).** Identification of additional players in the alternative biosynthesis pathway to isovaleryl-CoA in the myxobacterium *Myxococcus xanthus*. *ChemBioChem* 10, 128-140.

137. **Schley C., Altmeyer M.O., Swart R., Müller R., and Huber C.G. (2006).** Proteome analysis of *Myxococcus xanthus* by off-line two-dimensional chromatographic separation using monolithic poly-(styrene-divinylbenzene) columns combined with ion-trap tandem mass spectrometry. *J.Proteome Res.* 5, 2760-2768.

138. **Vos M. and Velicer G.J. (2006).** Genetic population structure of the soil bacterium *Myxococcus xanthus* at the centimeter scale. *Appl.Environ.Microbiol.* 72, 3615-3625.

139. **Dettmer K., Aronov P.A., and Hammock B.D. (2007).** Mass spectrometry-based metabolomics. *Mass Spectrom.Rev.* 26, 51-78.

140. **Dunn W.B. (2008).** Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Phys.Biol.* 5, 11001.

141. **Benton H.P., Wong D.M., Trauger S.A., and Siuzdak G. (2008).** XCMS2: Processing tandem mass spectrometry data for metabolite identification and structural characterization. *Anal.Chem.* 80, 6382-6389.

E. Appendix

1. Author's effort in publications from chapter B

Chapter B-1

DK performed purification of recombinant CmdF and aminomutase activity assays, as well as the determination of chondramide stereochemistry and mass spectrometric analysis throughout this study. DK also contributed to the experimental design and interpretation of results and wrote parts of the manuscript.

Chapter B-2

DK and RM planned all experiments. DK performed the practical work and wrote the manuscript.

Chapter B-3

DK participated in the analytical work during this study.

Chapter B-4

DK and RM devised the experimental design of this study. DK carried out most of the chemical, molecular biological and analytical procedures and wrote the manuscript.

Chapter B-5

DK performed purification of recombinant Sam8 and enzyme kinetic analysis. DK participated in writing the paper.

Chapter B-6

DK carried out the microbiological and chemical procedures, developed the experimental design together with RM and performed data evaluation. Instrumentation and software was provided by Bruker Daltonik. DK wrote the manuscript.

Chapter B-7

DK devised the experimental setup, performed data evaluation and wrote major parts of the publication.

Chapter B-8

DK contributed the analytical part to this manuscript.

2. Weitere Publikationen

Zusätzlich zu den Pulikationen, die Teil dieser Arbeit sind, habe ich auch an den folgenden Veröffentlichungen mitgewirkt:

1. **Meiser P., Weissman K.J., Bode H.B., Krug D., Dickschat J.S., Sandmann A., and Müller R. (2008).** DKxanthene biosynthesis – understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism. *Chem.Biol.* 15, 771-781.

2. **Krug D., Zurek G., Schneider B., Bässmann C., and Müller R. (2007).** Analysis of secondary metabolites from myxobacteria using ESI-TOF-MS and PCA. *LC-GC Europe, The Applications Book March 2007*, 41-42.

3. **Schneiker S., Perlova O., Kaiser O., Gerth K., Alici A., Altmeyer M.O., Bartels D., Bekel T., Beyer S., Bode E., Bode H.B., Bolten C.J., Choudhuri J.V., Doss S., Elnakady Y.A., Frank B., Gaigalat L., Goesmann A., Groeger C., Gross F., Jelsbak L., Jelsbak L., Kalinowski J., Kegler C., Knauber T., Konietzny S., Kopp M., Krause L., Krug D., Linke B., Mahmud T., Martinez-Arias R., McHardy A.C., Merai M., Meyer F., Mormann S., Munoz-Dorado J., Perez J., Pradella S., Rachid S., Raddatz G., Rosenau F., Ruckert C., Sasse F., Scharfe M., Schuster S.C., Suen G., Treuner-Lange A., Velicer G.J., Vorholter F.J., Weissman K.J., Welch R.D., Wenzel S.C., Whitworth D.E., Wilhelm S., Wittmann C., Blöcker H., Pühler A., and Müller R. (2007).** Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat.Biotechnol.* 25, 1281-1289.

4. **Bode H.B., Meiser P., Klefisch T., Socorro D.J.Cortina N., Krug D., Göhring A., Schwär G., Mahmud T., Elnakady Y.A., and Müller R. (2007).** Mutasynthesis-derived myxalamids and origin of the isobutyryl-CoA starter unit of myxalamid B. *ChemBioChem* 8, 2139-2144.

5. **Simunovic V., Zapp J., Rachid S., Krug D., Meiser P., and Müller R. (2006).** Myxovirescin biosynthesis is directed by hybrid polyketide synthases/ nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl CoA synthases and trans-acting acyltransferases. *ChemBioChem* 7, 1206-1220.

6. **Perlova O., Fu J., Kuhlmann S., Krug D., Stewart F., Zhang Y., and Müller R. (2006).** Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*. *Appl.Environ.Microbiol.* 72, 7485-7494.

7. **Stephan S., Heinzle E., Wenzel S.C., Krug D., Müller R., and Wittmann C. (2006).** Metabolic physiology of *Pseudomonas putida* for heterologous production of myxochromide. *Process Biochemistry* 41, 2146-2152.

3. Selected conference contributions

VAAM Fachgruppentagung, 2005, Dresden

Poster “Functional and biochemical characterisation of ammonium-lyase-type enzymes in myxobacteria”.

Chemical Biology: Directing Biosynthesis, 2006, Cambridge/UK

Poster “Functional and biochemical characterisation of ammonium-lyase-type enzymes in myxobacteria”.

VAAM Fachgruppentagung, 2006, Tübingen

Talk “Functional and biochemical characterisation of ammonium-lyase-type enzymes from myxobacteria”.

Bruker Daltonik Seminartour, 2007, Hannover und Heidelberg

Vortrag “Analyse von Naturstoffen aus Myxobakterien mit ESI-TOF - Massenspektrometrie und PCA (Principal Component Analysis)“.

International Conference on the Biology of Myxobacteria, 2007, Granada/Spain

Talk “Analysis of *Myxococcus xanthus* secondary metabolite profiles: a study of intraspecies variation”.

VAAM Fachgruppentagung, 2007, Saarbrücken

Talk “The incorporation of urocanic acid into a PKS secondary metabolite from *Stigmatella aurantiaca* Sg a15”.

Biofine Conference, 2008, Freiburg

Poster “The incorporation of urocanic acid into a PKS secondary metabolite from *Stigmatella aurantiaca* Sg a15”.

VAAM Fachgruppentagung, 2008, Berlin

Poster “A diversity-oriented apporach to mining myxobacterial secondary metabolomes”.