

# **Establishing and engineering heterologous production systems for argyrimycin and $\alpha$ -pyrone antibiotics**

Dissertation  
zur Erlangung des Grades  
des Doktors der Naturwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät  
der Universität des Saarlandes

von

**Domen Pogorevc**

Saarbrücken

2019

Tag des Kolloquiums: 02.05.2019

Dekan: Prof. Dr. G. Kickelbick

Berichterstatter: Prof. Dr. R. Müller

Prof. Dr. C. Wittmann

Vorsitz: Prof. Dr. U. Kazmaier

Akad. Mitarbeiter: Dr. Y. Rebets

Die vorliegende Arbeit wurde von November 2014 bis Januar 2019 unter Anleitung von Herrn Prof. Dr. Rolf Müller am Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS) angefertigt.

# Acknowledgments

First, I would like to thank Prof. Dr. Rolf Müller for the opportunity that he offered me with this PhD position in his group. I am grateful for his supervision during this time and for his professional advice and discussions.

I would like to thank Prof. Dr. Christoph Wittmann for being my second supervisor and for reviewing my thesis.

A special thanks goes to Dr. Silke Wenzel, who was the best supervisor and one that I am very grateful for. She taught me a lot about molecular biology and her exceptional knowledge and expert opinion always helped me to advance with my projects and career.

I would like to thank all of the current and past group members for the collaborations and all the time we spent together.

I would also like to thank my family for their constant support during my time abroad.

The biggest thank you goes to Sophia for her support during all this time.

## Abstract

Myxobacterial natural products have proven to be a valuable source of antibacterial and anticancer compounds in the last decades. This thesis covers the biosynthesis elucidation, heterologous production as well as structure and yield optimization of two potent myxobacterial compounds. The work on coralloypronin resulted in the design of an efficient heterologous production platform in the myxobacterial model organism *Myxococcus xanthus*. Research on the established heterologous producer led to identification of specific biochemical processes and resulted in isolation and structure elucidation of two novel coralloypronin derivatives. Furthermore, gene deletion experiments led to discovery of hypothetical self-resistance mechanism, which seems to be conserved in most myxobacterial RNA polymerase (RNAP) inhibitor biosynthetic gene clusters (BGC). The BGC of argyirin, a cyclic octapeptide produced by nonribosomal peptide synthetase (NRPS), was identified in *Cystobacter sp.* SBCb004. The heterologous expression of synthetic argyirin BGC was achieved in *M. xanthus*, which, after medium optimization, led to a significant yield improvement over the wild type producer. The production of argyirin could be even further improved by utilization of different promoter systems with a specifically optimized leader sequence. Furthermore, precursor directed biosynthesis was applied to produce a library of novel argyirin derivatives.

## Zusammenfassung

Myxobakterielle Naturstoffe haben sich in den letzten Jahrzehnten als wertvolle Ressource für Wirkstoffe gegen Bakterien und Krebs bewährt. Diese Dissertation umfasst die Biosyntheseaufklärung, die heterologe Produktion sowie Struktur und Ausbeuteoptimierung zweier wirksamer myxobakterieller Verbindungen. Die Arbeit über Corallopyronin führte zum Design einer effizienten myxobakteriellen Produktionsplattform im myxobakteriellen Modellorganismus *Myxococcus xanthus*. Die Erforschung dieses etablierten heterologen Produzenten führte zur Identifizierung spezifischer biochemischer Prozesse, die in der Isolation und Strukturaufklärung zweier neuer Corallopyroninderivate resultierte. Zusätzliche Gendeletionsexperimente führten zur Entdeckung eines hypothetischen Selbstresistenzmechanismus, der in myxobakteriellen RNA Polymerase (RNAP) Inhibitor Biosynthese-Genclustern (BGC) konserviert zu sein scheint. Das BGC des zyklischen Oktapeptids Argyrin wurde in einem Nichtribosomalen Peptid Synthetase (NRPS) Gencluster in *Cystobacter sp.* SBCb004 entdeckt. Die heterologe Expression eines synthetischen Argyrin BGC in *M. xanthus* führte nach Medienoptimierung zu einer signifikanten Steigerung der Ausbeute im Vergleich zum Wildtyp. Zusätzliche Steigerungen der Argyrinproduktion wurden durch Integration verschiedener Promotoren mit einer optimierten Leader Sequenz erreicht. Zudem konnte durch die Fütterung spezieller Vorläufermoleküle eine Bibliothek neuer Argyrinderivate produziert werden.

## Vorveröffentlichungen der Dissertation

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

### Publications

Sucipto, H.\*; **Pogorevc, D.\***; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, *44*, pp. 160–170.

**Pogorevc, D.\***; Tang, Y.\*; Hoffmann, M.; Zipf, G.; Bernauer, H. S.; Popoff, A.; Steinmetz, H.; Wenzel, S. C. Biosynthesis and heterologous production of argyrins. **2019**, *submitted manuscript*.

**Pogorevc, D.**; Panter, F.; Schillinger, C.; Jansen, R.; Wenzel, S. C.; Müller, R. Production optimization and biosynthesis revision of coralopyronin A, a potent anti-filarial antibiotic. **2019**, *submitted manuscript*.

(\*: Authors contributed equally to the work)

### Conference Contributions (Posters and Oral Presentations)

**D. Pogorevc**, H. Sucipto, E. Luxenburger, S. C. Wenzel and R.Müller. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*. **Poster Presentation**, 7<sup>th</sup> International HIPS Symposium, 2017: Saarbrücken, Germany.

**D. Pogorevc**, H. Sucipto, E. Luxenburger, S. C. Wenzel and R.Müller. Heterologous production of coralopyronin A in *Myxococcus xanthus*. **Poster Presentation**, 8<sup>th</sup> International HIPS Symposium, 2018: Saarbrücken, Germany.

A. Schiefer, A. Krome, M. Hübner, S. Kehraus, S. Hüttel, R. Jansen, **D. Pogorevc**, S. Bouhired, T. Schaberle, M. Roth, G. König, C. Keller, J. Rupp, R. Müller, K. Wagner, M. Stadler, K. Pfarr, A. Hoerauf. Pre-clinical development of the antibiotic Coralopyronin A. **Poster Presentation**, 8<sup>th</sup> International HIPS Symposium, 2018: Saarbrücken, Germany.

**D. Pogorevc**, H. Sucipto, E. Luxenburger, S. C. Wenzel and R.Müller. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*. **Poster Presentation**, Summer Symposium of Interdisciplinary Graduate School, 2017: Saarbrücken, Germany.

**D. Pogorevc**, H. Sucipto, E. Luxenburger, S. C. Wenzel and R.Müller. Heterologous production of coralopyronin A in *Myxococcus xanthus*. **Poster Presentation**, Summer Symposium of Interdisciplinary Graduate School, 2018: Saarbrücken, Germany.

## Table of Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction.....</b>  | <b>1</b>  |
| 1.1      | Natural products .....  | 1         |
| 1.2      | Myxobacteria as producers of natural products.....  | 3         |
| 1.3      | Multifunctional megasynthetases for natural product biosynthesis .....  | 6         |
| 1.3.1    | Polyketide synthases (PKS) .....  | 8         |
| 1.3.2    | Structural diversity in PKS biosynthesis.....   | 9         |
| 1.3.3    | Unusual biochemistry of trans-AT PKS systems.....   | 10        |
| 1.3.4    | Nonribosomal peptide synthetases (NRPS) .....   | 14        |
| 1.3.5    | Structural diversity in NRPS biosynthesis .....   | 15        |
| 1.4      | Heterologous expression of complex natural product pathways.....  | 16        |
| 1.5      | Combinatorial biosynthesis approaches .....   | 18        |
| 1.6      | Outline of the present work .....   | 19        |
| 1.7      | References .....  | 22        |
| <b>2</b> | <b>Heterologous production of myxobacterial <math>\alpha</math>-pyrone antibiotics in <i>Myxococcus xanthus</i> .....</b> | <b>33</b> |
| 2.1      | Abstract.....   | 33        |
| 2.2      | Introduction .....  | 33        |
| 2.3      | Material and methods .....  | 37        |
| 2.3.1    | Sequence analysis and <i>mxn</i> and <i>cor</i> gene cluster sequences.....   | 37        |
| 2.3.2    | Construction and engineering of plasmids.....   | 37        |
| 2.3.3    | Generation of <i>M. fulvus</i> Mx f50 mutant strains and subcloning of <i>mxn</i> gene cluster fragments .....            | 40        |
| 2.3.4    | Transfer and chromosomal integration of the expression constructs into the heterologous host.....                         | 42        |
| 2.3.5    | Initial cultivation experiments with the heterologous expression strains in comparison to the native producers.....       | 43        |
| 2.3.6    | Additional cultivation experiments with the heterologous expression strains in different production media.....            | 43        |
| 2.3.7    | Analysis of myxopyronin and coralopyronin production by HPLC-MS .....   | 45        |
| 2.4      | Results and discussion .....  | 45        |

---

|          |  |           |
|----------|--|-----------|
| 2.4.1    | Subcloning of <i>mxn</i> gene cluster fragments from <i>M. fulvus</i> Mx f50.....  | 46        |
| 2.4.2    | Construction of a <i>mxn</i> gene cluster expression vector .....  | 47        |
| 2.4.3    | Heterologous myxopyronin production in <i>M. xanthus</i> .....   | 48        |
| 2.4.4    | Reconstitution of the <i>corK-O</i> gene cluster part from cosmids.....  | 49        |
| 2.4.5    | Construction of a <i>mxn/cor</i> hybrid gene cluster expression vector .....   | 51        |
| 2.4.6    | Heterologous coralloyronin production in <i>M. xanthus</i> .....   | 52        |
| 2.4.7    | Improvement of myxopyronin/coralloyronin production yields.....  | 53        |
| 2.5      | Conclusions .....  | 55        |
| 2.6      | References .....   | 56        |
| 2.7      | Supplementary information .....  | 58        |
| <b>3</b> | <b>Production optimization and biosynthesis revision of coralloyronin A, a potent anti-filarial antibiotic.....</b>              | <b>71</b> |
| 3.1      | Abstract.....  | 71        |
| 3.2      | Introduction .....   | 72        |
| 3.3      | Materials and methods.....   | 73        |
| 3.3.1    | Sequence analysis of Coralloyronin biosynthetic gene clusters (BGC).....   | 73        |
| 3.3.2    | Construction and engineering of plasmids .....   | 73        |
| 3.3.3    | Transfer and chromosomal integration of the expression constructs into the heterologous host.....                                | 75        |
| 3.3.4    | Cultivation of the heterologous expression strains .....   | 76        |
| 3.4      | Results and discussion .....   | 77        |
| 3.4.1    | Construction of a transposon-based COR BGC expression construct.....   | 77        |
| 3.4.2    | Production of coralloyronin by an inducible system after transposition into the heterologous host <i>M. xanthus</i> DK1622 ..... | 79        |
| 3.4.3    | Evaluation of CorM function .....  | 80        |
| 3.4.4    | Deletion of <i>corO</i> .....  | 81        |
| 3.4.5    | Overexpression of CorO: An attempt to improve conversion of preCOR A to COR A and thus increase COR A production.....            | 82        |
| 3.4.6    | Evaluation of the function of CorN.....  | 85        |
| 3.4.7    | Evaluation of the function of CorP .....   | 89        |
| 3.5      | Conclusion.....  | 92        |

---

|          |  |            |
|----------|--|------------|
| 3.6      | References .....   | 93         |
| 3.7      | Supplementary information .....  | 97         |
| <b>4</b> | <b>Biosynthesis and heterologous production of argyrins.....</b>   | <b>129</b> |
| 4.1      | Abstract.....  | 129        |
| 4.2      | Introduction .....   | 130        |
| 4.3      | Results and Discussion .....   | 131        |
| 4.3.1    | Identification and characterization of the argyirin biosynthetic pathway .....   | 131        |
| 4.3.2    | Establishment of a heterologous expression system based on synthetic DNA   | 135        |
| 4.3.3    | Approaches to engineer the argyirin production profile .....   | 139        |
| 4.3.4    | Approaches to further improve the heterologous expression system.....  | 143        |
| 4.4      | Conclusion .....   | 146        |
| 4.5      | Materials and methods.....   | 146        |
| 4.5.1    | Identification and sequence analysis of the arg BGC from <i>Cystobacter sp.</i><br>SBCb004 .....                                 | 146        |
| 4.5.2    | Strains, transformation and cultivation conditions.....  | 147        |
| 4.5.3    | General methods for DNA manipulation, analysis and PCR.....  | 148        |
| 4.5.4    | Studies with the native argyirin producer strains .....  | 148        |
| 4.5.5    | Sequence design and assembly of synthetic <i>arg</i> BGCs .....  | 149        |
| 4.5.6    | Heterologous expression in <i>M. xanthus</i> and argyirin production analysis.....   | 150        |
| 4.5.7    | NRPS A1 domain engineering .....   | 150        |
| 4.6      | References .....   | 153        |
| 4.7      | Supplementary information .....  | 159        |
| <b>5</b> | <b>Production profile engineering and precursor directed<br/>biosynthesis for production of novel argyirin derivatives .....</b> | <b>211</b> |
| 5.1      | Introduction .....   | 211        |
| 5.2      | Results and discussion.....  | 212        |
| 5.2.1    | Argyirin production profile engineering .....  | 212        |
| 5.2.2    | Precursor directed biosynthesis for production of novel argyirin derivatives...  | 216        |
| 5.2.3    | Optimization of argyirin production in <i>M. xanthus</i> DK1622 by utilization of new<br>promoter systems.....                   | 220        |
| 5.3      | Conclusion.....  | 225        |

---

|          |   |            |
|----------|---|------------|
| 5.4      | References .....  | 226        |
| 5.5      | Supplementary information .....   | 229        |
| <b>6</b> | <b>Discussion .....</b>   | <b>275</b> |
| 6.1      | Yield improvement and production profile engineering of microbial natural products<br>275 |            |
| 6.1.1    | Heterologous expression of myxobacterial natural product megasynthetases.                 | 276        |
| 6.1.2    | Medium optimization .....   | 279        |
| 6.1.3    | Genetic engineering techniques .....  | 281        |
| 6.2      | Structure engineering of microbial natural products .....                                 | 285        |
| 6.2.1    | Precursor-directed biosynthesis and mutasynthesis approaches .....                        | 285        |
| 6.2.2    | Engineering of tailoring genes .....  | 290        |
| 6.2.3    | Engineering of PKS/NRPS megasynthetases.....  | 292        |
| 6.3      | Conclusion and outlook .....  | 296        |
| 6.4      | References .....  | 297        |



# 1 Introduction

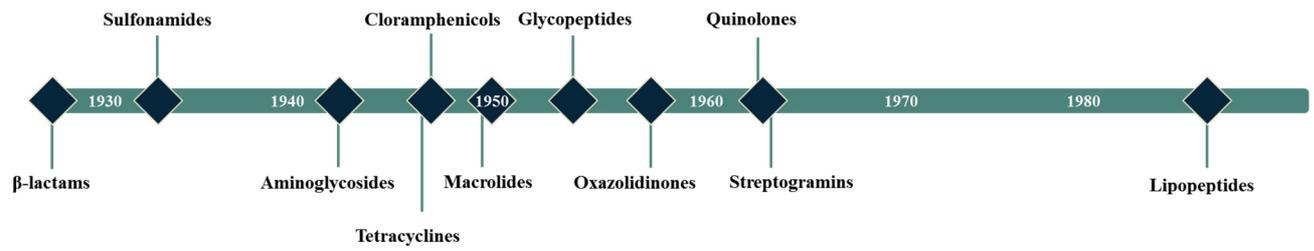
## 1.1 Natural products

By broad definition natural products (NP) include everything produced by living systems<sup>1</sup>, however a more detailed definition of the term usually refers to secondary metabolites, or those not required for growth of the producing organisms<sup>2</sup>. The history of NPs from plants, animals, bacteria or fungi extends back into the prehuman era. Millions of years of evolution resulted in a complex and ever-evolving variety of NPs that exists today. First records of NPs used by humans dated back into 2600 BC, documenting an extensive list of around 1000 plant-derived substances in Mesopotamia. NPs were used in their crude form, to treat a variety of illnesses ranging from colds to parasitic infections and inflammations since the first use of medicine by ancient human civilizations<sup>3</sup>.

With the development of modern medicine and transition into the “golden era of antibiotics”, an extensive investigation of new active microbial NPs began. The discovery of penicillin, the first  $\beta$ -lactam antibiotic, by Alexander Fleming in 1928 opened up new opportunities in the field of NPs as the vast source of yet undiscovered anti-infective agents began to unravel. Since then, until late 1970s a number of different antibiotic classes were revealed, majority of them originating from microorganisms (Fig. 1). Discovery of streptomycin, gentamicin, tetracycline and other antibiotics prompted the pharmaceutical companies to start expanding their NP research programs<sup>4</sup>. Microorganisms quickly became one of the main sources of active metabolites with a variety of different activities. Many of those metabolites were found to be highly potent due to the evolutionary pressure by which their chemical optimisation was driven<sup>4</sup>. NPs play a pivotal role in the development of novel anti-infectives and anticancer agents. A large portion of the small-molecule drugs approved between 1981 and 2014 (43.6 % of anti-infectives and 40.7 % of anticancer agents) was based on natural products and their derivatives<sup>5</sup>. Some of the most important NPs throughout the history exert various biological activities, including antibacterial activity (e.g. penicillin, vancomycin, erythromycin, rifamycin)<sup>6,7</sup>, cytotoxic activity (e.g. bleomycin and doxorubicin)<sup>6</sup>, antifungal activity (e.g. amphotericin<sup>7</sup>, griseofulvin<sup>6</sup>) as well as immunosuppressive effects (e.g. cyclosporin<sup>6</sup>) (Fig. 2).

## 2 | Introduction

---



**Figure 1.** Discovery timeline of novel antibiotic classes.

**Figure 2.** Chemical structures of some important natural products with various activity profiles.

As more and more antibiotics reached the market in the 1960 and 70s, the need for novel anti-infectives seemed to decrease and pharmaceutical companies began to lose interest in this field<sup>8,9</sup>. The discovery of novel antibiotics has therefore been in steady decline since the peak of anti-infectives discovery in the middle of the 20<sup>th</sup> century (Fig. 2). In the period from 1970 to 1990 only one broad spectrum class of antibiotics, the fluoroquinolones, was introduced to the market<sup>9</sup>. Due to rapidly emerging bacterial antibiotic resistance, novel pharmaceuticals are once again urgently needed for treating infectious diseases<sup>10,11</sup>.

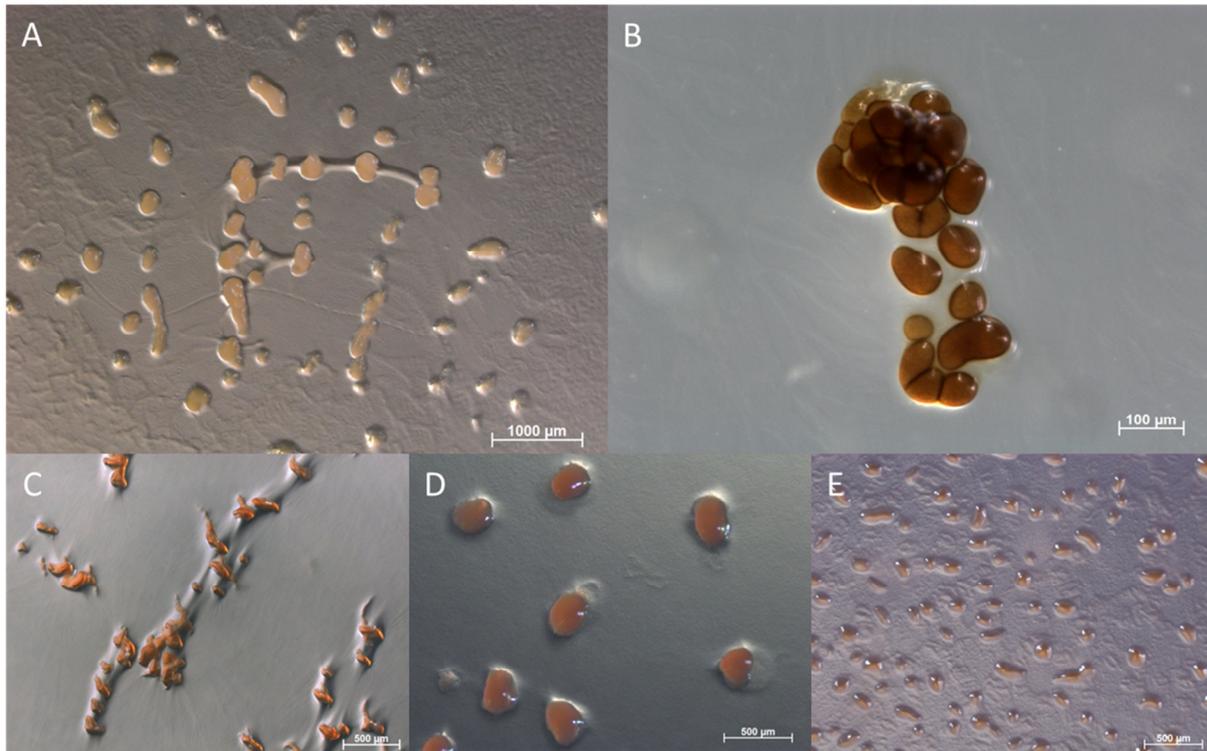
## 1.2 Myxobacteria as producers of natural products

Microorganisms have been the leading force in production of antibiotics. Especially important are actinobacteria which helped establish microorganisms as prominent anti-infective producers with the discovery of streptothricin and streptomycin in the 1940s<sup>12-14</sup>. To date, more than 5000 antibiotics have been identified from the order of actinomycetales, with over 90% of these produced solely by the *Streptomyces* genus<sup>12</sup>. Other microbial producers, like fungi and bacilli, have also been established as prolific sources for NPs<sup>15,16</sup>, the latter are known by producing a special class of antibiotics called lantibiotics<sup>12</sup>. In the last decades, myxobacteria acquired a reputation of prominent secondary metabolite producers by contributing a sizable share of NPs<sup>17,18</sup>. From the discovery of ambruticin in 1970s by Warner & Lambert<sup>19</sup> until 2010, more than 100 distinct core structures were isolated from myxobacteria and many of them were shown to exhibit intriguing bioactivities<sup>20-22</sup>.

Myxobacteria are Gram-negative bacteria belonging to the group of  $\delta$ -proteobacteria<sup>23</sup>. Similar to the Streptomycetes, they are also known as soil residing organisms, famous for their rich secondary metabolism, originating from a vast number of biosynthetic gene clusters (BGCs) encoded in their genomes<sup>24,17,25</sup>. They have one of the largest genomes of all known prokaryotes, ranging from 9 to 14.8 Mb<sup>24,26,25</sup>. It is speculated that those large genome sizes are necessary for their complex way of life<sup>26</sup> as it has been reported that bacteria showing such multicellular behavior typically possess large proteomes<sup>26,27</sup>.

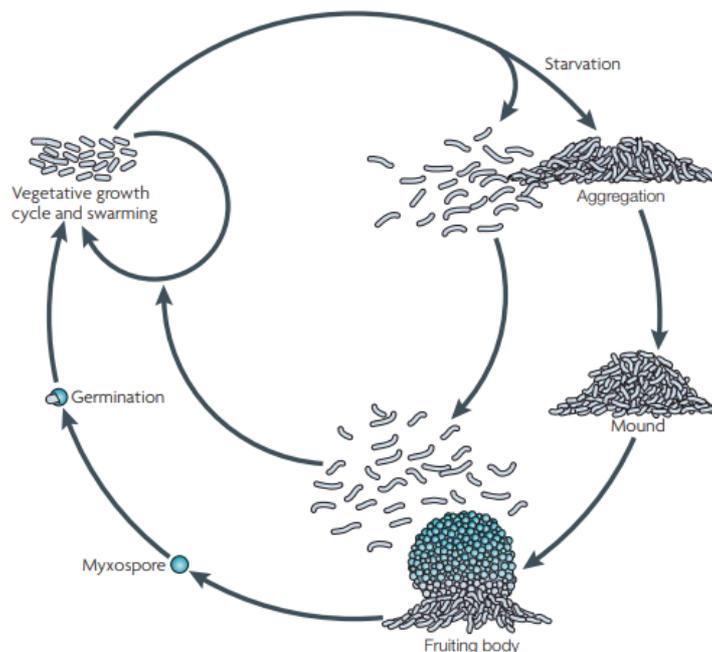
Myxobacteria are known for several intriguing multicellular, social behaviors. Some of those include motility in form of swarming and predatory behavior. Both behaviours are somewhat connected, since the detection of nutrients (e.g. prey or other metabolites) triggers gliding motility in the corresponding direction. As the swarm connects with the foreign colony, myxobacteria lyse opposing cells by releasing antibiotics and exoenzymes. Nutrients released from digested prey cells are taken up as source of energy for further growth and expansion of the colony<sup>26</sup>. When myxobacteria deplete their prey and are subjected to nutrient poor

conditions, they start a differentiation process that results in two types of cells. Some cells turn into a monolayer of cells called peripheral rods<sup>28–30</sup>, while most cells aggregate and form intriguing biofilm-like structures ranging from simple mounds to complicated three dimensional structures called fruiting bodies<sup>26,31</sup> (Fig. 3).



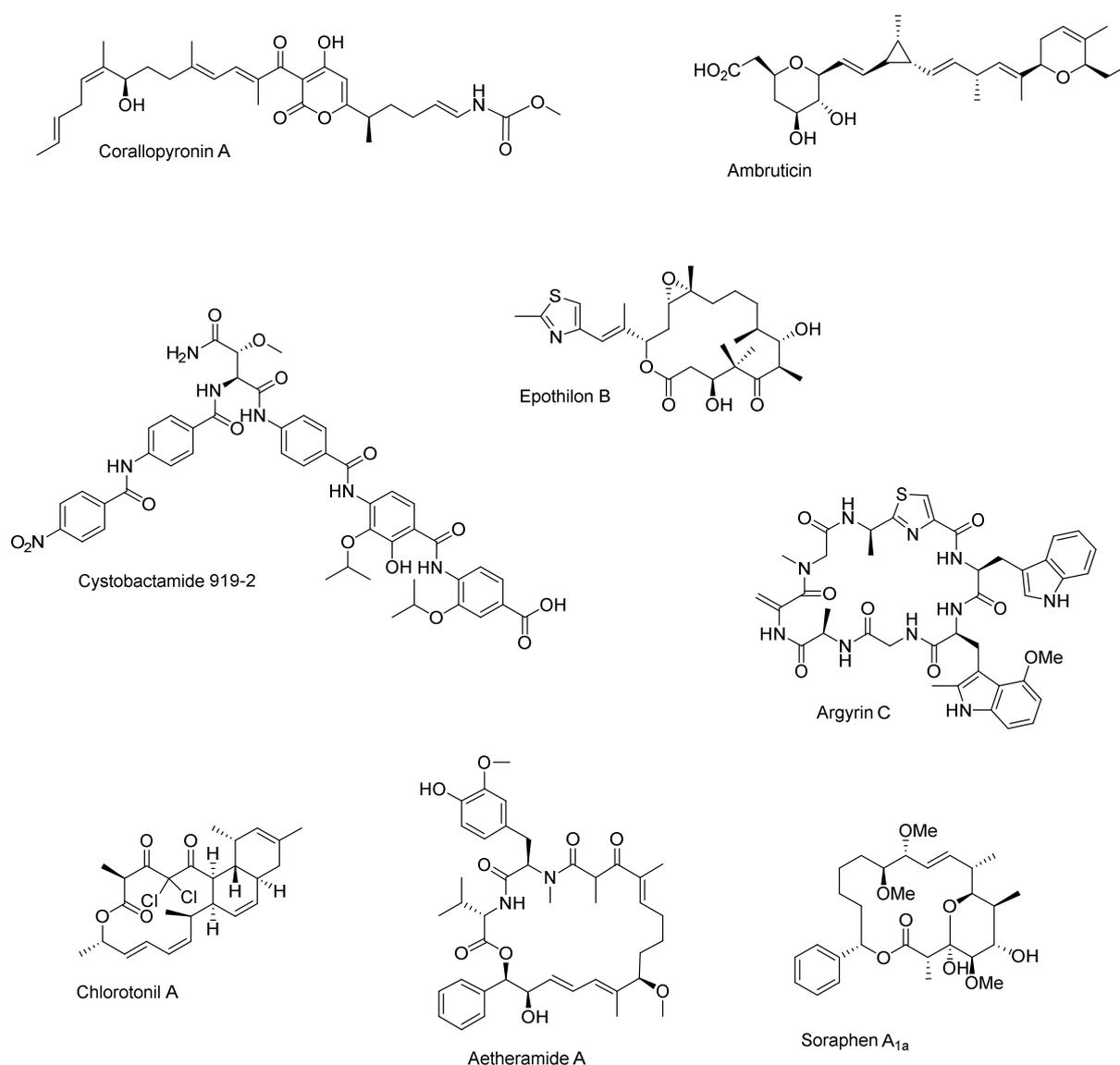
**Figure 3.** Fruiting bodies of *Myxococcus xanthus* DK1622 (A), *Cystobacter* SBCb004 (B), *Archangium cellulosum* Ar8082 (C), *Myxococcus fulvus* Mx f50 (D) and *Coralloccoccus coralloides* Cc c127 (E).

To cope with the depleted conditions, cells within fruiting bodies develop specific functions. Some develop into hibernating or non-reproductive cells while others differentiate into resistant, reproductive myxospores (Fig. 4). Sporulation in *M. xanthus* differs from endospore formation in *Bacillus* spp. in terms that the entire cell converts into a myxospore<sup>28</sup>. During the starvation conditions, cells that do not aggregate into fruiting bodies or form the surrounding monolayer of peripheral rods often lyse. It has been suggested that the released nutrients serve as source of energy for peripheral rods that act as a defensive line, protecting fruiting bodies from approaching microorganisms<sup>28</sup>.



**Figure 4.** Life cycle of *Myxococcus xanthus*<sup>32</sup>.

Myxobacteria are also metabolically very well equipped for fending off intruders or competitors in their natural habitat. With their rich secondary metabolism they are a great alternative to streptomycetes, which are still one of the biggest sources for microbial NPs today<sup>12</sup>. These natural products exhibit some common biological activities like antifungal (e.g. soraphen<sup>33</sup> and ambruticin<sup>19</sup>), antibacterial (e.g. cystobactamid<sup>34</sup>), antitumor (e.g. epothilone<sup>35</sup>) as well as some more unusual ones like antiviral (e.g. aetheramide<sup>36</sup>), antimalarial (e.g. chlorotonil<sup>37</sup>), immunomodulatory (e.g. argyirin<sup>38</sup>) and antifilarial (e.g. corallopyronin<sup>39</sup>) (Fig. 5). Due to their potent bioactivity profiles, myxobacterial NPs play an important role in drug development. One of the most important myxobacterial compounds is epothilon as its semisynthetic derivative ixabepilone has already been approved for breast cancer treatment by the FDA<sup>40</sup>. Some other myxobacterial anticancer agents include e.g. bengamide, a derivative of which was studied in anticancer phase I clinical trial<sup>41</sup> and tubulysin, which also showed promising results in preclinical studies<sup>42</sup>. Chlorotonil is a macrolactone antibiotic isolated from *Sorangium cellulosum* So ce1525. It shows potent *in vitro* activity against Gram-positive pathogens, however, even more interesting is the activity against *Plasmodium falciparum*. In an *in vivo* study with *P. berghei*, chlorotonil led to a substantial reduction of parasitemia in mice<sup>37</sup>. Another very potent antibacterial agent from myxobacteria, cystobactamide, was identified as an inhibitor of bacterial topoisomerase<sup>34</sup>. This antibiotic is of high interest as it showed high inhibitory activity against *E. coli* in *in vitro* tests<sup>21</sup>.



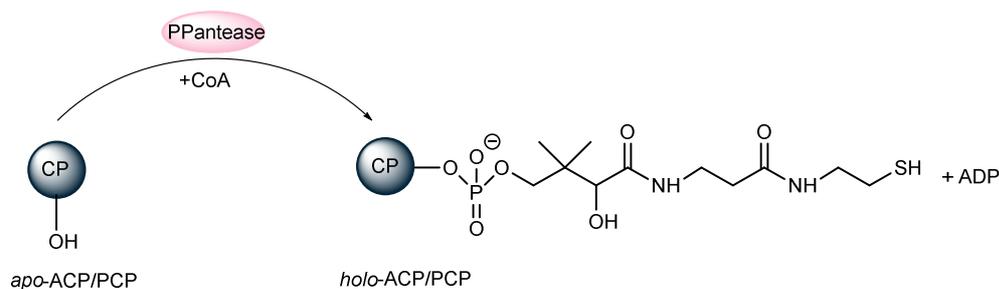
**Figure 5.** Myxobacterial secondary metabolites.

### 1.3 Multifunctional megasynthetases for natural product biosynthesis

Most of the myxobacterial NPs characterized so far are assembled by giant megasynthetases. Those can be classified as either polyketides (PKs), nonribosomal peptides (NRPs), or hybrids thereof (i.e. PK-NRP hybrids)<sup>43</sup>, while other structural types e.g. include phenyl-propanoids and alkaloids<sup>20</sup>. Modular PKs and NRPs are assembled in a similar manner. They consist of monomer units coupled into longer chains by large multimodular enzymes commonly referred to as megasynthetases<sup>44,43</sup>. Megasynthetases are composed of multifunctional subunits containing various catalytic domains, which are organized in so-called modules. Each of these modules is responsible for incorporation of corresponding building blocks into a growing

ketide or peptide chain, which is transferred along the entire assembly line until its release<sup>43</sup>. Modules consists of multiple domains and each domain has a specific catalytic function, which makes it responsible for a specific reaction step in the chain elongation cascade<sup>43</sup>. Polyketide synthases (PKS) use acyl-coenzyme A (acyl-CoA) precursors as building blocks, whereas nonribosomal peptide synthetases (NRPS) employ a significantly larger and more diverse group of monomers, by accepting proteinogenic and non-proteinogenic amino acids as well as aryl acids<sup>43,45</sup>. Due to structural and catalytic similarities of PKS and NRPS, it is biochemically possible for those systems to interact and form functional hybrid systems<sup>46</sup>. The results are so called hybrid NRPS/PKS megasynthetases, which represent a significant share of myxobacterial/microbial natural product biosynthetic pathways<sup>47</sup>.

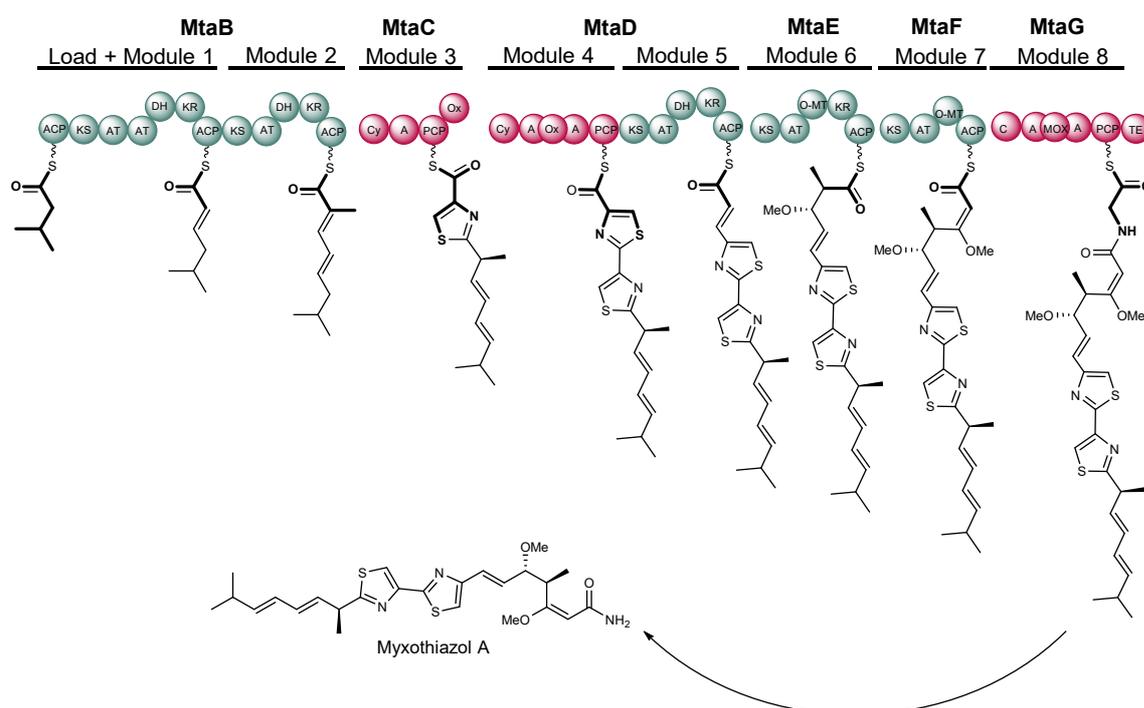
During the assembly process the precursors and growing intermediates are covalently bound to the megasynthetase. The main domain responsible for binding of the building blocks is called carrier protein (CP) domain. For the CP to perform its intended function, a posttranslational modification called phosphopantetheinylation is necessary, to convert the inactive *apo* form of the CP into its active *holo* form. In this process, a superfamily of proteins called phosphopantetheinyl transferases (PPantases) transfer the 4'-phosphopantetheine moiety (PPant), also called "PPant arm" or "prosthetic arm", from coenzyme A to a catalytic serine in the active site of the ACP<sup>48</sup>. Once the CP has been PPantenylation, the PPant group serves as an attachment site to covalently tether chain intermediates and shuttle them from/to other catalytic domains (e.i. AT, KS)<sup>48</sup> (Fig. 6).



**Figure 6.** Transformation of the carrier protein *apo* form to *holo* form by phosphopantetheinylation.

An example of myxobacterial compound produced by a megasynthetase is myxothiazol. It is a natural product with antifungal properties produced by *Angiococcus*, *Stigmatella*, and *Myxococcus* genera of myxobacteria. It acts on the cytochrome *bc<sub>1</sub>* complex and is a potent inhibitor of the respiratory chain. The megasynthetase responsible for production of myxothiazol is a hybrid PKS/NRPS with several unusual features. The biosynthesis begins with an incorporation of the unusual starter unit isovalerate, which is further condensed with malonate, methylmalonate and cysteine units<sup>49</sup>. The Ox domain in *mtaC* was shown to be

superfluous by the deletion experiment, therefore it's assumed that the Ox domain of *mtaD* is responsible for oxidation of both thazolines<sup>50</sup>. The assembly line consists of *mtaB*, *mtaE*, *mtaF*, encoding the PKS part, *mtaC* and *mtaG*, encoding NRPS part and the gene *mtaD* consists of NRPS and PKS elements<sup>49</sup> (Fig. 7). The terminal amide in myxothiazol structure is formed by an unusual release mechanism. It is predicted that the monooxygenase (MOX) domain present in MtaG oxidizes the last module intermediate. Dealkylation of the resulting  $\alpha$ -hydroxylated intermediate leads to the release of myxothiazol, whereas the terminal amino acid is cleaved from the PCP by subsequent action of the TE domain<sup>51,52</sup>.

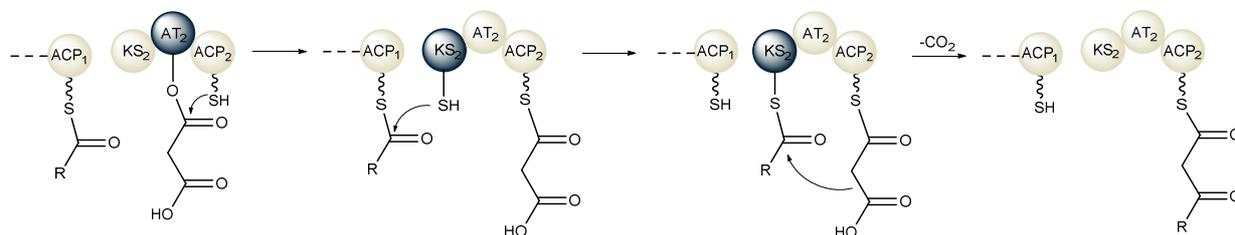


**Figure 7.** The assembly line of myxothiazol. PKS modules are highlighted in green and NRPS modules in red. Catalytic domains are illustrated as circles.

### 1.3.1 Polyketide synthases (PKS)

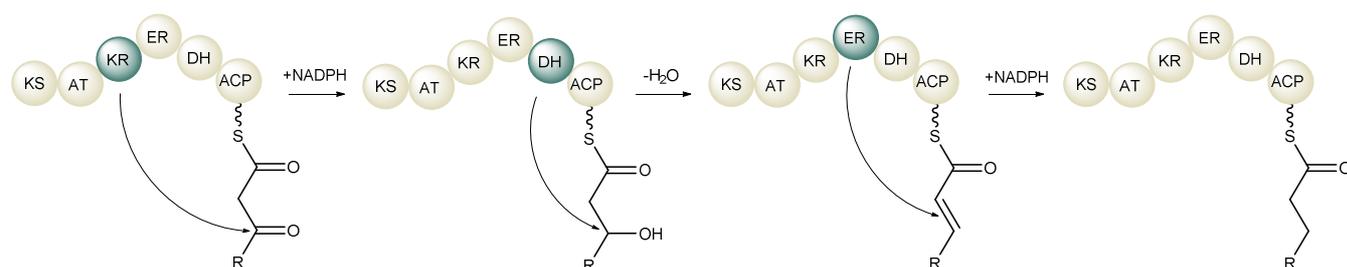
The biochemistry of PKS is similar to that of the fatty acid synthases (FAS) and both groups are evolutionarily connected<sup>53</sup>, which becomes quite clear once we take a closer look at their biosynthesis. In essence, both machineries follow the same concept of connecting monomer acyl units into a long ketide chain. The chain is formed in a cascade of Claisen condensation reactions responsible for interconnection of activated acyl starter units with malonyl-CoA-derived extender units<sup>44</sup>. Typically, the most common starter units are malonyl-CoA and methyl malonyl-CoA. The Claisen condensation is an essential reaction for chain elongation, which requires at least three catalytic domains: acyltransferase domain (AT) responsible for selection of starter and extender units, an acyl carrier protein domain (ACP)

posttranslationally activated via phosphopantetheinylation<sup>48</sup> to shuttle monomer units and elongated intermediates and a ketosynthase domain (KS) for C-C bond formation (Fig. 8). After the full length chain has been assembled, the thioesterase (TE) domain in the final module catalyzes its release from the assembly line by hydrolysis or lactonization<sup>54</sup>.



**Figure 8.** Polyketide chain elongation. The catalytic domain in each step is highlighted.

Three additional domains have to be employed in order to fully reduce the initially generated  $\beta$ -keto group after chain elongation. Those include the ketoreductase domain (KR) which generates a  $\beta$ -hydroxyl group, the dehydratase domain (DH) responsible for dehydration of the hydroxyl group to form a double bond which is further reduced by the enoylreductase domain (ER)<sup>55</sup> (Fig. 9). The main difference between FAS and PKS is that each incorporated monomer unit in FAS is always fully reduced<sup>55</sup>, whereas in modular PKS not all modifications necessarily take place<sup>44</sup>. This leads to higher diversity of PKS products, which compared to fatty acids differ in chain lengths and consist of structural differences such as various hydroxyl or keto groups and double bonds. Additional functional domains, enzymatic reactions and various system architectures can further increase the diversity of PKS<sup>44</sup>.



**Figure 9.** Optional  $\beta$ -keto processing steps of PKS.

### 1.3.2 Structural diversity in PKS biosynthesis

There are several factors contributing to the structural diversity of the PKS. Some of this remarkable diversity comes from combinatorial utilization of only a few simple building blocks. The major part of the monomer building blocks used in PKS chain elongation is represented by malonyl-CoA and methyl malonyl-CoA, however, the starter units show much

larger diversity in form of e.g. thioesters of monoacyl groups such as acetyl-, propionyl-, and benzoyl-CoAs, or structural variants, such as malonamyl-CoA<sup>56</sup>.

In modular PKS systems a set of KS, AT, and ACP domains, as well as optional  $\beta$ -keto processing domains, constitute one module. Each module is responsible for only one single elongation cycle, therefore the number of modules correlates with the number of elongation steps. This makes it possible to predict metabolite structure from enzyme architecture or vice versa. The correlation of the chain length with the number of catalytic modules is referred to as the co-linearity rule<sup>44</sup>. The unique structures of the PKs are additionally shaped by the actions of optional reduction and/or modification step. The complete or partial absence of  $\beta$ -keto processing domains is reflected in the structures of PKs by the presence of keto groups, hydroxyl groups and unsaturated double bonds<sup>44</sup>. The diversity of the PKs is further influenced by various stereochemistry orientations provided by different catalytic domains. For instance, KR domains can lead to generation of S or R hydroxyl groups, based on the type of the KR domain. Their specificity can be predicted by the presence of the important sequence motifs, which have been experimentally elucidated by mutagenesis and genetic engineering experiments<sup>57-59</sup>. Another level of stereo-chemical diversity is provided by the action of DH domains. Dehydration of the D configured moieties in most cases results in *trans* double bond configuration, whereas dehydration of the L isomers usually leads to generation of *cis* double bonds<sup>44</sup>. ER domains can lead to S or R methyl branches in the structures of the PKs. Ladley and co-workers were able to deduce an important residue in the ER domains. By mutation of the catalytic tyrosin to valin in the PKS derived from erythromycin biosynthesis, they were able to achieve a change in methyl branch stereochemistry from S to R. The reverse experiment on the ER from rapamycin was however not successful. This suggests that additional residues might play a role<sup>60</sup>.

### 1.3.3 Unusual biochemistry of *trans*-AT PKS systems

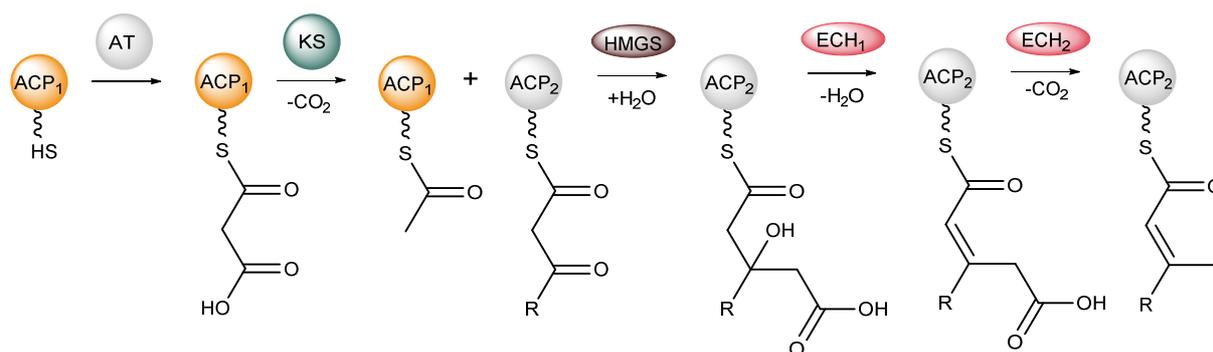
Typically, in standard PKS systems each AT is paired with its own ACP domain, which accepts the substrate activated by the corresponding AT. Such systems are commonly referred to as *cis*-AT PKS. However, there is a special class of PKS called *trans*-AT PKS, in which the AT domain is lacking in every module and is instead replaced by a free-standing AT domain<sup>61</sup>. Such systems are thus sometimes also referred to as AT-less PKS. Since the discovery of pederin, first reported *trans*-AT PKS with an attributed metabolite, it has become clear that alternatives to *cis*-AT PKS exist<sup>61</sup>. Despite the same modular organization and other similarities to *cis*-AT PKS, *trans*-AT PKS have evolved independently from standard PKS

systems. They initially failed to be recognized as a standalone class of metabolites due to their scarcity in streptomycetes. However, more and more of those compounds are being identified in less common types of microorganisms, which are harder to cultivate, including myxobacteria<sup>62</sup>. A recent study of PKS systems showed that *trans*-AT PKs are a major group of bacterial compounds, which attributed 38% of all sequenced bacterial genomes to *trans*-AT type<sup>62</sup>. This clearly shows that *trans*-AT PKs are important bacterial secondary metabolite pathways, which deserve to be classified as standalone class. There are several known cases where the biosynthesis of PKS deviates from the standard modular PKS biochemistry. Some of such peculiarities include e.g. module skipping, split modules located on two adjacent proteins, non-elongating modules, unusual system organizations and some catalytic functions performed in *trans*<sup>63,62</sup>. Those features are known to commonly occur in *trans*-AT PKS systems<sup>62</sup>.

A typical example of *trans*-AT PKS is bacillaene biosynthetic pathway which was used in a number of studies to investigate many fundamental principles of *trans*-AT PKSs<sup>61</sup> (Fig. 10). Comparison of the bacillaene structure with the PKS architecture reveals many non-canonical features. The cluster encodes 17 modules but the compound contains only 13 building blocks, indicating the presence of non-elongating modules. Modules 4, 8 and 14 are split and located on two separate proteins. As commonly observed in *trans*-AT PKSs, the AT functionality was shown to be provided in *trans* by the PksC, a homologue of BaeC from another producer<sup>64</sup>. Furthermore, *trans*-ER activity provided by PksE (BaeE) in module 2, was also identified<sup>65</sup>.



malonyl-CoA, a nonelongating KS domain responsible for decarboxylation of the aforementioned malonyl-CoA and finally a HMGS homolog responsible for interconnection of the generated acetate unit with the biosynthetic intermediate bound to the corresponding module of the assembly line. The three core proteins are normally accompanied by one or two additional enoyl-CoA hydratases (ECH) homologs, responsible for a subsequent dehydration and decarboxylation<sup>67</sup> (Fig. 11).

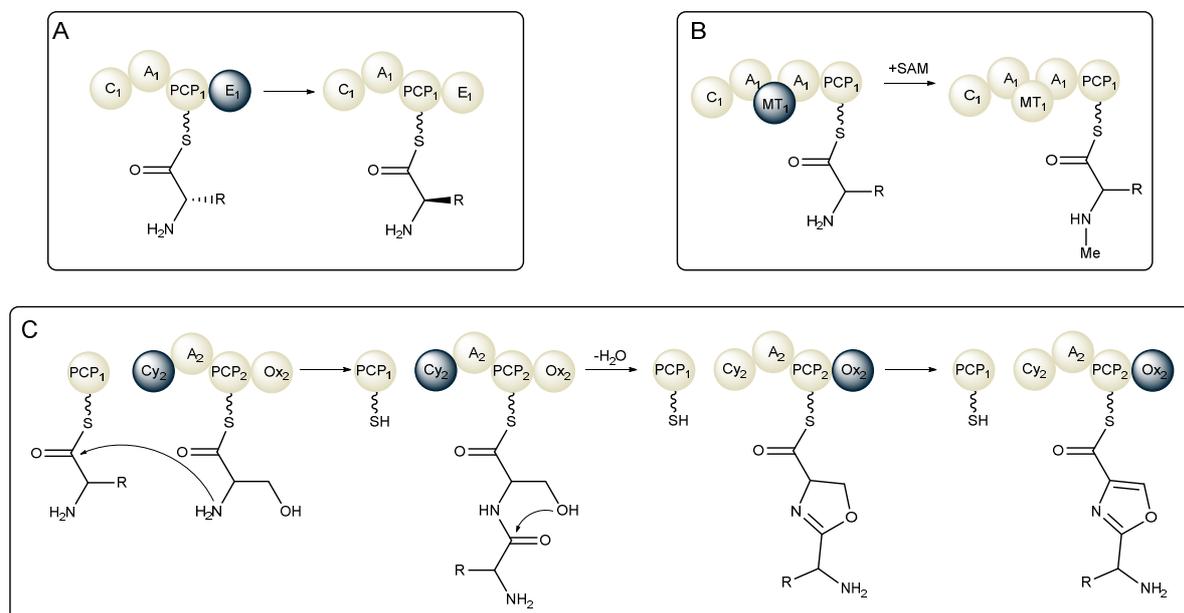


**Figure 11.** Catalytic reactions of the  $\beta$ -branching cassette.

Introduction of  $\beta,\gamma$  double bonds is yet another unusual element found in PKS systems. Most of the double bonds found in PKs are in  $\alpha,\beta$  position. However, several cases of  $\beta,\gamma$  double bonds have also been reported. This seems to be a common feature in *trans*-AT PKS since  $\beta,\gamma$  dehydration, as well as  $\alpha,\beta$  to  $\beta,\gamma$  double bond migration have been observed in bacillaene biosynthesis<sup>68</sup>. Additionally,  $\alpha,\beta$  to  $\beta,\gamma$  double bond migrations have been reported in rhizoxin<sup>69</sup>, nosperin<sup>70</sup> and coralopyronin (COR)<sup>71</sup> biosynthetic pathways. Both described reactions yield a  $\beta,\gamma$  double bond, however, their mechanisms are different. One is a direct  $\beta,\gamma$ -dehydration by a single, dedicated module and the other is canonical dehydration on one module followed by a double bond isomerization by a non-elongating downstream module containing an additional DH-like domain, termed DH\*<sup>72</sup>. Bacillaene possesses three  $\beta,\gamma$  double bonds, where the first two were proven to be installed by  $\beta,\gamma$  dehydration and the third one seems to be introduced by a non-elongating module (Fig. 10). This is in analogy with other non-elongating modules found in rhizoxin, nosperin and COR biosynthetic pathways. The non-elongating modules found in rhizoxin and nosperin biosynthetic pathway consist of  $KS_0$ -DH\*-ACP, where the  $KS_0$  is nonfunctional. The modules appear to be responsible solely for isomerization of the  $\alpha,\beta$  double bond, introduced by the previous module<sup>69</sup>. A similar situation is found in COR biosynthesis where a non-elongating, split module is found, consisting of the same  $KS_0$ -DH\*-ACP domain organization, yielding a  $\beta,\gamma$  double bond<sup>71</sup>.



incorporated amino acids<sup>46,51</sup>. Thiazoline/oxazoline rings can also be reduced by the reduction (R) domains to form thiazolidine/oxazolidine as observed in e.g. pyochelin<sup>83</sup>.



**Figure 13.** Optional domains of NRPS modules performing (A) epimerization, (B) methylation, (C) heterocyclization and oxidation.

### 1.3.5 Structural diversity in NRPS biosynthesis

As already mentioned for PKS, the selection of building blocks is an important factor that contributes to the structural diversity of NPs. This is even more so in NRPS, where the pool of monomer units is even larger and more diverse. NRPS accept anything from 20 proteinogenic amino acids to a large variety of nonproteinogenic amino acids and aryl acids<sup>43</sup>, which leads to practically unlimited diversity of NRPs. The A domains play a pivotal role in NRPS substrate selection and activation. It has been shown by several separate studies, that A domains are the specificity-mediating “gatekeeper” units, however, the exact mechanism of selectivity at first remained unclear<sup>84,74,85</sup>. This has changed with the elucidation of the first two crystal structures from the adenylate forming superfamily of enzymes. The first one was the crystal structure of firefly luciferase from *Photinus pyralis*<sup>86</sup>, followed by a crystal structure of A domain GrsA from gramicidin S pathway, in complex with phenylalanine and AMP<sup>87</sup>. By solving the crystal structure of the GrsA domain with its corresponding substrate and ATP, Stachelhaus and coworkers were able to identify the binding pocket of the A domain and develop ground rules for substrate specificity prediction of A domains, based on their primary protein sequence alignment. The binding pocket of the A domain is lined by ten amino acid residues, known as specificity-conferring code for substrate prediction of NRPSs,

now also commonly referred to as the nonribosomal code<sup>74,87,88</sup>. The nonribosomal code of ten residues was later expanded, to include important residues in range of 8 Å from the binding pocket of the A domain<sup>89</sup>. Those residues are predicted to interact with the substrate and thus influence the A domain specificity. Based on this knowledge, several A domain specificity predictor tools have been developed over the years (e.g. NRSPredictor2<sup>90</sup> or NRPSsp<sup>91</sup>). Those tools can identify the specificity conferring code in the given A domain sequence and predict the specificity of the A domain based on this code. Such data can be very useful when trying to predict a structure of a compound originating from a NRPS containing BGC as well as when trying to manipulate specificity of an A domain in order to initiate production of new derivatives. It has been shown by several studies that it is possible to change the specificity of the A domain by manipulating its nonribosomal code, where as little as one mutated residue was in some cases enough to completely change the A domain specificity<sup>74,92,93</sup>. When modifying the specificity of the A domain it is important to keep in mind, that the substrate also has to be processed by the downstream domains. Activation of the substrate, reflecting significant structural difference compared to the native building block, often leads to noncompliance with the downstream C domains which act as additional gatekeepers in the NRPS systems<sup>94,76,78</sup>. Furthermore, such specificity modifying approaches are even more difficult, as introduction of point mutations in complex megasynthetases is not trivial.

#### **1.4 Heterologous expression of complex natural product pathways**

Heterologous expression has proven to be an indispensable tool in the discovery, production, engineering and characterization of bacterial secondary metabolites<sup>95–97</sup>. Many of the NP producers are not easy to handle in the standard laboratory conditions without significant effort invested in method development. They are often slow in growth, requiring weeks to produce secondary metabolites. Despite all efforts to improve their growth, product yield sometimes remains too low for further studies. There are various methods that can be implemented to improve production yields of native or heterologous producers. Modifications include implementations of strong promoters<sup>98</sup> or modifying the expression levels of negative<sup>99,100</sup> and positive<sup>101</sup> regulators. As native producers are often not genetically accessible, expression of BGCs in well-established heterologous hosts provides a better alternative. Selected surrogate hosts are usually well studied organisms for which various genetic engineering techniques and molecular biology tools are readily available. Those tools can be employed to modify biosynthetic pathways introduced into genetically amendable

hosts for their characterization and modification in order to e.g. produce novel derivatives. Heterologous expression also provides means for expression of natural product BGCs uncovered by the genome sequencing projects. In the last decades, several hundred fully sequenced bacterial genomes have been acquired from well-known NP producers<sup>95</sup>. They present a rich source of BGCs for production of to date unknown compounds.

Genes that are transferred to a selected heterologous host have to be transcribed and translated, as well as the resulting protein has to fold correctly. For this purpose, the heterologous host has to be capable of performing appropriate posttranslational modifications. This is also the case in PKSs and NRPSs, as the heterologous host has to be able to posttranslationally modify the megasynthetases, e.g. via phosphopantetheinylation of ACPs or PCPs<sup>48</sup>. PKS and NRPS represent a major portion of secondary metabolites produced by actinobacteria and myxobacteria, which are also well known for their high GC content. Therefore, a heterologous host, selected for production of compounds from those organisms, also has to be able to express GC-rich genes<sup>95,102</sup>. Rare codons sometimes also represent a problem, as they can cause ribosomal stalling, resulting in truncated or degraded proteins. However, in practice it seems that rare codons are less problematic than initially believed<sup>98</sup>.

Once the appropriate heterologous host is selected, a desired BGC has to be appropriately modified, before successful expression in a heterologous host is possible. To obtain the BGC from native hosts, various techniques can be employed. Traditionally, large cosmid libraries were generated, which were then screened for constructs harboring desired genes. In most cases, the BGC has to be rearranged or reassembled on a special expression vector, harboring appropriate genetic elements like promoters and terminators as well as a good integration or replication system, specifically tailored to fit the host organism<sup>96</sup>. To assemble functional expression vectors, restriction-digestion and ligation based approaches were traditionally employed. Such methods are extremely laborious and time consuming, therefore faster and more versatile techniques were developed<sup>103</sup>.

Some of the more advanced methods include *in vitro* techniques like the Gibson assembly<sup>104</sup> or Golden Gate cloning<sup>105</sup> as well as *in vivo* methods like transformation-associated recombination (TAR) in yeast<sup>106</sup> or Red/ET recombination in *E. coli*<sup>107</sup>. Both *in vivo* methods can be employed to assemble DNA within the overlapping homologous regions via recombination, with the Red/ET being a bit more versatile and easily employed for further plasmid modifications in *E. coli*.

## 1.5 Combinatorial biosynthesis approaches

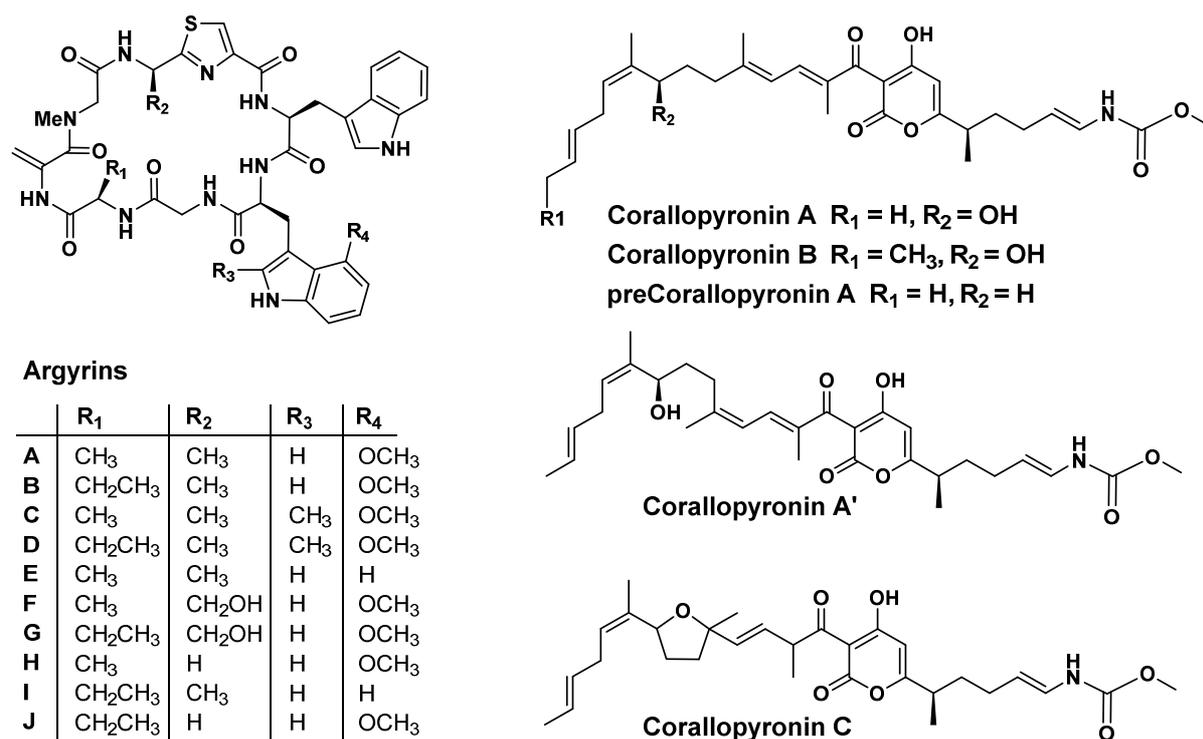
Another field, closely related to heterologous expression, is combinatorial biosynthesis. Swapping of entire domains, modules and subunits are some of the methods applied in combinatorial biosynthesis<sup>108</sup>. The approaches to this form of genetic engineering began with PKS systems, due to their predictable, modular organization<sup>109</sup>. The trait of certain PKS to follow co-linearity rules made it straightforward to predict the outcome of mix-matching modules from various biosynthetic pathways. The approach was limited to the modification of well characterized systems which resulted in a small library of novel non-natural derivatives<sup>103</sup>. One famous example of such approaches include swapping of AT domain from erythromycin biosynthetic pathway with its counterpart from rapamycin biosynthesis<sup>110</sup>. The method yielded 61 new analogues, many of which have not been naturally detected. Other similar examples of combinatorial biosynthesis in PKS include pikromycin<sup>111</sup>, tylosin<sup>111</sup> and daptomycin<sup>112</sup>, genes of which have also been combinatorially assembled for production of novel analogues. Alternatively to modifying preexisting pathways, *de novo* assembly of biosynthetic pathways by combining various enzymes in heterologous hosts has also been described<sup>103</sup> for flavonoids<sup>113,114</sup> and carotenoids<sup>115</sup>. Despite some successful examples of combinatorial biosynthesis, the method is by no means well established or without limitations. The idea of hybrid megasynthetases has been around for quite some time, however, it has thus far not been developed to the point of becoming a standardized method. One limitation of combinatorial studies comes from traditional restriction-digestion and ligation based cloning methods, which are tedious and time-consuming. As a result, traditional approaches are not appropriate to construct large libraries of combinatorially assembled pathways that are necessary for screening. A major challenge in assembly of chimeric synthases is their functional expression. Many of such enzymes are insoluble or functionally impaired resulting in abolished or very low production. The limitation is most often a result of incomplete understanding of protein-protein interactions necessary for efficient processing of the intermediates<sup>108</sup>.

A more modern approach for biosynthetic pathway assembly is *de novo* DNA synthesis, which has been steadily gaining importance in the last decade. The method is usually based on extensive *in silico* analysis of biosynthetic pathways, which allows their rearrangement or optimized design of genes. Unlike traditional approaches, methods based on DNA synthesis are not limited by the arrangement of native genes. Since the DNA can be *de novo* synthesized, it is possible to rearrange it as desired by employing previously mentioned DNA assembly techniques like TAR, Red/ET, Gibson assembly, Golden Gate assembly as well as

traditional restriction-digestion and ligation approach. This allows construction of more versatile biosynthetic pathways that enable easier genetic modifications later on. Nevertheless, the method is not without its limitations. The main issue at this point is still the cost of the synthesized DNA fragments. Moreover, the synthesis of fragments larger than 5 kb for now remains unreliable and impractical<sup>116</sup>, therefore development of advanced DNA assembly techniques is of importance.

## 1.6 Outline of the present work

The work of this thesis focuses on two distinctive classes of natural products, coralopyronins and argyrins (Fig. 14). Both compounds exert prominent biological activities. Coralopyronins were identified as potent antibiotics active against multiple Gram-positive and a few Gram-negative bacterial species. They recently showed strong potential for treatment of filariasis by targeting *Wolbachia*, endosymbiotic bacteria inside filarial nematodes. Agryrins were shown to possess antibacterial activity against several bacterial species including *Pseudomonas sp.*, opportunistic, Gram-negative pathogen. Furthermore, they exert immunomodulatory and anticancer activity. Argyrin F has already been evaluated as a preclinical candidate for anticancer therapy. Biosynthetic pathways for both compounds families were already identified and initial studies on biosynthesis were performed. In both cases the natural producer strains were not easily genetically amendable, which motivated the establishment of heterologous expression systems to facilitate structure and yield improvement approaches.



**Figure 14.** Structures of argyrimycin and corallopyronin compound families.

The first part of the thesis focuses on the  $\alpha$ -pyrone antibiotic COR, produced by a *trans*-AT hybrid PKS-NRPS system<sup>117</sup>. COR A was shown to target bacterial RNA polymerase (RNAP) at a novel binding site – the switch region<sup>118,119</sup>. The new binding site differs from the binding site of previously known RNAP inhibitors (e.g. rifamycins) and is thus expected to help overcome resistance issues of currently used clinical RNAP inhibitors<sup>120</sup>. COR A was recently also shown to deplete symbiotic endobacteria – *Wolbachia* from filarial nematodes<sup>121,39</sup>. It is therefore currently being investigated as the lead candidate for treatment of filariasis. The main goal of the COR project was to establish and optimize the production platform for COR in the heterologous host *M. xanthus* (see chapter 2). As the compound developed into a lead candidate for broad antifilarial treatment, the establishment of optimized production platform was of the highest priority as the wild type strain did not allow for significant production of the compound. Parallel studies on the structurally related  $\alpha$ -pyrone antibiotic myxopyronin (MXN) led to the assembly of the corresponding BGC expression construct, which could be successfully employed to generate a hybrid *mxn/cor* heterologous expression system. The platform was successfully expressed in a derivative of *M. xanthus* DK1622 and led to a significant increase in production of COR A, compared to the native producer. To improve the COR A yield, achieved by the production system, several additional modifications were performed (see chapter 3). Exchange of promoter and replacement of the genome integration system led to further increase in production. To confirm the function of certain proteins in the

COR BGC and try to generate novel COR derivatives, several gene deletion experiments were performed (see chapter 3). The information gathered from the experiments helped redefine certain aspects of the COR biosynthesis and cultivation of the gene deletion mutants led to isolation of two novel COR analogues (see chapter 3). These studies also included evaluation of the myxobacterial  $\alpha$ -pyrone antibiotic self-resistance, which provided data that hints towards a myxobacterial RNAP inhibitor specific self-resistance mechanism. This finding might help improve self-resistance of the COR heterologous producer and lead to increased COR production yields.

The second part of the thesis describes a macrocyclic octapeptide – argyrin, known for its interesting activity profile. It has recently been shown that elongation factor G (EF-G) is the cellular target for argyrin in *P. aeruginosa*<sup>122,123</sup>. Upon binding to EF-G, argyrin prevents protein synthesis and inhibits bacterial growth. By mutagenesis studies and structural analysis it could be shown that argyrin binds to a target site of EF-G that clearly differs from the binding site of previously known inhibitor - fusidic acid<sup>122,123</sup>. Argyrins were also shown to inhibit the eukaryotic elongation factor G1 (EF-G1)<sup>123</sup> and are described as proteasome inhibitors<sup>124</sup>, which makes them a promising candidate for the development of anticancer drugs. Besides their antibacterial and cytotoxic properties, argyrins also exhibit immunosuppressive activity and are currently studied as potent inhibitors of T - cell independent antibody formation<sup>125</sup>. The main focus of the argyrin project was to direct the biosynthesis towards production of specific argyrin derivatives, as different analogues exhibit a different activity profile and because the production strain typically generates a complex mixture of analogues (see chapter 4). To facilitate the production of different analogues we aimed at directing the argyrin production profiles via engineering the NRPS megasynthetase, targeting adenylation domains. As an alternative method for production of novel argyrin analogues, precursor directed biosynthesis was established. By employing a readily available technique, production of several tryptophan analogues was achieved and supplementation of heterologous producer cultures with those analogues led to production, isolation and structure elucidation of more than 14 new argyrin derivatives. Furthermore, studies of alternative promoter systems led to significant improvement in production of argyrins (see chapter 5).

## 1.7 References

- (1) The classes of natural product and their isolation. In *Natural products. The secondary metabolites / James R. Hanson*; Hanson, J. R., Ed.; Royal Society of Chemistry: Cambridge, 2003, pp. 1–34.
- (2) Pickens, L. B.; Tang, Y.; Chooi, Y.-H. Metabolic engineering for the production of natural products, *Annual review of chemical and biomolecular engineering*. **2011**, *2*, pp. 211–236.
- (3) Cragg, G. M.; Newman, D. J. Natural products: a continuing source of novel drug leads, *Biochim. Biophys. Acta*. **2013**, *1830*, pp. 3670–3695.
- (4) Baker, D. D.; Chu, M.; Oza, U.; Rajgarhia, V. The value of natural products to future pharmaceutical discovery, *Nat. Prod. Rep.* **2007**, *24*, pp. 1225–1244.
- (5) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014, *J. Nat. Prod.* **2016**, *79*, pp. 629–661.
- (6) Dewick, P. M. *Medicinal Natural Products*; Wiley, 2009.
- (7) Procópio, R. E. d. L.; Silva, I. R. d.; Martins, M. K.; Azevedo, J. L. d.; Araújo, J. M. d. Antibiotics produced by *Streptomyces*, *Braz. J. Infect. Dis.* **2012**, *16*, pp. 466–471.
- (8) Bush, K.; Courvalin, P.; Dantas, G.; Davies, J.; Eisenstein, B.; Huovinen, P.; Jacoby, G. A.; Kishony, R.; Kreiswirth, B. N.; Kutter, E.; Lerner, S. A.; Levy, S.; Lewis, K.; Lomovskaya, O.; Miller, J. H.; Mobashery, S.; Piddock, L. J. V.; Projan, S.; Thomas, C. M.; Tomasz, A.; Tulkens, P. M.; Walsh, T. R.; Watson, J. D.; Witkowski, J.; Witte, W.; Wright, G.; Yeh, P.; Zgurskaya, H. I. Tackling antibiotic resistance, *Nat. Rev. Microbiol.* **2011**, *9*, pp. 894–896.
- (9) Lewis, K. Antibiotics. Recover the lost art of drug discovery, *Nature*. **2012**, *485*, pp. 439–440.
- (10) Gould, I. M. Antibiotic resistance. The perfect storm, *Int. J. Antimicrob. Agents* **2009**, *34*, S2-S5.
- (11) World Health Organization Global Antimicrobial Resistance Surveillance System (GLASS) Report, 31.01.2018. <http://apps.who.int/iris/bitstream/handle/10665/259744/9789241513449-eng.pdf;jsessionid=97E045EBFE3B29F286CBA3EE8360F0D1?sequence=1>. Wednesday, January 31, 2018.
- (12) Challinor, V. L.; Bode, H. B. Bioactive natural products from novel microbial sources, *Ann. N. Y. Acad. Sci.* **2015**, *1354*, pp. 82–97.
- (13) Waksman, S. A. Production and Activity of Streptothricin, *J. Bacteriol.* 1943, *46*, pp. 299–310.
- (14) Schatz, A.; Bugle, E.; Waksman, S. A. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.\*, *Exp. Biol. Med.* 1944, *55*, pp. 66–69.
- (15) Stadler, M.; Hoffmeister, D. Fungal natural products-the mushroom perspective, *Front. Microbiol.* **2015**, *6*, p. 127.
- (16) Stein, T. Bacillus subtilis antibiotics: structures, syntheses and specific functions, *Mol. Microbiol.* **2005**, *56*, pp. 845–857.
- (17) Wenzel, S. C.; Müller, R. Myxobacteria--'microbial factories' for the production of bioactive secondary metabolites, *Mol. Biosyst.* **2009**, *5*, pp. 567–574.
- (18) Wenzel, S. C.; Müller, R. Myxobacteria - unique microbial secondary metabolite factories. In *Comprehensive Natural Products Chemistry II, Vol 2: Structural Diversity II -*

*Secondary Metabolite Sources, Evolution and Selected Molecular Structures*; Moore, B., Ed.; Elsevier: Oxford, 2010, pp. 189–222.

(19) Ringel, S. M.; Greenough, R. C.; Roemer, S.; Connor, D.; Gutt, A. L.; Blair, B.; Kanter, G.; von Strandtmann Ambruticin (W7783), a new antifungal antibiotic, *J. Antibiot.* **1977**, *30*, pp. 371–375.

(20) Weissman, K. J.; Müller, R. Myxobacterial secondary metabolites: bioactivities and modes-of-action, *Nat. Prod. Rep.* **2010**, *27*, pp. 1276–1295.

(21) Herrmann, J.; Fayad, A. A.; Müller, R. Natural products from myxobacteria: novel metabolites and bioactivities, *Nat. Prod. Rep.* **2017**, *34*, pp. 135–160.

(22) Schäberle, T. F.; Lohr, F.; Schmitz, A.; König, G. M. Antibiotics from myxobacteria, *Nat. Prod. Rep.* **2014**, *epub ahead of print*.

(23) Garcia, R.; Gerth, K.; Stadler, M.; Dogma Jr., I. J.; Müller, R. Expanded phylogeny of myxobacteria and evidence for cultivation of the 'unculturables', *Mol. Phylogenet. Evol.* **2010**, *57*, pp. 878–887.

(24) Wenzel, S. C.; Müller, R. The biosynthetic potential of myxobacteria and their impact on drug discovery, *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, pp. 220–230.

(25) Zaburannyi, N.; Bunk, B.; Maier, J.; Overmann, J.; Müller, R. Genome analysis of the fruiting body forming myxobacterium *Chondromyces crocatus* reveals high potential for natural product Biosynthesis, *Appl. Environ. Microbiol.* **2016**, *82*, pp. 1945–1957.

(26) Munoz-Dorado, J.; Marcos-Torres, F. J.; Garcia-Bravo, E.; Moraleda-Munoz, A.; Perez, J. Myxobacteria: Moving, Killing, Feeding, and Surviving Together, *Front. Microbiol.* **2016**, *7*, p. 781.

(27) Aravind, L.; Anantharaman, V.; Venancio, T. M. Apprehending multicellularity: Regulatory networks, genomics, and evolution, *Birth Defects Research Part C: Embryo Today: Reviews.* **2009**, *87*, pp. 143–164.

(28) Mauriello, E. M. F.; Mignot, T.; Yang, Z.; Zusman, D. R. Gliding motility revisited: how do the myxobacteria move without flagella?, *Microbiol. Mol. Biol. Rev.* **2010**, *74*, pp. 229–249.

(29) O'Connor, K. A.; Zusman, D. R. Behavior of peripheral rods and their role in the life cycle of *Myxococcus xanthus*, *J. Bacteriol.* **1991**, *173*, pp. 3342–3355.

(30) O'Connor, K. A.; Zusman, D. R. Development in *Myxococcus xanthus* involves differentiation into two cell types, peripheral rods and spores, *J. Bacteriol.* **1991**, *173*, pp. 3318–3333.

(31) Shimkets, L. J. Intercellular signaling during fruiting-body development of *Myxococcus xanthus*, *Annu. Rev. Microbiol.* **1999**, *53*, pp. 525–549.

(32) Zusman, D. R.; Scott, A. E.; Yang, Z.; Kirby, J. R. Chemosensory pathways, motility and development in *Myxococcus xanthus*, *Nat. Rev. Microbiol.* **2007**, *5*, pp. 862–872.

(33) Reichenbach, H.; Höfle, G. Discovery of a new antifungal mechanism of action: soraphen - an almost-success story. In *Scientific Annual Report*; Gesellschaft für Biotechnologische Forschung GmbH: Braunschweig, 1994, pp. 5–22.

(34) Baumann, S.; Herrmann, J.; Raju, R.; Steinmetz, H.; Mohr, K. I.; Hüttel, S.; Harmrolfs, K.; Stadler, M.; Müller, R. Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity, *Angew. Chem. Int. Ed.* **2014**, *53*, pp. 14605–14609.

(35) Mulzer, J.; Altmann, K. H.; Hofle, G.; Müller, R.; Prantz, K. Epothilones - a fascinating family of microtubule stabilizing antitumor agents, *Comptes Rendus Chimie.* **2008**, *11*, pp. 1336–1368.

- (36) Plaza, A.; Garcia, R.; Bifulco, G.; Martinez, J. P.; Hüttel, S.; Sasse, F.; Meyerhans, A.; Stadler, M.; Müller, R. Aetheramides A and B, potent HIV-inhibitory depsipeptides from a myxobacterium of the new genus "Aetherobacter", *Org. Lett.* **2012**, *14*, pp. 2854–2857.
- (37) Held, J.; Gebru, T.; Kalesse, M.; Jansen, R.; Gerth, K.; Müller, R.; Mordmüller, B. Antimalarial activity of the myxobacterial macrolide chlorotonil A, *Antimicrob. Agents Chemother.* **2014**, *58*, pp. 6378–6384.
- (38) Sasse, F.; Steinmetz, H.; Schupp, T.; Petersen, F.; Memmert, K.; Hofmann, H.; Heusser, C.; Brinkmann, V.; Matt, P. von; Höfle, G.; Reichenbach, H. Argyrins, immunosuppressive cyclic peptides from myxobacteria. I. Production, isolation, physico-chemical and biological properties, *J. Antibiot.* **2002**, *55*, pp. 543–551.
- (39) Schäberle, T. F.; Schiefer, A.; Schmitz, A.; König, G. M.; Hoerauf, A.; Pfarr, K. Coralopyronin A - a promising antibiotic for treatment of filariasis, *Int. J. Med. Microbiol.* **2014**, *304*, pp. 72–78.
- (40) Mulzer, J., Ed. The Epothilones, an Outstanding Family of Anti-Tumor Agents; Springer: New York, 2009.
- (41) Dumez, H.; Gall, H.; Capdeville, R.; Dutreix, C.; van Oosterom, A. T.; Giaccone, G. A phase I and pharmacokinetic study of LAF389 administered to patients with advanced cancer, *Anticancer. Drugs.* **2007**, *18*, pp. 219–225.
- (42) Kaur, G.; Hollingshead, M.; Holbeck, S.; Schauer-Vukasinovic, V.; Camalier, R. F.; Domling, A.; Agarwal, S. Biological evaluation of tubulysin A: a potential anticancer and antiangiogenic natural product, *Biochem. J.* **2006**, *396*, pp. 235–242.
- (43) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms, *Chem. Rev.* **2006**, *106*, pp. 3468–3496.
- (44) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, pp. 4688–4716.
- (45) Süßmuth, R. D.; Mainz, A. Nonribosomal peptide synthesis - Principles and prospects, *Angew. Chem. Int. Ed.* **2017**, *56*, pp. 3770–3821.
- (46) Du, L. H.; Sanchez, C.; Shen, B. Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules, *Metab. Eng.* **2001**, *3*, pp. 78–95.
- (47) Silakowski, B.; Kunze, B.; Müller, R. Multiple hybrid polyketide synthase/non-ribosomal peptide synthetase gene clusters in the myxobacterium *Stigmatella aurantiaca*, *Gene.* **2001**, *275*, pp. 233–240.
- (48) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life, *Nat. Prod. Rep.* **2013**, *31*, pp. 61–108.
- (49) Perlova, O.; Fu, J.; Kuhlmann, S.; Krug, D.; Stewart, F.; Zhang, Y.; Müller, R. Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2006**, *72*, pp. 7485–7494.
- (50) Weinig, S.; Mahmud, T.; Müller, R. Markerless mutations in the myxothiazol biosynthetic gene cluster: A delicate megasynthetase with a superfluous nonribosomal peptide synthetase domain, *Chem. Biol.* **2003**, *10*, pp. 953–960.
- (51) Silakowski, B.; Schairer, H. U.; Ehret, H.; Kunze, B.; Weinig, S.; Nordsiek, G.; Brandt, P.; Blöcker, H.; Höfle, G.; Beyer, S.; Müller, R. New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1, *J. Biol. Chem.* **1999**, *274*, pp. 37391–37399.

- (52) Müller, I.; Weinig, S.; Steinmetz, H.; Kunze, B.; Veluthoor, S.; Mahmud, T.; Müller, R. A unique mechanism for methyl ester formation via an amide intermediate found in myxobacteria, *ChemBioChem*. **2006**, *7*, pp. 1197–1205.
- (53) Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary implications of bacterial polyketide synthases, *Mol. Biol. Evol.* **2005**, *22*, pp. 2027–2039.
- (54) Kohli, R. M.; Walsh, C. T. Enzymology of acyl chain macrocyclization in natural product biosynthesis, *Chem. Commun.* **2003**, pp. 297–307.
- (55) Staunton, J.; Weissman, K. J. Polyketide biosynthesis: a millennium review, *Nat. Prod. Rep.* **2001**, *18*, pp. 380–416.
- (56) Moore, B. S.; Hertweck, C. Biosynthesis and attachment of novel bacterial polyketide synthase starter units, *Nat. Prod. Rep.* **2002**, *19*, pp. 70–99.
- (57) Baerga-Ortiz, A.; Popovic, B.; Siskos, A. P.; O'Hare, H. M.; Spiteller, D.; Williams, M. G.; Campillo, N.; Spencer, J. B.; Leadlay, P. F. Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase, *Chem. Biol.* **2006**, *13*, pp. 277–285.
- (58) Castonguay, R.; He, W.; Chen, A. Y.; Khosla, C.; Cane, D. E. Stereospecificity of ketoreductase domains of the 6-deoxyerythronolide B synthase, *J. Am. Chem. Soc.* **2007**, *129*, pp. 13758–13769.
- (59) Keatinge-Clay, A. T. A tylosin ketoreductase reveals how chirality is determined in polyketides, *Chem. Biol.* **2007**, *14*, pp. 898–908.
- (60) Kwan, D. H.; Sun, Y.; Schulz, F.; Hong, H.; Popovic, B.; Sim-Stark, J. C.; Haydock, S. F.; Leadlay, P. F. Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases, *Chem. Biol.* **2008**, *15*, pp. 1231–1240.
- (61) Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases, *Nat. Prod. Rep.* **2010**, *27*, pp. 996–1047.
- (62) Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases, *Nat. Prod. Rep.* **2016**, *33*, pp. 231–316.
- (63) Keatinge-Clay, A. T. The Uncommon Enzymology of Cis-Acyltransferase Assembly Lines, *Chem. Rev.* **2017**, *117*, pp. 5334–5366.
- (64) Calderone, C. T.; Kowtoniuk, W. E.; Kelleher, N. L.; Walsh, C. T.; Dorrestein, P. C. Convergence of isoprene and polyketide biosynthetic machinery: isoprenyl-S-carrier proteins in the pksX pathway of *Bacillus subtilis*, *Proc. Natl. Acad. Sci. USA*. **2006**, *103*, pp. 8977–8982.
- (65) Bumpus, S. B.; Magarvey, N. A.; Kelleher, N. L.; Walsh, C. T.; Calderone, C. T. Polyunsaturated fatty-acid-like trans-enoyl reductases utilized in polyketide biosynthesis, *J. Am. Chem. Soc.* **2008**, *130*, p. 11614.
- (66) Butcher, R. A.; Schroeder, F. C.; Fischbach, M. A.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Clardy, J. The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*, *Proc. Natl. Acad. Sci. USA*. **2007**, *104*, pp. 1506–1509.
- (67) Buchholz, T. J.; Rath, C. M.; Lopanik, N. B.; Gardner, N. P.; Hakansson, K.; Sherman, D. Polyketide  $\beta$ -branching in bryostatin biosynthesis: identification of surrogate acetyl-ACP donors for BryR, an HMG-ACP synthase, *Chem. Biol.* **2010**, *17*, pp. 1092–1100.
- (68) Moldenhauer, J.; Götz, D. C. G.; Albert, C. R.; Bischof, S. K.; Schneider, K.; Süßmuth, R. D.; Engeser, M.; Gross, H.; Bringmann, G.; Piel, J. The Final Steps of Bacillaene Biosynthesis in *Bacillus amyloliquefaciens* FZB42: Direct Evidence for  $\beta,\gamma$  Dehydration by a trans-Acyltransferase Polyketide Synthase, *Angew. Chem. Int. Ed.* **2010**, *49*, pp. 1465–1467.

- (69) Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. Functionally distinct modules operate two consecutive  $\alpha,\beta \rightarrow \beta,\gamma$  double-bond shifts in the rhizoxin polyketide assembly line, *Angew. Chem. Int. Ed.* **2010**, *49*, pp. 1460–1464.
- (70) Kampa, A.; Gagunashvili, A. N.; Gulder, T. A. M.; Morinaka, B. I.; Daolio, C.; Godejohann, M.; Miao, V. P. W.; Piel, J.; Andrésson, Ó. S. Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E3129–37.
- (71) Lohr, F.; Jenniches, I.; Fritzier, M.; Meehan, M. J.; Sylvester, M.; Schmitz, A.; Gütschow, M.; Dorrestein, P. C.; König, G. M.; Schäberle, T. F.  $\alpha,\beta / \beta,\gamma$  double bond migration in corallopyronin A biosynthesis, *Chem. Sci.* **2013**, *2013*, pp. 4175–4180.
- (72) Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases, *Nat. Prod. Rep.* **2016**, *33*, pp. 231–316.
- (73) Walsh, C. T.; Chen, H. W.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L. S.; Marshall, C. G.; Miller, D. A.; Patel, H. M. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines, *Curr. Opin. Chem. Biol.* **2001**, *5*, pp. 525–534.
- (74) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem. Biol.* **1999**, *6*, pp. 493–505.
- (75) Wenzel, S. C.; Müller, R. Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry, *Nat. Prod. Rep.* **2007**, *24*, pp. 1211–1224.
- (76) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution, *BMC Evol. Biol.* **2007**, *7*, pp. 78–92.
- (77) Finking, R.; Marahiel, M. A. Biosynthesis of nonribosomal peptides, *Annu. Rev. Microbiol.* **2004**, *58*, pp. 453–488.
- (78) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis, *Science*. **1999**, *284*, pp. 486–489.
- (79) Samel, S. A.; Czodrowski, P.; Essen, L.-O. Structure of the epimerization domain of tyrocidine synthetase A, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, pp. 1442–1452.
- (80) Labby, K. J.; Watsula, S. G.; Garneau-Tsodikova, S. Interrupted adenylation domains: unique bifunctional enzymes involved in nonribosomal peptide biosynthesis, *Nat. Prod. Rep.* **2015**, *32*, pp. 641–653.
- (81) Velkov, T.; Lawen, A. Mapping and molecular modeling of S-adenosyl-L-methionine binding sites in N-methyltransferase domains of the multifunctional polypeptide cyclosporin synthetase, *J. Biol. Chem.* **2003**, *278*, pp. 1137–1148.
- (82) Marshall, C. G.; Hillson, N. J.; Walsh, C. T. Catalytic mapping of the vibriobactin biosynthetic enzyme VibF, *Biochemistry*. **2002**, *41*, pp. 244–250.
- (83) Reimann, C.; Patel, H. M.; Serino, L.; Barone, M.; Walsh, C. T.; Haas, D. Essential PchG-dependent reduction in pyochelin biosynthesis of *Pseudomonas aeruginosa*, *J. Bacteriol.* **2001**, *183*, pp. 813–820.
- (84) Stachelhaus, T.; Marahiel, M. A. Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA, *J. Biol. Chem.* **1995**, *270*, pp. 6163–6169.
- (85) Mootz, H. D.; Marahiel, M. A. The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains, *J. Bacteriol.* **1997**, *179*, pp. 6843–6850.
- (86) Conti, E.; Franks, N. P.; Brick, P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes, *Structure*. **1996**, *4*, pp. 287–298.

- (87) Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S, *EMBO J.* **1997**, *16*, pp. 4174–4183.
- (88) Döhren, H. von; Dieckmann, R.; Pavela-Vrancic, M. The nonribosomal code, *Chem. Biol.* **1999**, *6*, pp. R273-R279.
- (89) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs), *Nucleic Acids Res.* **2005**, *33*, pp. 5799–5808.
- (90) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. NRPSpredictor2 - a web server for predicting NRPS adenylation domain specificity, *Nucleic Acids Res.* **2011**, *39*, pp. W362-W367.
- (91) Prieto, C.; García-Estrada, C.; Lorenzana, D.; Martín, J. F. NRPSsp. Non-ribosomal peptide synthase substrate predictor, *Bioinformatics (Oxford, England)*. **2012**, *28*, pp. 426–427.
- (92) Bian, X.; Plaza, A.; Yan, F.; Zhang, Y.; Müller, R. Rational and efficient site-directed mutagenesis of adenylation domain alters relative yields of luminide derivatives in vivo, *Biotechnol. Bioeng.* **2015**, *112*, pp. 1343–1353.
- (93) Yan, F.; Auerbach, D.; Chai, Y.; Keller, L.; Tu, Q.; Hüttel, S.; Glemser, A.; Grab, H. A.; Bach, T.; Zhang, Y.; Müller, R. Biosynthesis and Heterologous Production of Vioprolides: Rational Biosynthetic Engineering and Unprecedented 4-Methylazetidincarboxylic Acid Formation, *Angew. Chem. Int. Ed. Engl.* **2018**.
- (94) Gatto, G. J., Jr.; McLoughlin, S. M.; Kelleher, N. L.; Walsh, C. T. Elucidating the substrate specificity and condensation domain activity of FkbP, the FK520 pipecolate-incorporating enzyme, *Biochemistry*. **2005**, *44*, pp. 5993–6002.
- (95) Wenzel, S. C.; Müller, R. Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways, *Curr. Opin. Biotechnol.* **2005**, *16*, pp. 594–606.
- (96) Ongley, S.; Bian, X.; Neilan, B. A.; Müller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways, *Nat. Prod. Rep.* **2013**, *30*, pp. 1121–1138.
- (97) Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R. Heterologous expression of bacterial natural product biosynthetic pathways, *Nat. Prod. Rep.* **2019**.
- (98) Stevens, D. C.; Hari, T. P. A.; Boddy, C. N. The role of transcription in heterologous expression of polyketides in bacterial hosts, *Nat. Prod. Rep.* **2013**, *30*, pp. 1391–1411.
- (99) Sandmann, A.; Frank, B.; Müller, R. A transposon-based strategy to scale up myxothiazol production in myxobacterial cell factories, *J. Biotechnol.* **2008**, *135*, pp. 255–261.
- (100) Rachid, S.; Gerth, K.; Müller, R. NtcA-A negative regulator of secondary metabolite biosynthesis in *Sorangium cellulosum*, *J. Biotechnol.* **2008**, *140*, pp. 135–142.
- (101) Rachid, S.; Gerth, K.; Kochems, I.; Müller, R. Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium *Sorangium cellulosum* So ce56, *Mol. Microbiol.* **2007**, *63*, pp. 1783–1796.
- (102) Rodriguez, E.; Menzella, H. G.; Gramajo, H. Heterologous production of polyketides in bacteria, *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates*. **2009**, *459*, pp. 339–365.
- (103) Cobb, R. E.; Ning, J. C.; Zhao, H. DNA assembly techniques for next-generation combinatorial biosynthesis of natural products, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, pp. 469–477.

- (104) Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A., III; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat. Methods*. **2009**, *6*, pp. 343–345.
- (105) Engler, C.; Kandzia, R.; Marillonnet, S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability, *PLoS ONE*. **2008**, *3*, e3647.
- (106) Kouprina, N.; Larionov, V. Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology, *Chromosoma*. **2016**, *125*, pp. 621–632.
- (107) Zhang, Y.; Buchholz, F.; Muyrers, J. P.; Stewart, F. A. A new logic for DNA engineering using recombination in *Escherichia coli*, *Nat. Genet.* **1998**, *20*, pp. 123–128.
- (108) Sun, H.; Liu, Z.; Zhao, H.; Ang, E. L. Recent advances in combinatorial biosynthesis for drug discovery, *Drug Des. Devel. Ther.* **2015**, *9*, pp. 823–833.
- (109) Wong, F. T.; Khosla, C. Combinatorial biosynthesis of polyketides—a perspective, *Curr. Opin. Chem. Biol.* **2012**.
- (110) McDaniel, R.; Thamchaipenet, A.; Gustafsson, C.; Fu, H.; Betlach, M.; Ashley, G. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products, *Proc. Natl. Acad. Sci. USA*. **1999**, *96*, pp. 1846–1851.
- (111) Yoon, Y. J.; Beck, B. J.; Kim, B. S.; Kang, H. Y.; Reynolds, K. A.; Sherman, D. H. Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in *Streptomyces venezuelae*, *Chem. Biol.* **2002**, *9*, pp. 203–214.
- (112) Baltz, R. H. Combinatorial Biosynthesis of Cyclic Lipopeptide Antibiotics: A Model for Synthetic Biology To Accelerate the Evolution of Secondary Metabolite Biosynthetic Pathways, *ACS Synth. Biol.* **2014**, *3*, pp. 748–758.
- (113) Horinouchi, S. Combinatorial biosynthesis of non-bacterial and unnatural flavonoids, stilbenoids and curcuminoids by microorganisms, *J. Antibiot.* **2008**, *61*, pp. 709–728.
- (114) Horinouchi, S. Combinatorial biosynthesis of plant medicinal polyketides by microorganisms, *Curr. Opin. Chem. Biol.* **2009**, *13*, pp. 197–204.
- (115) Sandmann, G. Combinatorial biosynthesis of carotenoids in a heterologous host: a powerful approach for the biosynthesis of novel structures, *ChemBioChem*. **2002**, *3*, pp. 629–635.
- (116) Kosuri, S.; Church, G. M. Large-scale de novo DNA synthesis: technologies and applications, *Nat. Methods*. **2014**, *11*, pp. 499–507.
- (117) Erol, Ö.; Schäberle, T. F.; Schmitz, A.; Rachid, S.; Gurgui, C.; El Omari, M.; Lohr, F.; Kehraus, S.; Piel, J.; Müller, R.; König, G. M. Biosynthesis of the myxobacterial antibiotic coralopyronin A, *ChemBioChem*. **2010**, *11*, pp. 1235–1265.
- (118) Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H. The RNA polymerase "switch region" is a target of inhibitors, *Cell*. **2008**, *135*, pp. 295–307.
- (119) Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G. Transcription inactivation through local refolding of the RNA polymerase structure, *Nature*. **2008**.
- (120) Goldstein, B. P. Resistance to rifampicin: a review, *J. Antibiot.* **2014**, *67*, pp. 625–630.
- (121) Schiefer, A.; Schmitz, A.; Schäberle, T. F.; Specht, S.; Lämmer, C.; Johnston, K. L.; Vassilyev, D. G.; König, G. M.; Hoerauf, A.; Pfarr, K. Coralopyronin A Specifically Targets and Depletes Essential Obligate Wolbachia Endobacteria From Filarial Nematodes In Vivo, *J. Infect. Dis.* **2012**, *206*, pp. 249–257.

- (122) Bielecki, P.; Lukat, P.; Hüsecken, K.; Dötsch, A.; Steinmetz, H.; Hartmann, R. W.; Müller, R.; Häussler, S. Mutation in elongation factor G confers resistance to the antibiotic argyrimycin in the opportunistic pathogen *Pseudomonas aeruginosa*, *ChemBioChem*. **2012**, *13*, pp. 2339–2345.
- (123) Nyfeler, B.; Hoepfner, D.; Palestrant, D.; Kirby, C. A.; Whitehead, L.; Yu, R.; Deng, G.; Caughlan, R. E.; Woods, A. L.; Jones, A. K.; Barnes, S. W.; Walker, J. R.; Gaulis, S.; Haug, E.; Brachmann, S. M.; Krastel, P.; Studer, C.; Riedl, R.; Estoppey, D.; Aust, T.; Movva, N. R.; Wang, Z.; Salcius, M.; Michaud, G. A.; McAllister, G.; Murphy, L. O.; Tallarico, J. A.; Wilson, C. J.; Dean, C. R. Identification of elongation factor G as the conserved cellular target of argyrimycin B, *PLoS ONE*. **2012**, *7*, pp. e42657.
- (124) 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.; Malek, N. P. Argyrimycin A reveals a critical role for the tumor suppressor protein p27kip1 in mediating antitumor activities in response to proteasome inhibition, *Cancer Res*. **2008**, *14*, pp. 23–35.
- (125) Ley, S. V.; Priour, A.; Heusser, C. Total synthesis of the cyclic heptapeptide Argyrimycin B: a new potent inhibitor of T-cell independent antibody formation, *Org. Lett*. **2002**, *4*, pp. 711–714.



## Chapter 2

### **Heterologous production of myxobacterial $\alpha$ -pyrone antibiotics in *Myxococcus xanthus***

Hilda Sucipto <sup>†</sup>, Domen Pogorevc <sup>†</sup>, Eva Luxenburger, Silke C. Wenzel and Rolf Müller

<sup>†</sup> These authors contributed equally to this work

*Metabolic Engineering*, 2017, 44, 160-170

DOI: [10.1016/j.ymben.2017.10.004](https://doi.org/10.1016/j.ymben.2017.10.004)

Published online: 10.10.2017

## **Contributions to the Presented Work**

### **Author's Contribution**

The author significantly contributed to the conception of the study, designed and performed experiments, evaluated and interpreted the resulting data. The author contributed to the construction of the coralloyronin expression construct and its heterologous expression in the *Myxococcus xanthus* DK1622. Furthermore, the author contributed to the cultivation of the heterologous myxopyronin and coralloyronin production strains as well as the wild type strains. Preparation of the extracts for quantification of the production yields was performed by the author. The author contributed to the conception and writing of the manuscript.

### **Contribution by Co-Workers**

Hilda Sucipto significantly contributed to the conception of the study, designed and performed experiments, evaluated and interpreted the resulting data. She designed and constructed the myxopyronin expression vector. The heterologous expression of this construct in *Myxococcus xanthus* DK1622 was conducted by her. Furthermore, she contributed to the cultivation of the heterologous myxopyronin production strain. She also contributed to the design and assembly of the coralloyronin expression construct and the writing of the manuscript. Eva Luxenburger performed HPLC-MS analysis of the extracts and quantified the coralloyronin and myxopyronin yields. Silke C. Wenzel contributed to the design of both heterologous expression constructs and to the writing of this chapter. The project was supervised by Silke C. Wenzel and Rolf Müller who were responsible for the conception of the project and performed the proofreading of this chapter.

## 2 Heterologous production of myxobacterial $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*

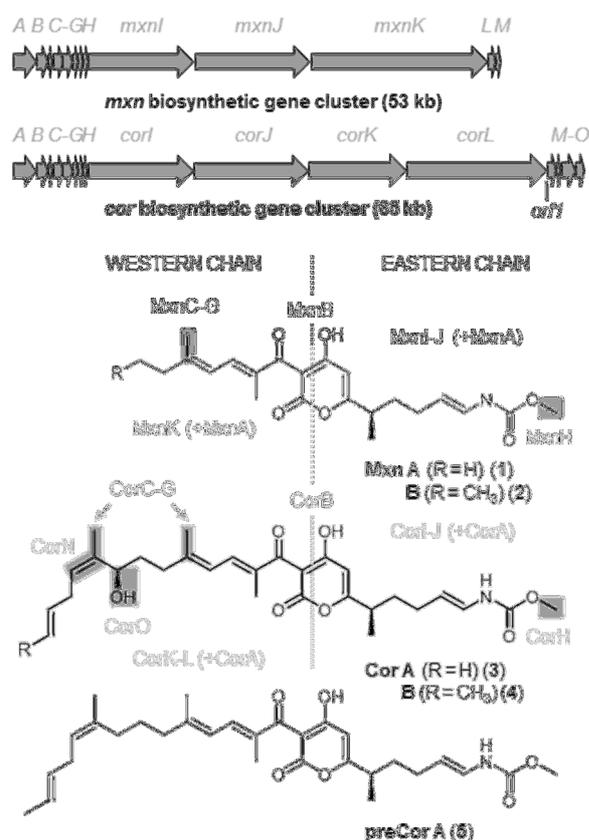
### 2.1 Abstract

Myxopyronins (MXN) and corallopyronins (COR) are structurally related  $\alpha$ -pyrone antibiotics from myxobacteria that represent a highly promising compound class for the development of broad-spectrum antibacterial therapeutic agents. Their ability to inhibit RNA polymerase through interaction with the “switch region”, a novel target, distant from previously characterized RNA polymerase inhibitors (e.g. rifampicin), makes them particularly promising candidates for further research. To improve compound supply for further investigation of MXN, COR and novel derivatives of these antibacterial agents, establishment of an efficient and versatile microbial production platform for myxobacterial  $\alpha$ -pyrone antibiotics is highly desirable. Here we describe design, construction and expression of a heterologous production and engineering platforms for MXN and COR to facilitate rational structure design and yield improvement approaches in the myxobacterial host strain *Myxococcus xanthus* DK1622. Optimization of the cultivation medium yielded significantly higher production titers of MXN A at around 41-fold increase and COR A at around 25-fold increase, compared to the standard CTT medium.

### 2.2 Introduction

Myxobacteria represent a prolific source of novel anti-infectives with unprecedented modes of action<sup>1-3</sup>. Prominent examples are myxopyronins (MXN) and corallopyronins (COR), structurally related  $\alpha$ -pyrone antibiotics discovered from *Myxococcus fulvus* and *Corallocooccus coralloides* strains, respectively<sup>4,5</sup> (Fig. 1). They both inhibit the prokaryotic RNA polymerase (RNAP), which is highly conserved across the bacterial domain and a proven target for broad-spectrum antibacterial therapy<sup>6</sup>. Importantly, the binding site of MXN and COR – the RNAP ‘switch region’<sup>7,8</sup> – is distant from targets of previously characterized RNAP inhibitors including rifamycins (Rif), which are currently used in clinics for treatment of Gram-negative and Gram-positive bacterial infections and represent first-line anti-tuberculosis agents. The clinical utility of Rif is threatened by the rising development and spread of resistant strains that typically harbor RNAP mutations at or adjacent to the Rif binding site<sup>9</sup>. Due to their different RNAP target site MXN and COR do not exhibit cross-

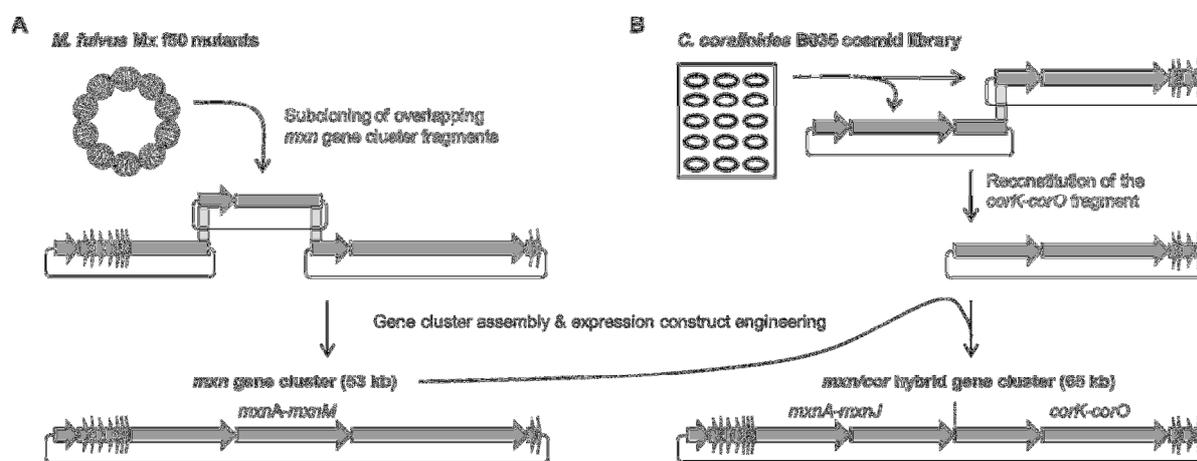
resistance with Rif, which makes them promising candidates for the development of novel broad-spectrum antibiotics<sup>10</sup>. Currently, MXN and analogs thereof are evaluated for treatment of *Mycobacterium tuberculosis*, *Staphylococcus aureus* and other infections with bacterial pathogens<sup>11–13</sup>, whereas COR is investigated for application in filariasis therapy as it was shown to specifically target and deplete obligate *Wolbachia* endobacteria from filarial nematodes<sup>14,15</sup>.



**Figure 1.** Myxopyronin (*mxn*) and corallopyronin (*cor*) biosynthetic gene clusters and their products described from native producer strains.

To improve compound supply for further exploitation of MXN, COR and novel derivatives thereof as antibacterial agents, establishment of an efficient and versatile microbial production platform for myxobacterial  $\alpha$ -pyrone antibiotics is highly desirable. A synthetic route was established but requires more than 22 steps with an overall yield of less than 1%, suggesting that total synthesis is currently not a viable alternative for compound production<sup>16</sup>. Rational bioengineering efforts require insight into the biosynthesis, which was investigated for both compound families in the past years. The corresponding *mxn* and *cor* biosynthetic gene clusters (Fig. 2) were identified from the native producer strains *M. fulvus* Mxf50<sup>22</sup> and *C. coralloides* B035<sup>31</sup>, respectively, and shown to encode multimodular polyketide synthase

(PKS) systems belonging to the growing class of *trans*-AT PKS<sup>17</sup>. A striking feature of MXN and COR biosynthesis is the assembly of two polyketide chains, designated as eastern and western chain, which are finally condensed by a pyrone-ring forming ketosynthase MxnB/CorB<sup>18,19</sup>. The chains are generated by two separate PKS assembly lines, which both depend on the acyltransferase (AT) and enoylreductase (ER) activity of the bifunctional MxnA/CorA provided *in trans*. As expected from the identical substructure, MXN and COR eastern chains are generated via the same biosynthetic route by homologous assembly lines. They consist of two PKS subunits (MxnI-J/CorI-J) and harbor a nonribosomal peptide synthetase (NRPS) module<sup>20</sup> for incorporation of an amino acid building block (glycine). The terminal eastern chain carboxy moiety is modified by the *O*-methyltransferase MxnH/CorH<sup>21</sup>. In contrast to eastern chain assembly, MXN and COR western chain biosynthesis differs significantly due to additional PKS elongation cycles and modifications in the case of COR revealing a more bulky western part<sup>22</sup>. Whereas the COR western chain PKS system consists of two subunits (CorK-L), the Mxn PKS (MxnK) is much smaller. In both cases, different starter units can be employed (acetyl-CoA or propionyl-CoA) resulting in the production of MXN A/B (1/2) or COR A/B (3/4)<sup>4,5</sup> (Fig. 1). Western chain assembly includes  $\beta$ -branching<sup>23</sup>, a typical feature of *trans*-AT PKS systems resulting in methyl group substituents. It is mediated by a set of five proteins (MxnC-MxnG/CorC-CorG) and occurs twice during COR biosynthesis. Additional modifying enzymes encoded by the *cor* gene cluster are responsible for double-bond isomerization (CorN) and hydroxylation (CorO). Deficiency of the latter modification catalyzed by a P450 enzyme results in the production of so-called ‘precorallopynons’, like preCOR A (5)<sup>24</sup>.



**Figure 2.** Strategies for the generation of expression constructs for heterologous myxopyronin (A) and corallopynonin (B) production.

In the present study we aimed at establishing heterologous production and engineering platforms for MXN and COR to facilitate rational structure and yield improvement approaches, which turned out to be challenging with the native producers (see e.g.<sup>25</sup>). The myxobacterial model strain *Myxococcus xanthus* DK1622<sup>26</sup> was selected as heterologous host due to its close phylogenetic relationship to the native producers plus additional advantageous features. These include established genetic tools, growth in suspension, short doubling time (~ 5 hours) compared to other myxobacteria as well as available fermentation know-how<sup>27</sup>. In addition, the strain is characterized as promising producer of natural products<sup>28</sup> indicating the availability of common biosynthetic precursors, and two broad spectrum phosphopantetheinyl transferases were identified<sup>29</sup>, which are required for posttranslational activation of PKS/NRPS megasynthetases. Thus, *M. xanthus* DK1622 represents a promising host for heterologous expression of PKS/NRPS biosynthetic pathways, especially those from other myxobacteria, which is underpinned by successful examples from previous work<sup>27</sup>. To engineer biosynthetic gene clusters for heterologous MXN and COR production in *M. xanthus* different strategies were envisaged (Fig. 2). In case of MXN, we aimed at subcloning of *mxn* gene cluster fragments by plasmid recovery from chromosomal DNA of *M. fulvus* Mx f50 mutants that can be generated via targeted integration of suicide vectors as established previously<sup>22</sup>. Employing recombination-based cloning technologies, like Red/ET recombineering<sup>30</sup>, the overlapping gene cluster fragments can be subsequently stitched to reconstitute the 53 kb *mxn* pathway on one expression vector (Fig. 2A). Regarding COR, the initial situation was different as methods for genetic manipulation of the native producer strain have not been established yet, but a cosmid library was available<sup>31</sup>. Based on this, we intended to reconstitute the downstream part of the *cor* gene cluster (*corK-corO*), which diverges from the *mxn* pathway as expressed by the different COR/MXN western chains<sup>22</sup>. Replacement of the corresponding *mxn* gene cluster region (*mxnK-mxnM*) with the *corK-corO* gene set should yield a 65 kb *mxn/cor* hybrid pathway for COR production (Fig. 2B). However, this depends on the effective interaction of the chimeric MXN/COR biosynthesis machinery, which we considered likely due to the close relationship of both pathways and the homology of the functionally equivalent *mxnA-mxnJ* and *corA-corJ* gene cluster regions<sup>22</sup>. Overall, the heterologous production platforms described here should facilitate access to MXN, COR and novel derivatives thereof to set the stage for further exploitation of myxobacterial  $\alpha$ -pyrone antibiotics for therapeutic applications.

## 2.3 Material and methods

### 2.3.1 Sequence analysis and *mxn* and *cor* gene cluster sequences

Sequence analysis and routine *in silico* DNA work was performed using Geneious v8-v10 software package (Biomatters Ltd.). The sequence of the myxopyronin biosynthetic gene cluster from *Myxococcus fulvus* Mx f50<sup>22</sup> and the coralloyronin biosynthetic gene cluster from *Coralloccoccus coralloides* B035<sup>31</sup> was retrieved from GenBank: KF356280 and HM071004. In addition to restriction analysis, integrity of the generated expression constructs pHSU-mxn43 (harboring the *mxn* gene cluster) and pDPO-mxn116 (harboring the *mxn/cor* hybrid gene cluster) as well as of the intermediate construct pHSU-mxn113b (harboring *corK-corO*) was verified by sequencing. Using the Illumina paired-end technology on a MiSeq platform a 1806 fold (pHSU-mxn43), a 480-fold (pDPO-mxn116) and 722-fold (pHSU-mxn113) mean sequencing coverage was achieved and no mutations within the gene cluster region were detected.

### 2.3.2 Construction and engineering of plasmids

Routine handling of nucleic acids, such as isolation of plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, was performed according to standard protocols<sup>32</sup>. *E. coli* HS996 (Invitrogen) was used as host for standard cloning experiments and *E. coli* SCS110 (Stratagene) for preparation of plasmid DNA free of Dam or Dcm methylation. *E. coli* strains were cultured in LB medium or on LB agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) at 30-37 °C (and 200 rpm) overnight. Antibiotics were used at the following final concentrations: 100  $\mu$ g/mL ampicillin, 50  $\mu$ g/mL kanamycin, 5  $\mu$ g/mL tetracycline, 20  $\mu$ g/mL zeocin, 12.5  $\mu$ g/mL chloramphenicol. Transformation of *E. coli* strains was achieved via electroporation in 0.1 cm-wide cuvettes at 1250 V, a resistance of 200  $\Omega$ , and a capacitance of 25  $\mu$ F. Plasmid, cosmid and BAC DNA was either purified by standard alkaline lysis<sup>32</sup> or by using the GeneJet Plasmid Miniprep Kit (Thermo Fisher Scientific) or the NucleoBond PC100 kit (Machery Nagel). Restriction endonucleases, alkaline phosphatase (FastAP) and T4 DNA ligase were obtained from Thermo Fisher Scientific. Oligonucleotides used for PCR and sequencing were obtained from Sigma-Aldrich and are listed in Table S1. PCR reactions were carried out in a peqSTAR 96 universal gradient thermocycler (Peqlab) or Mastercycler® pro (Eppendorf) using Phusion™ High-Fidelity or Taq DNA polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. For Taq: Initial denaturation (3 min, 95 °C); 30 cycles of denaturation (30 s, 95 °C),

annealing (30 s, 53 or 57 °C) and elongation (varied based on PCR product length 1 kb/min, 72 °C); and final extension (10 min, 72 °C). For Phusion™: Initial denaturation (30 s, 98 °C); 30 cycles of denaturation (20 s, 98 °C), annealing (25 s, 53 or 57 °C) and elongation (varied based on PCR product length 0.5 kb/min, 72 °C); and final extension (10 min, 72 °C). PCR products or DNA fragments from restriction digests were purified by agarose gel electrophoresis and isolated using the peqGold Gel Extraction (Pepq). The PCR products were cloned into pJET1.2/blunt vector using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific) or pCRII-TOPO vector using the TOPO TA Cloning Kit (Invitrogen) and sequenced using the primers pJET1.2For/pJET1.2Rev or M13For/M13Rev (Table S1). Red/ET recombineering experiments for plasmid modifications<sup>30</sup> using the strains *E. coli* HS996/pSC101-BAD-gbaA-tet or *E. coli* GB05-red were performed according to the manufacturer's protocol (Gene Bridges GmbH). After selection with suitable antibiotics, clones harboring correct recombination products were identified by plasmid isolation and restriction analysis with a set of different endonucleases. Details on the construction of all plasmids generated in this study are given in Table 1.

**Table 1.** Genetic constructs generated in this study.

| Plasmid name  | Construction details   | Features  |
|---------------|--|---|
| pJet-Tn5-mxnA | 1374 bp <i>P<sub>Tn5</sub>-mxnA'</i> (1191 bp from <i>mxnA</i> 5' end) fragment amplified via overlap extension (OE) PCR using primers mxn33/mxn30 and subcloned into pJET1.2/blunt; templates used for OE PCR: <i>P<sub>Tn5</sub></i> amplified from pCRII-TOPO (Invitrogen) using primers mxn33/41 and a part of <i>mxnA'</i> amplified from <i>M. fulvus</i> Mx f50 genomic DNA using primers mxn42/mxn30 | <i>P<sub>Tn5</sub>-mxnA'</i> , pMB1 ori, <i>amp<sup>R</sup></i>                 |
| p15A-Tet-Zeo  | 2946 bp <i>PvuI-zeo<sup>R</sup>-PvuI</i> fragment amplified from pCDNA-Zeo (Invitrogen) using primers mxn128/mxn129 and ligated into a p15A-Tet minimal vector backbone, which was amplified from pACYC184 <sup>33</sup> using primers mxn126/mxn127 and hydrolyzed with <i>AvrII</i>  | p15A ori, <i>tet<sup>R</sup></i> , <i>zeo<sup>R</sup></i>                       |
| pHSU-mxn16    | 1371 bp <i>PacI-P<sub>Tn5</sub>-mxnA'</i> (1191 bp from <i>mxnA</i> 5' end)- <i>EcoRV</i> fragment from pJet-Tn5-mxnA ligated into a p15A-Tet minimal vector backbone, which was amplified from pACYC184 <sup>33</sup> using primers mxn52/53 and hydrolyzed with <i>PacI/PvuII</i>  | <i>P<sub>Tn5</sub>-mxnA'</i> , p15A ori, <i>tet<sup>R</sup></i>                 |
| pHSU-mxn19    | 1491 bp <i>EcoRV-mxnM</i> plus flanking orf- <i>NheI-NdeI-SspI-PacI</i> fragment amplified from <i>M. fulvus</i> Mx f50 genomic DNA using primers mxn64/mxn65 and subcloned into pCRII-TOPO (Invitrogen)   | <i>mxnM</i> , pUC ori, <i>kan<sup>R</sup></i> , <i>amp<sup>R</sup></i>          |
| pHSU-mxn26    | Plasmid recovery of a 31.6 kb <i>EcoRV-mxnJ''</i> (6.2 kb from <i>mxnJ</i> 3' end)- <i>mxnKLM</i> -pUCori- <i>amp<sup>R</sup>-kan<sup>R</sup>-EcoRV</i> fragment from chromosomal DNA of the mutant <i>M. fulvus</i> Mxf50::pHSU-mxn16::pHSU-mxn19, which was hydrolyzed with <i>EcoRV</i> and re-ligated  | <i>mxnJ''-mxnK-M</i> , pUC ori, <i>kan<sup>R</sup></i> , <i>amp<sup>R</sup></i> |

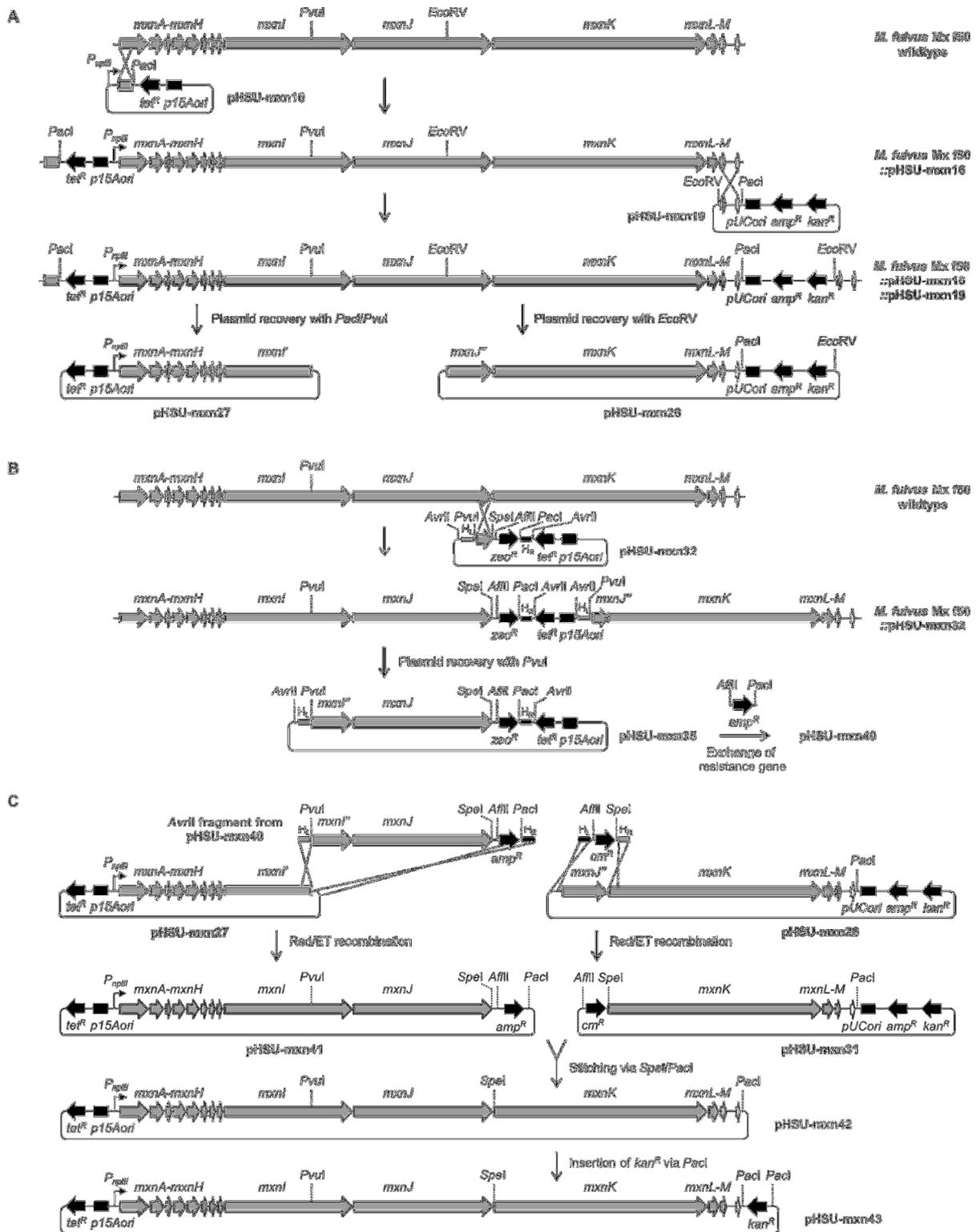
|              |   |   |
|--------------|---|---|
| pHSU-mxn27   | Plasmid recovery of a 17.1 kb <i>PacI-tet<sup>R</sup>-p15Aori-P<sub>Tn5</sub>-mxnABCDEFGH-mxnI'</i> (5.9 kb from <i>mxnI</i> 5' end)- <i>PvuI</i> fragment from chromosomal DNA of the mutant <i>M. fulvus</i> Mxf50::pHSU-mxn16::pHSU-mxn19, which was hydrolyzed with <i>PacI/PvuI</i> and re-ligated | <i>P<sub>Tn5</sub>-mxnA-H-mxnI'</i> ,<br>p15A ori, <i>tet<sup>R</sup></i>   |
| pHSU-mxn31   | Red/ET recombineering between pHSU-mxn26 and a 979 bp <i>AflIII-cm<sup>R</sup>-SpeI</i> linear fragment amplified from pACYC184 <sup>33</sup> using primers mxn122/mxn123   | <i>mxnK-M</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, cm<sup>R</sup></i>                               |
| pHSU-mxn32   | 1410 bp <i>PvuI-mxnJ''</i> (1380 bp from <i>mxnJ</i> 3' end)- <i>SpeI-mxnK'</i> - <i>AflIII</i> fragment amplified from <i>M. fulvus</i> Mx f50 genomic DNA using primers mxn124/mxn125 and ligated into p15A-Tet-Zeo hydrolyzed with <i>PvuI/AflIII</i>  | <i>mxnJ''</i><br>p15A ori, <i>tet<sup>R</sup>, zeo<sup>R</sup></i>  |
| pHSU-mxn35   | Plasmid recovery of a 20.3 kb <i>PvuI-mxnI''</i> (5.8 kb from <i>mxnI</i> 3' end)- <i>mxnJ-zeo<sup>R</sup>-tet<sup>R</sup>-p15Aori-PvuI</i> fragment from chromosomal DNA of the mutant <i>M. fulvus</i> Mxf50::pHSU-mxn32, which was hydrolyzed with <i>PvuI</i> and re-ligated                        | <i>mxnI''-mxnJ</i> ,<br>p15A ori, <i>tet<sup>R</sup>, zeo<sup>R</sup></i>   |
| pHSU-mxn37   | 949 bp <i>PacI-kan<sup>R</sup>-PacI</i> fragment amplified from pCRII-TOPO (Invitrogen) using primers mxn144/145 and subcloned into pJET1.2/blunt (ThermoFisher Scientific)   | pMB1 ori, <i>amp<sup>R</sup>, kan<sup>R</sup></i>   |
| pHSU-mxn38   | 1086 bp <i>AflIII-amp<sup>R</sup>-PacI</i> fragment amplified from pJET1.2/blunt (ThermoFisher Scientific) using primers mxn147/mxn148 and subcloned into pJET1.2/blunt   | pMB1 ori, <i>amp<sup>R</sup></i> ,  |
| pHSU-mxn40   | 1086 bp <i>AflIII-amp<sup>R</sup>-PacI</i> fragment from pHSU-mxn38 ligated into pHSU-mxn35 hydrolyzed with <i>AflIII/PacI</i>  | <i>mxnI''-J</i> ,<br>p15A ori, <i>tet<sup>R</sup>, amp<sup>R</sup></i>  |
| pHSU-mxn41   | Red/ET recombineering between pHSU-mxn27 and a 18.5 kb <i>AvrII-mxnI''</i> (5.8 kb from <i>mxnI</i> 3' end)- <i>mxnJ- amp<sup>R</sup>-AvrII</i> linear fragment obtained from <i>AvrII</i> hydrolysis of pHSU-mxn40   | <i>P<sub>Tn5</sub>-mxnA-J</i> ,<br>p15A ori, <i>tet<sup>R</sup>, amp<sup>R</sup></i>                              |
| pHSU-mxn42   | 21.5 kb <i>SpeI-mxnKLM-PacI</i> fragment derived from pHSU-mxn31 ligated into pHSU-mxn41 hydrolyzed with <i>SpeI/PacI</i>   | <i>P<sub>Tn5</sub>-mxnA-M</i> , p15A ori, <i>tet<sup>R</sup></i>  |
| pHSU-mxn43   | 949bp <i>PacI-kan<sup>R</sup>-PacI</i> fragment derived from pHSU-mxn37 ligated into pHSU-mxn42 hydrolyzed with <i>PacI</i>   | <i>P<sub>Tn5</sub>-mxnA-M</i> ,<br>p15A ori, <i>tet<sup>R</sup>, kan<sup>R</sup></i>                              |
| pHSU-mxn51   | Red/ET recombineering between cosmid BA5 and a 913 bp <i>AflIII-apra<sup>R</sup>-AflIII-PacI</i> amplified from pKC1132 <sup>34</sup> using primers mxn200/mxn201   | <i>corK''-corL-orf1-corM-O</i> , pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, apra<sup>R</sup></i>               |
| pHSU-mxn52   | Red/ET recombineering between cosmid AM24 and a 923 bp <i>NdeI-PmeI-apra<sup>R</sup>-PmeI</i> linear fragment amplified from pKC1132 <sup>34</sup> using primers mxn202/mxn203  | <i>corI''-corJ-corK''</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, apra<sup>R</sup></i>                 |
| pHSU-mxn53   | Red/ET recombineering between pHSU-mxn52 and a 933 bp <i>NdeI-cm<sup>R</sup>-SpeI</i> fragment amplified from pACYC184 <sup>33</sup> using primers mxn209/mxn210  | <i>corK'</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, apra<sup>R</sup>, cm<sup>R</sup></i>              |
| pHSU-mxn54   | Red/ET recombineering between pHSU-mxn51 and a 9.5 kb <i>NdeI-cm<sup>R</sup>-corK'</i> (8.5 kb from <i>corK</i> 5' end)- <i>NdeI</i> linear fragment obtained from <i>NdeI</i> hydrolysis of pHSU-mxn53   | <i>corK-L-orf1-corM-O</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, apra<sup>R</sup>, cm<sup>R</sup></i> |
| pHSU-mxn56   | pHSU-mxn54 hydrolyzed with <i>AflIII</i> and re-ligated (to eliminate <i>AflIII-apra<sup>R</sup>-AflIII</i> cassette)   | <i>corK-L-orf1-corM-O</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, cm<sup>R</sup></i>                   |
| pHSU-mxn113a | Red/ET recombineering between pHSU-mxn56 and a 1401 bp <i>PmeI-tet<sup>R</sup>-AflIII-PacI</i> fragment amplified from pACYC184 <sup>33</sup> using primers mxn330/mxn331   | <i>corK-L-orf1-corM-O</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, cm<sup>R</sup>, tet<sup>R</sup></i>  |
| pHSU-mxn113b | 1401 bp <i>PmeI-tet<sup>R</sup>-AflIII-PacI</i> fragment amplified from pACYC184 using primers Tet-for-AflIII-PacI/Tet-rev-PmeI and ligated into pHSU-mxn113a hydrolyzed with <i>PmeI/PacI</i>  | <i>corK-L-orf1-corM-O</i> ,<br>pUC ori, <i>kan<sup>R</sup>, amp<sup>R</sup>, cm<sup>R</sup>, tet<sup>R</sup></i>  |

|                |   |   |
|----------------|---|---|
| pBeloBac11-Kan | Red/ET recombineering between pBeloBAC11 (NEB) and a 1063 bp <i>PacI-kan<sup>R</sup>-AvrII-SwaI</i> linear fragment amplified from pCRII-TOPO using primers Kan-BeloBAC-for/Kan-BeloBAC-rev ( <i>cm<sup>R</sup></i> exchanged with <i>kan<sup>R</sup></i> ) | <i>ori2, sopA, sopB, sopC, repE, kan<sup>R</sup>,</i>   |
| pDPO-mxn115    | 33.2 kb <i>Bst1107I-P<sub>Tn5</sub>-mxnABCDEFGHIJ-amp<sup>R</sup>-PacI</i> fragment from pHSU-mxn41 ligated into pBeloBac11-Kan hydrolyzed with <i>Bst1107I/PacI</i>  | <i>P<sub>Tn5</sub>-mxnA-J, ori2, sopA, sopB, sopC, repE, kan<sup>R</sup>, amp<sup>R</sup></i>                   |
| pDPO-mxn116    | 34.8 kb <i>SpeI-corKL-orf1-corMNO-tet<sup>R</sup>-PacI</i> fragment from pHSU-mxn113b ligated into pDPO-mxn115 hydrolyzed with <i>SpeI/PacI</i>   | <i>P<sub>Tn5</sub>-mxnA-J-corKL-orf1-corM-O, ori2, sopA, sopB, sopC, repE, kan<sup>R</sup>, tet<sup>R</sup></i> |

### 2.3.3 Generation of *M. fulvus* Mx f50 mutant strains and subcloning of *mxn* gene cluster fragments

According to a previously established electroporation procedure<sup>22</sup> the native myxopyronin producer *M. fulvus* Mx f50<sup>4</sup> was transformed with the suicide plasmids pHSU-mxn16, pHSU-mxn19 and pHSU-mxn32 (Table 1). *M. fulvus* Mx f50 wild type and mutants were cultivated at 30 °C in Casitone Yeast (CY) medium or on CY agar (0.3% casitone, 0.1% yeast extract, 0.1% CaCl<sub>2</sub> x 2 H<sub>2</sub>O, (1.5% agar)) supplemented with 0.5 mg/L vitamin B12 after autoclaving. For liquid cultures, the strains were grown in Erlenmeyer flasks on an orbital shaker at 180 rpm for 3-5 days. Antibiotics for selection of *M. fulvus* mutants were used at the following final concentrations: 50 µg/mL kanamycin and/or 6.25 µg/mL oxytetracycline. Correct chromosomal integration of the suicide plasmids via homologous recombination was verified by PCR using genomic DNA of the transformants as template isolated with the Genra Puregene Genomic DNA Purification Kit (Qiagen) according to the manufacturer's protocol. For each suicide plasmid correct chromosomal integration was confirmed using two different primer combinations revealing PCR products of the expected sizes: pHSU-mxn16, primers mxn60/p15A-Tet1 (1342 bp) and mxn61/p15A-Tet2 (1473 bp); pHSU-mxn19, primers mxn94/pTOPO-out (1720 bp) and mxn95/pTOPO-in (1594 bp); pHSU-mxn32, primers mxn136/p15A-Tet1 (1992 bp) and p15A-Tet2/mxn137 (1527 bp); for primer sequences see Table S1. PCR reactions were performed according to the conditions described in chapter 2.2. using Taq polymerase. For plasmid recovery approaches, high molecular weight chromosomal DNA of the verified *M. fulvus* Mx f50::pHSU-mxn16::pHSU-mxn19 and *M. fulvus* Mx f50::pHSU-mxn32 mutants was isolated via the phenol chloroform isoamyl alcohol extraction method<sup>32</sup> and hydrolyzed with suitable endonucleases (see Fig. 3). Subsequent re-ligation and transformation of *E. coli* HS996 with the ligation mixture revealed clones harboring the expected plasmids pHSU-mxn26, pHSU-mxn27 and pHSU-mxn35

(Table 1), which were verified by restriction analysis. The subcloned chromosomal fragments of these three plasmids cover the entire *mxn* biosynthetic gene cluster region (Fig. 3).



**Figure 3.** Generation of an expression construct for the myxopyronin (*mxn*) biosynthetic gene cluster. A/B: Subcloning of three overlapping gene cluster fragments from the native producer *M. fulvus* Mx f50 via chromosomal integration of suicide plasmids and subsequent plasmid recovery. C: Reconstitution of the entire *mxn* gene cluster on a *p15Aori*-based expression vector.

### 2.3.4 Transfer and chromosomal integration of the expression constructs into the heterologous host

According to a previously established electroporation procedure for *Myxococcus xanthus* DK1622<sup>35</sup> the host strain *M. xanthus* DK1622  $\Delta$ *mchA-tet* (Wenzel et al., unpublished) was transformed with the generated expression constructs pHSU-mxn43 and pDPO-mxn116. *M. xanthus* DK1622  $\Delta$ *mchA-tet* is a mutant in which the myxochromide A gene cluster has been deleted. Myxochromide A is one of the major compounds produced by *M. xanthus* DK1622, and its gene cluster deletion was expected to provide an increased precursor pool for production of other secondary metabolites. In brief, the myxochromide A gene cluster was completely deleted from *M. xanthus* DK1622 and replaced by a tetracycline resistance gene, which served as the target site for the chromosomal integration of both expression constructs. *M. xanthus* DK1622 mutants were routinely cultivated at 30 °C in CTT medium or an CTT agar (1% casitone, 10 mM Tris buffer pH 7.6, 1 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.6, 8 mM MgSO<sub>4</sub> (1.5% agar) with final pH 7.6). For liquid cultures, the strains were grown in Erlenmeyer flasks on an orbital shaker at 180 rpm for 3-6 days. For selection of *M. xanthus* mutants 50  $\mu$ g/mL kanamycin was used. Correct chromosomal integration of the expression constructs via homologous recombination into the *tet<sup>R</sup>* locus was verified by PCR. Genomic DNA of the pHSU-mxn43 transformants was isolated using the Gentra Puregene Yeast/Bact. Genomic DNA Purification Kit (Qiagen) and in case of the pDPO-mxn116 transformants cells were lysed by incubating at 95 °C for 20 minutes prior of being added to the PCR reaction ('colony PCR'). For each expression construct correct chromosomal integration was confirmed using two different primer combinations revealing PCR products of the expected sizes: pHSU-mxn43, primers P4/P2 (1635 bp) and P3/P1 (1609 bp); pDPO-mxn116, primers P4/P6 (1632 bp) and P5/P1 (1519 bp) (Fig. S1). In both cases, genomic DNA of *M. xanthus* DK1622  $\Delta$ *mchA-tet* was used as negative control. A complementary experiment using primers P4/P1 revealed a 1461 bp PCR product for *M. xanthus* DK1622  $\Delta$ *mchA-tet*, but not for the expression strains *M. xanthus* DK1622  $\Delta$ *mchA-tet*::pHSU-mxn43 and *M. xanthus* DK1622  $\Delta$ *mchA-tet*::pDPO-mxn116. PCR reactions were performed according to the conditions described in chapter 2.2. using Taq polymerase; for primer sequences see Table S1.

### 2.3.5 Initial cultivation experiments with the heterologous expression strains in comparison to the native producers

To evaluate the production titers of MXN and COR the heterologous producers (*M. xanthus* DK1622  $\Delta mchA-tet::pHSU-mxn43$  and *M. xanthus* DK1622  $\Delta mchA-tet::pDPO-mxn116$ ), host strain (*M. xanthus* DK1622  $\Delta mchA-tet$ ) and native producers (*M. fulvus* Mx f50<sup>4</sup> and *C. coralloides* Cc c127<sup>5</sup>) were cultivated in parallel. Strains were inoculated from cryo stocks and grown on agar plates for several days until plates were mostly overgrown with cells. All of the cells were scraped from the plates to inoculate preculture medium (50 mL medium in 300 mL Erlenmeyer flask), which was cultivated at 30 °C, 180 rpm for 48 h. 5 mL of well grown preculture was used to inoculate 50 mL production medium in which the strain was grown at the same conditions for 6 days. In case of *M. xanthus* DK1622 derivatives, CTT medium (see 2.3.4) amended with suitable antibiotics was used (50  $\mu$ g/mL kanamycin for heterologous producers and 6  $\mu$ g/mL oxytetracycline for the host strain), *M. fulvus* Mx f50 was grown in the BTY medium (0.6% tryptone, 0.05% yeast extract, 0.2% MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 0.04% CaCl<sub>2</sub> x 2 H<sub>2</sub>O, with final pH 7.2)<sup>4</sup> and *C. coralloides* Cc c127 in MD1/4 medium (0.6% tryptone, 0.2% soluble starch, 0.2% MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 0.04% CaCl<sub>2</sub> x 2 H<sub>2</sub>O, with final pH 7.2)<sup>5</sup>. All cultivations were performed in triplicates at 30 °C and 180 rpm for 6 days. Cultures were centrifuged at 8000 rpm, 4 °C for 15 min to separate supernatant and cells. After addition of one aliquot (55 mL) MeOH to the supernatant and incubation in a culture flask on an orbital shaker (180 rpm) for 1 h, 2 mL of the supernatant/MeOH mixture was transferred to an Eppendorf tube and centrifuged at 15000 rpm, 4 °C for 15 minutes prior the LC-MS analysis. The cells were extracted separately by adding 50 mL MeOH to 50 mL falcon tube and rotating at the rotation wheel for 1 h. After filtration the extract was evaporated and re-dissolved in 0.5 mL of MeOH prior to the LC-MS analysis. Quantification of production yields was performed as described in chapter 2.3.7, results are shown in Table S6 (MXN A) and Table S7 (COR A).

### 2.3.6 Additional cultivation experiments with the heterologous expression strains in different production media

After initial evaluation of M7s medium for MXN production (Table S2), the medium composition was optimized to develop M7/s6 medium (see SI). To analyze the effect on MXN and COR production in comparison to the CTT (standard medium for *M. xanthus* DK1622) and the native producer media (MD1/4 and BTY) parallel cultivation experiments

were performed and production yields were quantified (Table 2). The strains were inoculated from glycerol stocks and grown for several days on CTT agar plates supplemented with 50  $\mu$ g/mL kanamycin, until plates were mostly covered with cell mass. All of the cells were scraped from the plates to inoculate preculture media; M7/s4 (0.5% soy flour, 0.5% corn starch, 0.2% glucose, 0.1% yeast extract, 0.1%  $\text{MgSO}_4 \times 7 \text{H}_2\text{O}$ , 0.1 %  $\text{CaCl}_2 \times 2 \text{H}_2\text{O}$ , 1% HEPES, with final pH 7.4 and supplemented with 0.1 mg/L of vitamin B12 and 5 mg/L of  $\text{FeCl}_3$  after autoclaving), CTT (see 2.3.4), BTY and MD1/4 (see 2.3.5) supplemented with 50  $\mu$ g/mL kanamycin (50 mL medium in 300 mL Erlenmeyer flask). Precultures were cultivated at 30 °C, 180 rpm for 48 h. Five mL of well grown preculture was used to inoculate 50 mL production medium supplemented with 50  $\mu$ g/mL kanamycin in which strain was grown at the same conditions for 6 days. M7/s6 (0.5% soy flour, 0.5% corn starch, 0.2% glucose, 0.1% yeast extract, 1% potassium acetate, 0.1%  $\text{MgSO}_4 \times 7 \text{H}_2\text{O}$ , 0.1%  $\text{CaCl}_2 \times 2 \text{H}_2\text{O}$ , 1% HEPES, with final pH 7.4 and supplemented with 0.1 mg/L of vitamin B12 and 5 mg/L of  $\text{FeCl}_3$  after autoclaving)) was used as production medium for M7/s4 preculture and for other precultures the same medium as for the preculture was used for production. All cultivations were done in triplicates. After 6 days of cultivation at 30 °C and 180 rpm, cultures were extracted by addition of 55 mL of MeOH directly to the culture flask which was then incubated at 180 rpm for 1 h on orbital shaker. Two mL of the culture broth/MeOH mixture was centrifuged in 2 mL Eppendorf tubes at 15000 rpm at 4 °C for 15 min. The supernatant was subjected to LC-MS analysis to determine MXN A and COR A production yields.

**Table 2.** MXN A and COR A production yields achieved with the heterologous expression strains in different media after 6 days of cultivation.

| Medium                              | MXN A [mg/L] <sup>a</sup> | COR A [mg/L] <sup>b</sup> |
|-------------------------------------|---------------------------|---------------------------|
| CTT                                 | 3.79 $\pm$ 0.09           | 1.48 $\pm$ 0.14           |
| Native producer medium <sup>c</sup> | 3.03 $\pm$ 0.11           | 0.98 $\pm$ 0.09           |
| M7/s6                               | 156 $\pm$ 12              | 37 $\pm$ 4                |

<sup>a</sup> *M. xanthus* DK1622  $\Delta$ mchA-tet::pHSU-mxn43

<sup>b</sup> *M. xanthus* DK1622  $\Delta$ mchA-tet::pDPO-mxn116

<sup>c</sup> BTY medium for MXN A; MD1/4 medium for COR A

### 2.3.7 Analysis of myxopyronin and coralloyronin production by HPLC-MS

All measurements were performed on a Dionex Ultimate 3000 RSLC system using a BEH C18, 100 x 2.1 mm, 1.7  $\mu$ m dp column equipped with a C18 precolumn (Waters, Germany). Separation of a 1  $\mu$ L sample was achieved by a linear gradient from (A) H<sub>2</sub>O + 0.1% FA to (B) ACN + 0.1% FA at a flow rate of 600  $\mu$ L/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step at 5% B, followed by an increase to 50% B in 1 min, to 65% B in 6 min and to 98% B in 0.5 min. After a 1 min step at 98% B the system was re-equilibrated to the initial conditions (5% B). UV spectra were recorded by a DAD in the range from 200 to 600 nm. For MS measurements on solariX XR (7T) FT-ICR mass spectrometer (Bruker Daltonics, Germany), the LC flow was split to 75  $\mu$ L/min before entering the mass spectrometer using the Apollo ESI source. In the source region, the temperature was set to 200 °C, the capillary voltage was 4500 V, the dry-gas flow was 4.0 L/min and the nebulizer was set to 1.1 bar. After the generated ions passed the quadrupole with a low cutoff at 150 m/z they were trapped in the collision cell for 80 ms and finally transferred within 0.9 ms through the hexapole into the ICR cell. Captured ions were excited by applying a frequency sweep from 150 to 950 m/z and detected in broadband mode by acquiring a 184 ms transient. For quantification of MXN A the peak area of [M+H]<sup>+</sup> 418.2225 m/z at 5.60 min was integrated and compared to a calibration curve obtained from serial dilutions of authentic MXN A reference material covering the concentration range from 0.5 - 100  $\mu$ g/mL (generated from 1 mg/L stock solution in acetonitrile). For quantification of COR A the peak area of [M+H-H<sub>2</sub>O]<sup>+</sup> 510.2850 m/z at 5.83 min was integrated and compared to a calibration curve obtained from serial dilutions of authentic COR A reference material covering the concentration range from 0.5-100  $\mu$ g/mL (generated from 1 mg/L stock solution in acetonitrile).

## 2.4 Results and discussion

As illustrated in Fig. 2 different strategies were pursued to generate expression constructs for the heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *M. xanthus*. We first aimed to achieve successful heterologous expression of the myxopyronin biosynthetic pathway (see 2.4.1-2.4.3) prior to developing an expression system based on a hybrid gene cluster for coralloyronin production (see 2.4.4-2.4.6). After establishment and analysis of heterologous production under standard cultivation conditions, different media were evaluated to improve myxopyronin and coralloyronin yields (see 2.4.7).

### 2.4.1 Subcloning of *mxn* gene cluster fragments from *M. fulvus* Mx f50

Based on a previously established directed mutagenesis procedure for the myxopyronin producer *M. fulvus* Mx f50<sup>22</sup> mutant strains were generated allowing for subcloning the *mxn* biosynthetic gene cluster (or fragments thereof) via plasmid recovery from chromosomal DNA. In our initial approach we aimed to capture the entire pathway in one step based on unique restriction sites (R-sites) introduced upstream and downstream of the *mxn* gene cluster. For this purpose, two suicide vectors with compatible selection markers, pHSU-*mxn*16 (*tet*<sup>R</sup>) and pHSU-*mxn*19 (*kan*<sup>R</sup>), were constructed. The pACYC184 derivative pHSU-*mxn*16 harbors a 1.2 kb fragment homologous to the 5' end of the *mxnA* gene (Fig. 3A). This fragment was fused with the 5'-untranslated region and *PnptII* promoter sequence from the Tn5 *kan*<sup>R</sup> resistance gene in order to replace the native transcriptional/translational initiation region of *mxnA* after chromosomal integration. In addition, a unique *PacI* site was introduced downstream of the *mxnA* homology region to enable plasmid recovery approaches. Transformation of *M. fulvus* Mx f50 with this suicide vector yielded mutant *M. fulvus* Mx f50::pHSU-*mxn*16. For genetic modification at the 3' end of the *mxn* gene cluster a 1.5 kb chromosomal fragment of *M. fulvus* Mx f50 including the *mxnM* gene plus downstream region was amplified and subcloned into pCRII-TOPO revealing pHSU-*mxn*19 (Fig. 3A). The homologous fragment was flanked with *EcoRV* and *PacI* R-sites to enable different plasmid recovery approaches after single cross over into *M. fulvus* Mx f50::pHSU-*mxn*16. The chromosome of the resulting mutant strain *M. fulvus* Mx f50::pHSU-*mxn*16::pHSU-*mxn*19 harbors two plasmid backbones integrated upstream and downstream of the 53 kb *mxn* gene cluster. Correct integration of both plasmids was verified via PCR analyses. Next, we aimed to recover the entire pathway together with the flanking *p15Aori-tet*<sup>R</sup> fragment in one step via *PacI* hydrolysis and re-ligation. However, as the expected plasmid could not be obtained (probably due to the large size of the target fragment) the strategy was adapted for subcloning of three overlapping gene cluster fragments. Based on chromosomal DNA from the *M. fulvus* Mx f50::pHSU-*mxn*16::pHSU-*mxn*19 mutant two separate plasmid recovery approaches were performed (Fig. 3A). Hydrolysis with *PacI/PvuI* (generating compatible cohesive ends) allowed for subcloning a 17.1 kb *mxnA-mxnI'* fragment revealing pHSU-*mxn*27, whereas a 31.6 kb *mxnJ''-mxnM* fragment was acquired via *EcoRV* hydrolysis and re-ligation yielding pHSU-*mxn*26. In order to access the missing genetic information from *mxnI-mxnJ* another suicide plasmid, pHSU-*mxn*32, was constructed. It harbors a 1.4 kb 3' end fragment of *mxnJ* including a *SpeI* R-site after the stop codon for later gene cluster reconstitution approaches and a *PvuI* R-site upstream of the homology region for the planned plasmid recovery

approach (Fig. 3B). In addition two homology regions, H<sub>L</sub> and H<sub>R</sub> (each 40 nt in size and flanked with an *AvrII* R-site), were engineered during construction of pHSU-mxn32 to enable later Red/ET recombineering for *mxnA-mxnJ* stitching (Fig. 3C). Transformation of *M. fulvus* Mx f50 yielded the mutant strain *M. fulvus* Mx f50::pHSU-mxn32, which was verified by PCR analysis. Hydrolysis with *PvuI* and re-ligation yielded pHSU-mxn35 harboring a 20.3 kb *mxnI''-mxnJ* fragment. Subsequent exchange of the *zeo<sup>R</sup>* selection marker with the ampicillin resistance gene (*amp<sup>R</sup>*) via the unique *AflIII* and *PacI* R-sites revealed pHSU-mxn40 (Fig. 3B). All constructs obtained from the plasmid recovery approaches were verified by restriction analysis.

#### 2.4.2 Construction of a *mxn* gene cluster expression vector

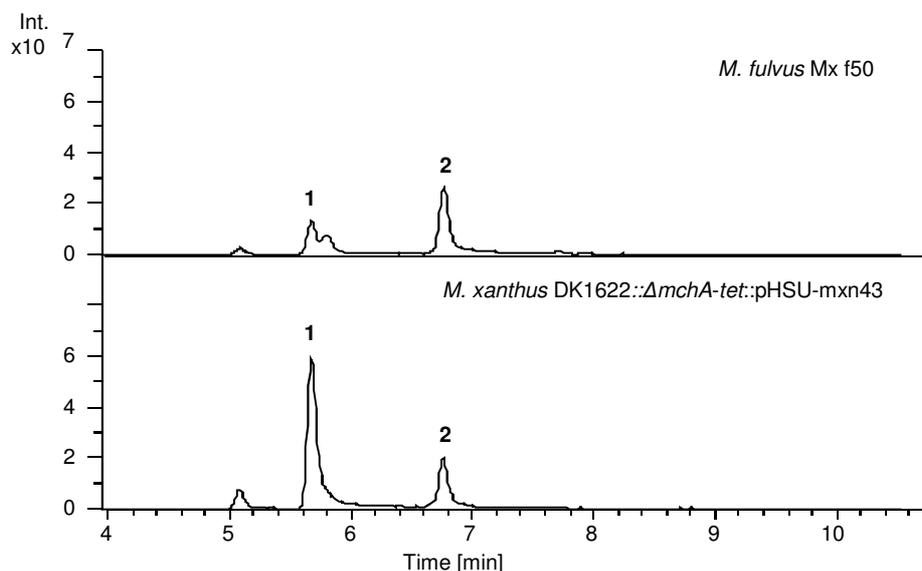
Based on the three plasmids harbouring fragments of the *mxn* gene cluster (pHSU-mxn26, pHSU-mxn27 and pHSU-mxn40) we aimed to reconstitute the entire biosynthetic pathway on one expression construct. The overall cloning strategy involves a combination of Red/ET recombineering<sup>30</sup> and conventional restriction and ligation methods (Fig. 3C). In order to join the *mxnA-mxnJ* gene cluster region, the 18.5 kb *mxnI''-mxnJ* fragment from pHSU-mxn40 was prepared via *AvrII* hydrolysis and subsequent gel-purification. The Red/ET proficient strain *E. coli* GB05-red harbouring plasmid pHSU-mxn27 was transformed with the linear gene cluster fragment, which is flanked by 40 nt homology arms for recombination that were already engineered during construction of the suicide plasmid pHSU-mxn32. Selection on ampicillin and tetracycline revealed clones propagating the expected recombination product pHSU-mxn41. In parallel, pHSU-mxn26 harbouring the downstream part of the gene cluster was modified by Red/ET recombineering to introduce a *SpeI* site at the 5' end of *mxnK*. The Red/ET proficient strain *E. coli* HS996/pSC101-BAD-gbaA-tet harboring plasmid pHSU-mxn26 was transformed with a linear fragment containing the chloramphenicol resistance gene (*cm<sup>R</sup>*), which is flanked by 40 nt homology arms introduced via PCR. Clones growing on chloramphenicol and kanamycin were shown to harbor the expected recombination product pHSU-mxn31. Insertion of the 21 kb *mxnKLM* fragment from pHSU-mxn31 into pHSU-mxn41 can be accomplished via different strategies: ligation of the *AflIII-cm<sup>R</sup>-SpeI-mxnKLM-PacI* fragment including the chloramphenicol resistance cassette, which can be subsequently released via *SpeI* hydrolysis and re-ligation, or direct ligation of the *mxnKLM* target fragment via *SpeI/PacI* without the additional *cm<sup>R</sup>* selection option (Fig. 3C). We first pursued the latter, one-step approach, and were indeed able to reconstitute the entire *mxn* gene cluster on a low copy vector (*p15Aori*). The resulting plasmid, pHSU-mxn42, was further modified by

insertion of a *kan<sup>R</sup>* gene via a unique *PacI* R-site. The resistance cassette was obtained from pHSU-mxn37, a pJET1.2 derivative harbouring the amplified *kan<sup>R</sup>* gene from the pCRII-TOPO vector. The generated expression construct pHSU-mxn43 features the entire 53 kb *mxn* gene cluster under control of the constitutive *PnptII* promoter, a suitable selection marker (*kan<sup>R</sup>*) as well as a homology fragment (*tet<sup>R</sup>*) for directed chromosomal integration into the heterologous host *M. xanthus* DK1622  $\Delta$ *mchA-tet*. The unique *SpeI* site engineered at the *mxnJ/mxnK* interface is not only employed for *mxn* gene cluster stitching, but also represents an important feature for the construction of a *mxn/cor* hybrid gene cluster (see 2.4.5). In addition to restriction analysis, the expression plasmid pHSU-mxn43 was verified via Illumina sequencing to exclude potential mutations that might have occurred during the cloning procedure, especially within the *PnptII* promoter and gene cluster region.

### 2.4.3 Heterologous myxopyronin production in *M. xanthus*

The generated *mxn* gene cluster expression construct pHSU-mxn43 (57 kb) was transferred to *M. xanthus* DK1622  $\Delta$ *mchA-tet* by electroporation. Clones growing on selection plates amended with kanamycin were analyzed for correct chromosomal integration of the expression plasmid by PCR (Fig. S1). After genotypic verification, the *M. xanthus* DK1622  $\Delta$ *mchA-tet*::pHSU-mxn43 mutant and *M. xanthus* DK1622  $\Delta$ *mchA-tet* were cultivated under standard conditions in CTT medium at 30 °C for 6 days. The cultures were harvested by centrifugation to analyze the cell pellet and supernatant metabolite profiles separately for the presence of myxopyronins (MXN). As expected, LC-MS analysis revealed no MXN production in case of *M. xanthus* DK1622  $\Delta$ *mchA-tet* (data not shown). In extracts from the mutant strain *M. xanthus* DK1622  $\Delta$ *mchA-tet*::pHSU-mxn43 production of MXN could be indeed confirmed, proving the successful and functional expression of the heterologous *mxn* gene cluster (Fig. S4). The majority of MXN was detected in the supernatant fraction indicating that the product is efficiently secreted into the medium (Table S6). The same was observed for the native producer *M. fulvus* Mx f50<sup>4</sup>, which was re-cultivated and analyzed in our study. In both cases, the heterologous and native producer, comparable MXN production yields of around 3-10 mg/L were achieved under the applied cultivation conditions. In the heterologous producer MXN A was the dominating myxopyronin derivative, whereas MXN B was produced in significantly lower amounts (Fig. 4). The latter originates from incorporation of a propionyl-CoA instead of an acetyl-CoA starter unit in western chain biosynthesis<sup>22</sup>. In case of *M. fulvus* Mx f50 slightly more MXN B than MXN A was detected in the performed

cultivation experiment, although in previous studies MXN A was found to be the major derivative<sup>4,22</sup>.

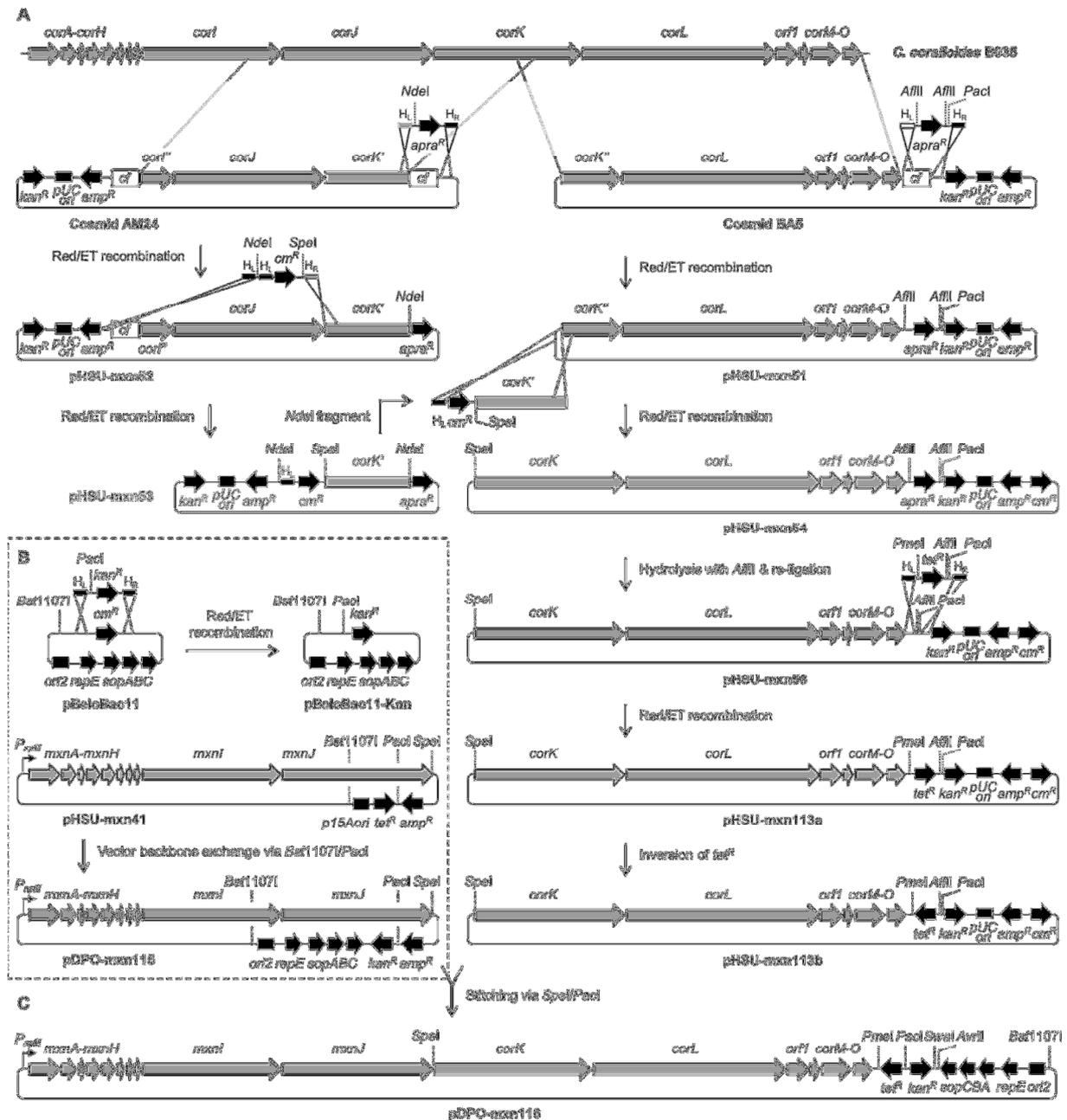


**Figure 4.** Comparative myxopyronin production analysis in the heterologous and native producer. HPLC-MS analysis of the supernatant showing extracted ion chromatograms (EIC)  $[M + H]^+ = 418.222$  (1);  $[M + H]^+ = 432.238$  (2).

#### 2.4.4 Reconstitution of the *corK-O* gene cluster part from cosmids

After establishment of heterologous MXN production we aimed to develop the expression system for biosynthesis of the related coralopyronin (COR). The structural differences in the western chain are caused by the polyketide synthases CorK and CorL as well as the enoyl CoA hydratase CorN and the cytochrome P450 CorO, homologues of which do not exist in the *mxn* biosynthetic pathway (Fig. 2). In order to reconstitute the *corK-corO* gene set we started from two cosmids that were previously used to clone the downstream part of the *cor* gene cluster sequence<sup>31</sup>. The overall strategy is illustrated in Fig. 5 and first involved the insertion of an apramycin resistance gene (*apra*<sup>R</sup>) at the 3' ends of the *cor* gene cluster fragments from cosmids AM24 and BA5 to introduce R-sites for subsequent cloning approaches and to delete part of the chimeric cosmid inserts. The Red/ET-proficient strains *E. coli* GB05-red/cosmid AM24 and *E. coli* GB05-red/cosmid BA5 were transformed with *apra*<sup>R</sup> fragments flanked by suitable R-sites and homology arms (40 nt) generated via PCR. Selection on apramycin revealed clones harboring the expected recombination products, pHSU-mxn52 and pHSU-mxn51, respectively. A second Red/ET recombineering step using *E. coli* GB05-red/pHSU-mxn52 was performed to insert a *cm*<sup>R</sup> gene upstream of *corK* thereby eliminating a large part of the cosmid insert including *corI''-corJ*. In addition to the homology regions for the immediate recombination event, the resistance cassette was flanked by an

additional homology arm plus *NdeI* R-site required for the subsequent Red/ET recombination step as well as a *SpeI* R-site, which was introduced at the 5' end of *corK*. Selection using chloramphenicol revealed the recombinant product pHSU-mxn53, which was next hydrolyzed with *NdeI*. The obtained 9.5 kb *cm<sup>R</sup>-SpeI-corK'* linear fragment is flanked by a 40 nt homologous region to the T7 end of the cosmid vector backbone and shares 577 bp overlap with the *corK* region from pHSU-mxn51. Red/ET recombineering with *E. coli* GB05-red/pHSU-mxn51 and subsequent selection on chloramphenicol revealed construct pHSU-mxn54 harboring the entire *corK-corO* gene cluster region. As initial attempts to stitch the *cor* gene cluster fragment with the *mxn* part from pHSU-mxn41 failed for unknown reasons, alternative cloning strategies were evaluated including various vector backbone engineering steps. One of the generated constructs, pHSU-mxn113b, was finally used for construction of the *mxn/cor* hybrid gene cluster (Fig. 5). It was obtained from pHSU-mxn54 via three cloning steps, even though a more straightforward engineering route would have been possible retrospectively. In the first step, the *apra<sup>R</sup>* gene was eliminated by *AflIII* hydrolysis and re-ligation yielding pHSU-mxn56. Next, Red/ET recombination using *E. coli* GB05-red/pHSU-mxn56 and a linear *tet<sup>R</sup>* fragment flanked by suitable homology arms and R-sites was performed. Selection on tetracycline and kanamycin revealed the recombinant product pHSU-mxn113a, which was modified by ligation of a *PmeI-tet<sup>R</sup>-AflIII-PacI* fragment with reverse *tet<sup>R</sup>* orientation. The opposite arrangement of *tet<sup>R</sup>* is preferred as it allows for chromosomal integration of the final expression construct in the same direction as achieved for the *mxn* expression construct pHSU-mxn43 (Fig. S1). The obtained construct, pHSU-mxn113b, was verified via Illumina sequencing before pursuing the final cluster stitching step.



**Figure 5.** Generation of an expression construct for a myxopyronin/corallopyronin (*mxn/cor*) hybrid biosynthetic gene cluster. **A:** Reconstitution and engineering of a *corK-corO* gene cluster fragment based on a cosmid library from the native producer *C. coralloides* B035 (cosmid inserts are chimeric as indicated by ‘*cf* boxes’). **B:** Engineering of a BAC derivative harboring the *mxnA-mxnJ* gene cluster fragment. **C:** Reconstitution of the *mxn/cor* hybrid gene cluster on a BAC-based expression vector.

### 2.4.5 Construction of a *mxn/cor* hybrid gene cluster expression vector

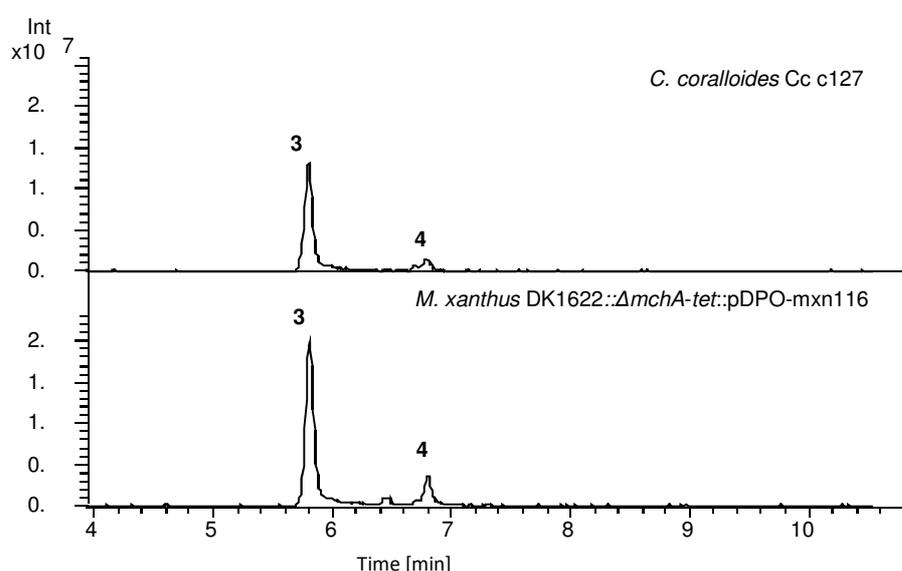
As previous attempts to reconstitute the 65 kb *mxn/cor* hybrid pathway by stitching the *mxnA-mxnJ* with the *corK-corO* gene cluster part on a *p15Aori* based vector backbone failed, we aimed to perform the final ligation step based on a single-copy bacterial artificial chromosome (BAC) vector. In order to achieve the intended vector backbone exchange of plasmid pHSU-mxn41, which harbors the upstream part of the *mxn* gene cluster (*mxnA-mxnJ*),

the commercial pBeloBAC11 vector was modified via Red/ET recombineering (Fig. 5B). For this, the Red/ET proficient strain *E. coli* GB05-red/pBeloBAC11 was transformed with a *kan<sup>R</sup>* gene fragment flanked by unique R-sites and suitable 45 nt homology arms to replace the pBeloBAC *cm<sup>R</sup>* gene. Subsequent selection on kanamycin revealed pBeloBAC-Kan as recombinant product. Making use of unique *Bst*1107I and *Pac*I R-sites the *p15Aori-tet<sup>R</sup>* fragment from pHSU-mxn41 was replaced with the linearized pBeloBAC11-Kan vector yielding the BAC derivative pDPO-mxn115, which harbors the upstream part of the *mxn* gene cluster (*mxnA-mxnJ*) that is homologous to *corA-corJ*. In analogy to the final *mxn* gene cluster stitching step (Fig. 3) the downstream part of the *cor* pathway was ligated via *Spe*I and *Pac*I R-sites to replace the *amp<sup>R</sup>* gene in pDPO-mxn115. The respective 34.8 kb *corK-corO* fragment including a *tet<sup>R</sup>* gene was obtained from pHSU-mxn113b after *Spe*I/*Pac*I hydrolysis. The resulting expression construct pDPO-mxn116 harbors the entire 65 kb *mxn/cor* hybrid pathway (*mxnA-mxnJ*, *corK-corO*) under control of the *PnptII* promoter and features the required genetic elements for selection and directed chromosomal integration into the expression host (*kan<sup>R</sup>* and *tet<sup>R</sup>*; Fig. 5C). The orientation of the *tet<sup>R</sup>* gene allows for implementation of the hybrid gene cluster into the former *mchA* locus in the same direction as achieved for the pure *mxn* pathway based on expression construct pHSU-mxn43. To exclude potential mutations from the cloning procedure, the expression construct pDPO-mxn116 was verified via Illumina sequencing before transfer into the host.

#### 2.4.6 Heterologous corallopyronin production in *M. xanthus*

The generated *mxn/cor* hybrid gene cluster expression construct pDPO-mxn116 (74 kb) was transferred to *M. xanthus* DK1622  $\Delta$ *mchA-tet* by electroporation. Only very few clones growing on selection plates amended with kanamycin were obtained, probably due to the large construct size. After PCR verification of chromosomal integration via the *tet<sup>R</sup>* gene (Fig. S1), the *M. xanthus* DK1622  $\Delta$ *mchA-tet*::pDPO-mxn116 mutant and *M. xanthus* DK1622  $\Delta$ *mchA-tet* were cultivated under standard conditions in CTT medium at 30 °C for 6 days. The cultures were harvested and extracted in the same way as described for MXN production analysis. As expected, LC-MS analysis revealed no COR production in the case of *M. xanthus* DK1622  $\Delta$ *mchA-tet* (data not shown). With *M. xanthus* DK1622  $\Delta$ *mchA-tet*::pDPO-mxn116 production of COR was indeed confirmed (Fig. S5). The yield was similar to the ones (1-1.5 mg/L COR) obtained with the native producer *C. coralloides* Cc c127<sup>5</sup>, which was re-cultivated in parallel. In both cases, COR A represents the main derivative besides low amounts of COR B (Fig. 6). The majority of COR was detected in the supernatant fraction

(Table S7) indicating that the product is efficiently secreted into the medium as also observed for MXN (*cf.* 2.4.3). The heterologous COR production achieved with *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pDPO-mxn116 demonstrates the successful interplay of the expressed MXN/COR hybrid enzymatic machinery. Whereas the COR eastern chain is exclusively synthesized by MXN proteins (MxnI-J interacting with MxnA), western chain biosynthesis involves several hybrid interactions of MXN proteins with the CorK/CorL assembly line. These include extender unit loading/reduction by MxnA, two  $\beta$ -branching events mediated by MxnC-G and the final interaction with the  $\alpha$ -pyrone forming enzyme MxnB, which is able to perform the condensation reaction using the structurally more complex COR western chain.



**Figure 6.** Comparative corallopyronin production analysis in the heterologous and native producer. HPLC-MS analysis of the supernatant showing extracted ion chromatograms (EIC)  $[M + Na]^+ = 550.278$ ,  $[M - H_2O + H]^+ = 510.285$  (3);  $[M + Na]^+ = 564.293$ ,  $[M - H_2O + H]^+ = 524.301$  (4)

#### 2.4.7 Improvement of myxopyronin/corallopyronin production yields

In parallel to our efforts towards establishing a COR pathway expression system, we aimed at optimizing cultivation conditions for heterologous MXN production. One important parameter, when it comes to production yield improvement, is the composition of the production medium. In addition to CTT, our standard medium for *M. xanthus* DK1622 experiments, the heterologous MXN producer was cultivated in M7s medium, which led to promising results in a previous heterologous expression study (Wenzel and Müller, unpublished). As MXN production in M7s was more than two-fold higher compared to CTT (Table S2), M7s medium was used as basis for further optimization approaches (see SI). Numerous variations with different carbon sources were analyzed after scaling down the

cultivation setup to 7 mL allowing for a higher throughput (Table S3-S5). The best yields were achieved with supplements of potassium acetate or different oils (soy oil, vegetable oil, methyl oleate). The latter was already observed for epothilone production<sup>36</sup>, but it might hamper downstream purification processes. Based on the obtained results M7/s6 medium was developed, which additionally contains 10 g/L potassium acetate and 5 g/L corn starch instead of the more expensive soluble starch. Compared to CTT medium the nutrition content in the complex M7/s6 medium is much higher, which allows for reaching increased biomass levels. In order to evaluate the optimized production medium in our standard 50 mL scale shake flask cultivation scenario, the heterologous MXN and COR producer (which was in the meantime available) were grown in M7/s6, CTT and the production medium of their native producer strains. Analysis of  $\alpha$ -pyrone antibiotic production revealed for both expression systems significant higher yields in M7/s6 compared to the two other media, 156 mg/L MXN A and 37 mg/L COR A (Table 2). One obvious reason for the lower COR A production yield compared to MXN A is inefficient COR western chain hydroxylation resulting in the accumulation of significant amounts of preCOR A (Figure S2; the exact yield was not quantified due to the lack of pure reference material). This bottleneck is probably caused by low expression levels or low activity of the cytochrome P450 enzyme CorO, which is encoded as last gene of the 65 kb *mxn/cor* operon and requires a so far unknown reduction partner that has to be provided by the expression strain. However, considering the by-product preCOR A, the overall COR production yield seems to be lower, compared to the MXN expression system, which might be related to the nature of the hybrid pathway and/or to COR biosynthesis in general. Probably, the genetic design of the *mxnJ-corK* hybrid interface has a negative impact on CorK expression/function as the translational initiation region is different (3' end of *mxnJ* instead of *corJ*) and a *SpeI* R-site was engineered at the 5' end of *corK* (which seems to be well tolerated in case of MxnK; Figure S6). It is also possible that COR western chain biosynthesis, performed by the two-subunit CorK-L assembly line, is in general less efficient compared to MxnK and/or that the interplay between the chimeric MXN/COR biosynthesis machinery (*cp.* 2.4.6) is not as proficient as in the native systems. However, compared to previously reported heterologous expression studies using *M. xanthus* as host (Table S8), very promising production yields for MXN and COR were already achieved with the developed M7/s6 medium, which might enable yield improvement in other cases as well.

## 2.5 Conclusions

In this study, *mxn* and hybrid *mxn/cor* biosynthetic gene clusters were successfully constructed and expressed in a derivative of *M. xanthus* DK1622. Optimization of the cultivation medium yielded high production titers of MXN A (156 mg/L) and COR A (37 mg/L plus significant amounts of preCOR A) in shake flask experiments. The applied cultivation conditions might be useful for other heterologous expression studies in *M. xanthus* as well, which represents the most promising host for myxobacterial biosynthetic pathways so far. The observed hydroxylation bottleneck in the COR expression system might be addressed by overexpression of the corresponding P450 enzyme CorO to efficiently direct production towards COR A. Alternatively, *corO* can be deleted from the gene cluster to produce preCOR A as major product, which was also shown to represent a promising compound for further development. This derivative is hardly accessible from the native producer since it is a minor product and genetic manipulation of *C. coralloides* B035 has yet to be established. Overall, the developed heterologous expression systems set the stage for future bioengineering approaches to further exploit this promising class of myxobacterial  $\alpha$ -pyrone antibiotics, e.g. including the generation of novel analogs with improved pharmaceutical properties.

## Acknowledgements

This work was supported by DZIF. We would like to thank Dr. Nestor Zaburanyi for Illumina sequencing analysis, Dr. Rolf Jansen for reference material and our collaborators at ACIES Bio for media improvement.

## Appendix A. Supplementary Information

Supplementary data associated with this article can be found in the online version at: <http://dx.doi.org/10.1016/j.ymben.2017.10.004>

## 2.6 References

- (1) Weissman, K. J.; Müller, R. Myxobacterial secondary metabolites: bioactivities and modes-of-action, *Nat. Prod. Rep.* **2010**, *27*, pp. 1276–1295.
- (2) Schäberle, T. F.; Lohr, F.; Schmitz, A.; König, G. M. Antibiotics from myxobacteria, *Nat. Prod. Rep.* **2014**, *epub ahead of print*.
- (3) Herrmann, J.; Fayad, A. A.; Müller, R. Natural products from myxobacteria: novel metabolites and bioactivities, *Nat. Prod. Rep.* **2017**, *34*, pp. 135–160.
- (4) Irschik, H.; Gerth, K.; Höfle, G.; Kohl, W.; Reichenbach, H. The myxopyronins, new inhibitors of bacterial RNA synthesis from *Myxococcus fulvus* (Myxobacterales), *J. Antibiot.* **1983**, *36*, pp. 1651–1658.
- (5) Irschik, H.; Jansen, R.; Höfle, G.; Gerth, K.; Reichenbach, H. The coralopyronins, new inhibitors of bacterial RNA synthesis from Myxobacteria, *J. Antibiot.* **1985**, *38*, pp. 145–152.
- (6) Ma, C.; Yang, X.; Lewis, P. J. Bacterial Transcription as a Target for Antibacterial Drug Development, *Microbiol. Mol. Biol. Rev.* **2016**, *80*, pp. 139–160.
- (7) Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H. The RNA polymerase "switch region" is a target of inhibitors, *Cell.* **2008**, *135*, pp. 295–307.
- (8) Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G. Transcription inactivation through local refolding of the RNA polymerase structure, *Nature.* **2008**.
- (9) Goldstein, B. P. Resistance to rifampicin: a review, *J. Antibiot.* **2014**, *67*, pp. 625–630.
- (10) Srivastava, A.; Talaue, M.; Liu, S.; Degen, D.; Ebright, R. Y.; Sineva, E.; Chakraborty, A.; Druzhinin, S. Y.; Chatterjee, S.; Mukhopadhyay, J.; Ebright, Y. W.; Zozula, A.; Shen, J.; Sen Gupta, S.; Niedfeldt, R. R.; Xin, C.; Kaneko, T.; Irschik, H.; Jansen, R.; Donadio, S.; Connell, N.; Ebright, R. H. New target for inhibition of bacterial RNA polymerase: 'switch region', *Curr. Opin. Microbiol.* **2011**, *14*, pp. 532–543.
- (11) Ebright, R. H.; Ebright, Y. W.; Shen, J.; Bacci, J.; Hiebel, A.-C.; Solvibile, W.; Self, C.; Olson, G. (2016) Antibacterial Agents: Aryl Myxopyronin Derivatives.
- (12) Moy, T. I.; Daniel, A.; Hardy, C.; Jackson, A.; Rehrauer, O.; Hwang, Y. S.; Zou, D.; Nguyen, K.; Silverman, J. A.; Li, Q.; Murphy, C. Evaluating the activity of the RNA polymerase inhibitor myxopyronin B against *Staphylococcus aureus*, *FEMS microbiology letters.* **2011**, *319*, pp. 176–179.
- (13) Srivastava, A.; Degen, D.; Ebright, Y. W.; Ebright, R. H. Frequency, spectrum, and nonzero fitness costs of resistance to myxopyronin in *Staphylococcus aureus*, *Antimicrob. Agents Chemother.* **2012**, *56*, pp. 6250–6255.
- (14) Schiefer, A.; Schmitz, A.; Schäberle, T. F.; Specht, S.; Lämmer, C.; Johnston, K. L.; Vassilyev, D. G.; König, G. M.; Hoerauf, A.; Pfarr, K. Coralopyronin A Specifically Targets and Depletes Essential Obligate Wolbachia Endobacteria From Filarial Nematodes In Vivo, *J. Infect. Dis.* **2012**, *206*, pp. 249–257.
- (15) Schäberle, T. F.; Schiefer, A.; Schmitz, A.; König, G. M.; Hoerauf, A.; Pfarr, K. Coralopyronin A - a promising antibiotic for treatment of filariasis, *Int. J. Med. Microbiol.* **2014**, *304*, pp. 72–78.

- (16) Rentsch, A.; Kalesse, M. The total synthesis of corallopyronin A and myxopyronin B, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, pp. 11381–11384.
- (17) Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases, *Nat. Prod. Rep.* **2016**, *33*, pp. 231–316.
- (18) Sucipto, H.; Sahner, J. H.; Prusov, E.; Wenzel, S. C.; Hartmann, R.; Koehnke, J.; Müller, R. *In Vitro* reconstitution of  $\alpha$ -Pyrone ring formation in myxopyronin biosynthesis, *Chem. Sci.* **2015**, *6*, pp. 5076–5085.
- (19) Zocher, G.; Vilstrup, J.; Heine, D.; Hallab, A.; Goralski, E.; Hertweck, C.; Stahl, M.; Schäberle, T. F.; Stehle, T. Structural basis of head to head polyketide fusion by CorB, *Chem. Sci.* **2015**, *6*, pp. 6525–6536.
- (20) Finking, R.; Marahiel, M. A. Biosynthesis of nonribosomal peptides, *Annu. Rev. Microbiol.* **2004**, *58*, pp. 453–488.
- (21) Schäberle, T. F.; Mohseni, M. M.; Lohr, F.; Schmitz, A.; König, G. M. Function of the loading module in CorI and of the O-methyltransferase CorH in vinyl carbamate biosynthesis of the antibiotic corallopyronin A, *Antimicrob. Agents Chemother.* **2014**, *58*, pp. 950–956.
- (22) Sucipto, H.; Wenzel, S. C.; Müller, R. Exploring chemical diversity of  $\alpha$ -pyrone antibiotics: molecular basis of myxopyronin biosynthesis, *ChemBioChem.* **2013**, *14*, pp. 1581–1589.
- (23) Calderone, C. T. Isoprenoid-like alkylations in polyketide biosynthesis, *Nat. Prod. Rep.* **2008**, *25*, pp. 845–853.
- (24) Schäberle, T. F.; Schmitz, A.; Zocher, G.; Schiefer, A.; Kehraus, S.; Neu, E.; Roth, M.; Vassilyev, D. G.; Stehle, T.; Bierbaum, G.; Hoerauf, A.; Pfarr, K.; König, G. M. Insights into Structure-Activity Relationships of Bacterial RNA Polymerase Inhibiting Corallopyronin Derivatives, *J. Nat. Prod.* **2015**, *78*, pp. 2505–2509.
- (25) Sahner, J. H.; Sucipto, H.; Wenzel, S. C.; Groh, M.; Hartmann, R. W.; Müller, R. Advanced mutasynthesis studies on the natural  $\alpha$ -pyrone antibiotic myxopyronin from *Myxococcus fulvus*, *ChemBioChem.* **2015**, *16*, pp. 946–953.
- (26) Kaiser, D. Social gliding is correlated with the presence of pili in *Myxococcus xanthus*, *Proc. Natl. Acad. Sci. USA.* **1979**, *76*, pp. 5952–5956.
- (27) Wenzel, S. C.; Müller, R. Host Organisms: Myxobacterium. In *Industrial biotechnology, Microorganisms Volume 3a and 3b*; Wittmann, C.; Liao, J., Eds.; Wiley-VCH: Weinheim, Germany, 2017, pp. 453–485.
- (28) Krug, D.; Zurek, G.; Revermann, O.; Vos, M.; Velicer, G. J.; Müller, R. Discovering the Hidden Secondary Metabolome of *Myxococcus xanthus*: a Study of Intraspecific Diversity, *Appl. Environ. Microbiol.* **2008**, *74*, pp. 3058–3068.
- (29) Meiser, P.; Müller, R. Two functionally redundant Sfp-type 4'-phosphopantetheinyl transferases differentially activate biosynthetic pathways in *Myxococcus xanthus*, *ChemBioChem.* **2008**, *9*, pp. 1549–1553.
- (30) Zhang, Y.; Buchholz, F.; Muyrers, J. P.; Stewart, F. A. A new logic for DNA engineering using recombination in *Escherichia coli*, *Nat. Genet.* **1998**, *20*, pp. 123–128.
- (31) Erol, Ö.; Schäberle, T. F.; Schmitz, A.; Rachid, S.; Gurgui, C.; El Omari, M.; Lohr, F.; Kehraus, S.; Piel, J.; Müller, R.; König, G. M. Biosynthesis of the myxobacterial antibiotic corallopyronin A, *ChemBioChem.* **2010**, *11*, pp. 1235–1265.

(32) Sambrook, J.; Russell, D. W. *Molecular cloning: A laboratory manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.

(33) Chang, A. C.; Cohen, S. N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid, *J. Bacteriol.* **1978**, *134*, pp. 1141–1156.

(34) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schonert, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp., *Gene.* **1992**, *116*, pp. 43–49.

(35) Kashefi, K.; Hartzell, P. L. Genetic suppression and phenotypic masking of a *Myxococcus xanthus* *frzF*- defect, *Mol. Microbiol.* **1995**, *15*, pp. 483–494.

(36) Lau, J.; Frykman, S.; Regentin, R.; Ou, S.; Tsuruta, H.; Licari, P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*, *Biotechnol. Bioeng.* **2002**, *78*, pp. 280–288.

## 2.7 Supplementary information

### 2.7.1 Evaluation and optimization of M7s medium for myxopyronin production

A comparative production experiment between M7s and CTT media (Table S2) was performed, to evaluate the potential of M7s medium for myxopyronin production. Frozen glycerol culture of *M. xanthus* DK1622  $\Delta$ *mchA-tet::pHSU-mxn43* was scraped from a cryo vial and mixed with 200  $\mu$ L of CTT medium in an Eppendorf tube. Five drops were then spotted on CTT agar plates supplemented with kanamycin 50  $\mu$ g/mL which were incubated at 30 °C for several days, until most of the agar surface was covered with cells. All of the cells were scraped from the plate and inoculated into the CTT seed medium supplemented with kanamycin 50  $\mu$ g/mL and incubated for 48 h in Erlenmeyer flasks. Grown seed culture was used (10 %) to inoculate 50 mL of M7s production medium supplemented with kanamycin 50  $\mu$ g/mL in Erlenmeyer flasks and production stage culture was incubated for 96 hours. After cultivation, the culture was mixed with the same amount of methanol, shaken vigorously, centrifuged and supernatant was subjected to the HPLC analysis to perform a relative quantification of MXN (MXN A and MXN B) production in the different media. The results showed around two fold increase of production in M7s medium compared to CTT (Table S2). Carbon source evaluation for M7s medium was performed in falcon tubes, using a down scaled volume of 7 mL, to achieve higher throughput. The seed culture was inoculated in an Erlenmeyer flask as already mentioned above and 10 % was again used to inoculate production cultures in falcon tubes. The cultures were extracted and analysed in the same way as already mentioned. The best carbon sources turned out to be potassium acetate and certain oils (Tables S3 – S5). Addition of potassium acetate to the M7s medium and replacement of

soluble starch with a corn starch as a cheaper alternative, led to the development of M7/s6 medium which was used for comparative production with CTT and native producer media (Table 2 and Table S8). Media components were purchased from the following suppliers: Bacto casitone (DB Biosciences), soy flour (Sojaprotein), soluble starch, maltose and fructose (Merck), glucose (Roquette), lactose (Fluka), sucrose (Gram Mol), dextrin and corn starch (Helios),  $MgSO_4 \times 7 H_2O$  and HEPES (Glentham life science), yeast extract,  $CaCl_2 \times 2 H_2O$ , vitamin B12, methyl oleate, K-acetate, vinyl propionate and  $FeCl_3$  (Sigma Aldrich), n-propanol (J.T.Baker), skim milk (Oxoid), soy oil and vegetable oil (Zvezda).

**Table S1.** List of primers used in this study

| Primer name | Sequence (5'→3')<br>(restriction sites in bold, homology arms underlined)                               | Restriction sites  |
|-------------|---|--|
| mxn30       | TGCTTAATTAATATTCATATGGCTAGCGAGCTCCTGGAT<br>GGCCTT   | <i>PacI</i> fused with<br><i>SspI</i> , <i>EcoRV</i> , <i>NheI</i> |
| mxn33       | CAGGTAGATATCTGGACAGCAAGCGAACC   | <i>EcoRV</i>   |
| mxn41       | AATCTGTACCTCCTTATCCTGTCTCTTGATCAGAT   |  |
| mxn42       | CAAGATCTGATCAAGAGACAGGATAAGGAGGTACAGATT<br>ATGACTTTCACCGTCGTT   |  |
| mxn52       | GATACACAGCTGTCCCTCCTGTTTCAGCTAC   | <i>PvuII</i>   |
| mxn53       | ATCTTAATTAAGGACGCGATGGATATGTT   | <i>PacI</i>  |
| mxn60       | TGCCATGAGCAGACAGTTA   |  |
| mxn61       | CGGGTGGTTGATATCGTG  |  |
| mxn64       | GCTGAGACGGTCTTCCTC  |  |
| mxn65       | TGCTTAATTAATATT <b>CATATG</b><br><b>GCTAGCACGCGATTC</b> ACTGTCTCA                                       | <i>PacI</i> fused with<br><i>SspI</i> , <i>EcoRV</i> , <i>NheI</i> |
| mxn94       | TCACGGCGAAGCACATCC  |  |
| mxn95       | AGGAGGCAGGGTCCGAAT  |  |
| mxn122      | <u>CCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGAT</u><br><b>CTTAAGTACCTGTGACGGAAGATCAC</b>                    | <i>AflIII</i>  |
| mxn123      | <u>AAGCCACCGCCTCCCGTGAGGACCCAGGCTTCTCGTGGGC</u><br>ACTAGTGGGCACCAATAACTGCCTT                            | <i>SpeI</i>  |
| mxn124      | CAATACGATCGCTGGACTGGTGAAG   |  |
| mxn125      | TGCTACTTAAG <b>TTATCA</b><br><b>ACTAGTCATGGCTTCGCTCCCGCC</b>  | <i>AflIII</i> , STOP codon,<br><i>SpeI</i>                         |
| mxn126      | GATACACCTAGGTCCCTCCTGTTTCAGCTAC   | <i>AvrII</i>   |
| mxn127      | TCAGGTCCTAGGGGACGCGATGGATATGTT  | <i>AvrII</i>   |
| mxn128      | ACATCGCCTAGGCGCGCCCTGAGCGCCCTGCGGAACAAG<br><u>GAGTGTGATGCGGCGATCGGACATACTTAAGGGATCTGA</u><br>TCAGCACGTG | <i>AvrII</i> , <i>PvuI</i> , <i>AflIII</i>                         |
| mxn129      | TCTATGCCTAGGTGCGCAAACCAACCCTTGGCAGAACATA<br><u>TCCATCGCGTCCCTTAATTAATCAGTCCTGCTCCTCGGC</u>              | <i>AvrII</i>   |
| mxn136      | AACATCGGCCACCTGGAG  |  |
| mxn137      | AGGACCCAGGCTTCTCGT  |  |
| mxn144      | GATACATTAATTAACCGGAATTGCCAGCTG  | <i>PacI</i>  |
| mxn145      | CTGTTAATTAATCAGAAGAAGCTCGTCAAGAAG   | <i>PacI</i>  |
| mxn147      | TATGTCCTTAAGATAATTTCGGCTGCAGGGG   | <i>AflIII</i>  |
| mxn148      | TATCTTAATTAATTACCAATGCTTAATCAGTG  | <i>PacI</i>  |

|                         |   |                     |
|-------------------------|---|---------------------|
| mxn200                  | <u>AGCAGCCGGGTGAGGCTGAGGTCCACCTGGCTCACGACG</u><br><u>ACTTAAGGGTTCATGTGCAGCTCCA</u>  | <i>AflIII</i>       |
| mxn201                  | <u>CATGAGAATTCGCGGCCGCATAATACGACTCACTATAGGG</u><br><u>TTAATTAAGGAGCTCAGCCAATCGACT</u>   | <i>PacI, AflIII</i> |
| mxn202                  | <u>GTCCGCCACCACGACGCGCGCTGGTGCTCACCGGCCGCT</u><br><u>CATATGGTTTAAACGGTTCATGTGCAGCTCCA</u>   | <i>NdeI, PmeI</i>   |
| mxn203                  | <u>TCTTCAAGAATTCGCGGCCGCAATTAACCCTCACTAAAGG</u><br><u>GTTTAAACGAGCTCAGCCAATCGACT</u>  | <i>PmeI</i>         |
| mxn209                  | <u>CATGAGAATTCGCGGCCGCATAATACGACTCACTATAGGG</u><br><u>CATATGGTCTTCAAGAATTCGCGGCCGCAATTAACCCTCA</u><br><u>CTAAAGTACCTGTGACGGAAGATCAC</u> | <i>NdeI</i>         |
| mxn210                  | <u>AGGCCATCCCGATGATGGCCACCCCGGGCGCCTGTTACG</u><br><u>ACTAGTGGGCACCAATAACTGCCTT</u>  | <i>SpeI</i>         |
| mxn330                  | <u>CGAAGCCGGGGACGCCGCGCGGATTTCAGCGAGGCGTCC</u><br><u>CGTTTAAACATTAATTCTCATGTTTGAC</u>   | <i>PmeI</i>         |
| mxn331                  | <u>GCCGCATAATACGACTCACTATAGGGTTAATTAACTTAA</u><br><u>GCCAAGGGTTGGTTTGCGC</u>  | <i>PacI, AflIII</i> |
| Tet-for-AflIII-<br>PacI | <u>GAGCTCTTAATTAAGTGCTTAAGCGTAATTCTCATGTTTG</u><br><u>ACAGC</u>   | <i>PacI, AflIII</i> |
| Tet-rev-PmeI            | <u>GATACAGTTTAAACCAAGGGTTGGTTTGCGCA</u>   | <i>PmeI</i>         |
| Kan-<br>BeloBAC-rev     | <u>TTCAGGCGTAGCAACCAGGCGTTTAAAGGCACCAATAACT</u><br><u>GCCTTCCTAGGACCATTAAATTCAGAAGAACTCGTCAA</u><br><u>GAAG</u>                         | <i>AvrII, SwaI</i>  |
| Kan-<br>BeloBAC-for     | <u>TCCTGTGCGACGGTTACGCCGCTCCATGAGCTTATCGCGA</u><br><u>ATAAATTAATTAATGGACAGCAAGCGAACCGG</u>  | <i>PacI</i>         |
| p15A-Tet1               | <u>GGAGAACTGTGAATGCGC</u>   |                     |
| p15A-Tet2               | <u>ACTCCGCTAGCGCTGATG</u>   |                     |
| pJET1.2For              | <u>CGACTCACTATAGGGAGAGCGGC</u>  |                     |
| pJET1.2Rev              | <u>AAGAACATCGATTTTCCATGGCAG</u>   |                     |
| pTOPO-in                | <u>CCTCTAGATGCATGCTCGAG</u>   |                     |
| pTOPO-out               | <u>TTGGTACCGAGCTCGGATCC</u>   |                     |
| P1                      | <u>CGAGCAATCCGCTATTGGC</u>  |                     |
| P2                      | <u>TCGGTTCAAAGAGTTGGTAGC</u>  |                     |
| P3                      | <u>GAGAACCTGCGTGCAATC</u>   |                     |
| P4                      | <u>CTGTGTCCTTCTGCGACGC</u>  |                     |
| P5                      | <u>GGACATCCCAGTGCCTAC</u>   |                     |
| P6                      | <u>ACCTGCGTGCAATCCATC</u>   |                     |

**Table S2.** Production of myxopyronin A and B in CTT and M7s media after 72 h of cultivation (10 % seed culture inoculum).

| Medium  | Relative MXN A/B<br>yield [%] |
|---|-------------------------------|
| <b>CTT</b>  | 100                           |
| -10 g/L casitone                                      |                               |
| -10 mL/L Tris-HCl (1 M, pH 7.6)                       |                               |
| -1 mL/L KH <sub>2</sub> PO <sub>4</sub> (1 M, pH 7.6) |                               |
| -10 mL/L MgSO <sub>4</sub> (0.8 M)                    |                               |
| <b>M7s</b>  |                               |
| - 5 g/L soya flour                                    |                               |
| - 5 g/L soluble starch                                | 233                           |
| - 2 g/L glucose                                       |                               |
| - 1 g/L yeast extract                                 |                               |
| - 1 g/L MgSO <sub>4</sub> x 7 H <sub>2</sub> O        |                               |
| - 1 g/L CaCl <sub>2</sub> x 2 H <sub>2</sub> O        |                               |
| - 10 g/L HEPES  |                               |
| - 0.1 mg/L vitamin B12                                |                               |
| - 5 mg/L FeCl <sub>3</sub>                            |                               |

**Table S3.** Effect of carbon source addition to M7s medium

| Variations of M7s     | Relative yield [%] |
|-----------------------|--------------------|
| <b>Control</b>        | <b>100</b>         |
| fructose 5 g/L        | 334                |
| fructose 10 g/L       | 157                |
| fructose 20 g/L       | 178                |
| glucose 5 g/L         | 102                |
| glucose 10 g/L        | 160                |
| lactose 5 g/L         | 119                |
| lactose 10 g/L        | 183                |
| lactose 20 g/L        | 110                |
| maltose 5 g/L         | 147                |
| maltose 10 g/L        | 145                |
| maltose 20 g/L        | 128                |
| sucrose 5 g/L         | 148                |
| sucrose 10 g/L        | 171                |
| sucrose 20 g/L        | 141                |
| dextrin 5 g/L         | 107                |
| dextrin 10 g/L        | 160                |
| dextrin 20 g/L        | 122                |
| corn starch 5 g/L     | 147                |
| corn starch 10 g/L    | 124                |
| corn starch 20 g/L    | 114                |
| soluble starch 10 g/L | 136                |
| soluble starch 20 g/L | 203                |
| methyl oleate 1 g/L   | 231                |
| methyl oleate 2 g/L   | 483                |
| methyl oleate 5 g/L   | 364                |
| methyl oleate 10 g/L  | 17                 |
| methyl oleate 20 g/L  | 72                 |

**Table S4.** Effect of carbon source addition to M7s medium

| Variations of M7s        | Relative yield [%] |
|--------------------------|--------------------|
| <b>Control</b>           | <b>100</b>         |
| K-acetate 0.5 g/L        | 114                |
| K-acetate 1 g/L          | 113                |
| K-acetate 5 g/L          | 471                |
| n-propanol 0.5 g/L       | 89                 |
| n-propanol 1 g/L         | 95                 |
| n-propanol 5 g/L         | 109                |
| vinyl propionate 0.5 g/L | 150                |
| vinyl propionate 1 g/L   | 104                |
| vinyl propionate 5 g/L   | 125                |
| soy oil 1 g/L            | 209                |
| soy oil 2 g/L            | 314                |
| soy oil 5 g/L            | 459                |
| soy oil 10 g/L           | 430                |
| vegetable oil 1 g/L      | 191                |
| vegetable oil 2 g/L      | 230                |
| vegetable oil 5 g/L      | 252                |
| vegetable oil 10 g/L     | 425                |

**Table S5.** Retesting of carbon source addition to M7s medium

| Variations of M7s                           | Relative yield [%] |
|---|--------------------|
| <b>Control</b>                              | <b>100</b>         |
| fructose 2 g/L                              | 128                |
| fructose 5 g/L                              | 109                |
| fructose 10 g/L                             | 112                |
| glucose 10 g/L                              | 126                |
| glucose 20 g/L                              | 160                |
| glucose 40 g/L                              | 219                |
| soluble starch 20 g/L                       | 74                 |
| soluble starch 30 g/L                       | 98                 |
| K-acetate 1 g/L                             | 163                |
| K-acetate 5 g/L                             | 444                |
| K-acetate 10 g/L                            | 521                |
| soy oil 2 g/L                               | 400                |
| soy oil 5 g/L                               | 488                |
| soy oil 10 g/L                              | 270                |
| soy oil 20 g/L                              | 284                |
| soy oil 5 g/L + feed after 48 h 5 g/L       | 474                |
| vegetable oil 5 g/L                         | 488                |
| vegetable oil 10 g/L                        | 333                |
| vegetable oil 20 g/L                        | 435                |
| vegetable oil 5 g/L + feed after 48 h 5 g/L | 426                |

**Table S6.** Myxopyronin A concentration in the pellet and supernatant of the native producer (*M. fulvus* Mx f50) and heterologous host (*M. xanthus* DK1622  $\Delta$ mchA-tet::pHSU-mxn43) after cultivation in BTY and CTT medium respectively.

| Strain  | Pellet avg. [mg/L] | Supernatant avg. [mg/L] |
|---|--------------------|-------------------------|
| <i>M. fulvus</i> Mx f50 <sup>a</sup>                                | 0.003 ± 0.001      | 1.800 ± 0.000           |
| <i>M. xanthus</i> DK1622 $\Delta$ mchA-tet::pHSU-mxn43 <sup>b</sup> | 0.023 ± 0.006      | 8.000 ± 0.346           |

<sup>a</sup> BTY medium<sup>b</sup> CTT medium**Table S7.** Corallopyronin A concentration in the pellet and supernatant of the native producer (*C. coralloides* Cc c127) and heterologous host (*M. xanthus* DK1622  $\Delta$ mchA-tet::pDPO-mxn116) after cultivation in MD1/4 and CTT medium respectively.

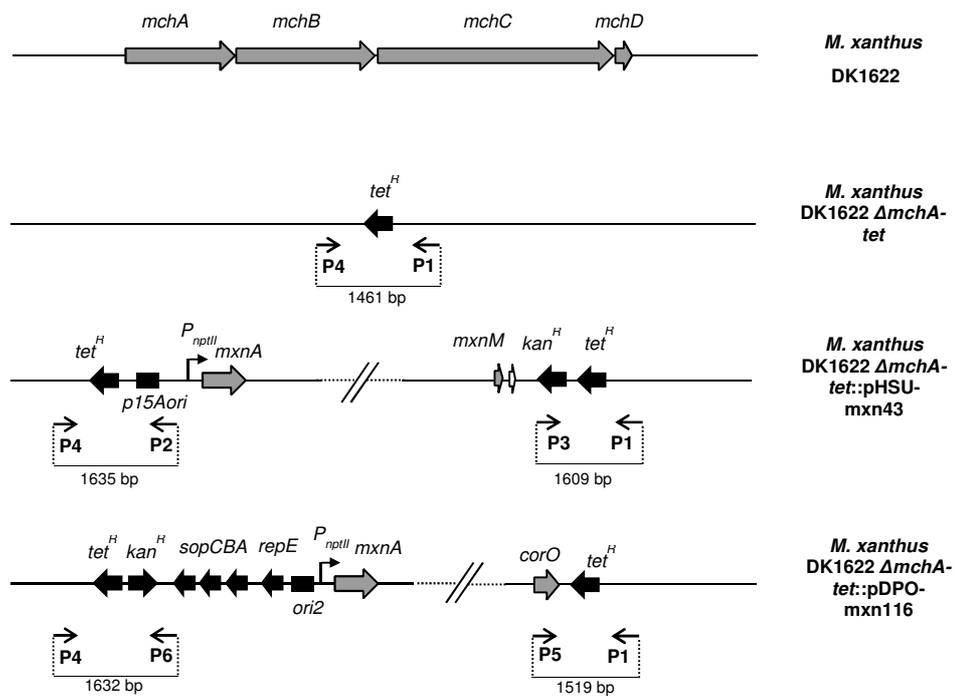
| Strain   | Pellet avg. [mg/L] | Supernatant avg. [mg/L] |
|--|--------------------|-------------------------|
| <i>C. coralloides</i> C cc127 <sup>a</sup>                           | 0.018 ± 0.006      | 0.867 ± 0.115           |
| <i>M. xanthus</i> DK1622 $\Delta$ mchA-tet::pDPO-mxn116 <sup>b</sup> | 0.030 ± 0.006      | 1.467 ± 0.115           |

<sup>a</sup> MD1/4 medium<sup>b</sup> CTT medium**Table S8.** Heterologous production of secondary metabolites in *Myxococcus xanthus* (Table adapted from<sup>1</sup>).

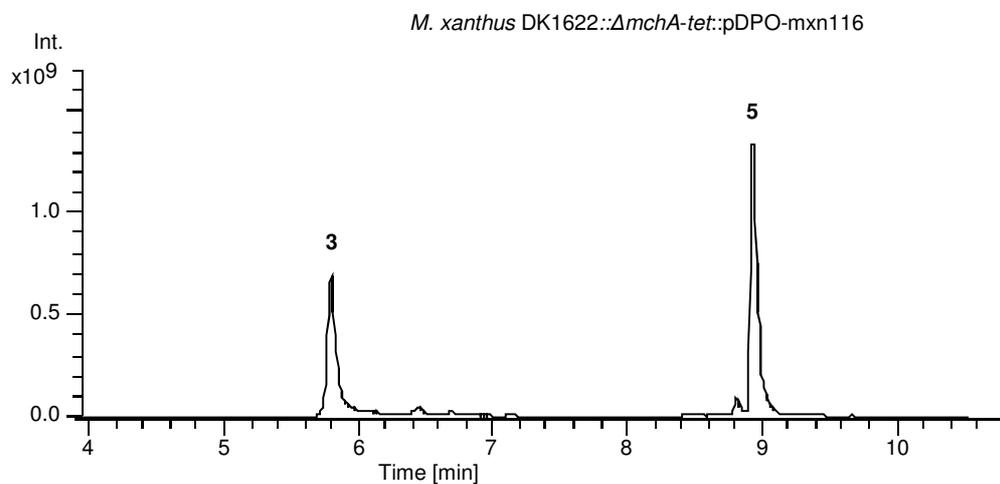
| Compound        | Native producer                   | Medium           | Promoter | Pathway-type [size] | Yield                | Ref          |
|-----------------|-----------------------------------|------------------|----------|---------------------|----------------------|--------------|
| Bengamide       | <i>Myxococcus virescens</i>       | CTT              | PnptII   | PKS/NRPS [25 kb]    | >10 mg/L             | 2            |
| Dawenol         | <i>Stigmatella aurantiaca</i>     | CTT              | native   | PKS [21 kb]         | n.d.                 | 3            |
| Epothilone      | <i>Sorangium cellulosum</i>       | CMM              | native   | PKS/NRPS [56 kb]    | 0.1 - 0.4 mg/L       | 4-6          |
| Epothilone      | <i>Sorangium cellulosum</i>       | CTS              | native   | PKS/NRPS [56 kb]    | 23 mg/L              | 5            |
| Myxochromide S  | <i>Stigmatella aurantiaca</i>     | CTT              | PnptII   | PKS/NRPS [29 kb]    | >1000 mg/L           | 7            |
| Myxothiazol     | <i>Stigmatella aurantiaca</i>     | CTT              | Pm       | PKS/NRPS [57 kb]    | 20 mg/L              | 8            |
| Oxytetracycline | <i>Streptomyces rimosus</i>       | CTTYE            | native   | PKS [32 kb]         | 10 mg/L              | 9            |
| Pretubulysin    | <i>Cystobacter species</i>        | CTT <sup>a</sup> | Ptet     | PKS/NRPS [40 kb]    | 0.2 mg/L             | 10           |
| PUFAs           | <i>Aetherobacter fasciculatus</i> | CTT              | Ptet     | PKS/FAS [18 kb]     | ~1 mg/CDW            | 11           |
| Myxopyronin     | <i>Myxococcus fulvus</i>          | CTT<br>M7/s6     | PnptII   | PKS/NRPS [53 kb]    | 3.8 mg/L<br>156 mg/L | (this study) |
| Corallopyronin  | <i>Coralloccoccus coralloides</i> | CTT<br>M7/s6     | PnptII   | PKS/NRPS [65 kb]    | 1.5 mg/L<br>37 mg/L  | (this study) |

n.d., not determined

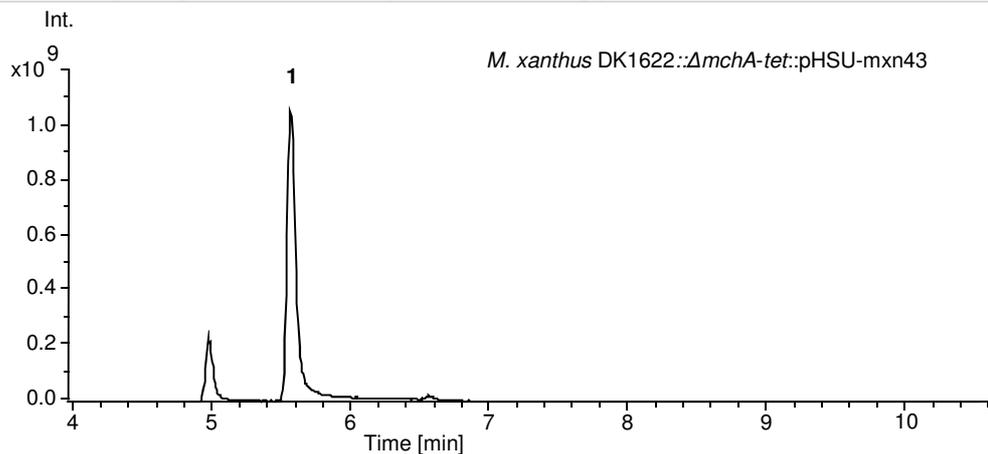
<sup>a</sup> 0.1 mg/L supplemental of racemic D,L-pipecolic acid



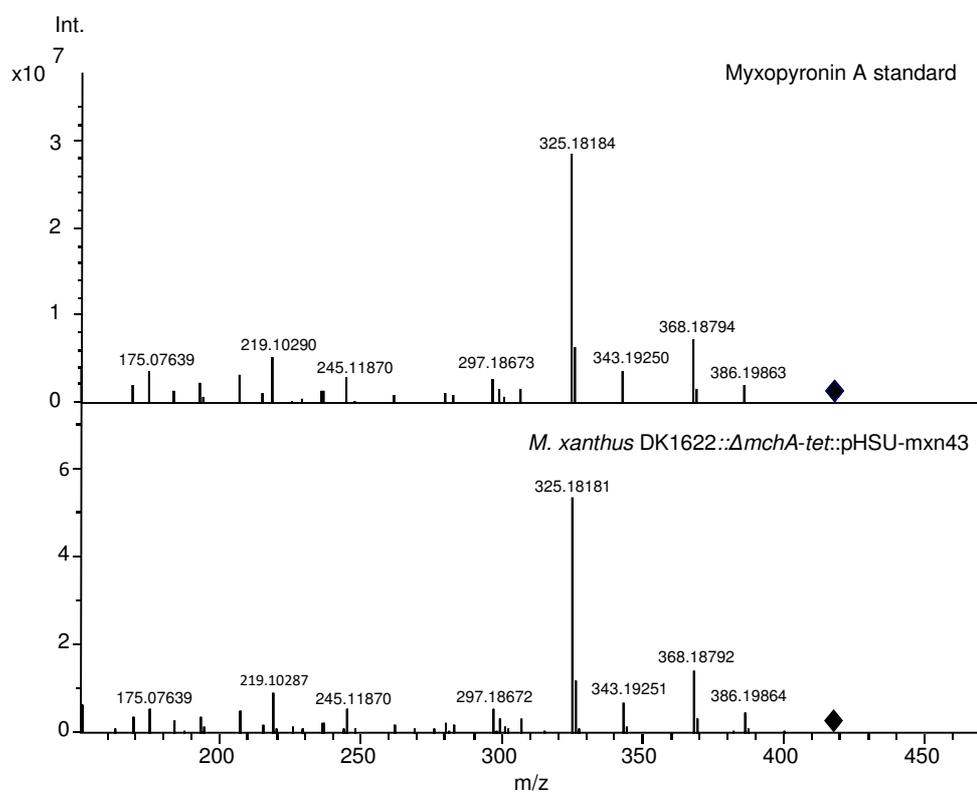
**Figure S1.** Genotypic verification of integration of both clusters into the heterologous host *M. xanthus* DK1622  $\Delta$ mchA-tet. Primer sets used for genotypic verification, their binding sites and amplicon sizes are illustrated. Nucleotide sequences of the primers used are listed in the table S1.



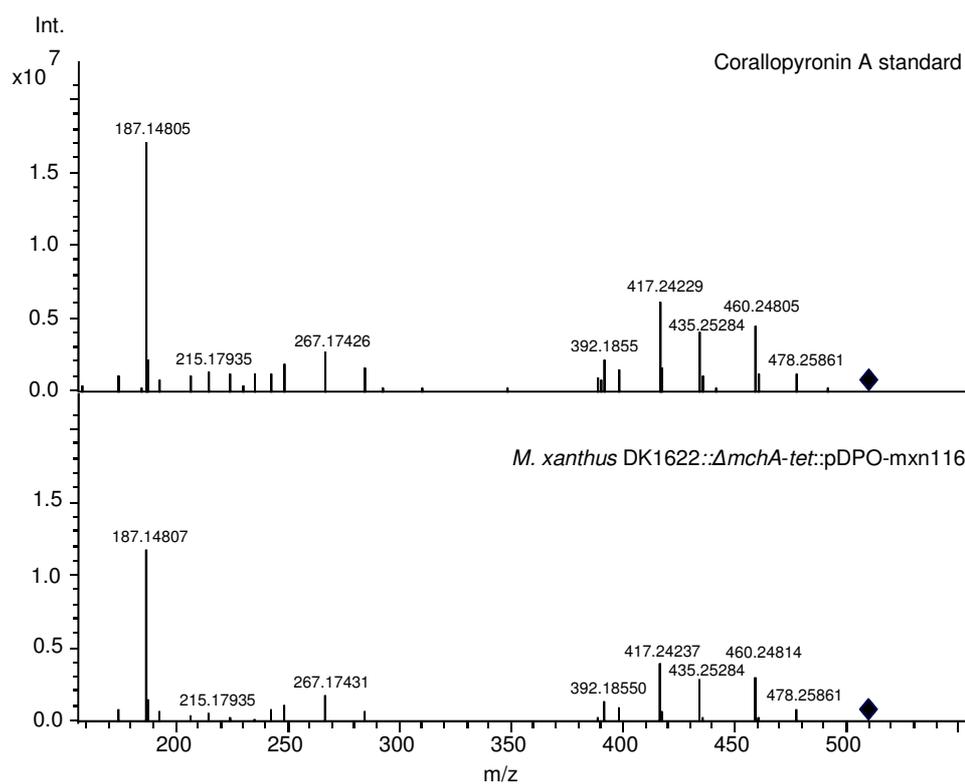
**Figure S2.** Production analysis of *M. xanthus* DK1622  $\Delta$ mchA-tet::pDPO-mxn116 cultivated in the M7/s6 medium, showing significant amount of preCOR A (5) being produced in comparison to the COR A (3). HPLC-MS analysis of supernatant showing extracted ion chromatograms (EIC)  $[M + Na]^+ = 550.278$ ,  $[M - H_2O + H]^+ = 510.285$  (3);  $[M + H]^+ = 512.296$ ,  $[M + Na]^+ = 534.278$ ,  $[M - H_2O + H]^+ = 494.285$  (5).



**Figure S3.** Production analysis of *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pHSU-mxn43 cultivated in the M7/s6 medium, showing significant production of MXN A (1) and no production of MXN B (2). HPLC-MS analysis of supernatant showing extracted ion chromatograms (EIC)  $[M + H]^+ = 418.222$  (1);  $[M + H]^+ = 432.238$  (2).



**Figure S4.** MS/MS spectrum of myxopyronin A ( $m/z = 418.222$ ) from pure myxopyronin A standard and from an extract of *M. xanthus* DK1622:: $\Delta mchA$ -*tet*::pHSU-mxn43, showing identical pattern.



**Figure S5.** MS/MS spectrum of corallopyronin A ( $m/z = 510.285$ ) from pure corallopyronin A standard and from an extract of *M. xanthus* DK1622:: $\Delta$ mchA-tet::pDPO-mxn116, showing identical pattern.



**Figure S6.** Translational alignment of 3' and 5' ends of PKS genes from the native and engineered MXN and COR biosynthetic pathways.

## 2.7.2 References

- (1) Wenzel, S. C.; Müller, R. Host Organisms: Myxobacterium. In *Industrial biotechnology, Microorganisms Volume 3a and 3b*; Wittmann, C.; Liao, J., Eds.; Wiley-VCH: Weinheim, Germany, 2017, pp. 453–485.
- (2) Wenzel, S. C.; Hoffmann, H.; Zhang, J.; Debussche, L.; Haag-Richter, S.; Kurz, M.; Nardi, F.; Lukat, P.; Kochems, I.; Tietgen, H.; Schummer, D.; Nicolas, J.-P.; Calvet, L.; Czepczor, V.; Vrignaud, P.; Mühlenweg, A.; Pelzer, S.; Müller, R.; Brönstrup, M. Production of the bengamide class of marine natural products in myxobacteria: biosynthesis and structure-activity relationships, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, pp. 15560–15564.
- (3) Oßwald, C.; Zaburannyi, N.; Burgard, C.; Hoffmann, T.; Wenzel, S. C.; Müller, R. A highly unusual polyketide synthase directs dawenol polyene biosynthesis in *Stigmatella aurantiaca*, *J. Biotechnol.* **2014**, *191*, pp. 54–63.
- (4) Julien, B.; Shah, S. Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*, *Antimicrob. Agents Chemother.* **2002**, *46*, pp. 2772–2778.
- (5) Lau, J.; Frykman, S.; Regentin, R.; Ou, S.; Tsuruta, H.; Licari, P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*, *Biotechnol. Bioeng.* **2002**, *78*, pp. 280–288.
- (6) Oßwald, C.; Zipf, G.; Schmidt, G.; Maier, J.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Modular Construction of a Functional Artificial Epothilone Polyketide Pathway, *ACS Synth. Biol.* **2014**, *3*, pp. 759–772.
- (7) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (8) Perlova, O.; Fu, J.; Kuhlmann, S.; Krug, D.; Stewart, F.; Zhang, Y.; Müller, R. Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2006**, *72*, pp. 7485–7494.
- (9) Stevens, D. C.; Henry, M. R.; Murphy, K. A.; Boddy, C. N. Heterologous expression of the oxytetracycline biosynthetic pathway in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2010**, *76*, pp. 2681–2683.
- (10) Chai, Y.; Shan, S.; Weissman, K. J.; Hu, S.; Zhang, Y.; Müller, R. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis, *Chem. Biol.* **2012**, *19*, pp. 361–371.
- (11) Gemperlein, K.; Rachid, S.; Garcia, R. O.; Wenzel, S. C.; Müller, R. Polyunsaturated fatty acid biosynthesis in myxobacteria. Different PUFA synthases and their product diversity, *Chem. Sci.* **2014**, *5*, pp. 1733–1741.



## **Chapter 3**

# **Production optimization and biosynthesis revision of corallopyronin A, a potent anti-filarial antibiotic**

Domen Pogorevc, Fabian Panter, Carolina Schillinger, Rolf Jansen, Silke C. Wenzel and  
Rolf Müller

(Submitted manuscript)

## **Contributions to the Presented Work**

### **Author's Contribution**

The author conceived and performed most of the experiments described in this chapter and interpreted the resulting data. The author designed and constructed gene deletion and co-expression constructs. Heterologous expression of these constructs in *Myxococcus xanthus* DK1622 was conducted by the author. Furthermore, the author conducted the cultivation of heterologous mutant strains, prepared extracts and quantified the coralopyronin A and precoralopyronin A production yields. The author conceived and wrote this chapter.

### **Contribution by Co-Workers**

Fabian Panter detected oxycoralopyronin A by principal component analysis and performed purification and NMR elucidation of precoralopyronin A, coralopyronin D and oxycoralopyronin A. Carolina Schillinger contributed to the *corN* and *corM* co-expression as well as generation of the transposon based coralopyronin expression construct. Silke C. Wenzel designed the transposase based coralopyronin expression construct. The project was supervised by Silke C. Wenzel and Rolf Müller who were responsible for the conception of the project and performed the proofreading of this chapter.

### **3 Production optimization and biosynthesis revision of corallopyronin A, a potent anti-filarial antibiotic**

#### **3.1 Abstract**

Corallopyronins (COR) are  $\alpha$ -pyrone antibiotics from myxobacteria representing highly promising lead structures for the development of antibacterial therapeutic agents. Their ability to inhibit RNA polymerase through interaction with the “switch region”, a novel target, distant from binding sites of previously characterized RNA polymerase inhibitors (e.g. rifampicin), makes them particularly promising as antibiotic candidates. Corallopyronin A is currently also investigated as a lead compound for the treatment of lymphatic filariasis because of its superb activity against the nematode symbiont *Wolbachia*. As total synthesis is not a valid production option biotechnological optimisation of compound supply is of utmost importance to further develop this highly potent compound class.

Here we describe decisive improvements of the previously reported heterologous COR production and engineering platform yielding production of ~100 mg/L COR A. Furthermore, we provide a revised model of COR biosynthesis shedding light on the function of several biosynthetic proteins, including an unusual ECH-like enzyme providing dehydration functionality in *trans* and an uncharacterized protein conferring COR self-resistance in the myxobacterial heterologous host *Myxococcus xanthus* DK1622. We also report two new COR derivatives, COR D and oxyCOR A discovered in genetically engineered strains.

### 3.2 Introduction

Filariasis and Onchocerciasis, also known as river blindness, are devastating parasitic diseases mainly present in tropical and subtropical countries. Caused by filarial nematodes endemic in more than 80 countries they infect more than 150 million people worldwide<sup>1</sup>. Current treatment options resort to use of ivermectin and diethylcarbamazine as main drugs of choice<sup>2</sup>. Although safe and effective they primarily target first stage larvae (microfilariae) causing temporary infertility in the female worms. Thus, the drugs have to be administered once to twice yearly to avoid microfilariae relapse<sup>2</sup>. Distribution and compliance issues as well as emerging resistance raise concerns for future treatment options<sup>3</sup>. An emerging new target for treatment against filarial nematodes are *Wolbachia*, Gram-negative bacterial endosymbionts of the nematodes<sup>4,5</sup>. Depletion of *Wolbachia* leads to various effects on the nematodes, including sterility of adult worms<sup>6</sup>. Therefore, new anti-filarial drugs that successfully deplete *Wolbachia* from filarial nematodes are highly desired.

One prominent candidate are the structurally related myxobacterial  $\alpha$ -pyrone antibiotics corallopyronins (COR) and myxopyronins (MXN), discovered in *Coralloccoccus coralloides* Cc c127<sup>7</sup> and *Myxococcus fulvus* Mx f50<sup>8</sup>, respectively. COR A was shown to specifically target and deplete obligate *Wolbachia* endobacteria from filarial nematodes<sup>1,2</sup>, while showing no toxicity against eukaryotic cells<sup>7</sup>. The main target of COR A is the prokaryotic RNA polymerase (RNAP), which is highly conserved across the bacterial domain and a proven target for broad-spectrum antibacterial treatment<sup>9</sup>. Importantly, the RNAP ‘switch region’<sup>10,11</sup>, the binding site of COR, is distant from targets of previously characterized RNAP inhibitors including rifamycins (Rif). Rif are currently used in the clinics for treatment of Gram-negative and Gram-positive bacterial infections and represent first-line anti-tuberculosis agents. Unfortunately, the clinical utility of Rif is already threatened by the rising resistance by RNAP mutations<sup>12</sup>. Utilizing Rif as an anti-wolbachial agent is thus problematic<sup>2</sup>. As COR binds to a different RNAP target site it does not exhibit cross-resistance with Rif making it a promising candidate for the development of novel broad-spectrum antibiotics<sup>13</sup>.

For further exploitation of COR as well as novel derivatives thereof as antifilarial and antibacterial agents, the improvement of compound supply is of utmost importance. Since the established total synthesis was shown to be tedious and cost-inefficient, requiring more than 22 steps with an overall yield of less than 1%<sup>14</sup>, a heterologous production platform in the myxobacterium *M. xanthus* DK1622 was developed as a more viable alternative<sup>15</sup>. By heterologous expression of *mxn* and hybrid *mxn/cor* biosynthetic gene clusters (BGCs) driven by the *P<sub>npiII</sub>* promoter, significant yield improvements of both MXN and COR could be

achieved, reaching 150 mg/L and 40 mg/L, respectively. Significant amounts of precorallopyronin A (preCOR A) were detected in the extracts of the heterologous COR producer, which might account for part of the decreased COR A yield in comparison to MXN production<sup>15</sup>.

In the present study we aimed to improve the COR heterologous production platform<sup>15</sup>, by co-expression of an additional CorO (cytochrome P450) copy in the heterologous host, to facilitate hydroxylation of preCOR A to COR A (see Figure 6). Furthermore, the effect of different promoters as well as integration of the whole biosynthetic pathway in an alternative genomic locus was evaluated. With the deletion of *corO* from the corallopyronin biosynthetic gene cluster (BGC), production of preCOR A could be highly improved which facilitates future drug development, as its antibiotic potential was shown not to lag behind that of COR A<sup>16</sup>. Several additional gene deletions were performed in order to produce novel COR analogues as well as to reveal surprising details about functions of specific proteins encoded in the BGC.

### 3.3 Materials and methods

#### 3.3.1 Sequence analysis of Corallopyronin biosynthetic gene clusters (BGC)

Sequence analysis and routine *in silico* DNA work was performed using Geneious v8-v11 software package (Biomatters Ltd.). In addition to restriction analysis, integrity of the generated expression constructs: pDPO-mxn116-Pvan-Tpase, pDPO-mxn116-Pvan-Tpase $\Delta$ CorN, pDPO-mxn116-Pvan-Tpase $\Delta$ CorO, pDPO-mxn116-Pvan-Tpase $\Delta$ CorN-O, pDPO-mxn116-Pvan-Tpase $\Delta$ CorM and pDPO-mxn116-Pvan-Tpase $\Delta$ CorP was verified by sequencing. Using the Illumina paired-end technology, on a MiSeq platform, 569-fold (pDPO-mxn116-Pvan-Tpase), 651-fold (pDPO-mxn116-Pvan-Tpase $\Delta$ CorN), 755-fold (pDPO-mxn116-Pvan-Tpase $\Delta$ CorO) and 513-fold (pDPO-mxn116-Pvan-Tpase $\Delta$ CorN-O) mean sequencing coverage was achieved and no mutations within the BGC region were detected. Constructs pDPO-mxn116-Pvan-Tpase $\Delta$ CorM and pDPO-mxn116-Pvan-Tpase $\Delta$ CorP were verified by sequence analysis of the modified region and detailed restriction analysis.

#### 3.3.2 Construction and engineering of plasmids

Routine handling of nucleic acids, such as isolation of plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, was performed according to

standard protocols<sup>17</sup>. *E. coli* HS996 (Invitrogen) was used as host for standard cloning experiments. *E. coli* strains were cultured in LB medium or on LB agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) at 30-37 °C (and 200 rpm) overnight. Antibiotics were used at the following final concentrations: 100 µg/mL ampicillin, 50 µg/mL kanamycin, 6 µg/mL oxytetracycline, 50 µg/mL zeocin, 12.5 µg/mL chloramphenicol and 40 µg/mL apramycin. Transformation of *E. coli* strains was achieved via electroporation in 0.1 cm-wide cuvettes at 1250 V, a resistance of 200 Ω, and a capacitance of 25 µF. Plasmid, cosmid and bacterial artificial chromosome (BAC) DNA was either purified by standard alkaline lysis<sup>17</sup> or by using the GeneJet Plasmid Miniprep Kit (Thermo Fisher Scientific) or the NucleoBond PC100 kit (Machery Nagel). Restriction endonucleases, alkaline phosphatase (FastAP) and T4 DNA ligase were purchased from Thermo Fisher Scientific. Oligonucleotides used for PCR and sequencing were obtained from Sigma-Aldrich and are listed in Table S1. PCR reactions were carried out in a Mastercycler® pro (Eppendorf) using Phusion™ High-Fidelity, Taq DNA polymerase or DreamTaq (Thermo Fisher Scientific) according to the manufacturer's protocol. For Taq and DreamTaq: Initial denaturation (3 min, 95 °C); 30 cycles of denaturation (30 s, 95 °C), annealing (30 s, 53 or 57 °C) and elongation (varied based on PCR product length 1 kb/min, 72 °C); and final extension (10 min, 72 °C). For Phusion™: Initial denaturation (30 s, 98 °C); 30 cycles of denaturation (20 s, 98 °C), annealing (25 s, 53 or 57 °C) and elongation (varied based on PCR product length 0.5 kb/min, 72 °C); and final extension (10 min, 72 °C). PCR products or DNA fragments from restriction digests were purified by agarose gel electrophoresis and isolated using the NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) or peqGold Gel Extraction (Peqlab). Red/ET recombineering experiments for plasmid modifications<sup>18</sup> using the strains *E. coli* HS996/pSC101-BAD-gbaA-tet or *E. coli* GB05-red were performed according to the manufacturer's protocol (Gene Bridges GmbH). After selection with suitable antibiotics, clones harbouring correct recombination products were identified by plasmid isolation and restriction analysis with a set of different endonucleases. Details on the construction of all plasmids generated in this study are given in Table S3.

### 3.3.3 Transfer and chromosomal integration of the expression constructs into the heterologous host

According to a previously established electroporation procedure the host strain *Myxococcus xanthus* DK1622<sup>19</sup> or derivatives thereof were transformed with the generated expression constructs (see Table 1). *M. xanthus* DK1622 mutants were routinely cultivated at 30 °C in CTT medium or on CTT agar (1% casitone, 10 mM Tris buffer pH 7.6, 1 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.6, 8 mM MgSO<sub>4</sub> (1.5% agar) with final pH 7.6). For liquid cultures, the strains were grown in Erlenmeyer flasks on an orbital shaker at 180 rpm for 3-6 days. For selection of *M. xanthus* mutants 50 µg/mL kanamycin, 50 µg/mL zeocin, 12 µg/mL oxytetracycline or combinations of 50 µg/mL kanamycin with 50 µg/mL zeocin or 50 µg/mL zeocin with 12 µg/mL oxytetracycline were used. Successful chromosomal integration of the expression constructs via transposase or *mx8* and *mx9* phage integrase into the host genome was verified by PCR. Transformants' cells were lysed by incubation at 95 °C for 20 minutes prior of being added to the PCR reaction ('colony PCR'). For each expression construct successful chromosomal integration was confirmed using suitable primer combinations revealing PCR products of the expected sizes: transposon based constructs: dpo-cor-1/dpo-cor-2 (667 bp) and dpo-cor-3/dpo-cor-4 (599 bp); *mx8* phage integrase based constructs: Mx8-attB-up2/Mx8-attP-down (427 bp) and Mx8-attP-up2/Mx8-attB-down (403 bp) (Fig. S2). A complementary experiment using primers mx8-attB-up2/Mx8-attB-down revealed a 449 bp PCR product for *M. xanthus* DK1622, but not for the *mx8* phage integrase based expression strains. Successful chromosomal integration by *mx9* phage integrase was verified according to the protocol used by Gemperlein et al.<sup>20</sup>. In all cases, genomic DNA of *M. xanthus* DK1622 was used as negative control.

PCR reactions were performed according to the conditions described in chapter 3.3.2 using Taq polymerase; for primer sequences see Table S1.

**Table 1.** *Myxococcus xanthus* (*Mx.*) expression strains generated in this study (for details on construction of expression plasmids see Table S3).

| Strain name  | Expressed genetic constructs                     | COR derivative yields                            |
|--|--|--|
| <i>Mx.</i> DK1622::pDPO-mxn116-Pvan-Tpase                              | pDPO-mxn116-Pvan-Tpase                           | COR A: +++, preCOR A: +++, COR D: +, oxyCOR A: + |
| <i>Mx.</i> DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corN                | pDPO-Mxn116-Pvan-Tpase $\Delta$ corN             | COR A: -, preCOR A: -, COR D: +++, oxyCOR A: ++  |
| <i>Mx.</i> DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corO                | pDPO-Mxn116-Pvan-Tpase $\Delta$ corO             | COR A: -, preCOR A: +++, COR D: ++, oxyCOR A: -  |
| <i>Mx.</i> DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O              | pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O           | COR A: -, preCOR A: -, COR D: +++, oxyCOR A: -   |
| <i>Mx.</i> DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corM                | pDPO-Mxn116-Pvan-Tpase $\Delta$ corM             | COR A: ++, preCOR A: ++, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corP                | pDPO-Mxn116-Pvan-Tpase $\Delta$ corP             | COR A: ++, preCOR A: ++, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622 $\Delta$ mchA::pDPO-mxn116::pUC18-Zeo-mx8-nptII-corO | pDPO-mxn116, pUC18-Zeo-mx8-nptII-corO            | COR A: +++, preCOR A: -, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622 $\Delta$ mchA::pDPO-mxn116::pUC18-Zeo-mx8-Pvan-corO  | pDPO-mxn116, pUC18-Zeo-mx8-Pvan-corO             | COR A: +++, preCOR A: -, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622 $\Delta$ mchA::pDPO-mxn116::pUC18-Zeo-mx8-nptII-corM | pDPO-mxn116, pUC18-Zeo-mx8-nptII-corM            | COR A: ++, preCOR A: ++, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622 $\Delta$ mchA::pDPO-mxn116::pUC18-Zeo-mx8-Pvan-corM  | pDPO-mxn116, pUC18-Zeo-mx8-Pvan-corM             | COR A: ++, preCOR A: ++, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622 $\Delta$ mchA::pDPO-mxn116::pUC18-Zeo-mx8-nptII-mxnM | pDPO-mxn116, pUC18-Zeo-mx8-nptII-mxnM            | COR A: ++, preCOR A: ++, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622 $\Delta$ mchA::pDPO-mxn116::pUC18-Zeo-mx8-Pvan-mxnM  | pDPO-mxn116, pUC18-Zeo-mx8-Pvan-mxnM             | COR A: ++, preCOR A: ++, COR D: +, oxyCOR A: +   |
| <i>Mx.</i> DK1622::pDPO-mxn116-Pvan-Tpase::pUC18-Zeo-mx8-nptII-corP    | pDPO-mxn116-Pvan-Tpase, pUC18-Zeo-mx8-nptII-corP | COR A: +++, preCOR A: +++, COR D: +, oxyCOR A: + |

+++ high production

++ low production

+ trace amounts

- no production

### 3.3.4 Cultivation of the heterologous expression strains

All cultivations were performed according to the previously established procedure<sup>15</sup>. Strains were inoculated from cryo stocks and grown on agar plates for several days until plates were mostly overgrown with cells. All of the cells were scraped from the plates to inoculate preculture medium M7/s4 (0.5% soy flour, 0.5% corn starch, 0.2% glucose, 0.1% yeast extract, 0.1% MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 0.1% CaCl<sub>2</sub> x 2 H<sub>2</sub>O, 1% HEPES, with final pH 7.4 and supplemented with 0.1 mg/L of vitamin B12 and 5 mg/L of FeCl<sub>3</sub> after autoclaving) (50 mL medium in 300 mL Erlenmeyer flask), which was cultivated at 30 °C, 180 rpm for 48 h. Five mL of well grown seed culture was used to inoculate 50 mL production medium M7/s6 (0.5% soy flour, 0.5% corn starch, 0.2% glucose, 0.1% yeast extract, 1% potassium acetate, 0.1%

MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 0.1% CaCl<sub>2</sub> x 2 H<sub>2</sub>O, 1% HEPES, with final pH 7.4 and supplemented with 0.1 mg/L of vitamin B12 and 5 mg/L of FeCl<sub>3</sub> after autoclaving) in which the strain was grown at the same conditions for 6 days. In case of *M. xanthus* DK1622 derivatives, media were amended with suitable antibiotics (50 µg/mL kanamycin, 50 µg/mL zeocin and 12 µg/mL oxytetracycline for heterologous producers and without antibiotics for the wild type host strain). All cultivations were performed in triplicates at 30 °C and 180 rpm for 6 days. After addition of an equal volume of MeOH (55 mL) to the culture broth and incubation in a culture flask on an orbital shaker (180 rpm) for 1 h, 2 mL of the supernatant/MeOH mixture was transferred to an Eppendorf tube and centrifuged at 15000 rpm, 4 °C for 15 minutes prior the HPLC-MS analysis. Quantification of production yields was performed as described in the supplementary information.

### 3.4 Results and discussion

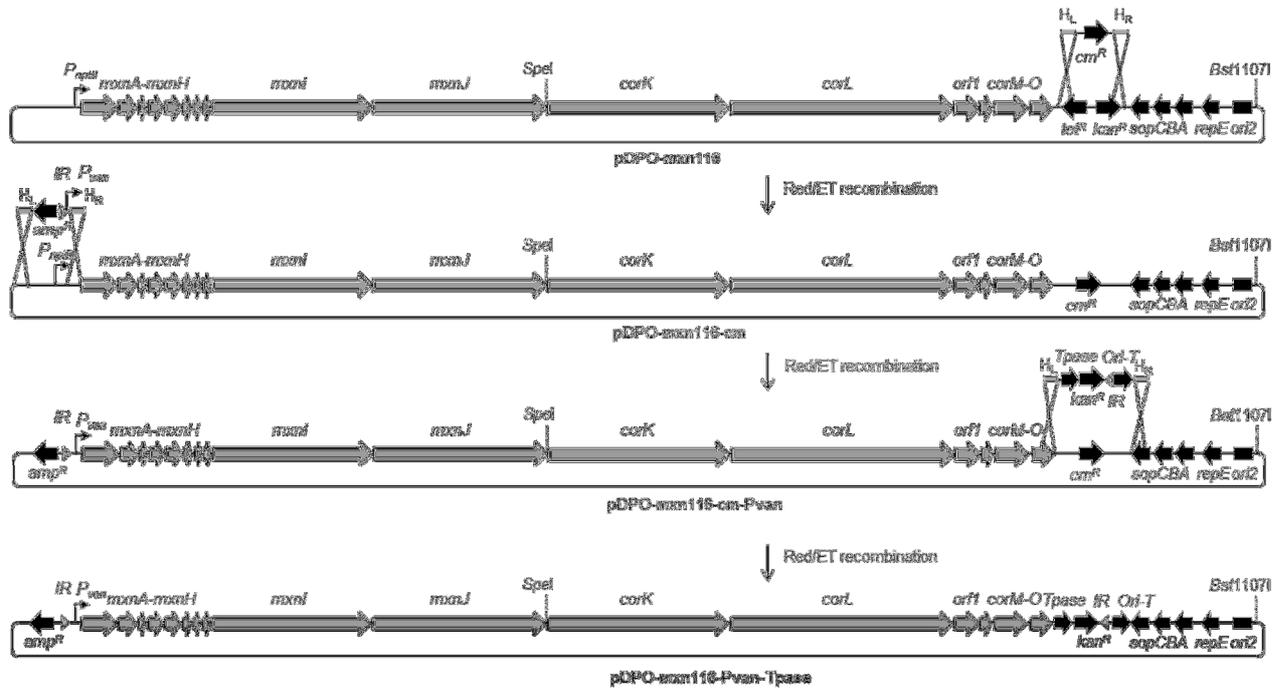
#### 3.4.1 Construction of a transposon-based COR BGC expression construct

One of the goals of this study was to improve the COR A production yield of the previously generated heterologous producer strain<sup>15</sup>. To achieve this goal additional genetic modifications were designed and implemented to the original heterologous expression construct pDPO-mxn116, as outlined in Figure 1. Two major modification attempts were pursued to evaluate different chromosomal integration sites as well as efficacy of different promoters. Promoter exchange is one way to improve transcription of biosynthetic pathways<sup>21</sup> as different promoters can have a big impact on gene expression levels, consequently leading to increased or decreased secondary metabolite production<sup>22,23</sup>. Furthermore, inducible promoters can have an advantage in initiating the production at a certain time point during fermentation, which can in some cases lead to increased production, especially if the product turns out to be toxic or inhibitory for the host organism<sup>24</sup>. Those aspects make selection of a suitable promoter a crucial part of BGC expression construct design. Another factor to consider when improving production yield is the genomic integration site, as positional effects can have large impact on the overall expression and yield<sup>25-27</sup>. Several methods are available for integration of BGC into the host genome, some of which are site-specific and lead to integration in a precisely defined locus (e.g. homologous recombination, phage integration), whereas others are unspecific and integrate randomly in the genome (e.g. transposition)<sup>22,28</sup>. In most cases, methods that allow integration in a specific locus are preferred. Nevertheless, we chose random genome integration by a transposition for this study, since it allows to

screen a high number of myxobacterial clones harboring the COR BGC at various genomic loci and subsequently compare their production levels. To evaluate additional chromosome integration sites in the *M. xanthus* DK1622 heterologous host, the *tet<sup>R</sup>* gene on the expression plasmid was exchanged with a more effective transposase integration system<sup>28</sup>. The *tet<sup>R</sup>* was previously used as locus for genome integration by homologous recombination, but the efficiency of homologous integration was found very low. This is in line with our previous results showing that transposon based integration systems for BGCs are significantly more efficient than chromosomal integration via homologous recombination<sup>28</sup>. Since the choice of known selection markers in *M. xanthus* is limited, the existing *tet<sup>R</sup>* and *kan<sup>R</sup>* genes were removed to make them available in later modification steps. Conveniently, both genes were located next to each other on the backbone of pDPO-mxn116 construct, making it possible to replace them with a *cm<sup>R</sup>* gene in one *in vivo*, Red/ET recombination step, resulting in construction of pDPO-mxn116-cm. The detailed procedure is provided in the supplementary information.

To further improve COR production, we aimed to evaluate the *P<sub>van</sub>* promoter system. By performing one Red/ET recombination step, the pDPO-mxn116-cm-Pvan construct was generated (Fig. 1). The construct had to be further modified to allow for its integration in *M. xanthus* genome, since the previous integration cassette, based on homologous recombination via *tet<sup>R</sup>*, was removed in the first modification step.

Employing the same Red/ET technology, a previously generated pDPO-mxn116-cm-Pvan construct was modified to introduce a transposase gene on its backbone. The resulting construct pDPO-mxn116-Pvan-Tpase now harbored all necessary elements for heterologous expression in *M. xanthus* DK1622 (Fig. 1). A detailed description of the heterologous expression construct design is provided in the supplementary information.



**Figure 1.** Sequence of RedET *in vivo* recombination steps required to construct pDPO-mxn116-Pvan-Tpase construct starting from pDPO-mxn116 expression plasmid.

### 3.4.2 Production of corallopyronin by an inducible system after transposition into the heterologous host *M. xanthus* DK1622

The generated *mxn/cor* gene cluster expression construct pDPO-mxn116-Pvan-Tpase (77 kb) was transferred to *M. xanthus* DK1622 by electroporation. Clones growing on selection plates amended with kanamycin were analyzed for successful chromosomal integration of the transposable cassette by PCR (Fig. S2). After genotypic verification, the *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase mutant and *M. xanthus* DK1622 were cultivated under standard conditions in M7/s6 medium for 6 days and culture extracts were analyzed for the presence of corallopyronins. By HPLC-MS analysis the production of COR could indeed be confirmed in the extracts from the mutant strain *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase (Fig. 2). This proves the successful and functional expression of the heterologous *mxn/cor* gene cluster under the control of the  $P_{van}$  promoter after random transposition into *M. xanthus* DK1622.

To evaluate the effect of different genome loci on the yield of COR, several transposon mutants (*M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase) were evaluated in comparison to the previously described heterologous producer *M. xanthus* DK1622  $\Delta mchA$ ::pDPO-mxn116<sup>15</sup>. One study from myxobacteria<sup>27</sup> and several more from other bacteria<sup>29,25,26</sup> could previously underpin the impact of the positional genomic effect on the gene expression. We screened five independent clones which were cultivated in parallel and the corresponding

culture extracts were analysed for the presence of COR. Interestingly, production of all five mutants was almost the same under the analyzed conditions and was quantified at around 100 mg/L. This is more than twice the production yield of the previous heterologous producer. Thus, either the genome locus or the promoter exchange greatly benefited the COR production. Since the transposase mediates chromosomal integration at random loci, and the achieved yield in all *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase clones was nearly the same under the analyzed conditions, it is very likely, that BGC location is not the main reason for the observed yield increase. Most likely the *P<sub>van</sub>* promoter is in this case superior to the previously used *P<sub>npIII</sub>*, probably because of better suited transcription rate.

Surprisingly, in addition to COR A, significant amounts of preCOR A (around 160 mg/L) were detected in *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase extracts (Fig. 2). PreCOR A (see Figure 6B) is assumed to be the biosynthetic precursor of COR A which is oxidized by the cytochrome P450 enzyme CorO. Therefore, the total yield of corallopyronins in this set of experiments was increased to more than 250 mg/L (cp. 120 mg/L in previous experiments) with an obvious bottleneck in preCOR A oxidation. This finding lead to the speculation that the chosen construct induces a misbalance of activity of the responsible oxidase CorO and the remaining biosynthetic pathway (see chapter 3.4.4 and 3.4.5).

### 3.4.3 Evaluation of CorM function

CorM was described as a type II TE domain<sup>30</sup>. These enzymes are known to have editing function, removing wrong substrates and nonreactive moieties attached to the carrier proteins blocking the assembly line, thus resulting in increased production<sup>31,32</sup>. Contrary to most of the other proteins in the corallopyronin BGC, CorM does not have a homolog in the structurally related myxopyronin BGC from *M. fulvus* Mx f50<sup>33</sup>. However, MxnM from the myxopyronin biosynthetic pathway is described as a type II AT domain<sup>33</sup>, an enzyme class that has also been described to provide proofreading functionality<sup>34</sup>, similar to the type II TE enzymes.

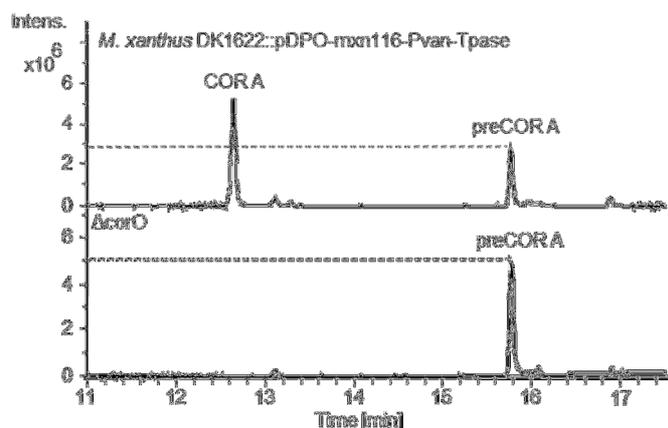
To evaluate the function of CorM a gene deletion experiment was performed. The *apra<sup>R</sup>* gene was PCR amplified and used to replace *corM* by a single *in vivo* recombination step in *E. coli* GB05-red. The resulting pDPO-Mxn116-Pvan-Tpase $\Delta$ corM-apra<sup>R</sup> plasmid was hydrolyzed by *Xma*JI restriction endonuclease and re-ligated, to excise *apra<sup>R</sup>*. The final construct pDPO-Mxn116-Pvan-Tpase $\Delta$ corM was transferred into *M. xanthus* DK1622 by electroporation. The details on genotypic verification, cultivation and quantification are provided in the Materials and Methods / Supplementary information. Product quantification revealed an average 77% yield decrease in the CorM deletion mutant compared to the control strain showing that the

type II TE CorM is important for high yield production of COR A. We thus hypothesized that an additional copy of *corM* might increase production yield, hence another copy of the gene was co-expressed in the heterologous COR producer *M. xanthus* DK1622  $\Delta mchA::pDPO-mxn116$ . The analogous question was in parallel addressed for *mxnM*. To study the effect of overexpression of these two genes two different promoter systems well established in *M. xanthus*, were tested: the constitutive  $P_{npIII}$  promoter of the neomycin phosphotransferase II resistance gene from Tn5<sup>35</sup> and the vanillate inducible  $P_{van}$  promoter<sup>36</sup>, which was implemented without repressor gene (to also achieve constitutive expression). Generated expression constructs (see Table 1 and S3) were transferred into *M. xanthus* DK1622  $\Delta mchA::pDPO-mxn116$  by electroporation. The obtained mutants (see Table 1) were cultivated in parallel with *M. xanthus* DK1622  $\Delta mchA::pDPO-mxn116$  control mutant. The details on genotypic verification, cultivation and quantification are provided in the Materials and Methods / Supplementary information. As expected, HPLC-MS analysis revealed production of both metabolites in all of the evaluated mutants, however a detailed quantification revealed around 50% decrease of overall COR production in CorM and MxnM overexpression mutants. The overexpression of either proofreading enzyme probably results in unspecific cleaving of correct polyketide intermediates from the assembly line, leading to a decreased overall production efficiency<sup>32,37</sup>. This result indicates that assembly line maintenance by CorM (MxnM) needs to be fine-tuned and is not a main bottleneck in the current expression system under the analyzed conditions.

#### 3.4.4 Deletion of *corO*

The previously generated heterologous expression system<sup>15</sup> enabled easy pathway modifications and gene function studies. Red/ET *in vivo* recombination<sup>18</sup> allowed to perform simple gene deletions in a quick two-step process. This method opened up the possibility to identify functions of deleted and overexpressed genes as well as to potentially produce novel derivatives in the process. As functions of most genes in the *mxn/cor* clusters had not been experimentally proven, it was essential to perform gene deletion experiments to assign functions to corresponding genes in the BGC. To exclude the possibility that part of the hydroxylation activity required to form COR A from preCOR A comes from an unidentified protein encoded in the host genome, we deleted *corO* from the expression construct. This experiment mostly aimed to generate a production strain providing access to pure preCOR A. To perform the *corO* deletion the same approach as for *corM* was applied (see chapter 3.4.3). The final construct pDPO-Mxn116-Pvan-Tpase $\Delta$ corO was transferred into *M. xanthus*

DK1622 by electroporation. The details on genotypic verification, cultivation and quantification are provided in the Materials and Methods / Supplementary information. As expected, HPLC-MS analysis revealed production of both compounds in *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase, however in *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corO only production of preCOR A was observed (Fig. 2). This finding confirms that CorO is the only enzyme responsible for hydroxylation of preCOR A at C-24 and thus its conversion to COR A in *M. xanthus* (Fig. 6B). *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corO can be employed for production of preCOR A as this mutant revealed 300-350 mg/L production of preCOR A, which confirmed the overall high corallopyronin yield reported in 3.4.2 and represents more than a twofold increase over the original expression strain (*M. xanthus* DK1622  $\Delta$ mchA::pDPO-mxn116-Pvan-Tpase). This result is important since preCOR A is an interesting candidate for drug development, as its antibiotic potential was shown to be comparable to that of the structurally more complex COR A<sup>16</sup> exhibiting increased lipophilicity important to achieve worm penetration.



**Figure 2.** Production analysis of *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase and *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corO cultivated in the M7/s6 medium. The HPLC-MS analysis of supernatant showing base peak chromatograms. Peak areas for preCOR A are 8.9 (Tpase) versus 19.6 ( $\Delta$ corO).

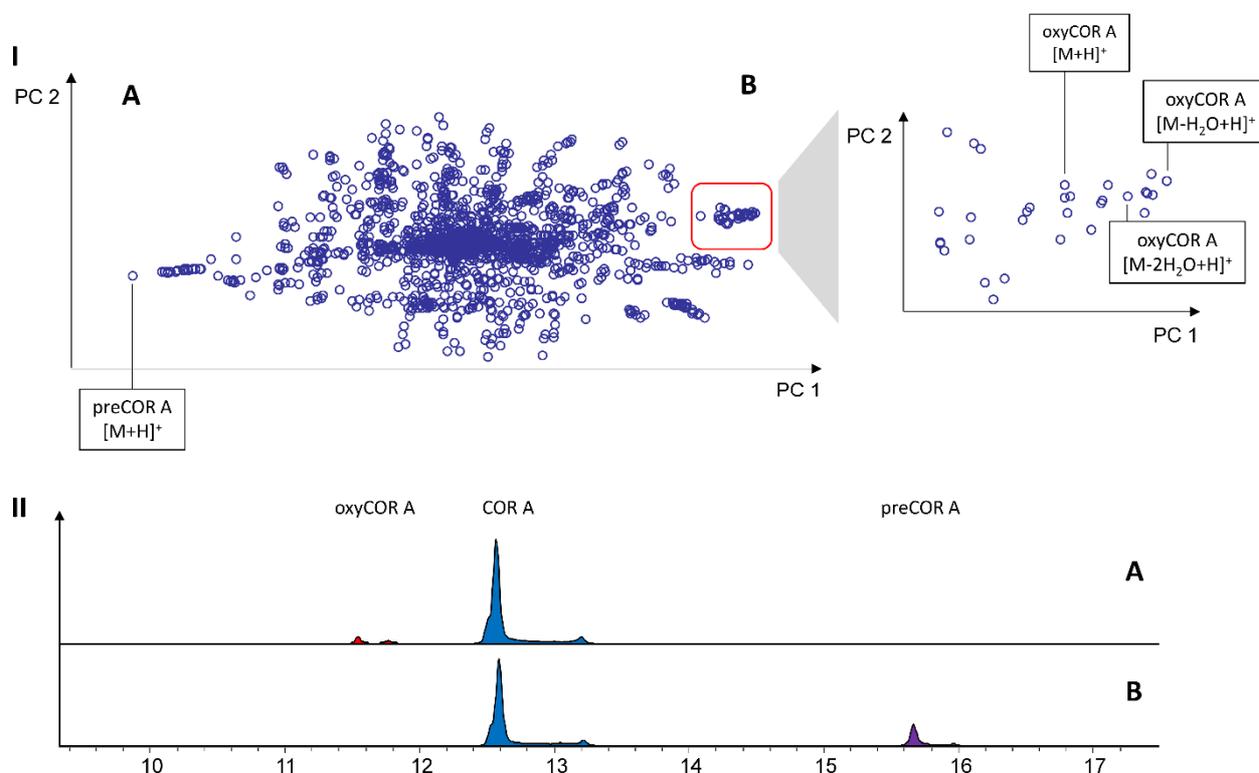
### 3.4.5 Overexpression of CorO: An attempt to improve conversion of preCOR A to COR A and thus increase COR A production

To further improve COR A production starting from the mutant in the optimized medium we aimed to improve conversion of preCOR A to COR A. As hypothesized in previous studies<sup>30,15</sup> and experimentally proven in this study (see chapter 3.4.4), the cytochrome P450 CorO is responsible for preCOR A hydroxylation to form COR A. During the cultivation of *M. xanthus* DK1622  $\Delta$ mchA::pDPO-mxn116 producer strain in M7/s6 medium, high levels of preCOR A were detected in addition to the COR A pointing towards incomplete conversion

via CorO leading to accumulation of preCOR A. Further supporting this hypothesis almost no preCOR A could be detected in extracts from the native producer *C. coralloides* Cc c127 or in the *M. xanthus* DK1622  $\Delta mchA::pDPO-mxn116$  heterologous producer when cultivated in CTT medium (data not shown). The overall COR yield, measured under those conditions, was significantly lower compared to the production in the optimized production medium M7/s6<sup>15</sup>. It is thus likely that CorO achieved complete turnover under the “low” production conditions, however, once the production was optimized, it could no longer hydroxylate all of the substrate, resulting in preCOR A accumulation. It is also possible that hydroxylation is inefficient because CorO is missing appropriate interaction partners (e.g. ferredoxin and/or ferredoxin reductase) in the heterologous producer. In this case corresponding homologous proteins present in *M. xanthus* DK1622 might be performing these functions. However, their efficiency might be lower, resulting in only partial conversion of preCOR A to COR A.

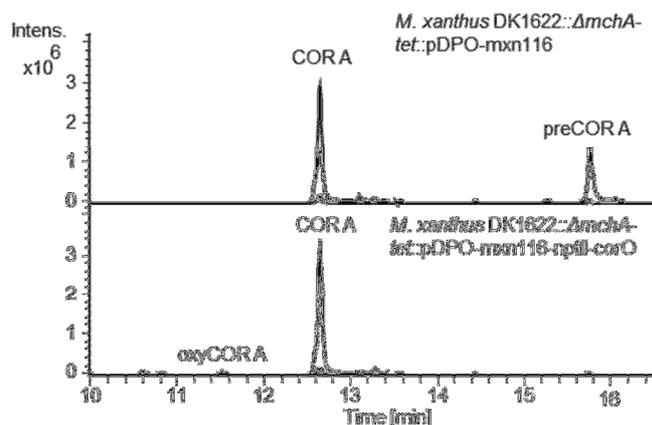
In order to facilitate hydroxylation of preCOR A to increase production of COR A, an additional copy of *corO* was co-expressed in *M. xanthus* DK1622  $\Delta mchA::pDPO-mxn116$ . The co-expression was performed in the same fashion as already described for the CorM/MxnM (see chapter 3.4.3). As observed before HPLC-MS analysis revealed production of both metabolites in the *M. xanthus* DK1622  $\Delta mchA::pDPO-mxn116$  mutant (Fig. 2 and 3). As expected in clones harboring a second copy of *corO* under the control of either promoter, preCOR A production was basically abolished (Fig. 2 and 3). Surprisingly, however, the yield of COR A remained approximately unchanged indicating conversion of preCOR A or COR A to an unknown product. In order to detect putative novel COR derivatives that might arise from the overexpression of *corO*, we analyzed the secondary metabolome of *M. xanthus* DK1622:: $pDPO-Mxn116-nptII-corO$  in more detail. Therefore, extracts were profiled against such of the control strain *M. xanthus* DK1622:: $pDPO-Mxn116$ , using principal component analysis (PCA) applied to the HPLC-MS chromatograms (see Supplementary information)<sup>38</sup>. The PCA revealed an additional minor derivative eluting after 11.5 min (Fig. 3) with an  $[M+H]^+$  of 544.290 m/z as well as slightly more prominent ions  $[M-H_2O+H]^+$  of 526.279 and  $[M-2H_2O+H]^+$  of 508.267. This monoisotopic mass indicates a mass addition of 16 respective to the mass of COR A, which indicated the presence an additional OH group. The yield of the novel derivative was very low at this point requiring significant effort to elucidate its structure by NMR. Fortunately, the same derivative was detected in another mutant in parallel studies, in noticeably higher amount (see chapter 3.4.6). Quantification of preCOR A yields revealed around ~80 mg/L of preCOR A in extracts of *M. xanthus* DK1622  $\Delta mchA-tet::pDPO-mxn116$

and total abolishment of preCOR A in extracts of *M. xanthus* DK1622  $\Delta mchA-tet::pDPO-mxn116-nptII-corO$ .



**Figure 3. I)** Identification of oxyCOR A by PCA based metabolome comparison of *M. xanthus* DK1622  $\Delta mchA-tet::pDPO-Mxn116-nptII-corO$  (A) and *M. xanthus* DK1622  $\Delta mchA-tet::pDPO-Mxn116$  (B) and **II)** comparative analysis of the extracted Ion chromatograms (EIC) for oxyCOR A  $[M-H_2O+H]^+ = 526.280$ , COR A  $[M-H_2O+H]^+ = 510.285$ , and preCOR A  $[M+H]^+ = 512.295$ .

Despite the detection of the newly identified doubly-hydroxylated derivative (see 3.4.6), the total yield difference between *M. xanthus* DK1622  $\Delta mchA-tet::pDPO-Mxn116$  and the CorO overexpression mutant cannot be explained as the newly identified peak represents a very minor derivative. It is therefore assumed that the overall yield of COR derivatives decreased after overexpression of CorO due to degradation starting from a second hydroxylation step. An unambiguous explanation for this phenomenon is currently not available but it may be possible that CorO acts during polyketide assembly and thus before pyrone ring formation takes place. Its overexpression might thus lead to more efficient hydroxylation of the western chain intermediate, possibly leading to hydroxylation at multiple positions which may also prevent those from further correct processing to COR A.



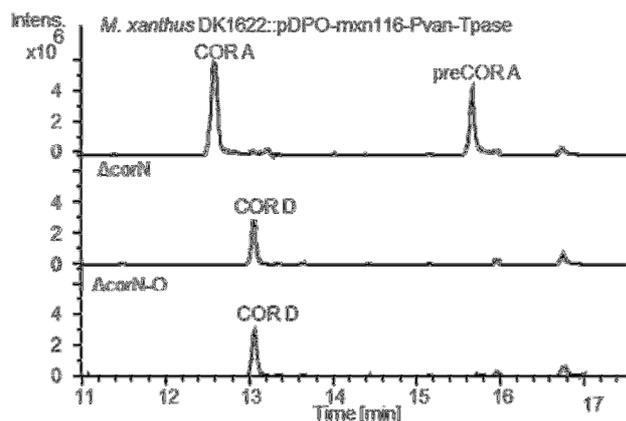
**Figure 4.** Production analysis of *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pDPO-mxn116 and *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pDPO-mxn116-nptII-corO cultivated in the M7/s6 medium, showing significant amount of preCOR A produced in *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pDPO-mxn116 and total abolishment of preCOR A in mutant with additional copy of CorO (*M. xanthus* DK1622  $\Delta mchA$ -*tet*::pDPO-mxn116-nptII-corO). A very small peak for oxyCOR A could also be detected in the latter mutant. HPLC-MS analysis of supernatant showing extracted ion chromatograms (EIC)  $[M + Na]^+ = 550.278$ ,  $[M - H_2O + H]^+ = 510.285$ ,  $[M + H]^+ = 528.295$  (COR A),  $[M + Na]^+ = 534.278$ ,  $[M - H_2O + H]^+ = 494.285$ ,  $[M + H]^+ = 512.295$  (preCOR A) and  $[M - 2H_2O + H]^+ = 508.267$ ,  $[M - H_2O + H]^+ = 526.280$ ,  $[M + H]^+ = 544.290$  (oxyCOR A)

### 3.4.6 Evaluation of the function of CorN

CorN was identified as enoyl-CoA hydratase according to a comprehensive BLAST analysis<sup>30</sup> and was proposed to be responsible for double bond isomerization from  $\Delta^{24,25}$  to the  $\Delta^{25,27}$  position. To confirm this hypothesis and produce novel COR derivatives, a *corN* gene deletion was performed as described below. To facilitate production of additional novel derivatives, we also aimed to construct a *corN/corO* deletion mutant.

The same approach as for *corM* (see chapter 3.4.4) was applied. The *apra*<sup>R</sup> gene was flanked by *Xma*II R-sites in all cases, however the flanking homology regions were adapted for each gene deletion specifically. Two more linear *apra*<sup>R</sup> PCR constructs were amplified, flanked by specific homology regions, designed for deletion of *corN*, as well as *corN-O* region covering both genes. Final constructs pDPO-Mxn116-Pvan-Tpase $\Delta$ corN and pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O were transferred into *M. xanthus* DK1622. The details on genotypic verification, cultivation and quantification are provided in the Materials and Methods / Supplementary information. The culture extracts were analyzed for the presence of COR A, preCOR A and isomers thereof. To our surprise, HPLC-MS analysis revealed a novel  $[M+H]^+$  peak at  $m/z$  528.294 at 13.1 min (Fig. 5) in both gene deletion mutants, whereas COR A and preCOR A were not detected. Chromatograms of the *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase in contrast showed standard production profile with COR A and preCOR A as major metabolites.

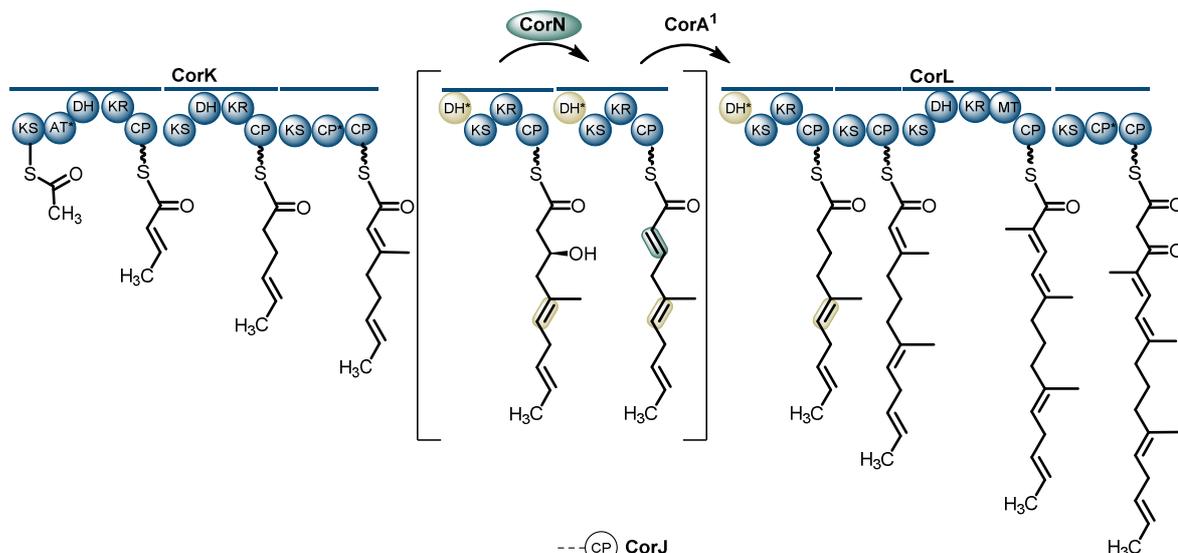
Purification and NMR structure elucidation of the novel peak exhibiting a retention time of 13.1 min. revealed the structure of the novel derivative COR D possessing a hydroxyl group at the C-23 instead of C-24 (Fig. 6C; Fig. S4). Details on the NMR based structure elucidation are provided in the supplementary information. A hydroxyl group at C-23 was thus far not observed in any COR derivative. Based on those results, the current coralopyronin biosynthesis model was revised and a different function was assigned to CorN. Instead of acting as tailoring enzyme involved in double bond isomerization from  $\Delta^{24, 25}$  to  $\Delta^{25, 27}$ <sup>30</sup>, CorN is assumed to eliminate the C-23 hydroxyl group by dehydration during western chain assembly. The double bond which should arise after the dehydration is assumed to be reduced by CorA which provides the *trans*-ER activity<sup>30</sup>. The double bond isomerization from  $\Delta^{24, 25}$  to  $\Delta^{25, 27}$  position might be mediated by the N-terminal DH\* domain of CorL, previously thought to be in charge of OH group dehydration at C-23. This DH domain would in this case fall into the clade of “shifting DH” domains (DH\*)<sup>39,40</sup>. In standard DH domains the H44 residue acts as a catalytic base to deprotonate the  $\alpha$  position of the reduced polyketide intermediate and the D206 residue promotes the elimination<sup>41</sup>. In contrast to standard DH domains, residue H44 was hypothesized to serve as a proton donor in DH\* domains, whereas a yet unknown basic residue should assist in the double bond migration<sup>39</sup>. To support the hypothesis of DH\* in CorL being a “shifting DH”, a protein alignment with known DH\* domains was performed and core motifs were extracted (Fig. S1). The catalytic residue H44 in the HxxxGxxxxP motif could indeed be assigned, whereas the catalytic D206 in DxxxQ/H motif was found to be mutated to N206, which is consistent with most of the other known shifting DH domains<sup>39,40,42</sup>. Nevertheless, in case of the N-terminal CorJ DH\*, which was previously shown to be involved in double bond isomerization of the coralopyronin eastern chain, the D206 in DxxxQ/H was found to be conserved and is thus assumed to be the catalytic residue for the double bond shifting<sup>43</sup>. However, after D206N (D211N<sup>43</sup>) mutation, the enzyme was shown to keep the double bond shifting activity *in vitro*<sup>43</sup>. The basic residue that acts as the proton acceptor in the double bond migration, thus remains unknown.



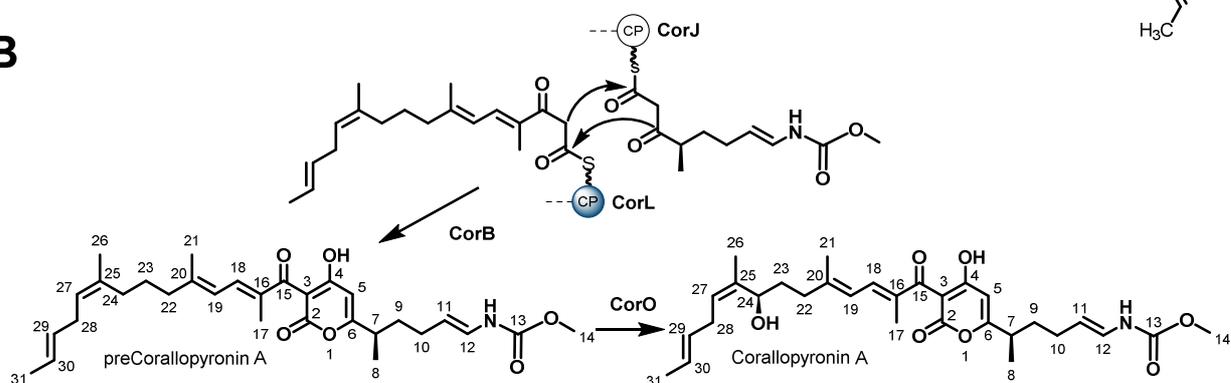
**Figure 5.** Production analysis of *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase, *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corN and *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corN-O cultivated in the M7/s6 medium, showing production profile of different gene deletions. HPLC-MS analysis of supernatant showing base peak chromatograms.

If our hypothesis about CorN function is correct, a downstream reduction activity is required to saturate the double bond between C-23 and C-24, which necessarily arises after the dehydration of the C-23 OH group (Fig. 6A). Such functionality could be provided by CorA exhibiting AT/ER type of activity in *trans*<sup>30</sup>. In the hybrid *mxn/cor* pathway the CorA function is mediated by the homolog MxnA<sup>33</sup>.

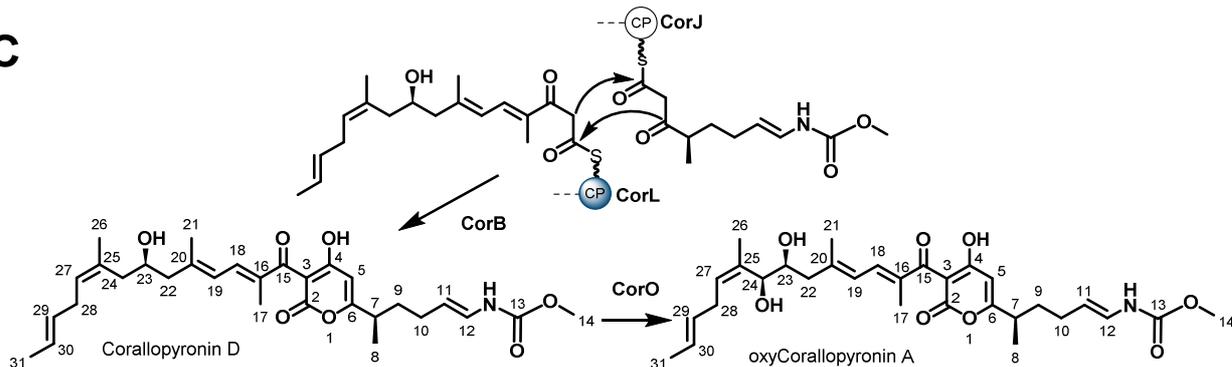
A



B



C

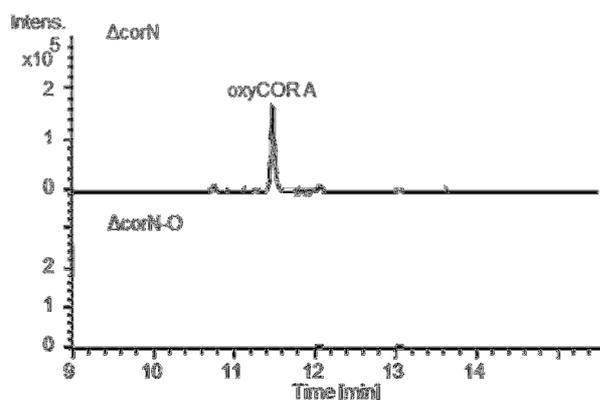


**Figure 6.** Proposed biosynthesis of the corallopyronin A western chain, showing the DH\* with the corresponding migrated double bond highlighted in tan and modifications implemented by CorN highlighted in green (A). Pyrone ring formation and tailoring of the produced intermediate in the native system (B). Pyrone ring formation and tailoring of the produced intermediate in a *corN* deletion mutant (C).

<sup>1</sup>The CorA function is mediated by its homolog MxnA from the myxopyronin BGC.

While screening all available COR cluster mutants, we realized that strain pDPO-Mxn116-Pvan-Tpase $\Delta$ corN produces higher amounts of the new derivative detected in the CorO overexpression mutant with the mass  $[M+H]^+$  of 544.290 m/z (see chapter 3.4.5, Fig. 7). We

thus chose to attempt isolation of said metabolite from pDPO-Mxn116-Pvan-Tpase $\Delta$ corN. Due to quite low production of this new derivative, a 12 L cultivation had to be performed, to isolate enough compound for NMR structure elucidation. Analysis indeed revealed the presence of two OH groups, present at the C-23 and C-24 (Fig. 6C, Fig. S5). The new derivative, named oxyCOR A, is thus probably biosynthesized from COR D via hydroxylation. The fact that oxyCOR A could only be detected in extracts of the  $\Delta$ corN mutant (*M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corN) and not in the extract of  $\Delta$ corN-O mutant (*M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O), can be interpreted in analogy with the previous statement that CorO is responsible for hydroxylation at position C-24 (see chapters 3.1 and 3.5). Addition of the OH group at the C-24 of COR D leads to production of oxyCOR A (Fig. 6C). Very low amounts for oxyCOR A detected in the extracts of the  $\Delta$ corN-O mutant could be caused by inefficient hydroxylation of COR D, which could be due to the steric hindrance of C-24 position by the C-23 OH group of COR D. This also correlates well with the results of principal component analysis performed on *M. xanthus* DK1622::pDPO-Mxn116-nptII-corO (see chapter 3.4.5), where an additional copy of the CorO is probably responsible for more efficient hydroxylation of minor COR D amounts present in the broth (see Table 1), leading to production of oxyCOR A.



**Figure 7.** Production analysis of *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corN and *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corN-O cultivated in the M7/s6 medium, showing oxyCOR A production in *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corN, whereas no oxyCOR A was detected in *M. xanthus* DK1622::pDPO-mxn116-Pvan-Tpase $\Delta$ corN-O. HPLC-MS analysis of supernatant showing extracted ion chromatograms (EIC)  $[M - H_2O + H]^+ = 526.279$ ,  $[M + H]^+ = 544.290$  (oxyCOR A).

### 3.4.7 Evaluation of the function of CorP

An open reading frame encoding a hypothetical protein (CorP), was found in the cluster, positioned between the assembly line and the three tailoring genes (*corM*, *corN* and *corO*)<sup>30</sup>. Its function remained elusive, as no homologue with assigned specific function could be

found in any database and no additional functionality necessary for corallopyronin production was missing from the identified genes. By screening the internal database of myxobacterial genomes using BLAST, several homologues of CorP with unknown function were identified in BGCs of myxopyronin, sorangicin and several yet unknown biosynthetic pathways. To elucidate whether the protein is essential for the biosynthesis and/or to finally decrypt its function, a gene deletion was performed according to the described method (see supplementary information). The final construct pDPO-Mxn116-Pvan-Tpase $\Delta$ corP was transferred into *M. xanthus* DK1622. The details on genotypic verification, cultivation and quantification are provided in the Materials and Methods / Supplementary information. Quantification revealed an average of 82% yield decrease in the *corP* deletion mutant compared to the control strain. As no additional enzymatic function was required for COR biosynthesis according to the current model, the role of CorP remained ambiguous. Taking into account the high yield of COR A and the unexplained high levels of preCOR A, together with the decrease in yield after the *corP* deletion, we assumed that the mutant might lack sufficient self-resistance. Intriguingly, the aforementioned homologs identified by BLAST analysis were all found in BGCs of other known RNAP inhibitors. Furthermore, a CorP homolog was found upstream of the ripostatin BGC from *Sorangium cellulosum* So ce377 which was not obviously involved in the biosynthesis<sup>44</sup>. Consequently, we hypothesized that the protein could function as the self-resistance protein in ripostatin, corallopyronin, myxopyronin, sorangicin and possibly other RNAP inhibitor BGCs. In order to evaluate this hypothesis, we first aimed to express *corP* in corallopyronin susceptible *E. coli*  $\Delta$ TolC, as a similar approach was successfully applied in other studies (e.g. bengamide study<sup>45</sup>). An *mx8* integrase based expression plasmid with *corP* driven by the *P<sub>nptII</sub>* promoter was constructed (pUC18-zeo-mx8-nptII-corP, Table S3). The final construct as well as the empty control plasmid (pUC18-Zeo-mx8-MCS, Table S3) were transferred into *E. coli*  $\Delta$ TolC. Clones were selected on LB agar supplemented with zeocin and the COR A MIC was evaluated. Results revealed no difference between *E. coli*  $\Delta$ TolC::pUC18-zeo-mx8-MCS and *E. coli*  $\Delta$ TolC::pUC18-zeo-mx8-nptII-corP, as both values were determined at 1 mg/L. The results showed that CorP does not confer resistance against COR A in *E. coli*  $\Delta$ TolC, potentially because the myxobacterial protein is not functionally expressed in *E. coli*. Therefore, we further aimed to express CorP in the myxobacterial host *M. xanthus* DK1622. Previously constructed plasmids pUC18-zeo-mx8-nptII-corP and pUC18-Zeo-mx8-MCS were transferred into *M. xanthus* DK1622 and clones growing on the selection plates amended with zeocin were analyzed for correct chromosomal integration of the expression plasmids by PCR

(Fig. S2). Generated mutants *M. xanthus* DK1622::pUC18-zeo-mx8-nptII-corP and *M. xanthus* DK1622::pUC18-zeo-mx8-MCS were grown in presence of COR A and their MIC was determined at ~64 mg/L and ~4 mg/L, respectively, confirming that CorP confers resistance against corallopyronin in *M. xanthus* DK1622. To provide additional evidence that CorP confers resistance to COR A in *M. xanthus* DK1622, we also evaluated the MIC of the CorP deletion mutant *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corP. The MIC was measured at ~4 mg/L which is the same value as measured for the control mutant harboring empty plasmid without CorP and for the wild type *M. xanthus* DK1622, providing solid evidence that CorP indeed mediates increased resistance to COR.

With this new information available, we further investigated if co-expression of an additional CorP copy in the COR A producer mutant further increases COR A production. The pUC18-zeo-mx8-nptII-corP plasmid was transferred into *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase. The details on genotypic verification, cultivation and quantification are provided in Materials and Methods / Supplementary information. Quantification of COR A production revealed no difference between both mutants (~ 100 mg/L) indicating that co-expression of an additional *corP* copy does not increase the level of self-resistance and/or other bottlenecks in COR biosynthesis limit production levels.

In order to determine if increased levels of CorP have a positive impact on COR resistance a compatible vector for co-expression of a second *corP* copy was constructed (pUC18-TetR-mx9-nptII-corP; Table S3). The construct harbored *corP* driven by the  $P_{nptII}$  and an *mx9* phage integrase<sup>46</sup> for integration in a different genomic locus than previously described for the *mx8* based construct. The plasmid was transferred into *M. xanthus* DK1622::pUC18-zeo-mx8-nptII-corP and clones growing on the selection plates amended with zeocin and oxytetracycline were analyzed for correct chromosomal integration of the expression plasmid by PCR<sup>20</sup>. Finally, clone *M. xanthus* DK1622::pUC18-zeo-mx8-nptII-corP::pUC18-TetR-mx9-nptII-corP was grown in presence of COR A and its MIC was evaluated using the standard protocol (see Supplementary information). Additional concentrations between 64 mg/L and 128 mg/L (e.g. 73, 83 and 100 mg/L) were also evaluated. The MIC was determined at ~64 mg/L, the same as for the strain with a single expressed *corP* copy. This result shows that the expression of an additional *corP* copy does not further increase the self-resistance to COR A in *M. xanthus* DK1622. The exact molecular basis of CorP conferred resistance to RNAP inhibitors remains to be determined in ongoing studies.

### 3.5 Conclusion

One of the main aspects of this study was the modification of the coralopyronin heterologous production platform leading to significantly increased production of ~100 mg/L COR A and ~160 mg/L preCOR A. Deletion of *corO* led to a strain generating preCOR A only at high production levels of 350 mg/L. Overexpression of the cytochrome P450 enzyme (CorO) did not result in stable conversion of the remaining preCOR A to COR A. However, this experiment led to the discovery of a novel dihydroxylated derivative – oxyCOR A. Another new derivative COR D was detected after the deletion of *corN*, which also helped to redefine the corresponding protein as an ECH-like enzyme performing a dehydration reaction in *trans* during the chain elongation in western chain biosynthesis. *In silico* analysis of the CorL N-terminal DH domain revealed a missing typical DH domain motif and showed some shared traits with DH\* domains. This led us to hypothesize that this domain is responsible for the isomerization of the double bond from  $\Delta^{24, 25}$  to  $\Delta^{25, 27}$ . Furthermore, deletion of *corP*, together with the MIC studies performed on the heterologous host and mutant strains, revealed the function of CorP to be a self-resistance protein. Interestingly, homologs of CorP were found in BGCs of several other RNAP inhibitors (e.g. myxopyronin, sorangicin, ripostatin). We assume that this type of protein represents an RNAP inhibitor specific self-resistance mechanism in myxobacteria, which might in the future serve as a great tool in the identification of novel RNAP inhibitor encoding BGCs. The exact function of CorP in COR resistance is subject of further studies.

The contributions to the elucidation of coralopyronin biosynthesis achieved in this study provide a better understanding of several biosynthetic steps and enable better control of COR production which will lead to even further increase of production yield in the future. Production optimization is currently of high priority in order to continue the ongoing evaluation as (pre)clinical candidate.

### Acknowledgments

We would like to thank Dr. Nestor Zaburanyi for bioinformatics support, Daniel Sauer for HPLC-MS measurements and Viktoria Schmitt for assistance with MIC testing.

This work was generously funded by a grant from DZIF (TTU09.807).

### 3.6 References

- (1) Schiefer, A.; Schmitz, A.; Schäberle, T. F.; Specht, S.; Lämmer, C.; Johnston, K. L.; Vassilyev, D. G.; König, G. M.; Hoerauf, A.; Pfarr, K. Corallopyronin A Specifically Targets and Depletes Essential Obligate Wolbachia Endobacteria From Filarial Nematodes In Vivo, *J. Infect. Dis.* **2012**, *206*, pp. 249–257.
- (2) Schäberle, T. F.; Schiefer, A.; Schmitz, A.; König, G. M.; Hoerauf, A.; Pfarr, K. Corallopyronin A - a promising antibiotic for treatment of filariasis, *Int. J. Med. Microbiol.* **2014**, *304*, pp. 72–78.
- (3) Osei-Atweneboana, M. Y.; Awadzi, K.; Attah, S. K.; Boakye, D. A.; Gyapong, J. O.; Prichard, R. K. Phenotypic evidence of emerging ivermectin resistance in *Onchocerca volvulus*, *PLoS Negl. Trop. Dis.* **2011**, *5*, e998.
- (4) Bandi, C.; McCall, J. W.; Genchi, C.; Corona, S.; Venco, L.; Sacchi, L. Effects of tetracycline on the filarial worms *Brugia pahangi* and *Dirofilaria immitis* and their bacterial endosymbionts *Wolbachia*, *Int. J. Parasitol.* **1999**, *29*, pp. 357–364.
- (5) Hoerauf, A.; Nissen-Pähle, K.; Schmetz, C.; Henkle-Dührsen, K.; Blaxter, M. L.; Büttner, D. W.; Gallin, M. Y.; Al-Qaoud, K. M.; Lucius, R.; Fleischer, B. Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility, *J. Clin. Invest.* **1999**, *103*, pp. 11–18.
- (6) Hoerauf, A.; Volkmann, L.; Hamelmann, C.; Adjei, O.; Autenrieth, I. B.; Fleischer, B.; Büttner, D. W. Endosymbiotic bacteria in worms as targets for a novel chemotherapy in filariasis, *The Lancet.* **2000**, *355*, pp. 1242–1243.
- (7) Irschik, H.; Jansen, R.; Höfle, G.; Gerth, K.; Reichenbach, H. The corallopyronins, new inhibitors of bacterial RNA synthesis from *Myxobacteria*, *J. Antibiot.* **1985**, *38*, pp. 145–152.
- (8) Irschik, H.; Gerth, K.; Höfle, G.; Kohl, W.; Reichenbach, H. The myxopyronins, new inhibitors of bacterial RNA synthesis from *Myxococcus fulvus* (Myxobacterales), *J. Antibiot.* **1983**, *36*, pp. 1651–1658.
- (9) Ma, C.; Yang, X.; Lewis, P. J. Bacterial Transcription as a Target for Antibacterial Drug Development, *Microbiol. Mol. Biol. Rev.* **2016**, *80*, pp. 139–160.
- (10) Mukhopadhyay, J.; Das, K.; Ismail, S.; Koppstein, D.; Jang, M.; Hudson, B.; Sarafianos, S.; Tuske, S.; Patel, J.; Jansen, R.; Irschik, H.; Arnold, E.; Ebright, R. H. The RNA polymerase "switch region" is a target of inhibitors, *Cell.* **2008**, *135*, pp. 295–307.
- (11) Belogurov, G. A.; Vassilyeva, M. N.; Sevostyanova, A.; Appleman, J. R.; Xiang, A. X.; Lira, R.; Webber, S. E.; Klyuyev, S.; Nudler, E.; Artsimovitch, I.; Vassilyev, D. G. Transcription inactivation through local refolding of the RNA polymerase structure, *Nature.* **2008**.
- (12) Goldstein, B. P. Resistance to rifampicin: a review, *J. Antibiot.* **2014**, *67*, pp. 625–630.
- (13) Srivastava, A.; Talaue, M.; Liu, S.; Degen, D.; Ebright, R. Y.; Sineva, E.; Chakraborty, A.; Druzhinin, S. Y.; Chatterjee, S.; Mukhopadhyay, J.; Ebright, Y. W.; Zozula, A.; Shen, J.; Sen Gupta, S.; Niedfeldt, R. R.; Xin, C.; Kaneko, T.; Irschik, H.; Jansen, R.; Donadio, S.; Connell, N.; Ebright, R. H. New target for inhibition of bacterial RNA polymerase: 'switch region', *Curr. Opin. Microbiol.* **2011**, *14*, pp. 532–543.
- (14) Rentsch, A.; Kalesse, M. The total synthesis of corallopyronin A and myxopyronin B, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, pp. 11381–11384.

- (15) Sucipto, H.; Pogorevc, D.; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, *44*, pp. 160–170.
- (16) Schäberle, T. F.; Schmitz, A.; Zoicher, G.; Schiefer, A.; Kehraus, S.; Neu, E.; Roth, M.; Vassilyev, D. G.; Stehle, T.; Bierbaum, G.; Hoerauf, A.; Pfarr, K.; König, G. M. Insights into Structure-Activity Relationships of Bacterial RNA Polymerase Inhibiting Corallopyronin Derivatives, *J. Nat. Prod.* **2015**, *78*, pp. 2505–2509.
- (17) Sambrook, J.; Russell, D. W. *Molecular cloning: A laboratory manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.
- (18) Zhang, Y.; Buchholz, F.; Muyrers, J. P.; Stewart, F. A. A new logic for DNA engineering using recombination in *Escherichia coli*, *Nat. Genet.* **1998**, *20*, pp. 123–128.
- (19) Kashefi, K.; Hartzell, P. L. Genetic suppression and phenotypic masking of a *Myxococcus xanthus* *frzF*- defect, *Mol. Microbiol.* **1995**, *15*, pp. 483–494.
- (20) Gemperlein, K.; Rachid, S.; Garcia, R. O.; Wenzel, S. C.; Müller, R. Polyunsaturated fatty acid biosynthesis in myxobacteria. Different PUFA synthases and their product diversity, *Chem. Sci.* **2014**, *5*, pp. 1733–1741.
- (21) Stevens, D. C.; Hari, T. P. A.; Boddy, C. N. The role of transcription in heterologous expression of polyketides in bacterial hosts, *Nat. Prod. Rep.* **2013**, *30*, pp. 1391–1411.
- (22) Ongley, S.; Bian, X.; Neilan, B. A.; Müller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways, *Nat. Prod. Rep.* **2013**, *30*, pp. 1121–1138.
- (23) Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R. Heterologous expression of bacterial natural product biosynthetic pathways, *Nat. Prod. Rep.* **2019**.
- (24) Wenzel, S. C.; Gross, F.; Zhang, Y.; Fu, J.; Stewart, F. A.; Müller, R. Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via red/ET recombineering, *Chem. Biol.* **2005**, *12*, pp. 349–356.
- (25) Bryant, J. A.; Sellars, L. E.; Busby, S. J. W.; Lee, D. J. Chromosome position effects on gene expression in *Escherichia coli* K-12, *Nucleic Acids Res.* **2014**, *42*, pp. 11383–11392.
- (26) Bilyk, B.; Horbal, L.; Luzhetskyy, A. Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in *Streptomyces albus* J1074, *Microb. Cell. Fact.* **2017**, *16*, p. 5.
- (27) Perlova, O.; Gerth, K.; Kuhlmann, S.; Zhang, Y.; Müller, R. Novel expression hosts for complex secondary metabolite megasynthetases: Production of myxochromide in the thermophilic isolate *Coralloccoccus macrosporus* GT-2, *Microb. Cell Fact.* **2009**, *8*.
- (28) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (29) Sousa, C.; Lorenzo, V. de; Cebolla, A. Modulation of gene expression through chromosomal positioning in *Escherichia coli*, *Microbiology (Reading, England)*. **1997**, *143* ( Pt 6), pp. 2071–2078.
- (30) Erol, Ö.; Schäberle, T. F.; Schmitz, A.; Rachid, S.; Gurgui, C.; El Omari, M.; Lohr, F.; Kehraus, S.; Piel, J.; Müller, R.; König, G. M. Biosynthesis of the myxobacterial antibiotic corallopyronin A, *ChemBioChem.* **2010**, *11*, pp. 1235–1265.

- (31) Heathcote, M. L.; Staunton, J.; Leadlay, P. F. Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units, *Chem. Biol.* **2001**, *8*, pp. 207–220.
- (32) Kotowska, M.; Pawlik, K. Roles of type II thioesterases and their application for secondary metabolite yield improvement, *Appl. Microbiol. Biotechnol.* **2014**, *98*, pp. 7735–7746.
- (33) Sucipto, H.; Wenzel, S. C.; Müller, R. Exploring chemical diversity of  $\alpha$ -pyrone antibiotics: molecular basis of myxopyronin biosynthesis, *ChemBioChem.* **2013**, *14*, pp. 1581–1589.
- (34) Jensen, K.; Niederkrüger, H.; Zimmermann, K.; Vagstadt, A. L.; Moldenhauer, J.; Brendel, N.; Frank, S.; Pöplau, P.; Kohlhaas, C.; Townsend, C. A.; Oldiges, M.; Hertweck, C.; Piel, J. Polyketide proofreading by an acyltransferase-like enzyme, *Chem. Biol.* **2012**, *19*, pp. 329–339.
- (35) Beck, E.; Ludwig, G.; Auerswald, E. A.; Reiss, B.; Schaller, H. Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5, *Gene.* **1982**, *19*, pp. 327–336.
- (36) Iniesta, A. A.; García-Heras, F.; Abellón-Ruiz, J.; Gallego-García, A.; Elías-Arnanz, M. Two systems for conditional gene expression in *Myxococcus xanthus* inducible by isopropyl- $\beta$ -D-thiogalactopyranoside or vanillate, *J. Bacteriol.* **2012**, *194*, pp. 5875–5885.
- (37) Kim, B. S.; Cropp, T. A.; Beck, B. J.; Sherman, D. H.; Reynolds, K. A. Biochemical evidence for an editing role of thioesterase II in the biosynthesis of the polyketide pikromycin, *J. Biol. Chem.* **2002**, *277*, pp. 48028–48034.
- (38) Hoffmann, T.; Krug, D.; Hüttel, S.; Müller, R. Improving natural products identification through targeted LC-MS/MS in an untargeted secondary metabolomics workflow, *Anal. Chem.* **2014**, *86*, pp. 10780–10788.
- (39) Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. Functionally distinct modules operate two consecutive  $\alpha,\beta \rightarrow \beta,\gamma$  double-bond shifts in the rhizoxin polyketide assembly line, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, pp. 1460–1464.
- (40) Moldenhauer, J.; Götz, D. C. G.; Albert, C. R.; Bischof, S. K.; Schneider, K.; Süßmuth, R. D.; Engeser, M.; Gross, H.; Bringmann, G.; Piel, J. The Final Steps of Bacillaene Biosynthesis in *Bacillus amyloliquefaciens* FZB42: Direct Evidence for  $\beta,\gamma$  Dehydration by a trans-Acyltransferase Polyketide Synthase, *Angew. Chem. Int. Ed.* **2010**, *49*, pp. 1465–1467.
- (41) Keatinge-Clay, A. Crystal structure of the erythromycin polyketide synthase dehydratase, *J. Mol. Biol.* **2008**, *384*, pp. 941–953.
- (42) Kampa, A.; Gagunashvili, A. N.; Gulder, T. A. M.; Morinaka, B. I.; Daolio, C.; Godejohann, M.; Miao, V. P. W.; Piel, J.; Andrésson, Ó. S. Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E3129–37.
- (43) Lohr, F.; Jenniches, I.; Fritzler, M.; Meehan, M. J.; Sylvester, M.; Schmitz, A.; Gütschow, M.; Dorrestein, P. C.; König, G. M.; Schäberle, T. F.  $\alpha,\beta / \beta,\gamma$  double bond migration in corallopyronin A biosynthesis, *Chemical Science.* **2013**, *2013*, pp. 4175–4180.
- (44) Fu, C.; Auerbach, D.; Li, Y.; Scheid, U.; Luxenburger, E.; Garcia, R.; Irschik, H.; Müller, R. Solving the puzzle of the one-carbon loss in ripostatin biosynthesis, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, pp. 2192–2197.

(45) Wenzel, S. C.; Hoffmann, H.; Zhang, J.; Debussche, L.; Haag-Richter, S.; Kurz, M.; Nardi, F.; Lukat, P.; Kochems, I.; Tietgen, H.; Schummer, D.; Nicolas, J.-P.; Calvet, L.; Czepczor, V.; Vrignaud, P.; Mühlenweg, A.; Pelzer, S.; Müller, R.; Brönstrup, M. Production of the bengamide class of marine natural products in myxobacteria: biosynthesis and structure-activity relationships, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, pp. 15560–15564.

(46) Julien, B. Characterization of the integrase gene and attachment site for the *Myxococcus xanthus* bacteriophage Mx9, *J. Bacteriol.* **2003**, *185*, pp. 6325–6330.

### 3.7 Supplementary information

#### Table of contents

|          |   |     |
|----------|---|-----|
| 3.7.1    | Construction of a transposon-based COR BGC expression construct .....   | 104 |
| 3.7.2    | Deletion of <i>corM</i> , <i>corN</i> , <i>corO</i> , <i>corN-O</i> and <i>corP</i> biosynthetic genes from COR BGC expression construct by in vivo recombination in <i>E. coli</i> ..... | 105 |
| 3.7.3    | Assembly of CorO, CorP, CorM and MxnM expression/co-expression constructs.....  | 105 |
| 3.7.4    | Coralopyronin MIC testing against <i>M. xanthus</i> DK1622, <i>E. coli</i> $\Delta$ TolC and mutants thereof .....  | 106 |
| 3.7.5    | Alignment of the conserved core motives of classical DH domains with the two DH* domains from the coralopyronin A BGC and several other DH* domains.....                                  | 107 |
| 3.7.6    | Standardized HPLC-MS method for sample analysis and application of principal component analysis .....   | 108 |
| 3.7.7    | HPLC-MS method for quantification of (pre)Cor A production yields.....  | 108 |
| 3.7.8    | Statistical data treatment using the principal component analysis.....  | 109 |
| 3.7.9    | Compound isolation and characterization .....   | 110 |
| 3.7.9.1  | Isolation of preCOR A.....  | 110 |
| 3.7.9.2  | NMR based characterization of preCOR A.....   | 111 |
| 3.7.9.3  | Isolation of COR D.....   | 112 |
| 3.7.9.4  | NMR based characterization of COR D.....  | 113 |
| 3.7.9.5  | Isolation of oxyCOR A.....  | 114 |
| 3.7.9.6  | NMR based characterization of oxyCOR A.....   | 115 |
| 3.7.10   | Investigation of the stereochemistry of COR D and oxyCOR A.....   | 116 |
| 3.7.10.1 | <i>In silico</i> evaluation of the ketoreductase stereochemistry .....  | 116 |
| 3.7.10.2 | Experimental evidence for an S,S-diol configuration in oxyCOR A.....  | 117 |
| 3.7.11   | 1D and 2D NMR spectra.....  | 118 |
| 3.7.11.1 | NMR spectra of preCOR A.....  | 118 |
| 3.7.11.2 | NMR spectra of COR D.....   | 120 |
| 3.7.11.3 | NMR spectra of oxyCOR A.....  | 122 |
| 3.7.12   | References .....  | 125 |

**List of figures**

|   |     |
|---|-----|
| Figure S1. Alignment of DH and DH' amino acid sequences. ....   | 107 |
| Figure S2. Genotypic verification of <i>M. xanthus</i> DK1622 heterologous expression strains. ....   | 110 |
| Figure S3. Structure formula of preCOR A with atom numbering according to the NMR data. ....  | 111 |
| Figure S4. Structure formula of COR D with atom numbering according to the NMR data. ....   | 113 |
| Figure S5. Structure formula of oxyCOR A with atom numbering according to the NMR data. ....  | 115 |
| Figure S6. Alignment of important KR motifs. The alignment is reproduced from the original alignment by Keatinge-Clay with example sequences for two different type A and type B KRs (Keatinge-Clay, 2007). The LDD motif, typically found in A type KRs, is highlighted in green. .... | 116 |
| Figure S7. Structure formula of R-corallopyronin A. ....  | 117 |
| Figure S8. Pseudo <i>S,S</i> form of oxyCOR A. ....   | 117 |
| Figure S9. Pseudo <i>meso</i> form of oxyCOR A. ....  | 117 |
| Figure S10. <i>R</i> -form of COR D. ....   | 118 |
| Figure S11. <sup>1</sup> H NMR spectrum of preCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 118 |
| Figure S12. <sup>1</sup> H NMR spectrum of preCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 119 |
| Figure S13. HSQC NMR spectrum of preCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 119 |
| Figure S14. HMBC NMR spectrum of preCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 120 |
| Figure S15. <sup>1</sup> H NMR spectrum of COR D in methanol- <i>d</i> <sub>4</sub> . ....  | 120 |
| Figure S16. COSY NMR spectrum of COR D in methanol- <i>d</i> <sub>4</sub> . ....  | 121 |
| Figure S17. HSQC NMR spectrum of COR D in methanol- <i>d</i> <sub>4</sub> . ....  | 121 |
| Figure S18. HMBC NMR spectrum of COR D in methanol- <i>d</i> <sub>4</sub> . ....  | 122 |
| Figure S19. <sup>1</sup> H NMR spectrum of oxyCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 122 |
| Figure S20. COSY NMR spectrum of oxyCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 123 |
| Figure S21. HSQC NMR spectrum of oxyCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 123 |
| Figure S22. HMBC NMR spectrum of oxyCOR A in methanol- <i>d</i> <sub>4</sub> . ....   | 124 |

**List of tables**

|  |     |
|--|-----|
| Table S1. List of primers used in this study. ....   | 99  |
| Table S2. Gene synthesis constructs obtained from ATG:biosynthetics GmbH. ....                         | 100 |
| Table S3. Plasmids and expression constructs generated in this study. ....                             | 101 |
| Table S4. List of all observed NMR correlations for preCOR A in methanol- <i>d</i> <sub>4</sub> . .... | 111 |
| Table S5. List of all observed NMR correlations for COR D in methanol- <i>d</i> <sub>4</sub> . ....    | 113 |
| Table S6. List of all observed NMR correlations for oxyCOR A in methanol- <i>d</i> <sub>4</sub> . .... | 115 |

**Table S1.** List of primers used in this study.

| Primer name       | Sequence (5'→3')<br>(restriction sites in bold, homology arms underlined)                  | Restriction sites |
|-------------------|--|-------------------|
| dpo-nptII-corO-F  | GATACAC <b>ATAT</b> GGAGGAGTGCACCATGAGTAGCC  | <i>NdeI</i>       |
| dpo-nptII-corO-R1 | CTATGT <b>ATGCAT</b> GGACGCCTCGCTGAATC   | <i>NsiI</i>       |
| dpo-corP-F        | GATACAC <b>ATAT</b> GTTGCTTGGGGAGATGTCATGACG   | <i>NdeI</i>       |
| dpo-corP-R        | CTATGT <b>ATGCAT</b> CCCACGCCGTCCATCAACTC  | <i>NsiI</i>       |
| dpo-corP-PstI-F   | GATAC <b>ACTGCAGT</b> GGACAGCAAGCGAACCGGAATTG  | <i>PstI</i>       |
| dpo-corP-PstI-R   | GATAC <b>ACTGCAG</b> CCCACGCCGTCCATCAACTC  | <i>PstI</i>       |
| dpo-ntpII-mxnM-F  | GATACAC <b>ATAT</b> GGGAGAAGGACGGGCG   | <i>NdeI</i>       |
| dpo-mxnM-R        | CTATGT <b>ATGCAT</b> GGGTCGATCACGTCAGAC  | <i>NsiI</i>       |
| dpo-Pvan-mxnM-F   | GATACAC <b>ATAT</b> GAGCGCTGAGACGGTCTTC  | <i>NdeI</i>       |
| dpo-ntpII-corM-F  | GATACAC <b>ATAT</b> GGAAAGCGGGCGGACG   | <i>NdeI</i>       |
| dpo-ntpII-corM-R  | CTATGT <b>ATGCAT</b> CCTCAGCTCCCCGTCCAG  | <i>NsiI</i>       |
| dpo-Pvan-corM-F   | GATACAC <b>ATAT</b> GATTGATCCGCGTTGG   | <i>NdeI</i>       |
| dpo-mx8-PacI-F    | GATAC <b>ATTAATTA</b> ACCAAGGCAGAGTTGGTG   | <i>PacI</i>       |
| dpo-mx8-PciI-R    | GATACA <b>ACATGT</b> CCCCACTGCAAGCTACCTG   | <i>PciI</i>       |
| dpo-mx8-1         | GGACGCAGGCTCCAAGTC   |                   |
| dpo-mx8-2         | GACTCGTGGACTGCAATTG  |                   |
| dpo-mx8-3         | GTCCTGAAGGCGACCGAG   |                   |
| dpo-mx8-4         | CACGATGCTTGCCAGTTG   |                   |
| dpo-mx8-5         | GTGGTGCCTCAGCAGATG   |                   |
| dpo-CorM-F        | TAGAAAGCCAGTCCGCAG   |                   |
| dpo-cor-TD1-R2    | TGCTGGCCTTTTGCTCAC   |                   |
| dpo-apra-corM-F   | <u>GCTTCGCCGAAGGACCGGCCACCGGAAAGGACGGGCGG</u><br><u>AGTTGACCTAGGCTCACGGTAACTGATGCC</u>     | <i>XmaI</i>       |
| dpo-apra-corM-R   | <u>GACGCGTTGGATGCGGGAGACCATGGGAGCCCCTCAGCT</u><br><u>CCCCGTCTAGGTCAGCCAATCGACTGGCGAGC</u>  | <i>XmaI</i>       |
| dpo-apra-corN-F   | <u>CTCCAGCTCTCCGTCAGGTCTCCAGCCCTGGACGGGGA</u><br><u>GCTGACCTAGGCTCACGGTAACTGATGCC</u>      | <i>XmaI</i>       |
| dpo-apra-corN-R   | <u>GAAGACCTCGGGTTCGTTCAAGGCCTGCCGTCCGCGCGG</u><br><u>GACGGCCCTAGGTCAGCCAATCGACTGGCGAGC</u> | <i>XmaI</i>       |
| dpo-apra-corO-F   | <u>CGGATGCAGTCGCTGTTGATGCACCGGCGCGTGCTGCGC</u><br><u>AACTGACCTAGGCTCACGGTAACTGATGCC</u>    | <i>XmaI</i>       |

|                 |  |               |
|-----------------|--|---------------|
| dpo-apra-R      | TAAACGGGACGCCTCGCTGAATCCGGCGCGGCGTCCCCG<br>GCTTCGCCCTAGGTCAGCCAATCGACTGGCGAGC              | <i>Xma</i> II |
| dpo-apra-orf1-F | <u>CCGGCCCCCGCCCGGAATGGCATCCACGAACGAGAGTCG</u><br><u>GCGTGACCTAGGCTCACGGTAACTGATGCC</u>    | <i>Xma</i> II |
| dpo-apra-orf1-R | <u>CGCGGATCAATCACCGTCCGCCCCGCTTCGCTCCCACGCC</u><br><u>GTCCACCTAGGTCAGCCAATCGACTGGCGAGC</u> | <i>Xma</i> II |
| P5              | TATCCGTTTCGCGAGGAAG  |               |
| P6              | TGCTGCGGATCCATGTTC   |               |
| P7              | ATCCCGGAGACGCTGAAG   |               |
| P8              | GCCTGCGTCGCTCTTCAC   |               |
| Mx8-attB-up2    | GCGCACTGGACCATCACGTC   |               |
| Mx8-attP-down   | GGCTTGTGCCAGTCAACTGCG  |               |
| Mx8-attP-up2    | CGACGGTGCCGACAAATAC  |               |
| Mx8-attB-down   | CGGATAGCTCAGCGGTAGAG   |               |
| dpo-puc-tet-F   | GGTTTTCCAGTCACGAC  |               |
| dpo-mx9-corP-1  | GCGGGGTGGCAAATGAGC   |               |
| dpo-mx9-corP-2  | GACGGCCAGCTTGAACAG   |               |

**Table S2.** Gene synthesis constructs obtained from ATG:biosynthetics GmbH

| Construct name               | Fragment size | Description   |
|------------------------------|---------------|---|
| pUC57-TD1-cm                 | 1031 bp       | <i>Pac</i> I-H <sub>L</sub> -TD1- <i>Bst</i> 1107I-cm <sup>R</sup> - <i>Bst</i> 1107I-H <sub>R</sub> - <i>Pci</i> I   |
| pUC57-kan-IR-Tpase           | 2684 bp       | <i>Pac</i> I-H <sub>L</sub> -TD1- <i>Bst</i> 1107I-kan <sup>R</sup> -IR-Tpase- <i>Bst</i> 1107I-H <sub>R</sub> - <i>Pci</i> I   |
| pUC57-Pvan-vanR-IR-ampR-oriT | 3085 bp       | <i>Kpn</i> I-H <sub>L</sub> -oriT- <i>Bgl</i> II- <i>Stu</i> I-amp <sup>R</sup> - <i>Nru</i> I- <i>Sal</i> I-IR- <i>Ksp</i> AI- <i>Xho</i> I- <i>P<sub>van</sub></i> - <i>Nde</i> I-H <sub>R</sub> - <i>Pme</i> I |
| pUC57-Pvan-corO              | 1459 bp       | <i>Pci</i> I- <i>P<sub>van</sub></i> -corO- <i>Nsi</i> II   |

**Table S3.** Plasmids and expression constructs generated in this study.

| Plasmid name                        | Construction details   | Features  |
|-------------------------------------|--|---|
| pDPO-mxn116-Cm                      | <i>kan<sup>R</sup>-tet<sup>R</sup></i> cassette of pDPO-mxn116 replaced with 1031 bp fragment from pUC57-TD1-cm by <i>in vivo</i> recombination.   | <i>P<sub>nptII</sub>-mxnA-J-corKL-</i><br><i>corP-corM-O</i> , <i>ori2</i> ,<br><i>sopA</i> , <i>sopB</i> , <i>sopC</i> ,<br><i>repE</i> , <i>cm<sup>R</sup></i>  |
| pDPO-mxn116-Cm-Pvan                 | <i>P<sub>nptII</sub></i> of pDPO-mxn116-Cm replaced with 3085 bp fragment from pUC57-Pvan-vanR-IR-amp <sup>R</sup> -oriT by <i>in vivo</i> recombination.  | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corP-corM-O</i> , <i>ori2</i> ,<br><i>sopA</i> , <i>sopB</i> , <i>sopC</i> ,<br><i>repE</i> , <i>cm<sup>R</sup></i> , <i>amp<sup>R</sup></i> , <i>oriT</i>   |
| pDPO-mxn116-Pvan-Tpase              | <i>cm<sup>R</sup></i> of pDPO-mxn116-Cm-Pvan replaced with 2684 bp fragment from pUC57-kan-IR-Tpase by <i>in vivo</i> recombination.   | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corP-corM-O</i> , <i>ori2</i> ,<br><i>sopA</i> , <i>sopB</i> , <i>sopC</i> ,<br><i>repE</i> , <i>cm<sup>R</sup></i> , <i>amp<sup>R</sup></i> ,<br><i>oriT</i> , <i>tpase</i>                                 |
| pDPO-Mxn116-Pvan-TpaseΔcorM-apraR   | <i>corM</i> of pDPO-mxn116-Cm-Pvan-Tpase replaced with 1082 bp <i>apra<sup>R</sup></i> fragment amplified from pOJ446 <sup>1</sup> using primers dpo-apra-CorM-F/dpo-apra-CorM-R, by <i>in vivo</i> recombination.   | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corP-corN-O</i> , <i>ori2</i> ,<br><i>sopA</i> , <i>sopB</i> , <i>sopC</i> ,<br><i>repE</i> , <i>cm<sup>R</sup></i> , <i>amp<sup>R</sup></i> ,<br><i>oriT</i> , <i>tpase</i> , <i>apra<sup>R</sup></i>       |
| pDPO-Mxn116-Pvan-TpaseΔcorN-apraR   | <i>corN</i> of pDPO-mxn116-Cm-Pvan-Tpase replaced with 1082 bp <i>apra<sup>R</sup></i> fragment amplified from pOJ446 <sup>1</sup> using primers dpo-apra-CorM-F/dpo-apra-CorM-R, by <i>in vivo</i> recombination.   | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corP-corM-corO</i> ,<br><i>ori2</i> , <i>sopA</i> , <i>sopB</i> ,<br><i>sopC</i> , <i>repE</i> , <i>cm<sup>R</sup></i> ,<br><i>amp<sup>R</sup></i> , <i>oriT</i> , <i>tpase</i> ,<br><i>apra<sup>R</sup></i> |
| pDPO-Mxn116-Pvan-TpaseΔcorO-apraR   | <i>corO</i> of pDPO-mxn116-Cm-Pvan-Tpase replaced with 1082 bp <i>apra<sup>R</sup></i> fragment amplified from pOJ446 <sup>1</sup> using primers dpo-apra-CorM-F/dpo-apra-CorM-R, by <i>in vivo</i> recombination.   | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corP-corM-N</i> , <i>ori2</i> ,<br><i>sopA</i> , <i>sopB</i> , <i>sopC</i> ,<br><i>repE</i> , <i>cm<sup>R</sup></i> , <i>amp<sup>R</sup></i> ,<br><i>oriT</i> , <i>tpase</i> , <i>apra<sup>R</sup></i>       |
| pDPO-Mxn116-Pvan-TpaseΔcorN-O-apraR | <i>corN-O</i> of pDPO-mxn116-Cm-Pvan-Tpase replaced with 1082 bp <i>apra<sup>R</sup></i> fragment amplified from pOJ446 <sup>1</sup> using primers dpo-apra-CorM-F/dpo-apra-CorM-R, by <i>in vivo</i> recombination. | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corP-corM</i> , <i>ori2</i> ,<br><i>sopA</i> , <i>sopB</i> , <i>sopC</i> ,<br><i>repE</i> , <i>cm<sup>R</sup></i> , <i>amp<sup>R</sup></i> ,<br><i>oriT</i> , <i>tpase</i> , <i>apra<sup>R</sup></i>         |
| pDPO-Mxn116-Pvan-TpaseΔcorP-apraR   | <i>corP</i> of pDPO-mxn116-Cm-Pvan-Tpase replaced with 1082 bp <i>apra<sup>R</sup></i> fragment amplified from pOJ446 <sup>1</sup> using primers dpo-apra-CorM-F/dpo-apra-CorM-R, by <i>in vivo</i> recombination.   | <i>P<sub>van</sub>-mxnA-J-corKL-</i><br><i>corM-O</i> , <i>ori2</i> , <i>sopA</i> ,<br><i>sopB</i> , <i>sopC</i> , <i>repE</i> ,<br><i>cm<sup>R</sup></i> , <i>amp<sup>R</sup></i> , <i>oriT</i> ,<br><i>tpase</i> , <i>apra<sup>R</sup></i>            |

|  |   |   |
|--|---|---|
| pDPO-Mxn116-Pvan-Tpase $\Delta$ corM   | pDPO-Mxn116-Pvan-Tpase $\Delta$ corM-apraR hydrolyzed with <i>Xma</i> II and re-ligated (to eliminate <i>apra</i> <sup>R</sup> cassette).   | <i>P<sub>van-mxnA-J-corKL</sub></i> - <i>corP-corN-O</i> , ori2, <i>sopA</i> , <i>sopB</i> , <i>sopC</i> , <i>repE</i> , <i>cm</i> <sup>R</sup> , <i>amp</i> <sup>R</sup> , <i>oriT</i> , <i>tpase</i>    |
| pDPO-Mxn116-Pvan-Tpase $\Delta$ corN   | pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-apraR hydrolyzed with <i>Xma</i> II and re-ligated (to eliminate <i>apra</i> <sup>R</sup> cassette).   | <i>P<sub>van-mxnA-J-corKL</sub></i> - <i>corP-corM-corO</i> , ori2, <i>sopA</i> , <i>sopB</i> , <i>sopC</i> , <i>repE</i> , <i>cm</i> <sup>R</sup> , <i>amp</i> <sup>R</sup> , <i>oriT</i> , <i>tpase</i> |
| pDPO-Mxn116-Pvan-Tpase $\Delta$ corO   | pDPO-Mxn116-Pvan-Tpase $\Delta$ corO-apraR hydrolyzed with <i>Xma</i> II and re-ligated (to eliminate <i>apra</i> <sup>R</sup> cassette).   | <i>P<sub>van-mxnA-J-corKL</sub></i> - <i>corP-corM-N</i> , ori2, <i>sopA</i> , <i>sopB</i> , <i>sopC</i> , <i>repE</i> , <i>cm</i> <sup>R</sup> , <i>amp</i> <sup>R</sup> , <i>oriT</i> , <i>tpase</i>    |
| pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O | pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O-apraR hydrolyzed with <i>Xma</i> II and re-ligated (to eliminate <i>apra</i> <sup>R</sup> cassette).   | <i>P<sub>van-mxnA-J-corKL</sub></i> - <i>corP-corM</i> , ori2, <i>sopA</i> , <i>sopB</i> , <i>sopC</i> , <i>repE</i> , <i>cm</i> <sup>R</sup> , <i>amp</i> <sup>R</sup> , <i>oriT</i> , <i>tpase</i>      |
| pDPO-Mxn116-Pvan-Tpase $\Delta$ corP   | pDPO-Mxn116-Pvan-Tpase $\Delta$ corP-apraR hydrolyzed with <i>Xma</i> II and re-ligated (to eliminate <i>apra</i> <sup>R</sup> cassette).   | <i>P<sub>van-mxnA-J-corKL</sub></i> - <i>corM-O</i> , ori2, <i>sopA</i> , <i>sopB</i> , <i>sopC</i> , <i>repE</i> , <i>cm</i> <sup>R</sup> , <i>amp</i> <sup>R</sup> , <i>oriT</i> , <i>tpase</i>         |
| pUC18-Zeo-mx8-MCS                      | 3185 bp <i>PacI-mx8-PciI</i> PCR fragment amplified from <i>M. xanthus</i> DK1622 genomic DNA by dpo-mx8- <i>PacI-F</i> /dpo-mx8- <i>PciI-R</i> primers, ligated into pUC18-Zeo-Hom-MCS hydrolyzed with the same enzymes. | <i>P<sub>nptII</sub></i> , MCS, pUC ori, <i>mx8</i> , <i>zeo</i> <sup>R</sup>   |
| pUC18-Zeo-mx8-nptII-corO               | 1426 bp <i>NdeI-nptII-corO-NsiI</i> PCR fragment amplified from pDPO-Mxn116 <sup>2</sup> by dpo-ntpII-corO- <i>F</i> /dpo-ntpII-corO- <i>R1</i> primers, ligated into pUC18-Zeo-Hom-MCS hydrolyzed with the same enzymes. | <i>P<sub>nptII-corO</sub></i> , pUC ori, <i>mx8</i> , <i>zeo</i> <sup>R</sup>   |
| pUC18-Zeo-mx8-Pvan-corO                | 1426 bp <i>NdeI-Pvan-corO-NsiI</i> PCR fragment amplified from pDPO-Mxn116 <sup>2</sup> by dpo-ntpII-corO- <i>F</i> /dpo-ntpII-corO- <i>R1</i> primers, ligated into pUC18-Zeo-Hom-MCS hydrolyzed with the same enzymes.  | <i>P<sub>van-corO</sub></i> , pUC ori, <i>mx8</i> , <i>zeo</i> <sup>R</sup>   |
| pUC18-Zeo-mx8-nptII-corM               | 794 bp <i>NdeI-nptII-corM-NsiI</i> PCR fragment amplified from pDPO-Mxn116 <sup>2</sup> by dpo-ntpII-   | <i>P<sub>nptII-corM</sub></i> , pUC ori, <i>mx8</i> , <i>zeo</i> <sup>R</sup>   |

|                          |  |  |
|--------------------------|--|--|
| pUC18-Zeo-mx8-Pvan-corM  | corM-F/dpo-ntpII-corM-R primers, ligated into pUC18-Zeo-Hom-MCS hydrolyzed with the same enzymes.<br>776 bp <i>NdeI</i> -Pvan-corM- <i>NsiI</i> PCR fragment amplified from pDPO-Mxn116 <sup>2</sup> by dpo-Pvan-corM-F/dpo-ntpII-corM-R primers, ligated into pUC18-Zeo-mx8-Pvan-corO hydrolyzed with the same enzymes. | <i>P<sub>van-corM</sub></i> , pUC ori, <i>mx8</i> , <i>zeo<sup>R</sup></i>   |
| pUC18-Zeo-mx8-nptII-mxnM | 1038 bp <i>NdeI</i> -nptII-mxnM- <i>NsiI</i> PCR fragment amplified from pHSU-mxn43 <sup>2</sup> by dpo-ntpII-mxnM-F/dpo-mxnM-R primers, ligated into pUC18-Zeo-mx8-nptII-corO hydrolyzed with the same enzymes.   | <i>P<sub>nptII-mxnM</sub></i> , pUC ori, <i>mx8</i> , <i>zeo<sup>R</sup></i> |
| pUC18-Zeo-mx8-Pvan-mxnM  | 1014 bp <i>NdeI</i> -Pvan-corM- <i>NsiI</i> PCR fragment amplified from pHSU-mxn43 <sup>2</sup> by dpo-Pvan-corM-F/dpo-ntpII-corM-R primers, ligated into pUC18-Zeo-mx8-Pvan-corO hydrolyzed with the same enzymes.  | <i>P<sub>van-mxnM</sub></i> , pUC ori, <i>mx8</i> , <i>zeo<sup>R</sup></i>   |
| pUC18-Zeo-mx8-nptII-corP | 1776 bp <i>NdeI</i> -nptII-corP- <i>NsiI</i> PCR fragment amplified from pDPO-mxn116-Pvan-Tpase by dpo-corP-F/dpo-corP-R primers, ligated into pUC18-Zeo-mx8-nptII-corO hydrolyzed with the same enzymes.  | <i>P<sub>nptII-corP</sub></i> , pUC ori, <i>mx8</i> , <i>zeo<sup>R</sup></i> |

### 3.7.1 Construction of a transposon-based COR BGC expression construct

Since the choice of known selection markers in *M. xanthus* was limited, the existing *tet<sup>R</sup>* and *kan<sup>R</sup>* genes were removed from the pDPO-mxn116 expression construct in order to later employ them in the downstream modification steps. Both genes were eliminated from the pDPO-mxn116 backbone in one *in vivo*, Red/ET recombination step. For this purpose a ~1 kb synthetic fragment pUC57-TD1-cm (Table S2), harboring TD1 terminator<sup>3</sup>, followed by a *cm<sup>R</sup>* gene, was designed and obtained from a gene synthesis company. The fragment was flanked by 50 bp homology regions (H<sub>L</sub> and H<sub>R</sub>, see Fig. 1), homologous to the *tet<sup>R</sup>* - *kan<sup>R</sup>* cassette in the pDPO-mxn116 construct and by unique restriction sites *PacI* and *PciI* to allow its excision from the pGH vector in which it was supplied. After the excision the linear fragment could directly be used for *in vivo* recombination. By employing previously described *E. coli* GB05-red (chapter 3.3.2), Red/ET recombination strain, *tet<sup>R</sup>* - *kan<sup>R</sup>* cassette was successfully exchanged with *cm<sup>R</sup>* gene, resulting in construction of pDPO-mxn116-cm construct (Fig. 1).

To improve COR production by *M. xanthus* heterologous host, previously generated pDPO-mxn116-cm construct was further modified to replace the constitutive P<sub>*nptII*</sub> with the P<sub>*van*</sub> promoter. To facilitate this process, a ~3 kb synthetic fragment pUC57-Pvan-vanR-IR-AmpR-oriT (Table S2) was designed harboring P<sub>*van*</sub> promoter, *amp<sup>R</sup>* gene, oriT and one inverted repeat (IR) element fragment, to be employed for genome integration, mediated by the transposase. The transposase and a second IR element, required for transposase-mediated chromosomal BGC integration<sup>4</sup>, was planned to be introduced in downstream modification steps. The fragment was flanked by suitable 50 bp homology regions and unique *KpnI* and *PmeI* R-sites, allowing its excision from the standard pGH vector in which it was supplied, in order to directly use the linear fragment for *in vivo* recombination. By performing one Red/ET recombination step, pDPO-mxn116-cm-Pvan construct was generated (Fig. 1), which had to be further modified in order to be able to integrate in *M. xanthus* genome, since the previous transfer cassette based on homologous recombination via *tet<sup>R</sup>* was already deleted.

In the last step a previously generated pDPO-mxn116-cm-Pvan construct was further modified by Red/ET to introduce the transposase gene on its backbone. To facilitate this process, a ~2,7 kb synthetic fragment pUC57-Kan-IR-Tpase was designed (Table S2), harboring TD1 terminator followed by a *kan<sup>R</sup>* gene, IR and the transposase gene. As previously mentioned, a second IR element was included on this construct in order to allow chromosomal integration via transposition. The fragment was flanked by suitable 50 bp homologies and *PacI* and *PciI* R-sites, allowing its excision from the standard pGH vector in

which it was supplied. The resulting construct pDPO-mxn116-Pvan-Tpase now harbored all the necessary elements for heterologous expression in *M. xanthus* DK1622 heterologous host (Fig. 1). All synthetic fragments were provided by ATG:biosynthetics GmbH and are provided in Table S2.

### 3.7.2 Deletion of *corM*, *corN*, *corO*, *corN-O* and *corP* biosynthetic genes from COR BGC expression construct by *in vivo* recombination in *E. coli*

A 804 bp long *apra<sup>R</sup>* gene was amplified from the pOJ446<sup>1</sup> plasmid using Phusion polymerase, by suitable primer pairs: *corM* (dpo-apra-corM-F/dpo-apra-corM-R), *corN* (dpo-apra-corN-F/dpo-apra-corN-R), *corO* (dpo-apra-corO-F/dpo-apra-R), *corN-O* (dpo-apra-corN-F/dpo-apra-R) and *corP* (dpo-apra-orf1-F/dpo-apra-orf1-R). The *apra<sup>R</sup>* gene was flanked by *Xma*II R-sites and 45 bp long homology regions, homologous to the corresponding deletion region in the *mxn/cor* BGC, provided on the primers. By single *in vivo* recombination step in *E. coli* GB05-red, the corresponding gene was replaced by the *apra<sup>R</sup>* gene and correct *E. coli* clones were selected on LB medium supplemented with apramycin. Resulting plasmids (pDPO-Mxn116-Pvan-Tpase $\Delta$ corM-apraR, pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-apraR, pDPO-Mxn116-Pvan-Tpase $\Delta$ corO-apraR, pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O-apraR, pDPO-Mxn116-Pvan-Tpase $\Delta$ corP-apraR) were hydrolyzed by *Xma*II restriction endonuclease and re-ligated, to excise *apra<sup>R</sup>*. Sequences of the final constructs (pDPO-Mxn116-Pvan-Tpase $\Delta$ corM, pDPO-Mxn116-Pvan-Tpase $\Delta$ corN, pDPO-Mxn116-Pvan-Tpase $\Delta$ corO, pDPO-Mxn116-Pvan-Tpase $\Delta$ corN-O and pDPO-Mxn116-Pvan-Tpase $\Delta$ corP) were confirmed by Illumina sequencing, before transformation to *M. xanthus* DK1622. All primer sequences are provided in the Table S1.

### 3.7.3 Assembly of CorO, CorP, CorM and MxnM expression/co-expression constructs

Phusion polymerase was used to amplify 1426 bp long nptII-*corO* fragment (dpo-nptII-corO-F/dpo-nptII-corO-R1) from pDPO-mxn116 plasmid, 1776 bp long nptII-*corP* fragment (dpo-corP-F/dpo-corP-R) from pDPO-mxn116-Pvan-Tpase template and 1921 bp long *P<sub>nptII</sub>-corP* fragment (dpo-corP-PstI-F/ dpo-corP-PstI-R) from pUC18-Zeo-mx8-nptII-corP. DreamTaq polymerase was used to amplify 1038 bp long nptII-*mxnM* fragment (dpo-ntpII-mxnM-F/dpo-mxnM-R) and 1014 bp long Pvan-*mxnM* fragment (dpo-Pvan-mxnM-F/dpo-mxnM-R) from pHSU-Mxn43 template. The 794 bp nptII-*corM* fragment (dpo-ntpII-corM-F/dpo-ntpII-corM-R) and 776 bp Pvan-*corM* fragment (dpo-Pvan-corM-F/dpo-ntpII-corM-R) were amplified by

Taq polymerase from pDPO-mxn116 template. Primer names used for PCR reactions are provided in the brackets. Additionally, 1459 bp long Pvan-*corO* fragment was ordered from a gene synthesis company (Table S2). To express those fragments in *M. xanthus* DK1622 heterologous host, a suitable expression vector had to be constructed. The 3185 bp long *mx8* phage integrase gene was PCR amplified from *M. xanthus* DK1622 genomic DNA by dpo-mx8-PacI-F/dpo-mx8-PciI-R primer pair, using Phusion polymerase. The fragment was hydrolyzed with *PacI* and *PciI* restriction endonucleases and stitched into a previously constructed expression vector pUC18-Zeo-Hom-MCS (Pogorevc et al., unpublished) to construct pUC18-Zeo-mx8-MCS expression construct. Sequence of the *mx8* phage integrase gene was verified by Sanger sequencing using dpo-mx8-1, dpo-mx8-2, dpo-mx8-3, dpo-mx8-4 and dpo-mx8-5 primers. The *nptII-corO* and Pvan-*corO* fragments were digested with suitable restriction endonucleases (*NdeI/NsiI* pair for *nptII-corO* and *PciI/NsiI* pair for Pvan-*corO*) and stitched into pUC18-Zeo-mx8-MCS construct, resulting in construction of pUC18-Zeo-mx8-nptII-corO and pUC18-Zeo-mx8-Pvan-corO expression plasmids. The pUC18-Zeo-mx8-nptII-corO and pUC18-Zeo-mx8-Pvan-corO were further used to stitch *nptII-corM* and Pvan-*corM* fragments by *NdeI/NsiI* restriction endonucleases, yielding pUC18-Zeo-mx8-nptII-corM and pUC18-Zeo-mx8-Pvan-corM expression plasmids respectively. The same method was used for introduction of *nptII-mxnM* and Pvan-*mxnM* fragments into pUC18-Zeo-mx8-nptII-corO and pUC18-Zeo-mx8-Pvan-corO using *NdeI/NsiI* restriction endonucleases, yielding pUC18-Zeo-mx8-nptII-mxnM and pUC18-Zeo-mx8-Pvan-mxnM expression plasmids, respectively. The *nptII-corP* fragment was stitched into pUC18-Zeo-mx8-nptII-corO using *NdeI/NsiI* restriction endonucleases, yielding pUC18-Zeo-mx8-nptII-corP expression plasmid. The *P<sub>nptII</sub>-corP* fragment was stitched into pUC18-TetR-mx9 (Schillinger and Pogorevc et al., unpublished) by *PstI* restriction endonuclease, yielding pUC18-TetR-mx9-nptII-corP expression plasmid. Inserted gene fragments were verified by Sanger sequencing using primers dpo-mx9-corP-1, dpo-mx9-corP-2 and dpo-puc-tet-F in case of pUC18-TetR-mx9-nptII-corP and primers dpo-CorM-F and dpo-cor-TD1-R2 in all other cases. All primer sequences are provided in the Table S1.

#### **3.7.4 Coralopyronin MIC testing against *M. xanthus* DK1622, *E. coli* $\Delta$ TolC and mutants thereof**

*M. xanthus* strains were inoculated from cryo stocks and grown on CTT agar supplemented with suitable antibiotics at 30 °C. A patch of cells was scraped to inoculate 20 mL CTT medium supplemented with suitable antibiotic. Culture was grown at 30 °C and 180 rpm for 1

to 2 days until OD<sub>600</sub> of around 1.0 was reached. Culture was diluted to OD<sub>600</sub> 0,1 before performing the MIC test. *E. coli* strains were inoculated from cryo stocks directly into 20 mL LB medium supplemented with suitable antibiotic. Culture was grown at 37 °C and 180 rpm. Overnight culture was diluted to OD<sub>600</sub> 0,01 before performing the MIC test. MIC test was performed in 96 well plates by adding 19.2 µL of COR A solution in ACN with concentration of 1 mg/mL into one well in the first row. Plate was left open under the laminar flow until the ACN evaporated, before addition of 150 µL of diluted culture to all wells. Additional 150 µL of diluted culture were added to all wells in the first row and mixed with multi pipet. After mixing 150 µL of culture was transferred to the next row of wells and again mixed by pipetting. Procedure was repeated until the last row after which the last 150 µL of culture were discarded. The plate was incubated at 30 °C with shaking at 180 rpm overnight in case of *E. coli* and 48 hours in case of *M. xanthus*, before MIC was determined by visual evaluation.

### 3.7.5 Alignment of the conserved core motives of classical DH domains with the two DH\* domains from the corallopyronin A BGC and several other DH\* domains

The conserved motives HxxxGxxxxP and DxxxQ/H are essential for the catalytic activity of a DH as a reductive enzyme<sup>5</sup>. These motives are usually mutated in DH\* domains which are not responsible for the dehydration of the hydroxyl group but rather for the isomerization of the double bond<sup>5</sup>.

In the N-terminal DH domain from CorL of the corallopyronin BGC the first core motif HxxxGxxxxP is still present, whereas the second motif lacks both conserved residues (Fig. S1), indicating that this DH is unable to exhibit dehydratase activity, but might act as “shifting DH”.

|           |                    |                |
|-----------|--------------------|----------------|
| CorL_DH*  | ... HEVFGRPLFPT... | ... NGLLM...   |
| CorJ_DH*  | ... HTVLGQRVLLG... | ... DGVIV...   |
| NspC_DH*  | ... HTLLGDRVLLG... | ... NSAFLL...  |
| RhiE_DH*  | ... HQFNHRRILLG... | ... NSAFLL...  |
| BaeR_DH*  | ... HQFSGEPVLVG... | ... NSAYL...   |
| DifK_DH*  | ... HLVFGKPALMG... | ... NSCYM...   |
| EryAII_DH | ... HVVGGRTLVPG... | ... DAVAQ...   |
| DH_Cons.  | ... HXXXGXXXXPG... | ... DXXXQ/H... |

**Figure S1.** Alignment of DH and DH\* amino acid sequences. The conserved catalytic residues for DH domains are highlighted in red. The corresponding accession numbers are as follows: ADI59534.1 (CorL), ADI59532.1 (CorJ), YP\_003748161.1 (RhiE), ABS74065.1 (BaeR), CAJ57411.1 (DifK), AAV51821.1 (EryAII) and NspC\_DH\*<sup>6</sup>.

DH\_Cons. = consensus sequence normally found in standard DH domains<sup>5</sup>.

### 3.7.6 Standardized HPLC-MS method for sample analysis and application of principal component analysis

HPLC-MS measurements for initial quantification of the samples as well as the application of principal component analysis have to be highly standardized to assure exact mass and retention time reproducibility. Measurements were made on a Dionex Ultimate 3000 RSLC system using a Waters BEH C<sub>18</sub> column (50 x 2.1 mm, 1.7 μm) equipped with a Waters VanGuard BEH C<sub>18</sub> 1.7 μm guard column. Separation of 1 μl sample was achieved by a linear gradient from (A) H<sub>2</sub>O + 0.1% FA to (B) ACN + 0.1 % FA at a flow rate of 600 μL/min and a column temperature of 45 °C. Gradient conditions were as follows: 0 – 0.5 min, 5% B; 0.5 – 18.5 min, 5 – 95% B; 18.5 – 20.5 min, 95% B; 20.5 – 21 min, 95 – 5% B; 21-22.5 min, 5% B. UV spectra were recorded by a DAD in the range from 200 to 600 nm. The LC flow was split to 75 μL/min before entering the Bruker Daltonics maXis 4G hr-qTOF mass spectrometer using the Apollo II ESI source. Mass spectra were acquired in centroid mode ranging from 150 – 2500 m/z at a 2 Hz full scan rate. Mass spectrometry source parameters were set to 500V as end plate offset; 4000 V as capillary voltage; nebulizer gas pressure 1 bar; dry gas flow of 5 L/min and a dry temperature of 200°C. Ion transfer and quadrupole settings were set to Funnel RF 350 Vpp; Multipole RF 400 Vpp as transfer settings and Ion energy of 5eV as well as a low mass cut of 300 m/z as Quadrupole settings. Collision cell was set to 5.0 eV and pre pulse storage time was set to 5 μs. Spectra acquisition rate was set to 2Hz. Calibration of the maXis4G qTOF spectrometer was achieved with sodium formate clusters before every injection to avoid mass drifts. All MS analyses were acquired in the presence of the lock masses C<sub>12</sub>H<sub>19</sub>F<sub>12</sub>N<sub>3</sub>O<sub>6</sub>P<sub>3</sub>, C<sub>18</sub>H<sub>19</sub>O<sub>6</sub>N<sub>3</sub>P<sub>3</sub>F<sub>2</sub> and C<sub>24</sub>H<sub>19</sub>F<sub>36</sub>N<sub>3</sub>O<sub>6</sub>P<sub>3</sub> which generate the [M+H]<sup>+</sup> Ions of 622.028960, 922.009798 and 1221.990638. The corresponding MS<sup>2</sup> method operating in automatic precursor selection mode picks up the two most intense precursors per cycle, applies smart exclusion after five spectra and performs CID and MS/MS spectra acquisition time ramping. CID Energy was ramped from 35 eV for 500 m/z to 45 eV for 1000 m/z and 60 eV for 2000 m/z. MS full scan acquisition rate was set to 2Hz and MS/MS spectra acquisition rates were ramped from one to four Hz for precursor Ion intensities of 10kcts to 1000kcts.

### 3.7.7 HPLC-MS method for quantification of (pre)Cor A production yields

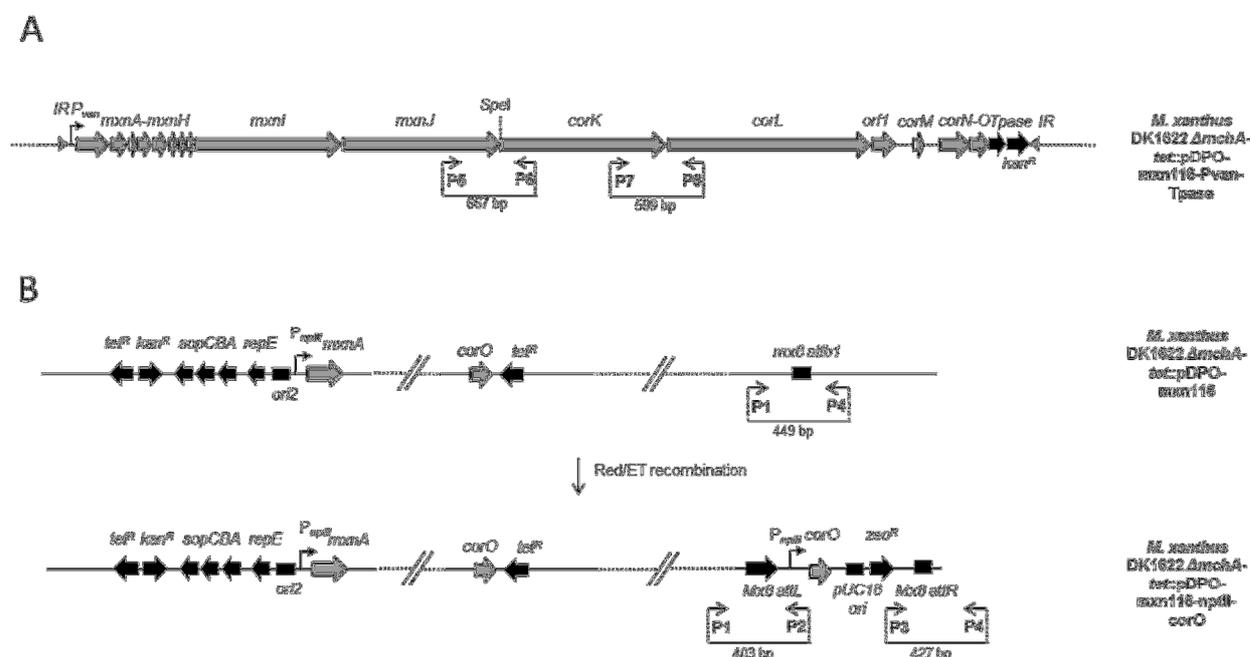
Detailed quantification of samples was performed on a Dionex Ultimate 3000 RSLC system using a BEH C<sub>18</sub>, 100 x 2.1 mm, 1.7 μm dp column equipped with a C<sub>18</sub> precolumn (Waters,

Germany). Separation of a 1  $\mu\text{L}$  sample was achieved by a linear gradient from (A)  $\text{H}_2\text{O} + 0.1\%$  FA to (B)  $\text{ACN} + 0.1\%$  FA at a flow rate of 600  $\mu\text{L}/\text{min}$  and 45  $^\circ\text{C}$ . The gradient was initiated by a 0.5 min isocratic step at 5% B, followed by an increase to 50% B in 1 min, to 65% B in 6 min and to 98% B in 0.5 min. After a 1 min step at 98% B the system was re-equilibrated to the initial conditions (5% B). UV spectra were recorded by a DAD in the range from 200 to 600 nm. For MS measurements on solariX XR (7T) FT-ICR mass spectrometer (Bruker Daltonics, Germany), the LC flow was split to 75  $\mu\text{L}/\text{min}$  before entering the mass spectrometer using the Apollo ESI source. In the source region, the temperature was set to 200  $^\circ\text{C}$ , the capillary voltage was 4500 V, the dry-gas flow was 4.0 L/min and the nebulizer was set to 1.1 bar. After the generated ions passed the quadrupole with a low cutoff at 150 m/z they were trapped in the collision cell for 80 ms and finally transferred within 0.9 ms through the hexapole into the ICR cell. Captured ions were excited by applying a frequency sweep from 150 to 950 m/z and detected in broadband mode by acquiring a 184 ms transient. For quantification of COR A and preCOR A the peak areas of  $[\text{M}+\text{H}-\text{H}_2\text{O}]^+$  510.2850 m/z at 5.85 min and  $[\text{M}+\text{H}]^+$  512.2850 m/z at 8.98 min, respectively, were integrated and compared to a calibration curve obtained from serial dilutions of authentic COR A and preCOR A reference material covering the concentration range from 0.5-100  $\mu\text{g}/\text{mL}$  (generated from 1 mg/L stock solution in acetonitrile).

### 3.7.8 Statistical data treatment using the principal component analysis

To assay the influence of additional *corO* copy and identify possible differences between the *M. xanthus* DK1622 *AmchA::pDPO-mxn116* mutant and the *M. xanthus* DK1622 *AmchA::pDPO-mxn116-nptII-corO* mutant, 6 LC-HRMS measurements acquired with our 18 min standard method on the UHPLC hyphenated maxis 4G spectrometer (Section 3.7.6) from extracts of three independent cultivations per bacterial strain were used for statistical data treatment. For preprocessing of MS data the molecular feature finder implemented in Bruker Compass Data Analysis 4.2 (Bruker, Bremen) was used with the compound detection parameters SN threshold 1; Correlation coefficient 0.9; minimum compound length 10 spectra and smoothing width of 3 spectra. Bucketing was done with Bruker Compass Profile Analysis 2.1 (Bruker, Bremen) with advanced bucketing and window parameters of 30s and 15 ppm. Bucket value was log transformed to avoid under-evaluating low intensity signals in the presence of high intensity signals. The PCA t-Test function was used in order to separate medium derived MS features from the metabolome derived MS features, since the t-Test table

that can be exported from Bruker Compass Profile Analysis 2.1 (Bruker, Bremen) contains information upon how many blanks and how many bacterial extracts contain said feature.



**Figure S2.** Genotypic verification of *M. xanthus* DK1622 heterologous expression strains. Genotypic verification of *mxn/cor* BGC (pDPO-*mxn116*-Pvan-Tpase) integration in *M. xanthus* DK1622 by transposition (A). Genotypic verification of *corO* integration in *M. xanthus* DK1622  $\Delta$ *mchA*::pDPO-*mxn116* by the *mx8* phage integrase. The same method was used to verify all other construct integrations in *mx8 attB1* site (B). Primer sets used for genotypic verification, their binding sites and amplicon sizes are illustrated. Nucleotide sequences of the primers used are listed in the table S1.

### 3.7.9 Compound isolation and characterization

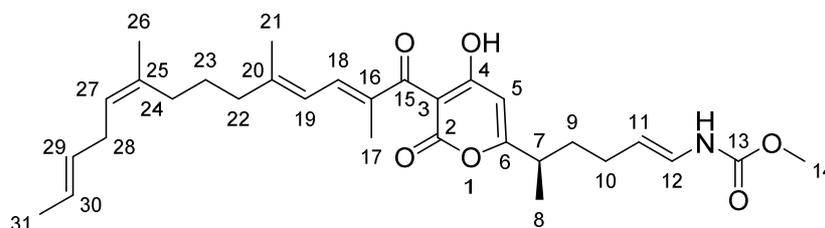
#### 3.7.9.1 Isolation of preCOR A

Strain *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ *corO* was grown in 6 L of M7/s6 medium (3 x 2 L medium in 5 L conical flask) supplemented with kanamycin 50  $\mu$ g/mL at 180 rpm and 30  $^{\circ}$ C. Corallopyronin production was induced with 1 mM vanillate at the beginning of the cultivation. After 6 days of cultivation cultures were harvested by centrifugation at 8000 rpm, 4  $^{\circ}$ C for 15 min. Supernatant was extracted by liquid/liquid extraction with 4 L of EtOAc and pellet was extracted separately with 1L of MeOH. Extracts were dried on the rotary evaporator prior to purification on HPLC-MS. Most of the preCOR A was obtained from the cell pellet. Purification of preCOR A from this crude extract was carried out on a Dionex Ultimate 3000 SDLC low pressure gradient system on an Agilent Zorbax C8 250x10 mm 5  $\mu$ m column with the eluents H<sub>2</sub>O + 0.1% FA as A and ACN + 0.1% FA as B, a flow rate of 5 mL/min and a column thermostated at 30  $^{\circ}$ C. Precorallopyronin A

can be detected by UV absorption at 312 nm. The gradient that starts with a plateau at 50% A for 3 minutes. Then A content is ramped to 5% A during 17 minutes and finally ramped to 95% A during 2 minutes. A content is kept stable at 95% for 1 minutes and then ramped back to 50% A during 30 seconds. The column is reequilibrated at 50% A for 3 minutes. Precorallopyronin A is obtained as weakly yellowish amorphous solid.

### 3.7.9.2 NMR based characterization of preCOR A

$^1\text{H}$  NMR,  $^{13}\text{C}$  NMR and 2D spectra were recorded at 700 MHz ( $^1\text{H}$ )/175 MHz ( $^{13}\text{C}$ ) conducting an Avance III Ascend 700 spectrometer using a cryogenically cooled triple resonance probe (Bruker Biospin, Rheinstetten, Germany). Samples were dissolved in methanol- $d_4$ . Chemical shifts are reported in ppm relative to TMS, the solvent was used as the internal standard.



**Figure S3.** Structure formula of preCOR A with atom numbering according to the NMR data.

**Table S4.** List of all observed NMR correlations for preCOR A in methanol- $d_4$ .

| Carbon number | $\delta$ $^1\text{H}$ [ppm] | Multiplicity, J [Hz] and proton number | $\delta$ $^{13}\text{C}$ [ppm] | COSY correlations [ppm] | HMBC correlations [ppm]          |
|---------------|-----------------------------|--|--------------------------------|-------------------------|----------------------------------|
| 2             | -                           | -                                      | 163.6                          | -                       | -                                |
| 3             | -                           | -                                      | 102.2                          | -                       | -                                |
| 4             | -                           | -                                      | n.d.                           | -                       | -                                |
| 5             | 6.11                        | s, 1H                                  | 100.6                          | 2.67                    | 38.9, 102.2, 163.6, 173.7        |
| 6             | -                           | -                                      | 173.7                          | -                       | -                                |
| 7             | 2.67                        | m, 1H                                  | 38.9                           | 1.25                    | 35.5, 38.9, 173.7                |
| 8             | 1.25                        | d, 6.9, 3H                             | 18.2                           | 2.67                    | 18.2, 35.5, 173.7                |
| 9a            | 1.58                        | m, 1H                                  | 35.5                           | 1.77, 2.02              | 18.2, 28.4, 38.9, 110.3, 173.0   |
| 9b            | 1.77                        | m, 1H                                  | 35.5                           | 1.58, 2.02              | 18.2, 28.4, 38.9, 110.3, 173.0   |
| 10            | 2.02                        | m, 2H                                  | 28.4                           | 1.58, 1.77, 5.04, 6.40  | 35.5, 39.9, 110.3, 125.9         |
| 11            | 5.04                        | dt, 14.3; 7.2, 1H                      | 110.3                          | 2.02, 6.40              | 28.4, 35.5, 125.9                |
| 12            | 6.40                        | d, 14.3, 1H                            | 125.9                          | 2.02, 5.04              | 28.4, 35.5, 110.3, 156.6         |
| 13            | -                           | -                                      | 156.6                          | -                       | -                                |
| 14            | 3.66                        | s, 3H                                  | 52.5                           | -                       | 156.6                            |
| 15            | -                           | -                                      | 198.4                          | -                       | -                                |
| 16            | -                           | -                                      | 134.8                          | -                       | -                                |
| 17            | 1.94                        | m, 3H                                  | 11.8                           | 6.28, 7.15              | -                                |
| 18            | 7.15                        | d, 11.5, 1H                            | 137.8                          | 1.94, 6.28              | 11.8, 121.9, 134.8, 151.0, 198.4 |

|    |      |             |       |                        |                                      |
|----|------|-------------|-------|------------------------|--------------------------------------|
| 19 | 6.28 | d, 11.5, 1H | 121.9 | 1.82, 1.94, 2.20, 7.15 | 17.2, 41.4, 134.0, 137.8, 151.0      |
| 20 | -    | -           | 151.0 | -                      | -                                    |
| 21 | 1.82 | s, 3H       | 17.2  | 2.20, 6.28, 7.15       | 41.4, 121.9, 137.8, 151.0            |
| 22 | 2.20 | t, 7.6, 2H  | 41.4  | 1.58, 1.82, 6.28       | 17.2, 27.1, 32.0, 41.4, 121.9, 151.0 |
| 23 | 1.58 | m, 2H       | 27.1  | 2.04, 2.20             | 32.0, 41.4, 136.2                    |
| 24 | 2.04 | m, 2H       | 32.0  |                        | 23.4, 27.4, 41.4, 124x, 136.2        |
| 25 | -    | -           | 136.2 | -                      | -                                    |
| 26 | 1.70 | m, 3H       | 23.4  | 2.04, 2.64, 5.15       | 17.9, 31.9, 124.5, 131.1, 136.2      |
| 27 | 5.15 | t, 7.2, 1H  | 124.5 | 1.70, 2.04, 2.64       | 23.4, 31.9, 131.1, 136.2             |
| 28 | 2.64 | m, 2H       | 31.9  | 1.71, 5.15, 5.38, 5.40 | 17.9, 23.4, 125.4, 131.1, 136.2      |
| 29 | 5.38 | m, 1H       | 131.1 | 1.62, 2.64, 5.40       | 17.9, 31.9, 125.4                    |
| 30 | 5.40 | m, 1H       | 125.4 | 1.62, 2.64, 5.38       | 17.9, 31.9, 131.1                    |
| 31 | 1.62 | d, 5.2, 3H  | 17.9  | 2.64, 5.38, 5.40       | 31.9, 125.4, 131.1                   |

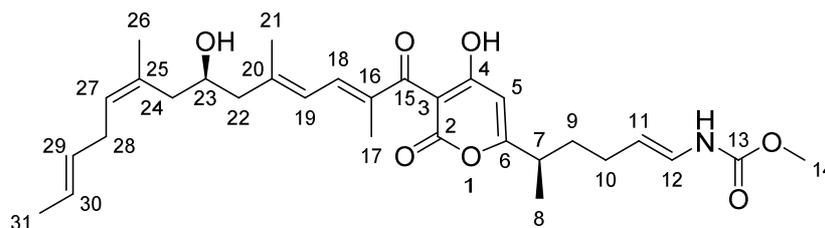
### 3.7.9.3 Isolation of COR D

Strain *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corN was grown in 6 L of M7/s6 medium (3 x 2 L medium in 5 L conical flask) supplemented with kanamycin 50  $\mu$ g/mL at 180 rpm and 30 °C. Corallopyronin production was induced with 1 mM vanillate at the beginning of the cultivation. After 6 days of cultivation cultures were harvested by centrifugation at 8000 rpm, 4 °C for 15 min and supernatant was extracted by liquid/liquid extraction with 4 L of EtOAc. Extracts were dried on the rotary evaporator prior to purification on HPLC-MS. Purification of corallopyronin D (Cor D) was carried out on a Waters preparative SFC system, equipped with a Thar fluid delivery module, 2767 sample manager, an ABPR 20, a GLS system featuring a 515 GLS pump, a column oven thermostated to 45 °C and a 2998 photodiode array detector coupled to a Waters Acquity QDA mass analyzer operated with a 515 Make up pump. Autopurification is done by setting the fraction collection trigger mass to 528.3 corresponding to Corallopyronin D  $[M+H]^+$ . QDA probe temperature is set to 600 °C, Cone voltage is set to 25 V while spray voltage was set to 1.75 kV. Separation was carried out on a Waters Torus 2EP 250x19mm 5 $\mu$ m column using MeOH as co solvent and a flow rate of 15 mL/min. GLS solvent was Methanol and make up solvent 80/20 Methanol/Water + 0.2% formic acid. Separation started by keeping 5% co solvent for 1 minute followed by a co solvent ramp to 55% during 10.5 minutes. Co solvent concentration is subsequently kept at 55% during 1 minute. Co solvent concentration is then ramped back to initial conditions during 30s and the column is reequilibrated for 1 minute prior to the next injection. Further purification is done using a Dionex Ultimate 3000 SDLC low pressure gradient system on a Agilent Zorbax C8 250x10 mm 5  $\mu$ m column with the eluents H<sub>2</sub>O + 0.1% FA as A and ACN + 0.1% FA as B, a flow rate of 5 mL/min and a

column thermostated at 30 °C. Corallopyronin D can be detected by UV absorption at 312 nm. The gradient that starts with a plateau at 60% A for 3 minutes. Then A content is ramped to 20% A during 23 minutes and finally ramped to 95% A during 2 minutes. A content is kept stable at 95% for 2 minutes and then ramped back to 60% A during 30 seconds. The column is reequilibrated at 60% A for 3 minutes. Corallopyronin D (COR D) is obtained as weakly yellowish amorphous solid.

### 3.7.9.4 NMR based characterization of COR D

$^1\text{H}$  NMR,  $^{13}\text{C}$  NMR and 2D spectra were recorded at 500 MHz ( $^1\text{H}$ )/125 MHz ( $^{13}\text{C}$ ) conducting an Avance III Ultrashield 500 spectrometer using a cryogenically cooled triple resonance probe (Bruker Biospin, Rheinstetten, Germany). Samples were dissolved in methanol- $d_4$ . Chemical shifts are reported in ppm relative to TMS, the solvent was used as the internal standard.



**Figure S4.** Structure formula of COR D with atom numbering according to the NMR data.

**Table S5.** List of all observed NMR correlations for COR D in methanol- $d_4$ .

| Carbon number | $\delta$ $^1\text{H}$ [ppm] | Multiplicity, J [Hz] and proton number | $\delta$ $^{13}\text{C}$ [ppm] | COSY correlations [ppm] | HMBC correlations [ppm]                 |
|---------------|-----------------------------|--|--------------------------------|-------------------------|---|
| 2             | -                           | -                                      | 141.3                          | -                       | -                                       |
| 3             | -                           | -                                      | 102.3                          | -                       | -                                       |
| 4             | -                           | -                                      | 174.5                          | -                       | -                                       |
| 5             | 6.09                        | s, 1H                                  | 101.0                          | -                       | 39.1, 102.3, 141.3, 173.5, 174.5, 198.8 |
| 6             | -                           | -                                      | 173.5                          | -                       | -                                       |
| 7             | 2.67                        | m, 1H                                  | 39.1                           | 1.25, 1.58, 1.75        | 18.3, 28.4, 35.3, 101.0, 173.5          |
| 8             | 1.25                        | d, 6.9, 3H                             | 18.3                           | 2.67                    | 35.3, 39.1, 173.5                       |
| 9a            | 1.58                        | m, 1H                                  | 35.3                           | 1.75, 2.02, 2.67        | 18.3, 28.4, 39.1, 110.5, 173.5          |
| 9b            | 1.75                        | m, 1H                                  | 35.3                           | 1.58, 2.02, 2.67        | 18.3, 28.4, 39.1, 110.5, 173.5          |
| 10            | 2.02                        | m, 2H                                  | 28.4                           | 1.58, 1.75, 5.05        | 35.5, 39.1, 110.5, 125.8                |
| 11            | 5.05                        | dt, 14.4; 7.2, 1H                      | 110.5                          | 2.02, 6.39              | 28.4, 35.3, 125.8                       |
| 12            | 6.39                        | d, 14.4, 1H                            | 125.8                          | 5.05, 2.02              | 28.4, 110.5, 156.9                      |
| 13            | -                           | -                                      | 156.9                          | -                       | -                                       |
| 14            | 3.66                        | s, 3H                                  | 52.5                           | -                       | 156.9                                   |
| 15            | -                           | -                                      | 198.8                          | -                       | -                                       |
| 16            | -                           | -                                      | 135.4                          | -                       | -                                       |

|     |      |             |       |                  |                                      |
|-----|------|-------------|-------|------------------|--------------------------------------|
| 17  | 1.95 | m, 3H       | 11.9  | 7.17             | 135.4, 137.6, 198.8                  |
| 18  | 7.17 | d, 11.5, 1H | 137.6 | 1.95, 6.33       | 11.9, 124.2, 135.4, 148.1, 198.8     |
| 19  | 6.33 | d, 11.5, 1H | 124.2 | 1.85, 7.17       | 17.7, 49.4, 137.6, 148.1             |
| 20  | -    | -           | 148.1 | -                | -                                    |
| 21  | 1.85 | s, 3H       | 17.7  | 2.25, 2.33, 6.33 | 49.4, 69.4, 124.2, 137.6, 148.1      |
| 22a | 2.25 | m, 1H       | 49.4  | 2.33, 3.93       | 17.7, 40.9, 49.4, 69.4, 124.2, 148.2 |
| 22b | 2.33 | m, 1H       | 49.4  | 2.25, 3.93       | 17.7, 40.9, 49.4, 69.4, 124.2, 148.2 |
| 23  | 3.93 | m, 1H       | 69.4  | 2.22, 2.25, 3.33 | 40.9, 49.1, 133.7, 148.1             |
| 24  | 2.22 | t, 8.1, 2H  | 40.9  | 3.93             | 24.1, 49.4, 69.4, 126.7, 133.7       |
| 25  | -    | -           | 133.7 | -                | -                                    |
| 26  | 1.75 | m, 3H       | 24.1  | 2.7, 5.25        | 40.9, 69.4, 126.7, 131.0, 133.7      |
| 27  | 5.25 | t, 7.3, 1H  | 126.7 | 1.75, 2.7        | 24.1, 32.1, 40.9, 131.0, 133.7       |
| 28  | 2.70 | m, 2H       | 32.1  | 1.75, 5.25, 5.38 | 17.8, 125.8, 126.7, 131.0, 133.7     |
| 29  | 5.39 | m, 1H       | 131.0 | 1.62, 2.7, 5.41  | 17.8, 32.1, 125.8                    |
| 30  | 5.41 | m, 1H       | 125.8 | 1.62, 2.7, 5.39  | 17.8, 32.1, 131.0                    |
| 31  | 1.62 | d, 6.2, 3H  | 17.9  | 2.7, 5.39, 5.41  | 32.2, 125.8, 131.0                   |

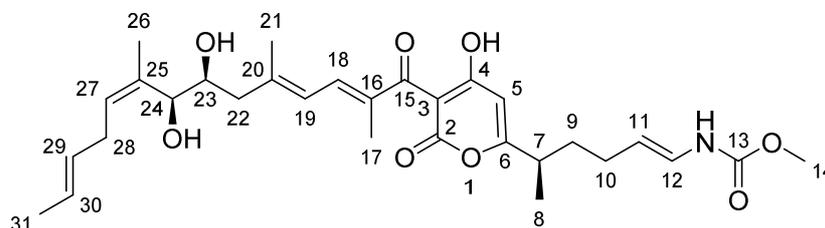
### 3.7.9.5 Isolation of oxyCOR A

Strain *M. xanthus* DK1622::pDPO-Mxn116-Pvan-Tpase $\Delta$ corN was grown in 6 L of M7/s6 medium (3 x 2 L medium in 5 L conical flask) supplemented with kanamycin 50  $\mu$ g/mL at 180 rpm and 30 °C. Coralopyronin production was induced with 1 mM vanillate at the beginning of the cultivation. After 6 days of cultivation cultures were harvested by centrifugation at 8000 rpm, 4 °C for 15 min and supernatant was extracted by liquid/liquid extraction with 4 L of EtOAc. Extracts were dried on the rotary evaporator prior to purification on HPLC-MS. Purification was carried out on a Waters Autopurifier (Eschborn, Germany) high pressure gradient system, equipped with 2545 binary gradient module, SFO system fluidics organizer, 2767 sample manager and a 2998 photodiode array detector coupled to a 3100 single quadrupole mass spectrometer operated in positive ion mode. Source and voltage settings for the MS were as follows: mass range, m/z 300 - 1000; scan duration, 1 s; points per Dalton, 4; capillary voltage, 3.5 kV; cone voltage, 30 V; extractor voltage 3 V; RF lens, 0.1 V; source temperature 150 C, desolvation temperature, 250 C; desolvation gas flow, 400 L/hr; cone gas flow, 50 L/hr; ion counting threshold, 30. Autopurification is done by setting time dependent fraction collection on the 2767 sample manager. Separation was carried out on Waters Xbridge BEH C18 150x20 mm column using MeOH + 0.1% FA as B and H<sub>2</sub>O + 0.1% FA as A with a flow rate of 25 mL/min. Separation is started with a plateau at 55% A for 2 minutes followed by a ramp to 35% A during 11 minutes and then to 5% A during 1 minute. The A content is kept at 5% A for 2 minutes. The

A content is ramped back to starting conditions during 30 seconds and the column is re equilibrated for 5 minutes. Oxycoralopyronin A (oxyCOR A) is obtained as slightly yellowish amorphous solid.

### 3.7.9.6 NMR based characterization of oxyCOR A

$^1\text{H}$  NMR,  $^{13}\text{C}$  NMR and 2D spectra were recorded at 500 MHz ( $^1\text{H}$ )/125 MHz ( $^{13}\text{C}$ ) conducting an Avance III Ultrashield 500 spectrometer using a cryogenically cooled triple resonance probe (Bruker Biospin, Rheinstetten, Germany). Samples were dissolved in methanol- $d_4$ . Chemical shifts are reported in ppm relative to TMS, the solvent was used as the internal standard.



**Figure S5.** Structure formula of oxyCOR A with atom numbering according to the NMR data.

**Table S6.** List of all observed NMR correlations for oxyCOR A in methanol- $d_4$ .

| Carbon number | $\delta$ $^1\text{H}$ [ppm] | Multiplicity, J [Hz] and proton number | $\delta$ $^{13}\text{C}$ [ppm] | COSY correlations [ppm] | HMBC correlations [ppm]        |
|---------------|-----------------------------|--|--------------------------------|-------------------------|--------------------------------|
| 2             | -                           | -                                      | 149.9                          | -                       | -                              |
| 3             | -                           | -                                      | 102.3                          | -                       | -                              |
| 4             | -                           | -                                      | 179.2                          | -                       | -                              |
| 5             | 5.81                        | s, 1H                                  | 105.8                          | -                       | 38.5, 102.3, 169.6, 179.2      |
| 6             | -                           | -                                      | 169.9                          | -                       | -                              |
| 7             | 2.56                        | m, 1H                                  | 38.5                           | 1.22, 1.55, 1.74        | 18.5, 28.4, 35.7, 169.6        |
| 8             | 1.22                        | d, 7.0, 3H                             | 18.5                           | 2.56                    | 35.7, 38.5, 169.6              |
| 9a            | 1.55                        | m, 1H                                  | 35.7                           | 1.74, 2.01, 2.56        | 18.5, 28.4, 38.5, 110.8, 169.9 |
| 9b            | 1.74                        | m, 1H                                  | 35.7                           | 1.55, 2.01, 2.56, 5.34  | 18.5, 28.4, 38.5, 110.8, 169.9 |
| 10            | 2.01                        | m, 2H                                  | 28.4                           | 1.55, 1.74, 5.07        | 35.7, 110.8, 125.5             |
| 11            | 5.07                        | dt, 13.9; 7.2, 1H                      | 110.8                          | 2.01, 6.39              | 28.3, 35.7, 125.5              |
| 12            | 6.39                        | d, 13.9, 1H                            | 125.5                          | 5.08, 2.01              | 28.3, 110.8, 156.6             |
| 13            | -                           | -                                      | 156.6                          | -                       | -                              |
| 14            | 3.66                        | s, 3H                                  | 52.3                           | -                       | 156.6                          |
| 15            | -                           | -                                      | 200.4                          | -                       | -                              |
| 16            | -                           | -                                      | 135.9                          | -                       | -                              |
| 17            | 1.93                        | m, 3H                                  | 11.51                          | 7.31                    | 135.9, 137.8, 200.4            |
| 18            | 7.33                        | d, 12.0, 1H                            | 137.8                          | 1.93, 6.32              | 11.51, 136.1, 171.2, 200.4     |
| 19            | 6.34                        | d, 12.0, 1H                            | 124.5                          | 7.33, 2.27, 1.83        | 17.3, 45.5, 136.1, 137.8       |
| 20            | -                           | -                                      | 146.5                          | -                       | -                              |
| 21            | 1.83                        | s, 3H                                  | 17.3                           | 2.27, 2.11, 3.76, 6.33  | 45.5, 124.5, 146.5, 151.1      |

|            |      |            |       |                  |                                  |
|------------|------|------------|-------|------------------|----------------------------------|
| <b>22a</b> | 2.27 | m, 1H      | 45.5  | 2.11, 3.76, 6.34 | 17.3, 72.6, 124.5, 146.2         |
| <b>22b</b> | 2.11 | m, 1H      | 45.5  | 2.27, 3.76, 6.34 | 17.3, 72.6, 73.9, 124.5, 146.2   |
| <b>23</b>  | 3.76 | m, 1H      | 72.6  | 4.30, 2.11       | 73.9, 146.2                      |
| <b>24</b>  | 4.3  | d, 7.5, 1H | 73.9  | 3.76             | 18.7, 45.5, 73.9, 128.3, 136.2   |
| <b>25</b>  | -    | -          | 136.2 | -                | -                                |
| <b>26</b>  | 1.72 | m, 3H      | 18.7  | 5.34             | 73.9, 128.3, 130.3, 136.2        |
| <b>27</b>  | 5.34 | t, 7.6, 1H | 128.3 | 1.72, 2.76       | 73.9, 18.7                       |
| <b>28</b>  | 2.76 | m, 2H      | 31.4  | 5.34, 5.39, 5.44 | 17.6, 125.8, 128.3, 130.3, 136.2 |
| <b>29</b>  | 5.39 | m, 1H      | 130.3 | 1.65, 2.76, 5.44 | 17.6, 31.4, 126.1, 128.3         |
| <b>30</b>  | 5.44 | m, 1H      | 126.1 | 1.65, 5.39       | 17.6, 31.4, 128.3, 130.3         |
| <b>31</b>  | 1.65 | m, 3H      | 17.6  | 2.76, 5.44, 5.39 | 130.3, 126.1                     |

### 3.7.10 Investigation of the stereochemistry of COR D and oxyCOR A

#### 3.7.10.1 *In silico* evaluation of the ketoreductase stereochemistry

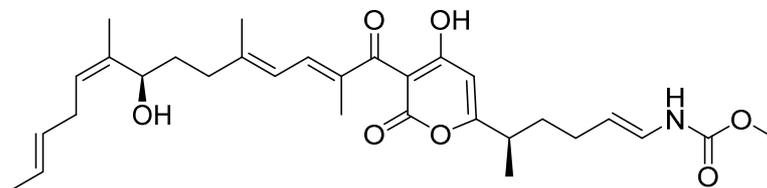
The additional hydroxyl group present in COR D and oxyCOR A originates from  $\beta$ -ketoreduction by the ketoreductase (KR) domain from module 4. The stereochemistry of the resulting OH group at C-23 depends on the type of the KR domain. The KR domains can be grouped in different types based on the specific core motifs present in the sequence<sup>7</sup>. The most obvious fingerprint to classify KR is the presence or absence of the LDD motif, in particular the presence of the second aspartate residue<sup>8</sup>. The LDD motif is absent in CorL\_KR\_M4' (Fig. S6) which indicates the A type KR. The A type KR should result in S alcohol stereochemistry, which is in agreement with the prediction from a software tool developed by Kitsche and Kalesse ([https://akitsche.shinyapps.io/profileHMM\\_App/](https://akitsche.shinyapps.io/profileHMM_App/))<sup>9</sup>. The S configuration of the OH group at C-23 has also been experimentally proven below.

| Sequence name | LDD   | KR type |
|---------------|---|---------|
| Ave1          | ...HTAGILDDATL...SSAAATFGAPGQANYAAANA...WGTWQGNGLADSDKARAYLDR---RG...   | Type B1 |
| Ery1          | ...HAAATLDDGTV...SSFASAFGAPGLGGYAPGNA...WGTWAGSGMAEGPVADRFRH----G...    | Type B2 |
| Amp1          | ...HTAAVIELAAL...SSTAGMWGSGVHAAYVAGNA...WGKWPDDLERE LADPHQIRR----SG...  | Type A2 |
| Ery2          | ...HAAGLPQQVAI...SSGAGVWGSARQGAYAAGNA...WGLWAAGGMTGDEEAVSFLR---ERG...   | Type A1 |
| CorI_KR_M2    | ...HAAGERRDARL...SSLVARTGNVQTDYAFANA...WPYWREGGMKVDAATEQFIRSFSGTE...    |         |
| CorJ_KR_M4    | ...HSALVLQDALL...SSAVGISGGAGQSHYAAASR...WGFWGTVGAVATPEYNQLLR---EEG...   |         |
| CorK_KR_M1    | ...HLGGVLSDRLL...SSI VALTGNVQADYAAAGNS...WPLWEGVGMAGQGLTGE--LRHLRPIA... |         |
| CorK_KR_M2    | ...HGAGVRRDAFL...SSVSAVLGNTGQGDYAYANA...WPLWRDGGMTVPASEALLAT--TMG...    |         |
| CorL_KR_M4    | ...HAAGNLR-SLQ...SSVSSVLGNAGQGDYAYANA...WPLWREGGLPVTAELERLVEKTLGMR...   |         |
| CorL_KR_M6    | ...HSAIVLKDATL...SSAESFSCDAGQSNYAAASR...WGYWGSVGI VASERYRRELAK-RGVG...  |         |

**Figure S6.** Alignment of important KR motifs. The alignment is reproduced from the original alignment by Keatinge-Clay with example sequences for two different type A and type B KR<sup>7</sup>. The LDD motif, typically found in B type KR, is highlighted in green.

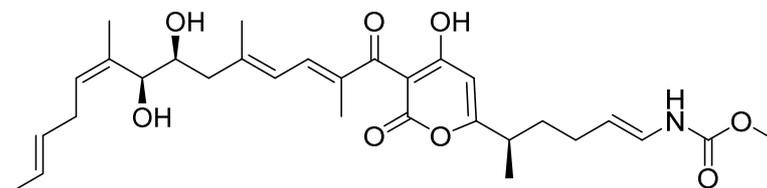
### 3.7.10.2 Experimental evidence for an *S,S*-diol configuration in oxyCOR A

The stereochemistry of the hydroxyl group in COR A has already been successfully determined to be *R*-configured<sup>10</sup>.

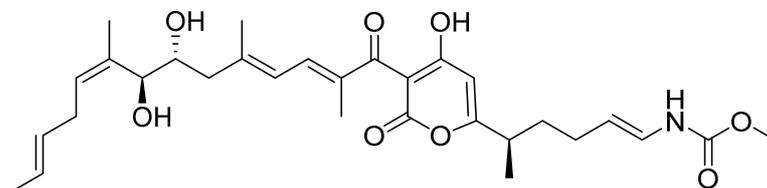


**Figure S7.** Structure formula of *R*-corallopyronin A.

In oxyCOR A the second OH group forms a vicinal diol together with the first OH group and therefore one can apply the coupling constant rules for vicinal diols. As vicinal diols form a strong hydrogen bond between the two OH their conformation is rather fixed via a hydrogen bond. This leads to the special case that the *meso* form of butane-2,3-diol has a much smaller coupling constant of about 3.5 Hz than the 6.8 Hz in the corresponding *S,S* or *R,R* isomers<sup>11</sup>. As we know from above, the configuration of the OH also present in COR A has to be *S* (inverted by IUPAC rules through the introduction of the second OH) while the configuration of the second hydroxyl group is still unknown. We therefore have the possibility of a pseudo-*meso* or a pseudo *S,S* arrangement.



**Figure S8.** Pseudo *S,S* form of oxyCOR A.



**Figure S9.** Pseudo *meso* form of oxyCOR A.

As the coupling constant of the two hydrogen atoms in this vicinal diol system equals 7.47 in methanol-*d*<sub>4</sub> one can safely assume the stereochemistry of the oxyCOR A diol to be *S,S*. Since the KR that creates this OH group is the same that creates the OH group in COR D and KR domains are enantioselective, one can assume COR D to have the OH group on the same molecule face (which in this case is *R*-configured as the second OH group is missing).

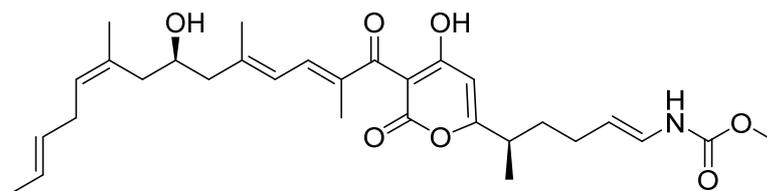


Figure S10. *R*-form of COR D.

### 3.7.11 1D and 2D NMR spectra

#### 3.7.11.1 NMR spectra of preCOR A

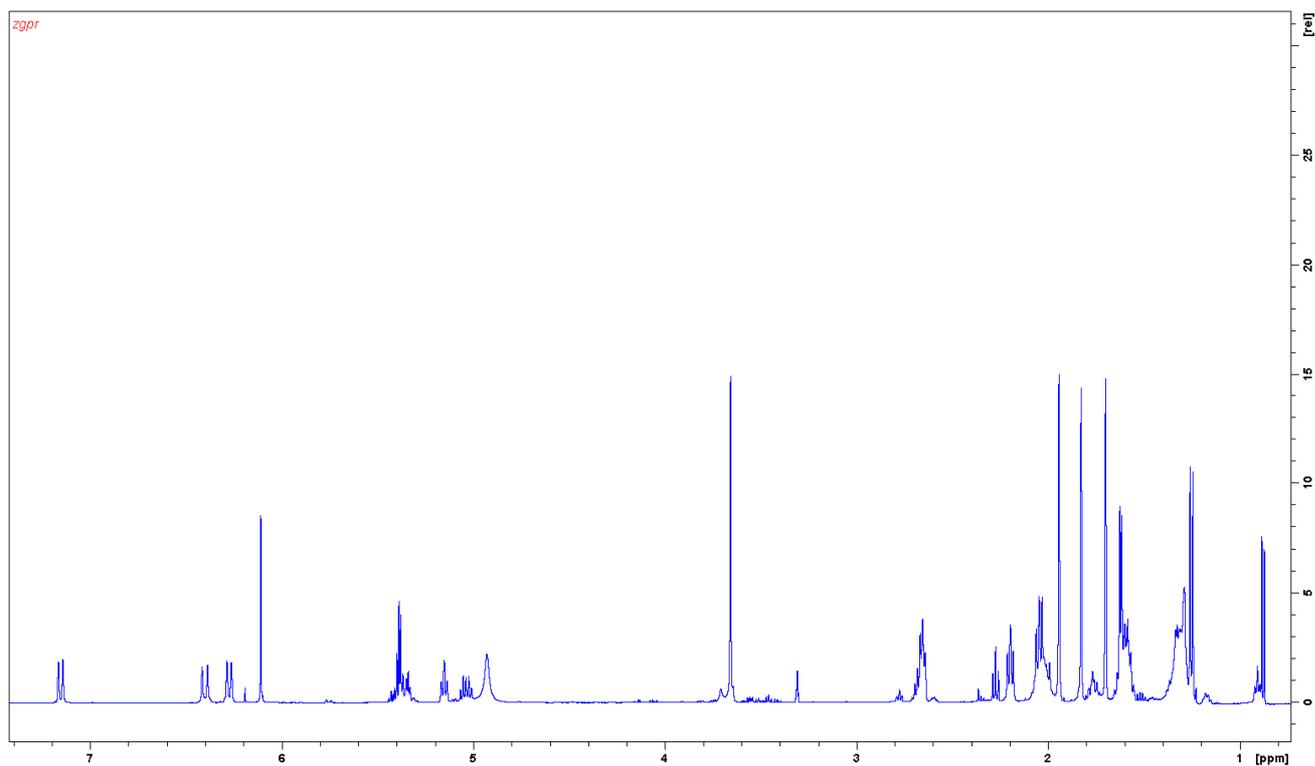
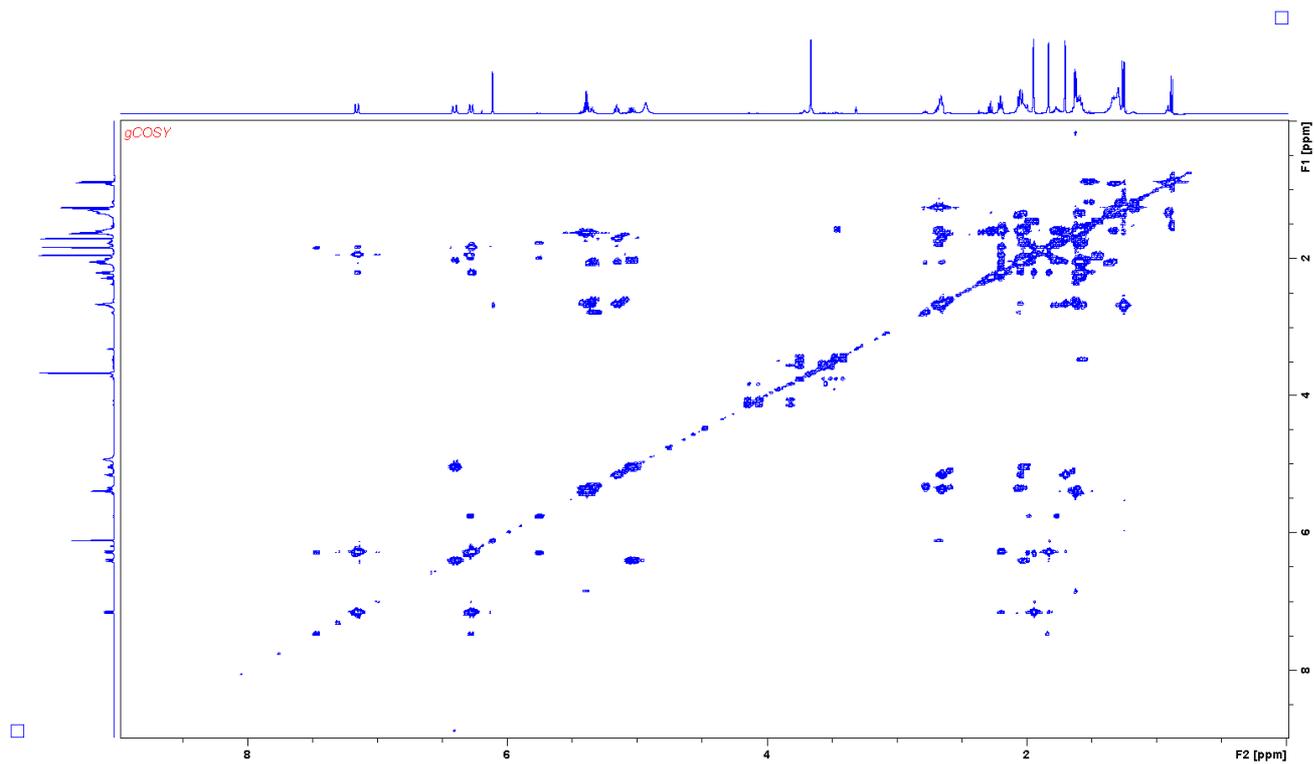
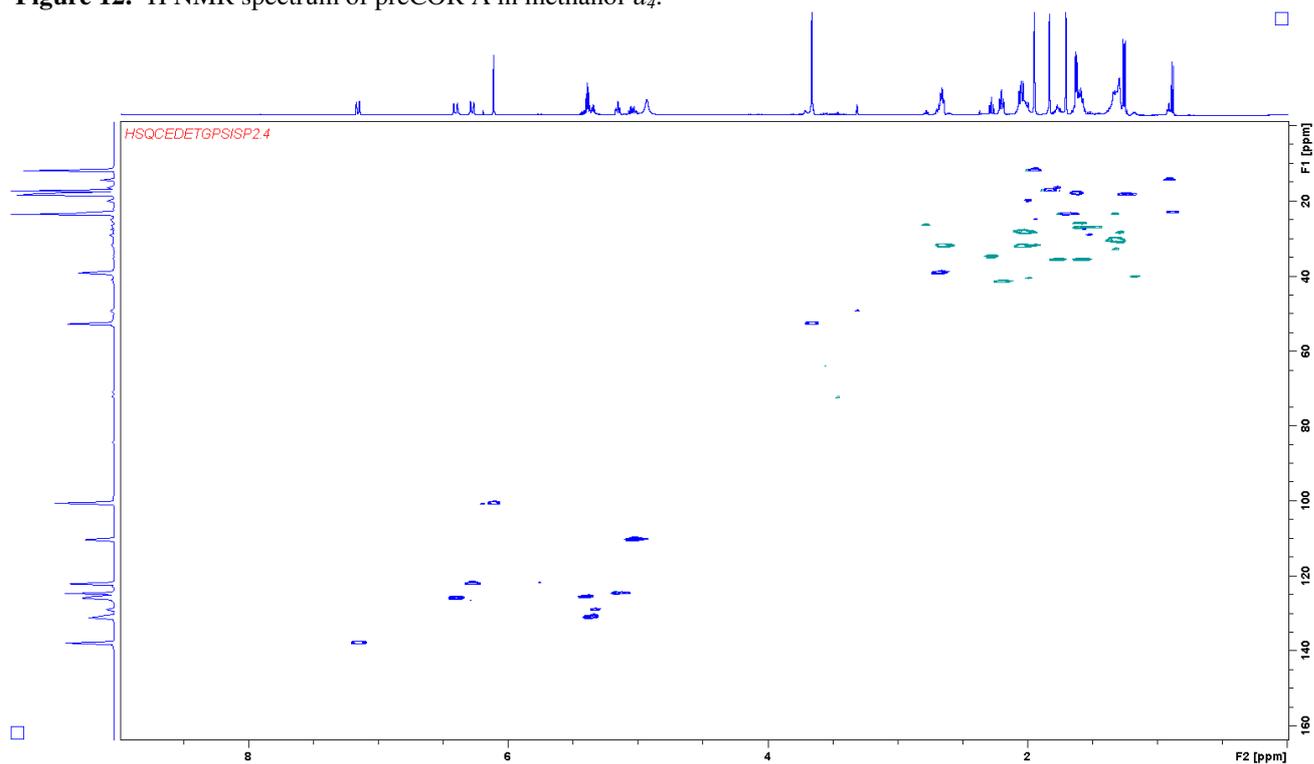


Figure S11.  $^1\text{H}$  NMR spectrum of preCOR A in methanol- $d_4$ .



**Figure 12.** <sup>1</sup>H NMR spectrum of preCOR A in methanol-*d*<sub>4</sub>.



**Figure S13.** HSQC NMR spectrum of preCOR A in methanol-*d*<sub>4</sub>.

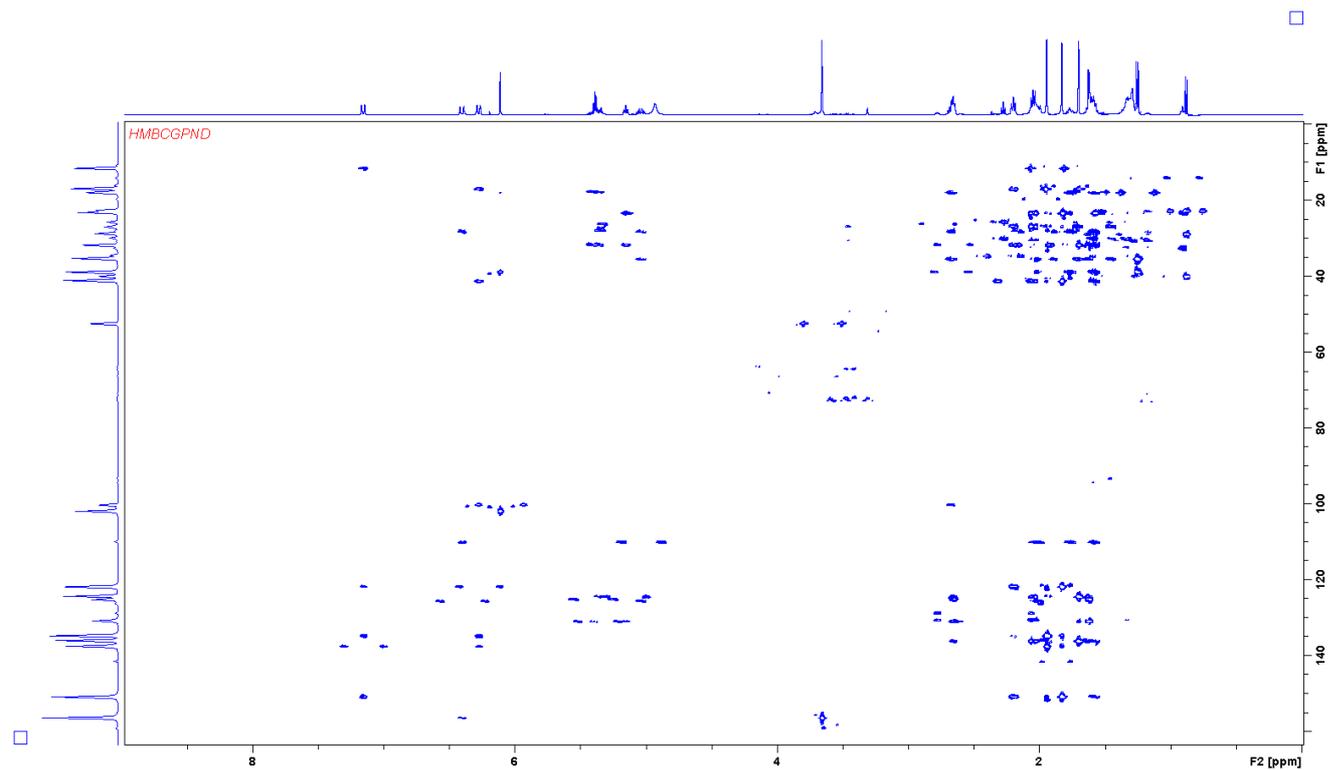


Figure S14. HMBC NMR spectrum of preCOR A in methanol- $d_4$ .

### 3.7.11.2 NMR spectra of COR D

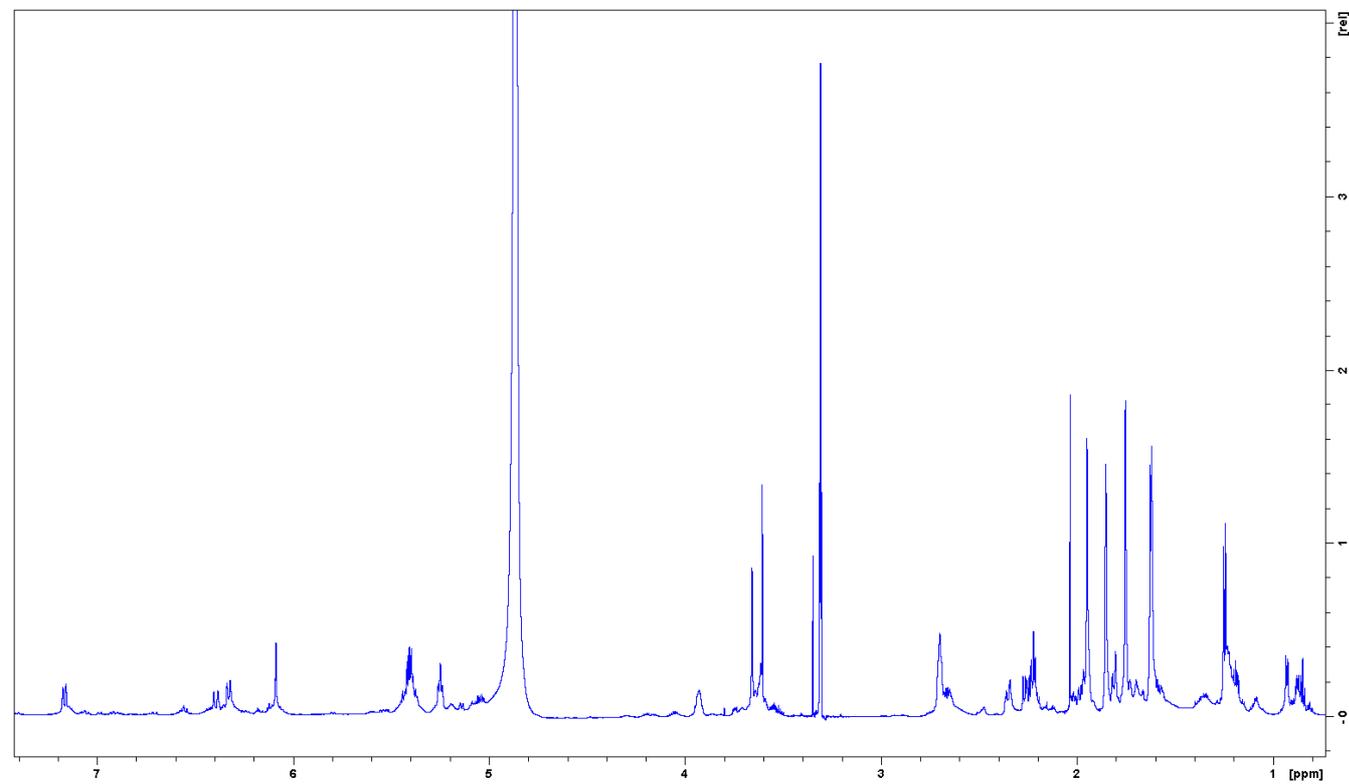
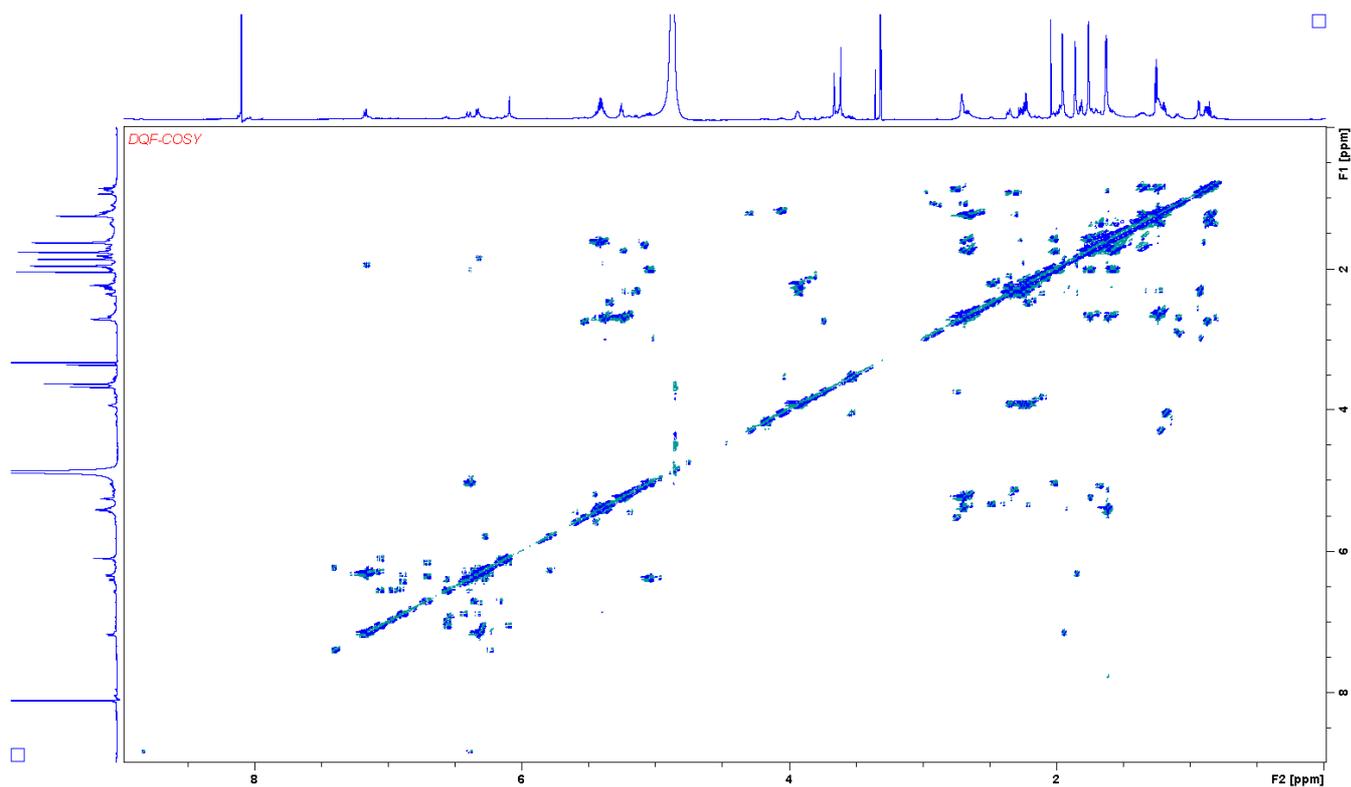
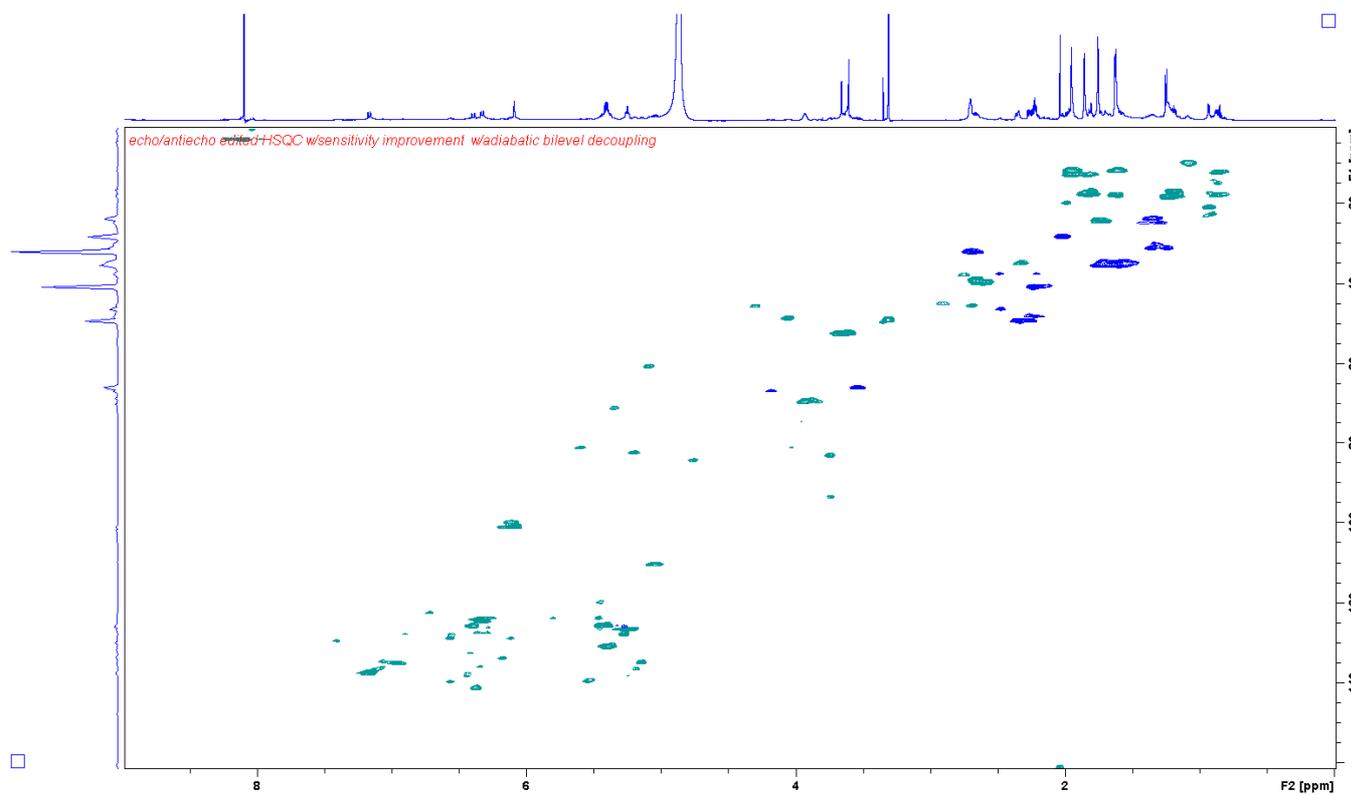


Figure S15.  $^1\text{H}$  NMR spectrum of COR D in methanol- $d_4$ .



**Figure S16.** COSY NMR spectrum of COR D in methanol- $d_4$ .



**Figure S17.** HSQC NMR spectrum of COR D in methanol- $d_4$ .

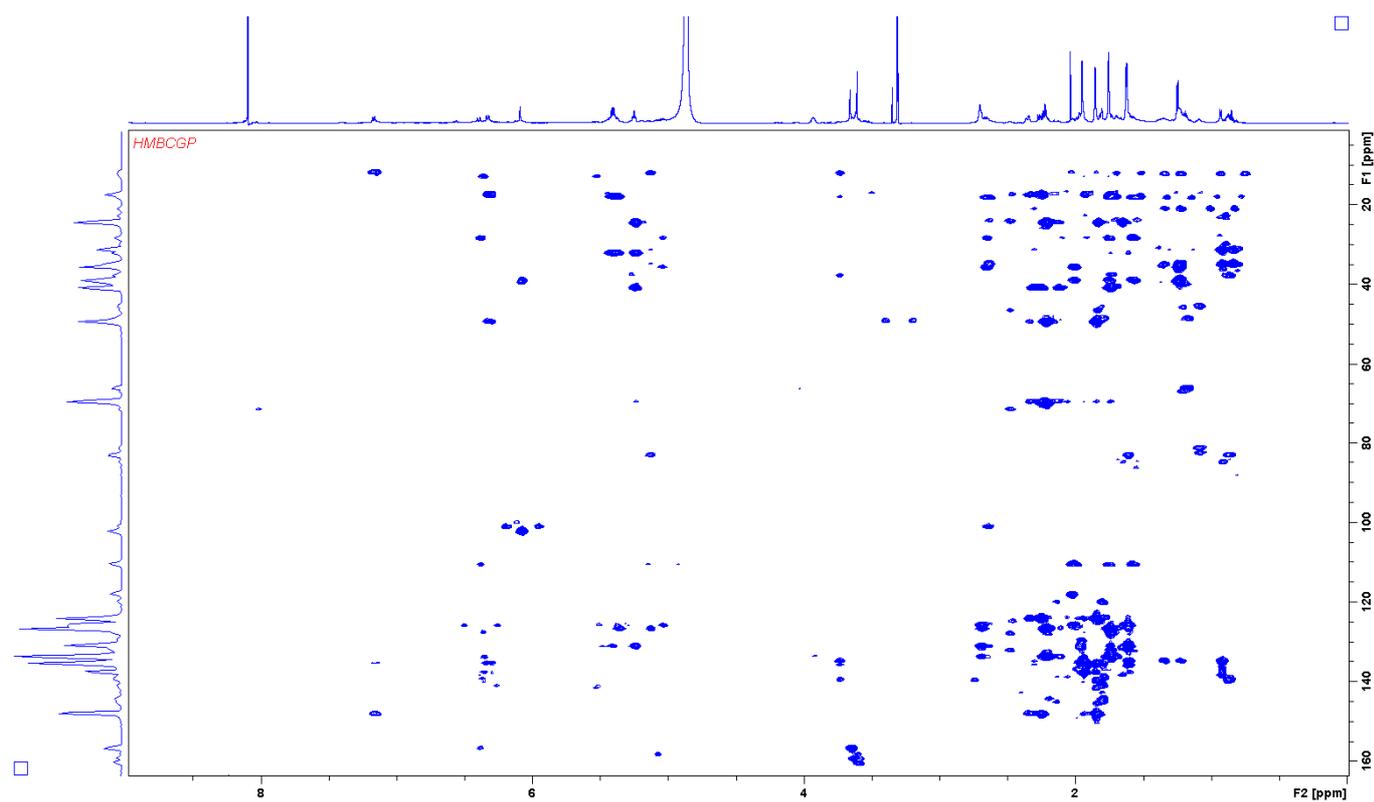


Figure S18. HMBC NMR spectrum of COR D in methanol- $d_4$ .

### 3.7.11.3 NMR spectra of oxyCOR A

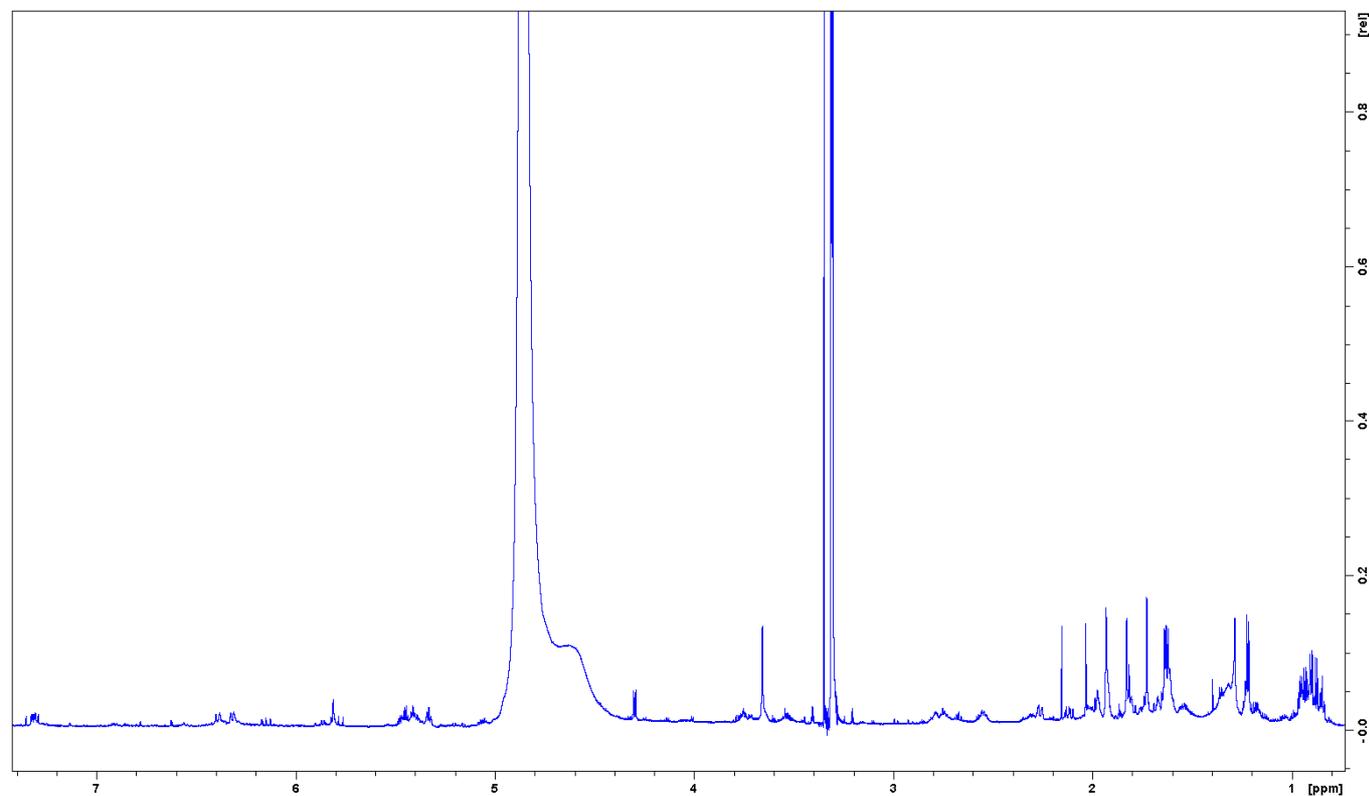
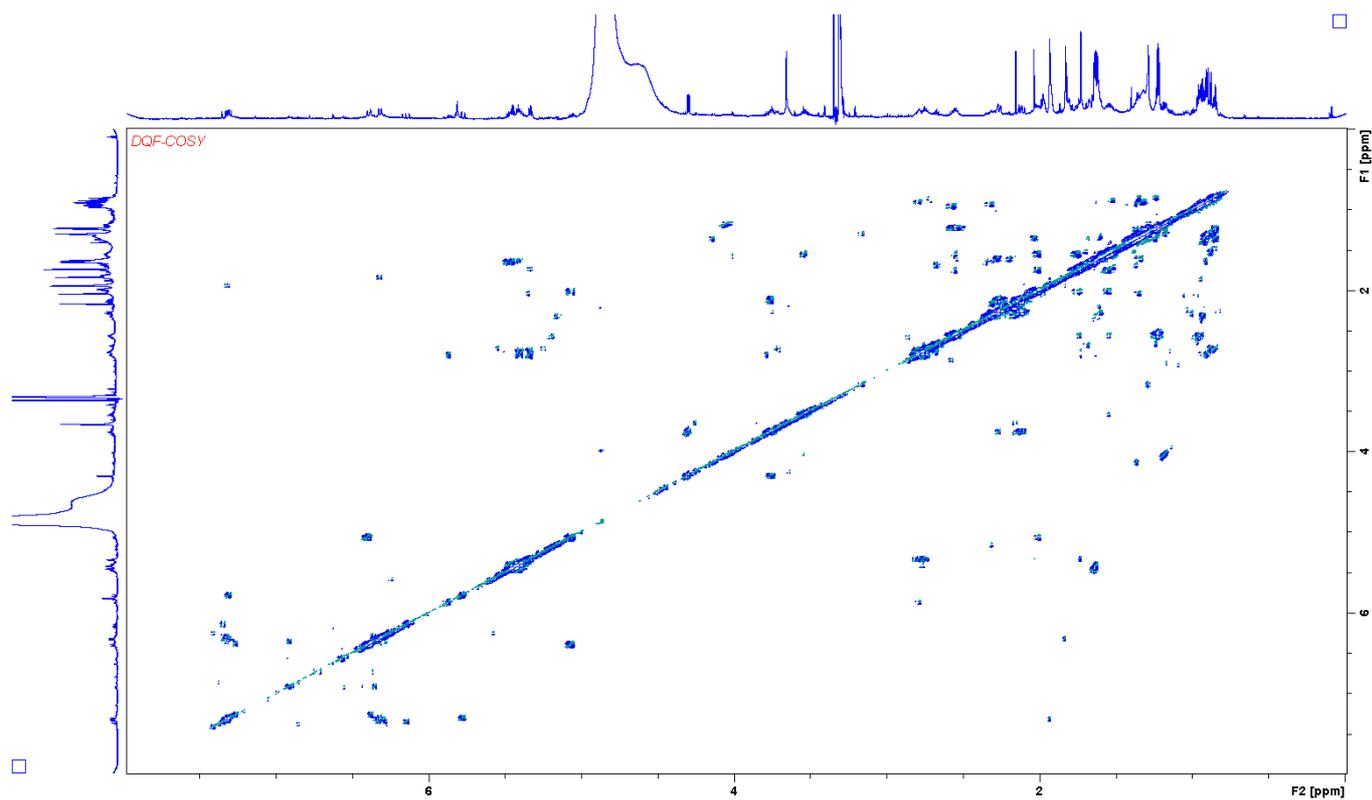
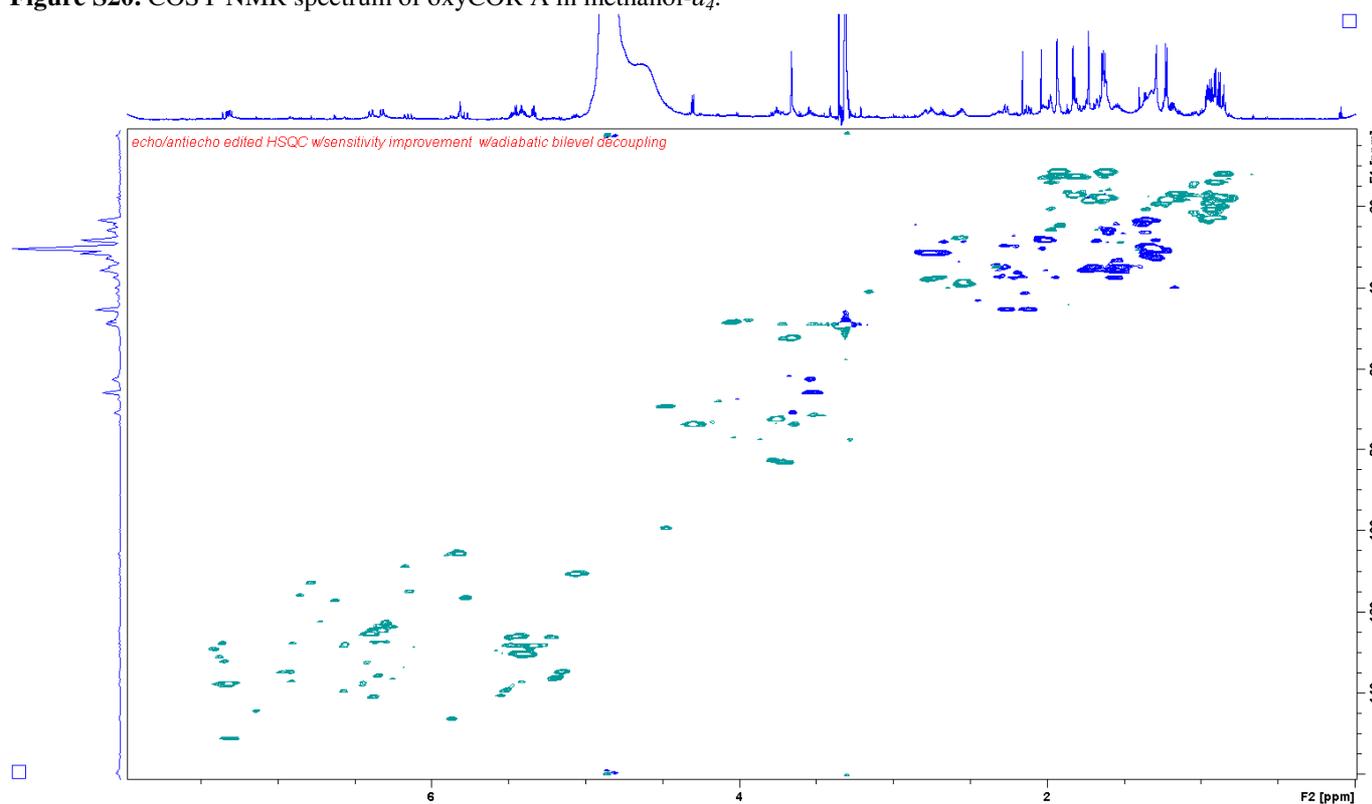


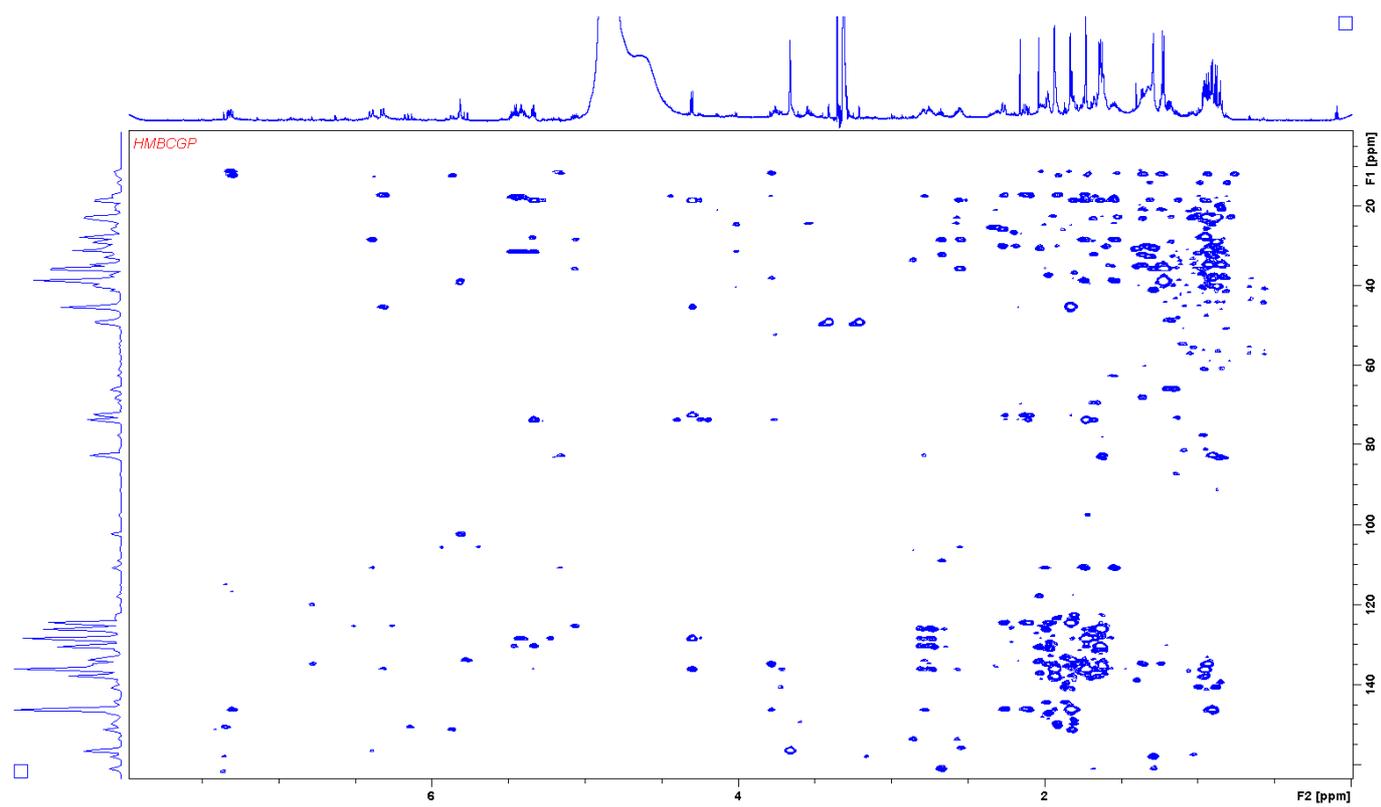
Figure S19.  $^1\text{H}$  NMR spectrum of oxyCOR A in methanol- $d_4$ .



**Figure S20.** COSY NMR spectrum of oxyCOR A in methanol- $d_4$ .



**Figure S21.** HSQC NMR spectrum of oxyCOR A in methanol- $d_4$ .



**Figure S22.** HMBC NMR spectrum of oxyCOR A in methanol- $d_4$ .

### 3.7.12 References

- (1) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp., *Gene*. **1992**, *116*, pp. 43–49.
- (2) Sucipto, H.; Pogorevc, D.; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, *44*, pp. 160–170.
- (3) Magrini, V.; Creighton, C.; Youderian, P. Site-specific recombination of temperate *Myxococcus xanthus* phage Mx8: Genetic elements required for integration, *J. Bacteriol.* **1999**, *181*, pp. 4050–4061.
- (4) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (5) Kusebauch, B.; Busch, B.; Scherlach, K.; Roth, M.; Hertweck, C. Functionally distinct modules operate two consecutive  $\alpha,\beta$ - $\rightarrow\beta,g$  double-bond shifts in the rhizoxin polyketide assembly line, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, pp. 1460–1464.
- (6) Kampa, A.; Gagunashvili, A. N.; Gulder, T. A. M.; Morinaka, B. I.; Daolio, C.; Godejohann, M.; Miao, V. P. W.; Piel, J.; Andrésson, Ó. S. Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E3129-37.
- (7) Keatinge-Clay, A. T. A tylosin ketoreductase reveals how chirality is determined in polyketides, *Chem. Biol.* **2007**, *14*, pp. 898–908.
- (8) Reid, R.; Piagentini, M.; Rodriguez, E.; Ashley, G.; Viswanathan, N.; Carney, J.; Santi, D. V.; Hutchinson, C. R.; McDaniel, R. A model of structure and catalysis for ketoreductase domains in modular polyketide synthases, *Biochemistry*. **2003**, *42*, pp. 72–79.
- (9) Kitsche, A.; Kalesse, M. Configurational assignment of secondary hydroxyl groups and methyl branches in polyketide natural products through bioinformatic analysis of the ketoreductase domain, *ChemBioChem*. **2013**, *14*, pp. 851–861.
- (10) Erol, Ö.; Schäberle, T. F.; Schmitz, A.; Rachid, S.; Gurgui, C.; El Omari, M.; Lohr, F.; Kehraus, S.; Piel, J.; Müller, R.; König, G. M. Biosynthesis of the myxobacterial antibiotic corallopyronin A, *ChemBioChem*. **2010**, *11*, pp. 1235–1265.
- (11) Lomas, J. S. <sup>1</sup>H NMR spectra of ethane-1,2-diol and other vicinal diols in benzene. GIAO/DFT shift calculations, *Magn. Reson. Chem.* **2013**, *51*, pp. 32–41.



## Chapter 4

### **Biosynthesis and heterologous production of argyrins**

Domen Pogorevc<sup>†</sup>, Ying Tang<sup>†</sup>, Michael Hoffmann, Gregor Zipf, Hubert S. Bernauer,  
Alexander Popoff, Heinrich Steinmetz and Silke C. Wenzel

<sup>†</sup> These authors contributed equally to this work

(Submitted manuscript)

## **Contributions to the Presented Work**

### **Author's Contribution**

The author significantly contributed to the conception of the study, designed and performed experiments, evaluated and interpreted the resulting data. The author constructed argyrim heterologous expression constructs with modified A1 domains as well as the constructs with modified A4 domains. The author also constructed the *Bsa*I free argyrim expression construct. Heterologous expression of those construct in *Myxococcus xanthus* DK1622 was conducted by the author. Furthermore, the author contributed to the cultivation of heterologous argyrim production strains and performed the mutasynthesis studies. Production yield optimization by medium improvement was also conducted by the author. The author contributed to the preparation of the extracts and quantification of the production yields. The author contributed to the conception and writing of the manuscript.

### **Contribution by Co-Workers**

Ying Tang significantly contributed to the conception of the study, designed and performed experiments, evaluated and interpreted the resulting data. She also contributed to the design and construction of argyrim expression constructs based on the synthetic DNA. Furthermore, she contributed to the cultivation of the heterologous argyrim production strains, preparation of the extracts and writing of this manuscript. Michael Hoffmann performed the HPLC-MS analysis of the samples. Gregor Zipf and Hubert S. Bernauer contributed to the analysis of the argyrim biosynthetic pathway and the design of the synthetic DNA constructs. Heinrich Steinmetz isolated argyrins I and J and performed the structure elucidation. Alexander Popoff assigned the chemical structures of argyrim I and J based on the NMR data. David C. B. Siebert, Roman Sommer and Alexander Titz performed the synthesis of argyrim mutasynthons. The project was supervised by Silke C. Wenzel. She contributed to the identification and characterization of the argyrim biosynthetic pathway, design of heterologous expression constructs based on the synthetic DNA and the design of synthetic sequences for the A1 domain engineering. She contributed to the writing of this manuscript, was responsible for the conception of the project and performed the proofreading of this chapter.

## 4 Biosynthesis and heterologous production of argyrins

### 4.1 Abstract

Argyrins represent a family of cyclic octapeptides exhibiting promising antimicrobial, antitumorigenic and immunosuppressant activities. They derive from a nonribosomal peptide synthetase pathway, which was identified and characterized in this study from the myxobacterial producer strain *Cystobacter sp.* SBCb004. Using the native biosynthetic gene cluster (BGC) sequence as template synthetic BGC versions were designed and assembled from gene synthesis fragments. A heterologous expression system was established after chromosomal deletion of a well-expressed lipopeptide pathway from the host strain *Myxococcus xanthus* DK1622. Different approaches were applied to engineer and improve heterologous argyirin production, which was finally increased to 160 mg/l, around 20-fold higher yields compared to the native producer. The optimized production system provides a versatile platform for future supply of argyrins and novel derivatives thereof.

## 4.2 Introduction

Microbial natural products have historically been proven as invaluable source for the development of therapeutic agents, especially in the field of anti-infective research<sup>1,2</sup>. They possess exceptional molecular diversity that cannot be matched by synthetic libraries of small molecules and exhibit a wide variety of biological activities<sup>3,4</sup>. A prolific source for novel natural products are myxobacteria,  $\delta$ -proteobacteria that are mostly soil-dwelling and can be found in various habitats<sup>5,6</sup>. Although the number of myxobacterial species known today is still limited<sup>7,8</sup>, numerous highly bioactive compounds have been already isolated<sup>9–11</sup>. Currently, one drug derived from a myxobacterial product, ixabepilone, has been approved for clinical use<sup>12,13</sup>, and several other compounds are being evaluated in preclinical studies due to their promising mode of action<sup>14–18</sup>. These also include the argyrins produced from *Archangium* and *Cystobacter sp.*<sup>19,20</sup>, which are also known from actinomycetes as Antibiotics A21459<sup>21,22</sup>.

Argyrins represent a family of cyclic octapeptides, for which eight derivatives, argyirin A-H, have been elucidated so far<sup>20</sup>. Initial bioactivity studies characterized argyrins as potent immunosuppressants, which also inhibit the growth of mammalian cell cultures and show selective antimicrobial activity by inhibition of bacterial protein synthesis<sup>19,22</sup>. More recently it was shown that argyrins target the elongation factor G (EF-G) in bacteria and its closest homologue, the mitochondrial elongation factor (EF-G1), in yeast and mammalian cells<sup>23,24</sup>. Elongation factor G is also inhibited by the clinically used antibiotic fusidic acid, however, argyirin B binds to EF-G at a different site revealing a new mode of protein synthesis inhibition<sup>23</sup>. Therefore, and considering the intriguing antibacterial spectrum<sup>23,24,19,25</sup>, which e.g. includes modest activity against the Gram-negative pathogen *Pseudomonas aeruginosa*, argyrins represent promising tool compounds for the development of novel antibiotics. In addition, argyrins are under investigation due to their immunosuppressive and antitumorigenic activities. Argyrin A was described to induce apoptosis and block angiogenesis by preventing the destruction of the cyclin kinase inhibitor p27<sup>kip1</sup> through proteasome inhibition<sup>26,27</sup>. Continuing structure-activity-relation studies identified argyirin F as most promising derivative for development of an antitumor drug<sup>28</sup>, which was further characterized by *in vivo* studies showing that argyirin F treatment might be useful as an additional therapy for pancreatic adenocarcinoma<sup>18</sup>.

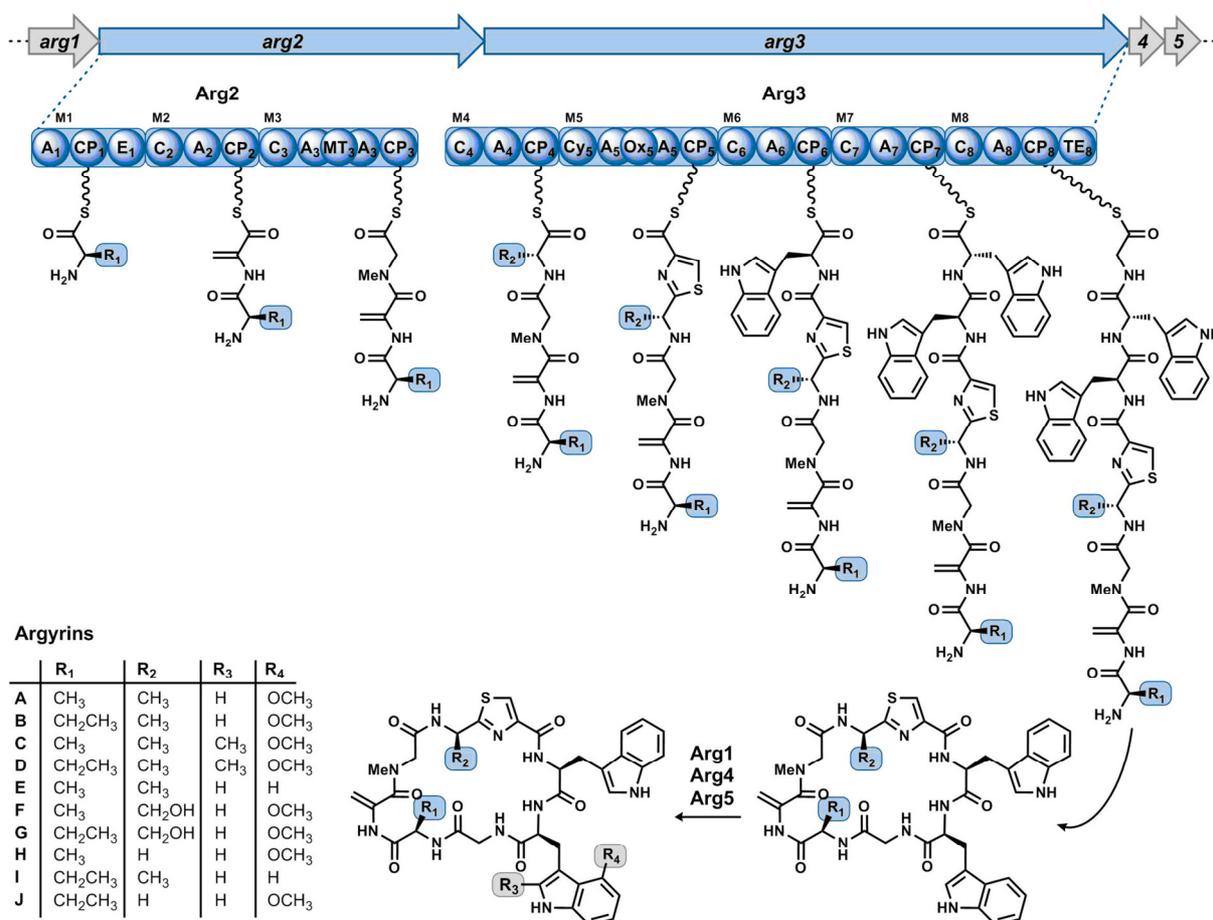
Most studies on elucidation of the biological activity of argyrins were pursued with the major natural derivatives argyirin A and B, whereat argyirin F, for example, represents a very minor

product in the native producer strain *Archangium gephyra* Ar8082<sup>19</sup>. Although total synthesis approaches have been already developed for generation of argyrins and novel derivatives<sup>28–30</sup>, the establishment of efficient and cost-effective supply routes is highly demanded for further investigation and development of this promising compound class. Argyrins biosynthetically derive from amino acids and were assumed to be synthesized by a nonribosomal peptide synthetase (NRPS) system. NRPS represent large multifunctional megasynthetases involved in the assembly of numerous structurally diverse compounds among them several marketed drugs, e.g. penicillin, vancomycin, cyclosporin<sup>31,32</sup>. The corresponding genes for production of a particular compound family are usually clustered within the chromosome of the microbial producer strain. In this study, we first aimed to elucidate argyirin biosynthesis by identification and analysis of the corresponding biosynthetic gene cluster (BGC). Next we envisaged the establishment of a versatile microbial production platform to facilitate future exploitation of argyrins and novel derivatives thereof. As genetic engineering of the native argyirin producer strains is difficult we intended to heterologously express the *arg* BGC in a derivative of the myxobacterial model strain *Myxococcus xanthus*, which turned out to be a suitable host for expression of other natural product pathways from myxobacteria in the past<sup>33</sup>. Synthetic biology approaches were applied to design and construct different synthetic versions of the argyirin BGC based on synthetic DNA and to engineer the argyirin production spectrum.

### 4.3 Results and Discussion

#### 4.3.1 Identification and characterization of the argyirin biosynthetic pathway

To identify the argyirin biosynthetic gene cluster (*arg* BGC) genome sequence data of *Cystobacter* sp. SBCb004<sup>34</sup>, a characterized argyirin producer from our myxobacterial strain collection, were analyzed. Using bioinformatics tools for detection of nonribosomal peptide synthetase (NRPS) encoding genes<sup>35,36</sup> a BGC encoding an expected octamodular NRPS assembly line was identified. The putative *arg* BGC is about 30 kb in size and consists of five genes, *arg1-5* (Fig. 1). Besides the two NRPS subunits Arg2 and Arg3, the gene products include three putative modifying enzymes: a radical SAM-dependent methyltransferase (Arg1), an *O*-methyltransferase (Arg4) and a tryptophan 2,3-dioxygenase (Arg5) (Table S1). To verify the involvement of the identified BGC in argyirin biosynthesis *arg3* from *Cystobacter* sp. SBCb004 was inactivated by integration of a suicide plasmid, whereby production of argyrins was abolished (Fig. S2/S3).



**Fig. 1.** Argyrin BGC, the encoded NRPS assembly line and corresponding products. The BGC is around 30 kb in size and comprises five genes (shown as arrows; Table S1). The NRPS assembly line consists of two subunits harboring eight modules. Catalytic domains are illustrated as circles and numbered according to module order. Domain abbreviations: A, adenylation domain; CP, carrier protein domain (also known as peptidyl carrier protein (PCP) or thiolation (T) domain); E, epimerization domain; MT, methyltransferase domain; Cy (=Cyc), heterocyclization domain; Ox, oxidation domain; TE, thioesterase domain. Argyrins A-H were described previously<sup>20</sup>, argyrins I and J were isolated and characterized in this study (Fig. S7-S8; Tables S6/S7).

Based on the NRPS domain arrangement and predicted substrate specificities of the adenylation (A) domains (Table S2) it was concluded that argyirin biosynthesis is initiated at amino acid residue R<sub>1</sub> (*D*-alanine or *D*- $\alpha$ -aminobutyrate; Fig. 1). C-terminal elongation with seven additional amino acids yields an octapeptide chain that is finally released from the NRPS megasynthetase by thioesterase (TE)-mediated macrolactamization. The structural diversity of the natural argyirin derivatives characterized so far, argyirin A-H (Fig. 1;<sup>20</sup>) results from different amino acids implemented by modules 1 and 4 as well as different modification stages at the tryptophan residue incorporated by module 7. The latter is assumed to be *C*-methylated via a radical mechanism by Arg1 (R<sub>3</sub> = CH<sub>3</sub>) and oxygenated by Arg5 followed by Arg4-mediated *O*-methylation (R<sub>4</sub> = OCH<sub>3</sub>). These tailoring steps probably occur after peptide core assembly, however, transformations at an earlier stage of the biosynthesis cannot

be excluded based on current data. Argyrins C and D feature a complete tryptophan substitution pattern ( $R_3 = \text{CH}_3$ ,  $R_4 = \text{OCH}_3$ ), whereas the peptide core from argyrin E is not altered in these positions ( $R_3, R_4 = \text{H}$ ). All other natural argyrin derivatives known so far harbor a partially modified tryptophan residue ( $R_4 = \text{H}$ ,  $R_3 = \text{OCH}_3$ ) indicating that radical C-methylation represents a bottleneck in the evaluated producer strains under the applied cultivation conditions.

In addition to the tryptophan derivative, argyrins harbor several non-proteinogenic amino acid building blocks including  $\alpha$ -aminobutyrate (Abu), dehydroalanine (Dha) as well as a thiazole unit that are also known from other NRPS biosynthesis pathways<sup>37</sup>. Thiazole rings are built after peptide chain elongation with cysteine via heterocyclization and subsequent oxidation<sup>38</sup>, the required heterocyclization (Cyc; Fig. S1) and oxidation (Ox) domains could be identified in module 5 of the argyrin megasynthetase. Dehydroalanine units are less common in NRPS products, and are often further converted during the biosynthesis through inter- or intramolecular Michael addition<sup>37,39</sup>. As recently shown for nocardicin they can be generated from thioester-bound seryl residues on the NRPS assembly line in *cis*<sup>40</sup>, which might be a possible scenario for argyrin biosynthesis as bioinformatics analysis of the *arg* BGC and flanking region did not point to candidate auxiliary enzymes known from other NRPS pathways (e.g. viomycin<sup>41</sup>, staphyloferrin<sup>42</sup>), or from posttranslational modifications in ribosomal peptide biosynthesis<sup>43</sup>. A feeding experiment with *Cystobacter sp.* SBCb004 confirmed the biosynthetic origin of the argyrin Dha unit from serine (Fig. S4), which corresponds with the serine substrate specificity predicted for the A domain from module 2 (Table S2). However, similar to other examples for which Dha unit biosynthesis remains elusive (e.g. bleomycin<sup>44</sup> the characteristic core motif described for the downstream condensation (C) domain in nocardicin biosynthesis<sup>40</sup> was not identified in the argyrin megasynthetase. Phylogenetic C domain analysis revealed that some examples of C domains from  $\alpha,\beta$ -dehydro amino acid forming pathways, including the C domain from module 3 of the argyrin NRPS, represent a distinct new clade besides the six C domain subtypes distinguished by Rausch et al.<sup>45</sup> (Fig. S1). This indicates that the module 3 C domain is involved in Dha unit generation when processing the seryl-dipeptide intermediate from module 2. In contrast to the Dha and thiazole units, which are present in all natural argyrin derivatives characterized so far, the Abu unit ( $R_1 = \text{CH}_2\text{CH}_3$ ) contributes to the structural diversity of argyrins. In order to confirm that this unit originates from activation and incorporation of Abu by module 1 instead of *L*-alanine ( $R_1 = \text{CH}_3$ ) that is subsequently methylated during biosynthesis (e.g. by the radical SAM-dependent methyltransferase Arg1),

a feeding experiment with *Cystobacter sp.* SBCb004 was performed. As isotopically labelled Abu was not available a metabolic precursor, *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N, was used<sup>46</sup>, which indeed resulted in an argyryn B labelling pattern expected for Abu incorporation (Fig. S5). Due to the presence of an epimerization (E) domain in the first NRPS module and a downstream *D*-specific C domain (<sup>D</sup>C<sub>L</sub>-type according to Rausch et al.<sup>45</sup>; Fig. S1) *D*-configured Abu and alanine units are finally incorporated into the argyryn peptide core.

The alanine (R<sub>2</sub> = CH<sub>3</sub>) and serine (R<sub>2</sub> = CH<sub>2</sub>OH) units incorporated by module 4 also exhibit *D*-configuration in the final argyryn derivatives. According to the argyryns known so far, alanine seems to be the preferred substrate, which correlates with the *in silico* prediction of the A<sub>4</sub> substrate specificity (Table S2), whereas serine and glycine (R<sub>2</sub> = H) were only observed in minor derivatives. However, in contrast to the first module, module 4 does not harbor an E domain for generation of *D*-configured intermediates. According to textbook NRPS biochemistry *D*-configured building blocks can also result from the presence of a dual C/E domain, e.g. as discovered for the arthrofactin megasynthetase<sup>47</sup>, or from direct incorporation of *D*-amino acids by gate-keeping A domains, e.g. as observed for cyclosporin<sup>48,49</sup>, which depends on sufficient metabolic supply of *D*-amino acids (*D*-alanine and *D*-serine for argyryn biosynthesis) as well as the presence of a <sup>D</sup>C<sub>L</sub>-type domain processing *D*-configured donor substrates. Both scenarios require downstream C domains generating and/or processing *D*-configured amino acids. Bioinformatics analysis according to Rausch et al.<sup>45</sup> did neither lead to the identification of dual C/E-type nor <sup>D</sup>C<sub>L</sub>-type domains in downstream modules of the argyryn NRPS (Fig. S1). The immediately downstream module 5 installs a thiazole unit and its *N*-terminal domain (Cy<sub>5</sub>) groups with Cyc domains in the applied phylogenetic analysis approach. It might be possible that this domain harbors a donor site accepting *D*-configured substrates required for both scenarios discussed above and probably additionally exhibits epimerization activity. However, modifications at a later stage of the biosynthesis cannot be excluded based on current data.

In accordance with the established biosynthesis model two additional argyryn derivatives, argyryn I and J (Fig. S6) were isolated during re-evaluation of extract fractions from previous large scale fermentation studies<sup>19</sup>. The planar structures were elucidated by NMR spectroscopy (Tables S6/S7; Fig. S7/S8) and it is proposed that the stereochemistry correlates with previously described derivatives as argyryn I and J represent the corresponding “Abu-derivatives” (R<sub>1</sub> = CH<sub>2</sub>CH<sub>3</sub>) of argyryns E and H. Argyryns A and B are produced as major products in *A. gephyra* Ar8082<sup>19</sup> and *Cystobacter sp.* SBCb004 (Fig. S3), whereas argyryns C-J represent minor derivatives.

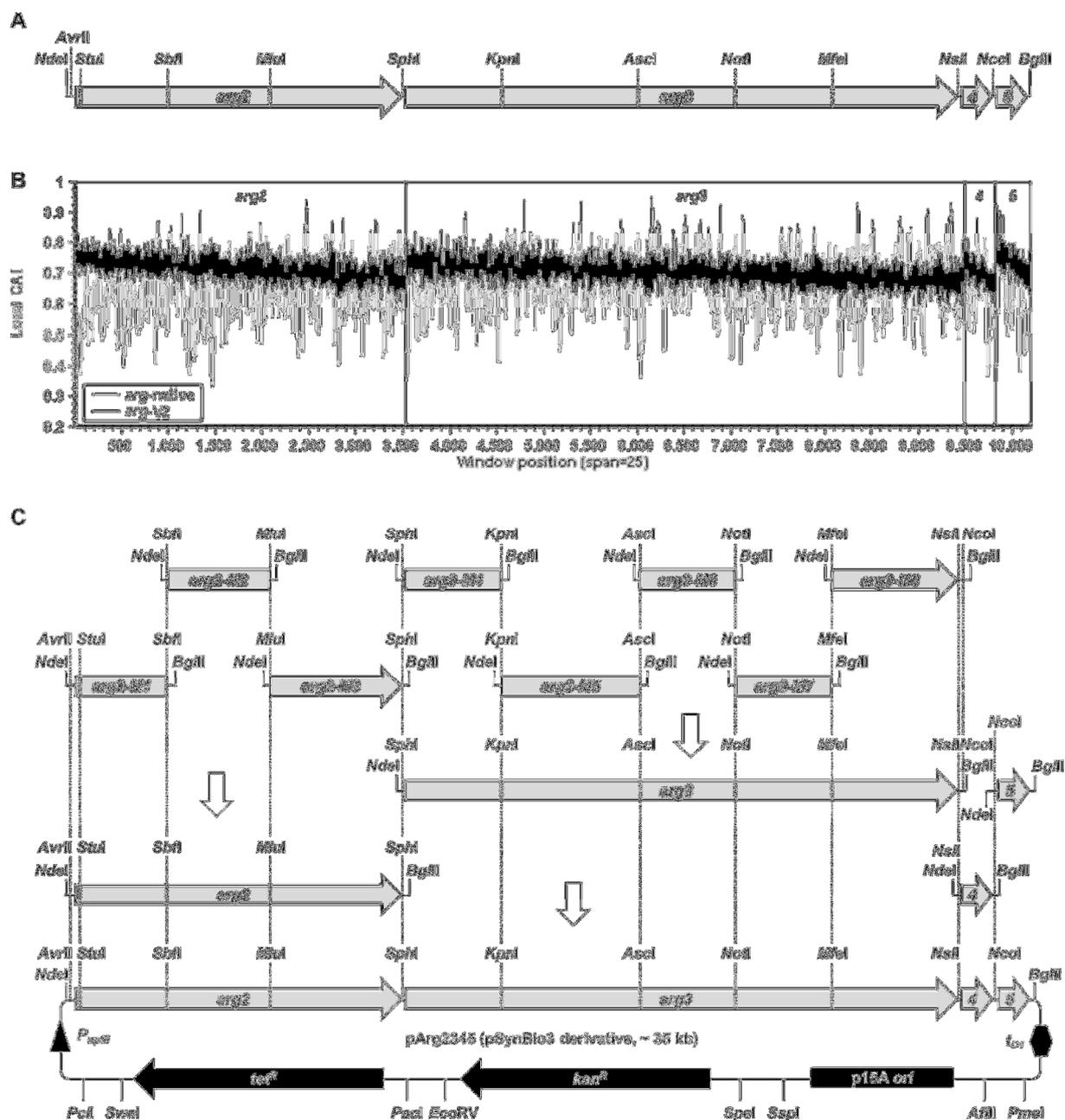
### 4.3.2 Establishment of a heterologous expression system based on synthetic DNA

In order to establish a versatile production platform for argyrins we aimed at heterologous expression of the pathway in a well-established host for myxobacterial megasynthetases, *Myxococcus xanthus*<sup>33</sup>. We intended to generate BGC DNA fragments via gene synthesis, which allows for modulation of *arg* sequences according to requirements relevant for BGC assembly, engineering and heterologous expression (Fig. 2).

For *arg* BGC design, the native pathway from *Cystobacter sp.* SBCb004 was used as blueprint, truncated to the *arg2-4* gene set to simplify the production spectra by avoiding partial methylation by Arg1. In analogy to the native pathway, the synthetic BGC was configured as single transcription unit by using promoter and terminator sequences established for *M. xanthus*, which were implemented in the expression vector backbone as described below. A key step in the constructional design process was the engineering of endonuclease restriction sites (R-sites; Fig. 2A) according to predefined requirements. This included the implementation of unique R-sites upstream and downstream of each gene as well as in between each module of the NRPS encoding genes *arg2* and *arg3* to facilitate BGC assembly and modifications. In addition, a unique R-site was introduced at the 5' end of *arg2* to open possibilities for leader sequence exchange and other engineering approaches. Interfering R-sites along the BGC sequence including those selected for vector backbone engineering (Table S8) were eliminated by point mutations, which were based on synonymous codon substitutions in coding regions. During the sequence modulations, two different versions of *arg* gene sets (*arg-V1* and *arg-V2*) were designed. For both versions the exact same R-sites engineering was applied, but *arg-V2* genes were additionally adapted according to sequence parameters that are usually addressed in formal gene optimization approaches. This included the avoidance of homopolymer stretches and DNA repeats, increase of stop codons in unused reading frames, reduction of Shine-Dalgarno (SD)-anti-SD interactions and codon adaptation to the heterologous environment comprising elimination of rare codon clusters as well as modulation of the local codon adaptation index (CAI). The resulting course of the local CAI shows a smoothed shape without any distinct peaks and a slight gradient along each coding DNA sequence in *arg-V2* compared to the native sequence (Fig. 2B). For codon usage optimization, a similar approach as described in our previous study on the heterologous expression of a synthetic epothilone BGC in *M. xanthus* was applied for *arg-V2*, using a set of PKS/NRPS encoding genes from the host as reference<sup>50</sup>.

In parallel, a suitable expression vector backbone, pSynBio3 (Fig. 2C), was designed harboring the low copy p15A origin of replication as well as kanamycin resistance gene (*kan<sup>R</sup>*) from pACYC177<sup>51</sup>, and the tetracycline resistance gene (*tet<sup>R</sup>*) from pACYC184<sup>52</sup> for integration into the host chromosome as described below. In addition the vector features a customized *P<sub>nptII</sub>*-MCS-*t<sub>DI</sub>* box for *arg* BGC assembly at the multiple cloning site (MCS) and subsequent heterologous expression in *M. xanthus* controlled by the *P<sub>nptII</sub>* promoter of the neomycin phosphotransferase II resistance gene from Tn5<sup>53</sup> and *t<sub>DI</sub>* terminator from a *M. xanthus* DK1622 tRNA gene<sup>54</sup>. Genetic elements in pSynBio3 were flanked by unique R-sites to enable future backbone engineering as well as the initial vector construction based on synthetic DNA fragments. In total, 22 restriction enzymes were selected for vector and *arg* BGC engineering (Table S8), and interfering R-sites were eliminated from the pSynBio3 sequence in similar manner as described for *arg* BGC modulation.

After completion of the design process, synthetic DNA fragments covering the two *arg* BGC versions as well as the pSynBio3 expression vector were generated via gene synthesis. The latter was constructed in one step from two synthetic DNA fragments (pGH-15A-kan and pGH-tet-nptII-MCS-TD1; Table S4) by applying conventional restriction and ligation methods (Table S5). The two modulated *arg* BGC sequences (*arg*-V1 and *arg*-V2) were dissected into ten fragments, whereat individual genes were separated and NRPS encoding genes *arg2/arg3* were fragmented according to their modular architecture. In addition to the unique R-sites engineered during the BGC design, each gene synthesis fragment was flanked by *Nde*I and *Bgl*II R-sites to facilitate the assembly process by applying conventional cloning techniques (Fig. 2C, Table S5). First of all, the NRPS encoding genes were reconstituted stepwise based on three (*arg2*) and five (*arg3*) gene synthesis fragments to yield constructs pGH-arg2-V1/V2 and pGH-arg3-V1/V2, which were verified by sequencing. Next, *arg* BGCs were sequentially assembled in the expression vector pSynBio3 based on the synthetic *arg* gene library revealing the final expression constructs pArg2345-V1/V2, which were verified by sequencing as well, before proceeding with the heterologous expression studies.



**Fig. 2.** Design and construction of synthetic *arg* BGCs and expression constructs based on the pSynBio3 vector. (A) Result of R-sites engineering applied for *arg*-V1/V2. Along the 31 kb native *arg2-arg5* BGC sequence 13 constructional R-sites were introduced and 59 interfering R-sites were eliminated (for restriction enzyme list see Table S8). (B) Comparison of the course of the local codon adaptation index (CAI) of the native (*arg*-native) and modulated (*arg*-V2) *arg2-arg5* gene sequences. The native sequence is shown in the background (gray), and the artificial sequence (*arg*-V2) is shown in black. Window width is 25 codons. In *arg*-V2, the gradient between start and stop codon of each coding DNA sequence (CDS) was adjusted to 0.05. Untranslated regions between the CDS were excluded from the graph. (C) Gene synthesis fragments and subsequent assemblies: first reconstitution of *arg2* and *arg3*, followed by BGC assembly at the multiple cloning site (MCS) of pSynBio3. The same assembly procedure was applied for generation of pArg2345-V1 and V2 (for details see Table S5). Expression vector features (not illustrated in scale): p15A *ori*, low copy origin of replication; *kan<sup>R</sup>*, kanamycin resistance gene (for selection); *tet<sup>R</sup>*, tetracyclin resistance gene (for chromosomal integration into *M. xanthus* DK1622  $\Delta$ *mchA-tet*); *P<sub>nptII</sub>*, constitutive promoter for *arg*-BGC expression; *t<sub>DJ</sub>*, terminator.

As host strain we aimed to employ the myxobacterial model strain *Myxococcus xanthus* DK1622, which led to the most promising yields in previous expression studies with

myxobacterial megasynthetases<sup>33</sup>. However, to simplify the production spectra and to probably cause additional beneficial effects on heterologous argyirin biosynthesis (e.g. reduce precursor competition) we first deleted the BGC that is in charge of lipopeptide production in the host (*mch<sub>A</sub>* BGC,<sup>55</sup>). The corresponding products, myxochromides A, represent one of the most abundant compound families produced in *M. xanthus* under standard cultivation conditions<sup>56</sup>. Deletion of the *mch<sub>A</sub>* BGC was achieved by replacement with a tetracycline resistance gene (*tet<sup>R</sup>*) via homologous recombination at flanking *mch<sub>A</sub>* BGC regions resulting in the mutant strain *M. xanthus* DK1622  $\Delta$ *mchA-tet* (Fig. S10). The engineered strain, which lacks myxochromide A production, was recently also employed in parallel expression studies on myxobacterial  $\alpha$ -pyrone antibiotics (myxopyronin, coralopyronin), different lipopeptides (myxochromides) as well as bioactive peptolides (vioprolides) underpinning its suitability as host<sup>57-59</sup>. The chromosomally integrated *tet<sup>R</sup>* gene provides a hub for integration of heterologous BGCs via homologous recombination of expression plasmids harboring a *tet<sup>R</sup>* gene (fragment), e.g. as considered in the design of the pSynBio3 expression vector (Fig. 2C). Transformation of the engineered *M. xanthus* strain with the generated *arg* BGC expression constructs pArg2345-V1/V2 resulted in the expected mutant strains (Table 1), which was verified by genotypic analysis (Fig. S10).

The generated expression strains *M. xanthus* DK1622  $\Delta$ *mchA-tet::pArg2345-V1/V2* as well as the native producer *Cystobacter sp.* SBCb004 were cultured under standard conditions for analysis of argyirin production. With both synthetic *arg* BGC versions successful production was achieved, yielding argyirin A and B as major products similar to the native producer. Production levels were significantly higher in the *arg*-V1 expression strain (35 mg/l argyirin A/B with pArg2345-V1) compared to *arg*-V2 (~100  $\mu$ g/l argyirin A/B with pArg2345-V2) for which additional parameters were applied during sequence modulation. The *arg*-V2 design represents the second example for a codon usage optimization strategy applied in a previous study on a synthetic epothilone BGC, for which production yields in similar range were observed (~100  $\mu$ g/l;<sup>50</sup>). This illustrates that the establishment of rational sequence modulation approaches for heterologous expression of giant megasynthetase complexes still remains a challenging task, which requires further insights on relevant criteria to improve production yields. However, the approach applied for *arg*-V1 focusing on the constructional design and replacement of the native promoter against *P<sub>npIII</sub>* revealed promising production yields, which are higher compared to the yields achieved with the native producer *Cystobacter sp.* SBCb004 (8 mg/l). In addition, the native producer is genetically less accessible and marker-less modifications for rational engineering approaches have not been established yet. Therefore,

the successful establishment of a heterologous production platform provides considerable progress for biotechnological production of argyrins and novel derivatives thereof. As expected, the minor derivatives argyrins C and D were not produced in the heterologous system due to the absence of *argI*. However, two novel argyirin derivatives, assigned as argyrins K and L, were detected in extracts of the expression strain *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pArg2345-V1 (Fig. 3). The corresponding MS data pointed to the respective valine derivatives of argyrins A and F (with  $R_1 = CH_2-(CH_3)_2$  instead of  $CH_3$ ), which is in accordance with the A1 domain specificity analysis (Table S2) and was further supported by feeding experiments with deuterium-labelled valine (Fig. S11).

**Table 1.** *Myxococcus xanthus* (*Mx.*) expression strains generated in this study (for details on construction of expression plasmids see Table S5).

| Strain name   | Expressed <i>arg</i> BGC   | Major (minor) derivatives                      |
|---|--|--|
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1      | $P_{npIII}(arg2-arg3-arg4-arg5)-V1-t_{D1}$                                     | Argyrins A, B (argyrins E-L)                   |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V2      | $P_{npIII}(arg2-arg3-arg4-arg5)-V2-t_{D1}$                                     | Argyrins A, B (n.a.) <sup>1</sup>              |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg23-V1        | $P_{npIII}(arg2-arg3)-V1-t_{D1}$   | Argyrins I, E (argyrins G3, F3)                |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg235-V1       | $P_{npIII}(arg2-arg3-arg5)-V1-t_{D1}$  | Argyrins I, E (G3, F3, A2)                     |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1-BsaI | $P_{npIII}(arg2-arg3-arg4-arg5)-V1-BsaI-t_{D1}$                                | Argyrins A, B (argyrins E-L)                   |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1-mod1 | $P_{npIII}(arg2^*-arg3-arg4-arg5)-V1-t_{D1}$<br>* A1 mutation mod1 (Table S11) | Argyrins A, B (n.a.) <sup>1</sup>              |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1-mod2 | $P_{npIII}(arg2^*-arg3-arg4-arg5)-V1-t_{D1}$<br>* A1 mutation mod2 (Table S11) | Argyrins A, B (n.a.) <sup>1</sup>              |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1-mod8 | $P_{npIII}(arg2^*-arg3-arg4-arg5)-V1-t_{D1}$<br>* A1 mutation mod8 (Table S11) | Argyirin A, B (n.a.) <sup>1</sup>              |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1-mod9 | $P_{npIII}(arg2^*-arg3-arg4-arg5)-V1-t_{D1}$<br>* A1 mutation mod9 (Table S11) | Argyirin A, B, K (n.a.) <sup>1</sup>           |
| <i>Mx.</i> DK1622 $\Delta mchA$ - <i>tet</i> ::pArg2345-V1-BsaI | $P_{npIII}(arg2-arg3-arg4-arg5)-V1-t_{D1}$                                     | Argyrins A, B, (I) <sup>2</sup> (argyrins E-L) |

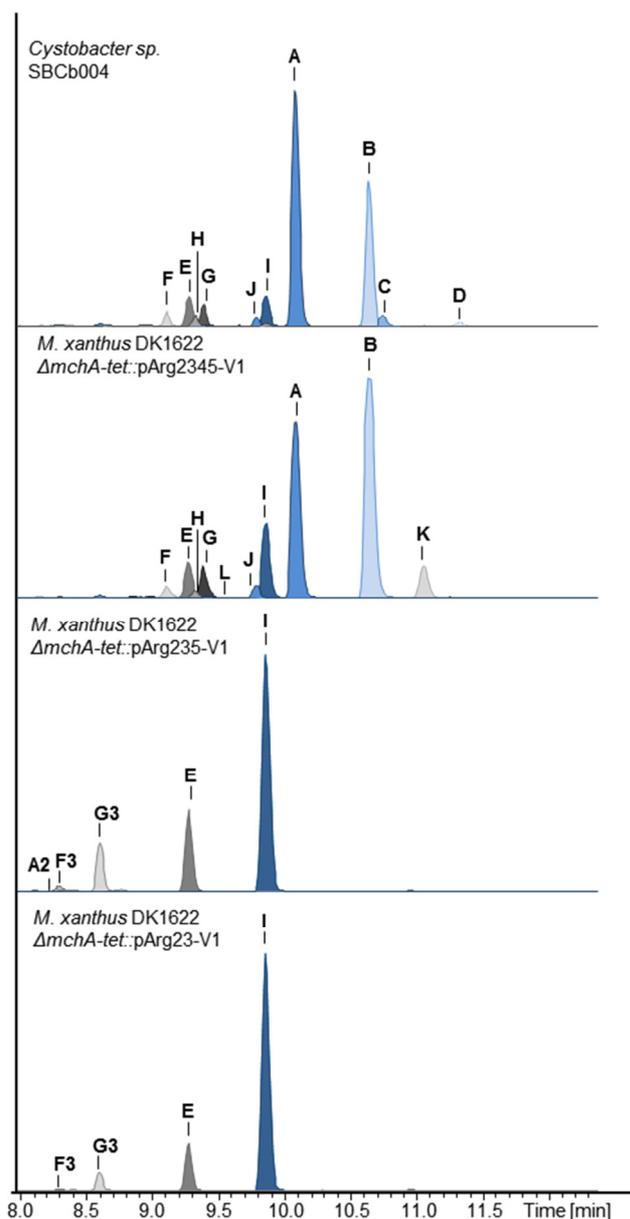
<sup>1</sup> Minor derivatives were not analyzed due to the very low production levels.

<sup>2</sup> Only detected in M7/s4 medium supplemented with amino acid solution as major derivative (see Fig. 5).

### 4.3.3 Approaches to engineer the argyirin production profile

The successful establishment of a heterologous expression system expands future opportunities for exploiting argyrins for potential therapeutic application, however, due to the complexity of the production profiles it would be highly desirable to develop strategies for directing argyirin biosynthesis towards certain derivatives. In our attempts to engineer argyirin biosynthesis we first focused on the expression of *arg*-V1 BGC variants lacking further tryptophan modifying genes. In the course of generation of the pArg2345-V1 expression

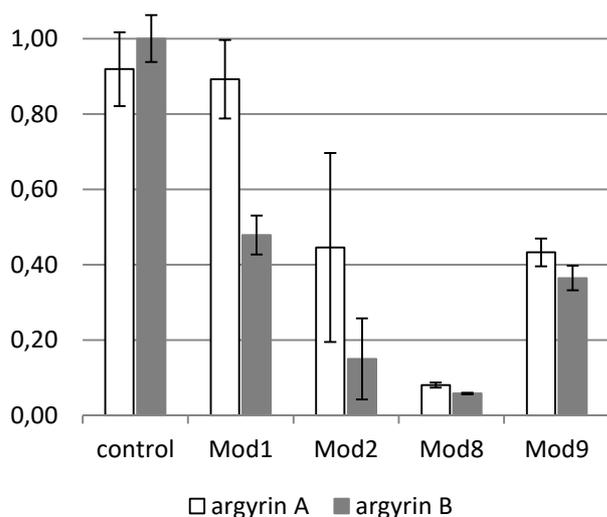
plasmid intermediate constructs with *arg* BGCs lacking *arg4* and/or *arg5* were generated. Transformation of the host with expression plasmids pArg23-V1 and pArg235-V1 resulted in the expected mutant strains (Table 1), which were subsequently analyzed for argyrim production (Fig. 3). Deletion of *arg4/arg5* directed biosynthesis towards argyrim derivatives with unmodified tryptophan rings yielding the respective equivalents of argyrins A and B, argyrins E and I ( $R_3/R_4 = H$ ), as main derivatives, which can be detected in the *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1$  expression strain and the native producer as well, at much lower levels. In addition, novel argyrim derivatives, assigned as argyrins G3 and F3, were produced by *M. xanthus* DK1622  $\Delta mchA-tet::pArg23-V1$  representing putative unmodified equivalents ( $R_3/R_4 = H$ ) of argyrins G and F as concluded from MS fragmentation data (Table S10). They harbor a serine residue in the fourth position of the peptide chain ( $R_2 = CH_2OH$ ), which is much less preferred than alanine ( $R_2 = CH_3$ ) during the NRPS assembly, resulting in the lower production of argyrins G3 and F3 compared to argyrins E and I. Heterologous expression based on pArg235-V1 harboring the modifying gene *arg5* was expected to yield novel hydroxylated argyrim derivatives ( $R_4 = OH$ ) through the action of the tryptophan 2,3-dioxygenase Arg5 and the absence of the *O*-methyltransferase Arg4. Indeed, one novel argyrim designated as argyrim A2 was detected representing a putative derivative of argyrim A with  $R_4 = OH$  instead of  $OCH_3$  as concluded from MS fragmentation data (Table S10). However, argyrins with unmodified tryptophan rings (argyrim E, I, G3 and F3) were still produced at higher yields indicating that tryptophan hydroxylation in the generated expression strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg23-V1$  is either not very efficient (e.g. due to the genetic design of *arg235-V1*) and/or the hydroxylated derivatives are less stable.



**Fig. 3.** Argyrimin production analysis of different *M. xanthus* DK1622 heterologous expression strains in comparison to the native producer *Cystobacter* sp. SBCb004 (HPLC-MS data of the detected argyrimin derivatives are listed in Table S10).

In a parallel approach to direct argyrimin production towards certain derivatives we aimed to engineer the NRPS assembly line. Our studies on argyrimin biosynthesis revealed that the first module (M1) incorporates either alanine (Ala;  $R_1 = \text{CH}_3$ ) or  $\alpha$ -aminobutyric acid (Abu;  $R_1 = \text{CH}_2\text{CH}_3$ ), at almost equal level, into the argyrimin peptide core (Fig. 1). To a much lower extent valine (Val;  $R_1 = \text{CH}_2(\text{CH}_3)_2$ ) is incorporated as well, yielding argyrins K and L (Fig. S11) as discovered during our heterologous expression studies. The first gatekeepers for amino acid incorporation in NRPS biosynthesis represent A domains, which are responsible for substrate selection and activation. Making use of the established expression system we

aimed to direct the biosynthetic capacity of the argyrin megasynthetase towards Ala, Abu or Val derived argyrins by engineering the A domain of M1 (A1). Several approaches to A domain engineering have been evaluated over the past three decades<sup>60</sup>, ranging from entire domain replacement<sup>61,62</sup>, subdomain replacement<sup>63,64</sup> and finally to modifications of the individual amino acid residues within the A domain binding pocket<sup>65,66</sup>. The latter was enabled after the A domain specificity conferring code defined by Stachelhaus et al.<sup>67</sup>, also known as “nonribosomal code”<sup>68</sup>, has been identified. In our study, we designed a panel of 14 modified A1 domains, in which residues of the substrate binding pocket were mutated based on the nonribosomal codes from other myxobacterial A domains, that activate Ala, Abu or Val as deduced from the corresponding NRPS products (see Table S11). Respective A1 domain encoding fragments were obtained via gene synthesis and used for generation of engineered expression plasmids (pArg2345-V1-mod[1-14], Table S5 and Fig. S12). After transformation of *M. xanthus* DK1622  $\Delta mchA-tet$  with the modified constructs and genotypic verification of the mutant strains according to the established procedure (Fig. S10) argyrin production was analyzed. For this, parallel cultivation studies including *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1$  as reference producer were performed. In all analyzed mutants, production of argyrin could indeed be confirmed, proving the successful and functional expression of the modified NRPS megasynthetases. However, as often observed in such bio-combinatorial approaches<sup>60,69,70</sup>, yields decreased significantly in many mutants, hampering the reliable analysis of argyrin production profiles. Nevertheless, four expression strains with different A1 mutations, DK1622  $\Delta mchA-tet::pArg2345-V1$ -[mod1;2;8;9] (see Table S11), yielded satisfactory argyrin levels and showed the desired shift in substrate selection towards alanine compared to the wild-type NRPS expressed in *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1$  (Fig. 4). Further optimization of production levels and profiles (including specificity shifts towards other substrates) might be achieved by evaluation of an extended series of NRPS mutants, e.g. also involving modifications of additional catalytic domains.



**Fig. 4.** Production levels of argyrin A and B evaluated in mutants expressing NRPS megasynthetases with engineered A1 domains (mod1, mod2, mod8 and mod9; see Table S11) compared to the host expressing the parent megasynthetase (control). Samples were analyzed by HPLC-MS and levels of argyrin A and B were detected by EIC  $[M + H]^+ = 825.313 \pm 0.05$  (arg A) and  $[M + H]^+ = 839.329 \pm 0.05$  (arg B). Corresponding peaks were integrated and the surface area values were used for yield comparison. The value of argyrin B produced by the control strain was normalized to 1.

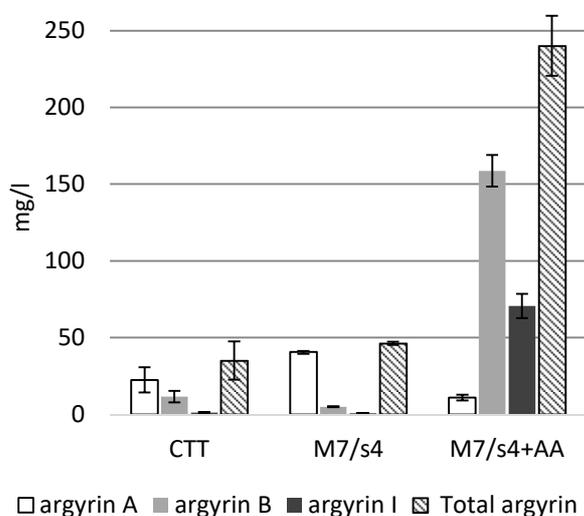
#### 4.3.4 Approaches to further improve the heterologous expression system

In the initial *arg-V1* design approach sequence modulation was focused on the engineering of a restricted number of R-sites to minimize the extent of (synonymous) codon substitutions. Although straightforward *arg-V1* assembly and subsequent BGC modification was achieved we aimed at improving the expression system by extending the engineering options. As shown for A1 modifications, NRPS engineering was feasible with *arg-V1*, but has proven to be tedious and time consuming as two-step cloning processes were required to (Fig. S12) avoid gene synthesis of entire module fragments. To reduce the experimental effort for such approaches we aimed to construct a derivative of *arg-V1*, in which all R-sites for the type IIS endonuclease *BsaI* were mutated. The BGC sequence was modulated accordingly eliminating in total 32 *BsaI* R-sites by silent point mutations (Table S9). An expression construct with the adapted BGC, designated as *arg-V1-BsaI*, was generated based on ten gene synthesis fragments using the established assembly approach (pArg2345-V1-BsaI; Table S5). Transformation of *M. xanthus* DK1622  $\Delta mchA-tet$  yielded the expression strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  as verified by PCR performed analogously to the previous mutant analysis (Fig. S10). Comparative production analysis with the parent *arg-V1* BGC from *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1$  under standard cultivation conditions revealed similar argyrin production yields (35 mg/l argyrin A/B for *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ ), indicating that the additional sequence

modifications had no negative impact on argyrim pathway expression. The absence of *BsaI* R-sites in the entire expression construct pArg2345-V1-*BsaI* greatly expands future engineering opportunities. Independently of the unique R-sites previously introduced in the synthetic *arg* BGC between genes and NRPS modules, any BGC region can now be directly targeted on the expression construct as illustrated in Fig. S13. After replacement of the target region with a selection marker flanked by (unique) *BsaI* R-sites via recombination technologies modified sequence(s) can be introduced scar-less via the *BsaI* R-sites by applying conventional restriction and ligation methods. The more versatile expression system already significantly reduced the cloning effort in ongoing argyrim NRPS engineering approaches (data not shown) and the design principle is broadly applicable on other natural product pathways.

Besides improvement of the *arg* BGC sequence to enable more versatile engineering procedures we aimed to further increase production yields of argyrins. As an alternative to the standard CTT medium, we evaluated argyrim production in M7/s4 and M7/s6 media, which were previously used for increased production of myxobacterial  $\alpha$ -pyrone antibiotics, myxopyronin and coralopyronin, in *M. xanthus* DK1622<sup>57</sup>. By comparative cultivation in both media we identified M7/s4 as superior medium for production of argyrins, reaching production levels of 45 mg/l argyrins A/B with argyrim A as major product (Fig. 5). Lower argyrim production in the M7/s6 medium can be explained from the perspective, that high amounts of potassium acetate do not directly benefit production of peptides in the same way as they do the polyketides. This prompted us to investigate if yield can be further improved by supplementation of M7/s4 medium with a mixture of amino acids required for argyrim peptide core assembly. Indeed argyrim A/B production could be increased significantly to around 160 mg/l by repetitive external supply of the precursor mix during cultivation. In contrast to the standard cultivation in M7/s4 medium, production significantly shifted towards argyrim B, probably due to the external supply of Abu from the precursor mix (Fig. 5). In addition, the feeding strategy led to a significant increase in production of argyrim I ( $R_3, R_4 = H$ ; Fig. 1) indicating a bottleneck of Arg4/Arg5-mediated tailoring biochemistry under the optimized production conditions (Fig. 5). Addressing this bottleneck in future studies, e.g. by co-expression of additional *arg4/5* gene copies, offers a great potential to further increase production yields of argyrins A and B. The argyrim A/B production yields achieved in the heterologous host so far (160 mg/l) are already much higher compared to the native argyrim producer (8 mg/l for *Cystobacter* sp. SBCb004). The established production system ranks with the most successful examples in heterologous expression studies on myxobacterial

megasynthetases in *M. xanthus* (Table 2) and again illustrates the suitability of *M. xanthus* DK1622  $\Delta$ *mchA-tet* (engineered in this study) as host.



**Fig. 5.** Production levels of argyrim A, B, I and total argyrim production evaluated in *M. xanthus* DK1622  $\Delta$ *mchA-tet::pArg2345-V1-BsaI*. Argyrim levels were analyzed by HPLC-MS and pure substance of each derivative was used as standard. Detailed method is provided in the supplementary information.

**Table 2.** Heterologous production of PKS/NRPS-derived natural products in *M. xanthus*.

| Compound         | Native producer                   | Pathway-type [size] | Yield ‡        | Ref        |
|------------------|-----------------------------------|---------------------|----------------|------------|
| Argyrim A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | 160 mg/l       | this study |
| Bengamide        | <i>Myxococcus virescens</i>       | PKS/NRPS [25 kb]    | >10 mg/l       | 71         |
| Corallopyronin A | <i>Corallococcus coralloides</i>  | PKS/NRPS [65 kb]    | 37 mg/l        | 57         |
| Dawenol          | <i>Stigmatella aurantiaca</i>     | PKS [21 kb]         | n.d.           | 72         |
| Epothilone       | <i>Sorangium cellulosum</i>       | PKS/NRPS [56 kb]    | 0.1 - 0.4 mg/l | 73,74,50   |
| Myxochromide A   | <i>Myxococcus xanthus</i>         | PKS/NRPS [29 kb]    | ~500 mg/l      | 58         |
| Myxochromide S   | <i>Stigmatella aurantiaca</i>     | PKS/NRPS [29 kb]    | >500 mg/l      | 75         |
| Myxopyronin A    | <i>Myxococcus fulvus</i>          | PKS/NRPS [53 kb]    | 156 mg/l       | 57         |
| Myxothiazol      | <i>Stigmatella aurantiaca</i>     | PKS/NRPS [57 kb]    | 20 mg/l        | 76         |
| Oxytetracycline  | <i>Streptomyces rimosus</i>       | PKS [32 kb]         | 10 mg/l        | 77         |
| Pretubulysin     | <i>Cystobacter species</i>        | PKS/NRPS [40 kb]    | 0.2 mg/l       | 78         |
| PUFAs            | <i>Aetherobacter fasciculatus</i> | PKS/FAS [18 kb]     | ~1 mg/CDW      | 79         |
| Vioprolide B, D  | <i>Cystobacter violaceus</i>      | NRPS [56 kb]        | 80 mg/l        | 80         |

‡ based on shake flask experiments

## 4.4 Conclusion

Argyriin biosynthesis is directed by a NRPS pathway that naturally produces at least ten derivatives, argyriins A-J, due to miscellaneous megasynthetase substrate specificities and inefficient tailoring chemistry. The dehydroalanine unit, a key structural feature of argyriins, is generated from serine by putative action of a specialized C domain from the downstream assembly line module. Based on our data and other recent examples, we suggest to expand the current C domain classification with a new subtype that generates  $\alpha,\beta$ -dehydro amino acid residues in addition to standard peptide bond formation. The identification of the *arg* BGC from *Cystobacter* sp. SBCb004 set the stage for the establishment of a heterologous expression system employing synthetic biology tools. Synthetic *arg* BGCs were constructed and successfully expressed in a derivative of the myxobacterial model strain *M. xanthus*. A gene sequence optimization approach related to a previous study on the epothilone pathway revealed similar production yields of around 100  $\mu\text{g/l}$  illustrating present challenges in rational gene design for complex megasynthetases. Based on the few examples evaluated so far, the most promising strategy for expression of such giant assembly lines originating from myxobacteria in *M. xanthus* seems to restrict gene sequence modulations to the constructional design. Using the respective *arg*-V1 and the advanced *arg*-V1-BsaI BGC version production yields of 35 mg/l were achieved, which were further enhanced to 160 mg/l argyriin A/B plus significant amounts of argyriin I (~70 mg/l) via media optimization guided by our insights on argyriin biogenesis. The established expression system provides a versatile platform for biotechnological production of argyriins and thereby supports further investigation of their potential medicinal application. Overall, several bottlenecks encountered with the native producer strain, e.g. low production yields and unavailable methods for scare-less chromosomal BGC engineering, were eluded via *arg* BGC re-design and efficient heterologous expression. In future work, we aim to use the improved platform to continue our studies on directing biosynthesis towards selected argyriins and novel derivatives thereof.

## 4.5 Materials and methods

### 4.5.1 Identification and sequence analysis of the *arg* BGC from *Cystobacter* sp. SBCb004

Shotgun genome sequence data of *Cystobacter* sp. SBCb004 were available from a previous study on tubulyisins<sup>34</sup>. The putative argyriin (*arg*) biosynthetic gene cluster (BGC) was identified using the antiSMASH 3.0<sup>35</sup> and PKS/NRPS analysis software tools<sup>36</sup>. Sequence

analysis of the identified BGC was performed with the Geneious software versions 5.3-11.1 (Biomatters, Auckland, New Zealand). Putative gene functions were assigned based on Pfam<sup>81</sup> and Blast<sup>82</sup> homology searches. By *in silico* analysis it was proposed that the *arg* BGC comprises five genes (Table S1). The annotated *arg* BGC sequence from *Cystobacter sp.* SBCb004 is accessible at the GenBank (accession number MK047651) and MIBiG (accession number BGC0001834). Catalytic domains of the two NRPS subunits Arg2 and Arg3 (Fig. 1) were analyzed with the antiSMASH 3.0 software<sup>35</sup>, predicted substrate specificities of adenylation (A) domains are listed in Table S2. C domains from the argyrin NRPS and selected examples from other pathways generating  $\alpha,\beta$ -dehydro amino acids<sup>83</sup> were aligned with a list of 525 C domain sequences from a previous study<sup>45</sup>, using MUSCLE. The alignment was used to reconstruct the phylogeny using phym1, employing the JTT model of amino acid substitution and a gamma-distributed rate variation with four categories. The support values were based on 100-fold bootstrapping, the resulting phylogenetic tree is illustrated in Fig. S1.

#### 4.5.2 Strains, transformation and cultivation conditions

*E. coli* strains HS996 (Invitrogen, Carlsbad, USA), SCS110 (Stratagene, La Jolla, USA) and HB101<sup>84</sup> were grown at 37 °C in Luria-Bertani (LB) medium<sup>85</sup>. Transformation of *E. coli* strains was achieved via electroporation in 0.1 cm-wide cuvettes at 1250 V, a resistance of 200  $\Omega$ , and a capacitance of 25  $\mu$ F. When required, the medium was amended with chloramphenicol (30  $\mu$ g/ml), ampicillin (100  $\mu$ g/ml), and/or kanamycin (60  $\mu$ g/ml) for selection of transformants. The myxobacterial wild-type strains and mutants thereof were grown at 30 °C on a rotary shaker at 180 rpm in standard Erlenmeyer flasks (e.g. 50 ml culture volume in a 300 ml flask). *Cystobacter sp.* SBCb004 strains were grown in M medium (10 g/l soy peptone (phytone), 10 g/l maltose, 1 g/l CaCl<sub>2</sub> x 2 H<sub>2</sub>O, 1 g/l MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 8 mg/l EDTA (FEIII), 11.9 g/l HEPES, pH adjusted to 7.2 with KOH), amended with 120  $\mu$ g/ml tobramycin (natural resistance) and 100  $\mu$ g/ml kanamycin for selection of transformants. Transformation was carried out via triparental conjugation (for details see Supplementary data). *Myxococcus xanthus* DK1622 strains were grown in CTT medium (10 g/l casitone; 10 mM Tris-HCl, pH 8.0; 8 mM MgSO<sub>4</sub>; 10 mM potassium phosphate, pH 7.6) amended with 50  $\mu$ g/ml kanamycin and/or 12.5  $\mu$ g/ml oxytetracycline for selection of transformants. Transformation was carried out via electroporation according to previously established methods<sup>86</sup>.

### 4.5.3 General methods for DNA manipulation, analysis and PCR

Plasmid DNA was either purified by standard alkaline lysis<sup>85</sup> or by using the GeneJet Plasmid Miniprep Kit (Thermo Fisher Scientific) or the NucleoBond PC100 kit (Machery Nagel). Oligonucleotides were obtained from Sigma-Aldrich and are listed in Table S3. PCR reactions were carried out in a Mastercycler® pro (Eppendorf) using Phusion™ High-Fidelity or Taq DNA polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. For Taq: Initial denaturation (5 min, 95 °C); 30 cycles of denaturation (30 s, 95 °C), annealing (30 s, 53 - 64 °C) and elongation (varied based on PCR product length 1 kb/min, 72 °C); and final extension (10 min, 72 °C). For Phusion™: Initial denaturation (1 min, 98 °C); 30 cycles of denaturation (20 s, 98 °C), annealing (25 s, 53 - 64 °C) and elongation (varied based on PCR product length 0.5 kb/ min, 72 °C); and final extension (5 min, 72 °C). Restriction enzymes were purchased from Thermo Fisher Scientific or New England Biolabs. Genomic DNA isolation was performed with the Gentra PURGENE® DNA Purification Kit (Qiagen, Hilden, Germany). DNA fragments were separated by agarose gel electrophoresis and purified by gel extraction by NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) or ethanol precipitation (for the fragments larger than 10 kb). T4 DNA ligase (Thermo Fisher Scientific) was used for ligation. Shrimp alkaline phosphatase (Thermo Fisher Scientific) was used to treat the accept vectors for avoiding self-ligation. Plasmid preparations, restriction digestions, gel electrophoresis, and ligation reactions were carried out according to standard methods<sup>85</sup>. Synthetic DNA fragments were obtained from ATG:biosynthetics GmbH and delivered in standard cloning vectors (Table S4). Details on the construction of plasmids generated in this study are given in Table S5. Sequence data of the expression vector pSynBio3 as well as the expression constructs pArg2345-V1, pArg235-V1, pArg23-V1, pArg2345-V2 and pArg2345-V1-BsaI are accessible at the GenBank (accession numbers: MK047657, MK047655, MK047653, MK047652, MK047654, MK047656).

### 4.5.4 Studies with the native argyirin producer strains

A gene inactivation experiment targeting the NRPS encoding gene *arg3* was performed using a derivative of the pSUP-Km suicide vector<sup>87</sup>. The constructed plasmid pArg1 was transferred into *Cystobacter sp.* SBCb004 via conjugation and integrated into the chromosome by homologous recombination (Fig. S2). The obtained mutant strain *Cystobacter sp.* SBCb004::pArg1 was cultivated in comparison to the wild-type and argyirin production was evaluated by HPLC-MS analysis of culture extracts (Fig. S3). Details on the inactivation

experiment are provided in the Supplementary data. For feeding studies, small-scale shake flask cultivations of *Cystobacter sp.* SBCb004 were performed under standard conditions. Labelled precursors (*L*-serine-2,3,3-D<sub>3</sub> and *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N) were fed to the cultures in aliquots and culture extracts were analyzed by HPLC-MS to evaluate the incorporation pattern (Fig. S4/S5). Details on the feeding experiments are provided in the Supplementary data. Re-evaluation of extract fractions from previous large-scale *Archangium sp.* fermentations<sup>19</sup> revealed two novel argyirin derivatives, argyirin I and argyirin J (Fig. S6). The planar structures were elucidated by 1D/2D NMR spectroscopy (Fig. S7/S8, Tables S6/S7) and high-resolution-MS data (for details see Supplementary data).

#### 4.5.5 Sequence design and assembly of synthetic *arg* BGCs

Using the native *arg* BGC sequence from *Cystobacter sp.* SBCb004 as template two versions of synthetic *arg* BGCs were designed (*arg*-V1 and *arg*-V2). Sequence modulations were performed with the evoMAG<sup>is</sup> software package (ATG:biosynthetics GmbH) as described previously<sup>50</sup>. The sequence design process included the engineering of endonuclease restriction sites (R-sites) for both *arg* BGC versions (Fig. 2A, Table S8). The *arg*-V2 sequence was further modified according to additional criteria: adaptation of the codon usage and the course of the local codon adaptation index (CAI; Fig. 2B), engineering of Shine-Dalgarno (SD) – anti-SD interactions (Fig. S9A), elimination of rare codon clusters (Fig. S9B), introduction of “hidden” stop codons in unused frames (Fig. S9C), as well as the elimination of sequence repeats and homopolymer stretches. The sequence modulation was applied for genes *arg2-arg4* based on synonymous codon substitutions (Table S9). A derivative of *arg*-V1 was generated by elimination of all *Bsa*I R-sites (*arg*-V1-*Bsa*I). Details on the sequence design are provided in the Supplementary data. The three *arg* BGC versions (*arg*-V1, *arg*-V2 and *arg*-V1-*Bsa*I) were assembled based on synthetic DNA fragments via conventional restriction and ligation methods making use of unique R-sites engineered during the sequence design (Fig. 2). Thirty DNA fragments (ten for each BGC version) were generated via gene synthesis (Table S4) as well as two fragments for construction of a suitable vector backbone (pSynBio3). Details on the construction of the expression plasmids pArg2345-V1, pArg2345-V2, pArg2345-V1-*Bsa*I, pArg23-V1 and pArg235-V1 are provided in Table S5. In addition to restriction analysis, integrity of the generated expression constructs was verified by Illumina sequencing prior transformation of the host strain. Using the Illumina paired-end technology on a MiSeq platform a mean sequencing coverage

of > 200 fold was achieved for each construct and no mutations within the *arg* BGC region were detected.

#### 4.5.6 Heterologous expression in *M. xanthus* and argyrim production analysis

A derivative of *M. xanthus* DK1622, *M. xanthus* DK1622  $\Delta mchA-tet$ , was constructed, in which the *mch<sub>A</sub>* BGC region<sup>55</sup> was replaced with a *tet<sup>R</sup>* gene (for details see Supplementary data). As illustrated in Fig. S10, expression constructs were chromosomally integrated via *tet<sup>R</sup>* after electroporation of the host. The obtained transformants were verified by PCR analysis (for details see Supplementary data). Expression strains generated in this study are listed in Table 1. After genotypic verification, heterologous argyrim production was analyzed. Small-scale cultivations were performed in shake flasks under standard conditions (CTT medium amended with 50  $\mu$ g/ml kanamycin and 2% XAD-16, 30 °C, 200 rpm for 5-6 days). Crude cell/XAD extracts were generated and concentrated aliquots were analyzed by HPLC-MS (for details see Supplementary data). Argyrim derivatives were identified based on their MS fragmentation spectra (Table S10). In addition to argyrins known from native producer strains, argyrins A-H<sup>20</sup> and argyrins I and J (see Supplementary data), novel derivatives were detected and assigned as argyrins K, L, A2, F3 and G3 (Fig. 3, Table S10). To confirm the incorporation of valine into the new derivatives argyrim K and L, a feeding experiment with *L*-valine-D8 was performed (Fig. S11). Further details are provided in the Supplementary data.

#### 4.5.7 NRPS A1 domain engineering

The A domain engineering experiment was performed according to the procedure illustrated in Figure S12. Synthetic fragments used to exchange the subdomain part of the A1 domain are provided in the Table S4. Modified argyrim expression constructs (see Table S5) were transferred into heterologous host *M. xanthus* DK1622  $\Delta mchA-tet$  by electroporation. The obtained mutants were cultivated in comparison to the control strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1$  and argyrim production was evaluated by HPLC-MS analysis of the culture extracts. Details of the A1 engineering experiment are provided in the Supplementary data.

## Acknowledgements

We would like to thank Dr. Ronald Garcia for provision and advice on the argyrim producer *Cystobacter* sp. SBCb004, Irene Kochems and Douglas Gatte Pichi for their efforts in SBCb004 engineering, Sara Andes and Jack Stanley for assistance in *arg* BGC cloning, Dr. Nestor Zaburanyi for bioinformatics support, and Eva Luxemburger, Daniel Sauer and Dr. Kirsten Harmrolfs for HPLC-MS and NMR measurements. Furthermore, we gratefully acknowledge Prof. Dr. Rolf Müller for constantly supporting and discussing this study. This work was generously funded by a grant from the German Federal Ministry of Education and Research (FKZ: 031A155).

## Database linking

The sequences of the native *arg* BGC from *Cystobacter* sp. SBCb004, the expression vector pSynBio3 as well as the expression constructs pArg2345-V1, pArg235-V1, pArg23-V1, pArg2345-V2 and pArg2345-V1-BsaI have been deposited in the GenBank database (accession numbers: MK047651, MK047657, MK047655, MK047653, MK047652, MK047654, MK047656 ). The information on the native *arg* BGC from *Cystobacter* sp. SBCb004 has also been deposited in the MIBiG repository (accession number BGC0001834).

## Appendix A. Supporting Information

Provided file includes:

- Supplementary text providing detailed description of experiments
- Figures show phylogenetic analysis of C domains, inactivation of *arg3* in *Cystobacter* sp. SBCb004 and genotypic analysis of the mutants, argyrim production analysis after *arg3* inactivation in *Cystobacter* sp. SBCb004, analysis of labelled L-serine and L-threonine feeding, structures and NMR spectra of argyrim I and J, effect of the *arg*-V2 sequence modulation on selected features in comparison to *arg*-native, construction and verification of expression strains, analysis of labelled L-valine feeding, strategy to exchange the subdomain part of A1 and modification strategy of the pArg2345-V1-BsaI expression construct.
- Tables show genes, proteins, proposed function in the *arg* BGC from *Cystobacter* sp. SBCb004, substrate specificity analysis of A domains from the argyrim NRPS from *Cystobacter* sp. SBCb004, oligonucleotides used in this study, gene synthesis

constructs obtained from ATG:biosynthetics GmbH, plasmids and expression constructs generated in this study, NMR spectroscopic data of argyrin I and J, restriction sites engineered within the *arg* BGC and pSynBio3 sequence, sequence features of synthetic *arg* BGC versions compared to the native sequence, HPLC-MS data of argyrin derivatives detected in the heterologous expression strains and nonribosomal codes of A domains.

## 4.6 References

- (1) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014, *J. Nat. Prod.* **2016**, *79*, pp. 629–661.
- (2) Katz, L.; Baltz, R. H. Natural product discovery: past, present, and future, *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, pp. 155–176.
- (3) Shen, B. A New Golden Age of Natural Products Drug Discovery, *Cell.* **2015**, *163*, pp. 1297–1300.
- (4) Bérdy, J. Thoughts and facts about antibiotics: Where we are now and where we are heading, *J. Antibiot. (Tokyo)*. **2012**, *65*, p. 441.
- (5) Shimkets, L.; Dworkin, M.; Reichenbach, H. The Myxobacteria. In *The Prokaryotes*; Dworkin, M.; Falkow, S.; Rosenberg, E.; Schleifer, K.; Stackebrandt, E., Eds.; Springer Verlag: New York, 2006, pp. 31–115.
- (6) Reichenbach, H. Order VIII. Myxococcales. Tchan, Pochon and Prévot 1948, 398AL. In *Bergey's manual of systematic bacteriology*, 2nd ed.; Brenner, D.; Krieg, N.; Staley, J., Eds.; Springer, 2005, pp. 1059–1144.
- (7) Garcia, R.; Krug, D.; Müller, R. Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods. In *Methods in Enzymology. Complex Enzymes in Microbial Natural Product Biosynthesis*; David A. Hopwood, Ed., 2009, pp. 59–91.
- (8) Mohr, K. I. Diversity of Myxobacteria-We Only See the Tip of the Iceberg, *Microorganisms*. **2018**, *6*.
- (9) Herrmann, J.; Fayad, A. A.; Müller, R. Natural products from myxobacteria: novel metabolites and bioactivities, *Nat. Prod. Rep.* **2017**, *34*, pp. 135–160.
- (10) Schäberle, T. F.; Lohr, F.; Schmitz, A.; König, G. M. Antibiotics from myxobacteria, *Nat. Prod. Rep.* **2014**, *epub ahead of print*.
- (11) Weissman, K. J.; Müller, R. Myxobacterial secondary metabolites: bioactivities and modes-of-action, *Nat. Prod. Rep.* **2010**, *27*, pp. 1276–1295.
- (12) Alvarez, R. H.; Valero, V.; Hortobagyi, G. N. Ixabepilone for the treatment of breast cancer, *Ann. Med.* **2011**, *43*, pp. 477–486.
- (13) Li, J.; Ren, J.; Sun, W. Systematic review of ixabepilone for treating metastatic breast cancer, *Breast cancer (Tokyo, Japan)*. **2017**, *24*, pp. 171–179.
- (14) Berod, L.; Friedrich, C.; Nandan, A.; Freitag, J.; Hagemann, S.; Harmrolfs, K.; Sandouk, A.; Hesse, C.; Castro, C. N.; Bähre, H.; Tschirner, S. K.; Gorinski, N.; Gohmert, M.; Mayer, C. T.; Huehn, J.; Ponimaskin, E.; Abraham, W.-R.; Müller, R.; Lochner, M.; Sparwasser, T. De novo fatty acid synthesis controls the fate between regulatory T and T helper 17 cells, *Nat. Med.* **2014**, *20*, pp. 1327–1333.
- (15) Held, J.; Gebru, T.; Kalesse, M.; Jansen, R.; Gerth, K.; Müller, R.; Mordmüller, B. Antimalarial activity of the myxobacterial macrolide chlorotonil A, *Antimicrob. Agents Chemother.* **2014**, *58*, pp. 6378–6384.
- (16) Baumann, S.; Herrmann, J.; Raju, R.; Steinmetz, H.; Mohr, K. I.; Hüttel, S.; Harmrolfs, K.; Stadler, M.; Müller, R. Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity, *Angew. Chem. Int. Ed.* **2014**, *53*, pp. 14605–14609.

- (17) Guenther, E. G.; Paulini, K. W.; Aicher, B.; Baasner, S.; Schmidt, P.; Teifel, M.; Schaefer, O.; Grundker, C.; Fister, S. Targeting Disorazol Z to LHRH-receptor positive tumors by the cytotoxic conjugate AEZS-125. In , 2008.
- (18) Chen, X.; Bui, K. C.; Barat, S.; Thi Nguyen, M. L.; Bozko, P.; Sipos, B.; Kalesse, M.; Malek, N. P.; Plentz, R. R. Therapeutic effects of Argyrin F in pancreatic adenocarcinoma, *Cancer letters*. **2017**, *399*, pp. 20–28.
- (19) Sasse, F.; Steinmetz, H.; Schupp, T.; Petersen, F.; Memmert, K.; Hofmann, H.; Heusser, C.; Brinkmann, V.; Matt, P. von; Hofle, G.; Reichenbach, H. Argyrins, immunosuppressive cyclic peptides from myxobacteria. I. Production, isolation, physico-chemical and biological properties, *J. Antibiot.* **2002**, *55*, pp. 543–551.
- (20) Vollbrecht, L.; Steinmetz, H.; Höfle, G.; Oberer, L.; Rihs, G.; Bovermann, G.; Matt, P. von Argyrins, immunosuppressive cyclic peptides from myxobacteria. II. Structure elucidation and stereochemistry, *J. Antibiot.* **2002**, *55*, pp. 715–721.
- (21) Ferrari, P.; Vekey, K.; Galimberti, M.; Gallo, G. G.; Selva, E.; Zerilli, L. F. Antibiotics A21459 A and B, new inhibitors of bacterial protein synthesis. II. Structure elucidation, *J. Antibiot.* **1996**, *49*, pp. 150–154.
- (22) Selva, E.; Gastaldo, L.; Saddler, G. S.; Toppo, G.; Ferrari, P.; Carniti, G.; Goldstein, B. P. Antibiotics A21459 A and B, new inhibitors of bacterial protein synthesis. I. Taxonomy, isolation and characterization, *J. Antibiot.* **1996**, *49*, pp. 145–149.
- (23) Nyfeler, B.; Hoepfner, D.; Palestrant, D.; Kirby, C. A.; Whitehead, L.; Yu, R.; Deng, G.; Caughlan, R. E.; Woods, A. L.; Jones, A. K.; Barnes, S. W.; Walker, J. R.; Gaulis, S.; Hauy, E.; Brachmann, S. M.; Krastel, P.; Studer, C.; Riedl, R.; Estoppey, D.; Aust, T.; Movva, N. R.; Wang, Z.; Salcius, M.; Michaud, G. A.; McAllister, G.; Murphy, L. O.; Tallarico, J. A.; Wilson, C. J.; Dean, C. R. Identification of elongation factor G as the conserved cellular target of argyirin B, *PLoS ONE*. **2012**, *7*, pp. e42657.
- (24) Bielecki, P.; Lukat, P.; Hüsecken, K.; Dötsch, A.; Steinmetz, H.; Hartmann, R. W.; Müller, R.; Häussler, S. Mutation in elongation factor G confers resistance to the antibiotic argyirin in the opportunistic pathogen *Pseudomonas aeruginosa*, *ChemBioChem*. **2012**, *13*, pp. 2339–2345.
- (25) Jones, A. K.; Woods, A. L.; Takeoka, K. T.; Shen, X.; Wei, J.-R.; Caughlan, R. E.; Dean, C. R. Determinants of Antibacterial Spectrum and Resistance Potential of the Elongation Factor G Inhibitor Argyrin B in Key Gram-Negative Pathogens, *Antimicrob. Agents Chemother.* **2017**, *61*.
- (26) Nিকেleit, 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.; Malek, N. P. Argyrin A reveals a critical role for the tumor suppressor protein p27kip1 in mediating antitumor activities in response to proteasome inhibition, *Cancer Res.* **2008**, *14*, pp. 23–35.
- (27) Stauch, B.; Simon, B.; Basile, T.; Schneider, G.; Malek, N. P.; Kalesse, M.; Carlomagno, T. Elucidation of the structure and intermolecular interactions of a reversible cyclic-peptide inhibitor of the proteasome by NMR spectroscopy and molecular modeling, *Angew. Chem. Int. Ed.* **2010**, *49*, pp. 3934–3938.
- (28) Bülow, L.; Nিকেleit, I.; Girbig, A.-K.; Brodmann, T.; Rentsch, A.; Eggert, U.; Sasse, F.; Steinmetz, H.; Frank, R.; Carlomagno, T.; Malek, N. P.; Kalesse, M. Synthesis and biological characterization of argyirin F, *ChemMedChem*. **2010**, *5*, pp. 832–836.

- (29) Ley, S. V.; Priour, A.; Heusser, C. Total synthesis of the cyclic heptapeptide Argyrin B: a new potent inhibitor of T-cell independent antibody formation, *Org. Lett.* **2002**, *4*, pp. 711–714.
- (30) Chen, C. H.; Genapathy, S.; Fischer, P. M.; Chan, W. C. A facile approach to tryptophan derivatives for the total synthesis of argyirin analogues, *Org. Biomol. Chem.* **2014**, *12*, pp. 9764–9768.
- (31) Finking, R.; Marahiel, M. A. Biosynthesis of nonribosomal peptides, *Annu. Rev. Microbiol.* **2004**, *58*, pp. 453–488.
- (32) Süssmuth, R. D.; Mainz, A. Nonribosomal peptide synthesis - Principles and prospects, *Angew. Chem. Int. Ed.* **2017**, *56*, pp. 3770–3821.
- (33) Wenzel, S. C.; Müller, R. Host Organisms: Myxobacterium. In *Industrial biotechnology, Microorganisms Volume 3a and 3b*; Wittmann, C.; Liao, J., Eds.; Wiley-VCH: Weinheim, Germany, 2017, pp. 453–485.
- (34) Chai, Y.; Pistorius, D.; Ullrich, A.; Weissman, K. J.; Kazmaier, U.; Müller, R. Discovery of 23 natural tubulysins from *Angiococcus disciformis* An d48 and *Cystobacter* SBCb004, *Chem. Biol.* **2010**, *17*, pp. 296–309.
- (35) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters, *Nucleic Acids Res.* **2015**, *43*, W237-W243.
- (36) Bachmann, B. O.; Ravel, J. Methods for in Silico Prediction of Microbial Polyketide and Nonribosomal Peptide Biosynthetic Pathways from Dna Sequence Data, *Methods Enzymol.* **2009**, *458*, pp. 181–217.
- (37) Walsh, C. T.; O'Brien, R. V.; Khosla, C. Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, pp. 7098–7124.
- (38) Walsh, C. T.; Chen, H. W.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L. S.; Marshall, C. G.; Miller, D. A.; Patel, H. M. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines, *Curr. Opin. Chem. Biol.* **2001**, *5*, pp. 525–534.
- (39) Kudo, F.; Miyanaga, A.; Eguchi, T. Biosynthesis of natural products containing  $\beta$ -amino acids, *Nat. Prod. Rep.* **2014**, *31*, pp. 1056–1073.
- (40) Gaudelli, N. M.; Long, D. H.; Townsend, C. A.  $\beta$ -Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis, *Nature.* **2015**, *520*, pp. 383–387.
- (41) Thomas, M. G.; Chan, Y. A.; Ozanick, S. G. Deciphering tuberactinomycin biosynthesis. Isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster, *Antimicrobial agents and chemotherapy.* **2003**, *47*, pp. 2823–2830.
- (42) Beasley, F. C.; Cheung, J.; Heinrichs, D. E. Mutation of L-2,3-diaminopropionic acid synthase genes blocks staphyloferrin B synthesis in *Staphylococcus aureus*, *BMC microbiology.* **2011**, *11*, p. 199.
- (43) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.;

Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süßmuth, R. E.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature, *Nat. Prod. Rep.* **2013**, *30*, pp. 108–160.

(44) Du, L.; Sanchez, C.; Chen, M. T.; Edwards, D. J.; Shen, B. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase, *Chem. Biol.* **2000**, *7*, pp. 623–642.

(45) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution, *BMC Evol. Biol.* **2007**, *7*, pp. 78–92.

(46) Tokarski, Z.; Klei, H. E.; Berg, C. M. Novel biotechnology process for the production of (+)-2-aminobutyrate, *Biotechnol. Lett.* **1988**, *10*, pp. 7–10.

(47) Balibar, C. J.; Vaillancourt, F. H.; Walsh, C. T. Generation of D-amino Acid residues in assembly of arthrofactin by dual condensation/epimerization domains, *Chem. Biol.* **2005**, *12*, pp. 1189–1200.

(48) Dittmann, J.; Wenger, R. M.; Kleinkauf, H.; Lawen, A. Mechanism of cyclosporin A biosynthesis. Evidence for synthesis via a single linear undecapeptide precursor, *J. Biol. Chem.* **1994**, *269*, pp. 2841–2846.

(49) Hoffmann, K.; Schneider-Scherzer, E.; Kleinkauf, H.; Zocher, R. Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis, *J. Biol. Chem.* **1994**, *269*, pp. 12710–12714.

(50) Oßwald, C.; Zipf, G.; Schmidt, G.; Maier, J.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Modular Construction of a Functional Artificial Epothilone Polyketide Pathway, *ACS Synth. Biol.* **2014**, *3*, pp. 759–772.

(51) Rose, R. E. The nucleotide sequence of pACYC177, *Nucleic Acids Res.* **1988**, *16*, p. 356.

(52) Rose, R. E. The nucleotide-sequence of pACYC184, *Nucleic Acids Res.* **1988**, *16*, p. 355.

(53) Beck, E.; Ludwig, G.; Auerswald, E. A.; Reiss, B.; Schaller, H. Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5, *Gene.* **1982**, *19*, pp. 327–336.

(54) Magrini, V.; Creighton, C.; Youderian, P. Site-specific recombination of temperate *Myxococcus xanthus* phage Mx8: Genetic elements required for integration, *J. Bacteriol.* **1999**, *181*, pp. 4050–4061.

(55) Wenzel, S. C.; Meiser, P.; Binz, T. M.; Mahmud, T.; Müller, R. Nonribosomal Peptide Biosynthesis: Point Mutations and Module Skipping Lead to Chemical Diversity, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, pp. 2296–2301.

(56) Krug, D.; Zurek, G.; Schneider, B.; Garcia, R.; Müller, R. 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.* **2008**, 624, pp. 97–106.

(57) Sucipto, H.; Pogorevc, D.; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, 44, pp. 160–170.

(58) Yan, F.; Burgard, C.; Popoff, A.; Zaburannyi, N.; Zipf, G.; Maier, J.; Bernauer, H. S.; Wenzel, S. C.; Müller, R. Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production, *Chem. Sci.* **2018**, 9, pp. 7510–7519.

(59) Yan, F.; Auerbach, D.; Chai, Y.; Keller, L.; Zhang, Y.; Müller, R. Biosynthesis and heterologous expression of vioprolides: C-domain catalyzed glycerate esterification and post-assembly maturation, *submitted manuscript.* **2017**.

(60) Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines, *Nat. Prod. Rep.* **2016**, 33, pp. 317–347.

(61) Stachelhaus, T.; Schneider, A.; Marahiel, M. A. Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains, *Science.* **1995**, 269, pp. 69–72.

(62) Doekel, S.; Marahiel, M. A. Dipeptide formation on engineered hybrid peptide synthetases, *Chem. Biol.* **2000**, 7, pp. 373–384.

(63) Crusemann, M.; Kohlhaas, C.; Piel, J. Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains, *Chem. Sci.* **2013**, 4, pp. 1041–1045.

(64) Kries, H.; Niquille, D. L.; Hilvert, D. A subdomain swap strategy for reengineering nonribosomal peptides, *Chem. Biol.* **2015**, 22, pp. 640–648.

(65) Bian, X.; Plaza, A.; Yan, F.; Zhang, Y.; Müller, R. Rational and efficient site-directed mutagenesis of adenylation domain alters relative yields of luminide derivatives in vivo, *Biotechnol. Bioeng.* **2015**, 112, pp. 1343–1353.

(66) Han, J. W.; Kim, E. Y.; Lee, J. M.; Kim, Y. S.; Bang, E.; Kim, B. S. Site-directed modification of the adenylation domain of the fusaricidin nonribosomal peptide synthetase for enhanced production of fusaricidin analogs, *Biotechnol. Lett.* **2012**, 34, pp. 1327–1334.

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

(68) Döhren, H. von; Dieckmann, R.; Pavela-Vrancic, M. The nonribosomal code, *Chem. Biol.* **1999**, 6, pp. R273–R279.

(69) Nguyen, K. T.; Ritz, D.; Gu, J. Q.; Alexander, D.; Chu, M.; Miao, V.; Brian, P.; Baltz, R. H. Combinatorial biosynthesis of novel antibiotics related to daptomycin, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, pp. 17462–17467.

(70) Miao, V.; Coeffet-Le Gal, M. F.; Nguyen, K.; Brian, P.; Penn, J.; Whiting, A.; Steele, J.; Kau, D.; Martin, S.; Ford, R.; Gibson, T.; Bouchard, M.; Wrigley, S. K.; Baltz, R. H. Genetic engineering in *Streptomyces roseosporus* to produce hybrid lipopeptide antibiotics, *Chem. Biol.* **2006**, 13, pp. 269–276.

(71) Wenzel, S. C.; Hoffmann, H.; Zhang, J.; Debussche, L.; Haag-Richter, S.; Kurz, M.; Nardi, F.; Lukat, P.; Kochems, I.; Tietgen, H.; Schummer, D.; Nicolas, J.-P.; Calvet, L.; Czepczor, V.; Vrignaud, P.; Mühlenweg, A.; Pelzer, S.; Müller, R.; Brönstrup, M. Production of the bengamide class of marine natural products in myxobacteria: biosynthesis and structure-activity relationships, *Angew. Chem. Int. Ed. Engl.* **2015**, 54, pp. 15560–15564.

- (72) Oßwald, C.; Zaburannyi, N.; Burgard, C.; Hoffmann, T.; Wenzel, S. C.; Müller, R. A highly unusual polyketide synthase directs dawenol polyene biosynthesis in *Stigmatella aurantiaca*, *J. Biotechnol.* **2014**, *191*, pp. 54–63.
- (73) Julien, B.; Shah, S. Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*, *Antimicrob. Agents Chemother.* **2002**, *46*, pp. 2772–2778.
- (74) Lau, J.; Frykman, S.; Regentin, R.; Ou, S.; Tsuruta, H.; Licari, P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*, *Biotechnol. Bioeng.* **2002**, *78*, pp. 280–288.
- (75) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (76) Perlova, O.; Fu, J.; Kuhlmann, S.; Krug, D.; Stewart, F.; Zhang, Y.; Müller, R. Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2006**, *72*, pp. 7485–7494.
- (77) Stevens, D. C.; Henry, M. R.; Murphy, K. A.; Boddy, C. N. Heterologous expression of the oxytetracycline biosynthetic pathway in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2010**, *76*, pp. 2681–2683.
- (78) Chai, Y.; Shan, S.; Weissman, K. J.; Hu, S.; Zhang, Y.; Müller, R. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis, *Chem. Biol.* **2012**, *19*, pp. 361–371.
- (79) Gemperlein, K.; Rachid, S.; Garcia, R. O.; Wenzel, S. C.; Müller, R. Polyunsaturated fatty acid biosynthesis in myxobacteria. Different PUFA synthases and their product diversity, *Chem. Sci.* **2014**, *5*, pp. 1733–1741.
- (80) Yan, F.; Auerbach, D.; Chai, Y.; Keller, L.; Tu, Q.; Hüttel, S.; Glemser, A.; Grab, H. A.; Bach, T.; Zhang, Y.; Müller, R. Biosynthesis and Heterologous Production of Vioprolides: Rational Biosynthetic Engineering and Unprecedented 4-Methylazetidincarboxylic Acid Formation, *Angew. Chem. Int. Ed. Engl.* **2018**.
- (81) Finn, R. D.; Bateman, A.; Clements, J.; Coghill, P.; Eberhardt, R. Y.; Eddy, S. R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonnhammer, E. L.; Tate, J.; Punta, M. Pfam: the protein families database, *Nucleic Acids Res.* **2014**, *42*, pp. D222–D230.
- (82) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool, *J. Mol. Biol.* **1990**, *215*, pp. 403–410.
- (83) Siodłak, D.  $\alpha,\beta$ -Dehydroamino acids in naturally occurring peptides, *Amino Acids.* **2015**, *47*, pp. 1–17.
- (84) Boyer, H. W.; Roulland-Dussoix, D. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*, *J. Mol. Biol.* **1969**, *14*, pp. 459–472.
- (85) Sambrook, J.; Russell, D. W. *Molecular cloning: A laboratory manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.
- (86) Kashefi, K.; Hartzell, P. L. Genetic suppression and phenotypic masking of a *Myxococcus xanthus* *frzF*- defect, *Mol. Microbiol.* **1995**, *15*, pp. 483–494.

(87) Ullrich, A.; Chai, Y.; Pistorius, D.; Elnakady, Y. A.; Herrmann, J. E.; Weissman, K. J.; Kazmaier, U.; Müller, R. Pretubulysin, a potent and chemically accessible tubulysin precursor from *Angiococcus disciformis*, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, pp. 4422–4425.

## 4.7 Supplementary information

### Table of Contents

|        |  |     |
|--------|--|-----|
| 4.7.1  | Inactivation of the argyrin biosynthetic pathway in <i>Cystobacter sp.</i> SBCb004<br>161                      |     |
| 4.7.2  | Feeding studies with <i>Cystobacter sp.</i> SBCb004 .....  | 162 |
| 4.7.3  | Isolation and structure elucidation of argyrins I and J .....  | 163 |
| 4.7.4  | Sequence design of synthetic arg BGCs (arg-V1, arg-V2, arg-V1-BsaI) .....                                      | 164 |
| 4.7.5  | Construction of <i>M. xanthus</i> DK1622 $\Delta$ <i>mchA-tet</i> and <i>arg</i> BGC expression strains<br>165 |     |
| 4.7.6  | Analysis of argyrin production in <i>M. xanthus</i> and valine feeding experiment                              | 167 |
| 4.7.7  | Engineering of the A1 domain of the argyrin NRPS megasynthetase .....  | 168 |
| 4.7.8  | Medium optimization for yield improvement .....  | 170 |
| 4.7.9  | Figures .....  | 171 |
| 4.7.10 | Tables .....   | 187 |
| 4.8    | Unpublished results .....  | 198 |
| 4.8.1  | Design, construction and evaluation of engineered argyrin A4 domains .....                                     | 198 |
| 4.8.2  | Development and evaluation of the argyrin expression system for mutasynthesis<br>studies                       | 201 |
| 4.8.3  | References .....   | 206 |

## List of Figures

|   |     |
|---|-----|
| Figure S1. Phylogenetic tree of different C domain subtypes.....  | 171 |
| Figure S2. Inactivation of <i>arg3</i> in <i>Cystobacter sp.</i> SBCb004 by insertional mutagenesis. ....                                     | 172 |
| Figure S3. Argyrin production analysis after <i>arg3</i> inactivation in <i>Cystobacter sp.</i> SBCb004. ....                                 | 172 |
| Figure S4. Analysis of the incorporation <i>L</i> -serine-2,3,3-D3 during argyirin biosynthesis. ....   | 173 |
| Figure S5. Analysis of the incorporation <i>L</i> -threonine- <sup>13</sup> C <sub>4</sub> , <sup>15</sup> N during argyirin biosynthesis. .. | 173 |
| Figure S6. Structures of argyirin I and argyirin J.....   | 174 |
| Figure S7. NMR spectra of argyirin I.....   | 178 |
| Figure S8. NMR spectra from argyirin J.....   | 182 |
| Figure S9. Effect of the <i>arg</i> -V2 sequence modulation on selected features in comparison to <i>arg</i> -native. ....                    | 183 |
| Figure S10. Construction and verification of expression strains. ....   | 184 |
| Figure S11. Analysis of the incorporation <i>L</i> -valine-D8 during argyirin biosynthesis. ....  | 185 |
| Figure S12. Strategy to exchange the subdomain part of A1.....  | 186 |
| Figure S13. Modification strategy of the pArg2345-V1-BsaI expression construct.....   | 186 |
| Figure S14. Strategy to exchange the subdomain part of A4.....  | 200 |
| Figure S15. Mutasynthesis approach for production of argyrins. ....   | 204 |
| Figure S16. Degradation kinetics of the D-Ala-Dha-Sar-SNAc mutasynthron in M7/S4 medium over 8 h period. ....                                 | 205 |

## List of Tables

|  |     |
|--|-----|
| Table S1. Genes, proteins, proposed function in the <i>arg</i> BGC from <i>Cystobacter sp.</i> SBCb004.....  | 187 |
| Table S2. Substrate specificity analysis of A domains from the argyirin NRPS from <i>Cystobacter sp.</i> SBCb004. ....   | 187 |
| Table S3. Oligonucleotides used in this study. ....  | 188 |
| Table S4. Gene synthesis constructs obtained from ATG:biosynthetics GmbH. ....   | 189 |
| Table S5. Plasmids and expression constructs generated in this study.....  | 190 |
| Table S6. NMR spectroscopic data of argyirin I.....  | 194 |
| Table S7. NMR spectroscopic data of argyirin J.....  | 195 |
| Table S8. Restriction sites (R-sites) engineered within the <i>arg</i> BGC and pSynBio3 sequence.....  | 196 |
| Table S9. Sequence features of synthetic <i>arg</i> BGC versions compared to the native sequence.....  | 196 |
| Table S10. HPLC-MS data of argyirin derivatives detected in the heterologous expression strains. ....  | 197 |
| Table S11. Nonribosomal codes of A domains.....  | 198 |
| Table S12. Specificity codes (based on the nonribosomal code) of 3 engineered argyirin A4 domains, aligned with GrsA specificity code and native argyirin A4 domain specificity code (Arg-A4)..... | 199 |
| Table S13. Serine specific A domain specificity codes (based on the 8 Å code) from published myxobacterial BGC, aligned with GrsA 8 Å code.....  | 199 |

#### 4.7.1 Inactivation of the argyirin biosynthetic pathway in *Cystobacter sp.* SBCb004

To verify that the identified BGC is indeed involved in argyirin production, one of the NRPS encoding genes (*arg3*) was inactivated by insertion of the suicide plasmid. For this purpose, the pArg1 plasmid was constructed and a conjugation protocol was established for transformation of *Cystobacter sp.* SBCb004. After genotypic verification of the obtained mutants secondary metabolite production was analyzed. In contrast to the wild type strain, all mutants were shown to lack production of argyrins. Experimental details are described below.

For the construction of pArg1 a 1.1 kb homologous fragment of *arg3* was amplified from genomic DNA using the oligonucleotides Arg1/Arg2 (Table S3) and Taq DNA polymerase according to the manufacturer's protocol. PCR conditions were as follows: initial denaturation 3 min at 95 °C; denaturation 30 s at 95 °C; annealing 50 s at 56 °C; extension 90 s at 72 °C; 30 cycles. The amplified fragment was ligated into the conjugable plasmid pSUP-Km<sup>1</sup>, using the enzymes *EcoRV/BamHI* to generate the suicide plasmid pArg1.

To transform *Cystobacter sp.* SBCb004 with the suicide plasmid pArg1 a suitable conjugation procedure was established as follows: Cultures of the methylation-deficient strain *E. coli* SCS110 carrying pArg1 (SCS110/pArg1), *E. coli* HB101 carrying the helper plasmid pRK600<sup>2</sup> (HB101/pRK600) as well as *Cystobacter sp.* SBCb004 were grown in LB medium (amended with kanamycin for pArg1 or chloramphenicol for pRK600) or rather M medium until an OD<sub>600</sub> of 0.6 was reached. Cells from 2 ml *Cystobacter sp.* SBCB004 culture were washed using 2 x 2 ml HEPES buffer (5 mM HEPES [pH 7.2], 0.5 mM CaCl<sub>2</sub>) and cells from the *E. coli* cultures (1 ml) were washed with LB medium (2 x 1 ml). After combining the cells in 1 ml M medium and shaking for 3 h at 30 °C, the cells were re-suspended in 250 µl fresh M medium and plated out on an M agar plate, which was incubated at 30 °C for 2 d. The cells were scraped from the plate, re-suspended in M medium and plated out on M agar amended with 100 µg/ml kanamycin and 120 µg/ml tobramycin. The plates were incubated at 30 °C until transconjugants appeared (about 3-4 d).

To verify the correct chromosomal integration of pArg1 the obtained transconjugants were analyzed by PCR (see Fig. S2). For this, the mutants were grown in M medium amended with kanamycin and genomic DNA was isolated using the Genra PURGENE® DNA Purification Kit (Qiagen). Using the genomic DNA as template PCR analysis with different sets of oligonucleotides (Arg10/Arg11, Arg10/pSUP\_B, Arg11/pSUP\_E; see Table S3) and Taq DNA polymerase was performed according to the manufacturer's protocol. PCR conditions were as follows: initial denaturation 3 min at 95 °C; denaturation 30 s at 95 °C; annealing 50 s

at 56 °C; extension 90 s at 72 °C; 30 cycles. PCRs using oligonucleotides Arg10/pSUP\_B and Arg11/pSUP\_E produced DNA fragments of expected size in the *Cystobacter* sp. SBCb004::pArg1 mutants but not in *Cystobacter* sp. SBCb004 verifying that pArg1 was successfully integrated into gene *arg3* on the chromosome (Fig. S2).

For phenotypic analysis the *Cystobacter* sp. SBCb004::pArg1 mutant strains as well as *Cystobacter* sp. SBCb004 wild-type were cultivated in 50 ml M liquid medium containing 2% XAD-16 for 4 days. The cells together with the XAD were harvested by centrifugation and extracted with methanol. To compare the spectrum of secondary metabolites produced by the mutant with that of the wild type strain, 50:1 concentrated methanol extracts were analysed by High performance liquid chromatography-mass spectrometry-diode array detection (HPLC-MS-DAD). An Agilent 1100 series solvent delivery system coupled to Bruker HCTplus ion trap mass spectrometer was used. Chromatographic separation was carried out on an RP column Nucleodur C18 (125 by 2 mm, 3 µm particle size; Macherey-Nagel) equipped with a precolumn C18 (8 x 3 mm, 5 µm). The mobile phase gradient (solvent A: water + 0.1% formic acid and solvent B: acetonitrile + 0.1% formic acid) was linear from 5% B at 2 min to 95% B at 32 min, followed by 4 min with 95% B at a flow rate of 0.4 ml/min. Diode array detection was carried out at 200-600 nm and mass detection was performed in the positive and negative ionization mode within a range of 100-1100 amu. As shown in Fig. S3, the wild type culture produced argyrim A, argyrim B, argyrim D, and argyrins E to H. On the contrary, the mutant did not produce any of these argyrins, demonstrating the involvement of enzyme encoded by *arg3* in the argyrim biosynthesis.

#### 4.7.2 Feeding studies with *Cystobacter* sp. SBCb004

To investigate the biosynthetic origin of the *L*-α-aminobutyrate (Abu) and dehydroalanine (Dha) units of argyrins, feeding studies with *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N and *L*-serine-2,3,3-D<sub>3</sub> were performed with the native producer *Cystobacter* sp. SBCb004. Experimental details and results are described below.

From a well-grown agar plate of *Cystobacter* sp. SBCb004 culture flasks with 50 ml M medium were inoculated and incubated on a rotary shaker (200 rpm) at 30 °C for 4 days. Solutions of isotopically labelled precursors were prepared by dissolving 25 mg of *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N (Isotec) and *L*-serine-2,3,3-D<sub>3</sub> (CIL) in 1.5 ml water followed by sterile filtration. Equal aliquots (300 µl) were fed to the production cultures after 24 h, 29 h, 32 h, 48 h and 56 h of incubation in parallel to a control experiment where no labelled precursor

was added. Amberlite XAD-16 adsorber resin (2%) was added after 72 h to all cultures. After 90 h of incubation the cultures were harvested by centrifugation and cells/XAD were extracted with 50 ml of methanol. Concentrated methanol extracts from the feeding experiments and the respective controls without feeding were analyzed by HPLC-MS. Separation was performed on a Dionex UltiMate 3000 RSLC system equipped with a Waters BEH C18 (100 × 2.1 mm, 1.7 μm dp) column using a linear gradient with solvent A (water + 0.1% formic acid) and B (acetonitrile + 0.1% formic acid) at a flow rate of 550 μl/min and 45 °C. The gradient was initiated by a 0.39 min isocratic step with 5% B followed by an increase to 95% B within 18 min. HPLC was coupled to a Thermo Fisher Orbitrap via an Advion Triversa NanoMate nano-ESI system. Mass spectra were acquired in positive ionization mode with a range of 200-2000  $m/z$  at a resolution of  $R=30000$ . Identities of the argyrins were confirmed by comparing with both retention time and  $MS^2$  fragmentation pattern of the authentic compounds. Results from the HPLC-MS analysis of the crude extracts are shown in Fig. S4/S5. Argyrin A and argyrin B exhibit in MS analysis three apparent isotopic signals with the monoisotopic signal  $[M+H]^+$  of 825.3143  $m/z$  (argyrin A) and 839.3300  $m/z$  (argyrin B). After feeding of the labelled precursors additional isotope signals appeared with a maximum intensity of a peak of +2.0125  $m/z$  for *L*-serine-2,3,3-D<sub>3</sub> (analyzed for argyrin A, Fig. S4) and +4.0134  $m/z$  for *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N (analyzed for argyrin B; Fig. S5). These data indicate that the Dha unit indeed originates from serine and that the ethyl group of the peptide core from argyrin B (as well as from argyrins D and G) results from incorporation of an Abu unit. The latter is supplied by primary metabolism through conversion of *L*-threonine by threonine deaminase and aminotransferase enzymes as illustrated in Fig. S5C<sup>3</sup>.

#### 4.7.3 Isolation and structure elucidation of argyrins I and J

During re-evaluation of extract fractions from large-scale fermentations of native argyrin producers<sup>4</sup>, two novel argyrin derivatives (argyrin I and J, Fig. S6) were isolated (chromatographic data not shown). Structure elucidation was achieved using NMR spectroscopy and the high-resolution mass spectrometry (HR-MS). High resolution mass spectra were acquired on an UPLC coupled Bruker Daltonics maXis 4G qToF mass spectrometer in a range from 150-2500  $m/z$  in positive ionization mode. Separation was carried out on a Waters BEH C18 100 x 2.1 mm, 1.7 μm dp column using H<sub>2</sub>O + 0.1% formic acid as A and acetonitrile + 0.1% formic acid as B. Separation was carried out on a linear

gradient from A to B over 16 minutes. HR-MS data of argyrin I yielded an  $[M+H]^+$  ion at  $m/z$  809.3193 consistent with a molecular formula  $C_{40}H_{44}N_{10}O_7S$  (809.3188 calculated for  $C_{40}H_{45}N_{10}O_7S$ ). HR-MS data of argyrin J yielded an  $[M+H]^+$  ion at  $m/z$  825.3204 consistent with a molecular formula  $C_{40}H_{44}N_{10}O_8S$  (825.3137 calculated for  $C_{40}H_{45}N_{10}O_8S$ ). NMR spectra were acquired in  $CDCl_3$  at a Bruker Ascend 700 MHz spectrometer. 1D  $^1H$  as well as 2D  $^1H$ - $^1H$  COSY, HSQC and HMBC spectra were recorded using standard pulse programs and are illustrated in Fig. S7/S8. Carbon chemical shifts were extracted from 2D NMR data. NMR spectroscopic data are listed in Tables S6/S7. The NMR data of argyrin I and J were in accordance with those of argyrin E and H, which harbor a methyl group instead of an ethyl group in position  $R_1$ . It is assumed that the absolute configuration of argyrin I and J corresponds to argyrin A-H as the argyrin megasynthetase very likely applies the same stereochemistry control.

#### 4.7.4 Sequence design of synthetic arg BGCs (arg-V1, arg-V2, arg-V1-BsaI)

To meet the constructive demands for *arg* BGC assembly based on synthetic building blocks and prospective modifications, endonuclease restriction sites (R-sites) were engineered as illustrated in Fig. 2A. This included the insertion of unique R-sites upstream and downstream of each gene, within the 5' end of *arg2* as well as in NRPS intermodule linker regions of *arg2* and *arg3*. Simultaneously, 59 interfering R-sites along the entire *arg* BGC sequence were eliminated. R-sites engineering was performed by point mutations (synonymous codon substitutions in coding regions) according to a list of 22 restriction enzymes selected for *arg* BGC and pSynBio3 vector backbone design (Table S8).

No further sequence modulation was performed in case of the *arg*-V1 BGC, whereat the *arg*-V2 sequence was additionally adapted by applying a similar strategy as previously reported for construction of a synthetic epothilone BGC<sup>5</sup>. An artificial host codon usage table (“myxo-PKS/NRPS”) was used as basis for formal sequence optimization with particular emphasis on parameters affecting translational elongation. The relative adaptiveness scoring system implemented in the codon adaptation index (CAI;<sup>6</sup>) was used to modulate the course of the local CAI in the four *arg*-V2 coding sequences (CDS). Due to the assumption that the translation rate of a single codon is to some degree correlated with its synonymous codon fraction, a slight gradient was engineered in the four *arg*-V2 CDSs (Fig. 2B), which theoretically should improve the ribosomal packing ratio along the whole message if the translation initiation rate is sufficiently high. The occurrence of internal Shine-Dalgarno (SD)

sequences was strongly suppressed to avoid unfavorable duplex formation with the anti-SD sequence of the 16S rRNA during translation<sup>7</sup>. As illustrated in Fig. S9A, the *arg-V2* CDSs show no strong signals indicating SD – ASD interactions in contrast to *arg-native*. In addition, the clustered occurrence of low frequency codons, which could cause ribosome stalling and abortion during translation, was avoided. The resulting *arg-V2* sequence lacks pronounced downward local CAI peaks according to a smaller 9 codon window<sup>8</sup> (Fig. S9B). As it is assumed that “hidden” stop codons can have a positive impact on gene expression by preventing off-frame gene reading<sup>9,10</sup>, the density of stop codons in neighbouring frames was increased in the *arg-V2* sequence (Fig. S9C). Furthermore, the emergence of homopolymeric stretches, which can cause transcriptional slippage, as well as DNA sequence repeats, was suppressed during *arg-V2* optimization and the GC-content was kept within preselected limits. Overall, sequence modifications were more extensive for *arg-V2* compared to *arg-V1*, for which only R-sites engineering was applied. To facilitate future BGC engineering procedures, the *arg-V1* BGC sequence was later on further modified by eliminating all 32 *BsaI* R-sites via synonymous codon substitutions (*arg-V1-BsaI*). Table S9 shows a comparison between *arg-native* and the adapted gene sequences of *arg-V1*, *arg-V2* and *arg-V1-BsaI*.

#### 4.7.5 Construction of *M. xanthus* DK1622 $\Delta$ *mchA-tet* and *arg* BGC expression strains

Prior to the transfer of the generated expression constructs into *M. xanthus*, the strain was engineered to delete the *mch<sub>A</sub>* BGC, which is in charge for production of the lipopeptide family myxochromides A<sup>11</sup>. A suitable suicide vector was constructed, which enables replacement of the *mch<sub>A</sub>* BGC region with a tetracycline resistance gene (*tet<sup>R</sup>*) via two successive homologous recombination events (Fig. S10). Two homologous fragments, each around 1 kb, were amplified by PCR from chromosomal DNA of *M. xanthus* DK1622 wild-type using the oligonucleotide pairs Mch140/Mch141 and Mch165/Mch166, respectively (Table S3). PCR was carried out with Taq DNA polymerase according to general procedures (see Materials and methods section 4.5.3). After gel purification, the PCR products were hydrolyzed using the restriction enzymes *SacI/XbaI* for the Mch140/Mch141 amplicon and *XbaI/BamHI* for the Mch165/mch166 amplicon. In parallel, an established suicide vector for *M. xanthus*, pSWU41<sup>12</sup>, was hydrolyzed with *SacI/BamHI*. The linearized vector backbone was ligated with the two homology fragments in one step (triple ligation) yielding the plasmid pMyx-Mch2. Next, a fragment harboring the tetracycline resistance gene (*tet<sup>R</sup>*), which confers resistance to oxytetracycline in *M. xanthus*, was amplified from pSWU22 (S. S. Wu and

D. Kaiser, unpublished results) using the oligonucleotides TetR-for/TetR-rev (Table S3) and Taq DNA polymerase. The resulting 1.4 kb *tetR* amplicon was gel-purified, hydrolyzed with *XbaI* and subsequently ligated with *XbaI* linearized pMyx-Mch2 yielding the construct pMyx-Mch2-tet. After transformation of *M. xanthus* DK1622 with pMyx-Mch2-tet via electroporation according to established procedures<sup>13</sup> transformants were selected on CTT medium amended with kanamycin (50 µg/ml). Chromosomal integration via either of the two homology regions was verified by Southern Blot analysis (data not shown). A verified mutant was selected and grown in CTT medium amended with oxytetracycline (12.5 µg/ml). After 3-4 days cultivation at 30 °C the well-grown culture was used to inoculate fresh medium. This procedure was repeated three times to increase the possibility of a second crossover event. After repetitive cultivation cycles different dilutions of the cell population were plated out on CTT medium amended with oxytetracycline (12.5 µg/ml) to select for the presence of the *tet<sup>R</sup>* gene and sucrose (6%) for counterselection based on the *sacB* gene from the pSWU41 vector backbone<sup>12</sup>. Under these selection conditions only mutants, in which the suicide plasmid was eliminated and the *mch<sub>A</sub>* BGC replaced with the *tet<sup>R</sup>* gene were expected. After 7-10 days the first colonies appeared, which were grown in CTT medium amended with oxytetracycline (12.5 µg/ml) to isolate genomic DNA for genotypic verification via Southern Blot analysis and PCR (data not shown). The resulting mutant was cultivated under standard conditions in parallel to *M. xanthus* DK1622 wild-type and culture extracts were subsequently analyzed by HPLC-MS as described in a previous study on myxochromide production<sup>14</sup>. Myxochromides A were detected in wild-type extracts, but not in extracts of the mutant *M. xanthus* DK1622 *ΔmchA-tet* confirming the expected phenotype after successful deletion of the *mch<sub>A</sub>* BGC.

The engineered host strain *M. xanthus* DK1622 *ΔmchA-tet* was transformed with the generated expression constructs by electroporation according to previously established procedures<sup>13</sup>. Transformants were selected on CTT agar containing kanamycin (50 µg/ml). Colonies usually appeared after 5-10 days. For each construct, several colonies were picked and transferred onto new plates. Correct chromosomal integration of the expression constructs via homologous recombination into the *tet<sup>R</sup>* locus was verified by ‘colony-PCR’. For this, cells were dissolved in 50 µl H<sub>2</sub>O and lysed by incubation at 95 °C for 30 min prior of 2 µl suspension being added to the PCR reaction. PCR was performed according to general procedures (see Materials and methods section 4.5.3). As illustrated in Fig S10, correct chromosomal integration of the expression constructs was confirmed using two primer combinations revealing PCR products of the expected sizes: P1/P2 (1458 bp) and P3/P4 (1461 bp). Genomic DNA of *M. xanthus* DK1622 *ΔmchA-tet* was used as negative control. A

complementary experiment using primers P1/P4 revealed a 1461 bp PCR product for *M. xanthus* DK1622  $\Delta$ *mchA-tet*, but not for the expression strains. In addition, primer sets to detect fragments of *arg* genes were designed to further verify the integrity of the *arg* BGC variants by additional PCRs (respective primer pair and expected PCR product size in brackets): *arg2-V1* (*arg2V1F/arg2V1R*, 428 bp), *arg3-V1* (*arg3V1F/arg3V1R*, 363 bp), *arg4-V1* (*arg4V1F/arg4V1R*, 582 bp), *arg5-V1* (*arg5V1F/arg5V1R*, 516 bp), *arg2-V2* (*arg2V2F/arg2V2R*, 512 bp), *arg3-V2* (*arg3V2F/arg3V2R*, 446 bp), *arg4-V2* (*arg4V2F/arg4V2R*, 476 bp) and *arg5-V2* (*arg5V2F/arg5V2R*, 463 bp). Primer sequences are listed in Table S3. Expression strains generated in this study are listed in Table 1.

#### 4.7.6 Analysis of argyirin production in *M. xanthus* and valine feeding experiment

Single colonies of the expression strains with verified genotype were grown in 50 ml CTT medium amended with 2% XAD-16 adsorber resin and 50  $\mu$ g/ml kanamycin for 5-6 days. After centrifugation, the cells and XAD were extracted with methanol and concentrated crude extracts were analyzed by HPLC-MS. For qualitative analysis of argyirin production profiles, the following HPLC-MS method was applied. Separation was performed on a Dionex UltiMate 3000 RSLC system equipped with a Waters reversed phase UPLC column (Acquity UPLC BEH C18 1,7 $\mu$ m; 2.1\*100mm) using a linear gradient with solvent A (water + 0.1% formic acid) and B (acetonitrile + 0.1% formic acid) at a flow rate of 600  $\mu$ l/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step with 5% B followed by an increase to 15% B within 1 min, 50% B within 11.5 min and 95% B within 1 min, which was kept for 1 min, before decreasing back to initial conditions of 5% B within 0.3 min and was kept for 1.7 min. HPLC was coupled to a Bruker Daltonics QqToF mass spec maxis4G system in a mass range from 150-2500 *m/z* in positive ionization mode. Structures of the new derivatives were characterized by comparing the MS<sup>2</sup> spectra with those of the known argyrins (Table S9). Argyrin production profiles of different *M. xanthus* expression strains in comparison to the native argyirin producer are illustrated in Fig. 3.

To quantify argyirin production levels in expression strains harboring the *arg2345* gene set (producing argyrins A and B as major derivatives), the following HPLC-MS method was applied. Separation was performed on a Dionex UltiMate 3000 RSLC system equipped with a Waters reversed phase UPLC column (Acquity UPLC BEH C18 1,7 $\mu$ m; 2.1\*100mm) using a linear gradient with solvent A (water + 0.1% formic acid) and B (acetonitrile + 0.1% formic acid) at a flow rate of 600  $\mu$ l/min and 45 °C. The gradient was initiated by a 0.5 min isocratic

step with 5% B followed by an increase to 15% B within 1 min, 50% B within 11.5 min and 95% B within 1 min, which was kept for 1 min, before decreasing back to initial conditions of 5% B within 0.3 min and was kept for 1.7 min. HPLC was coupled to a Bruker Daltonics ion trap mass spec ‘Amazon speed’ system. Mass spectra were acquired in positive ionization mode with a range of 200-2500  $m/z$  at a resolution of  $R=30000$ . Identities of the argyryns were confirmed by comparing with both retention time and  $MS^2$  fragmentation pattern of the authentic compounds.

To further verify the incorporation of valine into argyryns K and L ( $R_1 = CH_2(CH_3)_2$ ), a feeding experiment with *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-5-V1$  growing in 50 ml CTT medium (50  $\mu g/ml$  kanamycin) was performed. Around 29 mg *L*-valine-D8 (Deutero GmbH) was dissolved in 8 ml  $H_2O$ , sterile filtered and equal aliquots were added to the culture after 48 h, 60 h, 72 h, 84 h, 96 h, 108 h, 120 h, and 132 h of growth. Amberlite XAD-16 adsorber resin (2%) was added after 144 h. After 168 h of cultivation, cells and XAD were harvested, extracted with methanol and analyzed by HPLC-MS as described in chapter 2. Argyryns K and L exhibit in MS analysis the monoisotopic signal  $[M+H]^+$  of 853.3469  $m/z$  (argyryn K) and 869.3412  $m/z$  (argyryn L). As shown in Fig. S11, additional isotope signals appeared with a maximum intensity of a peak of +7.0439  $m/z$  indicating the incorporation of *L*-valine-D8 during biosynthesis of both derivatives (Deuterium label in  $\alpha$ -position of *L*-valine-D8 lost because of frequent H-D exchange).

#### 4.7.7 Engineering of the A1 domain of the argyryn NRPS megasynthetase

To rationally engineer argyryn the A1 specificity conferring code<sup>15</sup>, A domains from myxobacterial BGCs with known substrate specificity were used as template. Eight A domains, from published myxobacterial NRPS megasynthetases, with specificity for alanine,  $\alpha$ -aminobutyric acid or valine were selected and their nonribosomal codes were extracted (Table S11). A panel of modified A1 versions was designed, in which residues of the substrate-binding pocket were mutated according to the extracted nonribosomal codes.

In most cases the original code from “template” A domain was implemented, to adapt the original argyryn A1, but in some cases a hybrid code between original argyryn A1 and the “template” A domain was generated instead (Table S11). To avoid any bottlenecks, a codon usage bias of the selected surrogate host was taken into account, meaning that for each amino acid residue that was mutated, a codon with high prevalence in the *M. xanthus* DK1622

*AmchA-tet* host genome was employed. Synthetic sequences for 14 modified argyrim A1 domain fragments were designed based on the described procedure (Table S4).

Argyrim expression constructs with engineered A1 domains were constructed as shown in Fig. S12. Synthetic fragments of 14 modified argyrim A1 domain regions, flanked by *AarI* endonuclease restriction sites were obtained from a gene synthesis company (Table S4). *AarI* is a type IIS restriction endonuclease that cuts DNA outside of its recognition sequence, which allows construction of desired overhangs that are complementary to any selected target sequence. Construct pGH-arg2-M1-VI harboring first module of the Arg2 argyrim subunit was also obtained from a gene synthesis company for construction of the synthetic argyrim BGC (see chapter 4.5.5). The backbone of the construct was replaced with the pUC18 vector backbone by traditional restriction-ligation using *NdeI* and *SdaI* restriction endonucleases. Resulting construct pUC18-arg2-M1-V1, now harboring first module of the Arg2 argyrim subunit, was further modified using Red/ET *in vivo* recombination technique, to replace A1 subdomain region between residues 235 and 330 (based on the GrsA numbering<sup>15</sup>) with kanamycin resistance cassette, flanked by *AarI* endonuclease restriction sites. The cassette was generated by PCR amplification from pTpSmchS<sup>16</sup> template using Arg42 and Arg43 primer pair (Table S3). The resulting construct pUC18-arg2-M1-V1-Del-Pos235-330 as well as 14 synthetic A domain fragments, subcloned into standard gene synthesis vectors (pGH-A1-mod[1-14], see Table S4), were digested using *AarI* enzyme and subsequently ligated, to generate 14 modified constructs harboring modified first module (M1) of argyrim BGC. The resulting 14 modified M1 modules were sub-cloned in pArg2345-V1 argyrim expression construct by traditional restriction-ligation using *NdeI* and *SbfI* restriction endonucleases (Fig. S12). Finally 14 modified argyrim heterologous expression constructs (pArg2345-V1-mod[1-14], see Table S5) were generated and the modified region was verified by Sanger sequencing using the primer Arg51 (Table S3).

Generated argyrim BGC expression constructs with modified A1 were transferred into *M. xanthus* DK1622 *AmchA-tet* by electroporation. Clones growing on selection plates amended with kanamycin were analyzed for correct chromosomal integration of the expression plasmid by PCR (Fig. S10). To evaluate the production profile of argyrim, the heterologous producers (A1 domain modifications: *M. xanthus* DK1622 *AmchA-tet*::pArg2345-V1-mod[1-14]; see Table S5) and control strain *M. xanthus* DK1622 *AmchA-tet*::pArg2345-V1 were cultivated in parallel. Strains were inoculated from cryo stocks and grown on agar plates for several days until plates were mostly overgrown with cells. All of the cells were scraped from the plates to inoculate seed medium (50 ml medium in 300 ml

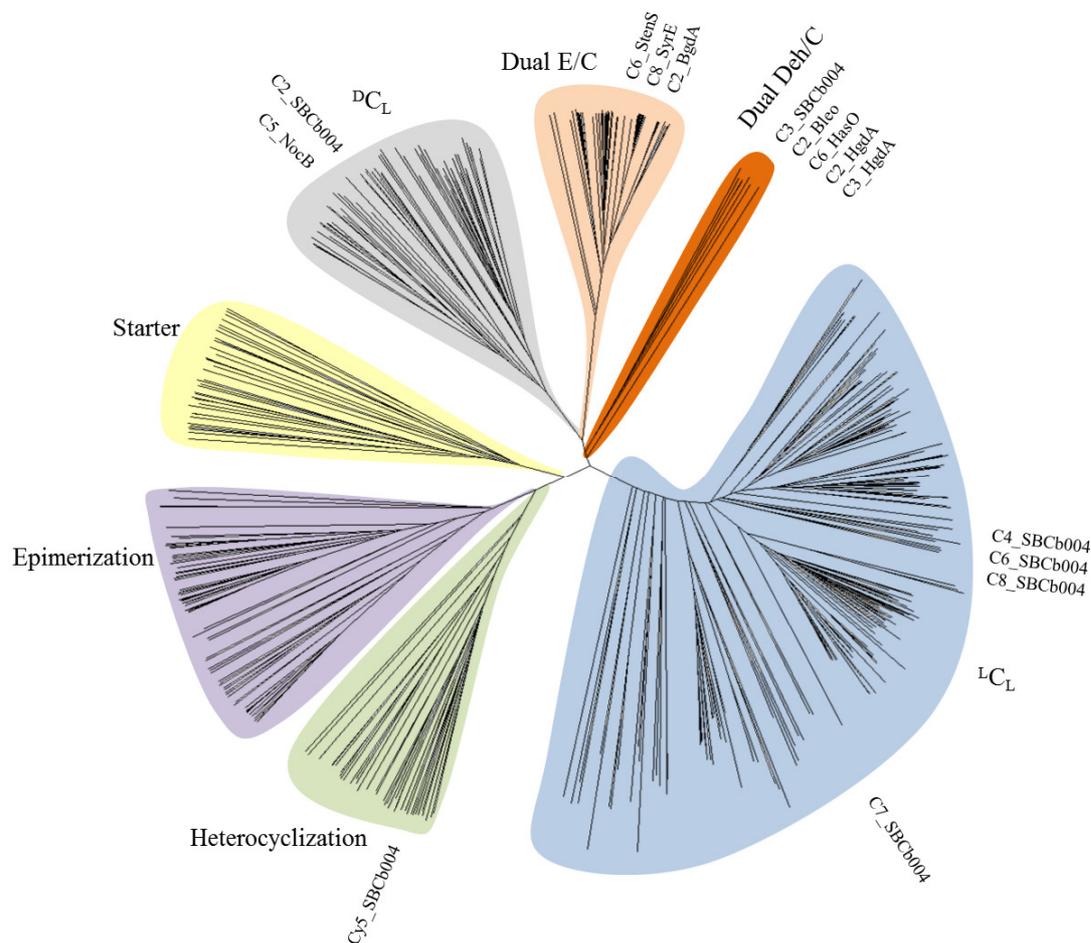
Erlenmeyer flask), which was incubated at 30 °C, 180 rpm for 48 h. Five ml of well grown seed culture were used to inoculate 50 ml production medium in which the strain was grown at the same conditions for 6 days. *M. xanthus* DK1622 derivatives were grown in CTT medium (see 4.5.2) supplemented with 50 µg/ml kanamycin and 2% Amberlite XAD-16 adsorber resin. All cultivations were performed in triplicates at 30 °C and 180 rpm for 6 days. Cultures were harvested by centrifugation at 8000 rpm and 4 °C for 15 min. Supernatant was removed and pelleted cells with XAD were extracted with 1:1 mixture of methanol and acetone (25 ml of methanol and 25 ml acetone). After filtration, the crude extracts were evaporated to dryness and subsequently re-dissolved in 1 ml of methanol for HPLC-MS analysis (see chapter 2). Argyrin peaks were identified by EIC  $[M + H]^+ = 825.313$  (arg A) and  $[M + H]^+ = 839.329$  (arg B) and their yields were evaluated by integration and relative comparison of the peak surface area.

#### 4.7.8 Medium optimization for yield improvement

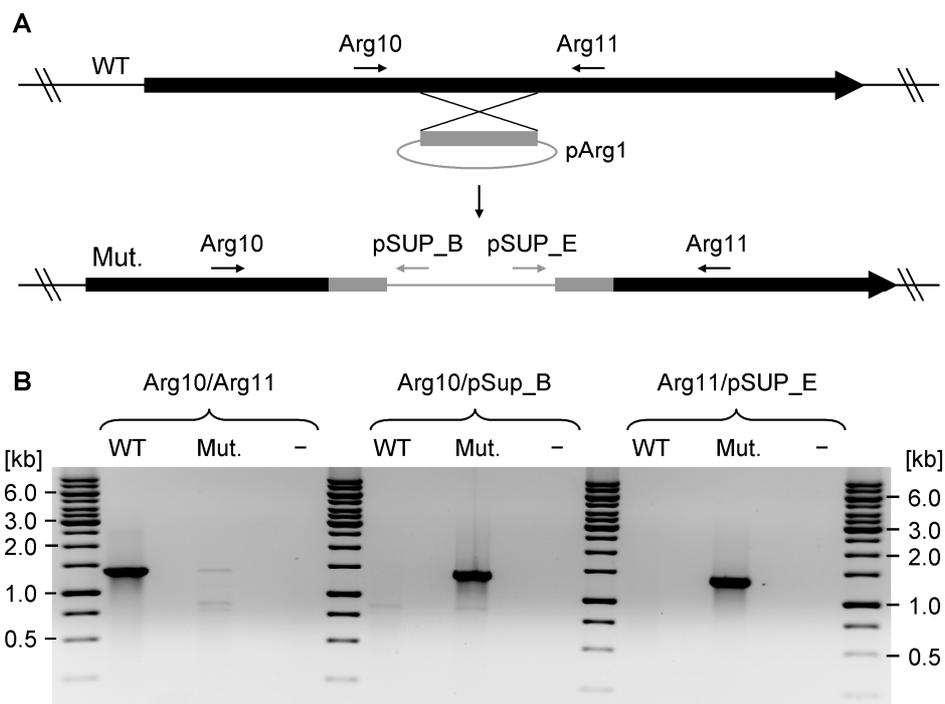
To further improve argyirin production levels, M7/s4 (0.5% soy flour, 0.5% corn starch, 0.2% glucose, 0.1% yeast extract, 0.1%  $MgSO_4 \times 7 H_2O$ , 0.1 %  $CaCl_2 \times 2 H_2O$ , 1% HEPES, with final pH 7.4 and supplemented with 0.1 mg/l of vitamin B12 and 5 mg/l of  $FeCl_3$  after autoclaving) and M7/s6 (0.5% soy flour, 0.5% corn starch, 0.2% glucose, 0.1% yeast extract, 1% potassium acetate, 0.1%  $MgSO_4 \times 7 H_2O$ , 0.1 %  $CaCl_2 \times 2 H_2O$ , 1% HEPES, with final pH 7.4 and supplemented with 0.1 mg/l of vitamin B12 and 5 mg/l of  $FeCl_3$  after autoclaving) media were investigated as alternative options. The M7/s4 medium was identified to be the best candidate for further optimization. Experimental details and results are described below.

From a well-grown agar plate of *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  culture, flasks with 50 ml M7/s4 medium were inoculated and incubated on a rotary shaker (180 rpm) at 30 °C for 6 days. Solution of amino acid mixture calculated for the final concentration of 5 mM serine, 5 mM cysteine, 5 mM alanine, 10 mM tryptophan, 10 mM glycine and 10 mM  $\alpha$ -aminobutyric acid was fed to the production cultures twice daily over the course of 6 days. The control experiment without supplementing the medium with the amino acid solution was performed in parallel. Amberlite XAD-16 adsorber resin (2%) was added after 24 h to all cultures. After 6 days of incubation the cultures were harvested by centrifugation and cells/XAD were extracted with 50 ml of ethyl acetate. Concentrated methanol extracts from the feeding experiments and the respective controls without feeding were analyzed by HPLC-MS and argyirin production was quantified as described in chapter 4.7.6.

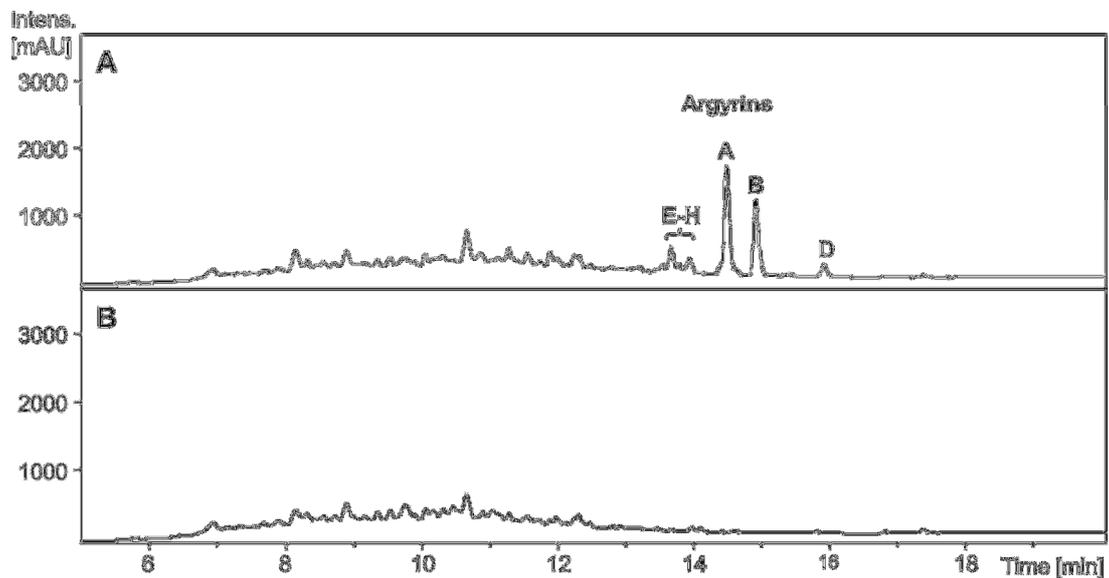
## 4.7.9 Figures



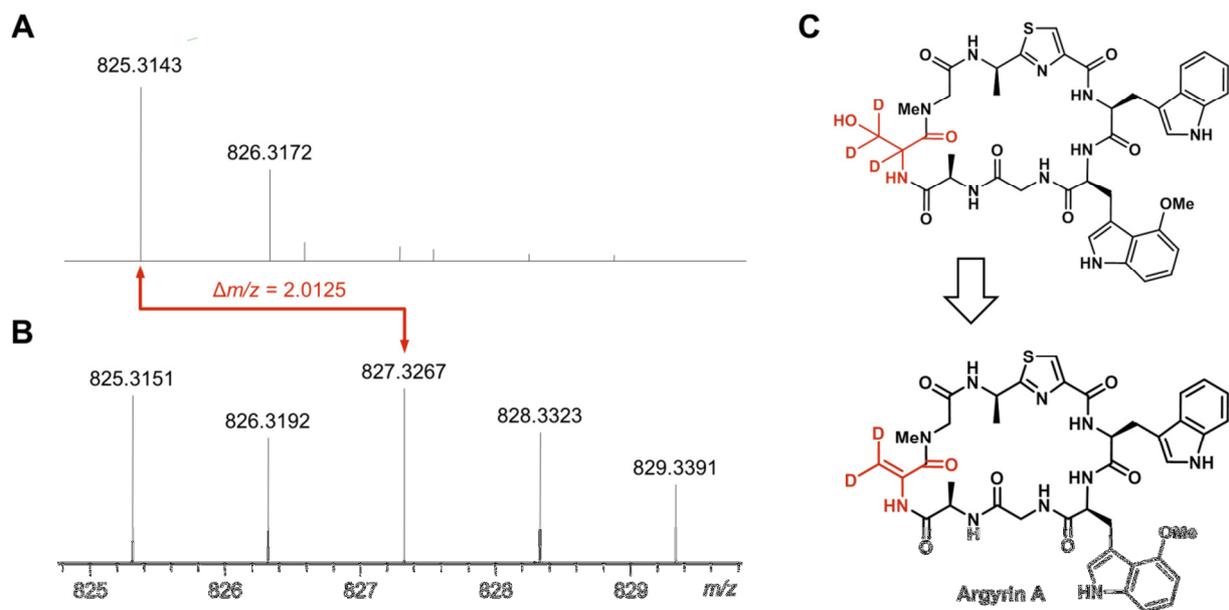
**Figure S1.** Phylogenetic tree of different C domain subtypes. Phylogenetic tree of all known C domain subtypes ( $^L$ C<sub>L</sub>,  $^D$ C<sub>L</sub>, Starter, Dual E/C, Epimerization and Heterocyclization domains), with the new proposed Dual Deh/C subtype highlighted in orange. The phylogeny was reconstructed using phymI, employing the JTT model of amino acid substitution and a gamma-distributed rate variation with four categories. The support values are based on 100-fold bootstrapping. The C domain list includes 525 domains from phylogenetic study by Rausch et. al.<sup>17</sup>, seven C domains from the argyirin NRPS (C2\_SBCb004, C3\_SBCb004, C4\_SBCb004, Cy5\_SBCb004, C6\_SBCb004, C7\_SBCb004 and C8\_SBCb004) as well as selected C domain examples from  $\alpha,\beta$ -dehydro amino acid forming pathways: bleomycin (blmVI; Q9FB23), burriogladin (BgdA; MH170348), haerogladin (HgdA, MH170356), nocardicin (NocB; Q5J1Q6), hassallidin (HasO; K7VZQ9), syringomycin (SyrE; O85168), stenothricin (StenS; EFE73312.1),



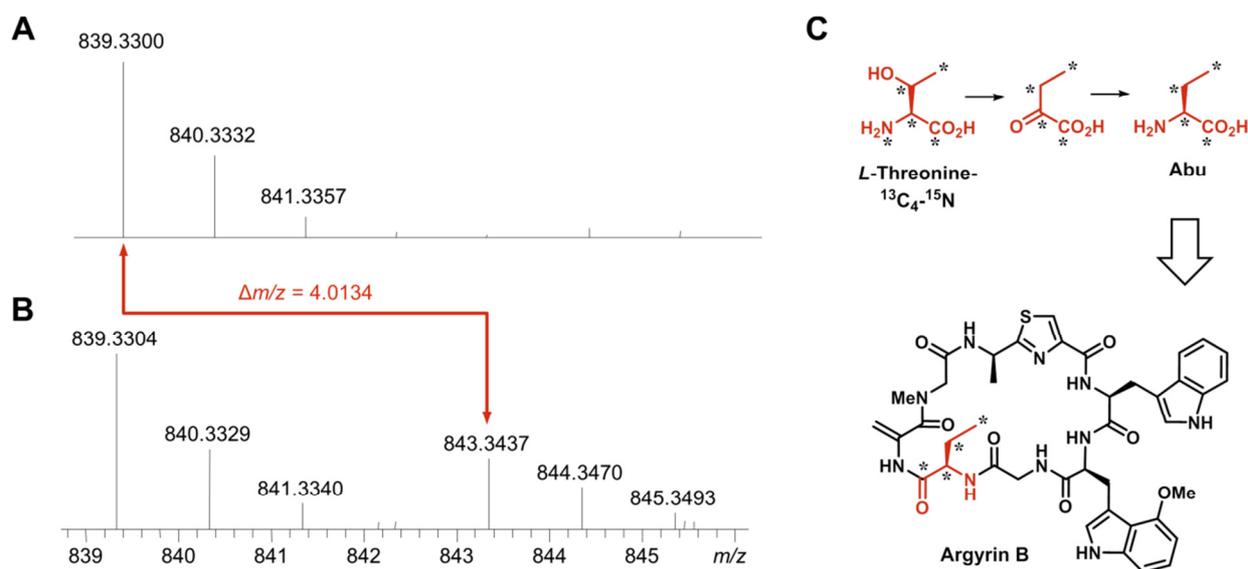
**Figure S2.** Inactivation of *arg3* in *Cystobacter sp.* SBCb004 by insertional mutagenesis. Integration of the suicide plasmid pArg1 via homologous recombination (A), and verification of the mutant strains via PCR analysis (B).



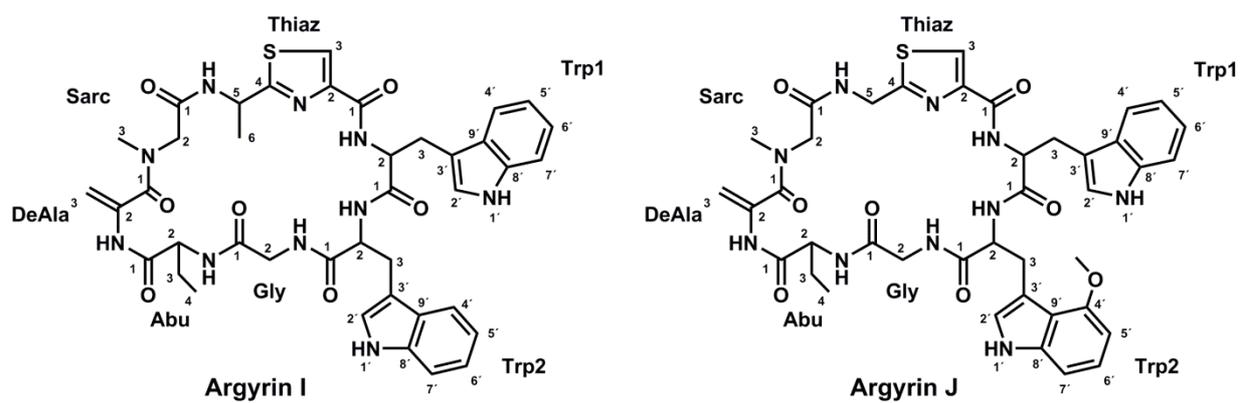
**Figure S3.** Argyrin production analysis after *arg3* inactivation in *Cystobacter sp.* SBCb004. Culture extracts of *Cystobacter sp.* SBCb004 wild-type (A) and the *arg3* inactivation mutant strain *Cystobacter sp.* SBCb004::pArg1 (B) were analyzed by HPLC-MS-DAD (DAD chromatograms at 200-600 nm shown). Argyrins A-E were detected in the wild-type, but not in the mutant extract.



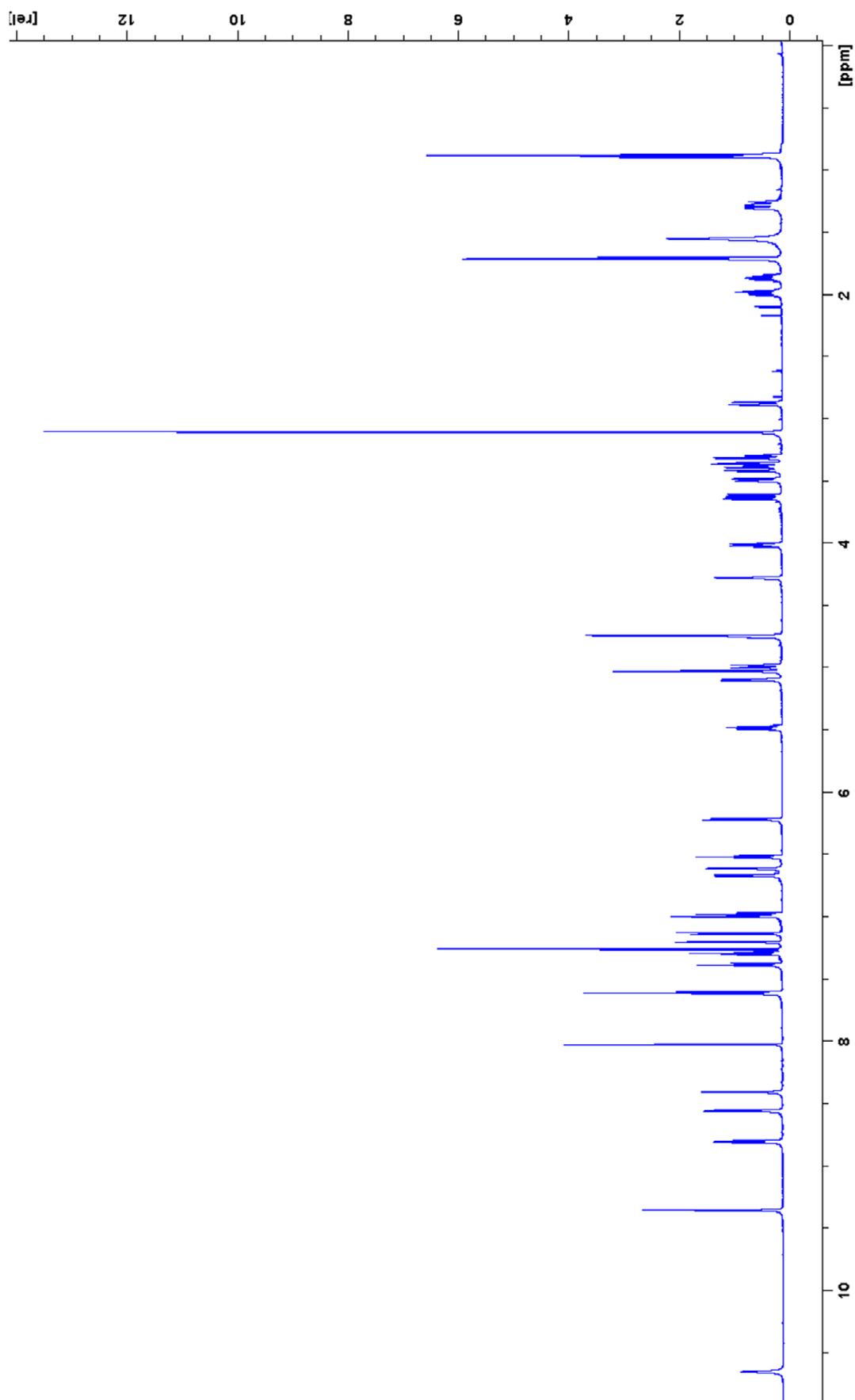
**Figure S4.** Analysis of the incorporation *L*-serine-2,3,3-D<sub>3</sub> during argyrin biosynthesis. MS-spectra of argyrin A after HPLC-MS analysis of culture extracts of *Cystobacter sp.* SBCb004. The monoisotopic signal of the argyrin A [M+H]<sup>+</sup> ion is at 825.3143 m/z (calc. 825.3137 m/z, C<sub>40</sub>H<sub>45</sub>N<sub>10</sub>O<sub>8</sub>S). (A) Control sample to which no labelled precursor was fed, (B) sample from feeding experiment with *L*-serine-2,3,3-D<sub>3</sub>, (C) expected labelling pattern of argyrin A after incorporation of *L*-serine-2,3,3-D<sub>3</sub> and Dha unit generation.

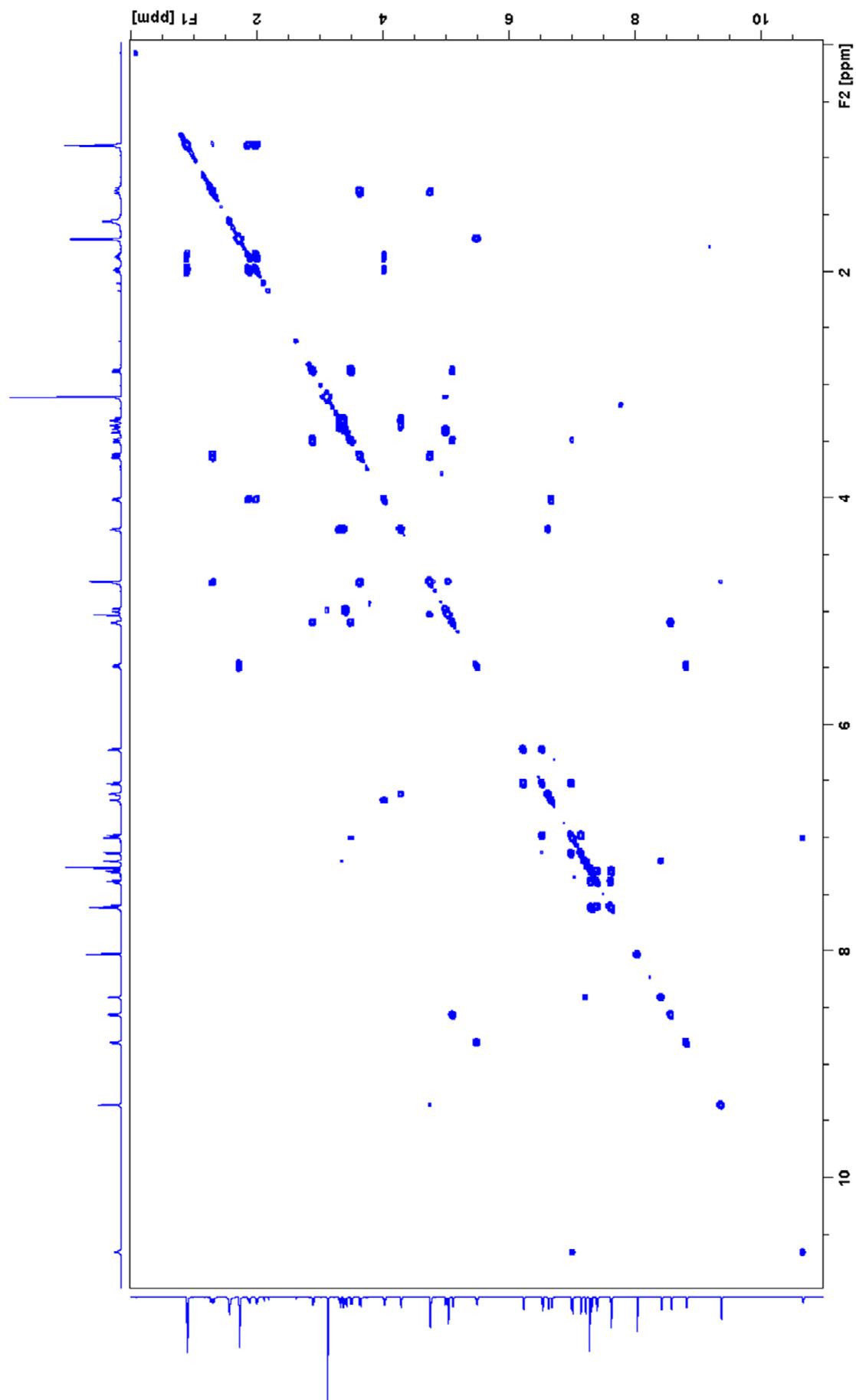


**Figure S5.** Analysis of the incorporation *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N during argyrin biosynthesis. MS-spectra of argyrin B after HPLC-MS analysis of culture extracts of *Cystobacter sp.* SBCb004. The monoisotopic signal of the argyrin B [M+H]<sup>+</sup> ion is at 839.3300 m/z (calc. 839.3294 m/z, C<sub>41</sub>H<sub>47</sub>N<sub>10</sub>O<sub>8</sub>S). (A) Control sample to which no labelled precursor was fed, (B) sample from feeding experiment with *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N, (C) expected labelling pattern of argyrin B after incorporation of Abu generated from *L*-threonine-<sup>13</sup>C<sub>4</sub>, <sup>15</sup>N metabolism by threonine deaminase and aminotransferase enzymes<sup>3</sup>.



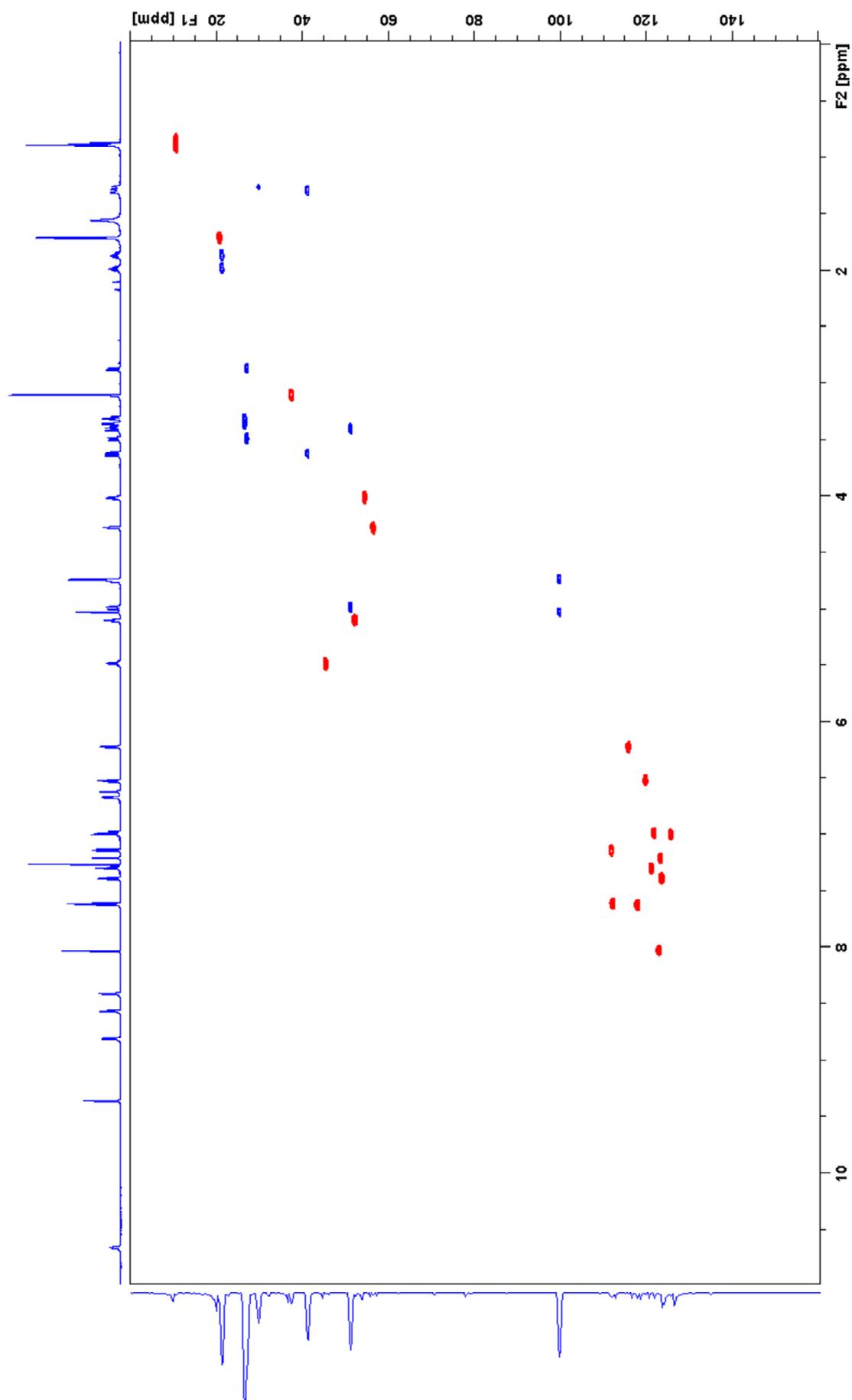
**Figure S6.** Structures of argyrin I and argyrin J. For NMR data see Tables S6/S7.

 $^1\text{H-NMR}$  spectrum of argyrin I in  $\text{CDCl}_3$  (700



$^1\text{H}$ - $^1\text{H}$ -COSY spectrum of argyrin I in  $\text{CDCl}_3$  (700)

Figure S7 (continued on next page)

HSQC spectrum of argyrin I in  $\text{CDCl}_3$  (700 MHz)

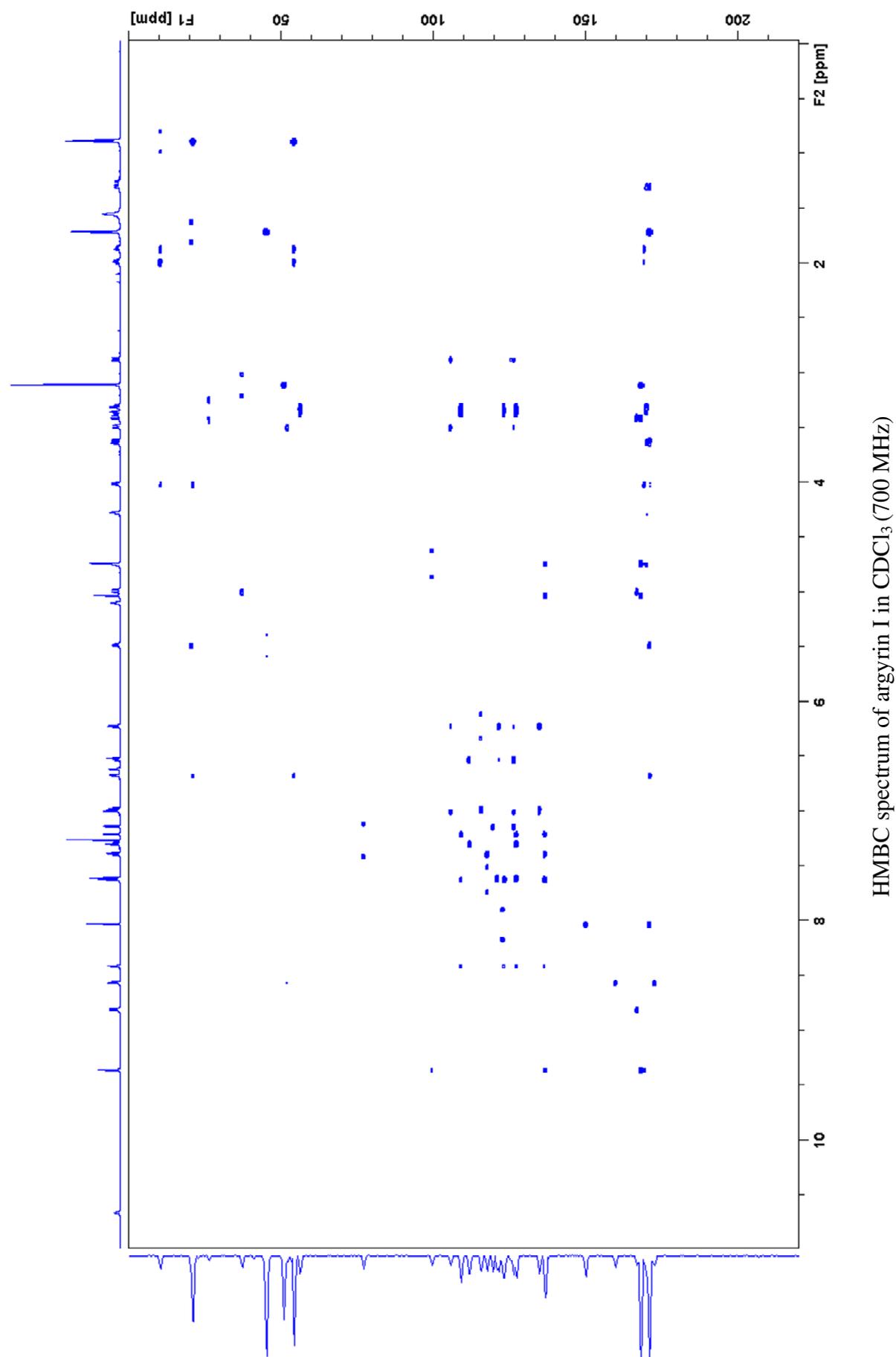
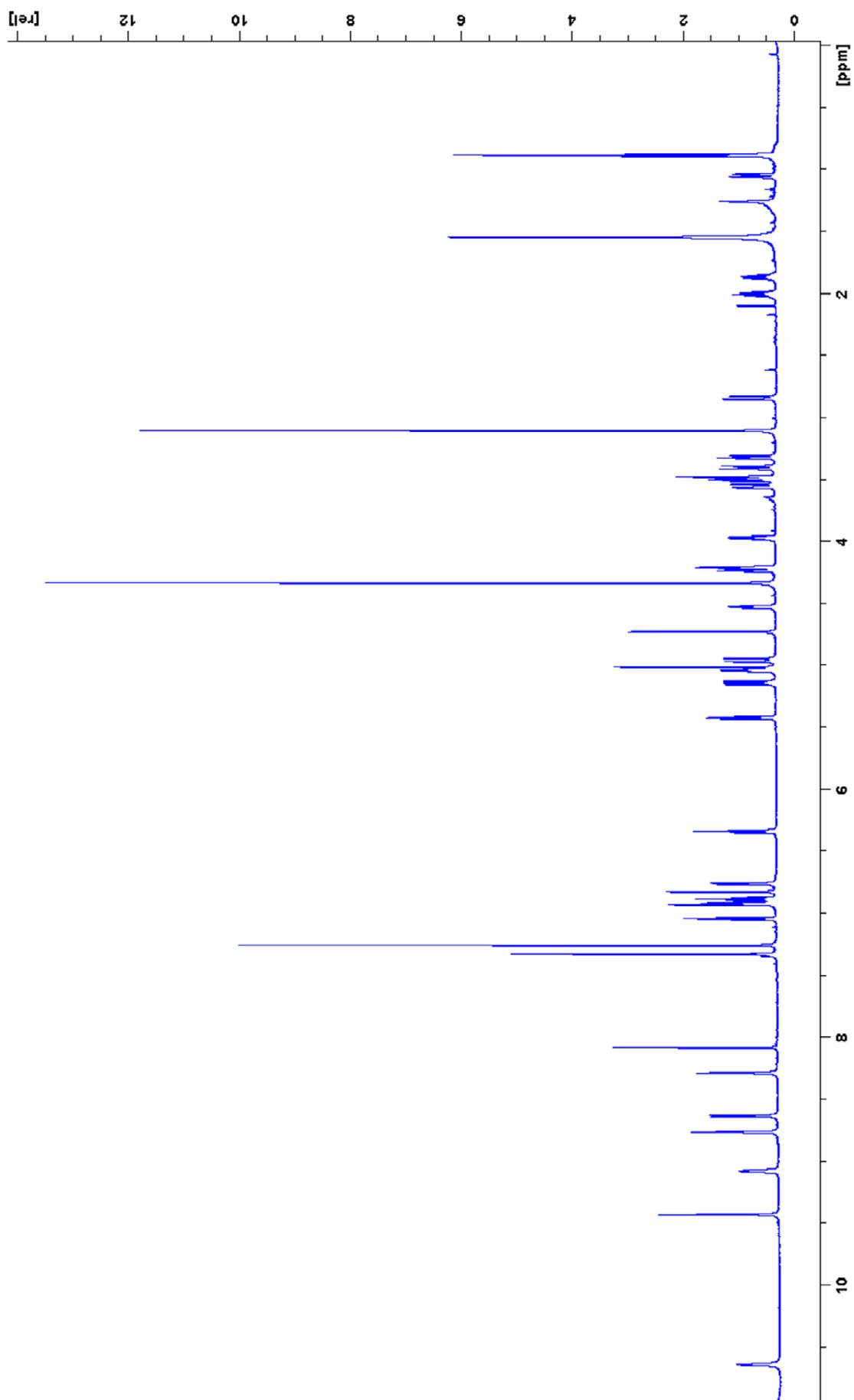
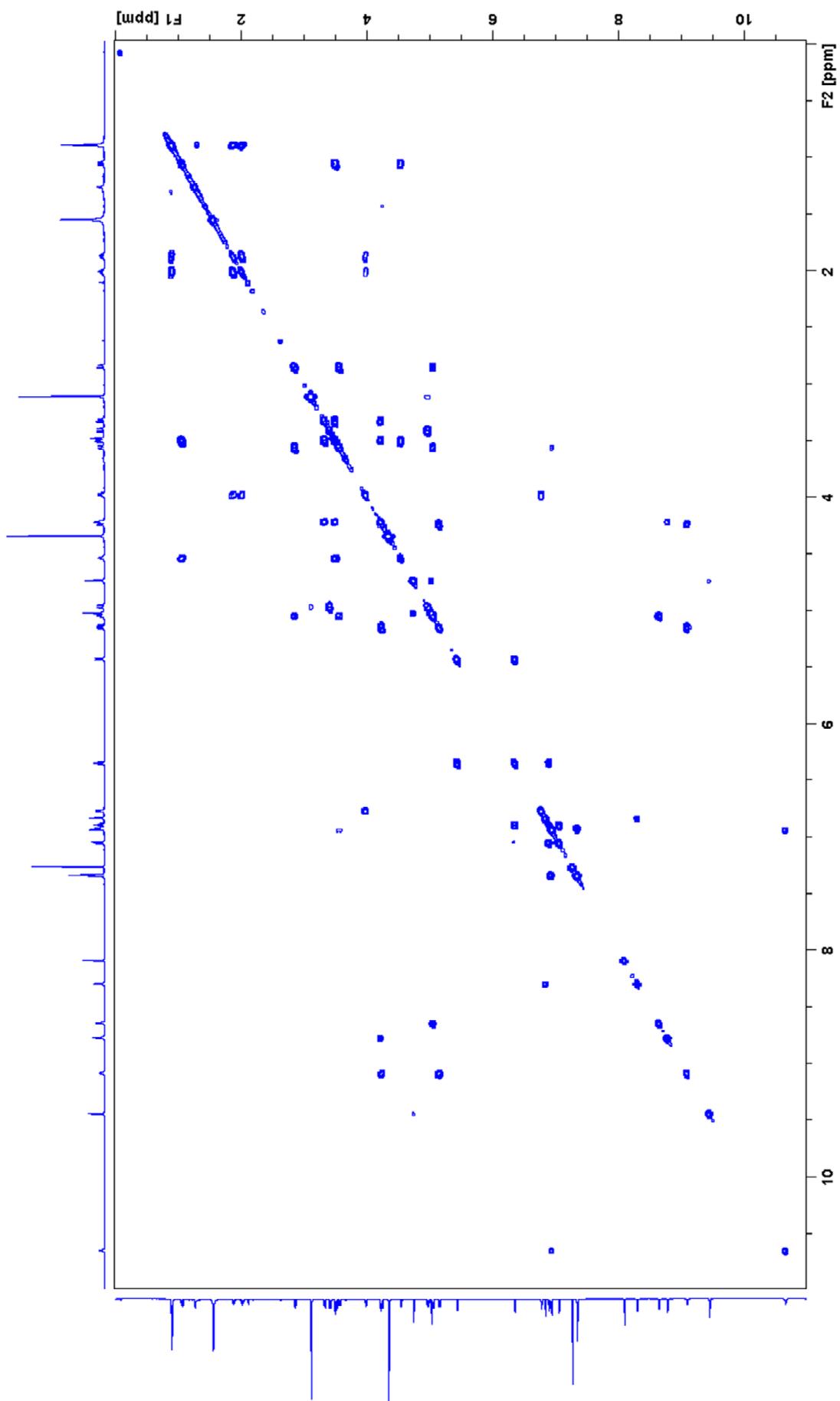


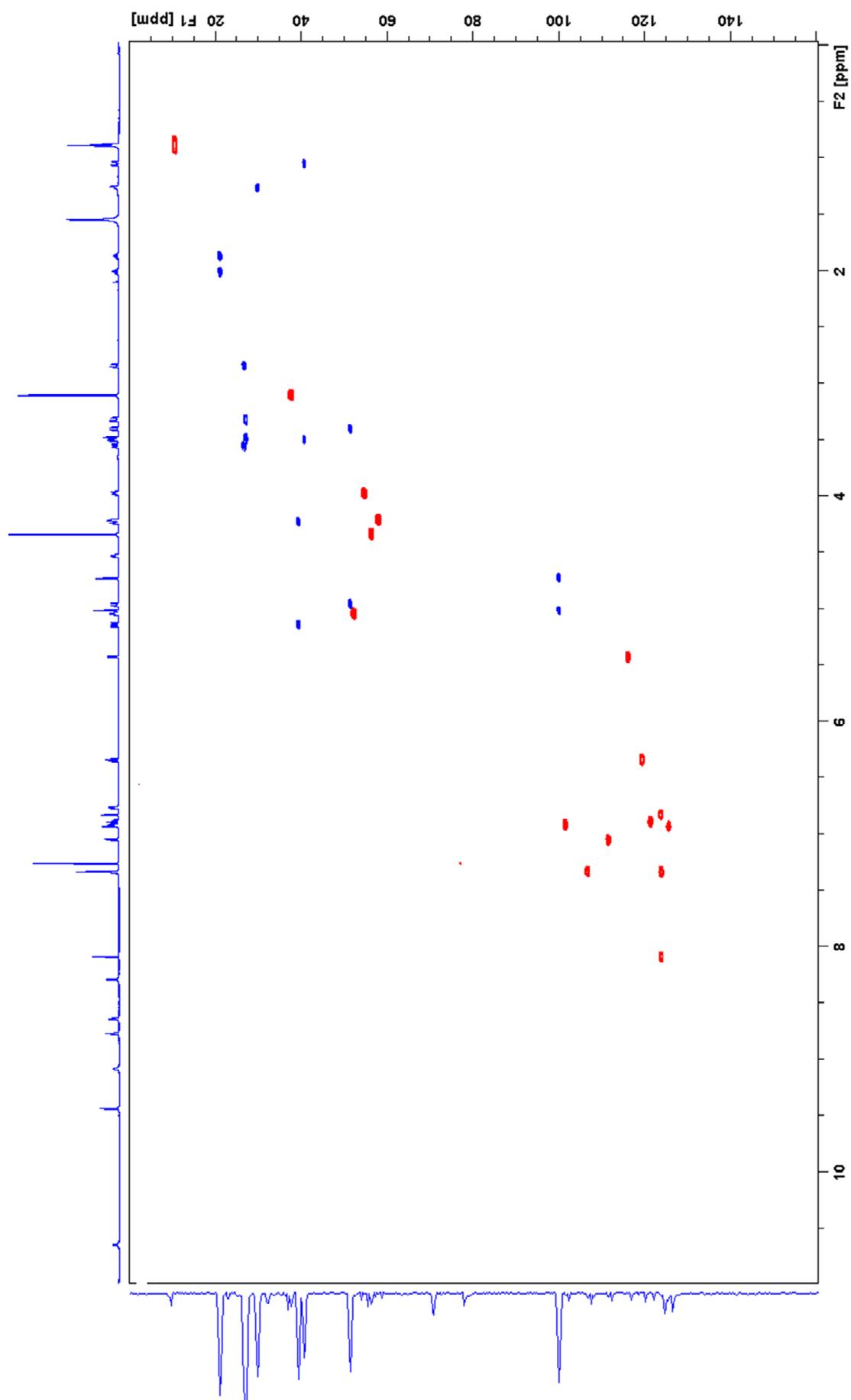
Figure S7. NMR spectra of argyrin I.

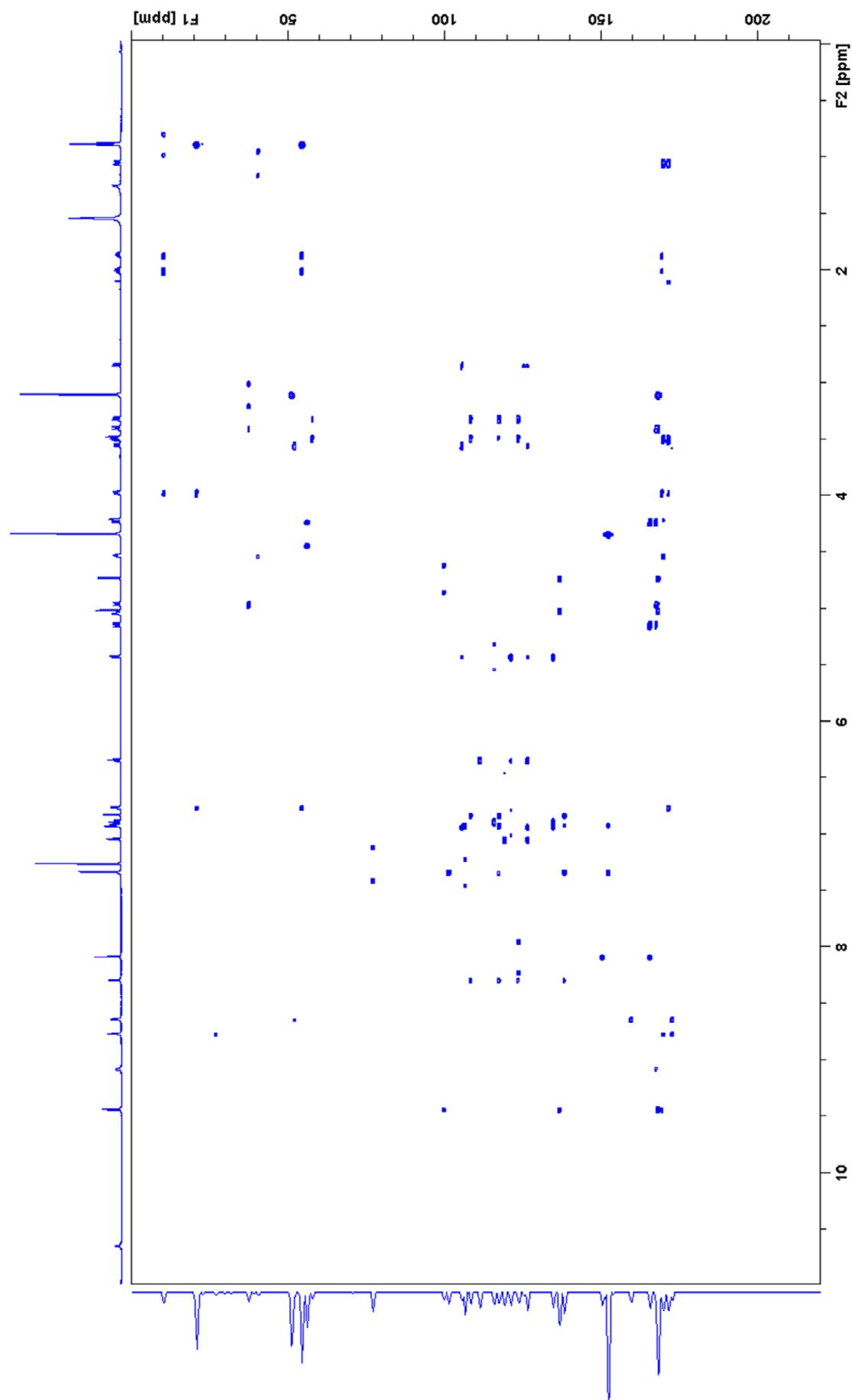
 $^1\text{H-NMR}$  spectrum of argyrin J in  $\text{CDCl}_3$  (700



$^1\text{H}$ - $^1\text{H}$ -COSY spectrum of argyrin J in  $\text{CDCl}_3$  (700)

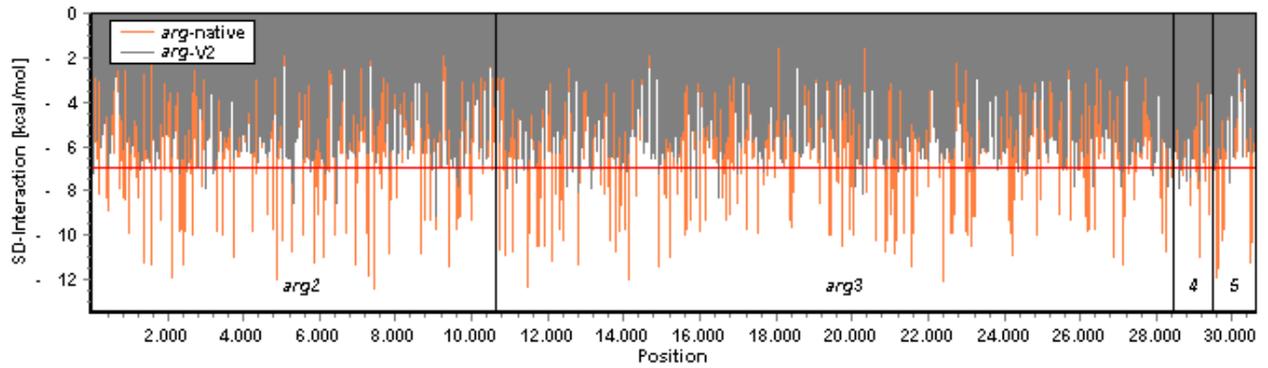
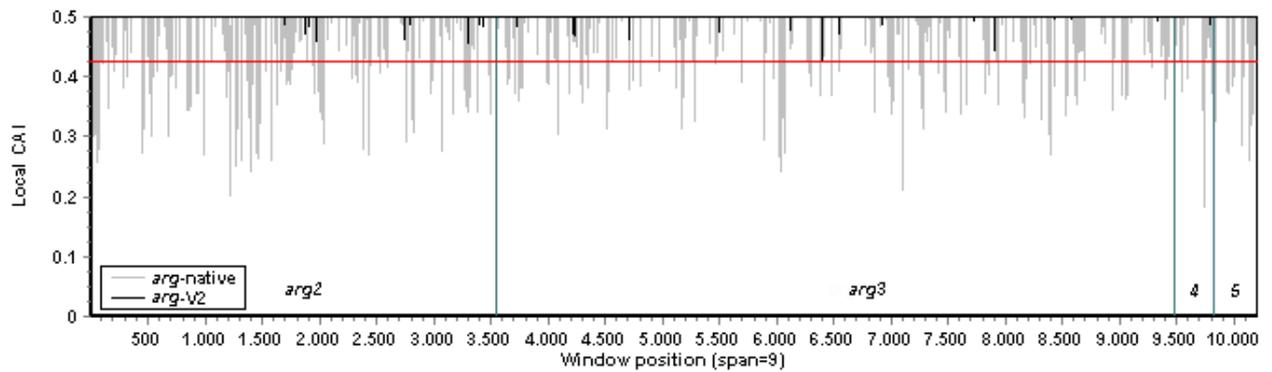
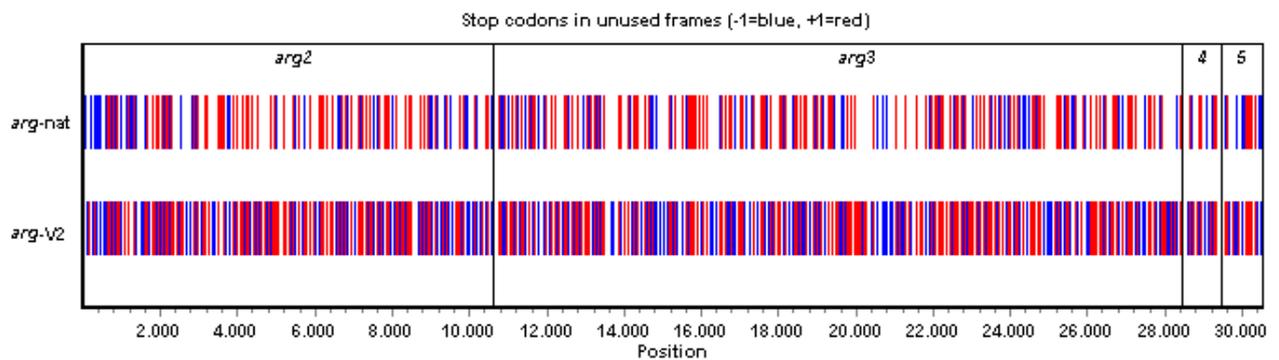
Figure S8 (continued on next page)

HSQC spectrum of argyrin J in  $\text{CDCl}_3$  (700 MHz)

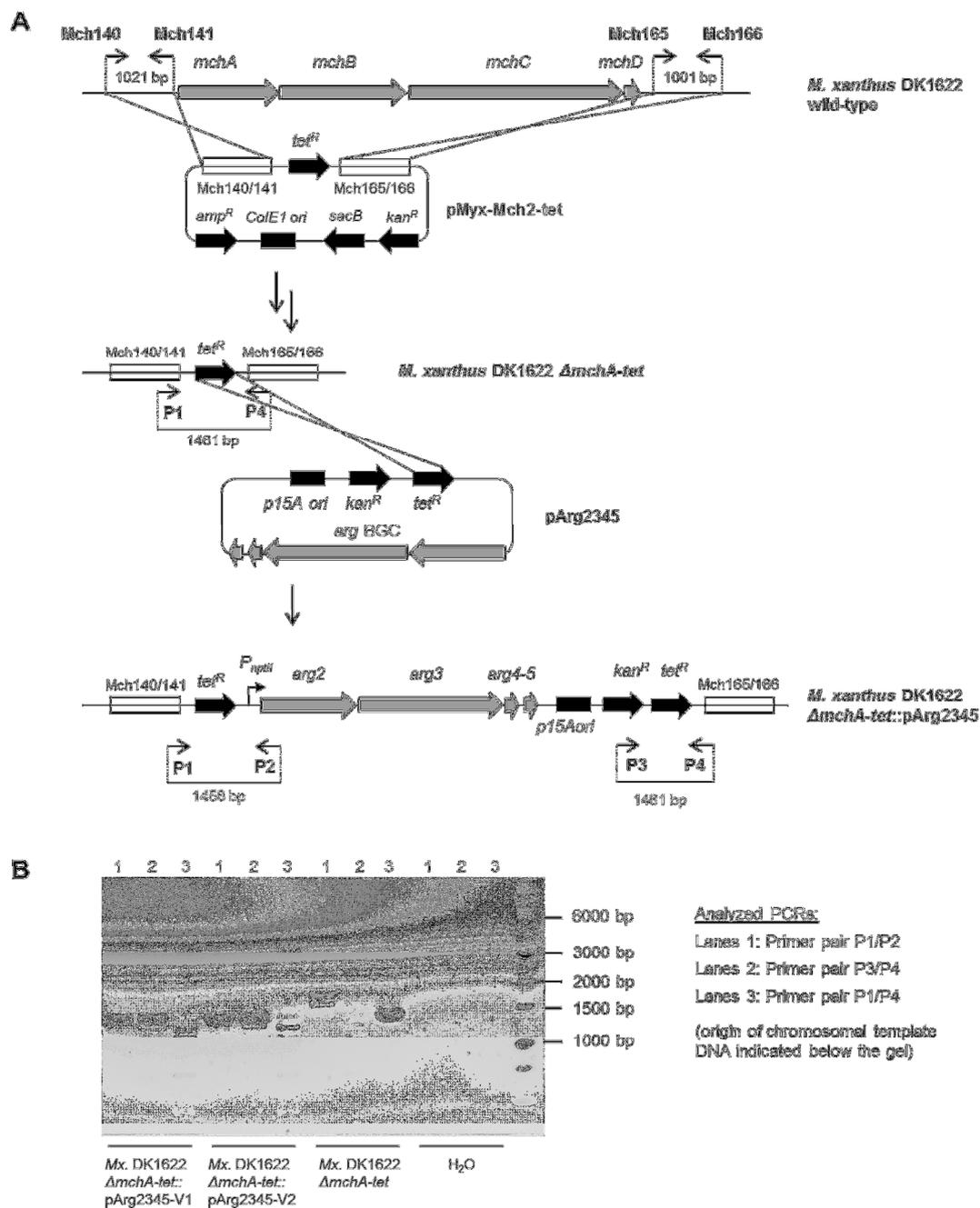


HMBC spectrum of argyrin J in  $\text{CDCl}_3$  (700 MHz)

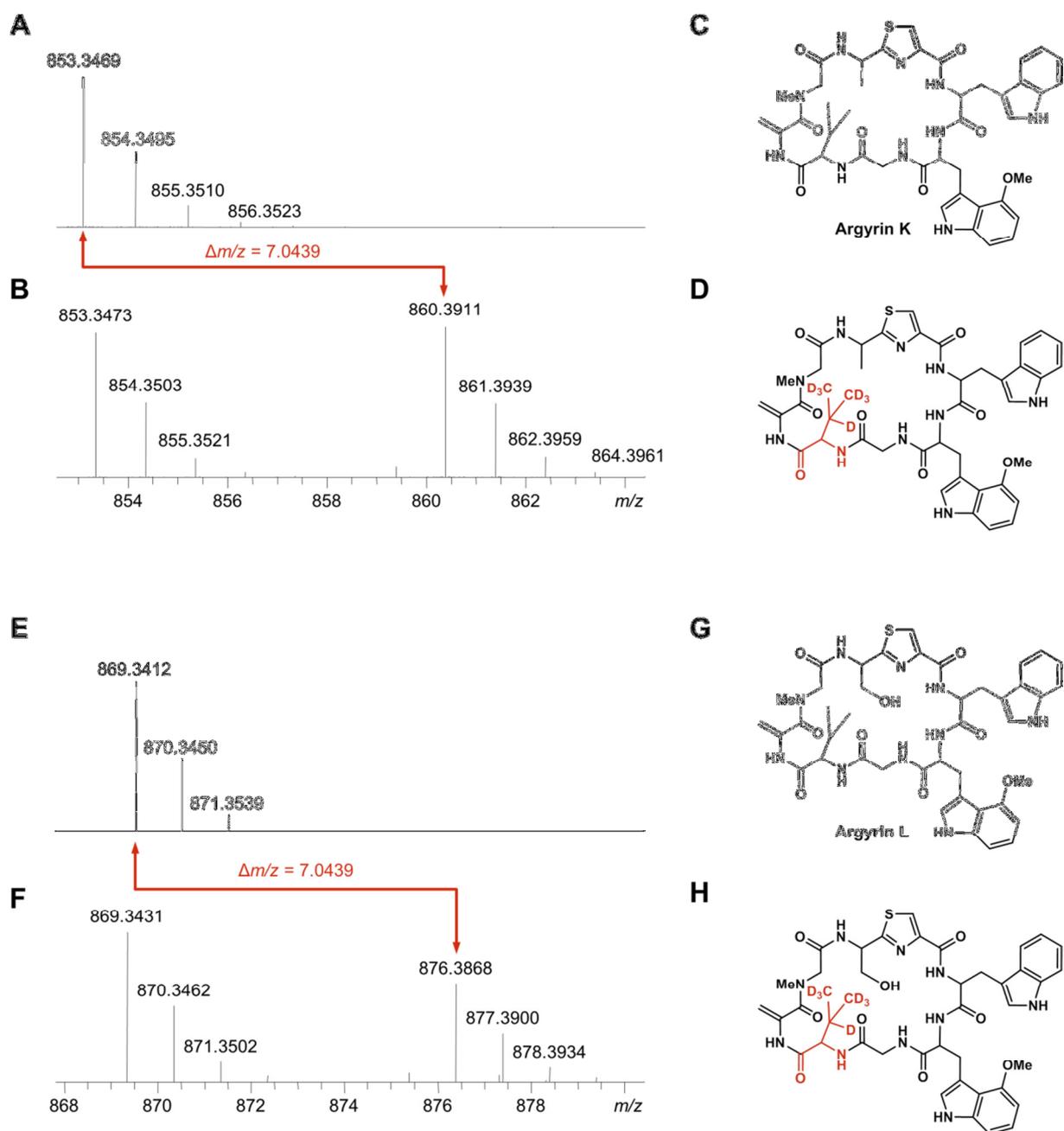
Figure S8. NMR spectra from argyrin J.

**A****B****C**

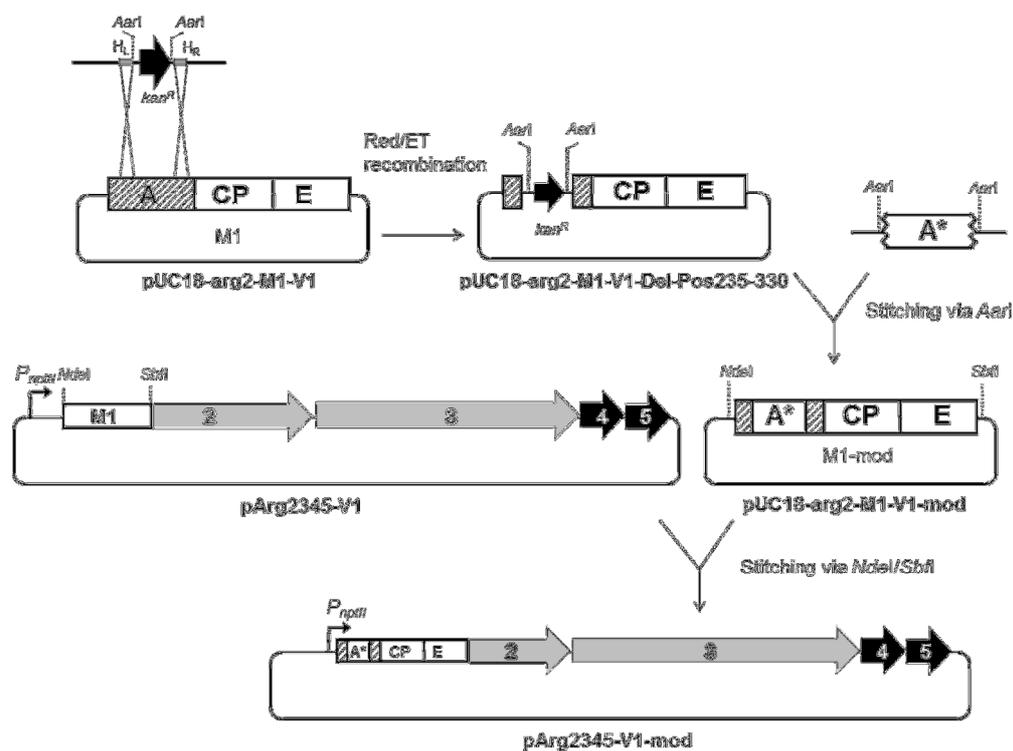
**Figure S9.** Effect of the *arg-V2* sequence modulation on selected features in comparison to *arg-native*. Untranslated regions between the protein coding sequences of *arg2-arg4* were excluded from the graphs. **(A)** Interaction of the 16s rRNA 3'-end (anti-Shine-Dalgarno (ASD) sequence) and internal Shine-Dalgarno (SD) within the *arg* CDSs. **(B)** Occurrence of rare codon clusters within the *arg* CDSs derived from 9 codon windows (including Trp and Met for CAI calculation). The modified BGC *arg-V2* shows no downward peaks below CAI=0.425, whereas the native BGC has values lower than CAI=0.2. **(C)** Occurrence patterns of hidden stop codons in neighbouring frames of *arg* CDSs.



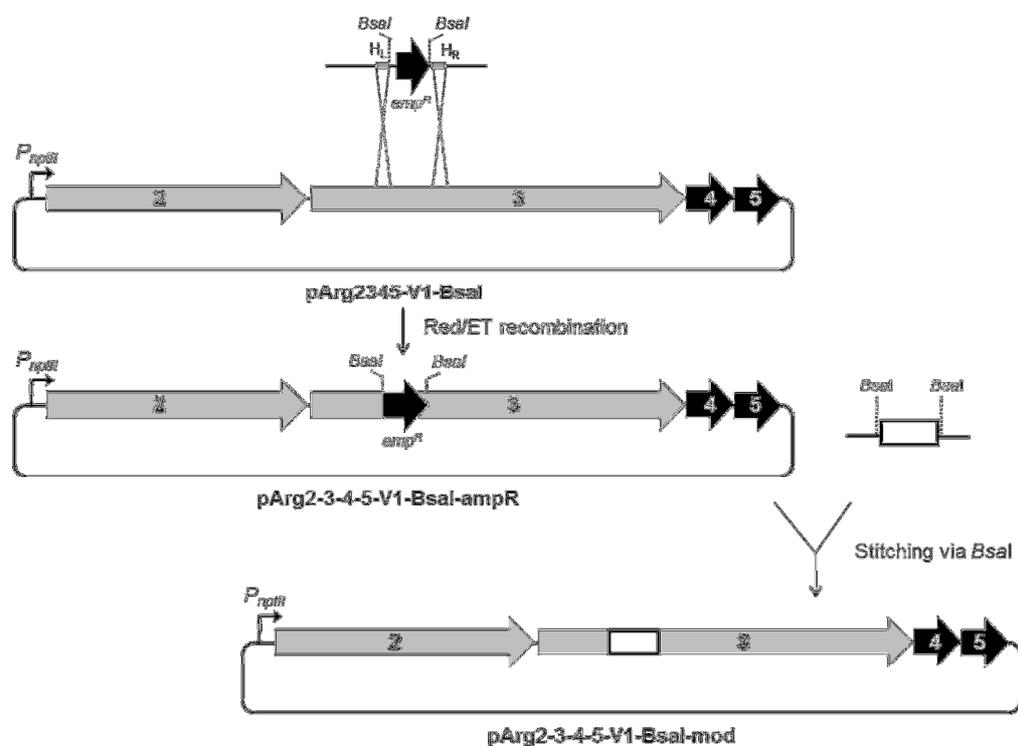
**Figure S10.** Construction and verification of expression strains. (A) After replacement of the *mch<sub>A</sub>* BGC region in *M. xanthus* DK1622 with the tetracycline resistance gene (*tet<sup>R</sup>*), *arg* BGC expression constructs were integrated by homologous recombination via *tet<sup>R</sup>*. Genetic elements are not illustrated in uniform scale (B) Correct chromosomal integration was verified by PCR using the oligonucleotides P1-P4 (Table S1) as shown for *M. xanthus* DK1622  $\Delta$ mchA-tet::pArg2345-V1/V2 in comparison to *M. xanthus* DK1622  $\Delta$ mchA-tet and water as negative control.



**Figure S11.** Analysis of the incorporation *L*-valine-D8 during argyirin biosynthesis. MS-spectra of argyirin K (C) and argyirin L (G) after HPLC-MS analysis of culture extracts of *Cystobacter* sp. SBCb004. The monoisotopic signal of the argyirin K  $[M+H]^+$  ion is at 853.3469  $m/z$  (calc. 853.3450  $m/z$ ,  $C_{42}H_{49}N_{10}O_8S$ ) and of the argyirin L  $[M+H]^+$  ion at 869.3412  $m/z$  (calc. 869.3399  $m/z$ ,  $C_{42}H_{49}N_{10}O_9S$ ). (A)/(E) Control sample to which no labelled precursor was fed, (B)/(F) sample from feeding experiment with *L*-valine-D8, (D)/(H) expected labelling pattern of argyirin K and L after incorporation of *L*-valine-D8 (D in  $\alpha$ -position lost because of frequent H-D exchange).



**Figure S12.** Strategy to exchange the subdomain part of A1. Using this approach 14 expression constructs with engineered A1 domains (pArg2345-V1-mod[1-14]) were generated.



**Figure S13.** Modification strategy of the pArg2345-V1-BsaI expression construct. General modification strategy of the pArg2345-V1-BsaI expression construct (e.g. Arg3 subunit), employing *BsaI* type IIS restriction sites.

## 4.7.10 Tables

**Table S1.** Genes, proteins, proposed function in the *arg* BGC from *Cystobacter sp.* SBCb004.

| Gene name   | Gene length | Protein name | Protein length | Putative function                                 |
|-------------|-------------|--------------|----------------|---|
| <i>arg1</i> | 2013 nt     | Arg1         | 670 aa         | radical SAM-dependent methyltransferase           |
| <i>arg2</i> | 10617 nt    | Arg2         | 3538 aa        | nonribosomal peptide synthetase (NRPS), subunit 1 |
| <i>arg3</i> | 17841 nt    | Arg3         | 5946 aa        | nonribosomal peptide synthetase (NRPS), subunit 2 |
| <i>arg4</i> | 1011 nt     | Arg4         | 336 aa         | <i>O</i> -methyltransferase                       |
| <i>arg5</i> | 1152 nt     | Arg5         | 383 aa         | tryptophan 2,3-dioxygenase                        |

nt = nucleotides, aa = amino acids

**Table S2.** Substrate specificity analysis of A domains from the argyrin NRPS from *Cystobacter sp.* SBCb004.

|           | Residue according to GrsA Phe numbering <sup>1</sup> |     |     |     |     |     |     |     |     |     | Specificity predictions <sup>2</sup> |     |     |     |
|-----------|--|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------------------------------------|-----|-----|-----|
|           | 235  | 236 | 239 | 278 | 299 | 301 | 322 | 330 | 331 | 517 |                                      |     |     |     |
| <b>A1</b> | D  | V   | L   | F   | L   | G   | V   | V   | A   | K   | (a)                                  | Ile | Ala | -   |
| <b>A2</b> | D  | V   | W   | H   | F   | S   | L   | I   | D   | K   | Ser                                  | Ser | Ser | Ser |
| <b>A3</b> | D  | I   | L   | Q   | L   | G   | M   | V   | W   | K   | Gly                                  | Gly | Gly | Gly |
| <b>A4</b> | D  | L   | W   | N   | M   | G   | E   | V   | W   | K   | (b)                                  | Asn | Ala | -   |
| <b>A5</b> | D  | L   | Y   | N   | M   | S   | L   | I   | W   | K   | Cys                                  | Cys | Cys | Cys |
| <b>A6</b> | D  | G   | W   | A   | L   | A   | A   | V   | T   | K   | (c)                                  | Trp | Tyr | -   |
| <b>A7</b> | D  | G   | W   | A   | T   | A   | V   | V   | T   | K   | Trp                                  | Trp | Ala | Trp |
| <b>A8</b> | D  | I   | L   | Q   | L   | G   | V   | I   | W   | K   | Gly                                  | Gly | Gly | Gly |

<sup>1</sup> Ten specificity-conferring residues defined by Stachelhaus et al.<sup>15</sup><sup>2</sup> Predicted substrate specificities were retrieved from reports of the applied antiSMASH 3.0 gene cluster analysis<sup>18</sup>. They include substrate predictions based on the NRPSpredictor2 method<sup>19</sup> (1st column) / Stachelhaus code<sup>15</sup> (2nd column) / method of Minowa *et al.*<sup>20</sup> (3rd column) / consensus of the three approaches (4th column; “-“ indicates no consensus). In some cases, no single substrates but (only) classes/clusters of several amino acids were predicted by NRPSpredictor2 as indicated by (a-c): (a) Hydrophobic aliphatic (Ala, Gly, Val, Leu, Ile, Abu, Iva, Ser, Thr, Hpg, Dhpg, Cys, Pro, Pip), (b) Apolar, aliphatic (Gly, Ala, Val, Leu, Ile, Abu, Iva), (c) Apolar aromatic ring (Phe, Trp).

**Table S3.** Oligonucleotides used in this study.

| Name                   | Sequence (5'→3')<br>(restriction sites in bold, homology arms underlined)                 | R-sites<br>(bold) |
|------------------------|---|-------------------|
| Arg1                   | <b>CTCGATAT</b> CCCAGCGCAAGAGCTATCG   | <i>EcoRV</i>      |
| Arg2                   | <b>CTCGGATCC</b> GGTCGGGAACCATGTACC   | <i>BamHI</i>      |
| Arg10                  | GAGCGTGCGGAAGAGCTC  |                   |
| Arg11                  | CCTCGGTGCTGTTCCAGG  |                   |
| Arg42                  | AACCCGAGGACCGCTTTCTCGCCCTGAGCTCGGTCTCCTTCG<br>ACGGCAGGTGTGGACAGCAAGCGAACCG                |                   |
| Arg43                  | TGCACACTGTCAACCGCAAGAGCTCTGGCGAAGTGACCGCG<br>ACGGCAGGTGTGAGAAGAAGTCTCGTCAAG               |                   |
| Arg44                  | TGACAT <b>CATATGC</b> CTAGGGACGAAAGGGCCTCGTGATAC  |                   |
| Arg45                  | ATGT <b>CAGCATGC</b> CCTAGGTTACCAATGCTTAATCAGTGAG   |                   |
| Arg47                  | <b>CACCAAGCTTT</b> GAATGCGCAAACCAACC  |                   |
| Arg48                  | CTAGGT <b>CAGGGCATATG</b> ATC   |                   |
| Arg51                  | CCCGCCTACTTCATGTACAC  |                   |
| Arg2V1F                | GGCTCTTCAATCTGGACGAG  |                   |
| Arg2V1R                | GATCTTCAGGTCTGGTTTGC  |                   |
| Arg3V1F (= MutArg3V1F) | CTCATGGCTTCTGGCGTCTG  |                   |
| Arg3V1R (= MutArg3V1R) | GGGACACCTTGTGCTTGTCC  |                   |
| arg4V1F                | GTTGATGCTCGTCCCTGGGTG   |                   |
| arg4V1R                | CAGGTCGCATCGTCCCAATC  |                   |
| arg5V1F                | CCGGCCTATATTCATGAGCG  |                   |
| arg5V1R                | CTTCAGATGCTCGAAGCCATC   |                   |
| Arg2V2F (= arg2F)      | GGTCCGATGTATCGCACC  |                   |
| Arg2V2R (= arg2R)      | GTGTAGCGCTCACAGTGC  |                   |
| Arg3V2F (= arg3F)      | GATGCCCAGCACTTCAACC   |                   |
| Arg3V2R (= arg3R)      | GCCTGACCCTCACCTTCC  |                   |
| Arg4V2F (= arg4F)      | CACCTCATCAATAGCACG  |                   |
| Arg4V2R (= arg4R)      | GCCTTACGCAGGAGTTCC  |                   |
| Arg5V2F (= arg5F)      | TCTTCGTGGCTGACGTGC  |                   |
| Arg5V2R (= arg5R)      | GCATGTTGCGGGAACAGC  |                   |
| P1 (= InMchP1)         | CGAGCAATCCGCTATTGGC   |                   |
| P2 (= InMchP2)         | CAGCTGGCAATCCGGTTCCG  |                   |
| P3 (= InMchP3)         | ACGGGACGGGATATCTGACC  |                   |
| P4 (= InMchP4)         | CTGTGTCTTCTGCGACGC  |                   |
| pSUP_B                 | CGCTCATCGTCATCCTCG  |                   |
| pSUP_E                 | GCTCATGAGCCCCGAAGTG   |                   |
| dpo-A4-ampR-F          | GGCCAGTCCGCGGGGAACCTGGACCTGGATTACGCCTCCGGG<br>GAT <b>GAGACCG</b> ACGAAAGGGCCTCGTGATAC     | <i>BsaI</i>       |
| dpo-A4-ampR-R          | AAGGGAGACGTGGACGAACCTTCTCCTTGTTTCAGCTCTGGCCG<br>CTTC <b>GAGACCT</b> TACCAATGCTTAATCAGTGAG | <i>BsaI</i>       |
| dpo-BsaI-A4-R          | GTCTTGTAGAGGCGGACGTTG   |                   |
| Ampint3                | GGACCACTTCTGCGCTCG  |                   |
| Ampint5                | GTCTCATGAGCGGATAC   |                   |
| Mch140                 | CAT <b>GAGCTCG</b> GAAACGGACGCGCCAATCC  | <i>SacI</i>       |
| Mch141                 | CATT <b>CTAGACT</b> ACGCCAATAGCGGATTGC  | <i>XbaI</i>       |
| Mch165                 | CATT <b>CTAGACG</b> ATGTCGCGCAGGTACG  | <i>XbaI</i>       |
| Mch166                 | CAT <b>GGATCCC</b> ATCCAGCAGAAGGCCATC   | <i>BamI</i>       |
| TetR-for               | CACGGT <b>TCTAGAC</b> GTAATTCTCATGTTTGACAGC   | <i>XbaI</i>       |
| TetR-rev               | CACGGT <b>TCTAGACA</b> AGGGTTGGTTTGCGCA   | <i>XbaI</i>       |

**Table S4.** Gene synthesis constructs obtained from ATG:biosynthetics GmbH.

| <b>Construct name</b>   | <b>Fragment size</b> | <b>Description</b>                                |
|-------------------------|----------------------|---|
| pGH-arg2-M1-V1          | 3129 bp              | <i>NdeI-AvrII-arg2-M1-V1-SbfI-BglIII</i>          |
| pGH-arg2-M2-V1          | 3374 bp              | <i>NdeI-SbfI-arg2-M2-V1-MluI-BglIII</i>           |
| pGH-arg2-M3-V1          | 4317 bp              | <i>NdeI-MluI-arg2-M3-V1-SphI-BglIII</i>           |
| pGH-arg3-M4-V1          | 3238 bp              | <i>NdeI-SphI-arg3-M4-V1-KpnI-BglIII</i>           |
| pGH-arg3-M5-V1          | 4446 bp              | <i>NdeI-KpnI-arg3-M5-V1-AscI-BglIII</i>           |
| pGH-arg3-M6-V1          | 3168 bp              | <i>NdeI-AscI-arg3-M6-V1-NotI-BglIII</i>           |
| pGH-arg3-M7-V1          | 3172 bp              | <i>NdeI-NotI-arg3-M7-V1-MfeI-BglIII</i>           |
| pGH-arg3-M8-V1          | 4037 bp              | <i>NdeI-MfeI-arg3-M8-V1-NsiI-NcoI</i>             |
| pGH-arg4-V1             | 1061 bp              | <i>NdeI-NsiI-arg4-V1-NcoI-BglIII</i>              |
| pGH-arg5-V1             | 1236 bp              | <i>NdeI-NcoI-arg5-V1-BglIII</i>                   |
| pGH-arg2-M1-V2          | 3129 bp              | <i>NdeI-AvrII-arg2-M1-V2-SbfI-BglIII</i>          |
| pGH-arg2-M2-V2          | 3374 bp              | <i>NdeI-SbfI-arg2-M2-V2-MluI-BglIII</i>           |
| pGH-arg2-M3-V2          | 4317 bp              | <i>NdeI-MluI-arg2-M3-V2-SphI-BglIII</i>           |
| pGH-arg3-M4-V2          | 3238 bp              | <i>NdeI-SphI-arg3-M4-V2-KpnI-BglIII</i>           |
| pGH-arg3-M5-V2          | 4446 bp              | <i>NdeI-KpnI-arg3-M5-V2-AscI-BglIII</i>           |
| pGH-arg3-M6-V2          | 3168 bp              | <i>NdeI-AscI-arg3-M6-V2-NotI-BglIII</i>           |
| pGH-arg3-M7-V2          | 3172 bp              | <i>NdeI-NotI-arg3-M7-V2-MfeI-BglIII</i>           |
| pGH-arg3-M8-V2          | 4037 bp              | <i>NdeI-MfeI-arg3-M8-V2-NsiI-NcoI</i>             |
| pGH-arg4-V2             | 1061 bp              | <i>NdeI-NsiI-arg4-V2-NcoI-BglIII</i>              |
| pGH-arg5-V2             | 1236 bp              | <i>NdeI-NcoI-arg5-V2-BglIII</i>                   |
| pGH-15A-kanR            | 2088 bp              | <i>PmeI-AflIII-p15Aori-kanR-EcoRV-PacI</i>        |
| pGH-tetR-PnptII-MCS-TD1 | 2002 bp              | <i>EcoRV-PacI-tetR-PnptII-MCS-TD1-PmeI-AflIII</i> |
| pGH-A1-mod1             | 324 bp               | <i>AarI-A1-mod1-AarI</i>                          |
| pGH-A1-mod2             | 324 bp               | <i>AarI-A1-mod2-AarI</i>                          |
| pGH-A1-mod3             | 324 bp               | <i>AarI-A1-mod3-AarI</i>                          |
| pGH-A1-mod4             | 324 bp               | <i>AarI-A1-mod4-AarI</i>                          |
| pGH-A1-mod5             | 324 bp               | <i>AarI-A1-mod5-AarI</i>                          |
| pGH-A1-mod6             | 324 bp               | <i>AarI-A1-mod6-AarI</i>                          |
| pGH-A1-mod7             | 324 bp               | <i>AarI-A1-mod7-AarI</i>                          |
| pGH-A1-mod8             | 324 bp               | <i>AarI-A1-mod8-AarI</i>                          |
| pGH-A1-mod9             | 324 bp               | <i>AarI-A1-mod9-AarI</i>                          |
| pGH-A1-mod10            | 324 bp               | <i>AarI-A1-mod10-AarI</i>                         |
| pGH-A1-mod11            | 324 bp               | <i>AarI-A1-mod11-AarI</i>                         |
| pGH-A1-mod12            | 324 bp               | <i>AarI-A1-mod12-AarI</i>                         |
| pGH-A1-mod13            | 324 bp               | <i>AarI-A1-mod13-AarI</i>                         |
| pGH-A1-mod14            | 324 bp               | <i>AarI-A1-mod14-AarI</i>                         |
| pGH-arg2-M1-V1-BsaI     | 3374 bp              | <i>NdeI-AvrII-arg2-M1-V1-BsaI-SbfI-BglIII</i>     |
| pGH-arg2-M2-V1-BsaI     | 4317 bp              | <i>NdeI-SbfI-arg2-M2-V1-BsaI-MluI-BglIII</i>      |
| pGH-arg2-M3-V1-BsaI     | 3226 bp              | <i>NdeI-MluI-arg2-M3-V1-BsaI-SphI-BglIII</i>      |
| pGH-arg3-M4-V1-BsaI     | 4446 bp              | <i>NdeI-SphI-arg3-M4-V1-BsaI-KpnI-BglIII</i>      |
| pGH-arg3-M5-V1-BsaI     | 3168 bp              | <i>NdeI-KpnI-arg3-M5-V1-BsaI-AscI-BglIII</i>      |
| pGH-arg3-M6-V1-BsaI     | 3172 bp              | <i>NdeI-AscI-arg3-M6-V1-BsaI-NotI-BglIII</i>      |
| pGH-arg3-M7-V1-BsaI     | 4037 bp              | <i>NdeI-NotI-arg3-M7-V1-BsaI-MfeI-BglIII</i>      |
| pGH-arg3-M8-V1-BsaI     | 1061 bp              | <i>NdeI-MfeI-arg3-M8-V1-BsaI-NsiI-NcoI</i>        |
| pGH-arg4-V1-BsaI        | 1236 bp              | <i>NdeI-NsiI-arg4-V1-BsaI-NcoI-BglIII</i>         |
| pGH-arg5-V1-BsaI        | 3374 bp              | <i>NdeI-NcoI-arg5-V1-BsaI-BglIII</i>              |

**Table S5.** Plasmids and expression constructs generated in this study.

| Construct name                            | Construction   | Description/Characteristics  |
|---|--|--|
| pSynBio3<br>(p15A-kan-tet-PnptII-MCS-TD1) | 1986 bp <i>PacI</i> -tetR-PnptII-MCS-TD1- <i>PmeI</i> - <i>Afl</i> III fragment from pGH-tetR-PnptII-MCS-TD1 ligated with 2070 bp <i>Afl</i> III-p15Aori-kanR- <i>EcoRV</i> - <i>PacI</i> fragment from pGH-15A-kanR | Expression vector backbone, p15A ori, <i>kan</i> <sup>R</sup> , <i>tet</i> <sup>R</sup>  |
| pGH-arg2-M1-M2-V1                         | 3359 bp <i>Sbf</i> I-arg2-M2-V1- <i>Bgl</i> III fragment derived from pGH-arg2-M2-V1 ligated into pGH-arg2-M1-V1 hydrolyzed with the same enzymes  | synthetic <i>arg2</i> -V1 fragment (encoding modules M1-M2 of NRPS Arg2), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg2-V1                               | 4303 bp <i>Mlu</i> I-arg2-M3-V1- <i>Bgl</i> III fragment derived from pGH-arg2-M3-V1 ligated into pGH-arg2-M1-M2-V1 hydrolyzed with the same enzymes   | complete synthetic <i>arg2</i> -V1 gene (encoding modules M1-M3 of NRPS Arg2), pUC ori, <i>amp</i> <sup>R</sup>  |
| pGH-arg3-M4-M5-V1                         | 4432 bp <i>Kpn</i> I-arg3-M5-V1- <i>Bgl</i> III fragment from pGH-arg3-M5-V1 ligated into pGH-arg3-M4-V1 hydrolyzed with the same enzymes  | synthetic <i>arg3</i> -V1 fragment (encoding modules M4-M5 of NRPS Arg3), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg3-M4-M5-M6-V1                      | 3153 bp <i>Asc</i> I-arg3-M6-V1- <i>Bgl</i> III fragment from pGH-arg3-M6-V1 ligated into pGH-arg3-M4-M5-V1 hydrolyzed with the same enzymes   | synthetic <i>arg3</i> -V1 fragment (encoding modules M4-M6 of NRPS Arg3), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg3-M4-M5-M6-M7-V1                   | 3157 bp <i>Not</i> I-arg3-M7-V1- <i>Bgl</i> III fragment from pGH-arg3-M7-V1 ligated into pGH-arg3-M4-M5-M6-V1 hydrolyzed with the same enzymes  | synthetic <i>arg3</i> -V1 fragment (encoding modules M4-M7 of NRPS Arg3), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg3-V1                               | 4023 bp <i>Mfe</i> I-arg3-M8-V1- <i>Bgl</i> III fragment from pGH-arg3-M8-V1 ligated into pGH-arg3-M4-M5-M6-M7-V1 hydrolyzed with the same enzymes   | complete synthetic <i>arg3</i> -V1 gene (encoding modules M4-M8 of NRPS Arg3), pUC ori, <i>amp</i> <sup>R</sup>  |
| pArg2-V1                                  | 10755 bp <i>Nde</i> I-arg2-V1- <i>Bgl</i> III fragment from pGH-arg2-V1 ligated into pSynBio3 hydrolyzed with the same enzymes   | expression vector with <i>P<sub>npIII</sub></i> - <i>arg2</i> -V1- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan</i> <sup>R</sup> , <i>tet</i> <sup>R</sup>   |
| pArg23-V1                                 | 17923 bp <i>Sph</i> I-arg3-V1- <i>Bgl</i> III fragment from pGH-arg3-V1 ligated into pArg2-V1 hydrolyzed with the same enzymes   | expression vector with <i>P<sub>npIII</sub></i> - <i>arg2</i> - <i>arg3</i> -V1- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan</i> <sup>R</sup> , <i>tet</i> <sup>R</sup>                             |
| pArg235-V1                                | 1222 bp <i>Nco</i> I-arg5-V1- <i>Bgl</i> III fragment from pGH-arg5-V1 ligated into pArg23-V1 hydrolyzed with the same enzymes   | expression vector with <i>P<sub>npIII</sub></i> - <i>arg2</i> - <i>arg3</i> - <i>arg5</i> -V1- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan</i> <sup>R</sup> , <i>tet</i> <sup>R</sup>               |
| pArg2345-V1                               | 1035 bp <i>Nsi</i> I-arg4-V1- <i>Nco</i> I fragment from pGH-arg4-V1 ligated into pArg235-V1 hydrolyzed with the same enzymes  | expression vector with <i>P<sub>npIII</sub></i> - <i>arg2</i> - <i>arg3</i> - <i>arg4</i> - <i>arg5</i> -V1- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan</i> <sup>R</sup> , <i>tet</i> <sup>R</sup> |
| pGH-arg2-LM-M1-V2                         | 3359 bp <i>Sbf</i> I-arg2-M2-V2- <i>Bgl</i> III fragment derived from pGH-arg2-M2-V2 ligated into pGH-arg2-M1-V2 hydrolyzed with the same enzymes  | synthetic <i>arg2</i> -V2 fragment (encoding modules M1-M2 of NRPS Arg2), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg2-V2                               | 4303 bp <i>Mlu</i> I-arg2-M3-V2- <i>Bgl</i> III fragment derived from pGH-arg2-M3-V2 ligated into pGH-arg2-M1-M2-V2 hydrolyzed with the same enzymes   | complete synthetic <i>arg2</i> -V2 gene (encoding modules M1-M3 of NRPS Arg2), pUC ori, <i>amp</i> <sup>R</sup>  |
| pGH-arg3-M4-M5-V2                         | 4432 bp <i>Kpn</i> I-arg3-M5-V2- <i>Bgl</i> III fragment from pGH-arg3-M5-V2 ligated into pGH-arg3-M4-V2 hydrolyzed with the same enzymes  | synthetic <i>arg3</i> -V2 fragment (encoding modules M4-M5 of NRPS Arg3), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg3-M4-M5-M6-V2                      | 3153 bp <i>Asc</i> I-arg3-M6-V2- <i>Bgl</i> III fragment from pGH-arg3-M6-V2 ligated into pGH-arg3-M4-M5-V2 hydrolyzed with the same enzymes   | synthetic <i>arg3</i> -V2 fragment (encoding modules M4-M6 of NRPS Arg3), pUC ori, <i>amp</i> <sup>R</sup>   |
| pGH-arg3-M4-M5-M6-M7-V2                   | 3157 bp <i>Not</i> I-arg3-M7-V2- <i>Bgl</i> III fragment from pGH-arg3-M7-V2 ligated into pGH-   | synthetic <i>arg3</i> -V2 fragment (encoding modules M4-M7 of NRPS Arg3),  |

|                                 |   |   |
|---------------------------------|---|---|
|                                 | arg3-M4-M5-M6-V2 hydrolyzed with the same enzymes   | pUC ori, <i>amp<sup>R</sup></i>   |
| pGH-arg3-V2                     | 4023 bp <i>MfeI</i> -arg3-M8-V2- <i>BglIII</i> fragment from pGH-arg3-M8-V2 ligated into pGH-arg3-M4-M5-M6-M7-V2 hydrolyzed with the same enzymes | complete synthetic <i>arg3-V2</i> gene (encoding modules M4-M8 of NRPS Arg3),<br>pUC ori, <i>amp<sup>R</sup></i>  |
| pArg2-V2                        | 10755 bp <i>NdeI</i> -arg2-V2- <i>BglIII</i> fragment from pGH-arg2-V2 ligated into pSynBio3 hydrolyzed with the same enzymes                     | expression vector with <i>P<sub>npIII</sub></i> -arg2-V2- <i>t<sub>D1</sub></i> transcription unit,<br>p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>                |
| pArg23-V2                       | 17923 bp <i>SphI</i> -arg3-V2- <i>BglIII</i> fragment from pGH-arg3-V2 ligated into pArg2-V2 hydrolyzed with the same enzymes                     | expression vector with <i>P<sub>npIII</sub></i> -arg2-arg3-V2- <i>t<sub>D1</sub></i> transcription unit,<br>p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>           |
| pArg235-V2                      | 1222 bp <i>NcoI</i> -arg5-V2- <i>BglIII</i> fragment from pGH-arg5-V2 ligated into pArg23-V2 hydrolyzed with the same enzymes                     | expression vector with <i>P<sub>npIII</sub></i> -arg2-arg3-arg5-V2- <i>t<sub>D1</sub></i> transcription unit,<br>p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>      |
| pArg2345-V2                     | 1035 bp <i>NsiI</i> -arg4-V2- <i>NcoI</i> fragment from pGH-arg4-V2 ligated into pArg235-V2 hydrolyzed with the same enzymes                      | expression vector with <i>P<sub>npIII</sub></i> -arg2-arg3-arg4-arg5-V2- <i>t<sub>D1</sub></i> transcription unit,<br>p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pUC18-arg2-M1-V1                | 3113 bp <i>NdeI</i> -arg2-M1-VI- <i>SdaI</i> fragment ligated into pUC18 hydrolyzed with the same enzymes   | cloning vector with arg2-M1-V1 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>  |
| pUC18-arg2-M1-V1-Del-Pos235-330 | A1 subdomain region of pUC18-arg2-M1-V1 replaced with 1052 bp <i>kan<sup>R</sup></i> fragment by <i>in vivo</i> recombination                     | cloning vector with arg2-M1-V1-Del-Pos235-330 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod1           | 302 bp <i>AarI</i> -A1-mod1- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod1 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod2           | 302 bp <i>AarI</i> -A1-mod2- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod2 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod3           | 302 bp <i>AarI</i> -A1-mod3- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod3 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod4           | 302 bp <i>AarI</i> -A1-mod4- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod4 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod5           | 302 bp <i>AarI</i> -A1-mod5- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod5 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod6           | 302 bp <i>AarI</i> -A1-mod6- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod6 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod7           | 302 bp <i>AarI</i> -A1-mod7- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod7 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod8           | 302 bp <i>AarI</i> -A1-mod8- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod8 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod9           | 302 bp <i>AarI</i> -A1-mod9- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                    | cloning vector with arg2-M1-V1-mod9 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod10          | 302 bp <i>AarI</i> -A1-mod10- <i>AarI</i> fragment  | cloning vector with arg2-M1-V1-mod10  |

|                        |   |   |
|------------------------|---|---|
| mod10                  | ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme  | synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>   |
| pUC18-arg2-M1-V1-mod11 | 302 bp <i>AarI</i> -A1-mod11- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                 | cloning vector with arg2-M1-V1-mod11 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>  |
| pUC18-arg2-M1-V1-mod12 | 302 bp <i>AarI</i> -A1-mod12- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                 | cloning vector with arg2-M1-V1-mod12 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>  |
| pUC18-arg2-M1-V1-mod13 | 302 bp <i>AarI</i> -A1-mod13- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                 | cloning vector with arg2-M1-V1-mod13 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>  |
| pUC18-arg2-M1-V1-mod14 | 302 bp <i>AarI</i> -A1-mod14- <i>AarI</i> fragment ligated into pUC18-arg2-M1-V1-Del-Pos235-330 hydrolyzed with the same enzyme                 | cloning vector with arg2-M1-V1-mod14 synthetic fragment, pUC18 ori, <i>amp<sup>R</sup></i>  |
| pArg2345-V1-mod1       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod1- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod1 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod1; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod2       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod2- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod2 ligated into pArg2345-V1 hydrolyzed with <i>NdeI/SbfI</i> | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod2; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod3       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod3- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod3 ligated pArg2345-V1 hydrolyzed with the same enzymes      | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod3; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod4       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod4- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod4 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod4; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod5       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod5- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod5 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod5; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod6       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod6- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod6 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod6; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod7       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod7- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod7 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod7; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod8       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod8- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod8 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod8; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod9       | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod9- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod9 ligated into pArg2345-V1 hydrolyzed with the same enzymes | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod9; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg2345-V1-mod10      | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod10- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod10 ligated pArg2345-V1 hydrolyzed with <i>NdeI/SbfI</i>    | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit (* A1 mutation mod10; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pArg2345-V1-mod11      | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod11- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-  | expression vector with <i>P<sub>npIII</sub>-arg2*-arg3-arg4-arg5-V1-t<sub>D1</sub></i> transcription unit   |

|                              |  |   |
|------------------------------|--|---|
|                              | mod11 ligated into pArg2345-V1 hydrolyzed with the same enzymes  | (* A1 mutation mod11; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>   |
| pArg2345-V1-mod12            | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod12- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod12 ligated into pArg2345-V1 hydrolyzed with the same enzymes              | expression vector with <i>P<sub>npII</sub></i> -arg2*-arg3-arg4-arg5-V1- <i>t<sub>D1</sub></i> transcription unit (* A1 mutation mod12; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pArg2345-V1-mod13            | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod13- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod13 ligated into pArg2345-V1 hydrolyzed with the same enzymes              | expression vector with <i>P<sub>npII</sub></i> -arg2*-arg3-arg4-arg5-V1- <i>t<sub>D1</sub></i> transcription unit (* A1 mutation mod13; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pArg2345-V1-mod14            | 3133 bp <i>NdeI</i> -arg2-M1-V1-mod14- <i>SbfI</i> fragment from pUC18-arg2-M1-V1-mod14 ligated into pArg2345-V1 hydrolyzed with the same enzymes              | expression vector with <i>P<sub>npII</sub></i> -arg2*-arg3-arg4-arg5-V1- <i>t<sub>D1</sub></i> transcription unit (* A1 mutation mod14; Table S11), p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pGH-arg2-M1-M2-V1-BsaI       | 3359 bp <i>SbfI</i> -arg2-M2-V1-BsaI- <i>BglIII</i> fragment derived from pGH-arg2-M2-V1 ligated into pGH-arg2-M1-V1 hydrolyzed with the same enzymes          | synthetic arg2-V1-BsaI fragment (encoding modules M1-M2 of NRPS Arg2), pUC ori, <i>amp<sup>R</sup></i>  |
| pGH-arg2-V1-BsaI             | 4303 bp <i>MluI</i> -arg2-M3-V1-BsaI- <i>BglIII</i> fragment derived from pGH-arg2-M3-V1 ligated into pGH-arg2-M1-M2-V1 hydrolyzed with the same enzymes       | complete synthetic arg2-V1-BsaI gene (encoding modules M1-M3 of NRPS Arg2), pUC ori, <i>amp<sup>R</sup></i>   |
| pGH-arg3-M5-M6-V1-BsaI       | 3153 bp <i>AscI</i> -arg3-M6-V1-BsaI- <i>BglIII</i> fragment from pGH-arg3-M6-V1-BsaI ligated into pGH-arg3-M5-V1-BsaI hydrolyzed with the same enzymes        | synthetic arg3-M5-M6-V1-BsaI fragment (encoding modules M5-M6 of NRPS Arg3), pUC ori, <i>amp<sup>R</sup></i>  |
| pGH-arg3-M5-M6-M7-V1-BsaI    | 3157 bp <i>NotI</i> -arg3-M7-V1-BsaI- <i>BglIII</i> fragment from pGH-arg3-M7-V1-BsaI ligated into pGH-arg3-M5-M6-V1-BsaI hydrolyzed with the same enzymes     | synthetic arg3-M5-M6-M7-V1-BsaI fragment (encoding modules M5-M7 of NRPS Arg3), pUC ori, <i>amp<sup>R</sup></i>   |
| pGH-arg3-M5-M6-M7-M8-V1-BsaI | 4023 bp <i>MfeI</i> -arg3-M8-V1-BsaI- <i>BglIII</i> fragment from pGH-arg3-M8-V1-BsaI ligated into pGH-arg3-M5-M6-M7-V1-BsaI hydrolyzed with the same enzymes  | synthetic arg3-M5-M6-M7-M8-V1-BsaI fragment (encoding modules M5-M8 of NRPS Arg3), pUC ori, <i>amp<sup>R</sup></i>  |
| pGH-arg3-V1-BsaI             | 3223 bp <i>NdeI</i> -arg3-M4-V1-BsaI- <i>KpnI</i> fragment from pGH-arg3-M4-V1-BsaI ligated into pGH-arg3-M5-M6-M7-M8-V1-BsaI hydrolyzed with the same enzymes | complete synthetic arg3-V1-BsaI fragment (encoding modules M4-M8 of NRPS Arg3), pUC ori, <i>amp<sup>R</sup></i>   |
| pArg2-V1-BsaI                | 10755 bp <i>NdeI</i> -arg2-V1-BsaI- <i>BglIII</i> fragment from pGH-arg2-V1 ligated into pSynBio3 hydrolyzed with the same enzymes                             | expression vector with <i>P<sub>npII</sub></i> -arg2-V1-BsaI- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>   |
| pArg23-V1-BsaI               | 17923 bp <i>SphI</i> -arg3-V1-BsaI- <i>BglIII</i> fragment from pGH-arg3-V1 ligated into pArg2-V1-BsaI hydrolyzed with the same enzymes                        | expression vector with <i>P<sub>npII</sub></i> -arg2-arg3-V1-BsaI- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>  |
| pArg234-V1-BsaI              | 1047 bp <i>NsiI</i> -arg4-V1-BsaI- <i>BglIII</i> fragment from pGH-arg4-V1-BsaI ligated into pArg23-V1-BsaI hydrolyzed with the same enzymes                   | expression vector with <i>P<sub>npII</sub></i> -arg2-arg3-arg4-V1-BsaI- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>                                   |
| pArg2345-V1-BsaI             | 1222 bp <i>NcoI</i> -arg5-V1-BsaI- <i>BglIII</i> fragment from pGH-arg5-V1-BsaI ligated into pArg234-V1-BsaI hydrolyzed with the same enzymes                  | expression vector with <i>P<sub>npII</sub></i> -arg2-arg3-arg4-arg5-V1-BsaI- <i>t<sub>D1</sub></i> transcription unit, p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>                              |

V1 = version 1, V2 = version 2, *amp<sup>R</sup>* = ampicillin resistance gene, *kan<sup>R</sup>* = kanamycin resistance gene, *tet<sup>R</sup>* = tetracycline resistance gene, *P<sub>npII</sub>* = *P<sub>npII</sub>* promoter, *t<sub>D1</sub>* = *t<sub>D1</sub>* terminator

**Table S6.** NMR spectroscopic data of argyrin I.

| Res                | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|--------------------|-----------------------------|----------------------|------|-----------|
| Thiaz NH           |                             | 8.81                 | d    | 8.6       |
| 1                  | 159.8                       |                      |      |           |
| 2                  | 150.1                       |                      |      |           |
| 3                  | 122.9                       | 8.03                 | s    |           |
| 4                  | 170.8                       |                      |      |           |
| 5                  | 45.5                        | 5.48                 | d,q  | 7.1, 7.6  |
| 6                  | 20.7                        | 1.71                 | d    | 7.2       |
| Trp1 NH            |                             | 8.56                 | d    | 7.3       |
| 1                  | 172.7                       |                      |      |           |
| 2                  | 52.2                        | 5.10                 | m    |           |
| 3                  | 27.0                        | 3.49                 | d,d  | 15.5, 3.3 |
|                    |                             | 2.88                 | d,d  | 15.3, 3.4 |
| 1' NH              |                             | 10.66                | s    |           |
| 2'                 | 125.9                       | 6.99 <sup>d</sup>    | d    | 2.6       |
| 3'                 | 105.5                       |                      |      |           |
| 4'                 | 115.9                       | 6.22                 | d    | 7.8       |
| 5'                 | 119.8                       | 6.52                 | d,d  | 7.4,7.4   |
| 6'                 | 121.5                       | 6.98 <sup>d</sup>    | d,d  | 7.6,7.6   |
| 7'                 | 111.8                       | 7.14                 | d    | 7.9       |
| 8'                 | 134.9                       |                      |      |           |
| 9'                 | 126.5                       |                      |      |           |
| Trp2 NH            |                             | 6.61                 | m    |           |
| 1                  | 169.9                       |                      |      |           |
| 2                  | 56.4                        | 4.28                 | m    |           |
| 3                  | 26.6                        | 3.37                 | d,d  | 15.3, 5.2 |
|                    |                             | 3.31                 | d,d  | 15.2, 4.9 |
| 1' NH              |                             | 8.41                 | s    |           |
| 2'                 | 123.4                       | 7.21                 | d    | 2.1       |
| 3'                 | 109.1                       |                      |      |           |
| 4'                 | 117.9                       | 7.62                 | d    | 7.8       |
| 5'                 | 121.1                       | 7.29                 | d,d  | 7.6,7.6   |
| 6'                 | 123.5                       | 7.39                 | d,d  | 8.0,7.7   |
| 7'                 | 112.1                       | 7.60                 | d    | 8.1       |
| 8'                 | 136.6                       |                      |      |           |
| 9'                 | 127.3                       |                      |      |           |
| Gly NH             |                             | 4.75                 | m    |           |
| 1                  | 171.1                       |                      |      |           |
| 2                  | 41.2                        | 3.62                 | d,d  | 17.1, 7.8 |
|                    |                             | 1.29                 | m    |           |
| Abu NH             |                             | 6.67                 | d    | 6.5       |
| 1                  | 169.2                       |                      |      |           |
| 2                  | 54.5                        | 4.01                 | t,d  | 8.8, 6.3  |
| 3                  | 21.3                        | 1.99                 | m    |           |
|                    |                             | 1.86                 | m    |           |
| 4                  | 10.7                        | 0.89                 | t    | 7.3       |
| DeAla NH           |                             | 9.36                 |      |           |
| 1                  | 168.2                       |                      |      |           |
| 2                  | 136.8                       |                      |      |           |
| 3                  | 99.7                        | 5.03                 | d    | 1.5       |
|                    |                             | 4.74                 | m    |           |
| Sarc               |                             |                      |      |           |
| 1                  | 166.8                       |                      |      |           |
| 2                  | 51.2                        | 4.99                 | d    | 17.0      |
|                    |                             | 3.41                 | d    | 16.9      |
| CH <sub>3</sub> -N | 37.4                        | 3.11                 | s    |           |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

**Table S7.** NMR spectroscopic data of argyrin J.

| Res                | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|--------------------|-----------------------------|----------------------|------|-----------|
| Thiaz NH           |                             | 9.08                 | d,d  | 8.0,4.4   |
| 1                  | 159.6                       |                      |      |           |
| 2                  | 150.3                       |                      |      |           |
| 3                  | 123.9                       | 8.08                 | s    |           |
| 4                  | 165.6                       |                      |      |           |
| 5                  | 39.4                        | 5.14                 | d,d  | 8.2, 14.4 |
| Trp1 NH            |                             | 8.64                 | d    | 7.3       |
| 1                  | 172.9                       |                      |      |           |
| 2                  | 52.3                        | 5.04                 | m    |           |
| 3                  | 26.8                        | 3.55                 | d,d  | 15.2, 3.7 |
|                    |                             | 2.84                 | d,d  | 15.4, 3.1 |
| 1' NH              |                             | 10.6                 | s    |           |
| 2'                 | 125.6                       | 6.99                 | d    | 2.6       |
| 3'                 | 105.7                       |                      |      |           |
| 4'                 | 116.1                       | 5.43                 | d    | 7.8       |
| 5'                 | 119.2                       | 6.34                 | d,d  | 8.0,8.0   |
| 6'                 | 121.4                       | 6.88                 | d,d  | 6.9,6.9   |
| 7'                 | 111.5                       | 7.04                 | d    | 8.1       |
| 8'                 | 134.8                       |                      |      |           |
| 9'                 | 126.5                       |                      |      |           |
| Trp2 NH            |                             | 8.77                 | m    |           |
| 1                  | 169.9                       |                      |      |           |
| 2                  | 57.9                        | 4.21                 | m    |           |
| 3                  | 27.2                        | 3.49                 | m    |           |
|                    |                             | 3.32                 | d,d  | 14.9,3.9  |
| 1' NH              |                             | 8.29                 | s    |           |
| 2'                 | 123.8                       | 6.84                 | d    | 2.2       |
| 3'                 | 108.3                       |                      |      |           |
| 4'                 | 152.3                       |                      |      |           |
| 5'                 | 101.5                       | 6.92                 | m    |           |
| 6'                 | 123.8                       | 7.34 <sup>d</sup>    | m    |           |
| 7'                 | 106.7                       | 7.34 <sup>d</sup>    | m    |           |
| 8'                 | 138.3                       |                      |      |           |
| 9'                 | 117.4                       |                      |      |           |
| 4'O-Me             | 56.4                        | 4.34                 | s    |           |
| Gly NH             |                             | 4.53                 | m    |           |
| 1                  | 171.5                       |                      |      |           |
| 2                  | 40.8                        | 3.49                 | m    |           |
|                    |                             | 1.05                 | d,d  | 17.1,5.2  |
| Abu NH             |                             | 6.76                 | d    | 6.5       |
| 1                  | 169.5                       |                      |      |           |
| 2                  | 54.7                        | 3.97                 | t,d  | 8.9, 6.2  |
| 3                  | 21.2                        | 2.00                 | m    |           |
|                    |                             | 1.87                 | m    |           |
| 4                  | 10.6                        | 0.88                 | t    | 7.5       |
| DeAla NH           |                             | 9.44                 |      |           |
| 1                  | 168.3                       |                      |      |           |
| 2                  | 136.8                       |                      |      |           |
| 3                  | 100.1                       | 5.02                 | d    | 1.4       |
|                    |                             | 4.73                 | m    |           |
| Sarc               |                             |                      |      |           |
| 1                  | 167.7                       |                      |      |           |
| 2                  | 51.4                        | 4.96                 | d    | 16.8      |
|                    |                             | 3.41                 | d    | 16.8      |
| CH <sub>3</sub> -N | 37.7                        | 3.10                 | s    |           |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

**Table S8.** Restriction sites (R-sites) engineered within the *arg* BGC and pSynBio3 sequence. R-sites introduced into coding sequences (CDS) are shown in bold. In total, 22 different restriction enzymes were selected for engineering including *EcoRI* as backup for future approaches. Introduced R-sites are unique within the *arg* BGC and pSynBio3, respectively. Interfering R-sites in other positions were eliminated during sequence modulation.

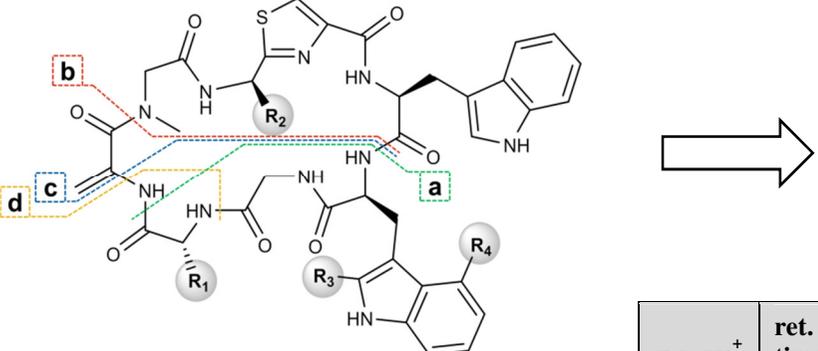
| <b>arg BGC</b>                          | <b>Position ‡</b> | <b>R-site</b>      | <b>Selected enzymes</b><br>(alphabetically) |
|---|-------------------|--------------------|---|
| 5' end <i>arg</i> BGC                   | 3                 | <i>NdeI</i>        |   |
| 5' end <i>arg2</i>                      | 14                | <i>AvrII</i>       |   |
| 5' end <i>arg2</i>                      | 181               | <b><i>StuI</i></b> |   |
| M1/M2 linker <i>arg2</i>                | 3116              | <b><i>SbfI</i></b> |   |
| M2/M3 linker <i>arg2</i>                | 6455              | <b><i>MluI</i></b> |   |
| 3' end <i>arg2</i> / 5' end <i>arg3</i> | 10746             | <i>SphI</i>        |   |
| M4/M5 linker <i>arg3</i>                | 13954             | <b><i>KpnI</i></b> |   |
| M5/M6 linker <i>arg3</i>                | 18365             | <b><i>AscI</i></b> |   |
| M6/M7 linker <i>arg3</i>                | 21501             | <b><i>NotI</i></b> |   |
| M7/M8 linker <i>arg3</i>                | 24642             | <b><i>MfeI</i></b> |   |
| 3' end <i>arg3</i> / 5' end <i>arg4</i> | 28641             | <i>NsiI</i>        |   |
| 3' end <i>arg4</i> / 5' end <i>arg5</i> | 29668             | <i>NcoI</i>        |   |
| 3' end <i>arg5</i> / <i>arg</i> BGC     | 30886             | <i>BglIII</i>      |   |
|   |                   | <i>NsiI</i>        |   |
|   |                   | <i>PacI</i>        |   |
|   |                   | <i>PciI</i>        |   |
|   |                   | <i>PmeI</i>        |   |
|   |                   | <i>SbfI</i>        |   |
|   |                   | <i>SpeI</i>        |   |
|   |                   | <i>SphI</i>        |   |
|   |                   | <i>SspI</i>        |   |
|   |                   | <i>StuI</i>        |   |
|   |                   | <i>SwaI</i>        |   |

‡ with *NdeI* = start of artificial *arg* BGC sequence

| <b>pSynBio3</b>                   | <b>Position</b> | <b>R-site</b> |
|-----------------------------------|-----------------|---------------|
| upstream <i>P<sub>npIII</sub></i> | 5               | <i>PciI</i>   |
| 5' end MCS                        | 151             | <i>NdeI</i>   |
| MCS                               | 162             | <i>AvrII</i>  |
| MCS                               | 179             | <i>SbfI</i>   |
| MCS                               | 188             | <i>MluI</i>   |
| MCS                               | 204             | <i>SphI</i>   |
| MCS                               | 216             | <i>KpnI</i>   |
| MCS                               | 225             | <i>AscI</i>   |
| MCS                               | 239             | <i>NotI</i>   |
| MCS                               | 252             | <i>MfeI</i>   |
| MCS                               | 268             | <i>NsiI</i>   |
| MCS                               | 276             | <i>NcoI</i>   |
| 3' end MCS                        | 288             | <i>BglIII</i> |
| downstream <i>t<sub>DJ</sub></i>  | 587             | <i>PmeI</i>   |
| upstream p15A <i>ori</i>          | 589             | <i>AflIII</i> |
| downstream p15A <i>ori</i>        | 1571            | <i>SspI</i>   |
| 5' end <i>kan<sup>R</sup></i>     | 1589            | <i>SpeI</i>   |
| 3' end <i>kan<sup>R</sup></i>     | 2645            | <i>EcoRV</i>  |
| 5' end <i>tet<sup>R</sup></i>     | 2659            | <i>PacI</i>   |
| 3' end <i>tet<sup>R</sup></i>     | 4044            | <i>SwaI</i>   |

**Table S9.** Sequence features of synthetic *arg* BGC versions compared to the native sequence. The coding sequences (CDS) are compared regarding three features: GC content (GC), codon adaptation index (CAI) and the degree of modification (Mod), which denotes the number of modified codons relative to the native sequence (percent change in parentheses).

| <b>Gene</b>    | <b><i>arg2</i>(3539 Codons)</b> |            |                 | <b><i>arg3</i>(5947 Codons)</b> |            |                 | <b><i>arg4</i>(337 Codons)</b> |            |                | <b><i>arg5</i>(384 Codons)</b> |            |                |
|----------------|---------------------------------|------------|-----------------|---------------------------------|------------|-----------------|--------------------------------|------------|----------------|--------------------------------|------------|----------------|
|                | <b>GC</b>                       | <b>CAI</b> | <b>Mod</b>      | <b>GC</b>                       | <b>CAI</b> | <b>Mod</b>      | <b>GC</b>                      | <b>CAI</b> | <b>Mod</b>     | <b>GC</b>                      | <b>CAI</b> | <b>Mod</b>     |
| <b>native</b>  | 64.8                            | 0.617      | 0               | 66.0                            | 0.655      | 0               | 63.7                           | 0.618      | 0              | 63.5                           | 0.630      | 0              |
| <b>V1</b>      | 64.9                            | 0.619      | 24<br>(0.7%)    | 66.1                            | 0.655      | 40<br>(0.7%)    | 64.1                           | 0.629      | 4<br>(1.2%)    | 63.5                           | 0.629      | 4<br>(1.0%)    |
| <b>V2</b>      | 66.9                            | 0.716      | 1745<br>(49.3%) | 66.9                            | 0.704      | 2946<br>(49.5%) | 65.2                           | 0.702      | 167<br>(49.6%) | 65.5                           | 0.732      | 185<br>(48.2%) |
| <b>V1-BsaI</b> | 64.9                            | 0.621      | 36<br>(1.0%)    | 66.1                            | 0.657      | 56<br>(0.9%)    | 64.1                           | 0.632      | 7<br>(2.1%)    | 63.5                           | 0.630      | 5<br>(1.3%)    |

**Table S10.** HPLC-MS data of argyrin derivatives detected in the heterologous expression strains.


|                   | $R_1$   | $R_2$              | $R_3$ | $R_4$            | $[M+H]^+$ | ret. time [min] | $m/z$       |          | $m/z$       |          | $m/z$       |          | $m/z$       |          |
|-------------------|---|--------------------|-------|------------------|-----------|-----------------|-------------|----------|-------------|----------|-------------|----------|-------------|----------|
|                   |   |                    |       |                  |           |                 | theoretical | observed | theoretical | observed | theoretical | observed | theoretical | observed |
| <b>Argyrin A</b>  | CH <sub>3</sub>                                 | CH <sub>3</sub>    | H     | OCH <sub>3</sub> | 825.3137  | 10.17           | 345.1557    | 345.1558 | 412.1438    | 412.1440 | 464.1387    | 464.1388 | 737.2500    | 737.2496 |
| <b>Argyrin A2</b> | CH <sub>3</sub>                                 | CH <sub>3</sub>    | H     | OH               | 811.2981  | 8.20            | 331.1401    | -        | 412.1438    | 412.1440 | 464.1387    | 464.1387 | 723.2344    | 723.2340 |
| <b>Argyrin B</b>  | CH <sub>2</sub> CH <sub>3</sub>                 | CH <sub>3</sub>    | H     | OCH <sub>3</sub> | 839.3294  | 10.72           | 359.1714    | 359.1715 | 412.1438    | 412.1437 | 464.1387    | 464.1388 | 737.2500    | 737.2496 |
| <b>Argyrin E</b>  | CH <sub>3</sub>                                 | CH <sub>3</sub>    | H     | H                | 795.3031  | 9.36            | 315.1452    | 315.1453 | 412.1438    | 412.1440 | 464.1387    | 646.1388 | 707.2395    | 707.2395 |
| <b>Argyrin F</b>  | CH <sub>3</sub>                                 | CH <sub>2</sub> OH | H     | OCH <sub>3</sub> | 841.3086  | 9.20            | 345.1557    | 345.1558 | 428.1387    | 428.1389 | 480.1336    | 480.1336 | 753.2450    | 753.2445 |
| <b>Argyrin F3</b> | CH <sub>3</sub>                                 | CH <sub>2</sub> OH | H     | H                | 811.2981  | 8.40            | 315.1452    | 315.1456 | 428.1387    | 428.1386 | 480.1336    | 480.1335 | 723.2344    | 723.2340 |
| <b>Argyrin G</b>  | CH <sub>2</sub> CH <sub>3</sub>                 | CH <sub>2</sub> OH | H     | OCH <sub>3</sub> | 855.3243  | 9.48            | 359.1714    | 359.1715 | 428.1387    | 428.1388 | 480.1336    | 480.1337 | 753.2450    | 753.2445 |
| <b>Argyrin G3</b> | CH <sub>2</sub> CH <sub>3</sub>                 | CH <sub>2</sub> OH | H     | H                | 825.3137  | 8.69            | 329.1608    | 329.1610 | 428.1387    | 428.1388 | 480.1336    | 480.1335 | 723.2344    | 723.2341 |
| <b>Argyrin H</b>  | CH <sub>3</sub>                                 | H                  | H     | OCH <sub>3</sub> | 811.2981  | 9.42            | 345.1557    | 345.1558 | 398.1281    | 398.1282 | 450.1231    | 450.1231 | 723.2344    | 723.2344 |
| <b>Argyrin I</b>  | CH <sub>2</sub> CH <sub>3</sub>                 | CH <sub>3</sub>    | H     | H                | 809.3188  | 9.88            | 329.1608    | 329.1609 | 412.1438    | 412.1440 | 464.1387    | 464.1388 | 707.2395    | 707.2395 |
| <b>Argyrin J</b>  | CH <sub>2</sub> CH <sub>3</sub>                 | H                  | H     | OCH <sub>3</sub> | 825.3137  | 9.95            | 359.1714    | 359.1715 | 398.1281    | 398.1282 | 450.1231    | 450.1231 | 723.2344    | 723.2343 |
| <b>Argyrin K</b>  | CH <sub>2</sub> (CH <sub>3</sub> ) <sub>2</sub> | CH <sub>3</sub>    | H     | OCH <sub>3</sub> | 853.3450  | 11.14           | 373.1870    | 373.1872 | 412.1438    | 412.1440 | 464.1387    | 464.1388 | 737.2500    | 737.2503 |
| <b>Argyrin L</b>  | CH <sub>2</sub> (CH <sub>3</sub> ) <sub>2</sub> | CH <sub>2</sub> OH | H     | OCH <sub>3</sub> | 869.3399  | 9.69            | 373.1870    | 373.1869 | 428.1387    | -        | 480.1336    | 480.1337 | 753.2450    | 753.2445 |

**Table S11.** Nonribosomal codes of A domains. Nonribosomal codes of A domains from different BGC, aligned with native argyris A1 code (Arg-A1) and 14 engineered A1 domain codes (modified residues are highlighted in bold). Residue numbering is adapted according to GrsA<sup>15</sup>.

| A domain                                | 235 | 236      | 239      | 278      | 299      | 301      | 322      | 330      | 331      | 517 |
|---|-----|----------|----------|----------|----------|----------|----------|----------|----------|-----|
| Arg-A1                                  | D   | V        | L        | F        | L        | G        | V        | V        | A        | K   |
| Mch <sub>A</sub> -A2 <sup>1</sup> (Ala) | D   | V        | W        | V        | I        | A        | A        | V        | I        | K   |
| A1-mod1                                 | D   | <b>V</b> | <b>W</b> | <b>V</b> | <b>I</b> | <b>A</b> | <b>A</b> | <b>V</b> | <b>A</b> | K   |
| Mch <sub>A</sub> -A5 <sup>2</sup> (Ala) | D   | V        | W        | V        | T        | A        | A        | I        | I        | K   |
| A1-mod2                                 | D   | <b>V</b> | <b>W</b> | <b>V</b> | <b>T</b> | <b>A</b> | <b>A</b> | <b>I</b> | <b>A</b> | K   |
| Mch <sub>S</sub> -A3 <sup>3</sup> (Ala) | D   | V        | M        | F        | I        | G        | I        | V        | A        | K   |
| A1-mod3                                 | D   | <b>V</b> | <b>M</b> | F        | <b>I</b> | <b>G</b> | <b>I</b> | <b>V</b> | <b>A</b> | K   |
| Mch <sub>S</sub> -A5 <sup>4</sup> (Ala) | D   | V        | W        | V        | L        | A        | A        | I        | I        | K   |
| A1-mod4                                 | D   | <b>V</b> | <b>W</b> | <b>V</b> | L        | <b>A</b> | <b>A</b> | <b>I</b> | <b>A</b> | K   |
| Arg-A4 (Ser/Gly)                        | D   | L        | W        | N        | M        | G        | E        | V        | W        | K   |
| A1-mod5                                 | D   | <b>L</b> | <b>W</b> | <b>N</b> | <b>M</b> | <b>G</b> | <b>E</b> | <b>V</b> | <b>A</b> | K   |
| A1-mod6                                 | D   | <b>L</b> | <b>W</b> | <b>V</b> | <b>M</b> | <b>G</b> | <b>E</b> | <b>V</b> | <b>A</b> | K   |
| A1-mod7                                 | D   | <b>L</b> | <b>W</b> | <b>N</b> | L        | <b>G</b> | <b>E</b> | <b>V</b> | <b>A</b> | K   |
| A1-mod8                                 | D   | <b>L</b> | <b>W</b> | <b>N</b> | <b>M</b> | <b>G</b> | <b>A</b> | <b>V</b> | <b>A</b> | K   |
| Sare0718 <sup>5</sup> (Ala)             | D   | M        | W        | I        | A        | A        | A        | I        | V        | K   |
| A1-mod9                                 | D   | <b>M</b> | <b>W</b> | <b>I</b> | <b>A</b> | <b>A</b> | <b>A</b> | <b>I</b> | <b>A</b> | K   |
| Cyc-A6 <sup>6</sup> (Abu)               | D   | A        | W        | F        | H        | A        | V        | A        | Y        | -   |
| A1-mod10                                | D   | <b>A</b> | <b>W</b> | F        | <b>H</b> | <b>A</b> | <b>V</b> | <b>A</b> | <b>A</b> | K   |
| A1-mod11                                | D   | <b>V</b> | <b>W</b> | F        | <b>H</b> | G        | <b>V</b> | <b>V</b> | <b>A</b> | K   |
| Tub-A3 <sup>7</sup> (Val)               | D   | A        | F        | W        | L        | G        | G        | T        | F        | -   |
| A1-mod12                                | D   | <b>A</b> | <b>F</b> | <b>W</b> | L        | <b>G</b> | <b>G</b> | <b>T</b> | <b>A</b> | K   |
| A1-mod13                                | D   | <b>A</b> | <b>F</b> | <b>W</b> | L        | <b>G</b> | <b>G</b> | <b>V</b> | <b>A</b> | K   |
| A1-mod14                                | D   | <b>A</b> | <b>F</b> | <b>W</b> | L        | <b>G</b> | <b>V</b> | <b>T</b> | <b>A</b> | K   |

<sup>1</sup>A2 domain from the myxochromide A assembly line of *M. xanthus* DK1622 (YP\_632256, aa 2028-2478)<sup>11</sup>

<sup>2</sup>A5 domain from the myxochromide A assembly line of *M. xanthus* DK1622 (YP\_632257, aa 2575-3030)<sup>11</sup>

<sup>3</sup>A3 domain from the myxochromide S assembly line of *S. aurantiaca* DW4/3-1 (CAG29032, aa 475-949)<sup>21</sup>

<sup>4</sup>A5 domain from the myxochromide S assembly line of *S. aurantiaca* DW4/3-1 (CAG29032, aa 2566-3036)<sup>21</sup>

<sup>5</sup>Alanine-activating A domain from *Salinispora arenicola* CNS-205 (CP000850)<sup>22</sup>

<sup>6</sup>A6 domain from the cyclosporin assembly line of *Tolypocladium inflatum* (CAA82227, aa 7572-7976),<sup>23</sup>

<sup>7</sup>A3 domain from the tubulylin assembly line of *Angiococcus disciformis* An d48 (CAF05648, aa 1593-1993)<sup>24</sup>

## 4.8 Unpublished results

### 4.8.1 Design, construction and evaluation of engineered argyris A4 domains

As already described in chapter 4.3.1, the fourth module of the argyris BGC is able to incorporate multiple amino acid residues into the core peptide structure, contributing to a higher diversity of argyrins. The A4 domain, responsible for amino acid substrate activation, was shown to incorporate alanine, serine and glycine, based on the panel of argyris derivatives detected in the production broth. Since water solubility is a major problem for argyris pharmacokinetics, a platform for production of derivatives with better water solubility was highly desired. Incorporation of serine in the fourth module, instead of alanine would result in production of argyris F, which shows better water solubility<sup>25</sup> as well as increased

proteasome inhibition activity<sup>26</sup>. To increase production of argyirin F, a similar genetic modification approach as in case of A1 (see chapter 4.3.3) was applied to A4 domain, to make it more specific for serine. In addition of relying purely on the nonribosomal code, to determine the specificity of the A domain, the so called 8 Å code was identified, which consists of 34 amino acid residues that confer the specificity of the A domain, 9 of which are also included in the nonribosomal code<sup>27</sup>.

For the A4 modifications three different strategies were applied. In addition to engineering of the nonribosomal code (Table S12), in some cases the 8 Å code was modified. The 8 Å code of argyirin A4 domain was extracted, and compared to five other serine specific A domains from myxobacterial BGCs with known substrate specificity, as well as argyirin A2 domain which is also serine specific (Table S13). Based on the serine specific 8 Å codes, several synthetic A domain fragments were designed and used to exchange a subdomain part of the original A4 domain, in a similar fashion as already described for A1 (see chapter 4.3.3). Alternatively in once case an entire subdomain of the A4 between core motifs A3 and A6<sup>15</sup> was replaced with the same subdomain region from argyirin A2.

**Table S12.** Specificity codes (based on the nonribosomal code) of 3 engineered argyirin A4 domains, aligned with GrsA specificity code and native argyirin A4 domain specificity code (Arg-A4).

| A domain | Observed specificity | 235 | 236 | 239 | 278 | 299 | 301 | 322 | 330 | 331 | 517 |
|----------|----------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg-A4   | Ala (Ser/Gly)        | D   | L   | W   | N   | M   | G   | E   | V   | W   | K   |
| A4-mod1  | -                    | D   | V   | W   | H   | F   | S   | L   | I   | D   | K   |
| A4-mod2  | -                    | D   | V   | W   | H   | F   | S   | L   | V   | D   | K   |
| A4-mod3  | -                    | D   | V   | W   | H   | V   | S   | L   | V   | D   | K   |

**Table S13.** Serine specific A domain specificity codes (based on the 8 Å code) from published myxobacterial BGC, aligned with GrsA 8 Å code.

| A domain             | Observed specificity | 8 Å code residues   |
|----------------------|----------------------|---|
| Arg-A4               | Ala/Ser              | L D P H F D L S V W E G N Q M L G G E Y N E Y G P A E C A V W T T H |
| Arg-A2               | Ser                  | R W M T F D V S V W E W H F F C S G E H N L Y G P T E A S I D V T A |
| Mxp-A1 <sup>1</sup>  | Ser                  | R W M T F D V S V W E W H F V C S G E H N L Y G P T E A A V D V T Y |
| Mxp-A4 <sup>2</sup>  | Ser                  | R W M T F D V S V W E W H F V C S G E H N L Y G P T E A A V D V T Y |
| Mxp-A6 <sup>3</sup>  | Ser                  | R W M T F D V S V W E W H F V C S G E H N L Y G P T E A A V D V T Y |
| Rhiz-A2 <sup>4</sup> | Ser                  | R W M T F D V S V W E W H F F T S G E H N L Y G P T E A A V D V T F |
| Chi-A1 <sup>5</sup>  | Ser                  | R W M T F D V S V W E W H F F F S G E H N L Y G P T E A A V D V S F |

<sup>1</sup>A1 domain from the myxoprincomide assembly line of *M. xanthus* DK1622 (Q1D5W2, aa 1,136 – 1,606)<sup>28</sup>

<sup>2</sup>A4 domain from the myxoprincomide assembly line of *M. xanthus* DK1622 (Q1D5W2, aa 4,727 – 5,253)<sup>28</sup>

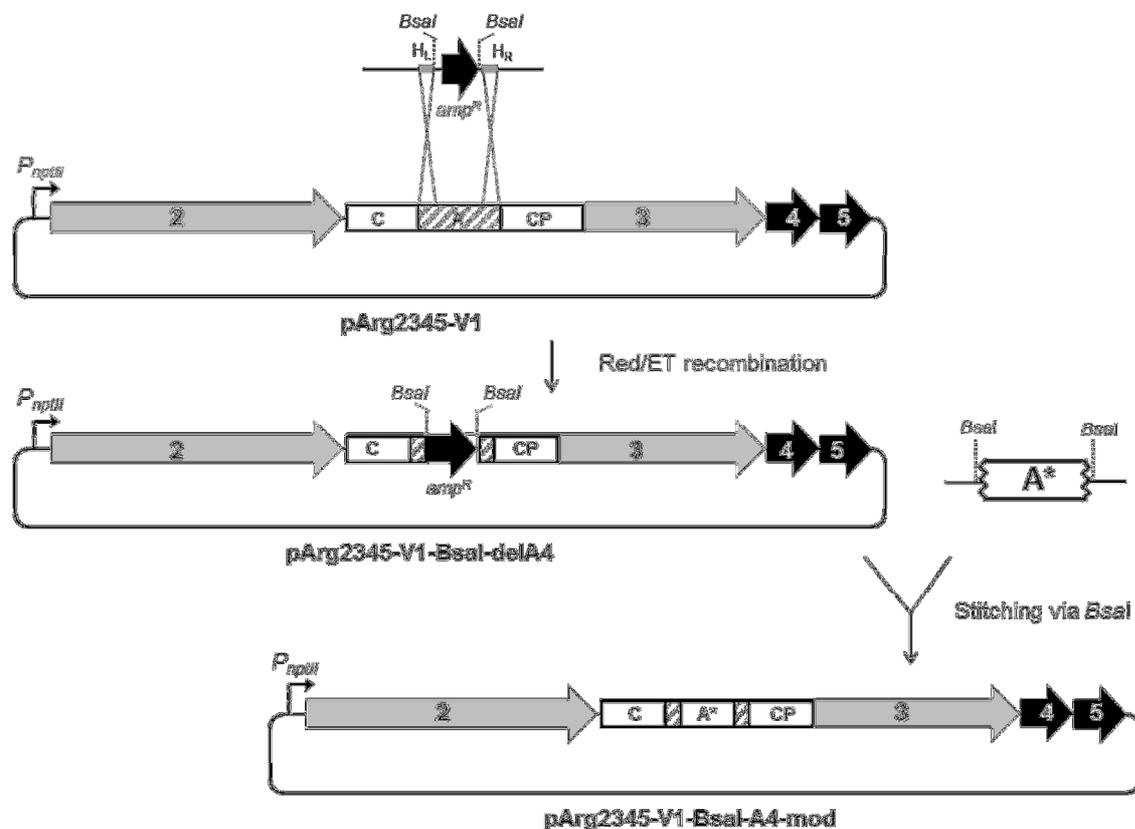
<sup>3</sup>A6 domain from the myxoprincomide assembly line of *M. xanthus* DK1622 (Q1D5W2, aa 9,142 – 9,668)<sup>28</sup>

<sup>4</sup>A2 domain from the rhizopodin assembly line of *S. aurantiaca* Sg a15 (G8YZM4, aa 928–1442)<sup>29</sup>

<sup>5</sup>A1 domain from the chivosazol assembly line of *S. cellulosum* soce56 (Q2N3S9, aa 1024–1505)<sup>30</sup>

The argyirin expression constructs with modified, synthetic A4 domains were designed based on the *BsaI* free argyirin BGC variant. This allowed us to skip one assembly step, compared to

the modifications of the A1 domain and perform modifications directly on the BGC construct (Fig. S14).



**Figure S14.** Strategy to exchange the subdomain part of A4. Using this approach 8 expression constructs with engineered A1 domains (pArg2345-V1-BsaI-A4-mod[1-3], pArg2345-V1-BsaI-A4-8A-[1-4], pArg2345-V1-BsaI-A4-argA2) were generated.

Synthetic sequences of 8 modified argyrian A4 subdomain regions, flanked by *BsaI* R-sites were obtained from a gene synthesis company. Argyrian *BsaI* free BGC construct pArg2345-V1-BsaI, was modified using Red/ET - *in vivo* recombination technique, to replace A4 subdomain region with kanamycin resistance cassette, flanked by *BsaI* R-sites (Fig. S14). Gene synthesis fragments were digested using *BsaI* enzyme, to excise the modified A subdomain regions, which were subsequently ligated into pArg2345-V1-BsaI-delA4 construct, generated by Red/ET, to construct eight modified final synthetic argyrian heterologous expression constructs (Fig. S14).

Generated argyrian BGC expression constructs with modified A4 domains (pArg2345-V1-BsaI-A4-mod[1-3], pArg2345-V1-BsaI-A4-8A-[1-4], pArg2345-V1-BsaI-A4-argA2) were transferred into *M. xanthus* DK1622  $\Delta mchA-tet$ . Obtained clones were genotypically verified by PCR (Fig. S10) and their production was evaluated by parallel cultivation with control

strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ , under the standard conditions, the same way as for the A1 modification experiment (see Supplementary information). For three of the generated expression constructs, correct transformants could not be obtained (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-A4-8A-1-BsaI$ , *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-A4-mod[2-3]-BsaI$ ). As expected, LC-MS analysis revealed standard argyirin production profile with around 1:1 ratio of argyirin A and argyirin B in case of *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ . Contrary to this, production of argyirin could this time not be detected in the extracts of any of the cultivated mutant clones. This is surprising, especially because it has already been proven, by detection of argyirin F in the wild type producer broth, that the incorporation of serine in the 4<sup>th</sup> module is indeed possible. This means that the subsequent C domain, which were also shown to act as the second gatekeeper<sup>31</sup>, should be able to process a serine containing intermediate. Most likely the modifications implemented on the extended specificity conferring code are in this case unfavorable and result in A domain sequence which is not able to fold correctly, therefore leading to its inactivity.

#### 4.8.2 Development and evaluation of the argyirin expression system for mutasynthesis studies

In order to access novel argyirin derivatives we aimed to establish a mutasynthesis approach based on the generated mutant *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1$  (Fig. S15). To verify that the approach works, we first tried to restore the argyirin production by feeding mimics of the native tripeptide intermediate activated as *N*-acetylcysteamine thioesters (SNAC), to the culture of the generated mutant *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1$ . Native tripeptide intermediate and additional variants thereof were synthesized according to organic synthesis procedures (unpublished data).

Based on the previously established heterologous expression system for argyirin production (see chapter 4.3.2) an *M. xanthus* mutant for mutasynthesis studies was constructed by deletion of the *arg2* gene from the expression construct. The expression construct pArg2345-V1 was hydrolyzed using *NdeI/SphI* restriction enzymes, to excise 10.1 kb *arg2* fragment, resulting in pArg345-V1 fragment. A 1097 bp *amp<sup>R</sup>* cassette was amplified from pUC18 plasmid using the oligonucleotides Arg44 and Arg45. The PCR fragment was hydrolyzed using *NdeI/SphI* restriction enzymes and ligated into previously generated pArg345-V1 fragment, producing pArg345-V1-*amp* plasmid, which was hydrolyzed using *AarI* restriction

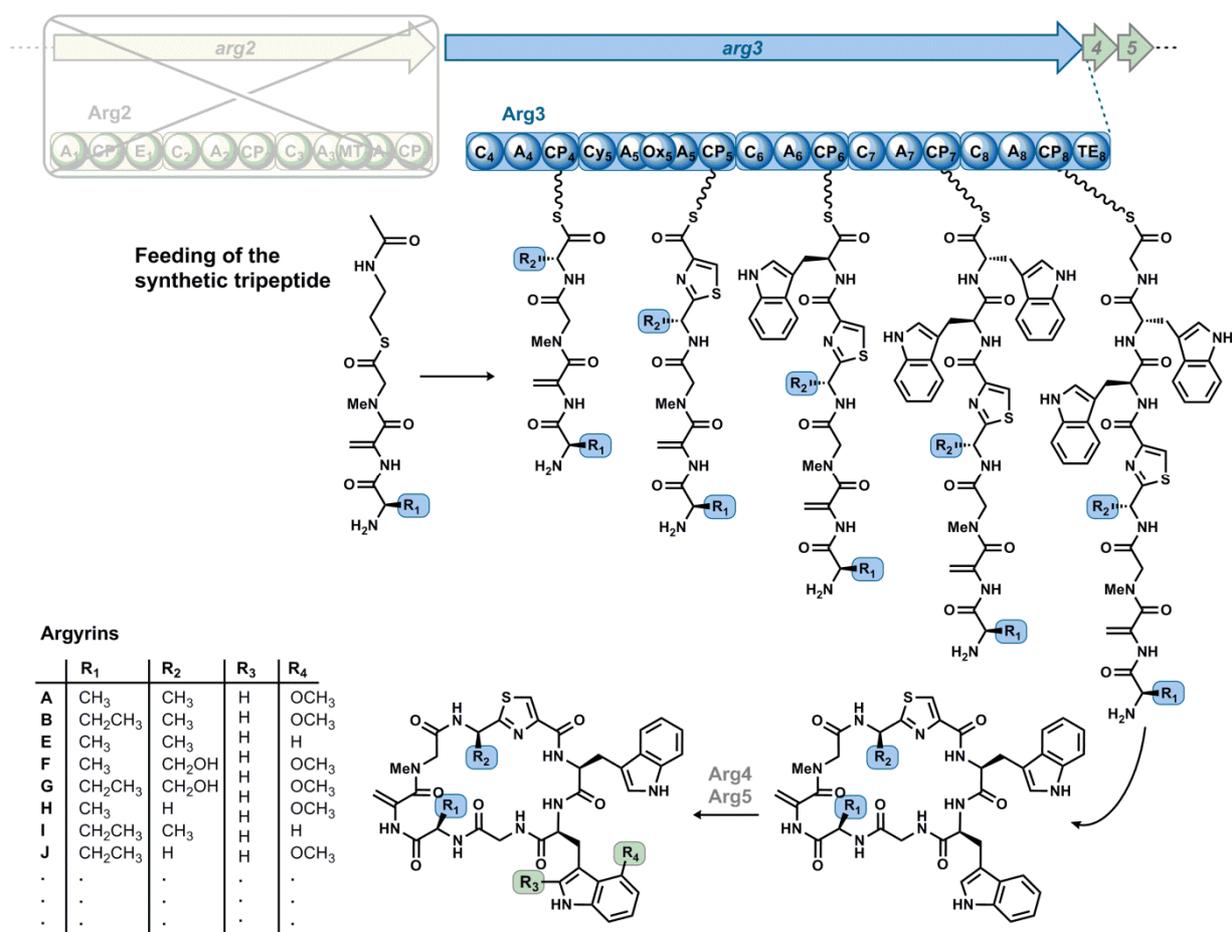
enzyme and subsequently re-ligated to construct pArg345-V1. The heterologous host *M. xanthus* DK1622  $\Delta mchA-tet$  was transformed with the modified expression construct by electroporation. Argyrin production was abolished in the generated expression strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1$  due to the absence of the first subunit (Arg2) from the argyirin megasynthetase. According to the argyirin biosynthesis model Arg2 initiates peptide core assembly by generating an enzyme bound tripeptide intermediate ( $H_2N-D-Ala-Dha-Sar$ -thioester or  $H_2N-D-Abu-Dha-Sar$ -thioester), which is subsequently transferred to the second NRPS subunit (Arg3) to extend the peptide chain with five additional amino acids followed by macrocyclization (see chapter 4.3.1).

To verify the functionality of the truncated argyirin biosynthesis operon (*arg3-arg4-arg5*) gene *arg2* was integrated into the host chromosome at a different locus. For this purpose a compatible expression vector had to be assembled. Plasmid pGH-arg2-V1 was hydrolyzed by *HindIII/XbaI* to excise *arg2* fragment which was subsequently ligated into pDKzeo1 plasmid hydrolyzed with the same enzymes, to produce pDKzeo1-arg2-V1. To introduce  $P_{nptIII}$  promoter upstream of the *arg2* gene, a 208 bp fragment was amplified from pArg2345-V1 using primers Arg47/Arg48. The resulting  $P_{nptIII}$  promoter fragment and the expression plasmid pDKzeo1-arg2-V1 were hydrolyzed by *HindIII/NdeI* restriction enzymes and ligated to produce pDKzeo1-PnptII-arg2-V1. The heterologous host *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1$  was transformed with the generated expression construct by electroporation. The obtained mutant strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1-nptII-arg2$  was shown to produce argyrins indicating that the engineered operon is functionally expressed via the integrated pArg345-V1 construct.

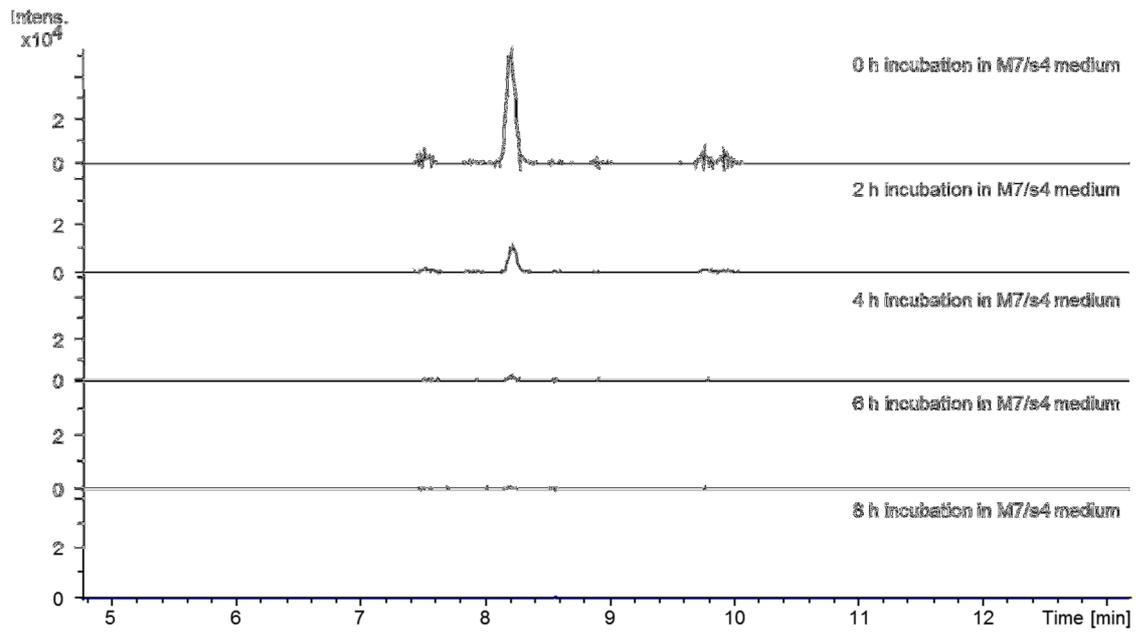
To restore argyirin production by mutasynthesis, mimics of the native tripeptide intermediates (SNAC-tripeptides) were fed to cultures of the mutant *M. xanthus* DK1622::pArg345-V1. Mutant strains of *Myxococcus xanthus* DK1622 were routinely cultivated in CTT medium amended with 50  $\mu\text{g/ml}$  kanamycin for *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1$  and 50  $\mu\text{g/ml}$  kanamycin plus 50  $\mu\text{g/ml}$  zeocin for *M. xanthus* DK1622  $\Delta mchA-tet::pArg345-V1/pDKzeo1-nptII-arg2$ . Cultivations were carried out in shake flasks on a rotary shaker at 200 rpm and at 30 °C. Feeding experiments were performed in 20 ml scale. 15 ml CTT medium was inoculated with 5 ml of a well grown pre-culture of *M. xanthus* DK1622::pArg345-V1. Mutasynthons were fed in equal portions (50  $\mu\text{L}$  each) after 24 h, 27 h, 30 h, 33 h and 36 h from a DMSO stock solution of the synthesized tripeptide thioesters (12 mg/mL). After 39 h the adsorber resin XAD-16 was added (2% final conc.) and cultivation was continued overnight. The cultures were harvested by centrifugation at 8.000 rpm for 10

min and cells/XAD were extracted with 40 ml MeOH (stirring for 60 min at room temperature followed by filtration). Evaporated extracts were re-dissolved in 200  $\mu$ l MeOH for HPLC-MS analysis. Feeding studies included parallel processing of a culture to which a pure DMSO solution (without mutasynthon) was supplied. Argyrin could be detected in pellet of *M. xanthus* DK1622  $\Delta$ *mchA-tet::pArg345-V1-nptII-arg2*, but not in pellet of *M. xanthus* DK1622  $\Delta$ *mchA-tet::pArg345-V1* fed with synthetic tripeptide.

In order to evaluate stability of the synthetic tripeptide mimic, its degradation kinetics was evaluated under the standard cultivation conditions. Two ml Eppendorf tube was filled with two ml of cultivation medium supplemented with 25  $\mu$ l of the tripeptide DMSO stock solution and incubated at 30 °C, 800 rpm in the heat block. A 100  $\mu$ l sample was taken every 2 h, centrifuged at 15000 rpm, 4 °C for 5 min and supernatant was subjected to HPLC-MS analysis. HPLC-MS analysis: column EC150/2 Nucleoshell HILIC 2.7  $\mu$ M (Macherey-Nagel, Germany), buffer A: 100 mM ammonium formate buffer B: MeCN / 0.1% HCO<sub>2</sub>H, gradient of 5% A to 12% A over 12 min (flow 800  $\mu$ L/min). The HPLC-MS analysis revealed a rather rapid degradation of the analysed mutasynthon, as after only 2 h of incubation time the majority was degraded and after 4 h of incubation almost complete degradation of the compound was observed (Fig. S16).



**Figure S15.** Mutasynthesis approach for production of argyrins. An expression strain harboring a truncated *arg* BGC lacking *arg1* and *arg2* (encoding the first subunit Arg2 of the argyrimycin NRPS megasynthetase) is fed with mutasynthons to restore argyrimycin production. Mutasynthons represent mimics of natural or unnatural tripeptide intermediates. For abbreviations of catalytic domains see Fig. 1.



**Figure S16.** Degradation kinetics of the D-Ala-Dha-Sar-SNAc mutasynthon in M7/S4 medium over 8 h period.

### 4.8.3 References

- (1) Ullrich, A.; Chai, Y.; Pistorius, D.; Elnakady, Y. A.; Herrmann, J. E.; Weissman, K. J.; Kazmaier, U.; Müller, R. Pretubulysin, a potent and chemically accessible tubulysin precursor from *Angiococcus disciformis*, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, pp. 4422–4425.
- (2) Sharma, S. B.; Signer, E. R. Temporal and spatial regulation of the symbiotic genes of *Rhizobium meliloti* in planta revealed by transposon Tn5-gusA, *Genes. Dev.* **1990**, *4*, pp. 344–356.
- (3) Tokarski, Z.; Klei, H. E.; Berg, C. M. Novel biotechnology process for the production of (+)-2-aminobutyrate, *Biotechnol. Lett.* **1988**, *10*, pp. 7–10.
- (4) Sasse, F.; Steinmetz, H.; Schupp, T.; Petersen, F.; Memmert, K.; Hofmann, H.; Heusser, C.; Brinkmann, V.; Matt, P. von; Hofle, G.; Reichenbach, H. Argyrins, immunosuppressive cyclic peptides from myxobacteria. I. Production, isolation, physico-chemical and biological properties, *J. Antibiot.* **2002**, *55*, pp. 543–551.
- (5) Oßwald, C.; Zipf, G.; Schmidt, G.; Maier, J.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Modular Construction of a Functional Artificial Epothilone Polyketide Pathway, *ACS Synth. Biol.* **2014**, *3*, pp. 759–772.
- (6) Sharp, P. M.; Li, W.-H. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications, *Nucleic Acids Res.* **1987**, *15*, pp. 1281–1295.
- (7) Li, G. W.; Oh, E.; Weissman, J. S. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria, *Nature.* **2012**, *484*, pp. 538–541.
- (8) Widmann, M.; Clairo, M.; Dippon, J.; Pleiss, J. Analysis of the distribution of functionally relevant rare codons, *BMC Gen.* **2008**, *9*, pp. 1–8.
- (9) Křížek, M.; Křížek, P. Why has nature invented three stop codons of DNA and only one start codon?, *J. Theor. Biol.* **2012**, *304*, pp. 183–187.
- (10) Seligmann, H.; Pollock, D. D. The ambush hypothesis. Hidden stop codons prevent off-frame gene reading, *DNA Cell Biol.* **2004**, *23*, pp. 701–705.
- (11) Wenzel, S. C.; Meiser, P.; Binz, T. M.; Mahmud, T.; Müller, R. Nonribosomal Peptide Biosynthesis: Point Mutations and Module Skipping Lead to Chemical Diversity, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, pp. 2296–2301.
- (12) Wu, S. S.; Kaiser, D. Markerless deletions of pil genes in *Myxococcus xanthus* generated by counterselection with the *Bacillus subtilis* sacB gene, *J. Bacteriol.* **1996**, *178*, pp. 5817–5821.
- (13) Kashefi, K.; Hartzell, P. L. Genetic suppression and phenotypic masking of a *Myxococcus xanthus* frzF- defect, *Mol. Microbiol.* **1995**, *15*, pp. 483–494.
- (14) Yan, F.; Burgard, C.; Popoff, A.; Zaburannyi, N.; Zipf, G.; Maier, J.; Bernauer, H. S.; Wenzel, S. C.; Müller, R. Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production, *Chem. Sci.* **2018**, *9*, pp. 7510–7519.
- (15) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem. Biol.* **1999**, *6*, pp. 493–505.
- (16) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.

- (17) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution, *BMC Evol. Biol.* **2007**, *7*, pp. 78–92.
- (18) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters, *Nucleic Acids Res.* **2015**, *43*, W237-W243.
- (19) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. NRSPredictor2 - a web server for predicting NRPS adenylation domain specificity, *Nucleic Acids Res.* **2011**, *39*, pp. W362-W367.
- (20) Minowa, Y.; Araki, M.; Kanehisa, M. Comprehensive Analysis of Distinctive Polyketide and Nonribosomal Peptide Structural Motifs Encoded in Microbial Genomes, *J. Mol. Biol.* **2007**, *368*, pp. 1500–1517.
- (21) Wenzel, S. C.; Kunze, B.; Höfle, G.; Silakowski, B.; Scharfe, M.; Blöcker, H.; Müller, R. Structure and Biosynthesis of Myxochromides S<sub>1-3</sub> in *Stigmatella aurantiaca*: Evidence for An Iterative Bacterial Type I Polyketide Synthase and for Module Skipping in Nonribosomal Peptide Biosynthesis, *ChemBioChem.* **2005**, *6*, pp. 375–385.
- (22) Xia, S.; Ma, Y.; Zhang, W.; Yang, Y.; Wu, S.; Zhu, M.; Deng, L.; Li, B.; Liu, Z.; Qi, C. Identification of Sare0718 As an Alanine-Activating Adenylation Domain in Marine Actinomycete *Salinispora arenicola* CNS-205, *PLoS ONE.* **2012**, *7*.
- (23) Weber, G.; Leitner, E. Disruption of the cyclosporin synthetase gene of *Tolypocladium niveum*, *Curr. Genet.* **1994**, *26*, pp. 461–467.
- (24) Sandmann, A.; Sasse, F.; Müller, R. Identification and analysis of the core biosynthetic machinery of tubulysin, a potent cytotoxin with potential anticancer activity, *Chem. Biol.* **2004**, *11*, pp. 1071–1079.
- (25) Bülow, L.; Nickeleit, I.; Girbig, A.-K.; Brodmann, T.; Rentsch, A.; Eggert, U.; Sasse, F.; Steinmetz, H.; Frank, R.; Carlomagno, T.; Malek, N. P.; Kalesse, M. Synthesis and biological characterization of argyirin F, *ChemMedChem.* **2010**, *5*, pp. 832–836.
- (26) Stauch, B.; Simon, B.; Basile, T.; Schneider, G.; Malek, N. P.; Kalesse, M.; Carlomagno, T. Elucidation of the structure and intermolecular interactions of a reversible cyclic-peptide inhibitor of the proteasome by NMR spectroscopy and molecular modeling, *Angew. Chem. Int. Ed.* **2010**, *49*, pp. 3934–3938.
- (27) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs), *Nucleic Acids Res.* **2005**, *33*, pp. 5799–5808.
- (28) Cortina, N. S.; Krug, D.; Plaza, A.; Revermann, O.; Müller, R. Myxoprincomide: a natural product from *Myxococcus xanthus* discovered by comprehensive analysis of the secondary metabolome, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, pp. 811–816.
- (29) Müller, R.; Pistorius, D. Discovery of the rhizopodin biosynthetic gene cluster in *Stigmatella aurantiaca* Sg a15 by genome mining, *ChemBioChem.* **2012**, *13*, pp. 416–426.
- (30) Perlova, O.; Gerth, K.; Hans, A.; Kaiser, O.; Müller, R. Identification and analysis of the chivosazol biosynthetic gene cluster from the myxobacterial model strain *Sorangium cellulosum* So ce56, *J. Biotechnol.* **2006**, *121*, pp. 174–191.

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

(32) Krug, D.; Müller, R. Discovery of additional members of the tyrosine aminomutase enzyme family and the mutational analysis of CmdF, *ChemBioChem*. **2009**, *10*, pp. 741–750.

## **Chapter 5**

### **Production profile engineering and precursor directed biosynthesis for production of novel argyrim derivatives**

Domen Pogorevc, Alexander Popoff, Antoine Abou Fayad, Silke C. Wenzel and Rolf Müller

(Unpublished results)

## **Contributions to the Presented Work**

### **Author's Contribution**

The author conceived and performed most of the experiments described in this chapter and interpreted the resulting data. The author performed sequence analysis, design and construction of the *arg1* gene expression construct. The author contributed to the design of synthetic promoter sequences and constructed argyrian expression constructs with replaced promoter regions. Heterologous expression of these constructs in *Myxococcus xanthus* DK1622 was conducted by the author. Furthermore, the author conducted the cultivation of the heterologous mutant strains, prepared extracts and quantified argyrian yields. The author also performed argyrian production profile engineering and analyzed the effect of different promoters on the argyrian production. The author conducted biotransformation of tryptophan derivatives, their extraction and purification. Furthermore the author performed precursor directed biosynthesis, extraction and purification of novel argyrian analogues. The author conceived and wrote this chapter.

### **Contribution by Co-Workers**

Alexander Popoff performed NMR analysis and structure elucidation of tryptophan and argyrian derivatives. Silke C. Wenzel contributed to the design of the *arg1* gene expression construct and synthetic promoter sequences. Antoine Abou Fayad contributed to the experimental design of the tryptophan biotransformation as well as the purification and the structure elucidation of the tryptophan and argyrian analogues. He also performed proofreading of this chapter. Rolf Müller was responsible for the conception of the project and performed the proofreading of this chapter.

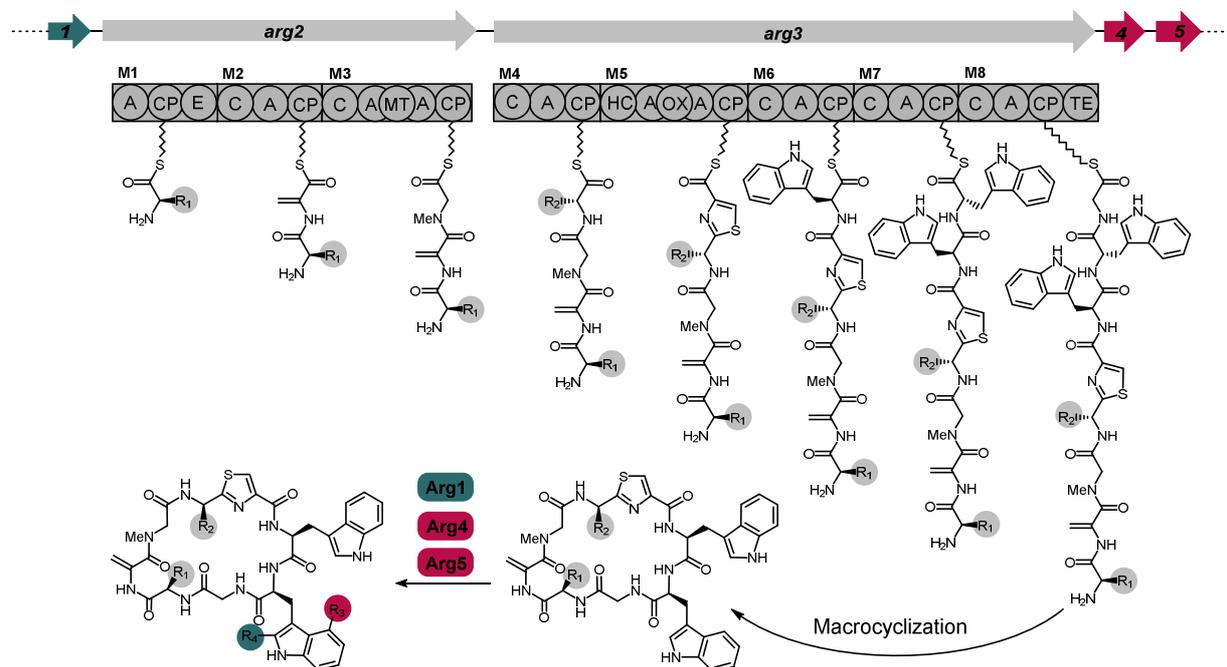
## 5 Production profile engineering and precursor directed biosynthesis for production of novel argyrin derivatives

### 5.1 Introduction

New antibiotics are urgently needed to combat more and more multidrug-resistant pathogens emerging every day. Especially hard to treat are Gram negative bacteria, since most of the antibiotics struggle with penetration of the bacterial wall to find their way to the cellular target site. One particularly troublesome multidrug-resistant Gram negative pathogens is *Pseudomonas aeruginosa*. This bacterium is a major cause for nosocomial infections with high mortality rate worldwide, therefore there is a high demand for novel, effective antibiotics targeting this pathogen<sup>1</sup>. Traditionally microbial natural products represent a major source of novel antibiotics<sup>2</sup>. A promising group of natural compounds discovered from myxobacteria are the argyrins, initially named antibiotics A21459<sup>3,4</sup>. These cyclic octapeptides exhibit an interesting antibacterial activity profile as they seem to be particularly effective against *Pseudomonas aeruginosa*, *Burkholderia multivorans* and *Stenotrophomonas maltophilia*<sup>5,6</sup>. Recently it was shown that elongation factor G (EF-G) is the cellular target for argyrin in *P. aeruginosa*<sup>7,8</sup>. Upon binding to EF-G argyrin prevents protein synthesis and inhibits bacterial growth. The same target is also addressed by the steroid antibiotic fusidic acid, which is in clinical use e.g. to treat *Staphylococcus* infections, but it is not effective against Gram-negative bacteria like pseudomonads. By mutagenesis studies and structural analysis it could also be shown that argyrin binds to an alternative site of EF-G that clearly differs from the one of fusidic acid<sup>7,8</sup>. Argyrins were also shown to inhibit the eukaryotic elongation factor G1 (EF-G1)<sup>8</sup> and are described as proteasome inhibitors<sup>9</sup>. Due to their cytotoxic activity profile they represent promising lead structures for the development of anticancer drugs<sup>10</sup>. Besides their antibacterial and cytotoxic properties argyrins also exhibit immunosuppressive activity and are currently studied as a potent inhibitor of T-cell independent antibody formation<sup>11</sup>.

Argyrin biosynthetic pathway from a myxobacterial producer strain *Cystobacter sp.* SBCb004 (Fig. 1) was recently identified and analyzed. Heterologous expression system based on synthetic genes and *Myxococcus xanthus* as host organism, led to identification of novel argyrin derivatives and enabled production yields of argyrin A and B of up to 160 mg/L<sup>12</sup>. The optimized production system provides a versatile platform for improved supply of argyrins and novel derivatives thereof. In this study we aimed to further optimize the production system, direct the biosynthesis towards certain analogues with improved

pharmaceutical properties as well as produce novel argyrin derivatives by precursor directed biosynthesis.



**Figure 1.** The biosynthetic pathway for argyrin production in *Cystobacter sp.* SBCb004 (Arg1, radical SAM-dependent methyltransferase; Arg2/Arg3, nonribosomal peptide synthetases; Arg4, O-methyltransferase; Arg5, tryptophan 2,3-dioxygenase).

## 5.2 Results and discussion

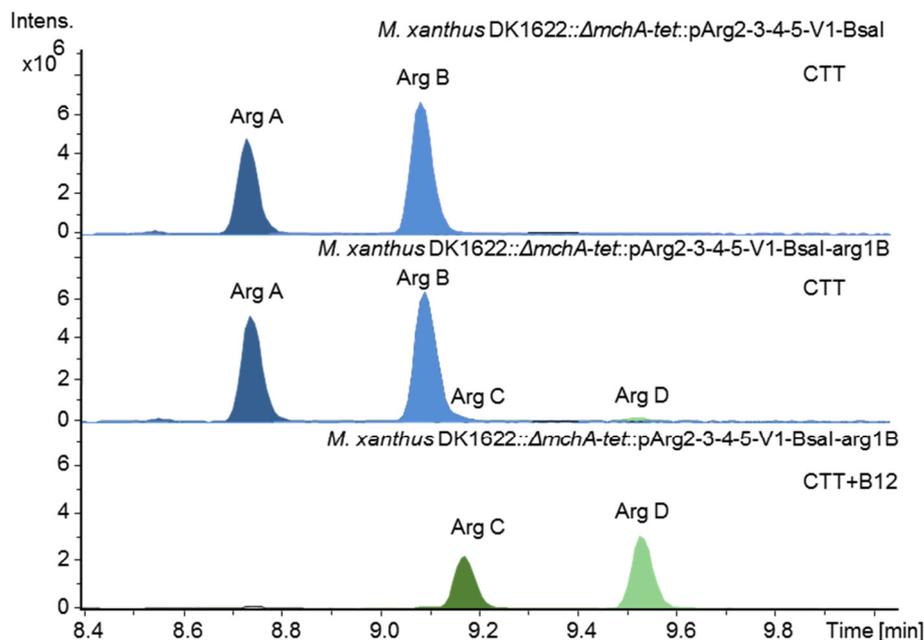
### 5.2.1 Argyrin production profile engineering

#### 5.2.1.1 Production of argyrins C and D by co-expression of radical SAM - Arg1

Gene *arg1* was excluded from the argyrin biosynthetic gene cluster (BGC) in our previous study<sup>12</sup>, to simplify the production spectra and avoid partial methylation by Arg1. In light of novel activities found for argyrins, especially immuno-modulation, it was important to also investigate the bioactivity of methylated argyrin variants such as argyrins C and D. To achieve production of the latter compounds, an expression vector for co-expression of *arg1* in the heterologous argyrin producer *M. xanthus* DK1622  $\Delta mchA$ -*tet*::pArg2345-V1-BsaI, was designed.

An extensive *in silico* analysis of the native producer's (*Cystobacter* SBCb 004) genome sequence was performed using FramePlot software package, to annotate the *arg1* gene sequence and use it as a template for construction of functional expression construct. To facilitate the construction of a suitable expression vector, a pGH-arg1-V1 plasmid, harboring

a truncated version of the *arg1* gene, was used. The plasmid was obtained from a gene synthesis company (Table S4), however the *arg1* synthetic sequence was truncated due to an error in the original annotation and was missing the 5' terminus of the gene. To reconstitute the gene, the 5'-*arg1* gene extension fragment was PCR amplified from the genomic DNA of wild type producer *Cystobacter* SBCb004. Since no clear ribosomal start site (RBS) sequence could be identified in the *Cystobacter* SBCb004 genome, an RBS from *mchA* (AAGGAGG<sup>13</sup>) was introduced on the forward primer (Table S3). The fragment was ligated into the pGH-*arg1*-V1 plasmid, yielding the pGH-*arg1*B plasmid harboring reconstituted *arg1* sequence. The *arg1*B fragment was released from pGH-*arg1*B plasmid and inserted into previously constructed pUC18-*zeo*-*Hom*-MCS expression vector (see Supplementary information) by traditional restriction-digestion and ligation to produce pUC18-*zeo*-*Hom*-*arg1*B expression construct. Detailed description of the plasmid construction is provided in the supplementary information. The construct was transformed into the host strain *M. xanthus* DK1622  $\Delta$ *mchA*-*tet*::pArg2345-V1-BsaI and clones growing on selection plates amended with kanamycin and zeocin were analyzed for correct chromosomal integration of the expression plasmid by PCR (Fig. S8). After genotypic verification, the generated mutant strain *M. xanthus* DK1622  $\Delta$ *mchA*-*tet*::pArg2345-V1-BsaI-*arg1*-B and *M. xanthus* DK1622  $\Delta$ *mchA*-*tet*::pArg2345-V1-BsaI were cultivated under standard conditions in CTT medium at 30 °C for 6 days (see Supplementary information). The cultures were then centrifuged and the cell mass with Amberlite XAD-16 resin were extracted with EtOAc and analyzed for argyrim production. As expected HPLC-MS analysis revealed successful production of argyrins A and B in case of *M. xanthus* DK1622  $\Delta$ *mchA*-*tet*::pArg2345-V1-BsaI. Similar amounts of argyrins A and B could be observed in the DK1622  $\Delta$ *mchA*-*tet*::pArg2345-V1-BsaI-*arg1*-B mutant, however, after a closer inspection, minor amounts of argyrins C and D could also be detected (Fig. 2). To improve methylation efficiency and aid conversion of argyrins A and B into their methylated counterparts, argyrins C and D, the cultivation was repeated and the production medium was supplemented with vitamin B<sub>12</sub>. Vitamin B<sub>12</sub> is an important co-factor for radical-SAM type of enzymes<sup>14</sup>, which has been shown to be a curtail cofactor for activity of radical SAMs<sup>15,16</sup>. The cultivation, extraction and sample analysis was performed the same way as previously described. HPLC-MS analysis revealed significant improvement in production levels of argyrins C and D, whereas argyrins A and B could no longer be detected (Fig. 2). Additional details on cultivation are provided in the supplementary information.



**Figure 2.** Chromatogram showing conversion of argyryns A and B by coexpression of *arg1* gene in *M. xanthus* DK1622  $\Delta mchA-tet::pArg2-3-4-5-V1-BsaI$  strain and further improvement of argyryn C and D production by addition of vitamin B12. HPLC-MS analysis of supernatant showing extracted ion chromatograms (EIC)  $[M + H]^+ = 825.313$  (Arg A),  $[M + H]^+ = 839.329$  (Arg B),  $[M + H]^+ = 839.329$  (Arg C),  $[M + H]^+ = 853.344$  (Arg D).

### 5.2.1.2 Argyrin production profile engineering by variation of cultivation conditions

Actions of the tailoring genes *arg1*, *arg4* and *arg5* combined with the relaxed specificity of the first and the fourth module of the argyryn biosynthesis, lead to high structural diversity of argyryns. A complex mixture of derivatives produced by the native or heterologous producer is inconvenient to work with and poses a challenge when trying to purify a specific derivative with high purity. We applied extensive optimization of medium and cultivation conditions, to decrease structural complexity of argyryns and simplify future purification processes.

By cultivating heterologous argyryn producer *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  in different media and/or by feeding  $\alpha$ -aminobutyric acid (Abu), we were able to shift production between argyryn A and argyryn B. By using the previously described mutant, with co-expressed *arg1* gene (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-arg1-B$ ), production could further be manipulated to produce methylated argyryn derivatives (argyryn C and D) (see chapter 5.2.1.1).

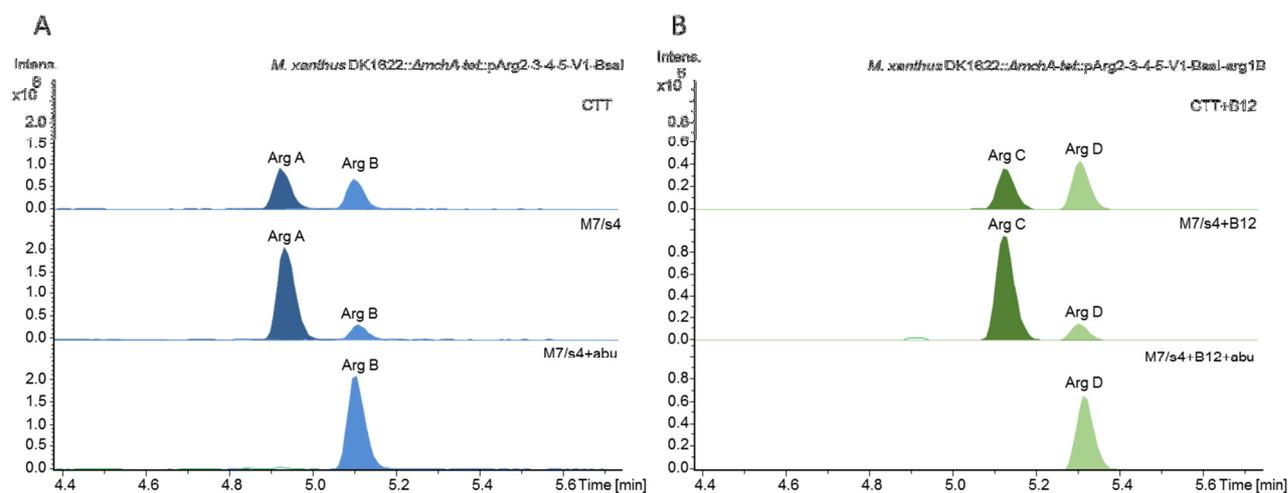
As reported in our previous study and seen in our initial cultivation experiments (see chapter 5.2.1.1), argyryn heterologous producers (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1$  and *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ ) produce around equal amounts of argyryn A and B when cultivated in standard CTT medium. This suggests that the first module of argyryn biosynthesis is promiscuous as it incorporates either alanine (Ala) or aminobutyric

acid (Abu), resulting in production of argyrin A or argyrin B, respectively<sup>12</sup>. However, when CTT medium was supplemented with 10 mM Abu, an almost 100% shift towards argyrin B was observed. This suggests that there is nevertheless a preference for incorporation of Abu, but most likely due to its low supply in the CTT medium, Ala will be incorporated after Abu depletion.

In an effort to increase argyrin production, effectiveness of previously established M7/s4 medium<sup>17</sup> was evaluated in our previous study<sup>12</sup>. This medium was specifically developed for high titer production of antibiotics myxopyronin and coralopyronin by *M. xanthus* heterologous host<sup>17</sup>. During the first cultivation in M7/s4 a significant difference in the production profile, compared to the CTT, was observed in addition to the increased argyrin yield. In M7/s4 argyrin A was produced almost exclusively (Fig. 3A), probably due to the low availability of Abu in this medium. In the following experiment the M7/s4 medium was supplemented with 10mM Abu and a near 100% production of argyrin B was observed (Fig. 3A), supporting the hypothesis of low Abu availability in M7/s4.

To prove that the same approach of derivative manipulation can be used for argyrins C and D the same cultivation experiments were performed with *M. xanthus* DK1622  $\Delta$ *mchA-tet::pArg2345-V1-BsaI-arg1B*. During the initial cultivation experiments with co-expressed *arg1* gene in CTT medium only small amounts of argyrins C and D could be detected (see chapter 5.2.1.1 and Fig. 2). However, once the medium was supplemented with 1 mg/mL of vitamin B12, a 100% turnover was observed (see chapter 5.2.1.1 and Fig. 1). Now that a complete shift to argyrins C and D was possible, the same production manipulation as described for argyrins A and B above, could be performed (Fig. 3B). Detailed description of the cultivation is provided in the supplementary information.

Medium and cultivation condition engineering in this study led to shifting ratios of major argyrin derivatives produced by the *M. xanthus* heterologous producer. Exchange of the standard CTT medium with M7/s4 medium allowed almost exclusive production of argyrin A, which could be entirely converted to argyrin B by supplementation with Abu. The same results were achieved for production of argyrins C and D, by applying the same conditions to the mutant with co-expressed *arg1* gene. Addition of vitamin B12 was in this case necessary, to improve the activity of *arg1* (Fig. 3).



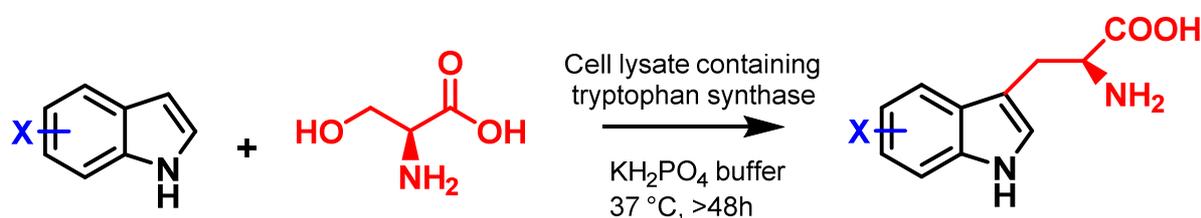
**Figure 3.** Variation in production of argyriin A and B in different media (A). Variation in production of argyriin C and D in different media (B). HPLC-MS analysis of supernatant showing extracted ion chromatograms (EIC)  $[M + H]^+ = 825.313$  (Arg A),  $[M + H]^+ = 839.329$  (Arg B),  $[M + H]^+ = 839.329$  (Arg C) and  $[M + H]^+ = 853.344$  (Arg D).

### 5.2.2 Precursor directed biosynthesis for production of novel argyriin derivatives

Precursor directed biosynthesis is based on the cellular uptake of modified biosynthetic intermediates and their incorporation into the core structure of the produced compound<sup>18</sup>. It relies on unspecific substrate incorporation and processing by the gatekeeping enzymes (e.g. adenylation and condensation domains) and is thus a fairly simple method to execute. It requires only rudimentary knowledge of the host organisms' growth conditions and the biosynthetic assembly process of the target compound.

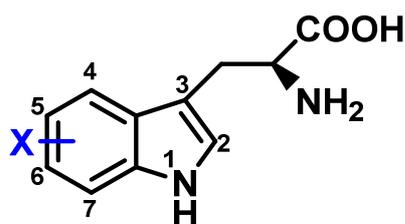
As some modules of argyriin BGC were already shown to act promiscuously, similar behavior could possibly be expected from the downstream modules. The sixth and the seventh module of the argyriin assembly line are responsible for incorporation of two consecutive tryptophan residues into the argyriin core structure<sup>12</sup>. Since argyriins have poor pharmacokinetic properties (e.g. low water solubility) it would be beneficial to introduce new functional groups on those residues for subsequent semisynthetic modifications. Introduction of new chemical groups can have a direct impact on the compounds pharmacokinetics, as observed in argyriin F where an additional hydroxyl group significantly increases the compound's water solubility<sup>10</sup> or in the study by García-Pindado et. al. where introduction of brominated tryptophan in cyclic peptides improved passive diffusion, permeability across membranes, biostability in human serum and cytotoxicity<sup>19</sup>. Furthermore, introduction of bromine or iodine groups can be employed for additional synthetic modifications<sup>19–21</sup>.

By utilizing a readily available procedure for biotransformation of tryptophan derivatives, which focuses on fusing commercially available indole derivatives and serine by tryptophan synthase<sup>22</sup> (Fig. 4), we were able to produce a library of nine tryptophan derivatives (Fig. 5) and use them for supplementation of culture broth of the argyirin heterologous producer. To evaluate the ability of the argyirin BGC to accept modified tryptophan derivatives, a set of easily accessible indoles were obtained from commercial vendors and used for production of their corresponding tryptophans (Fig. 5). One of the main obstacles during the reaction procedure was poor solubility of the indoles, which resulted in large amounts of precipitate in the reaction buffer, leading to lower yields. To overcome this issue, reaction buffer was supplemented with increasing amounts of DMSO and the reaction efficiency was evaluated. Results showed increasing reaction efficiency with higher amounts of DMSO added to the reaction buffer. The best result was achieved by addition of 20% DMSO, which had a beneficial effect on the indole solubility and was still tolerated by the tryptophan synthase contained in the dialysis bag (data not shown). The method was successfully used to produce the 7-I, 5-Br, 5-Cl, 4-F, 5-F, 6-F, 7-F, 6-NO<sub>2</sub> and 7-NO<sub>2</sub> tryptophans (Fig. 5). Detailed description of the biotransformation, isolation and structure elucidation of the tryptophan derivatives are provided in the supplementary information.



**Figure 4.** Tryptophan biotransformation reaction. Chemical reaction showing tryptophan biotransformation from indole and serine using tryptophan synthase. The X represents substitution on the indole ring (in this case F, Cl, Br, I or NO<sub>2</sub>).

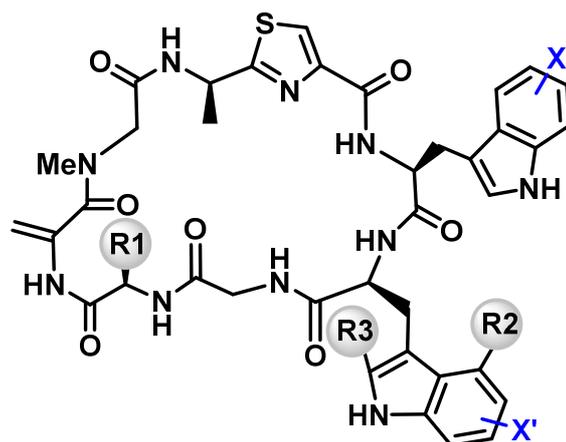
Tryptophan derivatives produced by the above described method were used to supplement the argyirin heterologous producer culture broth. We hoped that the modules 6 and 7 of the argyirin assembly line exert broader substrate specificity and would in this case be able to tolerate modification on the indole ring, thus accepting various tryptophan derivatives. The method could lead to production of novel argyirin analogs, provided that the incorporated tryptophan derivatives would also be processed by the rest of the assembly line.



| Name                   | Mass   | X               |
|------------------------|--------|-----------------|
| 4-F-Trp                | 222.08 | F               |
| 5-F-Trp                | 222.08 | F               |
| 6-F-Trp                | 222.08 | F               |
| 7-F-Trp                | 222.08 | F               |
| 7-I-Trp                | 329.99 | I               |
| 5-Cl-Trp               | 238.05 | Cl              |
| 5-Br-Trp               | 282.00 | Br              |
| 7-NO <sub>2</sub> -Trp | 249.07 | NO <sub>2</sub> |
| 6-NO <sub>2</sub> -Trp | 249.07 | NO <sub>2</sub> |

**Figure 5.** Table of produced tryptophan derivatives. The X represents substitution at the corresponding position of the indole ring, marked in the figure.

Both heterologous argyirin producer mutants were cultivated in M7/s4 medium supplemented with suitable antibiotics, vitamin B<sub>12</sub> and Amberlite XAD-16 resin under standard cultivation conditions (see Supplementary information). Tryptophan derivatives were fed to the cultures twice per day and Abu was added to each culture flask to help decrease the number of derivatives by directing the production towards analogs of argyirin B. After 6 days, Amberlite XAD-16 resin and cells were harvested by centrifugation, the supernatant was removed and the pellets were frozen and freeze dried. Dry pellets were extracted by ethyl acetate and crude extracts were dissolved in a mixture of acetonitrile and DMSO, before purification on HPLC. Using this method, a library of argyirin derivatives could be produced, with structures of 14 confirmed by NMR (Fig. 6, Supplementary information). Detailed description of the precursor directed biosynthesis, isolation and structure elucidation as well as absolute yields of the argyirin derivatives are provided in the supplementary information.



| Name                 | Mass    | R1                              | R2               | R3              | X  | X' |
|----------------------|---------|---------------------------------|------------------|-----------------|----|----|
| Argyrin A            | 824.30  | CH <sub>3</sub>                 | OCH <sub>3</sub> | H               | /  | /  |
| Argyrin B            | 838.32  | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | H               | /  | /  |
| Argyrin C            | 838.32  | CH <sub>3</sub>                 | OCH <sub>3</sub> | CH <sub>3</sub> | /  | /  |
| Argyrin D            | 852.33  | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | CH <sub>3</sub> | /  | /  |
| Argyrin I            | 808.30  | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | /  | /  |
| 4-F-argyrin C        | 856.31  | CH <sub>3</sub>                 | OCH <sub>3</sub> | CH <sub>3</sub> | F  | H  |
| 6'-F-argyrin B       | 856.31  | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | H               | H  | F  |
| 6'-F-argyrin C       | 856.31  | CH <sub>3</sub>                 | OCH <sub>3</sub> | CH <sub>3</sub> | H  | F  |
| 6,6'-di-F-argyrin C2 | 844.29  | CH <sub>3</sub>                 | H                | CH <sub>3</sub> | F  | F  |
| 7,7'-di-F-argyrin I  | 844.29  | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | F  | F  |
| 5-Cl-argyrin B       | 872.28  | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | H               | Cl | H  |
| 5'-Cl-argyrin I      | 842.27  | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | H  | Cl |
| 5,5'-di-Cl-argyrin I | 876.23  | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | Cl | Cl |
| 5-Br-argyrin B       | 916.23  | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | H               | Br | H  |
| 5'-Br-argyrin I      | 886.22  | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | H  | Br |
| 5,5'-di-Br-argyrin I | 964.13  | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | Br | Br |
| 7-I-argyrin B        | 964.22  | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | H               | I  | H  |
| 7,7'-di-I-argyrin B  | 1090.11 | CH <sub>2</sub> CH <sub>3</sub> | OCH <sub>3</sub> | H               | I  | I  |
| 7,7'-di-I-argyrin I  | 1060.10 | CH <sub>2</sub> CH <sub>3</sub> | H                | H               | I  | I  |

**Figure 6.** List of major native argyrin derivatives and novel argyrin derivatives. The R1, R2 and R3 positions are highlighted with spheres and the X represents the substitution at the corresponding position of the indole ring, marked in the argyrin structure.

### 5.2.3 Optimization of argyrim production in *M. xanthus* DK1622 by utilization of new promoter systems

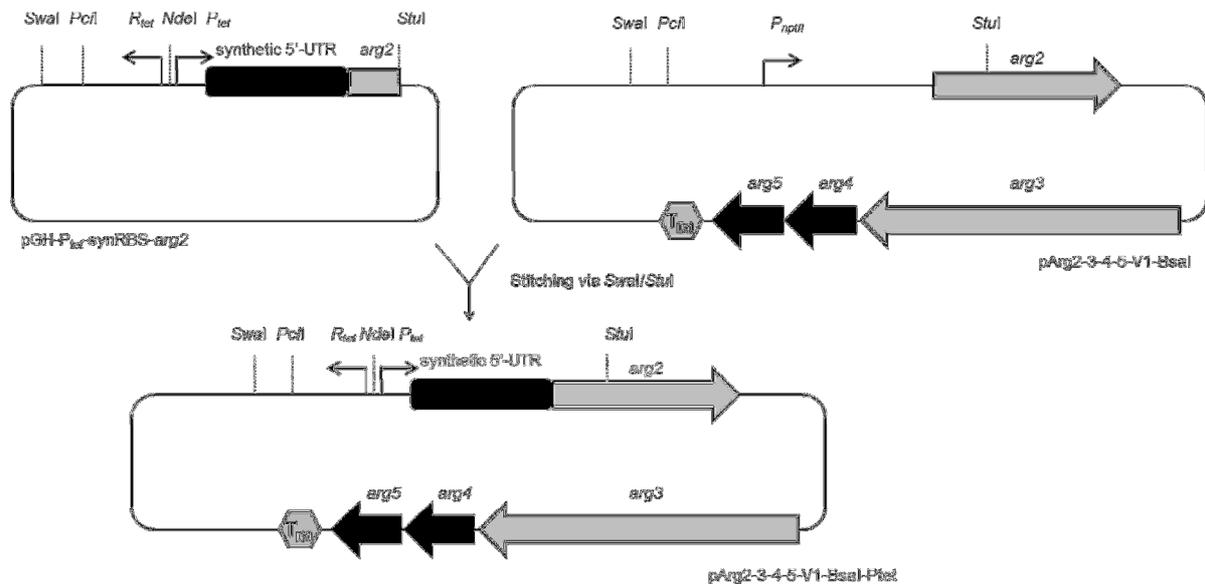
Promoters are the key regulators of gene expression that have the potential to increase or decrease of secondary metabolite production levels<sup>23</sup>. There are only a handful of promoter systems, which were successfully used for gene expression in myxobacteria, including  $P_{nptII}$ ,  $P_{tet}$ ,  $P_{pm}$  and  $P_{van}$  (see Table S2). Most of those systems were initially developed and evaluated in different host organisms and were later tested in myxobacteria due to a lack of highly efficient myxobacterial promoters.  $P_{nptII}$  and  $P_{van}$  are considered superior to other tested promoters systems, due to their high transcription rate and are thus commonly used for BGC activation and heterologous BGC expression. Furthermore,  $P_{van}$  is the only viable option when it comes to inducible promoters, as it was shown to be more efficient than the  $P_{IPTG}$ <sup>24</sup> and no other reliable inducible promoters are described for myxobacteria<sup>25</sup>. Since no data for direct comparison of the gene expression by the aforementioned promoters in myxobacteria was available, we decided, to evaluate their effect on expression of argyrim BGC. As argyrim heterologous production platform with high yield output was the main priority, this would provide a good variety of systems to determine the optimal candidate.

To evaluate different promoter systems for production of argyrins, 6 new promoter constructs, based on 4 different promoter systems were designed ( $P_{nptII}$ ,  $P_{van}$ ,  $P_{pm}$ ,  $P_{tet}$ ). Original promoter  $P_{nptII}$  from the pArg2345-V1-BsaI construct was exchanged with those 6 new designed promoter sequences, to evaluate the effect on argyrim production. To make results of all promoters consistent and comparable between each other, and to achieve the most optimal transcription and translation rate, a specific 5'-untranslated region (5'-UTR) was designed and used in all 6 designed constructs. This 5'-UTR sequence was generated using online RBS calculator tool<sup>26,27</sup> and synthesized by a gene synthesis company.

Synthetic fragments were designed for each of the promoter systems (Table S4), harboring specific promoter, followed by a 31 bp 5'-UTR region (TACGAGAGCAAAAACGAGGAG-AGGAGAAGAT) and first 61 bp of the downstream *arg2* gene, ending with a unique *StuI* restriction endonuclease at the 3' end. For cloning purposes all constructs were also flanked by the *SwaI* restriction endonucleases at the 5' end, allowing integration into pArg2345-V1-BsaI expression construct as outlined in the Figure 7.

In case of  $P_{nptII}$  of the neomycin phosphotransferase II resistance gene from Tn5<sup>28</sup> we used the same promoter sequence as originally used in pArg2345-V1-BsaI expression construct, to design nptII-synRBS-arg2 synthetic fragment. The only modification was the replacement of

the 5'-UTR region with the new, synthetic one and the introduction of the flanking restriction sites for the cloning purposes. The sequence of  $P_{van}$  promoter<sup>24</sup> was used to design the Pvan-synRBS-arg2 fragment and only the first 100 bp of the original  $P_{van}$  promoter sequence were implemented to design Pvan-wo\_vanR-synRBS-arg2 fragment with deleted repressor sequence in an attempt to make it function as a constitutive promoter system. The same approach was used in case of  $P_{tet}$  system<sup>29</sup>, to design Ptet-synRBS-arg2 and Ptet-wo\_tetR-synRBS-arg2 constructs. For version without repressor the first 63 bp of the *tet* promoter system were used. In case of the  $P_{pm}$  promoter, originating from *Pseudomonas putida*, the improved version of the promoter system, generated by random mutagenesis approach<sup>30</sup>, was used to design Ppm-synRBS-arg2 construct. For assembly purposes the *StuI* site, present in the promoter sequence, had to be mutated. This position is important for activity of the promoter as described in the promoter mutagenesis experiment<sup>30</sup>. The initial design was based on ML2-2 sequence that showed the best results in experiments performed by Bakke et. al., however due to the aforementioned mutation of *StuI* restriction site, a hybrid between ML2-2 and ML2-1 sequences, which both achieved good transcription levels, was generated. The  $P_{pm}$  promoter normally acts as an inducible promoter in pseudomonads, however we used it as a constitutive promoter, as it was shown to work constitutively in *M. xanthus*<sup>25</sup>.



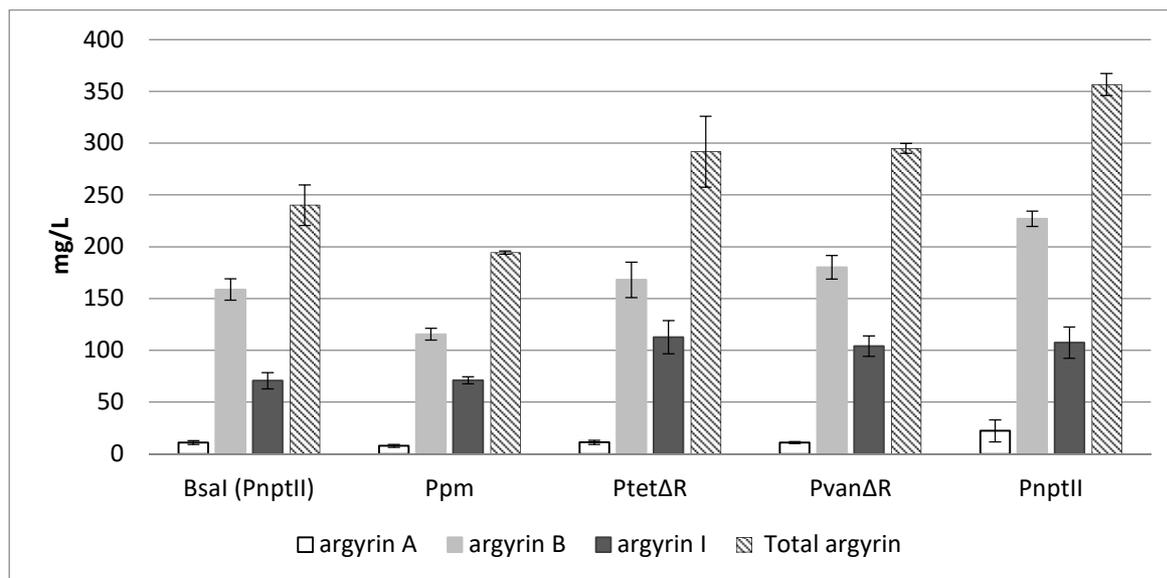
**Figure 7.** Assembly strategy for exchange of the promoter systems upstream of the *arg* operon. The same strategy was applied in case of all promoter fragments, the figure shows  $P_{tet}$  sequence as an example.

Generated argyirin BGC expression constructs with exchanged promoter (pArg2345-V1-BsaI-Ppm, pArg2345-V1-BsaI-Ptet, pArg2345-V1-BsaI-Ptet $\Delta$ R, pArg2345-V1-BsaI-PnptII,

pArg2345-V1-BsaI-Pvan and pArg2345-V1-BsaI-Pvan $\Delta$ R) were transferred into *M. xanthus* DK1622  $\Delta$ mchA-tet and clones growing on selection plates amended with kanamycin were analyzed for correct chromosomal integration of the expression plasmid by PCR (Fig. S1). Details on promoter exchange are provided in the supplementary information. After genotypic verification, the promoter mutant strains and *M. xanthus* DK1622  $\Delta$ mchA-tet::pArg2345-V1-BsaI were cultivated under standard conditions in M7/s4 medium at 30 °C for 6 days. The culture of *M. xanthus* DK1622  $\Delta$ mchA-tet::pArg2345-V1-BsaI-Ptet was induced with tetracycline and the culture of *M. xanthus* DK1622  $\Delta$ mchA-tet::pArg2345-V1-BsaI-Pvan was induced with vanillate. The M7/s4, supplemented with amino acid solution was selected to reduce possibility of bottlenecks by the lack of nutrients, since we already showed that the production in supplemented M7/s4 increases significantly compared to the CTT medium<sup>12</sup>. To evaluate production of argyrim by constitutive promoter systems, the promoter mutants and control strain were cultivated under standard conditions in M7/s4 medium continuously supplemented with the amino acid solution. The cultures were harvested by centrifugation to analyze the cell pellet with Amberlite XAD-16 resin for the presence of argyrins. HPLC-MS analysis revealed successful production of argyrins by all mutants, which revealed significant differences in production levels between different promoter systems. Supplementation with the amino acid solution, led to significant increase in production of argyrim I (R<sub>2</sub>, R<sub>3</sub> = H; Fig. 6), which has already been observed<sup>12</sup>. This indicates a bottleneck in the Arg5 (tryptophan-2,3-dioxygenase) and Arg4 (hydroxyindole O-MT) mediated tailoring biochemistry under the optimized production conditions. High amounts of argyrim, produced under those optimized conditions, probably exceed the limit of Arg5 and/or Arg4, which is unable to efficiently hydroxylate all of the produced argyrim, leading to increased production of argyrim I. The detailed description of the cultivation and argyrim quantification is provided in the supplementary information.

Detailed comparison of relative peak areas of argyrim I and argyrim B revealed that the 5'-UTR region used in the promoter system has a big impact on the production levels. We hypothesized that the system with the strongest promoter and the most optimal leader sequence will result in the best yield due to the highest transcription and translation rate, which leads to high levels of biosynthetic proteins<sup>31</sup>. The highest constitutive argyrim production was achieved with the use of  $P_{nptIII}$  promoter in combination with the optimized artificial 5'-UTR region. The same  $P_{nptIII}$  promoter was also used in the control construct (pArg2345-V1-BsaI) meaning that the efficient translation was probably the bottleneck in the control system. Once the 5'-UTR region was exchanged with the more optimal one, the

production increased by nearly 50% (Fig. 8). Promoters  $P_{tet\Delta R}$  and  $P_{van\Delta R}$  performed better than the control system, however their performance was still slightly below the one of the  $P_{npII}$ . The worst performing promoter was the  $P_{pm}$ , which even with the optimized 5'-UTR region resulted in lower production than the control (Fig. 8). This probably means that in this case our version of the  $P_{pm}$  promoter is weaker than  $P_{npII}$  and shows that for optimal production a combination of strong transcription and translation is critical.

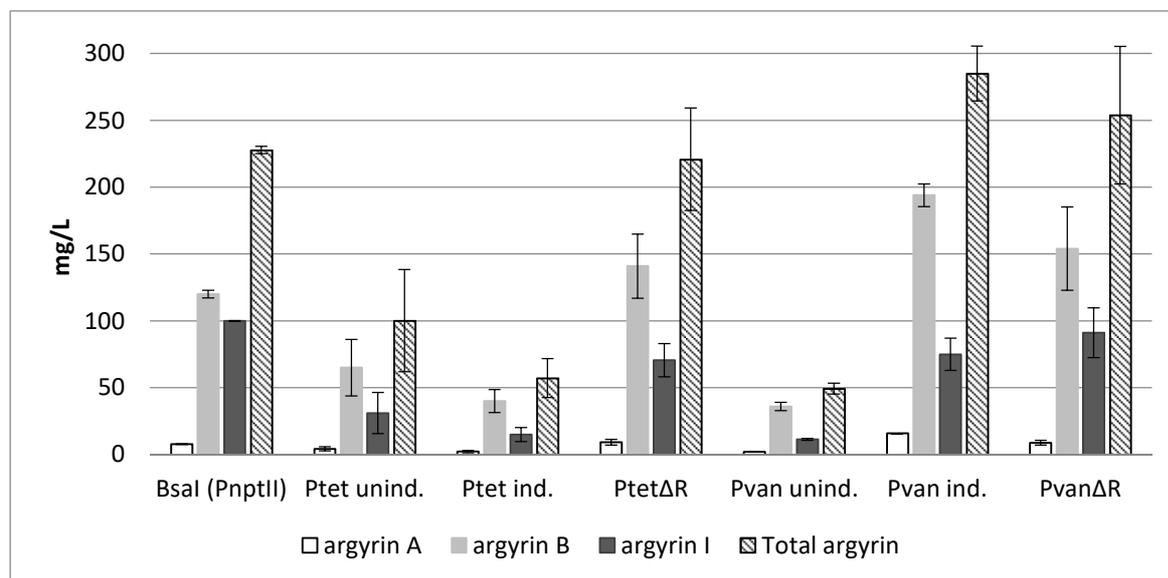


**Figure 8.** Production levels of argyrim I and B evaluated in heterologous argyrim producers with different promoter systems in comparison with control (BsaI). Samples were analyzed by HPLC-MS and levels of argyrim A, B and I were detected by  $[M + H]^+ = 825.314$  (arg A),  $[M + H]^+ = 839.329$  (arg B) and EIC  $[M + H]^+ = 809.318$  (arg I).

To evaluate the effect of the inducers on argyrim production, the cultivation was repeated using mutants harboring argyrim BGC driven by the inducible promoters  $P_{tet}$  and  $P_{van}$ . Performance of the promoters was evaluated under induced and uninduced conditions.  $P_{tet\Delta R}$  and  $P_{van\Delta R}$  promoter versions without repressor were also evaluated without addition of inducers and *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  was used as control. The detailed description of the cultivation is provided in the supplementary information.

In case of  $P_{tet}$  promoter induction with tetracycline did not result in increased production as one would expect. One reason could be that repression of  $P_{tet}$  is not efficient in *M. xanthus*. However, this theory is not very likely, since deletion of repressor in  $P_{tet\Delta R}$  led to significant increase in production (Fig. 9). *M. xanthus* is sensitive to tetracycline, therefore a possible explanation could be, that even addition of small amounts of tetracycline (e.g. 0,7  $\mu\text{g/mL}$ ) challenges the strain, which results in decreased production of secondary metabolites. As expected, in case of  $P_{van}$  promoter induction with vanillate led to near 6-fold increase in

production of argyrin. Surprisingly, deletion of repressor in  $P_{van\Delta R}$  did not lead to further improvement of production levels, but rather resulted in a slight decrease in production, however this could be attributed to the standard deviation (Fig. 9).



**Figure 9.** Production levels of argyrin A, B and I evaluated in heterologous argyrin producers with different promoter systems. The yields detected in mutants harboring argyrin BGC under the control of  $P_{tet}$  and  $P_{van}$  was evaluated under induced and uninduced conditions. Repressor deletion mutants and control (Bsal) were evaluated under uninduced conditions. Samples were analyzed by HPLC-MS and levels of argyrin A, B and I were detected by  $[M + H]^+ = 825.314$  (arg A),  $[M + H]^+ = 839.329$  (arg B) and EIC  $[M + H]^+ = 809.318$  (arg I).

After evaluation of results from both cultivation experiments, we observed a certain degree of variation between both experiments. The reason for this could be minor differences in cultivation conditions (e.g. medium variation, temperature variation, and other external factors). Nevertheless, we can conclude that  $P_{nptII}$  is the current best option for constitutive production of argyrins in *M. xanthus*, with the  $P_{tet\Delta R}$  and  $P_{van\Delta R}$  following closely behind. A similar degree of production can be achieved with inducible system by utilization of  $P_{van}$  induced with vanillate. This information can be used in the future, to increase production of other myxobacterial secondary metabolite systems by promoter exchange. However, it has to be taken into account that in addition to the promoter efficiency, a 5'-UTR region can have a significant effect on protein expression<sup>31</sup> which in turn influences metabolite yield. This is clearly visible from the comparative argyrin production by *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  and *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-PnptII$  (Fig. 8). The argyrin BGC in both strains is driven by the same  $P_{nptII}$  promoter, however the 5'-UTR region in the second strain was optimized using the RBS calculator tool<sup>27,26</sup>, which led to 43% improvement of argyrin production.

### 5.3 Conclusion

In this study we achieved a successful conversion of argyrins A and B to their methylated counterparts, argyrins C and D, by co-expression of the radical SAM Arg1. By utilizing precursor directed biosynthesis, we were able to produce more than 14 novel argyirin derivatives. Their structures were elucidated by NMR, however, the bioactivity testing still has to be performed.

Furthermore, total production of argyirin could be significantly improved by implementation of different promoter systems, leading to a 55% increase in total argyirin production. In addition to argyrins A and B, significant levels of argyirin I were also produced, which could possibly be converted to argyrins A and B in the future by coexpression of additional Arg4/Arg5 copies. The library of evaluated promoters provides a valuable source of information which could be used in the future studies, to optimize heterologous production of other secondary metabolites in *M. xanthus*.

## 5.4 References

- (1) Grossi, P.; Dalla Gasperina, D. Treatment of *Pseudomonas aeruginosa* infection in critically ill patients, *Expert Rev. Anti-infect. Ther.* **2006**, *4*, pp. 639–662.
- (2) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010, *J. Nat. Prod.* **2012**, *75*, pp. 311–335.
- (3) Selva, E.; Gastaldo, L.; Saddler, G. S.; Toppo, G.; Ferrari, P.; Carniti, G.; Goldstein, B. P. Antibiotics A21459 A and B, new inhibitors of bacterial protein synthesis. I. Taxonomy, isolation and characterization, *J. Antibiot.* **1996**, *49*, pp. 145–149.
- (4) Ferrari, P.; Vekey, K.; Galimberti, M.; Gallo, G. G.; Selva, E.; Zerilli, L. F. Antibiotics A21459 A and B, new inhibitors of bacterial protein synthesis. II. Structure elucidation, *J. Antibiot.* **1996**, *49*, pp. 150–154.
- (5) Sasse, F.; Steinmetz, H.; Schupp, T.; Petersen, F.; Memmert, K.; Hofmann, H.; Heusser, C.; Brinkmann, V.; Matt, P. von; Höfle, G.; Reichenbach, H. Argyrins, immunosuppressive cyclic peptides from myxobacteria. I. Production, isolation, physico-chemical and biological properties, *J. Antibiot.* **2002**, *55*, pp. 543–551.
- (6) Jones, A. K.; Woods, A. L.; Takeoka, K. T.; Shen, X.; Wei, J.-R.; Caughlan, R. E.; Dean, C. R. Determinants of Antibacterial Spectrum and Resistance Potential of the Elongation Factor G Inhibitor Argyrin B in Key Gram-Negative Pathogens, *Antimicrob. Agents Chemother.* **2017**, *61*.
- (7) Bielecki, P.; Lukat, P.; Hüsecken, K.; Dötsch, A.; Steinmetz, H.; Hartmann, R. W.; Müller, R.; Häussler, S. Mutation in elongation factor G confers resistance to the antibiotic argyrin in the opportunistic pathogen *Pseudomonas aeruginosa*, *ChemBioChem.* **2012**, *13*, pp. 2339–2345.
- (8) Nyfeler, B.; Hoepfner, D.; Palestrant, D.; Kirby, C. A.; Whitehead, L.; Yu, R.; Deng, G.; Caughlan, R. E.; Woods, A. L.; Jones, A. K.; Barnes, S. W.; Walker, J. R.; Gaulis, S.; Hauy, E.; Brachmann, S. M.; Krastel, P.; Studer, C.; Riedl, R.; Estoppey, D.; Aust, T.; Movva, N. R.; Wang, Z.; Salcius, M.; Michaud, G. A.; McAllister, G.; Murphy, L. O.; Tallarico, J. A.; Wilson, C. J.; Dean, C. R. Identification of elongation factor G as the conserved cellular target of argyrin B, *PLoS ONE.* **2012**, *7*, pp. e42657.
- (9) Nিকেleit, 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.; Malek, N. P. Argyrin A reveals a critical role for the tumor suppressor protein p27kip1 in mediating antitumor activities in response to proteasome inhibition, *Cancer Res.* **2008**, *14*, pp. 23–35.
- (10) Bülow, L.; Nিকেleit, I.; Girbig, A.-K.; Brodmann, T.; Rentsch, A.; Eggert, U.; Sasse, F.; Steinmetz, H.; Frank, R.; Carlomagno, T.; Malek, N. P.; Kalesse, M. Synthesis and biological characterization of argyrin F, *ChemMedChem.* **2010**, *5*, pp. 832–836.
- (11) Ley, S. V.; Priour, A.; Heusser, C. Total synthesis of the cyclic heptapeptide Argyrin B: a new potent inhibitor of T-cell independent antibody formation, *Org. Lett.* **2002**, *4*, pp. 711–714.
- (12) Pogorevc, D.; Tang, Y.; Hoffmann, M.; Zipf, G.; Bernauer, H. S.; Popoff, A.; Steinmetz, H.; Wenzel, S. C. Biosynthesis and heterologous production of argyrins. **2019**, *submitted manuscript*.

- (13) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (14) Zhang, Q.; van der Donk, W. A.; Liu, W. Radical-mediated enzymatic methylation: A tale of two SAMs, *Acc. Chem. Res.* **2012**, *45*, pp. 555–564.
- (15) Pierre, S.; Guillot, A.; Benjdia, A.; Sandstrom, C.; Langella, P.; Berteau, O. Thiostrepton tryptophan methyltransferase expands the chemistry of radical SAM enzymes, *Nat. Chem. Biol.* **2012**, *8*, pp. 957–959.
- (16) Kim, H. J.; McCarty, R. M.; Ogasawara, Y.; Liu, Y.-N.; Mansoorabadi, S. O.; LeVieux, J.; Liu, H.-w. GenK-catalyzed C-6' methylation in the biosynthesis of gentamicin. Isolation and characterization of a cobalamin-dependent radical SAM enzyme, *J. Am. Chem. Soc.* **2013**, *135*, pp. 8093–8096.
- (17) Sucipto, H.; Pogorevc, D.; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, *44*, pp. 160–170.
- (18) Kirschning, A.; Taft, F.; Knobloch, T. Total synthesis approaches to natural product derivatives based on the combination of chemical synthesis and metabolic engineering, *Org. Biomol. Chem.* **2007**, *5*, pp. 3245–3259.
- (19) García-Pindado, J.; Willemse, T.; Goss, R.; Maes, B. U. W.; Giralt, E.; Ballet, S.; Teixidó, M. Bromotryptophans and their incorporation in cyclic and bicyclic privileged peptides, *Biopolymers.* **2018**.
- (20) Haas, D.; Hammann, J. M.; Greiner, R.; Knochel, P. Recent Developments in Negishi Cross-Coupling Reactions, *ACS Catal.* **2016**, *6*, pp. 1540–1552.
- (21) Devendar, P.; Qu, R.-Y.; Kang, W.-M.; He, B.; Yang, G.-F. Palladium-Catalyzed Cross-Coupling Reactions. A Powerful Tool for the Synthesis of Agrochemicals, *J. Agric. Food Chem.* **2018**, *66*, pp. 8914–8934.
- (22) Corr, M. J.; Smith, D. R.; Goss, R. J. One-pot access to 1-5,6-dihalotryptophans and 1-alknyltryptophans using tryptophan synthase, *Tetrahedron.* **2016**, *72*, pp. 7306–7310.
- (23) Ongley, S.; Bian, X.; Neilan, B. A.; Müller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways, *Nat. Prod. Rep.* **2013**, *30*, pp. 1121–1138.
- (24) Iniesta, A. A.; García-Heras, F.; Abellón-Ruiz, J.; Gallego-García, A.; Elías-Arnanz, M. Two systems for conditional gene expression in *Myxococcus xanthus* inducible by isopropyl- $\beta$ -D-thiogalactopyranoside or vanillate, *J. Bacteriol.* **2012**, *194*, pp. 5875–5885.
- (25) Perlova, O.; Fu, J.; Kuhlmann, S.; Krug, D.; Stewart, F.; Zhang, Y.; Müller, R. Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2006**, *72*, pp. 7485–7494.
- (26) Salis, H. M.; Mirsky, E. A.; Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression, *Nat. Biotechnol.* **2009**, *27*, pp. 946–950.
- (27) Espah Borujeni, A.; Channarasappa, A. S.; Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites, *Nucleic Acids Res.* **2014**, *42*, pp. 2646–2659.

- (28) Beck, E.; Ludwig, G.; Auerswald, E. A.; Reiss, B.; Schaller, H. Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5, *Gene*. **1982**, *19*, pp. 327–336.
- (29) Chai, Y.; Shan, S.; Weissman, K. J.; Hu, S.; Zhang, Y.; Müller, R. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis, *Chem. Biol.* **2012**, *19*, pp. 361–371.
- (30) Bakke, I.; Berg, L.; Aune, T. E.; Brautaset, T.; Sletta, H.; Tondervik, A.; Valla, S. Random mutagenesis of the Pm promoter as a powerful strategy for improvement of recombinant gene expression, *Appl. Environ. Microbiol.* **2009**.
- (31) Dvir, S.; Velten, L.; Sharon, E.; Zeevi, D.; Carey, L. B.; Weinberger, A.; Segal, E. Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast, *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E2792-801.

## 5.5 Supplementary information

### Table of contents

|          |   |     |
|----------|---|-----|
| 5.5.1    | Materials and methods .....   | 231 |
| 5.5.1.1  | Construction and engineering of plasmids .....  | 231 |
| 5.5.1.2  | Assembly of <i>argI</i> expression construct.....   | 232 |
| 5.5.1.3  | Transfer and chromosomal integration of expression construct harboring <i>argI</i> into the heterologous host.....                    | 232 |
| 5.5.1.4  | Argyirin production profile engineering by variation of cultivation conditions<br>233   |     |
| 5.5.1.5  | Biotransformation method for production of tryptophan analogs .....   | 234 |
| 5.5.1.6  | Purification and structure elucidation of tryptophan derivatives.....   | 234 |
| 5.5.1.7  | Production and purification of novel argyirin derivatives by precursor-directed biosynthesis .....                                    | 235 |
| 5.5.1.8  | Purification and structure elucidation of novel argyirin derivatives .....  | 236 |
| 5.5.1.9  | Construction of new synthetic promoter systems to replace the $P_{ntpII}$ promoter in heterologous argyirin expression construct..... | 237 |
| 5.5.1.10 | Transfer and chromosomal integration of expression construct harboring argyirin BGC constructs with exchanged promoter system.....    | 237 |
| 5.5.1.11 | Analysis of argyirin production in <i>M. xanthus</i> using different promoter systems for heterologous argyirin production .....      | 238 |
| 5.5.2    | NMR tables and spectra .....  | 243 |
| 5.5.2.1  | Tryptophan NMR data.....  | 243 |
| 5.5.2.2  | Argyirin NMR data.....  | 248 |
| 5.5.3    | References .....  | 272 |

**List of figures**

|  |     |
|--|-----|
| Figure S1. Genotypic verification of integration of all argyrim cluster versions into the heterologous host <i>M. xanthus</i> DK1622 $\Delta$ mchA-tet. .... | 240 |
| Figure S2. Structure of 4-F-L-tryptophan.....  | 243 |
| Figure S3. Structure of 5-Br-L-tryptophan.....   | 243 |
| Figure S4. Structure of 5-Cl-L-tryptophan.....   | 244 |
| Figure S5. Structure of 5-F-L-tryptophan.....  | 244 |
| Figure S6. Structure of 6-F-L-tryptophan.....  | 245 |
| Figure S7. Structure of 6-NO <sub>2</sub> -L-tryptophan.....   | 246 |
| Figure S8. Structure of 7-F-L-tryptophan.....  | 246 |
| Figure S9. Structure of 7-I-L-tryptophan.....  | 247 |
| Figure S10. Structure of 7-NO <sub>2</sub> -L-tryptophan.....  | 247 |
| Figure S11. Structure formula of 5,5'-di-Br-argyrim I.....   | 248 |
| Figure S12. Structure formula of 5,5'-di-Cl-argyrim I.....   | 250 |
| Figure S13. Structure formula of 7,7'-di-F-argyrim I.....  | 251 |
| Figure S14. Structure formula of 7-I-argyrim B.....  | 253 |
| Figure S15. Structure formula of 4-F-argyrim C.....  | 255 |
| Figure S16. Structure formula of 6,6'-di-F-argyrim C2.....   | 256 |
| Figure S17. Structure formula of 6'-F-argyrim C.....   | 258 |
| Figure S18. Structure formula of 5-Cl-argyrim B.....   | 260 |
| Figure S19. Structure formula of 5-Br-argyrim B.....   | 261 |
| Figure S20. Structure formula of 7,7'-di-I-argyrim I.....  | 263 |
| Figure S21. Structure formula of 6'-F-argyrim B.....   | 265 |
| Figure S22. Structure formula of 5'-Cl-argyrim I.....  | 266 |
| Figure S23. Structure formula of 5'-Br-argyrim I.....  | 268 |
| Figure S24. Structure formula of 7,7'-di-I-argyrim B.....  | 270 |

**List of tables**

|   |     |
|---|-----|
| Table S1. Absolute and calculated yields of novel argyrim derivatives.....                                    | 236 |
| Table S2. Heterologous production of PKS/NRPS-derived natural products in <i>M. xanthus</i> .....             | 240 |
| Table S3. List of primers used in this study.....   | 241 |
| Table S4. Gene synthesis constructs obtained from ATG:biosynthetics GmbH.....                                 | 241 |
| Table S5. Constructs generated in this study.....   | 242 |
| Table S6. List of all observed NMR correlations for 4-F-L-tryptophan in D <sub>2</sub> O.....                 | 243 |
| Table S7. List of all observed NMR correlations for 5-Br-L-tryptophan in D <sub>2</sub> O.....                | 243 |
| Table S8. List of all observed NMR correlations for 5-Cl-L-tryptophan in D <sub>2</sub> O.....                | 244 |
| Table S9. List of all observed NMR correlations for 5-F-L-tryptophan in D <sub>2</sub> O.....                 | 245 |
| Table S10. List of all observed NMR correlations for 6-F-L-tryptophan in D <sub>2</sub> O.....                | 245 |
| Table S11. List of all observed NMR correlations for 6-NO <sub>2</sub> -L-tryptophan in D <sub>2</sub> O..... | 246 |
| Table S12. List of all observed NMR correlations for 7-F-L-tryptophan in D <sub>2</sub> O.....                | 246 |
| Table S13. List of all observed NMR correlations for 7-I-L-tryptophan in D <sub>2</sub> O.....                | 247 |
| Table S14. List of all observed NMR correlations for 7-NO <sub>2</sub> -L-tryptophan in D <sub>2</sub> O..... | 247 |
| Table S15. List of all observed NMR correlations for 5,5'-di-Br-argyrim I in CDCl <sub>3</sub> .....          | 248 |
| Table S16. List of all observed NMR correlations for 5,5'-di-Cl-argyrim I in CDCl <sub>3</sub> .....          | 250 |
| Table S17. List of all observed NMR correlations for 7,7'-di-F-argyrim I in CDCl <sub>3</sub> .....           | 252 |
| Table S18. List of all observed NMR correlations for 7-I-argyrim B in CDCl <sub>3</sub> .....                 | 253 |
| Table S19. List of all observed NMR correlations for 4-F-argyrim C in CDCl <sub>3</sub> .....                 | 255 |
| Table S20. List of all observed NMR correlations for 6,6'-di-F-argyrim C2 in CDCl <sub>3</sub> .....          | 257 |
| Table S21. List of all observed NMR correlations for 6'-F-argyrim C in CDCl <sub>3</sub> .....                | 258 |
| Table S22. List of all observed NMR correlations for 5-Cl-argyrim B in CDCl <sub>3</sub> .....                | 260 |

|   |     |
|---|-----|
| Table S23. List of all observed NMR correlations for 5-Br-argyrin B in CDCl <sub>3</sub> . .....      | 262 |
| Table S24. List of all observed NMR correlations for 7,7'-di-I-argyrin I in CDCl <sub>3</sub> . ..... | 263 |
| Table S25. List of all observed NMR correlations for 6'-F-argyrin B in CDCl <sub>3</sub> . .....      | 265 |
| Table S26. List of all observed NMR correlations for 5'-Cl-argyrin I in CDCl <sub>3</sub> . .....     | 267 |
| Table S27. List of all observed NMR correlations for 5'-Br-argyrin I in CDCl <sub>3</sub> . .....     | 268 |
| Table S28. List of all observed NMR correlations for 7,7'-di-I-argyrin B in CDCl <sub>3</sub> . ..... | 270 |

## 5.5.1 Materials and methods

### 5.5.1.1 Construction and engineering of plasmids

Routine handling of nucleic acids, such as isolation of plasmid DNA, restriction endonuclease digestions, DNA ligations, and other DNA manipulations, was performed according to standard protocols<sup>1</sup>. *E. coli* HS996 (Invitrogen) was used as host for standard cloning experiments. *E. coli* strains were cultured in LB medium or on LB agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, (1.5% agar)) at 30–37 °C (and 200 rpm) overnight. Antibiotics were used at the following final concentrations: 100 µg/mL ampicillin, 50 µg/mL kanamycin, 12 µg/mL tetracycline. Transformation of *E. coli* strains was achieved via electroporation in 0.1 cm-wide cuvettes at 1250 V, a resistance of 200 Ω, and a capacitance of 25 µF. Plasmid DNA was either purified by standard alkaline lysis<sup>1</sup> or by using the GeneJet Plasmid Miniprep Kit (Thermo Fisher Scientific) or the NucleoBond PC100 kit (Machery Nagel). Restriction endonucleases, alkaline phosphatase (FastAP) and T4 DNA ligase were obtained from Thermo Fisher Scientific. Oligonucleotides used for PCR and sequencing were obtained from Sigma-Aldrich and are listed in Table S3. PCR reactions were carried out in a Mastercycler® pro (Eppendorf) using Phusion™ High-Fidelity, Taq DNA polymerase (Thermo Fisher Scientific) or Dream Taq DNA polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. For Taq and Dream Taq: Initial denaturation (5 min, 95 °C); 30 cycles of denaturation (30 s, 95 °C), annealing (30 s, 53 - 64 °C) and elongation (varied based on PCR product length 1 kb/min, 72 °C); and final extension (10 min, 72 °C). For Phusion™: Initial denaturation (30 s, 98 °C); 30 cycles of denaturation (20 s, 98 °C), annealing (25 s, 53 - 64 °C) and elongation (varied based on PCR product length 0.5 kb/ min, 72 °C); and final extension (10 min, 72 °C). PCR products or DNA fragments from restriction digests were purified by agarose gel electrophoresis and isolated using the NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) or peqGold Gel Extraction (Peqlab). Red/ET recombineering experiments for plasmid modifications<sup>2</sup> using the strains *E. coli* HS996/pSC101-BAD-gbaA-tet or *E. coli* GB05-red were performed according to the manufacturers protocol (Gene Bridges GmbH). After selection with suitable antibiotics, clones harboring correct

recombination products were identified by plasmid isolation and restriction analysis with a set of different endonucleases. Synthetic DNA fragments were obtained from ATG:biosynthetics GmbH and delivered in standard cloning vectors (Table S4). Details on the construction of plasmids generated in this study are given in Table S5.

### 5.5.1.2 Assembly of *arg1* expression construct

Synthetic fragment zeo-Hom-MCS was ordered from a gene synthesis company and supplied in a standard pGH expression vector. Primers dpo-pUC18-F and dpo-pUC18-R were used to amplify 731 bp pUC18 ORI fragment from the pUC18 template plasmid using Phusion polymerase. The pUC18 ORI PCR fragment was ligated into pGH-zeo-Hom-MCS plasmid using *PmeI* and *SpeI* restriction endonucleases, resulting in pUC18-zeo-Hom-MCS expression vector. Primers dpo-arg1F and dpo-arg1R were used to amplify 238 bp *arg1* extension fragment from *Cystobacter sp.* SBCb004 genomic DNA using Phusion polymerase. The fragment was ligated into the pGH-arg1-V1 plasmid, obtained from a gene synthesis company, using *NdeI* and *EcoRI* endonucleases, to construct pGH-arg1B plasmid harboring reconstituted *arg1* sequence. The 2047 bp arg1B fragment was released from pGH-arg1B plasmid and inserted into previously constructed pUC18-zeo-Hom-MCS expression vector by traditional restriction-digestion and ligation using *NdeI* and *BglII* restriction endonucleases, yielding pUC18-zeo-Hom-arg1B expression construct (Table S5).

### 5.5.1.3 Transfer and chromosomal integration of expression construct harboring *arg1* into the heterologous host

According to a previously established electroporation procedure for *Myxococcus xanthus* DK1622 (Kashefi and Hartzell, 1995) the host strain *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI^3$  was transformed with the generated expression construct pUC18-zeo-Hom-arg1B (Table S5). *M. xanthus* DK1622 mutants were routinely cultivated at 30 °C in CTT medium or on CTT agar (1% casitone, 10 mM Tris buffer pH 7.6, 1 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.6, 8 mM MgSO<sub>4</sub> (1.5% agar) with final pH 7.6). For liquid cultures, the strains were grown in Erlenmeyer flasks on an orbital shaker at 180 rpm for 3–6 days. For selection of *M. xanthus* mutants 50 µg/mL kanamycin and 50 µg/mL zeocin were used. Correct chromosomal integration of the expression constructs via homologous recombination into the downstream locus of the previously integrated argyrin BGC<sup>3</sup> was verified by PCR. Transformant cells were lysed by incubating at 95 °C for 30 min prior of being added to the PCR reaction

(‘colony PCR’). For each expression construct correct chromosomal integration was confirmed using two primer combinations, revealing PCR products of the expected sizes: dpo-arg1B-1/dpo-arg1B-2 (1261 bp) and dpo-arg1B-5/dpo-arg1B-4 (1401 bp). Genomic DNA of *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  was used as negative control. Complementary experiment using primers dpo-arg1B-1/dpo-arg1B-4 revealed a 1112 bp PCR product for *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ , but not for the expression strain. PCR reactions were performed according to the conditions described in Materials and methods, chapter 5.5.1.1, using Taq polymerase; for primer sequences see Table S3.

#### 5.5.1.4 Argyrin production profile engineering by variation of cultivation conditions

To vary the production between different argyrin derivatives (e.g. arg A, arg B, arg C and arg D) strains *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  and *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-arg1B$  were cultivated in CTT and M7/s4 medium (0.5 % soy flour, 0.5 % corn starch, 0.2 % glucose, 0.1 % yeast extract, 0.1 % MgSO<sub>4</sub> x 7 H<sub>2</sub>O, 0.1 % CaCl<sub>2</sub> x 2 H<sub>2</sub>O, 1 % HEPES, with final pH 7.4 and supplemented with 0.1 mg/L of vitamin B12 and 5 mg/L of FeCl<sub>3</sub> after autoclaving) with and without supplementation of  $\alpha$ -aminobutyric acid. Strains were inoculated from cryo stocks and grown on CTT agar plates supplemented with kanamycin 50  $\mu$ g/mL (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ ) or with kanamycin 50  $\mu$ g/mL and zeocin 50  $\mu$ g/mL (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-arg1B$ ) for several days until plates were mostly overgrown with cells. All of the cells were scraped from the plates to inoculate preculture medium (50 mL medium in 300 mL Erlenmeyer flask), which was cultivated at 30 °C, 180 rpm for 48 h. Five mL of well grown preculture was used to inoculate 50 mL production medium in which the strain was grown at the same conditions for 6 days. All cultures were supplemented with suitable antibiotics as mentioned above as well as with 4 % Amberlite XAD-16 resin. All cultures were cultivated in triplicates for 6-8 days before harvesting by centrifugation at 8000 rpm for 15 min. Supernatant was removed and pelleted cells with Amberlite XAD-16 resin were extracted with 50 mL of ethyl acetate. Extracts were filtered and dried in round bottom flasks on rotary evaporator. Crude extracts were dissolved in 1 mL of methanol and subjected to HPLC-MS analysis.

### 5.5.1.5 Biotransformation method for production of tryptophan analogs

By utilizing readily available enzymatic biotransformation method<sup>4</sup>, a library of tryptophane derivatives was generated from several commercially available indoles. *Escherichia coli* pre-transformed with pSTB7, a high copy number plasmid encoding genes for tryptophan synthase from *Salmonella enterica*, was commercially obtained (ATCC 37845)<sup>5,6</sup>. Plasmid pSTB7 was isolated from overnight *E. coli* culture and retransformed into *E. coli* BL21 (DE3) using standard electroporation conditions (see Materials and methods, chapter 5.5.1.1). Five mL of *E. coli* BL21 (DE3)::pSTB7 culture were used to inoculate 500 mL of LB medium in 2 L Erlenmeyer flask and cultivated for 20-24 h at 37 °C. Fully grown culture was then centrifuged at 3500 g at 4 °C for 20 minutes. The pellet was washed twice with 80 mL of saturated salt solution and re-suspended in 80 mL of the lysis buffer (500 mM Tris-HCl, 10 mM  $\beta$ -mercaptoethanol, 5 mM EDTA and 3 mM pyridoxal phosphate (PLP)), adjusted to pH of 7.8. Cells were lysed by sonication and lysate was clarified by centrifugation at 20000 g, 4 °C for 20 minutes, before being aliquoted in 1.5 mL Eppendorf tubes, flash frozen in liquid nitrogen and stored at – 80 °C until further use.

The biotransformation reaction was prepared in 100 mL of 1M KH<sub>2</sub>PO<sub>4</sub> buffer (pH 7.8) with 20% DMSO in 1 L Erlenmeyer flask. DMSO improves indole solubility and 20% content was found to significantly increase yield. Three mL of *E. coli* BL21 (DE3)::pSTB7 cell lysate, containing tryptophan synthase, were thawed at the room temperature and transferred into a 3 mL dialyzer (D-Tube™ Dializer Maxi, MWCO 6-8 kDa from Novagen®, Merck Millipore) which was submerged into the reaction buffer. Substrates; serine (2.5 mmol) and desired indole derivative (2 mmol) were added to the reaction buffer and the flask was incubated at 37 °C for 48-72 hours. After 24h dialysis bags were withdrawn and filled with fresh cell lysate to increase reaction efficiency.

After three days of incubation at 37 °C, reaction was transferred into a separation funnel and extracted with equal volume of ethyl acetate. Water phase was frozen and freeze dried, to remove remaining DMSO, before purification on Biotage Isolera One.

Commercially available indoles were obtained from Alfa Aesar (4-F, 6-F, 7-F, 5-Br, 5-Cl, 6-NO<sub>2</sub> and 7-NO<sub>2</sub> indoles) and Sigma-Aldrich (7-I indole).

### 5.5.1.6 Purification and structure elucidation of tryptophan derivatives

Samples were redissolved in H<sub>2</sub>O before purification on a Biotage Isolera One using reverse-phase Biotage® SNAP Cartridge KP-C18-HS 30 g column cartridge. Water with 0.1% formic

acid (A) was used to remove PLP, serine and other impurities and methanol with 0.1% formic acid (B) was used to elute novel tryptophan derivatives. The purification was carried out using the following gradient: 25 mL/min flow, equilibration with 5/95% B/A over 3 column volumes, 5/95% B/A over 3 column volumes, increase to 100%B over 8 column volumes and keeping 100% B over 3 column volumes. The collection wavelength was set to UV1+UV2 (254, 280 nm). Collection of the peaks occurring from approx. 4.3 min to 12.0 min gave the resulting tryptophan analogs. Fractions containing the product were combined in a round bottom flask and dried on a rotary evaporator. Dry extract was treated twice with 20 mL of 0.1 M hydrochloric acid and dried on rotary evaporator, to form hydrochloric salt. Product was re-dissolved in 3 - 4 mL of H<sub>2</sub>O, frozen at -80 °C and dried overnight by freeze drying, to produce powdered product. Ethyl acetate phase, containing unreacted indole, was dried on rotary evaporator and recovered indole could be recycled for further reactions.

#### **5.5.1.7 Production and purification of novel argyirin derivatives by precursor-directed biosynthesis**

Two heterologous argyirin producer mutants were used to produce new argyirin derivatives. *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$  was used for production of argyirin B based derivatives and *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-arg1B$  was used to produce argyirin C based derivatives. Both strains were cultivated on CTT agar plates supplemented with kanamycin 50 µg/mL or kanamycin 50 µg/mL and zeocin 50 µg/mL, respectively. Once the plates were completely covered with cell mass, an entire plate was scraped and used to inoculate 50 mL of M7/s4 medium supplemented with suitable antibiotics. Seed culture was grown for 2 days after which 5 mL were used to inoculate main cultures, supplemented with suitable antibiotics and 2% Amberlite XAD-16 resin. In case of *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-arg1B$ , the medium was supplemented with additional 1 mg/L of vitamin B<sub>12</sub>, to improve methylation efficiency. Tryptophan derivatives were dissolved in H<sub>2</sub>O and fed to the cultures 2 times per day for 6 days, until final concentration of 1 mM was reached. Additionally  $\alpha$ -aminobutyric acid (10 mM) was added to each culture flask to help decrease the number of derivatives by shifting the production towards argyirin B. After 6 days, Amberlite XAD-16 resin and cells were harvested by centrifugation at 4 °C for 15 min. The Supernatant was removed and the pellets were frozen and freeze dried. Dry pellets were extracted by ethyl acetate via standard extraction procedure (2x with 50 mL of ethyl acetate). Crude extracts were dissolved in a mixture of acetonitrile and DMSO, before purification on HPLC (see Material and methods chapter 5.5.1.8).

### 5.5.1.8 Purification and structure elucidation of novel argyrim derivatives

First round of LC purification was performed on the Luna<sup>®</sup> 5 $\mu$ m C18 (2) 100 Å, LC (250 x 10 mm) semi prep HPLC column, for all samples. Separation of a sample was achieved by a linear gradient from (A) H<sub>2</sub>O + 0.1 % FA to (B) ACN + 0.1 % FA. The gradient was initiated by a 4 min isocratic step at 5 % B, followed by an increase to 30 % B in 6 min, to 70 % B in 20 min and to 95 % B in 0.5 min. After a 4 min step at 95 % B the system was re-equilibrated to the initial conditions of 5 % B, which were kept for 4 min before next injection. Column oven temperature was set to 45°C and separation was done at a flow rate of 5 mL/min.

Additional purification steps were necessary to separate minor analogs containing 6-F, 7-F, 5-Br, 7-NO<sub>2</sub> and 5-Cl tryptophan residues. Second round of purification was performed on the Kinetex<sup>®</sup> 5  $\mu$ m Biphenyl 100 Å LC column 250 x 10.0 mm. Separation of a sample was achieved by a linear gradient from (A) H<sub>2</sub>O + 0.1 % FA to (B) ACN + 0.1 % FA. The gradient was initiated by a 4 min isocratic step at 5 % B, followed by an increase to 45 % B in 11 min, to 47 % B in 6 min and to 95 % B in 1 min. After a 2.5 min step at 95 % B the system was re-equilibrated to the initial conditions of 5 % B, which were kept for 2 min before next injection. Column oven temperature was set to 45°C and separation was done at a flow rate of 5 mL/min.

Yields of purified argyrim derivatives are provided in table S1.

**Table S1.** Absolute and calculated yields of novel argyrim derivatives. The yields were achieved after purification from 150 mL of production broth.

| Name                 | Mass    | Absolute yield <sup>a</sup> [mg] | Calculated yield [mg/L] |
|----------------------|---------|----------------------------------|-------------------------|
| 4-F-argyrim C        | 856.31  | 1.00                             | 6.67                    |
| 6'-F-argyrim B       | 856.31  | 0.87                             | 5.80                    |
| 6'-F-argyrim C       | 856.31  | 0.86                             | 5.73                    |
| 6,6'-di-F-argyrim C2 | 844.29  | 1.45                             | 9.60                    |
| 7,7'-di-F-argyrim I  | 844.29  | 3.76                             | 25.07                   |
| 5-Cl-argyrim B       | 872.28  | 0.58                             | 3.86                    |
| 5'-Cl-argyrim I      | 842.27  | 0.73                             | 4.87                    |
| 5,5'-di-Cl-argyrim I | 876.23  | 4.07                             | 27.13                   |
| 5-Br-argyrim B       | 916.23  | 0.97                             | 6.47                    |
| 5'-Br-argyrim I      | 886.22  | 0.73                             | 4.87                    |
| 5,5'-di-Br-argyrim I | 964.13  | 3.30                             | 22.00                   |
| 7-I-argyrim B        | 964.22  | 2.21                             | 14.73                   |
| 7,7'-di-I-argyrim B  | 1090.11 | 0.18                             | 1.20                    |
| 7,7'-di-I-argyrim I  | 1060.10 | 0.52                             | 3.47                    |

<sup>a</sup> Yield after HPLC purification from 150 mL production broth.

\* Production levels of natural argyrim derivatives achieved under the same conditions are 40 mg/L (argyrim B) and 16 mg/L (argyrim C).

### 5.5.1.9 Construction of new synthetic promoter systems to replace the $P_{nptII}$ promoter in heterologous argyrimin expression construct

To evaluate other promoter systems for production of argyrins, 6 new promoter constructs, based on 4 different promoter systems were designed ( $P_{nptII}$ ,  $P_{van}$ ,  $P_{pm}$ ,  $P_{tet}$ ). Six synthetic promoter fragments were ordered from a gene synthesis company and supplied in a standard pGH expression vector. The promoter fragments with lengths of 233 bp ( $P_{pm}$ ), 803 bp ( $P_{tet}$ ), 169 bp ( $P_{tet\Delta R}$ ), 245 bp ( $P_{nptII}$ ), 1145 bp ( $P_{van}$ ) and 200 bp ( $P_{van\Delta R}$ ) were released from the pGH vector and ligated into pArg2345-V1-BsaI expression vector by *SwaI*, *StuI* restriction endonucleases, resulting in construction of the following argyrimin BGC expression constructs: pArg2345-V1-BsaI-Ppm, pArg2345-V1-BsaI-Ptet, pArg2345-V1-BsaI-Ptet $\Delta R$ , pArg2345-V1-BsaI-PnptII, pArg2345-V1-BsaI-Pvan and pArg2345-V1-BsaI-Pvan $\Delta R$ .

### 5.5.1.10 Transfer and chromosomal integration of expression construct harboring argyrimin BGC constructs with exchanged promoter system

The host strain *M. xanthus* DK1622  $\Delta mchA-tet^3$  was transformed with the generated expression constructs pArg2345-V1-BsaI-Ppm, pArg2345-V1-BsaI-Ptet, pArg2345-V1-BsaI-Ptet $\Delta R$ , pArg2345-V1-BsaI-PnptII, pArg2345-V1-BsaI-Pvan and pArg2345-V1-BsaI-Pvan $\Delta R$  (Table S5). Cultivation conditions were kept the same as described in Materials and methods, chapter 5.5.1.4. For selection of *M. xanthus* mutants 50  $\mu\text{g/mL}$  kanamycin was used. Correct chromosomal integration of the expression constructs via homologous recombination into the  $tet^R$  locus was verified by PCR. Transformant cells were lysed by incubating at 95 °C for 30 min prior of addition to the PCR reaction ('colony PCR'). For each expression construct correct chromosomal integration was confirmed using two primer combinations, revealing PCR products of the expected sizes: P1/P2 (1458 bp) and P3/P4 (1461 bp). Genomic DNA of *M. xanthus* DK1622  $\Delta mchA-tet$  was used as negative control. Complementary experiment using primers P1/P4 revealed a 1461 bp PCR product for *M. xanthus* DK1622  $\Delta mchA-tet$ , but not for the expression strains. PCR reactions were performed according to the conditions described in Materials and methods, chapter 5.5.1.1 using Taq polymerase. For primer sequences see Table S3.

### 5.5.1.11 Analysis of argyirin production in *M. xanthus* using different promoter systems for heterologous argyirin production

M7/s4 medium was used for parallel cultivation of the heterologous producers with replaced promoter (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-Ppm$ , *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-Ptet$ , *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-Ptet\Delta R$ , *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-PnptII$ , *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-Pvan$  and *M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI-Pvan\Delta R$ ) and control strain (*M. xanthus* DK1622  $\Delta mchA-tet::pArg2345-V1-BsaI$ ), to evaluate the effect of various promoters on the production profile of argyirin. Strains were cultivated on CTT agar plates supplemented with kanamycin 50  $\mu\text{g}/\text{mL}$  for several days, until the plates were completely covered with cell mass. An entire plate was scraped and used to inoculate 50 mL of M7/s4 medium supplemented with kanamycin 50  $\mu\text{g}/\text{mL}$ . Solution of amino acid mixture calculated for the final concentration of 5 mM serine, 5 mM cysteine, 5 mM alanine, 10 mM tryptophan, 10 mM glycine and 10 mM  $\alpha$ -aminobutyric acid was fed to the production cultures twice daily over the course of 6 days. Seed culture was grown for 2 days after which 5 mL were used to inoculate the main cultures, supplemented with 50  $\mu\text{g}/\text{mL}$  kanamycin and 4% Amberlite XAD-16 resin. In case of inducible  $P_{tet}$  or  $P_{van}$  promoters corresponding cultures were supplemented with 0,7  $\mu\text{g}/\text{mL}$  tetracycline or 1 mM vanillate, respectively. All mutants were cultivated in triplicates for 6 days before harvesting. Cultures were harvested by centrifugation at 8000 rpm for 15 min. Supernatant was removed and pelleted cells with Amberlite XAD-16 resin were extracted with 50 mL of ethyl acetate. Extracts were filtered and dried in round bottom flasks on rotary evaporator. Crude extracts were dissolved in 1 mL of methanol and subjected to HPLC-MS analysis.

Argyirin samples were routinely analyzed on a Dionex Ultimate 3000 RSLC system using a Waters BEH  $C_{18}$  column (50 x 2.1 mm, 1.7  $\mu\text{m}$ ) equipped with a Waters VanGuard BEH  $C_{18}$  1.7  $\mu\text{m}$  guard column. Separation of 1  $\mu\text{L}$  sample (Crude extract 1/5 dilutions in methanol) was achieved by a linear gradient from (A)  $\text{H}_2\text{O}$  + 0.1% FA to (B) ACN + 0.1% FA at a flow rate of 600  $\mu\text{L}/\text{min}$  and a column temperature of 45  $^\circ\text{C}$ . Gradient conditions were as follows: 0 – 0.5 min, 5% B; 0.5 – 18.5 min, 5 – 95% B; 18.5 – 20.5 min, 95% B; 20.5 – 21 min, 95 – 5% B; 21–22.5 min, 5% B. UV spectra were recorded by a DAD in the range from 200 to 600 nm. The LC flow was split to 75  $\mu\text{L}/\text{min}$  before entering the Bruker Daltonics maXis 4G hr-qToF mass spectrometer using the Apollo II ESI source. Mass spectra were acquired in centroid mode ranging from 150 – 2500  $m/z$  at a 2 Hz full scan rate. Mass spectrometry source parameters were set to 500V as end plate offset;

4000 V as capillary voltage; nebulizer gas pressure 1 bar; dry gas flow of 5 L/min and a dry temperature of 200°C. Ion transfer and quadrupole settings were set to Funnel RF 350 Vpp; Multipole RF 400 Vpp as transfer settings and Ion energy of 5eV as well as a low mass cut of 300 m/z as Quadrupole settings. Collision cell was set to 5.0 eV and pre pulse storage time was set to 5  $\mu$ s. Spectra acquisition rate was set to 2Hz. Calibration of the maXis4G qTOF spectrometer was achieved with sodium formate clusters before every injection to avoid mass drifts. All MS analyses were acquired in the presence of the lock masses  $C_{12}H_{19}F_{12}N_3O_6P_3$ ,  $C_{18}H_{19}O_6N_3P_3F_2$  and  $C_{24}H_{19}F_{36}N_3O_6P_3$  which generate the  $[M+H]^+$  Ions of 622.028960, 922.009798 and 1221.990638. The corresponding MS<sup>2</sup> method operating in automatic precursor selection mode picks up the two most intense precursors per cycle, applies smart exclusion after five spectra and performs CID and MS/MS spectra acquisition time ramping. CID Energy was ramped from 35 eV for 500 m/z to 45 eV for 1000 m/z and 60 eV for 2000 m/z. MS full scan acquisition rate was set to 2Hz and MS/MS spectra acquisition rates were ramped from one to four Hz for precursor Ion intensities of 10kcts to 1000kcts.

To quantify argyirin production levels described in chapter 5.2.3, the following HPLC-MS method was applied. Separation was performed on a Dionex UltiMate 3000 RSLC system equipped with a Waters reversed phase UPLC column (Acquity UPLC BEH C18 1,7 $\mu$ m; 2.1\*100mm) using a linear gradient with solvent A (water + 0.1% formic acid) and B (acetonitrile + 0.1% formic acid) at a flow rate of 600  $\mu$ l/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step with 5% B followed by an increase to 15% B within 1 min, 50% B within 11.5 min and 95% B within 1 min, which was kept for 1min, before decreasing back to initial conditions of 5% B within 0.3 min and was kept for 1.7 min. HPLC was coupled to a Bruker Daltonics ion trap mass spec 'Amazon speed' system. Mass spectra were acquired in positive ionization mode with a range of 200-2500 m/z at a resolution of R=30000. Identities of the argyrins were confirmed by comparing with both retention time and MS<sup>2</sup> fragmentation pattern of the authentic compounds.

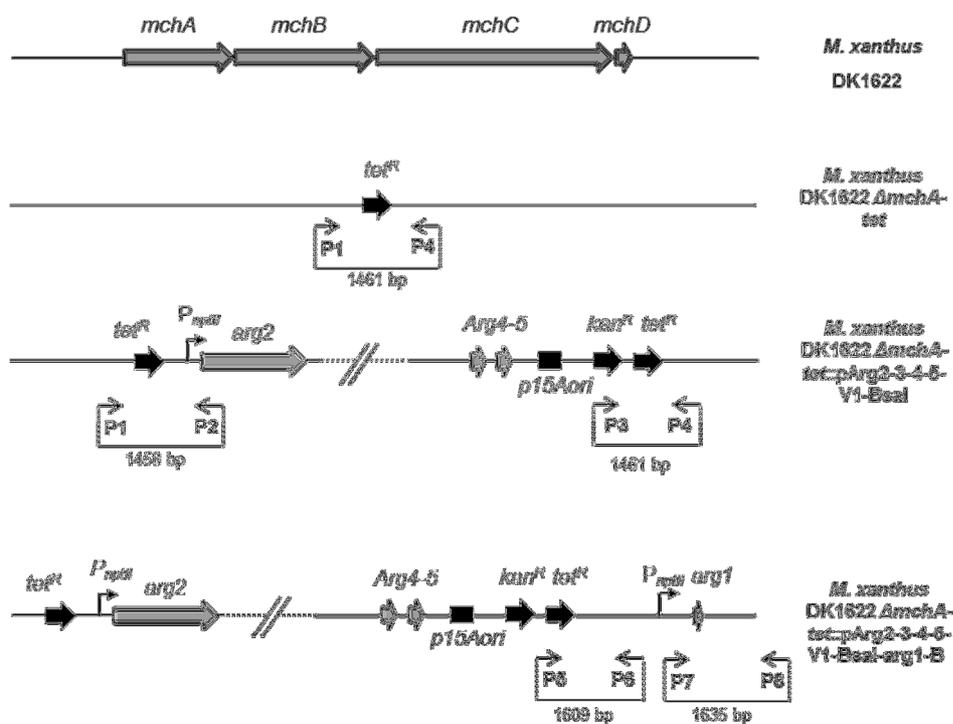
Quantification was performed using Bruker Daltonics Quant analysis software version 4.4. The following method was used to quantify the corresponding derivatives; argyirin I (retention time 9.83, MS<sup>2</sup> fragments: 707.2, 724.2), argyirin A (retention time 10.05, MS<sup>2</sup> fragments: 737.3; 754.3), argyirin B (retention time 10.60, MS<sup>2</sup> fragments: 737.3; 754.3), argyirin C (retention time 10.71, MS<sup>2</sup> fragments: 751.3; 768.3) and argyirin D (retention time 11.28, MS<sup>2</sup> fragments: 751.3; 768.3).

**Table S2.** Heterologous production of PKS/NRPS-derived natural products in *M. xanthus*.

| Compound         | Native producer                   | Pathway-type [size] | Promoter          | Yield ‡        | Ref              |
|------------------|-----------------------------------|---------------------|-------------------|----------------|------------------|
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{nptII}$       | 160 mg/L       | <sup>3</sup>     |
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{nptII}$       | 250 mg/L       | This study       |
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{van}$         | 250 mg/L       | This study       |
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{van\Delta R}$ | 200 mg/L       | This study       |
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{tet}$         | 50 mg/L        | This study       |
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{tet\Delta R}$ | 150 mg/L       | This study       |
| Argyrin A, B     | <i>Cystobacter species</i>        | NRPS [33 kb]        | $P_{pm}$          | 120 mg/L       | This study       |
| Bengamide        | <i>Myxococcus virescens</i>       | PKS/NRPS [25 kb]    | $P_{nptII}$       | >10 mg/L       | <sup>7</sup>     |
| Corallopyronin A | <i>Corallococcus coralloides</i>  | PKS/NRPS [65 kb]    | $P_{nptII}$       | 37 mg/L        | <sup>8</sup>     |
| Corallopyronin A | <i>Corallococcus coralloides</i>  | PKS/NRPS [65 kb]    | $P_{van}$         | 100 mg/L       | <sup>9</sup>     |
| Dawenol          | <i>Stigmatella aurantiaca</i>     | PKS [21 kb]         | native            | n.d.           | <sup>10</sup>    |
| Epothilone       | <i>Sorangium cellulosum</i>       | PKS/NRPS [56 kb]    | native            | 0.1 - 0.4 mg/L | <sup>11-13</sup> |
| Myxochromide A   | <i>Myxococcus xanthus</i>         | PKS/NRPS [29 kb]    | $P_{nptII}$       | ~500 mg/L      | <sup>14</sup>    |
| Myxochromide S   | <i>Stigmatella aurantiaca</i>     | PKS/NRPS [29 kb]    | $P_{nptII}$       | >500 mg/L      | <sup>15</sup>    |
| Myxopyronin A    | <i>Myxococcus fulvus</i>          | PKS/NRPS [53 kb]    | $P_{nptII}$       | 156 mg/L       | <sup>8</sup>     |
| Myxothiazol      | <i>Stigmatella aurantiaca</i>     | PKS/NRPS [57 kb]    | $P_{pm}$          | 20 mg/L        | <sup>16</sup>    |
| Oxytetracycline  | <i>Streptomyces rimosus</i>       | PKS [32 kb]         | native            | 10 mg/L        | <sup>17</sup>    |
| Pretubulysin     | <i>Cystobacter species</i>        | PKS/NRPS [40 kb]    | $P_{tet}$         | 0.2 mg/L       | <sup>18</sup>    |
| PUFAs            | <i>Aetherobacter fasciculatus</i> | PKS/FAS [18 kb]     | $P_{tet}$         | ~1 mg/CDW      | <sup>19</sup>    |
| Vioprolide B, D  | <i>Cystobacter violaceus</i>      | NRPS [56 kb]        | $P_{tet}$         | 80 mg/L        | <sup>20</sup>    |

‡ based on shake flask experiments

n.d., not determined

**Figure S1.** Genotypic verification of integration of all argyrin cluster versions into the heterologous host *M. xanthus* DK1622  $\Delta mchA$ -*tet*. Primer sets used for genotypic verification, their binding sites and amplicon sizes are illustrated. Nucleotide sequences of the primers used are listed in the Table S3.

**Table S3.** List of primers used in this study

| Primer name     | Sequence (5'→3')<br>(restriction sites in bold, homology arms underlined) | Restriction sites           |
|-----------------|---|-----------------------------|
| dpo-pUC18F      | <b>GACGTTTAAACTGACCACTTAAGCATGTGAGCAAAAGG</b><br>CCAGCA                   | <i>PmeI</i> , <i>AflIII</i> |
| dpo-pUC18R      | GTGAAACTAGTTCCGTGGCAAAGCAAATATTACCCCGTA<br>GAAAAGATCAAAGG                 | <i>SspI</i> , <i>SpeI</i>   |
| dpo-arg1F       | CATCATATGAAGGAGGCTTTTCTGATGGGCCAGAC                                       | <i>NdeI</i>                 |
| dpo-arg1R       | GTACATGAATTCCCACCATCCGGGGCCGATG   | <i>EcoRI</i>                |
| dpo-arg1B-1     | GAGAACTGTGAATGCGCAAAC   |                             |
| dpo-arg1B-2     | GGAAGATTGGATGCGTTCAG  |                             |
| dpo-arg1B-4     | GACTTTCGACGATCTTCAGCTTC   |                             |
| dpo-arg1B-5     | GACGACGTGACCCTGTTTCATC  |                             |
| dpo-arg2-prom-F | CTCGCTAACGGATTCACCAC  |                             |
| dpo-arg2-prom-R | GTTCACTCCTGGCGACAAG   |                             |
| InMchP1         | CGAGCAATCCGCTATTGGC   |                             |
| InMchP2         | CAGCTGGCAATTCCGGTTCG  |                             |
| InMchP3         | ACGGGACGGGATATCTGACC  |                             |
| InMchP4         | CTGTGTCTTCTGCGACGC  |                             |

**Table S4.** Gene synthesis constructs obtained from ATG:biotechnics GmbH.

| Plasmid name                 | Synthetic fragment name | Design   | Synthetic fragment size (bp) |
|------------------------------|-------------------------|--|------------------------------|
| pGH-arg1-V1                  | Seq1-arg1-V1            | <i>NdeI</i> -arg1- <i>AvrII</i> - <i>BglIII</i>  | 1945                         |
| pGH-zeo-Hom-MCS              | zeo-Hom-MCS             | <i>SspI</i> -zeo <sup>R</sup> -Hom-P <sub>nptII</sub> -MCS- <i>AflIII</i>                | 2007                         |
| pGH-Ppm-synRBS-arg2          | Ppm-synRBS-arg2         | <i>SwaI</i> - <i>PciI</i> - <i>NdeI</i> -P <sub>pm</sub> -synRBS-arg2- <i>StuI</i>       | 240                          |
| pGH-Ptet-synRBS-arg2         | Ptet-synRBS-arg2        | <i>SwaI</i> - <i>PciI</i> -TetR- <i>NdeI</i> -P <sub>tet</sub> -synRBS-arg2- <i>StuI</i> | 810                          |
| pGH-Ptet-wo_tetR-synRBS-arg2 | PtetΔR-synRBS-arg2      | <i>SwaI</i> - <i>PciI</i> - <i>NdeI</i> -P <sub>tet</sub> -synRBS-arg2- <i>StuI</i>      | 187                          |
| pGH-nptII-synRBS-arg2        | nptII-synRBS-arg2       | <i>SwaI</i> - <i>PciI</i> - <i>NdeI</i> -P <sub>nptII</sub> -synRBS-arg2- <i>StuI</i>    | 263                          |
| pGH-Pvan-synRBS-arg2         | Pvan-synRBS-arg2        | <i>SwaI</i> - <i>PciI</i> -VanR- <i>NdeI</i> -P <sub>van</sub> -synRBS-arg2- <i>StuI</i> | 1163                         |
| pGH-Pvan-wo_vanR-synRBS-arg2 | PvanΔR-synRBS-arg2      | <i>SwaI</i> - <i>PciI</i> - <i>NdeI</i> -P <sub>van</sub> -synRBS-arg2- <i>StuI</i>      | 218                          |

**Table S5.** Constructs generated in this study.

| Plasmid name            | Construction details  | Features   |
|-------------------------|---|--|
| pGH-arg1B               | 238 bp <i>NdeI-arg1-EcoRI</i> PCR fragment ligated into pGH-arg1-V1 hydrolyzed with the same enzymes.   | Reconstituted <i>arg1</i> gene, pUC ori, <i>amp<sup>R</sup></i>                                      |
| pUC18-zeo-Hom-MCS       | 731 bp <i>PmeI-pUC18-ORI-SpeI</i> PCR fragment ligated into pGH-zeo-Hom-MCS plasmid hydrolyzed with the same enzymes.   | pUC ori, <i>zeo<sup>R</sup></i> , homology region, MCS   |
| pUC18-zeo-Hom-arg1B     | 2047 bp <i>NdeI-arg1B-BglIII</i> fragment derived from pGH-arg1B plasmid ligated into pUC18-zeo-Hom-MCS expression vector hydrolyzed with the same enzymes.                     | pUC ori, <i>zeo<sup>R</sup></i> , homology region, reconstituted <i>arg1</i> gene                    |
| pArg2345-V1-BsaI-Ppm    | 233 bp <i>SwaI-P<sub>pm</sub>-StuI</i> fragment derived from pGH-Ppm-synRBS-arg2, ligated into pArg2345-V1-BsaI expression vector hydrolyzed with the same enzymes.             | <i>P<sub>pm</sub></i> , <i>arg2-5</i> , p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>    |
| pArg2345-V1-BsaI-Ptet   | 803 bp <i>SwaI-P<sub>tet</sub>-StuI</i> fragment derived from pGH-Ptet-synRBS-arg2, ligated into pArg2345-V1-BsaI expression vector hydrolyzed with the same enzymes.           | <i>P<sub>tet</sub></i> , <i>arg2-5</i> , p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>   |
| pArg2345-V1-BsaI-PtetΔR | 169 bp <i>SwaI-P<sub>tetΔR</sub>-StuI</i> fragment derived from pGH-Ptet-wo_tetR-synRBS-arg2, ligated into pArg2345-V1-BsaI expression vector hydrolyzed with the same enzymes. | <i>P<sub>tetΔR</sub></i> , <i>arg2-5</i> , p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pArg2345-V1-BsaI-PnptII | 245 bp <i>SwaI-P<sub>nptII</sub>-StuI</i> fragment derived from pGH-nptII-synRBS-arg2, ligated into pArg2345-V1-BsaI expression vector hydrolyzed with the same enzymes.        | <i>P<sub>nptII</sub></i> , <i>arg2-5</i> , p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |
| pArg2345-V1-BsaI-Pvan   | 1145 bp <i>SwaI-P<sub>van</sub>-StuI</i> fragment derived from pGH-Pvan-synRBS-arg2, ligated into pArg2345-V1-BsaI expression vector hydrolyzed with the same enzymes.          | <i>P<sub>van</sub></i> , <i>arg2-5</i> , p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i>   |
| pArg2345-V1-BsaI-PvanΔR | 200 bp <i>SwaI-P<sub>vanΔR</sub>-StuI</i> fragment derived from pGH-Pvan-wo_vanR-synRBS-arg2, ligated into pArg2345-V1-BsaI expression vector hydrolyzed with the same enzymes. | <i>P<sub>vanΔR</sub></i> , <i>arg2-5</i> , p15A ori, <i>kan<sup>R</sup></i> , <i>tet<sup>R</sup></i> |

## 5.5.2 NMR tables and spectra

## 5.5.2.1 Tryptophan NMR data

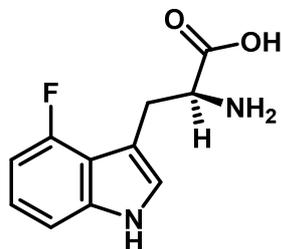


Figure S2. Structure of 4-F-L-tryptophan

Table S6. List of all observed NMR correlations for 4-F-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 54.0                                    | 4.12 d,d (8.3,5.2)                   |
| $\beta$        | 38.6                                    | 3.65 d,d (15.2,5.3)                  |
|                | 38.6                                    | 3.11 d,d (15.2,8.4)                  |
| $\alpha$ -COOH | 171.8                                   |                                      |
| 1              |   |                                      |
| 2              | 125.5                                   | 7.03 s                               |
| 3              | 104.8 d (11.8)                          |                                      |
| 4              | 156.3 d (242.5)                         |                                      |
| 4a             | 114.9 d (20.0)                          |                                      |
| 5              | 104.2 d (19.0)                          | 6.63 d,d (8.0,11.7)                  |
| 6              | 122.5 d (8.2)                           | 6.96 d,d,d (5.3,8.0,8.0)             |
| 7              | 108.1 d (3.5)                           | 7.09 d (8.2)                         |
| 7a             | 139.1 d (11.8)                          |                                      |

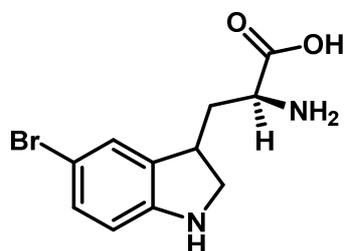
<sup>a</sup> acquired at 125 MHz<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.

Figure S3. Structure of 5-Br-L-tryptophan

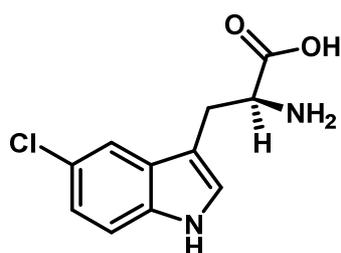
Table S7. List of all observed NMR correlations for 5-Br-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 54.0                                    | 4.06 d,d (7.4,5.04)                  |
| $\beta$        | 25.8                                    | 3.27 d,d (5.2,15.3)                  |
|                | 25.8                                    | 31.7 d,d (7.4,15.4)                  |
| $\alpha$ -COOH | 172.6                                   |                                      |
| 1              |   |                                      |

|    |       |                    |
|----|-------|--------------------|
| 2  | 120.6 | 7.13 s             |
| 3  | 106.3 |                    |
| 4  | 124.6 | 7.67 d (2.0)       |
| 4a | 128.2 |                    |
| 5  | 112.0 |                    |
| 6  | 126.4 | 7.17 d,d (8.8,1.8) |
| 7  | 113.5 | 7.23 d (8.5)       |
| 7a | 134.9 |                    |

<sup>a</sup> acquired at 125 MHz

<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.



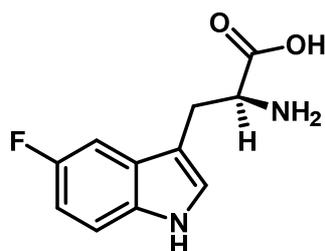
**Figure S4.** Structure of 5-Cl-L-tryptophan

**Table S8.** List of all observed NMR correlations for 5-Cl-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.8                                    | 4.03 d,d (7.4,5.3)                   |
| $\beta$        | 25.8                                    | 3.27 d,d (15.4,5.0)                  |
|                | 25.8                                    | 3.17 d,d (15.4,7.5)                  |
| $\alpha$ -COOH | 172.9                                   |                                      |
| 1              |   |                                      |
| 2              | 122.1                                   | 7.15 s                               |
| 3              | 106.5                                   |                                      |
| 4              | 117.5                                   | 7.51 d (1.8)                         |
| 4a             | 124.6                                   |                                      |
| 5              | 127.5                                   |                                      |
| 6              | 126.5                                   | 7.04                                 |
| 7              | 113.03                                  | 7.28                                 |
| 7a             | 137.5                                   |                                      |

<sup>a</sup> acquired at 125 MHz

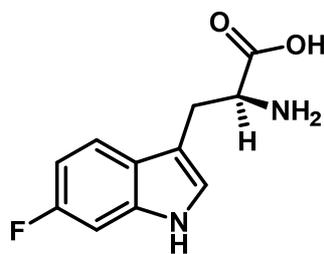
<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.



**Figure S5.** Structure of 5-F-L-tryptophan

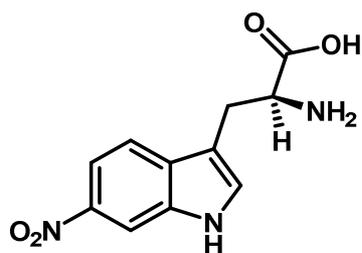
**Table S9.** List of all observed NMR correlations for 5-F-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.4                                    | 4.12 d,d (7.4, 5.4)                  |
| $\beta$        | 25.7                                    | 3.29 d,d (15.5, 5.3)                 |
|                | 25.7                                    | 3.21 d,d (15.4, 7.3)                 |
| $\alpha$ -COOH | 172.3                                   |                                      |
| 1              |   |                                      |
| 2              | 125.5                                   | 7.09 s                               |
| 3              | 106.6                                   |                                      |
| 4              | 119.0 d (9.9)                           | 7.42 d,d (8.7,5.5)                   |
| 4a             | 123.1                                   |                                      |
| 5              | 97.8 d (26.3)                           | 6.78 d,d,d (2.0,9.0,9.0)             |
| 6              | 159.6 d                                 | 6.84 d,d,d (2.5,9.3,9.3)             |
| 7              | 112.0 d (10)                            | 7.27 d,d (8.9,4.43)                  |
| 7a             | 132.8                                   |                                      |

<sup>a</sup> acquired at 125 MHz<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.**Figure S6.** Structure of 6-F-L-tryptophan**Table S10.** List of all observed NMR correlations for 6-F-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.4                                    | 4.12 d,d (7.3, 5.3)                  |
| $\beta$        | 25.7                                    | 3.29 d,d (15.5, 5.3)                 |
|                | 25.7                                    | 3.21 d,d (15.4, 7.3)                 |
| $\alpha$ -COOH | 172.3                                   |                                      |
| 1              |   |                                      |
| 2              | 125.5                                   | 7.09 s                               |
| 3              | 106.6                                   |                                      |
| 4              | 119.0 d (9.9)                           | 7.42 d,d (8.7, 5.5)                  |
| 4a             | 123.1                                   |                                      |
| 5              | 97.8 d (26.3)                           | 6.78 d,d,d (2.0,9.0,9.0)             |
| 6              | 159.6 d (235.2)                         |                                      |
| 7              | 108.0                                   | 7.04 d,d (10.3, 2.1)                 |
| 7a             | 136.2 d (12.7)                          |                                      |

<sup>a</sup> acquired at 125 MHz<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.

Figure S7. Structure of 6-NO<sub>2</sub>-L-tryptophanTable S11. List of all observed NMR correlations for 6-NO<sub>2</sub>-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.7                                    | 4.09 d,d (6.9,5.6)                   |
| $\beta$        | 25.5                                    | 3.29 d,d (15.4,5.4)                  |
|                | 25.5                                    | 2.26 d,d (15.6, 7.1)                 |
| $\alpha$ -COOH | 172.4                                   |                                      |
| 1              |   |                                      |
| 2              | 131.4                                   | 7.39 s                               |
| 3              | 107.9                                   |                                      |
| 4              | 118.0                                   | 7.48 d (8.7)                         |
| 4a             | 131.6                                   |                                      |
| 5              | 114.6                                   | 7.73 d,d (2.0,8.7)                   |
| 6              | 142.3                                   |                                      |
| 7              | 108.9                                   | 8.13 d (1.72)                        |
| 7a             | 134.5                                   |                                      |

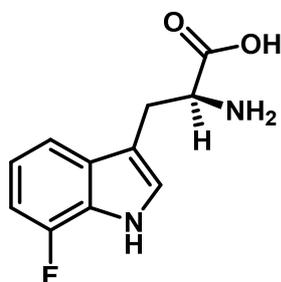
<sup>a</sup> acquired at 125 MHz<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.

Figure S8. Structure of 7-F-L-tryptophan

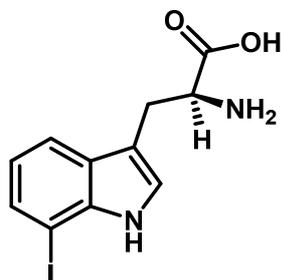
Table S12. List of all observed NMR correlations for 7-F-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.1                                    | 4.09 d,d (5.5,7.1)                   |
| $\beta$        | 25.6                                    | 3.18 d,d (15.4,5.2)                  |
|                | 25.6                                    | 3.11 d,d (15.3,7.3)                  |
| $\alpha$ -COOH | 171.8                                   |                                      |
| 1              |   |                                      |
| 2              | 125.8                                   | 7.05 s                               |
| 3              | 107.1                                   |                                      |
| 4              | 113.9 d (2.7)                           | 7.14 d (8.0)                         |
| 4a             | 130.2 d (5.4)                           |                                      |
| 5              | 119.7 d (5.4)                           | 6.82 d,d,d (7.8,7.8,4.7)             |
| 6              | 106.6 d (15.4)                          | 6.73 d,d (11.4,7.7)                  |
| 7              | 149.4 d (243.4)                         |                                      |

|    |                |  |
|----|----------------|--|
| 7a | 124.3 d (13.6) |  |
|----|----------------|--|

<sup>a</sup> acquired at 125 MHz

<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.



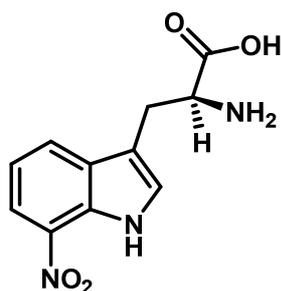
**Figure S9.** Structure of 7-I-L-tryptophan

**Table S13.** List of all observed NMR correlations for 7-I-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.8                                    | 4.04 d,d (7.4,5.4)                   |
| $\beta$        | 26.1                                    | 3.26 d,d (15.3,5.3)                  |
|                | 26.1                                    | 3.15 d,d (15.5,7.3)                  |
| $\alpha$ -COOH | 172.7                                   |                                      |
| 1              |   |                                      |
| 2              | 125.5                                   | 7.16 s                               |
| 3              | 108.3                                   |                                      |
| 4              | 118.4                                   | 7.49 d (7.0)                         |
| 4a             | 126.8                                   |                                      |
| 5              | 121.1                                   | 6.77 d,d (7.8,7.8)                   |
| 6              | 131.0                                   | 7.47 d (5.9)                         |
| 7              | 75.9                                    |                                      |
| 7a             | 133.2                                   |                                      |

<sup>a</sup> acquired at 125 MHz

<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.



**Figure S10.** Structure of 7-NO<sub>2</sub>-L-tryptophan

**Table S14.** List of all observed NMR correlations for 7-NO<sub>2</sub>-L-tryptophan in D<sub>2</sub>O.

| Res            | $\Delta^{13}\text{C}^a$ Mult. (j in Hz) | $\Delta^1\text{H}^b$ Mult. (j in Hz) |
|----------------|---|--------------------------------------|
| $\alpha$       | 53.8                                    | 4.04 d,d (7.2, 5.3)                  |
| $\beta$        | 25.5                                    | 3.29 d,d (15.5,5.3)                  |
|                | 25.5                                    | 3.21 d,d (15.5,7.3)                  |
| $\alpha$ -COOH | 172.7                                   |                                      |
| 1              |   |                                      |
| 2              | 127.1                                   | 7.21 s                               |

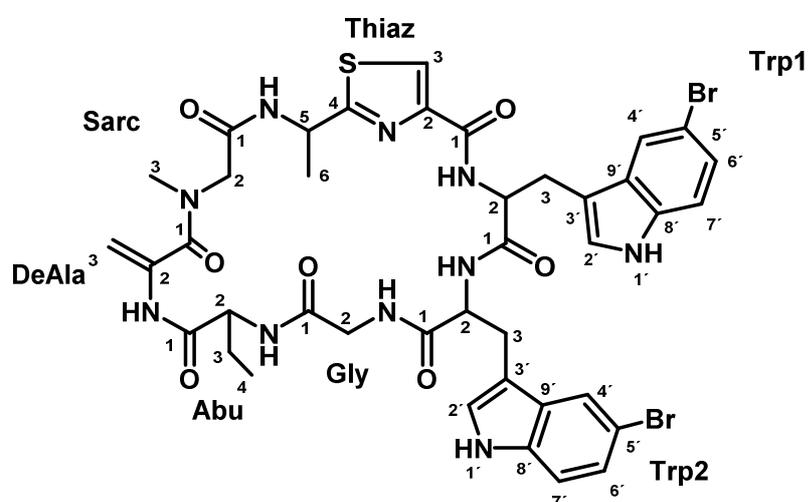
|    |       |                    |
|----|-------|--------------------|
| 3  | 108.9 |                    |
| 4  | 127.2 | 7.83 d (7.8)       |
| 4a | 128.9 |                    |
| 5  | 119.6 | 7.03 d,d (8.3,8.3) |
| 6  | 118.6 | 7.92 d (8.3)       |
| 7  | 132.3 |                    |
| 7a | 130.6 |                    |

<sup>a</sup> acquired at 125 MHz

<sup>b</sup> acquired at 500 MHz, referenced to solvent signal D<sub>2</sub>O at  $\delta$  4.79 ppm.

### 5.5.2.2 Argyrin NMR data

#### 5,5'-di-Br-argyrin I



**Figure S11.** Structure formula of 5,5'-di-Br-argyrin I. Atom numbers are assigned according to the NMR data.

**Table S15.** List of all observed NMR correlations for 5,5'-di-Br-argyrin I in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|----------|-----------------------------|----------------------|------|-----------|
| Thiaz NH |                             | 8.75                 | d    | 8.5       |
| 1        | 159.9                       |                      |      |           |
| 2        | 149.8                       |                      |      |           |
| 3        | 123.2                       | 7.94                 | s    |           |
| 4        | 171.3                       |                      |      |           |
| 5        | 45.5                        | 5.48                 | d,q  | 7.2, 7.5  |
| 6        | 20.7                        | 1.72                 | d    | 7.2       |
|          |                             |                      |      |           |
| Trp1 NH  |                             | 8.61                 | m    |           |
| 1        | 172.8                       |                      |      |           |
| 2        | 52.1                        | 5.08                 | m    |           |
| 3        | 27.3                        | 3.51                 | d,d  | 15.9, 3.7 |
|          |                             | 2.91                 | d,d  | 15.3, 3.4 |
| 1' NH    |                             | 10.84                | s    |           |
| 2'       | 127.1                       | 7.09                 | d    | 2.3       |
| 3'       | 105.8                       |                      |      |           |
| 4'       | 124.8                       | 7.26                 | s    |           |

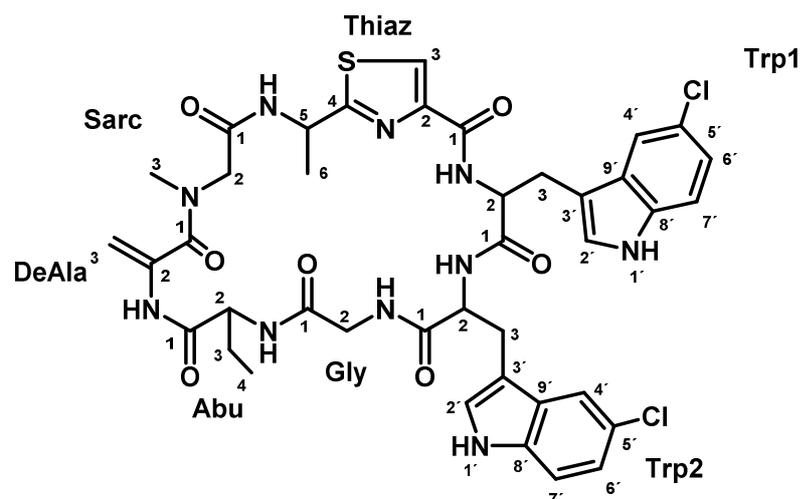
|                    |       |                   |     |           |
|--------------------|-------|-------------------|-----|-----------|
| 5'                 | 112.8 |                   |     |           |
| 6'                 | 119.2 | 7.19              | m   |           |
| 7'                 | 113.6 | 7.13              | d   | 8.5       |
| 8'                 | 133.6 |                   |     |           |
| 9'                 | 128.5 |                   |     |           |
|                    |       |                   |     |           |
| Trp2 NH            |       | 6.64              | s   |           |
| 1                  | 169.9 |                   |     |           |
| 2                  | 56.4  | 4.30              | m   |           |
| 3                  | 26.5  | 3.32              | d,d | 15.2, 5.2 |
|                    |       | 3.23              | d,d | 15.3, 6.9 |
| 1' NH              |       | 8.59              |     |           |
| 2'                 | 124.2 | 7.19              | m   |           |
| 3'                 | 108.3 |                   |     |           |
| 4'                 | 120.4 | 7.81              | s   |           |
| 5'                 | 113.9 |                   |     |           |
| 6'                 | 126.3 | 7.41 <sup>d</sup> | m   |           |
| 7'                 | 113.9 | 7.41 <sup>d</sup> | m   |           |
| 8'                 | 134.8 |                   |     |           |
| 9'                 | 129.0 |                   |     |           |
|                    |       |                   |     |           |
| Gly NH             |       | 4.97              | m   |           |
| 1                  | 170.8 |                   |     |           |
| 2                  | 41.6  | 3.79              | d,d | 17.5, 8.2 |
|                    |       | 1.70              | m   |           |
|                    |       |                   |     |           |
| Abu NH             |       | 6.84              | d   | 6.55      |
| 1                  | 169.3 |                   |     |           |
| 2                  | 54.5  | 4.06              | t,d | 8.7, 6.4  |
| 3                  | 21.4  | 1.99              | m   |           |
|                    |       | 1.88              | m   |           |
| 4                  | 10.6  | 0.89              | t   | 7.4       |
|                    |       |                   |     |           |
| DeAla NH           |       | 9.37              |     |           |
| 1                  | 168.3 |                   |     |           |
| 2                  | 136.6 |                   |     |           |
| 3                  | 100.2 | 5.07              | m   |           |
|                    |       | 4.76              | d   | 1.3       |
|                    |       |                   |     |           |
| Sarc               |       |                   |     |           |
| 1                  | 166.8 |                   |     |           |
| 2                  | 51.1  | 4.97              | m   |           |
|                    |       | 3.42              | d   | 17.0      |
| CH <sub>3</sub> -N | 37.6  | 3.13              | s   |           |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

**5,5'-di-Cl-argyrin I****Figure S12.** Structure formula of 5,5'-di-Cl-argyrin I. Atom numbers are assigned according to the NMR data.**Table S16.** List of all observed NMR correlations for 5,5'-di-Cl-argyrin I in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|----------|-----------------------------|----------------------|------|-----------|
| Thiaz NH |                             | 8.75                 | d    | 8.5       |
| 1        | 159.9                       |                      |      |           |
| 2        | 149.8                       |                      |      |           |
| 3        | 123.1                       | 7.93                 | s    |           |
| 4        | 171.2                       |                      |      |           |
| 5        | 45.5                        | 5.46                 | m    |           |
| 6        | 20.6                        | 1.72                 | d    | 7.2       |
| Trp1 NH  |                             | 8.65                 | d    | 7.4       |
| 1        | 172.9                       |                      |      |           |
| 2        | 52.2                        | 5.12                 | m    |           |
| 3        | 27.3                        | 3.53                 | d,d  | 15.9, 3.5 |
|          |                             | 2.93                 | d,d  | 15.7, 3.1 |
| 1' NH    |                             | 10.84                | s    |           |
| 2'       | 127.4                       | 7.11                 | d    | 2.4       |
| 3'       | 105.9                       |                      |      |           |
| 4'       | 116.0                       | 6.93                 | s    |           |
| 5'       | 125.4                       |                      |      |           |
| 6'       | 122.3                       | 7.07                 | d,d  | 8.4, 2.1  |
| 7'       | 113.2                       | 7.16                 | d    | 8.5       |
| 8'       | 133.3                       |                      |      |           |
| 9'       | 127.9                       |                      |      |           |
| Trp2 NH  |                             | 6.81                 | s    |           |
| 1        | 170.2                       |                      |      |           |
| 2        | 56.3                        | 4.31                 | m    |           |
| 3        | 26.3                        | 3.30                 | d,d  | 15.4, 4.9 |
|          |                             | 3.26                 | d,d  | 15.3, 6.4 |
| 1' NH    |                             | 8.57                 | s    |           |
| 2'       | 124.4                       | 7.23                 | d    | 2.0       |
| 3'       | 108.7                       |                      |      |           |
| 4'       | 117.3                       | 7.62                 | d    | 1.9       |
| 5'       | 126.7                       |                      |      |           |
| 6'       | 123.8                       | 7.27                 | m    |           |

|                    |       |      |     |           |
|--------------------|-------|------|-----|-----------|
| 7'                 | 113.5 | 7.44 | d   | 8.6       |
| 8'                 | 134.8 |      |     |           |
| 9'                 | 128.6 |      |     |           |
|                    |       |      |     |           |
| Gly NH             |       | 4.95 | m   |           |
| 1                  | 171.1 |      |     |           |
| 2                  | 41.5  | 3.74 | d,d | 17.4, 7.9 |
|                    |       | 1.60 | m   |           |
|                    |       |      |     |           |
| Abu NH             |       | 6.84 | d   | 6.7       |
| 1                  | 169.5 |      |     |           |
| 2                  | 54.5  | 4.05 | t,d | 9.1, 6.4  |
| 3                  | 21.4  | 1.99 | m   |           |
|                    |       | 1.88 | m   |           |
| 4                  | 10.5  | 0.89 | t   | 7.4       |
|                    |       |      |     |           |
| DeAla NH           |       | 9.37 | s   |           |
| 1                  | 168.3 |      |     |           |
| 2                  | 136.8 |      |     |           |
| 3                  | 100.2 | 5.07 | d   | 1.7       |
|                    |       | 4.76 | d   | 1.3       |
|                    |       |      |     |           |
| Sarc               |       |      |     |           |
| 1                  | 166.8 |      |     |           |
| 2                  | 51.1  | 4.96 | d   | 17.1      |
|                    |       | 3.42 | d   | 16.8      |
| CH <sub>3</sub> -N | 37.5  | 3.13 | s   |           |

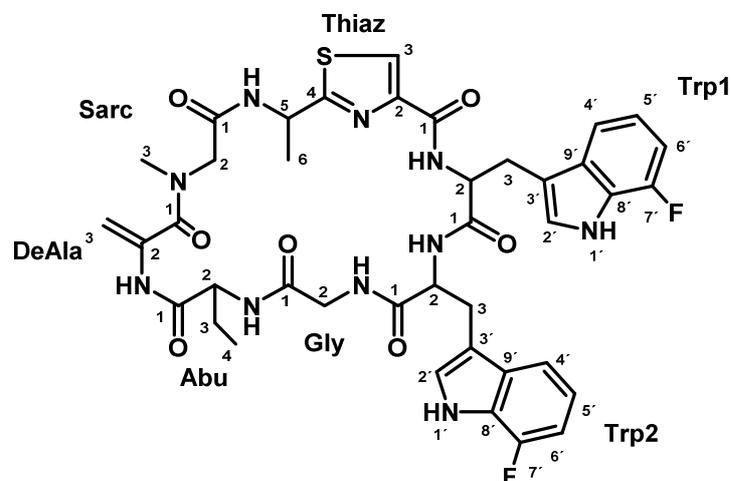
<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

### 7,7'-di-F-argyrin I



**Figure S13.** Structure formula of 7,7'-di-F-argyrin I. Atom numbers are assigned according to the NMR data.

**Table S17.** List of all observed NMR correlations for 7,7'-di-F-argyrin I in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult           | J(Hz)            |
|----------|-----------------------------|----------------------|----------------|------------------|
| Thiaz NH |                             | 8.85                 | d              | 8.3              |
| 1        | 160.2                       |                      |                |                  |
| 2        | 149.9                       |                      |                |                  |
| 3        | 122.9                       | 8.01                 | s              |                  |
| 4        | 171.4                       |                      |                |                  |
| 5        | 45.2                        | 5.46                 | m              |                  |
| 6        | 20.3                        | 1.73                 | d              | 6.9              |
|          |                             |                      |                |                  |
| Trp1 NH  |                             | 8.71                 | m              | 7.4              |
| 1        | 172.8                       |                      |                |                  |
| 2        | 52.1                        | 5.26                 | m              |                  |
| 3        | 27.4                        | 3.59                 | m              |                  |
|          |                             | 3.09                 | d,d            | 15.6, 3.5        |
| 1' NH    |                             | 11.1                 | s              |                  |
| 2'       | 126.2                       | 7.08                 | d              | 2.6              |
| 3'       | 106.6                       |                      |                |                  |
| 4'       | 111.3                       | 6.14                 | d              | 8.2              |
| 5'       | 119.7                       | 6.42                 | d,d,d          | 4.5,7.9,7.9      |
| 6'       | 106.2                       | 6.67                 | m              |                  |
| 7'       | 149.8                       |                      | d <sup>e</sup> | 245 <sup>e</sup> |
| 8'       | 123.2                       |                      |                |                  |
| 9'       | 130.4                       |                      |                |                  |
|          |                             |                      |                |                  |
| Trp2 NH  |                             | 7.11                 | m              |                  |
| 1        | 170.0                       |                      |                |                  |
| 2        | 56.3                        | 4.28                 | m              |                  |
| 3        | 26.3                        | 3.43                 | m              |                  |
|          |                             | 3.30                 | d,d            | 15.3, 4.9        |
| 1' NH    |                             | 8.70                 | s              |                  |
| 2'       | 124.0                       | 7.45                 | d              | 2.2              |
| 3'       | 110.2                       |                      |                |                  |
| 4'       | 113.9                       | 7.31                 | d              | 7.9              |
| 5'       | 121.0                       | 7.12                 | m              |                  |
| 6'       | 107.9                       | 7.06                 | d,d            | 7.8, 10.8        |
| 7'       | 149.9                       |                      | d <sup>e</sup> | 244 <sup>e</sup> |
| 8'       | 125.0                       |                      |                |                  |
| 9'       | 131.2                       |                      |                |                  |
|          |                             |                      |                |                  |
| Gly NH   |                             | 4.69                 | d,d            | 6.6, 6.3         |
| 1        | 171.0                       |                      |                |                  |
| 2        | 41.0                        | 3.57                 | m              |                  |
|          |                             | 1.20                 | m              |                  |
|          |                             |                      |                |                  |
| Abu NH   |                             | 6.68                 | m              |                  |
| 1        | 169.5                       |                      |                |                  |
| 2        | 54.4                        | 4.03                 | t,d            | 9.3, 6.2         |
| 3        | 21.3                        | 1.98                 | m              |                  |
|          |                             | 1.86                 | m              |                  |
| 4        | 10.4                        | 0.88                 | t              | 7.3              |
|          |                             |                      |                |                  |
| DeAla NH |                             | 9.27                 | s              |                  |
| 1        | 168.6                       |                      |                |                  |

|                    |       |      |   |      |
|--------------------|-------|------|---|------|
| 2                  | 136.1 |      |   |      |
| 3                  | 100.4 | 5.03 | m |      |
|                    |       | 4.74 | d | 1.6  |
|                    |       |      |   |      |
| Sarc               |       |      |   |      |
| 1                  | 167.0 |      |   |      |
| 2                  | 51.1  | 4.94 | d | 17.1 |
|                    |       | 3.43 | m |      |
| CH <sub>3</sub> -N | 37.5  | 3.12 | s |      |

<sup>a</sup> acquired at 125 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

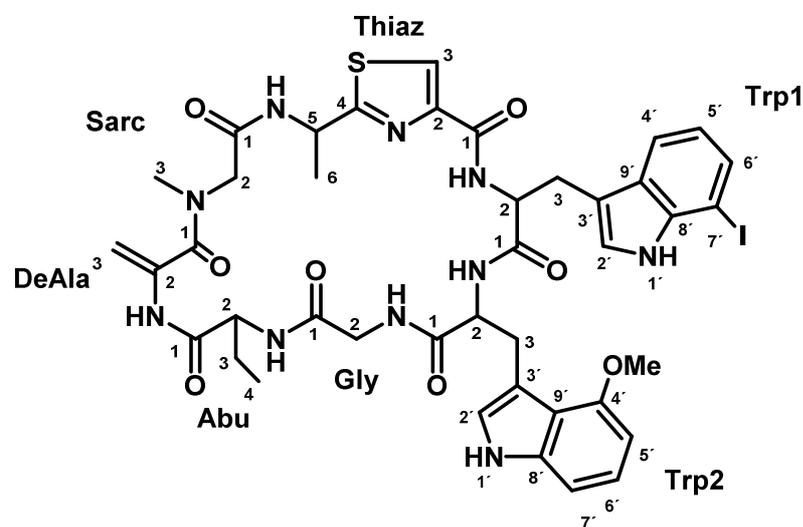
<sup>b</sup> acquired at 500 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

<sup>e</sup> obtained from <sup>13</sup>C NMR spectrum at 125 MHz

### 7-I-argyrin B



**Figure S14.** Structure formula of 7-I-argyrin B. Atom numbers are assigned according to the NMR data.

**Table S18.** List of all observed NMR correlations for 7-I-argyrin B in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|----------|-----------------------------|----------------------|------|-----------|
| Thiaz NH |                             | 8.69                 | d    | 8.7       |
| 1        | 160.0                       |                      |      |           |
| 2        | 150.3                       |                      |      |           |
| 3        | 122.8                       | 8.04                 | s    |           |
| 4        | 171.4                       |                      |      |           |
| 5        | 45.6                        | 5.40                 | m    |           |
| 6        | 20.2                        | 1.76                 | d    | 6.9       |
|          |                             |                      |      |           |
| Trp1 NH  |                             | 8.64                 | d    | 6.9       |
| 1        | 172.9                       |                      |      |           |
| 2        | 52.0                        | 5.07                 | m    |           |
| 3        | 27.2                        | 3.53                 | m    |           |
|          |                             | 2.78                 | d,d  | 15.4, 3.4 |
| 1' NH    |                             | 10.8                 | s    |           |
| 2'       | 125.9                       | 7.13                 | m    |           |
| 3'       | 107.8                       |                      |      |           |
| 4'       | 116.4                       | 5.37                 | d    | 7.9       |

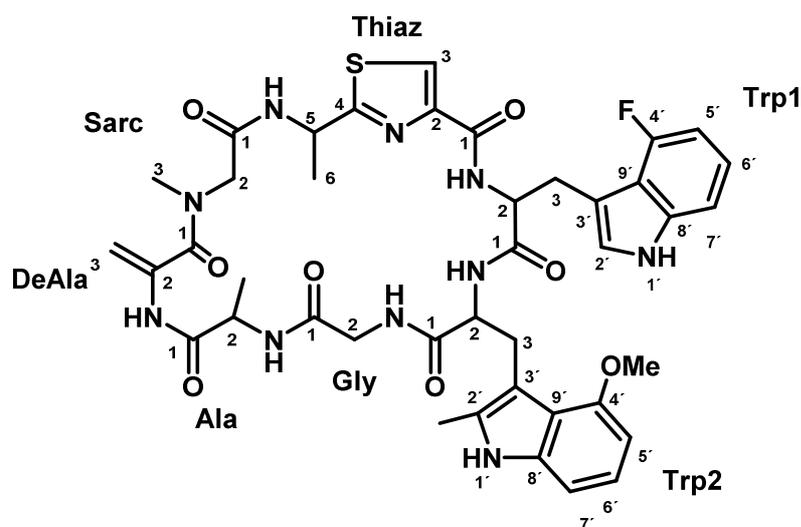
|                    |       |      |     |          |
|--------------------|-------|------|-----|----------|
| 5'                 | 121.1 | 6.09 | d,d | 7.8,7.8  |
| 6'                 | 130.1 | 7.26 | m   |          |
| 7'                 | 76.8  |      |     |          |
| 8'                 | 137.1 |      |     |          |
| 9'                 | 126.9 |      |     |          |
|                    |       |      |     |          |
| Trp2 NH            |       | 8.75 | m   |          |
| 1                  | 170.1 |      |     |          |
| 2                  | 57.8  | 4.20 | m   |          |
| 3                  | 26.9  | 3.46 | d,d | 3.8,14.9 |
|                    |       | 3.30 | d,d | 3.9,14.8 |
| 1' NH              |       | 8.29 | s   |          |
| 2'                 | 123.7 | 6.78 | s   | 2.2      |
| 3'                 | 108.5 |      |     |          |
| 4'                 | 152.4 |      |     |          |
| 5'                 | 101.5 | 6.92 | d   | 7.6      |
| 6'                 | 123.9 | 7.35 | m   |          |
| 7'                 | 106.7 | 7.32 | m   |          |
| 8'                 | 138.6 |      |     |          |
| 9'                 | 117.6 |      |     |          |
| 4'-OMe             | 56.3  | 4.35 | s   |          |
|                    |       |      |     |          |
| Gly NH             |       | 4.28 | d,d | 6.6,6.3  |
| 1                  | 171.5 |      |     |          |
| 2                  | 40.4  | 3.53 | m   |          |
|                    |       | 1.09 | d,d | 5.0,17.2 |
|                    |       |      |     |          |
| Abu NH             |       | 6.75 | m   |          |
| 1                  | 170.2 |      |     |          |
| 2                  | 54.6  | 4.04 | t,d | 9.1, 5.8 |
| 3                  | 21.4  | 2.02 | m   |          |
|                    |       | 1.91 | m   |          |
| 4                  | 10.6  | 0.90 | t   | 7.3      |
|                    |       |      |     |          |
| DeAla NH           |       | 9.36 | s   |          |
| 1                  | 168.6 |      |     |          |
| 2                  | 135.8 |      |     |          |
| 3                  | 101.8 | 5.01 | m   |          |
|                    |       | 4.74 | m   |          |
|                    |       |      |     |          |
| Sarc               |       |      |     |          |
| 1                  | 167.3 |      |     |          |
| 2                  | 50.9  | 4.95 | d   | 16.5     |
|                    |       | 3.39 | d   | 16.7     |
| CH <sub>3</sub> -N | 37.4  | 3.17 | s   |          |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

**4-F-argyrin C****Figure S15.** Structure formula of 4-F-argyrin C. Atom numbers are assigned according to the NMR data.**Table S19.** List of all observed NMR correlations for 4-F-argyrin C in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult           | J(Hz)              |
|----------|-----------------------------|----------------------|----------------|--------------------|
| Thiaz NH |                             | 8.82                 | d              | 8.6                |
| 1        | 159.8                       |                      |                |                    |
| 2        | 150.4                       |                      |                |                    |
| 3        | 123.1                       | 8.02                 | s              |                    |
| 4        | 170.8                       |                      |                |                    |
| 5        | 45.2                        | 5.46                 | m              |                    |
| 6        | 20.3                        | 1.72                 | d              | 7.1                |
|          |                             |                      |                |                    |
| Trp1 NH  |                             | 8.57                 | d              | 7.0                |
| 1        | 173.2                       |                      |                |                    |
| 2        | 52.3                        | 4.83                 | m              |                    |
| 3        | 28.0                        | 3.75                 | m              |                    |
|          |                             | 3.57                 | d,d            | 15.4, 3.3          |
| 1' NH    |                             | 11.17                | s              |                    |
| 2'       | 126.4                       | 7.02                 | m              |                    |
| 3'       | 104.5                       |                      |                |                    |
| 4'       | 156.2                       |                      | d <sup>e</sup> | 243.3 <sup>e</sup> |
| 5'       | 104.5                       | 6.69                 | d,d            | 7.0,11.5           |
| 6'       | 122.2                       | 7.03 <sup>d</sup>    | m              |                    |
| 7'       | 108.5                       | 7.03 <sup>d</sup>    | m              |                    |
| 8'       | 137.7                       |                      |                |                    |
| 9'       | 115.4                       |                      |                |                    |
|          |                             |                      |                |                    |
| Trp2 NH  |                             | 8.44                 | m              |                    |
| 1        | 171.0                       |                      |                |                    |
| 2        | 59.2                        | 4.04                 | d,t            | 9.7,2.8            |
| 3        | 26.4                        | 3.36                 | d,d            | 2.9, 14.8          |
|          |                             | 3.28                 | d,d            | 9.9,14.8           |
| 1' NH    |                             | 7.98                 | s              |                    |
| 2'       | 131.9                       |                      |                |                    |
| 3'       | 106.2                       |                      |                |                    |
| 4'       | 152.3                       |                      |                |                    |

|                    |       |      |     |         |
|--------------------|-------|------|-----|---------|
| 5'                 | 101.5 | 6.74 | d   | 7.8     |
| 6'                 | 122.3 | 7.13 | d,d | 8.0,8.0 |
| 7'                 | 105.3 | 7.02 | m   |         |
| 8'                 | 136.7 |      |     |         |
| 9'                 | 118.2 |      |     |         |
| 4'-OMe             | 56.2  | 4.27 | s   |         |
| 2'-Me              | 11.5  | 2.36 | s   |         |
|                    |       |      |     |         |
| Gly NH             |       | 5.42 | m   |         |
| 1                  | 171.1 |      |     |         |
| 2                  | 41.1  | 3.72 | m   |         |
|                    |       | 1.43 | m   |         |
|                    |       |      |     |         |
| Ala NH             |       | 6.96 | d   | 6.2     |
| 1                  | 169.8 |      |     |         |
| 2                  | 48.5  | 4.28 | m   |         |
| 3                  | 14.1  | 1.43 | d   | 7.1     |
|                    |       |      |     |         |
| DeAla NH           |       | 9.51 | s   |         |
| 1                  | 168.7 |      |     |         |
| 2                  | 136.8 |      |     |         |
| 3                  | 100.0 | 5.06 | d   | 1.5     |
|                    |       | 4.76 | d   | 1.5     |
|                    |       |      |     |         |
| Sarc               |       |      |     |         |
| 1                  | 166.9 |      |     |         |
| 2                  | 51.1  | 4.95 | d   | 17.1    |
|                    |       | 3.43 | d   | 16.9    |
| CH <sub>3</sub> -N | 37.6  | 3.13 | s   |         |

<sup>a</sup> acquired at 125 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

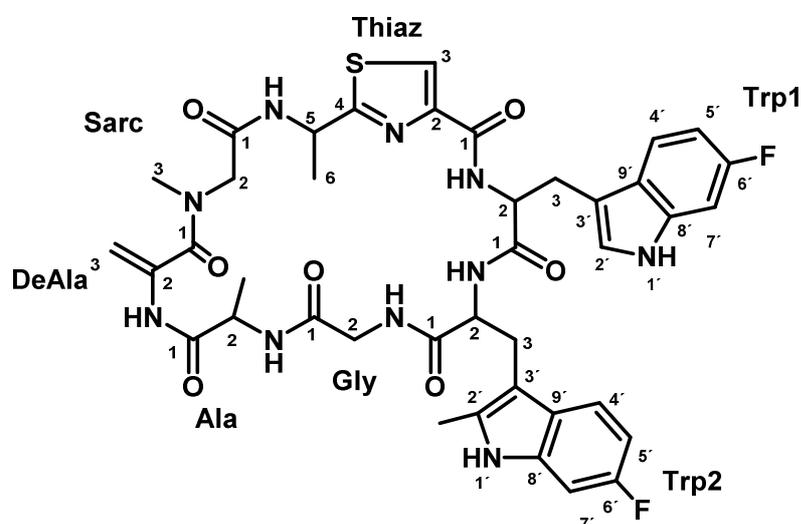
<sup>b</sup> acquired at 500 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

<sup>e</sup> obtained from <sup>13</sup>C NMR spectrum at 125 MHz

### 6,6'-di-F-argyrin C2



**Figure S16.** Structure formula of 6,6'-di-F-argyrin C2. Atom numbers are assigned according to the NMR data.

**Table S20.** List of all observed NMR correlations for 6,6'-di-F-argyrin C2 in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult           | J(Hz)            |
|----------|-----------------------------|----------------------|----------------|------------------|
| Thiaz NH |                             | 8.80                 | d              | 8.5              |
| 1        | 160.0                       |                      |                |                  |
| 2        | 150.2                       |                      |                |                  |
| 3        | 123.1                       | 8.05                 | s              |                  |
| 4        | 171.2                       |                      |                |                  |
| 5        | 45.5                        | 5.45                 | m              |                  |
| 6        | 20.6                        | 1.72                 | d              | 6.9              |
|          |                             |                      |                |                  |
| Trp1 NH  |                             | 8.48                 | d              | 7.0              |
| 1        | 172.7                       |                      |                |                  |
| 2        | 52.1                        | 5.05                 | m              |                  |
| 3        | 26.8                        | 3.41                 | m              |                  |
|          |                             | 2.58                 | d,d            | 15.5, 3.4        |
| 1' NH    |                             | 10.73                | s              |                  |
| 2'       | 126.0                       | 6.98                 | d              | 2.5              |
| 3'       | 106.1                       |                      |                |                  |
| 4'       | 116.7                       | 5.92                 | d,d            | 8.6, 5.1         |
| 5'       | 108.0                       | 6.45                 | m              |                  |
| 6'       | 159.4                       |                      | d <sup>e</sup> | 239 <sup>e</sup> |
| 7'       | 98.0                        | 6.84                 | d,d            | 9.4, 2.4         |
| 8'       | 134.7                       |                      |                |                  |
| 9'       | 123.1                       |                      |                |                  |
|          |                             |                      |                |                  |
| Trp2 NH  |                             | 6.43                 | m              |                  |
| 1        | 170.4                       |                      |                |                  |
| 2        | 56.6                        | 4.29                 | m              |                  |
| 3        | 24.9                        | 3.28                 | d,d            | 4.3,15.4         |
|          |                             | 3.22                 | d,d            | 5.6,15.4         |
| 1' NH    |                             | 8.27                 | s              |                  |
| 2'       | 134.1                       |                      |                |                  |
| 3'       | 104.7                       |                      |                |                  |
| 4'       | 117.1                       | 7.48                 | d,d            | 4.8,8.7          |
| 5'       | 109.8                       | 7.11                 | d,d,d          | 9.0,8.8,2.1      |
| 6'       | 160.1                       |                      | d              | 239              |
| 7'       | 98.2                        | 7.23                 | d,d            | 2.5,9.2          |
| 8'       | 135.6                       |                      |                |                  |
| 9'       | 125.4                       |                      |                |                  |
|          |                             |                      |                |                  |
| Gly NH   |                             | 4.69                 | m              |                  |
| 1        | 170.9                       |                      |                |                  |
| 2        | 41.3                        | 3.76                 | d,d            | 8.1,17.5         |
|          |                             | 1.56                 | m              |                  |
|          |                             |                      |                |                  |
| Ala NH   |                             | 6.70                 | d              | 6.1              |
| 1        | 169.8                       |                      |                |                  |
| 2        | 48.5                        | 4.30                 | m              |                  |
| 3        | 14.1                        | 1.42                 | d              | 7.2              |
|          |                             |                      |                |                  |
| DeAla NH |                             | 9.49                 | s              |                  |
| 1        | 168.3                       |                      |                |                  |
| 2        | 136.6                       |                      |                |                  |

|                    |       |      |   |      |
|--------------------|-------|------|---|------|
| 3                  | 100.5 | 5.07 | d | 1.6  |
|                    |       | 4.75 | d | 1.4  |
| Sarc               |       |      |   |      |
| 1                  | 167.0 |      |   |      |
| 2                  | 51.1  | 4.95 | d | 17.0 |
|                    |       | 3.41 | d | 17.0 |
| CH <sub>3</sub> -N | 37.6  | 3.13 | s |      |

<sup>a</sup> acquired at 125 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

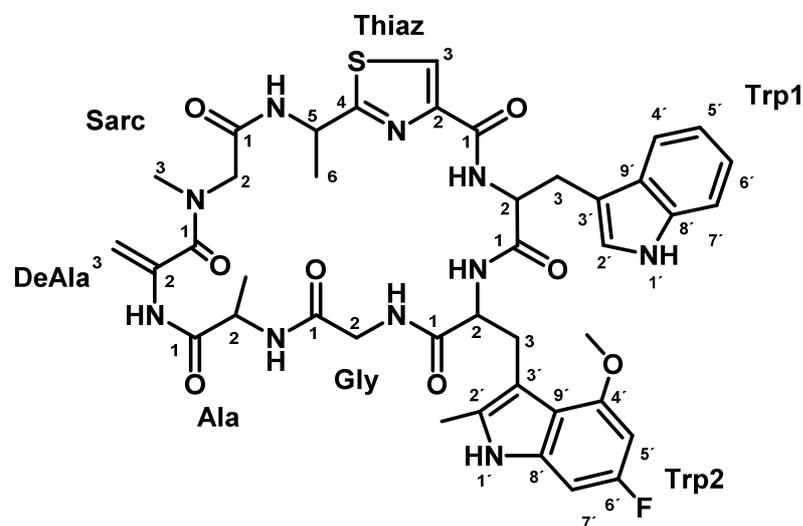
<sup>b</sup> acquired at 500 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

<sup>e</sup> obtained from <sup>13</sup>C NMR spectrum at 125 MHz

### 6'-F-argyrin C



**Figure S17.** Structure formula of 6'-F-argyrin C. Atom numbers are assigned according to the NMR data.

**Table S21.** List of all observed NMR correlations for 6'-F-argyrin C in CDCl<sub>3</sub>.

| Res      | $\Delta$ <sup>13</sup> C <sup>a,c</sup> | $\Delta$ <sup>1</sup> H <sup>b</sup> | Mult | J(Hz)     |
|----------|---|--------------------------------------|------|-----------|
| Thiaz NH |   | 8.82                                 | d    | 8.5       |
| 1        | 160.0                                   |                                      |      |           |
| 2        | 150.5                                   |                                      |      |           |
| 3        | 123.1                                   | 8.06                                 | s    |           |
| 4        | 170.9                                   |                                      |      |           |
| 5        | 45.5                                    | 5.48                                 | m    |           |
| 6        | 20.6                                    | 1.73                                 | d    | 7.0       |
| Trp1 NH  |   | 8.54                                 | m    |           |
| 1        | 172.7                                   |                                      |      |           |
| 2        | 52.2                                    | 5.03                                 | m    |           |
| 3        | 26.9                                    | 3.55                                 | d,d  | 15.3,3.2  |
|          |   | 2.81                                 | d,d  | 15.5, 3.1 |
| 1' NH    |   | 10.70                                | s    |           |
| 2'       | 125.7                                   | 6.98                                 | d    | 2.5       |
| 3'       | 105.9                                   |                                      |      |           |
| 4'       | 116.1                                   | 5.63                                 | m    |           |

|                    |       |      |                |                  |
|--------------------|-------|------|----------------|------------------|
| 5'                 | 119.1 | 6.44 | d,d            | 7.25, 7.25       |
| 6'                 | 121.4 | 6.94 | m              |                  |
| 7'                 | 111.6 | 7.09 | d              | 7.8              |
| 8'                 | 134.9 |      |                |                  |
| 9'                 | 126.6 |      |                |                  |
| Trp2 NH            |       | 6.86 | m              |                  |
| 1                  | 170.2 |      |                |                  |
| 2                  | 57.0  | 4.18 | m              |                  |
| 3                  | 25.6  | 3.39 | m              |                  |
|                    |       | 3.27 | d,d            | 15.4, 3.8        |
| 1' NH              |       | 8.09 | s              |                  |
| 2'                 | 133.5 |      |                |                  |
| 3'                 | 104.0 |      |                |                  |
| 4'                 | 114.9 |      |                |                  |
| 5'                 | 92.1  | 6.70 | d              | 11.0             |
| 6'                 | 160.1 |      | d <sup>e</sup> | 238 <sup>e</sup> |
| 7'                 | 91.9  | 6.94 | m              |                  |
| 8'                 | 151.3 |      |                |                  |
| 9'                 | 133.7 |      |                |                  |
| 4'-OMe             | 56.5  | 4.29 | s              |                  |
| 2'-Me              | 11.5  | 2.11 | s              |                  |
| Gly NH             |       | 4.60 | m              |                  |
| 1                  | 170.8 |      |                |                  |
| 2                  | 41.1  | 3.49 | d,d            | 17.3, 7.6        |
|                    |       | 1.21 | d,d            | 17.1, 5.1        |
| Ala NH             |       | 6.86 | d              | 6.1              |
| 1                  | 169.9 |      |                |                  |
| 2                  | 48.5  | 4.24 | m              |                  |
| 3                  | 14.1  | 1.42 | d              | 6.9              |
| DeAla NH           |       | 9.46 | s              |                  |
| 1                  | 168.2 |      |                |                  |
| 2                  | 136.9 |      |                |                  |
| 3                  | 99.8  | 5.02 | m              |                  |
|                    |       | 4.73 | m              |                  |
| Sarc               |       |      |                |                  |
| 1                  | 166.7 |      |                |                  |
| 2                  | 51.2  | 4.97 | d              | 16.9             |
|                    |       | 3.39 | m              |                  |
| CH <sub>3</sub> -N | 37.5  | 3.11 | s              |                  |

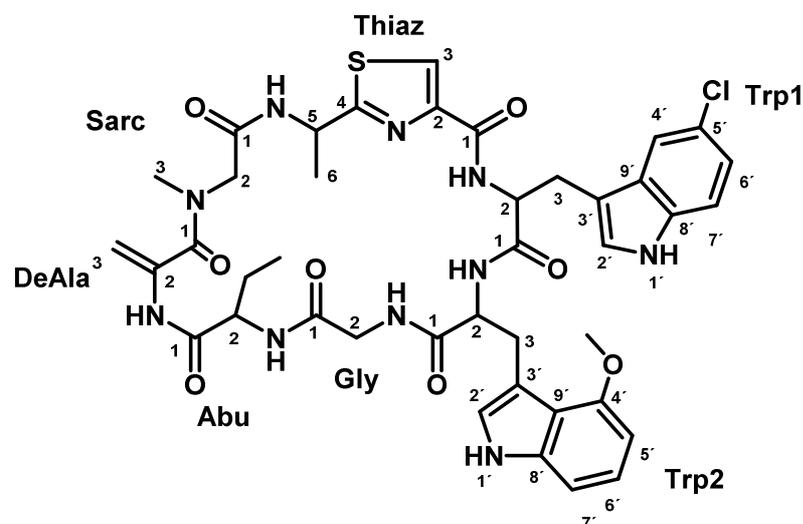
<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

<sup>e</sup> obtained from <sup>13</sup>C NMR spectrum at 175 MHz

**5-Cl-argyrin B****Figure S18.** Structure formula of 5-Cl-argyrin B. Atom numbers are assigned according to the NMR data.**Table S22.** List of all observed NMR correlations for 5-Cl-argyrin B in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|----------|-----------------------------|----------------------|------|-----------|
| Thiaz NH |                             | 8.75                 | d    | 8.3       |
| 1        | 159.7                       |                      |      |           |
| 2        | 150.1                       |                      |      |           |
| 3        | 123.0                       | 8.04                 | s    |           |
| 4        | 170.7                       |                      |      |           |
| 5        | 45.4                        | 5.46                 | m    |           |
| 6        | 20.5                        | 1.73                 | d    | 6.9       |
|          |                             |                      |      |           |
| Trp1 NH  |                             | 8.64                 | d    | 7.1       |
| 1        | 172.7                       |                      |      |           |
| 2        | 52.2                        | 4.99                 | m    |           |
| 3        | 26.9                        | 3.52                 | m    |           |
|          |                             | 2.62                 | d,d  | 15.4, 3.3 |
| 1' NH    |                             | 10.99                | s    |           |
| 2'       | 127.3                       | 7.01                 | m    |           |
| 3'       | 105.7                       |                      |      |           |
| 4'       | 115.2                       | 5.73                 | d    | 2         |
| 5'       | 127.6                       |                      |      |           |
| 6'       | 122.0                       | 6.93                 | m    |           |
| 7'       | 112.7                       | 7.00                 | m    |           |
| 8'       | 133.4                       |                      |      |           |
| 9'       | 125.2                       |                      |      |           |
|          |                             |                      |      |           |
| Trp2 NH  |                             | 8.80                 | s    |           |
| 1        | 170.1                       |                      |      |           |
| 2        | 57.8                        | 4.24                 | m    |           |
| 3        | 27.1                        | 3.52                 | m    |           |
|          |                             | 3.34                 | d,d  | 14.9, 3.6 |
| 1' NH    |                             | 8.34                 | s    |           |
| 2'       | 123.9                       | 6.88                 | m    |           |
| 3'       | 108.0                       |                      |      |           |



**Table S23.** List of all observed NMR correlations for 5-Br-argyrin B in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|----------|-----------------------------|----------------------|------|-----------|
| Thiaz NH |                             | 8.74                 | m    |           |
| 1        | 159.7                       |                      |      |           |
| 2        | 150.1                       |                      |      |           |
| 3        | 123.1                       | 8.04                 | s    |           |
| 4        | 170.8                       |                      |      |           |
| 5        | 45.3                        | 5.48                 | d,q  | 7.3, 7.6  |
| 6        | 20.4                        | 1.73                 | d    | 7.1       |
|          |                             |                      |      |           |
| Trp1 NH  |                             | 8.61                 | m    |           |
| 1        | 172.8                       |                      |      |           |
| 2        | 52.2                        | 4.98                 | m    |           |
| 3        | 27.1                        | 3.53                 | m    |           |
|          |                             | 2.62                 | d,d  | 15.2, 3.0 |
| 1' NH    |                             | 11.04                | s    |           |
| 2'       | 127.3                       | 7.00                 | m    |           |
| 3'       | 105.4                       |                      |      |           |
| 4'       | 118.2                       | 6.02                 | s    |           |
| 5'       | 112.7                       |                      |      |           |
| 6'       | 124.8                       | 7.08                 | d,d  | 8.7,1.8   |
| 7'       | 113.4                       | 6.98                 | d    | 8.5       |
| 8'       | 133.7                       |                      |      |           |
| 9'       | 128.0                       |                      |      |           |
|          |                             |                      |      |           |
| Trp2 NH  |                             | 8.76                 | m    |           |
| 1        | 170.2                       |                      |      |           |
| 2        | 57.4                        | 4.24                 | m    |           |
| 3        | 27.2                        | 3.52                 | m    |           |
|          |                             | 3.35                 | d,d  | 14.7, 4.0 |
| 1' NH    |                             | 8.37                 | m    |           |
| 2'       | 123.8                       | 7.90                 | m    |           |
| 3'       | 108.3                       |                      |      |           |
| 4'       | 152.2                       |                      |      |           |
| 5'       | 101.3                       | 6.87                 | d,d  | 2.2,6.1   |
| 6'       | 123.4                       | 7.32 <sup>d</sup>    | m    |           |
| 7'       | 108.4                       | 7.32 <sup>d</sup>    | m    |           |
| 8'       | 138.4                       |                      |      |           |
| 9'       | 117.2                       |                      |      |           |
| 4'OMe    | 56.4                        | 4.30                 | s    |           |
|          |                             |                      |      |           |
| Gly NH   |                             | 4.78                 | d,d  | 5.1,7.9   |
| 1        | 171.3                       |                      |      |           |
| 2        | 41.6                        | 3.69                 | d,d  | 7.9,17.1  |
|          |                             | 1.17                 | d,d  | 5.1,17.6  |
|          |                             |                      |      |           |
| Abu NH   |                             | 6.91                 | m    |           |
| 1        | 169.5                       |                      |      |           |
| 2        | 54.7                        | 3.98                 | t,d  | 8.6, 6.3  |
| 3        | 21.3                        | 1.99                 | m    |           |
|          |                             | 1.89                 | m    |           |
| 4        | 10.6                        | 0.88                 | t    | 7.4       |
|          |                             |                      |      |           |

|                    |       |      |   |      |
|--------------------|-------|------|---|------|
| DeAla NH           |       | 9.31 |   |      |
| 1                  | 168.7 |      |   |      |
| 2                  | 136.9 |      |   |      |
| 3                  | 99.7  | 5.00 | d | 1.5  |
|                    |       | 4.72 | d | 1.4  |
| Sarc               |       |      |   |      |
| 1                  | 166.9 |      |   |      |
| 2                  | 51.2  | 4.94 | d |      |
|                    |       | 3.41 | d | 17.0 |
| CH <sub>3</sub> -N | 37.5  | 3.09 | s |      |

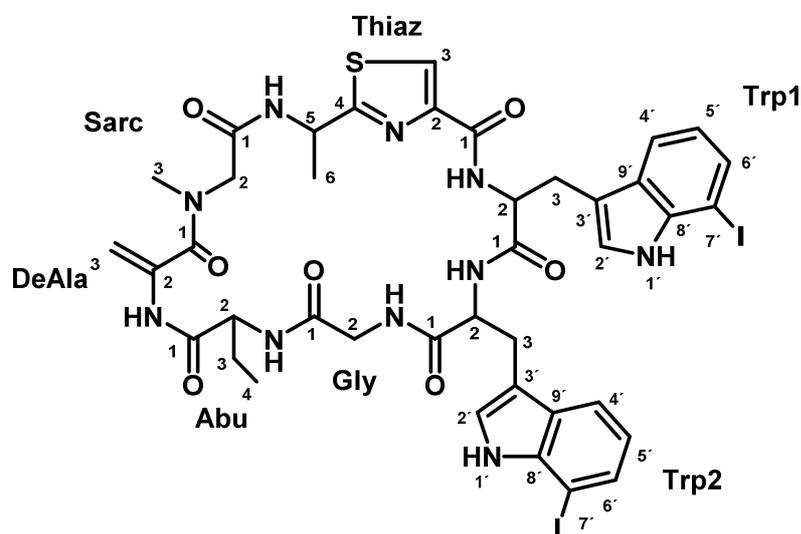
<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

### 7,7'-di-I-argyirin I



**Figure S20.** Structure formula of 7,7'-di-I-argyirin I. Atom numbers are assigned according to the NMR data.

**Table S24.** List of all observed NMR correlations for 7,7'-di-I-argyirin I in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)    |
|----------|-----------------------------|----------------------|------|----------|
| Thiaz NH |                             | 8.70                 | d    | 8.0      |
| 1        | 159.1                       |                      |      |          |
| 2        | 149.9                       |                      |      |          |
| 3        | 122.8                       | 8.02                 | s    |          |
| 4        | 171.3                       |                      |      |          |
| 5        | 45.6                        | 5.40                 | m    |          |
| 6        | 20.3                        | 1.75                 | d    | 7.2      |
|          |                             |                      |      |          |
| Trp1 NH  |                             | 8.61                 | d    | 7.0      |
| 1        | 172.5                       |                      |      |          |
| 2        | 52.0                        | 5.16                 | m    |          |
| 3        | 27.4                        | 3.49                 | d,d  | 15.9,2.9 |
|          |                             | 2.82                 | m    |          |
| 1' NH    |                             | 10.86                | s    |          |
| 2'       | 125.8                       | 7.17                 | d    | 2.2      |
| 3'       | 107.4                       |                      |      |          |

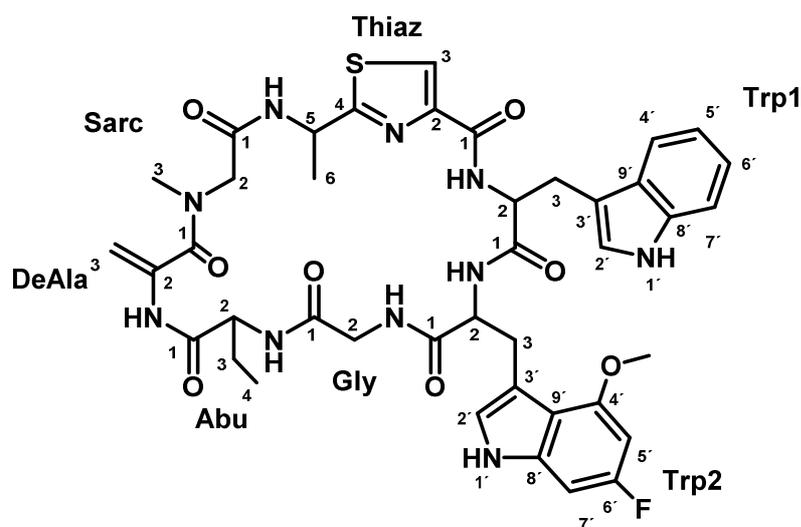
|                    |       |                   |     |           |
|--------------------|-------|-------------------|-----|-----------|
| 4'                 | 115.8 | 6.12              | d   | 7.2       |
| 5'                 | 121.5 | 6.36              | d,d | 7.7,7.7   |
| 6'                 | 130.5 | 7.36              | d   | 7.5       |
| 7'                 | 77.1  |                   |     |           |
| 8'                 | 136.9 |                   |     |           |
| 9'                 | 126.4 |                   |     |           |
| Trp2 NH            |       | 6.55 <sup>d</sup> | m   |           |
| 1                  | 169.8 |                   |     |           |
| 2                  | 56.3  | 4.28              | m   |           |
| 3                  | 26.7  | 3.36              | d,d | 4.6,15.3  |
|                    |       | 3.22              | d,d | 4.7,14.9  |
| 1' NH              |       | 8.29              | s   |           |
| 2'                 | 123.6 | 7.26 <sup>d</sup> | m   |           |
| 3'                 | 110.7 |                   |     |           |
| 4'                 | 117.9 | 7.55              | d   | 7.6       |
| 5'                 | 122.6 | 7.06              | d,d | 7.7,7.7   |
| 6'                 | 132.0 | 7.75              | d   | 7.4       |
| 7'                 | 77.4  |                   |     |           |
| 8'                 | 138.4 |                   |     |           |
| 9'                 | 127.4 |                   |     |           |
| Gly NH             |       | 4.42              | m   |           |
| 1                  | 170.8 |                   |     |           |
| 2                  | 41.4  | 3.63              | d,d | 17.3, 7.5 |
|                    |       | 1.27              | m   |           |
| Abu NH             |       | 6.55 <sup>d</sup> | m   |           |
| 1                  | 169.9 |                   |     |           |
| 2                  | 54.5  | 4.06              | t,d | 9.4, 6.1  |
| 3                  | 21.3  | 2.02              | m   |           |
|                    |       | 1.88              | m   |           |
| 4                  | 10.5  | 0.90              | t   | 7.3       |
| DeAla NH           |       | 9.35              | s   |           |
| 1                  | 168.3 |                   |     |           |
| 2                  | 135.5 |                   |     |           |
| 3                  | 101.9 | 5.04              | d   | 1.6       |
|                    |       | 4.76              | m   |           |
| Sarc               |       |                   |     |           |
| 1                  | 167.0 |                   |     |           |
| 2                  | 50.9  | 4.96              | d   | 16.9      |
|                    |       | 3.40              | d   | 16.9      |
| CH <sub>3</sub> -N | 37.9  | 3.18              | s   |           |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

**6'-F-argyrin B****Figure S21.** Structure formula of 6'-F-argyrin B. Atom numbers are assigned according to the NMR data.**Table S25.** List of all observed NMR correlations for 6'-F-argyrin B in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)    |
|----------|-----------------------------|----------------------|------|----------|
| Thiaz NH |                             | 8.81                 | d    | 8.6      |
| 1        | 159.9                       |                      |      |          |
| 2        | 150.5                       |                      |      |          |
| 3        | 122.9                       | 8.05                 | s    |          |
| 4        | 171.0                       |                      |      |          |
| 5        | 45.4                        | 5.49                 | m    |          |
| 6        | 20.7                        | 1.72                 | d    | 6.9      |
|          |                             |                      |      |          |
| Trp1 NH  |                             | 8.58                 | d    | 6.9      |
| 1        | 173.1                       |                      |      |          |
| 2        | 52.3                        | 5.05                 | m    |          |
| 3        | 26.4                        | 3.57                 | d,d  | 15.5,3.4 |
|          |                             | 2.84                 | d,d  | 15.3,3.0 |
| 1' NH    |                             | 10.68                | s    |          |
| 2'       | 125.8                       | 6.98                 | m    |          |
| 3'       | 105.5                       |                      |      |          |
| 4'       | 115.9                       | 5.65                 | d    | 7.9      |
| 5'       | 119.4                       | 6.41                 | d,d  | 7.4,7.4  |
| 6'       | 121.6                       | 6.92                 | d,d  | 7.5,7.5  |
| 7'       | 111.7                       | 7.08                 | d    | 8.0      |
| 8'       | 135.0                       |                      |      |          |
| 9'       | 126.4                       |                      |      |          |
|          |                             |                      |      |          |
| Trp2 NH  |                             | 8.43                 | s    |          |
| 1        | 170.1                       |                      |      |          |
| 2        | 57.9                        | 4.20                 | m    |          |
| 3        | 26.9                        | 3.44                 | d,d  | 14.9,3.6 |
|          |                             | 3.29                 | d,d  | 15.0,4.0 |
| 1' NH    |                             | 8.30                 | s    |          |
| 2'       | 123.8                       | 6.81                 | m    |          |
| 3'       | 108.8                       |                      |      |          |
| 4'       | 152.8                       |                      |      |          |

|                    |       |      |                |                  |
|--------------------|-------|------|----------------|------------------|
| 5'                 | 92.2  | 6.73 | d              | 10.8             |
| 6'                 | 160.8 |      | d <sup>e</sup> | 240 <sup>e</sup> |
| 7'                 | 92.7  | 7.02 | d              | 8.3              |
| 8'                 | 137.2 |      |                |                  |
| 9'                 | 114.2 |      |                |                  |
| 4'OMe              | 56.5  | 4.32 | s              |                  |
|                    |       |      |                |                  |
| Gly NH             |       | 4.58 | m              |                  |
| 1                  | 171.4 |      |                |                  |
| 2                  | 407   | 3.51 | d,d            | 17.1,7.8         |
|                    |       | 1.12 | d,d            | 17.2,5.1         |
|                    |       |      |                |                  |
| Abu NH             |       | 6.79 | m              |                  |
| 1                  | 169.8 |      |                |                  |
| 2                  | 54.5  | 3.98 | m              |                  |
| 3                  | 21.3  | 1.99 | m              |                  |
|                    |       | 1.88 | m              |                  |
| 4                  | 10.7  | 0.88 | t              | 7.3              |
|                    |       |      |                |                  |
| DeAla NH           |       | 9.35 | s              |                  |
| 1                  | 168.4 |      |                |                  |
| 2                  | 137.1 |      |                |                  |
| 3                  | 99.6  | 5.00 | m              |                  |
|                    |       | 4.72 | m              |                  |
|                    |       |      |                |                  |
| Sarc               |       |      |                |                  |
| 1                  | 166.9 |      |                |                  |
| 2                  | 51.1  | 4.98 | d              | 16.9             |
|                    |       | 3.39 | d              | 16.8             |
| CH <sub>3</sub> -N | 37.4  | 3.10 | s              |                  |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

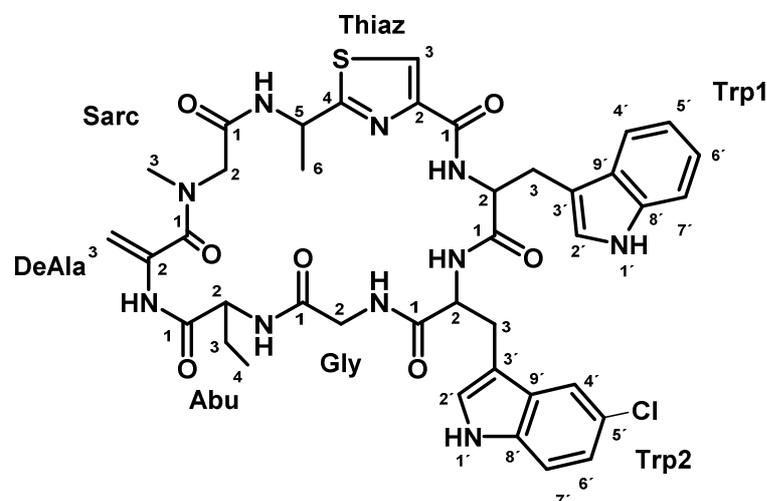
<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

<sup>e</sup> obtained from <sup>13</sup>C NMR spectrum at 125 MHz

### 5'-Cl-argyrin I



**Figure S22.** Structure formula of 5'-Cl-argyrin I. Atom numbers are assigned according to the NMR data.

**Table S26.** List of all observed NMR correlations for 5'-Cl-argyrin I in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)     |
|----------|-----------------------------|----------------------|------|-----------|
| Thiaz NH |                             | 8.84                 | d    | 8.6       |
| 1        | 159.9                       |                      |      |           |
| 2        | 150.0                       |                      |      |           |
| 3        | 122.7                       | 8.02                 | s    |           |
| 4        | 171.0                       |                      |      |           |
| 5        | 45.1                        | 5.48                 | d,q  | 7.5,7.1   |
| 6        | 20.4                        | 1.71                 | d    | 7.2       |
| Trp1 NH  |                             | 8.61                 | d    | 7.5       |
| 1        | 172.9                       |                      |      |           |
| 2        | 51.9                        | 5.14                 | m    |           |
| 3        | 27.0                        | 3.52                 | d,d  | 15.5,3.3  |
|          |                             | 3.05                 | d,d  | 15.4,3.1  |
| 1' NH    |                             | 10.70                | s    |           |
| 2'       | 125.4                       | 7.01                 | m    |           |
| 3'       | 105.6                       |                      |      |           |
| 4'       | 115.4                       | 6.45 <sup>d</sup>    | m    |           |
| 5'       | 119.4                       | 6.45 <sup>d</sup>    | m    |           |
| 6'       | 121.5                       | 7.00                 | m    |           |
| 7'       | 111.7                       | 7.15                 | d    | 8.1       |
| 8'       | 134.8                       |                      |      |           |
| 9'       | 126.4                       |                      |      |           |
| Trp2 NH  |                             | 6.63                 | s    |           |
| 1        | 170.1                       |                      |      |           |
| 2        | 55.5                        | 4.24                 | m    |           |
| 3        | 26.4                        | 3.34                 | d,d  | 15.1,5.4  |
|          |                             | 3.25                 | d,d  | 15.0,5.3  |
| 1' NH    |                             | 8.45                 | s    |           |
| 2'       | 124.5                       | 7.37                 | m    |           |
| 3'       | 109.0                       |                      |      |           |
| 4'       | 117.9                       | 7.49 <sup>d</sup>    | m    |           |
| 5'       | 126.5                       |                      |      |           |
| 6'       | 123.5                       | 7.32                 | d,d  | 8.8,1.7   |
| 7'       | 112.6                       | 7.49 <sup>d</sup>    | m    |           |
| 8'       | 134.9                       |                      |      |           |
| 9'       | 128.2                       |                      |      |           |
| Gly NH   |                             | 4.83                 | d,d  | 7.4,5.4   |
| 1        | 171.1                       |                      |      |           |
| 2        | 40.9                        | 3.65                 | d,d  | 17.4, 7.9 |
|          |                             | 1.29                 | d,d  | 17.4,5.0  |
| Abu NH   |                             | 6.69                 | d    | 6.7       |
| 1        | 169.2                       |                      |      |           |
| 2        | 54.2                        | 4.02                 | d,t  | 8.9,6.4   |
| 3        | 20.9                        | 1.97                 | m    |           |
|          |                             | 1.87                 | m    |           |
| 4        | 10.4                        | 0.88                 | t    | 7.3       |
| DeAla NH |                             | 9.37                 | s    |           |
| 1        | 168.2                       |                      |      |           |

|                    |       |      |   |      |
|--------------------|-------|------|---|------|
| 2                  | 136.8 |      |   |      |
| 3                  | 99.5  | 5.04 | d |      |
|                    |       | 4.75 | d |      |
|                    |       |      |   |      |
| Sarc               |       |      |   |      |
| 1                  | 166.9 |      |   |      |
| 2                  | 50.9  | 4.99 | d | 16.9 |
|                    |       | 3.42 | d | 16.9 |
| CH <sub>3</sub> -N | 37.2  | 3.11 | s |      |

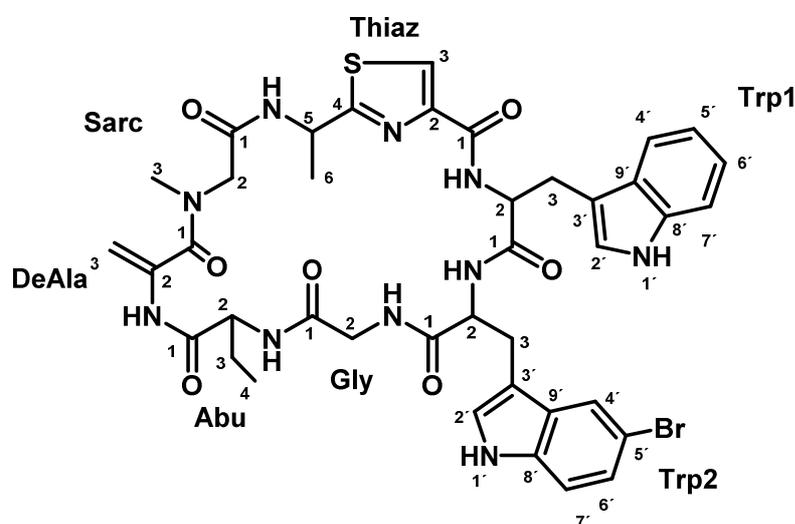
<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

### 5'-Br-argyrin I



**Figure S23.** Structure formula of 5'-Br-argyrin I. Atom numbers are assigned according to the NMR data.

**Table S27.** List of all observed NMR correlations for 5'-Br-argyrin I in CDCl<sub>3</sub>.

| Res      | $\Delta$ <sup>13</sup> C <sup>a,c</sup> | $\Delta$ <sup>1</sup> H <sup>b</sup> | Mult | J(Hz)    |
|----------|---|--------------------------------------|------|----------|
| Thiaz NH |   | 8.81                                 | d    | 8.6      |
| 1        | 159.9                                   |                                      |      |          |
| 2        | 150.0                                   |                                      |      |          |
| 3        | 123.0                                   | 8.01                                 | s    |          |
| 4        | 171.1                                   |                                      |      |          |
| 5        | 45.5                                    | 5.49                                 | d,q  | 7.6,7.3  |
| 6        | 20.8                                    | 1.71                                 | d    | 7.2      |
|          |   |                                      |      |          |
| Trp1 NH  |   | 8.58                                 | d    | 7.3      |
| 1        | 172.7                                   |                                      |      |          |
| 2        | 52.3                                    | 5.11                                 | m    |          |
| 3        | 27.2                                    | 3.55                                 | d,d  | 15.5,3.5 |
|          |   | 3.00                                 | d,d  | 15.4,3.2 |
| 1' NH    |   | 10.69                                | s    |          |
| 2'       | 125.6                                   | 7.04 <sup>d</sup>                    | m    |          |
| 3'       | 105.6                                   |                                      |      |          |
| 4'       | 115.8                                   | 6.46                                 | m    |          |
| 5'       | 119.9                                   | 6.48                                 | m    |          |

|                    |       |                   |     |           |
|--------------------|-------|-------------------|-----|-----------|
| 6'                 | 121.9 | 7.04 <sup>d</sup> | m   |           |
| 7'                 | 112.1 | 7.17              | d   | 8.0       |
| 8'                 | 134.9 |                   |     |           |
| 9'                 | 126.4 |                   |     |           |
|                    |       |                   |     |           |
| Trp2 NH            |       | 6.63              | s   |           |
| 1                  | 170.0 |                   |     |           |
| 2                  | 55.8  | 4.24              | m   |           |
| 3                  | 26.8  | 3.33              | d,d | 15.3,5.6  |
|                    |       | 3.23              | m   |           |
| 1' NH              |       | 8.48              | s   |           |
| 2'                 | 124.5 | 7.27              | m   |           |
| 3'                 | 109.2 |                   |     |           |
| 4'                 | 121.4 | 7.67              | m   |           |
| 5'                 | 113.9 |                   |     |           |
| 6'                 | 126.4 | 7.49 <sup>d</sup> | m   |           |
| 7'                 | 113.4 | 7.49 <sup>d</sup> | m   |           |
| 8'                 | 135.6 |                   |     |           |
| 9'                 | 128.8 |                   |     |           |
|                    |       |                   |     |           |
| Gly NH             |       | 4.84              | m   |           |
| 1                  | 171.1 |                   |     |           |
| 2                  | 41.3  | 3.68              | d,d | 17.5, 8.0 |
|                    |       | 1.37              | d,d | 17.1,4.2  |
|                    |       |                   |     |           |
| Abu NH             |       | 6.68              | d   | 6.6       |
| 1                  | 169.3 |                   |     |           |
| 2                  | 54.5  | 4.03              | d,t | 8.6,6.4   |
| 3                  | 21.3  | 1.98              | m   |           |
|                    |       | 1.86              | m   |           |
| 4                  | 10.7  | 0.88              | t   | 7.5       |
|                    |       |                   |     |           |
| DeAla NH           |       | 9.37              | s   |           |
| 1                  | 168.1 |                   |     |           |
| 2                  | 136.8 |                   |     |           |
| 3                  | 99.9  | 5.04              | d   | 1.4       |
|                    |       | 4.75              | d   |           |
|                    |       |                   |     |           |
| Sarc               |       |                   |     |           |
| 1                  | 166.8 |                   |     |           |
| 2                  | 51.4  | 4.99              | d   | 16.7      |
|                    |       | 3.41              | d   | 16.8      |
| CH <sub>3</sub> -N | 37.2  | 3.11              | s   |           |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

<sup>d</sup> overlapped signals.

## 7,7'-di-I-argyrin B

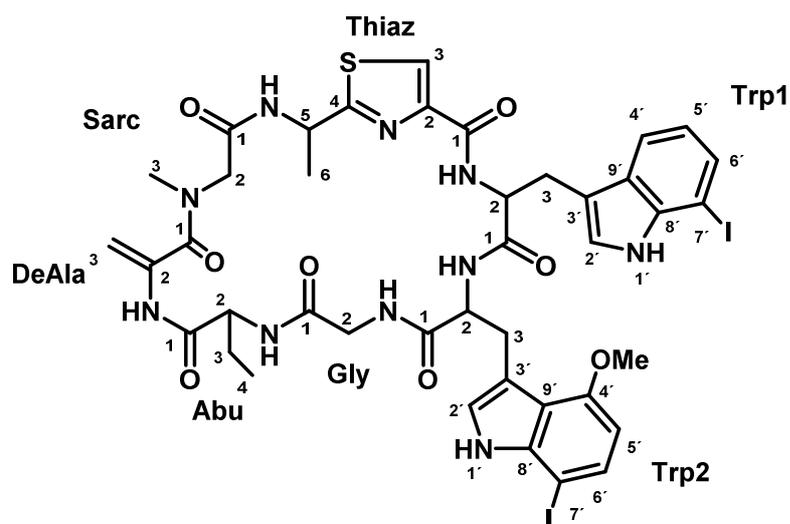


Figure S24. Structure formula of 7,7'-di-I-argyrin B. Atom numbers are assigned according to the NMR data.

Table S28. List of all observed NMR correlations for 7,7'-di-I-argyrin B in CDCl<sub>3</sub>.

| Res      | $\Delta^{13}\text{C}^{a,c}$ | $\Delta^1\text{H}^b$ | Mult | J(Hz)    |
|----------|-----------------------------|----------------------|------|----------|
| Thiaz NH |                             | 8.71                 | d    | 7.9      |
| 1        | 159.8                       |                      |      |          |
| 2        | 149.9                       |                      |      |          |
| 3        | 122.9                       | 8.04                 | s    |          |
| 4        | 171.1                       |                      |      |          |
| 5        | 45.6                        | 5.39                 | d,q  | 7.4,7.1  |
| 6        | 20.1                        | 1.75                 | d    | 7.2      |
|          |                             |                      |      |          |
| Trp1 NH  |                             | 8.64                 | d    | 6.9      |
| 1        | 172.7                       |                      |      |          |
| 2        | 52.0                        | 5.05                 | m    |          |
| 3        | 27.3                        | 3.52                 | m    |          |
|          |                             | 2.72                 | d,d  | 15.2,3.7 |
| 1' NH    |                             | 10.87                | m    |          |
| 2'       | 125.8                       | 7.13                 | d    | 2.7      |
| 3'       | 107.8                       |                      |      |          |
| 4'       | 116.0                       | 5.33                 | d    | 7.3      |
| 5'       | 121.1                       | 6.19                 | d,d  | 7.7,7.7  |
| 6'       | 130.4                       | 7.32                 | d    | 7.9      |
| 7'       | 76.7                        |                      |      |          |
| 8'       | 136.8                       |                      |      |          |
| 9'       | 126.5                       |                      |      |          |
|          |                             |                      |      |          |
| Trp2 NH  |                             | 8.57                 | s    |          |
| 1        | 169.8                       |                      |      |          |
| 2        | 57.6                        | 4.20                 | m    |          |
| 3        | 27.1                        | 3.43                 | d,d  | 15.3,4.3 |
|          |                             | 3.28                 | d,d  | 15.1,3.7 |
| 1' NH    |                             | 8.34                 | m    |          |
| 2'       | 123.8                       | 6.86                 | d    | 2.2      |
| 3'       | 110.0                       |                      |      |          |
| 4'       | 152.9                       |                      |      |          |

|                    |       |      |     |          |
|--------------------|-------|------|-----|----------|
| 5'                 | 103.6 | 6.78 | d   | 8.2      |
| 6'                 | 132.0 | 7.69 | d   | 8.2      |
| 7'                 | 69.2  |      |     |          |
| 8'                 | 139.5 |      |     |          |
| 9'                 | 117.2 |      |     |          |
| 4'-OMe             | 56.6  | 4.33 | s   |          |
|                    |       |      |     |          |
| Gly NH             |       | 4.27 | m   |          |
| 1                  | 171.0 |      |     |          |
| 2                  | 41.0  | 3.54 | d,d | 16.9,7.6 |
|                    |       | 1.09 | d,d | 16.8,5.4 |
|                    |       |      |     |          |
| Abu NH             |       | 6.72 | d   | 6.2      |
| 1                  | 170.1 |      |     |          |
| 2                  | 54.5  | 4.04 | d,t | 9.7,6.1  |
| 3                  | 21.4  | 2.01 | m   |          |
|                    |       | 1.89 | m   |          |
| 4                  | 10.7  | 0.89 | t   | 7.4      |
|                    |       |      |     |          |
| DeAla NH           |       | 9.37 | s   |          |
| 1                  | 168.4 |      |     |          |
| 2                  | 135.6 |      |     |          |
| 3                  | 102.0 | 5.01 | d   | 1.4      |
|                    |       | 4.75 | m   |          |
|                    |       |      |     |          |
| Sarc               |       |      |     |          |
| 1                  | 167.2 |      |     |          |
| 2                  | 50.9  | 4.94 | d   | 17.1     |
|                    |       | 3.39 | d   | 16.9     |
| CH <sub>3</sub> -N | 38.0  | 3.17 | s   |          |

<sup>a</sup> acquired at 175 MHz and assigned from 2D NMR spectra, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  77.16 ppm.

<sup>b</sup> acquired at 700 MHz, referenced to solvent signal CDCl<sub>3</sub> at  $\delta$  7.26 ppm.

<sup>c</sup> proton showing HMBC correlations to indicated carbons.

### 5.5.3 References

- (1) Sambrook, J.; Russell, D. W. *Molecular cloning: A laboratory manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.
- (2) Zhang, Y.; Buchholz, F.; Muyrers, J. P.; Stewart, F. A. A new logic for DNA engineering using recombination in *Escherichia coli*, *Nat. Genet.* **1998**, *20*, pp. 123–128.
- (3) Pogorevc, D.; Tang, Y.; Hoffmann, M.; Zipf, G.; Bernauer, H. S.; Popoff, A.; Steinmetz, H.; Wenzel, S. C. Biosynthesis and heterologous production of argyriins. **2019**, *submitted manuscript*.
- (4) Corr, M. J.; Smith, D. R.; Goss, R. J. One-pot access to 1-5,6-dihalotryptophans and 1-alkanyltryptophans using tryptophan synthase, *Tetrahedron.* **2016**, *72*, pp. 7306–7310.
- (5) Kawasaki, H.; Bauerle, R.; Zon, G.; Ahmed, S. A.; Miles, E. W. Site-specific mutagenesis of the alpha subunit of tryptophan synthase from *Salmonella typhimurium*. Changing arginine 179 to leucine alters the reciprocal transmission of substrate-induced conformational changes between the alpha and beta 2 subunits, *J. Biol. Chem.* **1987**, *262*, pp. 10678–10683.
- (6) Ruvinov, S. B.; Yang, X.-J.; Parris, K. D.; Banik, U.; Ahmed, S. A.; Miles, E. W.; Sackett, D. L. Ligand-mediated Changes in the Tryptophan Synthase Indole Tunnel Probed by Nile Red Fluorescence with Wild Type, Mutant, and Chemically Modified Enzymes, *J. Biol. Chem.* **1995**, *270*, pp. 6357–6369.
- (7) Wenzel, S. C.; Hoffmann, H.; Zhang, J.; Debussche, L.; Haag-Richter, S.; Kurz, M.; Nardi, F.; Lukat, P.; Kochems, I.; Tietgen, H.; Schummer, D.; Nicolas, J.-P.; Calvet, L.; Czepczor, V.; Vrignaud, P.; Mühlenweg, A.; Pelzer, S.; Müller, R.; Brönstrup, M. Production of the bengamide class of marine natural products in myxobacteria: biosynthesis and structure-activity relationships, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, pp. 15560–15564.
- (8) Sucipto, H.; Pogorevc, D.; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, *44*, pp. 160–170.
- (9) Pogorevc, D.; Panter, F.; Schillinger, C.; Jansen, R.; Wenzel, S. C.; Müller, R. Production optimization and biosynthesis revision of Coralopyronin A, a potent anti-filarial antibiotic. **2019**, *submitted manuscript*.
- (10) Oßwald, C.; Zaburanyi, N.; Burgard, C.; Hoffmann, T.; Wenzel, S. C.; Müller, R. A highly unusual polyketide synthase directs dawenol polyene biosynthesis in *Stigmatella aurantiaca*, *J. Biotechnol.* **2014**, *191*, pp. 54–63.
- (11) Julien, B.; Shah, S. Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*, *Antimicrob. Agents Chemother.* **2002**, *46*, pp. 2772–2778.
- (12) Lau, J.; Frykman, S.; Regentin, R.; Ou, S.; Tsuruta, H.; Licari, P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*, *Biotechnol. Bioeng.* **2002**, *78*, pp. 280–288.
- (13) Oßwald, C.; Zipf, G.; Schmidt, G.; Maier, J.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Modular Construction of a Functional Artificial Epothilone Polyketide Pathway, *ACS Synth. Biol.* **2014**, *3*, pp. 759–772.
- (14) Yan, F.; Burgard, C.; Popoff, A.; Zaburanyi, N.; Zipf, G.; Maier, J.; Bernauer, H. S.; Wenzel, S. C.; Müller, R. Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production, *Chem. Sci.* **2018**, *9*, pp. 7510–7519.

- (15) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (16) Perlova, O.; Fu, J.; Kuhlmann, S.; Krug, D.; Stewart, F.; Zhang, Y.; Müller, R. Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2006**, *72*, pp. 7485–7494.
- (17) Stevens, D. C.; Henry, M. R.; Murphy, K. A.; Boddy, C. N. Heterologous expression of the oxytetracycline biosynthetic pathway in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2010**, *76*, pp. 2681–2683.
- (18) Chai, Y.; Shan, S.; Weissman, K. J.; Hu, S.; Zhang, Y.; Müller, R. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis, *Chem. Biol.* **2012**, *19*, pp. 361–371.
- (19) Gemperlein, K.; Rachid, S.; Garcia, R. O.; Wenzel, S. C.; Müller, R. Polyunsaturated fatty acid biosynthesis in myxobacteria. Different PUFA synthases and their product diversity, *Chem. Sci.* **2014**, *5*, pp. 1733–1741.
- (20) Yan, F.; Auerbach, D.; Chai, Y.; Keller, L.; Tu, Q.; Hüttel, S.; Glemser, A.; Grab, H. A.; Bach, T.; Zhang, Y.; Müller, R. Biosynthesis and Heterologous Production of Vioprolides: Rational Biosynthetic Engineering and Unprecedented 4-Methylazetidinecarboxylic Acid Formation, *Angew. Chem. Int. Ed. Engl.* **2018**.



## 6 Discussion

The projects discussed in this thesis focus on several aspects of natural product (NP) research, with the main emphasis on the structure and yield improvement. The achieved results can be split in two parts. The first part, discussed in the chapters 2 and 3, focuses on the establishment of heterologous production platforms of myxobacterial  $\alpha$ -pyrone antibiotics myxopyronin (MXN) and coralopyronin (COR) and their yield optimization. Moreover, the heterologous expression of COR biosynthetic gene cluster (BGC) enabled detailed investigation of the biosynthesis, which led to the isolation and structure elucidation of two novel derivatives. The second part, discussed in chapters 4 and 5, describes the argyrimycin (ARG) biosynthetic machinery from the native producer *Cystobacter* sp. SBCb004 and its heterologous expression in *Myxococcus xanthus* DK1622. It also focuses on structure optimization of ARG by adenylation (A) domain engineering and precursor-directed biosynthesis. Furthermore, yield improvement by medium optimization and promoter engineering are discussed.

### 6.1 Yield improvement and production profile engineering of microbial natural products

Myxobacteria have steadily been gaining importance in the field of NPs over the last decades with more than 67 distinct core structures and some 500 derivatives reported from approximately 7500 myxobacterial strains in 2010<sup>1,2</sup>. Many of the isolated secondary metabolites show interesting activities that have the potential to be implemented in the field of medicine. However, the path from a NP to an applicable medicinal drug is not only lengthy and expensive, but very low production titers of the compounds often hamper the development process. Bacteria in general are known to naturally produce secondary metabolites (e.g. antibiotics) in minor amounts, which are sufficient to provide competitive advantage over rival strains growing in the same habitat. However, those low production levels are significantly below the desired threshold required for profitable industrial production, which is fundamental for pharmaceutical companies to show interest in the target compound and launch it to the market<sup>3</sup>. In addition, the total production level of a compound class is often partitioned between multiple derivatives, which further decreases the production level of a specific analogue. Engineering of production profile is thus often necessary, to direct biosynthesis towards more prominent analogues. Production bottlenecks of bacterial secondary metabolites can have different origin. Common reasons that influence the

production profile include e.g. chromosomal location of the BGC, gene dosage, transcription, translation, correct protein folding, posttranslational activation, availability of precursors and cofactors, self-resistance of the strain etc.. Improving production of myxobacterial secondary metabolites is not trivial, as the regulation of NP biosynthesis in myxobacteria is not well understood and has only been intensely studied in a few cases<sup>4-7</sup>. Studying those regulatory networks is additionally challenging, as the regulators of myxobacterial secondary metabolism are typically not co-localized with the biosynthetic genes<sup>8</sup>. Furthermore, it is hard to evaluate the bottlenecks on the protein level related to the correct folding and efficient activation of multifunctional giant megasynthetase subunits. Bottlenecks regarding biosynthetic precursors also require insights into the biogenesis of the target compounds as well as the knowledge about the metabolic networks, however, only a few studies in myxobacteria are available on this topic<sup>9-11</sup>. Several approaches have been developed to overcome the struggle of poor metabolite production and help with engineering of the microbial production profile. Those techniques range from more traditional methods like medium and cultivation conditions optimization, to more modern and advanced metabolic engineering and heterologous expression approaches.

### 6.1.1 Heterologous expression of myxobacterial natural product megasynthetases

Myxobacterial strains are in general not easy to handle in standard laboratory conditions. Most of them are slow growing and require significant time for production of secondary metabolites, often in very low yields. Those low yields can be improved by utilization of genetic engineering techniques. However, most myxobacteria are not genetically amendable which prevents the engineering of their secondary metabolite pathways. Heterologous expression is a method commonly used for production, engineering and characterization of bacterial secondary metabolites. It provides a new and extended set of genetic engineering tools, by transferring the NP pathway to a better studied and more genetically amendable host. An overview of myxobacterial pathways, heterologously expressed in different bacterial surrogate hosts, is provided in Table 1. A variety of surrogate hosts has been employed to study the effects of heterologous expression on the production of myxobacterial NRPS/PKS pathways. Most of the given examples report successful expression in different bacteria; however, from the production yields it is obvious that in this aspect myxobacterial strains have an advantage over other bacterial hosts (Table 1). Indeed the production in *Pseudomonas sp.* in some cases did reach a few mg/L<sup>12,13</sup>, however this is still significantly lower than the yields achieved for the same compound by heterologous expression in myxobacteria. One

example is the parallel expression of myxochromide S (Mch S) and epothilone BGCs in *M. xanthus* and *P. Putida*, where a substantially higher production yield was achieved in *M. xanthus* for both metabolites<sup>14</sup>. A similar results was achieved for Vioprolides, where heterologous production in *M. xanthus* led to significantly higher production yields, compared to heterologous expression in *Burkholderia* DSM 7029 and *Pseudomonas putida* KT2440<sup>15</sup>. Promising heterologous expression of Mch S was also achieved in thermophilic myxobacterial isolate *C. macrosporus* GT-2<sup>16</sup>. The benefit of this strain over the more commonly used *M. xanthus* is its faster generation time, which is beneficial for quick and efficient production of secondary metabolites. Production of Mch S in *C. macrosporus* was measured at 600 mg/L. A similar level of production was achieved in *M. xanthus*, however, the yields cannot be directly compared due to different expression constructs (e.g. different promoter)<sup>14,16</sup>. Based on the current status of myxobacterial heterologous hosts, *M. xanthus* seems to be the best option for expression of myxobacterial NRPS/PKS pathways. It enables expression of diverse myxobacterial NP pathways in relatively high yields while being the most genetically amendable myxobacterial strain with the biggest set of available genetic engineering tool at this time. Furthermore, *M. xanthus* is the only myxobacterial strain that was so far successfully used for heterologous expression of PKS from phylogenetically distant organism without the exchange of a promoter. Expression of oxytetracycline BGC from *Streptomyces rimosus* in this case led to the production of 10 mg/L of oxytetracycline<sup>17</sup>. Despite much higher oxytetracycline production levels achieved by heterologous expression in other actionbacteria<sup>18</sup>, this work shows potential for *M. xanthus* to be used as an expression host for production of metabolites from more distant bacterial strains.

Taken together, heterologous expression is a great tool, which allowed us to perform genetic engineering of MXN, COR and ARG biosynthetic machinery, discussed in the following sections.

**Table 1.** Heterologously produced myxobacterial compounds

| Compound         | Native producer                     | Heterologous host                    | Pathway-type [size] | Promoter  | Medium                      | Yield ‡        | Ref              |
|------------------|-------------------------------------|--------------------------------------|---------------------|---|-----------------------------|----------------|------------------|
| Argyran A, B     | <i>Cystobacter species</i>          | <i>M. xanthus</i> DK1622             | NRPS [33 kb]        | <i>P<sub>nptII</sub></i>                          | M7/s4 <sup>a</sup>          | 160 mg/L       | <sup>19</sup>    |
|                  |                                     |                                      |                     | <i>P<sub>nptII</sub></i> / <i>P<sub>van</sub></i> | M7/s4 <sup>a</sup>          | ~250 mg/L      | <sup>20</sup>    |
| Bengamide        | <i>Mycrococcus virescens</i>        | <i>M. xanthus</i>                    | PKS/NRPS [25 kb]    | <i>P<sub>nptII</sub></i>                          | CTT                         | >10 mg/L       | <sup>21</sup>    |
| Corallopyronin A | <i>Corallococcus coralloides</i>    | <i>M. xanthus</i> DK1622             | PKS/NRPS [65 kb]    | <i>P<sub>nptII</sub></i>                          | M7/s6                       | 37 mg/L        | <sup>22</sup>    |
|                  |                                     |                                      |                     | <i>P<sub>van</sub></i>                            | M7/s6                       | ~100 mg/L      | <sup>23</sup>    |
| Dawenol          | <i>Stigmatella aurantiaca</i>       | <i>M. xanthus</i>                    | PKS [21 kb]         | native  | CTT                         | n.d.           | <sup>24</sup>    |
| Epothilone       | <i>Sorangium cellulosum</i> So ce90 | <i>M. xanthus</i>                    | PKS/NRPS [56 kb]    | native  | CMM                         | 0.1 - 0.4 mg/L | <sup>25,26</sup> |
|                  | <i>S. cellulosum</i> SMP44          | <i>S. coelicolor</i> CH999           | PKS/NRPS [56 kb]    | <i>actI</i>                                       | R2YE                        | 50-100 µg/L    | <sup>27</sup>    |
|                  | <i>S. cellulosum</i> So ce90        | <i>S. venezuelae</i> DHS2001         | PKS/NRPS [56 kb]    | <i>pikAI</i>                                      | R2YE                        | 0.4 µg/L       | <sup>28</sup>    |
|                  | <i>S. cellulosum</i> So ce90        | <i>E. coli</i>                       | PKS/NRPS [54 kb]    | <i>P<sub>BAD</sub></i>                            | 2xYT                        | 1 µg/L         | <sup>29</sup>    |
|                  | <i>S. cellulosum</i>                | <i>M. xanthus</i> DK1622             | PKS/NRPS [58 kb]    | <i>P<sub>nptII</sub></i>                          | CTT                         | 100 µg/L       | <sup>30</sup>    |
|                  | <i>S. cellulosum</i>                | <i>M. xanthus</i> DK1622             | PKS/NRPS [60 kb]    | <i>P<sub>nptII</sub></i>                          | CTT                         | 100 µg/L       | <sup>14</sup>    |
|                  | <i>S. cellulosum</i>                | <i>P. putida</i> KT2440              | PKS/NRPS [60 kb]    | <i>P<sub>m</sub></i>                              | LB                          | -              | <sup>14</sup>    |
| Flaviolin        | <i>S. cellulosum</i> So ce56        | <i>P. putida</i> KT2440              | PKS [1.1 kb]        | <i>P<sub>m</sub></i>                              | LB                          | 6 mg/L         | <sup>12</sup>    |
|                  |                                     | <i>P. syringae</i> pv. <i>tomato</i> |                     |   |                             | trace amounts  |                  |
| Myxochromide A   | <i>Mycrococcus xanthus</i>          | <i>M. xanthus</i> DK1622             | PKS/NRPS [29 kb]    | <i>P<sub>nptII</sub></i>                          | CTT                         | ~500 mg/L      | <sup>31</sup>    |
| Myxochromide S   | <i>Stigmatella aurantiaca</i>       | <i>M. xanthus</i>                    | PKS/NRPS [29 kb]    | <i>P<sub>nptII</sub></i>                          | CTT                         | >500 mg/L      | <sup>14</sup>    |
|                  |                                     |                                      |                     | <i>P<sub>nptII</sub></i>                          | LB                          | 0.1 mg/L       | <sup>14</sup>    |
|                  |                                     |                                      |                     | <i>P<sub>m</sub></i>                              | LB                          | 8 mg/L         | <sup>13</sup>    |
|                  |                                     |                                      |                     | <i>P<sub>aphII</sub></i>                          | M                           | 600 mg/L       | <sup>16</sup>    |
| Myxopyronin A    | <i>Mycrococcus fulvus</i>           | <i>M. xanthus</i> DK1622             | PKS/NRPS [53 kb]    | <i>P<sub>nptII</sub></i>                          | M7/s6                       | 156 mg/L       | <sup>22</sup>    |
| Myxothiazol      | <i>Stigmatella aurantiaca</i>       | <i>M. xanthus</i> DZF1               | PKS/NRPS [57 kb]    | <i>P<sub>m</sub></i>                              | CTT                         | 20 mg/L        | <sup>32</sup>    |
|                  |                                     | <i>P. putida</i> FG2005              |                     | <i>P<sub>m</sub></i>                              | Minimal medium <sup>b</sup> | 0.6 mg/L       | <sup>33</sup>    |
| Oxytetracycline  | <i>Streptomyces rimosus</i>         | <i>M. xanthus</i>                    | PKS [32 kb]         | native  | CTTYE                       | 10 mg/L        | <sup>17</sup>    |
| Pretubulysin     | <i>Cystobacter species</i>          | <i>M. xanthus</i>                    | PKS/NRPS [40 kb]    | <i>P<sub>tet</sub></i>                            | CTT <sup>c</sup>            | 0.2 mg/L       | <sup>34</sup>    |
|                  |                                     | <i>P. putida</i>                     |                     | <i>P<sub>tet</sub></i>                            | LB                          | 1.76 µg/L      | <sup>34</sup>    |
| PUFAs            | <i>Aetherobacter fasciculatus</i>   | <i>M. xanthus</i>                    | PKS/FAS [18 kb]     | <i>P<sub>tet</sub></i>                            | CTT                         | ~1 mg/CDW      | <sup>35</sup>    |
| Soraphen A       | <i>S. cellulosum</i> So ce26        | <i>Streptomyces lividans</i> ZX7     | PKS [67.5 kb]       | <i>tipA</i>                                       | YEME                        | 0.3 mg/L       | <sup>36</sup>    |
| Vioprolide       | <i>Cystobacter violaceus</i>        | <i>M. xanthus</i> DK1622             | NRPS [56 kb]        | <i>P<sub>tet</sub></i>                            | CTT                         | 500 mg/L       | <sup>15</sup>    |

n.d., not determined

<sup>a</sup> supplementation with amino acid mixture<sup>b</sup> supplementation with L-leucine and vitamin B12<sup>c</sup> 0.1 mg/L supplementation of racemic D,L-pipecolic acid

### 6.1.2 Medium optimization

Medium optimization is one of the older, traditional yield improvement methods, which is still very relevant today. It is well known that the secondary metabolite production profile of a particular strain depends on the applied cultivation parameters<sup>37</sup>. Therefore, for every microbial production process, a rational medium optimization is necessary to maximize yield and minimize production costs. Before the 1970s, media optimization was performed by classical “One-factor-at-a-time” methods that are inefficient, expensive and time consuming. However, nowadays the modern statistical techniques drastically aid the optimization process, making it more reliable and economic<sup>38</sup>. The appropriate medium components (e.g. carbon and nitrogen sources) have to be selected based on the metabolic preferences of the producer organism. Enhanced production can usually be increased by medium engineering or by strain selection/improvement, however both approaches are somewhat connected as they can affect each other. Consequently, determining the optimal flow of the experiments remains a challenging task.

Using a well-studied myxobacterial strain in the MXN/COR project (see chapter 2), simplified the medium optimization due to a better knowledge of its metabolism. However, since the M7s medium was already established for MXN production and turned out to be a promising choice, the more traditional “One-factor-at-a-time” approach was in this case applied to test an extensive library of carbon sources. The most promising results were achieved by supplementation with various oils (e.g. soy oil, vegetable oil, methyl oleate), which already showed good results in epothilone production optimization<sup>26</sup>. As the costs of the entire process has to be kept at a minimum, while still maintaining high levels of production, the impact of the used metabolites on downstream processing plays an important role. In favor of simplified compound purification, potassium acetate was finally used as the main carbon source in M7/s6 medium. This led to a significant yield improvement of MXN and COR, respectively, compared to the standard CTT medium, which is traditionally used for *M. xanthus* DK1622 cultivation (see chapter 2).

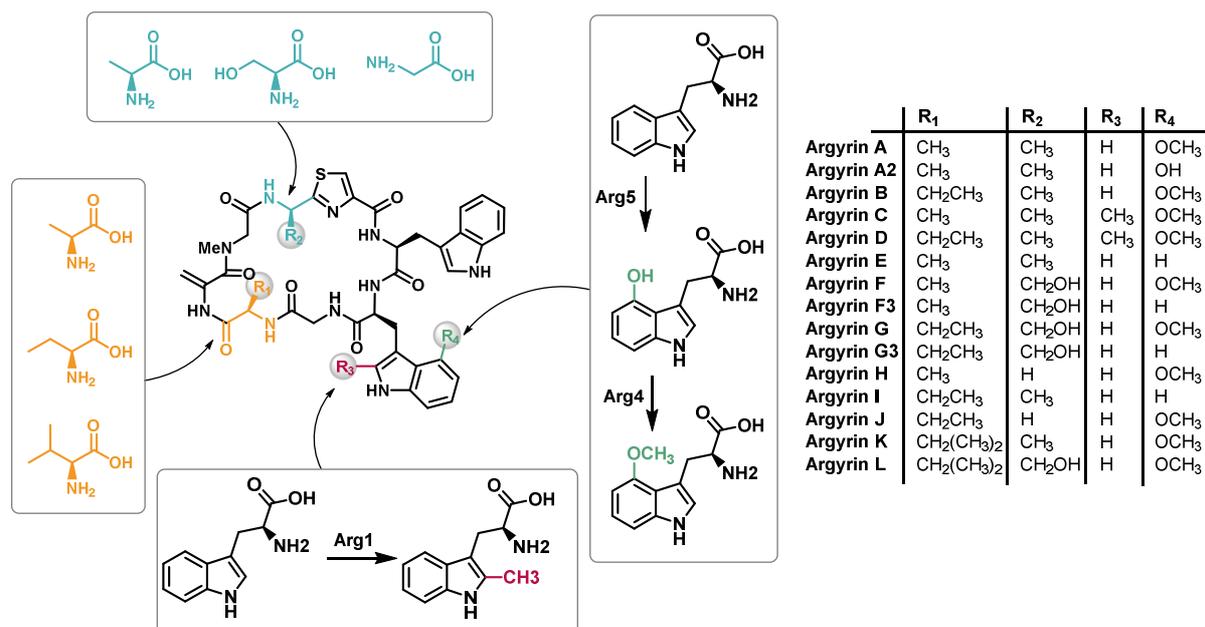
In the ARG project M7/s4 medium could be successfully improved by addition of specific amino acids, to grant a 3.5-fold increase in total ARG production by a heterologous expression system in *M. xanthus*. On the other hand, the same optimization of the potassium acetate containing M7/s6 medium resulted in a lower production, probably because potassium acetate does not directly benefit the biosynthesis of peptides such as ARG. This is a good example of how a change of a single component (e.g. amino acid source) can drastically affect

the production of fundamentally different metabolites like polyketides and peptides (see chapter 4).

Activity of PKS/NRPS megasynthetases often leads to production of various derivatives from a certain compound class, which results in a reduced production of one specific analogue (e.g. vioprolides<sup>15</sup>, disorazoles<sup>39</sup>, maltepolides<sup>40</sup>, myxalamids<sup>41</sup> epothilones<sup>42</sup> and argyins<sup>43</sup>). The main reasons for such events include e.g. relaxed substrate specificity of assembly line modules, inefficient modification steps on the assembly line as well as post-PKS/NRPS tailoring (Fig. 1). Multiple derivatives from a certain compound family are usually produced at different levels. Most often a few major derivatives, accompanied by several minor analogues are observed. In many cases production of minor derivatives is so low, that they can only be discovered from large scale fermentations. A fermentation of 700 L broth led to isolation of 37 natural epothilone variants, where only a few were produced in significant amounts<sup>42</sup>.

Incorporation of different amino acid precursors by two of the eight assembly line modules in the ARG biosynthesis led to production of different peptide cores. Incomplete tailoring reactions by the three modifying enzymes resulted in production of additional derivatives, further expanding the ARG structural diversity (see chapter 4, Fig. 1).

Such partitioning has a severe impact on the production of a particular target analogue, as the available precursors/cofactors and the biosynthetic capacity of the enzymatic machinery are spread amongst the entire spectrum of derivatives. Furthermore, downstream isolation processes are increasingly difficult with higher number of derivatives, especially when their structures, masses and chemical properties are so similar that they are in some cases almost impossible to separate by standard chromatographic methods. This also turned out to be a serious problem in the ARG precursor-directed biosynthesis study discussed below. The number of produced derivatives can often be manipulated by medium engineering, which was also addressed in the ARG project (see chapter 4). By using M7/s4 medium, we were able to decrease production of unwanted major and minor ARG derivatives in favor of significantly increased production of ARG A. Furthermore, the production could be almost exclusively directed towards ARG B by addition of the corresponding building block,  $\alpha$ -aminobutyric acid. Such medium engineering approaches are often a quick and easy way to increase the production and modify the production profile of microbial metabolites<sup>44-46</sup>.

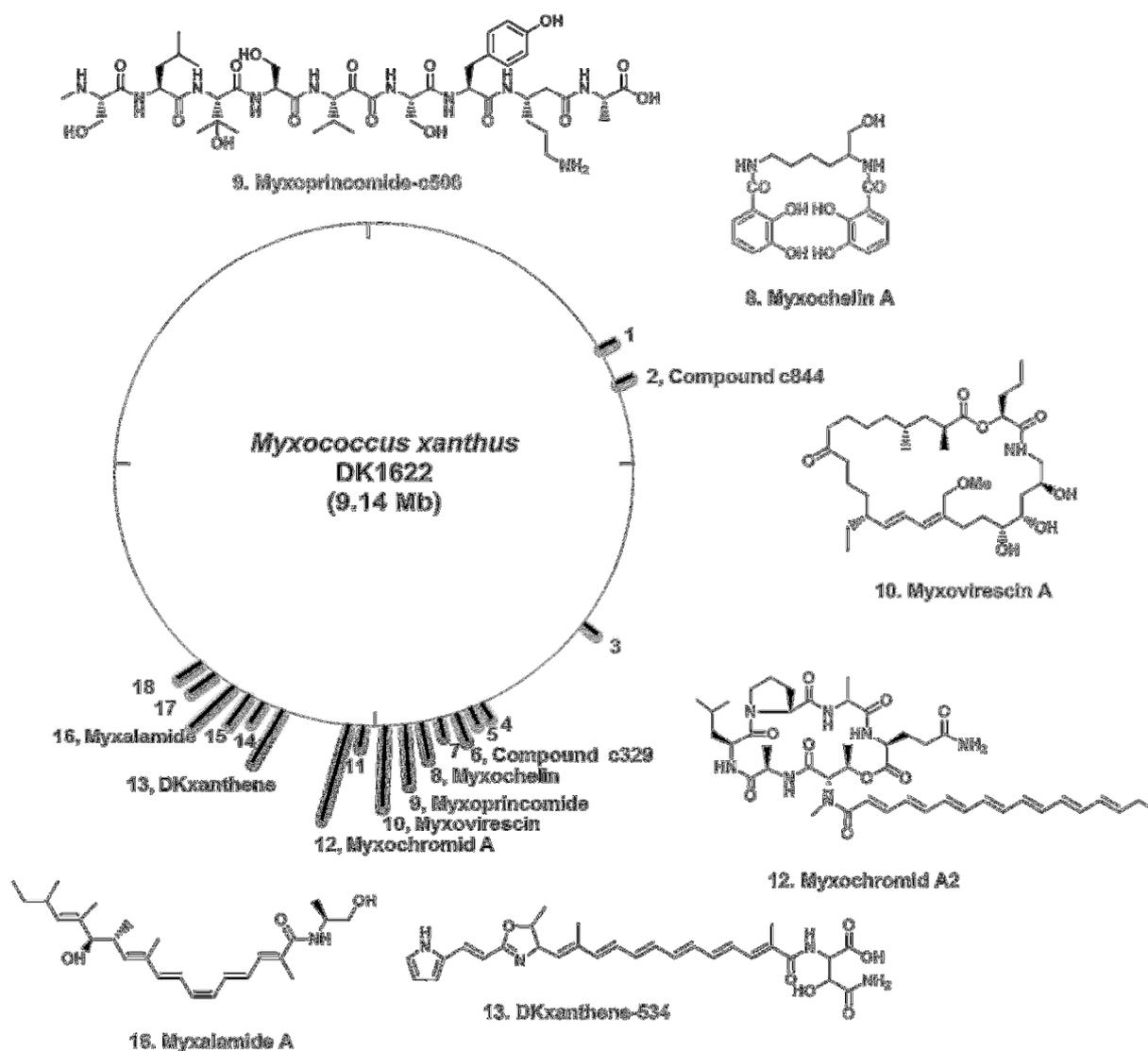


**Figure 1.** Structural diversity of argyrins. The residues contributing to structural diversity are highlighted in different colors and their origin is shown in the squares. It is unclear if the modification steps performed by *arg1*, *arg4* and *arg5* take place prior to assembly or after the release from the NRPS.

### 6.1.3 Genetic engineering techniques

In addition to traditional medium improvement approaches, genetic engineering techniques have been developing and becoming increasingly more efficient over the last decades. One of the important techniques for yield enhancement is the modification of the precursor supply. The main method usually pursued to achieve higher intracellular concentrations of building blocks is the regulation of primary metabolism pathways (e.g. increase of malonyl-CoA supply by manipulation of the acetyl-CoA carboxylase complex led to increased production of mithramycin<sup>47</sup>).<sup>3</sup> Elimination of unwanted competing pathways can potentially also increase the shared precursor pool<sup>48</sup>. The latter method was, to some extent, applied in MXN/COR as well as ARG projects, by deletion of one native BGC from the genome of *M. xanthus* (see chapters 2 and 4). However, the elimination of the myxochromide BGC in *M. xanthus* DK1622 *AmchA-tet* is only the first step towards the construction of a genome minimized expression strain<sup>49</sup>. *M. xanthus* DK1622 possesses 24 BGCs<sup>50</sup>, 16 of which represent PKS/NRPS biosynthetic pathways<sup>51</sup> (Fig. 2), therefore a substantial effort would be required to delete all of them. Heterologous expression of BGCs is one of the widely used approaches to improve compound supply, however it often needs to be combined with additional genetic engineering techniques (e.g. promoter exchange, changing gene orientation, codon adaptation to the selected surrogate host, etc.) to achieve the beneficial effect on the production<sup>52,53</sup>. A lack of good expression strains and limited genetic tools (e.g. very few suitable selection

markers, no reliable replicative plasmid, limited set of promoters) are the main bottlenecks for efficient heterologous expression in myxobacteria. This is mainly due to the fact that they are relatively poorly researched, in contrast to the streptomycetes, which have been extensively studied since the middle of the 1980s. This provided enough time for development of several advanced actinomyces expression strains, with deleted multiple secondary metabolite BGCs<sup>54–57</sup>. Recently, a first cluster-free *Streptomyces albus* heterologous expression strain has been developed and used as a platform for activation of several cryptic BGCs, which led to the isolation of a new compound fralnimycin<sup>58</sup>.



**Figure 2.** PKS/NRPS gene clusters of *Myxococcus xanthus* DK1622. The BGCs with known metabolites are labeled. Chemical structures of one representative from each metabolite family are shown. Chemical structures of Compound c844 and c329 were thus far not elucidated.

A common method to improve the expression of complex BGC for heterologous NP biosynthesis is the implementation of strong heterologous promoters<sup>59</sup>. Promoters driving the

target BGCs have always been a focus for engineering attempts in order to improve secondary metabolite production. A number of studies have been conducted in various bacterial strains in an effort to provide increasingly more efficient promoter systems. A random mutagenesis approach in *E. coli* has been performed by Bakke *et al.* to boost the performance of  $P_{pm}$  promoter from *Pseudomonas putida*, leading to a 14-fold increase of gene expression<sup>60</sup>. In another study by Siegel *et al.* a randomized promoter modification approach was applied to establish a library of 56 synthetic promoters<sup>61</sup>. The promoters yielded expression levels between 2 % and 319 %, providing a good selection for various future metabolic engineering or heterologous expression applications. We conducted several experiments with the emphasis on the promoter evaluation in COR and ARG projects. Myxobacteria are not nearly as well studied as some other bacterial strains, thus only a few different promoters have been evaluated in the heterologous expression studies of NP BGC in *M. xanthus* so far (Table 1). Most of the heterologous systems resorted to the use of a strong constitutive  $P_{nptII}$ <sup>14,21,22,30</sup> promoter, while the  $P_{tet}$ <sup>35,15,34</sup> and  $P_m$ <sup>32</sup> were employed only in a few cases. One study evaluated the inducible  $P_{van}$  and  $P_{IPTG}$  promoters as viable alternatives for inducible gene expression in *M. xanthus*<sup>62</sup>. Moreover, no study on systematic comparison of all different promoters available for *M. xanthus* was conducted thus far. Due to a lack of well described options, at first only a simple replacement of constitutive  $P_{nptII}$  with an inducible  $P_{van}$  promoter was performed in the COR project (see chapter 3). The exchange led to a twofold increase in COR A production, which showed good potential for further experiments in this direction. A more extensive promoter study has thus been performed in the ARG project, which focused on the comparison of several promoter systems and their impact on ARG production in *M. xanthus* DK1622 (see chapter 5). The applied approach was different from the ones described above, as no mutagenesis of the promoter sequences themselves has been performed. Instead, the native sequences of the promoters were compared directly to one another to provide a ranking of the most promising systems, which are currently available in myxobacteria. To eliminate other factors that could influence the production, the same 5'-UTR region had to be implemented in all cases. Comparison of the promoters in the argylin project showed significant differences between the tested systems. The  $P_{nptII}$  and  $P_{van}$  were shown to be superior to other compared promoters, which provided a good starting ground for further optimization approaches. In the future studies previously mentioned mutagenesis experiments could be applied to further improve efficiency of the selected promoter. Moreover, the implemented 5'-UTR region provided significant improvement over the initially used sequence, which could be interesting for future heterologous expressions of

other promising metabolites in *M. xanthus* DK1622. To provide further insight and determine if this is a promising approach for future BGC expressions in myxobacteria, more studies, evaluating specifically tailored artificial 5'-UTR regions are needed.

One of the approaches to improve the supply of COR A was the expression of multiple BGC copies in the heterologous host (unpublished results, data not shown). The method was already proven to work well in streptomycetes, in which a ~350% yield enhancement could be achieved by expressing three copies of the same BGC in *S. albus* and its derivatives<sup>52</sup>. We applied a similar method in *M. xanthus* DK1622 to enhance the yield of COR, while at the same time abolishing the production of the two unwanted major secondary metabolites myxochelin and myxochromide. The introduction of novel *mx8* and *mx9 attB* sites in place of the myxochelin and myxochromide BGCs resulted in complete abolishment of the corresponding metabolites and provided phage attachment sites for genome integration of multiple COR gene cluster copies. Unfortunately, integration into the novel *attB* sites was never achieved, due to so far unknown reasons. One explanation could be that a longer neighboring region of the native *attB* site from *M. xanthus* DK1622 genome is required for successful genome integration. Due to the unsuccessful approach with multiple phage integration sites, we simplified the strategy in order to achieve the expression of at least two COR BGC copies. A previously generated transposase-based mutant with improved COR production profile (see chapter 3) was additionally transformed with another *mx8*-based BGC copy. Contrary to the study published by Manderscheid *et al.*<sup>52</sup>, in our case an additional copy of the COR BGC did not provide any beneficial effect on the COR production. We later discovered that the reason for this was most likely the toxic effect of the COR on the host organism, as the COR MIC on the producer strain, harboring a single BGC copy, was in the same range as the achieved production level (see chapter 3). Compound toxicity is one of the common problems associated with heterologous expression, which should be always kept in mind when dealing with compounds possessing potent bioactive properties. This issue is usually addressed by overexpression of resistance genes, provided that those have previously been identified<sup>53</sup>. In the COR cluster no such genes were reported, however additional experiments lead to the identification of a putative resistance protein CorP. An expression of additional CorP copy, however did not provide additional resistance to COR, therefore further research on COR self-resistance mechanism is necessary to be finally able to evaluate the expression of multiple COR clusters and its effect on COR production levels.

## 6.2 Structure engineering of microbial natural products

Although many microbial NPs exhibit potent biological activities, they often lack the desired pharmaceutical properties required for medical application. To tackle this problem and generate structural variants of NPs, semisynthetic methods and genetic engineering approaches were developed. The structural complexity of many NPs (especially the compounds produced by NPRS and PKS megasynthetases) is the main hurdle when it comes to their engineering by standard chemical synthesis methods. Chemical synthesis of NPs has in most cases proven to be tedious and inefficient, often requiring many steps and producing total yields that are below the desired threshold<sup>63,64</sup>. Semi synthesis, however, showed some good potential with e.g. oritavacin and ixabepilone reaching market approval<sup>65,66</sup>. Oritavacin is a semi-synthetic derivative generated from the naturally occurring glycopeptide antibiotic chloroeremomycin, by addition of the N-alkyl-p-chlorophenylbenzyl substituent on its disaccharide sugar. This helps oritavacin to better bind its target, thus increasing its potency<sup>67</sup>. Ixabepilone, a semi-synthetic derivative of myxobacterial anticancer compound epothilone B, was generated by chemically substituting the naturally present lactone in epothilone with a lactam. This modification substantially improved stability of the compound in the *in vivo* conditions, which was one of the main problems of the epothilone B<sup>68</sup>.

Different approaches can be employed to engineer the microbial production systems in order to achieve biosynthesis of structural analogs. Some of such alternative methods include precursor-directed biosynthesis and mutasynthesis, which both focus on incorporation of surrogate building blocks into the core structure of the target compound (Fig. 3)<sup>69–71</sup>. Furthermore, the structure of NRPs/PKS can also be modified by utilizing exogenous tailoring enzymes from other pathways alongside the biosynthetic machinery to diversify the final structure or by modifying the biosynthetic machinery itself<sup>72,73</sup>.

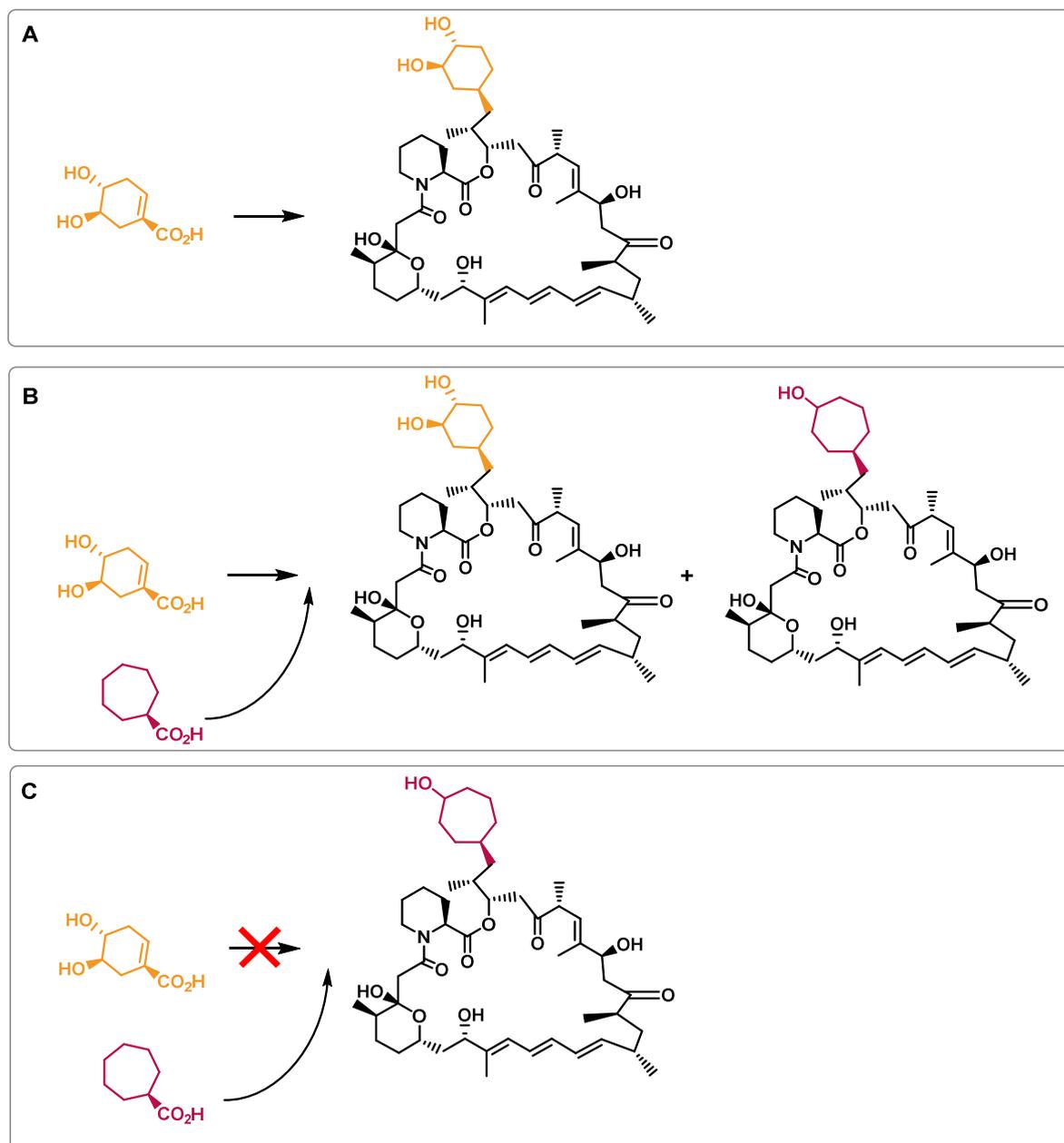
### 6.2.1 Precursor-directed biosynthesis and mutasynthesis approaches

Precursor-directed biosynthesis relies on unspecific substrate incorporation and processing by the gatekeeping enzymes (e.g. adenylation and condensation domains). It does not require any genetic engineering effort to be successfully implemented. Moreover, no extensive knowledge about the biosynthesis or specificity of the involved biosynthetic enzymes is required. However, some basic information about the biogenesis can help facilitate the selection of the appropriate precursors/building blocks. The method is hence a popular approach to generate structural analogues of NPs. The main drawback is the competition of the enzymes between

the native and the alternative substrate, where the enzymes typically prefer the natural substrate. Furthermore, integration of various substrates results in a complex mixture of produced analogues which are often difficult to separate<sup>69</sup>.

To avoid issues of competition with native building blocks, mutasynthesis is often applied instead of precursor-directed biosynthesis<sup>69</sup>. It provides selectivity by genetic engineering of the assembly line machinery or by mutation of genes involved in the supply of biosynthetic precursors and supplementing the corresponding biosynthetic precursor or missing intermediate with synthetic mimics. For successful integration of structurally diverse synthetic mimics, the gatekeeping enzymes have to exert sufficiently relaxed substrate specificity. With increasing structural differences of mutasynthons, compared to the native intermediate, the chances of successful integration are decreasing. It is very hard to predict which intermediates might still be accepted by the truncated biosynthetic pathway, thus a library of diverse mutasynthons is usually generated, to help determine the limits of structural complexity. The differences of precursor-directed biosynthesis and mutasynthesis are highlighted in the Figure 3.

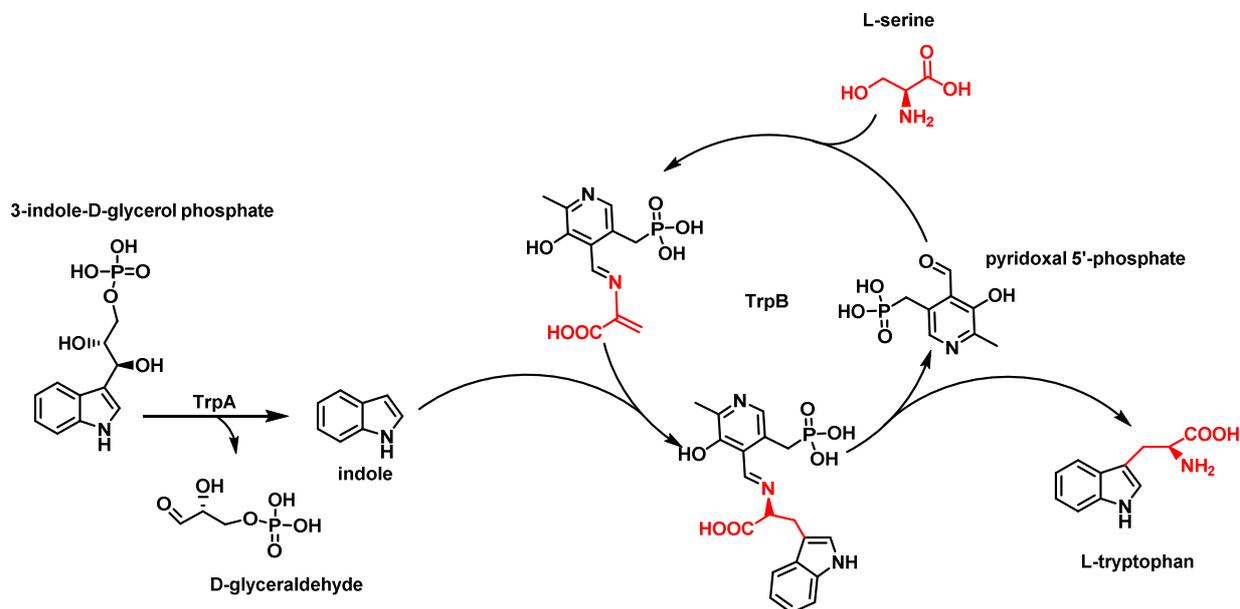
The first successful example of precursor-directed biosynthesis was reported in 1989 by the researchers of Sandoz<sup>74</sup>. By supplementing the medium of the cyclosporine producer with exogenous amino acids, they were able to produce novel structural derivatives, some of which exhibited potent immunosuppressive activity. Since then many more successful examples of precursor-directed biosynthesis were reported. By feeding a series of 21 carboxylic acids to a rapamycin producer *S. hygroscopicus*, researchers were able to generate several novel rapamycin analogues<sup>75</sup>. In another study a variety of non-proteinogenic amino acids was fed to *Streptomyces griseoflavus*, which led to isolation of novel hormaomycins<sup>76</sup>. Successful examples were also reported from the field of myxobacteria, e.g. feeding of halogenated phenylalanine and cinnamate to the myxobacterium *Sorangium cellulosum*, led to production of novel derivatives of potent antifungal compound Soraphen<sup>77</sup>.



**Figure 3.** Comparison of natural biosynthesis (A), precursor-directed biosynthesis (B) and mutasynthesis (C) shown on the example of pre-rapamycin biosynthetic model.

This strategy was also applied in the ARG project targeting the two tryptophan moieties present in the peptide core (see chapter 5). Many of the substituted tryptophan derivatives were either not commercially available at all or were very expensive to obtain. On the other hand, substituted indoles were cheaper and more accessible from commercial vendors. We thus combined a straightforward approach for the biotransformation of halogenated and nitrated tryptophan derivatives with precursor-directed biosynthesis to generate a library of corresponding ARG derivatives, 14 of which were so far confirmed by NMR. The tryptophan biotransformation utilizes the activity of the tryptophan synthase  $\beta$ -subunit, to combine the externally supplied indole derivatives with L-serine and produce the corresponding

tryptophanes<sup>78–81</sup> (Fig. 4). The system has potential to be developed further by e.g. expression of the tryptophan synthase in the host organism. This would enable direct feeding of the substituted indoles to the culture, eliminating the currently required biotransformation and workup steps and thus significantly speed up the process.



**Figure 4.** Biosynthesis of L-tryptophan by tryptophan synthase  $\alpha$ -subunit (TrpA) and  $\beta$ -subunit (TrpB).

Several modules of the ARG BGC were already proven to be promiscuous as they were shown to incorporate multiple building blocks by studies conducted in chapter 4. It is not surprising that the downstream modules are exerting similar features. As previously mentioned, precursor-directed biosynthesis studies show overall moderate success<sup>75–77,82</sup> due to the technique's shortcomings often limiting applicability in industrial processes. Competition of the enzymes between the native and the alternative substrate leads to partitioning of the resources, negatively impacting the production yield of the desired novel compound in addition to causing difficulties in the downstream processing. The same was true in the case of ARG, where HPLC purification of the novel derivatives turned out to be more complicated than expected. Several rounds of purification on different analytical scale columns had to be performed, to remove impurities in form of various minor ARG derivatives. Furthermore, as ARG possesses two tryptophan residues, the incorporation of the modified tryptophan on either position resulted in the production of two isobaric derivatives with very similar chemical features, which were in some cases impossible to separate by standard chromatographic methods.

Alternatively to precursor-directed biosynthesis different approaches of mutasynthesis are possible. Some simpler methods rely on disruption of starter unit biosynthesis and its

complementation with synthetic mimics that are usually easy to synthesize (e.g. generation of rapamycin analogues by mutasynthesis<sup>83</sup>). Other more complex methods rely on deletion/inactivation of larger parts of biosynthetic machinery (e.g deletion of several modules or entire megasynthetases subunit) and complementation of the corresponding biosynthetic intermediates by larger and more complex synthetic mutasynthons mimicking biosynthetic intermediates of the respective assembly line. An example of complex mutasynthesis is the study conducted on myxobacterial  $\alpha$ -pyrone antibiotics from *Myxococcus fulvus*<sup>84</sup>. Biosynthesis of the native compound was disrupted by mutation of the carrier protein (CP) in the first or the fourth module of the myxopyronin western chain assembly line. The cultures of the deletion mutants were supplemented with NAC-thio-esters of the corresponding first module precursors or longer fourth module intermediates. The disrupted myxopyronin biosynthetic machinery accepted most of the mutasynthons, however, the obtained production yield of generated derivatives was very low. Possible limitations in such complex approaches could be poor stability of the mutasynthons, or their poor permeability through the cell membrane. This is especially problematic if the synthetic intermediate is a peptide, which could probably be degraded by many proteases present in the production broth. Furthermore, complex synthetic intermediates are often difficult to synthesize, requiring significant time and effort, which often discourages more advanced mutasynthesis strategies.

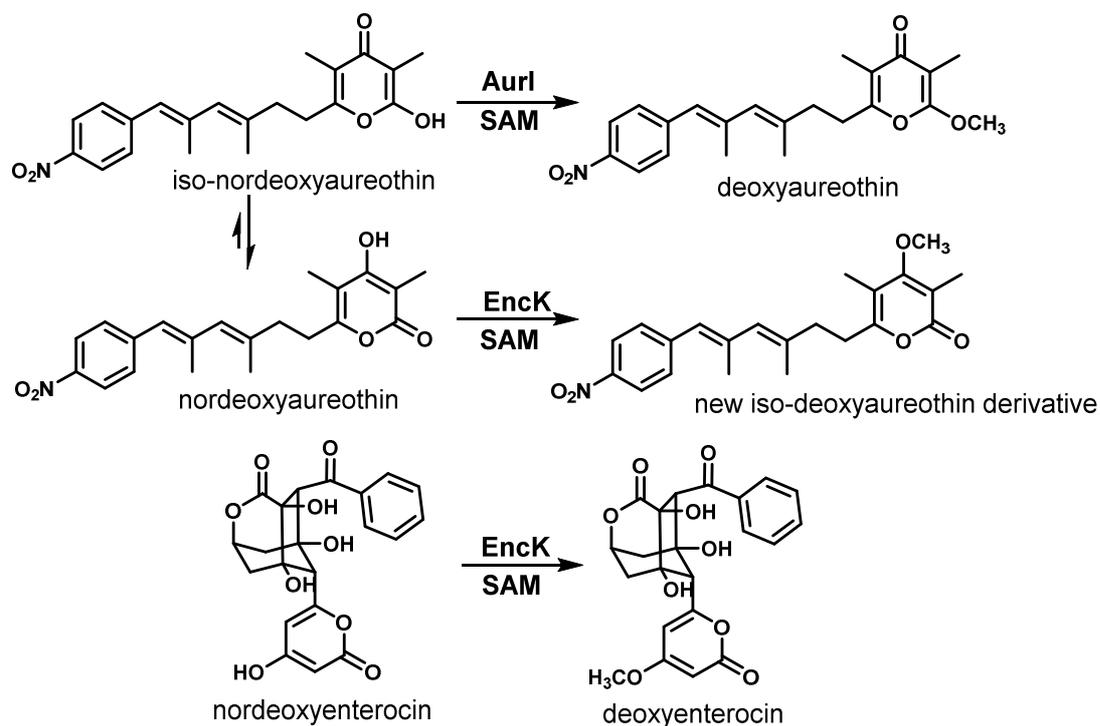
To access novel ARG analogues, a respective mutant with deleted *arg2* gene, encoding the first subunit of the NRPS assembly line, was generated (see chapter 4). A set of synthetic intermediates, mimicking the ARG tripeptides including the intermediate with the native amino acid sequence, were generated by total synthesis. The first big problem was already encountered in this step as the chemical synthesis of the complex mutasynthons turned out to be difficult. Despite significant efforts, production of ARG by mutasynthesis could never be achieved. To establish the production of novel derivatives and test the boundaries of structural complexity, various intermediates are usually employed. In case of ARG the chemical structure of mutasynthons is most likely not the culprit for unsuccessful mutasynthesis as not even the synthetic intermediate with the native amino acid sequence could be successfully integrated by the Arg3 NRPS subunit (see chapter 4). The most likely explanation is probably the instability of synthetic intermediates in the production medium, as degradation kinetics of the native intermediate mimic (D-Ala-Dha-Sar-SNAc) in the production medium showed almost complete degradation after 2 hours of incubation (see chapter 4). Inefficient incorporation of the mutasynthons by the biosynthetic process, however, cannot be excluded as additional issue.

### 6.2.2 Engineering of tailoring genes

In recent years tailoring enzyme engineering has begun to be exploited to bring additional structural diversity to NPs. One way the tailoring enzymes can modify the structure of the NP is by addition of chemical groups directly on the core compound after its release from the assembly line. They can also act by tailoring the biosynthetic precursor prior to its incorporation into the final NP structure or they can act *in trans* to modify the biosynthetic intermediates during the assembly process. In case of precursor tailoring the enzyme modifies the original building block *in vivo*, thus generating a pool of endogenous modified building blocks. Those modified precursors compete with the original precursors to be accepted by the biosynthetic machinery. The tailoring enzyme in this case modifies a simple biosynthetic precursor, which can often be found in other biosynthetic pathways. This makes it possible to introduce the genes for such tailoring enzymes into surrogate biosynthetic pathways to modify the corresponding metabolite. Since the enzyme is specific for the precursor molecule and not the final compound, such bio-combinatorial approaches usually lead to some degree of success even when combining enzymes from structurally diverse metabolite pathways. For example, expression of PrnA, a flavin-dependent tryptophan-7-halogenase from *Pseudomonas fluorescens* Pf-5, alongside the NRPS genes for antibiotic pacidamycin from *Streptomyces coeruleorubidus*, led to production of novel halogenated analogue<sup>85</sup>. One downside of such approaches, is the preference of the enzyme for the native substrate and the resulting mixture of the native and the modified NP, provided that the modified building block is accepted by the biosynthetic machinery. This was also the case in pacidamycin engineering approach, where the novel analogue was produced as a minor derivative alongside the native pacidamycin<sup>85</sup>.

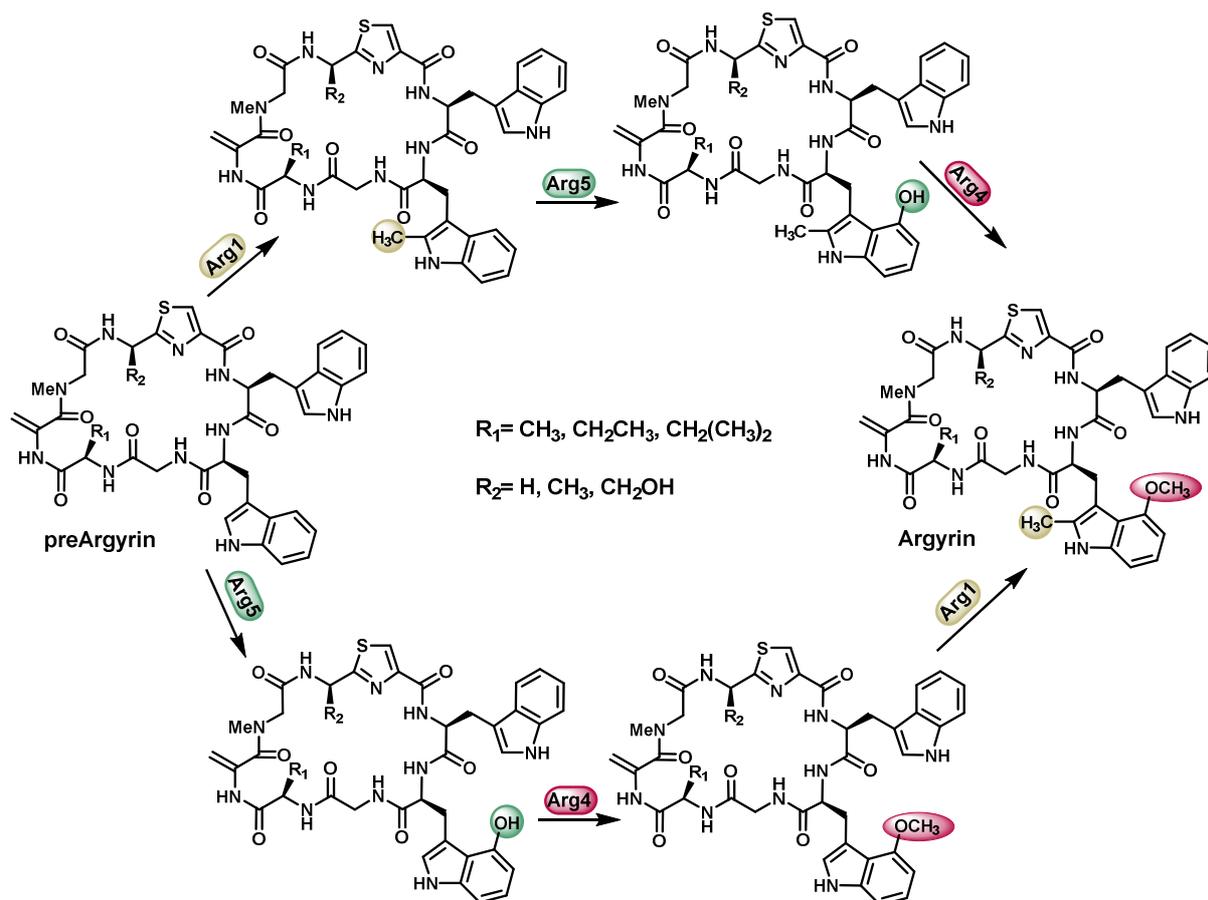
In contrast to the precursor tailoring, the enzymes that act on the post assembly line compound or *in trans* on the biosynthetic intermediates are usually more specific, making their mix-matching in most cases possible only in metabolic pathways with very similar products. For instance, expression of halogenase from ramoplanin BGC resulted in mono-chlorinated enduracidin<sup>86</sup>. Both compounds are closely related lipopeptides chlorinated at one of the six L-Hpg residues by a flavin-dependent halogenase. Some enzymes exhibit lower selectivity, making them more suitable for such bio-combinatorial approaches. For example, replacement of endogenous  $\gamma$ -regiospecific pyrone methyltransferase AurI with the  $\alpha$ -regiospecific EncK homolog led to production of iso-deoxyaureothin derivative, despite different polyketide core structures of aureothin and enterocin (Fig. 5). However,

complementation of EncK with AurI was unsuccessful, indicating that EncK exhibits more relaxed substrate preference<sup>87</sup>.



**Figure 5.** Regiospecific pyrone *O*-methylation by AurI and EncK (The figure is adapted from<sup>87</sup>).

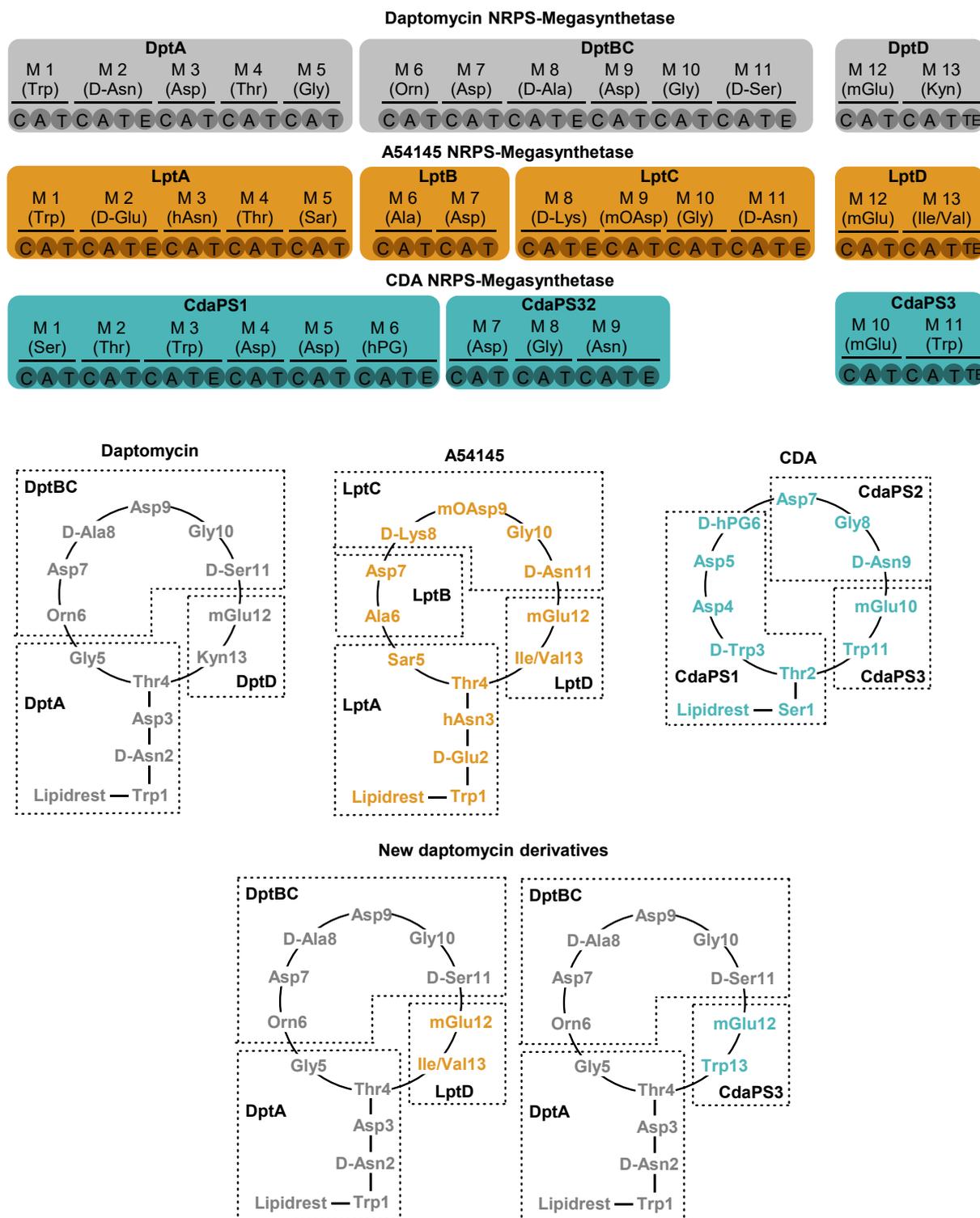
Basic engineering of tailoring genes was applied in the COR and ARG projects, to access novel derivatives (see chapters 3 and 4). Deletion of *corO* encoding cytochrome P450 in COR project led to complete redirection of production towards the non-hydroxylated derivative preCOR A. Furthermore, deletion of *corN*, which turned out to encode a *trans*-acting ECH-like enzyme, led to isolation of two novel derivatives bearing an OH group at a different position. Expression of ARG BGC variants lacking one or both of the tailoring genes *arg4* and *arg5*, in the ARG project resulted in production of various derivatives with different tailoring patterns at the second tryptophan residue. Expression of *arg1* additionally resulted in production of two more methylated variants, argyrins C and D. Furthermore, argyrim production optimization led to significantly increased production yield, which also resulted in significant production of an untailed core peptide (preArgyrim), probably due to the inadequate activity of the *arg4* and *arg5*. Expression of additional *arg4* and *arg5* copies could probably lead to conversion of the remaining untailed derivative to boost production of fully methoxylated derivatives e.g. ARG A and B (Fig. 6).



**Figure 6.** Conversion of preArgyirin to Argyrin by Arg1, Arg5 and Arg4 tailoring enzymes. List of argyirin derivatives is provided in Figure 1.

### 6.2.3 Engineering of PKS/NRPS megasynthetases

As described above, engineering of tailoring genes can be a prominent approach to increase structural diversity of naturally produced NRP/PK core structures. An alternative and more complex strategy is the engineering of the NRP/PK core directly. Modular megasynthetases, as the name suggests, consist of sequence of modules that incorporate individual building blocks into the growing peptide/polyketide chain. One possibility for the required megasynthetase engineering is deletion of NRPS/PKS subunits usually followed by their complementation with homologs from structurally similar product pathways. The method was successfully applied to exchange the *dptD* subunit from daptomycin BGC with *cdaPS3* and *lptD* from structurally related CDA and A54145 biosynthesis, respectively, which resulted in production of two novel daptomycin analogues<sup>88</sup> (Fig. 7). In another example the shuffling of the PKS subunits of tylosin, spiramycin and chalcomycin biosynthetic pathways resulted in production of hybrid molecules that varied in side chains at several positions<sup>89</sup>.

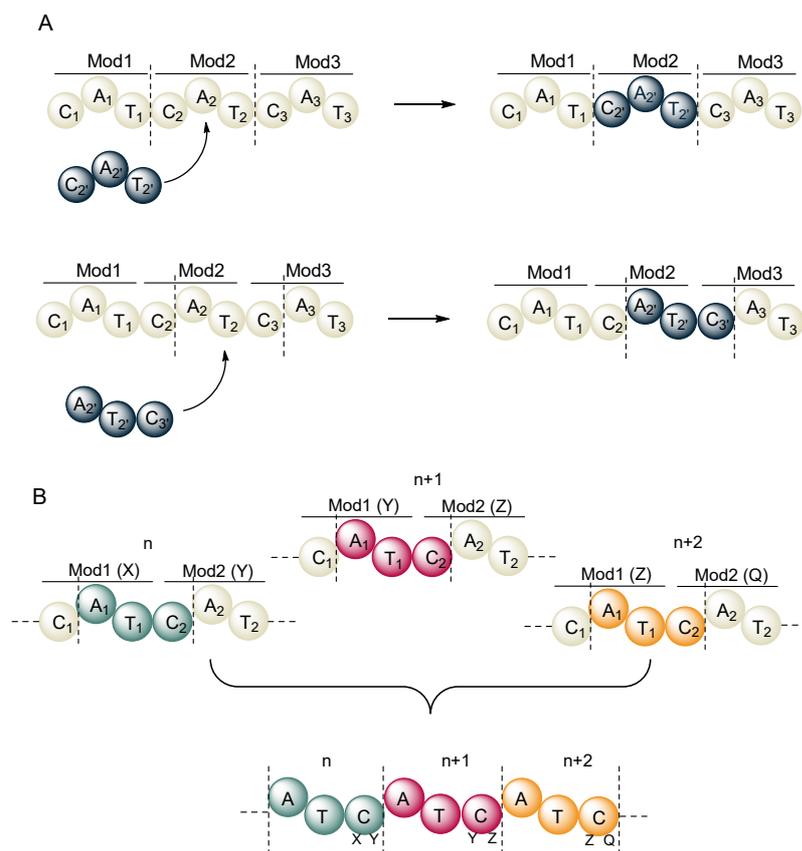


**Figure 7.** NRPS megasyntetases of Daptomycin, A54145 and CDA. The structures of the corresponding compounds and the resulting novel daptomycin derivatives are shown below.

One way to modify the structure of peptide/polyketide chains is to introduce changes within specific modules or directly delete/replace them. Exchange of highly homologous modules within daptomycin BGC or swapping of native modules with modules from structurally related A54145 BGC led to production of predicted daptomycin analogues with exchanged residues at the corresponding positions<sup>90</sup>. Experiments in this study also showed that

exchange of C-A-T module architecture leads to much better production compared to exchange of full C-A-T-E module, emphasizing the importance of the module-module linker regions<sup>90</sup>. In a combinatorial study with PKSs 14 modules from 8 PKS clusters were used to generate 154 bimodular chimeric PKSs, from which around half of the systems were functional<sup>91</sup>. Furthermore, module deletion led to reduction of surfactin ring size<sup>92</sup>, whereas insertion of an additional module resulted in an extended core in balhimycin<sup>93</sup> and erythromycin<sup>94</sup> biosynthesis. As described in the provided examples, many of such modifications often result in substantial production yield decrease or even total abolishment of the production. It is now common knowledge that terminal regions of the domains and modules within the NRPS/PKS systems often act as linker regions that facilitate protein-protein interactions<sup>72</sup>. By replacing or deleting modules/domains within those systems, those linker regions are often disrupted, preventing proper protein association and in turn leading to decreased or abolished production. Complications with linker region between NRPS domains led to the engineering of internal sub-domain part, to preserve the integrity of interactions with the neighboring domains. In hormaomycin study a sub-domain part of the A domain was replaced to generate hybrid A domains, some of which showed activity *in vitro*<sup>95</sup>. The technique was further improved in the gramicidin study<sup>96</sup>.

With the definition of the “nonribosomal code”<sup>97</sup> defined by Stachelahus et al.<sup>98</sup>, the foundations for engineering of the A domain binding pocket were set. Introducing point mutations in the active site of the A domains eliminates the problems of the previously mentioned disruption of the linker regions. However, the specificity of the downstream modules remains a challenge. In many cases changing the specificity of the A domain leads to incompliance of the downstream gatekeeping domains, which fail to process the new substrate. Furthermore, performing markerless modifications of complex BGCs is technically challenging and time consuming. Recently, *de novo* design and engineering of NRPSs was achieved by swapping the defined exchange units (A-T-C) fused at specific positions between C and A domains, rather than traditional C-A-T modules. The method better respects the original specificity of the downstream module to enable production of the desired peptides<sup>99</sup> (Fig. 8).



**Figure 8.** Traditional C-A-T module exchange in comparison with the alternative A-T-C unit (A). Schematic representation for recombination of three NRPS A-T-C units, taking into account the C domain specificities (B). The specificity of the C domains for upstream and downstream substrate is marked with letters (X,Y,Z,Q).

The A domains are the first and most obvious gatekeepers in NRPS biosynthesis, responsible for the building block incorporation. Due to their function of substrate selection and activation, they have been the preferred subject of genetic manipulation in the past<sup>72</sup>. Various approaches to A domain engineering have been evaluated over the last decades. They range from more robust (e.g. module and domain replacements)<sup>100,101,95,96</sup>, to more precise ones (e.g. targeted modifications of single amino acids lining the A domain active site)<sup>102,103</sup>. The latter approach was employed in the ARG project to engineer the production profile and direct the biosynthesis towards more desirable analogues (see chapter 4). Several modules of the ARG assembly line were shown to possess broader substrate specificity, as they are able to accept a few different building blocks. It was thus speculated that modifications of the nonribosomal codes in the A domains from those modules could make them more specific for one of the already accepted substrates. In case the substrate pool can provide enough of the preferred substrate, the modification should lead to increased production of the corresponding derivative. We also predicted that the downstream C domains should be able to process the integrated building block, as it already appears in the wild type production. The engineering of the A domain in the first module showed moderate success as several modifications led to a

higher incorporation of alanine compared to the control. The impact of the modification was not as successful as anticipated, due to the substantial reduction of ARG production in the majority of the mutants. Reduction of secondary metabolite production yields are not unusual following such assembly line modifications<sup>72,88,90</sup>. The main reason for this is probably incompliance of the downstream biosynthetic machinery with the newly incorporated substrate. To provide further insight into ARG A domain engineering, a similar approach was employed to modify the specificity of the A domain in the fourth module of the ARG assembly line. The modifications were adapted to include the extended region surrounding the A domain binding pocket, also known as the 8 Å code<sup>104</sup>. To our surprise, all modifications resulted in total abolishment of the ARG production, which shows that rational genetic engineering of the NRPS assembly line is a challenging and not a fully understood task<sup>72</sup>. The most logical explanation for the abolished production could be, that some of the mutations performed in the 8 Å code probably have an impact on structural integrity of the protein, resulting in its incorrect folding.

### 6.3 Conclusion and outlook

This thesis focused on an  $\alpha$ -pyrone polyketide antibiotic COR and the octapeptide nonribosomal peptide ARG from myxobacteria. Heterologous production as achieved for both compounds provided a significant improvement in production and simplified additional biosynthesis studies and structural engineering approaches. Studies on the COR biosynthetic gene cluster provided new insight into the RNAP self-resistance mechanism, which could lead to an engineering of an improved COR heterologous producer strain in the future. Medium optimization in ARG projects led to a significant yield improvement of major ARG derivatives and successfully employed precursor-directed biosynthesis yielded several novel ARG derivatives the bioactivity of which still needs to be evaluated. The combination of improved production and precursor-directed biosynthesis will hopefully lead to the production of novel ARG derivatives with enhanced antibiotic and/or immunoregulatory bioactivity profiles in the future. Several of the methods applied in this thesis provide new knowledge about engineering and heterologous expression of NP pathways in myxobacteria, which could be applied to other NP biosynthetic pathways in the upcoming studies.

## 6.4 References

- (1) Weissman, K. J.; Müller, R. Myxobacterial secondary metabolites: bioactivities and modes-of-action, *Nat. Prod. Rep.* **2010**, *27*, pp. 1276–1295.
- (2) Herrmann, J.; Fayad, A. A.; Müller, R. Natural products from myxobacteria: novel metabolites and bioactivities, *Nat. Prod. Rep.* **2017**, *34*, pp. 135–160.
- (3) Bilyk, O.; Luzhetskyy, A. Metabolic engineering of natural product biosynthesis in actinobacteria, *Curr. Opin. Biotechnol.* **2016**, *42*, pp. 98–107.
- (4) Rachid, S.; Gerth, K.; Kochems, I.; Müller, R. Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium *Sorangium cellulosum* So ce56, *Mol. Microbiol.* **2007**, *63*, pp. 1783–1796.
- (5) Rachid, S.; Gerth, K.; Müller, R. NtcA-A negative regulator of secondary metabolite biosynthesis in *Sorangium cellulosum*, *J. Biotechnol.* **2008**, *140*, pp. 135–142.
- (6) Rachid, S.; Sasse, F.; Beyer, S.; Müller, R. Identification of StiR, the first regulator of secondary metabolite formation in the myxobacterium *Cystobacter fuscus* Cb f17.1, *J. Biotechnol.* **2006**, *121*, pp. 429–441.
- (7) Sandmann, A.; Frank, B.; Müller, R. A transposon-based strategy to scale up myxothiazol production in myxobacterial cell factories, *J. Biotechnol.* **2008**, *135*, pp. 255–261.
- (8) Wenzel, S. C.; Müller, R. Myxobacteria--'microbial factories' for the production of bioactive secondary metabolites, *Mol. Biosyst.* **2009**, *5*, pp. 567–574.
- (9) Bock, T.; Volz, C.; Hering, V.; Scrima, A.; Müller, R.; Blankenfeldt, W. The AibR-isovaleryl coenzyme A regulator and its DNA binding site - a model for the regulation of alternative de novo isovaleryl coenzyme A biosynthesis in *Myxococcus xanthus*, *Nucleic Acids Res.* **2017**, *45*, pp. 2166–2178.
- (10) Mahmud, T.; Bode, H. B.; Silakowski, B.; Kroppenstedt, R. M.; Xu, M.; Nordhoff, S.; Höfle, G.; Müller, R. A novel biosynthetic pathway providing precursors for fatty acid biosynthesis and secondary metabolite formation in myxobacteria, *J. Biol. Chem.* **2002**, *277*, pp. 32768–32774.
- (11) Bolten, C. J.; Heinzle, E.; Müller, R.; Wittmann, C. Investigation of the central carbon metabolism of *Sorangium cellulosum*: metabolic network reconstruction and quantification of pathway fluxes, *J. Microbiol. Biotechnol.* **2009**, *19*, pp. 23–36.
- (12) Gross, F.; Luniak, N.; Perlova, O.; Gaitatzis, N.; Jenke-Kodama, H.; Gerth, K.; Gottschalk, D.; Dittmann, E.; Müller, R. Bacterial type III polyketide synthases: Phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads, *Arch. Microbiol.* **2006**, *185*, pp. 28–38.
- (13) Wenzel, S. C.; Gross, F.; Zhang, Y.; Fu, J.; Stewart, F. A.; Müller, R. Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via red/ET recombineering, *Chem. Biol.* **2005**, *12*, pp. 349–356.
- (14) Fu, J.; Wenzel, S. C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A. F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition, *Nucleic Acids Res.* **2008**, *36*, e113.
- (15) Yan, F.; Auerbach, D.; Chai, Y.; Keller, L.; Tu, Q.; Hüttel, S.; Glemser, A.; Grab, H. A.; Bach, T.; Zhang, Y.; Müller, R. Biosynthesis and Heterologous Production of Vioprolides:

Rational Biosynthetic Engineering and Unprecedented 4-Methylazetidinecarboxylic Acid Formation, *Angew. Chem. Int. Ed. Engl.* **2018**.

(16) Perlova, O.; Gerth, K.; Kuhlmann, S.; Zhang, Y.; Müller, R. Novel expression hosts for complex secondary metabolite megasynthetases: Production of myxochromide in the thermophilic isolate *Coralloccoccus macrosporus* GT-2, *Microb. Cell Fact.* **2009**, *8*.

(17) Stevens, D. C.; Henry, M. R.; Murphy, K. A.; Boddy, C. N. Heterologous expression of the oxytetracycline biosynthetic pathway in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2010**, *76*, pp. 2681–2683.

(18) Yin, S.; Li, Z.; Wang, X.; Wang, H.; Jia, X.; Ai, G.; Bai, Z.; Shi, M.; Yuan, F.; Liu, T.; Wang, W.; Yang, K. Heterologous expression of oxytetracycline biosynthetic gene cluster in *Streptomyces venezuelae* WVR2006 to improve production level and to alter fermentation process, *Appl. Microbiol. Biotechnol.* **2016**.

(19) Pogorevc, D.; Tang, Y.; Hoffmann, M.; Zipf, G.; Bernauer, H. S.; Popoff, A.; Steinmetz, H.; Wenzel, S. C. Biosynthesis and heterologous production of argyrins. **2019**, *submitted manuscript*.

(20) Pogorevc, D.; Popoff, A.; Fayad, A. A.; Wenzel, S. C.; Müller, R. Production profile engineering and precursor directed biosynthesis approaches for production of novel argyrim derivatives, *unpublished results*.

(21) Wenzel, S. C.; Hoffmann, H.; Zhang, J.; Debussche, L.; Haag-Richter, S.; Kurz, M.; Nardi, F.; Lukat, P.; Kochems, I.; Tietgen, H.; Schummer, D.; Nicolas, J.-P.; Calvet, L.; Czepczor, V.; Vrignaud, P.; Mühlenweg, A.; Pelzer, S.; Müller, R.; Brönstrup, M. Production of the bengamide class of marine natural products in myxobacteria: biosynthesis and structure-activity relationships, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, pp. 15560–15564.

(22) Sucipto, H.; Pogorevc, D.; Luxenburger, E.; Wenzel, S. C.; Müller, R. Heterologous production of myxobacterial  $\alpha$ -pyrone antibiotics in *Myxococcus xanthus*, *Metab. Eng.* **2017**, *44*, pp. 160–170.

(23) Pogorevc, D.; Panter, F.; Schillinger, C.; Jansen, R.; Wenzel, S. C.; Müller, R. Production optimization and biosynthesis revision of Corallopyronin A, a potent anti-filarial antibiotic. **2019**, *submitted manuscript*.

(24) Oßwald, C.; Zaburanyi, N.; Burgard, C.; Hoffmann, T.; Wenzel, S. C.; Müller, R. A highly unusual polyketide synthase directs dawenol polyene biosynthesis in *Stigmatella aurantiaca*, *J. Biotechnol.* **2014**, *191*, pp. 54–63.

(25) Julien, B.; Shah, S. Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*, *Antimicrob. Agents Chemother.* **2002**, *46*, pp. 2772–2778.

(26) Lau, J.; Frykman, S.; Regentin, R.; Ou, S.; Tsuruta, H.; Licari, P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*, *Biotechnol. Bioeng.* **2002**, *78*, pp. 280–288.

(27) Tang, L.; Shah, S.; Chung, L.; Carney, J.; Katz, L.; Khosla, C.; Julien, B. Cloning and heterologous expression of the epothilone gene cluster, *Science*. **2000**, *287*, pp. 640–642.

(28) Park, S. R.; Park, J. W.; Jung, W. S.; Han, A. R.; Ban, Y. H.; Kim, E. J.; Sohng, J. K.; Sim, S. J.; Yoon, Y. J. Heterologous production of epothilones B and D in *Streptomyces venezuelae*, *Appl. Microbiol. Biotechnol.* **2008**, *81*, pp. 109–117.

(29) Mutka, S. C.; Carney, J. R.; Liu, Y.; Kennedy, J. Heterologous Production of Epothilone C and D in *Escherichia coli*, *Biochemistry*. **2006**, *45*, pp. 1321–1330.

- (30) Oßwald, C.; Zipf, G.; Schmidt, G.; Maier, J.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Modular Construction of a Functional Artificial Epothilone Polyketide Pathway, *ACS Synth. Biol.* **2014**, *3*, pp. 759–772.
- (31) Yan, F.; Burgard, C.; Popoff, A.; Zaburannyi, N.; Zipf, G.; Maier, J.; Bernauer, H. S.; Wenzel, S. C.; Müller, R. Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production, *Chem. Sci.* **2018**, *9*, pp. 7510–7519.
- (32) Perlova, O.; Fu, J.; Kuhlmann, S.; Krug, D.; Stewart, F.; Zhang, Y.; Müller, R. Reconstitution of myxothiazol biosynthetic gene cluster by Red/ET recombination and heterologous expression in *Myxococcus xanthus*, *Appl. Environ. Microbiol.* **2006**, *72*, pp. 7485–7494.
- (33) Gross, F.; Ring, M. W.; Perlova, O.; Fu, J.; Schneider, S.; Gerth, K.; Kuhlmann, S.; Stewart, A. F.; Zhang, Y.; Müller, R. Metabolic engineering of *Pseudomonas putida* for methylmalonyl-CoA biosynthesis to enable complex heterologous secondary metabolite formation, *Chem. Biol.* **2006**, *13*, pp. 1253–1264.
- (34) Chai, Y.; Shan, S.; Weissman, K. J.; Hu, S.; Zhang, Y.; Müller, R. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis, *Chem. Biol.* **2012**, *19*, pp. 361–371.
- (35) Gemperlein, K.; Rachid, S.; Garcia, R. O.; Wenzel, S. C.; Müller, R. Polyunsaturated fatty acid biosynthesis in myxobacteria. Different PUFA synthases and their product diversity, *Chem. Sci.* **2014**, *5*, pp. 1733–1741.
- (36) Zirkle, R.; Ligon, J. M.; Molnar, I. Heterologous production of the antifungal polyketide antibiotic soraphen A of *Sorangium cellulosum* So ce26 in *Streptomyces lividans*, *Microbiology.* **2004**, *150*, pp. 2761–2774.
- (37) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. Big effects from small changes: possible ways to explore nature's chemical diversity, *ChemBioChem.* **2002**, *3*, pp. 619–627.
- (38) Singh, V.; Haque, S.; Niwas, R.; Srivastava, A.; Pasupuleti, M.; Tripathi, C. K. M. Strategies for Fermentation Medium Optimization. An In-Depth Review, *Front. Microbiol.* **2016**, *7*, p. 2087.
- (39) Hopkins, C. D.; Wipf, P. Isolation, biology and chemistry of the disorazoles: new anti-cancer macrodiolides, *Nat. Prod. Rep.* **2009**, *26*, pp. 585–601.
- (40) Irschik, H.; Washausen, P.; Sasse, F.; Fohrer, J.; Huch, V.; Müller, R.; Prusov, E. V. Isolation, structure elucidation, and biological activity of maltepolides: remarkable macrolides from myxobacteria, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, pp. 5402–5405.
- (41) Gerth, K.; Jansen, R.; Reifenstahl, G.; Höfle, G.; Irschik, H.; Kunze, B.; Reichenbach, H.; Thierbach, G. The myxalamids, new antibiotics from *Myxococcus xanthus* (Myxobacterales). I. Production, physico-chemical and biological properties, and mechanism of action, *J. Antibiot.* **1983**, *36*, pp. 1150–1156.
- (42) Hardt, I. H.; Steinmetz, H.; Gerth, K.; Sasse, F.; Reichenbach, H.; Höfle, G. New natural epothilones from *Sorangium cellulosum*, strains So ce90/B2 and So ce90/D13: isolation, structure elucidation, and SAR studies, *J. Nat. Prod.* **2001**, *64*, pp. 847–856.
- (43) Vollbrecht, L.; Steinmetz, H.; Höfle, G.; Oberer, L.; Rihs, G.; Bovermann, G.; Matt, P. von Argyrins, immunosuppressive cyclic peptides from myxobacteria. II. Structure elucidation and stereochemistry, *J. Antibiot.* **2002**, *55*, pp. 715–721.

- (44) Rateb, M. E.; Yu, Z.; Yan, Y.; Yang, D.; Huang, T.; Vodanovic-Jankovic, S.; Kron, M. A.; Shen, B. Medium optimization of *Streptomyces* sp. 17944 for tirandamycin B production and isolation and structural elucidation of tirandamycins H, I and J, *J. Antibiot.* **2014**, *67*, pp. 127–132.
- (45) Parra, R.; Aldred, D.; Magan, N. Medium optimization for the production of the secondary metabolite squalestatin S1 by a *Phoma* sp. combining orthogonal design and response surface methodology, *Enzyme Microb. Technol.* **2005**, *37*, pp. 704–711.
- (46) Jacob, J.; Rajendran, R. U.; Priya, S. H.; Purushothaman, J.; Saraswathy Amma, D. K. B. N. Enhanced antibacterial metabolite production through the application of statistical methodologies by a *Streptomyces nogalater* NIIST A30 isolated from Western Ghats forest soil, *PLoS ONE*. **2017**, *12*, e0175919.
- (47) Zabala, D.; Braña, A. F.; Flórez, A. B.; Salas, J. A.; Méndez, C. Engineering precursor metabolite pools for increasing production of antitumor mithramycins in *Streptomyces argillaceus*, *Metab. Eng.* **2013**, *20*, pp. 187–197.
- (48) Tanaka, Y.; Komatsu, M.; Okamoto, S.; Tokuyama, S.; Kaji, A.; Ikeda, H.; Ochi, K. Antibiotic overproduction by *rpsL* and *rsmG* mutants of various actinomycetes, *Appl. Environ. Microbiol.* **2009**, *75*, pp. 4919–4922.
- (49) Weber, T.; Charusanti, P.; Musiol-Kroll, E. M.; Jiang, X.; Tong, Y.; Kim, H. U.; Lee, S. Y. Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes, *Trends Biotechnol.* **2015**, *33*, pp. 15–26.
- (50) Korp, J.; Vela Gurovic, M. S.; Nett, M. Antibiotics from predatory bacteria, *Beilstein J. Org. Chem.* **2016**, *12*, pp. 594–607.
- (51) Wenzel, S. C.; Müller, R. Host Organisms: Myxobacterium. In *Industrial biotechnology, Microorganisms Volume 3a and 3b*; Wittmann, C.; Liao, J., Eds.; Wiley-VCH: Weinheim, Germany, 2017, pp. 453–485.
- (52) Manderscheid, N.; Bilyk, B.; Busche, T.; Kalinowski, J.; Paululat, T.; Bechthold, A.; Petzke, L.; Luzhetskyy, A. An influence of the copy number of biosynthetic gene clusters on the production level of antibiotics in a heterologous host, *J. Biotechnol.* **2016**, *232*, pp. 110–117.
- (53) Ongley, S.; Bian, X.; Neilan, B. A.; Müller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways, *Nat. Prod. Rep.* **2013**, *30*, pp. 1121–1138.
- (54) Gomez-Escribano, J. P.; Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters, *Microb Biotechnol.* **2011**, *4*, pp. 207–215.
- (55) Zhou, M.; Jing, X. Y.; Xie, P. F.; Chen, W. H.; Wang, T.; Xia, H. Y.; Qin, Z. J. Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of *Streptomyces coelicolor*, *FEMS Microbiol. Lett.* **2012**, *333*, pp. 169–179.
- (56) Ikeda, H.; Kazuo, S.-y.; Omura, S. Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters, *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, pp. 233–250.

- (57) Komatsu, M.; Uchiyama, T.; Omura, S.; Cane, D. E.; Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism, *Proc. Natl. Acad. Sci. USA*. **2010**, *107*, pp. 2646–2651.
- (58) Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A. Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters, *Metab. Eng.* **2018**, *49*, pp. 316–324.
- (59) Stevens, D. C.; Hari, T. P. A.; Boddy, C. N. The role of transcription in heterologous expression of polyketides in bacterial hosts, *Nat. Prod. Rep.* **2013**, *30*, pp. 1391–1411.
- (60) Bakke, I.; Berg, L.; Aune, T. E.; Brautaset, T.; Sletta, H.; Tondervik, A.; Valla, S. Random mutagenesis of the Pm promoter as a powerful strategy for improvement of recombinant gene expression, *Appl. Environ. Microbiol.* **2009**.
- (61) Siegl, T.; Tokovenko, B.; Myronovskyi, M.; Luzhetskyy, A. Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes, *Metab. Eng.* **2013**, *19*, pp. 98–106.
- (62) Iniesta, A. A.; García-Heras, F.; Abellón-Ruiz, J.; Gallego-García, A.; Elías-Arnanz, M. Two systems for conditional gene expression in *Myxococcus xanthus* inducible by isopropyl- $\beta$ -D-thiogalactopyranoside or vanillate, *J. Bacteriol.* **2012**, *194*, pp. 5875–5885.
- (63) Rentsch, A.; Kalesse, M. The total synthesis of corallopyronin A and myxopyronin B, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, pp. 11381–11384.
- (64) Hüttel, S.; Testolin, G.; Herrmann, J.; Planke, T.; Gille, F.; Moreno, M.; Stadler, M.; Brönstrup, M.; Kirschning, A.; Müller, R. Discovery and Total Synthesis of Natural Cystobactamid Derivatives with Superior Activity against Gram-Negative Pathogens, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, pp. 12760–12764.
- (65) Markham, A. Oritavancin. First global approval, *Drugs*. **2014**, *74*, pp. 1823–1828.
- (66) Conlin, A.; Fornier, M.; Hudis, C.; Kar, S.; Kirkpatrick, P. Ixabepilone, *Nat. Rev. Drug Discov.* **2007**, *6*, pp. 953–954.
- (67) Zhanel, G. G.; Calic, D.; Schweizer, F.; Zelenitsky, S.; Adam, H.; Lagacé-Wiens, P. R. S.; Rubinstein, E.; Gin, A. S.; Hoban, D. J.; Karlowsky, J. A. New lipoglycopeptides. A comparative review of dalbavancin, oritavancin and telavancin, *Drugs*. **2010**, *70*, pp. 859–886.
- (68) Lee, F. Y.; Borzilleri, R.; Fairchild, C. R.; Kim, S. H.; Long, B. H.; Reventos-Suarez, C.; Vite, G. D.; Rose, W. C.; Kramer, R. A. BMS-247550. A novel epothilone analog with a mode of action similar to paclitaxel but possessing superior antitumor efficacy, *Clin. Cancer Res.* **2001**, *7*, pp. 1429–1437.
- (69) Kirschning, A.; Taft, F.; Knobloch, T. Total synthesis approaches to natural product derivatives based on the combination of chemical synthesis and metabolic engineering, *Org. Biomol. Chem.* **2007**, *5*, pp. 3245–3259.
- (70) Weissman, K. J. Mutasyntesis - uniting chemistry and genetics for drug discovery, *Trends Biotechnol.* **2007**, *25*, pp. 139–142.
- (71) Kennedy, J. Mutasyntesis, chemobiosynthesis, and back to semi-synthesis: combining synthetic chemistry and biosynthetic engineering for diversifying natural products, *Nat. Prod. Rep.* **2008**, *25*, pp. 25–34.
- (72) Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines, *Nat. Prod. Rep.* **2016**, *33*, pp. 317–347.

(73) Klaus, M.; Grininger, M. Engineering strategies for rational polyketide synthase design, *Nat. Prod. Rep.* **2018**.

(74) Traber, R.; Hofmann, H.; Kobel, H. Cyclosporins--new analogues by precursor directed biosynthesis, *J. Antibiot.* **1989**, *42*, pp. 591–597.

(75) Lowden, P. A. S.; Böhm, G. A.; Metcalfe, S.; Staunton, J.; Leadlay, P. New Rapamycin Derivatives by Precursor-Directed Biosynthesis, *ChemBioChem.* **2004**, pp. 535–538.

(76) Zlatopolskiy, B. D.; Radzom, M.; Zeeck, A.; Meijere, A. de Synthesis and precursor-directed biosynthesis of new hormaomycin analogues, *Eur. J. Org. Chem.* **2006**, pp. 1525–1534.

(77) Hill, A. M.; Thompson, B. L. Novel soraphens from precursor directed biosynthesis, *Chem. Commun. (Camb.)*. **2003**, pp. 1360–1361.

(78) Goss, R. J. M.; Newill, P. L. A. A convenient enzymatic synthesis of L-halotryptophans, *Chem. Commun. (Camb.)*. **2006**, pp. 4924–4925.

(79) Winn, M.; Roy, A. D.; Grüschow, S.; Parameswaran, R. S.; Goss, R. J. M. A convenient one-step synthesis of L-aminotryptophans and improved synthesis of 5-fluorotryptophan, *Bioorg. Med. Chem. Lett.* **2008**, *18*, pp. 4508–4510.

(80) Smith, D. R. M.; Willemse, T.; Gkotsi, D. S.; Schepens, W.; Maes, B. U. W.; Ballet, S.; Goss, R. J. M. The first one-pot synthesis of L-7-iodotryptophan from 7-iodoindole and serine, and an improved synthesis of other L-7-halotryptophans, *Org. Lett.* **2014**, *16*, pp. 2622–2625.

(81) Corr, M. J.; Smith, D. R.; Goss, R. J. One-pot access to 1-5,6-dihalotryptophans and 1-alknlytryptophans using tryptophan synthase, *Tetrahedron.* **2016**, *72*, pp. 7306–7310.

(82) Ritacco, F. V.; Graziani, E. I.; Summers, M. Y.; Zabriskie, T. M.; Yu, K.; Bernan, V. S.; Carter, G. T.; Greenstein, M. Production of novel rapamycin analogs by precursor-directed biosynthesis, *Appl. Environ. Microbiol.* **2005**, *71*, pp. 1971–1976.

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

(84) Sahner, J. H.; Sucipto, H.; Wenzel, S. C.; Groh, M.; Hartmann, R. W.; Müller, R. Advanced mutasynthesis studies on the natural  $\alpha$ -pyrone antibiotic myxopyronin from *Myxococcus fulvus*, *ChemBioChem.* **2015**, *16*, pp. 946–953.

(85) Roy, A. D.; Gruschow, S.; Cairns, N.; Goss, R. J. M. Gene expression enabling synthetic diversification of natural products: chemogenetic generation of pacidamycin analogs, *J. Am. Chem. Soc.* **2010**, *132*, pp. 12243–12245.

(86) Yin, X.; Chen, Y.; Zhang, L.; Wang, Y.; Zabriskie, T. M. Enduracidin analogues with altered halogenation patterns produced by genetically engineered strains of *Streptomyces fungicidicus*, *J. Nat. Prod.* **2010**, *73*, pp. 583–589.

(87) Werneburg, M.; Busch, B.; He, J.; Richter, M. E.; Xiang, L.; Moore, B. S.; Roth, M.; Dahse, H. M.; Hertweck, C. Exploiting enzymatic promiscuity to engineer a focused library of highly selective antifungal and antiproliferative aureothin analogues, *J. Am. Chem. Soc.* **2010**, *132*, pp. 10407–10413.

(88) Miao, V.; Coeffet-Le Gal, M. F.; Nguyen, K.; Brian, P.; Penn, J.; Whiting, A.; Steele, J.; Kau, D.; Martin, S.; Ford, R.; Gibson, T.; Bouchard, M.; Wrigley, S. K.; Baltz, R. H. Genetic

- engineering in *Streptomyces roseosporus* to produce hybrid lipopeptide antibiotics, *Chem. Biol.* **2006**, *13*, pp. 269–276.
- (89) Reeves, C. D.; Ward, S. L.; Revill, W. P.; Suzuki, H.; Marcus, M.; Petrakovsky, O. V.; Marquez, S.; Fu, H.; Dong, S. D.; Katz, L. Production of Hybrid 16-Membered Macrolides by Expressing Combinations of Polyketide Synthase Genes in Engineered *Streptomyces fradiae* Hosts, *Chem. Biol.* **2004**, *11*, pp. 1465–1472.
- (90) Nguyen, K. T.; Ritz, D.; Gu, J. Q.; Alexander, D.; Chu, M.; Miao, V.; Brian, P.; Baltz, R. H. Combinatorial biosynthesis of novel antibiotics related to daptomycin, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, pp. 17462–17467.
- (91) Menzella, H. G.; Reid, R.; Carney, J. R.; Chandran, S. S.; Reisinger, S. J.; Patel, K. G.; Hopwood, D. A.; Santi, D. V. Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes, *Nat. Biotechnol.* **2005**.
- (92) Mootz, H. D.; Kessler, N.; Linne, U.; Eppelmann, K.; Schwarzer, D.; Marahiel, M. A. Decreasing the Ring Size of a Cyclic Nonribosomal Peptide Antibiotic by In-Frame Module Deletion in the Biosynthetic Genes, *J. Am. Chem. Soc.* **2002**, *124*, pp. 10980–10981.
- (93) Butz, D.; Schmiederer, T.; Hadatsch, B.; Wohlleben, W.; Weber, T.; Sussmuth, R. D. Module extension of a non-ribosomal peptide synthetase of the glycopeptide antibiotic balhimycin produced by *Amycolatopsis balhimycina*, *ChemBioChem.* **2008**, *9*, pp. 1195–1200.
- (94) Rowe, C.; Bohm, I.; Thomas, I.; Wilkinson, B.; Rudd, B.; Foster, G.; Blackaby, A.; Sidebottom, P.; Roddis, Y.; Buss, A.; Staunton, J.; Leadlay, P. Engineering a polyketide with a longer chain by insertion of an extra module into the erythromycin-producing polyketide synthase, *Chem. Biol.* **2001**, *8*, pp. 475–485.
- (95) Crusemann, M.; Kohlhaas, C.; Piel, J. Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains, *Chem. Sci.* **2013**, *4*, pp. 1041–1045.
- (96) Kries, H.; Niquille, D. L.; Hilvert, D. A subdomain swap strategy for reengineering nonribosomal peptides, *Chem. Biol.* **2015**, *22*, pp. 640–648.
- (97) Döhren, H. von; Dieckmann, R.; Pavela-Vrancic, M. The nonribosomal code, *Chem. Biol.* **1999**, *6*, pp. R273–R279.
- (98) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chem. Biol.* **1999**, *6*, pp. 493–505.
- (99) Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B. *De novo* design and engineering of non-ribosomal peptide synthetases, *Nature Chem.* **2018**, *10*, pp. 275–281.
- (100) Stachelhaus, T.; Schneider, A.; Marahiel, M. A. Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains, *Science.* **1995**, *269*, pp. 69–72.
- (101) Doekel, S.; Marahiel, M. A. Dipeptide formation on engineered hybrid peptide synthetases, *Chem. Biol.* **2000**, *7*, pp. 373–384.
- (102) Bian, X.; Plaza, A.; Yan, F.; Zhang, Y.; Müller, R. Rational and efficient site-directed mutagenesis of adenylation domain alters relative yields of luminide derivatives in vivo, *Biotechnol. Bioeng.* **2015**, *112*, pp. 1343–1353.
- (103) Han, J. W.; Kim, E. Y.; Lee, J. M.; Kim, Y. S.; Bang, E.; Kim, B. S. Site-directed modification of the adenylation domain of the fusaricidin nonribosomal peptide synthetase for enhanced production of fusaricidin analogs, *Biotechnol. Lett.* **2012**, *34*, pp. 1327–1334.

(104) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs), *Nucleic Acids Res.* **2005**, *33*, pp. 5799–5808.