Investigation and Genetic Engineering of Heterologous Hosts for the Production of Cyanobacterial Compounds

Zur Erlangung des akademischen Grades eines

Dr. rer. nat.

von der Fakultät Bio- und Chemieingenieurwesen
der Technischen Universität Dortmund
genehmigte Dissertation

vorgelegt von

M.Sc. Anna Tippelt

aus

Leipzig

Tag der mündlichen Prüfung: 06.09.2022

1. Gutachter: Prof. Dr. Markus Nett

2. Gutachter: Prof. Dr. Nadine Ziemert

Dortmund 2022

This work was conducted under the supervision of Prof. Dr. Markus Nett, TU Dortmund University, head of the research group „Technical Biology“.

The presented experimental investigations were conducted as part of the collaborative project “PRODIGY” (Process-Directed Drug Generation in Yeast), which was publicly funded by the German Federal Ministry of Education and Research - BMBF (German: Bundesministerium für Bildung und Forschung).

Parts of this work have been published by the author of this thesis and/or are based on data which were obtained in a BSc thesis performed at the Laboratory of Technical Biology under the supervision of the author. The following list gives a precise overview of the individual contributions.

Chapter 3.2.6 and 4.1.1 in parts modified from [I]

Chapter 1.2, 1.3, 1.3.2 and 4.2 in parts modified from [II]

Chapter 4.1.1 in parts modified from [III]

Chapter 3.1.2, 3.5.1, 4.4.1, 7.1 and 7.2 in parts based on data, strains and vectors from [a]

Relevant Publications

[I] Tippelt, A., Busche, T., Rückert, C., Nett, M. Complete Genome Sequence of the Cryptophycin-Producing Cyanobacterium *Nostoc* sp. Strain ATCC 53789. Microbiol Resour Announc, 9, e00040-20 (2020).

[II] Tippelt A., Nett, M. *Saccharomyces cerevisiae* as host for the recombinant production of polyketides and nonribosomal peptides. Microb Cell Fact 20, 161 (2021).

[III] Tippelt A., Nett, M. Charting the Layout of a Bacterial Factory for Anticancer Drugs - Genome sequencing of *Nostoc* sp. ATCC 53789. SCIENTIFIC HIGHLIGHTS 2020-Annual Report, TU Dortmund University (2020).

Relevant supervised student theses

[a] Schwing, M.D. Rekonstitution einer Phosphopantetheinyltransferase aus *Nostoc* sp. ATCC 53789 in *Escherichia coli*. Bachelor Thesis, Laboratory of Technical Biology, TU Dortmund University (2020)

Table of Contents

Table of Contents

Table of Contents

Zusammenfassung

Cyanobakterien sind eine wertvolle Ressource für die Entdeckung neuer, bioaktiver Sekundärmetabolite. Insbesondere im letzten Jahrzehnt wurde durch die rasante Entwicklung verbesserter Genomsequenzierungstechniken und bioinformatischer Analyse-Tools nicht nur ihre genetische Beschaffenheit, sondern auch ihr enormes Biosynthesepotential weiter erschlossen. Industriell bleiben diese wertvollen Ressourcen jedoch meist ungenutzt, was auf Kultivierungsschwierigkeiten, eingeschränkte Biosynthesekapazitäten sowie gentechnische Unzugänglichkeit der nativen Produzenten zurückzuführen ist. Biotechnologische Lösungskonzepte, wie die Anwendung heterologer Produktionssysteme eröffnen eine Möglichkeit zur industriellen Erschließung dieses Biosynthesepotentials.

Die in dieser Arbeit durchgeführte Genomsequenzierung ermöglichte faszinierende Einblicke in die genomische Struktur und das Biosynthesepotential des bisher unsequenzierten Cryptophycinproduzenten *Nostoc* sp. ATCC 53789. So ergab sich das Bild eines 8,7-Mb großen Genoms mit 13 Replikons, dessen Biosynthesepotential typische cyanobakterielle Charakteristika mit überwiegend peptidischen Sekundärmetaboliten aufwies. Die Mehrheit der Biosynthesegencluster (BGCs) entfiel auf das Chromosom. Insgesamt 17% der BGCs, darunter auch der Cryptophycin-Locus, erwiesen sich als Plasmid-assoziiert.

Zur Erschließung des cyanobakteriellen Biosynthesepotentials wurden unterschiedliche heterologe Wirtsorganismen, aber auch alternative, zellfreie Produktionssysteme untersucht, wobei sich letztere aufgrund unterschiedlicher Substratpräferenzen als ineffektiv erwiesen. Hingegen konnten die entwickelten *in vivo* Systeme, hier *S. cerevisiae* und *E. coli*, erfolgreich für die Rekonstitution und Genexpression kleiner Biosynthesegene und multimodularer Assemblierungslinien, wie Scytonemin und Cryptophycin, getestet werden. Diesbezüglich dokumentierte die effektive Rekonstruktion des Cryptophycin-Clusters erstmals die Genexpression einer bakteriellen, multimodularen NRPS-PKS-Assemblierungslinie in *S. cerevisiae*. Als zentrale Herausforderung gestaltete sich im Allgemeinen die Expression großer Biosntheseproteine in den heterologen Wirten sowie die Konstruktion eines eukaryotischen Plattformorganismus zur Expression großer bakterieller Gen-Cluster. Die Installation von Löslichkeits-Tags zur Vermeidung dysfunktionaler Proteinaggregate sowie die Etablierung eines vielseitig einsetzbaren Bakterien-Hefe-Klonierungssystems bildeten dabei erfolgreiche Lösungskonzepte. Die Metabolitenproduktion in den bisher nicht optimierten Chassisorganismen wurde anschließend analysiert, wobei eine hohe Zusatzbelastung des Zellstoffwechsels, eine geringe Präkursorenverfügbarkeit, die abweichende Codonverwendung und konkurrierende Stoffwechselwege als zentrale Engpässe identifiziert wurden.

Abstract

Cyanobacteria are a prolific resource for the discovery of valuable bioactive compounds. Also, within the last decade, the fast development of genome sequencing techniques and bioinformatics tools, significantly expanded the knowledge on their genomic and biosynthetic constitution. Yet, these resources remain industrially unexploited, due to cultivation difficulties, biosynthetic constraints and molecular genetic intractability of the native producers. Applying heterologous production systems is a viable biotechnological concept to access this potential on an industrial scale.

The herein conducted *de-novo* sequencing of the cryptophycin producer *Nostoc* sp. ATCC 53789 provided intriguing insights into its genomic and biosynthetic constitution with an 8.7-Mb genome, comprising 13 replicons and exhibiting a typical cyanobacterial secondary metabolome. It predominantly comprises peptide-associated biosynthetic gene clusters (BGCs), residing on the chromosome. Yet, 17% of the BGCs are plasmid-born, including the cryptophycin locus.

In the course of this dissertation, the biosynthetic potential of cyanobacteria was examined, developing heterologous hosts and alternative cell-free production systems, whereby the latter turned out ineffective due to restricted substrate preferences. However, the devised *in vivo* chassis, i.e. *S. cerevisiae* and *E. coli*, were successfully tested for the reconstruction of small BGCs and multimodular assembly lines (i.e. scytonemin and cryptophycin). Notably, the successful establishment of the cryptophycin BGC represent the first documentation on the efficacious reconstruction and expression of a large bacterial, multimodular NRPS-PKS assembly line in yeast.

In general, the reconstitution of large proteins and the construction of a eukaryotic chassis expressing large bacterial BGCs were central challenges in the chosen hosts. They were successfully addressed by the installation of solubility tags, avoiding inclusion-body formation and the establishment of a versatile bacteria-yeast cloning system. The encountered difficulties of final metabolite production in the non-engineered hosts were subsequently analyzed, identifying an increased metabolic burden, precursor limitations, codon usage and pathway competition as main drawbacks to devise a productive chassis.

1 Introduction

1.1 Cyanobacteria

Cyanobacteria represent an ancient phylum, which comprises a group of morphologically diverse Gram-negative bacteria performing oxygenic photosynthesis (Figure 1). As the evolutionary oldest phototrophs, they were crucial for the development of an oxygen-containing atmosphere on earth ~2.7 billion years ago, driving the evolution of life and higher organisms, which also led to the development of chloroplasts in plants. Today they are still the most abundant and important phototrophs on earth, contributing up to 80% of the marine and up to 35% of the world wide photosynthesis activity [1,2]. These pioneering phototrophs inhabit diverse, even extreme environments and niches, ranging from terrestrial, marine and freshwater habitats, up to hot springs, deserts and the Antarctica. Cyanobacteria, which can be classified into five morphological groups (Chroococcales, Pleurocapsales, Oscillatoriales, Nostocales, Stigonematales), proliferate as unicellular individuals (e.g. *Synechococcus* sp. and *Gloeothece* sp.) or by forming large consortia of filaments in (branched-)chain like trichomes and large bacterial mats (e.g. *Lyngbya* sp., *Nostoc* sp. and *Fischerella* sp.) . Some species are renowned for causing severe cyanobacterial blooms in eutrophic water bodies through mass-development events during summer season. Other species are capable of nitrogen (N_2) fixation, which makes them attractive symbiotic partners [3]. Symbiotic relations are established with a variety of organisms, including plants, fungi, animals and protists, whereby the cyanobiont resides obligately intracellular, or facultatively extracellular in (microaerophilic) cavities and tissues of the host. The symbiotic relation of the host is exclusive with only one cyanobacterial genus and can last from one up to perpetual generations. A close cell-cell contact with the interacting host is guaranteeing an optimal nutrient exchange between the partners. As such, nutrients are exchanged as fixed N_2 in the form of ammonia, citrulline and glutamine or as fixed carbon in the form of glucose and sucrose. Dependent on the partner organism the cyanobiont is either performing photosynthesis, which is the case for diatoms and fungal cyanobionts, or is mainly dedicated to N_2 fixation, which is common for photosynthetically active hosts. In case of N_2 fixation symbiosis, heterocyst-forming cyanobacteria, predominantly of the genus *Nostoc*, are involved as cyanobionts [4].

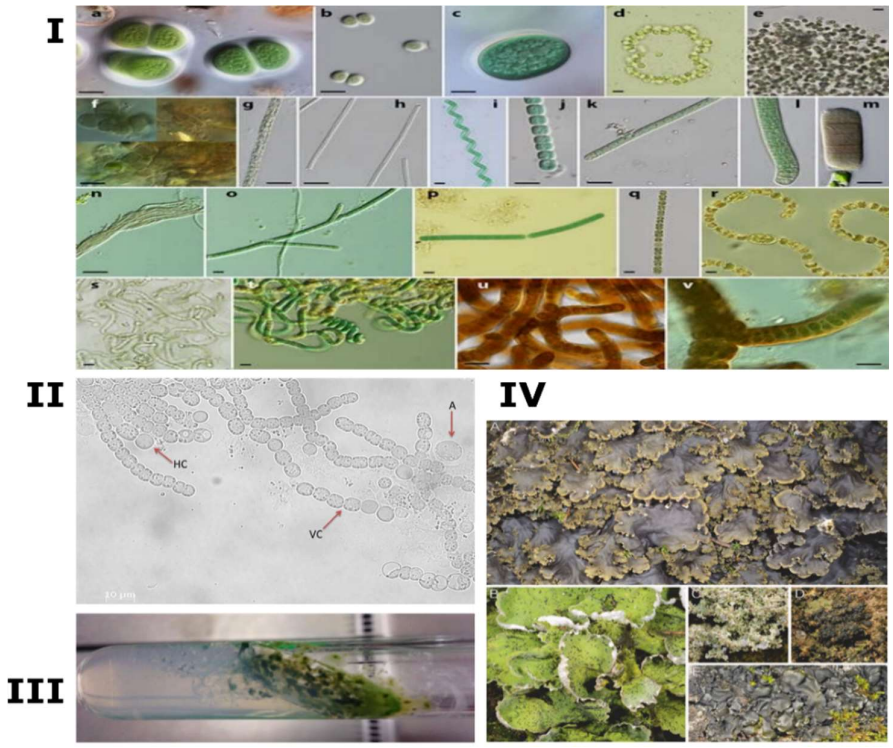


Figure 1. Diversity of cyanobacteria. I: Diverse morphology of cyanobacterial species [5] (Reprinted from Dvořák *et al.* [5] with permission from Springer International Publishing Switzerland, Copyright 2015) ; II: Microscopic view of the trichrome-forming cyanobacterium *Nostoc* sp. ATCC 53789; Bar: 10 µm, HC: heterocyst. Notation: VC, vegetative cell; A: akinete. Magnification: x1000; Microscope: Zeiss Axio Observer-D1; Zeiss AxioCam MRm; Software: Axioplan R4.8; III: Slant-agar culture of *Nostoc* sp. ATCC 53789; IV: Diverse species of cyanolichens [6] (Reprinted from Rikkinen [6] with permission from Springer International Publishing Switzerland, Copyright 2017).

Cyanobacteria exhibit a great structural and functional cell plasticity, which is also reflected in the multicellular organization of trichome forming species and their capability of a function-dedicated cell differentiation within these trichome structures. Trichomes are chain-like or branched chain-like filaments of photosynthetically active vegetative cells, some of which can undergo distinctive cell differentiation, forming e.g. heterocysts, akinetes or hormogonia (Figure 1) [4]. The basic structure of the thread-like trichomes is composed of a septa forming cell-chain filament, exhibiting four characteristic features, namely a cytoplasmic membrane, a peptidoglycan layer, an outer membrane generating the periplasmic space and an extracellular anionic envelope. The cytoplasmic membrane and the peptidoglycan layer surround each individual cell of the filament, whereby the peptidoglycan sheet is, unlike other Gram-negative bacteria, composed of multiple thick and highly cross-linked layers. This peptidoglycan mesh is densely connected at the septal cell junctions of the filament by septal disks which, in addition to the outer membrane, are responsible to maintain the filament integrity. The peptidoglycan disks are frequently traversed by nanopores at the emerging septa, which are channeled by microplasmodesmata to directly connect neighboring cells of the filament. This connection is vital

for a direct intercellular signaling and exchange of regulators and metabolites in between the vegetative as well as differentiated cells of the trichome. The outer membrane is enclosing the total filament to create a comprehensive periplasmic space, without entering the septa in between the cells. The resulting structural and functional continuous periplasm is facilitating intercellular communication, or the trafficking of metabolites, nutrients, regulatory molecules and proteins alongside the filament [7–9].

Many filamentous and unicellular cyanobacteria further exhibit an additional extracellular layer of mucilaginous sheaths, slime or capsules. These anionic structures are mainly composed of polysaccharides, monosaccharides, deoxy-sugars and uronic acid, some of which also contain sulphur, peptide, acetate and pyruvate groups. These extracellular structures are providing diverse advantages, such as adhesion to solid surfaces in demanding habitats (i.e. dunes, intertidal zones), promotion of cell motility of gliding cyanobacteria and structural stabilization of the cell and filament integrity. Furthermore, they are protecting the cells from toxic heavy metals, radicals, desiccation, or UV irradiation by additional accumulation of protective proteins and pigments within the sheaths, i.e. superoxide dismutase, scytonemin and mycosporine-like amino acids [10].

The morphological and functional cell differentiation of trichome forming cyanobacteria is significantly contributing to their ability to adapt to diverse habitats and niches [4,7]. Nitrogen deprivation or the development of symbiotic interactions with a host organism is triggering the differentiation of some vegetative cells of the trichome to photosynthetically inactive heterocysts. Heterocysts function as a microaerobic environment for the nitrogenase-facilitated, oxygen-sensitive fixation of N_2. As such, they serve as a source of bound nitrogen for the aerobic vegetative cells, which in exchange provide photosynthetic products such as ATP, NADPH and sugar to the heterocyst. In comparison to vegetative cells, the larger heterocysts are also no longer capable of cell division and exhibit a thickened, laminated cell wall, which is ensuring the integrity of a microaerophilic heterocyst environment. A decreased number of microplasmodesmata is also helping to maintain this milieu and regulates the necessary interchange of metabolites with adjacent vegetative cells, while preventing gas exchange. Heterocysts are regularly interspacing the trichome filament to guarantee an optimal transfer of nutrients with vegetative cells, whereby the explicit pattern of the intervals is strain-specific. The development of heterocysts enabled a spatial segregation of oxygenic photosynthesis and N_2 fixation during the evolution of cyanobacteria and for this enabled them to thrive in a tremendous variety of habitats [4,11].

Akinetes are enlarged, thick-walled, dormant cells with a multiple-layered extracellular envelope. They are typically abundant in the orders Nostocales and Stigonematales and are known to survive adverse environmental conditions, like desiccation, cold temperatures, or nutrient

depletion for several decades. Some akinetes accumulate granules of cyanophycin, an amino-acid polymer, as a nitrogen reservoir. Compared to vegetative cells they generally display a reduced metabolic activity. However, dependent on their age, they still exhibit some degree of protein and lipid biosynthesis as well as photosynthesis and respiration [4,11].

Hormogonia are vegetative cells that reversibly differentiate from trichome cells into short motile filaments, which can be gas-vacuolated in some cyanobacterial species. Triggered e.g. by light, nutrient availability, or the presence of symbiotic partners, they play a central role in cell chemotaxis and symbiosis development [4,11,12].

1.2 Natural products from cyanobacteria and mechanisms of secondary metabolite assembly

The following section contains modified passages from the author's previous publication [II] [13].

The morphological and functional adaptability of cyanobacteria is also reflected in the broad diversity of their secondary metabolites [14]. Secondary metabolites are low molecular weight compounds that are produced by bacteria, fungi, plants, and marine organisms. Unlike primary metabolites, they are not essential for the growth or replication of the producing organism. However, they enhance the survival of the producer in adverse and competitive environments [15]. Their natural functions comprise the adaptation to environmental changes, the defense of resources or habitats, self-protection against predators, herbivores or solar irradiation, as well as inter- and intraspecific communication [15,16]. Secondary metabolites are a remarkable resource of high-value pharmaceutical drugs to combat infectious diseases, cancer and (auto)immune disorders [17]. Furthermore, they are used as fine and commodity chemicals, e.g. in the food and cosmetic industry [13,18–20].

Together with actinomycetes and myxobacteria, cyanobacteria constitute one of the most important prokaryotic sources for drug discovery [21,22]. Chemical investigations have yielded numerous natural products, many of which exhibit strong pharmacological activities already at low nanomolar concentrations. Furthermore, the remarkable progress in genome sequencing and bioinformatics has illuminated the enormous biosynthetic potential of this phylum within the last decade. The available genome sequences from cyanobacteria represent a treasure trove for the identification of new BGCs with unique biosynthetic mechanisms [23–26]. Cyanobacterial secondary metabolites are structurally highly diverse. They encompass neurotoxic and cytotoxic (kalkitoxin, antillatoxin, nostophycin, microcystin), antibiotic (noscomin, hapalindole T), antimitotic (cryptophycins, curacins, dolastatins), antiviral (nostoflan, spirulan), antiprotozoal (viridamide A, venturamides) and UV-protective (scytonemin, shinorine) metabolites, whereby the well-

studied genera *Lyngbya*, *Microcystis*, *Nostoc* and *Hapalosiphon* account for more than two-thirds of the so far described chemistry [3,14,21,24].

Many cyanobacterial secondary metabolites originate from polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS) and hybrid PKS-NRPS pathways [21]. PKSs and NRPSs are modularly organized megaenzymes that are composed of multiple catalytic domains. These enzymes catalyze the linkage of simple acyl- or amino acid-derived monomers in a sequential or iterative fashion. As modular PKSs and NRPSs are analyzed in detail in this dissertation, their mode of action will be introduced in the following section.

The minimal set of catalytic domains present in one archetypal type I PKS module comprises an acyl transferase (AT), a ketoacyl synthase (KS) and an acyl carrier protein (ACP) domain. The posttranslational activation of the latter is a crucial prerequisite for PKS biosynthesis and is carried out by dedicated phosphopantetheinyl transferases (PPTases). This phosphopantetheinylation reaction is Mg^{2+}-dependent and involves the transfer of the phosphopantetheine moiety from coenzyme A (CoA) onto a conserved serine residue of the ACP. Hereby, a PPTase catalyzes the nucleophilic side chain attack of the ACP serine on the 5'-pyrophosphate bond of CoA (Figure 2). The covalently tethered phosphopantetheine arm facilitates the binding of substrates and biosynthetic intermediates via a reactive thioester bond to the ACPs. Moreover, the innate flexibility of the 18 Å phosphopantetheine arm allows the transport of bound substrates to distant catalytic centers of the megaenzymes. PPTases are ubiquitous to all domains of life. Although they share only low levels of sequence homology in general, they can be classified into two major groups known as the AcpS-type and the Sfp-type PPTases [27–30]. The AcpS-type is mainly connected to fatty acid biosynthesis, whereas the broad-substrate-range Sfp-type is closely associated with secondary metabolism [27,28].

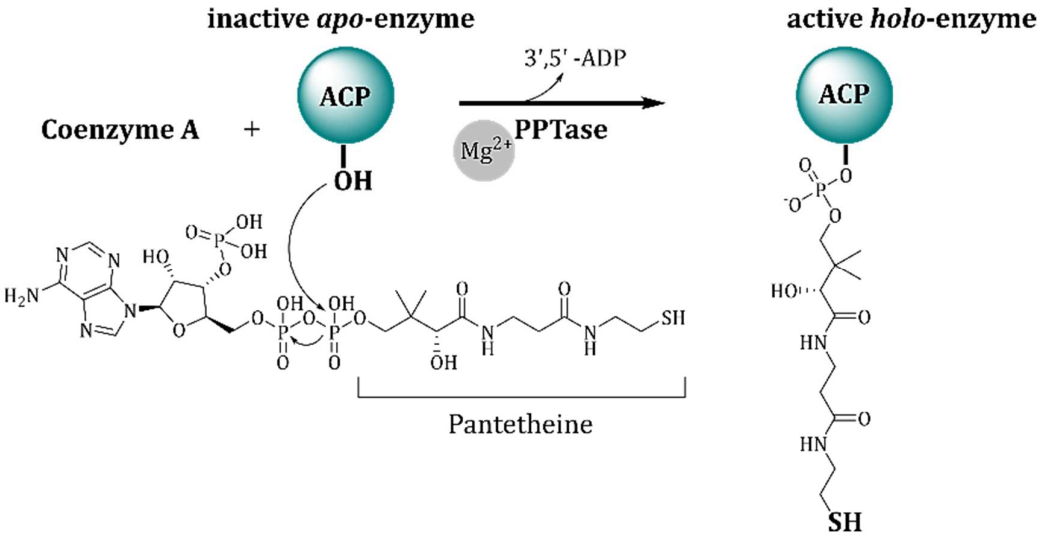


Figure 2. Phosphopantetheinylation of an acyl carrier protein. The phosphopantetheine unit of coenzyme A is transferred onto the conserved serine of the *apo*-enzyme, converting it to its active *holo*-form. Notation: PPTase, phosphopantetheinyl transferase; 3',5'-ADP, adenosine-3',5'-diphosphate.

In PKS biosynthesis (Figure 3), AT domains are responsible for selecting and loading acyl-CoA units onto ACP domains. Each single AT domain is selective for only one acyl-CoA unit according to its respective active site code [31]. The ACP is binding the selected substrate, shuttling it to distal catalytic domains of the assembly line like e.g. the KS domain. In multimodular PKSs, the ACP is further responsible to transfer the growing polyketide to the next module, likewise. The KS catalyzes the linkage of the extender unit to the growing polyketide chain in a decarboxylative *thio*-Claisen condensation reaction. In addition to this basic domain architecture, a subset of accessory, reductive domains can be part of a PKS module. The ketoreductase (KR) domain is performing a NADPH-dependent reduction of the β-carbonyl group in the Claisen product yielding a hydroxyl moiety. The latter can be further processed by a dehydratase (DH) domain, leading to the formation of a carbon-carbon double bond. An additional enoylreductase (ER) domain is eventually reducing the double bond to give a fully saturated moiety. However, this reductive cycle is not always completely represented or active in a PKS module. Dependent on the presence, activity and combination of the aforementioned accessory domains, a variety of fully, partially or non-reduced compounds can be formed, contributing to the tremendous diversity of polyketides.

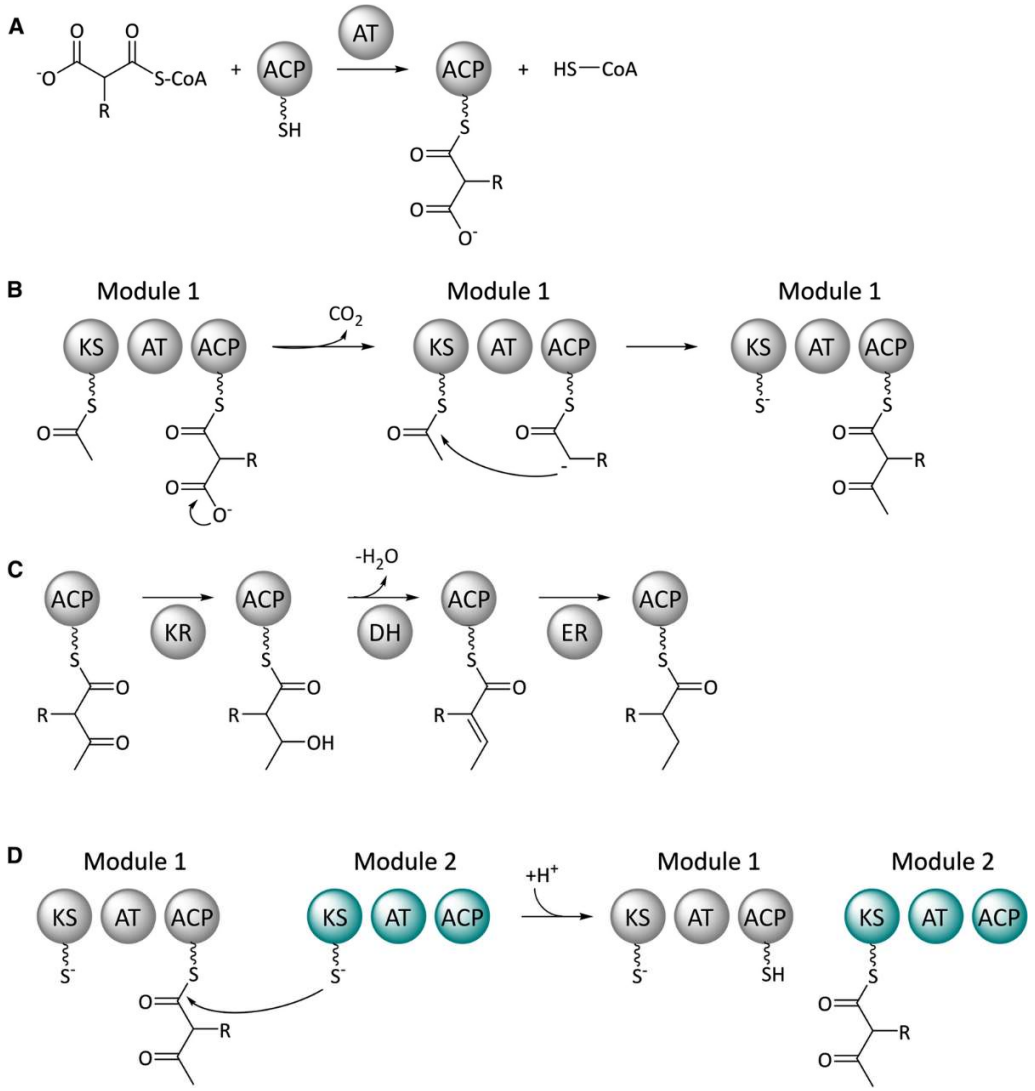


Figure 3. Schematic view of substrate selection (A), chain elongation (B), reductive processing (C), and intermediate transfer (D) by type I PKSs. Domain notation: KS β-ketoacylsynthase, AT acyltransferase, ACP acyl carrier protein, KR ketoreductase, DH dehydratase, ER enoyl reductase. Reprinted with permission from Tippelt and Nett [13], Copyright (2021) BMC, Springer Nature, Creative Commons Attribution 4.0 International license (CC BY 4.0)

In situ methylation by an accessory S-adenosylmethionine (SAM) dependent O- or C-methyltransferase (OMT, CMT) domain can further modify the bound polyketide. The release of the final product is mediated by a C-terminal thioesterase (TE) domain and can occur as simple intermolecular hydrolysis or by intramolecular macrocyclization, facilitated by an internal nucleophilic attack of e.g. a hydroxyl moiety [32–36]. The interrelation between domain organization and polyketide assembly is known as collinearity. This feature, in combination with the substrate preference of AT domains, allows the structural prediction of polyketides from their biosynthetic assembly lines[37–40].

In analogy to PKSs, modular NRPSs exhibit a functionally comparable domain architecture and assembly logic. Archetypal NRPSs also follow the collinearity rule. They feature an adenylation (A), a condensation (C) and a peptidyl carrier protein (PCP) domain. Similar to the AT domain, the A domain is responsible for substrate recognition. Unlike the former, however, the A domain also activates its substrate by the formation of an acyl-adenylate in an ATP-driven and Mg^{2+} dependent fashion. One single substrate is selected in conjunction with a specificity-conferring active site code [41]. Apart from proteinogenic L-amino acids, also non-proteinogenic amino acids and keto as well as aryl carboxylic acids can serve as substrates of A domains. The C domain corresponds to the KS domain of PKSs and catalyzes peptide or ester bond formation with an extender unit. The PCP has the same function as an ACP domain. It facilitates the binding and shuttling of substrates and biosynthetic intermediates. Unlike PKS assembly lines, the substrate monomers stay bound to the PCP during the assembly process. This minimal set of NRPS domains can be complemented by accessory domains, e.g. epimerization (E), oxidase (Ox), cyclization (Cy), or MT (O-, N-, C-linked) domains, which enhance the structural diversity of nonribosomal peptides. For instance, E domains offer an *in situ* epimerization of activated L-amino acids into their respective D-form prior to condensation with the growing peptide chain. Cy domains generate heterocyclic ring systems like thiazolines and oxazolines, which can be further processed by Ox domains, converting them to thiazole and oxazole moieties. The NRPS assembly is terminated through TE domains similar to PKS systems. Occasionally, the TE domain is missing in an NRPS assembly line and instead a C-terminal reductase (Red) or C domain catalyzes product liberation. Red domains mediate a reductive release, which results in aldehyde or alcohol formation, while dedicated C domains facilitate a direct nucleophilic attack to split the thioester bond or initiate cyclization to release the product [34–36,41–51].

Further product diversity can be achieved by post-release modifications through tailoring enzymes. Typical modifications include epoxidation, glycosylation (N-, C-, O-linked), oxygenation, hydroxylation and alkylation of the polyketide or peptide. In addition, carbon-bond cleavage rearrangements and cyclizations give rise to a tremendous structural diversity [27,52,53]. Another aspect contributing to the structural diversity of nonribosomal peptides is the promiscuity of their substrate selecting domains. Many A domains exhibit an extended substrate tolerance, while AT domains are in general highly selective [33,40,41,43,54,55]. The promiscuity of NRPSs is often exploited for precursor-directed biosynthesis and mutasynthesis approaches generating "unnatural" natural products with pharmaceutically desired properties [56].

Hybrid assembly lines feature a mixed PKS and NRPS module architecture. They are following the previously introduced logic of assembly and domain organization and increase the structural

diversity of natural products even further. Many clinically important compounds with anticancer (bleomycin, epothilone, cryptophycin), immunomodulatory (rapamycin) and antibiotic (pristinamycin, myxovirescin A) properties originate from hybrid PKS/NRPS pathways [34].

1.2.1 Cryptophycin – a cyanobacterial anticancer agent

A large number of cyanobacterial compounds exhibit potent anticancer properties and, therefore, have already been introduced into (pre)clinical trials or harnessed as promising drug leads (e.g. soblidotin [57], tasidotin [58], cryptophycin [59,60]). Among these agents, curacins, dolastatins, and cryptophycins were found to be active against various multidrug-resistant tumor cell lines [61–63] Of particular interest are the cryptophycins, comprising more than 25 native derivatives, some of which were used to develop synthetic analogues for (pre)clinical testing [62,64]. Among them, the prominent example cryptophycin 52, a synthetic congener of the native cryptophycin 1, has even been applied to phase II clinical trials in cancer therapy [59,60].

The antiproliferative efficiency of cryptophycin 1 exceeds approved antitumor agents, such as paclitaxel, vinblastine or vincristine [65]. Through inhibition of microtubule dynamics and (de)polymerization already at picomolar concentrations (e.g. 4.58-7.63 pM for KB and LoVo cell lines) [66], cryptophycin 1 and cryptophycin 52 are causing an arrest of the cell cycle in the transition of the mitotic metaphase/anaphase [65,67]. In addition it was found that both derivatives are irreversibly binding in close vicinity of the vinca alkaloid binding site at the ends of the microtubule. However, unlike vinblastine, both cryptophycin 1 and its congener cryptophycin 52 are causing conformational changes of the tubulin, which is ultimately affecting its association and dissociation potential [68]. Like many other microtubule-inhibiting agents, cryptophycin 1, cryptophycin 52 and another synthetic chlorohydrin-derivative, cryptophycin 55, were also found to additionally trigger an apoptotic cascade by causing a hyper-phosphorylation and inactivation of the Bcl-2 protein, a repressor protein in apoptosis regulation [66,69]. Due to their high affinity, the conformational change of the tubulin and the poor reversibility of the tubulin binding, the cryptophycins are hardly affected by resistance mechanisms of multidrug resistant tumor cell lines, which are mediated by the permeability-glycoprotein pumps (ABC transporters). As an example, cryptophycin 52 displayed an unaffected activity against vinca alkaloid and paclitaxel resistant tumor cell lines [68]. In sum, the low inhibitory concentration even against multidrug resistant cell lines classify the depsipeptide and its derivatives as some of the most potent antimitotic agents described to date. Therefore they represent some of the most promising sources for new chemotherapeutics [3,21,64].

To date two cyanobacterial strains, namely the terrestrial *Nostoc* sp. GSV224 and the lichen-isolated *Nostoc* sp. ATCC 53789, were reported to be capable of cryptophycin biosynthesis [64,70]. The cryptophycin (*crp*) operon (40.3 kb) comprises eight genes encoding the biosynthetic

enzymes CrpA-H, which are directing the biosynthesis of the more than 25 native cryptophycin derivatives (Figure 4) [64,71–73].

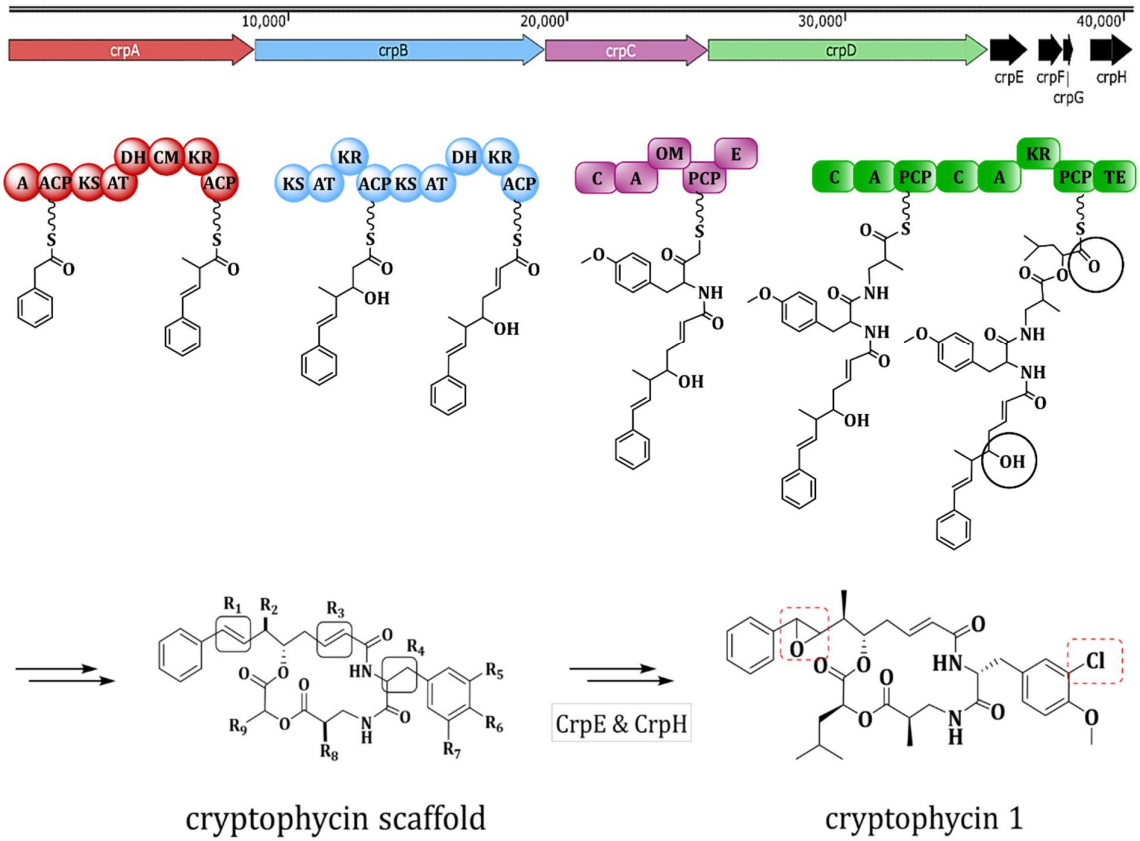


Figure 4. Cryptophycin biosynthesis. The cryptophycin locus comprises eight genes (*crpA-H*) with a total size of 40.3 kb. The enzymes CrpA-D are responsible for the formation of the cryptophycin scaffold. CrpE-H are enzymes, which introduce epoxide and halogen moieties (red squares) into the cryptophycin scaffold or provide specific precursor molecules for the biosynthesis. Variable residues are indicated by R1-9 [64]. Domain notation: A, adenylation; ACP/PCP, acyl/peptidyl carrier protein; KS, ketoacyl synthase; AT, acyl transferase; DH, dehydratase; CM, C-methyltransferase; KR, ketoreductase; C, condensation; OM, O-methyltransferase; E, epimerization; TE, thioesterase. The asterisk is indicating a hypothetical starter unit.

The cryptophycin assembly line encompasses the two type I PKSs CrpA and CrpB (loading module and modules 1-3) and the two NRPSs CrpC and CrpD (modules 4-6). The latter includes an usual PKS-associated KR domain [64,74]. The biosynthesis is initiated by the selection and tethering of an elusive starter unit through the CrpA loading module. Up to now, the identity of this starter unit

has not been clarified. Incorporation experiments with labelled precursors suggest the starter unit to derive from phenylalanine. However, only eight out of nine carbons of this amino acid could be detected in cryptophycin. Furthermore, feeding studies excluded phenylacetate as a direct biosynthetic precursor [64]. A similar situation was reported for microcystin biosynthesis [75], which is also assumed to start with a C6-C2 building block and involves a starter unit selecting A domain with a specificity conferring substrate code that is in accordance with CrpA [64,76]. Biochemical analyses of the microcystin A domain revealed *trans*-cinnamic acid and 3-phenylpropionic acid as the preferred substrates and not a phenylethyl or phenylacetate unit as expected. The exact mechanism by which the selected substrate is shortened and also the timing of this reaction remain elusive and require further exploration (cf. section 4.1.1.3) [64,76]. Following the loading process, module 1 extends the hypothetical starter unit by decarboxylative Claisen condensation with malonyl-CoA. A subsequent reductive processing (KR, DH) and SAM-dependent methylation (CMT) give rise to hypothesized intermediate 2-methyl-4-phenyl-3-butenoic acid (MPBA), which is enzyme-bound via a thioester bond (Figure 4). The biosynthesis is continued by CrpB. Two further rounds of polyketide chain extension and reductive processing are executed by modules 2 and 3. CrpC is responsible for the attachment of tyrosine in module 4, whereby the E domain is catalyzing the conversion of the selected L-configured substrate to D-tyrosine. Methylation of the tyrosine hydroxyl group is meanwhile carried out by the associated OM domain. The last assembly line enzyme, CrpD, comprises the modules 5 and 6, which successively incorporate 2*R*-methyl-β-alanine and α-ketoisocaproic acid into the growing peptide-polyketide chain. While the unusual amino acid 2*R*-methyl-β-alanine derives from methylaspartate, which is decarboxylated by CrpG, the α-keto acid is generated by CrpF, which catalyzes the oxidative deamination of leucine. The KR domain of CrpD is responsible for the reduction of the α-ketoisocaproic acid building block. The resulting α-hydroxyisocaproic acid allows an ester bond linkage with the previously incorporated 2*R*-methyl-β-alanine. This condensation reaction is performed by the C domain of module 6. The TE domain of module 6 catalyzes the concluding lactonization, whereby the hydroxyl group introduced in module 2 attacks the carbonyl group of the thioester to release a cyclic depsipeptide. The flavin-dependent halogenase CrpH and the P450 epoxidase CrpE are responsible for the halogenation and epoxidation of the product following its liberation from the assembly line [64]. CrpE was found to exhibit a significant substrate tolerance. Although the presence of an epoxide moiety is no prerequisite for bioactivity, it increases the potency of the resulting cryptophycin by two orders of magnitude [64,74,77].

The promiscuity of the Crp assembly line gives rise to various cryptophycin derivatives. Overall, nine variable residues (R1-9) have been found in naturally occurring cryptophycins (Figure 4). Apart from the mentioned epoxide replacement with a *trans*-styrene moiety (R1), halogenation

is also known for R7 and some natural cryptophycins exhibit a β-alanine instead of a 2-methyl-β-alanine moiety at R8 (Figure 4). The observation that a single NRPS pathway yields multiple product analogs is a common feature of cyanobacterial secondary metabolism and was also found in laxaphycin [78], nostopeptolide [70] and microcystin [79] biosynthesis.

1.2.2 Scytonemin – a cyanobacterial sunscreen

Cyanobacteria are versatile producers of structurally diverse secondary metabolites with a variety of bioactivities [3,80]. As cyanobacteria are living naturally in environments with extreme UV conditions, these phototrophic organisms evolved several protection strategies in response to detrimental UV irradiation, ranging from DNA repair mechanisms [81], via UV evasion behavior [82] up to the biosynthesis of UV-absorbing pigments [83,84]. Amongst the latter is the extracellular sheath pigment scytonemin, which exhibits effective sunscreen properties against UV-A, UV-B and UV-C irradiation and radical-scavenging activities [85]. UV-A and UV-B directly and indirectly damage cells on protein, nucleic acid, and lipid level, e.g. by causing the formation of reactive oxygen intermediates [86], or through direct lesion of the cyanobacterial phycobilisomes [87]. To avoid lethal cell damage by solar radiation, the yellow-green scytonemin pigment is accumulated in the extracellular sheaths upon UV-A light exposure [83,88]. Its broad, major absorption spectrum is ranging from 325-425 nm within the UV-A spectral region via the UV-B spectrum of 280-320 nm, down to the UV-C spectral region with a maximum of 250 nm [89,90].

The scytonemin molecule is a dimer. Its two monomers represent condensation products of tryptophan- and tyrosine-derived metabolites (Figure 5) [83,89]. Scytonemin has been isolated from a variety of cyanobacterial species living in aqueous and terrestrial habitats [89]. Its biosynthetic pathway is encoded by a cluster of 18 open reading frames (orfs). Only three of these genes, i.e. *scyC* (NpR1274), *scyB* (NpR1275) and *scyA* (NpR1276), are putatively essential for product formation [83,91,92]. The biosynthesis (Figure 5) is initiated by the leucine dehydrogenase homologue ScyB, which converts L-tryptophan to indole-3-pyruvate. Subsequently, the thiamin diphosphate (ThDP)-dependent enzyme ScyA condenses this intermediate with 4-hydroxyphenyl pyruvate into a labile β-ketoacid. Decarboxylation and oxidative cyclisation by ScyC generate a ketone, which eventually dimerizes to scytonemin. The function of the remaining genes is either elusive or dedicated to tryptophan and tyrosine precursor formation. For example the encoded prephenate dehydrogenase TyrA (NpR1269) is assumed to be responsible for the generation of 4-hydroxyphenyl pyruvate directly from prephenate instead of tyrosine [83,91,93,94].

Up to now, it is not clear whether the dimerization process occurs spontaneously or if an enzymatic process is involved. Currently, the locus-encoded ScyE (NpR1272) and an adjacent gene cluster, the *ebo* cluster, are under investigation for being involved in the dimerization process. Both, Δ*scyE* and Δ*ebo* mutants exhibited a scytonemin deficient phenotype and the

accumulation of scytonemin monomers in periplasmic extracts [92,95].

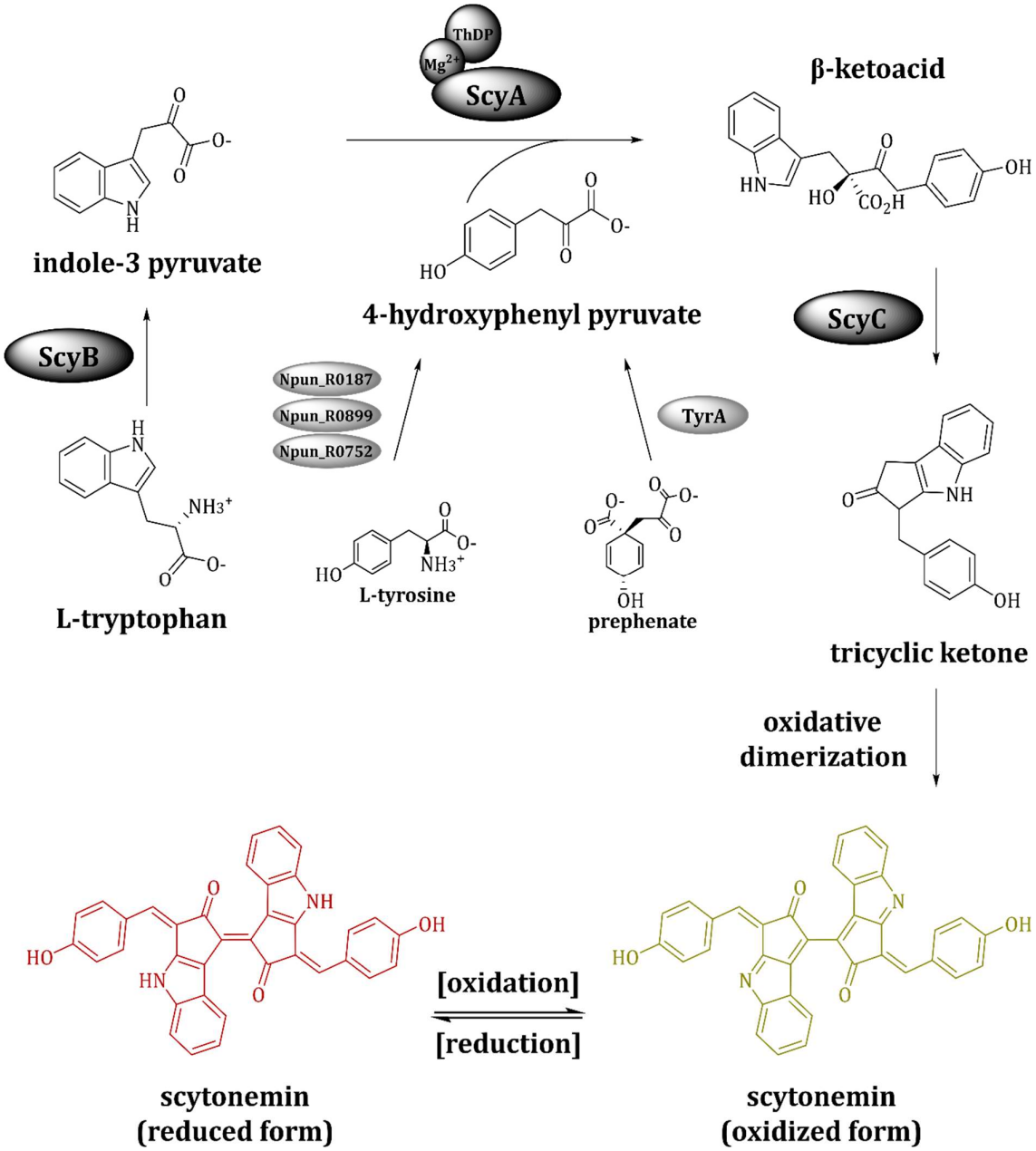


Figure 5. Scytonemin biosynthesis from L-tryptophan and 4-hydroxyphenyl pyruvate in *Nostoc punctiforme* ATCC 29133. Notation: ScyA, thiamine diphosphate (ThDP) and manganese (Mn^{2+}) dependent enzyme; ScyB, putative leucine dehydrogenase; ScyC, putative cyclase/decarboxylase; Npun_R0187/Npun_R0899/Npun_R752, native aminotransferases of *N. punctiforme*, which catalyze the transamination of L-tyrosine; TyrA, putative prephenate dehydrogenase [83,91,92].

1.3 Heterologous host systems for secondary metabolite biosynthesis

The following section was adapted from the author's previous publication [II] [13].

The development of economically feasible manufacturing processes is an ambitious challenge for many polyketides and (non-ribosomal) peptides. Chemical syntheses are often impractical, as the structural and stereochemical complexity of these molecules often requires multiple transformation steps, diminishing the overall yield. In addition it requires the laborious introduction and removal of protective groups and the usage of precious metal catalysts, or toxic reagents, which is affecting production costs, as well as environmental sustainability [13,96–99]. Biotechnological approaches, e.g. *in vivo* and *in vitro* biocatalysis, are often advantageous in catalyzing the required multistep transformations to generate complex secondary metabolites [98]. They provide high stereoselectivity and lead to enantiopure compounds [97,98,100,101]. In addition, biotechnological applications use milder, non-toxic operation conditions (e.g. ambient temperature, atmospheric pressure, aqueous solution). Starting materials for the production process include cheap and renewable substrate sources, e.g. glucose or sucrose. Since the use of sugars and other potential nutrient sources in biotechnology is highly controversial, processes are increasingly developed to deploy biomass from non-food resources (e.g., lignocellulose, glycerol, xylose) [102,103].

Overall, biotechnological processes can offer an alternative, environmentally more attractive route for the synthesis of bulk and fine chemicals [104,105]. In any case the environmental sustainability or impact of a process must be assessed individually e.g. taking resource efficiency, waste to product yield ratio (E factor), or resource efficiency and life cycle assessment (e.g. energy usage, eutrophication, ecotoxicity) into account [97,101,106].

Even though biotechnological processes are an attractive alternative to chemical synthesis, they can be hindered by low product yields and purification difficulties. In many cases, the natural producers of secondary metabolites show unfavorable process properties, which impede their industrial application [98]. Many microorganisms are difficult to cultivate under defined laboratory conditions, not amenable to classical experimental design, exhibit slow growth with doubling times of several days and show low productivity capacities [61,98,107]. In addition, high sensitivity against shear stress in stirred tank reactors, as observed for many filamentous and mycelia-forming microorganisms [108], may reduce productivity. Furthermore, downstream processing may be hindered by structurally similar molecules in the fermentation broth [98].

Apart from these process-associated limitations, another aspect is that many biosynthesis genes are silent or only weakly expressed in the native producer. In general, natural

product biosynthesis is tightly regulated by both, global and local transcription regulators that are controlled by various biotic and abiotic factors as well as environmental conditions. Factors that are known to affect secondary metabolite production include radiation, light, oxidative stress, habitat/nutrient limitations, as well as positive and negative interactions with other organisms. Especially cyanobacteria are known to exhibit a high evolutionary adaptability in response to environmental changes, which can lead to a genetic instability (e.g. sequence mutations) of BGCs under laboratory conditions [109–111]. Another drawback is that most of the native secondary metabolite producers are also not amenable to genetic engineering techniques, which would be required in order to activate these mostly silent BGCs and to increase the product yield of a desired compound. Yet, the development of such techniques and the establishment of suitable fermentation conditions for the native producer often require laborious empirical efforts [13,15,16,23,83,107,109,112–120].

One approach to overcome these limitations is the development and application of heterologous hosts for the biosynthesis of secondary metabolites. For this purpose, industrially proven host organisms are harnessed to reconstruct the biosynthesis of valuable target molecules. Methods of synthetic biology and metabolic engineering, such as pathway refactoring, support the rational development of highly productive strains [104,105]. *Escherichia coli*, *Saccharomyces cerevisiae*, or *Aspergillus nidulans* are frequently exploited as heterologous hosts to establish production processes for secondary metabolites [121–124]. In case of cyanobacterial compounds, heterologous (host) systems are also used to probe or manipulate BGCs [23,119]. Microorganisms like *E. coli*, *S. cerevisiae*, *Streptomyces venezuelae* and certain cyanobacteria, such as *Anabaena* sp. PCC 7120 have been successfully used as production hosts for cyanobacterial secondary metabolites [23]. Examples cover e.g. the heterologous production of microviridins [125], microcystin [126] and a scytonemin precursor [93] in *E. coli*, or shinorine [127] biosynthesis reconstruction in *S. cerevisiae. S. venezuelae* was applied to produce 4-O-demethylbarbamide [128], while cryptomaldamide [129] and lyngbyatoxin [119] biosynthesis were reconstructed in *Anabaena* sp. PCC 7120.

1.3.1 *Escherichia coli* as heterologous host

Since the 1980s *Escherichia coli* represents one of the most important and most frequently used workhorses in academia and industry. Favorable process properties, such as short doubling time, simple scale-up, high cell density growth, and high production yields together with its well-characterized biochemistry and physiology make *E. coli* an extremely versatile host for protein expression. In addition, its genetic tractability in combination with the availability of numerous genetic manipulation tools have promoted its industrial application for the production of recombinant biopharmaceuticals [130,131]. Examples include pharmaceuticals for the treatment of diabetes (human recombinant insulin, Eli Lilly), chronic hepatitis C infections (interferon α-2b,

Schering-Plough), multiple sclerosis (Interferon β-1b, Novartis), or chronic gout (Pegloticase, Savient Pharms) [130,132]. Moreover, *E. coli* is increasingly harnessed as a platform organism for the production of biofuels (bioethanol [133], 1-propanol [134], isopropanol [135], hydrogen [136]), sugar alcohols (xylitol [137], mannitol [138]), amino acids (L-threonine [139], L-phenylalanine [140]), organic acids (lactic acid [141], pyruvate [142]), diols and biopolymers (1,3-propanediol [143], polylactic acid [144]). These compounds find application in the energy, food, agricultural, cosmetics, fabric/textile and polyester industry [145]. *E. coli* has also been employed to produce a variety of secondary metabolites, including isoprenoids, nonribosomal peptides and polyketides (Figure 6) [145–147].

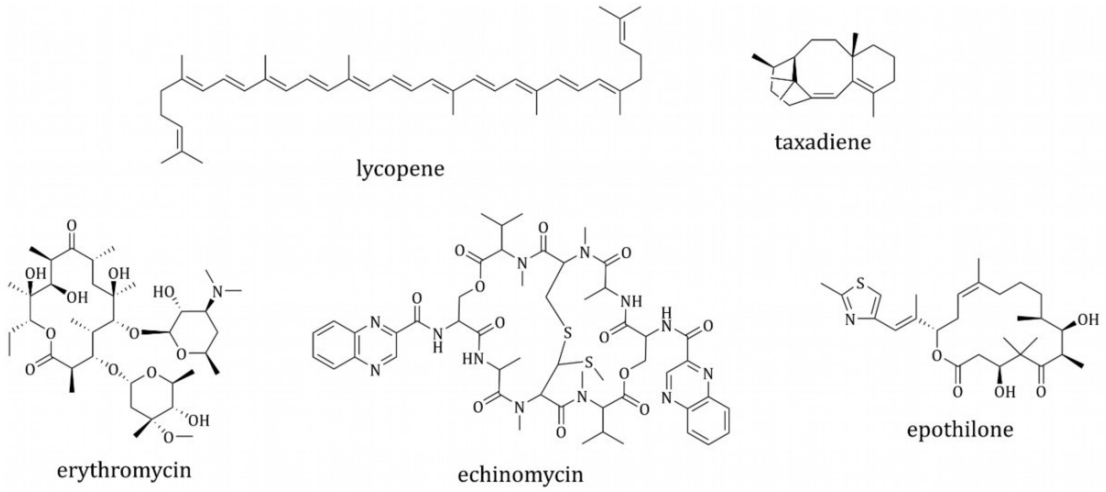


Figure 6. Examples of secondary metabolites heterologously produced in *E. coli*.

Cyanobacterial natural products, which have been heterologously produced in *E. coli* include the NRPS/PKS-derived hepatotoxin microcystin and the alkaloid and sunscreen pigment scytonemin [23].

The reconstitution of microcystin biosynthesis [126] was performed by reassembling the 55 kb-sized *mcy* BGC from a *Microcystis aeruginosa* PCC 7806 fosmid library, harnessing Red/ET recombineering. This included the use of the native microcystin promoter to direct the biosynthesis in the non-native host. For sufficient posttranslational activation of the assembly line, *E. coli* was further equipped with a chromosomal copy of the promiscuous PPTase MtaA from *Stigmatella aurantiaca* DW4/3-1. However, initial heterologous expression attempts did not result in any detectable microcystin production in *E. coli*. Analyses of the native microcystin promoter and mRNA profiling of the heterologous host and *M. aeruginosa*, revealed a lack of gene expression in *E. coli*. This was reasoned with an innate nitrogen-responsive regulation of

the native microcystin promoter. In *M. aeruginosa*, the nitrogen level-dependent transcription factor NtcA is directing the activation of the *mcy* promoter. Unlike the native producer, *E. coli* is not harboring an NtcA homologue, impeding a transcription of the heterologous *mcy* cluster. An exchange of the native *mcy* promoter with the bidirectional tetracycline-inducible promoter PtetO solved the innate transcription factor dependency. Together with an optimization of the cultivation conditions, this culminated in the successful biosynthesis of up to 162±23 µg/l microcystin. Due to the fast growth rate of *E. coli*, the achieved microcystin space-time yields (1.3–32.5 µg/l/day) were even superior to the native producer *M. aeruginosa* (5.7–16.7 µg/l/day).

The reassembly of scytonemin biosynthesis [93] (cf. section 1.2.2 and Figure 5) in *E. coli* was performed by assembly of all putative pathway genes (*scyA-F*, *tyrA*, *tyrP*, *dsbA*, *Npr1270*, *aroB*, *aroG*, *trpA-E*, *Npr1259*) from *N. punctiforme* ATCC 29133 in three plasmids. Although scytonemin formation could not be observed in the subsequent expression study, the biosynthesis of the monomeric ketone precursor and various pathway shunt products could be demonstrated. After comparative genomic analyses suggested that the genes *scyA-F* are sufficient for scytonemin biosynthesis, minimal expression vectors were tested. According to these analyses the production of the monomeric precursor can already be achieved with the three genes *scyA-C*. Supplementation of the recombinant production cultures with tyrosine and tryptophan increased the titer of the monomeric precursor up to 8.9 mg/l. However, the final dimerization to scytonemin could not be realized in *E. coli*, which was reasoned with either inactivity of the responsible, so far unidentified dimerization enzymes, or with the lack of genes encoding such enzymes in the assembled locus.

1.3.2 *Saccharomyces cerevisiae* as heterologous host

The following section contains passages from the author's previous publication [II] [13].

"The single-celled ascomycete *S. cerevisiae* is one of the most prominent microbial workhorses in academia and industry. As a robust, fast growing and safe organism, encoding no toxic or viral genes, budding yeast is of particular interest for biotechnological applications. The ease of transformation with exogenous DNA in conjunction with extremely efficient homologous recombination capabilities make it a primary choice for the recombinant production of pharmaceutical drugs and other high-value chemicals. Over the years, a number of techniques have been developed for the genetic engineering of *S. cerevisiae*, which exploit homologous recombination and are used for genome editing and pathway reassembly. Examples include transformation-associated recombination cloning [148–152], long terminal repeat-guided cloning (LTR-recombination) [153,154] and CRISPR/Cas9 [148,155–157]. This arsenal has been expanded by many

plug-and-play tools, which facilitate the assembly and expression of large DNA fragments [152,154,156,158–166]. Other scientific breakthroughs include the development of bidirectional expression plasmids [167,168] and synthetic minimal expression systems for *S. cerevisiae* [169–171], which make pathway refactoring in this model organism feasible.

The use of *S. cerevisiae* as a heterologous host was pioneered in the 1980s when strains were constructed for the manufacturing of pharmaceutical proteins, like interferon-α [172,173] and insulin [173,174]. At this time, the exploration and comprehension of yeast's fundamental secretory expression pathways paved the way to a successful maturation and secretion of recombinant proteins. Of particular relevance in this context is the leader sequence of the mating type peptide pheromone α-factor in *S. cerevisiae*, which was found to convey secretory competence to a multitude of heterologously expressed fusion proteins [172,174,175]."

Apart from human polypeptides, yeast was also harnessed for the expression of viral proteins in the field of recombinant vaccine biosynthesis. One prominent example is the hepatitis B surface antigen [176]. Since the beginning of the 21st century yeast has also been "explored as a production platform for small molecules, including chemical feedstocks, biofuels, food additives, flavors and cosmetics [177]. Due to its innate metabolism, yeast produces several intermediates of commercial value, such as ethanol and glycerol."

"Its long established industrial use as well as the available *omics* data facilitate rational metabolic engineering strategies on the basis of mathematical models [177–184]. The implementation of artificial pathways is also feasible and allows the production of molecules that do not naturally occur in yeast. An illustrative example is the biosynthesis of enantiopure lactic acid, which serves as raw material for the production of polylactide polymers [185]. Significant efforts were invested to establish competitive production rates in recombinant *S. cerevisiae* strains and to develop a manufacturing process at a commercial scale [186,187]."

In recent years, the reconstitution and reconstruction of pathways to secondary metabolites (Figure 7) underwent significant progress [105,188,189]. Examples include terpenoids, such as the flavor and insecticide limonene [190], the anti-inflammatory hormone hydrocortisone [191,192] and artemisinic acid, a precursor of the antimalarial drug artemisinin [193,194]. The latter drug was even produced at industrial scale [195]. Recombinantly produced alkaloids include the analgesic opioids thebaine and hydrocodone [196], or the antitussive noscapine [197], while phenylpropanoids encompass the antioxidant resveratrol [198] and naringenin [199]. Recent advances also focus on heterologous biosynthesis approaches of cannabinoids, using *S. cerevisiae* as a potential platform organisms for the production of e.g. tetrahydrocannabidiol [200–202]. Polyketides and nonribosomal peptides that were heterologously produced in yeast comprise predominantly compounds of fungal origin, such as 6-methylsalicylic acid [203–207], penicillin [165,180], or the depsipeptides

beauvericin and bassianolide [208,209]. In contrast, the reconstruction of PKSs and NRPSs from bacteria has been barely investigated [13]. Among the few exceptions are the unusual, C-domain-lacking indigoidine NRPS from *Streptomyces lavendulae* [210] and the NRPS-like, D-alanine-D-alanine ligase, which is involved in the biosynthesis of the sunscreen pigment shinorine in *N. punctiforme* [127,211].

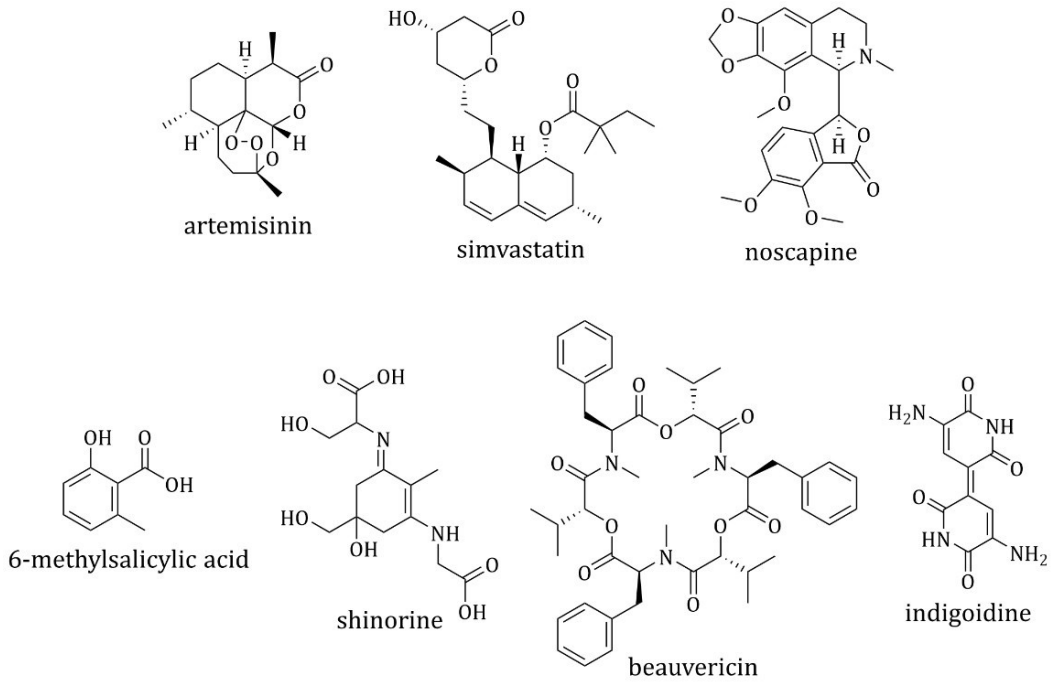


Figure 7. Examples of secondary metabolites heterologously produced with *S. cerevisiae*.

Combinatorial biosynthesis was also explored in *S. cerevisiae*, as exemplified by the production of triketide lactone (TKL) [212] and an artificial dipeptide [213]. The TKL polyketide was made with a fusion protein integrating the deoxyerythronolide B megasynthase (DEBS) module 2 and the DEBS-TE domain. The artificial dipeptide was produced after modules of two bacterial NRPSs, namely the tyrocidine synthetase (TycA) and the surfactin (SrfAC) synthetase, were expressed in yeast. Both NRPSs had been previously endowed with compatible communication-mediating (COM) domains to mediate the required protein-protein interactions [13].

1.3.2.1 Novel expression elements for *S. cerevisiae*

The heterologous expression of bacterial BGCs in *S. cerevisiae* is challenging, due to the incompatibility of polycistronic transcription units. In order to express bacterial BGCs in *S. cerevisiae*, it is obligatory to refactor each individual gene and to equip it with a suitable yeast promoter and terminator sequence [170].

For years, only a limited set of native yeast promoters and terminators was available for this task. Examples include the galactose inducible promoter Gal1, the strong constitutive TDH3 promoter, or the endogenous CYC1 terminator. Due to the innate complexity of eukaryotic expression systems and a large quantity of regulating *cis*-elements, e.g. the 5' upstream activating sequence (UAS) elements, transcription factor binding sites, or positioning elements and polyadenylation signals of the 3'-untranslated region, these promoter and terminator sequences usually span hundreds of base pairs up to 1 kb. Therefore, the reconstruction of large bacterial BGCs in yeast is hampered by the need to include the additional extensive cargo of these obligatory regulating DNA elements [169,170].

Recently, minimal synthetic promoters and terminators were developed for *S. cerevisiae*, which facilitate pathway refactoring [169,170]. The minimal synthetic promoters (Figure 8) take a space of less than 120 bp in total and are composed of a short hybrid assembled upstream activating sequence (UAS), a small AT-rich spacer and a core element. The core element is including the TATA-box, a 30 bp core spacer and the transcription start site (TSS), respectively. In total, six UAS elements (e.g. UAS A-F) and nine compatible generic core elements were designed, which are applicable to generate a set of constitutive promoters for *S. cerevisiae* [170].

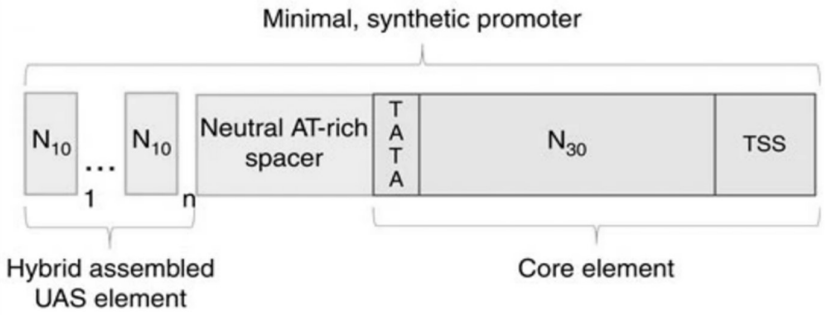


Figure 8. Architecture of minimal synthetic promoters for *S. cerevisiae*. The corresponding DNA is composed of an UAS element, a neutral AT-rich spacer and a core element. Notation: N_{10}, 10 nucleotide sequence; n, n-fold, indicating any natural number; UAS, upstream activating sequence; TATA, TATA-box; TSS, transcription start site. (Reprinted adapted with permission from Redden et. *al.* [170], Copyright (2015) Nature Communications, Creative Commons Attribution 4.0 International license. (CC BY 4.0)).

The minimal synthetic terminators (Figure 9) exhibit a size of only 35–70 bp. They comprise three key elements, i.e. an efficiency element, a positioning element and a polyA site, which are interspaced by short linker regions. Optional features include upstream and downstream poly(T)tracts, a variable linker of 3-20 bp and a 3'-located poly(T)tract. In total, 30 terminators are currently available, which can be used for gene expression in yeast [169].

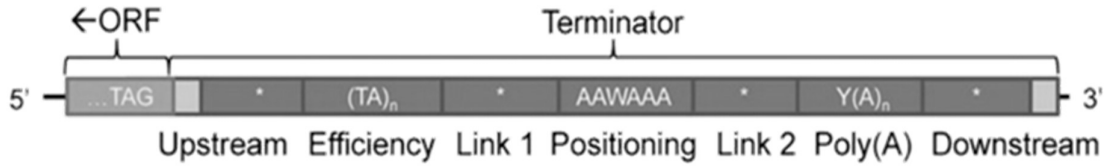


Figure 9. Architecture of minimal synthetic terminators for *S. cerevisiae*. The corresponding DNA is composed of an efficiency element, a positioning element and a Poly(A)-site. Notation: ORF, open reading frame including the TAG stop codon; Upstream, optional 5' poly(T) tract; Efficiency, efficiency element; Link1, 3-20 bp spacer sequence; Positioning, positioning element; Link2, 3-20 bp spacer sequence; Poly(A), polyA-site; Downstream, optional 3' poly(T) tract. (Reprinted adapted with permission from Curran *et al.* [169], Copyright 2015 American Chemical Society)

2 Scope of the thesis

Cyanobacteria are versatile producers of bioactive secondary metabolites and are therefore a prolific resource for drug discovery. The remarkable progress in genome sequencing and bioinformatics has uncovered the enormous biosynthetic potential of this phylum (cf. section 1.2). However, this potential frequently remains unexploited, which is due to cultivation and production constraints (cf. section 1.3 and 1.2). The phototrophic cyanobacteria often grow slowly with doubling times of several days, or they are non-culturable at all. Furthermore, many BGCs remain silent under laboratory conditions or are only weakly expressed, which results in low production titers. The overexpression of BGCs is also rarely feasible, as many strains are not amenable to genetic engineering. Overall, this impedes the exploitation of their great biosynthetic potential. The use of heterologous hosts in order to probe or manipulate cyanobacterial BGCs was already successful (cf. section 1.3). However, a heterologous platform organism, which could be applied to produce different cyanobacterial secondary metabolites is yet to be found.

The major aim of this thesis was to evaluate *S. cerevisiae* as heterologous host for the biosynthesis of selected cyanobacterial compounds. Another goal was to expand the biosynthetic and genomic knowledge of the cyanobacterial strain *Nostoc* sp. ATCC 53789.

(i) *S. cerevisiae* is a proven host organism for the heterologous production of secondary metabolites (cf. section 1.3). In this thesis, the potential of this microorganism for the reconstruction of the pathways to the cyanobacterial compounds cryptophycin and scytonemin was explored. For this purpose, a suitable method to transfer large bacterial BGCs into *S. cerevisiae* had to be developed and a viable expression system in this host had to be established. The respective results and challenges are described and analyzed in chapter 4.2 and 4.3 of this thesis.

(ii) The challenges associated with the reconstruction approach of cryptophycin biosynthesis were further investigated. In particular, questions regarding the identity and the processing of the starter unit had to be addressed. To this end, the biosynthesis enzyme CrpA was expressed in *E. coli*. The corresponding study is covered in section 4.4 of this thesis.

The biosynthetic investigations were complemented by the sequencing and annotation of the genome of the cryptophycin-producing *Nostoc* strain ATCC 53789. The results of this study are described in chapter 4.1 of this thesis.

Last but not least, an *in vitro* enzymatic bioconversion route was devised to generate putative intermediate derivatives of the cryptophycin biosynthesis pathway, which in perspective could help to clarify the biosynthetic model. The results are displayed and discussed in section 4.5 of this thesis.

3 Material and Methods

In the following section, the experimental design and used strains are described in detail, covering the applied general methodologies of strain cultivation, extraction, general molecular biological operations, as well as specific methods and information regarding the different research projects.

3.1 Strains, cultivation and general methodologies

The herein performed research projects required the cultivation of diverse microorganisms, including *Nostoc* spp., as well as multiple *E. coli* and *S. cerevisiae* strains. The respective cultivation methodologies are described in the following.

3.1.1 *Nostoc* spp.

Nostoc sp. ATCC 53789 was obtained from the American Type Culture Collection (ATCC) and cultivated in cell culture flasks (stored horizontally) for 1 month under diurnal illumination at room temperature (RT), using BG-13 medium [214]. In order to increase the titer of cryptophycin 1 for the development of a cancer cell line assay in cooperation with the LDC (Lead Discovery Center, Dortmund), *Nostoc* sp. ATCC 53789 was continuously cultivated for 6 month to scale-up the total culture volume. Likewise, *N. punctiforme* ATCC 29133 (PCC 73102) was obtained from ATCC and cultivated for two weeks, using ATCC medium 819. This medium corresponds to ATCC Medium 616 (BG-11) without $NaNO_3$. For both strains, initially 10 ml of medium were inoculated with the obtained culture. Scale-up was performed up to a volume of 200 ml medium. For this purpose, the total cell culture was harvested by centrifugation (4000 rpm, 15 min, RT) and dissolved in fresh medium by vigorous shaking. Volume scale-up was performed stepwise from 10 ml, via 50 ml, 100 ml up to 200 ml culture volume.

3.1.2 *E. coli*

If not specified otherwise, *E. coli* cells were cultivated in Erlenmeyer flasks using LB medium (Table 1) at 37 °C and 180 rpm. For selection, antibiotics were added to the medium at the following concentrations: 100 µg/ml ampicillin (Amp), 50 µg/ml kanamycin (Kan), or 25 µg/ml chloramphenicol (Cm). Regeneration of transformants was conducted in SOB medium (Table 1). The used *E. coli* plasmids and strains are listed in Table 2 and Table 3, respectively.

Table 1. *E. coli* medium composition.

Medium	Components	Amount
LB (pH 7)	Tryptone	10 g/l
	Yeast extract	5 g/l
	Sodium chloride	10 g/l
LB-Agar	Bacteriological Agar	15 g/l
SOB (pH 7)	Tryptone	20 g/l
	Yeast extract	5 g/l
	Sodium chloride	0.6 g/l
	Potassium chloride (25 mM)	1% (v/v)

Table 2. List of *E. coli* plasmids used in this thesis. All pMA-RQ and pYPKpw vectors harbor a unique, minimal, artificial *S. cerevisiae* expression system based on the work of Redden *et al.* [170] and Curran *et al.* [169]. Designated homologous locus flanks for subsequent recombination events were designed according to the work of Reider Apel *et al.* [156]. Additional information and vector maps regarding the applied plasmids are provided in the supplementary information (Table SI I). Notation: His-tag, hexahistidine tag; CmR, chloramphenicol resistance; MBP-tag, N-terminal maltose binding protein (solubility tag); Amp^R, ampicillin resistance; Kan^R/G418, kanamycin/G418 resistance; SUMO, N-terminal small ubiquitin-related modifier (solubility tag); sfp, Sfp-type phosphopantetheinyl transferase *B. subtilis*; sfpN, Sfp-type phosphopantetheinyl transferase *Nostoc* sp. ATCC 53789; Zeo^R, zeocin resistance; Hyg^R, hygromycin resistance.

Plasmid	Origin	Relevant features	Reference
MPBA biosynthesis project (*in vivo*)			
pACYCDuet	Novagen	Expression plasmid; His-tag; T7 promoter; Cm^R	
pMalc2x	New England Biolabs	Expression plasmid; MBP-tag; tac promoter; Amp^R	Walker et *al.* [215]
pET28a (+)	Novagen	Expression plasmid; His-tag; T7 promoter; Kan^R	
pET28a-tmmrMcGAS-SUMO	Regine Siedentop, Laboratory of Bioprocess Engineering, TU Dortmund University	SUMO-tag	Rolf et *al.* [216]
pET28a-crpA	pET28a (+)	Expression plasmid; His-tagged CrpA	This study
pMal-crpA	pMalc2x	Expression plasmid; MBP-tagged CrpA	This study
pACYC-sfp	pACYCDuet	*sfp* expression plasmid	This study
pACYC-sfpN	Marvin D. Schwing, Laboratory of Technical Biology TU Dortmund University	*sfpN* expression plasmid	Schwing [a]

pET28a-SUMO-crpA	pET28a-crpA, pET28a-tmmrMcGAS-SUMO	Expression plasmid; SUMO-tagged CrpA;	This study
pMal-crpA-eryAIII$_{TE}$	pMalc2x	Expression plasmid, MBP-tagged CrpA fused to the TE domain of EryAIII	This study
PBA biosynthesis project (*in vitro*)			
pET28a-admH	pET28a (+)	Expression plasmid; His-tagged AdmH	This study, Jin et *al.* [217]
Cryptophycin biosynthesis project			
pYTK080	Addgene	Zeo^R cassette	Lee et *al.* [158]
pYTK079	Addgene	Hyg^R cassette	Lee et *al.* [158]
pYTK077	Addgene	Kan^R/G418 cassette	Lee et *al.* [158]
pMA-RQ-PB2-T27	Invitrogen	Subcloning vector, minimal expression system; constitutive promoter PB2; terminator T27; target locus 308a (ARS308a) chromosome III;	This study
pMA-RQ-PB2-crpA-T27	pMA-RQ-PB2-T27	Subcloning vector; *crpA;* PB2-T27 expression system	This study
pMA-RQ-PA3-T30	Invitrogen	Subcloning vector, minimal expression system; constitutive promoter PA3; terminator T30	This study
pMA-RQ-PA3-crpC-T30	pMA-RQ-PA3-T30	Subcloning vector, *crpC*; PA3-T30 expression system	This study
pYPKpw-PC8-crpB-T3	ATG:biosynthetics GmbH	Synthetic vector; *crpB*; minimal expression system; constitutive promoter PC8; terminator T3	This study
pYPKpw-PD1-crpD-T8	ATG:biosynthetics GmbH	Synthetic vector; *crpD*; minimal expression system; constitutive promoter PD1; terminator T8	This study
pMA-RQ-PC8-T3	Invitrogen	Subcloning vector, minimal expression system; constitutive promoter PC8; terminator T3; target locus 416d (ARS416/ARS1) chromosome IV	This study
pMA-RQ-PC8-crpH-T3	pMA-RQ-PC8-T3	Subcloning vector; *crpH*; PC8-T3 expression system	This study
Scytonemin biosynthesis project			
pMA-RQ-PC8-T3	Invitrogen	Equivalent to cryptophycin biosynthesis project	This study
pMA-RQ-PA3-T30	Invitrogen	Equivalent to cryptophycin biosynthesis project	This study

pMA-RQ-PC8-scyA-T3	pMA-RQ-PC8-T3	Subcloning vector; *scyA*; PC8-T3 expression system	This study
pMA-RQ-PA3-scyC-T30	pMA-RQ-PA3-T30	Subcloning vector; *scyC*; PA3-T30 expression system	This study
pMA-RQ-PA3-scyB-T30	pMA-RQ-PA3-T30	Subcloning vector; *scyB*; PA3-T30 expression system	This study
pYTK079	Addgene	Hyg^R cassette	Lee et *al.* [158]
pYTK077	Addgene	Kan^R/G418 cassette	Lee et *al.* [158]

Table 3. *E. coli* strains used in this thesis. Notation: MBP-tag, N-terminal maltose binding protein (solubility tag); Sfp, Sfp-type phosphopantetheinyl transferase *Bacillus subtilis*; SUMO-tag, N-terminal small ubiquitin-related modifier (solubility tag); BGC, biosynthetic gene cluster; SfpN; Sfp-type phosphopantetheinyl transferase *Nostoc* sp. ATCC 53789.

Strain	Origin	Genotype	Note
***E. coli* DH5α**	ThermoFisher Scientific	F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG	subcloning host; expression system construction and preparation
***E. coli* TOP10**	Invitrogen	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG	subcloning host; expression system construction and preparation
***E. coli* BL21(DE3)**	ThermoFisher Scientific	F– ompT hsdSB (rB–, mB–) gal dcm (DE3)	B-strain expression host; IPTG inducible T7 system; high recombinant protein expression level; control strain in MPBA biosynthesis project
***E. coli* KRX**	Promega	[F´, traD36, ΔompP, proA+B+, lacIq, Δ(lacZ)M15] ΔompT, endA1, recA1, gyrA96 (Nalr), thi-1, hsdR17 (rk–, mk+), e14– (McrA–), relA1, supE44, Δ(lac-proAB), Δ(rhaBAD)::T7 RNA polymerase	K-strain expression host; rhamnose inducible T7 system; toxic protein expression; low pre-induced expression level [218]
***E. coli* ArcticExpress (DE3)**	Agilent	F– ompT hsdS(rB–mB–) dcm+ Tetr gal λ(DE3) endA Hte [cpn10 cpn60 Gentr]	BL21 GOLD expression host; 4-12 °C cultivation; chaperonin expression (*Oleispira antarctica*); IPTG inducible T7 system [219]

***E. coli* C41(DE3)**	Sigma Aldrich	F – ompT hsdSB (rB- mB-) gal dcm (DE3)	BL21(DE3) expression host; IPTG inducible T7 system; toxic protein expression [220]
***E. coli* C43(DE3)**	Sigma Aldrich	F – ompT hsdSB (rB- mB-) gal dcm (DE3)	C41(DE3) expression host; IPTG inducible T7 system; toxic protein expression [220],
Cryptophycin biosynthesis project			
Strain	**Origin**	**Plasmids**	**Note**
E. coli* DH5α-*crpA	*E. coli* DH5α	pMA-RQ-PB2-crpA-T27	Subcloning strain; *crpA*; PB2-T27
E. coli* DH5α-*crpC	*E. coli* DH5α	pMA-RQ-PA3-crpC-T30	Subcloning strain; *crpC*; PA3-T30
E. coli* DH5α-*crpH	*E. coli* DH5α	pMA-RQ-PC8-crpH-T3	Subcloning strain *crpH*; PC8-T3
MPBA biosynthesis project (*in vivo*)			
Strain	**Origin**	**Plasmids**	**Note**
E. coli* DH5α-MBP-*crpA	*E. coli* DH5α	pMal-crpA	MBP-tag; *crpA*; control strain
E. coli* BL21 (DE3)-*crpA-sfp	*E. coli* BL21 (DE3)	pET28a-crpA, pACYC-sfp	*crpA*; Sfp
E. coli* ArcticExpress (DE3)-MBP-*crpA-sfp	*E. coli* ArcticExpress (DE3)	pMal-crpA, pACYC-sfp	MBP-tag; *crpA*; *sfp*
E. coli* ArcticExpress (DE3)-SUMO-*crpA-sfp	*E. coli* ArcticExpress (DE3)	pET28a-SUMO-crpA, pACYC-sfp	SUMO-tag; *crpA*; *sfp*
E. coli* C43(DE3)-MBP-*crpA-sfp	*E. coli* C43(DE3)	pMal-crpA, pACYC-sfp	MBP-tag; *crpA*; *sfp*
E. coli* C43(DE3)-SUMO-*crpA-sfp	*E. coli* C43(DE3)	pET28a-SUMO-crpA, pACYC-sfp	SUMO-tag; *crpA*; *sfp*
E. coli* C41(DE3)-MBP-*crpA-sfp	*E. coli* C41(DE3)	pMal-crpA, pACYC-sfp	MBP-tag; *crpA*; *sfp*
E. coli* C41(DE3)-SUMO-*crpA-sfp	*E. coli* C41(DE3)	pET28a-SUMO-crpA, pACYC-sfp	SUMO-tag; *crpA*; *sfp*
E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp	*E. coli* BL21 (DE3)	pMal-crpA-eryAIII$_{TE}$, pACYC-sfp	MBP-tag; *crpA* fused to *eryAIII*$_{TE}$; *sfp*
E. coli* BL21 (DE3)-MBP-*crpA-sfp	*E. coli* BL21 (DE3)	pMal-crpA, pACYC-sfpN	MBP-tag; *crpA*; *sfpN* [a]
E. coli* BL21 (DE3)-MBP-*crpA-sfp	*E. coli* BL21 (DE3)	pMal-crpA, pACYC-sfp	MBP-tag; *crpA*; *sfp*
E. coli* BL21 (DE3)-SUMO-*crpA-sfp	*E. coli* BL21 (DE3)	pET28a-SUMO-crpA, pACYC-sfp	SUMO-tag; *crpA*; *sfp*

E. coli* BL21 (DE3)-*sfp	*E. coli* BL21 (DE3)	pACYC-sfp	*sfp*; control strain
E. coli* BL21 (DE3)-MBP-*crpA	*E. coli* BL21 (DE3)	pMal-crpA	MBP-tag; *crpA*; control strain
***E. coli* BL21 (DE3)-*sfp*-C2x**	*E. coli* BL21 (DE3)	pACYC-sfp, pMalc2x	*sfp*; empty pMalc2x; control strain
E. coli* KRX-MBP-*crpA-sfp	*E. coli* KRX	pMal-crpA, pACYC-sfp	MBP-tag; *crpA*; *sfp*
E. coli* KRX-SUMO-*crpA-sfp	*E. coli* KRX	pET28a-SUMO-crpA, pACYC-sfp	SUMO-tag; *crpA*; *sfp*
PBA biosynthesis project (*in vitro*)			
Strain	**Origin**	**Plasmids**	**Note**
E. coli* BL21 (DE3)-*admH	*E. coli* BL21 (DE3)	pET28a-admH	*admH*
Scytonemin biosynthesis project			
Strain	**Origin**	**Plasmids**	**Note**
***E. coli* DH5α**	*E. coli* DH5α	pMA-RQ-PC8-scyA-T3	Subcloning strain; *scyA*; PC8- T3
***E. coli* DH5α**	*E. coli* DH5α	pMA-RQ-PA3-scyC-T30	Subcloning strain; *scyC*; PA3-T30
***E. coli* TOP10**	*E. coli* TOP10	pMA-RQ-PA3-scyB-T30	Subcloning strain; *scyB*; PA3-T30

3.1.3 *S. cerevisiae*

If not specified otherwise, *S. cerevisiae* was cultivated in baffled Erlenmeyer flasks at 30 °C and 180-200 rpm, using YPD medium or Yeast Nitrogen Base (YNB) drop-out medium with selective amino acid and nutrient supplementation (Sigma-Aldrich)[221] (Table 4). For the selection of plasmids or positive transformants in homologous recombination events, depletion of medium supplements (uracil depletion) or antibiotic selection was performed, using hygromycin (200 µg/ml), zeocin (150 µg/ml), or G418 (210 µg/ml), respectively.

In case of cryptophycin biosynthesis, precursor uptake might be limited due to catabolite-repression and thus restricted to simple diffusion [222]. Hence, during expression cultivation, the cells were grown in acetate-buffered YPD (pH 4) to assure a sufficient uptake of the putative CrpA-precursor 3-phenylpropionic acid (3-PPA).

The used *S. cerevisiae* plasmids and strains are listed in Table 5 and Table 6, respectively.

Table 4. *S. cerevisiae* medium composition.

Medium	Components	Amount
YPD (pH 6-7)	Bacto Yeast Extract	10 g/l
	Bacto Peptone	20 g/l
	Glucose (50% w/v)	40 ml/l
YPD-Agar	Bacteriological Agar	20 g/l
YNB (Sigma-Aldrich)	Yeast Nitrogen Base (without supplements)	6.7 g/l
	Glucose (50% w/v)	40 ml/l
	Supplement Mix [221]	1.39 - 1.92 g/l Defined according to selection
YNB-Agar (Sigma-Aldrich)	YNB + Bacteriological Agar	20 g/l

Table 5. List of *S. cerevisiae* plasmids used in this thesis. Additional information and vector maps regarding the applied plasmids are provided in the supplementary information (Table SI II).

Plasmid	Features
Scytonemin biosynthesis project	
pCAS9-416 [156]	*S. cerevisiae-E. coli* shuttle plasmid; Cas9; gRNA for CRISPR/CAS genome editing; target locus ARS416/ARS1 chromosome IV
Cryptophycin biosynthesis project	
pCAS9-308 [156]	*S. cerevisiae-E. coli* shuttle plasmid; Cas9; gRNA for CRISPR/CAS genome editing; target locus ARS308a chromosome III

Table 6. *S. cerevisiae* strains used in this thesis. Notation: Sfp, Sfp-type phosphopantetheinyl transferase of *B. subtilis*; Kan^R/G418, kanamycin resistance; Zeo^R, zeocin resistance; Hyg^R, hygromycin resistance; *bna2*, *S. cerevisiae* gene encoding tryptophan 2,3-dioxygenase in kynurenine /anthranilate biosynthesis pathway.

Strain	Origin	Genotype	Note
Cryptophycin biosynthesis project			
***S. cerevisiae* CEN.PK2-1C**	Oliver Schiwy, Laboratory of Technical Biochemistry, TU Dortmund University	Δpep4 Δgal1 Δgal80 MATa gal1::loxP pep4::loxP gal80::loxP ura3-52 trp1-289 leu2-3,112 his3Δ1 MAL2-8C SUC2	Parental strain
S. cerevisiae-sfp	Oliver Schiwy, see *S. cerevisiae* CEN.PK2-1C	see *S. cerevisiae* CEN.PK2-1C ΔARS 720::Gal1-sfp-Cyc1	chromosomal Sfp; galactose-inducible promoter Gal1; terminator Cyc1

S. cerevisiae-sfp-crpA : pCAS9-308	*S. cerevisiae -sfp*	Cas9 plasmid pCAS9-308, Δ308::PB2-crpA-T27	Sfp; constitutive promoter PB2; terminator T27; crpA; Cas9 plasmid targeting locus 308a chromosome III
S. cerevisiae-sfp-crpA	*S. cerevisiae -sfp*	Δ308::PB2-crpA-T27	Sfp; PB2; T27; *crpA*
S. cerevisiae-sfp-crpA-crpB	*S. cerevisiae -sfp-crpA*	ΔYPRCΔ15::PC8-crpB-T3-KanR	Sfp; constitutive promoter PC8; terminator T3; Kan^R/G418; *crpA*; *crpB*
S. cerevisiae-sfp-crpA-crpB-crpD	*S. cerevisiae -sfp-crpA-crpB*	ΔYPRCτ3::PD1-crpD-T8-ZeoR	Sfp; constitutive promoter PD1; terminator T8; Kan^R/G418; ZeoR; *crpA*; *crpB*; *crpD*
***S. cerevisiae-sfp-crpA-crpB-crpC-crpD* (*S. cerevisiae* ATi01)**	*S. cerevisiae -sfp-crpA-crpB-crpD*	ΔYORWΔ22::PA3-crpC-T30-HygR	Sfp; constitutive PA3; terminator T30; Kan^R/G418; ZeoR; Hyg^R; *crpA*; *crpB*; *crpD*; *crpC*
***S. cerevisiae-sfp-crpA-crpB-crpC-crpD-crpH* : pCAS9-416**	*S. cerevisiae -sfp-crpA-crpB-crpC-crpD*	Cas9 plasmid pCAS9-416, Δ416::PC8-crpH-T3	Sfp; PC8; T3; Kan^R/G418; ZeoR; Hyg^R; *crpA*; *crpB*; *crpD*; *crpC*; *crpH*; Cas9 plasmid targeting locus 416d chromosome IV
***S. cerevisiae-sfp-crpA-crpB-crpC-crpD-crpH* (*S. cerevisiae* ATi02)**	*S. cerevisiae -sfp-crpA-crpB-crpC-crpD* (*S. cerevisiae* ATi01)	Δ416::PC8-crpH-T3	Sfp; PC8; T3; Kan^R/G418; ZeoR; Hyg^R; *crpA*; *crpB*; *crpD*; *crpC*; *crpH*
Scytonemin biosynthesis project			
Strain	**Origin**	**Genotype**	**Note**
***S. cerevisiae-scyA* : pCAS9-416**	*S. cerevisiae* CEN.PK2-1C	Cas9 plasmid pCAS9-416, Δ416::PC8-scyA-T3	Constitutive promoter PC8; terminator T3; *scyA*; Cas9 plasmid targeting locus 416d chromosome IV
S. cerevisiae-scyA	*S. cerevisiae* CEN.PK2-1C	Δ416::PC8-scyA-T3	PC8; T3; *scyA*
S. cerevisiae-scyA-scyC	*S. cerevisiae -scyA*	ΔYORWΔ22::PA3-scyC-T30-HygR	Constitutive promoter PA3; terminator T30; Hyg^R; *scyA*; *scyC*
S. cerevisiae-scyA-scyB-scyC	*S. cerevisiae -scyA-scyC*	Δbna2::PA3-scyB-T30-KanR	PA3; T30; Hyg^R; Kan^R/G418; *scyA*; *scyC*; *scyB*; *bna2*/ locus YJR078W knock-down

3.2 General molecular biological operations

The following sections describe molecular biological operations that were used in this PhD project. This includes cloning methods and procedures used for DNA isolation, *de-novo* sequencing, as well as bioinformatics analyses.

3.2.1 Isolation of genomic and plasmid DNA

Genomic DNA of *S. cerevisiae* was isolated using the extraction kit "Genomic DNA from microorganisms" (Macherey-Nagel) for yeast.

Preparation of genomic DNA of *N. punctiforme* and *Nostoc* sp. ATCC 53789 for PCR applications was performed likewise, applying the manufacturer's instructions for DNA extraction from yeast. For sequencing purposes genomic DNA of *Nostoc* sp. ATCC 53789 was isolated manually by phenol-chloroform extraction, following the instructions of Cohen *et al.* with minor modifications [223,224]. In brief, *Nostoc* sp. ATCC 53789 was cultivated as described previously (cf. section 3.1.1). In total 10 ml of the cultivated cells were harvested by centrifugation (4000 rpm, 10 min) and the isolation procedure was conducted accordingly, without the addition of CTAB (cetyltrimethylammoniumbromid). Cell lysis was performed at 65 °C for 10 min. After centrifugation the supernatant of the cell lysate was transferred to a new reaction tube and treated with two volumes of phenol/chloroform/isoamyl alcohol solution (Carl Roth) for phase separation. After centrifugation, the aqueous phase was transferred into a new reaction tube, followed by two consecutive phenol-chloroform extractions of the aqueous phase. Subsequently, the genomic DNA was precipitated by addition of one volume of chilled isopropanol and washed three times with 70% ethanol. Finally, the extracted genomic DNA was dried and resuspended in ultrapure double deionized water (ddH_2O).

Plasmid isolation from *E. coli* strains was performed using the "NucleoSpin Plasmid (NoLid), Mini kit for plasmid DNA purification" (Macherey-Nagel), following the manufacturer's instructions.

3.2.2 Gel-electrophoretic analyses and purification of DNA

Analysis of DNA fragments was performed by agarose gel electrophoretic separation in TAE buffer (pH 8.3 – 8.6), which is composed of 0.04% (w/v) EDTA, 0.1% (v/v) glacial acetic acid and 0.5% (w/v) Tris. Electrophoretic separation of plasmid DNA, PCR products and restriction fragments was conducted in 0.1% (w/v) agarose gels at 120 V, while 80 V and 0.8% gels were applied for genomic DNA separation. If required, subsequent gel extraction and purification of the separated DNA fragments (e.g. PCR products, restriction fragments) was performed using the "QIAquick Gel Extraction Kit" (Qiagen), following the manufacturer's instructions. The general quality of the extracted DNA/plasmids was assessed by spectral photometric absorption using a NanoDrop photometer (NanoDrop 2000, PeqLab).

3.2.3 Restriction of DNA

Restriction of plasmid DNA and PCR products was conducted using FastDigest restriction enzymes (ThermoFisher Scientific) according to the manufacturer's instructions. If required, alkaline phosphatase (FastAP, ThermoFisher Scientific) was added to the restriction digest reaction of plasmid DNA for dephosphorylation, following the manufacturer's instructions for combined restriction-dephosphorylation applications.

3.2.4 Ligation and vector assembly

Ligation or assembly of DNA fragments and plasmids was performed either by application of a T4 DNA ligase (ThermoFisher Scientific) or by Gibson assembly [225] (NEB), following the manufacturer's instructions. T4-based ligation was carried out with DNA fragments featuring either blunt or sticky ends that had been generated by restriction digest, if necessary. In deviation from the manual's instructions, 0.1 volumes of PEG_{4000} were added to the reaction mixture. In case of Gibson assembly, linear DNA fragments with terminal sequence overlaps were joined. The required insert and vector fragments were generated by PCR, or by adequate linearization of the vector using restriction enzymes. The assembly temperature during the Gibson assembly is defined by the designed overlapping regions and has to be identical for all generated overlaps. It was generally adjusted to ~50-65 °C. The assembly reaction was subsequently used to transform an *E. coli* subcloning host (cf. section 3.3.1), in order to propagate the constructed vector for subsequent PCR, sequencing and restriction verification. Afterwards, the verified plasmids were used to transform the expression host.

3.2.5 Sequencing, primer generation and general software tools

Sanger sequencing was used to validate the constructed DNA fragments, PCR products and plasmids. The sequencing was conducted externally by Microsynth Seqlab. Samples were prepared according to the provider's instructions.

The general visualization, annotation and *in silico* design of plasmid maps, nucleotide, as well as amino acid sequences and alignments, was performed with SnapGene [226] and the assistance of BLAST (Basic Local Alignment Search Tool), which is provided by the NCBI (National Center for Biotechnology Information) service [227]. Primer sequences were designed manually, using the SnapGene software tool [226]. The designed sequences were further analyzed for self-annealing sites, secondary or dimer structures, applying the online platform OligoCalc [228]. All primers used in this study were synthesized by Sigma-Aldrich.

ChemDraw Professional [229] was used to visualize chemical structures and NRPS/PKS assembly lines. The software tool Inkscape [230] was used to generate high-resolution graphics.

3.2.6 Sequencing of *Nostoc* sp. ATCC 53789

The following section contains adapted extracts from the author's previous publications [I] [223].

For sequencing purposes, *Nostoc* sp. ATCC 53789 was cultivated in BG-13 for 1 month (cf. section 3.1.1) and the genomic DNA was isolated by phenol-chloroform extraction as described previously (cf. section 3.2.1). Whole genome *de-novo* sequencing of *Nostoc* sp. ATCC 53789 was performed in cooperation with C. Rückert (IIT Biotech GmbH, Bielefeld University), who reconstructed the genome from short- and long-read DNA data sets obtained by Illumina and Nanopore sequencing [223]. In detail, "library preparation involved a TruSeq DNA PCR-free high-throughput library prep kit (Illumina) and the SQK-LSK109 ligation sequencing kit (Oxford Nanopore Technologies [ONT]). Illumina sequencing was performed using a MiSeq reagent kit v3 (600 cycle) in a 2 x 300-nucleotide run. For Nanopore sequencing, a GridION platform with an R9.4.1 flow cell was used. Base calling and demultiplexing were performed using Guppy v3.1.5. Illumina data were assembled with Newbler, while Nanopore data were processed with Canu v1.8 [231]. Subsequently, the Canu contigs were polished with Racon v1.3.3 [232], medaka v0.11.0 [233] and Pilon v1.22 [234]. Minimap2 v2.17, BWA-MEM v2 [235] and Bowtie 2 v2.3.2 [236] were used for mapping. Unicycler v0.4.6 [237] was used for the subsequent hybrid assembly of the Illumina data and the contigs from the polished Canu assembly. The assemblies were combined manually in Consed v27.0 [238]. Next, the chromosome was reoriented based on the gene *dnaA*. Overlapping ends from the polished Canu assembly were trimmed by assembly in Consed. Ambiguities in these regions as well as all other repeat regions were corrected based on the contigs produced by the Newbler assembly. Finally, all differences between the contigs of the three assemblies, as well as low-quality regions marked in Consed, were resolved by manual curation using IGV v2.4.14 [239] for visualization of the ONT data."

The provided genomic data of the *de-novo* sequencing approach were investigated, using SnapGene [226] to visualize and analyze the genomic constitution of the strain. In order to assess the strain's biosynthetic potential and identify the location of the cryptophycin cluster, the obtained data were further analyzed with antiSMASH (version 5.0.0 and 6.0) [240]. The software tool DNAPlotter [241] was used to visualize the genome and to mark the BGCs identified by antiSMASH. In addition to the antiSMASH evaluation, the cryptophycin BGC and its genomic neighborhood was also manually inspected to verify the functional integrity of the encoded PKS and NRPS domains and to obtain further information on the biosynthetic assembly. These analyses were conducted with the help of SnapGene, BLASTx (BLAST performing a protein database enquiry with a translated nucleotide query), BLASTp (BLAST performing a protein database enquiry using a protein query) and BLASTn (BLAST performing a nucleotide database

enquiry using a nucleotide query). Furthermore, the work of Stachelhaus *et al.* [41], Minowa *et al.* [242], Yadav *et al.* [243] and Reeves *et al.* [31], as well as already published data of the cryptophycin assembly line by Magarvey *et al.* [64], were used to assess the domain integrity.

3.3 General molecular biological operations in *E. coli*

Specifications on the general molecular biological operations used for the herein conducted *E. coli*-based investigations are described in the following chapters. This covers experimental details on the preparation of competent cells, transformation procedure, protein expression, purification and analyses.

3.3.1 Generation of chemically competent cells and transformation of *E. coli*

Chemically competent *E. coli* cells, applicable to heat-shock transformation, were prepared by inoculating a 5 ml preculture in LB medium followed by overnight incubation. Subsequently, 10 ml of fresh LB medium were inoculated with the overnight culture to an OD_{600} of 0.1 and incubated up to an OD_{600} of 0.4. The culture, as well as solution A (14.7 g/l calcium chloride-dihydrate) and solution B (14.7 g/l calcium chloride-dihydrate, 189 g/l glycerol), were then chilled on ice for 30 minutes. Aliquots of 1 ml cell solution were transferred into an Eppendorf cup and harvested by centrifugation at 5600 rpm for 10 minutes at 4 °C. The pellet was resuspended in solution A and centrifuged, accordingly. Subsequently, the resulting pellet was resuspended in 100 µl solution B. The chemically competent cells were stored at -80 °C until further usage. Subsequent transformation with expression plasmids or sub-cloning vectors was performed following the NEB protocol for transformation of chemically competent *E. coli* cells [244]. Regeneration of the transformed cells was conducted in SOB medium (0,5% (w/v) yeast extract, 0,02% (w/v) KCl, 0,06% (w/v) NaCl, 2,0% (w/v) tryptone, pH 7,0) for 1.5 h at 37 °C and 180 rpm using an incubator. In case of transformation with *crpA*-containing plasmids, regeneration was performed at 30 °C for 2 h and 180 rpm, respectively. For selection, agar plates were supplemented with the corresponding antibiotics.

3.3.2 Production and purification of proteins

The pET28a- and pMal-derived expression plasmids (Table 2) were used to produce the heterologous proteins CrpA and AdmH for SDS-PAGE analysis, as well as for the biosynthesis of MPBA and *in vitro* studies (cf. section 3.5 and 3.6). For this purpose, LB medium containing antibiotics was inoculated with the respective *E. coli* strain (Table 3) and incubated overnight at 37 °C and 180 rpm. Subsequently, a main culture was inoculated to an OD_{600} of 0.2 and incubated up to an OD_{600} of 0.6-0.8. The culture broth was chilled on ice for 15 min, followed by supplementation with 0.1-0.5 mM IPTG inducing the heterologous protein production. The induced cultures were further incubated overnight at 16 °C and 180 rpm, in order to produce a

sufficient amount of proteins. Subsequently, the cells were harvested (4000 rpm, 15 min, 4 °C) and prepared for cell lysis by sonication. In brief, sonication was conducted by resuspension of the cell pellet in lysis buffer (pH 8.0), which is composed of 10% (v/v) glycerol, 10 mM imidazole, 300 mM NaCl and 50 mM NaH_2PO_4. The freshly prepared suspension was supplemented with β-mercaptoethanol (1.407 ml/l). Sonication was performed at 4 °C (specifications: amplitude 10%, pulse on 0.5 sec., pulse off 1.0 sec., cycle time duration 30 sec.). In total, the sonication cycle was repeated seven times. A subsequent centrifugation step at 4000 rpm and 4 °C for 30 min was performed to remove the cell debris from the cell lysate. The resulting cell lysate supernatant, containing soluble proteins, and the cell lysate pellet, comprising cell debris and insoluble proteins, were then analyzed by SDS-PAGE, or subjected to chromatographic protein purification.

In order to directly evaluate the successful production of the recombinant proteins in *E. coli*, the cell lysate supernatant was directly applied to SDS-PAGE analysis (3.3.3), as the recombinant protein can be identified by comparative analysis with a control strain (compare Table 3). To evaluate if insoluble recombinant proteins were produced, the pellet phase of the obtained sonication lysate was simultaneously analyzed by SDS-PAGE. This procedure was pursued for the expression performance analyses concerning the production of CrpA in different *E. coli* host strains (see section 3.5).

The hexahistidine-tagged (His6-tagged) proteins His6-AdmH and His6-CrpA were purified from the cell lysate using affinity chromatography. For this purpose, a polypropylene column (Qiagen) was filled with a stationary phase composed of Protino® Ni-NTA agarose (Macherey-Nagel), consisting of an agarose-bead matrix. The matrix is precharged with Ni^{2+}-ions and is embedded in nitrilotriacetic acid (NTA). Freshly prepared lysis buffer with increasing imidazole concentrations (20-250 mM) was used to purify the matrix-bound proteins. The elution of the recombinant enzymes occurred at an imidazole concentration of 250 mM. The resulting chromatography fractions were analyzed by SDS-PAGE (3.3.3), in order to verify the successful production of His6-AdmH and His6-CrpA, respectively. The purified His6-AdmH was further prepared for the application in an *in vitro* assay (see section 3.6). For this purpose, the eluted samples were desalted, using a PD-10 column (GE Healthcare Life Sciences) and concentrated by centrifugation applying an Amicon Ultra-15 Centrifugal Filter Unit (Merck) according to the manufacturer's instructions. The purified His6-AdmH was dissolved in phosphate buffer (50 mM, pH 8.0), containing 5% (v/v) glycerol [245].

3.3.3 SDS-PAGE electrophoretic protein separation

To analyze the production of CrpA in *E. coli* strains featuring different expression plasmids (Table 3), the sonified cell lysate fractions (supernatant and pellet) were subjected to SDS-PAGE.

For this, a Criterion™ chamber (Bio-Rad) and 3-8% precast Criterion™ XT Tris-Acetate gradient gels were used. The LC5699-HiMark™ (Thermo Fisher Scientific) served as protein ladder. Following the manual's instructions, 30 µl of each sample were mixed with 25 µl XT sample buffer (Bio-Rad) and 5 µl reducing agent (Bio-Rad). The prepared samples were incubated for 5 min at 95 °C and centrifuged for 10 min at 12000 rpm, prior to application. A run of 90 min was performed according to the manufacturer's instructions. Subsequently, visualization of the separated protein samples was conducted by incubation of the gels in staining solution (0.05% (w/v) Coomassie brilliant blue R250, 5% (v/v) acetic acid, 50% (v/v) methanol) for 30 min at room temperature on a rocking platform shaker. Destaining was performed overnight at room temperature on a rocking platform shaker, using a water bath (dH_2O).

3.4 Cryptophycin biosynthesis in *Nostoc* sp. ATCC 53789

In order to verify the production of cryptophycins in *Nostoc* sp. ATCC 53789 and to develop a suitable method for the isolation of these compounds, the strain was cultivated at a 200 ml scale, as described earlier (cf. section 3.1.1), and harvested by centrifugation at 4 °C and 4000 rpm for 30 min. The metabolites in the supernatant were recovered by adsorption onto the resin Amberlite® XAD7 (Sigma-Aldrich), which was added at a concentration of 2% (w/v). The resin was separated from the culture broth by filtration, using a DURAN® filter funnel and the bound compounds were eluted from the resin with dichloromethane (10% (v/v) of the total broth volume). The elution step was repeated three times. At the same time, the obtained cell pellet was frozen at -80 °C for 24 h, followed by lyophilization (24 h). Subsequently, the freeze-dried cells were grinded to a powder, using a mortar and a pestle and poured into a beaker for extraction. The extraction was conducted three times (10 ml solvent/1 g pellet dry weight), administering a mixture of dichloromethane and acetonitrile (1:4 v/v). The extracts were filtered using a cellulose filter (Macherey-Nagel, No. 1670), concentrated to dryness on a rotary evaporator and redissolved in 1 ml acetonitrile for liquid chromatography-mass spectrometry (LC/MS) analysis.

The analysis was carried out using an Agilent 1260 Infinity LC system coupled to a UV detector and a Compact Q-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). Ionization was performed by electrospray ionization (ESI). All analyses were performed in positive ion mode. The high pressure liquid chromatography (HPLC) flow rate was adjusted to 0.4 ml/min using an RP EC-Nucleodur C18 column (100 x 2 mm, 2.7 µm; Macherey-Nagel). All samples were analyzed applying a gradient from 5 to 98% acetonitrile/0.1% formic acid in water for 10 min, followed by 5 min at 98% acetonitrile.

3.5 *In vivo* biosynthesis of MPBA

In the following sections, methods and relevant background information regarding the reconstruction and *in vivo* biosynthesis of MPBA production in *E. coli* is described in detail.

3.5.1 *In vivo* biosynthesis of MPBA in *E. coli*

The following section contains work that was conducted by Marvin Daney Schwing [a] during his Bachelor thesis, which was designed and supervised by the author of this dissertation. Contents of this work were translated and summarized to be presented in an adapted form.

To reconstruct MPBA biosynthesis in the heterologous host *E. coli*, a 2-plasmid-based expression system was constructed, co-expressing CrpA and an assembly line activating PPTase. For this purpose a set of expression vectors, carrying either *crpA* (i.e. pET28a-crpA, pET28a-SUMO-crpA, pMal-crpA and pMal-crpA-eryAIII$_{TE}$), or a PPTase gene (i.e. pACYC-sfp, pACYC-sfpN) were constructed as described in the following chapters (3.5.1.1, 3.5.1.2 and 3.5.1.3).

Subsequently, *E. coli* BL21 (DE3) was co-transformed (cf. section 3.3.1) with pACYC-sfp and pET28a-crpA, to generate the expression strain *E. coli* BL21 (DE3)-*crpA-sfp*. The recombinant strain was grown in LB medium containing chloramphenicol and kanamycin (cf. section 3.3.2). To induce the production of proteins, 0.1 mM IPTG were added, when an OD_{600} of 0.6-0.8 was reached. In order to heterologously produce MPBA, the putative CrpA substrate 3-PPA (200 mg/l) was fed upon induction of the culture, which was further cultivated for 24 h. Next, the cell culture was lysed by sonication and subjected to Ni-NTA chromatography (cf. section 3.3.2). The production of recombinant His6-CrpA (327 kDa) was assessed by SDS-PAGE (3.3.3) of the cell lysate (pellet fraction) and the cell lysate supernatant fraction. Cell lysates of the parental strain, *E. coli* BL21 (DE3), served as negative controls for SDS-PAGE.

Furthermore, *E. coli* BL21 (DE3) was tested for the co-expression of pMal-crpA or pET28a-SUMO-crpA in combination with pACYC-sfp, or pACYC-sfpN respectively. Transformation of the parental strain with the respective plasmid combination led to the construction of *E. coli* BL21 (DE3)-MBP-*crpA-sfp*, *E. coli* BL21 (DE3)-SUMO-*crpA-sfp* and *E. coli* BL21 (DE3)-MBP-*crpA-sfpN* [a]. Strains harboring pACYC-sfp, or pACYC-sfpN and pMal-crpA were grown in the presence of chloramphenicol and ampicillin, while those with pET28a-SUMO-crpA were cultured in the presence of chloramphenicol and kanamycin. As described for the previous strains, the protein formation of SUMO-CrpA (337 kDa) and MBP-CrpA (366 kDa) was analyzed by SDS-PAGE, directly applying the prepared cell lysates. A Ni-NTA purification step was omitted in this case.

In addition *E. coli* BL21 (DE3) was co-transformed with pACYC-sfp and the TE-fusion plasmid pMal-crpA-eryAIII$_{TE}$, yielding *E. coli* BL21 (DE3)-MBP-*crpA$_{TE}$-sfp*. The production of the corresponding protein MBP-CrpA$_{TE}$ (390 kDa) and the metabolite MPBA was conducted as described earlier, followed by a comparative SDS-PAGE and HPLC analysis (cf. section 3.5.1.4), respectively.

Apart from *E. coli* BL21 (DE3), additional *E. coli* expression strains were evaluated for the

production of SUMO-CrpA (337 kDa) and MBP-CrpA (366 kDa), as well as for MPBA formation. Apart from *E. coli* BL21 (DE3), the strains *E. coli* KRX, *E. coli* ArcticExpress (DE3), *E. coli* C41 (DE3) and *E. coli* C43 (DE3) were co-transformed with pACYC-sfp and either pET28a-SUMO-crpA, or pMal-crpA. The recombinant strains were cultivation and protein expression was performed as described above. The biosynthesis of recombinant proteins was analyzed by SDS-PAGE. Following this, sample extracts were prepared and investigated by comparative LC-MS analysis in order to identify newly synthesized metabolites with a mass corresponding to MPBA (*m/z* 177.0910 $[M+H]^+$). Likewise, *E. coli* DH5α-MBP-*crpA,* a comparative host strain that is not co-expressing the PPTase gene *sfp*, was created and tested for the biosynthesis of MBP-CrpA.

Concluding the comparative analyses of all investigated recombinant expression strains, i.e. *E. coli* BL21 (DE3)-MBP-*crpA*-*sfp*, *E. coli* BL21 (DE3)-MBP-*crpA*-*sfpN* [a] and *E. coli* BL21 (DE3)-MBP-$crpA_{TE}$-*sfp*, a fractionation of new emerging HPLC signals (cf. section 4.4.1) was conducted, followed by successive LC-MS and NMR analyses (cf. section 3.5.1.4) using a MPBA standard for comparison.

Detailed information on the herein described vector constructs and designated strains can be found in the following chapters, covering the individual vector constructions, as well as in

Table 2, Table 3 and Table SI I, accordingly.

3.5.1.1 Vector construction of pET28a-crpA, pET28a-SUMO-crpA and pMal-crpA

The gene *crpA* (locus_tag: GJB62_33220) was amplified by overhang PCR with primers P1 and P2 (Table 7) from *Nostoc* sp. ATCC 53789 genomic DNA, applying the Phusion Flash PCR master mix (Thermo Fisher Scientific). The PCR program was set to the following conditions: 98 °C for 2 min; 30 cycles of 98 °C for 15 sec, 50 °C for 30 sec, 72 °C for 15 sec/kb; and a final elongation step at 72 °C for 10 min. The amplicon was subjected to NdeI-NotI restriction and ligated into the equally linearized pET28a (+), yielding pET28a-crpA.

pET28a-SUMO-crpA, carrying the small ubiquitin-related modifier (SUMO) tag, was constructed amplifying the SUMO-tag from pET28a-tmmrMcGAS-SUMO [216] by PCR. For PCR, the primers P7 and P8 (Table 7) were used, which carry homologous overlaps targeting pET28a-crpA. The PCR was conducted as described above, with the following specifications for the 30 amplification cycles: 98 °C for 15 sec, 63 °C for 30 sec, 72 °C for 15 sec/kb; and a final elongation step at 72 °C for 10 min. Subsequently, the amplified SUMO-tag fragment (355 bp) was cloned into the NdeI-linearized pET28a-crpA by Gibson assembly [225] (cf. section 3.2.4), yielding pET28a-SUMO-crpA.

Simultaneously, pMal-crpA, carrying the maltose binding protein solubility tag, was constructed. For this, *crpA* was amplified from *Nostoc* sp. ATCC 53789 genomic DNA by PCR using primers P9

and P10 (Table 7), which carry homologous overlaps, targeting pMalc2x. The PCR was conducted as described for *crpA* above. Subsequently, the amplified *crpA* fragment (8880 bp) was cloned into the Eco53kI-linearized expression vector pMalc2x by Gibson assembly [225] (cf. section 3.2.4), yielding pMal-crpA.

3.5.1.2 Vector construction of pACYC-sfp and pACYC-sfpN

The *sfp* gene [246,247] was synthesized by ATG:biosynthetics. The gene was amplified by overhang PCR using primers P3 and P4 to introduce EcoRI and KpnI restriction sites (Table 7). The PCR was conducted with the following specifications for the 30 amplification cycles: 98 °C for 15 sec, 59 °C for 30 sec, 72 °C for 15 sec/kb. The amplicon was subjected to restriction and ligated into the equally linearized pACYCDuet, yielding pACYC-sfp.

The gene *sfpN* (locus_tag: GJB62_00165) was amplified from isolated genomic DNA of *Nostoc* sp. ATCC 53789, by overhang PCR using primers P5 and P6 to introduce EcoRI and KpnI restriction sites (Table 7). The PCR was conducted with the following specifications for the 30 amplification cycles: 98 °C for 45 sec, 60 °C for 30 sec, 72 °C for 15 sec/kb; and a final elongation step at 72 °C for 10 min. The amplicon was subjected to restriction and ligated into the equally linearized pACYCDuet, yielding pACYC-sfpN [a].

3.5.1.3 Vector construction of pMal-crpA-eryAIII$_{TE}$

In parallel, the vector pMal-crpA-eryAIII$_{TE}$ was constructed for the generation of a CrpA-TE fusion protein. For this purpose, the gene encoding the TE domain (TE$_{eryAIII}$) of the erythromycin producing 6-deoxyerythronolide B synthase (DEBS) [212,248–251] was linked to the pMal-crpA-encoded *crpA*, by fusing the sequence of the N-terminal ACP-TE-linker region of *eryAIII* to the C-terminal ACP linker of *crpA* (Figure 10).

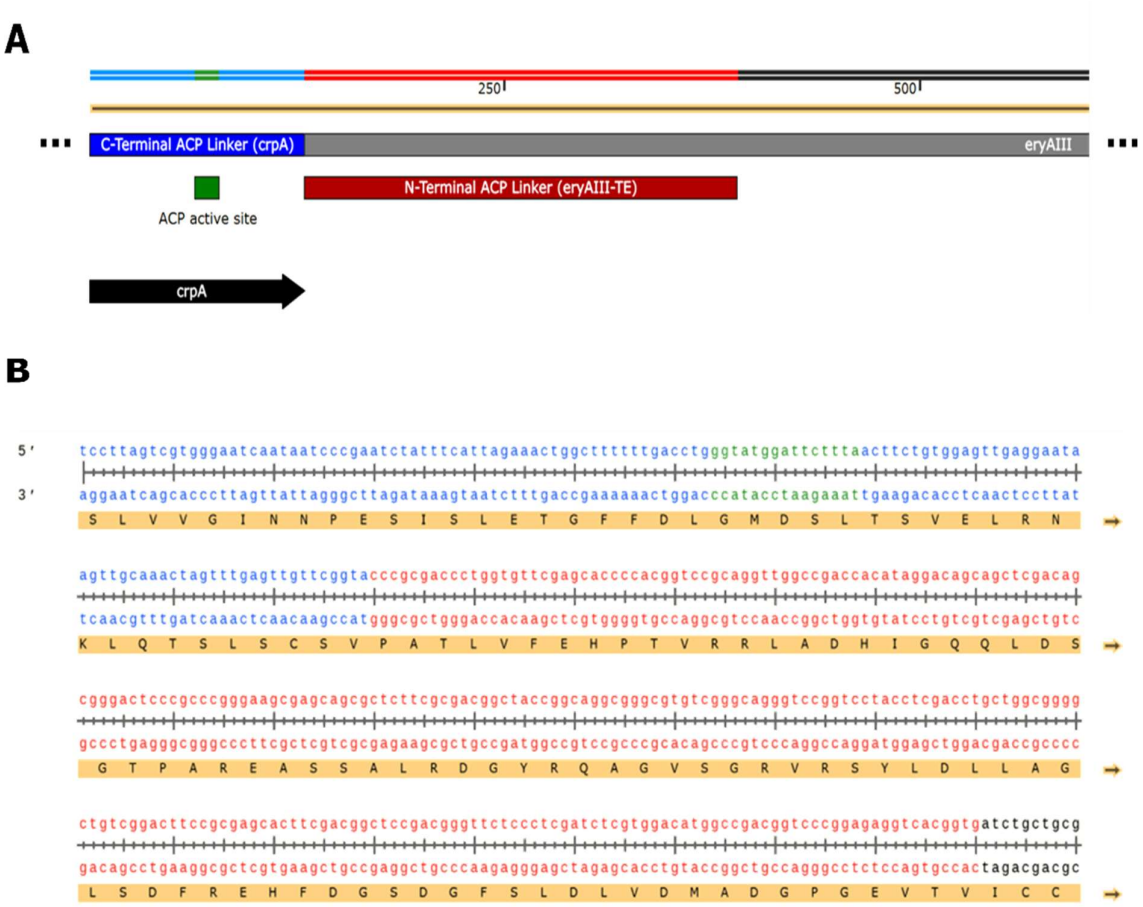


Figure 10. TE domain fusion concept. Intrinsic TE-fusion approach. The N-terminal ACP-TE linker region (red) of the DEBS3 (*eryAIII, grey*) was fused to the C-terminal ACP-linker encoding region of *crpA* (blue), while respecting the ACP active site (green) integrity of the CrpA-ACP domain (A). The TE domain sequence was cloned in frame to *crpA*, resulting in a continuously fused sequence (B).

To generate the expression vector, the $TE_{eryAIII}$ nucleotide sequence was amplified by PCR from genomic DNA of *Saccharopolyspora erythraea* DSM 40517, which was purchased from the DSMZ (German Collection of Microorganisms and Cell Cultures GmbH). For PCR amplification of the $TE_{eryAIII}$ fragment (955 bp), the primers P11 and P12 (Table 7) were used, carrying homologous overlaps to target pMal-crpA. The PCR was conducted with the following specifications for the 30 amplification cycles: 98 °C for 15 sec, 60 °C for 30 sec, 72 °C for 15 sec/kb; and a final elongation step at 72 °C for 10 min. Following this, the fragment was assembled with the KpnI-BamHI-linearized vector pMal-crpA using Gibson assembly [225], to yield the expression vector pMal-crpA-eryAIII$_{TE}$.

Table 7. Primers used for the reconstruction of MPBA biosynthesis in *E. coli*. Overlapping homologous flanks as part of the Gibson assembly are underlined. Capital letters indicate the primer annealing sites. Restriction sites are marked in italics.

Primer	Sequence
P1	5' – ggga attc*catATG*ATTACACCTTCACATGA – 3'
P2	5' – aaggaaaaaa*gcggccgc*TCAGTAACTAATATTTCTAAG – 3'
P3	5' – cc*gaattc*gATGAAGATTTACGGAATTTATATGG – 3'
P4	5' – cg*ggtacc*TTATAAAAGCTCTTCGTACGAG – 3'
P5	5' – cc*gaattc*gATGACCGCTACTCATCATTTC – 3'
P6	5' – cg*ggtacc*TCAATACTGCCAACACTTTAAG - 3'
P7	5' – gcggcctggtgccgcgcggcagccatatgcatatgTCGGACTCAGAAGTCAATC – 3'
P8	5' – aaattttcatgtgaaggtgtaatcatATGACCACCAATCTGTTCTCTGTG – 3'
P9	5' – aaagacgcgcagactaattcgagctctATGATTACACCTTCACATG – 3'
P10	5' – aggttgttgttattgttattgttgttgttgttcgagTCAGTAACTAATATTTCTAAG – 3'
P11	5' – gttgcaaactagtttgagttgttcggtaCCCGCGACCCTGGTGTTCG – 3'
P12	5' – cttgcctgcaggtcgactctagagTCATGAATTCCCTCCGCCCAGC – 3'

3.5.1.4 MPBA extraction and analytical evaluation

The expression cultures were harvested by centrifugation for 30 min at 4000 rpm. The metabolites in the supernatant were recovered as described earlier (cf. section 3.4), using the adsorber resin Amberlite® XAD7 (Sigma-Aldrich) and methanol as elution solvent. The resulting extract was applied to LC-MS analyses with the specifications described in section 3.4.

Subsequently, the sample extracts and a MPBA standard (1 mg/ml) were analyzed by comparative HPLC analysis.

The purification of compounds from the extracts followed the protocol of Malla *et al.* [93]. For this, a single-step reverse-phase HPLC was conducted on a Sphinx column (Macherey-Nagel, 10 x 250 mm, 5.0 µm particle size, 2.0 ml/min flow rate), using a Shimadzu HPLC-system (LC-20AD) with a diode array detector (SPD-M20A). The separation was performed using a linear gradient of methanol (MeOH) in water, supplemented with 0.1% trifluoroacetic acid (TFA): 20% to 100% MeOH within 20 min, 100% over 5 min and 100% to 20% in 2 min.

NMR measurements were carried out at the Department of Chemistry and Chemical Biology (TU Dortmund), using a 500, 600 or 700 MHz Avance III HD spectrometer (Bruker Biospin GmbH). For this purpose, the samples were dissolved in methanol-d_4. The obtained NMR datasets were analyzed by Prof. Markus Nett.

The MPBA standard used was kindly provided by the LDC (Lead Discovery Center, Dortmund).

3.6 *In vitro* biosynthesis of PBA

The aminomutase gene *admH* of *P. agglomerans* (Accession No. AY192157.1) was obtained by gene synthesis (ATG:biosynthetics GmbH) and amplified by PCR using primers P13 and P14 (Table 8). The PCR was conducted with the following specifications for the 30 amplification cycles: 98 °C for 15 sec, 64 °C for 30 sec, 72 °C for 15 sec/kb. The amplicon (1638 bp) was subjected to NheI-BamHI restriction and ligated into the equally linearized pET28a (+), yielding pET28a-admH. Subsequently, *E. coli* BL21 (DE3) was transformed with pET28a-admH (cf. section 3.3.1, 3.1.2 and Table SI I). The correct vector assembly was confirmed by PCR and restriction analyses. The hexahistidyl-tagged protein His6-AdmH (61.4 kDa) was heterologously produced with *E. coli* BL21 (DE3):pET28a-admH using LB medium supplemented with kanamycin (cf. section 3.3.2 and 3.5.1). For induction IPTG (0.5 mM) was added. The recombinant protein was purified by Ni-NTA chromatography (cf. section 3.3.2) and directly applied to SDS-PAGE to verify the proper production of soluble His6-AdmH.

Table 8. Primers *in vitro* PBA biosynthesis. Sequences annealing to the DNA template are indicated in capital letters, while restriction recognition sites, introduced to the primer tails, are written in italics.

Primer	Sequence
P13	5' – gatgat*gctagc*ATGAGTATTGTAAACGAAAGCG – 3'
P14	5' – cgccgc*ggatcc*TTAAATATCGCTATTTTTAG – 3'

3.6.1 Aminomutase assay

The reaction mixture was prepared in Eppendorf cups and contained purified enzyme (0.881 – 1.879 µmol) in sodium phosphate buffer [245] (50 mM, sodium phosphate containing 5% (v/v) glycerol and 300 mM NaCl, pH 8.0. In total four reaction cups were prepared. The reaction was started by the addition of substrate, i.e. 2.232 – 4.575 µmol α-homophenylalanine and allowed to proceed for a fixed time (0, 1, 2, 4 hours and overnight) at 31 °C and 150 rpm, using an Eppendorf ThermoMixer. In addition, negative control reactions were prepared, containing either the enzyme (0.881 µmol) solution without substrate, or a buffer-substrate mix without His6-AdmH. The latter was prepared containing 1.5 ml of either L-α-homophenylalanine (0.5 mg/ml), or L-β-homophenylalanine (0.5 mg/ml), or a mixture of both substrates (50% v/v), respectively. As a positive control the enzymatic conversion of L-phenylalanine (4.54 µmol) by His6-AdmH was tested.

Following the addition of acetonitrile (15% final concentration) to stop the reaction, the solution was centrifuged for 5 min at 11000 rpm. The supernatant was analyzed for product formation by HPLC on a Sphinx column as described earlier (cf. section 3.5.1.4). In an initial study the assay

samples were analyzed, using a gradient, which was set from 10 to 100% MeOH within 15 min, 100% MeOH over 5 min and from 100% to 10% MeOH in 2 min. In a subsequent study, the method was optimized for peak separation, using the same column and method with a gradient set from 10 to 100% MeOH within 20 min.

3.7 General molecular biological operations in *S. cerevisiae*

Techniques for the genetic engineering of *S. cerevisiae* are described in the following.

3.7.1 Generation of chemically competent cells and transformation of *S. cerevisiae*

If not specified otherwise, *S. cerevisiae* was transformed using the lithium acetate method. Chemically competent yeast cells, applicable to lithium acetate transformation, were generated according to a modified protocol from Gietz *et al.* [252]. A preculture grown in YPD overnight was used to inoculate 50 ml of fresh YPD medium to an OD_{600} of 0.2. Incubation was then continued until an OD_{600} of 0.5-0.6 was reached. The cells were harvested by centrifugation at 4000 rpm for 5 min at room temperature and washed twice with sterile, deionized H_2O (dH_2O). For the washing step, the pellet was first dissolved in 25 ml dH_2O, centrifuged and resuspended in 0.5 ml dH_2O. The suspension was transferred into an Eppendorf cup and centrifuged likewise. Subsequently, the cell pellet was resuspended in 0.5 ml sterile filtered frozen competent cell solution (FCCS), composed of 0.5 ml 5% (v/v) glycerol, 1 ml 10% (v/v) DMSO and 10 ml dH_2O. Aliquots of 50 µl FCCS dissolved cells were frozen at −80 °C until further usage.

For each transformation, 50 µl FCCS dissolved cells were thawed at 37 °C for 15-30 sec and centrifuged at 13000 rpm for 2 min. The pellet was dissolved in a freshly prepared transformation mix containing 240 µl PEG_{3350} (50% (w/v), 36 µl LiAc (1.0 M) and 50 µl single-stranded carrier DNA (2.0 mg/ml, salmon sperm) [252]. Successively, the DNA of interest was added for transformation. The samples were dissolved completely by harsh vortexing and applied to a heat shock at 42 °C for 40 min. The cells were centrifuged at 13000 rpm for 30 sec and resuspended in 1 ml YPD for regeneration at 30 °C and 180 rpm for 3 h. Finally, the cells were centrifuged, washed with dH_2O, resuspended in 200 µl dH_2O and plated for screening on selection plates at 30 °C for three days. In case of genome editing (*crpA*, *crpH*, *scyA*) using CRISPR-Cas9, uracil deficient YNB Drop Out Medium agar (Sigma-Aldrich) was used for selection, while in case of LTR-recombination (*crpB-D*, *scyC, scyB*) YPD agar with antibiotics was applied.

3.7.2 Generation of expression cassettes

For the heterologous expression of cyanobacterial biosynthesis genes in *S. cerevisiae*, a set of universal minimal expression systems was constructed according to the work of Redden *et al.* [170] and Curran *et al.* [169] (cf. section 1.3.2.1). The individual sets of required promoter and terminator sequences were designed according to the provided libraries [169,170] and combined to generate

several vectors, including the pMA and the pYPK series (Table 2, Table SI I), which encode universal expression cassettes, respectively.

The pMA vectors were equipped with two unique restriction recognition sites to facilitate the cloning of genes of interest (GOIs) between the yeast promoter and a terminator sequence. Two further unique NotI restriction sites enable the liberation of the GOI-expression cassette (promoter-GOI-terminator) as individual fragments. This linear NotI-excised fragments can be directly applied to LTR-facilitated genome editing (cf. section 3.7.4) or CRISPR/Cas9-directed genome editing (cf. section 3.7.3) in *S. cerevisiae*. The latter required the additional integration of homologous flanks into the vector design, which would mediate the chromosomal integration into *S. cerevisiae* (Figure SI I). Both applications are described in detail in the individual research projects (cf. section 3.8 and 3.9). The same design principle of restriction-liberation was applied to the pYPK vectors (i.e. pYPKpw-PC8-crpB-T3 and pYPKpw-PD1-crpD-T8), which already include the expression cassettes PC8-T3 and PD1-T8, as well as the large cryptophycin genes *crpB* and *crpD*.

In total three pMA and two pYPK vectors (Table 2, Table SI I), encoding the previously outlined design of expression systems, were constructed *in silico* and ordered from Invitrogen or ATG:biosynthetics.

The subcloning vector pMA-RQ-PC8-T3 exhibits the unique restriction recognition sites Bsp120I and BglII. In addition, upstream and downstream “inter-fragment homology” flanking regions targeting locus 416d on yeast chromosome IV, were integrated in the design for subsequent CRISPR/Cas9 applications. The vector was applied for the chromosomal integration of *crpH* and *scyA* and their associated expression system (cf. section 3.8 and 3.9) in the yeast’s genome. pMA-RQ-PA3-T30 was equipped with unique NsiI and KpnI restriction sites. It was subsequently used for the chromosomal integration of *crpC* and *scyC* in an LTR-directed homologous recombination approach, by providing the Not-liberated *crpC*- or *scyC*-expression cassette (cf. section 3.8 and 3.9). The same expression system was used to reconstruct the expression of *scyB* in *S. cerevisiae*. However, instead of the NotI-facilitated restriction, a TauI restriction was used to liberate the total *scyB*-expression cassette from the subcloning vector. Subsequently the cassette was integrated into locus YJRO78W, which is encoding the gene *bna2*, instead of an LTR-site (cf. section 3.9). The subcloning vector pMA-RQ-PB2-T27 was equipped with the unique restriction sites MluI and BglII. In addition, upstream and downstream “inter-fragment homology” flanking regions, targeting locus 308a on yeast chromosome III, were integrated in the design for subsequent CRISPR/Cas9 application. The vector was used for the reconstruction of the *crpA*-expression cassette (cf. section 3.8).

These pYPK vectors were used as template vectors in order to insert the *crpB*- and *crpD*-

expression cassettes into the yeast genome loci YPRCΔ15/Ty1 (chromosome XVI) and YPRCτ3/Ty4 (chromosome XVI) via LTR recombination. The *crpB*-expression cassette was liberated using a pair of uniqueBsp1407I restriction sites, while for the *crpD* expression cassette an XhoI restriction site was used, accordingly.

Detailed information on the direct application of the described expression systems is given in the individual research projects (cf. section 3.8 and 3.9). Detailed sequence information and featured elements of the constructed universal minimal expression systems are illustrated in Figure SI I and Figure SI II.

3.7.3 CRISPR/Cas9-directed genome editing

CRISPR/Cas9 was used for the chromosomal integration of *crpA*, *crpH* and *scyA*, adapting the CASdesigner tool and following the already published performance protocol, as described by Reider Apel *et al.* [156].

In this thesis the CASdesigner software was used to manually design the 1 kb locus homology fragments (upstream/downstream) and the "inter-fragment homology" flanking regions (upstream/downstream) with overlaps (30-60 bp) to the 1-kb homology fragments. Instead of using the promoter and terminator bricks suggested by the CASdesigner tool, the design was adapted to incorporate the individual minimal expression systems, constructed according to Redden *et al.* and Curran *et al.* [169,170] (cf. section 3.7.2). Hence, deviating from the original design, the herein used CASdesign (Figure 11) was realized, harnessing the universal subcloning vectors pMA-RQ-PC8-T3 and pMA-RQ-PB2-T27 (Table 2, Table SI I) for CRISPR/Cas9 genome editing. As previously described, the vectors are already encoding the minimal promoter and terminator sequences and the CASdesigner-suggested upstream and downstream "inter-fragment homology" flanking regions. In this case, the homology flanks are targeting the loci 416d (ARS416/ARS1, chromosome IV) and 308a (ARS308a, chromosome III). As the subcloning vectors contain unique restriction sites, it is possible to directly clone any GOI between the promoter and the terminator sequence. By NotI restriction, the GOI, including a functional expression cassette and the locus-targeting "inter-fragment homology" flanking regions can be liberated from the vectors and directly applied to the CRISPR/Cas procedure (cf. section 3.7.2, 3.7.4 and Figure SI I). The predefined donor DNA transformation mix, i.e. the NotI-liberated fragment and the PCR-generated 1-kb locus homology fragments (upstream/downstream), together with the locus-corresponding pCut-X plasmid of the toolkit (here pCAS9-416 and pCAS9-308, Table 5), can be applied to subsequent genome editing in yeast.

In brief, preparation of the strain for CRISPR/Cas9 genome editing was conducted, as previously published [156]. A preculture was cultivated overnight in YPD, followed by inoculation of the main culture to an OD_{600} of 0.15. The main culture was incubated up to an OD_{600} of 0.8-1.2 and

centrifuged for 3 min at 3000 rpm. The subsequent washing step was successively performed by resuspension of the cell pellet in 45 ml ice-cold ddH_2O and 45 ml ice-cold yZap buffer (1 M sorbitol, 10 mM $CaCl_2$) with an intermediate centrifugation step. Following this, the resulting cell pellet was resuspended in 20 ml yCondition buffer (0.1 M LiAc, 1 mM DTT) and incubated for 30 min at 30 °C and 250 rpm. A second washing step was performed with 45 ml ice-cold yZap buffer, followed by centrifugation and resuspension in 200 µl ice-cold yZap buffer. Transformation of the prepared strain with the transformation mix was conducted by electroporation in 2 mm cuvettes with 2.5 kV, 200 mA and 25 µF, followed by regeneration in 0.5 ml YPD at 30 °C and 180 rpm for 30 min. The regenerated cells were harvested by centrifugation for 1 min at 10,000 rpm. The supernatant was removed and the pellet was washed once with ddH_2O. Finally, the pellet was resuspended in 100 µl ddH_2O. Subsequent selection was performed on selective uracil deficient YNB Drop Out Medium agar (Sigma-Aldrich) for 2-3 days.

After PCR-guided confirmation, all positive yeast transformants, carrying the successfully integrated expression cassette and the GOI, were cured from the used pCAS9 plasmid applying a 5-fluoroorotic acid (5-FOA) counter selection for uracil deficient strains, as described by Flagfeldt *et. al.* [253]**.** Selective replica plating on YPD and YNB with uracil depletion identified the successfully cured transformants, which were confirmed again by PCR analysis.

Information on the used primer sequences and the exact procedure are given in the individual research projects (cf. section 3.8 and 3.9).

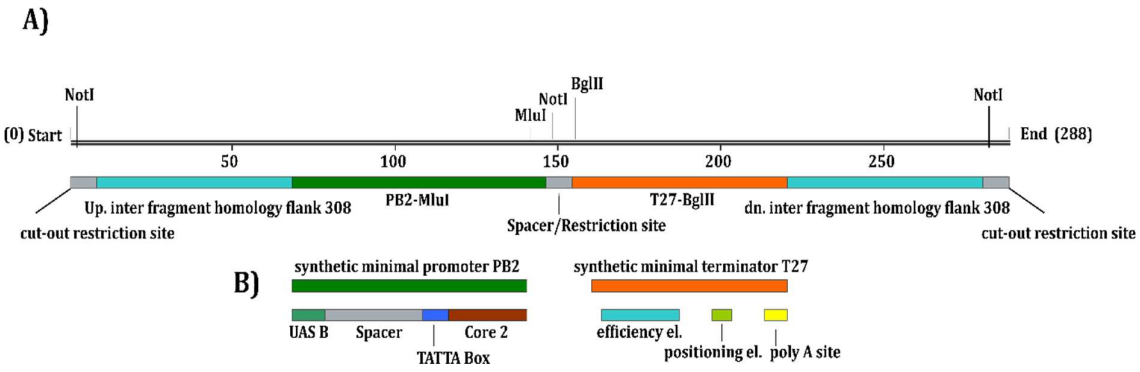


Figure 11. Schematic outline of CASdesign (here: universal empty vector pMA-RQ-PB2-T27) for CRISPR/Cas9-guided genome editing in *S. cerevisiae*. (A) The design is applicable to integrate any gene of interest into locus 308a (chromosome III) and enable its functional expression, by the integration of the gene between the minimal synthetic promoter and terminator sequence, using the unique restriction sites MluI and BglII. The integrated NotI restriction sites (cut-out restriction sites, Spacer/Restriction site) allow for the individual liberation of the synthetic minimal promoter and terminator cassette and any subsequently cloned gene of interest as a discrete cassette, equipped with a functional yeast expression system. (B) The synthetic minimal expression systems were designed according to Redden *et al.* [170] and Curran *et al.* [169]. Exemplarily, the exact composition of the pMA-RQ-PB2-T27-encoded expression system is depicted. The total CASdesign was constructed adapting the CASdesigner software [156]. Notation: PB2-MluI refers to the total synthetic minimal promoter PB2, including a MluI restriction recognition site; T27-BglII indicates the total synthetic minimal terminator T27, including a BglII restriction recognition site; up., upstream; UAS, upstream activating sequence; efficiency el., efficiency element; positioning el., positioning element; dn., downstream.

3.7.4 LTR-recombination

Long terminal repeat directed recombination (LTR-recombination) is targeting native, genomic retrotransposon terminal repeats in the yeast chromosomes and can be harnessed for the genomic reconstruction of genes, based on homologous recombination [253,254] (Figure 12). Consequently, the expression cassettes of *crpB-D* and *scyC* were reconstructed in locus YORWΔ22/Ty1 (chromosome XV), YPRCτ3/Ty4 (chromosome XVI), or YPRCΔ15/Ty1 (chromosome XVI), adapting the protocol from Flagfeldt *et al*. [253]. For selection purposes, the expression cassettes were equipped with hygromycin, zeocin, or G418 (kanamycin) resistance genes. The LTR method was expanded for the genomic integration of *scyB* and its associated expression system. However, instead of LTR sites, the integration of *scyB* was coupled to the knock-down of *bna2*/locus YJR078W (chromosome X), which is involved in anthranilate biosynthesis. In general, LTR-recombination of the respective yeast strain was performed by LiAc transformation (cf. section 3.7.1) with a set of linear, gel-purified fragments equipped with overlapping homologous regions to provide subsequent *in vivo* homologous assembly of adjacent fragments and recombination into the target locus. Each set of fragments contained an antibiotic resistance cassette, the locus-targeting upstream- and downstream fragments (LTR-flanks) and a linear expression-cassette fragment, including the GOI and the minimal promoter and terminator sequences. The individual upstream/downstream LTR flanks were amplified by PCR from previously isolated genomic DNA of *S. cerevisiae*. In case of *scyB* the LTR flanks were substituted by flanks targeting *bna2*/locus YJR078W (chromosome X). The corresponding resistance cassette was generated by PCR from plasmid pYTK077, pYTK079, or pYTK080. The expression cassettes were provided by Bsp1407I-, NotI-, TauI-, or XhoI-mediated liberation from the designed subcloning and pYPK vectors, i.e. pYPKpw-PC8-crpB-T3, pMA-RQ-PA3-crpC-T30, pYPKpw-PD1-crpD-T8, pMA-RQ-PA3-scyB-T30 and pMA-RQ-PA3-scyC-T30. The overlapping homologous regions, which are required for the correct *in vivo* assembly of adjacent fragments (upstream LTR flank, GOI cassette, resistance gene, downstream LTR flank), were incorporated during PCR amplification by integration into the primer tails. Information on expression cassettes are specified in section 3.7.2. Primer sequences, resistance marker, as well as homologous upstream and downstream flanks and overlapping regions are listed and specified in the individual research projects (cf. section 3.8 and 3.9).

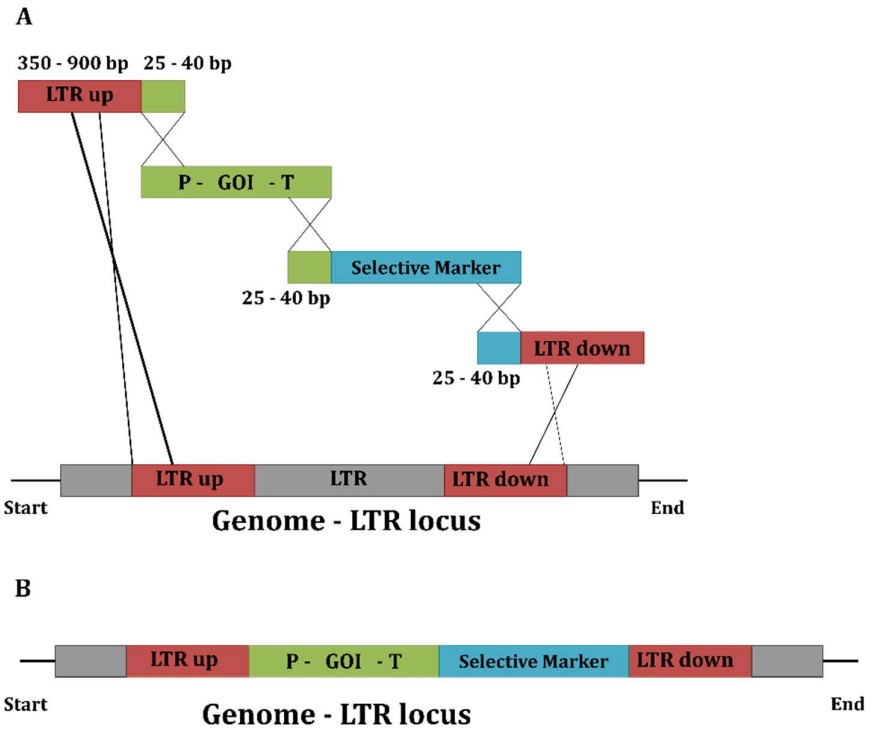


Figure 12. Schematic view of LTR-recombination. (A) During PCR amplification, the individual fragments are equipped with homologous regions (25-40 bp) equivalent to the sequence of the adjacent fragment. The amplified LTR up and LTR down regions (350-900 bp) further convey sequence homology, targeting the site of genomic integration. Upon transformation of *S. cerevisiae* and based on subsequent *in vivo* homologous recombination, the individual linear fragments are assembled and (B) integrated into the dedicated genomic locus, accordingly. Notation: LTR, long terminal repeat of retrotransposon; Start, defines the start of the LTR sequence; LTR up, defines the upstream fragment of the integration locus within the LTR sequence; P, promoter; GOI, gene of interest; T, terminator; LTR down, defines the downstream fragment of the integration locus within the LTR sequence; End, defines the end of the LTR sequence. Homologous recombination events are indicated by crossing lines.

3.8 Reconstruction of cryptophycin biosynthesis in *S. cerevisiae*

The four genes *crpA-D* were PCR amplified from genomic DNA of *Nostoc* sp. ATCC 53789 and subsequently equipped with yeast compatible expression elements. For this purpose, each single gene was cloned into a predesigned subcloning vector (cf. section 3.7.2) using *E. coli* as subcloning host (compare Table 3). In order to assemble the cryptophycin BGC in the heterologous host, the *S. cerevisiae* parental strain (*S. cerevisiae-sfp*) was successively transformed with the liberated individual GOI-expression cassettes, applying either a CRISPR/Cas9 (cf. section 3.7.3) or an LTR-assisted (cf. section 3.7.4) genome editing approach.

In detail, the first gene of cryptophycin biosynthesis, *crpA*, was integrated into the yeast's genome using CRISPR/Cas9-directed genome editing, targeting locus 308a on yeast chromosome III. For this purpose, *crpA* was amplified by PCR using primers P15 and P16 (Table 9) cloned into the

pMA-RQ-PB2-T27 subcloning vector (Table 2, Table SI I), using MluI-BglII-directed restriction-ligation cloning. The functional *crpA*-expression cassette, including the upstream and downstream inter-fragment homology regions (see Figure SI I) was liberated by NotI restriction and applied to the subsequent transformation procedure (cf. section 3.7.3), together with the plasmid pCAS9-308 and the 1 kb locus homology fragments (upstream and downstream). The latter were amplified by PCR, using P17-P20, (Table 9). Selection was performed on YNB agar with uracil depletion. Subsequently, positive transformants were cured from pCAS9-308, applying the 5-FOA counter selection, followed by selective replica plating. Positive clones, carrying the GOI and lacking the pCAS9 plasmid were identified by PCR, targeting *crpA*.

Subsequently, *crpB* was integrated using pYPKpw-PC8-crpB-T3. The *crpB*-expression cassette was liberated by Bsp1407I restriction and directly applied to LTR-recombination, targeting locus YPRCΔ15/Ty1 (chromosome XVI). The required locus-specifying LTR flanks were amplified by PCR using primers P21-P24 (Table 9), containing a homologous sequence with the promoter PC8 of the *crpB*-expression cassette (LTR flank upstream). P25 and P26 (Table 9) were used to amplify the selective kanamycin (G418) marker sequence from pYTK077, including a homologous region with the T3-terminator of the *crpB*-expression cassette, respectively. Subsequently, *S. cerevisiae-sfp-crpA* was transformed with this fragment set and screened for positive transformants on YPD agar containing G418 for the selection of a positive *crpB*-Kan^R integration. Verification of a successful integration was assessed by PCR of the isolated genomic DNA from the selected yeast clones, targeting *crpB*. In addition, the resulting mutant *S. cerevisiae-sfp-crpA-crpB* was tested for the presence of *crpA*.

The pathway gene *crpD* was integrated into the host's genome using pYPKpw-PD1-crpD-T8. The *crpD*-expression cassette was excised by XhoI restriction and subsequently applied to LTR-recombination, targeting locus YPRCτ3/Ty4 (chromosome XVI). The required locus-specifying LTR flanks were amplified by PCR using the primers P27-P30 (Table 9), which include a homologous sequence with the promoter PD1 of the *crpD*-expression cassette (LTR flank upstream, P27 and P28), or the Zeo^R terminator sequence (LTR flank downstream, P29 and P30). P31 and P32 were used to amplify the selective zeocin marker sequence from pYTK080, including a homologous region with the T8-terminator of the *crpD*-expression cassette, respectively. Subsequently, *S. cerevisiae-sfp-crpA-crpB* was transformed with this fragment set and screened for positive transformants on YPD agar containing zeocin. To maintain the selective pressure for *crpB*-Kan^R during the recombination event G418 was added to the selective medium. Verification of a successful integration was assessed by PCR of the isolated genomic DNA from the selected yeast clones, targeting *crpD*. In addition, the resulting mutant *S. cerevisiae-sfp-crpA-crpB-crpD* was tested for the presence of *crpA* and *crpB*, respectively.

In order to genomically integrate *crpC*, pMA-RQ-PA3-T30 was applied to equip the GOI with suitable yeast expression elements. For this purpose, the gene was amplified by PCR using primers P33 and P34 (Table 9) and cloned into the subcloning vector by KpnI-NsiI-directed restriction-ligation cloning to yield pMA-RQ-PA3-crpC-T30. Subsequently, the resulting *crpC*-expression cassette was liberated by NotI-excision and used in LTR-recombination (cf. section 3.7.4) for the genomic integration into locus YORWΔ22/Ty1 on chromosome XV. Hence, the *crpC*-expression cassette was liberated by NotI-excision. The primers P37-P40 (Table 9) were used to amplify the locus-specifying LTR flanks, containing homologous regions with either the promoter PA3 of the *crpC*-expression cassette (LTR flank upstream, P37 and P38), or the Hyg^R terminator sequence (LTR flank downstream, P39 and P40). For the PCR amplification of the selective hygromycin marker sequence from pYTK079, the primers P35 and P36 (Table 9) were applied, including a homologous region with the T30-terminator of the *crpC*-expression cassette, respectively. Next, *S. cerevisiae-sfp-crpA-crpB-crpD* was transformed with this fragment set and screened for positive transformants on YPD agar containing hygromycin, as well as the antibiotics G418 and zeocin to maintain the selective pressure for *crpB*-Kan^R and *crpD*-Zeo^R during the recombination procedure. Verification of a successful integration was assessed by PCR of the isolated genomic DNA from the selected yeast clones, targeting *crpC*. Likewise, the resulting strain *S. cerevisiae* ATi01 was simultaneously tested for the presence of the PPTase-encoding *sfp*, as well as for the total set of the cryptophycin BGC-genes *crpA*, *crpB* and *crpD*.

The cryptophycin-specific halogenase CrpH was heterologously expressed in *S. cerevisiae* to facilitate the analytical identification of a recombinantly produced metabolite. Therefore, *crpH* was integrated into locus 416d (ARS416/ARS1) of *S. cerevisiae* chromosome IV using CRISPR/Cas9. For this purpose, *crpH* was amplified by PCR using primers P41 and P42 (Table 9) and cloned by Bsp120I-BglII-directed restriction-ligation cloning into the subcloning vector pMA-RQ-PC8-T3, which includes the upstream and downstream inter-fragment homology regions of locus 416d (Figure SI I). The resulting vector pMA-RQ-PC8-crpH-T3 served to liberate the *crpH*-expression cassette by NotI restriction and subsequent application to the genome editing procedure. Following the previously described protocol (cf. section 3.7.3), *S. cerevisiae* ATi01 was transformed with the *crpH*-expression cassette, the plasmid pCAS9-416 and the 1 kb locus homology fragments. The latter were amplified by PCR using P43-P46 (Table 9).

The subsequent selection was performed on uracil-deficient YNB agar. Positive transformants were cured from pCAS9-308, applying the 5-FOA counter selection, followed by selective replica plating. Overall, clones carrying the GOI and lacking the pCAS9 plasmid were identified by PCR, targeting *crpH*. Likewise, the resulting strain *S. cerevisiae* ATi02 was tested by PCR for the presence of the cryptophycin BGC-genes *crpA-D* and the PPTase-encoding gene *sfp*.

A detailed overview of the used primer pair sequences is given in Table 9.

Table 9. Primers of cryptophycin biosynthesis reconstruction. Sequences annealing to the DNA template are indicated in capital letters, while restriction recognition sites are symbolized by italics lower case letters. Primer tails including homologous regions are underlined.

Primer	Sequence
P15	5' – acagtcgc*acgcgt*ATGATTACACCTTCACATG – 3'
P16	5' – catctcga*agatct*TCAGTAACTAATATTTCTAAG – 3'
P17	5' – CAAAACTAAGAAGAATGAGATGATAAG – 3'
P18	5' – TTAGATAAAAAGAAAAAAATTCGAAGTTAATGT – 3'
P19	5' – TCTTTGCTACATATTGCTACCACTTC – 3'
P20	5' – TGATAGAACGAGTACAACACCCG – 3'
P21	5' – GCCAGGCGCCTTTATATCAT – 3'
P22	5' – <u>cttttataacttgaattacaacgttacaagttaattaatcacatgctactgc</u>GTTTGCGAAACCCTATGCTCT – 3'
P23	5' – AATGGAAGGTCGGGATGAG – 3'
P24	5' – ATAAAGCAGCCGCTACCAAA – 3'
P25	5' – <u>gaaataaagagtatcatctttcaaagaacacccgagatcgtccaGTATAC</u>TTAGCTTGCCTCGTCCCCG – 3'
P26	5' – cagtatagcgaccagcattcacata – 3'
P27	5' – AAAGGAGGTGCACGCATTAT – 3'
P28	5' – <u>tataacttgaattacaacgttacaagttaattaagcccctctgtgtctgc</u>TTCCAAGGAGGTGAAGAACG – 3'
P29	5' – <u>agaaagtaatatcatgcgtcaatcgtatgtgaatgctggtcgctatactg</u>GATGGGACGTCAGCACTGTA – 3'
P30	5' – CGGTATTACTCGAGCCCGTA – 3'
P31	5' – <u>caaactgccatggaaaatcgatgttcttgtggtgaatgtgtgaaGTCGAC</u>TTAGCTTGCCTCGTCCCCG – 3'
P32	5' – cagtatagcgaccagcattcacata – 3'
P33	5' – cctgacca*atgcat*ATGAATTTAATCGAATTTTTACA – 3'
P34	5' – ggattcgg*ggtacc*CTATTTGCCCAAGTTTTTT – 3'
P35	5' – <u>aactaaaaaaatataaggaagaaaaaaatagctaatttttccggcagaa</u>GTTTAGCTTGCCTCGTCCC – 3'
P36	5' – CAGTATAGCGACCAGCATTCA – 3'
P37	5' – CACCGGAGCTTGGATATGAT – 3'
P38	5' – <u>cacaaccgacgatccggggtcatggcggctatttttttttttttttttt</u>TTCGCGGGCTGTTACTTATC – 3'
P39	5' – <u>agaaagtaatatcatgcgtcaatcgtatgtgaatgctggtcgctatactg</u>GGACCAACTATCATCCGCTAA – 3'
P40	5' – CGTGATAAACGATCGCCATA – 3'
P41	5' – gaagaa*gggccc*ATGTCTACACTGCCTAATTC – 3'
P42	5' – ggagga*agatct*CTAAGTGTGGAGCAAAGAG – 3'
P43	5' – CCAACTGCATGGAGATGAGTC – 3'
P44	5' – GGGTCCGGTTAAACGGATC – 3'
P45	5' – CCGAACATGCTCCTTCACTATTTTAAC – 3'
P46	5' – CAATTGAGGAAACTTGAAAGGTGTGG – 3'

3.8.1 Heterologous expression of the cryptophycin pathway

In order to analyze the transcription profile and to assess the production of cryptophycin derivatives of the transgenic yeast strain *S. cerevisiae* ATi01, an expression cultivation was performed. For this purpose, a preculture of *S. cerevisiae* ATi01 was grown overnight in YPD at 30 °C and 180 rpm. A main culture of 1 l YPD (pH 4) was inoculated to an OD_{600} of 0.2 and incubated at 30 °C and 180 rpm up to an OD_{600} of 0.4-0.6. Following this, the culture was adjusted to the expression conditions by incubation at 15 °C and 180 rpm for 30 min. Subsequently, activation of the heterologous pathway was accomplished by the addition of 2% (v/v) galactose, which is initiating the PGal1 (galactose promoter)-directed *sfp* expression. As the yeast is no native producer of cryptophycin, the addition of the relevant and non-native biosynthetic precursor 3-PPA (100 mg/l) (cf. section 1.2.1) was conducted concurrent with the galactose induction. Likewise, leucine was added in order to support the generation of the vital precursor α-ketoisocaproic acid, which is an intermediate of yeast's native leucine metabolism [255]. Subsequently, the incubation of the expression cultures was continued for three days at 15 °C and 180 rpm.

In a parallel attempt, an expression culture of *S. cerevisiae* ATi01 was fed with 3-fluorotyrosine (50 mg/l) upon induction of the BGC to introduce the fluorinated precursor into the cryptophycin biosynthesis. In case of a successful incorporation into the final product, the synthesized cryptophycin could be detected by ^{19}F-NMR spectroscopy. In addition, the transgenic strain *S. cerevisiae* ATi02 was analyzed for the biosynthesis of chlorinated cryptophycin derivatives. In this case, the culture was supplemented with 0.5% (v/v) NaCl (1 M) upon induction.

In all experiments, the strain *S. cerevisiae* CEN.PK2-1C was used as a negative control.

3.8.2 SDS-PAGE

In order to analyze the presence of the cryptophycin biosynthesis proteins in *S. cerevisiae* ATi01, aliquots (1 ml) of cell lysate from the recombinant strain and *S. cerevisiae* CEN.PK2-1C were subjected to protein isolation and SDS-PAGE analysis. For this purpose, the samples were taken 24 h after induction of the strains. The suspension was centrifuged at 4000 rpm and 4 °C for 10 min. The resulting pellet was resuspended in 50 mM Tris-HCl buffer (pH 7.5). Subsequently, cell lysis was performed applying the bead tube principle of the Kit "Genomic DNA from microorganisms" (Macherey-Nagel) for yeast. The dissolved cell suspension was transferred to the provided bead tubes (type C) and lysis was performed by bead-beating in a vortexer. Following this, the lysate was centrifuged (4000 rpm, 4 °C, 10 min) and the samples (supernatant and pellet fraction) were prepared in analogy to the *E. coli* experiments (cf. section 3.3.3). Likewise, a Criterion™ chamber (Bio-Rad) and 3-8% precast Criterion™ XT Tris-Acetate gradient gels were used for SDS-PAGE analysis. The LC5699-HiMark™ (Thermo Fisher Scientific) served

as protein ladder. In addition to the undissolved samples, a series of diluted (1:10) samples was analyzed in parallel to the undissolved protein samples. The expected size of the cryptophycin proteins was calculated with SnapGene (CrpA: 325.6 kDa; CrpB: 383.2 kDa; CrpC: 219.6 kDa; CrpD: 377.3 kDa).

3.8.3 RNA isolation and reverse transcription-profiling

To assess the functionality of the applied minimal artificial expression systems, the transcription of the genes *crpA-D* was analyzed. For this purpose, the total RNA was isolated from the transgenic yeast strain *S. cerevisiae* ATi01 after induction, followed by reverse transcription of the transcriptome. The resulting cDNA was analyzed by PCR (Table 9) and subsequent gel electrophoresis to verify the transcription of the individual cryptophycin genes. In detail, the expression cultures were harvested by centrifugation (4000 rpm, 10 min) in Eppendorf cups, 24 h after induction of the cryptophycin BGC. Subsequently, the resulting cell pellets were frozen at -80 °C overnight. Next, the isolation of RNA was conducted using the "NucleoSpin RNA Plus, Mini kit" (Macherey-Nagel), according to the manufacturer's instructions, which is including an RNA purification protocol by the application of DNA removal columns during the isolation procedure. In this case, cell lysis was performed applying the recommended concept of a combined application of the kit for "Genomic DNA from microorganisms" (Macherey-Nagel), using the bead tube cell lysis protocol (without RNAse) and the "NucleoSpin RNA Plus" kit (Macherey-Nagel). For this purpose, the obtained frozen cell pellets were resuspended in buffer RA1 ("Genomic DNA Kit") and cell lysis was performed using the recommended bead tubes (Type B, "Genomic DNA kit"). However, due to the subsequent application of the lysate in RNA isolation, the addition of RNAse was omitted during the total process. After lysis, the RNA isolation procedure was continued by transferring the homogenized cell-lysate to the gDNA removal column ("RNA Plus Kit") and following the manual's instructions, respectively. The final elution step was performed with 2x 30 µl RNase-free ddH_2O. Subsequently, reverse transcription of the isolated RNA was performed according to the "ProtoScript® II First Strand cDNA Synthesis Kit" (NEB), following the manufacturer's instructions. The resulting cDNA was analyzed by PCR, with primers (Table 10) targeting for the amplification of *sfp* (P57 and P58, 675 bp) and *crpA* (P47 and P48, 489 bp), *crpB* (P49 and P50, 569 bp), *crpC* (P51 and P52, 397 bp) and *crpD* (53 and P54, 443 bp). As positive control the ubiquitously expressed housekeeping gene *acs1*, encoding for the acetyl-CoA synthase, was targeted by PCR amplification (P55 and P56, 574 bp) to verify a successful RNA isolation and reverse transcription procedure. In parallel, the procedure was conducted with A/T1-RNAse Mix (ThermoFisher Scientific)-treated RNA of *S. cerevisiae* ATi01. This negative control served to verify the origin of the PCR amplification products from RNA-generated cDNA and not from remnants of co-isolated genomic DNA.

Table 10. Primers of transcription-profiling in heterologous cryptophycin biosynthesis.

Primer	Sequence
P47	5' – ATGATTACACCTTCACATG – 3'
P48	5' – CTGCTCTGATTGAAATAGT – 3'
P49	5' – ATGGACATGAATATTAATAGG – 3'
P50	5' – CCTGATACAACATAAGCATC – 3'
P51	5' – ATGAATTTAATCGAATTTTTACA – 3'
P52	5' – AATTATTTACCTGTTGGATTG – 3'
P53	5' – CTCTTCAATTTATCCACTTTC – 3'
P54	5' – TGGTGATGACTCCAGATAAAT – 3'
P55	5' – ATGTCGCCCTCTGCCGTA – 3'
P56	5' – CAGTATCGCCCTTGCGAAC – 3'
P57	5' – ATGAAGATTTACGGAATTTATATGG – 3'
P58	5' – TTATAAAAGCTCTTCGTACGAG – 3'

3.8.4 Extraction of cryptophycins from *S. cerevisiae* and metabolite analyses

The expression cultures were subjected to metabolite extraction, followed by chemical analyses of the extracts for the biosynthesis of cryptophycin derivatives. For this purpose, the cells were harvested at room temperature at a speed of 4000 rpm for 15 min. In analogy to the established extraction procedure for cryptophycin 1 (cf. section 3.4), the cell pellet was frozen for 24 h at -80 °C. Subsequently, the cell pellet was lysed applying a standard glass bead method that was adapted from Conzelmann *et al.* [256]. Accordingly, the cells were resuspended in an equal amount (w/v) of 50 mM Tris-HCL buffer (pH 7.5) and glass beads, followed by vortexing at maximum speed for 2 x 30 min at room temperature to crack the cell wall. The extraction of the pellet lysate was conducted as previously stated for the extraction of cryptophycin from *Nostoc* sp. ATCC 53789 (cf. section 3.4).

The extracts were dried by rotary evaporation and lyophilization. Next, the dried extracts were redissolved in acetonitrile and applied to comparative LC-MS analyses, as described earlier (cf. section 3.4). After the conducted LC-MS analyses, the samples were provided to the LDC for activity testing. For this purpose, a cryptophycin-sensitive cell-line assay was used, which was established in the course of this project consortium on the basis of cryptophycin 1. Based on this assay, the presence of active cryptophycin derivatives can be analyzed in every sample extract.

In case of fluorinated cryptophycins, the raw extract samples (pellet and supernatant) were

directly applied to an HPLC-directed purification and fractionation procedure, which was coupled to a ^{19}F-NMR-assisted analysis for the detection of fluorinated compounds in the total raw extract. For this purpose, all sample raw extracts were analyzed using a Shimadzu HPLC-system (LC-20AD) with a diode array detector (SPD-M20A). The initial separation was performed using an acetonitrile (ACN)/0.1% trifluoroacetic acid (TFA) in water gradient of 10-100% in 20 min on an EC Nucleodur 100-5 C_{18} VarioPrep column (Macherey-Nagel, 21.0 x 125 mm, 5 µm particle size, 4 ml/min flow rate). After lyophilization (24 h), the fractionated samples were submitted to ^{19}F-NMR spectroscopy. The sample fractions showing a fluoro signal in addition to the referenced TFA signal at -76.55 ppm [257] were further purified by HPLC using an isocratic method of 30% ACN/TFA (0.1%) for 20 min with the same column and flow rate specifications.

All ^{19}F-NMR analyses were performed at the Department of Chemistry and Chemical Biology. The samples were dissolved in deuterated chloroform. Subsequent to this perpetual HPLC-^{19}F-NMR-assisted compound identification, the fractions of interest were analyzed by LC-MS analysis.

3.9 Reconstruction of scytonemin biosynthesis in *S. cerevisiae*

In order to reconstruct the scytonemin biosynthesis in *S. cerevisiae*, the three genes *scyA-C* were amplified from isolated genomic DNA of *N. punctiforme* and equipped with yeast-compatible promoters and terminators by cloning them into the prepared subcloning vectors (cf. section 3.7.2). First, *scyA* was integrated into locus 416d on yeast chromosome IV, using CRISPR/Cas9-directed genome editing (cf. section 3.7.3). For this purpose, *scyA* was amplified by PCR using primer P59 and P60 (Table 11) and cloned into pMA-RQ-PC8-T3 (Table 2, Table SI I) by Bsp120I-BglII directed restriction–ligation cloning. The functional *scyA*-expression cassette, including the upstream and downstream inter-fragment homology regions (Figure SI I), was liberated by NotI restriction. The 1 kb locus homology fragments were amplified by PCR. P69 and P70 were applied for the amplification of the upstream homology fragment, while P71 and P72 were used for the downstream homology fragment (Table 11). Together with the plasmid pCAS9-416, *S. cerevisiae* was transformed with the fragment mix following the protocol of Reider Apel *et al.* [156]. Selection was performed on YNB agar with uracil depletion. Subsequently, positive transformants were cured from pCAS9-416 applying the 5-FOA counter selection, followed by selective replica plating. Positive clones, carrying the GOI and lacking the pCAS9 plasmid were identified by PCR, targeting *scyA*.

Successively, *scyC* and *scyB* were integrated into the host's genome, using the pMA-RQ-PA3-T30 (Table 2, Table SI I). *scyC* was amplified by PCR with the primers P61 and P62 (Table 11). *scyB* was amplified using the primers P79 and P80 (Table 11). Both genes were cloned by KpnI-NsiI restriction-ligation into the linearized subcloning vector.

In order to adapt the dissimilar cyanobacterial to yeast codon usage, the start codon of *scyC* (GTG)

was altered to ATG. Subsequently, *scyC* was integrated into locus YORWΔ22/Ty1, applying LTR-recombination (cf. section 3.7.4). For this, selective LTR flanks of the locus were amplified by PCR (P65-P68, Table 11), containing homologous regions with either the promoter PA3 of the *scyC*-expression cassette (LTR flank upstream, P65 and P66), or the Hyg^R terminator (LTR flank downstream, P67 and P68). The selective marker Hyg^R was amplified by PCR (P63 and P64, Table 11) from pYTK079, including a homologous region with the T30-terminator of the *scyC*-expression cassette. The *scyC*-expression cassette was liberated from pMA-RQ-PA3-T30 by NotI restriction. Next, *S. cerevisiae-scyA* was transformed with this fragment set and screened for positive transformants on YPD agar containing hygromycin. Verification of a successful integration was assessed by genomic PCR of the isolated genomic DNA from the selected yeast clones, targeting *scyC*.

In case of *scyB*, the LTR method was expanded. Instead of LTR sites, the integration of *scyB* was coupled to the knock-down of *bna2* (locus YJR078W, chromosome X), which is involved in kynurenine and anthranilate formation, encoding the biosynthesis of the enzyme tryptophan 2,3-dioxygenase.

The functional *scyB*-expression cassette was liberated from the subcloning vector by TauI restriction. The homologous flanks, targeting the *bna2* locus, were amplified by PCR using the primers P73-P76 (Table 11). The upstream flank of 355 bp (coding sequence 189-543 bp of *bna2*) was amplified by P73 and P74 and included a homologous region with the promoter PA3 of the *scyB*-expression cassette. The downstream flank of 290 bp (coding sequence 1046-1335 bp of *bna2)* was amplified by P75 and P76 and contained a homologous sequence with the Kan^R terminator. The Kan^R (G418) selective marker sequence was amplified from pYTK077, using the primers P77 and P78 (Table 11), which included a homologous overlap with the T30-terminator of the *scyB*-expression cassette. The resulting fragments were used for the transformation of *S. cerevisiae-scyA-scyC*. In order to maintain the selective pressure for *scyC*-Hyg^R during the recombination event, the transformants were screened for the positive integration of *scyB*-Kan^R on YPD agar containing G418 and hygromycin. Verification of a successful integration and disruption of *bna2* was assessed by PCR of the isolated genomic DNA from the selected yeast clones, targeting *scyB* by a combination of the primers P80 and P81 (Table 11). P81 is binding to a region 174 bp upstream of the integration locus for amplification. Furthermore, the resulting strain, *S. cerevisiae-scyA-scyB-scyC*, was again tested for the presence of *scyA* and *scyC* by genomic PCR.

The primers used in this study are listed in Table 11. For cloning purposes restriction recognition sites and homologous overlapping regions were integrated by PCR amplification, through the addition of the sequences into the primer tails. Codon usage was also adapted during PCR

amplification.

Table 11. Primers of scytonemin biosynthesis reconstruction. Sequences annealing to the DNA template are indicated in capital letters, while restriction recognition sites are symbolized by italics lower case letters. Primer tails including homologous regions are underlined. The start codon of *scyC* was adapted to yeast codon usage and is indicated in red bold capital letters.

Primer	Sequence
P59	5' – ctgcagaa*gggccc*ATGAGTCAAAACTATACTG – 3'
P60	5' – ggagga*agatct*TCAAACCATTGGAAATGAAA – 3'
P61	5' – ccacca*atgcat***ATG**GAAAAAAATACTTTTGCAAC – 3'
P62	5' – ggcggc*ggtacc*TTAGTTGGGAACTAGGGAT – 3'
P63	5' – aactaaaaaaatataaggaagaaaaaaatagctaatttttccggcagaaGTTTAGCTTGCCTCGTCCC – 3'
P64	5' – CAGTATAGCGACCAGCATTCA – 3'
P65	5' – CACCGGAGCTTGGATATGAT – 3'
P66	5' – cacaaccgacgatccggggtcatggcggctatttttttttttttttttttTTCGCGGGCTGTTACTTATC – 3'
P67	5' – agaaagtaatatcatgcgtcaatcgtatgtgaatgctggtcgctatactgGGACCAACTATCATCCGCTAA – 3'
P68	5' – CGTGATAAACGATCGCCATA – 3'
P69	5' – CCAACTGCATGGAGATGAGTC – 3'
P70	5' – GGGTCCGGTTAAACGGATC – 3'
P71	5' – CCGAACATGCTCCTTCACTATTTTAAC – 3'
P72	5' – CAATTGAGGAAACTTGAAAGGTGTGG – 3'
P73	5' – GCAAGAGGATAACAAGGTGC – 3'
P74	5' – ttcttttataacttgaattacaacgttacaagttaattaacaccgcccccCAACTTTCGTCAACGGTTCC – 3'
P75	5' – agaaagtaatatcatgcgtcaatcgtatgtgaatgctggtcgctatactgCTCGTAACGCTTCTAACCGTGC – 3'
P76	5' – CGCAGTGGCGACAGTCTCG – 3'
P77	5' – tatatatatatataaactcatttacttatgtaggaataaatttttttcaaaTTAGCTTGCCTCGTCCCCG – 3'
P78	5' – CAGTATAGCGACCAGCATTCACATA – 3'
P79	5' – ccacca*atgcat*CATGCTGCTATTTGAAACTGTTAG – 3'
P80	5' – ggcggc*ggtacct*CCACTTCTCACTCCTTAAGATA – 3'
P81	5' – CATAACCGGACCACAAGTAC – 3'

3.9.1 Heterologous expression of the scytonemin pathway

For heterologous expression precultures of *S. cerevisiae-scyA-scyC* and *S. cerevisiae-scyA-scyB-scyC* were incubated in YPD at 30 °C and 180 rpm overnight. The resulting seed cultures were used to inoculate the main expression cultures to an OD_{600} of 0.04, which were continuously cultivated for three days at 15 °C and 180 rpm. In case of *S. cerevisiae-scyA-scyC* an additional main culture was prepared in parallel, supplementing the medium with L-tryptophan (50 mg/l), while its congener *S. cerevisiae-scyA-scyB-scyC* was also cultivated in the presence of L-tryptophan (50 mg/l) and a combination of L-tryptophan and L-tyrosine (50 mg/l each).

3.9.2 Extraction of scytonemin from *S. cerevisiae* and metabolite analyses

In order to analyze the metabolic profile of the transgenic strains and to assess the production of scytonemin, the constructed *S. cerevisiae* strains (i.e. *S. cerevisiae-scyA-scyC*, *S. cerevisiae-scyA-scyB-scyC*) were compared with *S. cerevisiae* CEN.PK2-1C. Following a 60-h incubation of the (transgenic) yeast strains, the cultures were harvested by centrifugation at room temperature and 4000 rpm for 15 min. The obtained supernatant and pellet fractions were extracted and analyzed for the formation of scytonemin and pathway related products, which were previously described by Malla *et al.* [93]. The extraction was performed as described earlier (cf. section 3.8.4 and 3.4). A literature survey revealed the usage of various solvents, e.g. acetone, ethanol, ethyl acetate and methanol, which have been successfully used to isolate scytonemin and its pathway intermediates from different sources before, including native and heterologous producers [93,94,258,259]. Therefore, different extraction solvents were tested in this PhD project. In the initial study, analyzing *S. cerevisiae-scyA-scyC*, methanol (MeOH) was used for the extraction of the pellet lysate and the XAD-bound compounds of the supernatant broth. In the consecutive study of *S. cerevisiae-scyA-scyB-scyC*, the solvents dichloromethane (DCM) and MeOH were used to extract the pellet and the supernatant fraction. For this purpose, first the fractions were extracted with DCM, followed by a second extraction with MeOH. In a parallel approach, a culture of *S. cerevisiae-scyA-scyB-scyC* was extracted with acetone instead of MeOH.

Subsequent to extraction, the metabolic profiles were evaluated by comparative HPLC analyses. Metabolites that were not observed in the control, were purified and subjected to further analyses and structure elucidation by LC-MS and NMR.

In general, all sample extracts were analyzed using a Shimadzu HPLC-system (LC-20AD) with a diode array detector (SPD-M20A). The raw extracts were initially separated using an acetonitrile (ACN)/0.1% trifluoroacetic acid (TFA) in water gradient on an RP EC-NUCLEODUR PFP column (Macherey-Nagel, 4.6 x 250 mm, 5.0 µm particle size, 0.5 ml/min flow rate). The time parameters of the separation were set to 0-5 min at 25% ACN and 25-42 min at 75% ACN, followed by a re-equilibration to 25% ACN. For the extract of *S. cerevisiae-scyA-scyC*, a succeeding HPLC purification using an RP VP-NUCLEODUR C18 Gravity column (Macherey-Nagel, 10x250 mm, 3 µm, 2 ml/min flow rate) was conducted for a newly detected compound. The latter exhibited a distinctive signal at 334 nm and increased in signal intensity upon tryptophan supplementation (cf. section 4.2). A gradient method of ACN/TFA was executed with 0-5 min at 5% ACN, 25-40 min at 100% ACN and 43-50 min at 5% ACN. The final purification at 334 nm was achieved by application of an RP VP-NUCLEODUR C18 Isis column separation (Macherey-Nagel, 10 x 250 mm, 5.0 µm particle size, 2 ml/min flow rate). The ACN/TFA gradient method was applied from 0-5 min at 5% ACN, via 60% ACN at 30 min, up to 100% ACN at 32-34 min and a final re-equilibration

to 5% ACN at 35-38 min.

Following this, an LC-MS analysis and NMR-guided structure elucidation were performed for the obtained purified compound from *S. cerevisiae-scyA-scyC*. The conducted LC-MS analysis was carried out using an Agilent 1260 Infinity LC system connected to a UV detector and a 6120 quadrupole mass spectrometer (Agilent, Santa Clara, USA). Ionization was performed by electrospray ionization (ESI). The analysis was performed in positive ion mode. The HPLC flow rate was adjusted to 0.4 ml/min using an RP EC-Nucleodur C18 column (100 x 2 mm, 2.7 µm; Macherey-Nagel). The MeOH-dissolved samples were analyzed applying a gradient from 5 to 98% ACN/0.1% formic acid in water for 10 min, followed by 5 min at 98% ACN.

A subsequent study investigated the metabolic profile of *S. cerevisiae-scyA-scyB-scyC*, administering the same HPLC specifications as described for the previously used PFP column. The initial separation of the raw extracts (XAD7-DCM and XAD7-MeOH extracts), resulted in the identification of several new metabolites for the transgenic strain in the XAD7-DCM extract-fractions (cf. section 4.2), which were consequently purified, by a repetitive usage of the identical column and parameter specifications. The obtained purified compound fractions were directly analyzed by NMR spectroscopy.

The XAD7-acetone extracts were analyzed by HPLC in comparison to the control strain and a commercially acquired indole-3-pyruvate standard. Subsequently, the extracts were directly applied to comparative NMR analysis, recording differences between the metabolic profile of *S. cerevisiae-scyA-scyB-scyC* and the control strain. In this case, a preceding HPLC-based fractionation of the extracts was not conducted.

In order to elucidate the final structure of the purified compounds, the sample extracts were dissolved in MeOH-d_4 and subjected to NMR analyses.

4 Results and Discussion

In the following chapters, the obtained results of the conducted experiments are presented and analyzed. This covers the investigation of the cyanobacterial strain *Nostoc* sp. ATCC 53789 from a genomic perspective and an inspection of its biosynthetic potential. Furthermore, the heterologous production of two cyanobacterial secondary metabolites, scytonemin and cryptophycin, and their biosynthetic intermediates is evaluated in different hosts and systems.

4.1 Cryptophycin biosynthesis and *de-novo* genome sequencing of *Nostoc* sp. ATCC 53789

4.1.1 Results

The following section contains adapted passages from the author's previous publications [I] and [III] [223,260].

4.1.1.1 Production of cryptophycin by *Nostoc* sp. ATCC 53789

Genetic and biosynthetic information about the cryptophycin producing cyanobacterium *Nostoc* sp. ATCC 53789 are scarce. Although genes involved in cryptophycin biosynthesis were identified some years ago, knowledge on their integration into the metabolic network of the producing strain is limited [223,260]. This lack of information also applies to the overall biosynthetic potential of the strain, which apart from cryptophycin [64] and nostocyclopeptides [73] has not been systematically analyzed. In order to complement the available data and estimate the genetic potential of the strain for the biosynthesis of secondary metabolites, the genome of *Nostoc* sp. ATCC 53789 was subjected to a *de-novo* sequencing analysis.

Initially, cultures of *Nostoc* sp. ATCC 53789 were tested for the production of cryptophycin in order to verify the functional expression of the corresponding BGC. The obtained culture extracts were analyzed by LC-MS analyses. A detected molecular ion of *m/z* 655.2 $[M+H]^+$ indicated the biosynthesis of cryptophycin 1 (Figure 13). Testing of the culture extracts against selected cancer cell lines revealed strong antiproliferative activities, thus corroborating the results from the chemical analysis (data not shown). Noteworthy, the production of cryptophycin 1 was no longer detectable after continuous cultivation of *Nostoc* sp. ATCC 53789 for 6 month.

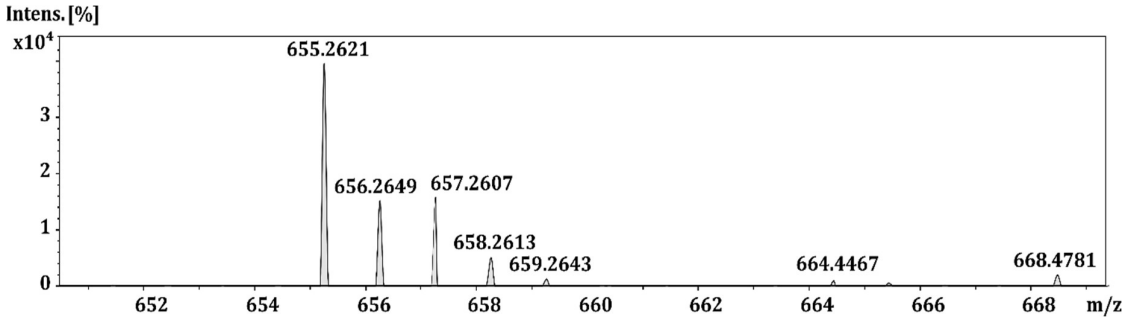


Figure 13. ESI spectrum of cryptophycin 1 from culture extracts (1 l in total) of *Nostoc* sp. ATCC 53789, which was cultivated for 4 weeks. The spectrum was recorded in positive ion mode. The detected compound mass of *m/z* 655.2 $[M+H]^+$ indicated the biosynthesis of cryptophycin 1.

4.1.1.2 Biosynthetic potential of *Nostoc* sp. ATCC 53789

In order to assess its secondary metabolome, the genomic DNA of the cryptophycin-producing strain was isolated and subjected to genome sequencing at the IIT Biotech GmbH. "The genome of *Nostoc* sp. ATCC 53789 was reconstructed from short- and long-read DNA data sets obtained by Illumina and Nanopore sequencing. This approach revealed 13 discrete replicons with a total size of 8.7 Mb and a medium G+C content of 40-42%. In addition to one circular chromosome of 7.34 Mb (Figure 14), the replicons comprise two linear and ten circular plasmids, ranging in size from 34.75 kb to 337.07 kb (Table 12). The automated genome annotation resulted in the assignment of 7408 genes, 7300 protein-coding sequences, 88 tRNAs, 12 rRNAs and 8 ncRNAs. Overall, the genomic features of *Nostoc* sp. ATCC 53789 were found to be consistent with other members of this genus in respect to genome size and number of replicons." With a total of 13 replicons the strain exhibits the second largest number of extrachromosomal replicons within the genus *Nostoc*, only surpassed by its closest taxonomic relative *Nostoc* sp. C057 (Accession No. CP040281.1, symmetrical genome identity of 76%, Figure SI X), which possesses 16 replicons [261–263].

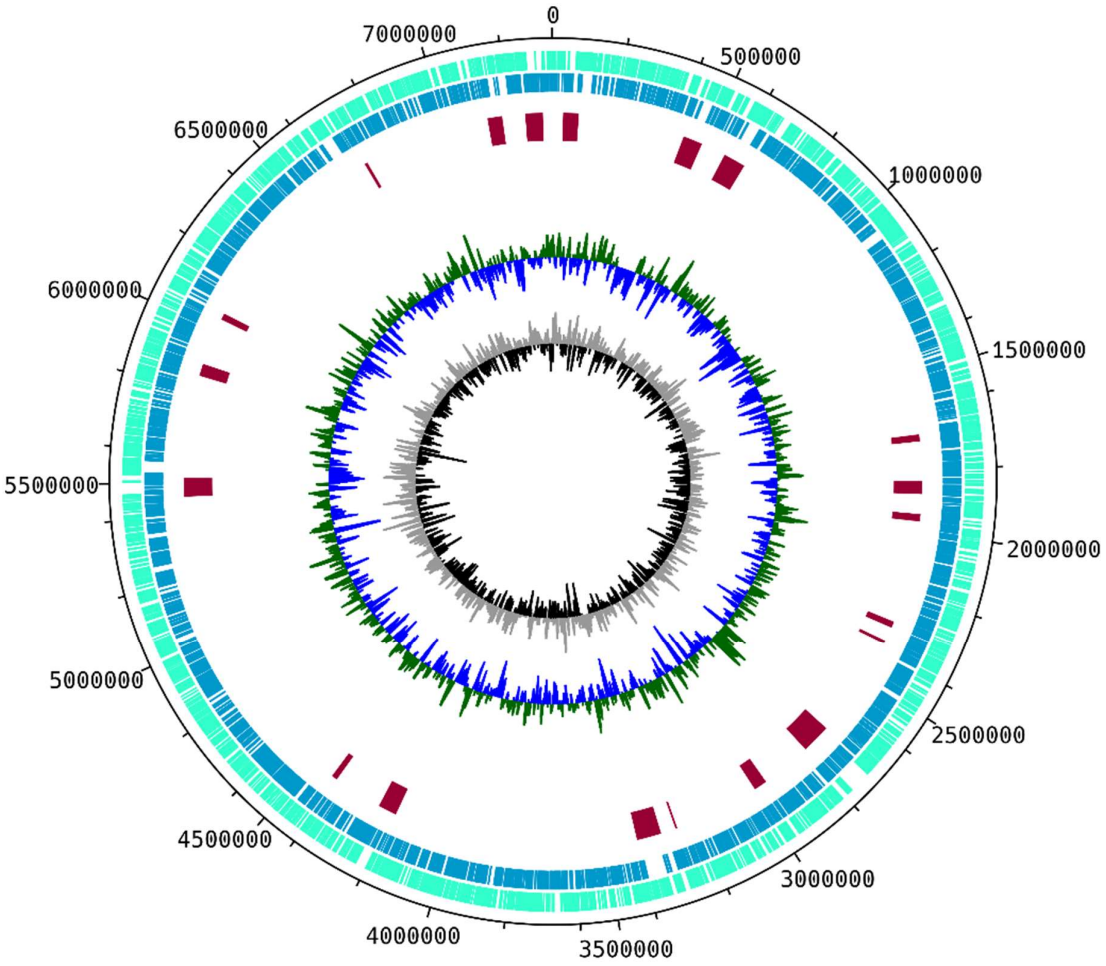


Figure 14. Chromosome map of *Nostoc* sp. ATCC 53789. The outer size-scale is given in 0.5 Mb intervals. Notation: Circle 1 (mint), forward strand open reading frames; Circle 2 (blue), reverse strand open reading frames; Circle 3 (red), secondary metabolite biosynthetic gene clusters detected by antiSMASH; Circle 4, G+C content (dark green, above average; dark blue, below average); Circle 5, G+C skew (grey, above average; black, below average). The genome map was designed using DNAPlotter. [241,260]

An inspection of the genome with antiSMASH 5.0 and 6.0 [240] revealed a high diversity of BGCs with a unique combination of in total 24 BGCs and a characteristic profile of cyanobacterial secondary metabolite types. Overall the high diversity, the detected types of secondary metabolite clusters and thus the biosynthetic potential of *Nostoc* sp. ATCC 53789, is similar to those of taxonomically related strains, e.g. *Nostoc* sp. C057, the *Peltigera membranacea*-cyanobiont *Nostoc* sp. N6 (Accession No.: CP026681.1), or *N. punctiforme* ATCC 29133 (Accession No.: CP001037).

Table 12. Genomic features of *Nostoc* sp. strain ATCC 53789.

Replicon	Size [bp]	Topology	G+C content [%]	Accession No.
chromosome	7,340,101	Circular	41.4	CP046703
pNsp_a	337,072	Circular	41.6	CP046704
pNsp_b	325,114	Circular	40.9	CP046705
pNsp_c	219,529	Circular	41.0	CP046706
pNsp_d	65,222	Linear	41.8	CP046707
pNsp_e	57,504	Circular	42.5	CP046708
pNsp_f	56,077	Linear	41.7	CP046709
pNsp_g	54,032	Circular	42.3	CP046710
pNsp_h	49,561	Circular	39.6	CP046711
pNsp_i	40,105	Circular	41.2	CP046712
pNsp_j	38,437	Circular	41.4	CP046713
pNsp_k	36,221	Circular	40.7	CP046714
pNsp_l	34,754	Circular	42.0	CP046715

In detail, the predicted BGCs in *Nostoc* sp. ATTCC 53789 are involved in the assembly of different natural product classes, including nonribosomal peptides, polyketides, PKS-NRPS hybrid-compounds, terpenes and RiPPs (ribosomally synthesized posttranslationally modified peptide), e.g. bacteriocins, lantipeptides, lasso peptides, or linear azol(in)e-containing peptides (Table 13). "The majority of these loci reside on the chromosome (Figure 14 and Table 13), while four BGCs are located on plasmids" [260], a genetic constitution that is not unusual among cyanobacteria [264]. Each of the four largest plasmids was found to harbor one BGC (Figure 15 and Table 13). Three of the extrachromosomal BGCs are assigned to the biosynthesis of unidentified RiPPs (pNsp_a, pNsp_b, pNsp_d). The only exception is the cryptophycin BGC, which resides on pNsp_c.

Table 13. Secondary metabolome of *Nostoc* ATCC 53789. The *de-novo* sequencing approach revealed the distribution, locus and size of biosynthetic gene clusters in the genome of *Nostoc* sp. ATCC 53789. The 24 BGCs were detected using antiSMASH v. 5.0 and v. 6.0. Further, antiSMASH was used to predict their corresponding type of secondary metabolite clusters, product, including the similarity of the BGC to known database entries. Notation: nt, nucleotide; kb, kilo base pair; n.i., not identified; hglE-KS, heterocyst glycolipid synthase-like PKS; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; RiPP-LAP, ribosomally synthesized posttranslationally modified-linear azol(in)e-containing peptides; RiPP, ribosomally synthesized posttranslationally modified peptide.

No.	Replicon	Locus [nt]	Size [kb]	Type	Predicted product and database similarity
1	chromosome	32262-82288	50.027	hglE-KS	heterocyst glycolipid (100%)
2	chromosome	422953-485535	62.583	NRPS	nostocyclopeptide (100%)
3	chromosome	565089-636478	71.390	NRPS	aeruginoside (41%)
4	chromosome	1686632-1708851	22.220	RiPP-LAP	n.i.
5	chromosome	1834244-1876803	42.560	NRPS-like	n.i.
6	chromosome	1940618-1963941	23.324	RiPP	bacteriocin-like, lathipeptide, n.i.
7	chromosome	2292921-2313706	20.786	terpene	hexose-palythine-serine/ hexose-shinorine (28%)
8	chromosome	2360082-2370315	10.234	RiPP	bacteriocin-like, n.i.
9	chromosome	2696002-2789507	93.506	NRPS-PKS	merocyclophane (22%)
10	chromosome	2952035-2993696	41.662	NRPS-like	aeruginoside (17%)
11	chromosome	3319988-3401502	81.515	NRPS-PKS	nostophycin (27%)
12	chromosome	3995007-4036689	41.683	RiPP	lathipeptide (n.i.)
13	chromosome	4184772-4244316	59.545	NRPS-PKS	anatoxin (20%)
14	chromosome	4401435-4420197	18.763	terpene	n.i.
15	chromosome	5455333-5514751	59.419	NRPS	anabaenopeptin (100%)
16	chromosome	5843529-5882982	39.454	terpene	geosmin (100%)
17	chromosome	5929435-5971787	42.353	RiPP	lanthipeptide, cyanobactin (14%)
18	chromosome	6035023-6055853	20.831	terpene	n.i.
19	chromosome	7131152-7178345	47.194	hglE-KS	heterocyst-glycolipid (57%)
20	chromosome	7251830-7307970	56.141	NRPS-PKS	nostocyclopeptide (28%)
21	pNsp_a	305982-330267	24.286	RiPP	bacteriocin-like, lassopeptide, n.i.
22	pNsp_b	32964-56278	23.315	RiPP	bacteriocin-like, lanthipeptid, n.i.
23	pNsp_c	193220-8850	40.311	NRPS-PKS	cryptophycin (100%)
24	pNsp_d	51874-62164	10.291	RiPP	bacteriocin-like, n.i.

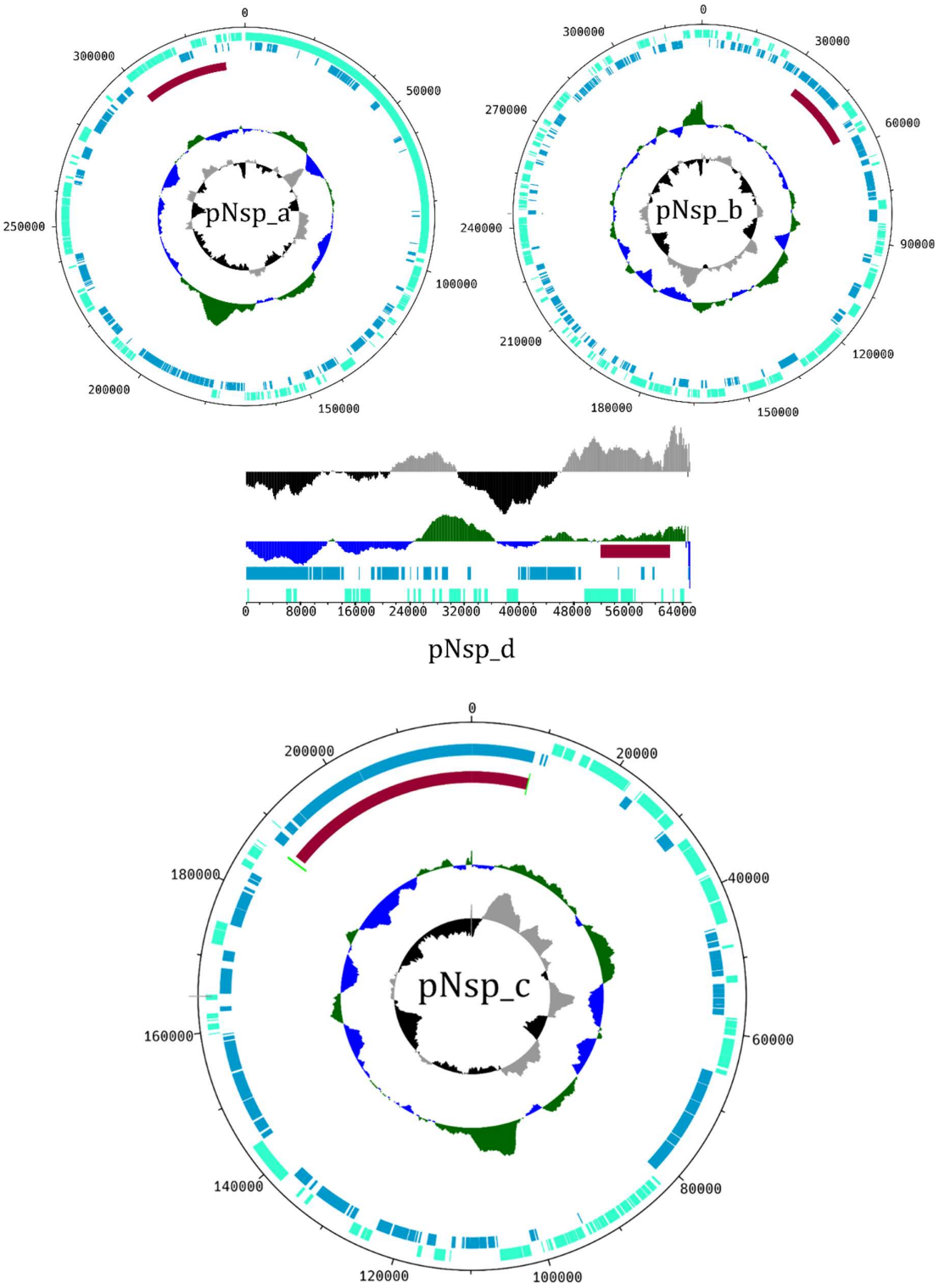


Figure 15. Plasmid maps of the secondary metabolite-encoding extrachromosomal replicons of *Nostoc* sp. ATCC 53789. pNsp_a, pNsp_b and pNsp_c are circular plasmids, while pNsp_d was identified as linear plasmid. pNsp_a, pNsp_b and pNsp_d encode RiPP BGCs. pNsp_c encodes the cryptophycin BGC, which is flanked by transposase genes. The two directly flanking transposases are highlighted in bright green. The outer size-scale is given in 0.05 - 0.008 Mb intervals. Notation: Circle 1 (mint), forward strand open reading frames; Circle 2 (blue), reverse strand open reading frames; Circle 3 (red), secondary metabolite BGCs detected by antiSMASH, i.e. the cryptophycin and the RiPP BGCs; Circle 4, G+C content (dark green, above average; dark blue, below average); Circle 5, G+C skew (grey, above average; black, below average). The genome map was designed using DNAPlotter. [241]

According to the antiSMASH analysis, *Nostoc* sp. ATCC 53789 should be capable to produce further typical cyanobacterial metabolites apart from cryptophycin. This includes the sequiterpene geosmin [265], the non-ribosomal peptide anabaenopeptin [266] and heterocyst glycolipids [267], whereby the latter are exclusive to heterocystous members of the phylum [267]. The corresponding BGCs are highly conserved, as indicated by a 100% similarity score in antiSMASH (Table 13). As many cyanobacterial species produce the phosphatase inhibitor anabaenopeptin [3,266], the biosynthesis of this compound by *Nostoc* sp. ATCC 53789 is not surprising. However, the production of anabaenopeptin was hitherto not known for the cryptophycin producer. In addition, the biosynthetic analysis indicated that the strain is capable of nostocyclopeptide biosynthesis [268], complementing the existing data of Golakoti *et al.* [73] and Becker *et al.* [269], who isolated and identified this metabolite from *Nostoc* sp. ATCC 53789 and characterized its biosynthetic assembly already a while ago [73,269]. The entire genome sequence of *Nostoc* sp. ATCC 53789 now discloses the chromosomal positioning of the nostocyclopeptide locus.

Aside from the mentioned BGCs, the vast majority of the detected loci in *Nostoc* sp. ATCC 53789 could not be associated with the production of known secondary metabolites. The majority of the predicted BGCs are involved in the biosynthesis of peptidic compounds, including RiPPs and NRPS-, or NRPS-PKS-derived metabolites (17 of 24 BGCs). This is consistent with literature reports, which indicate that cyanobacterial secondary metabolism is dominated by peptides (66 % of known secondary metabolites) [23]. Altogether, the herein conducted genome analysis revealed a far greater genomic potential of the strain for the production of secondary metabolites than initially expected by chemical investigations. Especially under the premise of urgently required new anti-infectives to fight modern infectious diseases, the potential of the strain to produce a high, yet unidentified variety of RiPPs (8 of 24 detected BGCs) and NRPS-, or PKS-NRPS-derived compounds (9 of 24 detected BGCs), clearly bears a great chance to enlarge the available toolbox of "therapeutic instruments" by the isolation and identification of new bioactive metabolites. As a matter of fact, both of the prevalently detected secondary metabolite classes are renowned to comprise promising anti-infective candidates for treating e.g. viral, protozoal, fungal and microbial infectious diseases that are caused by (multidrug-resistant) pathogens [23,34,270–272].

4.1.1.3 Inspection of the cryptophycin BGC

In order to further complement the existing information on the plasmid-associated cryptophycin locus of *Nostoc* sp. ATCC 53789, a closer inspection of the data was conducted under the premise of (i) an assessment of its broader genetic context on the plasmid, (ii) the integrity and functionality of the associated domains with a focus on the loading and substrate-selecting domains of CrpA and (iii) genetic indications for an enzymatic truncation of a hypothetical starter unit.

(i) Apart from its plasmid-associated localization, the comprehensive review of the obtained annotated sequencing data in combination with a manual SnapGene-BLASTx-coupled analysis of the BGC neighborhood revealed that the cryptophycin locus is flanked by several upstream and downstream transposase genes. An additional transposase was identified to be located within the BGC itself (compare Figure SI III). This "BGC-internal" transposase (locus tag: GJB62_RS34220) is positioned in the gap between the biosynthetic genes *crpG* and *crpH*, respectively. The localization of the cluster in a potentially highly mobile genomic area of the plasmid was not known so far and, apart from complementing the existing data of its genetic neighborhood, may shed light on the observed loss of cryptophycin production during the elongated fermentation of *Nostoc* sp. ATCC 53789. Therefore, it is reasonable that the transposase-near localization of the BGC may have contributed to the observed production loss over time (cf. section 4.1.2.1).

(ii) Consistent with the previously published data of the assembly line [64], the antiSMASH-based inspection of the cluster confirmed the integrity and activity of the single cluster domains. As CrpA was a central object of the conducted experiments in this doctoral thesis, a subsequent manual investigation of the specificity-conferring code of the encoded loading A domain and the substrate-selecting AT domains was also conducted. These manual analyses were based on general findings of Stachelhaus *et al.* [41], Minowa *et al.* [242], Yadav *et al.* [243] and Reeves *et al.* [31], as well as on already existing specificity-code information that were identified by Magarvey *et al.* [64]. Consequently, the CrpA-A domain was aligned to the A domain of the microcystin synthetase McyG, which was presupposed to contain an identical active site motive [64]. The analysis reconfirmed the integrity of the active site motive V G V W V A A S G K, attesting the domain's functionality (compare Figure SI V). In parallel the AT domain of CrpA was analyzed using the compiled information of Minowa, Yadav and Reeves, which likewise confirmed the functional integrity of the substrate-selecting AT domain in CrpA, showing the typical active site motive Q Q G H S G R H V with the catalytic S92 and a selectivity for malonyl-CoA (compare Figure SI IV).

(iii) The unresolved one-carbon truncation of the CrpA-selected phenylpropanoid substrate remains a subject of ongoing investigations. In order to identify genetic indications for the observed truncation of the starter unit, the vicinity of the genetic locus was screened for additional *orfs* that could be associated with substrate or intermediate modifications. Hence, the obtained annotated sequencing data of pNsp_c were reviewed again and additionally analyzed by antiSMASH and BLAST. In close proximity of the cluster only transposase encoding genes and one unidentified *Nostoc*-specific hypothetical protein (locus tag: GJB62_RS33410) were detected. While the function of the transposase genes can be assumed as unrelated to the observed substrate truncation, the role of the hypothetical protein still remains elusive, as no homolog with a known function could be identified via BLAST analysis. Yet, if the gene was involved in this

unusual truncation, a homolog would have been expected from the microcystin BGC of *M. aeruginosa*, which shows a similar behavior of substrate truncation by McyG. However, BLASTn, BLASTp and BLASTx-analyses against the *M. aeruginosa* genome and proteome did not reveal any likely homologs. Consequently, the protein, although specific to *Nostoc* spp., is supposed not to be involved in a truncation of the starter unit. Apart from these closely located genes, another small locus was identified ~2.2 kb upstream of *crpA*. This locus comprises a putative NAD(P)-dependent oxidoreductase gene (locus tag: GJB62_RS33420), concomitant to a glycosyltransferase gene (locus tag: GJB62_RS33425) and a preceding response regulator transcription factor (locus tag: GJB62_RS33415). Yet, these enzymes can also be assumed as not related to the observed carbon loss of the starter unit. Altogether, the conducted genetic survey did not point to the presence of any gene within or near the cryptophycin BGC that could be associated with the observed substrate modifications. However, it is still possible that a so far unidentified *trans*-acting modification enzyme, which is encoded elsewhere in the genome of *Nostoc* sp. ATCC 53789, may be involved in the substrate truncation. Another plausible explanation for the observed carbon loss may have been found in a malfunction of any of the PKS-associated AT domains. Therefore, in addition to the CrpA-AT domain the two remaining CrpB-AT domains of the assembly line were inspected for their integrity and functionality using a sequence alignment (compare Figure SI IV). The conducted analysis confirmed the functional integrity of all involved AT domains of the assembly line, thus corroborating the findings of Magarvey *et al.* [64] for a full functionality of all substrate-selecting AT domains and excluding the hypothesis of a truncation due to their dysfunctionality.

Altogether, the conducted bioinformatics analyses of the BGC neighborhood, as well as the relevant loading and substrate-selecting domains of CrpA and CrpB, did not reveal any indications for the observed truncation of the starter unit.

4.1.2 Discussion

4.1.2.1 Loss of cryptophycin biosynthesis

The *de-novo* sequencing of *Nostoc* sp. ATCC 53789 disclosed the intriguing genomic constitution of this cyanobacterial strain and provided valuable insights into its biosynthetic potential. Although the bioinformatics-based analysis of the cryptophycin BGC and its vicinity did not elucidate the mystery of the carbon truncation of the starter unit, the functional integrity of the biosynthetic enzymes and their catalytic domains could be verified. However, the sequencing approach revealed interesting indications, which could be connected with the loss of cryptophycin biosynthesis during the course of an elongated cultivation period. A fact that may be linked to (i) a repression or genetic instability of the BGC that is caused by a general evolutionary adaptation in response to environmental changes, and (ii) the herein detected

genetic constitution of the cluster to be localized in the proximity of transposons. Both of these assumptions could have contributed to the observed instable production of cryptophycin during the fermentation of *Nostoc* sp. ATCC 53789.

(i) Cyanobacteria are known to exhibit a high evolutionary adaptability in response to environmental changes and in fact structural genetic instability (e.g. sequence mutations) is frequently detected in strains that are cultivated under laboratory conditions [109,120]. As the routine maintenance of the strain under laboratory conditions is differing from the natural environmental pressure, or inter-species competition, the repression of an in this case "superfluous" BGC would not be surprising, aiming for the saving of resources. Likewise, mutational changes of the BGC, or the segregational loss of the plasmid would not result in an environmental fitness reduction of the strain, when maintained in the laboratory. In the contrary, the applied "unnatural" selective conditions in the lab may favor the dysfunctionality or loss of the BGC in terms of optimizing resource management. As a result, all of these changes could culminate in the (functional) loss of cryptophycin biosynthesis, giving a reasonable explanation for the observed biosynthetic behavior of the strain. In fact, these assumptions are consistent with the accepted concept of natural product diversity and evolution, whereby the production of a natural product is always subjected to a cost-fitness benefit relation. The biosynthesis of a natural product is increasing the metabolic costs of the cell in order to produce the compound and will only increase the organism's fitness, if the effort is compensated by the acquired benefits. Consequently, it is therefore assumed that non-beneficial, cost-intensive natural product BGCs will be lost from the population [273]. It is reasonable that this basic evolutionary concept could have caused the loss, or repressive regulation of the cryptophycin BGC under laboratory conditions. In perspective, re-sequencing of the strain could reveal probable genetic instability issues that may occur during an elongated cultivation time.

(ii) Another conceivable explanation may be found in the transposons, which are located close to the cryptophycin locus. The BGC was even found to include two transposase genes between *crpG* and *crpH*.

Transposable elements, which self-encode their respective transposase gene, are mobile genetic elements that are able to move themselves within the host's genome by the principle of excision and insertion, using a transposase-directed non-replicative (cut-and-paste), or replicative (copy-and-paste) mechanism. They are abundant to all domains of life and have evolutionary shaped their host's genomes by replicating and relocating in between genomic loci [274–276], contributing to speciation and adaptation to changing environments [277]. Yet, they can deleteriously influence the integrity, functionality, or expression of genes and integration into coding regions may cause detrimental mutations, leading to a functional loss of the genetic information at the target

site [276–278]. Likewise, surrounding genes at the excision locus can be affected by remnants of their characteristically associated flanking direct repeats that are remaining at the original transposon site [278]. Moreover, the mechanism of excision and insertion of transposable elements is known to be imprecise and hence contributing to the destruction of the original genetic locus integrity or the insertional target site [275]. With the high abundance of such mobile genetic elements within and in the vicinity of the cryptophycin locus, the genetic information of the BGC might be located in a highly active and "unsettled" genomic area, which might be prone to undergo over-proportional genetic rearrangements or alternations in transcription regulation by the neighboring transposases. Therefore, it is possible that transposition events negatively affect the BGC's integrity, consequently causing the observed loss of cryptophycin biosynthesis during the elongated cultivation period. Therefore, in perspective a recultivation and a thorough, continuous monitoring of metabolic and genomic changes over time, e.g., by metabolic profiling and PCR, or sequencing of the cryptophycin locus, is necessary to clarify the integrity of the cryptophycin locus, elucidate probable sequence mutations, or detect deviations due to transposase activities. This may reveal interesting information about active evolution processes within the cluster, or genomic changes due to probable transposition events. Hints on probable transposase activities could be found e.g. in an increased number of the same transposon over time within the genome, or the detection of the transposon at a different locus. In this context, RNA sequencing may further provide information on the active expression of the transposable elements.

4.2 Reconstruction of cryptophycin biosynthesis in *S. cerevisiae*

The following section contains adapted extracts from the author's previous publication [II] [13].

4.2.1 Results

In this thesis the industrially established host *S. cerevisiae* was tested for the biosynthesis of the cyanobacterial metabolite cryptophycin, which is assembled by a multimodular hybrid PKS-NRPS system. This work package included (i) the construction of a strain with the genes needed for the biosynthesis of cryptophycins. For this, refactoring of the bacterial polycistronic gene expression system into a monocistronic, yeast-compatible design was necessary. (ii) The functionality of the utilized minimal, synthetic expression system had to be verified in the transgenic strain by the analysis of its transcription profile. Furthermore, the formation of the biosynthesis proteins was assessed by SDS-PAGE analysis. (iii) The production of cryptophycins by the transgenic yeast strain had to be tested. This also included feeding studies as well as analytical analyses, involving LC-MS- and NMR-assisted compound identification.

(i) The conducted reconstruction approach resulted in the construction of the two transgenic strains *S. cerevisiae* ATi01 and *S. cerevisiae* ATi02 (Figure 16), carrying the genomic predisposition for cryptophycin biosynthesis. The expression systems for *crpA-D* were

successfully integrated into locus ARS 308a on chromosome III, YPRCΔ15 on chromosome XVI, YORWΔ22 on chromosome XV and YPRCτ3 on chromosome XVI, respectively. The transgenic strain *S. cerevisiae* ATi02 further contained the expression system for *crpH*, which was successfully integrated into locus ARS 416d on chromosome IV. In addition, all strains contained a genomic copy of *sfp*, encoding the broad spectrum PPTase of *B. subtilis* for the activation of the reconstituted cryptophycin assembly line. The *sfp*-expression cassette, which had previously been reconstructed by Oliver Schiwy (Laboratory of Technical Biochemistry, TU Dortmund University) resides on locus ARS 720 of chromosome VII.

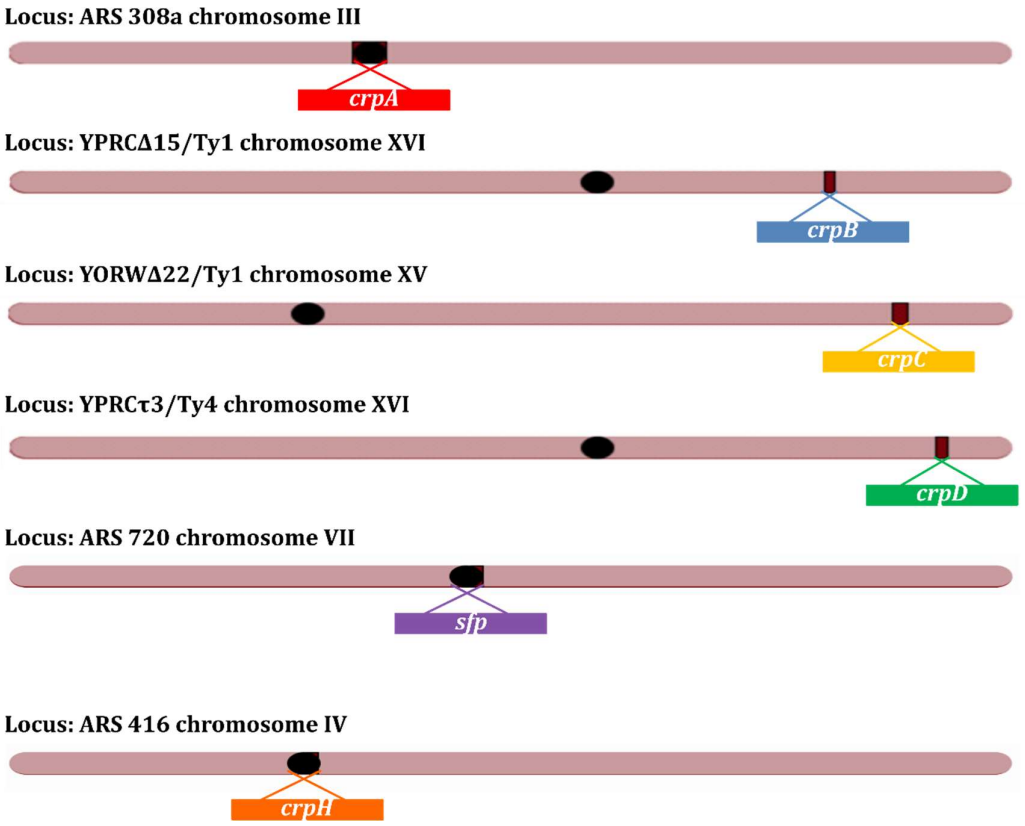


Figure 16. Schematic overview of the genomic map of *S. cerevisiae* ATi01 and *S. cerevisiae* ATi02. Both strains carry the genetic predisposition for the expression of the pathway-activating PPTase gene *sfp* (purple), which was previously integrated into locus ARS 720 on chromosome VII **. Subsequently, the genes of the cryptophycin BGC were inserted into different loci of the yeast's genome, using CRISPR/Cas9 or LTR-assisted genome editing approaches. *S. cerevisiae* ATi01 harbors the cryptophycin biosynthesis genes *crpA* (red) on chromosome III, *crpB* (blue) on chromosome XVI, *crpC* (yellow) on chromosome XV and *crpD* (green) on chromosome XVI. In addition to that, *S. cerevisiae* ATi02 also contains the halogenase gene of the cryptophycin BGC, *crpH* (orange) on chromosome IV, including a minimal synthetic expression system to drive its proper expression. ** The strain, solely harboring the *sfp*-expression cassette, was kindly donated by Oliver Schiwy (Laboratory of Technical Biochemistry, TU Dortmund University).

(ii) Although both strains contain the genetic capacity of cryptophycin biosynthesis, the functionality of the newly designed minimal synthetic expression systems for the cryptophycin biosynthesis genes had to be tested. In order to analyze the transcription of the cryptophycin genes, the RNA was isolated from a culture of *S. cerevisiae* ATi01 and subjected to RT-PCR analysis. The transcription analysis of the generated cDNA for *S. cerevisiae* ATi01 (Figure 17), resulted in a positive PCR product band (Figure 17 A) for *sfp* and *crpA-D*, as well as for the positive control of the house-keeping gene *acs1*. On the contrary, the analyzed negative control, which was treated with A/T1-RNAse during the cDNA-generation procedure, did not give any PCR product in the subsequent analysis (Figure 17 B). The result confirmed that the detected signals indeed originated from isolated mRNA of the transgenic strain. Consequently, the transcription profiling confirmed the functionality of the established minimal synthetic expression system in yeast and demonstrated the transcription of all necessary genes required for cryptophycin biosynthesis in *S. cerevisiae* ATi01 and *S. cerevisiae* ATi02.

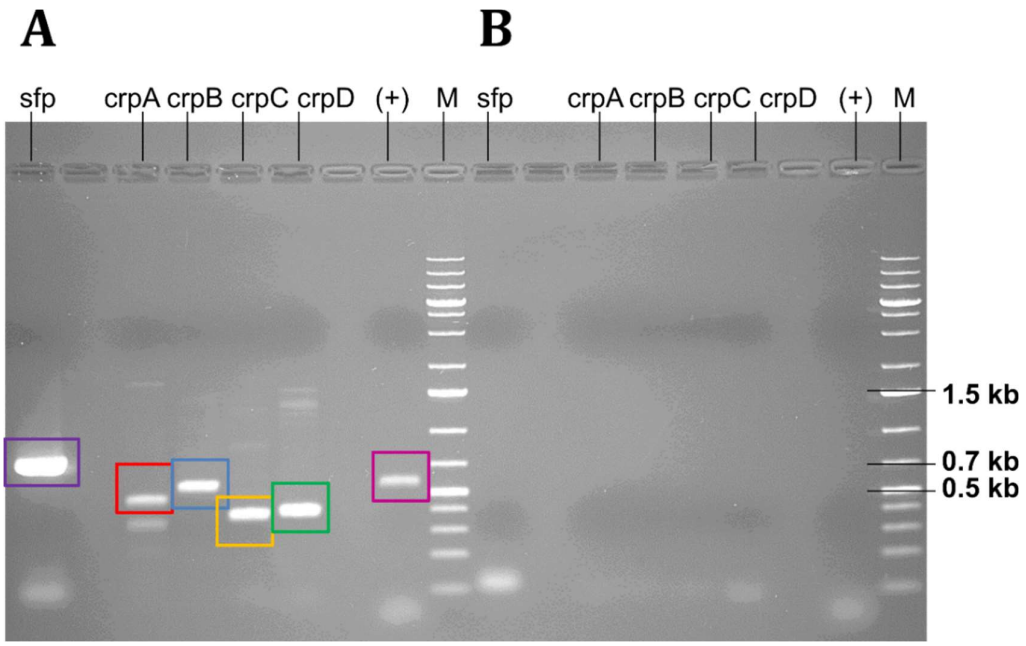


Figure 17. Transcription profile analysis of *S. cerevisiae* ATi01. The obtained cDNA was analyzed for the presence of the cryptophycin biosynthesis genes *crpA* (red), *crpB* (blue), *crpC* (yellow) and *crpD* (green) as well as the pathway-activating PPTase gene *sfp* (purple), using PCR. The *acs1* gene, encoding the acetyl-CoA synthase 1 in *S. cerevisiae* (magenta), served as positive control (A). The negative controls were treated with A/T1-RNAse during the cDNA-generation procedure (B). M, GeneRuler 1 kb Plus DNA ladder (ThermoFisher Scientific).

To analyze the presence of the recombinant biosynthesis enzymes, a comparative SDS-PAGE analysis of *S. cerevisiae* ATi01 and its parental strain lacking the *crp* genes was conducted. The resulting SDS gel (Figure 18) revealed a high abundance and intensity of low- to mid-molecular

weight proteins (up to 117 kDa), while the intensity and quantity of high molecular weight proteins was generally low. Yet, the comparative analysis of the cell pellet lysate (pellet and the supernatant fraction), did not show any detectably new protein signals for the transgenic strain in proportion to the control. Also, none of the expected proteins of the heterologous assembly line, ranging in size from 219.6 kDa up to 383.2 kDa, could be identified. Although a number of protein signals are visible in the range of <117 kDa, the enormous number of proteins of this size impedes the clear identification of any new signals in this area, which could correspond to truncated proteins arising from the assembly line. Also, the generally weak signal intensity of high molecular weight proteins in the range of interest (>171 kDa) hampers a clear conclusion regarding the production of the recombinant cryptophycin proteins in yeast. The experiment was repeated with an increased amount of sample material, using the pellet of a total 1 l culture of the recombinant strain (data not shown). However, the results could not be improved.

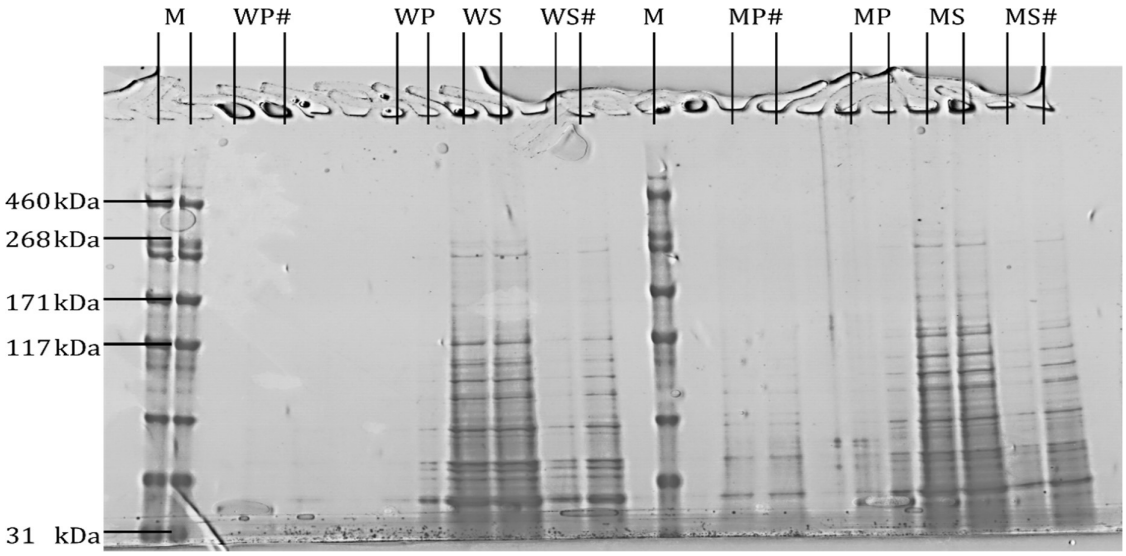


Figure 18. SDS-PAGE analysis of *S. cerevisiae* ATi01. The protein content of the prepared cell lysate supernatant and pellet fraction of the *S. cerevisiae* control strain and the designed transgenic strain *S. cerevisiae* ATi01 were analyzed using a Criterion™ chamber (Bio-Rad) and 3-8% precast Criterion™ XT Tris-Acetate gradient gels. Notation: M, LC5699-HiMark™ (Thermo Fisher Scientific) protein ladder; WP, control strain pellet fraction; WS, control strain supernatant fraction; MP, mutant pellet fraction; MS, mutant supernatant fraction; # 1:10 dilution of the respective sample.

(iii) In parallel to these molecular biological investigations, the transgenic strains were investigated on an analytical scale for the production of cryptophycin derivatives. For this purpose, their metabolic profiles were analyzed by HPLC, LC-MS and NMR spectroscopy.

In accordance with the biosynthetic logic of the partially reconstituted assembly line and the fed substrates, the recombinant yeast strains were expected to produce various derivatives of cryptophycin 1. These analogs are in the following referred to as cryptophycin Ys (Figure 19).

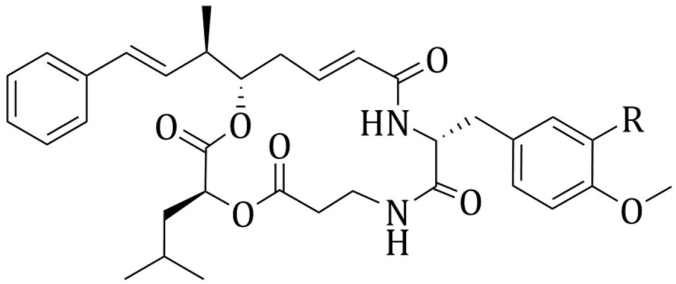


Figure 19. Expected cryptophycin derivatives produced by the transgenic yeasts *S. cerevisiae* ATi01 and ATi02. Three different analogs are expected to be produced by the transgenic strains, with a variable residue (R) in the tyrosine derived unit of the molecule. The derivatives are cryptophycin Y, R = H, cryptophycin Y-F, R = F, cryptophycin Y-Cl, R = Cl. For the biosynthesis of cryptophycin Y-F, the expression culture of *S. cerevisiae* ATi01 was supplemented with 3-fluorotyrosine. The derivative Y-Cl is exclusively produced by the transgenic strain *S. cerevisiae* ATi02, containing the halogenase CrpH.

Comparative LC-MS analyses were conducted, screening the metabolic profile of the recombinant strain *S. cerevisiae* ATi01 and its parental strain (Figure 20). Although a variety of metabolites were detected, none of the recorded chromatograms, neither for the supernatant nor for the pellet extract fraction, revealed any differences in the metabolic profile of the two strains. An extracted ion chromatogram (EIC) of the mass corresponding to cryptophycin Y (*m/z* 591.3065 $[M+H]^+$) did not show any signals. As the metabolic profile of *S. cerevisiae* ATi01 was lacking a detectable signal for the *in vivo* biosynthesis of cryptophycin, *S. cerevisiae* ATi02, expressing CrpH, was analyzed in parallel to facilitate the analytical identification of halogenated cryptophycin derivatives. However, similar results were obtained for *S. cerevisiae* ATi02 (data not shown) and the expected putatively chlorinated cryptophycin Y-Cl (*m/z* 625.2675 $[M+H]^+$). Subsequent testing of the sample extracts in the cell assay of the LDC confirmed the absence of active cryptophycin derivatives.

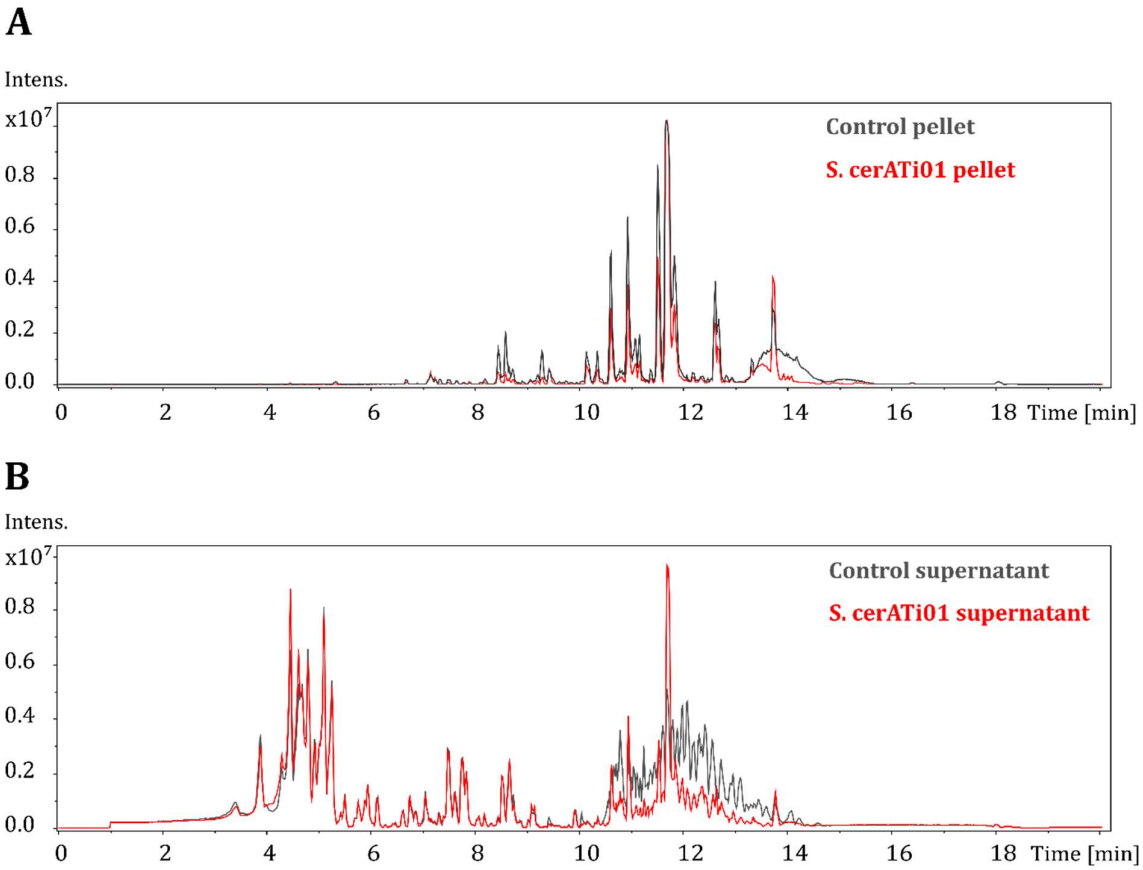


Figure 20. Metabolic profiles of *S. cerevisiae* strains. The sample extracts obtained from *S. cerevisiae* ATi01 (red) and the control strain *S. cerevisiae* CEN.PK2-1C (grey) were subjected to comparative LC-MS analysis. The recorded BPC-MS chromatograms of the pellet (A) and supernatant (B) fractions were analyzed separately in order to identify new metabolites in the transgenic strain.

We then explored the labeling of yeast-derived cryptophycins with fluorine atoms to facilitate their identification in the presence of abundant primary metabolites and media components. Precursor-directed biosynthesis was used for the incorporation of fluorine atoms into the cryptophycin scaffold. For this, cultures of *S. cerevisiae* ATi01 were supplemented with the amino acid 3-fluorotyrosine, which is known to be accepted as alternative substrate by the cryptophycin assembly line [64]. The culture extracts that were obtained from this feeding experiment, were fractionated by HPLC (Figure 21, pellet extract example). Every collected fraction was subjected to ^{19}F NMR analysis. The resulting spectra revealed that fraction 3 contains a fluorine signal, which could correspond to a putative, fluorinated cryptophycin derivative. However, fraction 3 did not represent a single compound and, hence, another HPLC separation was carried out to further purify the sample (Figure 21). The ^{19}F NMR signal was subsequently traced to fraction CDE. The supernatant was investigated in analogy to the pellet fraction (data not shown). To identify new compound signals in the mutant extracts, the results were compared with data from the control strain, *S. cerevisiae* CEN.PK2-1C, which was cultivated and analyzed in the same way.

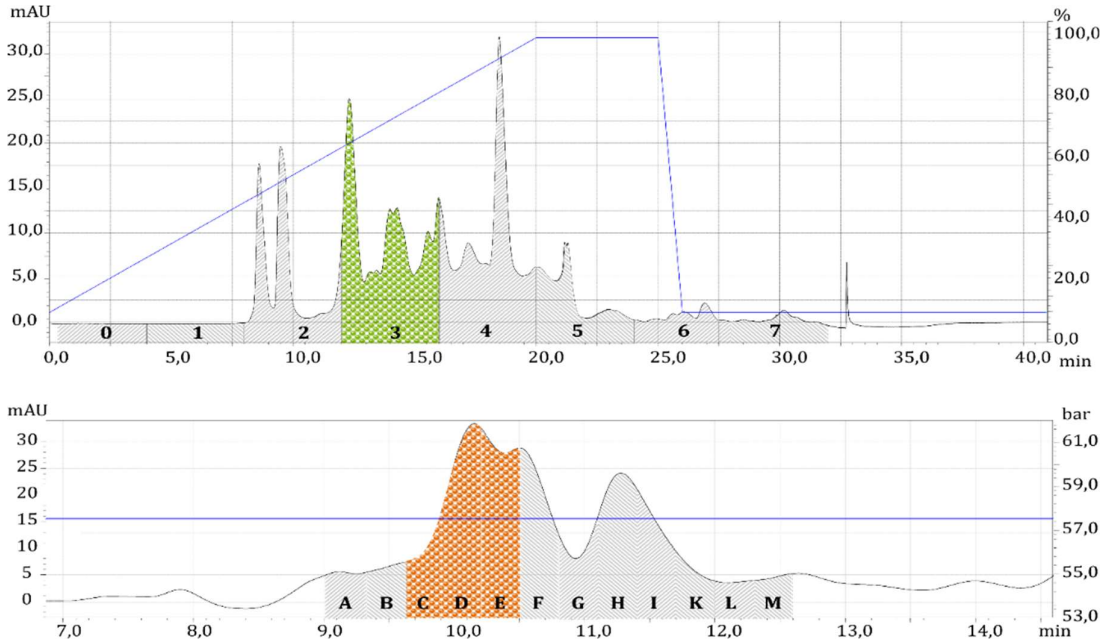


Figure 21. Metabolic profile of a pellet extract from *S. cerevisiae* ATi01 fed with 3-fluorotyrosine. The extract was initially fractionated by HPLC using gradient method (10-100% acetonitrile/TFA, blue line) and an EC Nucleodur 100-5 C_{18} VarioPrep column (top). Fraction 3 (green), which showed ^{19}F NMR signals, was further purified by HPLC using an isocratic method (30% acetonitrile/TFA, blue line) and the same column. Fractions C-E (orange) were positively identified for fluorinated compounds.

The ^{19}F NMR spectra of the different CDE extract fractions (mutant and control strain) are depicted in Figure 22. All spectra showed a signal corresponding to TFA, which had been used during the HPLC purification. The TFA signal was referenced to -76.55 ppm [257]. No difference was observed between the supernatant of the control strain and *S. cerevisiae* ATi01. However, the pellet fraction of the expression strain showed a signal at approximately -74.5 ppm, which was missing in the pellet control. However, a corresponding signal might be found in the control supernatant fraction. Therefore, to clarify the identity of the potentially new signal, the *S. cerevisiae* ATi01 pellet fraction was again analyzed by LC-MS. However, no further evidence was obtained for the presence of cryptophycin Y-F (*m/z* 609.2971 $[M+H]^+$) or cryptophycin Y (*m/z* 609.2971 $[M+H]^+$) (data not shown).

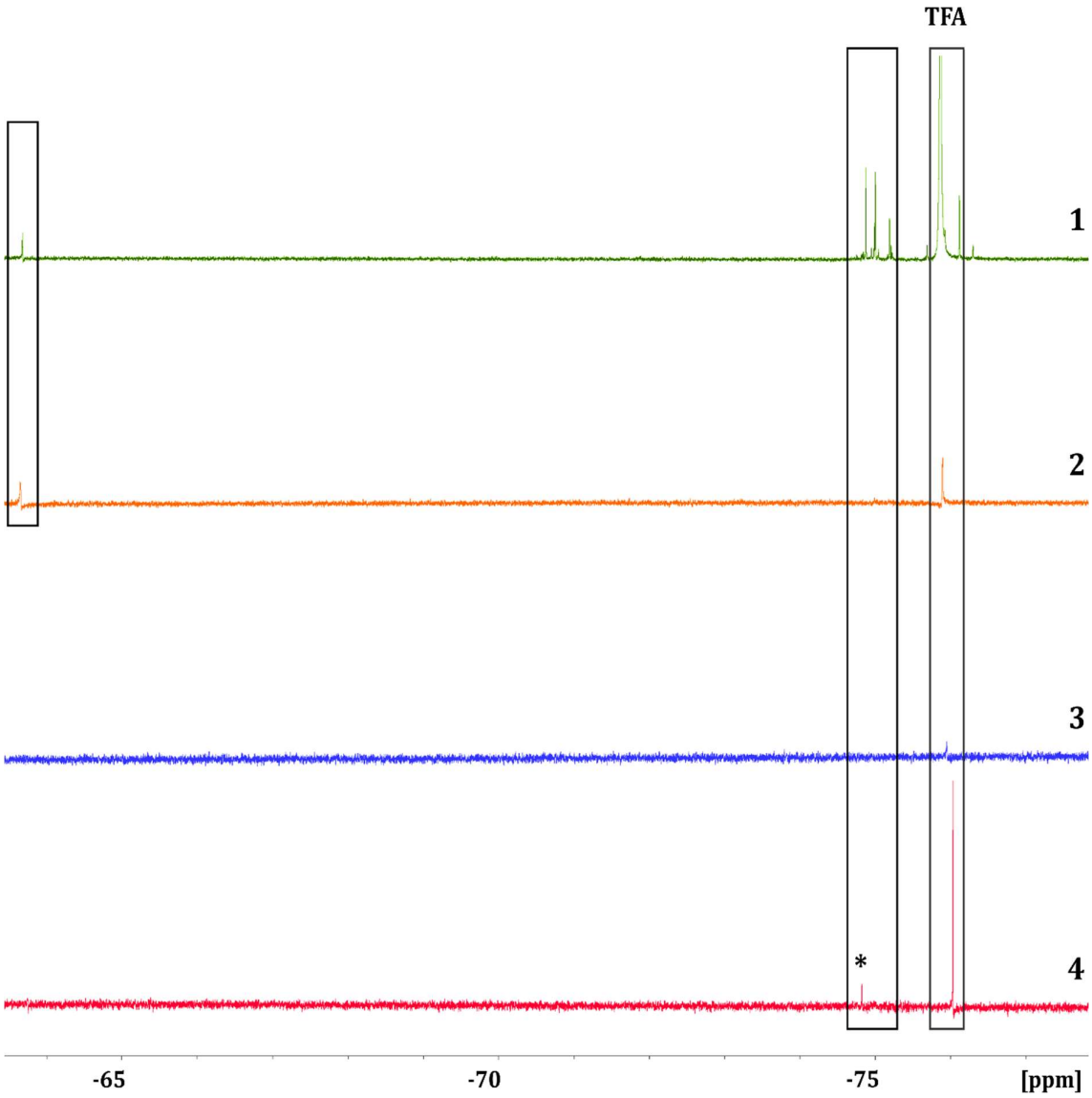


Figure 22. Comparative ^{19}F NMR signal analysis. The pellet and supernatant fractions of *S. cerevisiae* ATi01, which had been fed with 3-fluorotyrosine, were subjected to ^{19}F NMR analysis and compared with the corresponding fractions of a control strain lacking the *crp* genes. A potentially unique signal is indicated by an asterisk. The TFA signal is referenced to -76.55 ppm.

4.2.2 Discussion

The results of the transcription analyses confirm that the minimal synthetic expression units used in this study, successfully direct the heterologous expression of the *crp* genes in yeast. Yet, the presence of the cryptophycin biosynthesis enzymes could not be confirmed by SDS-PAGE analysis, which may indicate a translational issue with multiple causes resulting in low or even no recombinant protein titers. They reach from an increased metabolic burden, unfavorable codon usage and erroneous protein biosynthesis, to misfolding events, overloading of the assisting chaperon system and premature termination events during translation [123,279–281].

In fact, it is a known phenomenon that high level heterologous gene expression, encoding resource extensive proteins, is often concomitant with inclusion body formation, increased amino acid misincorporation, erroneous protein formation, and a higher number of misfolding events [123,280]. In addition, yeast is known to rapidly adapt to intracellular stress conditions, such as an increased metabolic burden due to elevated resource demands, which can result in the down-regulation of protein biosynthesis to match this intracellular burden [280,282]. Overall, this results in increased proteolysis of erroneous recombinant proteins and reduces the overall protein quantity.

Codon usage is playing a particularly critical role in heterologous gene expression. Differences in codon usage and bias can contribute to difficulties during protein biosynthesis in the heterologous hosts, as they can increase the translational error frequency, especially for low-frequented and less-preferential tRNAs [279,283]. This is reflected in an increased occurrence of frameshifts [284] and misincorporation of amino acids [281], or deceleration and stalling of the translation process, which ultimately results in premature translation termination, initiation of mRNA degradation and proteolysis of erroneous (i.e. misfolded, truncated) recombinant proteins by quality control mechanisms [285–287]. A conducted analysis of the *S. cerevisiae* and the *Nostoc* sp. ATCC 53789 codon usage and bias (Figure 23) revealed clear differences, especially in the frequency of use (red and grey squares). Of particular concern are very rare codons (red squares), which have a frequency twice as high in the native producer. In particular, arginine (R) codons were found to be rarely frequented in yeast, when compared to the frequency of the cyanobacterial strain. In comparison to the cyanobacterial strain, overall, 50 % of these codons (CGC, CGG, CGA) are rarely used in yeast. The codon bias (green squares) revealed only minor differences. However, again the already rarely frequented arginine (R) codon exhibited major differences. In yeast the amino acid is preferentially encoded by the triplet AGA (bias: 0.48) instead of CGC (bias: 0.06). In contrast, *Nostoc* sp. ATCC 53789 prefers the codon CGC (bias: 0.28), which in turn is also one of the rarely frequented codons in yeast. Overall, both a high demand of low-frequented codons and the additional bias divergence could cause severe bottlenecks for the selection of cognate tRNAs during protein biosynthesis, as they are less available in the heterologous host. Consequently, this could lead to erroneous protein biosynthesis, stalling of the ribosome and proteolysis. Especially for arginine the observed limitation might be of particular concern. A closer inspection of the abundance and distribution of rare arginine codons in the cryptophycin assembly line proteins (Table SI III) revealed that in total 3.7% of all codons are arginine codons, whereby overall 41.7% of these argenines are ascribed to the rarely used yeast codons CGC, CGG and CGA. In case of the reconstructed assembly line, CrpA-CrpD and CrpH, also 3.6-3.7% are ascribed to R codons, whereby 42.2% are attributed to the rarely used arginine codons. Therefore, limited availability of compatible rare arginine codons during protein

biosynthesis is quite possible and could give an explanation for the observed lack of detectable recombinant proteins.

In the view of the performed high-level gene expression and the large size of the recombinant proteins, the effect of metabolic stress is also of concern. As such, the high number of high-molecular weight proteins may impose a metabolic burden for the cell, as they compete for limited resources (i.e. ribosomes, tRNAs, amino acids) during protein biosynthesis, both with each other and with the primary metabolism [123,288]. In particular large enzyme complexes, such as those expressed in this research project, are assumed to be energetically more costly, disrupt this intracellular balance of the cell's metabolism and tend to be more error-prone in aspects of misfolding and subsequent proteolysis [13,280,282]. Consequently, this may further contribute to an overall reduced quantity of recombinant proteins [280], resulting in concentrations below the detection limit of SDS-PAGE (~0.1-0.5 µg protein [289]).

To clarify the whereabouts of the recombinant proteins, the fusion of protein tags that can be used in affinity chromatography (e.g. His-tag) could be advantageous, as these methods allow for the concentration of the desired protein and facilitate the detection in subsequent SDS-PAGE analyses. Furthermore, the application of reporter proteins, such as GFP, to verify and visualize the translated protein, or a Western blot analysis with individually designed probes, might be another option. An alternative strategy might be the application of MS-based proteomics or whole-proteome sequencing.

***Saccharomyces cerevisiae* [gbpln]: 14411 CDS's (6534504 codons)**

fields: [triplet] [amino acid] [fraction] [frequency: **per thousand**] ([number])

UUU F 0.59 26.1 (170666)	UCU S 0.26 23.5 (153557)	UAU Y 0.56 18.8 (122728)	UGU C 0.63 8.1 (52903)
UUC F 0.41 18.4 (120510)	UCC S 0.16 14.2 (92923)	UAC Y 0.44 14.8 (96596)	UGC C 0.37 4.8 (31095)
UUA L 0.28 26.2 (170884)	UCA S 0.21 18.7 (122028)	UAA * 0.47 1.1 (6913)	UGA * 0.30 0.7 (4447)
UUG L 0.29 27.2 (177573)	UCG S 0.10 8.6 (55951)	UAG * 0.23 0.5 (3312)	UGG W 1.00 10.4 (67789)
CUU L 0.13 12.3 (80076)	CCU P 0.31 13.5 (88263)	CAU H 0.64 13.6 (89007)	CGU R 0.14 6.4 (41791)
CUC L 0.06 5.4 (35545)	CCC P 0.15 6.8 (44309)	CAC H 0.36 7.8 (50785)	CGC R 0.06 2.6 (16993)
CUA L 0.14 13.4 (87619)	CCA P 0.42 18.3 (119641)	CAA Q 0.69 27.3 (178251)	CGA R 0.07 3.0 (19562)
CUG L 0.11 10.5 (68494)	CCG P 0.12 5.3 (34597)	CAG Q 0.31 12.1 (79121)	CGG R 0.04 1.7 (11351)
AUU I 0.46 30.1 (196893)	ACU T 0.35 20.3 (132522)	AAU N 0.59 35.7 (233124)	AGU S 0.16 14.2 (92466)
AUC I 0.26 17.2 (112176)	ACC T 0.22 12.7 (83207)	AAC N 0.41 24.8 (162199)	AGC S 0.11 9.8 (63726)
AUA I 0.27 17.8 (116254)	ACA T 0.30 17.8 (116084)	AAA K 0.58 41.9 (273618)	AGA R 0.48 21.3 (139081)
AUG M 1.00 20.9 (136805)	ACG T 0.14 8.0 (52045)	AAG K 0.42 30.8 (201361)	AGG R 0.21 9.2 (60289)
GUU V 0.39 22.1 (144243)	GCU A 0.38 21.2 (138358)	GAU D 0.65 37.6 (245641)	GGU G 0.47 23.9 (156109)
GUC V 0.21 11.8 (76947)	GCC A 0.22 12.6 (82357)	GAC D 0.35 20.2 (132048)	GGC G 0.19 9.8 (63903)
GUA V 0.21 11.8 (76927)	GCA A 0.29 16.2 (105910)	GAA E 0.70 45.6 (297944)	GGA G 0.22 10.9 (71216)
GUG V 0.19 10.8 (70337)	GCG A 0.11 6.2 (40358)	GAG E 0.30 19.2 (125717)	GGG G 0.12 6.0 (39359)

***Nostoc sp. ATCC 53789* [gbbct]: 7 CDS's (10657 codons)**

fields: [triplet] [amino acid] [fraction] [frequency: **per thousand**] ([number])

UUU F 0.67 24.9 (265)	UCU S 0.26 17.5 (186)	UAU Y 0.59 18.9 (201)	UGU C 0.64 6.3 (67)
UUC F 0.33 12.0 (128)	UCC S 0.10 6.9 (74)	UAC Y 0.41 12.9 (137)	UGC C 0.36 3.5 (37)
UUA L 0.35 45.3 (483)	UCA S 0.17 11.2 (119)	UAA * 0.29 0.2 (2)	UGA * 0.43 0.3 (3)
UUG L 0.21 26.8 (286)	UCG S 0.06 3.8 (41)	UAG * 0.29 0.2 (2)	UGG W 1.00 13.4 (143)
CUU L 0.10 13.5 (144)	CCU P 0.25 12.8 (136)	CAU H 0.67 13.6 (145)	CGU R 0.21 9.9 (106)
CUC L 0.08 10.2 (109)	CCC P 0.24 12.1 (129)	CAC H 0.33 6.8 (72)	CGC R 0.28 13.2 (141)
CUA L 0.12 16.0 (171)	CCA P 0.43 22.1 (235)	CAA Q 0.68 47.0 (501)	CGA R 0.14 6.7 (71)
CUG L 0.13 17.1 (182)	CCG P 0.08 4.3 (46)	CAG Q 0.32 22.5 (240)	CGG R 0.11 5.2 (55)
AUU I 0.54 36.4 (388)	ACU T 0.39 23.3 (248)	AAU N 0.61 26.0 (277)	AGU S 0.23 15.3 (163)
AUC I 0.24 16.5 (176)	ACC T 0.18 10.8 (115)	AAC N 0.39 16.5 (176)	AGC S 0.18 11.6 (124)
AUA I 0.22 14.8 (158)	ACA T 0.32 19.0 (203)	AAA K 0.77 26.6 (283)	AGA R 0.17 8.0 (85)
AUG M 1.00 14.2 (151)	ACG T 0.10 5.8 (62)	AAG K 0.23 8.0 (85)	AGG R 0.08 3.7 (39)
GUU V 0.28 18.7 (199)	GCU A 0.36 24.8 (264)	GAU D 0.71 31.1 (331)	GGU G 0.33 19.2 (205)
GUC V 0.18 12.2 (130)	GCC A 0.15 10.4 (111)	GAC D 0.29 12.5 (133)	GGC G 0.16 9.6 (102)
GUA V 0.28 18.6 (198)	GCA A 0.35 24.5 (261)	GAA E 0.63 43.4 (462)	GGA G 0.32 18.3 (195)
GUG V 0.26 17.5 (186)	GCG A 0.14 9.6 (102)	GAG E 0.37 25.4 (271)	GGG G 0.19 11.0 (117)

Figure 23. Codon usage and bias comparison. The codon usage and frequency of *S. cerevisiae* (top) and *Nostoc* sp. ATCC 53789 (bottom). Differences of more than twice as high in frequency are highlighted in red, while variations of more than plus six in frequency are highlighted in grey. Notable differences in bias for a specific amino acid codon are highlighted in green. Amino acids are given in one letter code. The fraction of codon usage represents the bias for a respective codon among all codons, encoding for the same amino acid. The codon table was constructed using the online Codon Usage Database tool available at http://www.kazusa.or.jp/codon/ [290].

In accordance with the results of protein analysis, analytical investigations of the recombinant strains did also not confirm a successful biosynthesis of the expected cryptophycin derivatives. The lack of cryptophycin production in the generated *S. cerevisiae* strains might hence be due to the limited availability of recombinant biosynthesis enzymes. Yet there are many additional pitfalls with regard to the production of PKS/NRPS-derived secondary metabolites in yeast. A "literature analysis exposed four central issues, which are particularly relevant for producing polyketides and nonribosomal peptides in yeast. They can be summarized as (i) sufficient availability of biosynthetic precursors, (ii) adequate phosphopantetheinylation of the PKSs and/or NRPSs, (iii) balanced expression of tailoring enzymes, and (iv) the efficient expression of PKSs and NRPSs."

While the issue of associated tailoring enzymes (iii) is not applicable for the herein conducted reconstruction approach, the other mentioned factors are worth further consideration.

The supply of sufficient precursors (i) is of special importance for cryptophycin biosynthesis in *S. cerevisiae*. The heterologous pathway can be anticipated to compete with yeast's innate metabolism on multiple levels. During cryptophycin biosynthesis several metabolites are consumed. Some of these compounds represent known bottlenecks, such as malonyl-CoA and the shikimate pathway product L-tyrosine (cf. section 1.2.1) [13]. The consumption of β-alanine and α-ketoisocaproic acid by CrpD may cause further deficiencies in pantothenate and CoA [291], as well as leucine [292] biosynthesis. None of the transgenic strains was optimized with regard to the relevant precursor pools. Without adequate supplementation or metabolic engineering, the available resources can be limiting for cryptophycin biosynthesis, leading to insignificant product levels, or metabolic stress of the heterologous host. Therefore, strain engineering would be required in future approaches, improving the precursor supply under the aspect of shikimate and malonyl-CoA availability. Further optimization approaches regarding β-alanine and α-ketoisocaproic acid supply should also be considered to further reduce metabolic bottlenecks and the metabolic burden on the cell.

Adequate phosphopantetheinylation (ii) depends on the available PPTase(s). The activity of Sfp, which had previously been shown to successfully activate bacterial PKSs and NRPSs in *S. cerevisiae* [13], can be assumed to be satisfactory. However, to further improve the phosphopantetheinylation of the carrier proteins of the assembly line, it is advisable to test the effect of overexpression of *sfp*, compared to the *crp* genes, on the biosynthesis of putative cryptophycins. This increases the copy number and stability of the *sfp* insert [293], which could improve phosphopantetheinylation of the assembly line, as previously demonstrated for the heterologous production of indigoidine in yeast [13,210]. Further it has to be noted that the beneficial effect of PPTase overexpression on secondary metabolite production in general, has already been successfully demonstrated for native natural product producers, e.g. actinomycetes [294].

The expression of the cryptophycin PKSs and NRPSs (iv) might still leave room for optimization. The expression of the *crp* genes using constitutive synthetic promoters was successfully demonstrated by the transcription profiling experiments in this PhD project, yet investigations on a more balanced expression of the single pathway enzymes might be a considerable approach to improve product titers. In doing so the translational and enzymatic burden could be minimized. In fact, “screening of different promoter combinations driving the expression of the involved biosynthetic enzymes is a straightforward approach to identify expression conditions that might lead to higher product titers.”

Another option to produce cryptophycin is to synthesize the recombinant assembly line proteins

in yeast for the subsequent application in an *in vitro* biocatalysis approach, using either the cell lysate or the isolated proteins. In this way, precursor supply restrictions during product formation could be prevented by the sufficient supplementation of all required precursors for the biosynthesis of cryptophycins. An illustrative example is given by simvastatin biosynthesis [295], whereby a cell lysis approach resolved the separation issue of the externally located precursors and the intracellularly located biosynthetic enzyme.

In order to lay the foundations for a successful cryptophycin production in yeast, it will be necessary to improve the flux of precursor pathways by metabolic engineering. Moreover, a balanced heterologous gene expression and sufficient translation process of the recombinant proteins must be achieved. The latter will require optimizations in tRNA pool composition, coping with codon bias, as well as with rare and highly frequented codons. In this case, co-expression of rare and overexpression of highly frequented cognate tRNAs, as well as codon adaptation of the heterologous genes themselves, might be an option. The overexpression of chaperones might also support the folding of heterologous proteins and also reduce the burden of excessive protein folding requirements [296]. In perspective, however, it might be advantageous to investigate the sufficient biosynthesis of each biosynthetic protein in an individual host cell and analyze its ideal expression conditions. Subsequently, a concerted cell-lysate *in vitro* study including all four enzymes and Sfp could be used to test for cryptophycin biosynthesis.

4.3 Reconstruction of scytonemin biosynthesis in *S. cerevisiae*

4.3.1 Results

Apart from the heterologous expression of a large NRPS-PKS hybrid cluster (cf. section 4.2), *S. cerevisiae* was also tested as host for the biosynthesis of a secondary metabolite, which is exclusively made from two products of the shikimate pathway, i.e. L-tyrosine and L-tryptophan. The cyanobacterial alkaloid scytonemin is a natural sunscreen and therefore of biotechnological interest.

The initial heterologous expression of the pathway genes *scyA* and *scyC* resulted in clear differences of the HPLC-recorded metabolic profile in comparison to the control strain (Figure 24). Further, it revealed a distinct compound with an absorption maximum of 334 nm, which was exclusively found in the transgenic strain *S. cerevisiae-scyA-scyC*, but completely absent in the control strain. In addition, the concentration of this new metabolite evidently increased upon tryptophan supplementation during the cultivation, indicating a possible biosynthetic connection. In order to identify the new metabolite, the compound was purified by HPLC with a product titer of 20.3 mg/l and subjected to NMR analysis. Interpretation of its NMR spectra revealed the compound as anthranilic acid (data not shown). Its identity was further confirmed by LC-MS analysis, which detected a molecular ion with *m/*z 138.1 $[M+H]^+$ [297] (Figure 25).

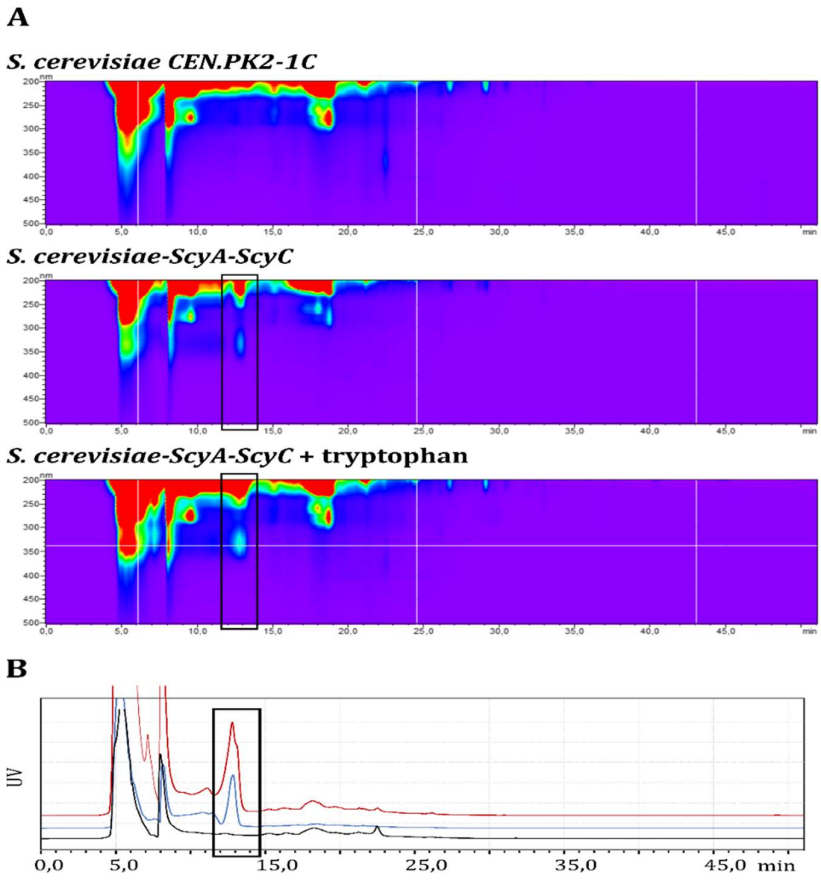


Figure 24. Metabolic profile of *S. cerevisiae* CEN.PK2-1C and *S. cerevisiae-scyA-scyC*. (A) HPLC analysis (200-500 nm), comparing the parental strain with the *S. cerevisiae-scyA-scyC* mutant, which was cultivated with and without L-tryptophan supplementation. The box is indicating new arising peaks within the mutant (λ-UV-max. 334 nm). (B). Distinctive UV-profile at 334 nm of *S. cerevisiae* CEN.PK2-1C (black), *S. cerevisiae-scyA-scyC* (blue) and *S. cerevisiae-scyA-scyC*, supplemented with L-tryptophan (red), respectively (B).

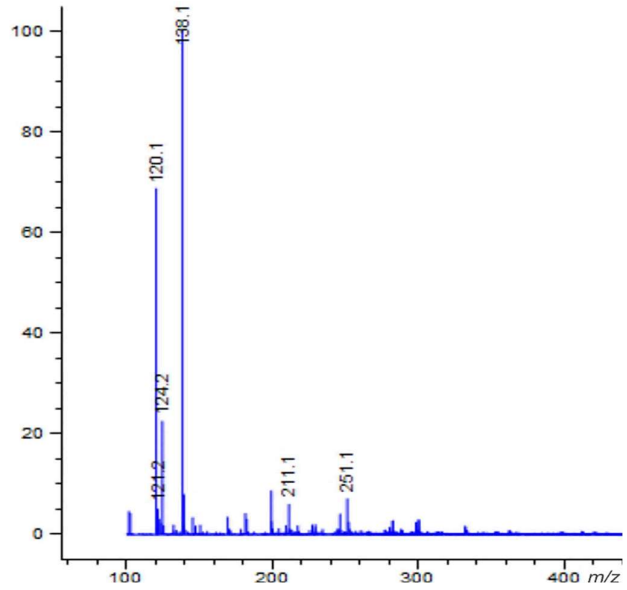


Figure 25. MS spectrum of the isolated anthranilic acid with *m/z* 138.1 $[M+H]^{+\cdot}$

The accumulation of anthranilate in *S. cerevisiae-scyA-scyC*, especially after supplementation of L-tryptophan, could indicate a possible bottleneck in scytonemin biosynthesis. Obviously, L-tryptophan is preferentially metabolized in yeast via *N*-formylkynurenine to anthranilate under the experimental cultivation conditions, whereas the alternative degradation route to indole-3-pyruvate is less preferred (Figure 26).

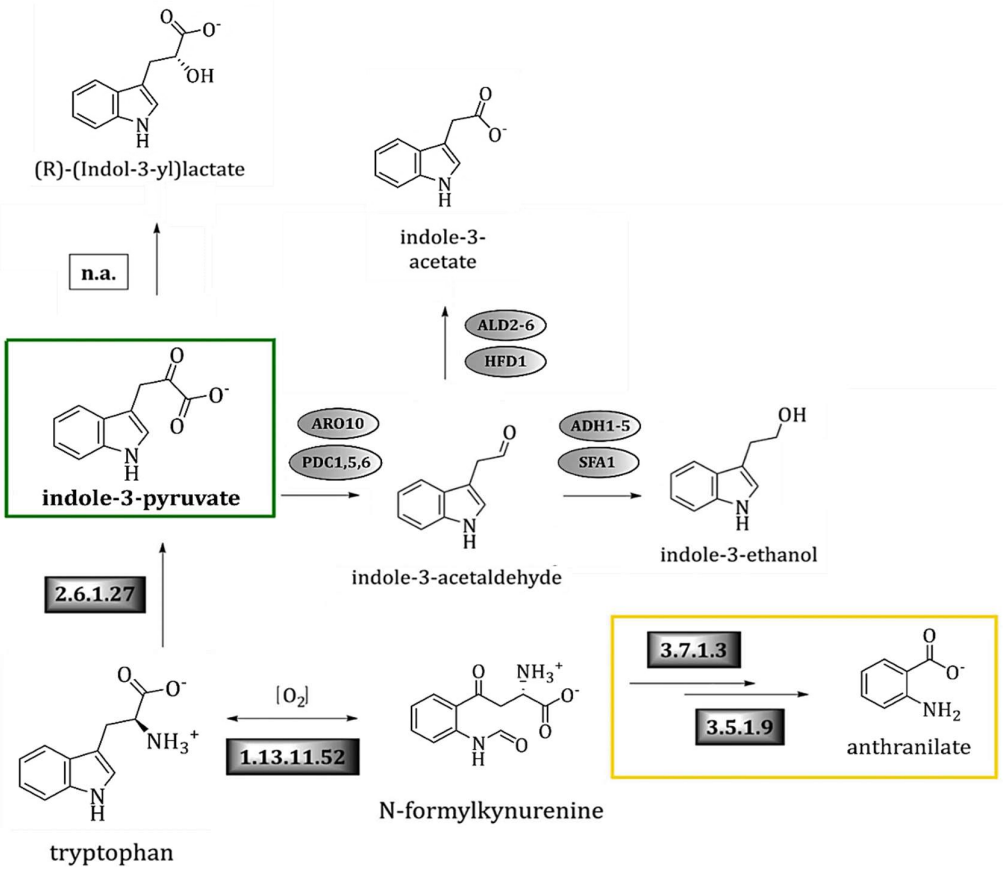


Figure 26. Tryptophan metabolism in *S. cerevisiae*. Schematic pathway illustration of the natively competing indole-3-pyruvate (green box) and kynurenine/anthranilate (yellow box) pathway in yeast, starting from tryptophan. Further indole-3-pyruvate can be degraded to indole-3-acetaldehyde, indole-3-ethanol, indole-3-acetate, or (*R*)-(indol-3-yl)lactate. Notation: 2.6.1.27 (Aro8), aminotransferase/transaminase; 1.13.11.52 (BNA2), tryptophan 2,3-dioxygenase; 3.7.1.3 (BNA5), kynureninase; 3.5.1.9 (BNA7); ARO10 and PDC1,5,6, pyruvate decarboxylases; ALD 2-6 and HFD1, aldehyde dehydrogenases; ADH1-5 and SFA1, alcohol dehydrogenases. Main pathway enzymes are indicated by boxes (grey), including the numerical enzyme nomenclature number. Oval boxes indicate indole-3-pyruvate degradation pathway enzymes. A white enzyme box, labeled with n.a., indicates that an intrinsic enzyme to catalyze the conversion is natively not available *S. cerevisiae*. The map was adapted from KEGG, SGD and BIOCYC reference pathways [298–300].

The existence of two alternative pathways in yeast for tryptophan catabolism is relevant to scytonemin biosynthesis, which relies on the availability of indole-3-pyruvate. The latter results from a transamination of L-tryptophan, which is catalyzed by the yeast enzyme Aro8 (EC: 2.6.1.27). In cyanobacterial scytonemin producers, the conversion of tryptophan into indole-3-

pyruvate is carried out by a pathway-specific leucine dehydrogenase homologue, i.e. ScyB [83].

To improve the flux to indole-3-pyruvate in *S. cerevisiae-scyA-scyC*, the competing pathway to anthranilate had to be shut down. For this purpose, the indoleamine 2,3-dioxygenase gene *bna2* was replaced with *scyB*, yielding the recombinant strain *S. cerevisiae-scyA-scyB-scyC*. Subsequently, XAD7-DCM-, XAD7-MeOH and XAD7-acetone extracts from different culture batches were prepared and analyzed by HPLC (cf. section 3.9.2). The HPLC analyses confirmed the absence of anthranilate in the *bna2::scyB*-engineered strain. While the XAD7-MeOH and XAD7-acetone extracts did not show any new signals (data not shown), several new metabolites were detected in the XAD7-DCM extract (Figure 27). The fractions corresponding to new metabolite signals (F0-F6 in Figure 27), were isolated and subjected to NMR analyses. However, the concentrations of the samples were too low for a structural identification. An additional, LC-MS inspection of the sample extracts also did not reveal any new compound signals (data not shown). An HPLC-based inspection of the extracts for the intermediate indole-3-pyruvate, in comparison to a standard compound, did not indicate the presence of this intermediate in extracts of the transgenic strain (data not shown).

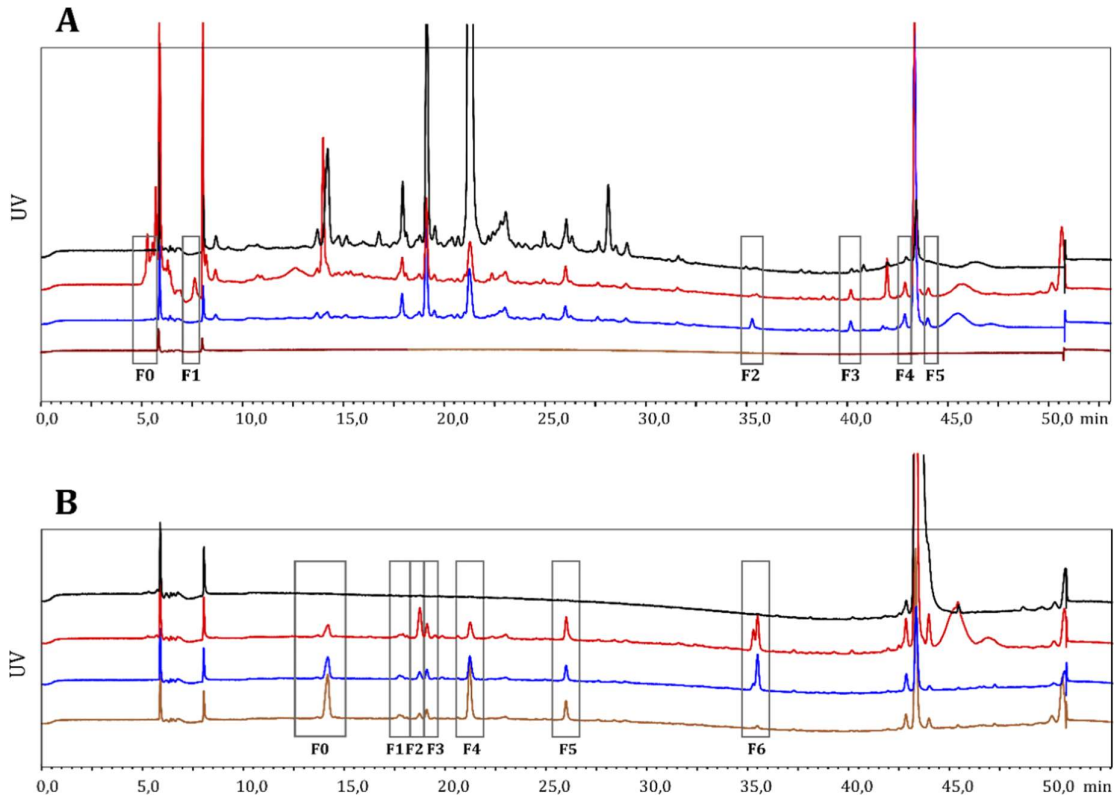


Figure 27. HPLC analysis (280 nm) of supernatant (A) and pellet (B) extracts of *S. cerevisiae* CEN.PK2-1C (black), as well as *S. cerevisiae-scyA-scyB-scyC* (red). The blue and brown chromatograms were obtained from culture extracts of *S. cerevisiae-scyA-scyB-scyC* upon supplementation with L-tryptophan and an equal mixture of L-tryptophan and L-tyrosine, respectively.

4.3.2 Discussion

In summary, the expression of selected scytonemin biosynthesis genes in *S. cerevisiae*, triggered the production of various new compounds which, however, could not be linked with the sunscreen pigment or its pathway intermediates up to now. Yet, new information was obtained that is highly relevant for heterologous scytonemin production in yeast. This includes evidence for (i) competitive primary metabolic pathways, and (ii) a limited supply of the required precursor indole-3-pyruvate, as well as (iii) enzymatic bottlenecks in the scytonemin pathway.

(i) A possible interference of scytonemin biosynthesis with endogenous yeast pathways was already diagnosed in the *S. cerevisiae-scyA-scyC* strain. This was visible by the accumulation of the tryptophan catabolite anthranilate upon the installation of the ScyA-ScyC-pathway. The obtained results suggest that the metabolic flux of the kynurenine/anthranilate pathway is intensified in the recombinant strain (Figure 26). Obviously, the two enzymes BNA2 and Aro8 compete for the degradation of L-tryptophan, whereby the route via BNA2 is clearly favored. Therefore, it can be assumed that the increased metabolic flux through the kynurenine pathway is also detrimental for the heterologous production of scytonemin in yeast, since it directly affects the availability of the scytonemin precursor indole-3-pyruvate (Figure 26).

This issue was already addressed in this PhD project by the construction of a *bna2::scyB* expression strain, i.e. *S. cerevisiae-scyA-scyB-scyC*. Although the disruption of *bna2* represents an intervention into the yeast's L-tryptophan detoxification pathway and could lead to accumulation of the toxic metabolite [301], Ohashi *et al.* showed that the knock-out of *bna2* does not increase the tryptophan-sensitivity in *Δbna2* mutants at elevated tryptophan levels [302]. Consequently, the disruption of *bna2* by *scyB*, was likely not affecting the strain's tryptophan sensitivity. Indeed, growth defects indicating a possible tryptophan intoxication, were not observed for *S. cerevisiae-scyA-scyB-scyC*. The abolishment of anthranilate production was successful, as evidenced by the chemical analysis of the constructed mutant strain. Despite this promising outcome, neither scytonemin, nor indole-3-pyruvate formation could be detected in the mutant strain.

(ii) Limited substrate availability is a frequent concern in heterologous natural product biosynthesis. As already noted during the pathway competition analyses, the natural conversion of L-tryptophan to indole-3-pyruvate depends on the action of Aro8 in yeast. Therefore, the generation of indole-3-pyruvate by Aro8 is resembling the function of the scytonemin pathway-specific enzyme ScyB that was not expressed at first. It was assumed that Aro8 could substitute for the transaminase functionality of ScyB, but the obtained results indicate that Aro8 alone is not capable to supply sufficient indole-3-pyruvate for scytonemin biosynthesis. This problem was addressed by replacing *bna2* with *scyB*. Although the allelic exchange was successful, the constructed mutant strain did not produce scytonemin or indole-3-pyruvate. This raised the

question of further constrains in the reconstructed pathway regarding precursor supply and potential enzymatic backlogs. One possible explanation for the absence of indole-3-pyruvate would be the biosynthesis of scytonemin or pathway intermediates (cf. Figure 5). Yet, none of the conceivable compounds was identified in the mutant extracts. Therefore, probable degradation or conversion processes in the heterologous host must also be considered, when investigating the plausible whereabouts of the compound. A literature survey revealed that indole-3-pyruvate is highly instable in an aqueous environment, as it would be present in the cell's cytosol. Under these conditions, indole-3-pyruvate is prone to spontaneous enzymatic and non-enzymatic conversion processes, that lead to indole-3-acetaldehyde, indole-3-acetate and indole-3-lactate [303,304]. Known conversion processes are depicted in Figure 26. Although *S. cerevisiae* is not known to possess an indole pyruvate decarboxylase, other enzymes, such as the decarboxylases Aro10, PDC6, PDC5 and PDC1, which are part of the tryptophol metabolism, can substitute this functionality, even though the catalysis is less efficient [299,300,305]. The resulting indole-3-acetaldehyde can subsequently be converted to indole-3-acetate and tryptophol by various aldehyde (ALD2-6, HFD1) and alcohol (ADH1-5, SFA1) dehydrogenases, respectively (see Figure 26) [298–300]. Furthermore, indole-3-pyruvate can be non-enzymatically converted to the degradation product indole-3-lactate [303,306]. Overall, the herein described degradation processes may indeed explain the detected lack of indole-3-pyruvate and consequently scytonemin in this heterologous expression attempt to some extent. In perspective, the formation of the possible degradation products must be systematically analyzed in *S. cerevisiae*.

(iii) The expression of the Aro8 homolog ScyB did not lead to the biosynthesis of scytonemin. Therefore, the obtained results were again analyzed to identify potential enzymatic bottlenecks in the scytonemin pathway. The accumulation of anthranilate in *S. cerevisiae-scyA-scyC*, even without the addition of tryptophan, could imply the degradation of a compound, which is structurally related to tryptophan and may be accepted as a substrate by BNA2. Thus, this compound would be channeled through the kynurenine pathway forming anthranilate. As the control strain did not show this accumulation, it can be assumed that the compound derives from the scytonemin pathway and may be linked to the degradation of indole-3-pyruvate or the labile β-ketoacid by BNA2. Furthermore, the supplementation-independent anthranilate accumulation may give rise to the suspicion of a functional backlog in regard to the scytonemin biosynthesis enzyme ScyC, which is responsible for the sufficient conversion of the unstable β-ketoacid, forming the more stable tricyclic ketone intermediate (cf. section 1.2.2). As this conversion may be hampered due to a slow conversion rate or a dysfunctionality of the enzyme in the heterologous host, high levels of the unstable β-ketoacid may be produced by ScyA. The β-ketoacid is readily degraded and culminates in the lack of the subsequent more stable ketone intermediate or the scytonemin product. Also, both the β-ketoacid itself and its probable

degradation products, could be readily processed via the kynurenine pathway and result in the additional formation of anthranilate. This hypothesis is supported by a close inspection of the first enzyme, i.e. BNA2, of the described tryptophan catabolism. As illustrated in Figure 26, BNA2 catalyzes the conversion of tryptophan towards *N*-formylkynurenine by an oxidative ring opening of the indole moiety, forming e.g. *N*-formylkynurenine, which is likewise present in the β-ketoacid and its degradation products. Therefore, it might be assumed that the resulting tryptophan-like β-ketoacid derivatives may also be catabolized by this enzyme, yielding the detected additional formation of the anthranilate by-product.

These enzymatic constraints could explain the lack of scytonemin in addition to the previously discussed issues. In order to elucidate probable enzymatic bottlenecks in the scytonemin pathway a subsequent *in vitro* assay analysis is recommended to investigate the substrate conversion of the recombinant enzymes and their bottlenecks in detail.

4.4 *In vivo* biosynthesis of MPBA in *E. coli*

The following section contains work that was conducted by Marvin Daney Schwing during his Bachelor thesis, which was designed and supervised by the author of this dissertation. Contents of this work were translated and summarized to be presented in an adapted form.

4.4.1 Results

The cryptophycin biosynthesis enzyme CrpA should be used for the biotechnological production of MPBA, which is a derivative of the industrially and pharmaceutically relevant bulk chemical PBA. For the expression of *crpA*, different plasmids and *E. coli* host strains were tested. The activation of the carrier protein domains in CrpA should be implemented by co-expression of PPTases and the premature off-loading of the desired product by a TE-domain fusion approach. Initially, the strain *E. coli* BL21 (DE3)-*crpA-sfp* was constructed and analyzed for the production of the recombinant His6-CrpA in comparison to *E. coli* BL21 (DE3). SDS-PAGE (Figure 28) revealed a strong band with a size corresponding to the expected His-tagged protein of 327 kDa. However, this band was only detected in the pellet fraction, indicating the formation of insoluble inclusion-bodies.

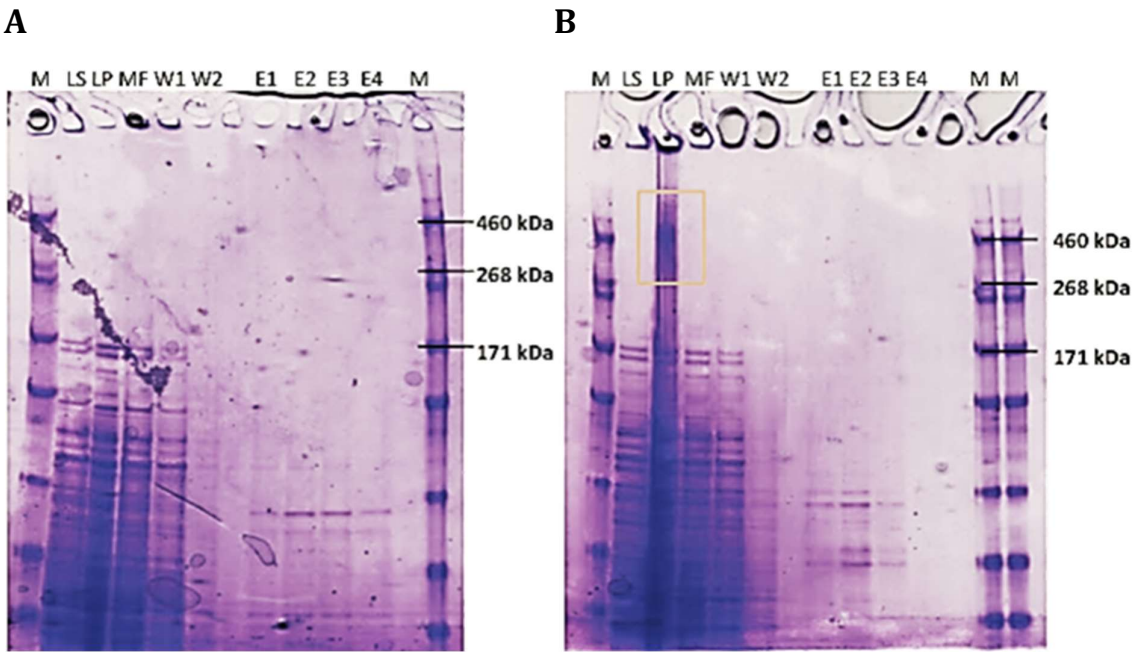


Figure 28. SDS-PAGE analysis of the cell lysate pellets and supernatants from (A) *E. coli* BL21 (DE3) and (B) *E. coli* BL21 (DE3) expressing pET28a-crpA and pACYC-sfp. Notation: LP, lysed cell pellet fraction after sonication; MF, Ni-NTA-column purification matrix flow through; W1, 1st column wash; W2, 2nd column wash; E1-E4, elution fractions; M, LC5699-HiMark™ (Thermo Fisher Scientific). His6-CrpA is indicated by a box, corresponding to the expected 327 kDa protein in the pellet lysate fraction of the mutant strain.

To solubilize CrpA different fusion strategies were pursued, including the use of an N-terminal SUMO or MBP solubility tag. Each gene fusion construct was introduced into five different *E. coli* host strains and expressed together or without Sfp. Afterwards recombinant protein production was analyzed by SDS-PAGE (Figure 29, Figure SI VI). This study confirmed the production of soluble MBP-CrpA (366 kDa) by all five host strains, although insoluble protein fractions were still detected in the lysed cell pellets (Figure SI VI). On the contrary, no evidence for the production of SUMO-CrpA (337 kDa) was found in any of the five test strains. Obviously, the SUMO tag is not suitable for the heterologous production of CrpA in *E. coli* (Figure SI VI). In total, all *E. coli* expression strains showed a successful production of the recombinant MBP-CrpA (Figure 29, Figure SI VI). *E. coli* BL21 (DE3)-MBP-*crpA-sfp, E. coli* C41(DE3)-MBP-*crpA-sfp* and *E. coli* C43(DE3)-MBP-*crpA-sfp* exhibited the strongest band in the conducted SDS-PAGE analysis, while the detected protein band was less intense in *E. coli* ArcticExpress and *E. coli* KRX-MBP-*crpA-sfp* (Figure SI VI), which indicates that the amount of synthesized protein seems to depend on the selected host strain as well as on the utilized fusion tag. Noteworthy, the formation of soluble CrpA appears to be independent from the presence of a functional PPTase in *E. coli*, as evidenced by the analysis of *E. coli* DH5α-MBP-*crpA.*

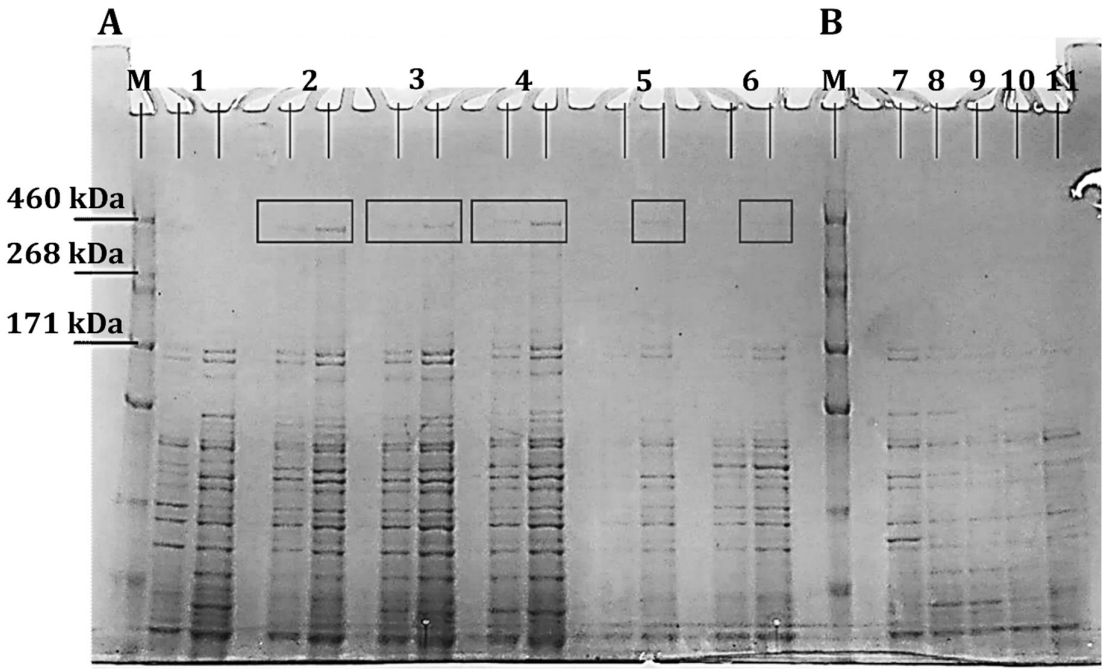


Figure 29. SDS-PAGE analysis of supernatant cell lysate fractions from five different *E. coli* strains harboring pMal-crpA +/- pACYC-sfp (A) in comparison to negative control strains (B). Notation: M, LC5699-HiMark™ (Thermo Fisher Scientific); 1, *E. coli* KRX-MBP-*crpA-sfp*; 2, *E. coli* BL21 (DE3)-MBP-*crpA-sfp*; 3, *E. coli* C41(DE3)-MBP-*crpA-sfp*; 4, *E. coli* C43(DE3)-MBP-*crpA-sfp*; 5, *E. coli* ArcticExpress (DE3)-MBP-*crpA-sfp*; 6, *E. coli* DH5α-MBP-*crpA*; 7, *E. coli* ArcticExpress (DE3)*;* 8, *E. coli* C41(DE3)*;* 9, *E. coli* C43(DE3); 10, *E. coli* BL21(DE3); 11, *E. coli* KRX. The formation of MBP-CrpA is indicated by a box, corresponding to the 366 kDa protein.

After the solubility issue of CrpA had been solved, the different *E. coli* host strains, expressing pMal-crpA and pACYC-sfp, were tested for the production of MPBA by LC-MS. A peak with a retention time of 7.3 min and a molecular ion at *m/z* 177.1 $[M+H]^+$, which might correspond to MPBA, was consistently observed in the culture extracts of the recombinant expression strains. The same signal was absent in the extracts of the respective control strains (Figure 30; Figure SI VII). The highest signal intensity of the newly detected metabolite was achieved by *E. coli* BL21 (DE3)-MBP-*crpA-sfp* (Figure SI VII), which is consistent with the previously observed protein formation. Unfortunately, the titer of the metabolite was too low (overall maximum intensity of only 1.8×10^4) for a thorough chemical characterization.

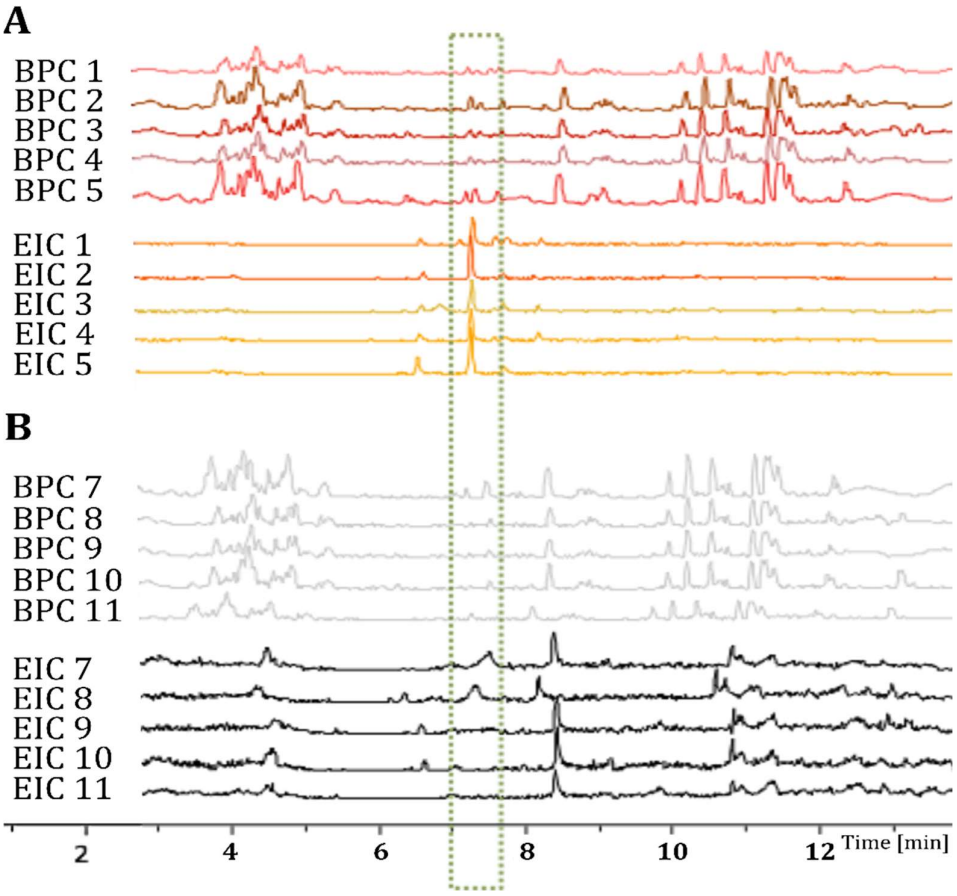


Figure 30. Base peak chromatograms (BPC) and extracted ion chromatograms (EIC) of the culture extracts from five different *E. coli* strains harboring pMal-crpA +/- pACYC-sfp (A) in comparison to negative control strains (B). The green box indicates the peak of a compound with a mass corresponding to MPBA. Notation: 1, *E. coli* KRX-MBP-*crpA-sfp*; 2, *E. coli* BL21 (DE3)-MBP-*crpA-sfp*; 3, *E. coli* C41(DE3)-MBP-*crpA-sfp*; 4, *E. coli* C43(DE3)-MBP-*crpA-sfp*; 5, *E. coli* ArcticExpress (DE3)-MBP-*crpA-sfp*; 7, *E. coli* ArcticExpress (DE3)*;* 8, *E. coli* C41(DE3)*;* 9, *E. coli* C43(DE3); 10, *E. coli* BL21(DE3); 11, *E. coli* KRX.

To verify the identity of the newly detected compound, a large-scale cultivation of *E. coli* BL21 (DE3)-MBP-*crpA-sfp* was performed and the strain extracts were analyzed by HPLC in comparison to a commercially obtained MPBA standard. Further, the control strains *E. coli* BL21 (DE3)-*sfp*, *E. coli* BL21 (DE3)-MBP-*crpA* and *E. coli* BL21 (DE3)-*sfp*-C2x were simultaneously analyzed, to identify increased, or new metabolite signals that do not result from a successful MPBA production, but solely from the discrete expression of pACYC-sfp, pMal-crpA, or a combination of pACYC-sfp and the empty pMalc2x vector (Figure 31).

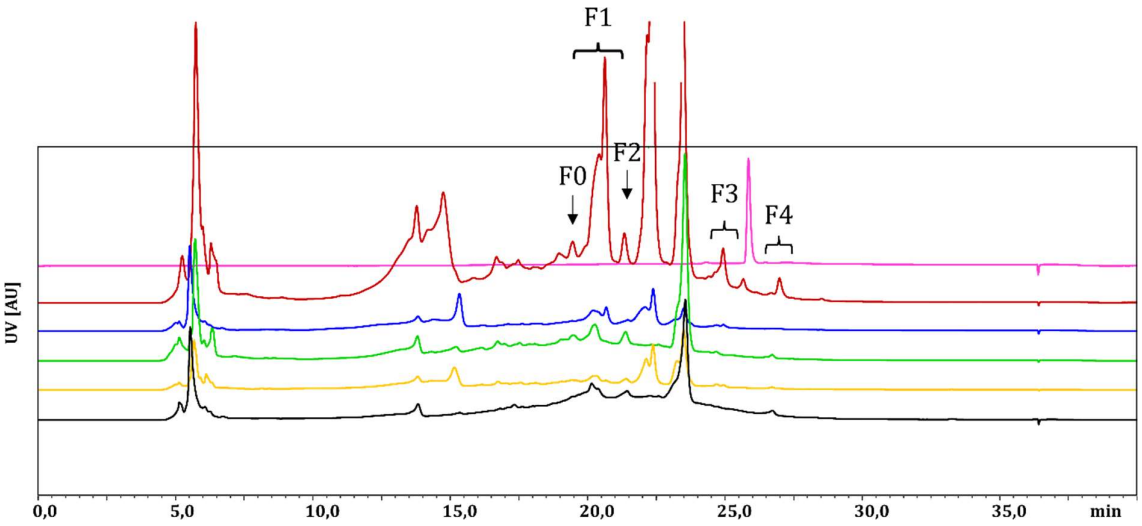


Figure 31. Comparative HPLC analysis of obtained *E. coli* extracts. The production strain *E. coli* BL21 (DE3)-MBP-crpA-sfp and control strains were tested for the *in vivo* biosynthesis of MPBA in comparison to a commercially acquired standard. Notation: Black, *E. coli* BL21 (DE3) control; Yellow, *E. coli* BL21 (DE3)-*sfp* control; Green, *E. coli* BL21 (DE3)-MBP-*crpA* control; Dark blue, *E. coli* BL21 (DE3)-*sfp*-C2x control; Red, *E. coli* BL21 (DE3)-MBP-*crpA-sfp* expression strain; Pink, MPBA standard. F0-F4 indicate fractions, corresponding to new metabolite signals that were isolated for the succeeding compound purification.

As a result, the comparative HPLC analysis revealed no signal corresponding to the MPBA standard in any of the analyzed production strains. Consequently, it could not be confirmed that the results from the previous LC-MS analysis (see Figure 30) correspond to a MPBA-related metabolite. Yet new HPLC-signals were observed at a retention time of 19.5 - 27.0 min (Figure 31 F0-F4). The new signals were fractionated and analyzed by LC-MS and NMR analyses (data not shown), which, however, also did not confirm the presence of MPBA or a related compound. The acquired LC-MS spectra of the obtained MPBA standard (Figure SI IX) affirmed a corresponding mass signal of *m/z* 177.0914 $[M+H]^+$ with a retention time of 7.1 min and a mass difference of +0.0004. In contrast, the previous results showed a corresponding mass of *m/z* 177.0642 $[M+H]^+$ at 7.3 min and a mass deviation of -0.0268. These results ultimately confirmed the previous LC-MS signals (compare Figure 30 and Figure SI VII) as not identical to the desired compound.

One possible explanation for the observed lack of MPBA in the recombinant strain could be reasoned with a low production titer of MPBA that is caused by insufficient phosphopantetheinylation of recombinant CrpA by Sfp. Therefore, a BLASTp search for Sfp-type proteins in the genome of *Nostoc* sp. ATCC 53789 (Accession No.: CP046703) was conducted to identify the native PPTase of the assembly line, which resulted in one hit. Although the sequence identity of this putative PPTase was only 27% to Sfp, an alignment of the two proteins revealed the conserved sequence motives [W (T/C) K E A] and [G K $P_{11\text{-}17}$ S H] [29] (Figure 32). In the following, the PPTase, which was detected in *Nostoc* sp. ATCC 53789 is referred to as SfpN. Since SfpN is the

only PPTase encoded in the genome of the cryptophycin-producing strain, it was consequently assumed to be responsible for the activation of CrpA under natural conditions [a].

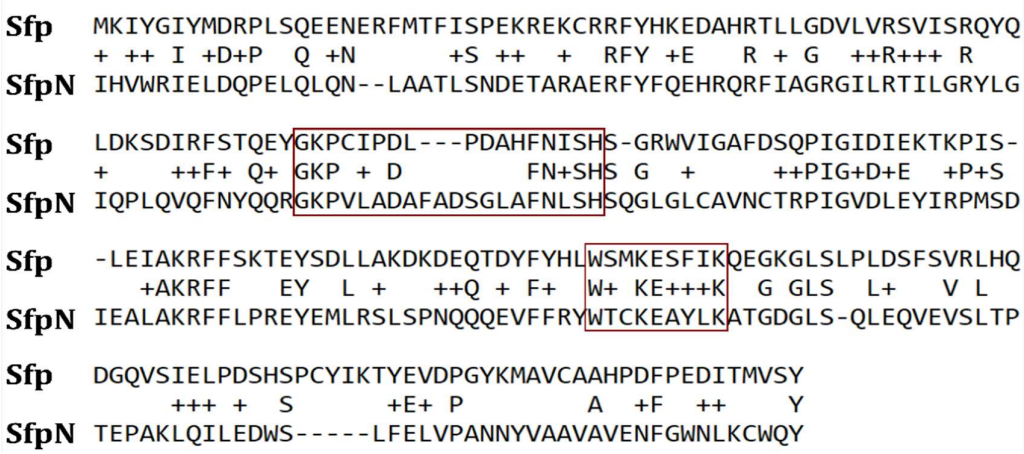


Figure 32. Protein sequence alignment of Sfp and SfpN. The amino acid sequences of Sfp, originating from *B. subtilis,* supported the identification of the cyanobacterial SfpN, exhibiting the conserved sequence motives [W(T/C)KEA] and [$GKP_{11\text{-}17}SH$] [29]. SfpN was consequently assigned to the Sfp-type PPTase of the W/KEA branch. The graphical presentation of the alignment was adapted from the Bachelor thesis of M.D. Schwing [a]

To test the biosynthesis of MPBA with SfpN the recombinant strain *E. coli* BL21 (DE3)-MBP-*crpA*-*sfpN* [a] was constructed. In analogy to the previously investigated strains, an expression study was performed. An SDS-PAGE analysis was conducted, confirming the successful formation of soluble MBP-CrpA (Figure SI VIII). Sample extracts of the expression culture were prepared and the formation of MPBA was analyzed by a comparative HPLC inspection with a commercially obtained MPBA standard, the initial expression strain *E. coli* BL21 (DE3)-MBP-*crpA*-*sfp* and the control strains (Figure 33). The fractionated extracts (Figure 33 F0-F4) were further analyzed by LC-MS and NMR analyses (data not shown). Overall, the recorded HPLC chromatograms confirmed the previously obtained results, whereby the metabolic profile of the recombinant SfpN strain was found to be identical to that of *E. coli* BL21 (DE3)-MBP-*crpA*-*sfp*.

Another factor contributing to low MPBA yields might be an insufficient product release. In NRPS and PKS assembly lines, the final product is typically released from the synthetizing complex by a terminal TE domain (cf. section 1.2) through hydrolysis or cyclization. In cryptophycin biosynthesis, the innate TE domain, encoded by *crpD*, is responsible for the macrocyclization of the processed cryptophycin intermediate, initiating the release of the functional cryptophycin core product [64] (cf. section 1.2.1). By releasing the substrate, the assembly line is available to traverse a new round of product formation, without any backlog. Yet, the recombinant CrpA, as

being the first enzyme of cryptophycin biosynthesis, is lacking a TE domain functionality that could take over the required off-loading of the synthesized product. This is causing a direct dependency of the MPBA concentration on the amount of available CrpA. As the liberation of the product thus depends on spontaneous hydrolysis, the biosynthetic activity of CrpA would be stalled during the assembly process, unable to run another circle of MPBA biosynthesis and reducing the overall product yield. Therefore the effect of the constructed recombinant strain *E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp*, comprising a C-terminally installed TE domain was tested for sufficient product formation. An expression cultivation was performed. The successful formation of soluble MBP-CrpA$_{TE}$ in *E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp* was confirmed by SDS-PAGE analysis (Figure SI VIII). Subsequently, MPBA product formation was analyzed by HPLC, alongside the previously tested strains and the MPBA standard (Figure 33). The obtained HPLC results are consistent with the data obtained for *E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp* and *E. coli* BL21 (DE3)-MBP-*crpA*-*sfpN* [a], revealing an identical metabolic profile. Also the subsequently conducted LC-MS and NMR analyses of the HPLC-purified extract fractions (data not shown) did not show a signal corresponding to MPBA.

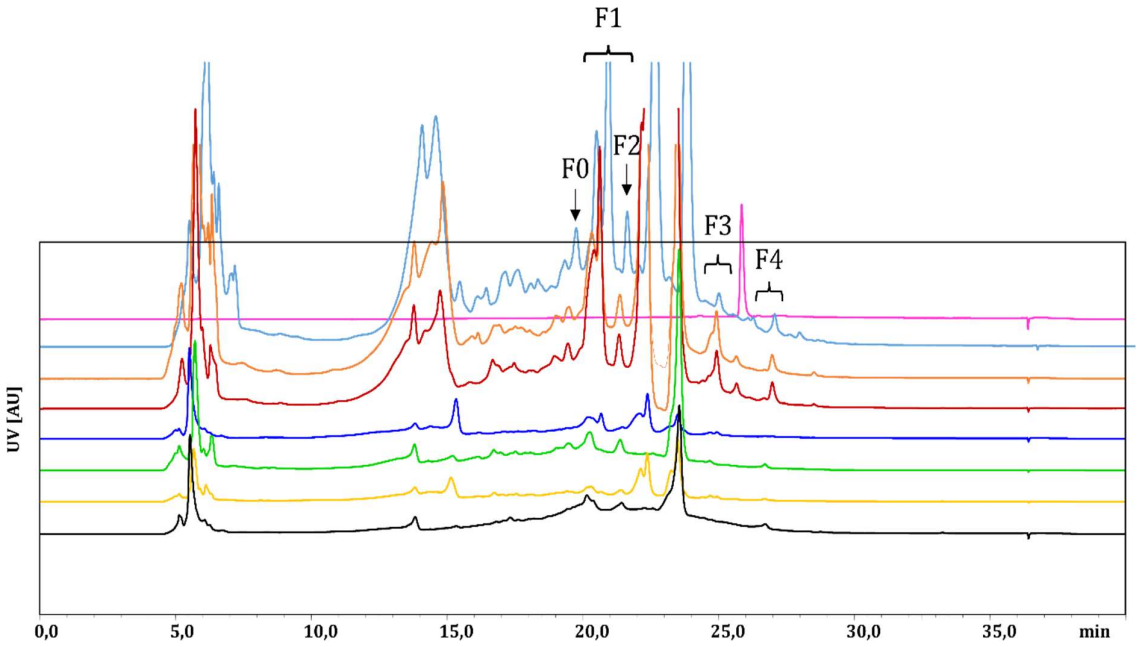


Figure 33. Comparative HPLC analysis of obtained *E. coli* extracts. Diverse production and control strains were tested for the *in vivo* biosynthesis of MPBA in comparison to a commercially acquired standard. Notation: Black, *E. coli* BL21 (DE3) control strain; Yellow, *E. coli* BL21 (DE3)-*sfp*; Green, *E. coli* BL21 (DE3)-MBP-*crpA*; Dark blue, *E. coli* BL21 (DE3)-*sfp*-C2x; Red, *E. coli* BL21 (DE3)-MBP-*crpA*-*sfp*; Orange, *E. coli* BL21 (DE3)-MBP-*crpA*-*sfpN* [a]; Light blue, *E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp*; Pink, MPBA standard. F0-F4 are indicating fractions, which were isolated for the succeeding compound purification.

4.4.2 Discussion

Overall, the *in vivo* biosynthesis of MPBA could not be realized in *E. coli*. None of the obtained results yielded compelling evidence for the production of this compound. Although new signals could be detected for the investigated recombinant expression strains, these metabolites could not be confirmed as MPBA or as any related compound.

As the problem of soluble protein formation was evidently solved by the installation of a solubility tag, this issue can be excluded as a cause for the lack of product formation. Likewise, the obtained data suggest that phosphopantetheinylation by the foreign PPTase Sfp is not an issue, as the coexpression of the naturally associated PPTase SfpN did not improve MPBA biosynthesis either. Furthermore, it has to be noted that Sfp is a PPTase, which was previously demonstrated to activate a variety of ACP and PCP domains, including those of enterobactin [246], erythromycin, epothilone [307–310], hapalosin and anabaenopeptin biosynthesis [23,310–312].

Other factors that could influence heterologous MPBA production are (i) the substrate selection and truncation by CrpA, (ii) precursor supply, and (iii) insufficient termination of the assembly process.

(i) The process of CrpA-loading and the subsequent unusual truncation of the selected precursor remains enigmatic, which might also affect the heterologous production of MPBA. So far, it is unknown how the carbon truncation of the assumed phenylpropanoid substrate is carried out by CrpA, or if other *in-trans* acting, cyanobacterial enzymes are involved in this process. Although the genomic investigation of the cryptophycin locus did not give any evidence for a possible mechanism (cf. section 4.1.1.3), the question arises whether the loading and subsequent truncation of the substrate are feasible in the recombinant *E. coli*. It is reasonable that improper loading and a lack of enzymes involved in the truncation mechanism would directly hamper MPBA biosynthesis in the heterologous host. Yet, *in vitro* and *in vivo* studies of the cryptophycin-related microcystin synthetase and its loading domain showed that proper loading of the assembly line is indeed feasible in *E. coli* and that the heterologous strain is capable to perform the unusual processing of the starter unit [76,126] (cf. sections 1.2.1, 1.3.1, Figure 4). Accordingly, it can be deduced that neither insufficient loading, nor elongation or the unusual substrate truncation should cause a problem for heterologous MPBA biosynthesis in *E. coli*.

(ii) Precursor supply is a considerable issue for the reconstitution of heterologous pathways, especially when the host strain is no native producer of secondary metabolites *per se*. It is a known fact that substrate limitations have a detrimental effect on product formation and frequently results in the failure of heterologous natural product biosynthesis [126]. In this study, the supply of a suitable CrpA precursor was achieved with the supplementation of 3-PPA to the cultivation medium, which has also been shown to suffice the effective *in vivo* loading of the reminiscent

McyG module in microcystin biosynthesis [76]. However, 3-PPA can be catabolized by *E. coli*, yielding succinate, pyruvate and acetyl-CoA [313]. The metabolic breakdown of 3-PPA could cause a severe precursor constraint during MPBA biosynthesis. In perspective, pathway engineering might be considered to interrupt the degradation of 3-PPA. However, negative effects on primary metabolic pathways, such as glycolysis and citrate cycle, have to be considered and tested.

(iii) Although a TE domain fusion approach was conducted, the functionality of the resulting protein remains elusive. Further experiments, investigating the actual activity of CrpA-TE and the product release of MPBA, would be required. For this purpose, an *in vitro* assay testing, coupled to the direct or indirect analytical detection of the desired assembly line-liberated or enzyme-bound metabolite could provide insights into the productivity, as well as the functionality of the assembly line. In this way, also production improvements that are based on the installation of the TE domain could be assessed without the interference of substrate limitations in the host. Moreover, any backlogs during product assembly and release could be identified. A direct method would involve the comparative investigation of the product formation in CrpA and CrpA-TE *in vitro* assays and a subsequent assessment of MPBA production and concentration by LC-MS or HPLC-based methods. To observe the actual off-loading functionality of the installed TE-eryAIII, a comparative enzyme reaction profile, analogous to the exemplary experiments of Schwarzer *et al.* [314], could represent a suitable indirect option. In this case, the researchers used a comparative (TE-fused and without-TE) two-stage *in vitro* assay of an artificial assembly line in combination with radioactively labeled precursors and a coupled time-dependent measurement of radioactivity for trichloroacetic acid-precipitated enzymes. Detected signals correspond to enzyme-bound substrates, whereby a decrease of the signal intensity over time symbolizes a release of the compound from the assembly line. In parallel, the formation of assembly line intermediates and products was analytically observed [314].

Furthermore, the investigation of different TE domains is worth considering in order to identify the most suitable enzyme for the cleavage of the desired product. Although the $TE_{eryAIII}$ domain of erythromycin biosynthesis seems to be a suitable candidate, the testing of alternative enzymes would be worthwhile. In this regard, one intriguing candidate is the TE domain of CrpD (TE_{crp}). This thioesterase was successfully tested as a robust enzyme, accepting a broad diversity of cryptophycin intermediates in the form of activated *N*-acetylcysteamine thioesters (SNAC-esters) to yield linear and cyclized products [315]. Even though the flexibility of the TE_{crp} for variations in the Western, phenyl ring-containing part of the molecule is restricted, the acceptance and liberation of SNAC esters with early assembly line intermediates reminiscent of those of CrpA or CrpB, was not tested so far. In addition, the investigation of TE domains that naturally function as exclusive hydrolases and not as cyclases, which is the case for TE_{crp} and $TE_{eryAIII}$, would be possible

for future TE fusion approaches.

The enzyme CrpA is a promising candidate as a catalyst for the biosynthesis of MPBA and *E. coli* has been successfully used for heterologous natural product biosynthesis before. Yet, the microorganism may not be suitable for the biosynthesis of MPBA due to precursor and product release restrictions of the recombinant CrpA. Other hosts, which are already in possession of sufficient precursor pathways, are worthwhile to consider for a heterologous expression approach. In addition, a general *in vitro* attempt, using CrpA as a catalyst to produce MPBA, may support the elucidation of actual production bottlenecks and serve as an alternative approach to achieve the biosynthetic production of the desired compound as a biosynthetic intermediate in order to clarify remaining question of the cryptophycin initial substrate selection and assembly process.

4.5 *In vitro* biosynthesis of PBA

4.5.1 Results

For the production of PBA, an artificial pathway starting from L-α-homophenylalanine was designed. This pathway comprises the aminomutase AdmH from *Pantoea agglomerans* and the ammonia lyase Npun_R2068 from *Nostoc punctiforme* as biocatalysts (Figure 34). The two enzymes are naturally involved in the metabolism of L-phenylalanine, but they were also reported to exhibit a broad substrate promiscuity [245,316–319]. However, the utilization of aromatic substrates with a four-carbon side-chain has not been tested so far.

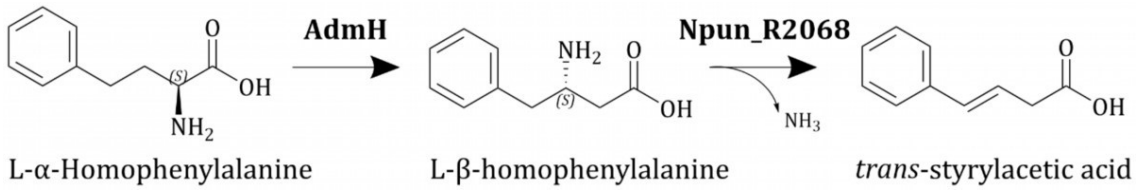


Figure 34. Artificial pathway to *trans*-styrylacetic acid (PBA).

The first step in the development of this synthetic pathway involved the testing of AdmH for the conversion of L-α-homophenylalanine. For this purpose, the strain *E. coli* BL21 (DE3)-*admH* was constructed to produce a hexahistidyl-tagged AdmH and subsequently analyze its enzymatic activity in an *in vitro* assay (cf. section 3.6).

SDS-PAGE analysis confirmed the recombinant production of His6-AdmH, as indicated by a protein band at 61.4 kDa (Figure 35).

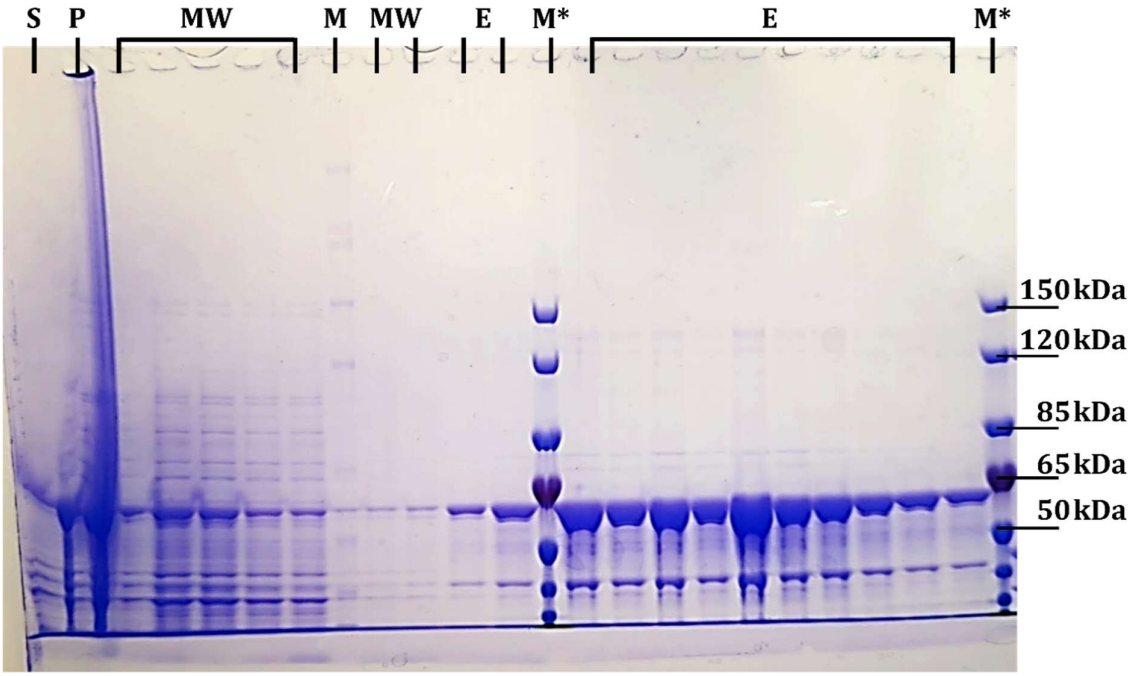
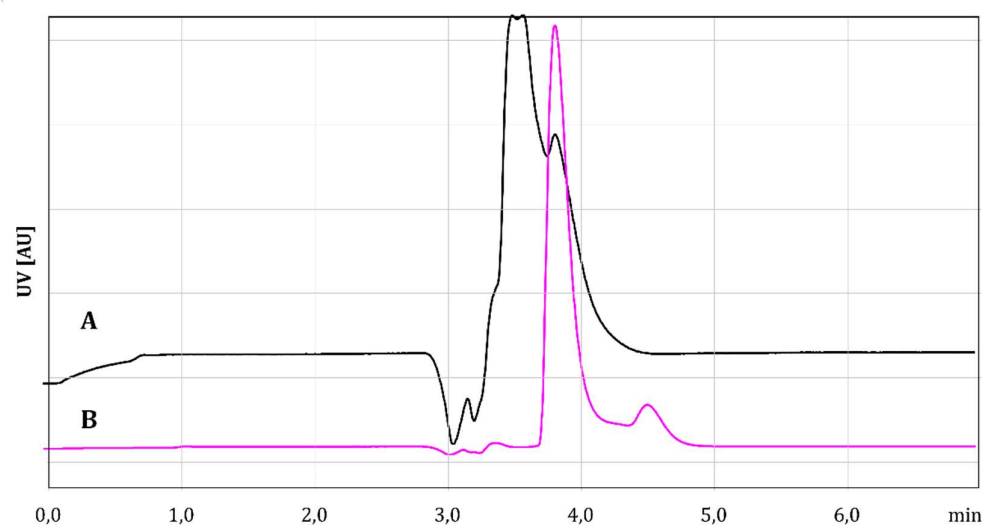


Figure 35. SDS-PAGE of His6-AdmH. The proteins of the cell lysate supernatant and pellet fraction of *E. coli* BL21 (DE3)-*admH* were purified by Ni-NTA chromatography and analyzed by SDS-PAGE, using a Criterion™ chamber (Bio-Rad) and 3-8% precast Criterion™ XT Tris-Acetate gradient gels. Notation: S, supernatant fraction of cell lysate; P, pellet fraction of cell lysate; MW, matrix wash after supernatant lysate application; M, LC5699-HiMark™ (Thermo Fisher Scientific); E, AdmH elution fraction; M*, PageRuler™ (Thermo Fisher Scientific).

The purified protein was assayed with L-α-homophenylalanine or L-α-phenylalanine, respectively. Product formation was analyzed by HPLC (Figure 37). While L-α-phenylalanine was successfully converted into its corresponding β-amino acid, as evidenced by the occurrence of a new peak, L-α-homophenylalanine was not processed by His6-AdmH. However, the conversion of L-α-homophenylalanine by His6-AdmH revealed a new, weak peak signal at a retention time of 3.7 min. Therefore, a closer inspection of the assay samples was conducted, using an optimized HPLC method for an improved peak separation with an extended gradient from 10 to 100% MeOH within 20 min instead of 15 min. The examination of the assay samples against the purified His6-AdmH and the prepared standards confirmed the previous result and showed that the newly detected peak corresponds to the isolated enzyme. (Figure 37). The same result was obtained when the assay incubation time was prolonged from 4 hours to overnight (data not shown).

I



II

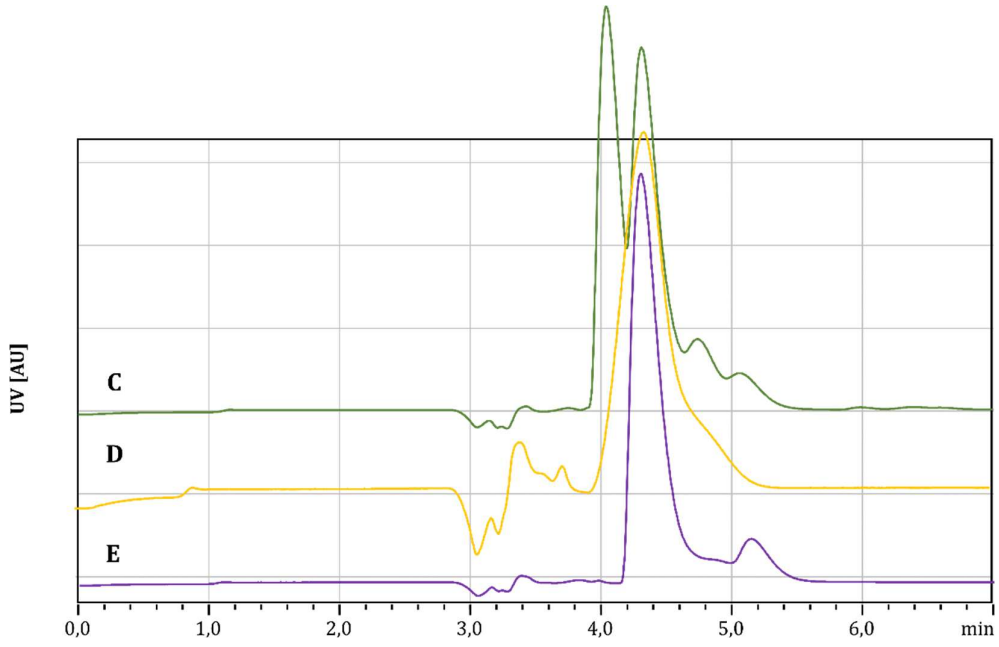


Figure 36. HPLC analysis of L-phenylalanine (I), L-α-homophenylalanine and L-β-homophenylalanine (II). 6xHis-AdmH was incubated with L-phenylalanine (A) or L-α-homophenylalanine (D) and the obtained chromatograms were compared to an L-phenylalanine (B), or L-α-homophenylalanine standard (E) and a standard mix, prepared from L-α-homophenylalanine and L-β-homophenylalanine (C). All chromatograms were recorded at 206 nm on a Sphinx column with a gradient set from 10 to 100% MeOH within 15 min, 100% MeOH over 5 min and from 100% to 10% MeOH in 2 min.

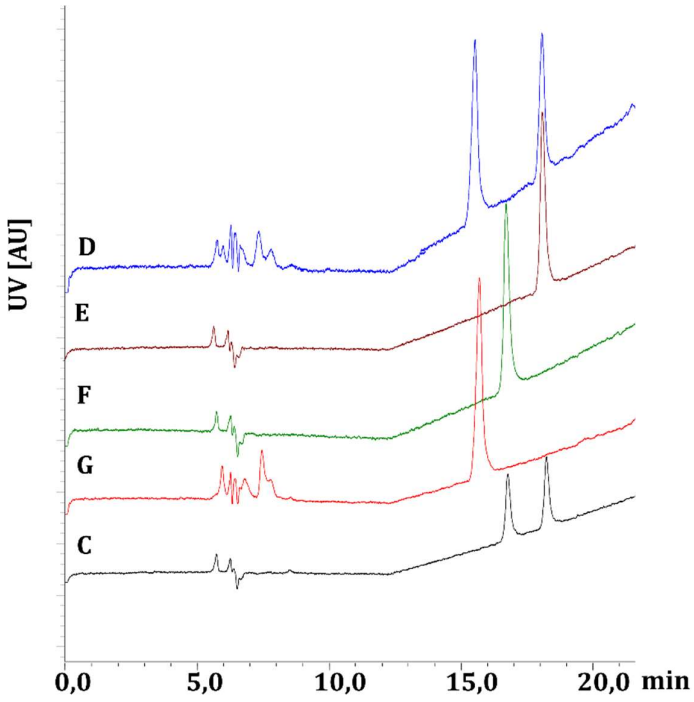


Figure 37. HPLC analysis of L-α-homophenylalanine and 6xHis-AdmH. 6xHis-AdmH was incubated with L-α-homophenylalanine (D) and the obtained chromatogram was compared to an L-α-homophenylalanine standard (E), an L-β-homophenylalanine standard (F), the purified 6xHis-AdmH without substrate addition (G) and a standard mix, prepared from L-α-homophenylalanine and L-β-homophenylalanine (C). All chromatograms were recorded at 260 nm on a Sphinx column with a gradient set from 10 to 100% MeOH within 20 min, 100% MeOH over 5 min and from 100% to 10% MeOH in 2 min.

4.5.2 Discussion

The aminomutase AdmH was not able to convert L-α-homophenylalanine into the corresponding L-β-amino acid. Although aminomutases were previously shown to possess a broad substrate promiscuity in respect to modified aromatic moieties [245,316,320], the herein conducted experiments revealed that substrates with longer carbon side-chains might not be tolerated by AdmH. Therefore, the devised pathway for the *in vitro* production of MPBA derivatives is not feasible. The testing of the ammonia lyase of *N. punctiforme* for the conversion of β-homophenylalanine into PBA was hence not further pursued.

In perspective, alternative routes must be considered for the biosynthesis of PBA or other MPBA derivatives. A possible alternative might be found in enzyme engineering of AdmH for the acceptance of the alternative substrate homophenylalanine. The crystal structure of AdmH was elucidated some years ago [321] and engineering approaches were already successful, increasing the stability and productivity of AdmH for the production of β-amino acids at higher temperatures (50 °C) [322]. A mutagenesis approaches in the active site of a related phenylalanine aminomutase

from *Taxus chinensis* further improved the selectivity of the enzyme and resulted in the acceptance of alternative substrates with a methoxy-, or methyl-substituted phenyl ring [323].

Another option may be the testing of a phenylalanine ammonia lyase to accept and directly catalyze the conversion of a provided L-β-homophenylalanine substrate to form PBA. Naturally, phenylalanine ammonia lyases catalyze the reversible deamination of L-phenylalanine to *trans*-cinnamic acid and exhibit a broad substrate specificity for halogenated phenylalanines [317,319]. Although L-β-homophenylalanine has not been tested as alternative substrate so far, enzyme engineering based on the available crystal structures of various phenylalanine ammonia lyases [324,325], e.g. from *Nostoc punctiforme*, or *Anabaena variabilis* [324], is also possible. Engineering approaches, targeting active site residues, or the aryl binding pocket, showed good results, e.g. for an increased catalytic efficiency and the acceptance of alternative substrates, yielding e.g. alkyl and methoxy phenylalanine derivatives [326], or 3-nitro cinnamic acid and diverse bromo- and fluor-substituted cinnamic acids [327].

The application of CrpA for an *in vitro* production of MPBA may also be an option, as soluble enzyme could successfully be produced in the herein conducted *E. coli* studies (cf. section 4.4). However, it has to be noted that compared to AdmH (61.4 kDa), CrpA is relatively large (e.g. MBP-CrpA: 366 kDa), requires the assistance of an activating PPTase and the supplementation of cofactors (e.g. ATP, coenzyme A, Mg^{2+}, NADPH) (cf. section 1.2). Further, a crystal structure of the enzyme is hitherto not available and its exact reaction mechanism is still elusive. Although interesting insights into its functioning could be acquired by an *in vitro* attempt, altogether these circumstances do not favor the application of CrpA for the *in vitro* production of MPBA.

5 Final remarks

Cyanobacteria exhibit a high biosynthetic potential and are a prolific resource for the discovery of novel drug leads and valuable industrial compounds, which was also shown in the course of the herein conducted sequencing project. Yet these resources remain rather unexploited in terms of (pharma-)industrial applications. This is mainly due to the slow growth of many cyanobacteria, cultivation difficulties, low production titers, silent BGCs, or genetic instabilities of the encoding BGCs. Furthermore, the restricted amenability to genetic engineering, as well as a reduced availability and applicability of molecular biological tools impede the development of suitable producers to target the activation or genetic optimization of these BGCs producing the required amounts of a desired compound on an industrial scale [23,109–111].

In the course of this thesis *S. cerevisiae* and *E. coli* were examined for the heterologous expression of cyanobacterial BGCs. This included the assembly and engineering of small cyanobacterial BGCs, up to multimodular NRPS-PKS systems. Moreover, *in vivo* systems were devised and applied successfully to reconstitute various enzymatic conversions. My studies showed that extensive engineering is required to reconstruct and establish a biosynthetic pathway in a new host system, especially when traversing the borders of prokaryotic and eukaryotic life. They also showcased bottlenecks in the chosen host organisms.

In this PhD project a plasmid-based cloning system was developed, which facilitated the assembly of large BGCs of bacterial origin in yeast. For this purpose, minimal-artificial expression cassettes were combined with established genome editing techniques, e.g. CRISPR-Cas9 and LTR-recombination. This led to an efficiently minimized cloning effort, which was illustrated by successful reconstruction of the >35 kb-sized cryptophycin BGC (*crpA-D*, *crpH*) in yeast. The heterologous production of the target compound was extremely challenging, possibly due to an increased metabolic burden, the scarcity of biosynthetic precursors, or translational limitations resulting from deviations in codon usage. To improve cryptophycin production in yeast, further efforts must be directed towards precursor pool optimization, tRNA pool adaptation, and gene expression balancing.

The heterologous expression of the much smaller scytonemin BGC demonstrated the multilayered interplay of an integrated pathway with the primary network of the host strain. Engineering of yeast's tryptophan metabolism was necessary to shut down a competing precursor flow towards anthranilate. Even though the engineering effort turned out successful, the biosynthesis of the desired sunscreen pigment could not be realized. Further bottlenecks, e.g. the degradation of pathway intermediates and the adequate functionality of all heterologous enzymes, must be addressed in the future. Engineering of the biosynthesis genes was necessary in case of the conducted *E. coli* studies. Here, inclusion-body formation initially affected the

functional reconstruction of the protein CrpA. However, the application of solubility tags solved this problem. Another challenge during the reconstruction of CrpA was the cleavage of the protein-bound product, which was addressed by the fusion of a TE domain.

At the end of this project the question arises, if the chosen host systems should still be considered as candidates for the expression of cyanobacterial BGCs, or if alternative host systems might be more suitable. Neither *E. coli*, nor *S. cerevisiae* are native producers of secondary metabolites. Prior to this study, however, *E. coli* was already successfully used for the heterologous production of secondary metabolites from cyanobacteria [23]. Therefore, it can be assumed that the difficulties encountered in this PhD project are specific for the cryptophycin assembly line. This means that *E. coli* represents a potential host for cyanobacterial metabolites. On the contrary, the eukaryote *S. cerevisiae* was hardly tested for the biosynthesis of cyanobacterial compounds [13,23]. Although the use of this host is attractive from an industrial perspective, extensive engineering might be necessary to enable the production of secondary metabolites from multimodular PKS and NRPS assembly lines.

Recently some cyanobacteria demonstrated their usefulness as hosts for the reconstruction and expression of BGCs from other cyanobacteria, which may be reasoned with a similar genetic background [23,119]. Examples are the filamentous, non-branching *Anabaena* PCC 7120, or the unicellular *Synechococcus* sp. UTEX 2973 and *Synechocystis* sp. PCC6803, which were successfully applied for the production of lyngbyatoxin A, cryptomaldamide, hapalindole H and shinorine [23]. The close phylogenetic relation and a similar cellular organization may support the biosynthesis of compounds with specific biosynthetic requirements, such as posttranslational modifications, or cellular site-specific biosynthesis mechanisms (e.g. periplasmic-transport of intermediates and assembly of final metabolites). Therefore, in particular for cyanobacterial BGCs that feature biosynthetic peculiarities, such as the unusual truncation of the CrpA starter unit (cf. section 1.2.1), a cyanobacterial host may be ideally suited. In perspective it might be worthwhile to design a cyanobacterial chassis organism as a workhorse for industrial applications, harnessing the biosynthetic potential of this interesting phylum in terms of secondary metabolite production. A few companies, e.g. Cyano Biotech (2004, Germany) [328,329], Living Ink Technology (2013, USA) [329,330] and Photanol (2008, Netherlands) [329,331] are already applying diverse cyanobacteria for the low-scale industrial production of cyanobacterial secondary metabolites as analytical standards (e.g. microcystin, nodularin, anatoxin, aeruginosin, microginin) [332,333], or large scale production of eco-friendly dyes and carbon-compounds (e.g. lactic acid, glycolic acid, ethylene, propylene, ethanol) [331]. Although a commercial cyanobacterial chassis is currently not available, the development of "recombinant microbial producer strains" [334] is already in the industrial focus.

6 Literature

(1) Madigan, M. T.; Martinko, J. M.; Bender, K. S.; Buckley, D. H.; Stahl, D. A.; Brock, T. Cyanobacteria. In *Brock Biology of Microorgansims*; Pearson, 2014; pp 436–443.

(2) Fuchs, G.; Schlegel, H.-G.; Eitinger, T.; Heider, J.; Kemper, B.; Kothe, E.; Schink, B.; Schneider, E.; Unden, G. Phototrophe Lebensweise. In *Allgemeine Mikrobiologie*; Fuchs, G., Schlegel, H.-G., Eds.; Georg Thieme Verlag: Stuttgart, 2007; pp 407–416. https://doi.org/10.1055/b-0034-45836.

(3) Kehr, J.-C.; Gatte Picchi, D.; Dittmann, E. Natural Product Biosyntheses in Cyanobacteria: A Treasure Trove of Unique Enzymes. *Beilstein J. Org. Chem.* **2011**, *7* (1), 1622–1635. https://doi.org/10.3762/bjoc.7.191.

(4) Rai, A. N.; Söderbäck, E.; Bergman, B. Tansley Review No. 116, Cyanobacterium-Plant Symbioses. *New Phytol.* **2000**, *147* (3), 449–481. https://doi.org/10.1046/j.1469-8137.2000.00720.x.

(5) Dvořák, P.; Poulíčková, A.; Hašler, P.; Belli, M.; Casamatta, D. A.; Papini, A. Species Concepts and Speciation Factors in Cyanobacteria, with Connection to the Problems of Diversity and Classification. *Biodiversity and Conservation*. Kluwer Academic Publishers April 1, 2015, pp 739–757. https://doi.org/10.1007/s10531-015-0888-6.

(6) Rikkinen, J. Cyanobacteria in Terrestrial Symbiotic Systems. In *Modern Topics in the Phototrophic Prokaryotes*; Springer International Publishing: Cham, 2017; pp 243–294. https://doi.org/10.1007/978-3-319-46261-5_8.

(7) Camargo, S.; Leshkowitz, D.; Dassa, B.; Mariscal, V.; Flores, E.; Stavans, J.; Arbel-Goren, R. Impaired Cell-Cell Communication in the Multicellular Cyanobacterium Anabaena Affects Carbon Uptake, Photosynthesis, and the Cell Wall. *iScience* **2021**, *24* (1), 101977. https://doi.org/10.1016/j.isci.2020.101977.

(8) Flores, E.; Nieves-Morión, M.; Mullineaux, C. Cyanobacterial Septal Junctions: Properties and Regulation. *Life* **2018**, *9* (1), 1. https://doi.org/10.3390/life9010001.

(9) Mariscal, V.; Herrero, A.; Flores, E. Continuous Periplasm in a Filamentous, Heterocyst-Forming Cyanobacterium. *Mol. Microbiol.* **2007**, *65* (4), 1139–1145. https://doi.org/10.1111/j.1365-2958.2007.05856.x.

(10) Pereira, S.; Zille, A.; Micheletti, E.; Moradas-Ferreira, P.; De Philippis, R.; Tamagnini, P. Complexity of Cyanobacterial Exopolysaccharides: Composition, Structures, Inducing Factors and Putative Genes Involved in Their Biosynthesis and Assembly. *FEMS Microbiol. Rev.* **2009**, *33* (5), 917–941. https://doi.org/10.1111/j.1574-6976.2009.00183.x.

(11) Adams, D. G.; Duggan, P. S. Tansley Review No. 107, Heterocyst and Akinete Differentiation in Cyanobacteria. *New Phytol.* **1999**, *144* (1), 3–33. https://doi.org/10.1046/j.1469-8137.1999.00505.x.

(12) Marsac, N. T. Differentiation of Hormogonia and Relationships with Other Biological Processes. In *The Molecular Biology of Cyanobacteria*; Springer Netherlands: Dordrecht, 1994; pp 825–842. https://doi.org/10.1007/978-94-011-0227-8_28.

(13) Tippelt, A.; Nett, M. Saccharomyces Cerevisiae as Host for the Recombinant Production of Polyketides and Nonribosomal Peptides. *Microb. Cell Fact.* **2021**, *20* (1), 161. https://doi.org/10.1186/s12934-021-01650-y.

(14) Singh, R. K.; Prakash Tiwari, S.; Rai, A. K.; Mohapatra, T. M. Cyanobacteria: An Emerging Source for Drug Discovery. *J. Antibiot. (Tokyo).* **2011**, *64*, 401–412. https://doi.org/10.1038/ja.2011.21.

(15) Demain, A. L.; Fang, A. The Natural Functions of Secondary Metabolites. *Adv. Biochem. Eng. Biotechnol.* **2000**, *69*, 1–39. https://doi.org/10.1007/3-540-44964-7_1.

(16) Hibbing, M. E.; Fuqua, C.; Parsek, M. R.; Peterson, S. B. Bacterial Competition: Surviving and Thriving in the Microbial Jungle. *Nat. Rev. Microbiol.* **2010**, *8* (1), 15–25. https://doi.org/10.1038/nrmicro2259.

(17) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

(18) Barrios-González, J. Secondary Metabolites Production: Physiological Advantages in Solid-State Fermentation. In *Current Developments in Biotechnology and Bioengineering*; Pandey, A., Larroche, C., Soccol, C. R., Eds.; Elsevier, 2018; pp 257–283. https://doi.org/10.1016/b978-0-444-63990-5.00013-x.

(19) Thirumurugan, D.; Cholarajan, A.; Raja, S. S. S.; Vijayakumar, R. An Introductory Chapter: Secondary Metabolites. In *Secondary Metabolites - Sources and Applications*; Vijayakumar, R., Raja, S. S. S., Eds.; InTech, 2018; pp 3–21. https://doi.org/10.5772/intechopen.79766.

(20) Verpoorte, R.; van der Heijden, R.; Memelink, J. Engineering the Plant Cell Factory for Secondary Metabolite Production. *Transgenic Res.* **2000**, *9* (4–5), 323–343; discussion 321. https://doi.org/10.1023/a:1008966404981.

(21) Tan, L. T. Bioactive Natural Products from Marine Cyanobacteria for Drug Discovery. *Phytochemistry* **2007**, *68* (7), 954–979. https://doi.org/10.1016/J.PHYTOCHEM.2007.01.012.

(22) Korp, J.; Vela Gurovic, M. S.; Nett, M. Antibiotics from Predatory Bacteria. *Beilstein J. Org. Chem* **2016**, *12*, 594–607. https://doi.org/10.3762/bjoc.12.58.

(23) Dhakal, D.; Chen, M.; Luesch, H.; Ding, Y. Heterologous Production of Cyanobacterial Compounds. *J. Ind. Microbiol. Biotechnol.* **2021**, *48* (3–4), 3. https://doi.org/10.1093/jimb/kuab003.

(24) Dittmann, E.; Gugger, M.; Sivonen, K.; Fewer, D. P. Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends in Microbiology*. Elsevier Ltd October 1, 2015, pp 642–652. https://doi.org/10.1016/j.tim.2015.07.008.

(25) Ziemert, N.; Alanjary, M.; Weber, T. The Evolution of Genome Mining in Microbes – a Review. *Nat. Prod. Rep.* **2016**, *33* (8), 988–1005. https://doi.org/10.1039/C6NP00025H.

(26) Nett, M. *Genome Mining: Concept and Strategies for Natural Product Discovery.*; 2014; Vol. 99. https://doi.org/10.1007/978-3-319-04900-7_4.

(27) Finking, R.; Marahiel, M. A. Biosynthesis of Nonribosomal Peptides. *Annu. Rev. Microbiol.* **2004**, *58* (1), 453–488. https://doi.org/10.1146/annurev.micro.58.030603.123615.

(28) Mofid, M. R.; Finking, R.; Essen, L. O.; Marahiel, M. A. Structure-Based Mutational Analysis of the 4'-Phosphopantetheinyl Transferases Sfp from Bacillus Subtilis : Carrier Protein Recognition and Reaction Mechanism „. *Biochemistry* **2004**, *43* (14), 4128–4136. https://doi.org/10.1021/bi036013h.

(29) Copp, J. N.; Neilan, B. A. The Phosphopantetheinyl Transferase Superfamily: Phylogenetic Analysis and Functional Implications in Cyanobacteria. *Appl. Environ. Microbiol.* **2006**, *72* (4), 2298–2305. https://doi.org/10.1128/AEM.72.4.2298-2305.2006.

(30) 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.* **2014**, *31* (1), 61–108.

https://doi.org/10.1039/C3NP70054B.

(31) Reeves, C. D.; Murli, S.; Ashley, G. W.; Piagentini, M.; Hutchinson, C. R.; McDaniel, R. Alteration of the Substrate Specificity of a Modular Polyketide Synthase Acyltransferase Domain through Site-Specific Mutations. *Biochemistry* **2001**, *40* (51), 15464–15470. https://doi.org/10.1021/bi015864r.

(32) Shen, B. Polyketide Biosynthesis beyond the Type I, II and III Polyketide Synthase Paradigms. *Curr. Opin. Chem. Biol.* **2003**, *7* (2), 285–295. https://doi.org/10.1016/S1367-5931(03)00020-6.

(33) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chemie Int. Ed.* **2009**, *48* (26), 4688–4716. https://doi.org/10.1002/anie.200806121.

(34) Rath, C. M.; Scaglione, J. B.; Kittendorf, J. D.; Sherman, D. H. NRPS/PKS Hybrid Enzymes and Their Natural Products. In *Comprehensive Natural Products II - Chemistry and Biology*; Elsevier, 2010; pp 453–492. https://doi.org/10.1016/B978-008045382-8.00725-5.

(35) Keating, T. A.; Walsh, C. T. Initiation, Elongation, and Termination Strategies in Polyketide and Polypeptide Antibiotic Biosynthesis. *Curr. Opin. Chem. Biol.* **1999**, *3* (5), 598–606. https://doi.org/10.1016/S1367-5931(99)00015-0.

(36) Du, L.; Lou, L. PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* **2010**, *27* (2), 255–278. https://doi.org/10.1039/B912037H.

(37) Walsh, C. T. Polyketide and Nonribosomal Peptide Antibiotics: Modularity and Versatility. *Science (80-.).* **2004**, *303* (5665), 1805–1810. https://doi.org/10.1126/science.1094318.

(38) Donadio, S.; Staver, M.; McAlpine, J.; Swanson, S.; Katz, L. Modular Organization of Genes Required for Complex Polyketide Biosynthesis. *Science (80-.).* **1991**, *252* (5006), 675–679. https://doi.org/10.1126/science.2024119.

(39) Mcdaniel, R.; Thamchaipenet, A.; Gustafsson, C.; Fu, H.; Betlach, M.; 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. U. S. A.* **1999**, *96* (5), 1846–1851. https://doi.org/https://doi.org/10.1007/978-3-319-04900-7_4.

(40) Koryakina, I.; McArthur, J. B.; Draelos, M. M.; Williams, G. J. Promiscuity of a Modular Polyketide Synthase towards Natural and Non-Natural Extender Units. *Org. Biomol. Chem.* **2013**, *11* (27), 4449. https://doi.org/10.1039/c3ob40633d.

(41) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The Specificity-Conferring Code of Adenylation Domains in Nonribosomal Peptide Synthetases. *Chem. Biol.* **1999**, *6* (8), 493–505. https://doi.org/10.1016/S1074-5521(99)80082-9.

(42) Strieker, M.; Tanović, A.; Marahiel, M. A. Nonribosomal Peptide Synthetases: Structures and Dynamics. *Curr. Opin. Struct. Biol.* **2010**, *20* (2), 234–240. https://doi.org/10.1016/j.sbi.2010.01.009.

(43) Weber, T.; Marahiel, M. A. Exploring the Domain Structure of Modular Nonribosomal Peptide Synthetases. *Structure*. Elsevier Ltd January 1, 2001, p R3. https://doi.org/10.1016/S0969-2126(00)00560-8.

(44) Pang, B.; Chen, Y.; Gan, F.; Yan, C.; Jin, L.; Gin, J. W.; Petzold, C. J.; Keasling, J. D. Investigation of Indigoidine Synthetase Reveals a Conserved Active-Site Base Residue of Nonribosomal Peptide Synthetase Oxidases. *J. Am. Chem. Soc.* **2020**, *142* (25), 10931–10935. https://doi.org/10.1021/jacs.0c04328.

(45) Keating, T. A.; Ehmann, D. E.; Kohli, R. M.; Marshall, C. G.; Trauger, J. W.; Walsh, C. T. Chain Termination Steps in Nonribosomal Peptide Synthetase Assembly Lines: Directed Acyl-S-Enzyme Breakdown in Antibiotic and Siderophore Biosynthesis. *ChemBioChem* **2001**, *2* (2), 99–107. https://doi.org/10.1002/1439-7633(20010202)2:2<99::AID-CBIC99>3.0.CO;2-3.

(46) Hur, G. H.; Vickery, C. R.; Burkart, M. D. Explorations of Catalytic Domains in Non-Ribosomal Peptide Synthetase Enzymology. *Natural Product Reports*. Royal Society of Chemistry 2012, pp 1074–1098. https://doi.org/10.1039/c2np20025b.

(47) Süssmuth, R. D.; Mainz, A. Nonribosomal Peptide Synthesis-Principles and Prospects. *Angew. Chemie Int. Ed.* **2017**, *56* (14), 3770–3821. https://doi.org/10.1002/anie.201609079.

(48) Schneider, T. L.; Shen, B.; Walsh, C. T. Oxidase Domains in Epothilone and Bleomycin Biosynthesis: Thiazoline to Thiazole Oxidation during Chain Elongation. *Biochemistry* **2003**, *42* (32), 9722–9730. https://doi.org/10.1021/bi034792w.

(49) Finking, R.; Marahiel, M. A. Biosynthesis of Nonribosomal Peptides. *Annu. Rev. Microbiol.* **2004**, *58* (1), 453–488. https://doi.org/10.1146/annurev.micro.58.030603.123615.

(50) May, J. J.; Kessler, N.; Marahiel, M. A.; Stubbs, M. T. Crystal Structure of DhbE, an Archetype for Aryl Acid Activating Domains of Modular Nonribosomal Peptide Synthetases. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (19), 12120–12125. https://doi.org/10.1073/pnas.182156699.

(51) Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C.; Luo, L.; 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* (5), 525–534. https://doi.org/10.1016/s1367-5931(00)00235-0.

(52) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chemie Int. Ed.* **2009**, *48* (26), 4688–4716. https://doi.org/10.1002/anie.200806121.

(53) Rix, U.; Fischer, C.; Remsing, L. L.; Rohr, J. Modification of Post-PKS Tailoring Steps through Combinatorial Biosynthesis. *Nat. Prod. Rep.* **2002**, *19* (5), 542–580. https://doi.org/10.1039/b103920m.

(54) Pohl, N. L.; Gokhale, R. S.; Cane, D. E.; Khosla, C. Synthesis and Incorporation of an N -Acetylcysteamine Analogue of Methylmalonyl-CoA by a Modular Polyketide Synthase. *J. Am. Chem. Soc.* **1998**, *120* (43), 11206–11207. https://doi.org/10.1021/ja9830290.

(55) Christiansen, G.; Philmus, B.; Hemscheidt, T.; Kurmayer, R. Genetic Variation of Adenylation Domains of the Anabaenopeptin Synthesis Operon and Evolution of Substrate Promiscuity. *J. Bacteriol.* **2011**, *193* (15), 3822–3831. https://doi.org/10.1128/JB.00360-11.

(56) Winand, L.; Sester, A.; Nett, M. Bioengineering of Anti-Inflammatory Natural Products. *ChemMedChem* **2021**, *16* (5), 767–776. https://doi.org/10.1002/cmdc.202000771.

(57) Natsume, T.; Watanabe, J. I.; Koh, Y.; Fujio, N.; Ohe, Y.; Horiuchi, T.; Saijo, N.; Nishio, K.; Kobayashi, M. Antitumor Activity of TZT-1027 (Soblidotin) against Vascular Endothelial Growth Factor-Secreting Human Lung Cancer in Vivo. *Cancer Sci.* **2003**, *94* (9), 826–833. https://doi.org/10.1111/j.1349-7006.2003.tb01526.x.

(58) Cunningham, C.; Appleman, L. J.; Kirvan-Visovatti, M.; Ryan, D. P.; Regan, E.; Vukelja, S.; Bonate, P. L.; Ruvuna, F.; Fram, R. J.; Jekunen, A.; Weitman, S.; Hammond, L. A.; Eder, J. P. Phase I and Pharmacokinetic Study of the

Dolastatin-15 Analogue Tasidotin (ILX651) Administered Intravenously on Days 1, 3, and 5 Every 3 Weeks in Patients with Advanced Solid Tumors. *Clin. Cancer Res.* **2005**, *11* (21), 7825–7833. https://doi.org/10.1158/1078-0432.CCR-05-0058.

(59) Edelman, M. J.; Gandara, D. R.; Hausner, P.; Israel, V.; Thornton, D.; DeSanto, J.; Doyle, L. A. Phase 2 Study of Cryptophycin 52 (LY355703) in Patients Previously Treated with Platinum Based Chemotherapy for Advanced Non-Small Cell Lung Cancer. *Lung Cancer* **2003**, *39* (2), 197–199. https://doi.org/10.1016/S0169-5002(02)00511-1.

(60) D'Agostino, G.; Del Campo, J.; Mellado, B.; Izquierdo, M. A.; Minarik, T.; Cirri, L.; Marini, L.; Perez-Gracia, J. L.; Scambia, G. A Multicenter Phase II Study of the Cryptophycin Analog LY355703 in Patients with Platinum-Resistant Ovarian Cancer. *Int. J. Gynecol. Cancer* **2006**, *16* (1), 71–76. https://doi.org/10.1111/j.1525-1438.2006.00276.x.

(61) Rohr, J. Cryptophycin Anticancer Drugs Revisited. *ACS Chem. Biol.* **2006**, *1* (12), 747–750. https://doi.org/10.1021/cb6004678.

(62) Liang, J.; Moore, R. E.; Moher, E. D.; Munroe, J. E.; Al-awar, R. S.; Hay, D. A.; Varie, D. L.; Zhang, T. Y.; Aikins, J. A.; Martinelli, M. J.; Shih, C.; Ray, J. E.; Gibson, L. L.; Vasudevan, V.; Polin, L.; White, K.; Kushner, J.; Simpson, C.; Pugh, S.; Corbett, T. H. Cryptophycins-309, 249 and Other Cryptophycin Analogs: Preclinical Efficacy Studies with Mouse and Human Tumors. *Invest. New Drugs* **2005**, *23* (3), 213–224. https://doi.org/10.1007/s10637-005-6729-9.

(63) Hearn, B. R.; Shaw, S. J.; Myles, D. C. Microtubule Targeting Agents. In *Comprehensive Medicinal Chemistry II*; Elsevier, 2007; pp 81–110. https://doi.org/10.1016/B0-08-045044-X/00205-4.

(64) Magarvey, N. A.; Beck, Z. Q.; Golakoti, T.; Ding, Y.; Huber, U.; Hemscheidt, T. K.; Abelson, D.; Moore, R. E.; Sherman, D. H. Biosynthetic Characterization and Chemoenzymatic Assembly of the Cryptophycins. Potent Anticancer Agents from Nostoc Cyanobionts. *ACS Chem. Biol.* **2006**, *1* (12), 766–779. https://doi.org/10.1021/cb6004307.

(65) Panda, D.; Himes, R. H.; Moore, R. E.; Wilson, L.; Jordan, M. A. Mechanism of Action of the Unusually Potent Microtubule Inhibitor Cryptophycin 1. *Biochemistry* **1997**, *36* (42), 12948–12953. https://doi.org/10.1021/bi971302p.

(66) Eißler, S.; Stoncius, A.; Nahrwold, M.; Sewald, N. The Synthesis of Cryptophycins. *Synthesis (Stuttg).* **2006**, *2006* (22), 3747–3789. https://doi.org/10.1055/s-2006-950332.

(67) Panda, D.; DeLuca, K.; Williams, D.; Jordan, M. A.; Wilson, L. Antiproliferative Mechanism of Action of Cryptophycin-52: Kinetic Stabilization of Microtubule Dynamics by High-Affinity Binding to Microtubule Ends. *Proc. Natl. Acad. Sci.* **1998**, *95* (16), 9313–9318. https://doi.org/10.1073/pnas.95.16.9313.

(68) Panda, D.; Ananthnarayan, V.; Larson, G.; Shih, C.; Jordan, M. A.; Wilson, L. Interaction of the Antitumor Compound Cryptophycin-52 with Tubulin. *Biochemistry* **2000**, *39* (46), 14121–14127. https://doi.org/10.1021/bi0010827.

(69) Lu, K.; Dempsey, J.; Schultz, R. M.; Shih, C.; Teicher, B. A. Cryptophycin-Induced Hyperphosphorylation of Bcl-2, Cell Cycle Arrest and Growth Inhibition in Human H460 NSCLC Cells. *Cancer Chemother. Pharmacol.* **2001**, *47* (2), 170–178. https://doi.org/10.1007/s002800000210.

(70) Golakoti, T.; Yoshida, W. Y.; Chaganty, S.; Moore, R. E. Isolation and Structures of Nostopeptolides A1, A2 and

A3 from the Cyanobacterium Nostoc Sp. GSV224. *Tetrahedron* **2000**, *56* (46), 9093–9102. https://doi.org/10.1016/S0040-4020(00)00764-X.

(71) Subbaraju, G. V; Golakoti, T.; Patterson, G. M. L.; Moore, R. E. Three New Cryptophycins from Nostoc Sp. GSV 224.

(72) Chaganty, S.; Golakoti, T.; Heltzel, C.; Moore, R. E.; Yoshida, W. Y. Isolation and Structure Determination of Cryptophycins 38, 326, and 327 from the Terrestrial Cyanobacterium Nostoc Sp. GSV 224 †. **2004**. https://doi.org/10.1021/np0499665.

(73) Golakoti, T.; Yoshida, W. Y.; Chaganty, S.; Moore, R. E. Isolation and Structure Determination of Nostocyclopeptides A1 and A2 from the Terrestrial Cyanobacterium Nostoc Sp. ATCC53789. *J. Nat. Prod.* **2001**, *64* (1), 54–59. https://doi.org/10.1021/np000316k.

(74) Eggen, M.; Georg, G. I. The Cryptophycins: Their Synthesis and Anticancer Activity. *Med. Res. Rev.* **2002**, *22* (2), 85–101. https://doi.org/10.1002/med.10002.

(75) Tillett, D.; Dittmann, E.; Erhard, M.; von Döhren, H.; Börner, T.; Neilan, B. A. Structural Organization of Microcystin Biosynthesis in Microcystis Aeruginosa PCC7806: An Integrated Peptide-Polyketide Synthetase System. *Chem. Biol.* **2000**, *7* (10), 753–764. https://doi.org/10.1016/s1074-5521(00)00021-1.

(76) Hicks, L. M.; Moffitt, M. C.; Beer, L. L.; Moore, B. S.; Kelleher, N. L. Structural Characterization of in Vitro and in Vivo Intermediates on the Loading Module of Microcystin Synthetase. *ACS Chem. Biol.* **2006**, *1* (2), 93–102. https://doi.org/10.1021/cb500007v.

(77) Ding, Y.; Seufert, W. H.; Beck, Z. Q.; Sherman, D. H. Analysis of the Cryptophycin P450 Epoxidase Reveals Substrate Tolerance and Cooperativity. https://doi.org/10.1021/ja710520q.

(78) Frankmolle, W. P.; Knubel, G.; Moore, R. E.; Patterson, G. M. L. Antifungal Cyclic Peptides from the Terrestrial Blue-Green Alga Anabaena Laxa II. Structures of Laxaphycins a, b, d and E. *J. Antibiot. (Tokyo).* **1992**, *45* (9), 1458–1466. https://doi.org/10.7164/antibiotics.45.1458.

(79) Christiansen, G.; Fastner, J.; Erhard, M.; Börner, T.; Dittmann, E. Microcystin Biosynthesis in Planktothrix: Genes, Evolution, and Manipulation. *J. Bacteriol.* **2003**, *185* (2), 564–572. https://doi.org/10.1128/JB.185.2.564-572.2003.

(80) Tan, L. T. Bioactive Natural Products from Marine Cyanobacteria for Drug Discovery. *Phytochemistry* **2007**, *68* (7), 954–979. https://doi.org/10.1016/J.PHYTOCHEM.2007.01.012.

(81) Levine, E.; Thiel, T. UV-Inducible DNA Repair in the Cyanobacteria Anabaena Spp. *J. Bacteriol.* **1987**, *169* (9), 3988–3993. https://doi.org/10.1128/JB.169.9.3988-3993.1987.

(82) Bebout, B. M.; Garcia-Pichel, F. UV B-Induced Vertical Migrations of Cyanobacteria in a Microbial Mat. *Appl. Environ. Microbiol.* **1995**, *61* (12), 4215 LP – 4222.

(83) Balskus, E. P.; Walsh, C. T. Investigating the Initial Steps in the Biosynthesis of Cyanobacterial Sunscreen Scytonemin. *J. Am. Chem. Soc.* **2008**, *130* (46), 15260–15261. https://doi.org/10.1021/ja807192u.

(84) COCKELL, C. S.; KNOWLAND, J. Ultraviolet Radiation Screening Compounds. *Biol. Rev. Camb. Philos. Soc.* **1999**, *74* (3), S0006323199005356. https://doi.org/10.1017/S0006323199005356.

(85) Matsui, K.; Nazifi, E.; Hirai, Y.; Wada, N.; Matsugo, S.; Sakamoto, T. The Cyanobacterial UV-Absorbing Pigment Scytonemin Displays Radical-Scavenging Activity. *J. Gen. Appl. Microbiol.* **2012**, *58* (2), 137–144.

https://doi.org/10.2323/jgam.58.137.

(86) Dillon, J. G.; Tatsumi, C. M.; Tandingan, P. G.; Castenholz, R. W. Effect of Environmental Factors on the Synthesis of Scytonemin, a UV-Screening Pigment, in a Cyanobacterium (Chroococcidiopsis Sp.). *Arch. Microbiol.* **2002**, *177* (4), 322–331. https://doi.org/10.1007/s00203-001-0395-x.

(87) Pandey, R.; Chauhan, S.; Singhal, G. S. UVB-Induced Photodamage to Phycobilisomes of Synechococcus Sp. PCC 7942. *J. Photochem. Photobiol. B Biol.* **1997**, *40* (3), 228–232. https://doi.org/10.1016/S1011-1344(97)00060-2.

(88) Sorrels, C. M.; Proteau, P. J.; Gerwick, W. H. Organization, Evolution, and Expression Analysis of the Biosynthetic Gene Cluster for Scytonemin, a Cyanobacterial UV-Absorbing Pigment. *Appl. Environ. Microbiol.* **2009**, *75* (14), 4861–4869. https://doi.org/10.1128/AEM.02508-08.

(89) Proteau, P. J.; Gerwick, W. H.; Garcia-Pichel, F.; Castenholz, R. The Structure of Scytonemin, an Ultraviolet Sunscreen Pigment from the Sheaths of Cyanobacteria. *Experientia* **1993**, *49* (9), 825–829. https://doi.org/10.1007/BF01923559.

(90) Garcia-Pichel, F.; Castenholz, R. W. CHARACTERIZATION AND BIOLOGICAL IMPLICATIONS OF SCYTONEMIN, A CYANOBACTERIAL SHEATH PIGMENT1. *J. Phycol.* **1991**, *27* (3), 395–409. https://doi.org/10.1111/j.0022-3646.1991.00395.x.

(91) Balskus, E. P.; Walsh, C. T. An Enzymatic Cyclopentyl[b]Indole Formation Involved in Scytonemin Biosynthesis. *J. Am. Chem. Soc.* **2009**, *131* (41), 14648–14649. https://doi.org/10.1021/ja906752u.

(92) Ferreira, D.; Garcia-Pichel, F. Mutational Studies of Putative Biosynthetic Genes for the Cyanobacterial Sunscreen Scytonemin in Nostoc Punctiforme ATCC 29133. *Front. Microbiol.* **2016**, *7*. https://doi.org/10.3389/fmicb.2016.00735.

(93) Malla, S.; Sommer, M. O. A. A Sustainable Route to Produce the Scytonemin Precursor Using Escherichia Coli. *Green Chem.* **2014**, *16* (6), 3255–3265. https://doi.org/10.1039/C4GC00118D.

(94) Balskus, E. P.; Case, R. J.; Walsh, C. T. The Biosynthesis of Cyanobacterial Sunscreen Scytonemin in Intertidal Microbial Mat Communities. *FEMS Microbiol. Ecol.* **2011**, *77* (2), 322–332. https://doi.org/10.1111/j.1574-6941.2011.01113.x.

(95) Klicki, K.; Ferreira, D.; Hamill, D.; Dirks, B.; Mitchell, N.; Garcia-Pichel, F. The Widely Conserved Ebo Cluster Is Involved in Precursor Transport to the Periplasm during Scytonemin Synthesis in Nostoc Punctiforme. *MBio* **2018**, *9* (6). https://doi.org/10.1128/mBio.02266-18.

(96) Reisman, S. E.; Maimone, T. J. Total Synthesis of Complex Natural Products: More Than a Race for Molecular Summits. *Acc. Chem. Res.* **2021**, *54* (8), 1815–1816. https://doi.org/10.1021/acs.accounts.1c00184.

(97) Sheldon, R. A.; Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **2018**, *118* (2), 801–838. https://doi.org/10.1021/acs.chemrev.7b00203.

(98) Siddiqui, M. S.; Thodey, K.; Trenchard, I.; Smolke, C. D. Advancing Secondary Metabolite Biosynthesis in Yeast with Synthetic Biology Tools. *FEMS Yeast Res.* **2012**, *12* (2), 144–170. https://doi.org/10.1111/j.1567-1364.2011.00774.x.

(99) Kinner, A.; Nerke, P.; Siedentop, R.; Steinmetz, T.; Classen, T.; Rosenthal, K.; Nett, M.; Pietruszka, J.; Lütz, S. Recent Advances in Biocatalysis for Drug Synthesis. *Biomedicines* **2022**, *10* (5), 964. https://doi.org/10.3390/biomedicines10050964.

(100) Choi, W. J. Biotechnological Production of Enantiopure Epoxides by Enzymatic Kinetic Resolution. *Appl. Microbiol. Biotechnol.* **2009**, *84* (2), 239–247. https://doi.org/10.1007/S00253-009-2110-9.

(101) Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Keßeler, M.; Stürmer, R.; Zelinski, T. Industrial Methods for the Production of Optically Active Intermediates. *Angew. Chemie Int. Ed.* **2004**, *43* (7), 788–824. https://doi.org/10.1002/ANIE.200300599.

(102) Jongedijk, E.; Cankar, K.; Buchhaupt, M.; Schrader, J.; Bouwmeester, H.; Beekwilder, J. Biotechnological Production of Limonene in Microorganisms. *Appl. Microbiol. Biotechnol.* **2016**, *100* (7), 2927–2938. https://doi.org/10.1007/s00253-016-7337-7.

(103) Rumbold, K.; van Buijsen, H. J. J.; Overkamp, K. M.; van Groenestijn, J. W.; Punt, P. J.; van der Werf, M. J. Microbial Production Host Selection for Converting Second-Generation Feedstocks into Bioproducts. *Microb. Cell Fact.* **2009**, *8* (1), 64. https://doi.org/10.1186/1475-2859-8-64.

(104) Lin, B.; Tao, Y. Whole-Cell Biocatalysts by Design. *Microbial Cell Factories*. BioMed Central Ltd. June 13, 2017, p 106. https://doi.org/10.1186/s12934-017-0724-7.

(105) Kampranis, S. C.; Makris, A. M. Developing a Yeast Cell Factory for the Production of Terpenoids. *Comput. Struct. Biotechnol. J.* **2012**, *3* (4), e201210006. https://doi.org/10.5936/csbj.201210006.

(106) Becker, M.; Lütz, S.; Rosenthal, K. Environmental Assessment of Enzyme Production and Purification. *Molecules* **2021**, *26* (3), 573. https://doi.org/10.3390/molecules26030573.

(107) Lin, Z.; Nielsen, J.; Liu, Z. Bioprospecting Through Cloning of Whole Natural Product Biosynthetic Gene Clusters. *Front. Bioeng. Biotechnol.* **2020**, *8*, 526. https://doi.org/10.3389/fbioe.2020.00526.

(108) Mitrović, I.; Lukić, N.; Grahovac, M.; Jokić, A.; Dodić, J.; Grahovac, J. Optimization of Streptomyces Hygroscopicus Cultivation Parameters in a Lab-scale Bioreactor. *Chem. Eng. Technol.* **2021**, *44* (2), 349–358. https://doi.org/10.1002/ceat.202000380.

(109) Jones, P. R. Genetic Instability in Cyanobacteria - an Elephant in the Room? *Front. Bioeng. Biotechnol.* **2014**, *2* (MAY), 12. https://doi.org/10.3389/fbioe.2014.00012.

(110) Gordon, G. C.; Pfleger, B. F. Regulatory Tools for Controlling Gene Expression in Cyanobacteria. In *Advances in Experimental Medicine and Biology*; 2018; pp 281–315. https://doi.org/10.1007/978-981-13-0854-3_12.

(111) Videau, P.; Wells, K. N.; Singh, A. J.; Eiting, J.; Proteau, P. J.; Philmus, B. Expanding the Natural Products Heterologous Expression Repertoire in the Model Cyanobacterium Anabaena Sp. Strain PCC 7120: Production of Pendolmycin and Teleocidin B-4. *ACS Synth. Biol.* **2020**, *9* (1), 63–75. https://doi.org/10.1021/acssynbio.9b00334.

(112) Galm, U.; Shen, B. Expression of Biosynthetic Gene Clusters in Heterologous Hosts for Natural Product Production and Combinatorial Biosynthesis. *Expert Opinion on Drug Discovery*. Taylor & Francis October 2006, pp 409–437. https://doi.org/10.1517/17460441.1.5.409.

(113) Yaegashi, J.; Oakley, B. R.; Wang, C. C. C. Recent Advances in Genome Mining of Secondary Metabolite Biosynthetic Gene Clusters and the Development of Heterologous Expression Systems in Aspergillus Nidulans. *J. Ind. Microbiol. Biotechnol.* **2014**, *41* (2), 433–442. https://doi.org/10.1007/s10295-013-1386-z.

(114) Baral, B.; Akhgari, A.; Metsä-Ketelä, M. Activation of Microbial Secondary Metabolic Pathways: Avenues and Challenges. *Synthetic and Systems Biotechnology*. KeAi Communications Co. September 1, 2018, pp 163–178. https://doi.org/10.1016/j.synbio.2018.09.001.

(115) Reen, F. J.; Romano, S.; Dobson, A. D. W.; O'Gara, F. The Sound of Silence: Activating Silent Biosynthetic Gene Clusters in Marine Microorganisms. *Marine Drugs*. MDPI AG August 1, 2015, pp 4754–4783. https://doi.org/10.3390/md13084754.

(116) Netzker, T.; Fischer, J.; Weber, J.; Mattern, D. J.; König, C. C.; Valiante, V.; Schroeckh, V.; Brakhage, A. A. Microbial Communication Leading to the Activation of Silent Fungal Secondary Metabolite Gene Clusters. *Frontiers in Microbiology*. Frontiers Media S.A. 2015, p 299. https://doi.org/10.3389/fmicb.2015.00299.

(117) Adnani, N.; Rajski, S. R.; Bugni, T. S. Symbiosis-Inspired Approaches to Antibiotic Discovery. *Nat. Prod. Rep.* **2017**, *34* (7), 784–814. https://doi.org/10.1039/C7NP00009J.

(118) Sester, A.; Korp, J.; Nett, M. Secondary Metabolism of Predatory Bacteria. In *The Ecology of Predation at the Microscale*; Springer International Publishing, 2020; pp 127–153. https://doi.org/10.1007/978-3-030-45599-6_5.

(119) Videau, P.; Wells, K. N.; Singh, A. J.; Gerwick, W. H.; Philmus, B. Assessment of Anabaena Sp. Strain PCC 7120 as a Heterologous Expression Host for Cyanobacterial Natural Products: Production of Lyngbyatoxin A. *ACS Synth. Biol.* **2016**, *5* (9), 978–988. https://doi.org/10.1021/acssynbio.6b00038.

(120) Hitchcock, A.; Hunter, C. N.; Canniffe, D. P. Progress and Challenges in Engineering Cyanobacteria as Chassis for Light-Driven Biotechnology. *Microb. Biotechnol.* **2020**, *13* (2), 363–367. https://doi.org/10.1111/1751-7915.13526.

(121) Lubertozzi, D.; Keasling, J. D. Developing Aspergillus as a Host for Heterologous Expression. *Biotechnol. Adv.* **2009**, *27* (1), 53–75. https://doi.org/10.1016/j.biotechadv.2008.09.001.

(122) Gómez, S.; Fernández, F. J.; Vega, M. C. Heterologous Expression of Proteins in Aspergillus. In *New and Future Developments in Microbial Biotechnology and Bioengineering: Aspergillus System Properties and Applications*; Gupta, V. K., Ed.; Elsevier, 2016; pp 55–68. https://doi.org/10.1016/B978-0-444-63505-1.00004-X.

(123) Pfeifer, B. A.; Khosla, C. Biosynthesis of Polyketides in Heterologous Hosts. *Microbiol. Mol. Biol. Rev.* **2001**, *65* (1), 106–118. https://doi.org/10.1128/MMBR.65.1.106-118.2001.

(124) Zhang, J. J.; Tang, X.; Moore, B. S. Genetic Platforms for Heterologous Expression of Microbial Natural Products. *Natural Product Reports*. 2019, pp 1313–1332. https://doi.org/10.1039/c9np00025a.

(125) Ziemert, N.; Ishida, K.; Liaimer, A.; Hertweck, C.; Dittmann, E. Ribosomal Synthesis of Tricyclic Depsipeptides in Bloom-Forming Cyanobacteria. *Angew. Chem. Int. Ed. Engl.* **2008**, *47* (40), 7756–7759. https://doi.org/10.1002/anie.200802730.

(126) Liu, T.; Mazmouz, R.; Ongley, S. E.; Chau, R.; Pickford, R.; Woodhouse, J. N.; Neilan, B. A. Directing the Heterologous Production of Specific Cyanobacterial Toxin Variants. *ACS Chem. Biol.* **2017**, *12* (8), 2021–2029. https://doi.org/10.1021/acschembio.7b00181.

(127) Park, S.-H.; Lee, K.; Jang, J. W.; Hahn, J.-S. Metabolic Engineering of Saccharomyces Cerevisiae for Production of Shinorine, a Sunscreen Material, from Xylose. *ACS Synth. Biol.* **2019**, *8* (2), 346–357. https://doi.org/10.1021/acssynbio.8b00388.

(128) Kim, E. J.; Lee, J. H.; Choi, H.; Pereira, A. R.; Ban, Y. H.; Yoo, Y. J.; Kim, E.; Park, J. W.; Sherman, D. H.; Gerwick, W. H.; Yoon, Y. J. Heterologous Production of 4-O-Demethylbarbamide, a Marine Cyanobacterial Natural Product. *Org. Lett.* **2012**, *14* (23), 5824–5827. https://doi.org/10.1021/ol302575h.

(129) Taton, A.; Ecker, A.; Diaz, B.; Moss, N. A.; Anderson, B.; Reher, R.; Leão, T. F.; Simkovsky, R.; Dorrestein, P. C.;

Gerwick, L.; Gerwick, W. H.; Golden, J. W. Heterologous Expression of Cryptomaldamide in a Cyanobacterial Host. *ACS Synth. Biol.* **2020**, *9* (12), 3364–3376. https://doi.org/10.1021/acssynbio.0c00431.

(130) Baeshen, M. N.; Al-Hejin, A. M.; Bora, R. S.; Ahmed, M. M. M.; Ramadan, H. A. I.; Saini, K. S.; Baeshen, N. A.; Redwan, E. M. Production of Biopharmaceuticals in E. Coli: Current Scenario and Future Perspectives. *J. Microbiol. Biotechnol.* **2015**, *25* (7), 953–962. https://doi.org/10.4014/jmb.1412.12079.

(131) Selas Castiñeiras, T.; Williams, S. G.; Hitchcock, A. G.; Smith, D. C. E. Coli Strain Engineering for the Production of Advanced Biopharmaceutical Products. *FEMS Microbiol. Lett.* **2018**, *365* (15), 162. https://doi.org/10.1093/femsle/fny162.

(132) Huang, C.-J.; Lin, H.; Yang, X. Industrial Production of Recombinant Therapeutics in Escherichia Coli and Its Recent Advancements. *J. Ind. Microbiol. Biotechnol.* **2012**, *39* (3), 383–399. https://doi.org/10.1007/s10295-011-1082-9.

(133) Trinh, C. T.; Unrean, P.; Srienc, F. Minimal Escherichia Coli Cell for the Most Efficient Production of Ethanol from Hexoses and Pentoses. *Appl. Environ. Microbiol.* **2008**, *74* (12), 3634–3643. https://doi.org/10.1128/AEM.02708-07.

(134) Atsumi, S.; Liao, J. C. Directed Evolution of Methanococcus Jannaschii Citramalate Synthase for Biosynthesis of 1-Propanol and 1-Butanol by Escherichia Coli. *Appl. Environ. Microbiol.* **2008**, *74* (24), 7802–7808. https://doi.org/10.1128/AEM.02046-08.

(135) Inokuma, K.; Liao, J. C.; Okamoto, M.; Hanai, T. Improvement of Isopropanol Production by Metabolically Engineered Escherichia Coli Using Gas Stripping. *J. Biosci. Bioeng.* **2010**, *110* (6), 696–701. https://doi.org/10.1016/j.jbiosc.2010.07.010.

(136) Seol, E.; Manimaran, A.; Jang, Y.; Kim, S.; Oh, Y.-K.; Park, S. Sustained Hydrogen Production from Formate Using Immobilized Recombinant Escherichia Coli SH5. *Int. J. Hydrogen Energy* **2011**, *36* (14), 8681–8686. https://doi.org/10.1016/j.ijhydene.2010.05.118.

(137) Akinterinwa, O.; Cirino, P. C. Anaerobic Obligatory Xylitol Production in Escherichia Coli Strains Devoid of Native Fermentation Pathways. *Appl. Environ. Microbiol.* **2011**, *77* (2), 706–709. https://doi.org/10.1128/AEM.01890-10.

(138) Kaup, B.; Bringer-Meyer, S.; Sahm, H. Metabolic Engineering of Escherichia Coli: Construction of an Efficient Biocatalyst for D-Mannitol Formation in a Whole-Cell Biotransformation. *Appl. Microbiol. Biotechnol.* **2004**, *64* (3), 333–339. https://doi.org/10.1007/s00253-003-1470-9.

(139) Lee, K. H.; Park, J. H.; Kim, T. Y.; Kim, H. U.; Lee, S. Y. Systems Metabolic Engineering of Escherichia Coli for L-Threonine Production. *Mol. Syst. Biol.* **2007**, *3*, 149. https://doi.org/10.1038/msb4100196.

(140) Báez-Viveros, J. L.; Flores, N.; Juárez, K.; Castillo-España, P.; Bolivar, F.; Gosset, G. Metabolic Transcription Analysis of Engineered Escherichia Coli Strains That Overproduce L-Phenylalanine. *Microb. Cell Fact.* **2007**, *6* (1), 1–20. https://doi.org/10.1186/1475-2859-6-30/FIGURES/6.

(141) Zhou, L.; Zuo, Z.-R.; Chen, X.-Z.; Niu, D.-D.; Tian, K.-M.; Prior, B. A.; Shen, W.; Shi, G.-Y.; Singh, S.; Wang, Z.-X. Evaluation of Genetic Manipulation Strategies on D-Lactate Production by Escherichia Coli. *Curr. Microbiol.* **2011**, *62* (3), 981–989. https://doi.org/10.1007/s00284-010-9817-9.

(142) Zhu, Y.; Eiteman, M. A.; Altman, R.; Altman, E. High Glycolytic Flux Improves Pyruvate Production by a Metabolically Engineered Escherichia Coli Strain. *Appl. Environ. Microbiol.* **2008**, *74* (21), 6649–6655.

https://doi.org/10.1128/AEM.01610-08.

(143) Emptage, M.; Haynie, S. L.; Laffend, L. A.; Pucci, J. P.; Whited, G. Process for the Biological Production of 1,3-Propanediol with High Titer. US: Patent 6514733, 2003.

(144) Jung, Y. K.; Lee, S. Y. Efficient Production of Polylactic Acid and Its Copolymers by Metabolically Engineered Escherichia Coli. *J. Biotechnol.* **2011**, *151* (1), 94–101. https://doi.org/10.1016/j.jbiotec.2010.11.009.

(145) Chen, X.; Zhou, L.; Tian, K.; Kumar, A.; Singh, S.; Prior, B. A.; Wang, Z. Metabolic Engineering of Escherichia Coli: A Sustainable Industrial Platform for Bio-Based Chemical Production. *Biotechnol. Adv.* **2013**, *31* (8), 1200–1223. https://doi.org/10.1016/j.biotechadv.2013.02.009.

(146) Challis, G. L. Engineering Escherichia Coli to Produce Nonribosomal Peptide Antibiotics. *Nat. Chem. Biol.* **2006**, *2* (8), 398–400. https://doi.org/10.1038/nchembio0806-398.

(147) Watanabe, K.; Oikawa, H. Robust Platform for de Novo Production of Heterologous Polyketides and Nonribosomal Peptides in Escherichia Coli. *Org. Biomol. Chem.* **2007**, *5* (4), 593–602. https://doi.org/10.1039/b615589h.

(148) Lee, N. C. O.; Larionov, V.; Kouprina, N. Highly Efficient CRISPR/Cas9-Mediated TAR Cloning of Genes and Chromosomal Loci from Complex Genomes in Yeast. *Nucleic Acids Res.* **2015**, *43* (8), e55–e55. https://doi.org/10.1093/nar/gkv112.

(149) Kouprina, N.; Larionov, V. TAR Cloning: Insights into Gene Function, Long-Range Haplotypes and Genome Structure and Evolution. *Nat. Rev. Genet.* **2006**, *7* (10), 805–812. https://doi.org/10.1038/nrg1943.

(150) Larionov, V.; Kouprina, N.; Graves, J.; Chen, X. N.; Korenberg, J. R.; Resnick, M. A. Specific Cloning of Human DNA as Yeast Artificial Chromosomes by Transformation-Associated Recombination. *Proc. Natl. Acad. Sci.* **1996**, *93* (1), 491–496. https://doi.org/10.1073/pnas.93.1.491.

(151) Zhang, B.; Tian, W.; Wang, S.; Yan, X.; Jia, X.; Pierens, G. K.; Chen, W.; Ma, H.; Deng, Z.; Qu, X. Activation of Natural Products Biosynthetic Pathways via a Protein Modification Level Regulation. *ACS Chem. Biol.* **2017**, *12* (7), 1732–1736. https://doi.org/10.1021/acschembio.7b00225.

(152) Zhang, J. J.; Yamanaka, K.; Tang, X.; Moore, B. S. Direct Cloning and Heterologous Expression of Natural Product Biosynthetic Gene Clusters by Transformation-Associated Recombination. In *Methods in enzymology*; Academic Press Inc., 2019; Vol. 621, pp 87–110. https://doi.org/10.1016/bs.mie.2019.02.026.

(153) Bai Flagfeldt, D.; Siewers, V.; Huang, L.; Nielsen, J. Characterization of Chromosomal Integration Sites for Heterologous Gene Expression in Saccharomyces Cerevisiae. *Yeast* **2009**, *26* (10), 545–551. https://doi.org/10.1002/yea.1705.

(154) Maury, J.; Germann, S. M.; Baallal Jacobsen, S. A.; Jensen, N. B.; Kildegaard, K. R.; Herrgård, M. J.; Schneider, K.; Koza, A.; Forster, J.; Nielsen, J.; Borodina, I. EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in Saccharomyces Cerevisiae. *PLoS One* **2016**, *11* (3), e0150394. https://doi.org/10.1371/journal.pone.0150394.

(155) Horwitz, A. A.; Walter, J. M.; Schubert, M. G.; Kung, S. H.; Hawkins, K.; Platt, D. M.; Hernday, A. D.; Mahatdejkul-Meadows, T.; Szeto, W.; Chandran, S. S.; Newman, J. D. Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell Syst.* **2015**, *1* (1), 88–96. https://doi.org/10.1016/j.cels.2015.02.001.

(156) Reider Apel, A.; D'Espaux, L.; Wehrs, M.; Sachs, D.; Li, R. A.; Tong, G. J.; Garber, M.; Nnadi, O.; Zhuang, W.;

Hillson, N. J.; Keasling, J. D.; Mukhopadhyay, A. A Cas9-Based Toolkit to Program Gene Expression in Saccharomyces Cerevisiae. *Nucleic Acids Res.* **2017**, *45* (1), 496–508. https://doi.org/10.1093/nar/gkw1023.

(157) DiCarlo, J. E.; Norville, J. E.; Mali, P.; Rios, X.; Aach, J.; Church, G. M. Genome Engineering in Saccharomyces Cerevisiae Using CRISPR-Cas Systems. *Nucleic Acids Res.* **2013**, *41* (7), 4336–4343. https://doi.org/10.1093/nar/gkt135.

(158) Lee, M. E.; DeLoache, W. C.; Cervantes, B.; Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **2015**, *4* (9), 975–986. https://doi.org/10.1021/sb500366v.

(159) Yuan, J.; Ching, C. B. Combinatorial Assembly of Large Biochemical Pathways into Yeast Chromosomes for Improved Production of Value-Added Compounds. *ACS Synth. Biol.* **2015**, *4* (1), 23–31. https://doi.org/10.1021/sb500079f.

(160) Tsunematsu, Y.; Ishiuchi, K.; Hotta, K.; Watanabe, K. Yeast-Based Genome Mining, Production and Mechanistic Studies of the Biosynthesis of Fungal Polyketide and Peptide Natural Products. *Nat. Prod. Rep.* **2013**, *30* (8), 1139–1149. https://doi.org/10.1039/c3np70037b.

(161) Shao, Z.; Zhao, H.; Zhao, H. DNA Assembler, an in Vivo Genetic Method for Rapid Construction of Biochemical Pathways. *Nucleic Acids Res.* **2009**, *37* (2), e16. https://doi.org/10.1093/nar/gkn991.

(162) Shao, Z.; Luo, Y.; Zhao, H. Rapid Characterization and Engineering of Natural Product Biosynthetic Pathways via DNA Assembler. *Mol. Biosyst.* **2011**, *7* (4), 1056. https://doi.org/10.1039/c0mb00338g.

(163) Jansen, G.; Wu, C.; Schade, B.; Thomas, D. Y.; Whiteway, M. Drag&Drop Cloning in Yeast. *Gene* **2005**, *344*, 43–51. https://doi.org/10.1016/j.gene.2004.10.016.

(164) Naesby, M.; Nielsen, S. V.; Nielsen, C. A. F.; Green, T.; Tange, T. O.; Simón, E.; Knechtle, P.; Hansson, A.; Schwab, M. S.; Titiz, O.; Folly, C.; Archila, R. E.; Maver, M.; van Sint Fiet, S.; Boussemghoune, T.; Janes, M.; Kumar, A. S. S.; Sonkar, S. P.; Mitra, P. P.; Benjamin, V. A. K.; Korrapati, N.; Suman, I.; Hansen, E. H.; Thybo, T.; Goldsmith, N.; Sorensen, A. S. Yeast Artificial Chromosomes Employed for Random Assembly of Biosynthetic Pathways and Production of Diverse Compounds in Saccharomyces Cerevisiae. *Microb. Cell Fact.* **2009**, *8* (1), 45. https://doi.org/10.1186/1475-2859-8-45.

(165) Blount, B. A.; Gowers, G.-O. F.; Ho, J. C. H.; Ledesma-Amaro, R.; Jovicevic, D.; McKiernan, R. M.; Xie, Z. X.; Li, B. Z.; Yuan, Y. J.; Ellis, T. Rapid Host Strain Improvement by in Vivo Rearrangement of a Synthetic Yeast Chromosome. *Nat. Commun.* **2018**, *9* (1), 1932. https://doi.org/10.1038/s41467-018-03143-w.

(166) Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. Direct Cloning and Refactoring of a Silent Lipopeptide Biosynthetic Gene Cluster Yields the Antibiotic Taromycin A. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (5), 1957–1962. https://doi.org/10.1073/pnas.1319584111.

(167) Li, A.; Liu, Z.; Li, Q.; Yu, L.; Wang, D.; Deng, X. Construction and Characterization of Bidirectional Expression Vectors in Saccharomyces Cerevisiae. *FEMS Yeast Res.* **2008**, *8* (1), 6–9. https://doi.org/10.1111/j.1567-1364.2007.00335.x.

(168) Öztürk, S.; Ergün, B. G.; Çalık, P. Double Promoter Expression Systems for Recombinant Protein Production by Industrial Microorganisms. *Appl. Microbiol. Biotechnol.* **2017**, *101* (20), 7459–7475. https://doi.org/10.1007/s00253-017-8487-y.

(169) Curran, K. A.; Morse, N. J.; Markham, K. A.; Wagman, A. M.; Gupta, A.; Alper, H. S. Short Synthetic Terminators

for Improved Heterologous Gene Expression in Yeast. *ACS Synth. Biol.* **2015**, *4* (7), 824–832. https://doi.org/10.1021/sb5003357.

(170) Redden, H.; Alper, H. S. The Development and Characterization of Synthetic Minimal Yeast Promoters. *Nat. Commun.* **2015**, *6* (1), 7810. https://doi.org/10.1038/ncomms8810.

(171) Blazeck, J.; Garg, R.; Reed, B.; Alper, H. S. Controlling Promoter Strength and Regulation in Saccharomyces Cerevisiae Using Synthetic Hybrid Promoters. *Biotechnol. Bioeng.* **2012**, *109* (11), 2884–2895. https://doi.org/10.1002/bit.24552.

(172) Singh, A.; Lugovoy, J. M.; Kohr, W. J.; Perry, L. J. Synthesis, Secretion and Processing of Alpha-Factor-Interferon Fusion Proteins in Yeast. *Nucleic Acids Res.* **1984**, *12* (23), 8927–8938. https://doi.org/10.1093/nar/12.23.8927.

(173) Kjeldsen, T. Yeast Secretory Expression of Insulin Precursors. *Appl. Microbiol. Biotechnol.* **2000**, *54* (3), 277–286. https://doi.org/10.1007/s002530000402.

(174) Jan Markussen, Ulrik Damgaard, Ivan Diers, Niels Pill, Mogens Trier Hansen, Per Larsen, Fanny Norris, Kjeld Norris, Ole Schou, Leo Snel, Lars Thim, H. O. V. *Porto Carras, Chalkidiki, Greece, Aug. 31–Sept. 5, 1986*; Theodoropoulos, D., Ed.; De Gruyter: Berlin, 1987. https://doi.org/10.1515/9783110864243.

(175) Brake, A. J.; Merryweather, J. P.; Coit, D. G.; Heberlein, U. A.; Masiarz, F. R.; Mullenbach, G. T.; Urdea, M. S.; Valenzuela, P.; Barr, P. J. Alpha-Factor-Directed Synthesis and Secretion of Mature Foreign Proteins in Saccharomyces Cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (15), 4642–4646. https://doi.org/10.1073/pnas.81.15.4642.

(176) Valenzuela, P.; Medina, A.; Rutter, W. J.; Ammerer, G.; Hall, B. D. Synthesis and Assembly of Hepatitis B Virus Surface Antigen Particles in Yeast. *Nature* **1982**, *298* (5872), 347–350. https://doi.org/10.1038/298347a0.

(177) Kim, I.-K.; Roldão, A.; Siewers, V.; Nielsen, J. A Systems-Level Approach for Metabolic Engineering of Yeast Cell Factories. *FEMS Yeast Res.* **2012**, *12* (2), 228–248. https://doi.org/10.1111/j.1567-1364.2011.00779.x.

(178) Kowald, A.; Wierling, C. Standards, Tools, and Databases for the Analysis of Yeast 'Omics Data. In *Methods in molecular biology (Clifton, N.J.)*; Humana Press, 2011; Vol. 759, pp 345–365. https://doi.org/10.1007/978-1-61779-173-4_20.

(179) Bond, C.; Tang, Y.; Li, L. Saccharomyces Cerevisiae as a Tool for Mining, Studying and Engineering Fungal Polyketide Synthases. *Fungal Genet. Biol.* **2016**, *89*, 52–61. https://doi.org/10.1016/j.fgb.2016.01.005.

(180) Awan, A. R.; Blount, B. A.; Bell, D. J.; Shaw, W. M.; Ho, J. C. H.; McKiernan, R. M.; Ellis, T. Biosynthesis of the Antibiotic Nonribosomal Peptide Penicillin in Baker's Yeast. *Nat. Commun.* **2017**, *8* (1), 15202. https://doi.org/10.1038/ncomms15202.

(181) Awan, A. R.; Shaw, W. M.; Ellis, T. Biosynthesis of Therapeutic Natural Products Using Synthetic Biology. *Adv. Drug Deliv. Rev.* **2016**, *105* (Pt A), 96–106. https://doi.org/10.1016/j.addr.2016.04.010.

(182) Siddiqui, M. S.; Thodey, K.; Trenchard, I.; Smolke, C. D. Advancing Secondary Metabolite Biosynthesis in Yeast with Synthetic Biology Tools. *FEMS Yeast Res.* **2012**, *12* (2), 144–170. https://doi.org/10.1111/j.1567-1364.2011.00774.x.

(183) Walker, R. S. K.; Pretorius, I. S. Applications of Yeast Synthetic Biology Geared towards the Production of Biopharmaceuticals. *Genes (Basel).* **2018**, *9* (7), 340. https://doi.org/10.3390/genes9070340.

(184) Nevoigt, E. Progress in Metabolic Engineering of Saccharomyces Cerevisiae. *Microbiol. Mol. Biol. Rev.* **2008**, *72* (3), 379–412. https://doi.org/10.1128/MMBR.00025-07.

(185) Sauer, M.; Porro, D.; Mattanovich, D.; Branduardi, P. 16 Years Research on Lactic Acid Production with Yeast - Ready for the Market? *Biotechnol. Genet. Eng. Rev.* **2010**, *27* (1), 229–256. https://doi.org/10.1080/02648725.2010.10648152.

(186) Miller, C.; Fosmer, A.; Rush, B.; McMullin, T.; Beacom, D.; Suominen, P. Industrial Production of Lactic Acid. In *Comprehensive Biotechnology*; Elsevier, 2011; Vol. 3, pp 179–188. https://doi.org/10.1016/B978-0-08-088504-9.00177-X.

(187) Rush, B. Turning a Novel Yeast into a Platform Host for Industrial Production of Fuels and Chemicals. In *Metabolic Engineering IX*; E. Heinzle, Saarland Univ.; P. Soucaille, INSA; G. Whited, Danisco Eds; ECI Symposium Series, 2013.

(188) Zhang, Y.; Nielsen, J.; Liu, Z. Engineering Yeast Metabolism for Production of Terpenoids for Use as Perfume Ingredients, Pharmaceuticals and Biofuels. *FEMS Yeast Res.* **2017**, *17* (8). https://doi.org/10.1093/femsyr/fox080.

(189) Chen, R.; Yang, S.; Zhang, L.; Zhou, Y. J. Advanced Strategies for Production of Natural Products in Yeast. *iScience* **2020**, *23* (3), 100879. https://doi.org/10.1016/j.isci.2020.100879.

(190) Ignea, C.; Raadam, M. H.; Motawia, M. S.; Makris, A. M.; Vickers, C. E.; Kampranis, S. C. Orthogonal Monoterpenoid Biosynthesis in Yeast Constructed on an Isomeric Substrate. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-11290-x.

(191) Szczebara, F. M.; Chandelier, C.; Villeret, C.; Masurel, A.; Bourot, S.; Duport, C.; Blanchard, S.; Groisillier, A.; Testet, E.; Costaglioli, P.; Cauet, G.; Degryse, E.; Balbuena, D.; Winter, J.; Achstetter, T.; Spagnoli, R.; Pompon, D.; Dumas, B. Total Biosynthesis of Hydrocortisone from a Simple Carbon Source in Yeast. *Nat. Biotechnol.* **2003**, *21* (2), 143–149. https://doi.org/10.1038/nbt775.

(192) Chen, J.; Fan, F.; Qu, G.; Tang, J.; Xi, Y.; Bi, C.; Sun, Z.; Zhang, X. Identification of Absidia Orchidis Steroid 11β-Hydroxylation System and Its Application in Engineering Saccharomyces Cerevisiae for One-Step Biotransformation to Produce Hydrocortisone. *Metab. Eng.* **2020**, *57*, 31–42. https://doi.org/10.1016/j.ymben.2019.10.006.

(193) Ro, D.-K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, M. C. Y.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D. Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast. *Nature* **2006**, *440* (7086), 940–943. https://doi.org/10.1038/nature04640.

(194) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, A.; Main, A.; Eng, D.; Polichuk, D. R.; Teoh, K. H.; Reed, D. W.; Treynor, T.; Lenihan, J.; Fleck, M.; Bajad, S.; Dang, G.; Dengrove, D.; Diola, D.; Dorin, G.; Ellens, K. W.; Fickes, S.; Galazzo, J.; Gaucher, S. P.; Geistlinger, T.; Henry, R.; Hepp, M.; Horning, T.; Iqbal, T.; Jiang, H.; Kizer, L.; Lieu, B.; Melis, D.; Moss, N.; Regentin, R.; Secrest, S.; Tsuruta, H.; Vazquez, R.; Westblade, L. F.; Xu, L.; Yu, M.; Zhang, Y.; Zhao, L.; Lievense, J.; Covello, P. S.; Keasling, J. D.; Reiling, K. K.; Renninger, N. S.; Newman, J. D. High-Level Semi-Synthetic Production of the Potent Antimalarial Artemisinin. *Nature* **2013**, *496* (7446), 528–532. https://doi.org/10.1038/nature12051.

(195) Peplow, M. Synthetic Biology's First Malaria Drug Meets Market Resistance. *Nature* **2016**, *530* (7591), 389–390. https://doi.org/10.1038/530390a.

(196) Galanie, S.; Thodey, K.; Trenchard, I. J.; Interrante, M. F.; Smolke, C. D. Complete Biosynthesis of Opioids in Yeast. *Science (80-.).* **2015**, *349* (6252), 1095–1100. https://doi.org/10.1126/science.aac9373.

(197) Li, Y.; Li, S.; Thodey, K.; Trenchard, I.; Cravens, A.; Smolke, C. D. Complete Biosynthesis of Noscapine and Halogenated Alkaloids in Yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (17), E3922–E3931. https://doi.org/10.1073/pnas.1721469115.

(198) Li, M.; Kildegaard, K. R.; Chen, Y.; Rodriguez, A.; Borodina, I.; Nielsen, J. De Novo Production of Resveratrol from Glucose or Ethanol by Engineered Saccharomyces Cerevisiae. *Metab. Eng.* **2015**, *32*, 1–11. https://doi.org/10.1016/j.ymben.2015.08.007.

(199) Levisson, M.; Araya-Cloutier, C.; De Bruijn, W. J. C.; Van Der Heide, M.; Salvador López, J. M.; Daran, J. M.; Vincken, J. P.; Beekwilder, J. Toward Developing a Yeast Cell Factory for the Production of Prenylated Flavonoids. *J. Agric. Food Chem.* **2019**, *67* (49), 13478–13486. https://doi.org/10.1021/acs.jafc.9b01367.

(200) Zirpel, B.; Stehle, F.; Kayser, O. Production of Δ9-Tetrahydrocannabinolic Acid from Cannabigerolic Acid by Whole Cells of Pichia (Komagataella) Pastoris Expressing Δ9-Tetrahydrocannabinolic Acid Synthase from Cannabis Sativa L. *Biotechnol. Lett.* **2015**, *37* (9), 1869–1875. https://doi.org/10.1007/s10529-015-1853-x.

(201) Thomas, F.; Schmidt, C.; Kayser, O. Bioengineering Studies and Pathway Modeling of the Heterologous Biosynthesis of Tetrahydrocannabinolic Acid in Yeast. *Appl. Microbiol. Biotechnol.* **2020**, *104* (22), 9551–9563. https://doi.org/10.1007/s00253-020-10798-3.

(202) Carvalho, Â.; Hansen, E. H.; Kayser, O.; Carlsen, S.; Stehle, F. Designing Microorganisms for Heterologous Biosynthesis of Cannabinoids. *FEMS Yeast Res.* **2017**, *17* (4), 37. https://doi.org/10.1093/femsyr/fox037.

(203) Hitschler, J.; Boles, E. De Novo Production of Aromatic M-Cresol in Saccharomyces Cerevisiae Mediated by Heterologous Polyketide Synthases Combined with a 6-Methylsalicylic Acid Decarboxylase. *Metab. Eng. Commun.* **2019**, *9*, e00093. https://doi.org/10.1016/j.mec.2019.e00093.

(204) Kealey, J. T.; Liu, L.; Santi, D. V; Betlach, M. C.; Barr, P. J. Production of a Polyketide Natural Product in Nonpolyketide-Producing Prokaryotic and Eukaryotic Hosts. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (2), 505–509. https://doi.org/10.1073/pnas.95.2.505.

(205) Wattanachaisaereekul, S.; Lantz, A. E.; Nielsen, M. L.; Nielsen, J. Production of the Polyketide 6-MSA in Yeast Engineered for Increased Malonyl-CoA Supply. *Metab. Eng.* **2008**, *10* (5), 246–254. https://doi.org/10.1016/j.ymben.2008.04.005.

(206) Wattanachaisaereekul, S.; Lantz, A. E.; Nielsen, M. L.; Andrésson, Ó. S.; Nielsen, J. Optimization of Heterologous Production of the Polyketide 6-MSA in Saccharomyces Cerevisiae. *Biotechnol. Bioeng.* **2007**, *97* (4), 893–900. https://doi.org/10.1002/bit.21286.

(207) Choi, J. W.; Da Silva, N. A. Improving Polyketide and Fatty Acid Synthesis by Engineering of the Yeast Acetyl-CoA Carboxylase. *J. Biotechnol.* **2014**, *187*, 56–59. https://doi.org/10.1016/j.jbiotec.2014.07.430.

(208) Yu, D.; Xu, F.; Zi, J.; Wang, S.; Gage, D.; Zeng, J.; Zhan, J. Engineered Production of Fungal Anticancer Cyclooligomer Depsipeptides in Saccharomyces Cerevisiae. *Metab. Eng.* **2013**, *18*, 60–68. https://doi.org/10.1016/j.ymben.2013.04.001.

(209) Yu, D.; Xu, F.; Zhang, S.; Zhan, J. Decoding and Reprogramming Fungal Iterative Nonribosomal Peptide Synthetases. *Nat. Commun.* **2017**, *8*, 15349. https://doi.org/10.1038/ncomms15349.

(210) Wehrs, M.; Prahl, J.-P.; Moon, J.; Li, Y.; Tanjore, D.; Keasling, J. D.; Pray, T.; Mukhopadhyay, A. Correction to:

Production Efficiency of the Bacterial Non-Ribosomal Peptide Indigoidine Relies on the Respiratory Metabolic State in *S. cerevisiae*. *Microb. Cell Fact.* **2019**, *18* (1), 218. https://doi.org/10.1186/s12934-019-1262-2.

(211) Balskus, E. P.; Walsh, C. T. The Genetic and Molecular Basis for Sunscreen Biosynthesis in Cyanobacteria. *Science* **2010**, *329* (5999), 1653–1656. https://doi.org/10.1126/science.1193637.

(212) Mutka, S. C.; Bondi, S. M.; Carney, J. R.; Da Silva, N. A.; Kealey, J. T. Metabolic Pathway Engineering for Complex Polyketide Biosynthesis in Saccharomyces Cerevisiae. *FEMS Yeast Res.* **2006**, *6* (1), 40–47. https://doi.org/10.1111/j.1567-1356.2005.00001.x.

(213) Siewers, V.; San-Bento, R.; Nielsen, J. Implementation of Communication-Mediating Domains for Non-Ribosomal Peptide Production in Saccharomyces Cerevisiae. *Biotechnol. Bioeng.* **2010**, *106* (5), 841–844. https://doi.org/10.1002/bit.22739.

(214) Schwartz, R. E.; Hirsch, C. F.; Sesin, D. F.; Flor, J. E.; Chartrain, M.; Fromtling, R. E.; Harris, G. H.; Salvatore, M. J.; Liesch, J. M.; Yudin, K. Pharmaceuticals from Cultured Algae. *J. Ind. Microbiol.* **1990**, *5* (2–3), 113–123. https://doi.org/10.1007/BF01573860.

(215) Walker, I. H.; Hsieh, P.; Riggs, P. D. Mutations in Maltose-Binding Protein That Alter Affinity and Solubility Properties. *Appl. Microbiol. Biotechnol.* **2010**, *88* (1), 187–197. https://doi.org/10.1007/s00253-010-2696-y.

(216) Rolf, J.; Siedentop, R.; Lütz, S.; Rosenthal, K. Screening and Identification of Novel CGAS Homologues Using a Combination of in Vitro and In Vivo Protein Synthesis. *Int. J. Mol. Sci.* **2019**, *21* (1), 105. https://doi.org/10.3390/ijms21010105.

(217) Jin, M.; Beer, S.; Clardy, J. Pantoea agglomerans andrimid biosynthetic gene cluster, complete sequence https://www.ncbi.nlm.nih.gov/nuccore/AY192157 (accessed May 4, 2022).

(218) Promega. Single Step KRX Competent Cells. **2006**.

(219) Agilent Technologies. ArcticExpress Competent Cells and ArcticExpress (DE3) Competent Cells Instruction Manual. **2015**.

(220) Lucigen. OverExpress™ E coli C43(DE3) & C41(DE3) Comp Cells | LG https://www.lucigen.com/OverExpress-C41-DE3-and-C43-DE3-Competent-Cells/ (accessed May 9, 2019).

(221) Sigma-Aldrich. *Yeast Synthetic Drop-out Media Supplements - Datasheet*.

(222) Casal, M.; Paiva, S.; Queirós, O.; Soares-Silva, I. Transport of Carboxylic Acids in Yeasts. *FEMS Microbiol. Rev.* **2008**, *32* (6), 974–994. https://doi.org/10.1111/j.1574-6976.2008.00128.x.

(223) Tippelt, A.; Busche, T.; Rückert, C.; Nett, M. Complete Genome Sequence of the Cryptophycin-Producing Cyanobacterium Nostoc Sp. Strain ATCC 53789. *Microbiol. Resour. Announc.* **2020**, *9* (14), e00040-20. https://doi.org/10.1128/MRA.00040-20.

(224) Cohen, M. F.; Wallis, J. G.; Campbell, E. L.; Meeks, J. C. Transposon Mutagenesis of Nostoc Sp. Strain ATCC 29133, a Filamentous Cyanobacterium with Multiple Cellular Differentiation Alternatives. *Microbiology* **1994**, *140* (12), 3233–3240. https://doi.org/10.1099/13500872-140-12-3233.

(225) Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6* (5), 343–345. https://doi.org/10.1038/nmeth.1318.

(226) Insightful Science. SnapGene | Software for everyday molecular biology https://www.snapgene.com/ (accessed Aug 23, 2021).

(227) Madden, T. The BLAST Sequence Analysis Tool. **2003**.

(228) Kibbe, W. A. OligoCalc: An Online Oligonucleotide Properties Calculator. *Nucleic Acids Res.* **2007**, *35* (Web Server issue), W43-6. https://doi.org/10.1093/nar/gkm234.

(229) PerkinElmer. ChemDraw https://perkinelmerinformatics.com/products/research/chemdraw/ (accessed Aug 23, 2021).

(230) Draw Freely | Inkscape https://inkscape.org/de/ (accessed Aug 23, 2021).

(231) Koren, S.; Walenz, B. P.; Berlin, K.; Miller, J. R.; Bergman, N. H.; Phillippy, A. M. Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation. *Genome Res.* **2017**, *27* (5), 722–736. https://doi.org/10.1101/gr.215087.116.

(232) Vaser, R.; Sović, I.; Nagarajan, N.; Šikić, M. Fast and Accurate de Novo Genome Assembly from Long Uncorrected Reads. *Genome Res.* **2017**, *27* (5), 737–746. https://doi.org/10.1101/gr.214270.116.

(233) GitHub - nanoporetech/medaka: Sequence correction provided by ONT Research https://github.com/nanoporetech/medaka (accessed Nov 5, 2019).

(234) Walker, B. J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C. A.; Zeng, Q.; Wortman, J.; Young, S. K.; Earl, A. M. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS One* **2014**, *9* (11), e112963. https://doi.org/10.1371/journal.pone.0112963.

(235) Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *Preprint* **2013**, *arxiv:1303*, 1–3.

(236) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9* (4), 357–359. https://doi.org/10.1038/nmeth.1923.

(237) Wick, R. R.; Judd, L. M.; Gorrie, C. L.; Holt, K. E. Unicycler: Resolving Bacterial Genome Assemblies from Short and Long Sequencing Reads. *PLOS Comput. Biol.* **2017**, *13* (6), e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

(238) Gordon, D.; Abajian, C.; Green, P. Consed: A Graphical Tool for Sequence Finishing. *Genome Res.* **1998**, *8* (3), 195–202. https://doi.org/10.1101/gr.8.3.195.

(239) Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P. Integrative Genomics Viewer. *Nat. Biotechnol.* **2011**, *29* (1), 24–26. https://doi.org/10.1038/nbt.1754.

(240) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(241) Carver, T.; Thomson, N.; Bleasby, A.; Berriman, M.; Parkhill, J. DNAPlotter: Circular and Linear Interactive Genome Visualization. *Bioinforma. Appl. NOTE* **2009**, *25* (1), 119–120. https://doi.org/10.1093/bioinformatics/btn578.

(242) 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* (5), 1500–1517. https://doi.org/10.1016/j.jmb.2007.02.099.

(243) Yadav, G.; Gokhale, R. S.; Mohanty, D. Computational Approach for Prediction of Domain Organization and Substrate Specificity of Modular Polyketide Synthases. *J. Mol. Biol.* **2003**, *328* (2), 335–363. https://doi.org/10.1016/s0022-2836(03)00232-8.

(244) NEB. Transformation Protocol E. coli| NEB https://international.neb.com/protocols/2012/05/21/transformation-protocol (accessed Aug 20, 2021).

(245) Cox, B. M.; Bilsborrow, J. B.; Walker, K. D. Enhanced Conversion of Racemic α-Arylalanines to (R)-β-Arylalanines by Coupled Racemase/Aminomutase Catalysis. *J. Org. Chem.* **2009**, *74* (18), 6953–6959. https://doi.org/10.1021/jo9009563.

(246) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus Subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases †. *Biochemistry* **1998**, *37* (6), 1585–1595. https://doi.org/10.1021/bi9719861.

(247) Nakano, M. M.; Corbell, N.; Besson, J.; Zuber, P. Isolation and Characterization of *Sfp*: A Gene That Functions in the Production of the Lipopeptide Biosurfactant, Surfactin, in *Bacillus Subtilis*. *Mol. Gen. Genet. MGG* **1992**, *232* (2), 313–321. https://doi.org/10.1007/bf00280011.

(248) Cortes, J.; Wiesmann, K.; Roberts, G.; Brown, M.; Staunton, J.; Leadlay, P. Repositioning of a Domain in a Modular Polyketide Synthase to Promote Specific Chain Cleavage. *Science (80-.).* **1995**, *268* (5216), 1487–1489. https://doi.org/10.1126/science.7770773.

(249) Tosin, M.; Betancor, L.; Stephens, E.; Li, W. M. A.; Spencer, J. B.; Leadlay, P. F. Synthetic Chain Terminators Off-Load Intermediates from a Type I Polyketide Synthase. *Chembiochem* **2010**, *11* (4), 539–546. https://doi.org/10.1002/cbic.200900772.

(250) Hagen, A.; Poust, S.; De Rond, T.; Fortman, J. L.; Katz, L.; Petzold, C. J.; Keasling, J. D. Engineering a Polyketide Synthase for in Vitro Production of Adipic Acid. *ACS Synth. Biol.* **2016**, *5* (1), 21–27. https://doi.org/10.1021/acssynbio.5b00153.

(251) Ranganathan, A.; Timoney, M.; Bycroft, M.; Cortés, J.; Thomas, I. P.; Wilkinson, B.; Kellenberger, L.; Hanefeld, U.; Galloway, I. S.; Staunton, J.; Leadlay, P. F. Knowledge-Based Design of Bimodular and Trimodular Polyketide Synthases Based on Domain and Module Swaps: A Route to Simple Statin Analogues. *Chem. Biol.* **1999**, *6* (10), 731–741. https://doi.org/10.1016/S1074-5521(00)80020-4.

(252) Gietz, R. D.; Schiestl, R. H. High-Efficiency Yeast Transformation Using the LiAc/SS Carrier DNA/PEG Method. **2007**. https://doi.org/10.1038/nprot.2007.13.

(253) Flagfeldt, D. B.; Siewers, V.; Huang, L.; Nielsen, J. Characterization of Chromosomal Integration Sites for Heterologous Gene Expression in Saccharomyces Cerevisiae. *Yeast* **2009**, *26* (10), 545–551. https://doi.org/10.1002/yea.1705.

(254) Maury, J.; Germann, S. M.; Baallal Jacobsen, S. A.; Jensen, N. B.; Kildegaard, K. R.; Herrgård, M. J.; Schneider, K.; Koza, A.; Forster, J.; Nielsen, J.; Borodina, I. EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in Saccharomyces Cerevisiae. *PLoS One* **2016**, *11* (3), e0150394. https://doi.org/10.1371/journal.pone.0150394.

(255) Kingsbury, J. M.; Sen, N. D.; Cardenas, M. E. Branched-Chain Aminotransferases Control TORC1 Signaling in Saccharomyces Cerevisiae. *PLOS Genet.* **2015**, *11* (12), e1005714. https://doi.org/10.1371/journal.pgen.1005714.

(256) Conzelmann, A.; Riezman, H.; Desponds, C.; Bron, C. A Major 125-Kd Membrane Glycoprotein of Saccharomyces Cerevisiae Is Attached to the Lipid Bilayer through an Inositol-Containing Phospholipid. *EMBO J.* **1988**, *7* (7), 2233–2240. https://doi.org/10.1002/j.1460-2075.1988.tb03063.x.

(257) ACS Division of Organic Chemistry. Hans Reich's Collection. NMR Spectroscopy https://organicchemistrydata.org/hansreich/resources/nmr/?index=nmr_index%2F19F_shift#f-data00 (accessed Oct 25, 2021).

(258) Rastogi, R. P.; Sonani, R. R.; Madamwar, D. The High-Energy Radiation Protectant Extracellular Sheath Pigment Scytonemin and Its Reduced Counterpart in the Cyanobacterium Scytonema Sp. R77DM. *Bioresour. Technol.* **2014**, *171*, 396–400. https://doi.org/10.1016/j.biortech.2014.08.106.

(259) Helms, G. L.; Moore, R. E.; Niemczura, W. P.; Patterson, G. M. L.; Tomer, K. B.; Gross, M. L. Scytonemin A, a Novel Calcium Antagonist from a Blue-Green Alga. *J. Org. Chem.* **1988**, *53* (6), 1298–1307. https://doi.org/10.1021/jo00241a033.

(260) Tippelt, A.; Nett, M. *SCIENTIFIC HIGHLIGHTS 2020 - Charting the Layout of a Bacterial Factory for Anticancer Drugs - Genome Sequencing of Nostoc Sp. ATCC 53789*; Dortmund, 2020.

(261) NCBI. Nostoc Genomes - NCBI https://www.ncbi.nlm.nih.gov/genome/browse#!/prokaryotes/13531/Nostoc (accessed Jan 28, 2021).

(262) NCBI. Nostoc sp. ATCC 53789 genome neighbors https://www.ncbi.nlm.nih.gov/genome/neighbors/13531?genome_assembly_id=389916 (accessed Dec 15, 2021).

(263) NCBI. Nostoc sp. C057 (ID 13531) - Genome https://www.ncbi.nlm.nih.gov/genome/13531?genome_assembly_id=917144 (accessed Dec 15, 2021).

(264) Popin, R. V.; Alvarenga, D. O.; Castelo-Branco, R.; Fewer, D. P.; Sivonen, K. Mining of Cyanobacterial Genomes Indicates Natural Product Biosynthetic Gene Clusters Located in Conjugative Plasmids. *Front. Microbiol.* **2021**, *12*, 3353. https://doi.org/10.3389/FMICB.2021.684565/BIBTEX.

(265) Nandagopal, P.; Steven, A. N.; Chan, L.-W.; Rahmat, Z.; Jamaluddin, H.; Mohd Noh, N. I. Bioactive Metabolites Produced by Cyanobacteria for Growth Adaptation and Their Pharmacological Properties. *Biology (Basel).* **2021**, *10* (10), 1061. https://doi.org/10.3390/biology10101061.

(266) Rouhiainen, L.; Jokela, J.; Fewer, D. P.; Urmann, M.; Sivonen, K. Two Alternative Starter Modules for the Non-Ribosomal Biosynthesis of Specific Anabaenopeptin Variants in Anabaena (Cyanobacteria). *Chem. Biol.* **2010**, *17* (3), 265–273. https://doi.org/10.1016/J.CHEMBIOL.2010.01.017.

(267) Bauersachs, T.; Compaoré, J.; Hopmans, E. C.; Stal, L. J.; Schouten, S.; Sinninghe Damsté, J. S. Distribution of Heterocyst Glycolipids in Cyanobacteria. *Phytochemistry* **2009**, *70* (17–18), 2034–2039. https://doi.org/10.1016/J.PHYTOCHEM.2009.08.014.

(268) Shishido, T. K.; Jokela, J.; Fewer, D. P.; Wahlsten, M.; Fiore, M. F.; Sivonen, K. Simultaneous Production of Anabaenopeptins and Namalides by the Cyanobacterium *Nostoc* Sp. CENA543. *ACS Chem. Biol.* **2017**, *12* (11), 2746–2755. https://doi.org/10.1021/acschembio.7b00570.

(269) Becker, J. E.; Moore, R. E.; Moore, B. S. Cloning, Sequencing, and Biochemical Characterization of the Nostocyclopeptide Biosynthetic Gene Cluster: Molecular Basis for Imine Macrocyclization. *Gene* **2004**, *325*, 35–42. https://doi.org/10.1016/j.gene.2003.09.034.

(270) Cao, L.; Do, T.; Link, A. J. Mechanisms of Action of Ribosomally Synthesized and Posttranslationally Modified Peptides (RiPPs). *J. Ind. Microbiol. Biotechnol.* **2021**, *48* (3–4), 5. https://doi.org/10.1093/jimb/kuab005.

(271) 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.; 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.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; 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. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/c2np20085f.

(272) Schwarzer, D.; Finking, R.; Marahiel, M. A. Nonribosomal Peptides: From Genes to Products. *Nat. Prod. Rep.* **2003**, *20* (3), 275–287. https://doi.org/10.1039/B111145K.

(273) Firn, R. D.; Jones, C. G. Natural Products--a Simple Model to Explain Chemical Diversity. *Nat. Prod. Rep.* **2003**, *20* (4), 382–391. https://doi.org/10.1039/b208815k.

(274) Skipper, K. A.; Andersen, P. R.; Sharma, N.; Mikkelsen, J. G. DNA Transposon-Based Gene Vehicles - Scenes from an Evolutionary Drive. *J. Biomed. Sci.* **2013**, *20* (1). https://doi.org/10.1186/1423-0127-20-92.

(275) Bourque, G.; Burns, K. H.; Gehring, M.; Gorbunova, V.; Seluanov, A.; Hammell, M.; Imbeault, M.; Izsvák, Z.; Levin, H. L.; Macfarlan, T. S.; Mager, D. L.; Feschotte, C. Ten Things You Should Know about Transposable Elements. *Genome Biol.* **2018**, *19* (1), 199. https://doi.org/10.1186/s13059-018-1577-z.

(276) Drongitis, D.; Aniello, F.; Fucci, L.; Donizetti, A. Roles of Transposable Elements in the Different Layers of Gene Expression Regulation. *Int. J. Mol. Sci.* **2019**, *20* (22). https://doi.org/10.3390/ijms20225755.

(277) Belyayev, A. Bursts of Transposable Elements as an Evolutionary Driving Force. *J. Evol. Biol.* **2014**, *27* (12), 2573–2584. https://doi.org/10.1111/JEB.12513.

(278) Pray, L. A. Transposons: The Jumping Genes. *Nature Education*. 2008, p 204.

(279) Drummond, D. A.; Wilke, C. O. The Evolutionary Consequences of Erroneous Protein Synthesis. *Nat. Rev. Genet.* **2009**, *10* (10), 715–724. https://doi.org/10.1038/nrg2662.

(280) Glick, B. R. Metabolic Load and Heterologous Gene Expression. *Biotechnol. Adv.* **1995**, *13* (2), 247–261. https://doi.org/10.1016/0734-9750(95)00004-A.

(281) Kurland, C.; Gallant, J. Errors of Heterologous Protein Expression. *Curr. Opin. Biotechnol.* **1996**, *7* (5), 489–493. https://doi.org/10.1016/s0958-1669(96)80050-4.

(282) Kastberg, L. L. B.; Ard, R.; Jensen, M. K.; Workman, C. T. Burden Imposed by Heterologous Protein Production in Two Major Industrial Yeast Cell Factories: Identifying Sources and Mitigation Strategies. *Front. Fungal Biol.* **2022**, *3*, 1. https://doi.org/10.3389/ffunb.2022.827704.

(283) Kramer, E. B.; Farabaugh, P. J. The Frequency of Translational Misreading Errors in E. Coli Is Largely Determined by TRNA Competition. https://doi.org/10.1261/rna.294907.

(284) Kane, J. F. Effects of Rare Codon Clusters on High-Level Expression of Heterologous Proteins in Escherichia Coli. *Curr. Opin. Biotechnol.* **1995**, *6* (5), 494–500. https://doi.org/10.1016/0958-1669(95)80082-4.

(285) Amrani, N.; Jacobson, A. All Termination Events Are Not Equal: Premature Termination in Yeast Is Aberrant and Triggers NMD. **2013**.

(286) Yang, Q.; Yu, C.-H.; Zhao, F.; Dang, Y.; Wu, C.; Xie, P.; Sachs, M. S.; Liu, Y. ERF1 Mediates Codon Usage Effects on MRNA Translation Efficiency through Premature Termination at Rare Codons. *Nucleic Acids Res.* **2019**, *47* (17), 9243. https://doi.org/10.1093/NAR/GKZ710.

(287) Karamyshev, A. L.; Karamysheva, Z. N. Lost in Translation: Ribosome-Associated MRNA and Protein Quality Controls. *Front. Genet.* **2018**, *0* (OCT), 431. https://doi.org/10.3389/FGENE.2018.00431.

(288) Eguchi, Y.; Makanae, K.; Hasunuma, T.; Ishibashi, Y.; Kito, K.; Moriya, H. Estimating the Protein Burden Limit of Yeast Cells by Measuring the Expression Limits of Glycolytic Proteins. *Elife* **2018**, *7*. https://doi.org/10.7554/ELIFE.34595.

(289) Brunelle, J. L.; Green, R. Coomassie Blue Staining. In *Methods in Enzymology*; Lorsch, J., Ed.; Elsevier Inc, 2014; Vol. 541, pp 161–167. https://doi.org/10.1016/B978-0-12-420119-4.00013-6.

(290) Nakamura, Y.; Gojobori, T.; Ikemura, T. Codon Usage Tabulated from International DNA Sequence Databases: Status for the Year 2000. *Nucleic Acids Res.* **2000**, *28* (1), 292. https://doi.org/10.1093/nar/28.1.292.

(291) KEGG PATHWAY: Pantothenate and CoA biosynthesis - Saccharomyces cerevisiae (budding yeast) https://www.genome.jp/kegg-bin/show_pathway?sce00770 (accessed Nov 2, 2021).

(292) KEGG PATHWAY: Valine, leucine and isoleucine degradation - Saccharomyces cerevisiae (budding yeast) https://www.genome.jp/kegg-bin/show_pathway?sce00280 (accessed Nov 2, 2021).

(293) Lee, F. W. F.; Da Silva, N. A. Improved Efficiency and Stability of Multiple Cloned Gene Insertions at the Delta Sequences of Saccharomyces Cerevisiae. *Appl. Microbiol. Biotechnol.* **1997**, *48* (3), 339–345. https://doi.org/10.1007/s002530051059.

(294) Zhang, B.; Tian, W.; Wang, S.; Yan, X.; Jia, X.; Pierens, G. K.; Chen, W.; Ma, H.; Deng, Z.; Qu, X. Activation of Natural Products Biosynthetic Pathways via a Protein Modification Level Regulation. *ACS Chem. Biol.* **2017**, *12* (7), 1732–1736. https://doi.org/10.1021/acschembio.7b00225.

(295) Bond, C. M.; Tang, Y. Engineering Saccharomyces Cerevisiae for Production of Simvastatin. *Metab. Eng.* **2019**, *51*, 1–8. https://doi.org/10.1016/j.ymben.2018.09.005.

(296) Huang, M.; Bao, J.; Hallström, B. M.; Petranovic, D.; Nielsen, J. Efficient Protein Production by Yeast Requires Global Tuning of Metabolism. *Nat. Commun. 2017 81* **2017**, *8* (1), 1–12. https://doi.org/10.1038/s41467-017-00999-2.

(297) National Institute of Advanced Industrial Science and Technology (AIST). Spectral Database for Organic Compounds SDBS - anthranilic acid https://sdbs.db.aist.go.jp/sdbs/cgi-bin/direct_frame_top.cgi (accessed May 17, 2022).

(298) Kanehisa Laboratories. KEGG PATHWAY: Tryptophan metabolism - Saccharomyces cerevisiae (budding yeast) https://www.genome.jp/kegg-bin/show_pathway?sce00380 (accessed Dec 2, 2021).

(299) Krieger, C. Saccharomyces cerevisiae tryptophan degradation https://pathway.yeastgenome.org/YEAST/NEW-IMAGE?object=PWY3O-214 (accessed May 18, 2022).

(300) Caspi, R. Saccharomyces cerevisiae S288c tryptophan degradation VIII (to tryptophol) https://biocyc.org/YEAST/NEW-IMAGE?type=PATHWAY&object=PWY-5081 (accessed Dec 2, 2021).

(301) Ohashi, K.; Chaleckis, R. High Levels of Tryptophan Reduce Cell Wall or Membrane Stress Tolerance in Saccharomyces Cerevisiae. *Biosci. Biotechnol. Biochem.* **2021**, *85* (10), 2131–2136. https://doi.org/10.1093/BBB/ZBAB142.

(302) Ohashi, K.; Chaleckis, R.; Takaine, M.; Wheelock, C. E.; Yoshida, S. Kynurenine Aminotransferase Activity of Aro8/Aro9 Engage Tryptophan Degradation by Producing Kynurenic Acid in Saccharomyces Cerevisiae. *Sci. Reports 2017 71* **2017**, *7* (1), 1–8. https://doi.org/10.1038/s41598-017-12392-6.

(303) Sardar, P.; Kempken, F. Characterization of Indole-3-Pyruvic Acid Pathway-Mediated Biosynthesis of Auxin in Neurospora Crassa. *PLoS One* **2018**, *13* (2), e0192293. https://doi.org/10.1371/journal.pone.0192293.

(304) Koga, J.; Adachi, T.; Hidaka, H. Purification and Characterization of Indolepyruvate Decarboxylase. A Novel Enzyme for Indole-3-Acetic Acid Biosynthesis in Enterobacter Cloacae. *J. Biol. Chem.* **1992**, *267* (22), 15823–15828.

(305) Koga, J. Structure and Function of Indolepyruvate Decarboxylase, a Key Enzyme in Indole-3-Acetic Acid Biosynthesis. *Biochim. Biophys. Acta* **1995**, *1249* (1), 1–13. https://doi.org/10.1016/0167-4838(95)00011-I.

(306) Koga, J.; Adachi, T.; Hidaka, H. IAA Biosynthetic Pathway from Tryptophan via Indole-3-Pyruvic Acid in Enterobacter Cloacae. *Agric. Biol. Chem.* **1991**, *55* (3), 701–706. https://doi.org/10.1080/00021369.1991.10870678.

(307) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. Coli. *Science (80-.).* **2001**, *291* (5509), 1790–1792. https://doi.org/10.1126/science.1058092.

(308) Gao, X.; Wang, P.; Tang, Y. Engineered Polyketide Biosynthesis and Biocatalysis in Escherichia Coli. *Appl. Microbiol. Biotechnol.* **2010**, *88* (6), 1233–1242. https://doi.org/10.1007/s00253-010-2860-4.

(309) Mutka, S. C.; Carney, J. R.; Liu, Y.; Kennedy, J. Heterologous Production of Epothilone C and D in Escherichia Coli. *Biochemistry* **2006**, *45* (4), 1321–1330. https://doi.org/10.1021/bi052075r.

(310) Yang, G.; Zhang, Y.; Lee, N. K.; Cozad, M. A.; Kearney, S. E.; Luesch, H.; Ding, Y. Cyanobacterial Sfp-Type Phosphopantetheinyl Transferases Functionalize Carrier Proteins of Diverse Biosynthetic Pathways. *Sci. Rep.* **2017**, *7* (1), 11888. https://doi.org/10.1038/s41598-017-12244-3.

(311) D'Agostino, P. M.; Gulder, T. A. M. Direct Pathway Cloning Combined with Sequence- and Ligation-Independent Cloning for Fast Biosynthetic Gene Cluster Refactoring and Heterologous Expression. *ACS Synth. Biol.* **2018**, *7* (7), 1702–1708. https://doi.org/10.1021/ACSSYNBIO.8B00151/SUPPL_FILE/SB8B00151_SI_001.PDF.

(312) Greunke, C.; Duell, E. R.; D'Agostino, P. M.; Glöckle, A.; Lamm, K.; Gulder, T. A. M. Direct Pathway Cloning (DiPaC) to Unlock Natural Product Biosynthetic Potential. *Metab. Eng.* **2018**, *47*, 334–345. https://doi.org/10.1016/j.ymben.2018.03.010.

(313) KEGG PATHWAY: Phenylalanine metabolism - Escherichia coli BL21(DE3) https://www.genome.jp/kegg-bin/show_pathway?ebl00360 (accessed Nov 19, 2021).

(314) Schwarzer, D.; Mootz, H. D.; Marahiel, M. A. Exploring the Impact of Different Thioesterase Domains for the Design of Hybrid Peptide Synthetases. *Chem. Biol.* **2001**, *8* (10), 997–1010. https://doi.org/10.1016/S1074-5521(01)00068-0.

(315) Beck, Z. Q.; Aldrich, C. C.; Magarvey, N. A.; Georg, G. I.; Sherman, D. H. Chemoenzymatic Synthesis of

Cryptophycin/Arenastatin Natural Products. *Biochemistry* **2005**, *44* (41), 13457–13466. https://doi.org/10.1021/bi051140u.

(316) Klettke, K. L.; Sanyal, S.; Mutatu, W.; Walker, K. D. Beta-Styryl- and Beta-Aryl-Beta-Alanine Products of Phenylalanine Aminomutase Catalysis. *J. Am. Chem. Soc.* **2007**, *129* (22), 6988–6989. https://doi.org/10.1021/ja071328w.

(317) Turner, N. J. Ammonia Lyases and Aminomutases as Biocatalysts for the Synthesis of α-Amino and β-Amino Acids. *Curr. Opin. Chem. Biol.* **2011**, *15* (2), 234–240. https://doi.org/10.1016/j.cbpa.2010.11.009.

(318) D'cunha, G. B.; Satyanarayan, V.; Nair, P. M. Novel Direct Synthesis of L-Phenylalanine Methyl Ester by Using Rhodotorula Glutinis Phenylalanine Ammonia Lyase in an Organic-Aqueous Biphasic System. *Enzyme Microb. Technol.* **1994**, *16* (4), 318–322. https://doi.org/10.1016/0141-0229(94)90173-2.

(319) Gloge, A.; Zoń, J.; Kövári, A.; Poppe, L.; Rétey, J. Phenylalanine Ammonia-Lyase: The Use of Its Broad Substrate Specificity for Mechanistic Investigations and Biocatalysis--Synthesis of L-Arylalanines. *Chemistry* **2000**, *6* (18), 3386–3390. https://doi.org/10.1002/1521-3765(20000915)6:18<3386::aid-chem3386>3.0.co;2-5.

(320) Varga, A.; Bánóczi, G.; Nagy, B.; Bencze, L. C.; Toşa, M. I.; Gellért, Á.; Irimie, F. D.; Rétey, J.; Poppe, L.; Paizs, C. Influence of the Aromatic Moiety in α- and β-Arylalanines on Their Biotransformation with Phenylalanine 2,3-Aminomutase from Pantoea Agglomerans. *RSC Adv.* **2016**, *6* (61), 56412–56420. https://doi.org/10.1039/C6RA02964G.

(321) Strom, S.; Wanninayake, U.; Ratnayake, N. D.; Walker, K. D.; Geiger, J. H. Insights into the Mechanistic Pathway of the Pantoea Agglomerans Phenylalanine Aminomutase. *Angew. Chem. Int. Ed. Engl.* **2012**, *51* (12), 2898–2902. https://doi.org/10.1002/anie.201108525.

(322) Zhou, L.; Wang, Y.; Liu, H.; Han, L.; Zhang, W.; Cui, W.; Liu, Z.; Zhou, Z. Surface Engineering of a Pantoea Agglomerans-Derived Phenylalanine Aminomutase for the Improvement of (S)-β-Phenylalanine Biosynthesis. *Biochem. Biophys. Res. Commun.* **2019**, *518* (2), 204–211. https://doi.org/10.1016/j.bbrc.2019.08.031.

(323) Zhu, L.; Ge, F.; Li, W.; Song, P.; Tang, H.; Tao, Y.; Liu, Y.; Du, G. One Step Synthesis of Unnatural β-Arylalanines Using Mutant Phenylalanine Aminomutase from Taxus Chinensis with High β-Regioselectivity. *Enzyme Microb. Technol.* **2018**, *114*, 22–28. https://doi.org/10.1016/j.enzmictec.2018.03.006.

(324) Moffitt, M. C.; Louie, G. V.; Bowman, M. E.; Pence, J.; Noel, J. P.; Moore, B. S. Discovery of Two Cyanobacterial Phenylalanine Ammonia Lyases: Kinetic and Structural Characterization. *Biochemistry* **2007**, *46* (4), 1004–1012. https://doi.org/10.1021/bi061774g.

(325) Calabrese, J. C.; Jordan, D. B.; Boodhoo, A.; Sariaslani, S.; Vannelli, T. Crystal Structure of Phenylalanine Ammonia Lyase: Multiple Helix Dipoles Implicated in Catalysis. *Biochemistry* **2004**, *43* (36), 11403–11416. https://doi.org/10.1021/bi049053+.

(326) Ahmed, S. T.; Parmeggiani, F.; Weise, N. J.; Flitsch, S. L.; Turner, N. J. Engineered Ammonia Lyases for the Production of Challenging Electron-Rich l -Phenylalanines. *ACS Catal.* **2018**, *8* (4), 3129–3132. https://doi.org/10.1021/ACSCATAL.8B00496/SUPPL_FILE/CS8B00496_SI_001.PDF.

(327) Rowles, I.; Groenendaal, B.; Binay, B.; Malone, K. J.; Willies, S. C.; Turner, N. J. Engineering of Phenylalanine Ammonia Lyase from Rhodotorula Graminis for the Enhanced Synthesis of Unnatural L-Amino Acids. *Tetrahedron* **2016**, *72* (46), 7343–7347. https://doi.org/10.1016/j.tet.2016.06.026.

(328) Cyano Biotech GmbH http://www.cyano-biotech.com/content/home/index.php (accessed Jan 17, 2022).

(329) Algade Amadu, A.; Amoako Atta deGraft-Johnson, K.; Komla Ameka, G. Industrial Applications of Cyanobacteria. In *Cyanobacteria - Recent Advances in Taxonomy and Applications*; IntechOpen, 2022. https://doi.org/10.5772/intechopen.98859.

(330) LivingInkTech — About Living Ink https://livingink.co/about (accessed Jan 17, 2022).

(331) Photanol https://photanol.com/#technology (accessed Jan 17, 2022).

(332) Cyano Biotech GmbH | Cyanotoxins http://www.cyano-biotech.com/content/cyanotoxins/index.php (accessed Jan 17, 2022).

(333) Cyano Biotech GmbH | Bioreagents http://www.cyano-biotech.com/content/bioreagents/index.php (accessed Jan 17, 2022).

(334) Cyano Biotech GmbH | Strain development http://www.cyano-biotech.com/content/straindevelopment/index.php (accessed Jan 17, 2022).

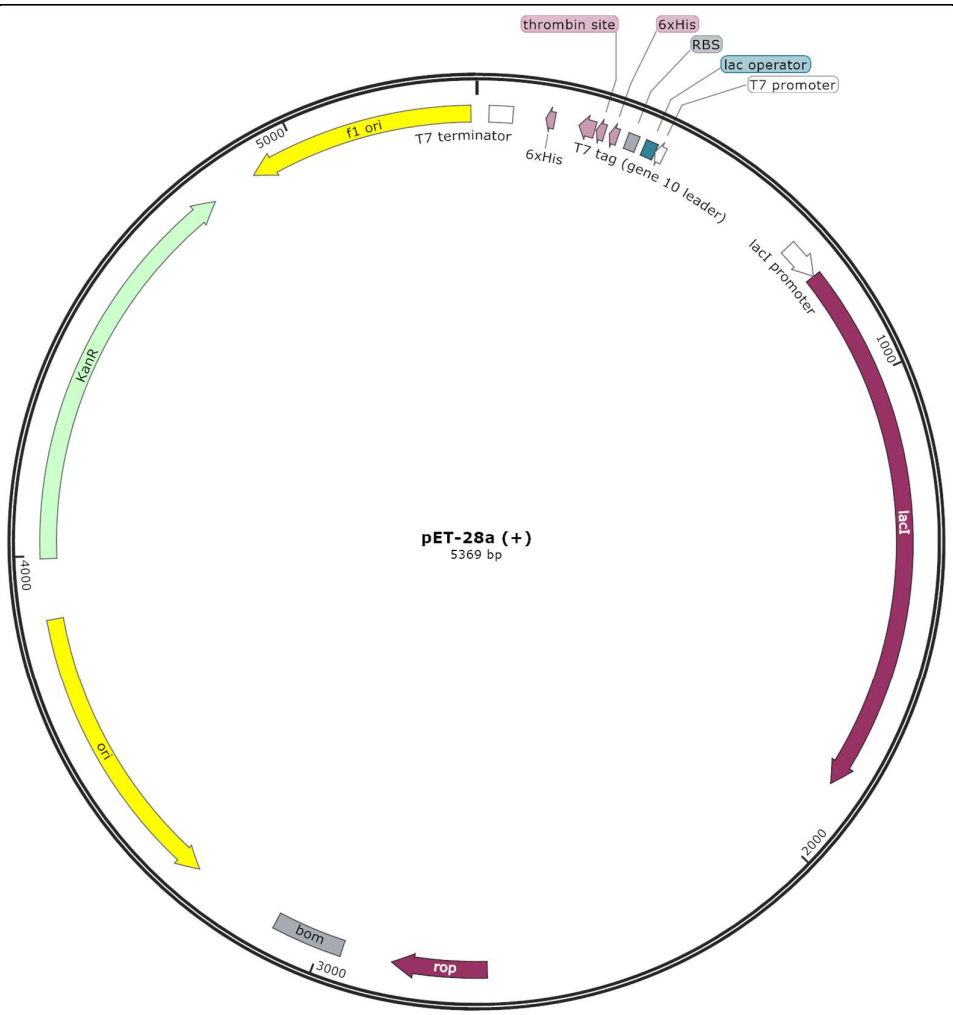
(335) Malakhov, M. P.; Mattern, M. R.; Malakhova, O. A.; Drinker, M.; Weeks, S. D.; Butt, T. R. SUMO Fusions and SUMO-Specific Protease for Efficient Expression and Purification of Proteins. *J. Struct. Funct. Genomics* **2004**, *5* (1/2), 75–86. https://doi.org/10.1023/B:JSFG.0000029237.70316.52.

7 Supplementary information

7.1 Supplementary tables

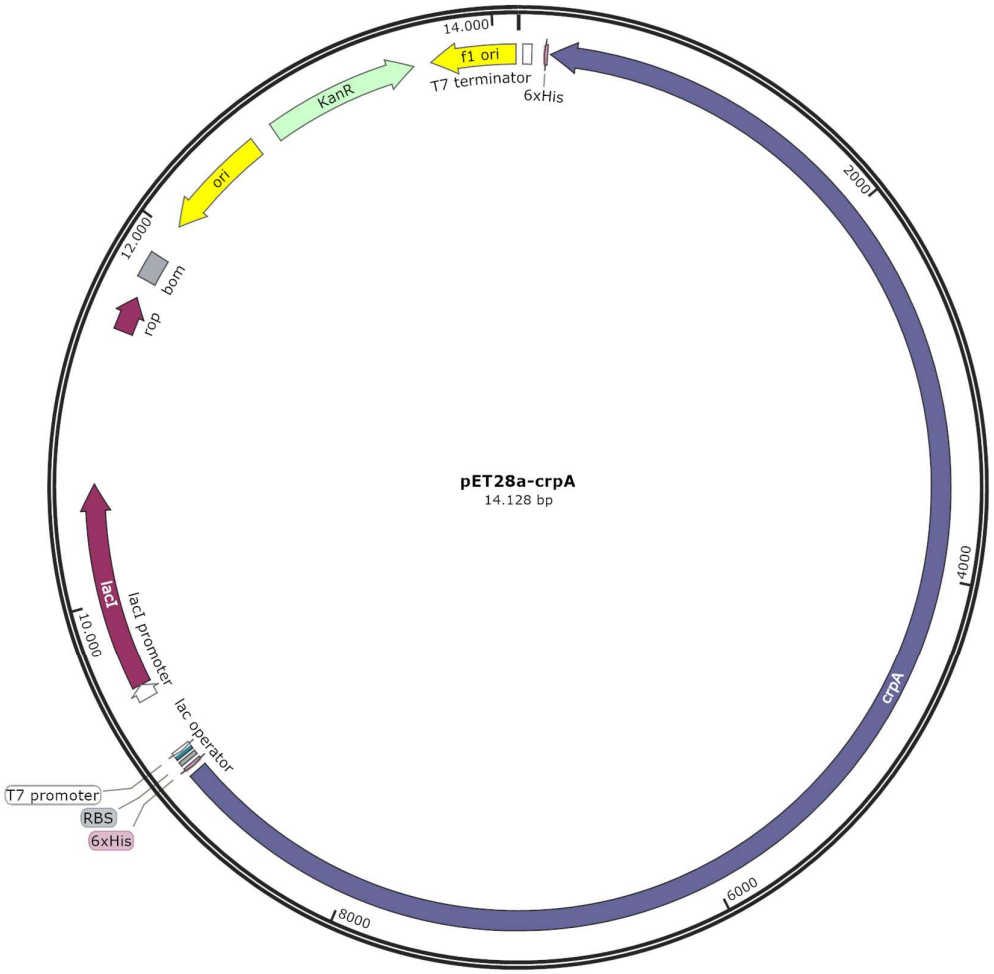
Table SI I. *E. coli* plasmids, vector maps and relevant specifications. Vector Maps were designed and annotated with SnapGene.

pET28a(+) (Novagen)



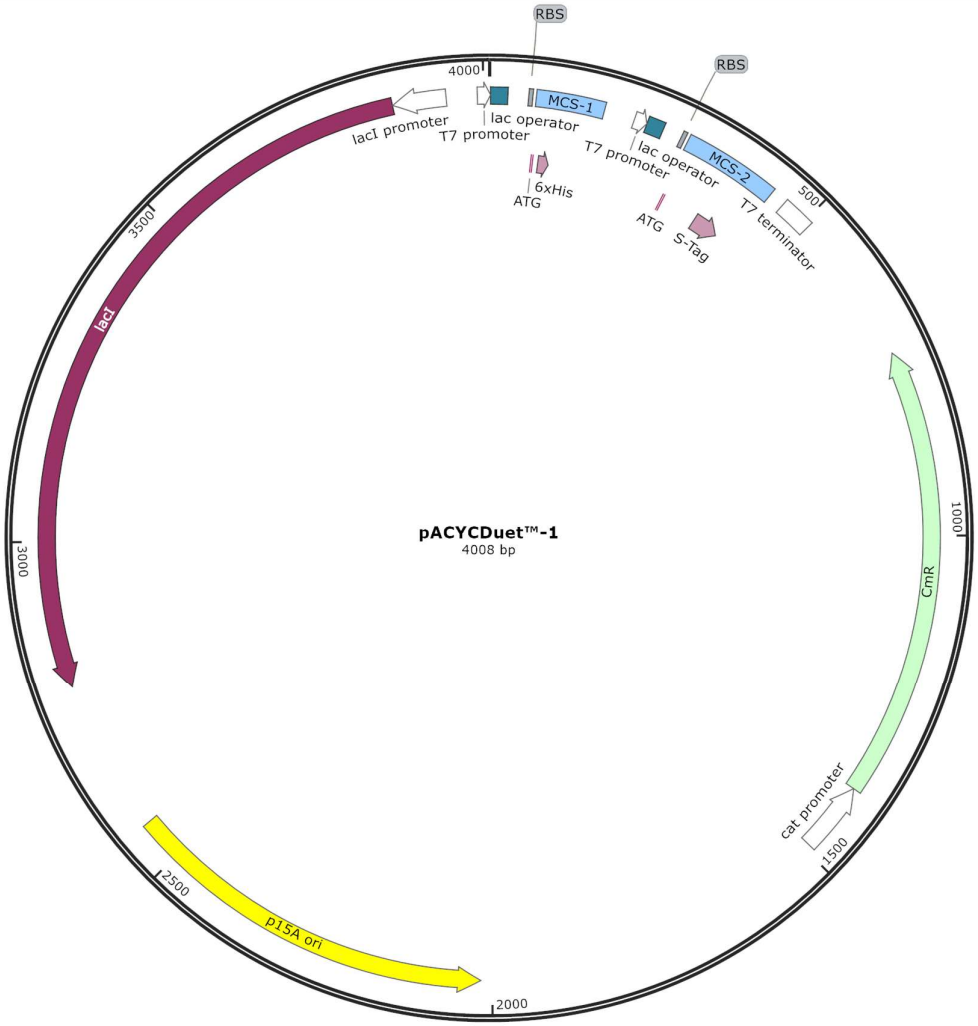
f1 ori: origin of replication
KanR: gene encoding kanamycin resistance
Ori: high-copy origin of replication
bom: site of mobilization of pBR322
rop: gene encoding rop-regulator protein for plasmid replication
lacI: gene encoding the lac-repressor protein
T7 Promotor: T7 Promotor of RNA Polymerase
RBS: ribosomal binding site
6xHis: hexahistidine -tag
thrombin-site: thrombine recognition and restriction sequence
T7 tag (gene 10 leader): encoding lead peptide of T7 bacteriophage gene10
T7 Terminator: terminator of T7 bacteriophage RNA Polymerase

pET28a-crpA



Equivalent to pET28a (+)
crpA: gene encoding the first protein of cryptophycin biosynthesis in *Nostoc* sp. ATCC 53789

pACYCDuet (Novagen)

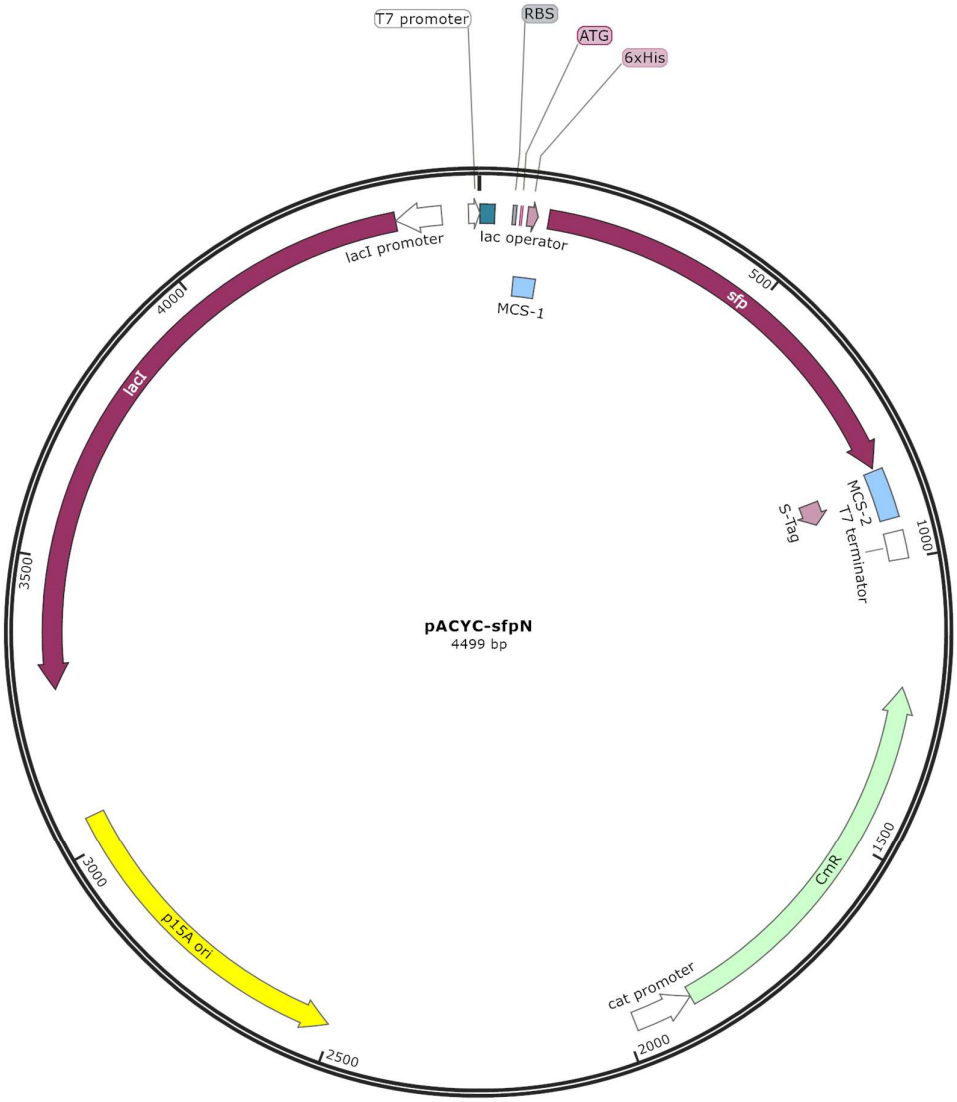


T7 Promotor: T7 Promotor of RNA Polymerase
T7 Terminator: terminator of T7 bacteriophage RNA Polymerase
MCS 1/ MCS 2 multiple cloning site
CmR: gene encoding chloramphenicol resistance
Cat Promoter: Promoter of CmR resistance gene
6xHis: hexahistidine -tag
S-Tag: affinity and epitope –tag of pancreatic ribonuclease A
ATG: Start codon
lac Operator: binding site of the lac-repressors regulating expression of downstream encoded genes
P15A ori: origin of replication
bom: site of mobilization of pBR322
rop: gene encoding rop-regulator protein for plasmid replication
lacI: gene encoding the lac-repressor protein
RBS: ribosomal binding site

pACYC-sfp

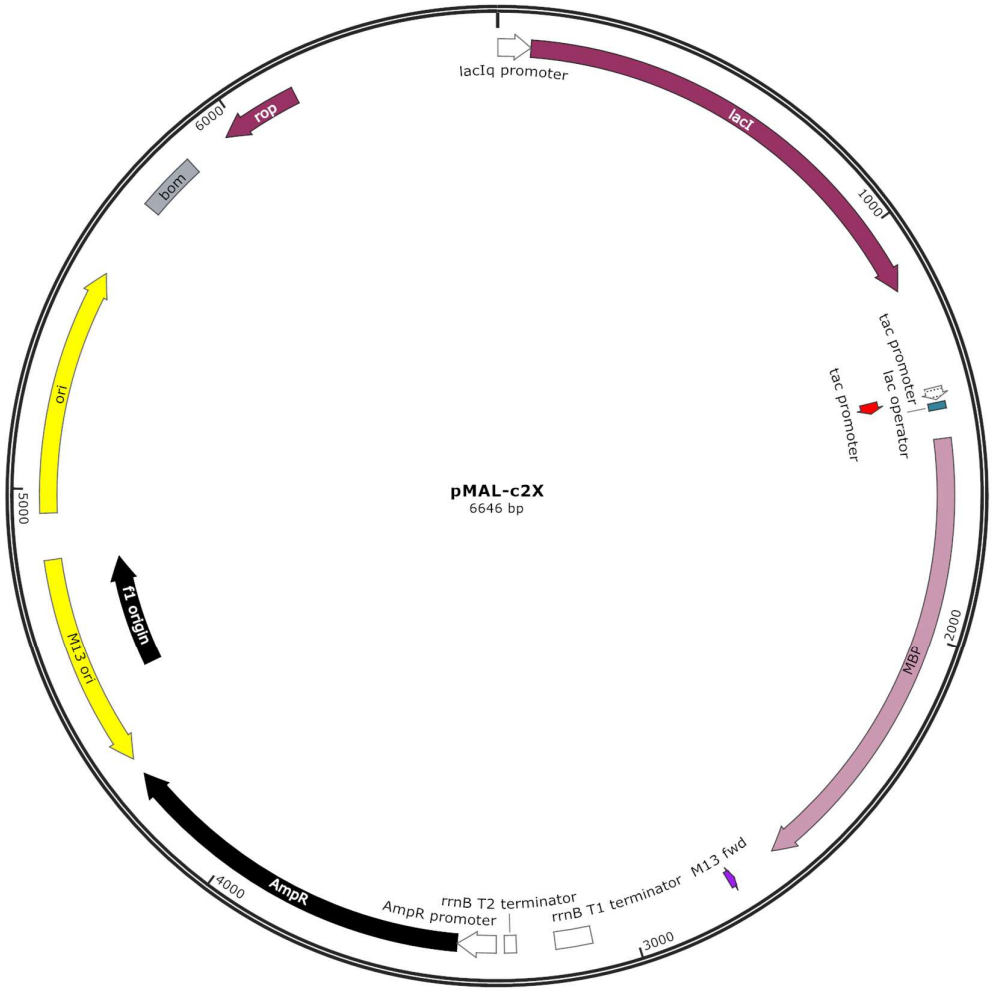
Equivalent to pACYCDuet
Sfp: gene encoding Sfp-type phosphopantetheinyl transferase from *Bacillus subtilis*

pACYC-sfpN [a]



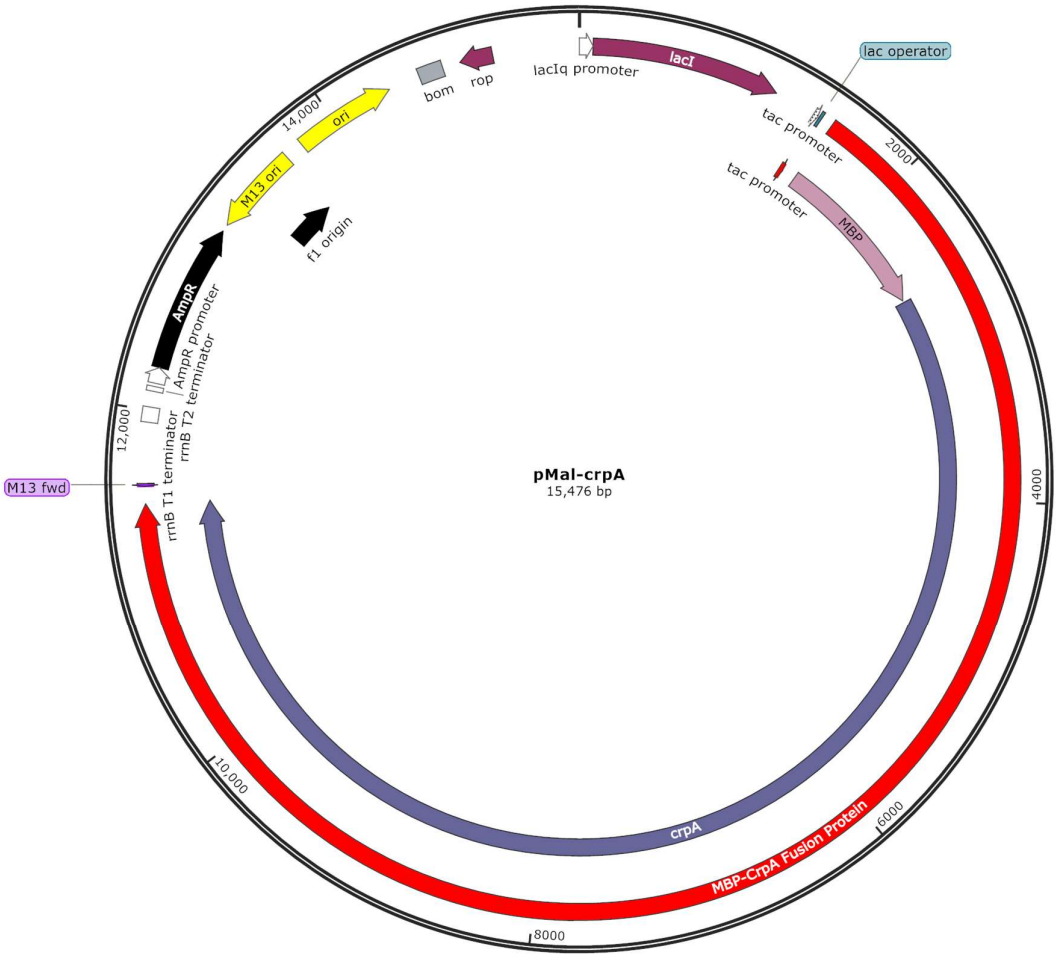
Equivalent to pACYCDuet
SfpN: gene encoding Sfp-type phosphopantetheinyl transferase from *Nostoc* sp. ATCC 53789

pMalc2x (Addgene)



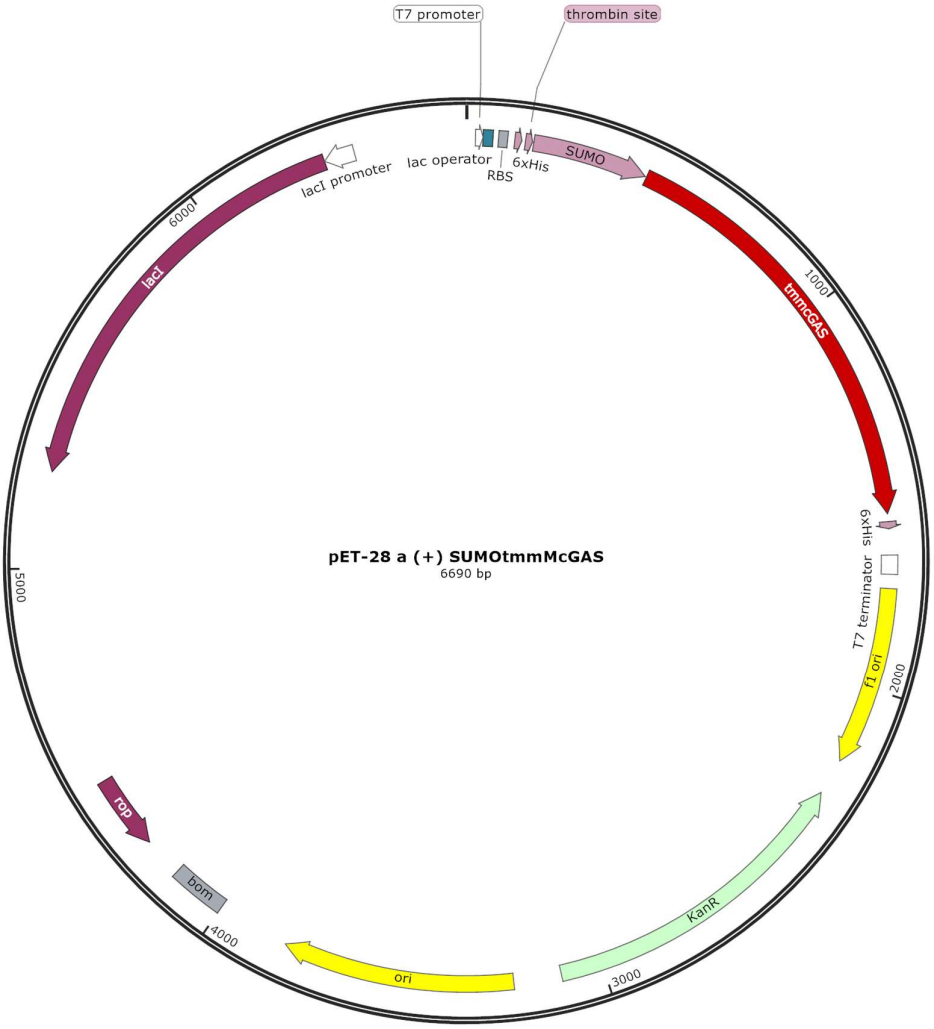
lacI: gene encoding the lac-repressor protein
lacIq promoter: Promoter of *lacI*
tac promoter hybrid trp/lac-UV5 promoter
lac Operator: binding site of the lac-repressors regulating expression of downstream encoded genes
MBP: encoding N-terminal maltose binding protein to improve translational expression and protein solubility [215]
M13 fwd: sequencing primer binding site
rrnB T1 terminator: terminator T1 of *rrnB E. coli* gene
rrnB T2 terminator: terminator T2 of *rrnB E. coli* gene
AmpR promoter: Ampicillin resistance gene promoter
AmpR: gene encoding ampicillin resistance
M13 ori: M13 bacteriophage origin of replication
f1 origin: origin of replication
ori: high-copy origin of replication ColE1/pMB1/pBR322/pUC
bom: site of mobilization of pBR322
rop: gene encoding rop-regulator protein for plasmid replication

pMal-crpA



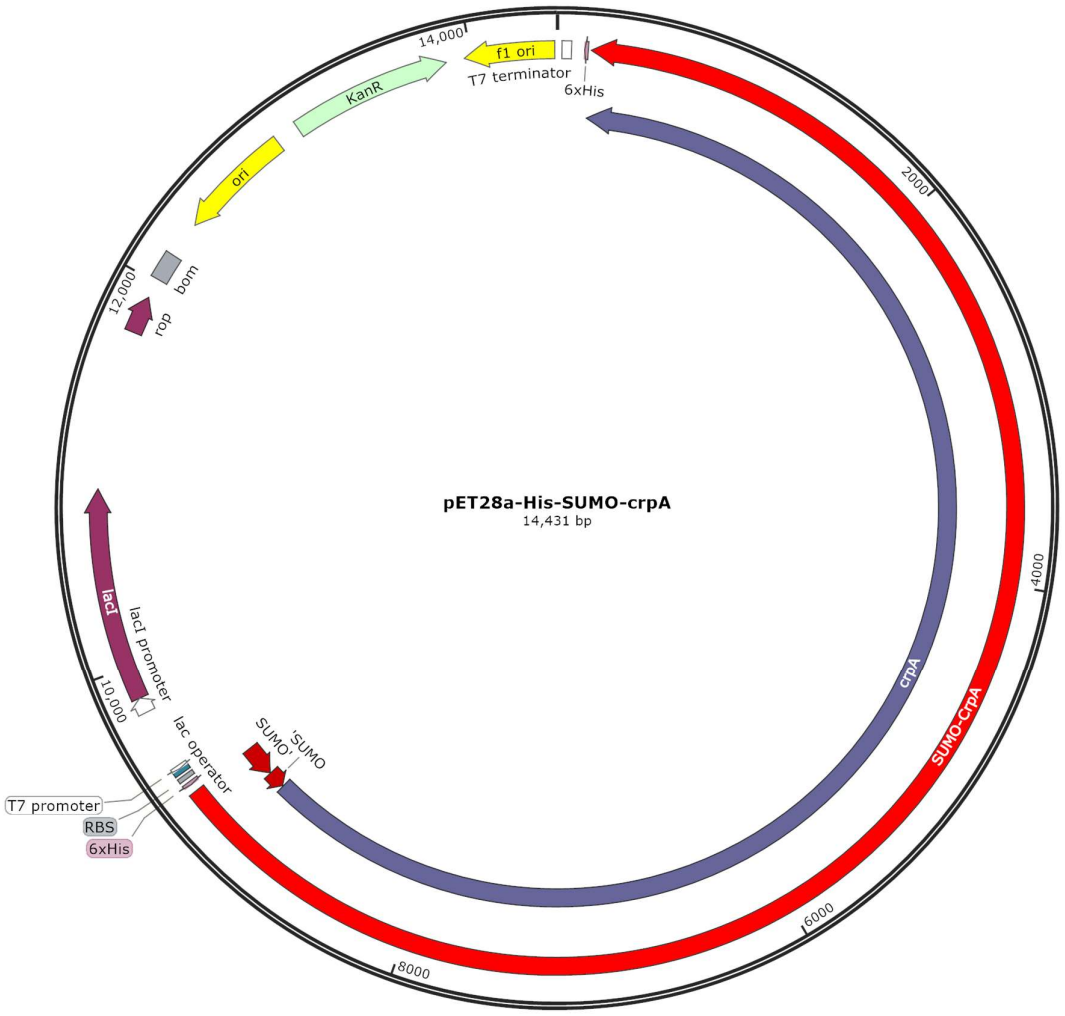
Equivalent to pMalc2x
crpA: gene encoding the first protein of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789
MBP-CrpA Fusion Protein: Sequence locus of MBP-CrpA fusion construct

pET28a-tmmrMcGAS-SUMO [216]



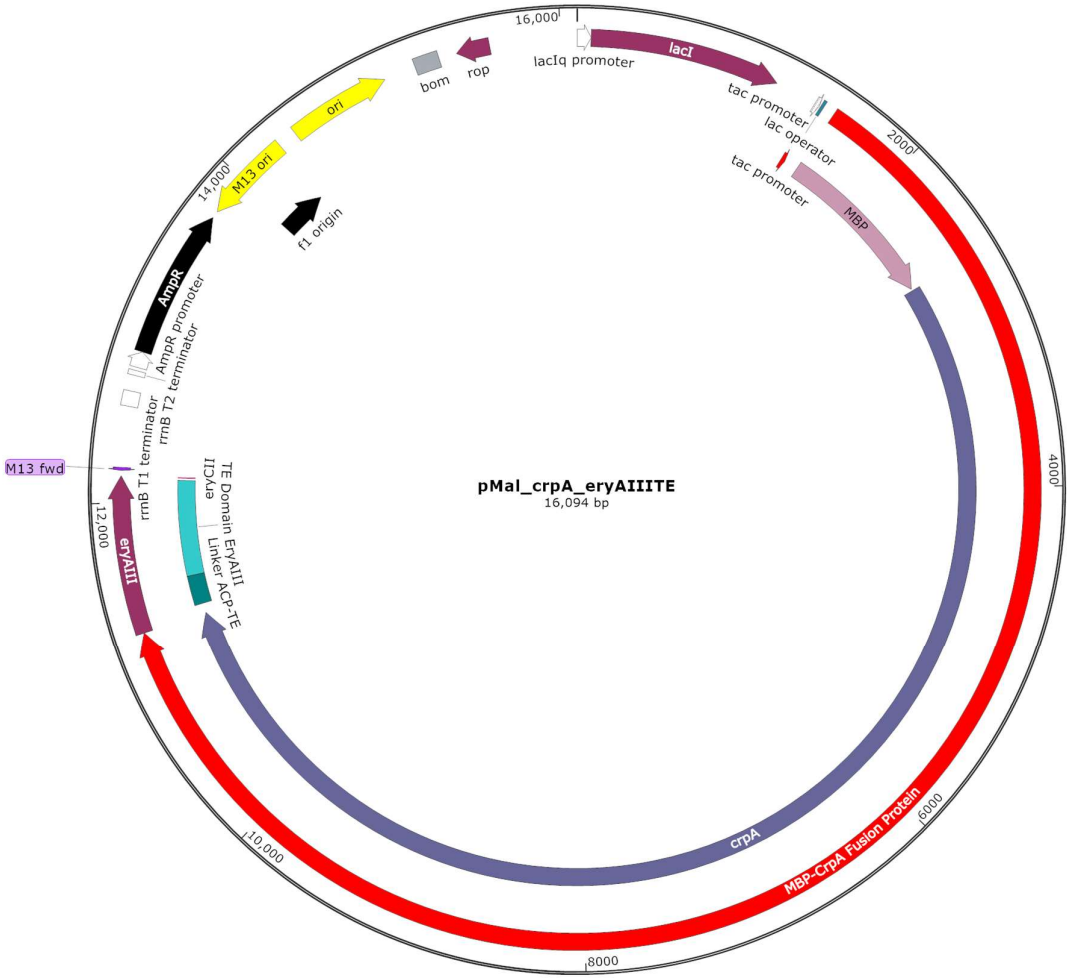
Equivalent to pET28a (+)
Original plasmid providing SUMO-tag.
SUMO: N-terminal small ubiquitin-related modifier modulates protein structure by covalent binding to the lysine side enhancing protein expression and solubility [335]; the vector was donated by Regine Siedentop [216]

pET28a-SUMO-crpA



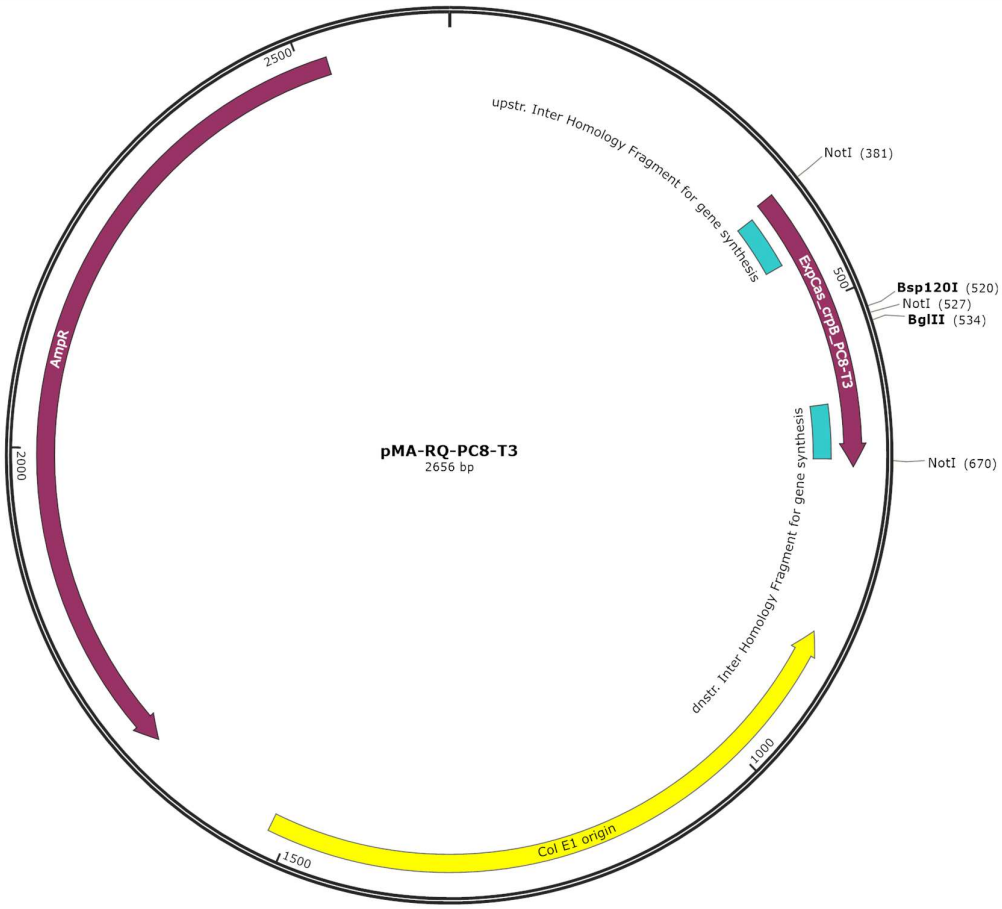
Equivalent to pET28a (+)
crpA: gene encoding the first protein of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789
SUMO-crpA: Sequence locus of SUMO-crpA fusion construct
SUMO: N-terminal small ubiquitin-related modifier modulates protein structure by covalent binding to the lysine side enhancing protein expression and solubility [335]; the SUMO-tag was acquired from pET28a-tmmrMcGAS-SUMO [216]

pMal-crpA-eryAIII$_{TE}$



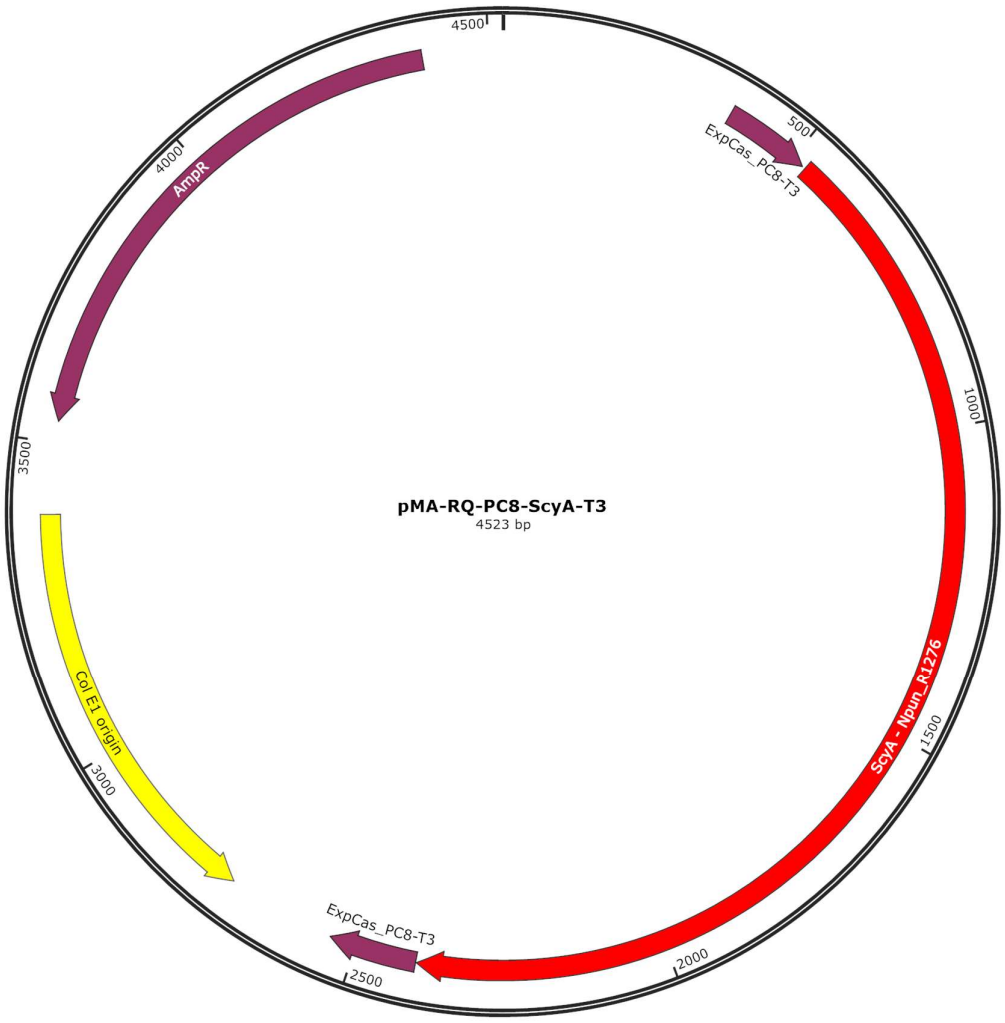
Equivalent to pMalc2x
crpA: gene encoding the first protein of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789
eryAIII: gene encoding the TE domain (thioesterase) of erythromycin biosynthesis
TE Domain: coding sequence of the erythromycin TE domain
MBP-CrpA Fusion Protein: Sequence locus of MBP-CrpA fusion construct
Linker ACP-TE: sequence locus of the TE-ACP (Acyl-Carrier-Protein) linker domain

pMA-RQ-PC8-T3 (invitrogen)



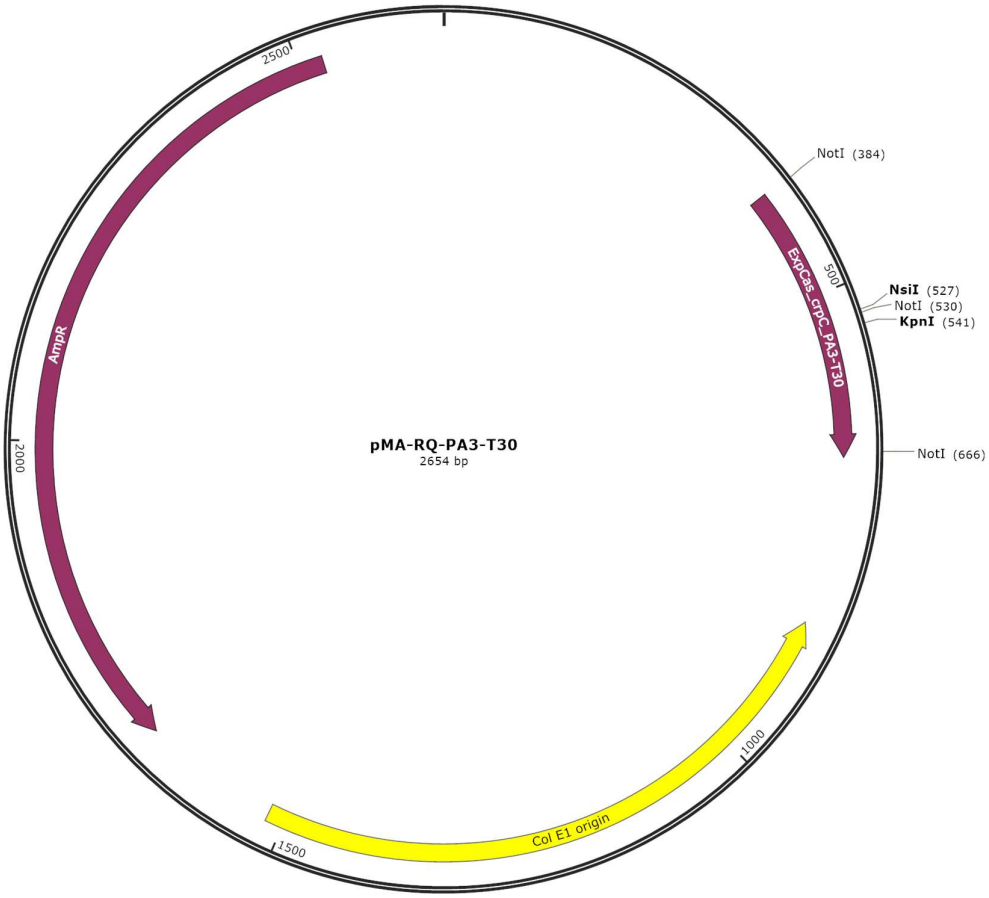
AmpR: gene encoding ampicillin resistance
Col E1 origin: origin of replication
ExpCas_PC8-T3: synthetic minimal promoter and terminator cassette for expression in *S. cerevisiae* [169,170]
upstr. Inter Homology Fragment for gene synthesis: upstream synthetic homologous flank of locus 416d (ARS416/ARS1) of *S. cerevisiae* chromosome IV for genomic integration by homologous recombination
dnstr. Inter Homology Fragment for gene synthesis: downstream synthetic homologous flank of locus 416d (ARS416/ARS1) of *S. cerevisiae* chromosome IV for genomic integration by homologous recombination
containing homologous flanking regions for yeast locus ARS416/ARS1 chromosome IV [156]
Containing unique restriction recognition sites Bsp120I and BglII to functionally clone any gene of interest into the artificial minimal expression system. Contains unique restriction recognition site NotI to individually liberate the promoter and terminator cassette and any subsequently cloned gene of interest as an individual cassette equipped with a functional expression system.

pMA-RQ-PC8-ScyA-T3



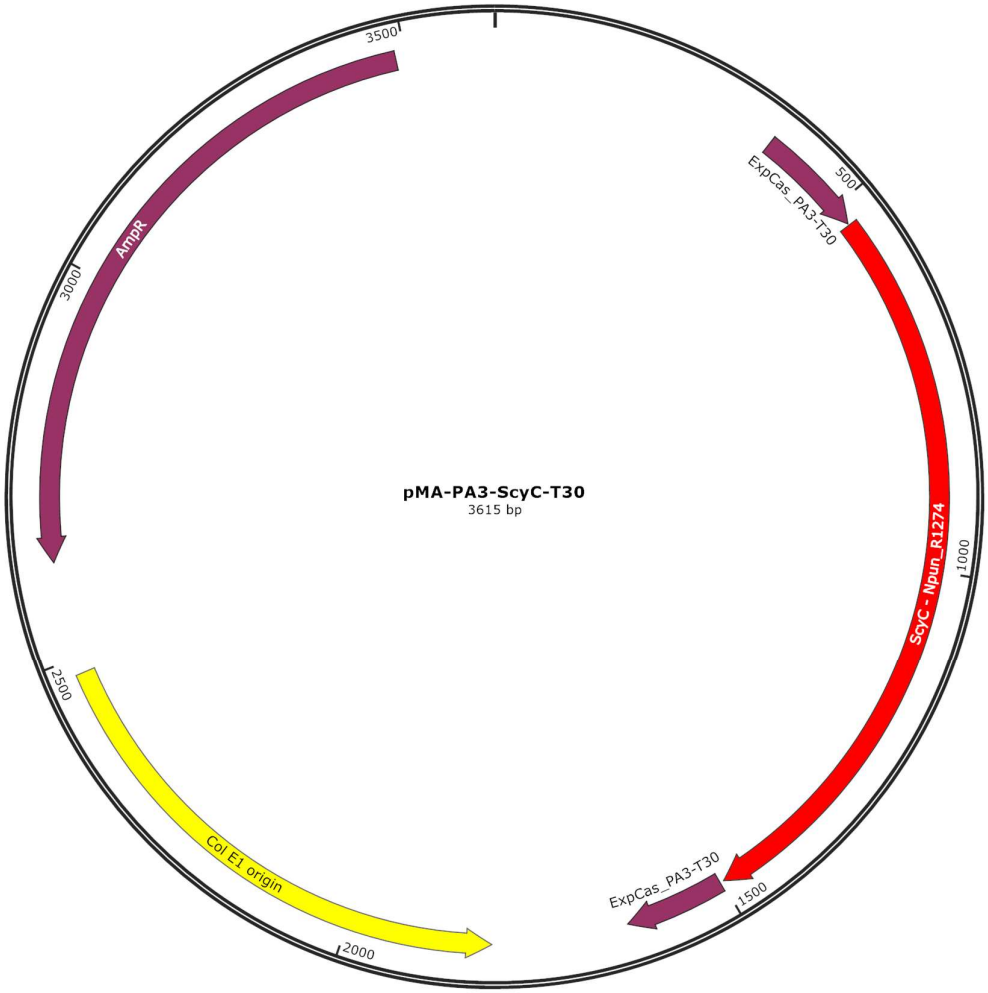
Equivalent to pMA-RQ-PC8-T3
ScyA-Npun_R1276: Gene encoding the first protein of scytonemin biosynthesis in *Nostoc punctiforme* cloned between the promoter PC8 and the terminator T3

pMA-RQ-PA3-T30 (invitrogen)



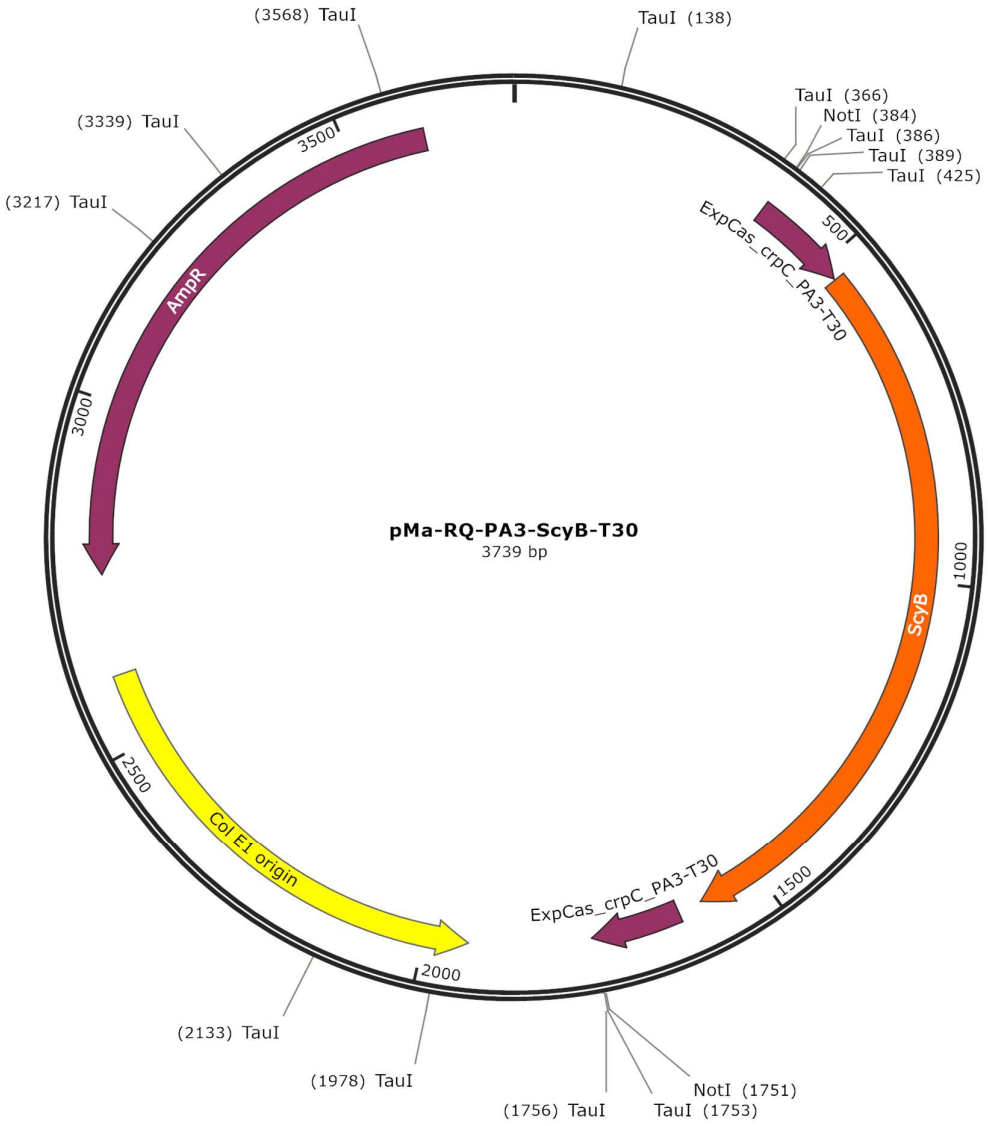
AmpR: gene encoding ampicillin resistance
Col E1 origin: origin of replication
ExpCas_PA3-T30: synthetic minimal promoter and terminator cassette for expression in *S. cerevisiae* [169,170]. Containing unique restriction recognition sites NsiI and KpnI to functionally clone any gene of interest into the artificial minimal expression system. Contains unique restriction recognition site NotI to individually liberate the promoter and terminator cassette and any subsequently cloned gene of interest as an individual cassette equipped with a functional expression system.

pMA-RQ-PA3-scyC-T30



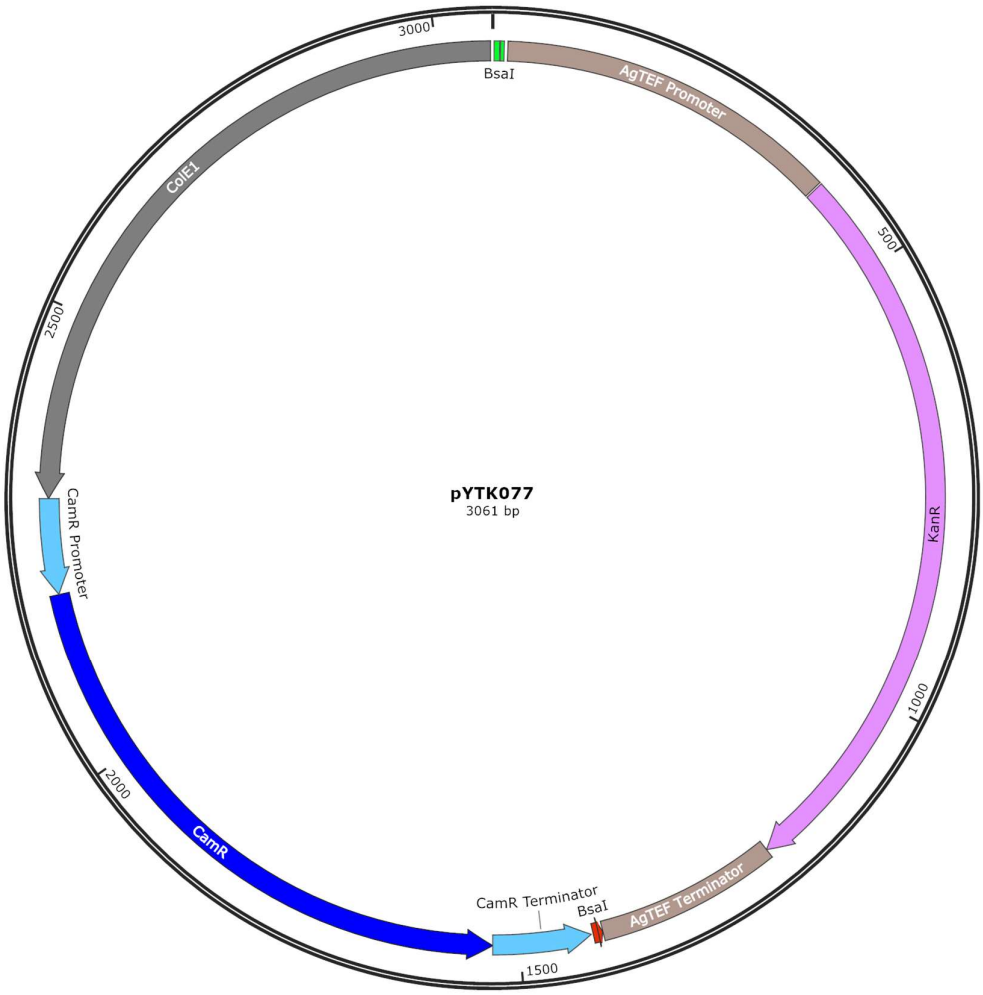
Equivalent to pMA-RQ-PA3-T30
ScyC-Npun_R1274: Gene encoding the third protein of scytonemin biosynthesis in *Nostoc punctiforme* cloned between the promoter PA3 and the terminator T30

pMA-RQ-PA3-scyB-T30



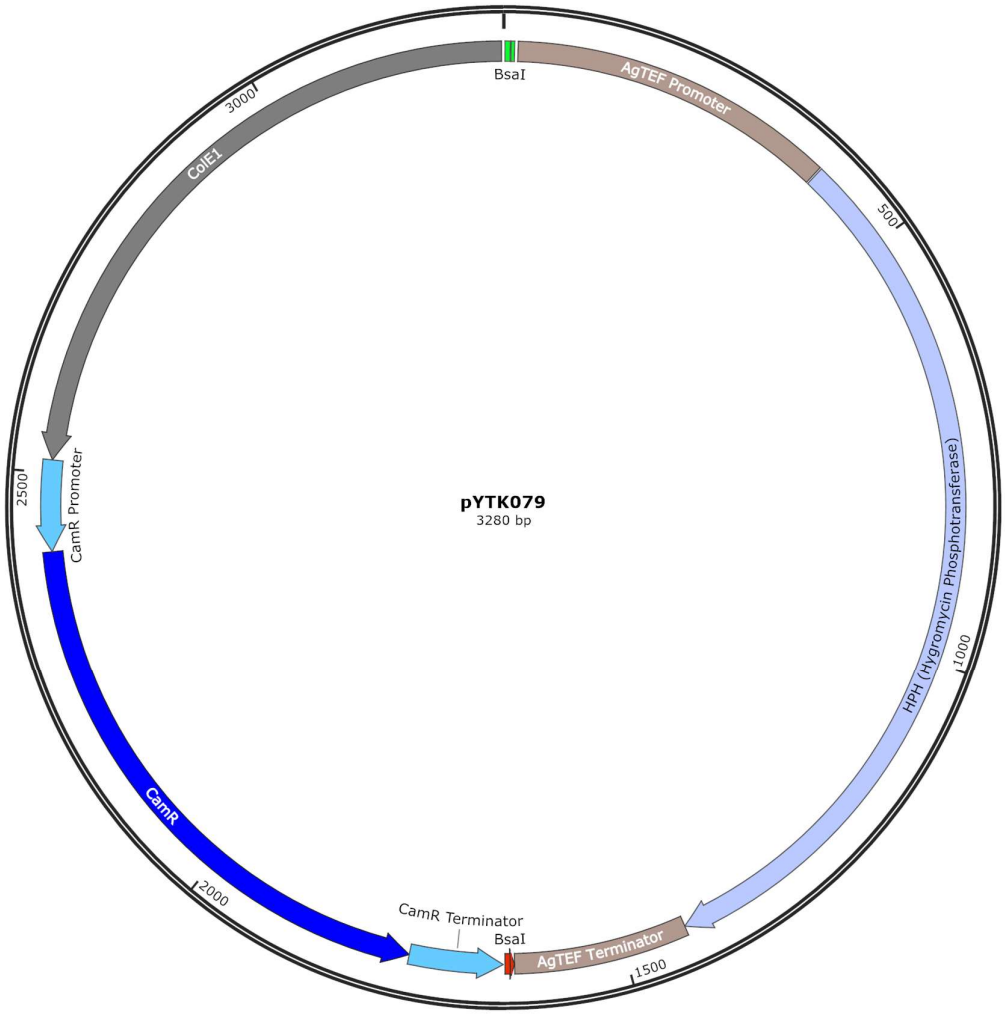
Equivalent to pMA-RQ-PA3-T30
ScyB: Gene encoding the second protein of scytonemin biosynthesis in *Nostoc punctiforme* cloned between the promoter PA3 and the terminator T30
NotI: restriction recognition sequence
TauI: restriction recognition sequence

pYTK077 (Addgene) [158]



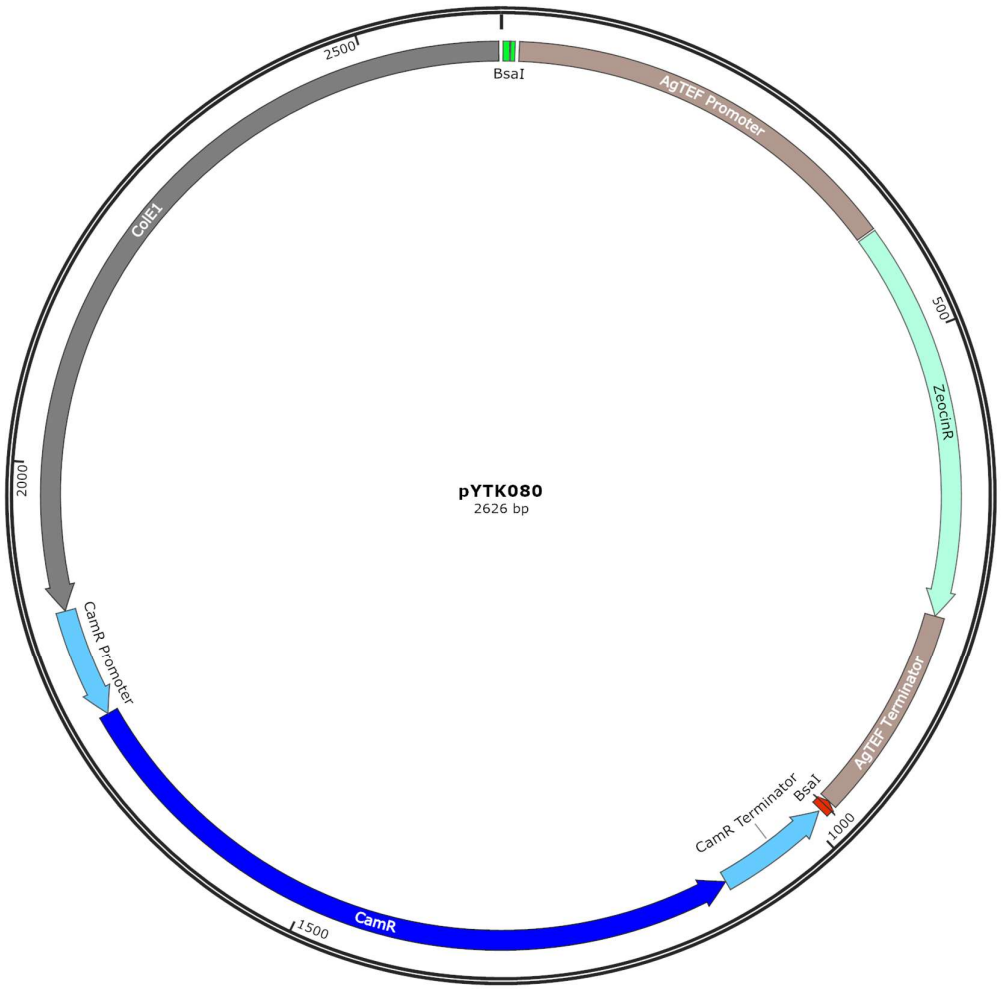
Yeast Toolkit containing diverse building bricks
Col E1 origin: origin of replication
BsaI: recognition sequence for BsaI restriction enzyme
AgTEF Promoter: *S. cerevisiae* promoter
KanR: Kanamycin/G418 resistance gene for *S. cerevisiae*
AgTEF Terminator: *S. cerevisiae* terminator
CamR Terminator: *E. coli* terminator of CamR
CamR: chloramphenicol resistance gene for *E. coli*
CamR Promoter: *E. coli* promoter of CamR

pYTK079 (Addgene) [158]



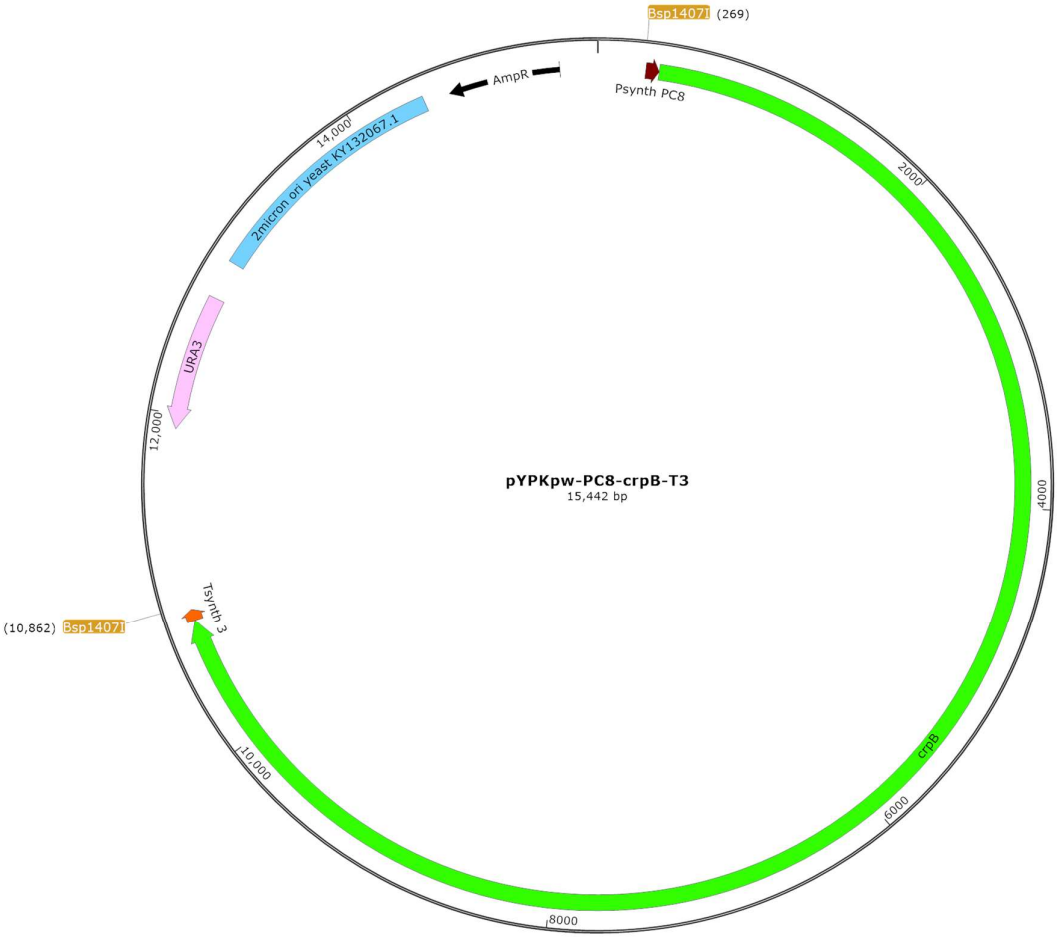
Yeast Toolkit containing diverse building bricks
Equivalent to pYTK077 except for yeast specific toolkit building brick
HPH: encoding hygromycin phosphotransferase gene for hygromycin resistance in *S. cerevisiae*

pYTK080 (Addgene) [158]



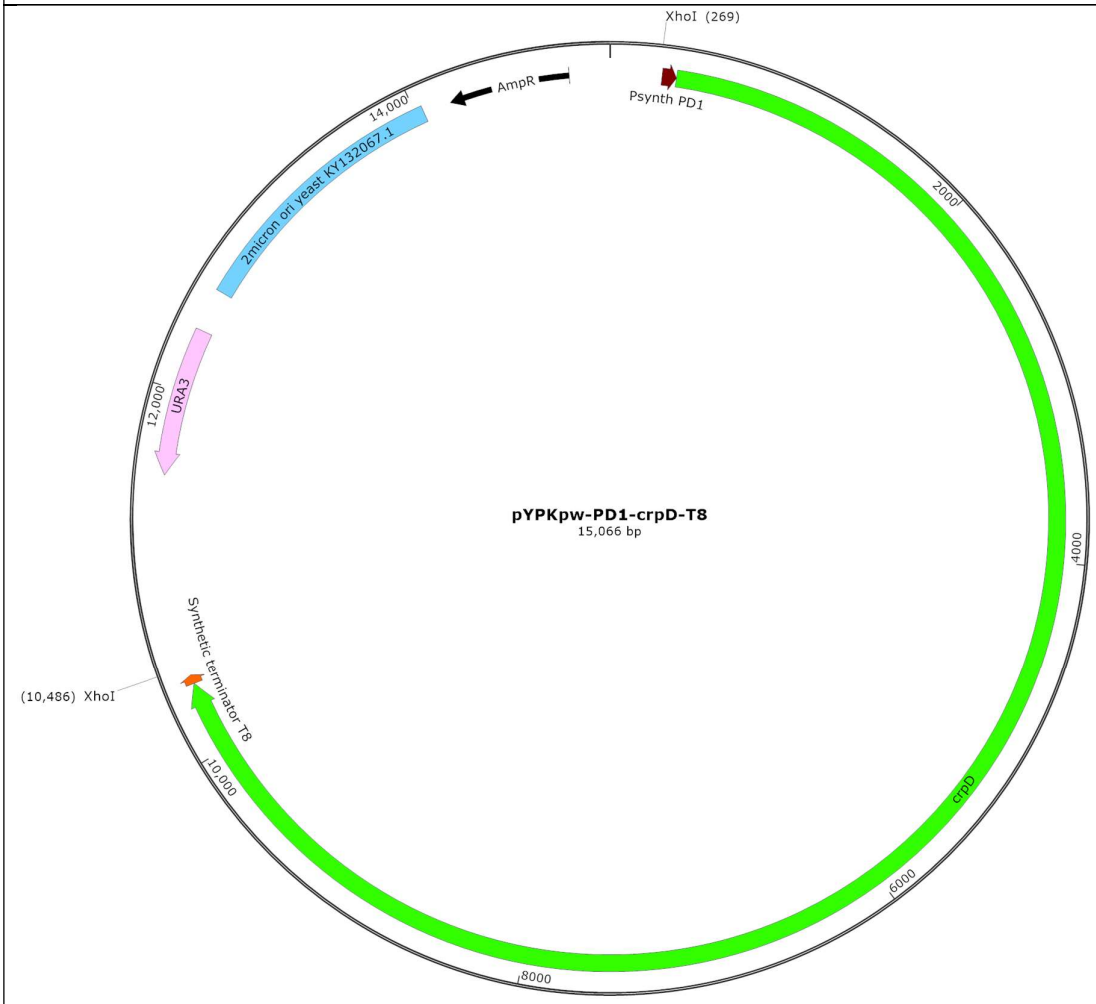
Yeast Toolkit containing diverse building bricks
Equivalent to pYTK077 and pYTK079 except for yeast specific toolkit building brick
ZeocinR: encoding gene for zeocin resistance in *S. cerevisiae*

pYPKpw-PC8-crpB-T3 (ATG:biosynthetics GmbH)



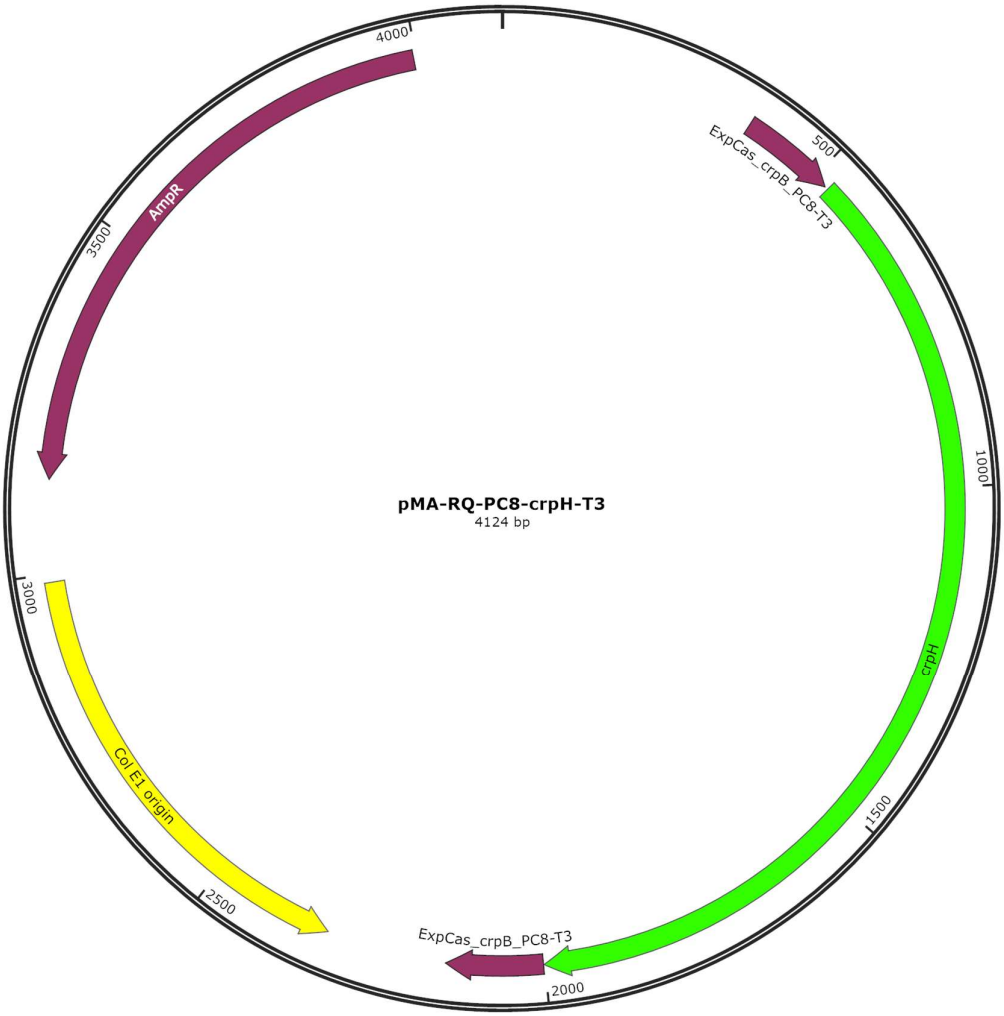
Psynth PC8: minimal synthetic promoter for *S. cerevisiae*
Tsynth 3: minimal synthetic terminator for *S. cerevisiae*
crpB: encoding the second gene of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789
URA3: selective marker encoding orotidine-5'-phosphate decarboxylase, required for uracil biosynthesis in *S. cerevisiae*
2micron ori yeast KY132067.1: *S. cerevisiae* 2µ plasmid origin of replication (multicopy)
AmpR: gene encoding ampicillin resistance in *E. coli*
ColEIori: *E. coli* origin of replication
Bsp1407I: unique pair of restriction recognition sites to liberate the *crpB*-expression cassette

pYPKpw-PD1-crpD-T8 (ATG:biosynthetics GmbH)



Equivalent to pYPKpw-PC8-crpB-T3 except for cryptophycin gene, minimal synthetic expression system and homologous flanks for recombination
Psynth PD1: minimal synthetic promoter for *S. cerevisiae*
Synthetic terminator T8: minimal synthetic terminator for *S. cerevisiae*
crpD: encoding the fourth gene of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789
XhoI: unique pair of restriction recognition sites to liberate the *crpD*-expression cassette

pMA-RQ-PC8-crpH-T3

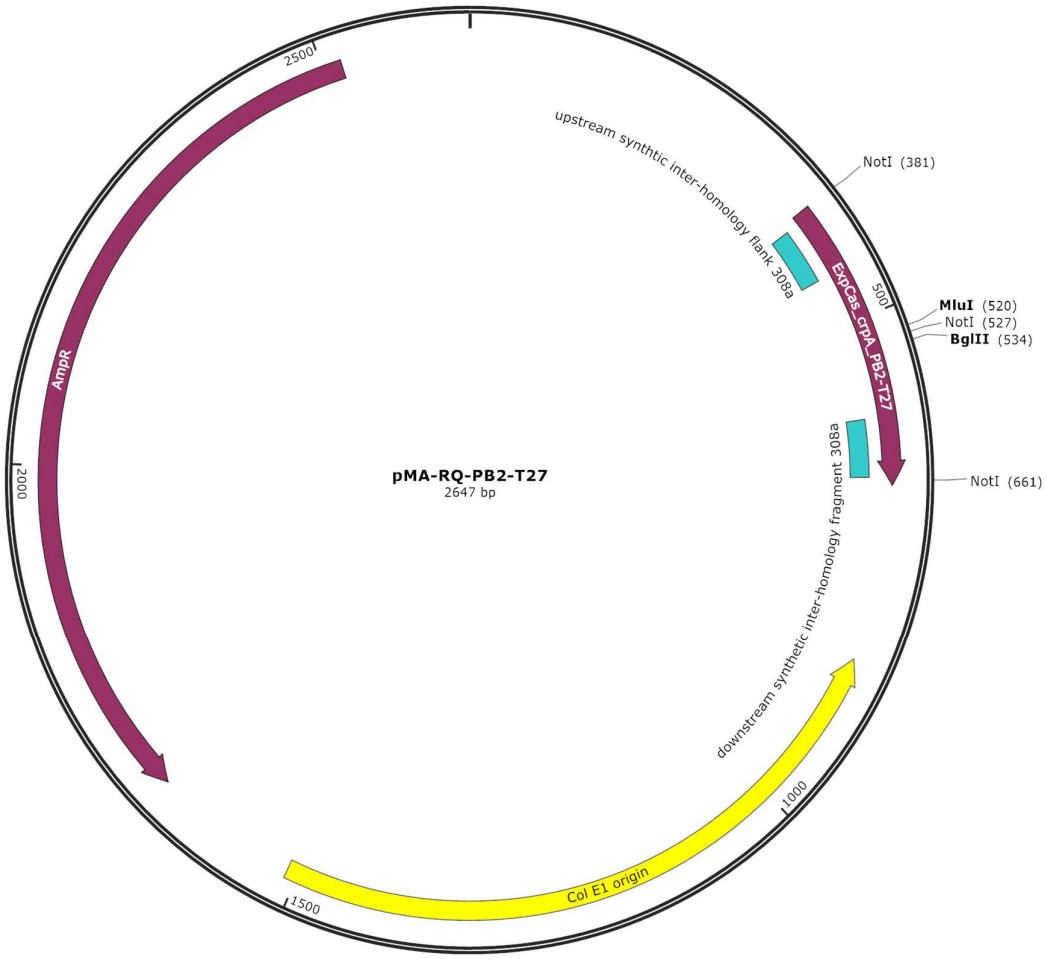


Equivalent to pMA-RQ-PC8-T3
crpH: encoding the last gene/ a halogenase of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789

pMA-RQ-PA3-crpC-T30

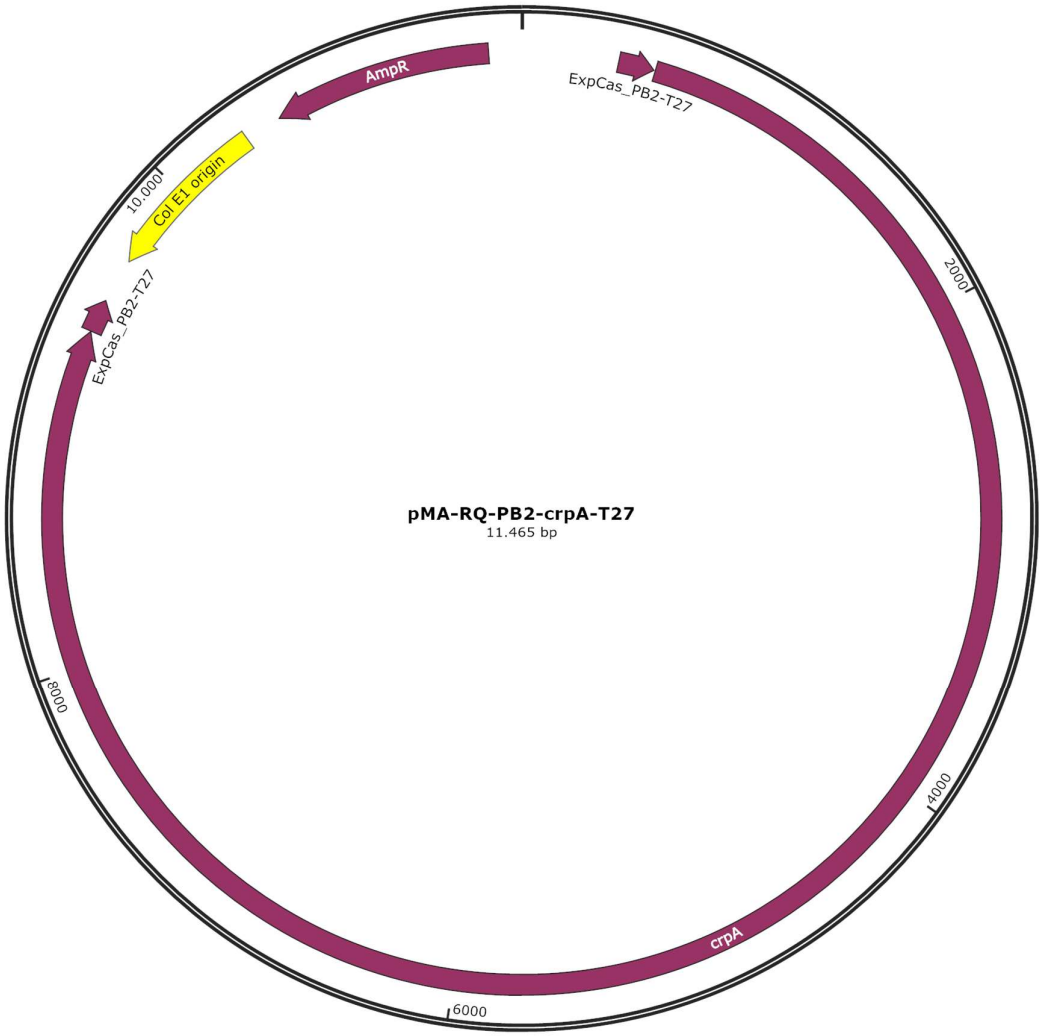
Equivalent to pMA-RQ-PA3-T30
crpC: encoding the third gene of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789

pMA-RQ-PB2-T27 (invitrogen)



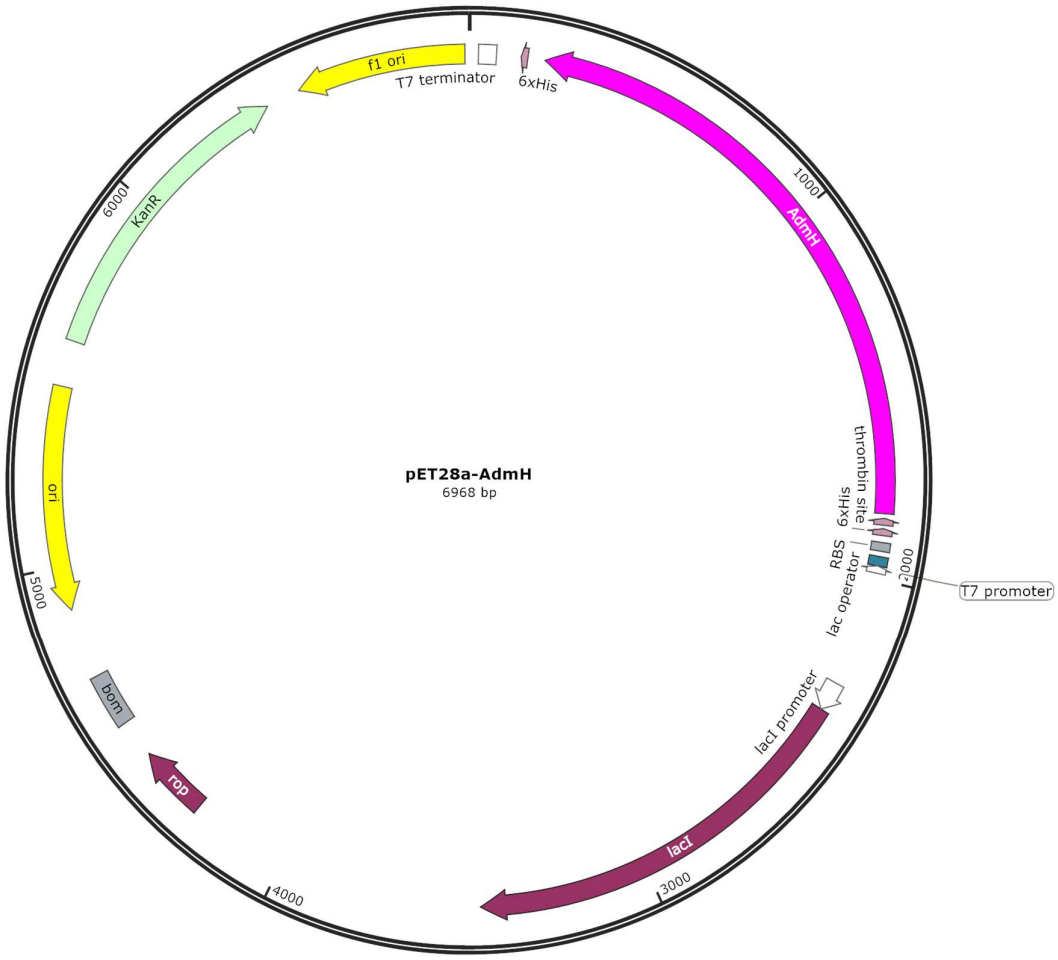
AmpR: gene encoding ampicillin resistance
Col E1 origin: origin of replication
ExpCas_PB2-T27: synthetic minimal promoter and terminator cassette for expression in *S. cerevisiae* [169,170]
upstream synthetic inter-homology flank 308a: upstream synthetic homologous flank of locus 308a (ARS308a) of *S. cerevisiae* chromosome III for genomic integration by homologous recombination
downstream synthetic inter-homology flank 308a: downstream synthetic homologous flank of locus 308a (ARS308a) of *S. cerevisiae* chromosome III for genomic integration by homologous recombination
Containing homologous flanks, targeting locus 308a (ARS308a) of *S. cerevisiae* chromosome III for genomic integration [156]
Containing unique restriction recognition sites MluI and BglII to functionally clone any gene of interest into the artificial minimal expression system. Contains unique restriction recognition site NotI to individually liberate the promoter and terminator cassette and any subsequently cloned gene of interest as an individual cassette equipped with a functional expression system.

pMA-RQ-PB2-crpA-T27



Equivalent to pMA-RQ-PB2-T27
crpA: encoding the first gene of cryptophycin biosynthesis in *Nostoc* sp. ATCC53789

pET28a-admH



Equivalent to pET28a (+)
AdmH: synthetic gene construct, encoding the phenylalanine 2,3-aminomutase of *Pantoea agglomerans* (gene purchased from ATG:biosynthetics GmbH) [217]

Table SI II. *S. cerevisiae* plasmids, vector maps and relevant specifications. Vector Maps were designed and annotated with SnapGene.

pCAS9-416 (p426_Cas9_gRNA-ARS416d) [156]

E. coli-S. cerevisiae shuttle plasmid
URA3 Promoter: *S. cerevisiae* promoter of URA3
URA3: selective marker encoding orotidine-5'-phosphate decarboxylase, required for uracil biosynthesis in *S. cerevisiae*
f1 ori: *E. coli* origin of replication
T7 promoter: *E. coli* T7 Promotor of RNA Polymerase
CYC1 terminator: *S. cerevisiae* terminator for Cas9
SV40 NLS: nuclear localization signal of SV40 (simian virus 40) large T antigen
CAS9: codon optimized Cas9 (Csn1) endonuclease of *Streptococcus pyogenes* Type II CRISPR/Cas system
ADH1 promoter: *S. cerevisiae* promoter for Cas9
gRNA scaffold: guide RNA scaffold for the *Streptococcus pyogenes* CRISPR/Cas9 system for locus ARS416/ARS1 of *S. cerevisiae* chromosome IV
T3 promoter: *E. coli* T3 promoter of RNA polymerase
lac operator: binding site of the lac-repressors regulating expression of downstream encoded genes
lac promoter: *E. coli* promoter for the lac operon
CAP binding site: CAP activator protein binding site of the lac operon
ori: *E. coli* origin of replication
AmpR: gene encoding ampicillin resistance in *E. coli*
AmpR promoter: Ampicillin resistance gene promoter
2µ ori: *S. cerevisiae* 2µ plasmid origin of replication (multicopy)

pCAS9-308 (p426_Cas9_gRNA-ARS308a) [156]

Equivalent to pCAS9-416 except for gRNA scaffold sequence
gRNA scaffold: guide RNA scaffold for the *Streptococcus pyogenes* CRISPR/Cas9 system for locus ARS308a of *S. cerevisiae* chromosome III

Table SI III. Arginine codon abundance and distribution of rare arginine codons in cryptophycin assembly line proteins. Notation: R, arginine.

Enzyme	Total no. of amino acids	Total no. of R codons	R codons in total amino acids [%]	Total no. of rarely used R codons	Rare arginine codons in yeast			Rare R codons in total R codons [%]
					CGC	CGG	CGA	
CrpA	2941	105	3.57	39	14	7	18	37.14
CrpB	3469	116	3.34	52	12	17	23	44.83
CrpC	1943	76	3.91	35	11	7	17	46.05
CrpD	3344	125	3.74	52	16	20	16	41.6
CrpE	450	22	4.89	10	1	2	7	45.45
CrpF	294	9	3.06	2	0	1	1	22.22
CrpG	113	3	2.65	0	0	0	0	0
CrpH	491	26	5.30	11	7	3	1	42.31
Total	**13045**	**482**	**3.7**	**201**	**61**	**57**	**83**	**41.7**

7.2 Supplementary figures

Figure SI I. pMA-RQ-based minimal expression systems for *S. cerevisiae*. The universal, minimal synthetic expression systems were constructed according to Curran *et al.* and Redden *et al.* [169,170]. The promoter PA3 was combined with the terminator T30, generating the expression system PA3-T30. PB2-T27 was generated by combining promoter PB2 and terminator sequence T27, while PC8-T3 was composed of promoter PC8 and terminator T3, respectively. The promoter and terminator regions are separated by a Cloning-Box, which is comprising unique restriction recognition sites, which are separated by a NotI restriction site. These unique sequences can be used to integrate any gene of interest (GOI) by restriction-ligation cloning. Notation: NotI, restriction recognition site for liberation of the total expression cassette construct including the cloned GOI; UAS, upstream activating sequence; Efficiency, efficiency element; Link1, 3-20 bp spacer sequence; Positioning, positioning element; Link2, 3-20 bp spacer sequence; PolyA, polyA-site. UAS, TATA-Box, efficiency element and PolyA-site are indicated in bold capital letters, while the core sequence and the positioning element are depicted as bold lower case letters. The Cloning-Box is indicated by a squared box. The spacer sequences and the link regions (Link1, Link2) are indicated by regular capital letters. Integrated upstream and downstream homology flanks, applicable to CRISPR/CAS applications are indicated by up./dn. homology flank, respectively. PC8-T3 contained homology flanks, directing the integration into locus 416d (ARS416/ARS1) on chromosome IV. PB2-T27 exhibited homology flanks, targeting the integration into locus 308a (ARS308a) on chromosome III, respectively.

Supplementary information

Supplementary figures

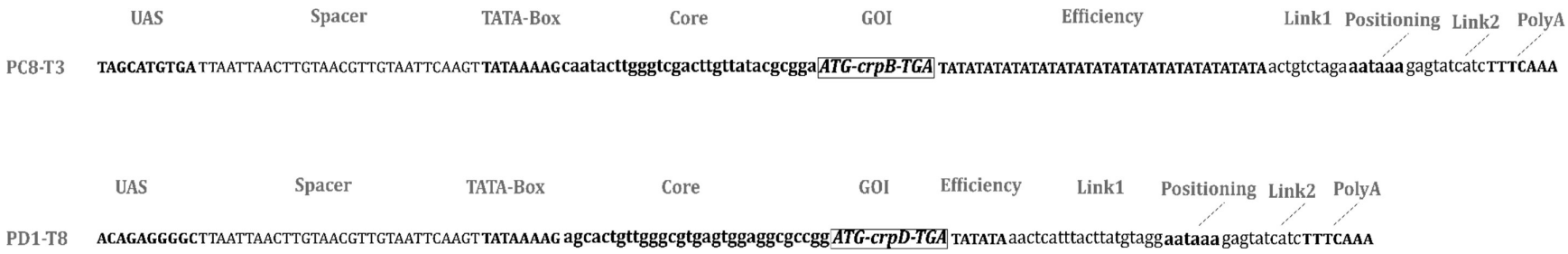


Figure SI II. pYPKpw-based minimal expression systems for *S. cerevisiae.* The minimal synthetic expression systems for *crpB* and *crpD* were constructed according to Curran *et al.* and Redden *et al.* [169,170]. The promoter PC8 was chosen for *crpB*, while PD1 was used to express *crpD*. The terminator T3 was used for *crpB*. Terminator T8 was chosen for *crpD*. Notation: UAS, upstream activating sequence; GOI, gene of interest, representing the total sequence of *crpB* and *crpD*, respectively; Efficiency, efficiency element; Link1, 3-20 bp spacer sequence; Positioning, positioning element; Link2, 3-20 bp spacer sequence; PolyA, polyA-site. UAS, TATA-Box, efficiency element and PolyA-site are indicated in bold capital letters, while the core sequence and the positioning element are depicted as bold regular letters. The GOI is indicated by a squared box. The spacer sequences and the link regions (Link1, Link2) are indicated by regular capital letters.

Supplementary information

Supplementary figures

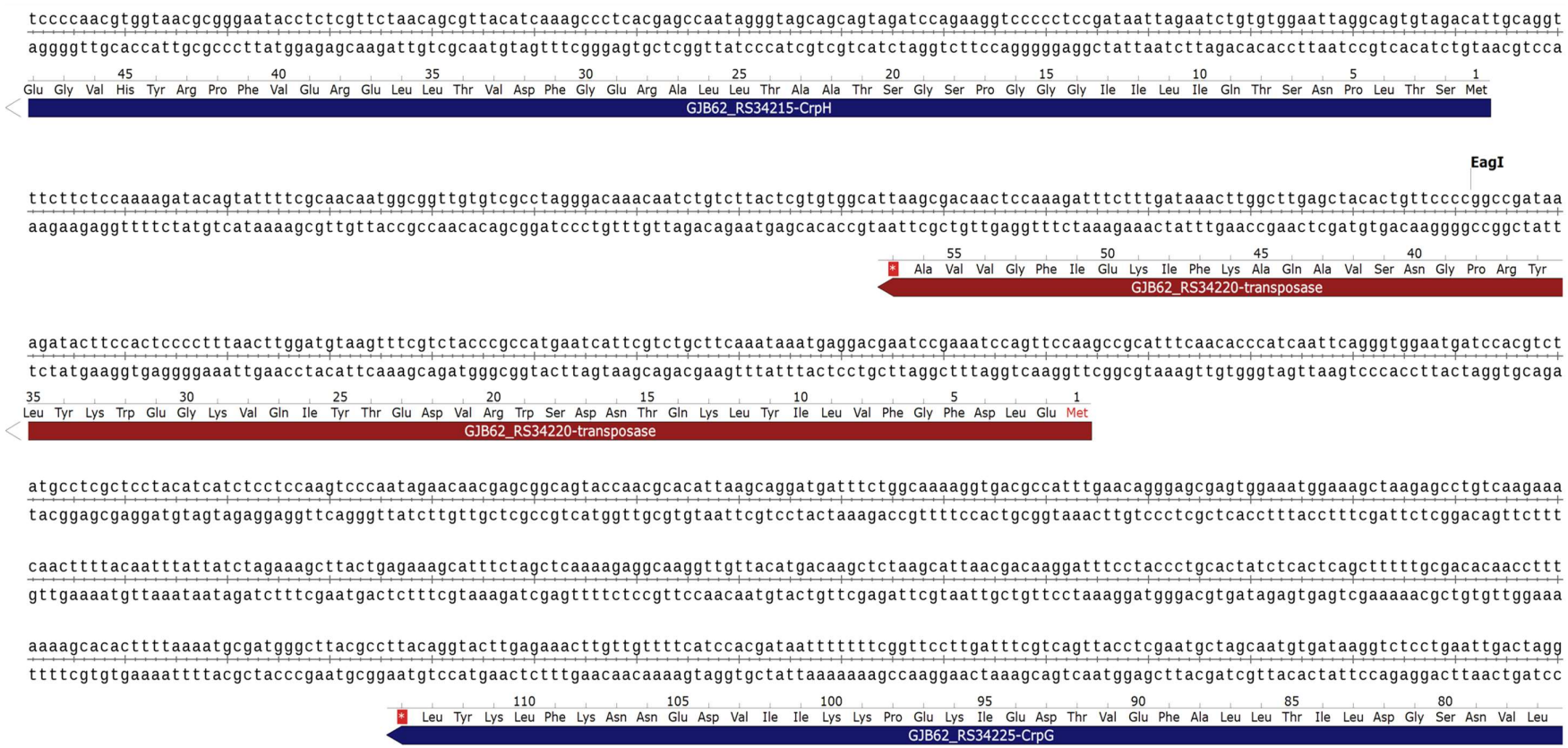


Figure SI III. Cryptophycin BGC-internal transposase. A cluster-internal transposase gene (red) was detected in the gap region between the two biosynthesis genes *crpG* and *crpH* (dark blue).

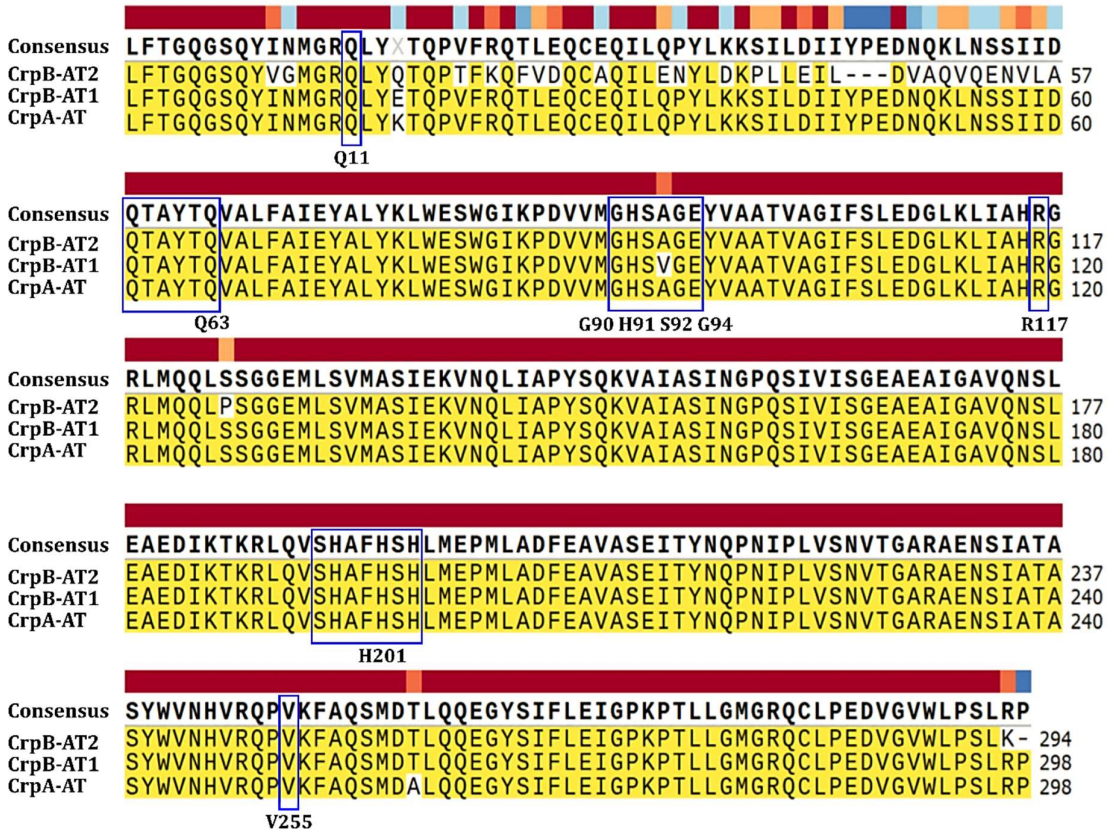


Figure SI IV. AT domain alignment. The protein sequences of the *crpA* (CrpA-AT)- and *crpB* (CrpB-AT1, CrpB-AT2)-encoded AT domains were compared by a manual alignment and examined for the characteristic active site motive of malonate-selective AT domains (i.e. Q11 Q63 G90 H91 S92 G94 R117 H201 V255, including the active site core motive G H S x G E with the catalytic S92) to evaluate their functional integrity (blue boxes). Identical sequences are highlighted in yellow. A consensus sequence is given, showing the overall similarities between the three different AT domains. The corresponding consensus sequence color code is highlighting identical amino acids in red. Amino acid deviations within the same group are highlighted in orange, while amino acids from different groups are shown in blue. The alignment was conducted with SnapGene.

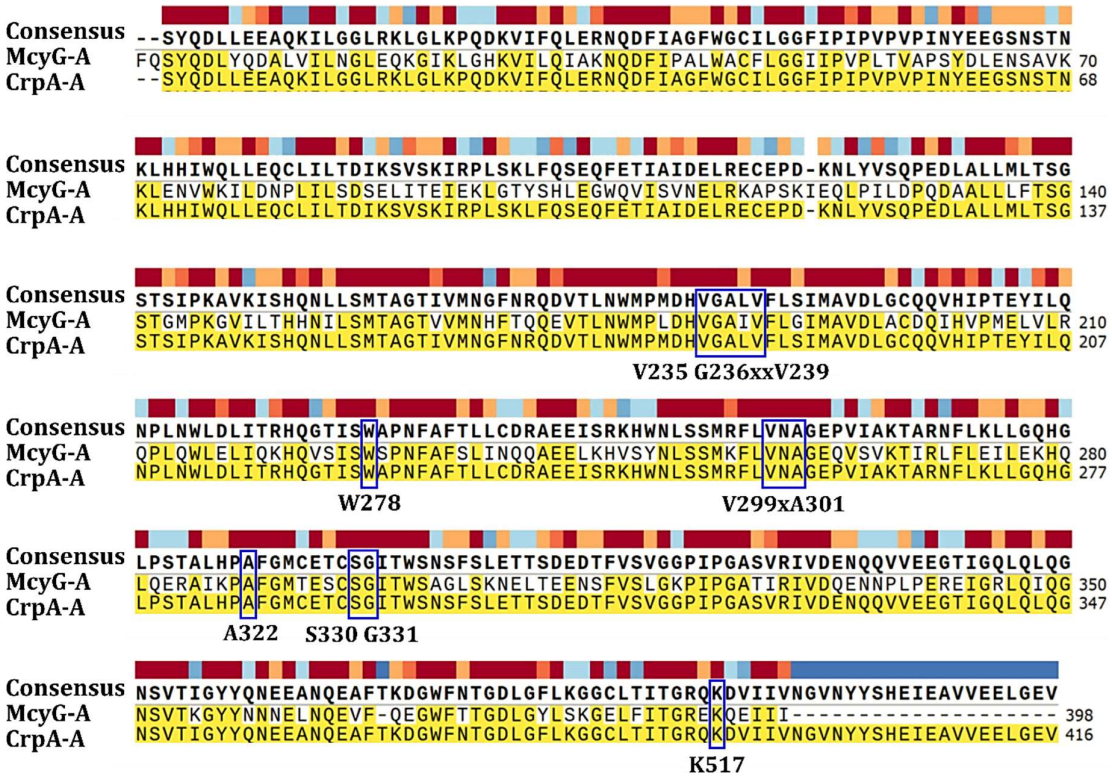


Figure SI V. A domain alignment. The protein sequences of the *crpA* (CrpA-A)- and *mcyG* (McyG-A)-encoded A domains were compared by a manual alignment and examined for the characteristic active site motives (i.e. V235, G236, V239, W278, V299, A301, A322, S330, G331, K517) to evaluate their functional integrity (blue boxes). Identical sequences are highlighted in yellow. A consensus sequence is given, showing the overall similarities between the two A domains. The corresponding consensus sequence color code is highlighting identical amino acids in red. Amino acid deviations within the same group are highlighted in orange, while amino acids from different groups are shown in blue. The alignment was conducted with SnapGene.

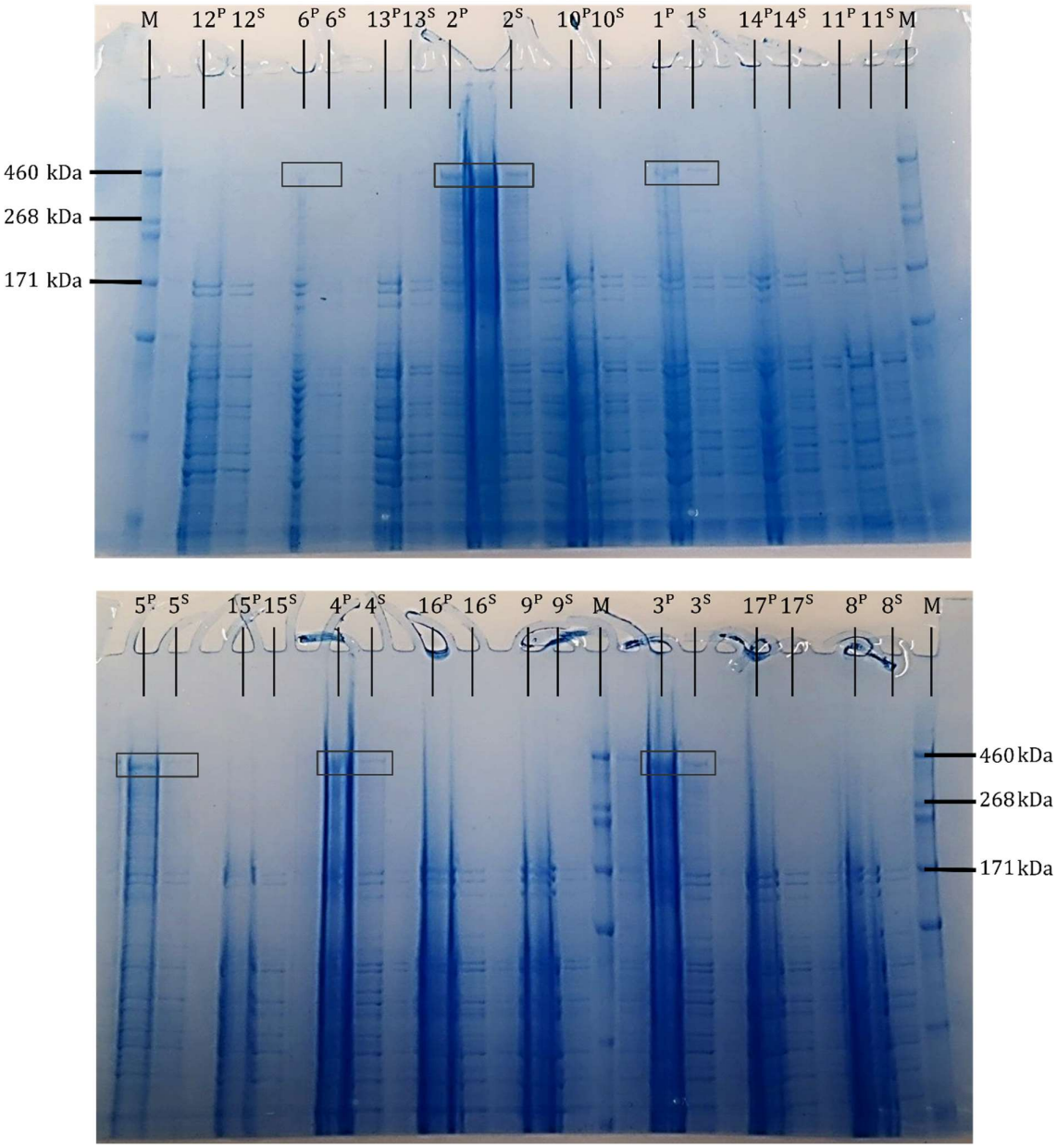


Figure SI VI. SDS-PAGE analysis of recombinant MBP-CrpA and SUMO-CrpA from diverse *E. coli* host strains. The SDS-PAGE analysis was performed for the cell lysate fractions (supernatant and pellet fraction) from five different host strains, expressing a combination of pACYC-sfp and pMal-crpA or pET28a-SUMO-crpA, respectively. Further, the control strain *E. coli* DH5α harboring pMal-crpA was analyzed for the formation of MBP-CrpA without the co-expression of Sfp. For comparison, the individual control strains were tested, likewise. Notation: M, LC5699-HiMark™ (Thermo Fisher Scientific); 12, *E. coli* ArcticExpress (DE3); 6, *E. coli* DH5α-MBP-crpA; 13, *E. coli* BL21 (DE3)-SUMO-crpA-sfp; 2, *E. coli* BL21 (DE3)-MBP-crpA-sfp; 10, *E. coli* BL21(DE3); 1, *E. coli* KRX-MBP-crpA-sfp; 14, *E. coli* KRX-SUMO-crpA-sfp; 11, *E. coli* KRX; 5, *E. coli* ArcticExpress (DE3)-MBP-crpA-sfp; 15, *E. coli* ArcticExpress (DE3)-SUMO-crpA-sfp; 4, *E. coli* C43(DE3)-MBP-crpA-sfp; 16, *E. coli* C43(DE3)-SUMO-crpA-sfp; 9, *E. coli* C43(DE3); 3, *E. coli* C41(DE3)-MBP-crpA-sfp; 17, *E. coli* C41(DE3)-SUMO-crpA-sfp; 8, *E. coli* C41(DE3). Superscript letters are referring to the pellet fraction (P) or supernatant fraction (S) of the analyzed cell lysate. The formation of the recombinant proteins SUMO-CrpA (337 kDa) and MBP-CrpA (366 kDa) is indicated by a box.

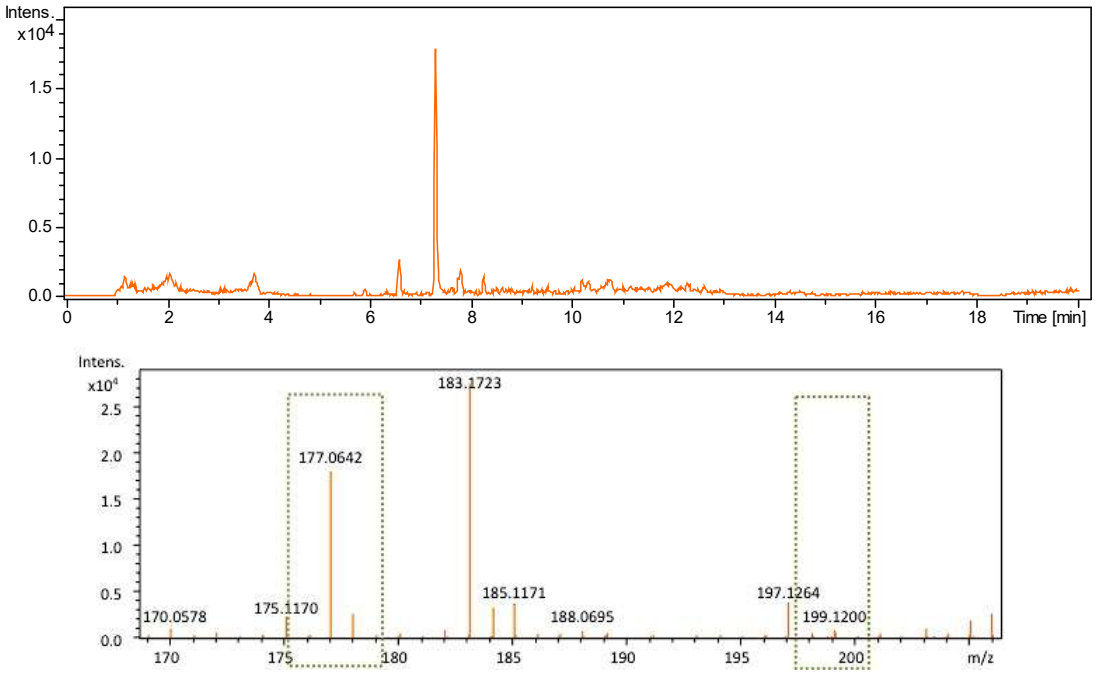


Figure SI VII. LC-MS Analysis of *E. coli* BL21 (DE3)-MBP-*crpA-sfp* expression strain. Individual extracted ion chromatogram (EIC) of the mutant strain (top) for *m/z* 177.0910 ± 0.05 $[M+H]^+$ at 7.3 min and its subjacent mass signals (bottom). The green boxes highlight the putative MPBA signal and its corresponding sodium adduct.

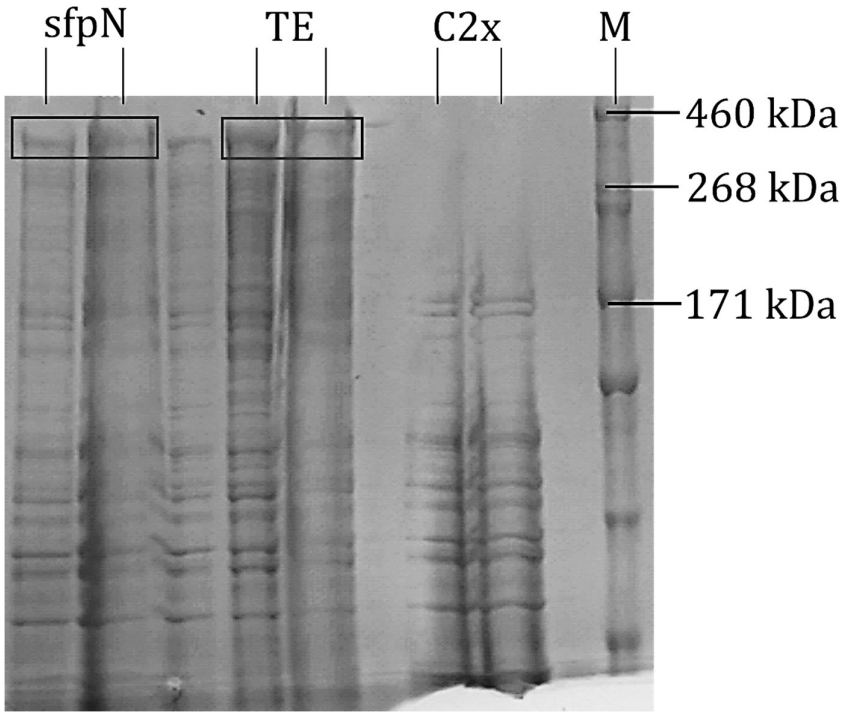


Figure SI VIII. SDS-PAGE analysis of *E. coli* BL21 (DE3)-MBP-*crpA-sfpN* [a] and *E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp*. The obtained supernatant fractions of the cell lysates were analyzed for the biosynthesis of the recombinant proteins MBP-CrpA (366 kDa) and MBP-CrpA$_{TE}$ (390 kDa). In comparison the strain *E. coli* BL21 (DE3)-*sfp*-C2x, which is incapable of CrpA formation was tested. Notation: sfpN, *E. coli* BL21 (DE3)-MBP-*crpA-sfpN* [a]; TE, *E. coli* BL21 (DE3)-MBP-*crpA*$_{TE}$-*sfp*; C2x, *E. coli* BL21 (DE3)-*sfp*-C2x; M, LC5699-HiMark™ (Thermo Fisher Scientific).

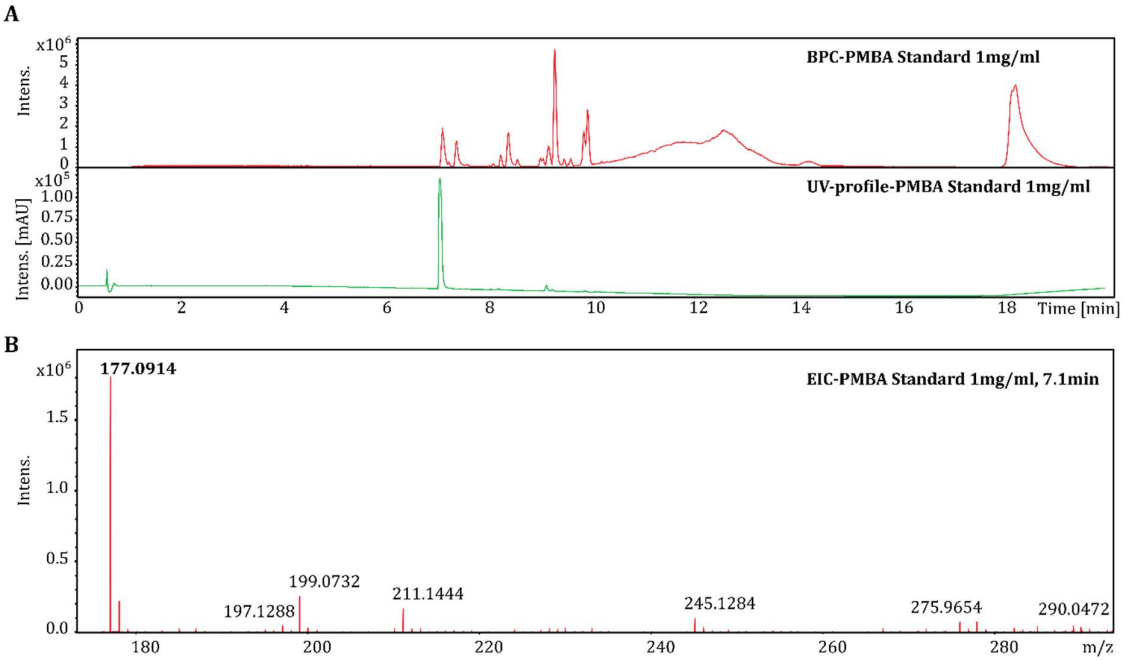


Figure SI IX. LC-MS analysis of MPBA standard. The acquired MPBA standard was subjected to LC-MS analysis, recording its base peak chromatogram (BPC) and the corresponding UV-profile (A). The extracted ion chromatogram (EIC) revealed a subjacent mass signal of *m/z* 177.0910 ± 0.0004 $[M+H]^+$ at 7.1 min. (B), matching to the expected mass of *m/z* 177.0910 $[M+H]^+$.

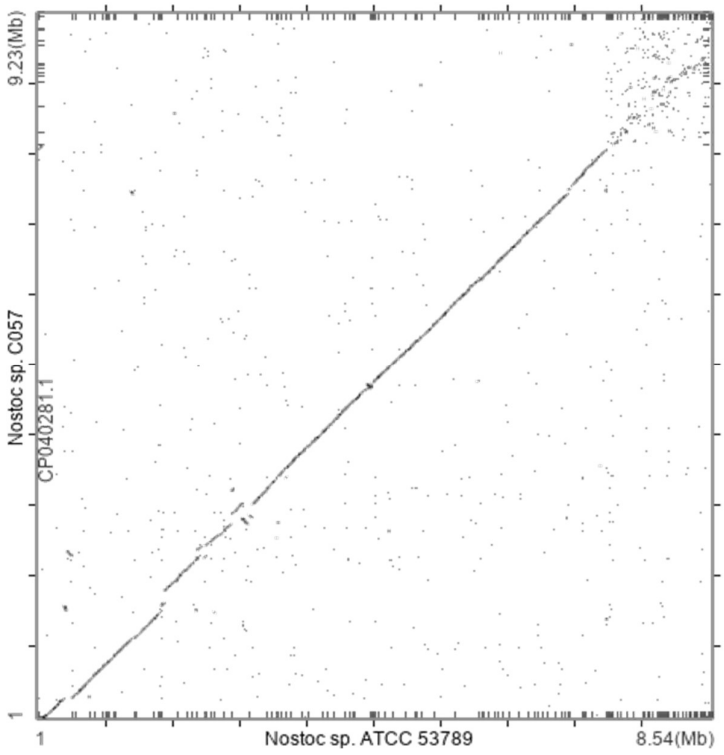


Figure SI X. Genome alignment of *Nostoc* sp. ATCC 53789 and *Nostoc* sp. C057. The *in silico* alignment of both NCBI-deposited genome sequences shows the symmetry between the two strains on a genomic basis. The symmetric identity of both cyanobacteria is 76%. The alignment was conducted by the NCBI genome neighbor alignment report [262]

8 Abbreviations

µg	microgram
3-PPA	3-phenylpropionic acid
5-FOA	5-fluoroorotic acid
A domain	adenylation domain
ACN	acetonitrile
ACP	acyl carrier protein
AdmH	aminomutase of *Pantoea agglomerans*
admH	gene encoding AdmH of *Pantoea agglomerans*
Amp	ampicillin
Amp^R	ampicillin resistance
antiSMASH	antibiotic and secondary metabolite analysis shell
ARO8	bifunctional aminotransferase/transaminase of *S. cerevisiae*
AT	acyl transferase
ATCC	American Type Culture Collection
BG-11	ATCC culture medium 819 for cyanobacteria
BGC(s)	biosynthetic gene cluster(s)
BLASTn	basic local alignment search tool for nucleotides
BLASTp	basic local alignment search tool for proteins
BLASTx	basic local alignment search tool for translated nucleotides
BMBF	German Federal Ministry of Education and Research - BMBF (Ger. Bundesministeriums für Bildung und Forschung)
BNA2	tryptophan 2,3-dioxygenase
bp	base pair
BPC	base peak chromatogram
C	condensation domain
$CaCl_2$	calcium chloride
Cm	chloramphenicol
Cm^R	chloramphenicol resistance
CM(T)	C-methyltransferase, methylating carbon
CO_2	Carbon dioxide
CoA	coenzyme A
COM	communication-mediating domain
crp	cryptophycin operon
CrpA-A	A domain of the cryptophycin biosynthetic enzyme CrpA
CrpA-AT	AT domain of the cryptophycin biosynthetic enzyme CrpA
CrpA-H	biosynthetic enzymes A-H of the cryptophycin assembly line
crpA-H	biosynthesis genes *a-h* of the cryptophycin operon
CrpB-AT1	first AT domain of the cryptophycin biosynthetic enzyme CrpB
CrpB-AT2	second AT domain of the cryptophycin biosynthetic enzyme CrpB
Cy domain	cyclization domain
CYC1	endogenous terminator for *S. cerevisiae*
Da	Dalton
DCM	dichloromethane
DCW	dry cell weight
ddH_2O	ultrapure double deionized water
DEBS	erythromycin producing 6-deoxyerythronolide B synthase

DH	dehydrogenase
dH_2O	deionized water
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E domain	epimerization domain
EIC	extracted ion chromatogram
ESI	electrospray ionization
ER	enoylreductase
FCCS	frozen competent cell solution
g/l	gram per liter
GOI	gene of interest
His6-AdmH	AdmH fused to a hexahistidine-tag
His6-tag	hexahistidine-tag
HPLC	high pressure liquid chromatography
HR	homologous recombination
HygR	hygromycin resistance
Kan	kanamycin
Kan^R (G418)	kanamycin (G418) resistance
kb	kilo base pair
kDa	kilo Dalton
KR	ketoreductase
KS	ketoacyl synthase
LB	lysogeny broth medium
LC/MS	liquid chromatography-mass spectrometry
LiAc	lithiumacetat dihydrate
LTR-recombination	long terminal repeat-guided recombination
M	mole per liter
m/z	charge ratio
Mb	mega base pair
MBP(-tag)	maltose binding protein(-tag)
MBP-CrpA	CrpA fused to a maltose binding protein
mcy	microcystin open reading frame
McyG	biosynthetic enzyme G of microcystin assembly line
McyG-A	A domain of the microcystin biosynthetic enzyme McyG
MeOH	methanol
methanol-d4	deuterierium methanol
mg/l	milligram per liter
mg/ml	milligram per milliliter
Mg^{2+}	manganese
min	minutes
mM	millimole per liter
MT (O-, N-, C-linked)	methyltransferase, methylating oxygen, nitrogen or carbon
N_2	nitrogen
NaCl	sodium chloride
NAD(P)	nicotinamide adenine dinucleotide (phosphate)
$NaNO_3$	sodium-nitrate
NCBI	National Center for Biotechnology Information
ncRNA	non-coding RNA

NMR	Nuclear magnetic resonance spectroscopy
NRPS	nonribosomal peptide synthetase
NTA	nitrilotriacetic acid
NTS	non-transcribed spacer
OMT	O-methyltransferase, methylating oxygen
orf(*s*)	open reading frame(s)
Ox domain	oxidase domain
PAM	peptidylglycine α-amidating monooxygenase
PBA	4-phenyl-3-butenoic acid
PCP	peptidyl carrier protein
PGal1/Gal1	galactose inducible promoter for *S. cerevisiae*
PKS	polyketide synthase
PKS-NRPS	hybrid pathway containing a polyketide synthase and peptide synthetase
MPBA	2-Methyl-4-Phenyl-3-butenoic acid
PPTases	phosphopantetheinyl transferases
PRODIGY	Process-Directed Drug Generation in Yeast
Red	reductase domain
RiPP(s)	ribosomally synthesized posttranslationally modified peptide(s)
RNAse	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT	room temperature
S7P	sedoheptulose 7-phosphate
SAM	S-adenosylmethionine
scy	scytonemin-*orf*
Sfp	Sfp-type phosphopantetheinyl transferases of *Bacillus subtilis*
sfp	gene encoding the Sfp-type phosphopantetheinyl transferase of *Bacillus subtilis*
sgRNA	single-guide RNA
SNAC-ester(s)	N-acetylcysteamine thioester(s)
SOB	Hanahan′s broth/ medium
SrfAC	biosynthetic enzyme of surfactin assembly line
SUMO-tag	small ubiquitin-related modifier-tag
TDH3	strong constitutive promoter for *S. cerevisiae*
TE	thioesterase
TE_{crp}	TE domain of CrpD in cryptophycin assembly line
$TE_{eryAIII}$	TE domain of DEBS AIII biosynthesis enzyme
TFA	Trifluoroacetic acid
ThDP	thiamin diphosphate
TKL	triketide lactone
tRNA	transfer RNA
TSS	transcription start site
TycA	biosynthetic enzyme of tyrocidine assembly line
UAS	5’ upstream activating sequence
v/v	volume per volume fraction
w/v	mass per volume fraction
YNB	Yeast Nitrogen Base medium
YPD	yeast extract peptone dextrose medium
ZeoR	zeocin resistance

9 Index of Tables

10 Index of Figures

11 Index of Supplementary Tables

12 Index of Supplementary Figures

13 Acknowledgements

My cordial and sincere thanks go to Prof. Dr. Markus Nett, who gave me the opportunity to perform my doctoral studies in the Laboratory of Technical Biology at TU Dortmund University. His professional advice, kind and supportive guidance and enthusiasm for science have always been very motivating to me. Thank you for always having an open door.

Further I would like to express my cordial thanks to the doctoral commission and the examination committee for reviewing my dissertation. In particular I would like to thank Prof. Nadine Ziemert for being my second reviewer and Prof. Oliver Kayser and PD Dr. Ing. Christoph Held for being part of the examination committee.

In addition, I would like to cordially thank the PRODIGY consortium: Prof. Lars Blank, Prof. Oliver Kayser and Prof. Stephan Lütz, Dr. Raffaella Di Lucrezia, Dr. Peter Nussbaumer, Dr. Katrin Rosenthal, Mariam Dianat and Oliver Schiwy for the great collaboration and discussions. Furthermore, I would like to thank the German Federal Ministry of Education and Research (BMBF) for providing financial support.

My special gratitude goes to Dr. Angela Sester, Sebastian Kruth, Katharina Kuhr, Jenny Schwarz, Andrea Steinmann and Lea Winand, for the great support, advice, motivating discussions and rocking the lab together. Special thanks also go to Kristine Hemmer, Jörg Fischer and Sabine Vogt. In addition, I would like to thank the entire TBL group, my office crew and BCI colleagues for the great support.

Last but not least, I would like to express my deepest and heartfelt gratitude to Flo, my family and friends. I could not have undertaken this adventure without you.