

Synthetic biotechnology to study and engineer natural product biosynthesis in actinomycetes

Dissertation

Zur Erlangung des Grades

Des Doktors der Naturwissenschaften

Der Naturwissenschaftlich-Technischen Fakultät III

Chemie, Pharmazie, Bio- und Werkstoffwissenschaften

Der Universität des Saarlandes

Von

Liujie Huo

Saarbrücken

2014

Tag des Kolloquiums: 23. Juni. 2014

Dekan: Prof. Dr. Volkhard Helms

Berichterstatter: Prof. Dr. Rolf Müller

Prof. Dr. Marc Stadler

Vorsitz: Prof. Dr. Rolf. W. Hartmann

Akad. Mitarbeiter: Dr. Andriy Luzhetskyy

To the losses and the gains.

Acknowledgement

First and foremost I would like to express my sincere gratitude to Prof. Dr. Rolf Müller for giving me the opportunity to work in the remarkable research environment and providing me the intriguing research projects in the laboratory of the Department of Pharmaceutical Biotechnology at Saarland University. Thanks to his inspirational support during these years, I have learned so much not only from his scientific ability and attitude but also his magnetic personality which will benefit my whole research career and my life. I also own my thanks to his thorough review of my thesis.

I would like to thank Prof. Dr. Marc Stadler for his enthusiastic help in the collaborative projects over the years and also for his being the second reviewer of this thesis.

I owe my special gratitude to Dr. Silke C. Wenzel. She is a wonderful researcher and mentor, who practically co-advised me to go through the tough research process, especially at the beginning of the Ph.D study. Her interesting scientific ideas and experimental designs always enlighten the research over these years. She also helped a lot for writing of the thesis and proof-reading.

At the same time I would like to thank Prof. Dr. Uli Kazmaier and Philipp Barbie from Department of Organic Chemistry at Saarland University. Over the past years, we cooperated for many projects, and had very pleasant and successful team works. Besides, I also want to thank Prof. Kazmaier for allowing me to join the Graduate school on Natural product research and benefit me a lot out of the lectures, scientific seminars and conferences.

I want to especially thank Prof. Dr. Youming Zhang from State Key Laboratory of Microbial Technology at Shandong University in China. He provided me so many advices and guidance for both my research and career.

I would like to express thanks to all the members of the research group who have helped and taught me during this period and their contributions to the good working atmosphere. Special thanks would go to the analytic team in our working group including Dr. Daniel Krug, Thomas Hoffmann, Eva Luxenburger, and Michael Hoffmann for their help and advices in LC-MS measurement and analysis. I would also like to thank Dr. Kevin Sours for proof-

Acknowledgement

reading of my thesis. I owe my special thanks to my colleague and friend Dr. Xiaoying Bian for his useful discussions advices over the years. I also want to thank my colleagues Qiang Tu, Anja Shwarz, Dr. Chengzhang Fu, Fu Yan for all their support and friendship over the years.

Last but not least, my thanks would also go to my beloved parents for their boundless love and financial support over all these years. I owe my sincere appreciation to my dear wife Yuan Li for her patience, support and love to me. Finally I want to appreciate all my friends for their friendship and love to me.

Without any of you mentioned above, this thesis would never have been possible.

Liujie Huo

23rd, April, 2014

Veröffentlichungen der Dissertation

Teilergebnisse dieser Arbeit wurden mit Genehmigung der Naturwissenschaftlich-Technischen Fakultät III, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen:

N. Quade⁺, **L. Huo**⁺, S. Rachid, D. W. Heinz, R. Müller, Unusual carbon fixation giving rise to diverse polyketide extender units. *Nature Chemical Biology*, 2012, 1, 117-124.

L. Huo⁺, S. Rachid⁺, M. Stadler, Silke C. Wenzel, R. Müller, Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chemistry&Biology*, 2012, 19, 1278-1287.

P. Barbie, **L. Huo**, R. Müller, U. Kazmaier, Stereoselective synthesis of deuterium-labeled (2S)-cyclohexenyl alanines, biosynthetic intermediates of cinnabaramide, *Organic Letters*, 2012, 14(23), 6064.

S. Rachid, **L. Huo**, J. Herrmann, M. Stadler, B. Köpcke, J. Bitzer, R. Müller, Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. *ChemBioChem*, 2011, 12, 922-931.

(⁺: Authors contributed equally to the work)

Tagungsbeiträge:

Vorträge:

Huo, L., Quade, N., Stadler, M., Heinz, D. W., Müller, R. Mining the cinnabaramide biosynthesis pathway and characterization an unusual carbon fixation enzyme catalyzing formation of unusual extender units for polyketides. Meeting: **2011 annual VAAM meeting "Biology of Bacteria Producing Natural Compounds"**. Bonn, Germany, September, 2011.

Huo, L., Stadler, M., Wenzel, S. C., Müller, R. Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. Meeting: **VAAM International Workshop 2012 on the "Biology and Chemistry of Antibiotic-Producing Bacteria and Fungi"**. Braunschweig, Germany, September, 2012.

Posters:

Huo, L., Stadler, M., Wenzel, S. C., Müller, R. Characterization and Engineering of Cinnabaramide Biosynthesis. Meeting: **1st European Conference on Natural Products**. Frankfurt am Main, Germany, September, 2013.

Huo, L., Quade, N., Rachid, S., Herrmann, J., Stadler, M., Köpcke, B., Bitzer, J., Heinz, D. W., Müller, R. Mining the cinnabaramide biosynthetic pathway and biochemical and structural characterization of involved octenoyl-CoA carboxylase/reductase. Meeting: **ESF-EMBO Symposium: Synthetic Biology of Antibiotic Production**. Sant Feliu de Guixols, Spain, October, 2011.

Huo, L., Rachid, S., Herrmann, J., Stadler, M., Köpcke, B., Bitzer, J., Müller, R. Mining the cinnabaramide biosynthetic pathway and characterization of the involved octenoyl-CoA reductase/carboxylase (CinF). Meeting: **Frontiers in Medicinal Chemistry**. Saarbrücken, Germany, March, 2011.

Huo, L., Rachid, S., Stadler, M., Wenzel, S. C., Müller, R. Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. June, 2012. Meeting: **2nd International HIPS Symposium**. Saarbrücken, Germany, September, 2012.

Huo, L., Rachid, S., Stadler, M., Wenzel, S. C., Müller, R. Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. July, 2012. Meeting: **Sommer-Symposium des Naturstoffzentrums Rheinland-Pfalz**. Mainz, Germany, September, 2012.

Zusammenfassung

Bakterielle Naturstoffe wie von den gut studierten Aktinomyceten weisen eine hohe strukturelle Diversität und interessante Bioaktivitäten auf und stellen daher attraktive Zielstrukturen für Forschung und klinische Anwendung dar.

Die vorliegende Arbeit beschäftigt sich mit der Identifizierung, Aufklärung und Manipulation der Biosynthesewege zu Cinnabaramiden aus *Streptomyces* sp. JS360 und Bottromycinen aus *Streptomyces* sp. BC16019. Über Synthetische Biotechnologie wurde die Produktion verbessert und neue Derivate dieser Naturstoffe erzeugt.

Mittels Vorläufer-dirigierter Biosynthese und Mutasynthese wurden neue Cinnabaramid-Derivate mit interessanten Bioaktivitäten hergestellt. Für die Bottromycine wurde eine heterologe Produktionsplattform etabliert, die Produktionsausbeuten optimiert und nach rationaler Modifikation des Biosyntheseweges neue Bottromycin-Analoga identifiziert.

Zudem wurde ein Enzym, welches entscheidend an der Bildung einer ungewöhnlichen Polyketid-Verlängerungseinheit für die Cinnabaramid-Biosynthese beteiligt ist, detailliert biochemisch und strukturell charakterisiert. Über Mutationen in der Bindetasche gelang es, die Substratspezifität des Enzyms zu verändern, welches ausgehend von Enolat-Derivaten über reduktive Carboxylierung Dicarbonsäuren generiert. Die erhaltenen Resultate lieferten tiefere Einblicke in die Biochemie dieser Enzymfamilie und stellen eine Grundlage zur Erzeugung veränderter Polyketide mit verbesserten pharmazeutischen Eigenschaften dar.

Abstract

Bacterial natural products like from the well-studied actinomycetes display diverse structural features and interesting bioactivities making these secondary metabolites attractive targets for scientific research and clinical application.

The present thesis deals with the identification, elucidation and engineering of biosynthetic pathways of the natural products cinnabaramides from *Streptomyces* sp. JS360 and bottromycins from *Streptomyces* sp. BC16019. Synthetic biotechnology tools were employed to improve production yields and to produce novel derivatives via biosynthetic engineering.

Precursor-directed biosynthesis and mutasynthesis approaches allowed the generation of novel cinnabaramide derivatives exhibiting intriguing bioactivities. For bottromycins, a heterologous production platform was established, optimized for higher production yields and novel analogues were identified after rational pathway modification.

Moreover, detailed biochemical and structural characterization of a key enzyme involved in the formation of an unusual polyketide extender unit during cinnabaramide biosynthesis was performed. By active site mutations it was possible to change the substrate specificity of the enzyme, which generates dicarboxylic acids by reductive carboxylation of enolate derivatives. The obtained results provided deeper insights into the biochemistry of this enzyme family and set the stage for generation of deliberately altered polyketides with improved pharmaceutical properties.

Table of Contents

A. Introduction

1. Natural products in drug discovery	1
2. Streptomyces, proficient producers of natural products	2
2.1 Ecology of Streptomyces	2
2.2 Streptomyces as key players in microbial natural product research	3
2.3 Alternative natural product producers	3
3. Biosynthetic logic of microbial natural products	4
3.1 Biochemistry of nonribosomal peptide synthetases	6
3.2 Biochemistry of polyketide synthases	7
3.3 PKS/NRPS or NRPS/PKS hybrids	8
3.4 Biochemistry of ribosomally synthesized and post-translationally modified peptides	9
4. Ways to produce and manipulate natural products	10
4.1. Heterologous expression: General aspects	10
4.2. Genetic engineering of expression constructs	11
5. Generation of “unnatural” natural products via synthetic biotechnology	13
6. Outline of the dissertation	15

B. Publications

I. Mining the cinnabaramide Biosynthetic pathway to Generate Novel Proteasome Inhibitors	18
II. Unusual carbon fixation gives rise to diverse polyketide extender unit	19
III. Stereoselective Synthesis of Deuterium-Labeled (2S)-Cyclohexenyl Alanines, Biosynthetic Intermediates of Cinnabaramides	20
IV. Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis	21

C. Final Discussion

1. General scope of this work	22
--------------------------------------	-----------

Table of contents

2. Cinnabaramides biosynthesis in <i>Streptomyces</i> sp. JS360	22
2.1 Cinnabaramides biosynthesis is directed by a hybrid type I PKS/NRPS incorporating unusual building blocks	22
2.2 Biosynthetic mechanism of the hexylmalonyl-CoA pathway	29
2.3 Creation of molecular diversity based on the cinnabaramide scaffold via synthetic biotechnology	32
3. Bottromycin biosynthesis in <i>Streptomyces</i> sp. BC16019	35
3.1 Bottromycin biosynthesis: Heavily modified peptides exhibiting unusual biosynthetic features	35
3.2 Engineering of the bottromycin biosynthetic pathway in a heterologous host	38
D. References	42
E. Appendix	

A. Introduction

1. Natural products in drug discovery

Over a thousand years one fact has been recognized and utilized by humans, nature produces an enormous number of metabolites with an amazing variety of structural diversity and bioactivities. The first insight into this nature's creativity occurred more than 200 years ago when Friedrich Sertürner isolated morphine from *Papaver somniferum*,^[1] which is regarded as the initiation of modern natural product research. Later on, inspired by the discovery of the β -lactam antibiotic penicillin from *Penicillium notatum* by A. Fleming, H. W. Florey and R. B. Chian,^[2] pharmaceutical research underwent a "golden era" including a massive screening of microorganisms resulting in discoveries of plenty of new antibiotics, many of which are still used in clinic today.^[3] To date, approximately 40% of drugs in clinical use are either natural products or derivatives thereof.^[4] They are broadly applied either in clinics, such as antibiotics, antitumor and immunosuppressive agents, or in agriculture and veterinary applications, such as growth promoters, insecticides, herbicides and antiparasitic agents (**Figure A1**).^[5-8]

Accompanying the diverse biological activities of natural products is tremendous structural diversity which cannot be obtained from alternative sources.^[9] In order to deal with the increasing development of the multi-drug resistant organisms, the lack of effective therapy for many severe diseases and the side effects caused by current medicines, searching for novel natural products is not only necessary but urgent.

Compared to other origins of natural products, such as plants or insects, microorganisms are predicted to be the most promising source of novel natural products for two reasons: 1) many past evidences have proven that microorganisms, especially Actinomycetes, possess a great ability to produce pharmaceutically relevant metabolites; 2) in the prokaryotic world only a very small percentage of microorganisms have been explored for production of natural products and consequently the potential to discover novel natural compounds is very promising.^[10] Thus, further exploration of microorganisms in order to identify novel bioactive compounds has excellent prospects.

A. Introduction

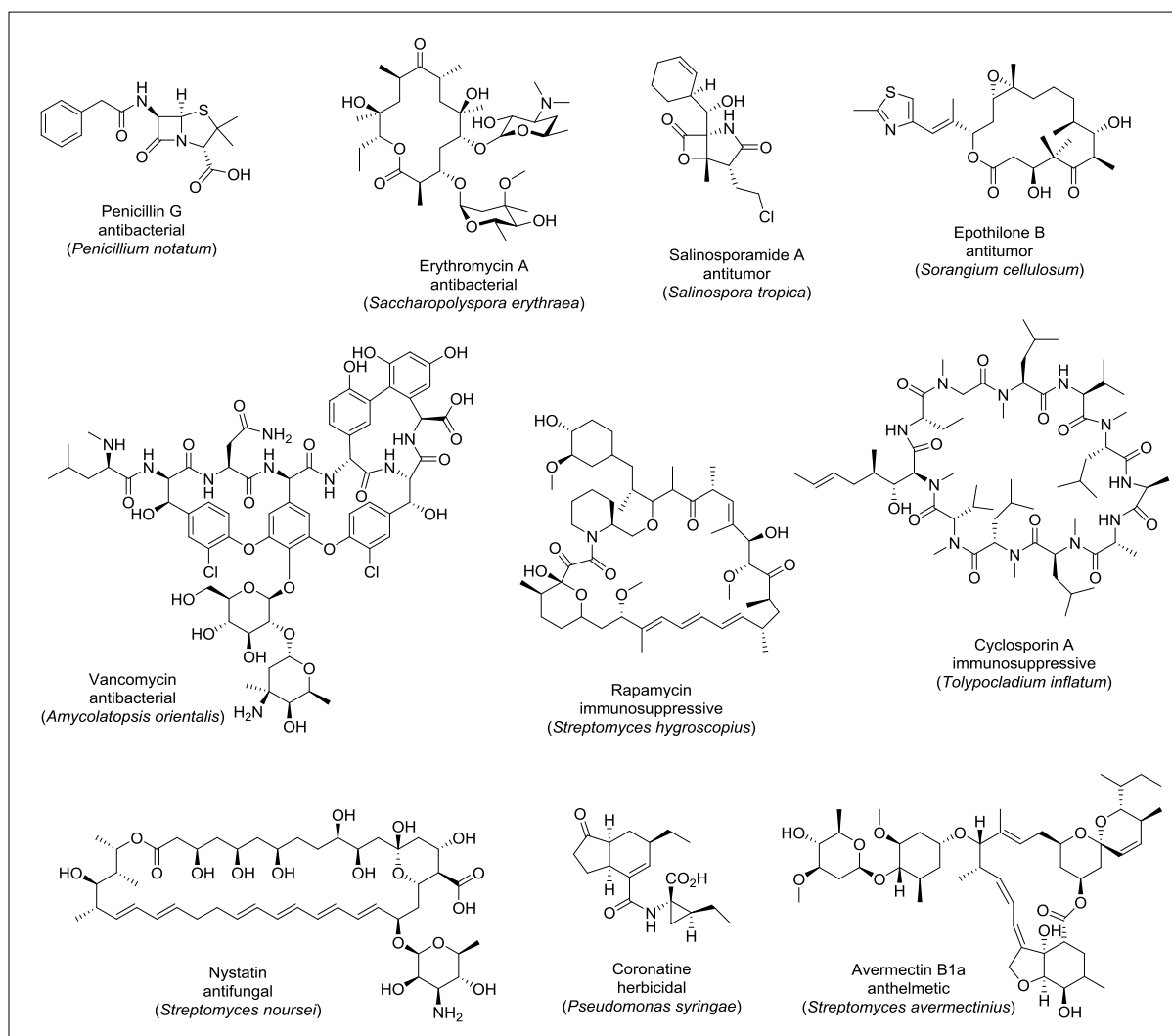


Figure A1. Selected examples of natural products derived from microbes used in the clinic/clinical trials, agriculture, and veterinary applications. The structure, name, original producer and type of application are given.

2. Streptomycetes, proficient producers of natural products

2.1 Ecology of Streptomycetes

Streptomycetes, which belong to the bacterial order Actinomycetales, are ubiquitous in nature. They are soil-dwelling and able to colonize new areas through growth as a vegetative hyphal mass which can differentiate into spores that assist in spread and persistence.^[11] Most of the time the relatively high numbers of Streptomycetes exist in soil as inactive spores, which are a semi-dormant stage in the life cycle and can survive for long periods.^[12] Because the spores can impart resistance to low nutrient and poor water availability, *Streptomyces* cultures can even be recovered from 70 year old soil samples. As one of the common characteristics of Actinomycetales, Streptomycetes are able to produce many extracellular enzymes (cellulases, xylanases, amylases, maltases, etc.) in soil, which can decompose various complex mixtures of polymers from dead plants, animals and fungi.^[13;14] Through this they play a very important role in biodegradation by recycling of nutrients associated with recalcitrant polymers.^[15]

2.2 Streptomycetes as key players in microbial natural products research

For over 60 years Streptomycetes have been well known as one of the most important producers of pharmaceutically relevant bioactive metabolites including antibiotics, antitumor agents, immunosuppressants, antiparasitic agents, herbicides and enzyme-inhibiting agents.^[16] Since the groundbreaking discovery of streptomycin as the first Actinomycete antibiotic in 1944,^[17] more than two-thirds of naturally derived antibiotics are produced by Actinomycetes, most of which belong to *Streptomyces* species. By 1995, 55% of ca.12000 known antibiotics originated from Streptomycetes and another 11% from other Actinomycetes.^[18] Amongst them there are a number of commercially important drugs (**Figure A2**) including the antibiotics erythromycin A, produced by *Saccharopolyspora erythraea*, vancomycin, produced by *Amycolatopsis orientalis*, tetracycline, produced by *Streptomyces aureofaciens* and anticancer agent doxorubicin, produced by *Streptomyces peucetius*.^[19] Due to the very successful history of *Streptomyces* spp. as an exceptional resource of pharmaceutically relevant bioactive metabolites, continued investigation of novel bioactive compounds with new chemical structures and/or biological activities from Streptomycetes is definitely still highly interesting to both natural product researchers and the pharmaceutical industry.

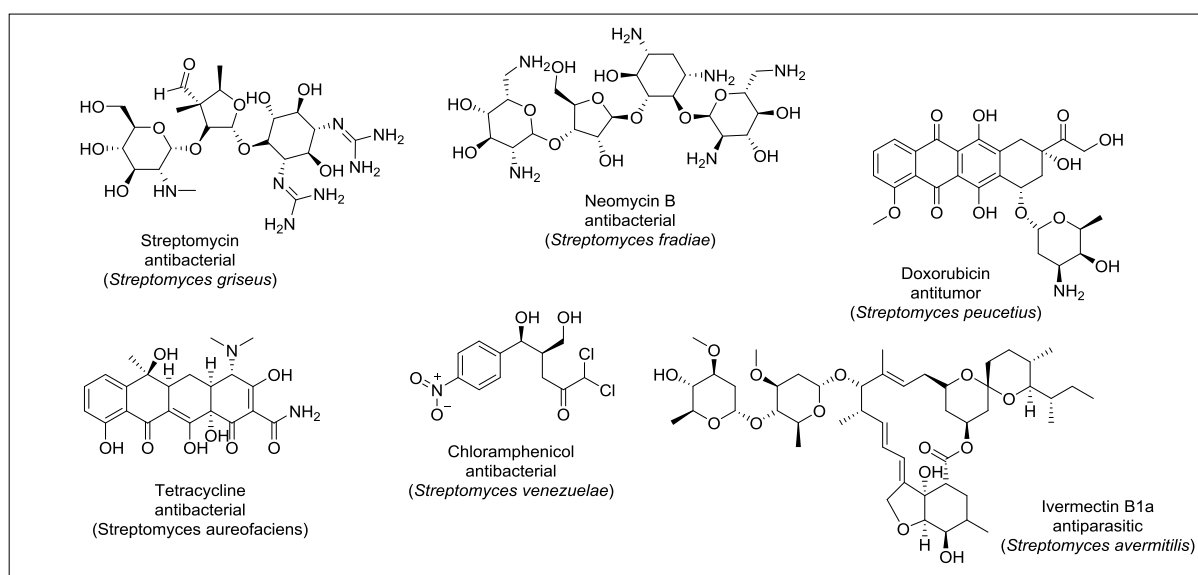


Figure A2. Drugs originating from Streptomycetes. The structure, name, original producer and type of application are given.

2.3 Alternative natural product producers

On the other hand, because of the relatively high probability of hitting a known compound originating from the resources having close phylogenetic relationships or living in similar environmental conditions, it is assumed that the novel producer classes, such as marine resources and other not intensively explored microbes, could correlate to the potentials to chemical novelty. Thus, intensive efforts have also been made on discovery of bioactive metabolites originating from other microorganisms in the last years. These studies mainly deal with marine bacteria (e.g. marine Actinomycetes), cyanobacteria, myxobacteria, and pathogenic bacteria (e.g. Pseudomonads). Among

them, myxobacteria showed especially promising potential as proficient producers of novel bioactive compounds.^[20] The most prominent myxobacterial secondary metabolites by far are the cytotoxic epothilones from *Sorangium* species.^[21] In 2007, a semisynthetic epothilone B analogue ixabepilone (Ixempra[®]) was approved by the American Food and Drug Administration (FDA) for the treatment of breast cancer.^[22]

3. Biosynthetic logic of microbial natural products

Up to present, natural products can be mostly classified into five groups based on their structural and biosynthetic commonality: terpenoids, alkaloids, polyketides (PKs), nonribosomal peptides (NRPs) and ribosomally synthesized and post-translationally modified peptides (RiPPs). Because the latter three groups are involved in the present work, they will be introduced in detail below.

A great majority of microbial natural products are biosynthesized from simple monomeric building blocks such as amino acids or short chain carboxylic acids on giant multifunctional enzymes named nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs), respectively.^[23;24] Although they have differences in utilization of building blocks, substrate activation and condensation mechanisms, both classes of enzymes share notable similarities in the modular architecture of various catalytic domains and assembly-line like mechanism. Based on the so-called “multiple carrier thio-template mechanism”,^[23] PKSs and NRPSs exhibit a multimodular organization consisting of repetitive catalytic units called modules, each of which is responsible for incorporation of one specific residue into the elongating “ketide” or “peptide” chain. Furthermore, each module can be subdivided into domains which represent the catalytic subunits in charge of individual biosynthetic steps of precursor loading, condensation, optional modifications and release of final products from the assembly line. The central tenet of this model is that chain elongation proceeds while the intermediates are tethered via covalent tethering of all the substrates, intermediates and products. For this, each carrier protein (CP) domain must be post-translationally modified by a dedicated phosphopantetheinyltransferase (PPTase) enzyme which catalyzes the *in trans* transfer of a 4'-phosphopantetheine (Ppant) moiety from coenzyme A to a highly conserved serine residue in the CP resulting in the conversion of the catalytically inactive *apo*-CP to active *holo*-CP (**Figure A3**).^[25-27] The choice of this thioester chemistry for activation of both acyl monomers and aminoacyl monomers provides the thermodynamic driving force and kinetically accessible nucleophiles for the Claisen (in case of polyketide) or amide bond (in case of nonribosomal peptide) condensations.^[25]

A. Introduction

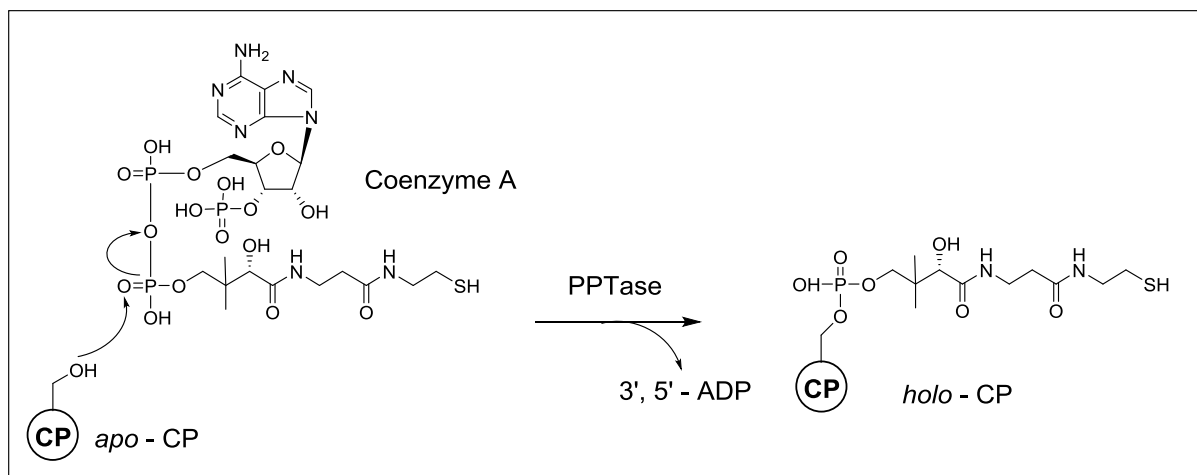


Figure A3. The post-translational modification of *apo*-CP converted to *holo*-CP by PPTase.

Apart from polyketides (PKs) and nonribosomal peptides (NRPs), ribosomally synthesized and post-translationally modified peptides (RiPPs) are another major class of natural products that have become increasingly more accessible thanks to recent genome sequencing efforts.^[28;29] RiPPs are produced by all three domains of life, such as lanthipeptides, thiopeptides and microcins from bacteria, amatoxins and phallotoxins from fungi, cyclotides and orbitides from plants, conopeptides from cone snails.^[29] Typically, RiPPs are extensively post-translationally modified and thus possess a restricted conformational flexibility which endows them with a better target recognition, metabolic stability and chemical functionality.^[29] Compared to the biosynthetic concept of PKSs and NRPSs, RiPPs biosynthesis is initiated by a ribosomally generated precursor peptide encoded by a structure gene. After the precursor peptide undergoes a variety of post-translational modifications (PTMs) upon binding to the corresponding post-translationally modifying enzymes, a final proteolytic cleavage of the leader peptide occurs, leading to the conversion of the precursor peptide into the final product which exhibits high chemical diversity and various biological activities.

In most cases, the genes encoding the biosynthetic pathway of a particular natural product are clustered together within the microbial genome. The first proof of this viewpoint was provided by Malpartida and Hopwood in 1984.^[30] They isolated and cloned a continuous DNA fragment from *Streptomyces coelicolor* which contained complete genetic information required for biosynthesis of the antibiotic actinorhodin. It turned out that the cloned DNA fragment could not only complement the production of actinorhodin in the non-producing mutants, but also be heterologously expressed in *Streptomyces parvulus*. Based on this rule, researchers could easily identify the biosynthetic gene cluster of natural products through single gene localization by straightforward chromosomal walking. Typically these gene clusters also harbour transcriptional regulators and genes responsible for self-resistances in upstream or downstream areas.^[31] However, many examples of split gene clusters are firstly found in myxobacteria.^[32-35]

3.1 Biochemistry of nonribosomal peptide synthetases

A minimal functional NRPS module requires three essential domains: an adenylation (A) domain, a peptidyl carrier protein (PCP) domain, also referred to a thiolation (T) domain and a condensation (C) domain. Every cycle of chain elongation is initiated by the action of an A domain which specifically selects, activates and transfers an amino acid to the PCP domain in the same module.^[36] In addition to 20 proteinogenic amino acids, various nonproteinogenic amino acids and aryl acids can also be recognized and processed as monomeric building blocks. This fact endows nonribosomal peptides (NRPs) with a large diversity of chemical structures and biological activities. The specificity of A domains towards corresponding amino acids play a crucial role in the peptide sequence of the final natural product and the biological activity as well. A successful crystallization and structural analysis of a prototype A domain with its amino acid substrate and ATP revealed the structural basis of substrate recognition and activation.^[37] Combined with primary sequence alignment analyses to other A domains, 8-10 amino acids within the substrate binding pocket could be identified. These residues could be defined as specificity-conferring code for substrate prediction of NRPSs.^[38] After the amino acid residue is transferred to the PCP domain, the C domain subsequently catalyzes the peptide bond formation between an amino acyl or peptidyl-S-PCP intermediate from the upstream module to the amino acyl moiety attached to the respective downstream module. As a result, the peptide chain is elongated via one amino acid residue and the product is tethered covalently to the downstream PCP domain. The peptide chain extension can be followed to the last module in the same fashion. In most cases, a fourth domain, the thioesterase (TE) domain, turns out to be essential for final product release from the biosynthetic template. A TE domain is usually located at the C-terminus of the last biosynthetic module and catalyzes the release of product from the assembly line in linear, cyclic or branched cyclic forms.^[39;40] All reactions catalyzed by essential NRPS domains are schematically summarized in **Figure A4**.

In addition to the domains described above, a number of optional domains have also been found in NRPSs which can contribute to structural diversity of NRPs. For example, the methylation (MT) domain can catalyze the transfer of a methyl group from *S*-adenosyl-L-Methionine (SAM) to the nitrogen or carbon atom of a PCP-bound amino acid.^[41] The epimerization (E) domain can convert the thioester-bound amino acid of an amino acyl-S-PCP or a peptidyl-S-PCP from L- into D-configuration.^[42;43] Heterocyclization (HC) domain is a substitution of the usual C domain in modules incorporating threonine, serine and cysteine residues. HC domains catalyze not only the peptide bond formation but also mediate the formation of the five-membered thiazoline or oxazoline heterocycles.^[44;45] Thiazoline or oxazoline heterocycles are often further oxidized to thiazoles or oxazoles by oxidation (Ox) domains using flavin mononucleotide (FMN) as a cofactor.^[46] It is noteworthy that Ox domains are strictly associated with the presence of HC-domains.^[23]

A. Introduction

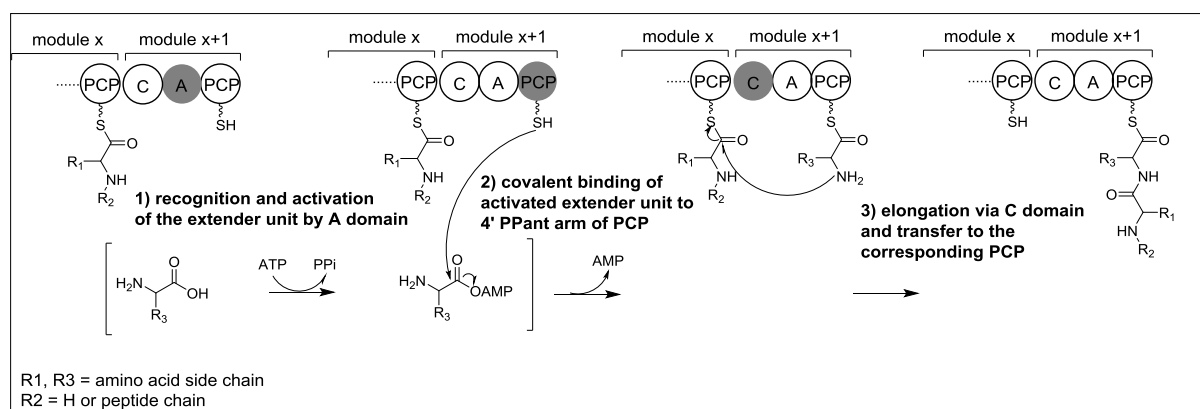


Figure A4. Schematic display of NRPS biochemistry.

3.2 Biochemistry of polyketide synthases

Analogous to NRPSs, a minimal functional PKS module also requires three essential domains: an acyl-transferase (AT) domain, an acyl carrier protein (ACP) domain and a α,β -ketoacyl synthase (KS) domain. Typically, a PKS module contains all the three domains except the loading module, which is responsible for initiating the selection of starter units of modular PKS (referred to as type I PKS, other types of PKSs are introduced below). Two architectures have been found to date: one type of loading module contains all the three domains, of which the KS domain is catalytically inactive because the active site cysteine residue is replaced by a glutamine residue, commonly referred to as a KS^Q domain.^[47] In this case, the AT-KS-ACP architecture selects malonyl-CoA or methylmalonyl-CoA followed by decarboxylation catalyzed by a KS^Q domain yielding acetate or propionate as a starter unit. Another type of loading module consists of an AT and an ACP domain. In this case, the AT domain directly selects a coenzyme A activated short chain (e.g. acetyl-CoA, propionyl-CoA), a branched-chain or an aromatic monocarboxylic acid.^[48-50] After the processing of loading modules, the starter unit is delivered onto the KS domain of the next extension module. In extension modules, different from the A domain of NRPS, by which the selected building block is activated via consumption of ATP, the AT domain prefers to select the pre-activated CoA thioesters of short dicarboxylic acids (e.g. malonyl-CoA and methylmalonyl-CoA) as extender units which then get transferred to the corresponding ACP domain of the same module. However, there is growing evidence for the biosynthesis and incorporation of unusual extender units such as hydroxymalonyl-CoA, methoxymalonyl-CoA, aminomalonyl-CoA, ethylmalonyl-CoA and chloroethylmalonyl-CoA.^[51-55] Similar to the PCP domain of NRPSs, the ACP domain is responsible for the transport of building blocks and is involved in the elongation of intermediates while the KS domain catalyzes the elongation reaction via the condensation between extender unit and the growing chain. The reaction is initiated with the transfer of the acyl-S-ACP intermediate of an upstream module to the conserved cysteine residue of the KS domain. Meanwhile, the KS domain catalyzes the decarboxylation of the extender unit tethered to the ACP of the corresponding module, resulting in a nucleophilic attack on the thioester of the ketide chain tethered to the cysteine residue of the KS domain. Via this condensation step a C₂-unit extended β -keto-acyl-S-ACP intermediate is formed.^[23;56] Following the same path, this

A. Introduction

prolonged intermediate is further transferred to the KS domain of next module and undergoes a chain elongation reaction again. All reactions catalyzed by essential PKS domains are schematically summarized in **Figure A5**.

Similar to NRPs, to enhance the structural diversity PKSs also contain a variety of optional domains acting on β -keto groups in certain modules after the chain elongation step.^[57] The auxiliary domains of the PKS mainly refer to the ketoreductase (KR) domain, the dehydratase (DH) domain and the enoylreductase (ER) domain. All three domains can stepwise reduce the β -keto moiety before transfer onto the KS domain of next module. The KR domain catalyzes the reduction of a β -keto group to a β -hydroxyl group by using the cofactor nicotinamide adenine dinucleotide phosphate (NADPH). Subsequently, this β -hydroxyl intermediate can be further dehydrated resulting in an α,β -enoyl intermediate which can be fully reduced into a saturated acyl chain by the ER domain which also uses NADPH as cofactor.

In addition to type I modular PKS, type II and type III PKS systems are also found in nature. Type II PKS contains a series of discrete monofunctional enzymes forming a multienzyme complex, which is employed iteratively to produce polycyclic, aromatic natural products.^[57] Type III PKS use a single catalytic activity including decarboxylation, condensation, cyclization and aromatization to accomplish the whole biosynthesis process.^[58;59] Another distinct difference between type III PKS and other two types lies in that they use CoA-esters as substrates instead of acyl-S-ACP intermediates.

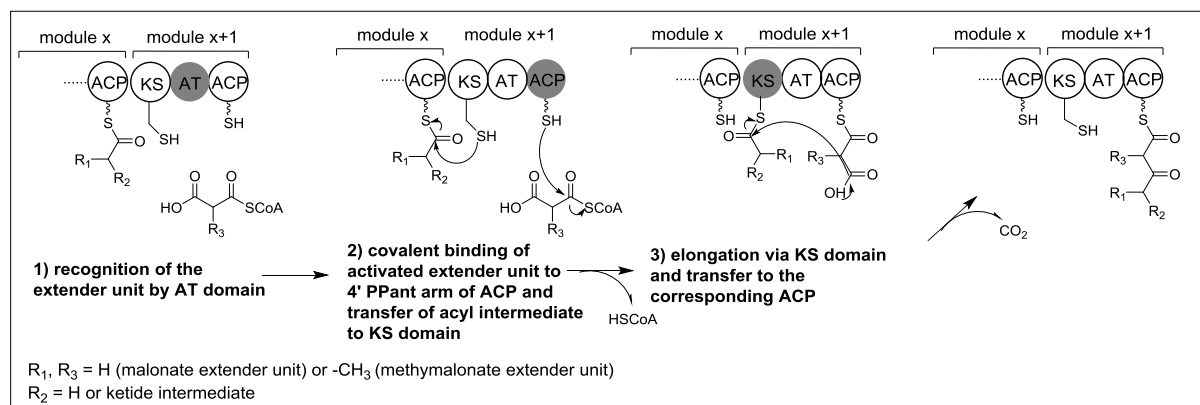


Figure A5. Schematic display of PKS biochemistry.

3.3 PKS/NRPS or NRPS/PKS hybrids

The similar structural organization and catalytic logic between PKS and NRPS make the cooperation of these multimodular enzymes possible. The combination of PKS and NRPS allows for immense variety of structures. The biosynthetic mechanisms of these PKS and NRPS hybrids could be classified into two types.^[60;61] The first type refers to no functional interactions between PKS and NRPS. In this case, the PKS and NRPS moieties are independently biosynthesized and finally coupled together by discrete enzymes. The phytotoxic coronatine biosynthetic pathway in *Pseudomonas syringae* is a representative for this type of hybrid system.^[62-64] A putative ligase is proposed to catalyze the final combination of coronafacic acid (PK part) and the coronamic acid (NRP part)

which are separately synthesized. Another type of PKS/NRPS hybrid system refers to the functional interactions between PKS and NRPS modules, meaning that a NRPS-bound peptidyl intermediate is directly transferred and elongated further by a PKS module or a PKS-bound acyl intermediate is directly transferred and subsequently elongated by a NRPS module. To date, quite a number of natural products have been discovered to follow such a type of PKS and NRPS hybrid pathway.^[55;65-68] In this case, the intermodular communication is critical for the functionality of the hybrid systems. For a PKS/NRPS interface, the C domain of NRPS must recognize the upstream acyl intermediate and catalyze the condensation between the ketide chain bound to the upstream ACP domain and the amino acid bound to the downstream PCP domain. Analogous, for a NRPS/PKS interface, the KS domain must accept the peptidyl intermediate bound to upstream PCP and condense it with the downstream acyl-S-ACP. Besides, another crucial factors for a functional hybrid system lie in the availability of a PPTas with a broad substrate specificity which can activate both types of CPs (ACPs and PCPs). To date, several biochemically characterized PPTases (MtaA, Sfp, Svp) turned out to be able to activate both ACPs and PCPs.^[69-71]

3.4 Biochemistry of ribosomally synthesized and post-translationally modified peptides

In most cases, biosynthesis of ribosomally synthesized and post-translationally modified peptides (RiPPs) starts with a linear precursor peptide 20 - 110 residues in length which is encoded by a so-called structure gene.^[29] Precursor peptides can be further subdivided into different functional segments: core peptides, leader peptides, optional recognition sequences and signal sequences.^[29] Core peptides comprise all the amino acid residues that will be transformed into a mature natural product. Leader peptides, which are predominately appended to the N-terminus of core peptides, guide the biosynthetic process by recognition of post-translational modification (PTM) enzymes and export of final products.^[72] In eukaryotes, an additional signal sequence located at the C-terminus of the leader peptide is often responsible for directing intermediates to appropriate cellular compartments where the PTMs occur. Sometimes the precursor peptide also contains an optional recognition sequence at the C-terminus for excision and cyclization to form the final product.^[73-75] A typical RiPPs biosynthetic pathway is displayed in **Figure A6**.

The tremendous post-translational modifications are the outstanding feature of RiPPs. Along with the explosion of genome data, it has been revealed that many RiPP biosynthetic pathways employ the common enzymes to realize post-translational modifications. Just to name some of them, radical-SAM dependent methyltransferases encoding genes exist in thiopeptides^[76] and proteusins^[77] biosynthetic pathways, oxazoline/oxazole and thiazoline/thiazole forming enzymes are found to participate in biosynthesis of thiopeptides^[76;78-81] and linear azole-containing peptides (LAPs),^[81-83] serine/threonine dehydratase that is commonly involved in lanthipeptides biosynthesis can also been used in proteusins,^[77] thiopeptides and LAPs biosynthesis,^[79;80;84;85] N- to C- (head to tail) heterocyclization occurs in biosynthesis of cyanobactins,^[86] amatoxins,^[87] bacteriocins,^[88] cyclotides^[89] and

A. Introduction

lanthipeptides.^[90] Besides, various sulfur chemistry including the conversion of thiols of cysteine to disulfides, thioethers or sulfoxides are involved in biosynthetic processes of lanthipeptides, cyclotides, thiopeptides, LAPs and cyanobactins, etc.^[29] The great variety of post-translational modifications endow the RiPPs with immense structural stability and diversity as well as high potential to valuable bioactivities.

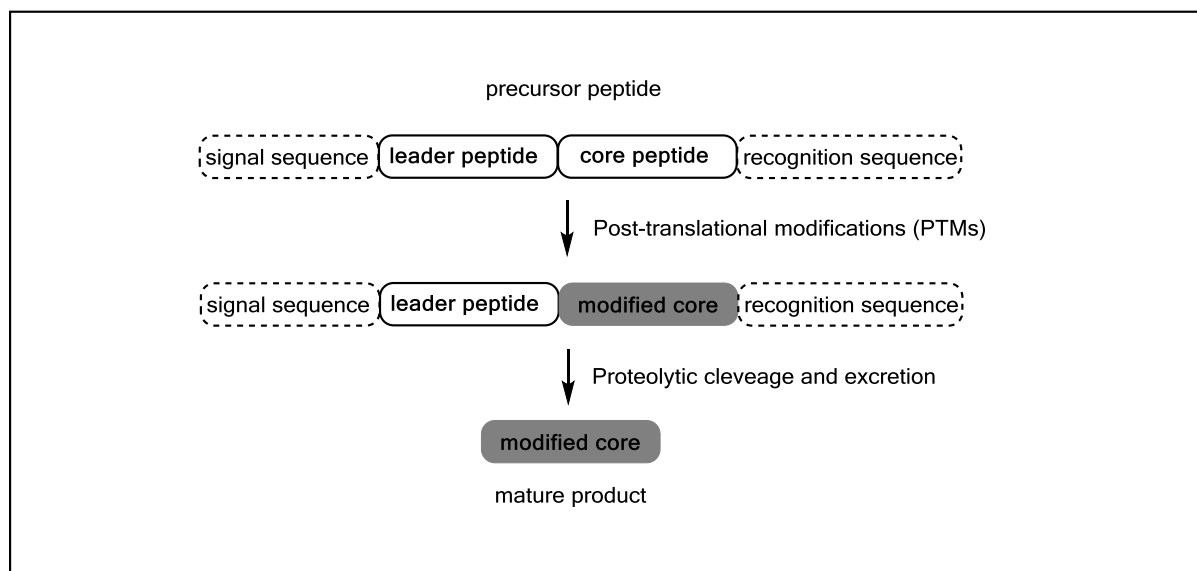


Figure A6. Schematic overview of a general RiPP biosynthetic pathway.

4. Ways to produce and manipulate natural products

Despite the simplicity of building blocks needed for assembly of PKs, NRPs or RiPPs, such as short chain carboxylic acids, amino acids and aryl acids, the structural complexity including numerous oxygen-containing substituents and an abundance of stereo centers extremely constrains their chemical syntheses route.^[91] Thus, microbial fermentation to generate these natural products is obviously much more economical and time-saving. However, this approach also faces various difficulties, e.g., low yield of target product in the native producer, tough cultivation conditions of the native producer or lack of genetic manipulation accessibility of the native producers.^[92;93] All of those leave the mobilization and expression of the corresponding biosynthetic gene cluster into an amenable heterologous host as a sufficient alternative to generate the desired natural product with improved yield and further allow the characterization and genetic engineering of the biosynthetic pathway for the generation of novel derivatives. In addition, establishment of a heterologous expression system can also elucidate the role of cryptic or silent gene clusters uncovered by genome sequencing projects in the post-genomics era.^[93]

4.1 Heterologous expression: General aspects

With respect to heterologous expression of complex natural products, several general considerations should be taken into account: 1) identification of the corresponding biosynthetic pathway in the

A. Introduction

original producer. In the past decade, plenty of natural product biosynthetic gene clusters have been revealed from the increasing amount of genome sequencing data.^[94] *In silico* identification and verification of the corresponding biosynthetic gene clusters from the genome sequences becomes a prerequisite for performance of heterologous expression; 2) a suitable vector containing the genetic elements of DNA transfer and integration into the host chromosome, functional promoter architecture, and the corresponding regulatory elements in the related host are also required. Traditionally, genomic libraries are constructed by using cosmids/fosmids that can bear an insert of size 30-35 kb or a bacterial artificial chromosome (BAC) that can accept an insert over 100 kb; 3) an amenable heterologous host is the most important aspect to be considered. First, the phylogenetic distance between the native producer and heterologous host should be considered. Generally, the more closely related the heterologous host is to the original producer, the more efficient is the transcriptional functionality of the biosynthetic pathway. Besides, closely related species may also share similar codon usage to the native producer, which may assure translational efficiency. With respect to heterologous expression of PKs and NRPs, the host must be able to post-translationally modify the giant multifunctional enzymes resulting in their active *holo* form, which is catalyzed by a suitable PPTase with broad specificity (**Figure A3**). Additionally, a pool of required precursors comprising CoA-activated short chain carboxylic acids, proteinogenic and nonproteinogenic acids must be supplied in adequate quantities to the host. Lastly, the aspect of self-resistance has to be considered, especially in cases where the target product exhibits antibacterial activity corresponding genes responsible for self-resistance are necessary to be transferred into host strain.^[93;95] All these aspects have been elaborated on in recent reviews.^[93]

It is noteworthy to state that even if all the elements described above have been taken into account, a successful and sufficient expression of a complex heterologous biosynthetic pathway still cannot be guaranteed in reality. This could be explained by several reasons, e.g., the expression levels of enzymes involved in the pathway in the heterologous host are different and they could require very diverse optimal expression conditions,^[93] or that the complex regulatory system does not work as well as in the native producer, etc. Thus, further genetic engineering is desired to optimize the heterologous expression levels and eventually generate novel derivatives.

4.2 Genetic engineering of expression constructs

First of all, the common problem for genetic manipulation of the large natural product pathways is that they cannot be modified by classical restriction enzymes. A very efficient and straightforward method Red/ET technology is quite beneficial to solve this problem. Red/ET recombineering is an *in vivo* technique that can be employed to specifically and precisely target the DNA molecules by homologous recombination using short homologous arm sequences in *Escherichia coli* mediated by phage-derived protein pairs, RecE/RecT from Rec phage or Red α /Red β from lambda phage (**Figure A7**).^[96;97] Compared to conventional cloning technologies, Red/ET recombineering is completely

A. Introduction

independent of DNA restriction enzymes. Furthermore, since the short homologous arm sequences can be synthesized chemically, in principle, any nucleotide of target DNA (cosmid, fosmid and BAC) can be altered by a variety of modifications, e.g., insertion, deletion, fusion, point mutation, sub-cloning or direct cloning.^[98-100]

A number of successful examples have been accomplished in our group using Red/ET technique for expression and engineering of large biosynthetic pathways. In 2005, a method was developed by Wenzel, *et al.* to reconstitute the hybrid NRPS and PKS myxochromide S pathway from myxobacterium *Stigmatella aurantiaca* on one expression construct and heterologously express myxochromide in *Pseudomonas putida* resulting a much higher yield of native producer.^[92] Perlova, *et al.* also used Red/ET to stitch together two cosmids containing the entire myxothiazole biosynthetic pathway and heterologously expressed myxothiazole in *Myxococcus xanthus*^[101] as well as *Pseudomonas putida*.^[102] In last ten years, further examples showed how efficient and promising this technique is for construction of large complex biosynthetic pathways and subsequent heterologous expression.^[103-106]

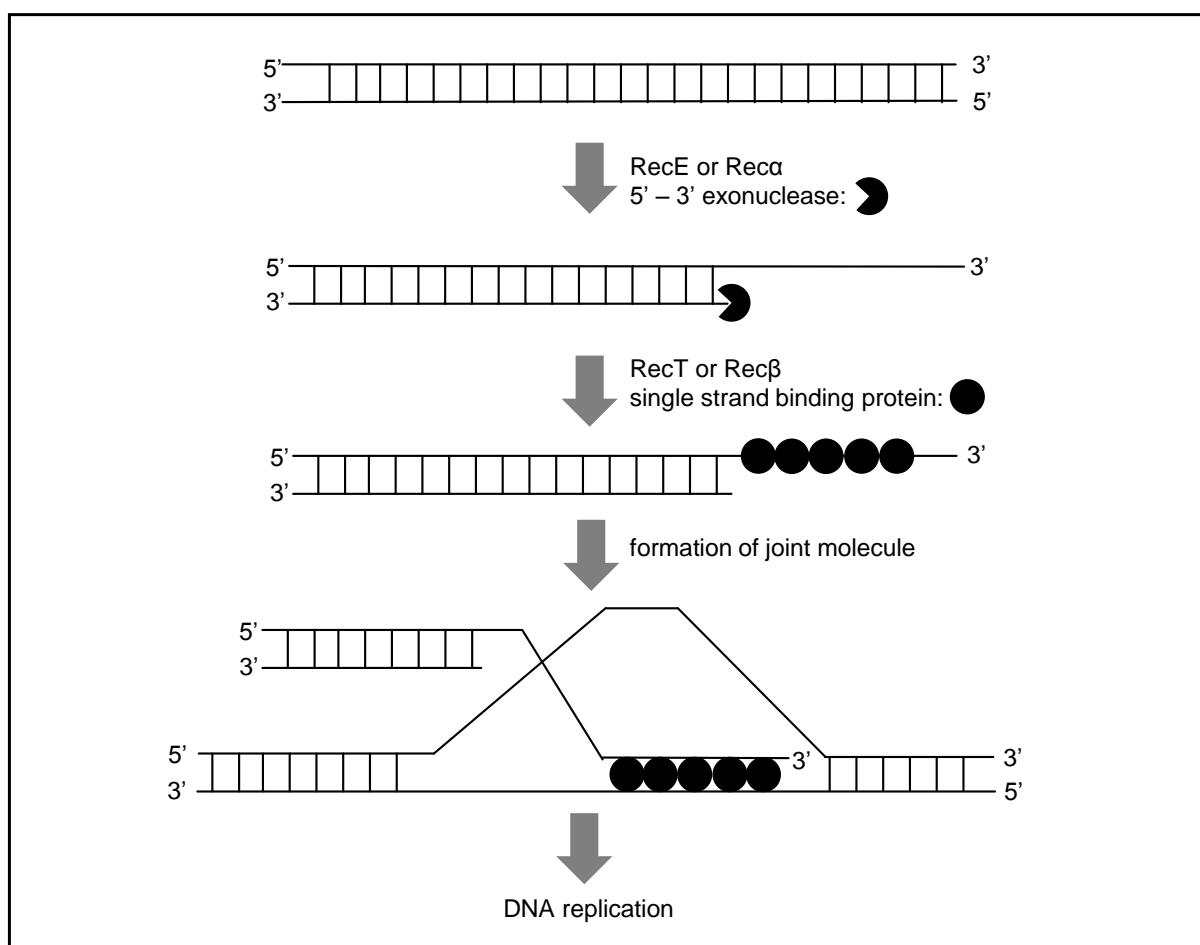


Figure A7. Mechanism of Red/ET recombination.

5. Generation of “unnatural” natural products via synthetic biotechnology

Oftentimes natural products are a very promising starting points for drug discovery; however, due to their physicochemical properties they are usually far from being ideal for clinical application.^[107] Thus, analogue development of natural products is necessary to improve the bioavailability as well as the bioactivity. In addition, to fully understand the natural products' mode of action based on the structure-activity relationship (SAR) studies a library of natural product analogues are often required to determine the essential functional groups for activity. Whilst the increasingly released genome data and fast development of genetic engineering and manipulation methods the combined use of biological and chemical tools to establish a natural products library are increasingly approached in comparison to the traditional total synthesis.^[108;109] There are several successful approaches that have been employed to generate a variety of natural product analogues. Among them, the precursor-directed biosynthesis and mutasynthesis that have been used in this thesis will be introduced in detail below (**Figure A8**).

The application of precursor-directed biosynthesis to natural products analogues generation, although an old concept, has been increasingly utilized over the past decade.^[107] It makes use of the tolerant flexibility of natural product biosynthetic enzymes enabling the acceptance of unnatural building blocks resulting in altered various unnatural products. Once the biosynthetic logic and its substrate specificity are well understood, unnatural precursors can be rationally designed, chemically synthesized and subsequently fed into the growth media of corresponding microorganisms. They are subsequently incorporated in the biosynthetic pathways due to substrate flexibility of the biosynthetic enzymes. By this approach, natural product analogues can be obtained through feeding of designed unnatural building blocks into the culture of the wild type producer. Therefore, the most limiting factor of this approach lies in the flexibility of substrate specificity. Moreover, the production of those analogues is always competing with that of the natural products which often leads to an insufficient yield of the unnatural analogues. Furthermore, owing to the structural similarity between the natural and unnatural products, the alongside produced native natural products make the downstream chromatographic separation more challenging and time-consuming. A representative example is the generation of analogues of antibiotic diazepinomycin. Ratnayake, Carter and coworkers demonstrated that the diazepinomycins are biosynthesized from tryptophan and anthranilate, which inspired researchers to explore whether the biosynthetic enzymes could be flexible to incorporate non-native analogues of these two precursors.^[110] Thus, they conducted the precursor-directed biosynthesis approach by feeding 4-, 5-, and 6-F-tryptophans and various 5-halo-anthranilates individually to the cultures of *Micromonospora* sp. DPJ15. It turned out that the biosynthetic assembly line could sufficiently incorporate the 5-, 6-F-tryptophans and 5-F-anthranilate into the assembly line while the administrated 5-Cl-anthranilate, 5-Br-anthranilate and 5-I-anthranilate were not utilized as substrates.^[110] The selective incorporation of fed non-native precursors and relatively low incorporation rates corroborates that the precursor-directed biosynthesis approach is dependent on the

A. Introduction

flexibility of the biosynthetic assembly line and the yield of generated novel derivatives is limited by the competition from the native compound.

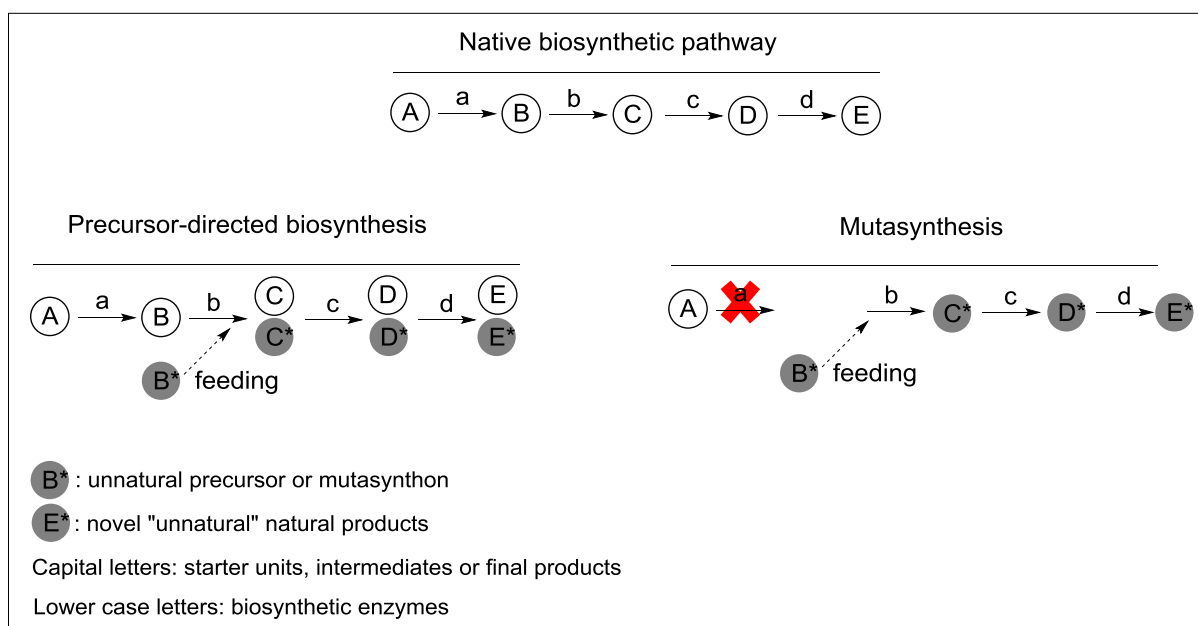


Figure A8. Schematic display of concepts of precursor-directed biosynthesis and mutasynthesis.

Another often applied strategy for the generation of natural product analogues is mutasynthesis. This concept was first raised over 40 years ago using mutants of antibiotic neomycin producer *Streptomyces fradiae* to generate novel antibiotics.^[111] Since then, this strategy has been applied to many other classes of compounds including siderophores,^[112] nucleosides,^[113] nonribosomal peptides^[114;115] and polyketides.^[116] In this case, the target biosynthetic pathway is blocked at a crucial step in corresponding mutants, which are able to employ unnatural building blocks into the biosynthetic pathway to jump-start the assembly line yielding unnatural analogues of interest. The major approach to mutasynthesis is the disruption of the essential gene(s) responsible for biosynthesis combined with feeding of the corresponding intermediates (so-called mutasynthons) to complement the production of the novel unnatural compounds upon the fed substrates.^[117] Similar to precursor-directed biosynthesis mutasynthesis is also dependent on the substrate flexibility of the natural biosynthetic enzymes. However, compared to precursor-directed biosynthesis the advantage of this approach lies in that there is no competition with the production of natural compounds at all, which allows a sufficient yield of the unnatural products and a simplified downstream compound separation process. This approach potentially allows the assessment to generation of a library of unnatural compounds. With respect to this, Moore and his colleagues employed this strategy to generate the fluorinated analogues to salinosporamides, a family of chlorinated compounds from *Salinispora tropica* CNB-440 exhibiting a highly potential proteasome inhibitory activity from.^[118-120] Moore and colleagues disrupted the *salL* gene encoding the chlorinase and subsequently complemented the *salL*-deficient mutant with chemically synthesized 5'-fluoro-5'-deoxyadenosine (5'-FDA). The resultant

fluorosalinosporamide was able to serve as a reversible inhibitor with a much greater activity than deschlorosalinosporamide.^[121;122]

In addition to precursor-directed biosynthesis and mutasynthesis, semi-synthetic approaches, combinatorial biosynthesis, post-biosynthetic modification and chemogenetics, etc. have also been increasingly applied for the generation of natural product analogues.^[107] A combinatorial application of these biological and chemical tools is becoming more popular and productive than using a single approach.

6. Outline of the dissertation

The overall goal of the work described in this thesis dealt with study and engineering of natural product biosynthetic pathways from Streptomycetes. Two novel bioactive compound families have been intensively studied: the cytotoxic and antifungal cinnabaramides from *Streptomyces* sp. JS360 and the antibacterial bottromycins from *Streptomyces* sp. BC16019, both of which represent promising lead structures for potential drug development.^[123-126]

The cinnabaramides (**Figure A9**) are mixed PKS and NRPS natural products, which were recently isolated from a terrestrial Streptomyce *Streptomyces* sp. JS360.^[123] They interfere with the proteasome and thus potentially inhibit the growth of cancer cells. The compounds exhibit a γ -lactam- β -lactone bicyclic ring structure attached to a cyclohexenyl unit and a hexyl side chain originating from an unusual PKS extender unit.^[123] Apart from the hexyl side chain, significant structural similarity was found between cinnabaramide and the salinosporamides containing a chlorinated ethyl side chain instead (**Figure A9**). Salinosporamides are a class of antitumor compounds produced by *Salinispora tropica* CNB-440.^[127-129] The salinosporamides exhibit potent cancer cell cytotoxicity and exert their effects via binding to the 20S proteasome core particle, thereby inducing cell-cycle arrest and programmed cell death (apoptosis).^[130] The mode of action studies of salinosporamides demonstrated that the key to the superior activity lies in the chlorine atom attached to the side chain which is required for the irreversible binding to the target.^[131]

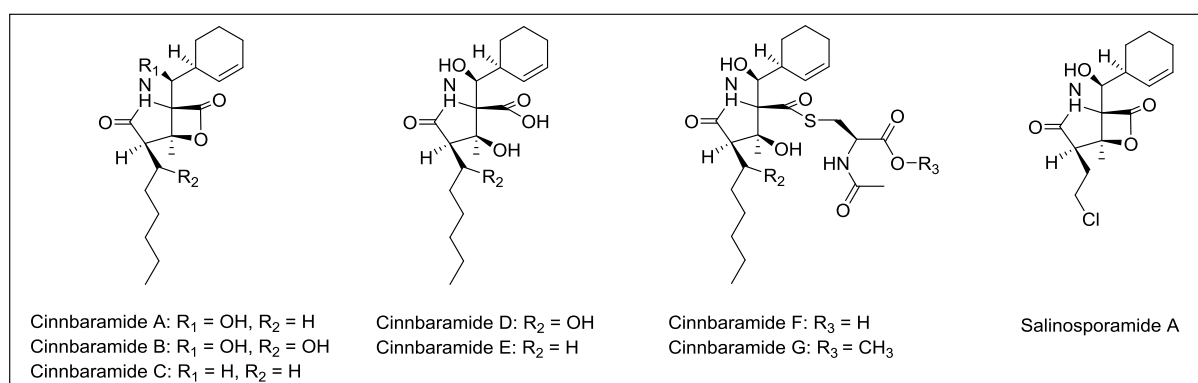


Figure A9. Structures of cinnabaramides produced by *Streptomyces* sp. JS360 and salinosporamide A produced by *Salinispora tropica*.

A. Introduction

Thus, we intended to generate chlorinated analogues of the cinnabaramides and screen their bioactivities after mining the cinnabaramide gene cluster with subsequent genetic and molecular engineering approaches. These results were described in **publication I and publication III**.

Meanwhile, we intended to determine the formation of the unusual extender unit incorporated into the PKS pathway. One key enzyme encoded by *cinF*, a gene involved in the downstream processing of the cinnabaramide biosynthetic gene cluster, was supposed to be responsible for the formation of the unusual extender unit hexylmalonyl-CoA in the PKS assembly line. Therefore, biochemical and structural characterization of this enzyme was planned to figure the detailed reaction mechanism out. All the data and analyses were reported in **publication II**.

The second part of this thesis refers to the in-depth studies on the biosynthetic pathway of the highly modified octapeptide bottromycins (**Figure A10**). Bottromycins were first discovered as antibacterial peptides with promising activity against Gram-positive bacterial and mycoplasma from the fermentation broth of *Streptomyces bottropensis*.^[126;132-135] Later on it was shown that their antibacterial ability extends to methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE).^[136] Mode of action studies revealed the aminoacyl-*t*RNA binding site (A site) on the 50S ribosome as the target of bottromycins ultimately leading to the inhibition of protein synthesis.^[137-139] Bottromycins represent octapeptides exhibiting an internal tetrapeptide cycle formed via a unique amidine linkage. Furthermore, the compound harbors an *exo*-thiazole and several unnatural amino acids which carry methyl-groups at non-activated carbons posing additional challenges for synthetic approaches aimed towards drug development. In this work, we intended to improve the structure and production yield using synthetic biotechnological approaches. To achieve this goal, the biosynthetic genes from *Streptomyces sp.* BC16019 were firstly identified by Dr. Shwan Rachid. This work uncovered a ribosomal locus required for bottromycin biosynthesis and identified bottromycins as a novel class of RiPPs exhibiting a variety of unprecedented features exemplified by the absence of a N-terminal leader sequence on the structure gene. After establishment of a heterologous expression platform, the yield was optimized and novel bottromycin analogues were identified using synthetic biotechnology approaches. In addition, some evidences for the formation of the *exo*-thiazole ring and a rare macrocyclodehydration mechanism leading to amidine ring formation was also provided. All the results were presented in **publication IV**.

A. Introduction

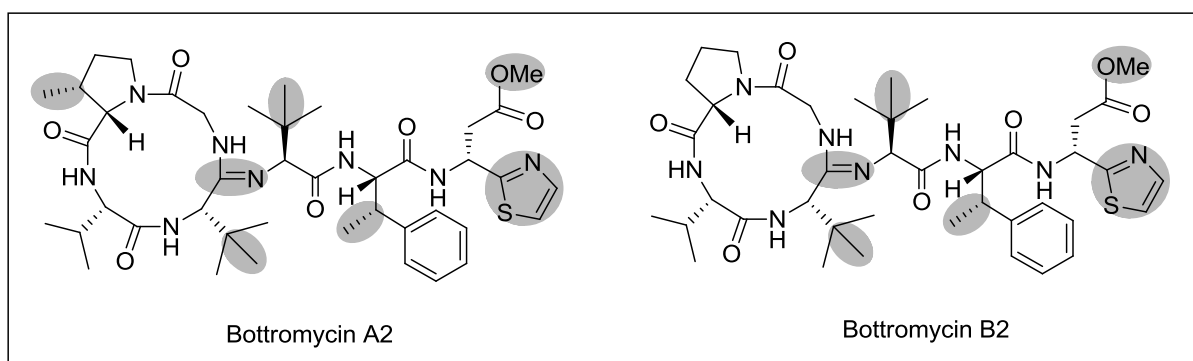


Figure A10. Structures of bottromycin A2 and B2 produced by *Streptomyces* sp. BC16019. Intriguing structure features were highlighted in gray.

I

Mining the cinnabaramide Biosynthetic pathway to Generate Novel Proteasome Inhibitors

Shwan Rachid,^[a,b,c] Liujie Huo,^[a,b] Jennifer Herrmann,^[a,b] Marc Stadler,^[d]
Bärbel Köpcke,^[d] Jens Bitzer,^[d] and Rolf Müller*^[a,b]

[a] Dr. S. Rachid, L. Huo, J. Herrmann, Prof. Dr. R. Müller, Department of Microbial Natural Products (MINS), Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI), Saarland University, Campus C2.3, 66123 Saarbrücken, Germany.

[b] Dr. S. Rachid, L. Huo, J. Herrmann, Prof. Dr. R. Müller*, Department of Pharmaceutical Biotechnology, Saarland University, Campus C2.3, P.O. Box 15 11 50, 66041, Saarbrücken, Germany. Fax: (+49)681-302-70202, Email: rom@mx.uni-saarland.de

[b] Dr. S. Rachid, Current address: College of Pharmacy, University of Sulaimani, Sulaiamni, Iraq.

[b] Dr. M. Stadler, Dr. B. Köpcke, Dr. J. Bitzer, InterMed Discovery GmbH. Otto-Hahn-Straße 15, 44227, Dortmund, Germany.

Mining the Cinnabaramide Biosynthetic Pathway to Generate Novel Proteasome Inhibitors

Shwan Rachid,^[a, b, c] Liujie Huo,^[a, b] Jennifer Herrmann,^[a, b] Marc Stadler,^[d] Bärbel Köpcke,^[d] Jens Bitzer,^[d] and Rolf Müller^{*[a, b]}

The cinnabaramides and salinosporamides are mixed PKS/NRPS natural products isolated from a terrestrial streptomycete and a marine actinomycete, respectively. They interfere with the proteasome and thus potentially inhibit the growth of cancer cells. The compounds exhibit a γ -lactam- β -lactone bicyclic ring structure attached to a cyclohexenyl unit and a PKS side chain. As a first step towards improving anticancer activity and permitting genetic approaches to novel analogues, we have cloned and characterized the cinnabaramide biosynthetic genes from *Streptomyces* sp. JS360. In addition to the expected PKS and NRPS genes, the cluster encodes functionalities for the assembly of the hexyl side chain precursor. The corre-

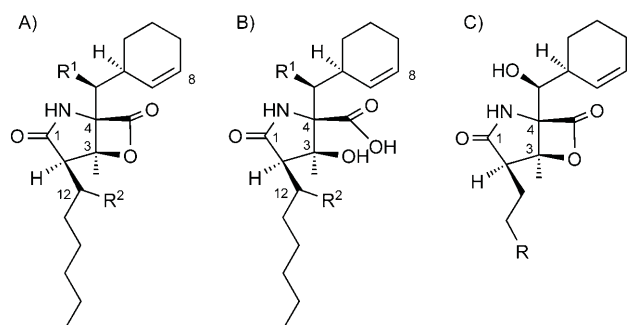
sponding enzymes exhibit relaxed substrate specificities towards a number of synthesized precursors, enabling production of novel chlorinated cinnabaramides. These were isolated and analyzed for activity, revealing that derivatives bearing a chlorine atom in the PKS side chain show higher inhibitory potentials towards the proteasome's proteolytic subunits (especially the trypsin and chymotrypsin units) and higher cytotoxicities towards human tumor cell lines than the parent cinnabaramide A. Although their activities towards the proteasome were weaker than that of salinosporamide A, the cinnabaramides were found to inhibit the growth of various fungi with greater potency.

Introduction

Natural products continue to fulfill their enormous potential in the development of future drugs and to serve as compounds of interest both in their natural forms and as semisynthetic or synthetic derivatives. Drugs derived from natural products are used for therapy for more than 87% of human diseases, and include anticancer, antifungal, antibacterial, anticoagulant, antiparasitic, and immunosuppressant agents, amongst others.^[1,2]

An important role in drug discovery in the past century has been played by microorganisms, in particular bacteria of the genus *Streptomyces*, which produce secondary metabolites in unprecedented chemical diversity, exhibiting a remarkable array of biological activity. Many of them are clinically valuable agents.^[3,4] The metabolites isolated from this genus are mostly end products of the enzymes of two major types of biosynthetic pathways: the non-ribosomal peptide synthetases (NRPSs) and the polyketide synthases (PKSs).^[4] NRPSs and PKSs are made up of so-called modules, which are sets of distinct active sites for catalyzing each condensation and chain-elongation step.^[5,6] Each module in an NRPS or a PKS consists of certain obligatory or core domains for addition of each peptide or ketide unit and a variable number of optional domains responsible for modification of the peptide/ketide backbone.^[7,8]

The terrestrial *Streptomyces* sp. JS360 is known to produce cinnabaramides A–G, each exhibiting a γ -lactam- β -lactone bicyclic ring structure attached to a cyclohexenyl unit, together with a hexyl side chain.^[9] With the exception of the hexyl side chain, significant structure similarity was found between these compounds and the salinosporamides (Scheme 1), a class of PKS/NRPS antitumor metabolites isolated from the marine actinomycetes *Salinispora tropica* CNB-440^[10–12] and *Salinispora*



Scheme 1. Structures of cinnabaramides and salinosporamides. A) Cinnabaramide A ($R^1 = \text{OH}$, $R^2 = \text{H}$), cinnabaramide B ($R^1 = R^2 = \text{OH}$), and cinnabaramide C ($R^1 = R^2 = \text{H}$). B) Cinnabaramide D ($R^1 = R^2 = \text{OH}$) and cinnabaramide E ($R^1 = \text{OH}$, $R^2 = \text{H}$). C) Salinosporamide A ($R = \text{Cl}$) and salinosporamide B ($R = \text{H}$).

[a] Dr. S. Rachid, L. Huo, J. Herrmann, Prof. Dr. R. Müller
Department of Microbial Natural Products (MINS)
Helmholtz Institute for Pharmaceutical Research Saarland (HIPS)
Helmholtz Centre for Infection Research (HZI), Saarland University
Campus C2.3, 66123 Saarbrücken (Germany)

[b] Dr. S. Rachid, L. Huo, J. Herrmann, Prof. Dr. R. Müller
Department of Pharmaceutical Biotechnology, Saarland University
Campus C2.3, P.O. Box 15 11 50, 66041 Saarbrücken (Germany)
Fax: (+49) 6181-302-70202
E-mail: rom@mx.uni-saarland.de

[c] Dr. S. Rachid
Current address: College of Pharmacy
University of Sulaimani, Sulaimani (Iraq)

[d] Dr. M. Stadler, Dr. B. Köpcke, Dr. J. Bitzer
InterMed Discovery GmbH
Otto-Hahn-Strasse 15, 44227 Dortmund (Germany)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201100024>.

pacifica CNT-113.^[13] Both cinnabaramides and salinosporamides exhibit potent cancer cell cytotoxicities and exert their effects through binding to the 20S proteasome core particle, thereby inducing cell cycle arrest and programmed cell death (apoptosis).^[9,14–16] The key to the superior activity of salinosporamide A versus the cinnabaramides is assumed to be the presence of the chlorine atom attached to the side chain, which is required for the drug's irreversible binding to its biological target and is thus a major reason for the compound's effectiveness against cancer.^[17]

We anticipated that the factor behind the varying side chains in the two compound classes should be a biosynthetic enzyme responsible for forming a hypothetical and unusual hexylmalonyl-CoA extender unit. By mining the substrate tolerance of this enzyme towards chlorinated and unsaturated fatty acid-CoA esters and using feeding experiments we succeeded in generating chlorinated analogues of the cinnabaramides and in analyzing their bioactivities.

Results and Discussion

To achieve our goal, we first set out to identify the biosynthetic machinery of cinnabaramides. Because the overall structure of cinnabaramides is suggestive of a biosynthetic origin based upon PKS and NRPS machinery and because the gene cluster for salinosporamides (*sal*) is known,^[18] a cosmid library was constructed from the genome of strain JS360 and screened for the corresponding gene cluster in an approach based on hybridization with degenerated PKS and NRPS (as described in the Experimental Section).

Within the identified *cin* gene cluster, at least six likely structural genes were designated *cinA–F*, encoding two PKS modules and a split NRPS module distributed over the first two genes (*cinA* and *cinB*), followed by a standalone ketosynthase, a P450 monooxygenase, a standalone thioesterase, and a homologue of crotonyl-CoA carboxylase/reductase (*Ccr*; *cinF*) (Table 1 and Scheme 2A), which can be assigned to roles in cinnabaramide core biosynthesis. Further genes are encoded upstream and downstream of the gene cluster and would be expected to be involved in production of the biosynthesis precursors (*cinQ* and *cinP*), post-assembly-line modification of the compound (*cinL*), bacterial resistance against the compound (*cinJ*), and regulation of the gene cluster (*cinRI* and *cinRII*) (Table 1). It is likely that the biosynthesis of cinnabaramides is carried out in a thiotemplated fashion, in which intermediates are covalently tethered to phosphopantetheinyl arms of carrier proteins (Scheme 2B).^[19] *CinA* consists of an atypically arranged starter module and one extension module (ACP-KS-AT₁-AT₂-ACP-C). We assume that the initiation step is selection and loading to the ACP of the starter unit acetyl-CoA by the first AT domain. According to the chemical structure of the compound, the second PKS module catalyzes chain elongation of acetyl-S-ACP with the unusual extender unit hexylmalonyl-S-ACP and results in the incorporation of the hexyl side chain. A bioinformatics-based approach to identify the substrate specificity of *CinA*-ATs, in comparison with the ATs of other biosynthetic pathways,^[20] indeed showed conserved residues in the

ATs implicated in binding of dicarboxylated extender units (Table 1). Interestingly, the substrate specificity of AT1 is predicted to be for acetyl-CoA and the conserved residues are almost identical to the salinosporamide domain. AT2, however, is different from AT2 in its salinosporamide counterpart. Unexpectedly, it also differs from the hexylmalonyl-CoA-specific TugC-AT3 (Table S1B in the Supporting Information) involved in the biosynthesis of thuggacins—macrolide antibiotics bearing hexyl side chains—from the myxobacterium *Sorangium cellulosum* So ce895, which are also produced through a NRPS/PKS biosynthetic pathway.^[21–23]

The C-terminal domains of *CinA* and *CinB* form the terminal NRPS module of the assembly line, and the highly homologous domain set in the salinosporamide gene cluster is known to activate the nonproteinogenic amino acid β -hydroxy-2'-cyclohexenylalanine.^[12,24] Because of the structural similarity of the two compounds, and because the deduced substrate specificity code of the *CinB* A domain (DLMNVGGV; determined as described previously^[25,26]) showed a significant degree of similarity to that of the *SalB* A domain (DLLSNGGV), we propose elongation of the cinnabaramide intermediate by extension with β -hydroxy-2'-cyclohexenylalanine (Scheme 2B). The resulting PCP-tethered cinnabaramide intermediate is then thought to undergo an intramolecular condensation promoted by the ketosynthase (*CinC*) to generate the γ -lactam ring, which provides the *CinB*-bound alcohol substrate that gives rise to the offloaded β -lactone product. It is assumed that the standalone *CinE* catalyzes release and cyclization of the intermediate to generate the β -lactone ring of the compound. The gene cluster harbors genes (*cinQ* and *cinP*) that encode putative L-amino acid aminotransferases and a prephenate dehydratase, respectively (Table 1). Similar enzymes have also been described in the salinosporamide biosynthesis pathway—it was suggested that these were involved in cyclohexenyl-alanine building block formation—and inactivation of a prephenate dehydratase homologue gene (*salX*) in the *sal* cluster led to the elimination of all natural salinosporamides.^[12,27]

It has been reported that the salinosporamide biosynthetic enzymes responsible for the activation and incorporation of the native *sal*-specific amino acid exhibit relaxed substrate specificities towards a number of aliphatic amino acids.^[28] In that work, Moore and co-workers used a comprehensive approach that combined chemical synthesis with metabolic engineering to generate a series of salinosporamide analogues with altered bioactivities against human cancer cell lines. Two salinosporamide derivatives with saturated cyclohexane or cyclopentane rings were prepared by mutasynthesis, through administration of derivatives and branched-chain amino acids to a culture of the *S. tropica salX* mutant strain.^[29]

To establish the function of *cinQ* in the salinosporamide gene cluster we inactivated the gene by insertion of an inactivation plasmid (pCN2) into the *Streptomyces* JS360 genome, leading to complete elimination of cinnabaramide biosynthesis in the mutant. When we attempted to restore product formation by feeding cyclopentyl-DL-alanine to the culture, new compounds exhibiting the expected masses were indeed detected. NMR spectroscopic examination of the purified cinna-

Table 1. Deduced functions of the ORFs identified within the cinnabaramide gene cluster based on percentage sequence similarity/identity to previously known proteins as well as to the salinosporamide biosynthetic gene cluster determined by use of BLAST searching and clustalW alignment.

Gene	Protein	Residues [aa]	Proposed function	Sequence identity/similarity to [%] ^[a]	Accession number	Sequence identity/similarity to <i>Salinispora tropica</i> CNB-440 ^[18] [%] ^[b]
<i>cinA</i>	CinA	2040	PKS/NRPS (ACP ₁ -KS-AT ₁ -AT ₂ -ACP ₂ -C)	SalA, PKS/NRPS, <i>Salinispora tropica</i> CNB-440, 62/72	ABP73645	SalA, PKS/NRPS, 60/71
<i>cinB</i>	CinB	585	NRPS (A-ACP)	SalB, NRPS, <i>Salinispora tropica</i> CNB-440, 66/76	ABP73646	SalB, NRPS, 61/71
<i>cinC</i>	CinC	596	ketosynthase	SalC, erythronolide synthase, <i>Salinispora tropica</i> CNB-440, 66/79	YP_001157874	SalC, erythronolide synthase, 63/76
<i>cinD</i>	CinD	414	cytochrome P450	SalD, cytochrome P450 hydroxylase, <i>Salinispora tropica</i> CNB-440, 76/88	ABP73648	SalD, cytochrome P450 hydroxylase, 72/84
<i>cinE</i>	CinE	277	thioesterase	SalF, thioesterase, <i>Salinispora tropica</i> CNB-440, 70/80	ABP73650	SalF, thioesterase, 64/74
<i>cinF</i>	CinF	448	octenoyl-CoA reductase/carboxylase	crotonyl-CoA reductase, <i>Streptomyces hygroscopicus</i> , 93/96	AAR32675	SalG, 4-chlorocrotonyl-CoA reductase/carboxylase, 67/81
<i>cinRl</i>	CinRl	333	regulatory protein, LysR family	LysR family transcriptional regulator, <i>Streptomyces griseus subsp. griseus</i> NBRC 13 350, 48/60	YP_001828298	–
<i>cinH</i>	CinH	733	NRPS	NRPS, <i>Streptomyces avermitilis</i> MA-46 801, 50/60	NP_822023	–
<i>cinI</i>	CinI	506	acyl-CoA dehydrogenase	acyl-CoA dehydrogenase, <i>Streptomyces bingchengensis</i> BCW-1, 40/55	ADI05331	–
<i>cinJ</i>	CinJ	281	20S proteasome β-subunit	20S proteasome β-subunit, <i>Streptomyces avermitilis</i> MA-4680, 74/88	NP_823988	SalI, proteasome β-subunit, 40/60
<i>cinK</i>	CinK	66	ferredoxin-2	ferredoxin-2, <i>Streptomyces diastaticus</i> , 55/73	P18 325	–
<i>cinL</i>	CinL	402	cytochrome P450	putative cytochrome P450, <i>Streptosporangium roseum</i> DSM 43021, 49/63	YP_003340485	–
<i>cinM</i>	CinM	280	phenazine biosynthesis (oxidizing protein)	PhzF family phenazine biosynthesis protein, <i>Salinispora tropica</i> CNB-440, 69/80	YP_001157895	PhzF family phenazine biosynthesis protein, 69/80
<i>cinN</i>	CinN	95	hypothetical protein	hypothetical protein Strop_1042, <i>Salinispora tropica</i> CNB-440, 71/78	YP_001157894	hypothetical protein strop_1042, 71/78
<i>cinO</i>	CinO	329	ornithine cyclodeaminase	ornithine cyclodeaminase, <i>Salinispora tropica</i> CNB-440, 71/83	YP_001157893	ornithine cyclodeaminase, 71/83
<i>cinP</i>	CinP	205	prephenate dehydratase	BacA, prephenate dehydratase, <i>Bacillus amylo-liquefaciens</i> , 36/58	Q8KWT1	–
<i>cinQ</i>	CinQ	331	L-amino acid aminotransferase	SalW, L-amino acid transferase, <i>Salinispora tropica</i> CNB-440, 70/80	YP_001157891	SalW, L-amino acid transferase, 36/79
<i>cinR</i>	CinR	447	phenyl-acetate-CoA ligase	hypothetical protein Strop_1038, <i>Salinispora tropica</i> CNB-440, 70/81	YP_001157890	hypothetical protein Strop_1038, 70/81
<i>cinS</i>	CinS	352	3-oxoacyl-(acyl-carrier protein) synthase III	3-oxoacyl-(acyl-carrier protein), <i>Streptomyces hygroscopicus</i> , 90/93	AAR32677	–
<i>cinT</i>	CinT	591	acyl-CoA ligase	acyl-CoA ligase, <i>Streptomyces hygroscopicus</i> , 87/92	AAR32676	–
<i>cinRll</i>	CinRll	323	regulatory protein, LuxR family	LuxR transcriptional regulator, <i>Streptomyces flavogriseus</i> , 49/65	ZP_05802671	–

[a] Protein Blast (Database: nonredundant protein sequences (nr)). [b] Geneious alignment (cost matrix: BLOSUM62, Gap open penalty: 11, Gap extension penalty: 3)

baramide A analogue ($C_{18}H_{29}NO_4$; m/z : 324.2 $[M+H]^+$) revealed that the derivatives made by mutasynthesis were novel cinnabaramides each carrying a cyclopentyl instead of cyclohexenyl ring (Table S1 and Figures S1 and S4).

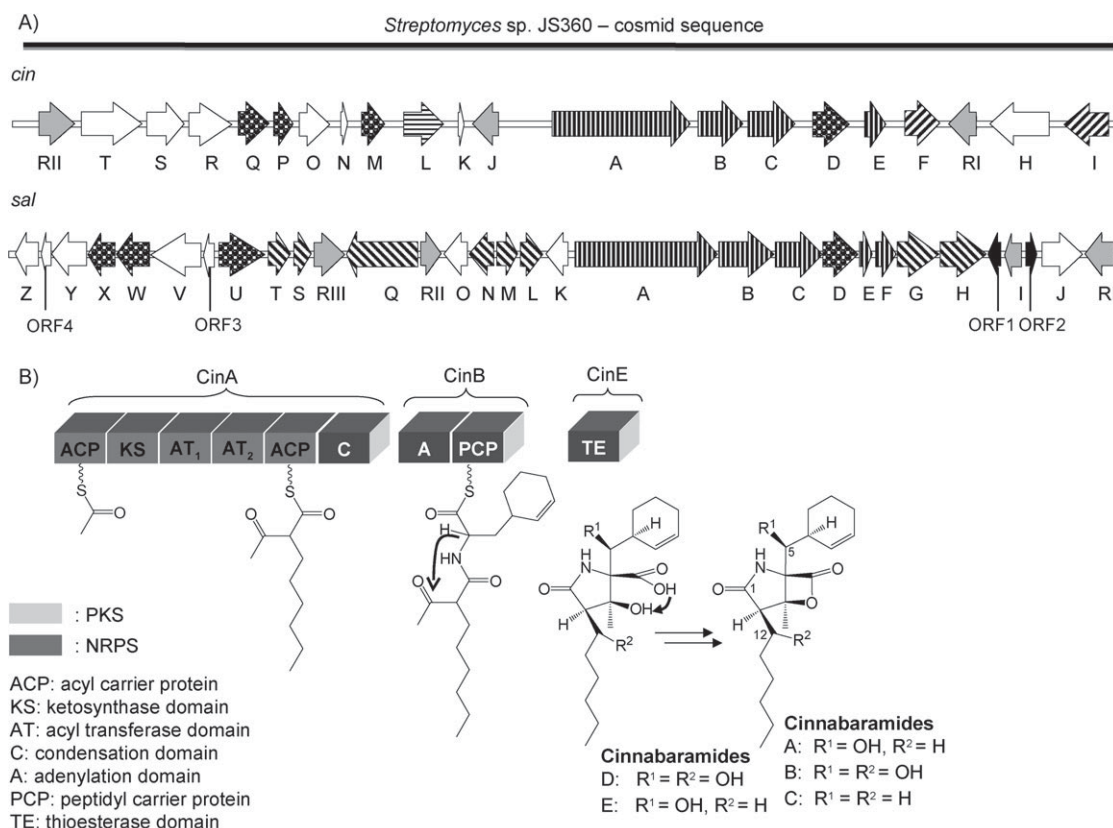
Because the *cinD* gene encodes a putative cytochrome P450 enzyme we suggest that this enzyme is involved in post-assembly-line hydroxylation of the compound to produce the β-hydroxycyclohex-2-enylalanine residue in cinnabaramides A, B, D, and E.

Administration of deuterated cyclohexenyl-serine to the culture of *Streptomyces* JS360 showed no incorporation of the compound into the cinnabaramides, which indicates that the cinnabaramide biosynthesis enzymes exhibit no tolerance toward alteration in the β-position of the amino acid precursor. Interestingly, the salinosporamide megasynthetase was shown to tolerate both secondary and tertiary carbons in the γ-posi-

tion of the amino acid precursor, but substitution in the β-position was also not accepted.^[12] The authors suggested that the latter substitution might directly affect the function of the cytochrome P450 enzyme (SalD) in the pathway.

A second copy of a cytochrome P450 is encoded by *cinL* in the upstream region of the *cin* gene cluster (Table 1). Because *cinL* shows no counterpart in the salinosporamide pathway we suggest that the corresponding protein plays a role in hydroxylation of cinnabaramides B and D at C12, whereas *cinD* is responsible for production of the C5 hydroxy functionality in A, B, and D (Scheme 2B).

After identifying the biosynthetic pathway of cinnabaramides, we predicted that the hexyl side chain of the compound might originate from a hexylmalonyl-CoA extender unit (Schemes 2B and 3A). This assumption was based on previous observations in other biosynthetic pathways involving unusual

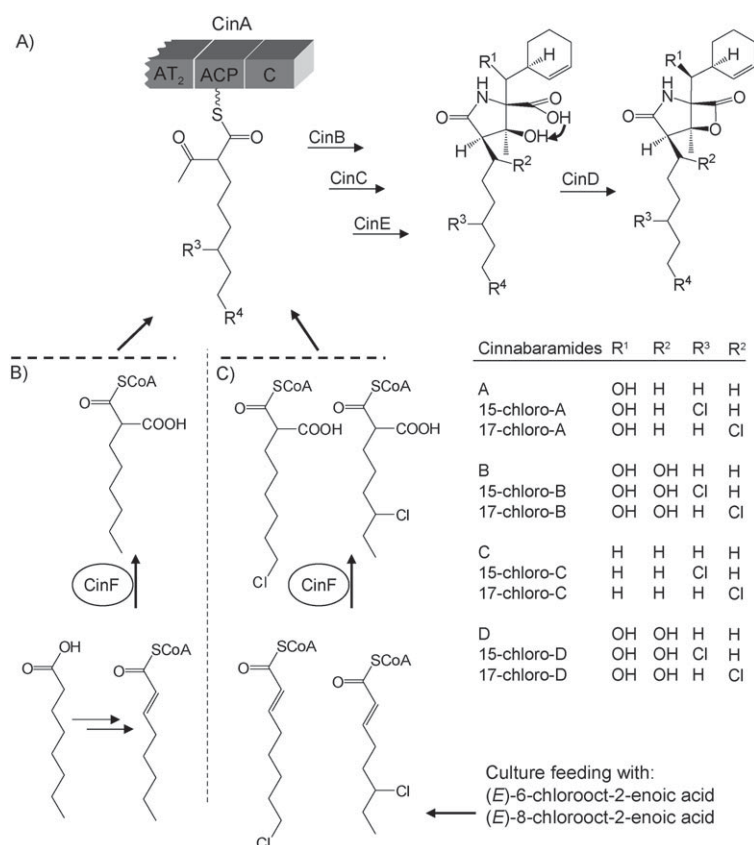


Scheme 2. A) Organization of the *cin* biosynthetic gene cluster from *Streptomyces* sp. JS360 and of the *sal* biosynthetic gene cluster from *Salinospora tropica*. Genes of similar function in both salinosporamide and cinnabaramide biosynthesis: genes involved in construction of the core structure γ -lactam- β -lactone bicyclic ring (□), genes involved in assembly of the nonproteinogenic amino acid β -hydroxycyclohex-2'-enylalanine (▨), genes involved in regulation and resistance (▩). Specific genes existing only in the cinnabaramide biosynthesis cluster: genes involved in the 2-carboxyl-CoA pathway (▤) and genes involved in hydroxylation of C12 in cinnabaramides B and D (▥). Specific genes existing only in the salinosporamide biosynthesis cluster: partial transposases (■), genes involved in the chloroethylmalonyl-CoA pathway (▧). Genes of unknown function in biosynthesis of cinnabaramides and salinosporamides (□). B) Proposed PKS/NRPS biosynthesis pathway for production of cinnabaramides.

extender units of polyketides and predominantly the work relating to the ethylmalonyl-CoA pathway in *Rhodobacter sphaeroides*. In this pathway, reductive carboxylation of (*E*)-crotonyl-CoA by a crotonyl-CoA carboxylase/reductase (Ccr) enzyme^[30] with use of the purified enzyme *in vitro* has been described. Similar biochemical steps have also been reported to be involved in the production of the PKS extender units ethylmalonyl-CoA and chloroethylmalonyl-CoA in the salinosporamide biosynthesis pathway.^[31] Another example has been reported in the biosynthesis of the NRPS/PKS macrolide thuggacin, isolated from *Sorangium cellulosum* So ce895, in which it was proposed that the hexyl side chain of the compound originates from extension by the unusual PKS extender hexylmalonyl-CoA, which is thought to be generated from octenyl-CoA by the Ccr homologue TgaD.^[21] In the cinnabaramide gene cluster we found a gene (*cinF*) encoding for a protein showing 90% sequence identity (at the protein level) to a putative Ccr from *Streptomyces hygroscopicus* (Accession No. AAR32 675). However, sequence identity to TgaD is very low. We inactivated *cinF* by insertion of the inactivation plasmid pCN1 into the genome of *Streptomyces* sp. JS360 through homologous recombination, resulting in a mutant incapable of producing cinnabaramides. We thus hypothesize that CinF catalyzes the reductive carboxy-

lation of octenoyl-CoA or -ACP, leading to 2-carboxy-octanoyl-CoA or 2-carboxy-octanoyl-ACP, respectively (Schemes 3B and 4). Notably, the octenoyl-CoA and octenoyl-ACP precursors to the proposed building blocks are intermediates furnished either from the fatty acid degradation or the fatty acid biosynthesis pathways.

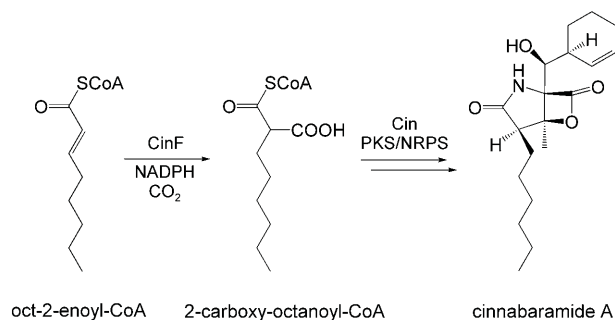
We speculated that the corresponding free acids obtained from administration to the growth media might be activated in the cell and thus enable mutasynthesis of analogues, and consequently next turned our attention towards the production of chlorinated cinnabaramides. A precursor-directed biosynthesis strategy was applied by feeding of the bacterial culture with chlorinated precursors. Accordingly, (*E*)-6-chlorooct-2-enoic acid, (*E*)-8-chlorooct-2-enoic acid, and their corresponding *N*-acetylcysteamine (NAC) thioesters were chemically synthesized (Figure S1) and added to the culture of the bacteria (Scheme 3C). Organic extracts of the culture supernatants were analyzed by HPLC-MS, which unambiguously showed incorporation of these precursors into cinnabaramide A (and into other cinnabaramide derivatives; data not shown) bearing a chlorine atom (Figure 1). After scale-up of the precursor feeding experiment, these compounds were also isolated to purity by preparative HPLC (details provided in the Supporting Infor-



Scheme 3. A) Biosynthetic pathway producing the cinnabaramide hexyl side chain. B) Putative role of CinF (carboxylase/reductase) in the pathway. C) Culture feeding with 6- or 8-chlorooct-2-enoic acid as an analogue precursor for the biosynthesis of chlorinated cinnabaramides.

mation) in order to allow their unambiguous structure elucidation.

1D and 2D NMR spectroscopy for structure elucidation (Table S4 and Figure S4), combined both with LC-MS data and with high-resolution ESI-MS (for comparison with cinnabaramide A; Scheme 4 and Table S2), identified the structures as novel chlorinated cinnabaramides with chlorine atoms at C15 or C17. From these results, we reason that the cinnabaramide PKS extender module is able to accept chlorinated substrates and to incorporate them into the assembly line. No production



Scheme 4. The presumed mechanism for production of the PKS extender unit hexylmalonyl-CoA, which is required for the formation of the hexyl side chain in cinnabaramides.

of chlorinated cinnabaramides was found when the bacterial culture was supplemented with (*E*)-4-chlorooct-2-enoic acid, so we postulate that the presence of the chlorine atom at this position might influence the substrate in such a way that it cannot be properly positioned in the CinA-AT₂ active site. Alternatively, CinF might not accept this substrate for reductive carboxylation.

Novel cinnabaramide A derivatives were isolated and screened against the human 20S proteasome (more precisely, against the proteolytic subunits with chymotrypsin-like (CT-L), trypsin-like (T-L), and caspase-like (C-L) activities) in vitro. As shown in Table 2, salinosporamide A has a higher inhibitory potential towards all three proteolytic subunits than the tested cinnabaramides. The IC₅₀ values determined for the CT-L activity were all in the low nanomolar range, as reported previously,^[9,32–34] although the chloro derivatives are more potent than the parent compound. With regard to inhibition at the T-L active site, 15-chlorocinnabaramide A showed the highest activity amongst the derivatives. The additional chlorine might serve as a leaving group and in this way enhance cinnabaramide's potency for inhibiting the β2 proteolytic subunit, as described earlier for inhibition at the CT-L site by salinosporamide A.^[17] In the case of C-L activity, cinnabaramide A and its analogues exhibited IC₅₀ values in the high nanomolar range.

As shown in Table 3, the cytotoxic activities of the novel 15-chloro derivative were superior to those of cinnabaramide A, whereas the 17-chloro derivative showed enhanced bioactivity only towards human myeloma RPMI-8226 cells. These findings are consistent with the results from the proteasome inhibition experiments. In general, the cytotoxic effects are still significantly weaker than those of salinosporamide A.

In addition to proteasome inhibition, the cinnabaramides showed strong antifungal activity. Although *Botrytis cinerea* was found to be insensitive towards the compounds, both *Mucor plumbeus* and *Pyricularia oryzae* were affected by low concentrations of the cinnabaramides, as well as by salinosporamide A (as shown in Table S3). Interestingly, the last compound showed the weakest antifungal activity, despite being much more potent than the cinnabaramides against the concurrently tested cancer cell lines. The chlorinated cinnabaramides and the cyclopentyl analogue of cinnabaramide A, however, were slightly less active than their parent compound.

Experimental Section

General microbiological and molecular biology methods: The cinnabaramide producer JS360 was grown in Q6 medium, which led to maximal production rate as described previously.^[9] Cells of *Escherichia coli* were grown in Luria–Bertani (LB) medium at 37 °C. dNTPs, restriction enzymes, T4 DNA ligase, isopropyl β-D-thiogalactopyranoside (IPTG), and X-gal were purchased from Fermentas. Ammonium persulfate and gel electrophoresis materials were ob-

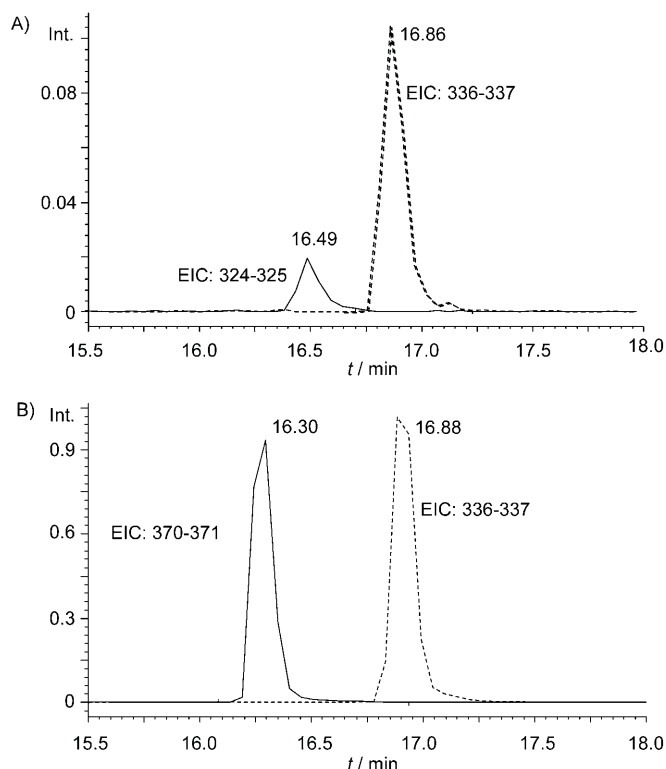


Figure 1. HPLC-MS analyses with extracted ion chromatograms (EICs) for m/z values of cinnabaramide A ($t_R = 16.9$ min; dashed line) and new cinnabaramide A derivatives. A) Feeding of strain JS360 with cyclopentyl-alanine resulting in the production of cyclopentyl-alanine cinnabaramide A ($t_R = 16.49$ min). B) Feeding of JS360 with 6- or 8-chlorooct-2-enoic acid or their SNAC derivatives, resulting in the production of cinnabaramide A analogues chlorinated at C15 or C17 ($t_R = 16.30$ min), respectively. The two chlorinated compounds had identical retention times on the applied system and showed similar levels of production.

Table 2. Inhibition of the $\beta 1$, $\beta 2$, and $\beta 5$ proteolytic subunits of human 20S proteasome by salinosporamide A and by cinnabaramide A and its analogues.

Compound	C-L ($\beta 1$)	IC ₅₀ [nM] ^[a] T-L ($\beta 2$)	CT-L ($\beta 5$)
salinosporamide A	32.7 ± 14.7	18.7 ± 10.5	3.1 ± 4.8
cinnabaramide A	248.6 ± 8.4	216.0 ± 12.3	11.9 ± 7.4
17-chlorocinnabaramide A	265.8 ± 10.7	442.5 ± 6.8	8.8 ± 5.5
15-chlorocinnabaramide A	402.9 ± 10.5	88.9 ± 8.3	9.3 ± 5.9
cyclopentyl-cinnabaramide A	301.2 ± 7.8	301.2 ± 9.1	64.9 ± 5.3

[a] IC₅₀ values were obtained by linear regression of kinetics measurements and secondary hyperbolic regression.

tained from Roth, Germany. Oligonucleotides were purchased from Sigma. Standard methods for DNA isolation and manipulation were used.^[36,37] DNA fragments were isolated from agarose gels by use of the NucleoSpin Extract gel-extraction kit (Macherey-Nagel). Southern genomic library analysis was performed with the DIG DNA labeling and detection kit (Roche). Hybridization was carried out at 42 °C for homologous probes with stringent washing at 68 °C. For heterologous probes, hybridization was performed at 42 °C with stringent washing at 60 °C. PCR reactions were performed with Taq DNA polymerase (Fermentas) or Pfu polymerase

Table 3. Cytotoxicities of salinosporamide A and of cinnabaramide A and its analogues towards human tumor cell lines^[a] as evaluated by MTT assay^[b] or by PI staining^[c].

Compound	HCT-116 ^[b]	RPMI-8226 ^[c] IC ₅₀ [μ M] ^[d]	SW480 ^[b]
salinosporamide A	0.047 ± 0.003	0.014 ± 0.001	0.013 ± 2.3 × 10 ⁵
cinnabaramide A	5.906 ± 0.301	1.940 ± 0.490	0.931 ± 0.141
17-chlorocinnabaramide A	7.286 ± 0.050	0.841 ± 0.190	0.977 ± 0.011
15-chlorocinnabaramide A	2.376 ± 0.433	0.341 ± 0.001	0.457 ± 0.007
cyclopentylcinnabaramide A	16.893 ± 0.235	n.d.	n.d.

[a] HCT-116: colon carcinoma. RPMI-8226: myeloma. SW480: colon adenocarcinoma. [d] Values represent the average of two independent measurements. Incubation times: [c] 4 d, or [b] 5 d. The observed IC₅₀ values were of the same order of magnitude as those previously reported.^[10,33,35] n.d.: not determined.

(Stratagene). Conditions for amplification with an Eppendorf Mastercycler were as follows: denaturation, 30 s at 95 °C; annealing, 30 s at 48–65 °C; extension, 45 s at 72 °C; 30 cycles and a final extension for 10 min at 72 °C. PCR products were purified by use of the High Pure PCR Product Purification kit (Boehringer Mannheim).

Construction of *Streptomyces* JS360 genomic cosmid library: Chromosomal DNA of the bacteria was prepared as described previously.^[37] Partial digestion of the bacterial genome was performed with *Sau*3AI enzyme to provide fragments of about 35–40 kbp. The fragments were dephosphorylated and ligated into pOJ436 cosmid pretreated with *Pvu*II, dephosphorylated, and restricted with *Bam*HI. The ligation mixture was packaged with Gigapack III Gold (Stratagene) packaging extract and transduction of the resulting phages into *E. coli* HS966 created a genomic library of 2304 clones. For colony hybridization, the colonies were transferred twice onto a 22.2 × 22.2 cm nylon membrane with a QPix (Hampshire, UK) colony picker.^[38]

Screening of the cosmid library for the cinnabaramide gene cluster: From the chemical structures of the cinnabaramides we expected involvement of a PKS/NRPS gene cluster in biosynthesis of the compound. The bacterial cosmid library was therefore first screened for PKS genes by hybridization with a homologous ketosynthase (KS) probe generated from the bacterial genome under low-stringency conditions, with use of the degenerate oligonucleotides KS1UP and KSD1.^[39] Furthermore, an adenylation probe (NRPS) was generated from the bacterial genome with the aid of the degenerate primers NRPS-A1-up: (5'-CGG CTC CAC CGG CAC [ACGT]CC [ACGT]AA [AG]GG [ACGT]G-3') and NRPS-H1-dn: (5'-CGG CCG AGG TCG CC[ACGT] GT[ACGT] C[GT][AG] TA-3'), and used for hybridization of the cosmid library. Seven of the KS-positive cosmids were positively cross-hybridized with the NRPS probe and were subjected to PCR analysis for amplification of KS and NRPS-adenylation fragments. The PCR fragments were cloned into pCR2.1TOPO vector (Invitrogen) and sequenced. The PCR fragments generated from cosmid C:A17 showed significant homology to KS and adenylation-domain sequences from the salinosporamide gene cluster. After cosmid end-sequencing from the T3 and the T7 termini of the pOJ436, the cosmid was thus sequenced on both strands, by use of a shotgun library as described previously,^[40] and the sequence was introduced into the EMBL nucleotide database (accession number: FR687018).

Bioinformatics analysis: All sequence similarity searches were carried out at the amino acid level in the Gene Bank database with use of the BLAST program (release 2.0). Amino acid and nucleotide

sequences were aligned with the aid of the Vector NTI advance 11 and Geneious 4.8.4 software packages (Invitrogen, USA, and Bio-matters, New Zealand, respectively).

Inactivation in the cinnabaramide gene cluster: Inactivation of *cinF*, encoding octenoyl-CoA reductase/carboxylase (Ocr), and of *cinQ*, encoding a putative branched-chain amino acid aminotransferase, was performed by integration of the inactivation plasmids pCN1 and pCN2, respectively, into the *Streptomyces* JS360 genome. Internal fragments of the *cinF* (1044 bp) and *cinQ* (847 bp) with frame shift mutation in the 5'-end sequence were generated by PCR by use of the oligonucleotides CinF-up (5'-ATG TAG CAG CTT GTC CCG AGC CGT A-3') and CinF-dn (5'-TTG TCG AAG GTG TGC AGG TAG-3'), and of CinQ-up (5'-CAG TGA CAG CGC GAC CCT CAT C-3') and CinQ-dn (5'-CAC GTC CCT GCC GGA GAG-3'; inserted nucleotides highlighted in bold). Next, the PCR products were cloned into the pCR2.1-TOPO vector (Invitrogen) for sequence analysis, and excised by use of the restriction enzymes EcoRV and HindIII. The resulting inserts were then subcloned into EcoRV and HindIII cloning sites of the conjugative plasmid pKC1132,^[41] creating the inactivation plasmids pCN1 and pCN2. The plasmids were then introduced into *E. coli* ET12567/pUZ8002. Appropriate transformants were then used for conjugation with JS360 spores as described previously.^[42] Plates containing putative exconjugants were overlaid with apramycin (60 µg mL⁻¹) and nalidixic acid (25 µg mL⁻¹). Because pKC1132 cannot replicate in *Streptomyces*, apramycin-resistance strains could be produced only by homologous recombination between the conjugative plasmids and the genes in the JS360 genome. To confirm integration of the plasmid into the genome, and thus disruption of the genes, PCR analysis with the aid of the primer pTOPO-Hind (5'-GAG CTC GGA TCC ACT AGT-3') in combination with each of CinF-out (5'-ATC ATT TCA AGG TAC TTC-3'), CinQ-out (5'-CTC GGT GTA CAG CTC GCG-3'), and CinQ-out (5'-TTC TCA CGG AAG TCG ATC T-3') of the bacterial genome located upstream or downstream regions of the cloned fragments was performed. PCR fragments 1163 bp and 980 bp were only amplified from genome of the mutants JS360-*cinF*⁻ and JS360-*cinQ*⁻.

Feeding experiments: JS360 was grown in YMG medium at 30 °C for 24 h and was then transferred to Q6 medium (25 mL) supplemented variously with DL-3-cyclopentyl-alanine, deuterated cyclohexenyl-serine, 4-chlorooct-2-enoic acid, 4-chlorooct-2-enoyl-SNAC, 6-chlorooct-2-enoic acid, 6-chlorooct-2-enoyl-SNAC, 8-chlorooct-2-enoic acid, 8-chlorooct-2-enoyl-SNAC, crotonic acid, pent-2-enoic acid, hex-2-enoic acid, or hept-2-enoic acid (all 0.1 mM). These compounds were each added in three portions to the bacterial culture after 0, 6, and 12 h, and further incubated for 24 h. To investigate for production of cinnabaramides or their derivatives, culture supernatant (20 mL) was extracted three times with ethyl acetate (20 mL). After removal of the solvent in vacuo, the remainder was dissolved in methanol (1 mL) and 5 µL of the solution was injected for HPLC-MS analysis.

Synthesis of precursors for feeding experiments

A) Synthesis of (*E*)-4-chlorooct-2-enoic acid: (*E*)-Oct-3-enoic acid (5 g, 35.2 mmol, Sigma—Aldrich) was added dropwise at RT under nitrogen to a solution of phenylselenium chloride (3.4 g, 17.6 mmol) in dry acetonitrile (200 mL). Molecular sieves (4 Å) were added, followed over 20 h by a solution of *N*-chlorosuccinimide (5.15 g, 38.5 mmol) in dry acetonitrile (100 mL). After addition, the mixture was stirred at RT for 2 d. The solvent was removed under reduced pressure and the residue was diluted with diethyl ether (100 mL). The solid formed was removed by filtration and the filtrate was washed with water (4 × 25 mL). The organic layer was

dried over MgSO₄ and the solvent was removed in vacuo. After purification by flash column chromatography over silica gel (with cyclohexane/ethyl acetate as mobile phase), the target compound (1.6 g) was obtained. Because NMR analysis indicated that the sample still contained oct-2-enoic acid (about 10%) as a side product, the compound was further purified by distillation at 0.1 mbar/175 °C to yield 4-chlorooct-2-enoic acid (93% purity by ¹H NMR, 1.3 g, 21%, Figures S2 and S3).

B) Synthesis of (*E*)-6-chlorooct-2-enoic acid and its corresponding SNAC derivative

Methyl 4-chlorohexanoate: A solution of γ-ethylbutyrolactone (25 g, 0.22 mol) in dry MeOH (250 mL) was saturated with HCl (g). The reaction mixture was stirred at RT in a sealed tube for 4 d. Excess HCl was removed and the reaction mixture was concentrated to dryness. The crude product was diluted with water (100 mL) and extracted with CH₂Cl₂ (3 × 75 mL). The organic layer was washed with a saturated solution of NaHCO₃, dried over MgSO₄, and concentrated to dryness. The product was purified by distillation (bp = 72–74 °C at 11 mbar) to yield 20.4 g (60%).

4-Chlorohexanal: A solution of DIBAL-H in CH₂Cl₂ (1 M, 130 mL) was added at –78 °C to a solution of methyl 4-chlorohexanoate (20.4 g, 0.12 mol) in dry CH₂Cl₂ (112 mL). The reaction mixture was stirred at this temperature for 1 h, after which it was poured into ice/conc. HCl (140 g/28 mL). The aqueous layer was extracted with CH₂Cl₂ (2 × 50 mL). The combined organic layers were dried over MgSO₄ and the solvent was removed under reduced pressure. The desired product (17.8 g) was obtained and used without further purification.

(*E*)-6-Chlorooct-2-enoic acid: A mixture of 4-chlorohexanal (17.8 g, 0.12 mol) and malonic acid (16.5 g, 0.16 mol) in dry pyridine (37 mL) was stirred at 30 °C for 4 d, at 60 °C for 1 h, and finally at 95 °C for 1 h. The reaction mixture was poured into ice/6 M HCl solution (200 mL) and extracted with Et₂O (3 × 100 mL). The organic layer was extracted with NaOH solution (3 M, 2 × 50 mL). The aqueous layer was acidified with HCl (10%) to pH 2 and then extracted with Et₂O (3 × 75 mL). The new organic layer was dried over MgSO₄ and the solvent was removed under reduced pressure to yield the desired compound (16.3 g, 77% over two steps), which was used without further purification.

Corresponding NAC thioester: EDAC (3.33 g, 17.4 mmol) was added to a solution of (*E*)-6-chlorooct-2-enoic acid (2 g, 11.3 mmol) and *N*-acetylcysteamine (1.64 g, 13.7 mmol) in dry DMF (115 mL). The reaction mixture was stirred at RT overnight. HCl solution (10%) was added (50 mL), and the mixture was diluted with water (100 mL) and extracted with CH₂Cl₂ (2 × 150 mL). The organic layer was washed with brine, dried over MgSO₄, and concentrated to dryness. The product was purified by flash column chromatography (silica gel, cyclohexane/ethyl acetate) and preparative HPLC (H₂O/acetonitrile). The desired compound (295 mg, 93% yield) was obtained (Figures S2 and S3).

C) Synthesis of (*E*)-8-chlorooct-2-enoic acid and its corresponding SNAC derivative

6-Chlorohexanal: A solution of sodium bichromate dihydrate (25 g, 0.084 mol) in water (312 mL) was added dropwise at reflux to a solution of 6-chlorohexanol (25 g, 0.18 mol) and H₂SO₄ (25 mL) in water (156 mL), and at the same time the same amount of water was distilled off. More water was added (30 mL) and simultaneously distilled off from the reaction mixture. The distillate was extracted with Et₂O (3 × 100 mL). The organic layer was dried over MgSO₄ and concentrated to dryness. The product obtained (17.8 g) was a

mixture of the starting alcohol and the desired corresponding aldehyde (1:1), and it was used without further purification for the next step.

(E)-8-Chlorooct-2-enoic acid: The mixture of 6-chlorohexanol and 6-chlorohexanal (1:1, 17.8 g, about 7.9 g/66 mmol of the aldehyde), together with malonic acid (8.25 g, 79.2 mmol) in dry pyridine (18 mL), was stirred at 30 °C for 4 d, at 60 °C for 1 h, and finally at 95 °C for 1 h. The reaction mixture was poured into ice/6 M HCl solution (100 mL) and extracted with Et₂O (3 × 50 mL). The organic layer was extracted with NaOH solution (3 M, 2 × 25 mL). The aqueous layer was acidified with HCl (10%) to pH 2 and extracted with Et₂O (3 × 50 mL). The new organic layer was dried over MgSO₄ and the solvent was removed under reduced pressure to yield the desired compound (10.8 g), which was used without further purification.

Corresponding NAC thioester: EDAC (3.12 g, 16.3 mmol) was added to a solution of (E)-8-chlorooct-2-enoic acid (1.9 g, 10.8 mmol) and N-acetylcysteamine (1.64 g, 12.9 mmol) in dry DMF (108 mL). The reaction mixture was stirred at RT overnight. HCl solution (10%, 50 mL) was added, and the mixture was diluted with water (100 mL) and extracted with CH₂Cl₂ (2 × 150 mL). The organic layer was washed with brine, dried over MgSO₄, and concentrated to dryness. The product was purified by flash column chromatography (silica gel, cyclohexane/ethyl acetate) and afterward by preparative HPLC (H₂O/acetonitrile). The desired compound (460 mg, about 15%) was obtained (Figures S2 and S3).

D) Investigation of the synthesis of (E)-5-chlorooct-2-enoic acid: Unfortunately, attempts to synthesize (E)-5-chlorooct-2-enoic acid by a strategy similar to that successfully applied for 6- and 8-chlorooct-2-enoic acids failed. Spontaneous HCl elimination of the desired product, leading to the corresponding octa-2,4-dienoic acid is unavoidable.

Ethyl 3-chlorohexanoate: Benzotriazole (29.5 g) was added to a solution of SOCl₂ (18.06 mL) in dry CH₂Cl₂ (50 mL). More CH₂Cl₂ was added until 165.4 mL of a 1.5 M solution of SOCl₂/benzotriazole were obtained. This solution was added dropwise to a solution of ethyl 3-hydroxyhexanoate (15.9 g, 99.2 mmol) in dry CH₂Cl₂ (500 mL), and the mixture was stirred at RT overnight. The solid formed was removed by filtration and the filtrate was washed with water (2 × 200 mL), aq. Na₂CO₃ (10%, 150 mL), and NaHCO₃ (150 mL). The organic layer was dried over MgSO₄ and concentrated to dryness. The product was purified by flash column chromatography (silica gel, cyclohexane/ethyl acetate) to give the target compound (5.5 g). Starting ethyl 3-hydroxyhexanoate (10 g) was recovered (yield 83%, conversion 37%).

3-Chlorohexanal: A solution of DIBAL-H in CH₂Cl₂ (1 M, 33.8 mL) was added at –78 °C to a solution of ethyl 3-chlorohexanoate (5.5 g, 30.8 mmol) in dry CH₂Cl₂ (28 mL). The reaction mixture was stirred at this temperature for 1 h, after which it was poured into 40 g ice/7.5 mL conc. HCl solution. The aqueous layer was extracted with CH₂Cl₂ (2 × 20 mL). The combined organic layer was dried over MgSO₄ and the solvent was removed under reduced pressure. The desired product (5 g, theory: 100% = 4.14 g) was obtained and used without further purification.

(E)-5-Chlorooct-2-enoic acid/(2E,4E)-octa-2,4-dienoic acid: A mixture of 3-chlorohexanal (4.1 g, about 30 mmol) and malonic acid (4.54 g, 43.4 mmol) in dry pyridine (10 mL) was stirred at 30 °C for 4 d, at 60 °C for 1 h, and finally at 95 °C for 1 h. The reaction mixture was poured into ice/6 M HCl solution (100 mL) and extracted with Et₂O (3 × 25 mL). The organic layer was extracted with NaOH

solution (3 M, 2 × 15 mL). The aqueous layer was acidified with HCl (10%) to pH 2 and extracted with Et₂O (3 × 25 mL). The new organic layer was dried over MgSO₄ and the solvent was removed under reduced pressure. The crude reaction product was purified by flash column chromatography (silica gel, cyclohexane/ethyl acetate) to yield (2E,4E)-octa-2,4-dienoic acid as reaction product. Any attempts to avoid HCl elimination and to obtain the desired (E)-5-chlorooct-2-enoic acid failed (Figures S2 and S3).

Analytical methods and structure elucidation: HPLC-MS was performed with a HPLC-DAD system (Agilent 1100) coupled to an HCT ultra ESI-MS ion trap apparatus (Bruker Daltonics) operating in positive ionization mode. Separation was achieved by use of a Luna RP-C18 column with a solvent system consisting of a water to acetonitrile gradient. For separation of cinnabaramides and their derivatives and of oct-2-enoyl-SNAC and 2-carboxyl-octenoyl-SNAC esters, a solvent system based on water (A)/acetonitrile (B) containing formic acid (0.1%) was used. Gradient: 0–2 min 5% B; 2–22 min linear from 5% to 95% B; 22–25 min isocratic at 95% B; 25–27 min linear from 95–5% B.

High-resolution measurements were performed with an Accela UPLC-system (Thermo-Fisher) coupled to an LTQ-Orbitrap (linear trap-FT-Orbitrap combination) operating in positive ionization mode with a Waters BEH-C18 column with a solvent system consisting of a water (A)/acetonitrile (B) gradient containing formic acid (0.1%).

NMR spectra were recorded in [D₆]DMSO with a Bruker DRX 500 spectrometer at 303 K, operating at 500.13 MHz proton frequency. The solvent peak was used as internal reference ($\delta_{\text{H}} = 2.50$ ppm, $\delta_{\text{C}} = 39.5$ ppm). Chemical shifts are given in ppm, coupling constants in Hertz. In the cases of the synthetic precursors used for the mutasynthesis projects, NMR spectra were obtained with a Bruker AVANCE spectrometer at 293 K in CDCl₃ (300 MHz proton frequency), which was also used as internal reference ($\delta_{\text{H}} = 7.26$ ppm). Analytical HPLC-MS was performed with an Agilent 1100 HPLC system coupled with an Agilent DAD detector, an evaporating light scattering detector (ELSD), and an LCT mass spectrometer (Micromass, Manchester, UK) operating in positive and negative ESI modes. A Waters Symmetry C18 column (3.5 μm , 150 × 2.1 mm) and a linear gradient from 0 to 100% MeCN (0.1% HCOOH, flow rate 0.4 mL min^{–1}) in 21 min were used. HR-ESIMS data were obtained with a Bruker MicroTOF instrument, coupled with an HPLC system as described above and with use of sodium formate as internal reference. Optical rotation values were measured in MeOH with a Schmidt + Haensch Polartronic HH8 polarimeter; concentrations are given in g per 100 mL. Structure elucidation of the novel cinnabaramide derivatives was achieved by thorough interpretation of 1D and 2D NMR spectra, combined with LC-MS data including extracted UV as well as positive and negative mode ESI spectra. Cinnabaramide A was used for comparison. Additionally, the molecular formulas and elemental compositions of all cinnabaramide derivatives were confirmed by high-resolution ESI-MS. The NMR dataset used consisted of 1D proton and carbon, ¹H,¹H gCOSY, ¹H,¹³C gHSQC, and ¹H,¹³C gHMBC spectra (Figures S3 and S4).

Inhibition assay with human 20S proteasome: All chemicals were of reagent grade and buffers were made with deionized distilled water. Inhibition experiments were performed in HEPES buffer (25 mM, pH 7.5) containing EDTA (0.5 mM), Triton-X100 (0.05%, v/v), and SDS (0.001%, w/v). Inhibition was monitored by measuring the release of AMC-coupled substrate peptides for the C-L, CT-L, and T-L activities, respectively. Compounds were pre-incubated for 15 min at 37 °C in serial dilutions from 0.03 to 3000 nM with

human 20S proteasome (2 nM, Biomol) dissolved either in buffer A (HEPES (pH 8.0, 20 mM), EDTA (0.5 mM), SDS (0.035%, w/v)) for CT-L and C-L activities or in buffer B (HEPES (pH 8.0, 20 mM), EDTA (0.5 mM)) for T-L activity. Boc-LRR-AMC-2HCl, suc-LLVY-AMC, and Z-LLE-AMC (Bachem) were made up to 50 mM DMSO stocks in order to monitor T-L, CT-L, and C-L activities, respectively. Reactions were initiated by addition of AMC substrates to final concentrations of 2 μ M. The emission at 460 nm (5 nm cutoff; excitation at 380 nm) was measured with a Spectramax M5 fluorescence reader (Molecular Devices) in black 96-well plates (BD Falcon) at 37 °C over 90 min. The measured AMC production was in the linear range of an AMC standard curve.

MTT assay: Cells were seeded at 6×10^3 cells per well on 96-well plates (Corning CellBind®) in complete medium (180 μ L) and directly treated with compounds dissolved in methanol in a serial dilution. After 5 d incubation, MTT [3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl tetrazolium bromide, 20 μ L of 5 mg mL⁻¹ stock] in PBS (phosphate-buffered saline, pH 7.4) was added to each well and further incubation at 37 °C was carried out for 2 h. The medium was then discarded and cells were washed with PBS before addition of propan-2-ol/10 N HCl (250:1, 100 μ L) in order to dissolve formazan granules. The absorbance at 570 nm was measured with a microplate reader (EL808, Bio-Tek Instruments, Inc.), and cell viability was expressed as percentage relative to the corresponding methanol control. IC₅₀ values were obtained by sigmoidal curve fitting.

Cytotoxicity assay (propidium iodide staining): Cells were harvested from exponential phase cultures, counted, and plated in 96-well microtiter plates at cell densities of 10 000 to 80 000 for hematological cancer cell lines. After a 24 h recovery period to allow cells to resume exponential growth, culture medium (four control wells/plate) or culture medium with the test compound (10 μ L) was added with use of a liquid handling robotic system and treatment was continued for 4 d. Compounds were applied in half-log increments at ten concentrations in duplicate. For hematological cancer cell lines (growing in suspension), 30 min after addition of PI (propidium iodide, 7 μ g mL⁻¹), fluorescence (FU1) was measured to quantify dead cells. Triton X-100 (0.1%, v/v) was added next, resulting in permeability of all cells (viable + dead cells). After a second fluorescence measurement (FU2), the amount of viable cells was calculated by subtraction of FU1 from FU2. Growth inhibition/cytotoxicity is expressed as Test/Control $\times 100$ (% T/C). IC₅₀ values were calculated by nonlinear regression (log [inhibitor] vs response (% T/C)) with the aid of the analysis software GraphPad Prism, Prism 5 for Windows, version 5.01 (GraphPad Software, Inc.). For calculation of mean IC₅₀ values the geometric means were used.

Determination of antifungal activities: Minimal inhibitory concentrations (MICs) in the serial dilution assay were determined in 96-well microtiter plates in a manner similar to that described previously,^[43] with use of *Botrytis cinerea* (DSM 5144), *Pyricularia oryzae* (DSM 62938), and *Mucor plumbeus* (MUCL 49355) as test organisms. Spore suspensions of the fungi were prepared from well-grown YMG (glucose 0.4%, malt extract 1%, yeast extract 0.4%, pH 6.3) agar plates and the initial titer for the biological assay was adjusted to initial titers of 105 spores per mL YMG medium.

The chlorinated cinnabaramides arising from the precursor feeding experiments and the cyclopentyl analogue of cinnabaramide A were tested for comparison with cinnabaramide A, salinosporamide A, and actinomycin C complex. All test compounds were dissolved in methanol and supplied to the microtiter plates in serial

dilutions (100, 50, 25, 12.5, 6.3, 3.1, 1.6, and 0.8 μ g mL⁻¹). The MICs were determined after 24 h of incubation by OD measurement.

Acknowledgements

Research in R.M.'s laboratory was funded by the Bundesministerium für Bildung und Forschung (BMBF, FKZ: 0315385A). We gratefully acknowledge expert technical assistance by Andrea Rademacher and Dirk Müller (InterMed Discovery GmbH). Furthermore, we thank Oncotest (Freiburg, Germany) for determination of the cytotoxic activities against multiple myeloma cells.

Keywords: biosynthesis • mutasynthesis • natural products • precursor-directed biosynthesis • proteasome inhibitors

- [1] D. J. Newman, G. M. Cragg, K. M. Snader, *J. Nat. Prod.* **2003**, 66, 1022–1037.
- [2] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2007**, 70, 461–477.
- [3] J. Berdy, *J. Antibiot.* **2005**, 58, 1–26.
- [4] R. H. Baltz, *Nat. Biotechnol.* **2006**, 24, 1533–1540.
- [5] K. J. Weissman, P. F. Leadlay, *Nat. Rev. Microbiol.* **2005**, 3, 925–936.
- [6] K. J. Weissman, R. Müller, *ChemBioChem* **2008**, 9, 826–848.
- [7] C. T. Walsh, *Acc. Chem. Res.* **2007**, 40, 4–10.
- [8] S. A. Samel, M. A. Marahiel, L. O. Essen, *Mol. Biosyst.* **2008**, 4, 387–393.
- [9] M. Stadler, J. Bitzer, A. Mayer-Bartschmid, H. Müller, J. Buchholz, F. Gantner, H. V. Tichy, P. Reinemer, K. B. Bacon, *J. Nat. Prod.* **2007**, 70, 246–252.
- [10] R. H. Felting, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen, W. Fenical, *Angew. Chem.* **2003**, 115, 369–371; *Angew. Chem. Int. Ed.* **2003**, 42, 355–357.
- [11] L. L. Beer, B. S. Moore, *Org. Lett.* **2007**, 9, 845–848.
- [12] R. P. McGlinchey, M. Nett, A. S. Eustaquio, R. N. Asolkar, W. Fenical, B. S. Moore, *J. Am. Chem. Soc.* **2008**, 130, 7822–7823.
- [13] A. S. Eustaquio, S. J. Nam, K. Penn, A. Lechner, M. C. Wilson, W. Fenical, P. R. Jensen, B. S. Moore, *ChemBioChem* **2010**.
- [14] J. Adams, *Cancer Cell* **2004**, 5, 417–421.
- [15] B. S. Moore, A. S. Eustaquio, R. P. McGlinchey, *Curr. Opin. Chem. Biol.* **2008**, 12, 434–440.
- [16] T. A. M. Gulder, B. S. Moore, *Angew. Chem. Int. Ed.* **2010**, 49, 9346–9367.
- [17] M. Groll, R. Huber, B. C. Potts, *J. Am. Chem. Soc.* **2006**, 128, 5136–5141.
- [18] A. S. Eustaquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2009**, 106, 12295–12300.
- [19] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, 106, 3468–3496.
- [20] G. Yadav, R. S. Gokhale, D. Mohanty, *J. Mol. Biol.* **2003**, 328, 335–363.
- [21] K. Buntin, H. Irschik, K. J. Weissman, E. Luxemburger, H. Blöcker, R. Müller, *Chem. Biol.* **2010**, 17, 342–356.
- [22] H. Irschik, H. Reichenbach, G. Höfle, R. Jansen, *J. Antibiot.* **2007**, 60, 733–738.
- [23] H. Steinmetz, H. Irschik, B. Kunze, H. Reichenbach, G. Höfle, R. Jansen, *Chem. Eur. J.* **2007**, 13, 5822–5832.
- [24] D. W. Udway, L. Zeigler, R. N. Asolkar, V. Singan, A. Lapidus, W. Fenical, P. R. Jensen, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2007**, 104, 10376–10381.
- [25] G. L. Challis, J. Ravel, *FEMS Microbiol. Lett.* **2000**, 187, 111–114.
- [26] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, 97, 2651–2674.
- [27] S. Mahlstedt, E. N. Fielding, B. S. Moore, C. T. Walsh, *Biochemistry* **2010**, 49, 9021–9023.
- [28] M. Nett, B. S. Moore, *Pure Appl. Chem.* **2009**, 81, 1075–1084.
- [29] M. Nett, T. A. Gulder, A. J. Kale, C. C. Hughes, B. S. Moore, *J. Med. Chem.* **2009**, 52, 6163–6167.
- [30] T. J. Erb, I. A. Berg, V. Brecht, M. Müller, G. Fuchs, B. E. Alber, *Proc. Natl. Acad. Sci. USA* **2007**, 104, 10631–10636.
- [31] Y. Liu, C. Hazzard, A. S. Eustaquio, K. A. Reynolds, B. S. Moore, *J. Am. Chem. Soc.* **2009**, 131, 10376–10377.

- [32] W. Fenical, P. R. Jensen, M. A. Palladino, K. S. Lam, G. K. Lloyd, B. C. Potts, *Bioorg. Med. Chem.* **2009**, *17*, 2175–2180.
- [33] A. S. Eustaquio, B. S. Moore, *Angew. Chem.* **2008**, *120*, 4000–4002; *Angew. Chem. Int. Ed.* **2008**, *47*, 3936–3938.
- [34] R. R. Manam, K. A. McArthur, T. H. Chao, J. Weiss, J. A. Ali, V. J. Palombella, M. Groll, G. K. Lloyd, M. A. Palladino, S. T. Neuteboom, V. R. Macherla, B. C. Potts, *J. Med. Chem.* **2008**, *51*, 6711–6724.
- [35] K. A. Reed, R. R. Manam, S. S. Mitchell, J. Xu, S. Teisan, T.-H. Chao, G. Deyanat-Yazdi, S. T. Neuteboom, K. S. Lam, B. C. M. Potts, *J. Nat. Prod.* **2007**, *70*, 269–276.
- [36] J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY **2001**.
- [37] T. Kieser, M. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, The John Innes Foundation, Norwich, **2000**, pp. 613.
- [38] I. Dunham in *Genome Analysis: A Laboratory Manual: Cloning Systems* (Ed.: B. Birren), Cold Spring Harbor Laboratory Press, New York **1997**, pp. 1–86.
- [39] S. Beyer, B. Kunze, B. Silakowski, R. Müller, *Biochim. Biophys. Acta Gene Struct. Expression* **1999**, *1445*, 185–195.
- [40] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blöcker, G. Höfle, S. Beyer, R. Müller, *J. Biol. Chem.* **1999**, *274*, 37391–37399.
- [41] M. Bierman, R. Logan, K. Brien, E. T. Sena, R. N. Rao, B. E. Schoner, *Gene* **1992**, *116*, 43–49.
- [42] S. Hoyt, G. H. Jones, *J. Bacteriol.* **1999**, *181*, 3824–3829.
- [43] M. Stadler, D. N. Quang, A. Tomita, T. Hashimoto, Y. Asakawa, *Mycol. Res.* **2006**, *110*, 811–820.

Received: January 14, 2011

Published online on March 8, 2011

Supporting Information

© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2011

Mining the Cinnabaramide Biosynthetic Pathway to Generate Novel Proteasome Inhibitors

Shwan Rachid,^[a, b, c] Liujie Huo,^[a, b] Jennifer Herrmann,^[a, b] Marc Stadler,^[d] Bärbel Köpcke,^[d]
Jens Bitzer,^[d] and Rolf Müller*^[a, b]

cbic_201100024_sm_miscellaneous_information.pdf

Precursor-Directed Biosynthesis with *Streptomyces* JS360 and Isolation of Novel Cinnabaramid Derivatives

A) General methods

Precultures of cinnabaramide producer JS360 were done in YM medium (D-glucose 4 g/L; malt extract 10 g/L, yeast extract 4 g/L; pH 7.2). 150 mL medium 150 mL in 500 mL Erlenmeyer flasks (without baffles) were inoculated with 2 mL spore suspension of JS 360. The precultures were grown for 80 h at 28 °C on a rotary shaker at 240 rpm.

Main cultures were done in Q6 medium (D-Glucose 5 g/L, glycerine 20 g/L, cottonseed flour 10 g/L, CaCO₃ 1 g/L; no pH adjustment) supplemented with the precursor as described below. 150 mL medium in 500 mL Erlenmeyer flasks (without baffles) were inoculated with 2 mL of well-grown precultures. The main cultures were grown at 28 °C for 40 h on a rotary shaker at 240 rpm.

Workup and extraction: Mycelium and supernatant were separated by centrifugation. The wet mycelium was extracted twice with 300 mL acetone each (30 min, ultrasound). The combined and filtered organic phases were concentrated in vacuo until an aqueous residue (ca. 150 mL) was obtained. The aqueous residue was extracted twice with 300 mL ethyl acetate each. The combined organic phases were dried over Na₂SO₄ and concentrated to dryness under reduced pressure to obtain a crude extract.

Preparative HPLC was done with an Inertsil ODS-3 column (10 µm, 30 x 250 mm) using a water (A) / acetonitril (B) gradient with 0.1% TFA (40-50% B in 15 min. 50-65% B in 10 min, 65% B isocratic for 10 min, 65-75% B in 10 min, 75% B isocratic for 10 min, 75-100% B in 30 min; flow rate 10 mL/min). The UV signal was detected at 210 and 254 nm.

B) Production and isolation of cyclopentyl-cinnabaramide A

Q6 medium was supplemented with 0.5 mM 3-cyclopentyl-DL-alanine. After 12 h of growth, the cultures (20 flasks with 150 mL) were fed with additional 0.5 mM 3-cyclopentyl-DL-alanine. The crude mycelium extract (812 mg) obtained as described above was split into 5 aliquots, which were successively subjected to preparative HPLC to yield 2.3 mg of cyclopentyl-cinnabaramide A.

C) Production and isolation of 15-chlorocinnabaramide A

The strain JS360 was cultivated in Q6 medium (15 flasks with 150 mL). After 14 h, the medium was supplemented with 0.75 mM (*E*)-6-chloro-oct-2-enoic acid (addition of 100 µL / flask of a solution of 264 mg precursor dissolved in 1 mL MeOH). After another 4 h, additional 0.75 mM of precursor were added to the growing culture. The crude mycelium extract (1830 mg) obtained as described above was split into 5 aliquots, which were successively subjected to preparative HPLC to yield 6.3 mg of 15-chlorocinnabaramide A.

D) Production and isolation of 17-chlorocinnabaramide A

The strain JS360 was cultivated in Q6 medium (10 flasks with 150 mL). After 14 h, the medium was supplemented with 1.0 mM (*E*)-8-chloro-oct-2-enoic acid (addition of 100 µL / flask of a solution of 264 mg precursor dissolved in 1 mL MeOH). The crude mycelium extract (737 mg) obtained as described above was subjected to preparative HPLC to yield 4.3 mg of 17-chlorocinnabaramide A.

E) Feeding of (*E*)-4-chlorooct-2-enoic acid

With the aim to obtain 13-chlorocinnabaramide A, strain JS360 was cultivated as described above but with Q6 nutrient medium supplemented with 1.5 mM (*E*)-4-chloro-oct-2-enoic acid, which was added after 4 h of growth. No chlorinated cinnabaramide derivatives could be detected by analytical HPLC-MS after 9, 24, 33, 48, and 57 h cultivation time.

Figures and Tables

Table S1: A) Comparison of the amino acid sequences of acyltransferase (AT) domains from the cinnabaramide gene cluster (CinA-AT1 and AT2, salinosporamide Sal-AT1, and AT2, and thuggacin TgaC-AT). The first two columns show the region around the active site Ser, and the conserved arginine implicated in binding of the dicarboxylated extender units, respectively. B) The table shows the active site and additional conserved residues which have been shown to correlate with domain specificity.^[1-3]

A)

CinA-AT1	G Q S I G	L A V A R S R L I
SalA-AT1	Q Q S V G	V A A T H A R L V
CinA-AT2	G H S L G	I A V A R S R L I
SalA-AT2	G H S I G	V T I A R S R L I
TgaC-AT	G N S Q G	V V A R R S R L L

B)

Domain	11	63	90	91	92	93	94	117	200	201	231	250	255	15	58	59	60	61	62	70	72	197	198	199
CinA-AT ₁	T	Q	G	Q	S	I	G	H	A	H	T	N	S	W	D	M	E	V	V	Q	A	G	Y	A
SalA-AT ₁	T	Q	G	Q	S	V	G	H	S	H	T	N	S	W	D	M	E	I	V	Q	A	G	Y	A
CinA-AT ₂	S	Q	G	H	S	L	G	R	G	H	T	N	V	W	E	G	D	I	Q	H	S	E	R	A
SalA-AT ₂	G	Q	G	H	S	I	G	R	G	H	T	N	V	W	A	T	D	I	Q	Q	S	D	R	P
TgaC-AT ₃	Q	Q	G	N	S	Q	G	R	T	H	T	N	V	W	R	I	D	V	M	M	S	G	D	A

Table S2: HR-ESIMS data for cinnabaramide A and its novel derivatives.

	Formula	<i>m/z</i> , found	<i>m/z</i> , calcd. for [M+H] ⁺	Error [ppm]
Cinnabaramide A	C ₁₉ H ₂₉ NO ₄	336.2178	336.2171	2.1
Cyclopentyl-Cinnabaramide A	C ₁₈ H ₂₉ NO ₄	324.2313	324.2311	0.9
15-Chlorocinnabaramide A	C ₁₉ H ₂₈ ClNO ₄	370.1783	370.1780	0.9
17-Chlorocinnabaramide A	C ₁₉ H ₂₈ ClNO ₄	370.1782	370.1780	0.7

Table S3: Minimal inhibitory concentrations (MICs) of cinnabaramide A, its novel derivatives, and reference compounds against fungi.

	MIC [μg/mL]		
	<i>Botrytis cinerea</i>	<i>Mucor plumbeus</i>	<i>Pyricularia oryzae</i>
Cinnabaramide A	>100	0.8	0.8
17-Chlorocinnabaramide A	>100	6.3	3.1
15-Chlorocinnabaramide A	>100	6.3	1.6
Cyclopentyl-Cinnabaramide A	>100	6.3	0.8
Salinosporamide A	>100	25	6.3

Table S4: NMR data assignment for cinnabaramide A and its novel derivatives

Cinnabaramide A			Cyclopentyl-Cinnabaramide ^[a]		15-Chlorocinnabaramide		17-Chlorocinnabaramide	
atom	δ_c , Mult.	δ_H (J in Hz)	δ_c , Mult.	δ_H (J in Hz)	δ_c , Mult.	δ_H (J in Hz)	δ_c , Mult.	δ_H (J in Hz)
1	175.9, qC	-	175.9, qC	-	175.8, qC	-	175.9, qC	-
2	47.7, CH	2.41, dd (7.6, 5.8)	47.7, CH	2.42, t(6.0)	47.5, CH	2.46, m	47.6, CH	2.44, t (6.5)
3	86.1, qC	-	86.6, qC	-	86.0, qC	-	86.1, qC	-
4	78.6, qC	-	78.6, qC	-	78.6, qC	-	78.6, qC	-
5	69.1 CH	3.67, dd (9.0, 7.8)	69.4, CH	3.75, t (7.6)	69.1, CH	3.67, t (8.2)	69.1, CH	3.68, d (9.8)
6	37.7, CH	2.29, m	42.1, CH	1.98, m	37.7, CH	2.29, m	37.7, CH	2.29, m
7	128.5, CH	5.81, m	29.8, CH ₂	1.29, 1.78, m	128.5, CH	5.80, m	128.5, CH	5.81, m
8	127.7, CH	6.73, m	24.1, CH ₂	1.44, 1.54, m	127.7, CH	5.72, m	127.7, CH	5.73, m
9	24.6, CH ₂	1.92, m	25.2, CH ₂	1.46, 1.58, m	24.5, CH ₂	1.91, m	24.5, CH ₂	1.92, m
10	21.0, CH ₂	1.41, 1.70, m	28.9, CH ₂	1.23, 1.69, m	21.0, CH ₂	1.41, 1.70, m	21.0, CH ₂	1.41, 1.70, m
11	25.3, CH ₂	1.24, 1.82, m	-	-	25.3, CH ₂	1.24, 1.81, m	25.3, CH ₂	1.24, 1.82, m
12	24.7, CH ₂	1.49, 1.60, m	24.7, CH ₂	1.48, 2.58, m	24.1, CH ₂	1.57, 1.77, m	24.5, CH ₂	1.51, 1.60, m
13	27.0, CH ₂	1.45, 1.53, m	27.0, CH ₂	1.44, 1.52, m	24.1, CH ₂	1.54, 1.59, m	26.9, CH ₂	1.46, 1.55, m
14	28.8, CH ₂	1.29, m	28.8, CH ₂	1.28, m	37.4, CH ₂	1.69, 1.76, m	28.3, CH ₂	1.33, m
15	31.0, CH ₂	1.28, m	31.0, CH ₂	1.28, m	66.0, CH	4.01, m	25.9, CH ₂	1.40, m
16	22.0, CH ₂	1.28, m	22.0, CH ₂	1.28, m	30.9, CH ₂	1.66, 1.82, m	31.9, CH ₂	1.72, m
17	13.9, CH ₂	0.87, t (6.7)	13.9, CH ₃	0.87, t (6.3)	10.6, CH ₃	0.97, t (7.2)	45.3, CH ₂	3.64, m
18	20.0, CH ₂	1.74, s	20.0, CH ₃	1.74, s	19.9, CH ₃	1.75, s	20.0, CH ₃	1.74, s
19	168.9, qC	-	169.3, qC	-	168.8, qC	-	168.9, qC	-
NH	-	8.91, s	-	8.79, s	-	8.92, s	-	8.92, s
OH	-	5.49, d (7.9)	-	5.34, d (7.1)	-	5.50, d (7.6)	-	n.d.

[a] The same atom numbering was used for the cyclopentyl derivatives as for other cinnabaramides. Consequently, C-11 is lacking here.

cinnabaramide A: $R^1 = OH$, $R^2 = H$, cyclopentyl-cinnabaramide B: $R^1 = R^2 = OH$, and cyclopentyl-cinnabaramide C: $R^1 = R^2 = H$. B) Cyclopentyl-cinnabaramide D: $R^1 = R^2 = OH$, cyclopentyl-cinnabaramide E: $R^1 = OH$, $R^2 = H$

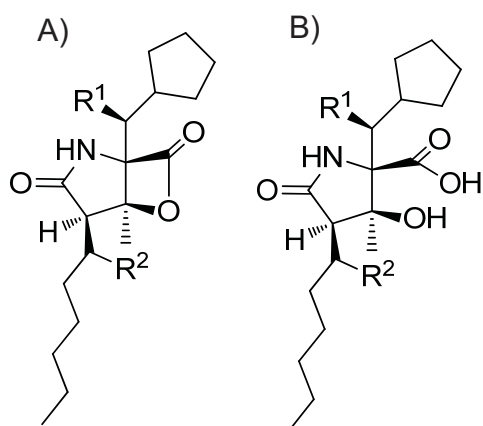
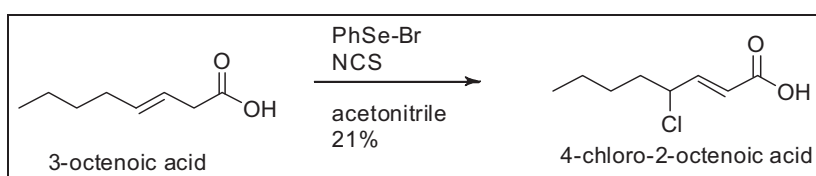
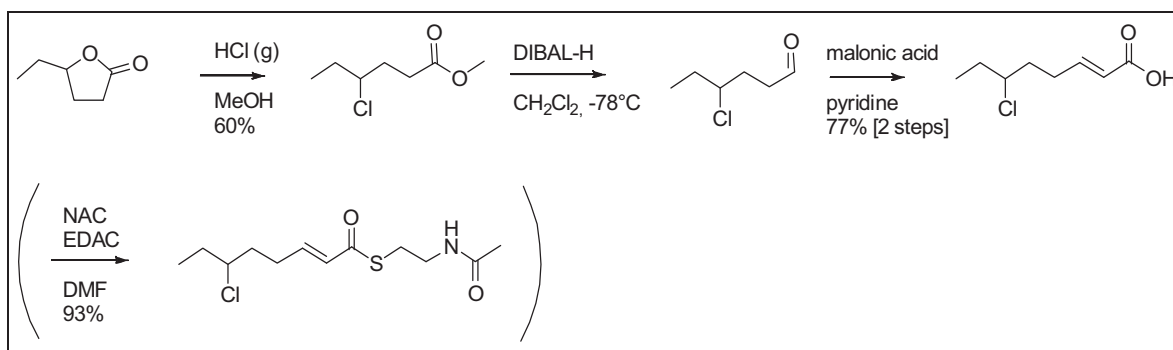


Figure S2: Synthesis of precursors for feeding experiments.

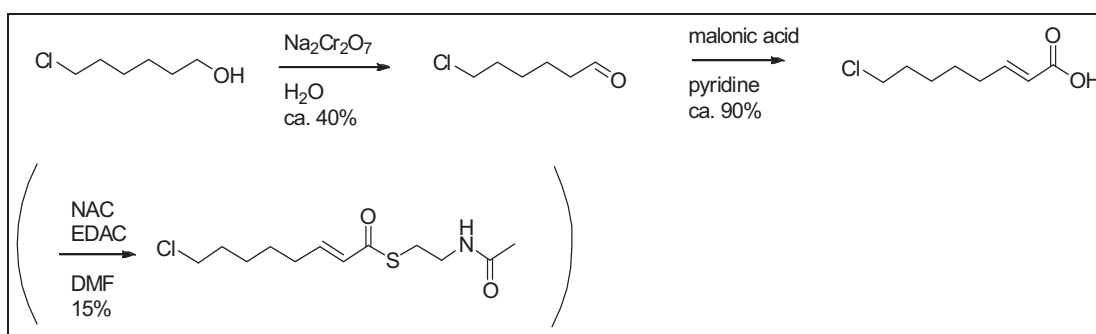
A) Scheme of (*E*)-4-chloro-2-octenoic acid synthesis.



B) Scheme of (*E*)-6-chloro-2-octenoic acid and its corresponding SNAC derivative synthesis.



C) Scheme of (*E*)-8-chloro-2-octenoic acid and its corresponding SNAC derivative synthesis.



D) Scheme of (*E*)-5-chloro-2-octenoic acid synthesis.

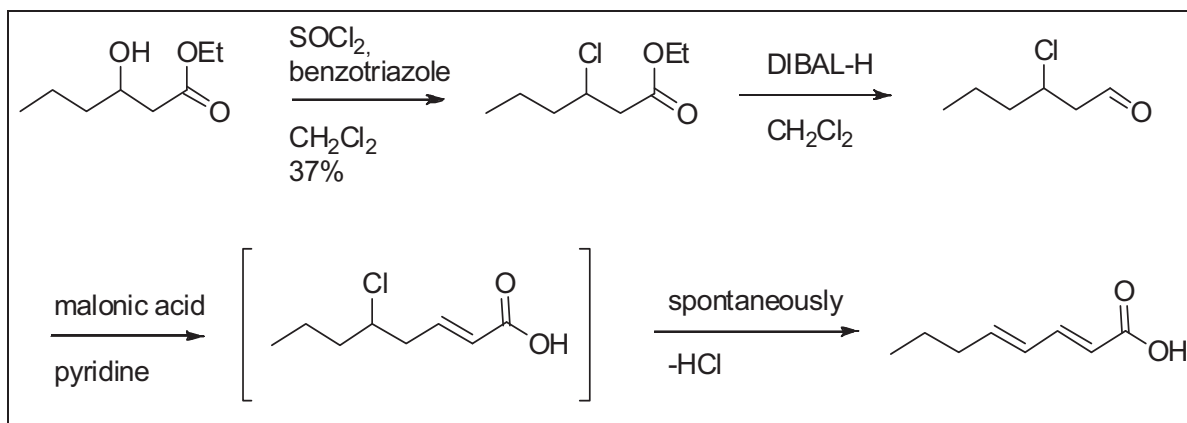
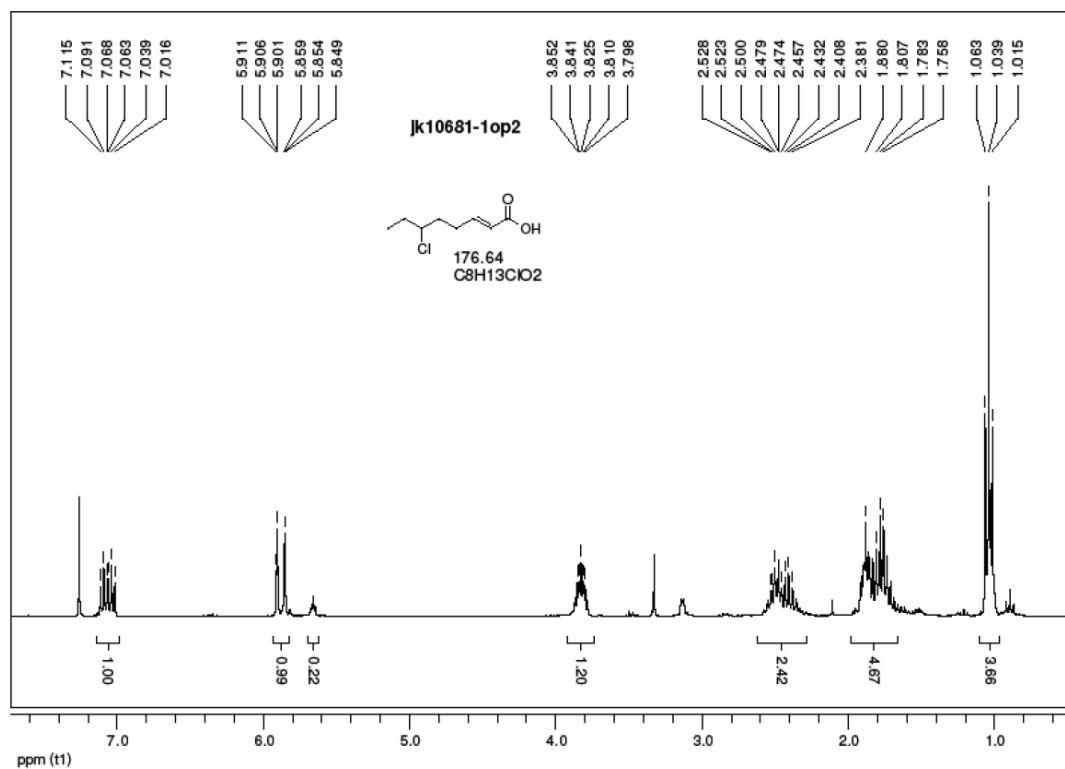
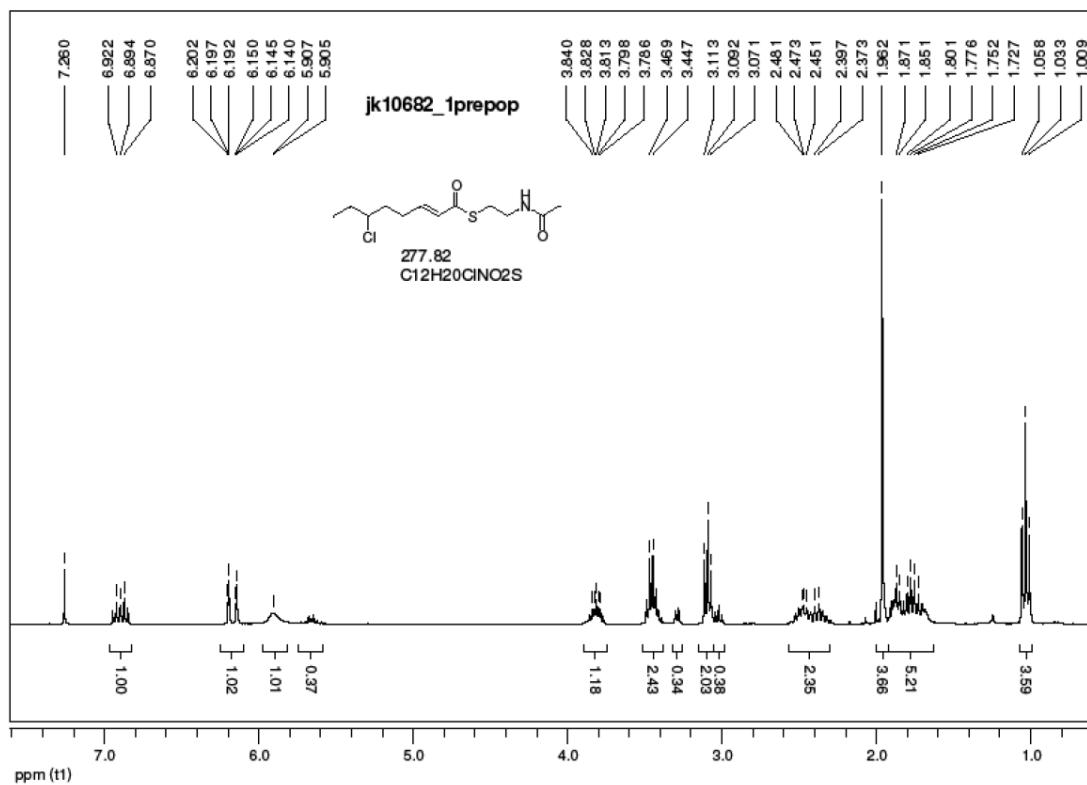


Figure S3: ^1H NMR spectra of the cinnabaramide hexyl side chain analog precursors.

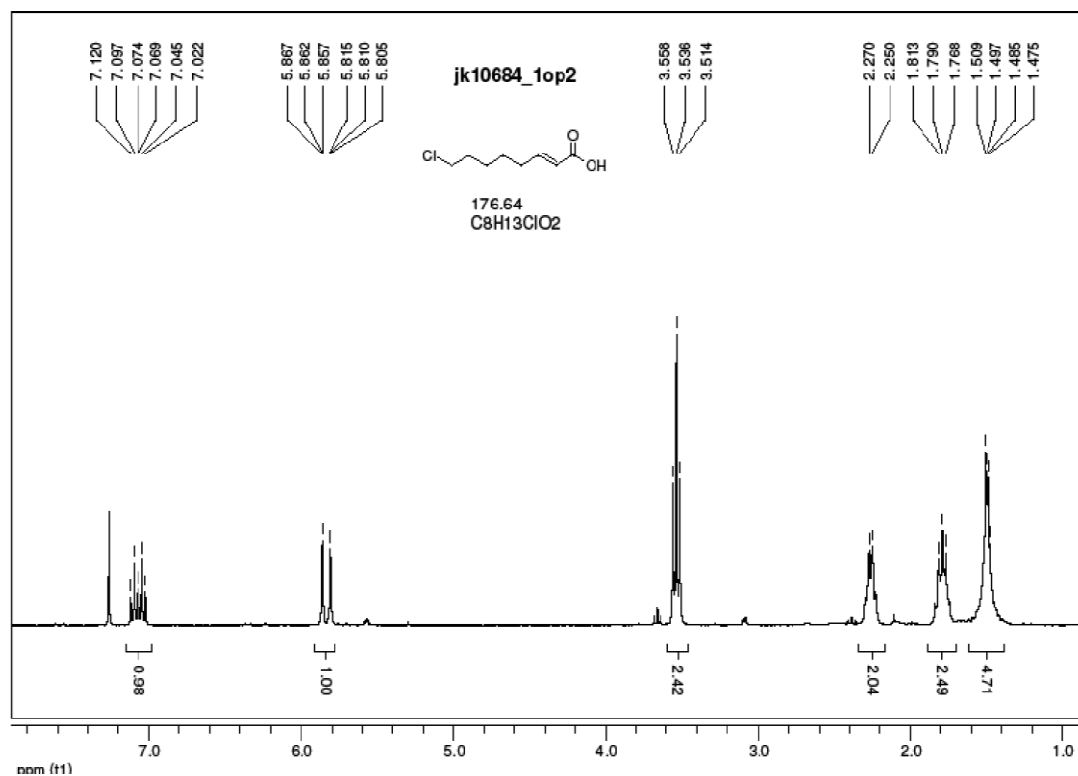
^1H NMR spectrum (CDCl_3 , 300 MHz) of (E)-6-chloro-2-octenoic acid



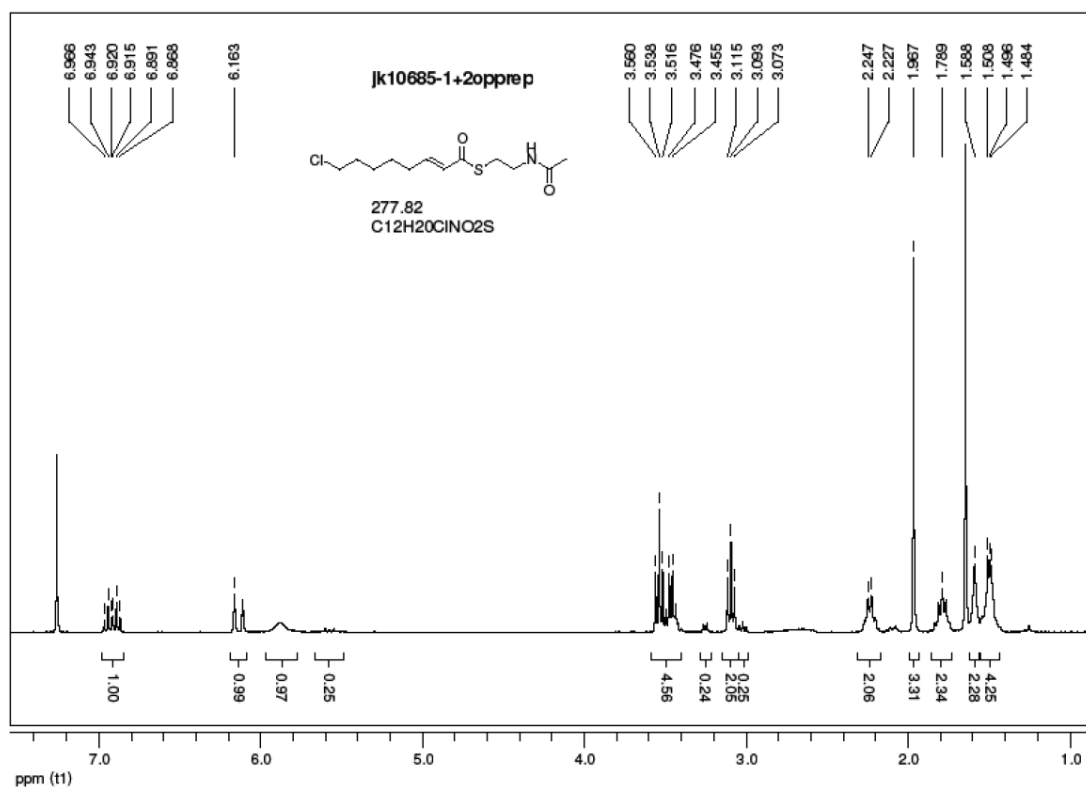
^1H NMR spectrum (CDCl_3 , 300 MHz) of (E)-6-chloro-2-octenoyl-SNAC



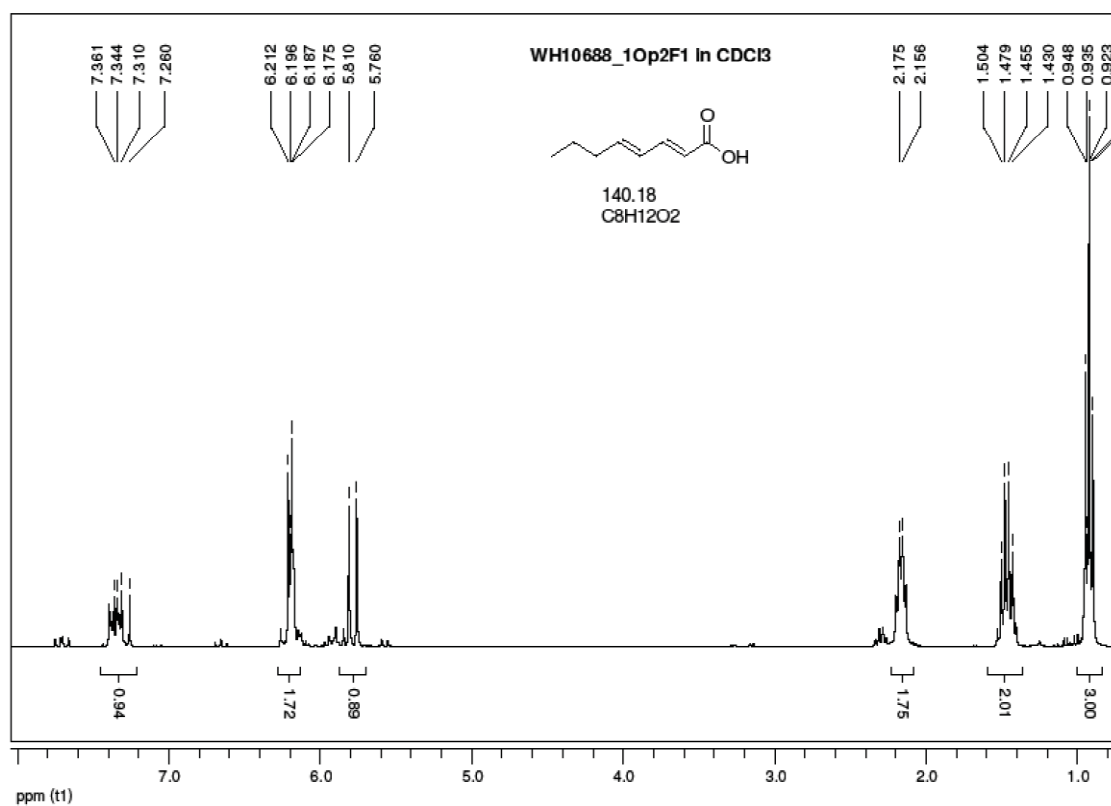
¹H NMR spectrum (CDCl₃, 300 MHz) of (*E*)-8-chloro-2-octenoic acid



¹H NMR spectrum (CDCl₃, 300 MHz) of (*E*)-8-chloro-2-octenoyl-SNAC



¹H NMR spectrum (CDCl₃, 300 MHz) of **(2*E*,4*E*)-2,4-octadienoic acid**



¹H NMR spectrum (CDCl₃, 300 MHz) of **(*E*)-4-chloro-2-octenoic acid**

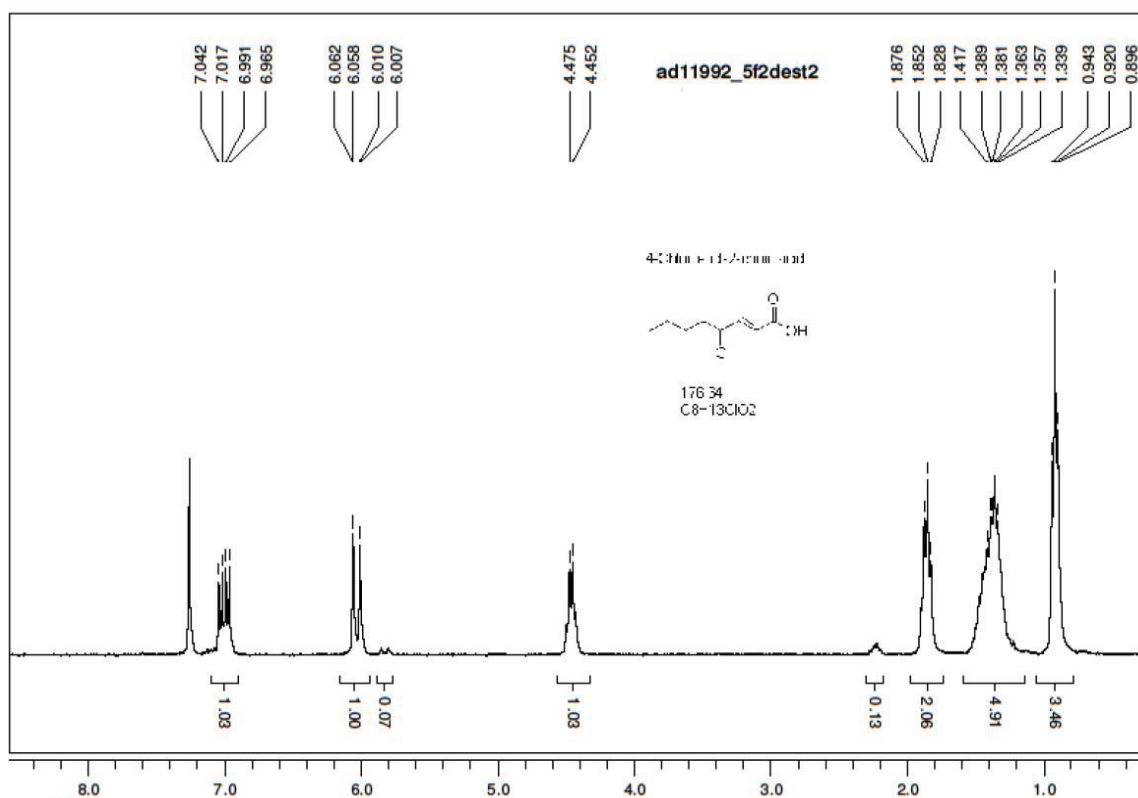
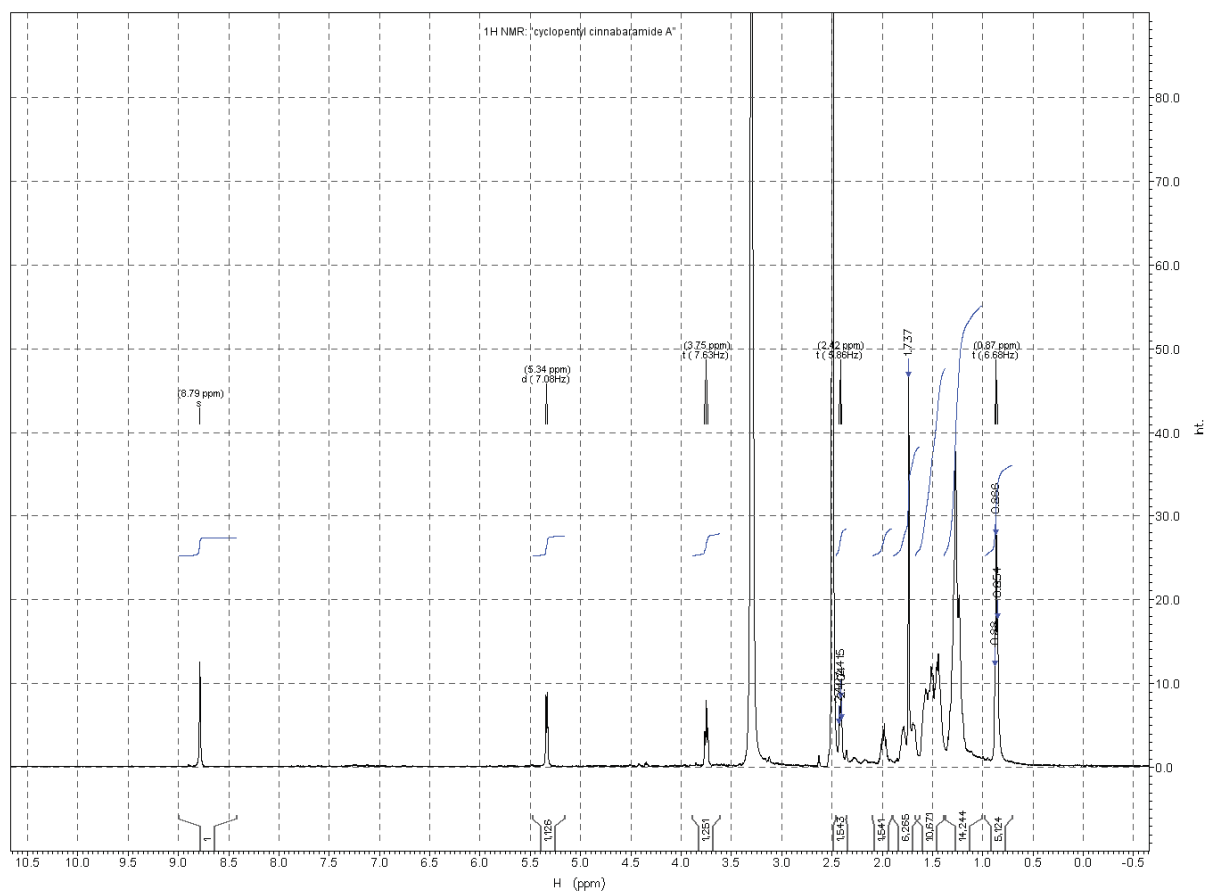
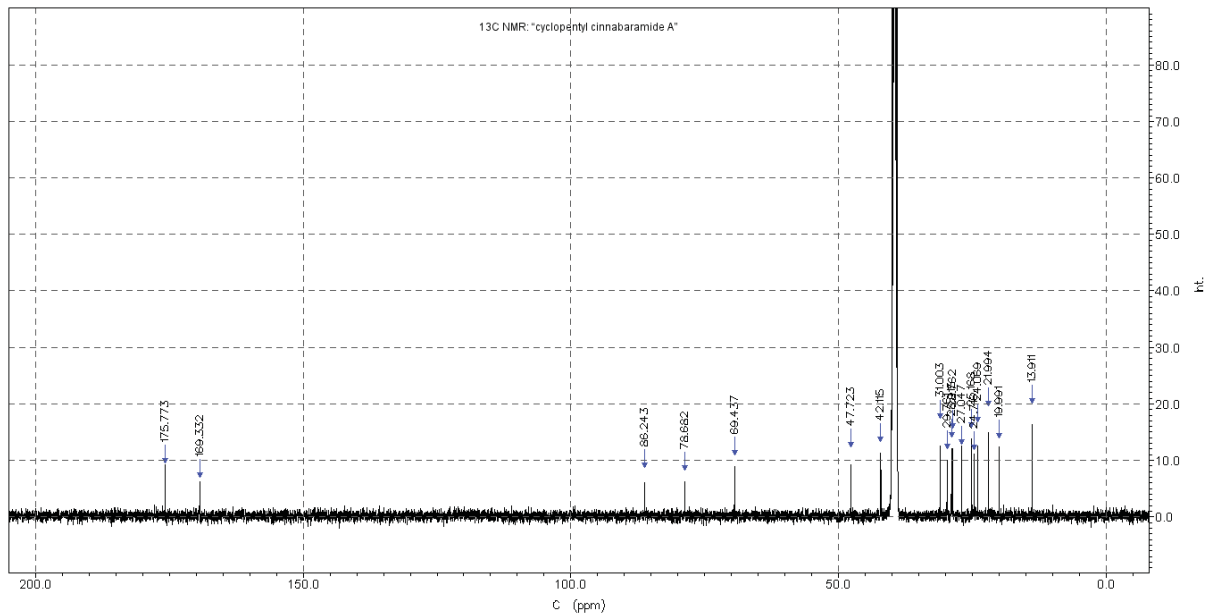


Figure S4: ¹H and ¹³C NMR spectra of the novel cinnabaramides.

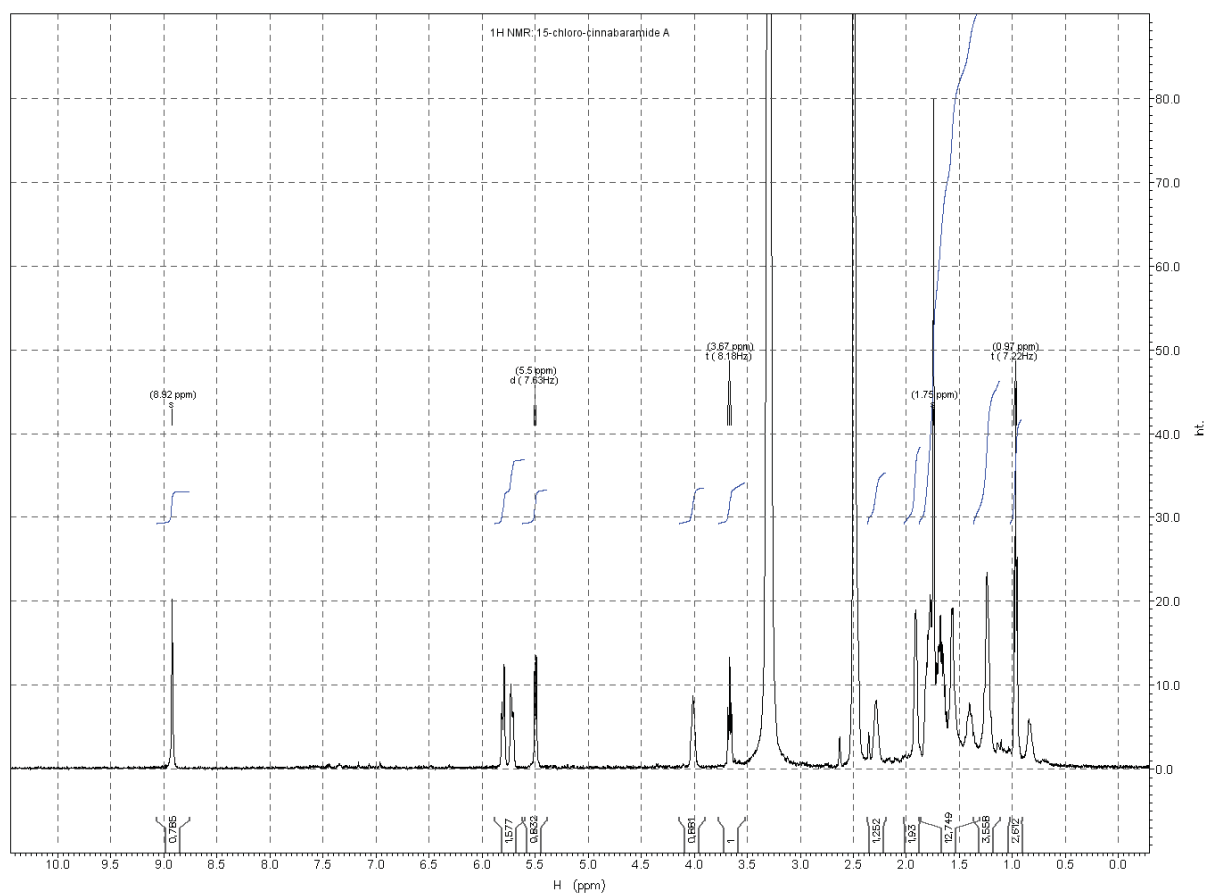
¹H NMR spectrum (DMSO-*d*₆, 500 MHz) of **cyclopentyl-cinnabaramide A**



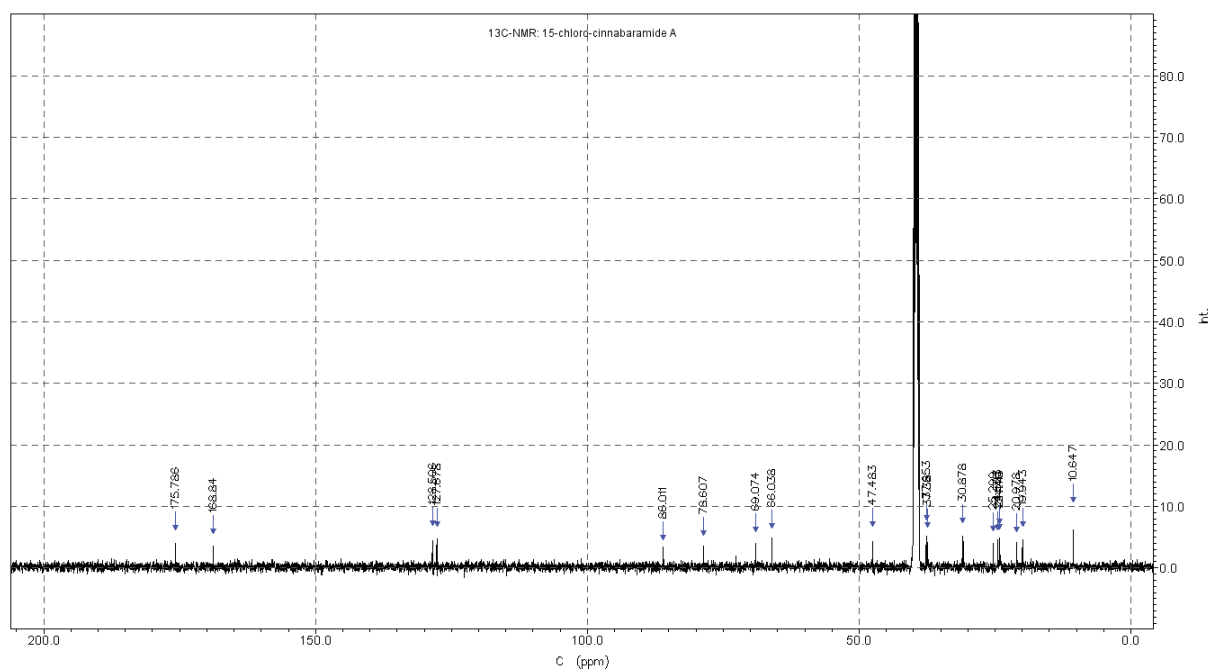
¹³C NMR spectrum (DMSO-*d*₆, 125 MHz) of cyclopentyl-cinnabaramide A



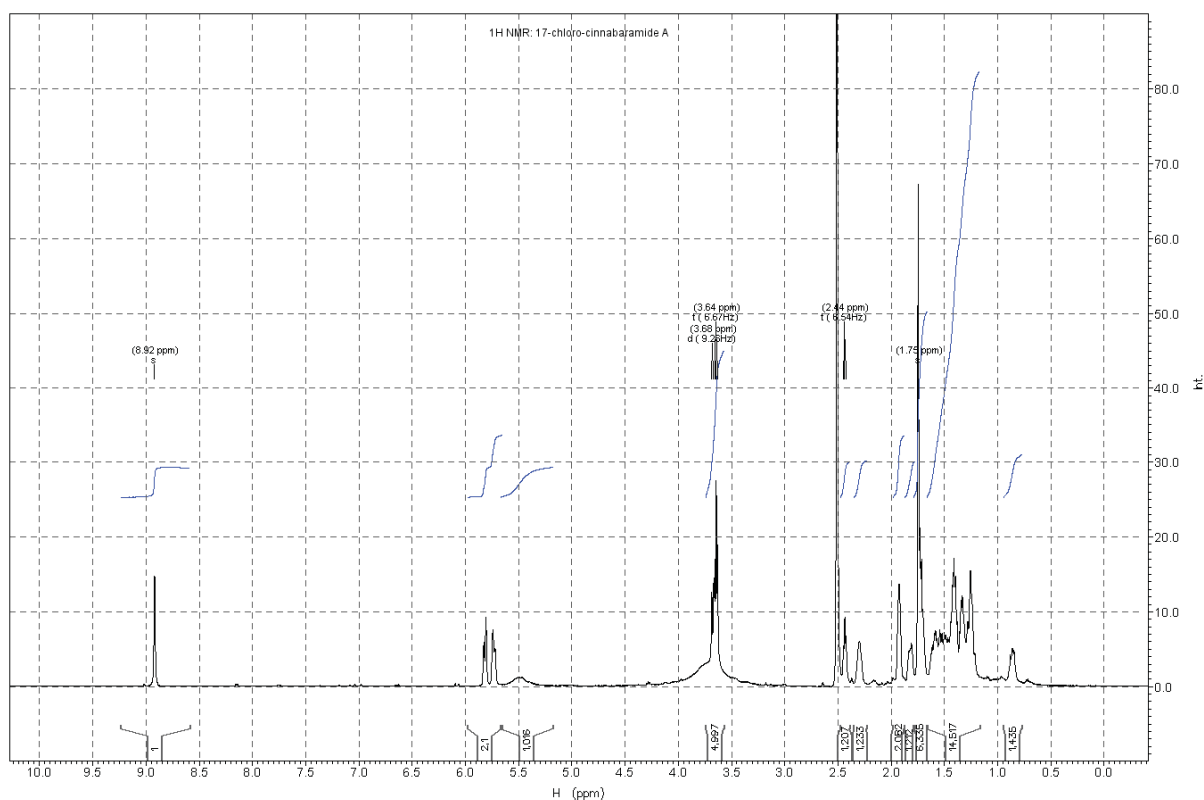
¹H NMR spectrum (DMSO-*d*₆, 500 MHz) of 15-chlorocinnabaramide A



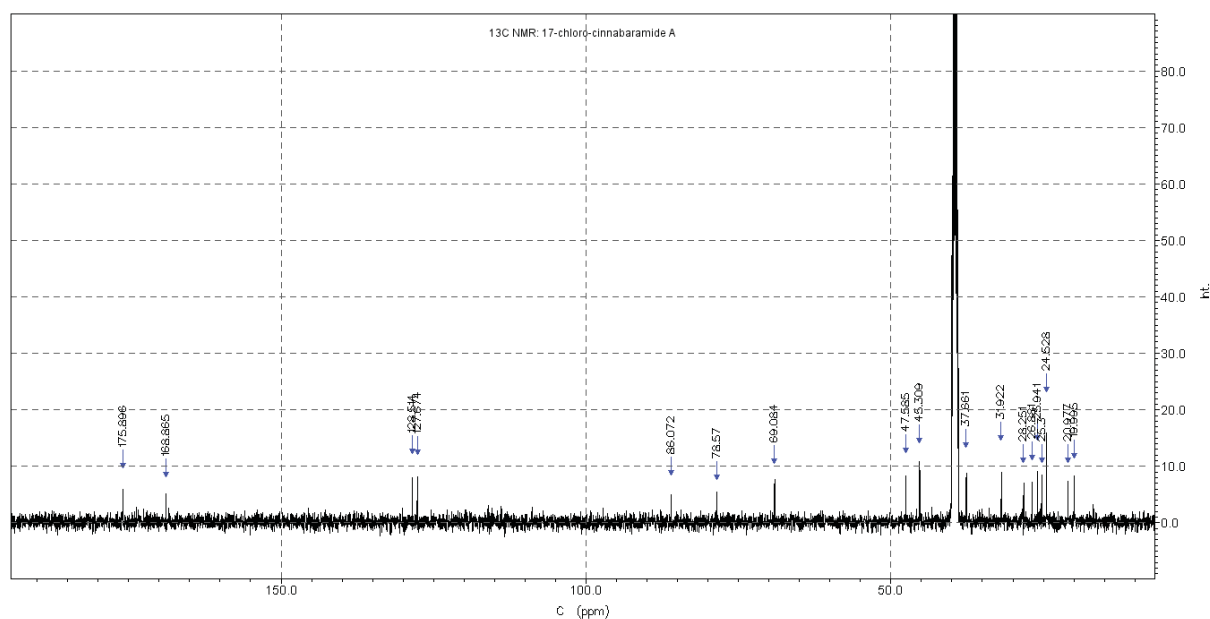
¹³C NMR spectrum (DMSO-*d*₆, 125 MHz) of **15-chlorocinnabaramide A**



¹H NMR spectrum (DMSO-*d*₆, 500 MHz) of **17-chlorocinnabaramide A**



¹³C NMR spectrum (DMSO-*d*₆, 125 MHz) of **17-chlorocinnabaramide A**



Reference List

- [1.] G. Yadav, R. S. Gokhale, D. Mohanty, *Journal of Molecular Biology* **2003**, 328 335-363.
- [2.] F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. Rudd, J. Staunton, P. F. Leadlay, *J.Ind.Microbiol.Biotechnol.* **2003**, 30 489-494.

- [3.] S. F. Haydock, J. F. Aparicio, I. Molnar, T. Schwecke, A. König, A. F. A. Marsden, I. S. Galloway, J. Staunton, P. F. Leadley, *FEBS Lett.* **1995**, 374 246-248.

II

Unusual carbon fixation gives rise to diverse polyketide extender unit

Nick Quade^{1,+}, Liujie Huo^{2,+}, Shwan Rachid², Dirk W. Heinz^{1,*} & Rolf Müller^{2,*}

¹Department of Molecular Structural Biology, Helmholtz Centre for Infection Research, D-38124, Braunschweig, Germany

²Department of Microbial Natural Products, Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Campus C2 3, D-66123, Saarbrücken, Germany.

⁺These authors contributed equally to this work.

* Corresponding authors: dirk.heinz@helmholtz-hzi.de; rom@helmholtz-hzi.de

Unusual carbon fixation gives rise to diverse polyketide extender units

Nick Quade^{1,4}, Liujie Huo^{2,4}, Shwan Rachid^{2,3}, Dirk W Heinz^{1*} & Rolf Müller^{2*}

Polyketides are structurally diverse and medically important natural products that have various biological activities. During biosynthesis, chain elongation uses activated dicarboxylic acid building blocks, and their availability therefore limits side chain variation in polyketides. Recently, the crotonyl-CoA carboxylase-reductase (CCR) class of enzymes was identified in primary metabolism and was found to be involved in extender-unit biosynthesis of polyketides. These enzymes are, in theory, capable of forming dicarboxylic acids that show any side chain from the respective unsaturated fatty acid precursor. To our knowledge, we here report the first crystal structure of a CCR, the hexylmalonyl-CoA synthase from *Streptomyces* sp. JS360, in complex with its substrate. Structural analysis and biochemical characterization of the enzyme, including active site mutations, are reported. Our analysis reveals how primary metabolic CCRs can evolve to produce various dicarboxylic acid building blocks, setting the stage to use CCRs for the production of unique extender units and, consequently, altered polyketides.

Microbial polyketides represent a major class of secondary metabolites formed by tremendously large multimodular enzymatic machineries termed polyketide synthases¹. Their importance is best demonstrated by referring to the large number of polyketides applied in the clinic, many of which are regarded as blockbuster drugs². The enormous structural variety shown by natural products of the polyketide is regarded as the major reason why these molecules can evolve to interact with a large number of targets.

Much of the chemical diversity stems from the multimodularity of each assembly line in which each module not only determines the choice of the starter or extender unit but also selectively specifies the oxidation status (keto, hydroxyl or enoyl) of each position in the structure, selectively³. Additional diversity is generated by the methylation of hydroxyl groups or carbon atoms to form methoxy, methyl or *gem*-dimethyl partial structures as well as by the tailoring of reactions that occur after polyketide assembly such as glycosylations, carbamoylations or hydroxylations.

Acyltransferases are in charge of loading the extender dicarboxylic acid onto the carrier protein, and thus each acyltransferase domain in every module of the polyketide synthase (PKS) has the potential to incorporate malonyl-CoA derivatives with any side chain (in theory, even two side chains) that metabolism can generate. For a long time, however, it seemed that the choice of extender units in polyketide biosynthesis was rather limited. Polyketides mostly incorporate malonyl-CoA from fatty acid metabolism and methylmalonyl-CoA, which in several secondary-metabolite producers is derived from other well-established pathways such as the succinyl-CoA mutase-epimerase pathway. Owing to the limited choice of extender units available from primary metabolism, side chain variability in polyketides seemed hardly possible.

The origins of methoxymalonyl-CoA giving rise to methoxy branches, as found in soraphen and ansamitocin, has been elucidated. It turns out that a glycerol-derived precursor is bound to a carrier protein and is further processed to form this extender unit^{4,5}. It was later reported that hydroxymalonyl-CoA and aminomalonyl-CoA can be formed and used in a similar fashion⁶.

However, the literature describes numerous polyketides with side chains that cannot be explained retrobiosynthetically using any of the mechanisms described above (Fig. 1). Their formation remained enigmatic until the recent report of the biochemistry of the CCRs⁷. These enzymes reductively carboxylate crotonyl-CoA^{7,8} to ethylmalonyl-CoA within the acetyl-CoA assimilation pathway, involving an unprecedented reaction mechanism. It was further suggested that crotonyl-CoA homologs may be involved in the generation of polyketide extender units such as ethylmalonyl-CoA.

Indeed, homologs of CCRs have been found to be closely associated with a growing number of secondary-metabolite biosynthetic pathways. The occurrence of these homologs can be logically connected to the structure of the respective compounds in a way that enables the prediction of the biochemistry of several of these enzymes. To date, CCRs were found in the pathways of cinnabar-amides⁹, thuggacins¹⁰, salinosporamides¹¹, FK506 (refs. 12,13), sanglifehrins¹⁴, leupyrins¹⁵, germicidins¹⁶, divergolides¹⁷, ansalactams¹⁸ and filipin^{19,20} (Fig. 1 and Supplementary Results, Supplementary Fig. 1). In addition, CCRs have been predicted to be involved in the biosynthesis of other secondary metabolites for which no biosynthetic genes have been reported yet (for example, icumazole). However, most of the enzymes lack biochemical characterization, and no structural information is available for them. So far, only SalG, an enzyme involved in the formation of chloroethylmalonyl-CoA from chlorocrotonyl-CoA in salinosporamide biosynthesis¹¹, and PteB, an 2-octenoyl-CoA reductase-carboxylase involved in filipin biosynthesis, have been biochemically characterized^{19,20}. We were intrigued by the fact that the myxobacterial metabolite thuggacin as well as the actinomycete-derived cinnabaramide (Fig. 1) contain hexyl side chains. Both biosynthetic gene clusters turned out to have CCR homologs (CinF and TgaD), which, interestingly, had very low sequence homology^{9,10} despite their assumed identical function. Both enzymes were hypothesized to be involved in the formation of hexylmalonyl-CoA from 2-octenoyl-CoA.

In this report, we biochemically characterized CinF and solved the crystal structure of the enzyme in complex with NADP⁺ and its substrate 2-octenoyl-CoA at high resolution. We were able to show

¹Department of Molecular Structural Biology, Helmholtz Centre for Infection Research, Braunschweig, Germany. ²Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Germany. ³Faculty of Science and Health, University of Koya, Koya, Kurdistan, Iraq. ⁴These authors contributed equally to this work. *e-mail: dirk.heinz@helmholtz-hzi.de or rom@helmholtz-hzi.de

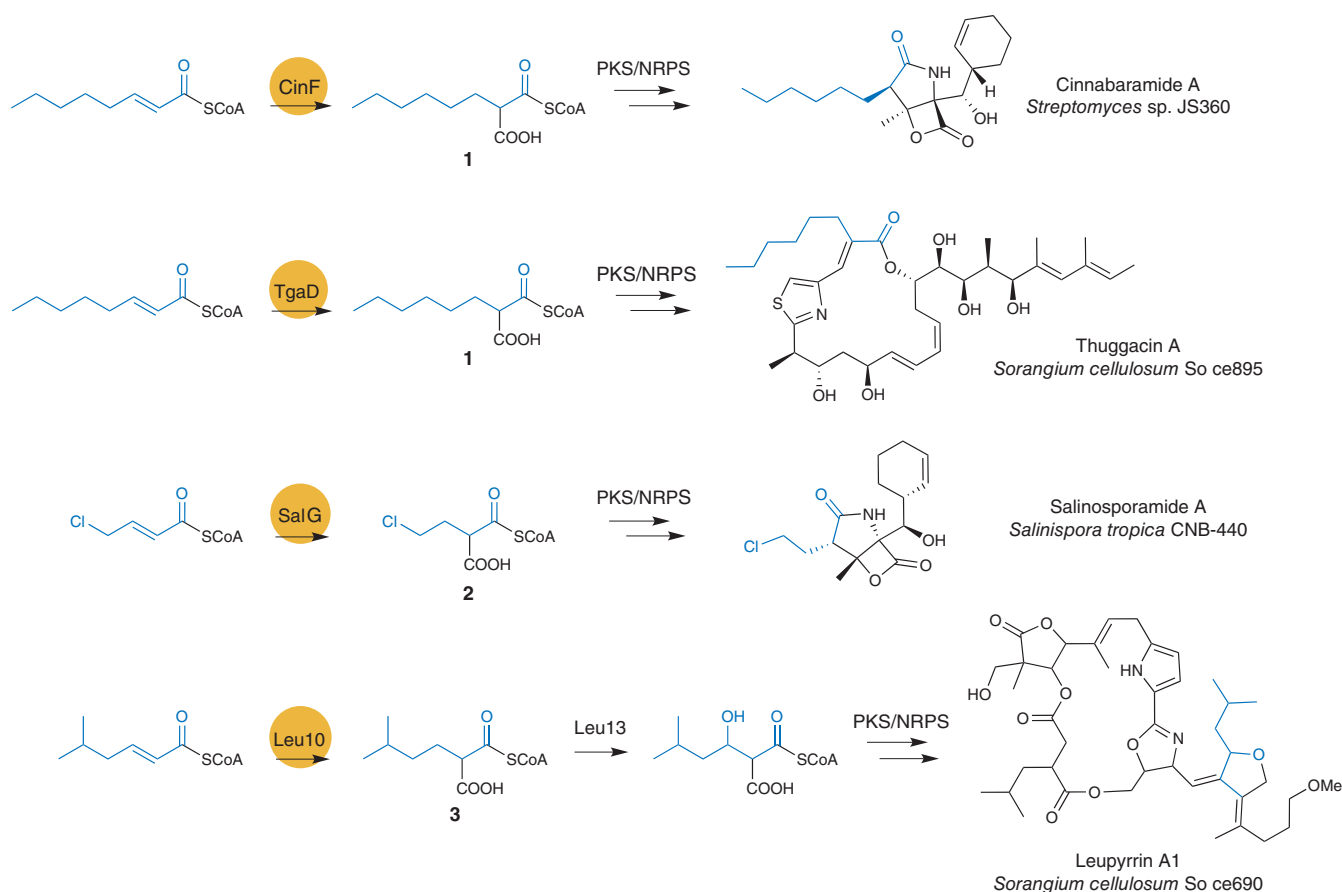


Figure 1 | Unusual alkylmalonyl thioester building blocks generated by reductive carboxylation and incorporated into bacterial polyketide metabolites. CinF homologs are circled in dark yellow and their likely products are framed in black: **1**, hexyl-malonyl-CoA; **2**, chloroethyl-malonyl-CoA; **3**, 2-carboxyl-5-methylhexanoyl-CoA.

that two residues in the active site are critical for the carboxylation reaction catalyzed by CinF. Additionally, the structure allowed us to identify residues important for substrate specificity. By point mutagenesis of these residues, it was possible to change the substrate specificity of the protein. To our knowledge, the structure of CinF provides the first structural model for the CCR class of enzymes, allowing insight into a unique way to establish structural diversity in polyketides.

RESULTS

CinF contains two distinct domains and forms a tetramer

We solved the structure of the apo-2-octenoyl-CoA carboxylase-reductase CinF at 2.25-Å resolution and in complex with NADP⁺ and 2-octenoyl-CoA at 1.9-Å resolution. The asymmetric unit contains four CinF monomers (chains A–D) (Fig. 2a). For the apo structure, residues 1–347 were modeled with some gaps due to weak electron density, which is probably brought about by flexibility in the absence of the ligands. These parts were well defined in the structure of the ligand-bound complex, which comprises residues 1–445. Both structures align quite well, with an average r.m.s. deviation of 0.63 Å for C_α atoms. Unless stated otherwise, we refer to the structure of the complex in the remainder of the text.

CinF can be divided into two domains: the catalytic domain present at the N and C termini (residues 1–201 and 361–445, respectively) and the central cofactor-binding domain (residues 202–360) (Fig. 2b). The latter has a Rossmann fold, which contains two mononucleotide-binding β α β α motifs and an additional strand, β14. The catalytic domain consists of a seven-stranded antiparallel β-sheet, a separate three-stranded antiparallel β-sheet and eleven α-helices.

In the asymmetric unit, four CinF monomers were found in a dimer-of-dimers assembly (Fig. 2a). Size-exclusion chromatography of CinF suggests a molecular mass of 176 kDa (calculated mass, 48 kDa), corresponding to a tetramer (Supplementary Fig. 5). Analysis of the most probable assembly using the protein interfaces, surfaces and assemblies (PISA) server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) also suggests a tetramer covering an interface area of 25,410 Å² (Supplementary Fig. 6)²¹. The strongest interaction between two monomers is between the cofactor binding domains of chains A and D as well as between chains B and C, with an average dimer interface area of 1,475 Å² (Fig. 2a). This interaction is mainly mediated by β15 of both subunits, thus creating the continuous 12-stranded β-sheet β14 that interacts with the same β-strand from the other chain, and helix α16, which is inserted between the two domains of the other chain. The two dimers interact via two interfaces to form the tetramer. The first tetramer interface is between the catalytic domains of chains A and B and chains C and D, covering an average area of 1,252 Å², whereas the second is located between the cofactor binding domains of chains A and C and chains B and D and is relatively weak, covering an area of only 629 Å².

Structure and activity define NADP(H) as the cofactor

Well-defined electron density between the two domains of CinF allowed unambiguous placement of the NADP⁺ molecule in all four subunits. The NADP⁺ molecule assumes an extended conformation and is only slightly bent around the loop preceding β14 and the loop between β12 and α15 (Fig. 3a). These loops make several interactions with the NADP⁺ molecule and are disordered in the apo structure. The adenine group is bound mostly via

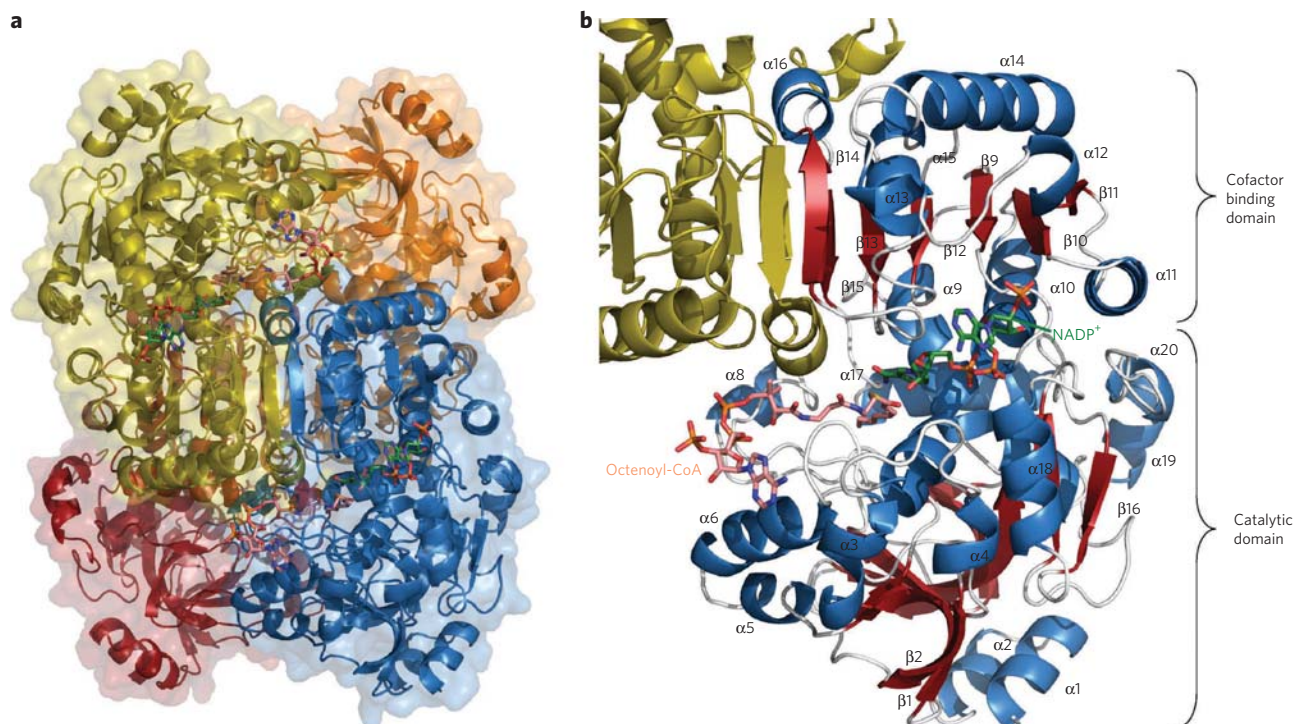


Figure 2 | Structure of CinF. (a) Image of the quaternary structure of CinF showing the dimer-of-dimers fold. The monomers are shown as cartoons, and the transparent surface is depicted in orange, blue, yellow and red. The ligands are shown as sticks. (b) Cartoon representation of the interface of the CinF monomer and dimer. Helices of one monomer are colored in blue, β -sheets in red and loops in white. The ligands are shown as sticks, NADP⁺ is in green, 2-octenoyl-CoA in pink and the other monomer in yellow.

water-mediated contacts and hydrophobic interactions with Ser338 and Ser339. The 2'-phosphate group is involved in hydrogen bonds with Lys257 and with the OH group and the backbone nitrogen of Ser253 as well as by several water-mediated interactions, for example with Arg272 and Ser230. This shows that NADPH is indeed the cofactor, not NADH, as the use of NADH instead of NADPH abolished any activity in the reaction assay (data not shown). The pyrophosphate group is tightly bound via hydrogen bonds with Gln405 and with the backbone nitrogens of Gly232 and Leu233 as well as through several water-mediated interactions. The last two residues are located at the N-terminal end of helix α 10. On the other hand, the nicotinamide group, which assumes an *anti* conformation, is inserted in a hydrophobic pocket lined by Leu201, Thr205, Leu233 and Cys335 and is held in position by hydrogen bonds with Thr205 and His361.

CinF positions 2-octenoyl-CoA for catalysis

Density for the 2-octenoyl-CoA was much clearer in monomers A and B than in C and D, suggesting only partial occupancy of the ligand in the latter two. This results in a high B factor for chains C and D and is probably due to subtle differences caused by crystal packing. 2-Octenoyl-CoA was built into the monomer with the best defined density (chain A) and was then placed in the other monomers using noncrystallographic symmetry. The mode of 2-octenoyl-CoA binding is thus described for chain A. The long and flexible ligand is bound in an extended conformation in which the acyl group is buried within the catalytic domain, whereas the CoA is bound on the surface of the monomer and held in place by interactions with the dimer partner of chain A (chain D) (Fig. 3b). The acyl group is accommodated by a hydrophobic pocket created by residues Pro141, Ala142, Met156, Ala163, Leu201, His361 and Gly362 as well as by the nicotinamide group of the NADP⁺ molecule (Fig. 3c). This places the C3 of the octenoyl chain in perfect position for the hydride transfer from an NADPH molecule.

The pantetheine part of the CoA is bound in the interface between the catalytic domain of chain A and α 14 and α 16 of chain D (Fig. 3b). It makes most of its contacts through hydrophobic interactions with Trp80, Pro86, Leu87, Phe92 and Phe166 from chain A as well as with Tyr349 and Met352 from chain D. Additionally, there is a hydrogen bond from Arg348, and there are some water-mediated interactions. The remainder of the CoA molecule winds around helix α 6 and the preceding loop. The CoA phosphates make contact with His91 from chain A, Arg286 and Arg293 from α 14, and Tyr349 from α 16 of chain D. These residues are not visible in the apo structure.

The ribose and the adenine parts of the CoA molecule are positioned on the surface of monomer A. They are only held in place by hydrophobic interactions with Leu87, Pro88 and His91 and by a water bridge to Asp36.

In silico docking suggests a CO₂ binding site

CCRs use CO₂ to carboxylate their substrates from the *re* face, in a position *anti* to the hydride transfer⁸. We aimed to confirm this by solving the structure of CinF in complex with CO₂. However, cocrystallization with as much as 100 mM Na₂CO₃ or CS₂, to act as CO₂ mimics, or aeration of the crystals with up to 40 bar CO₂ in a pressure chamber for 15 min right before freezing did not result in detectable CO₂ (or CS₂) binding. Therefore, we docked CO₂ into the structure *in silico*²². Several very similar solutions were found, and all of them had relatively low affinities (Fig. 3d and Supplementary Fig. 7). The CO₂ molecule is held in place by hydrogen bonds with residues Asn77 and Glu167 as well as by hydrophobic interactions with the underlying Phe166. The CO₂ is in an *anti* position relative to the NADP⁺, and the distance from the C atom of the CO₂ to the C2 of 2-octenoyl-CoA is about 3.3 Å. Therefore, we deem this a very probable orientation as the CO₂ is optimally positioned for a nucleophilic attack from C2 of 2-octenoyl-CoA.

A comparison with other protein structures in complex with CO₂ notably showed a similar CO₂ binding site as proposed for

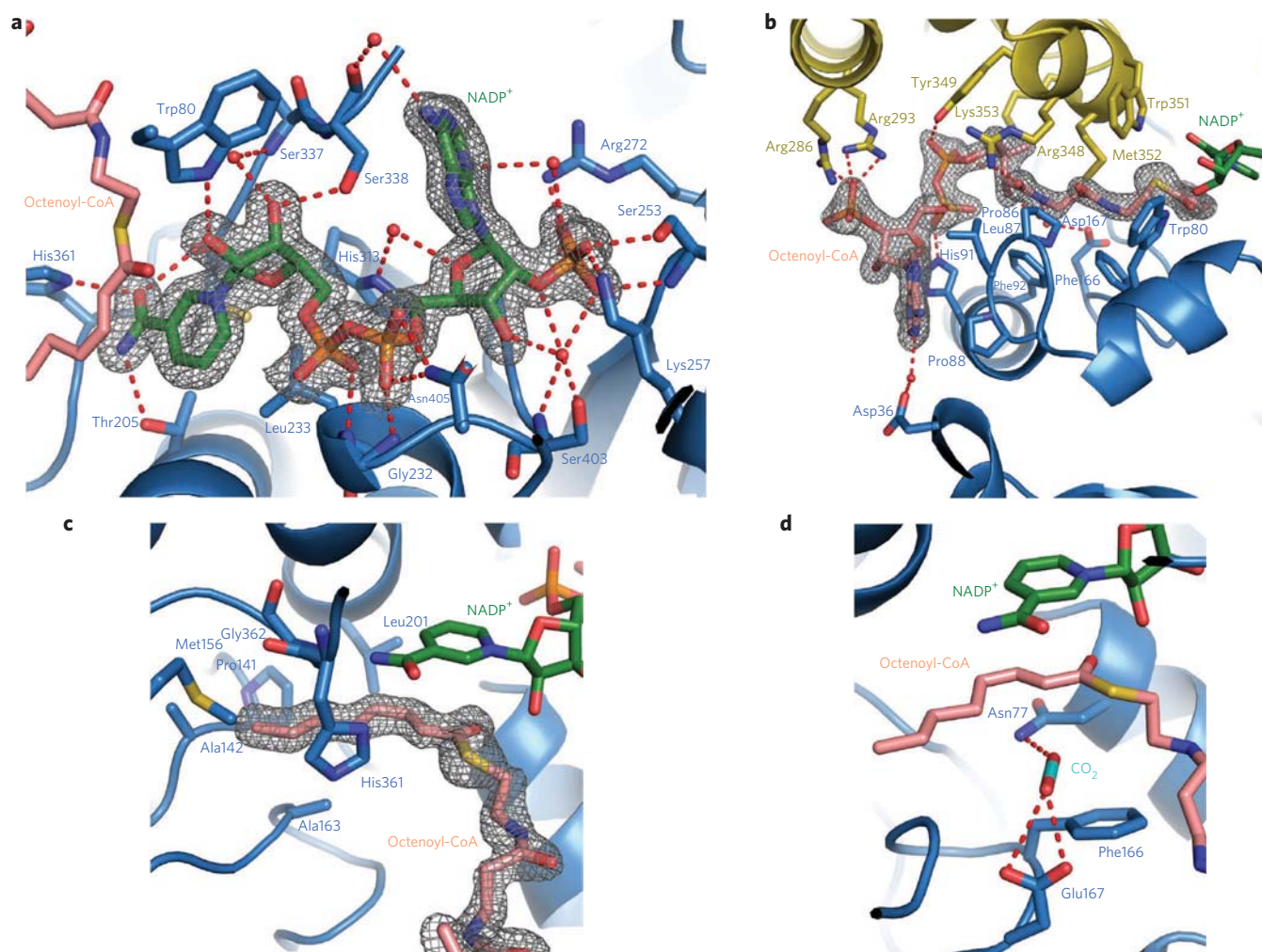


Figure 3 | Ligand binding by CinF. CinF is shown as a cartoon and in sticks; chain A is colored blue and chain D is yellow. NADP⁺ and 2-octenoyl-CoA are shown as green and pink sticks, respectively. The electron density is shown at 1.0 σ as gray mesh. **(a)** Interactions between NADP⁺ and CinF showing the residues responsible for cofactor specificity. **(b)** Binding of 2-octenoyl-CoA at the interface between chains A and D. **(c)** Close-up of the hydrophobic substrate binding pocket of the 2-octenoyl-chain. **(d)** Model of CO₂ docked into the structure of CinF. CO₂ is shown as cyan sticks. Dashed lines indicate hydrogen bonds.

CinF in an Rh protein (PDB code 3B9Z; *Nitrosomonas europaea*)²³ (Supplementary Fig. 8). In this structure, the CO₂ is bound by an asparagine on one side and an aspartate on the other, whereas a phenylalanine residue is found below the CO₂. This further corroborates our hypothesis that the CO₂ is bound as suggested by our docking results.

CCRs show several insertions compared to other reductases

The results from size exclusion chromatography and PISA analysis suggest that CinF forms a tetramer. The closest homologous structures of CinF, two putative CCRs (unpublished data; PDB codes: 3HZZ and 3KRT, both with 51% sequence identity), appear to form a similar tetramer. Tetrameric assemblies have also been shown for related proteins such as SsADH (PDB code 1R37, 24% sequence identity)²⁴, whereas others, such as PIG3 (PDB code 2J8Z, 25% sequence identity)²⁵, are dimeric. These dimers correspond to the strong dimer interface observed in CinF, which creates the continuous β -sheet between the cofactor binding domains.

Comparison of CinF to reductases such as PIG3 reveal some remarkable differences between the otherwise similar structures (Fig. 4a). The insertion of several amino acids in the cofactor binding domain of CinF leads to an elongation and twist of helix α 14 in CinF compared to the same helix in PIG3. This twist places the residues (Arg286 and Arg293 in CinF and PIG3, respectively) in

proximity to the CoA of the dimer partner. Another insertion in CinF exists between helix α 5 and sheet β 5, forming the two additional helices α 6 and α 7 as well as another short sheet, β 4. This insertion also has an important role in substrate binding. Another short insertion involved in CoA binding can be found after sheet β 1, leading to the formation of a helix, α 3. CinF contains additional amino acids at its N and C termini, which form two helices each; neither helix, however, is involved in ligand binding or in tetramer formation.

Two residues define the substrate specificity

The three structures of the carboxylase-reductase family are overall very similar, with r.m.s. deviations of 1.3 Å and 1.5 Å for common C α atoms of chain A when CinF is compared with the CCRs from *Streptomyces coelicolor* (PDB code 3KRT) and *Streptomyces collinus* (PDB code 3HZZ), respectively (Fig. 4b). These two proteins are proposed to catalyze the reductive carboxylation of crotonyl-CoA, whereas CinF catalyzes the reaction with 2-octenoyl-CoA, which has a longer acyl chain. CinF has an elongated hydrophobic pocket for the binding of the octenoyl chain, whereas in the other members of the carboxylase-reductase family no such pocket is detectable. Instead, the potential pocket is blocked by the large residues Phe370 and Ile171 in the CCR from *S. coelicolor*, whereas in CinF the corresponding residues Gly362 and Ala163 are small (Fig. 4c).

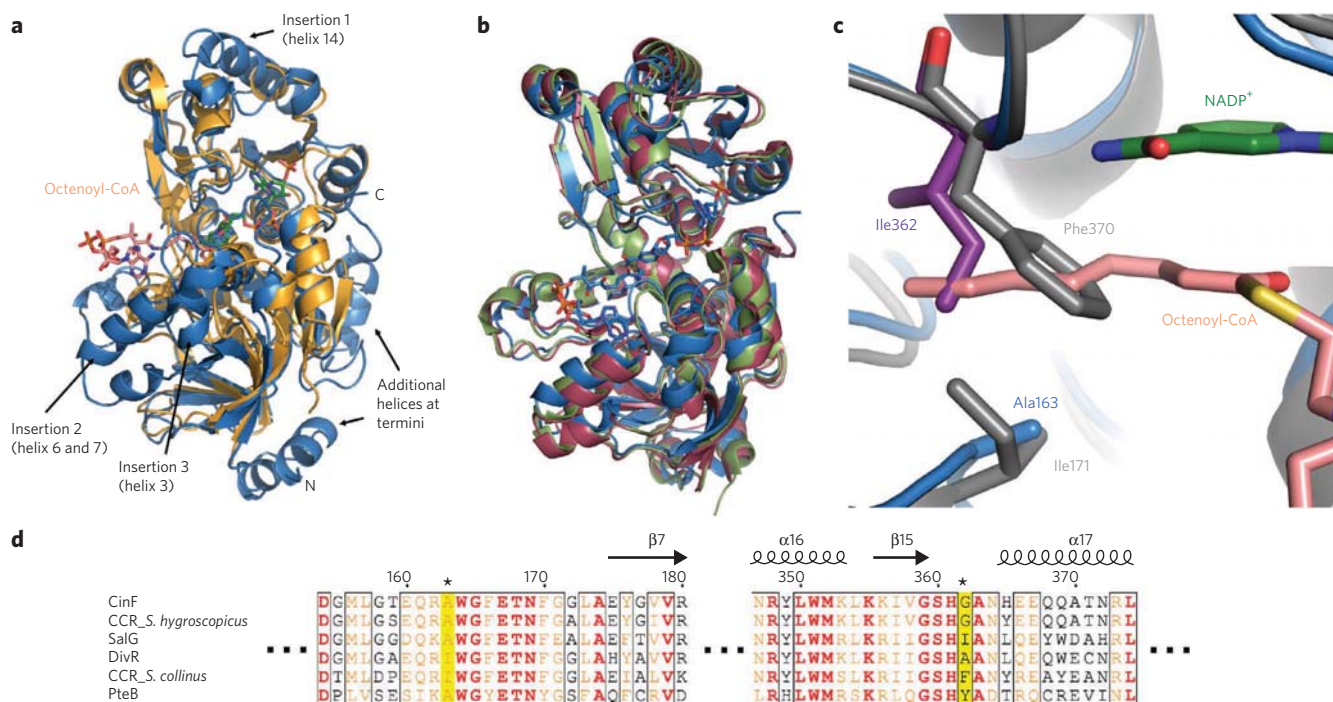


Figure 4 | Comparison of CinF with other structures. (a) Comparison of CinF (blue) with PIG3 (yellow) showing the insertions of CinF. (b) Superposition of CinF (blue) with the CCRs from *S. collinus* (green; PDB code 3HZZ) and *S. coelicolor* (maroon; PDB code 3KRT). The ligands of CinF are shown as blue sticks. (c) Comparison of the substrate binding sites of CinF (blue) and the CCR from *S. collinus* (gray). The binding site of CinF is lined by small residues (Gly362 and Ala163), whereas this binding site is blocked by Ile171 and Phe370 in the CCR from *S. collinus*. This explains why CinF is able to use 2-octenoyl-CoA as a substrate whereas the other CCRs are not. In SalG, instead of the glycine residue an isoleucine is found. This residue was modeled into the ligand-binding site of CinF (purple). It explains why SalG is able to use the longer Cl[−]-crotonyl-CoA instead of the usual crotonyl-CoA. (d) Truncated structure-based sequence alignment of CinF with other carboxylase-reductase proteins. The sequences of CinF (*S. coelicolor*), a putative 2-octenoyl-CoA carboxylase-reductase (*S. hygroscopicus*), SalG (*S. tropica*), DivR (*Streptomyces* sp. HK10576) and a putative CCR (*S. collinus*, PDB code 3KRT) and the 2-octenoyl-CoA carboxylase-reductase PteB (*S. avermitilis*) were aligned with ClustalW1 and displayed with ESPr2.22. The positions of the residues determining the substrate specificity are marked in yellow and with an asterisk.

Thus, these two positions seem to determine the substrate specificity in carboxylases and reductases.

A putative CCR from *Streptomyces hygroscopicus* (GenBank Protein accession number AAR32675), which has 90% sequence identity with CinF, is probably also able to use 2-octenoyl-CoA as a substrate as none of the differing amino acids are involved in ligand binding (Fig. 4d and Supplementary Fig. 9).

A recent report describes PteB from *Streptomyces avermitilis* as a 2-octenoyl-CoA carboxylase-reductase with relaxed substrate specificity¹⁹. In its ligand-binding site it shares the small alanine with CinF (Ala163). However, instead of also sharing a glycine with CinF, PteB has the large amino acid tyrosine (Fig. 4d). Therefore, it is probable that the geometry of the substrate binding pocket of PteB deviates slightly from that of CinF because it has several insertions and deletions relative to CinF that accommodate the relatively long 2-octenoyl-CoA.

Another protein with a substrate other than crotonyl-CoA is SalG (70% sequence identity) from *Salinispora tropica*, which is involved in the production of salinosporamides¹¹. SalG also deviates in the two positions previously described to be important for substrate specificity (Fig. 4d and Supplementary Fig. 9). SalG also has an alanine in the same position as CinF (Ala163), but instead of the Gly362 in CinF or the Phe370 in the CCR from *S. coelicolor*, it has an isoleucine. *In silico* modeling of this mutation shows that the size of the substrate binding pocket of CinF is reduced, so that instead of 2-octenoyl-CoA only pentenoyl-CoA will fit into the pocket (Fig. 4c). This agrees well with the data of Eustaquio *et al.* (2009)¹¹ that show chlorocrotonyl-CoA is the native substrate and is the same size as pentenoyl-CoA.

Recently, it has also been proposed that branched extender units can be introduced into polyketides by the CCR DivR (GenBank

Table 1 | Apparent kinetic constants of CinF, CinF^{G362F} and CinF^{A163I} for reductive carboxylation activity

Enzymes	Substrates	k_{cat} (min ^{−1})	K_{M} (mM)	$k_{\text{cat}}/K_{\text{M}}$ (min ^{−1} mM ^{−1})	Relative catalytic efficiency (%)
CinF	Crotonyl-CoA	82.6 ± 4.4	0.954 ± 0.050	86.6	100
	2-Octenoyl-CoA	17.4 ± 1.1	0.230 ± 0.010	75.7	75.7
	2-Octenoyl-SNAC	16.6 ± 1.1	0.387 ± 0.030	42.9	42.9
CinF ^{G362F}	Crotonyl-CoA	24.1 ± 2.2	0.728 ± 0.046	33.1	38.3
CinF ^{A163I}	Crotonyl-CoA	23.2 ± 1.8	0.998 ± 0.058	23.25	26.8

Original kinetic data are shown in Supplementary Figure 11. Each measurement was performed in three independent experiments. Data represent mean values ± s.d.

Table 2 | Apparent kinetic constants of CinF and its mutated variants for reductive activity versus 2-octenoyl-CoA

Enzymes	k_{cat} (min^{-1})	K_{M} (mM)	$k_{\text{cat}}/K_{\text{M}}$ ($\text{min}^{-1} \text{mM}^{-1}$)	Relative catalytic efficiency (%)
CinF	19.4 ± 1.3	0.883 ± 0.050	30.0	100.0
CinF ^{E167A}	20.1 ± 1.8	0.796 ± 0.09	25.3	84.1
CinF ^{E167K}	20.5 ± 1.1	0.807 ± 0.12	25.4	84.2
CinF ^{E167Y}	21.0 ± 2.3	0.812 ± 0.04	25.9	85.8
CinF ^{N77A}	22.5 ± 1.4	1.200 ± 0.10	18.8	62.1
CinF ^{N77E}	24.1 ± 1.3	1.200 ± 0.08	20.1	66.6
CinF ^{N77Y}	22.7 ± 1.1	0.980 ± 0.08	23.2	76.9
CinF ^{G362F}	18.8 ± 1.2	0.532 ± 0.030	35.3	117.6
CinF ^{A163I}	19.4 ± 0.5	0.621 ± 0.020	31.2	104.0

Original kinetic data are shown in **Supplementary Figure 11**. Each measurement was performed in three independent experiments. Data represent mean values \pm s.d.

Gene accession number NP_824066, 68% sequence identity)¹⁷. DivR deviates from the typical CCRs in the positions that are responsible for substrate specificity (**Fig. 4d** and **Supplementary Fig. 9**). Instead of a phenylalanine, DivR has an alanine, Ala363, but it has an isoleucine at position 164 as in typical CCRs. Therefore, it should be possible for DivR to accept larger aliphatic chains in the substrate, such as the isobutenoyl moiety it has been proposed to use.

CinF converts 2-octenoyl-CoA to hexylmalonyl-CoA

It has been proposed that CinF is responsible for providing hexylmalonyl-CoA for the synthesis of cinnabaramides⁹. To test this hypothesis, we incubated purified protein with 2-octenoyl-CoA, crotonyl-CoA and 2-octenoic acid N-acetylcysteamine thioester (2-octenoyl-SNAC) in the presence of hydrogen carbonate and NADPH and analyzed the corresponding products. High-Resolution MS (HRMS) analysis (**Supplementary Fig. 10**) showed that after incubation of crotonyl-CoA with CinF, a new compound with a retention time of 4.9 min and a m/z $[M+H]^+$ of 882.155 could be detected, corresponding to the expected product ethylmalonyl-CoA (calculated m/z : 882.150). After incubation with 2-octenoyl-CoA, CinF generated a new compound with a retention time of 11.4 min and a m/z $[M+H]^+$ of 938.194 corresponding to the expected mass of hexylmalonyl-CoA (calculated m/z : 938.193). Incubation of 2-octenoyl-SNAC yielded a new substance with a retention time of 5.2 min and a m/z $[M+H]^+$ of 290.142 corresponding to hexylmalonyl-SNAC (calculated m/z : 290.141).

To determine the specificity of CinF toward different substrates, we conducted kinetic studies with CinF (**Supplementary Fig. 11**). Comparison of the k_{cat} and K_{M} values for the different substrates showed that CinF is able to process all three substrates with similar efficiency. However, it has a higher substrate specificity for crotonyl-CoA than for the other substrates. CinF had a lower K_{M} value and thus a higher affinity for 2-octenoyl-CoA and 2-octenoyl-SNAC than for crotonyl-CoA (**Table 1**). It is interesting to note that 2-octenoyl-CoA and 2-octenoyl-SNAC are converted at nearly the same rate. It therefore seems that the CoA only increases affinity for the protein by establishing further interactions, whereas the reaction velocity is mostly determined by the acyl of the substrate.

CinF mutants show altered substrate specificity

To investigate the importance of the Gly362 located within the ligand-binding pocket for substrate specificity, we mutated this residue to phenylalanine as found in typical CCRs or to isoleucine as found in SalG¹¹. Another residue, Ala163, was also mutated to isoleucine as can be found in typical CCRs. Kinetic studies were performed with these three CinF variants using the preferred CinF substrate 2-octenoyl-CoA, 2-octenoyl-SNAC and the native CCR substrate crotonyl-CoA. All variants lost their ability to carboxylate

2-octenoyl-CoA and 2-octenoyl-SNAC completely, revealing the crucial role of these amino acids for substrate specificity (**Table 1**). CinF^{G362F} and CinF^{A163I} had weak catalytic activity toward crotonyl-CoA (**Table 1**), whereas CinF^{G362I} was found inactive in this assay against all of the applied substrates. It was therefore possible to suppress the 2-octenoyl-CoA-converting activity with one single mutation in the active site while retaining some of its crotonyl-CoA-converting activity.

Docking of CO₂ into the structure of CinF suggested that residues Glu167 and Asn77 are involved in the binding and activation of CO₂. To verify this hypothesis, we tested several mutants of these residues (CinF^{E167A}, CinF^{E167K}, CinF^{E167Y}, CinF^{N77A}, CinF^{N77E} and CinF^{N77Y}) for their ability to perform the reduction and carboxylation reactions. The activity assays showed that all six of these mutated CinF variants completely lost their reductive carboxylation activity but partially retained their reductive activity for 2-octenoyl-CoA (**Supplementary Fig. 12**). In the absence of CO₂, the reductive activity of all variants was kinetically characterized with 2-octenoyl-CoA only (**Table 2**). Remarkably, CinF^{G362F} showed 117% catalytic efficiency in comparison to native CinF. It is possible that CinF^{G362F}, owing to flexibility in the active site, is still able to bind the substrate 2-octenoyl-CoA, albeit with reduced affinity. The resulting shorter residence time of the substrate in the protein may be sufficient to allow transfer of the hydride from the NADPH but not addition of the CO₂, which seems to bind only weakly to the protein. Therefore, a decrease in CinF's affinity for the substrate may abolish carboxylation but not reduction.

These data strongly support our hypothesis regarding the CO₂ binding site as determined by *in silico* analysis on the basis of the crystal structure. All CinF mutants seemed to be properly folded as tested by CD spectroscopy (**Supplementary Fig. 13**).

DISCUSSION

Intrigued by the chemical diversity nature generates in secondary metabolism, we set out to analyze the molecular basis for structural diversity found in the side chains of polyketides. Recent studies suggest that, in addition to known ways to biosynthesize the building blocks required for Claisen ester condensation, there must exist another common way to generate activated malonic acid derivatives that vary in their C2 substituents. The key step on this route is reminiscent of the crotonyl-CoA carboxylase-reductase found in primary metabolism. As this enzyme has not been structurally described, we set out to study CinF and present here, to our knowledge, the first detailed structure analysis of a carboxylase-reductase.

Our structures of CinF showed how its substrate is bound in the active site and allowed the identification of two residues (Gly362 and Ala163) that are responsible for the substrate specificity of this class of enzymes. Mutations (G362F, G362I and A163I) of these

residues render CinF unable to use 2-octenoyl-CoA as a substrate, underlining their importance. It thus seems that by varying the size of the substrate-binding pocket, different substrates can be accommodated, leading to different substrate specificities of these enzymes. These findings are highly interesting for the generation of engineered CCRs to produce drugs incorporating uncommon extender units (described further below).

It has been demonstrated that CCRs carboxylate the substrate on the *re* side to yield a product with (2S) stereochemistry, whereas the hydride transfer from the NADPH is pro-(4R) specific to the *re* face of the substrate⁸. The arrangement of the substrates in the CinF structure clearly agrees with these findings, suggesting a common evolutionary origin of the carboxylase-reductase family and medium-chain reductases. However, CCRs have evolved to bind CO₂, and reductive carboxylation is preferred over sole reduction of the substrate⁸. We used *in silico* docking to identify a potential CO₂ binding site. CO₂ is bound by hydrogen bonds to a glutamate and an asparagine and by hydrophobic interactions with a phenylalanine (Fig. 3d). The hydrogen bonds would increase the electrophilicity of the molecule, thus making it more susceptible to nucleophilic attack from the 2-octenoyl-CoA double bond. The hydrophobic phenylalanine, on the other hand, probably reduces the binding of water in this pocket to prevent the reduction of the 2-octenoyl-CoA without carboxylation. This hypothesis is strongly corroborated by the results of our mutational analysis, leading to the discovery of enzymes that do not carboxylate but can still reduce.

In conclusion, to our knowledge we have for the first time solved a structure of a protein from the CCR family in complex with NADP⁺ and its substrate 2-octenoyl-CoA, of which until now only two apo structures were available. We have been able to show that two residues in the active site are critical for binding and activation of CO₂. In addition, we have demonstrated that CinF is responsible for providing the unusual polyketide building block hexylmalonyl-CoA. Furthermore, we have identified important residues that determine the size of the ligand-binding pocket and thereby the substrate specificity of these proteins. In light of the increasing recognition that polyketide compounds incorporate unusual building blocks, these findings provide a solid basis for attempts to modify the substrate specificity of CCR enzymes toward the generation of new polyketide extender units and, therefore, eventually the production of altered polyketide structures.

However, incorporation of those extender units into polyketides is dependent on acyl-transferase domains required for recognition and loading onto the corresponding carrier protein of the PKS. This issue could be addressed by swapping the respective acyl-transferase domains found in the native PKS to change its specificity, an experimental approach that has increasingly met with success in recent years²⁶. In addition, the subsequent PKS elongation modules need to be capable of processing the altered intermediate, and the unsaturated substrate for the CCR reaction needs to be available. Given these obstacles, the success of biosynthetic engineering may seem unlikely at first sight. However, as has been shown for the cinnabaramide biosynthetic pathway, such engineering approaches are indeed likely to result in the expected altered structures⁹. Acceptance of the unnatural dicarboxylic acid by the PKS and subsequent processing to produce chlorinated analogs was successful, thus demonstrating the validity of the approach.

On the basis of the structural insight presented in this manuscript, the stage is now set for directed engineering of CCRs to accept additional substrates. In the future, identification of even more CCR homologs can be expected in the course of characterizing biosynthetic pathways for polyketides with various side chains (Fig. 1). The knowledge gained will pave the way for learning more about this intriguing class of enzymes and will eventually enable advanced biosynthetic engineering of polyketide pathways to produce new drugs.

METHODS

Amplification, cloning and heterologous expression of *cinF* and its variants. *cinF* and its variants were cloned into the pSTW42 expression vector (a version of the pGEX-6P-1 vector in which the GST tag was replaced by SUMO3 tag from the pSUMO3 vector and the ampicillin resistance gene was replaced by the kanamycin resistance gene) (S. Werner, Helmholtz Institute for Pharmaceutical Research, and R.M., unpublished) and heterologously expressed in *Escherichia coli* BL21 (DE3) cells as described in **Supplementary Methods**.

Cell lysis, protein purification and characterization. The cells were harvested by centrifugation (7,000g, 10 min) and resuspended in binding buffer (50 mM Tris, 200 mM NaCl, 10% glycerol and 10 mM imidazole, pH 7.8). Cell lysis was performed using a cold French Press cell (SLM-Aminco) for two times at 900 p.s.i. Immediately after the lysis, 1/50 volume of 50× protease inhibitor (Roche) was added to the suspension, which was then centrifuged at 4 °C and 12,000g for 30 min.

The soluble fraction was filtered using a 1.2-μm pore filter (Sartorius) and was applied to a 5-ml Histrap (GE Healthcare) column equilibrated with binding buffer. The column was washed with binding buffer and eluted with elution buffer (50 mM Tris, 200 mM NaCl, 10% glycerol and 500 mM imidazole, pH 7.8). The purity of the elution fractions was determined by SDS gel electrophoresis. After immobilized metal-ion affinity chromatography, the buffer of the fractions with the recombinant protein was exchanged with 20 mM Tris pH 8 containing 150 mM NaCl using a PD-10 desalting column (GE Healthcare). To remove the SUMO tag from the recombinant fusion proteins, the samples were treated with SUMO-protease 2 (LifeSensor) according to manufacturer protocol (LifeSensor). After the cleavage, the protein was purified by nickel-nitrilotriacetic acid affinity chromatography. Both fusion protein (SUMO-CinF) and cleaved protein (CinF) were verified by MALDI-TOF-MS (**Supplementary Fig. 2**).

Finally, a calibrated Superdex 200 16/60 size exclusion chromatography column (GE Healthcare), pre-equilibrated with 20 mM Tris pH 8, 100 mM NaCl and 5 mM DTT, was used. The calibration using the gel-filtration calibration kits (GE healthcare) also allowed an estimation of the oligomeric state of the protein. The purity of the CinF fractions was analyzed by SDS-PAGE, and pure fractions were pooled and concentrated up to 10 mg ml⁻¹ with Vivaspinn 30-kDa molecular-weight cutoff concentrators (Sartorius). Further characterization of the protein was performed as described in **Supplementary Methods**.

***In vitro* activity assay.** The catalytic activity of CinF was tested *in vitro* with three different substrates: crotonyl-CoA, 2-octenoyl-CoA and 2-octenoyl-SNAC. Crotonyl-CoA (purity ~90%, HPLC) was purchased from Sigma-Aldrich; 2-octenoyl-CoA and 2-octenoyl-SNAC were chemically synthesized (**Supplementary Methods**), their structures were confirmed by NMR spectroscopy (**Supplementary Figs. 3 and 4**), and their purity was determined as >90% via HPLC as described in **Supplementary Methods**. The reactions were performed using a mixture of 2 mM of each substrate, 33 mM NaHCO₃, 4 mM NADPH, 100 mM Tris pH 7.4 and 0.43 nM CinF (control: heat-inactivated CinF). The reactions were incubated at 30 °C overnight, and the detection of the products was done via HPLC-MS analysis (analytical procedures are detailed in **Supplementary Methods**). The protein was denatured by addition of 0.1 vol trichloroethane and heating at 95 °C for 5 min, and it was removed by centrifugation at 12,000g for 5 min.

Kinetic studies of CinF. Kinetic studies of CinF were performed with three different substrates: crotonyl-CoA, 2-octenoyl-CoA and 2-octenoyl-SNAC. Gradient concentrations from 0.2 mM to 2.0 mM were used for each substrate. The kinetic study was followed spectrophotometrically at 340 nm using a 96-well plate. The reaction mixture (60 μl) contained 100 mM Tris-HCl buffer pH 8, 4 mM NADPH, 33 mM NaHCO₃ and the corresponding substrates at varying concentrations. The reaction was started by adding approximately 0.1 nM of purified CinF. The measurement was performed for 1 h. In addition, a standard curve of NADPH was made from 0.125 mM to 4 mM. This allowed us to follow the NADPH consumption, from which the velocity of the reaction and Michaelis-Menten curves can be calculated and generated via the software SigmaPlot Enzyme kinetics 1.3.

CD spectroscopy. The correct folding of CinF and its mutants was verified by CD using 0.2 mg ml⁻¹ protein in 10 mM Na₂HPO₄ pH 8, 10 mM NaCl and 2 mM dithiothreitol. Spectra were recorded from 190 nm to 260 nm and showed that CinF and all mutants were most likely properly folded.

Crystallization of CinF. 96-well sitting-drop crystal screens were set up at 291 K. Initial thin plate-like crystals were found with 0.1 M *N*-cyclohexyl-2-aminoethanesulfonic acid (CHES) buffer pH 9.5 and 10% PEG 3000. Reproduction of the plate-like crystals was only made possible by increasing the PEG concentration to 20% and streak seeding from the initial crystals. Streak seeding into the drops of the initial screen that remained clear produced several new crystallization conditions. Of these conditions, 0.1 M Tris pH 7, 20% glycerol and 12% PEG 8000 was chosen for reproduction of crystals in hanging-drop vapor-diffusion plates. To obtain large crystals, a mixture of 1 μl protein solution (10 mg ml⁻¹ in 20 mM Tris pH 8, 100 mM NaCl and 5 mM DTT) and 1 μl reservoir was used. The drop

containing the protein reservoir mixture was equilibrated against 500 μ l of reservoir solution. After 1 d, the drops were streak seeded using a horse-tail hair. Initial crystals grew after 1 d and reached their final size within 3 d. The crystals of CinF in complex with its substrates were produced by cocrystallization using a solution containing 7 mg ml⁻¹ CinF, 10 mM NADP⁺, 10 mM 2-octenoyl-CoA and 50 mM NaHCO₃.

Data collection, processing and structure determination. For data collection, the crystals were flash cooled in liquid nitrogen. An apo-CinF X-ray diffraction dataset up to 2.0 Å was collected at 100 K with the European Synchrotron Radiation Facility (Grenoble, France) beamline ID 29, equipped with an ADSC Q315R detector at a wavelength of 0.97 Å. A Micromax-007 HF rotating copper anode X-ray generator (Rigaku) with a Saturn 944+ CCD detector (Rigaku) was used to measure the dataset of CinF with NADP⁺ and 2-octenoyl-CoA diffracting to 1.9 Å. The structure was solved by molecular replacement using the unpublished structure of a CCR (PDB code 3KRT, submitted by V.N. Malashkevich, Y. Patskovsky, R. Toro, M.J. Sauder, S.K. Burley, S.C. Almo) as a model.

Detailed data collection and refinement methods are described in the **Supplementary Methods**, and statistics are listed in **Supplementary Table 2**. The final *R* values were 21.8%/27.6% (*R*_{work}/*R*_{free}) for the apo form and 19.9%/24.6% for the ligand-bound form.

Accession codes. Protein Data Bank: the structures for apoCinF and CinF with ligands are deposited under accession codes 4A10 and 4A0S. The previously published structures of the Rh protein from *N. europaea*, a CCR from *S. coelicolor*, a CCR from *S. collinus*, SsADH and PIG3 are deposited under accession codes 3B9Z, 3KRT, 3HZZ, 1R37 and 2J8Z, respectively.

Received 16 May 2011; accepted 21 September 2011;
published online 4 December 2011

References

- Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Edn Engl.* **48**, 4688–4716 (2009).
- Newman, D.J. & Cragg, G.M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **70**, 461–477 (2007).
- Staunton, J. & Weissman, K.J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18**, 380–416 (2001).
- Wenzel, S.C. *et al.* On the biosynthetic origin of methoxymalonyl-acyl carrier protein, the substrate for incorporation of “glycolate” units into ansamitocin and soraphen A. *J. Am. Chem. Soc.* **128**, 14325–14336 (2006).
- Dorrestein, P.C. *et al.* The bifunctional glycyl transferase/phosphatase OzmB belonging to the HAD superfamily that diverts 1,3-bisphosphoglycerate into polyketide biosynthesis. *J. Am. Chem. Soc.* **128**, 10386–10387 (2006).
- Chan, Y.A. *et al.* Hydroxymalonyl-acyl carrier protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units. *Proc. Natl. Acad. Sci. USA* **103**, 14349–14354 (2006).
- Erb, T.J. *et al.* Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc. Natl. Acad. Sci. USA* **104**, 10631–10636 (2007).
- Erb, T.J., Brecht, V., Fuchs, G., Muller, M. & Alber, B.E. Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Natl. Acad. Sci. USA* **106**, 8871–8876 (2009).
- Rachid, S. *et al.* Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. *ChemBioChem* **12**, 922–931 (2011).
- Buntin, K. *et al.* Biosynthesis of thuggacins in myxobacteria: comparative cluster analysis reveals basis for natural product structural diversity. *Chem. Biol.* **17**, 342–356 (2010).
- Eustáquio, A.S. *et al.* Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine. *Proc. Natl. Acad. Sci. USA* **106**, 12295–12300 (2009).
- Mo, S. *et al.* Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J. Am. Chem. Soc.* **133**, 976–985 (2011).
- Goranovic, D. *et al.* Origin of the allyl group in FK506 biosynthesis. *J. Biol. Chem.* **285**, 14292–14300 (2010).
- Qu, X. *et al.* Cloning, sequencing and characterization of the biosynthetic gene cluster of sanglifehrin A, a potent cyclophilin inhibitor. *Mol. Biosyst.* **7**, 852–861 (2011).
- Kopp, M. *et al.* Insights into the complex biosynthesis of the leupyrrins in *Sorangium cellulosum* So ce690. *Mol. Biosyst.* **7**, 1549–1563 (2011).
- Song, L. *et al.* Type III polyketide synthase I²-ketoacyl-ACP starter unit and ethylmalonyl-CoA extender unit selectivity discovered by *Streptomyces coelicolor* genome mining. *J. Am. Chem. Soc.* **128**, 14754–14755 (2006).
- Xu, Z., Ding, L. & Hertweck, C. A branched extender unit shared between two orthogonal polyketide pathways in an endophyte. *Angew. Chem. Int. Edn Engl.* **50**, 4667–4670 (2011).
- Wilson, M.C. *et al.* Structure and biosynthesis of the marine *Streptomyces* ansamycin ansalactam A and its distinctive branched chain polyketide extender unit. *J. Am. Chem. Soc.* published online, doi:10.1021/ja109226s (19 January 2011).
- Yoo, H.G. *et al.* Characterization of 2-octenoyl-CoA carboxylase/reductase utilizing pteB from *Streptomyces avermitilis*. *Biosci. Biotechnol. Biochem.* **75**, 1191–1193 (2011).
- Xu, L.H. *et al.* Regio- and stereospecificity of filipin hydroxylation sites revealed by crystal structures of cytochrome P450 105P1 and 105D6 from *Streptomyces avermitilis*. *J. Biol. Chem.* **285**, 16844–16853 (2010).
- Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
- Verdonk, M.L. *et al.* Modeling water molecules in protein-ligand docking using GOLD. *J. Med. Chem.* **48**, 6504–6515 (2005).
- Li, X., Jayachandran, S., Nguyen, H.H. & Chan, M.K. Structure of the *Nitrosomonas europaea* Rh protein. *Proc. Natl. Acad. Sci. USA* **104**, 19279–19284 (2007).
- Espósito, L. *et al.* Crystal structure of a ternary complex of the alcohol dehydrogenase from *Sulfolobus solfataricus*. *Biochemistry* **42**, 14397–14407 (2003).
- Porté, S. *et al.* Three-dimensional structure and enzymatic function of proapoptotic human p53-inducible quinone oxidoreductase PIG3. *J. Biol. Chem.* **284**, 17194–17205 (2009).
- Weissman, K.J. & Leadlay, P.F. Combinatorial biosynthesis of reduced polyketides. *Nat. Rev. Microbiol.* **3**, 925–936 (2005).

Acknowledgments

D.W.H. acknowledges support from the Fonds der Chemischen Industrie. Research in R.M.'s laboratory was funded by the Deutsche Forschungsgemeinschaft and the Bundesministerium für Bildung und Forschung (BMBF). We thank V. Wray for critically reviewing this manuscript. We also would like to thank T. Hoffmann for technical support, K. Harmrolfs for the NMR measurement and S.C. Wenzel for advice during the course of this project. In addition, we appreciate the work of A. Ullrich and K. Schultz from U. Kazmaier's work group for the syntheses of 2-octenoyl-CoA and 2-octenoyl-SNAC.

Author contributions

N.Q. performed protein crystallization, crystallographic data collection and structure determination. L.H. performed gene cloning, protein heterologous expression and purification experiments as well as *in vitro* characterization of proteins and generation of corresponding point mutations. S.R. contributed initial advice regarding protein expression. R.M. and D.W.H. designed the study, and R.M. wrote the manuscript together with N.Q., L.H. and D.W.H. All authors discussed the results and commented on the manuscript.

Competing financial interests

The authors declare no competing financial interests

Additional information

Supplementary information is available online at <http://www.nature.com/naturechemicalbiology/>. Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Correspondence and requests for materials should be addressed to D.W.H. or R.M.

Unusual Carbon Fixation Giving Rise to Diverse Polyketide Extender Units

Nick Quade^{1,+}, Liujie Huo^{2,+}, Shwan Rachid², Dirk W. Heinz^{1,*} & Rolf Müller^{2,*}

¹Department of Molecular Structural Biology, Helmholtz Centre for Infection Research, D-38124 Braunschweig, Germany

²Department of Microbial Natural Products, Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Campus C2 3, D-66123 Saarbrücken, Germany.

⁺These authors contributed equally to this work.

* Corresponding authors: dirk.heinz@helmholtz-hzi.de; rom@helmholtz-hzi.de

Supplementary Material

-Supplementary Results

-Supplementary Tables

-Supplementary Figures

-Supplementary Methods

Supplementary Results

Supplementary Tables

Supplementary Table 1 Schematic display of CinF and its functional homologs

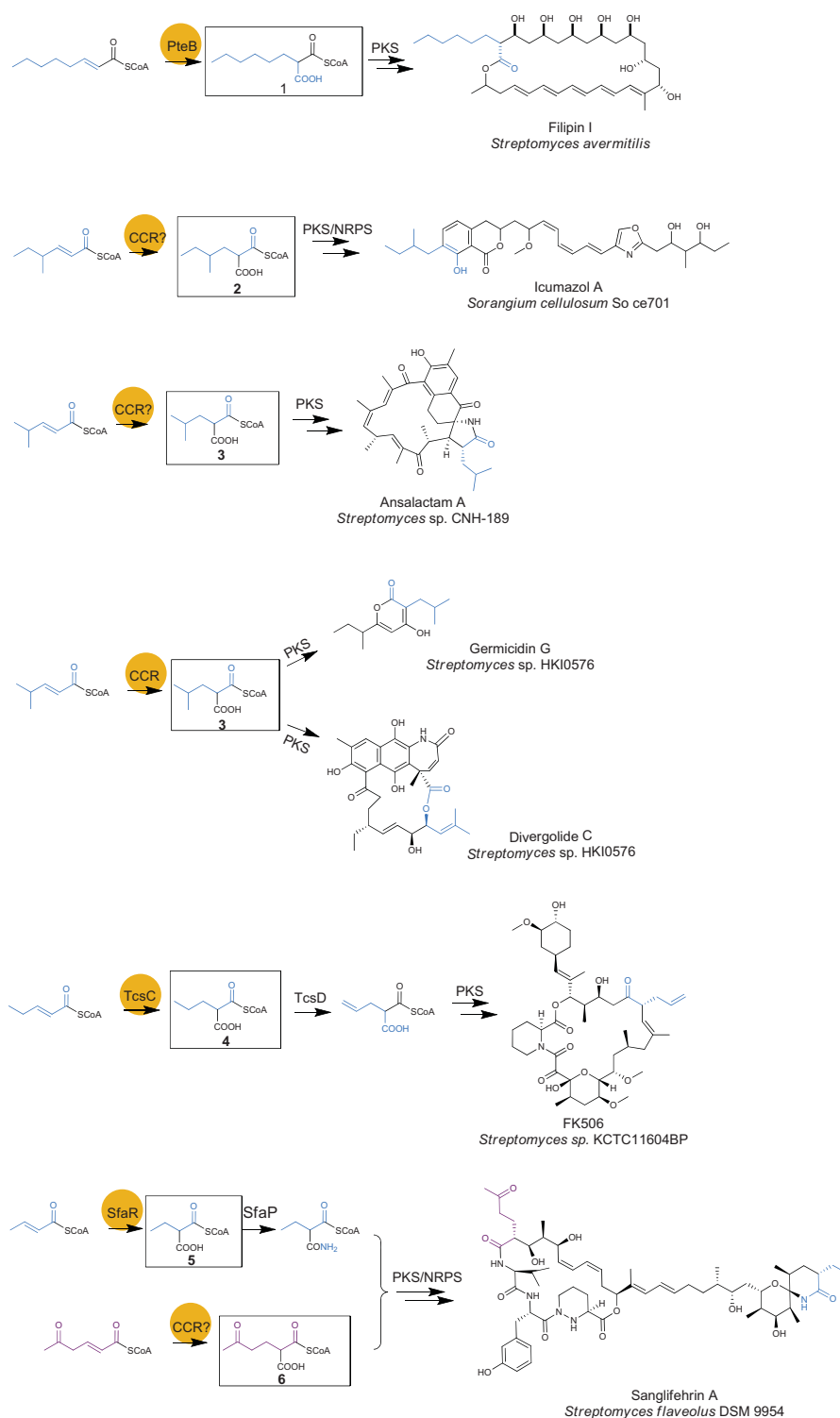
Protein	Residues (aa)	Proposed function	Sequence similarity, identity to CinF*	Accession number
CinF	449	2-Octenoyl-CoA reductase/carboxylase	-	CBW54676
TgaD	388	2-Octenoyl-CoA reductase/carboxylase	30%, 16%	ADH04642
Leu10	399	2-methylhexenoyl-CoA reductase/carboxylase	5%, 10%	ADZ24994
SalG	461	Chlorocrotonyl-CoA reductase/carboxylase	81%, 67%	ABP73651
TcsC	445	Pentenoyl-CoA reductase/carboxylase	63%, 51%	ADU56308
SfaR	454	Crotonyl-CoA reductase/carboxylase	66%, 53%	ACY06299

*: Geneious alignment (cost matrix: BLOSUM62, Gap open penalty: 11, Gap extension penalty: 3)

Supplementary Table 2 CinF data collection and refinement statistics

	Apo	2-Octenoyl-CoA/NADP complex
Data collection		
Space group	P2 ₁	P2 ₁
Unit cell dimensions		
a, b, c (Å)	101.6, 85.2, 113.9	96.0, 83.3, 122.7
α, β, γ (°)	90, 107.4, 90	90, 110.9, 90
Resolution (Å)	48.5-2.25 (2.4-2.25)	47.2-1.9 (2.0-1.9)
R _{merge} (%)	14.48 (47.7)	8.7 (42.6)
I/ σ _I	9.97 (5.49)	13.85 (3.6)
Completeness (%)	98.76 (98.85)	98.8 (94.1)
Redundancy	3.3 (3.2)	3.6 (3.2)
Refinement		
R _{work} / R _{free} (%)	20.9/27.3	19.9/24.6
Unique reflections	87221 (15214)	140724 (18977)
Atom numbers		
- Protein	12550	13389
- Ligand	-	400
- Solvent	1515	1905
Average B-factor (Å ²)		
- Protein (chains A, B, C, D)	5.7, 6.4, 9.5, 6.9	3.4, 4.3, 4.1, 4.9
- NADP	-	15.3, 16.4, 25.2, 22.8
- Oct-CoA	-	33.6, 39.1, 75.9, 55.8
- Solvent	18.4	14.2
R.m.s.d. from ideal geometry		
- Bond lengths (Å)	0.0103	0.0106
- Bond angles (°)	1.31	1.61

Supplementary Figures

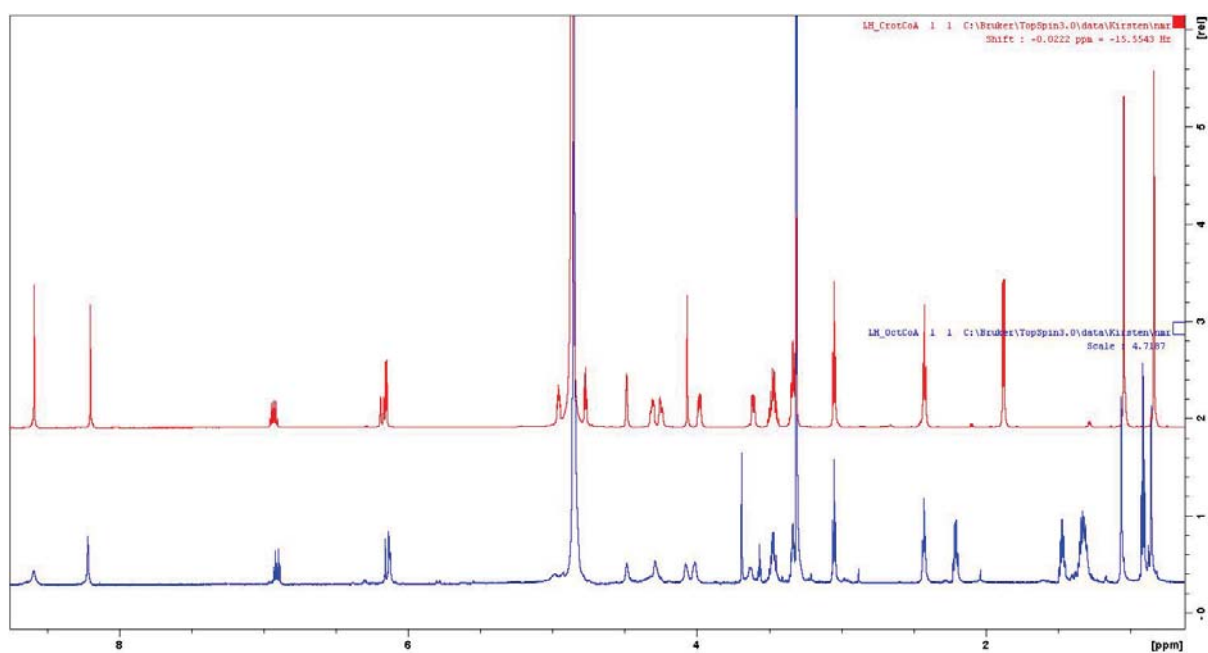


Supplementary Figure 1: Complementary examples of unusual alkylmalonyl thioester building blocks generated by reductive carboxylation and incorporated into bacterial polyketide metabolites. CinF homologues are circled in dark yellow and their likely products are framed in black: 1: hexyl-malonyl-CoA; 2: 2-(2-methylbutyl)-malonyl-CoA; 3: 2-isobutyryl-malonyl-CoA; 4: propyl-malonyl-CoA; 5: ethyl-malonyl-CoA; 6: 2-(2-oxobutyl)-malonyl-CoA.

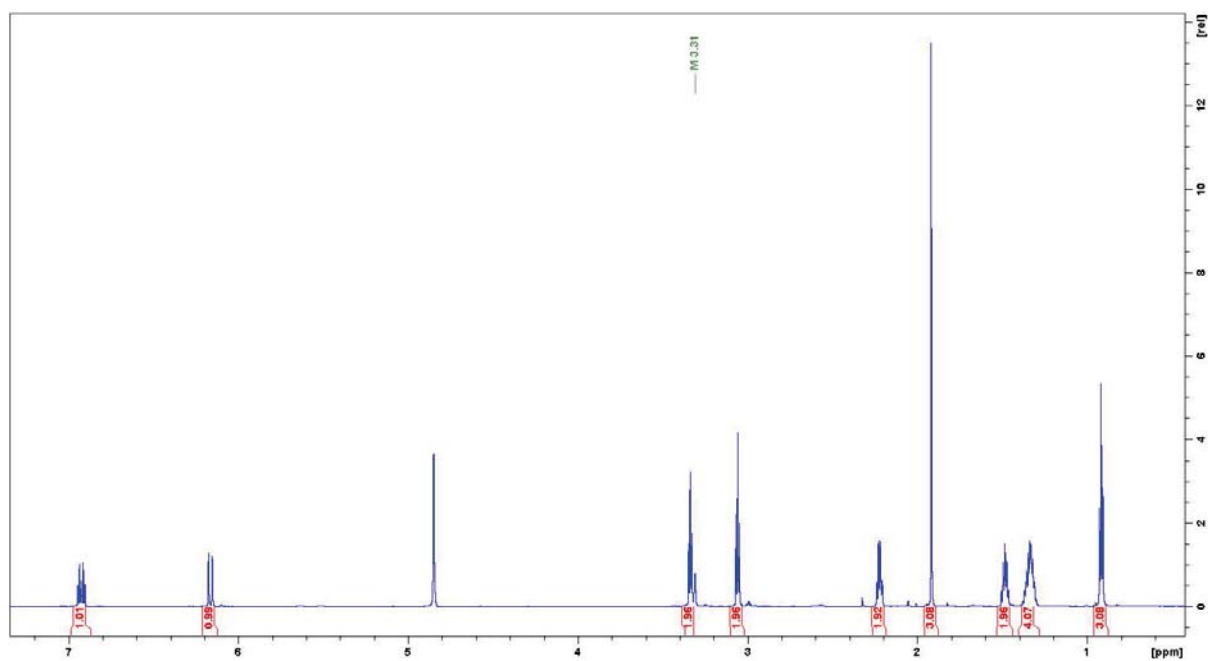
Rank	Protein Name	Accession No.	Protein MW	Protein PI	Pep. Count	Protein Score	Protein Score C. I. %	Total Ion C. I. %
1	CinG_1*	CinG_1	47940,6	6,36	12	78	99,985	
Peptide Information								
	Calc. Mass	Obsrv. Mass	\pm da	\pm ppm	Start Seq.	End Sequence Seq.	Modification	
	816,4322	816,4442	0,012	0	429	435 ARLGEDR		
	1182,6589	1182,6244	-0,0345	0	431	440 LGEDRLNPLR		
	1235,6235	1235,5858	-0,0377	0	209	219 MLVSDRGAQMK		
	1298,7539	1298,7098	-0,0441	0	436	447 LNPLRGLTATSR		
	1500,7177	1500,6704	-0,0473	0	273	286 AELGITDDIADDP		
	1583,7925	1583,7418	-0,0507	0	303	316 AGREPDIVFEHTGR		
	1656,8187	1656,7622	-0,0565	0	273	287 AELGITDDIADDP		
	1847,8743	1847,8234	-0,0509	0	357	373 IVGSHGANHEEQATNR		
	1910,0453	1909,9719	-0,0734	-1047	243	262 NGGGIPVAVVSSAQKEAAVR		
	1972,9551	1972,8865	-0,0686	0	163	180 AWGFETNFGGLAEYGVVR		
	2663,4038	2663,3235	-0,0803	0	6	31 AVLDGASAAEIEAAPVDPDYLALHL		
	3543,6821	3543,5698	-0,1123	0	131	162 WKPGDHVIVHPAHVDEQEPAATHGD GMLGTEQR		
2	SUMO-CinG_1*	SUMO-CinG_1	59565,2	6,32	12	69	99,886	
Peptide Information								
	Calc. Mass	Obsrv. Mass	\pm da	\pm ppm	Start Seq.	End Sequence Seq.	Modification	
	816,4322	816,4442	0,012	0	531	537 ARLGEDR		
	1182,6589	1182,6244	-0,0345	0	533	542 LGEDRLNPLR		
	1235,6235	1235,5858	-0,0377	0	311	321 MLVSDRGAQMK		
	1298,7539	1298,7098	-0,0441	0	538	549 LNPLRGLTATSR		
	1500,7177	1500,6704	-0,0473	0	375	388 AELGITDDIADDP		
	1583,7925	1583,7418	-0,0507	0	405	418 AGREPDIVFEHTGR		
	1656,8187	1656,7622	-0,0565	0	375	389 AELGITDDIADDP		
	1847,8743	1847,8234	-0,0509	0	459	475 IVGSHGANHEEQATNR		
	1910,0453	1909,9719	-0,0734	-1047	345	364 NGGGIPVAVVSSAQKEAAVR		

Supplementary Figure 2: Results of MALDI-TOF-MS measurements for both fusion protein (SUMO-CinF) and cleaved protein (CinF). Both proteins were identified with a high score.

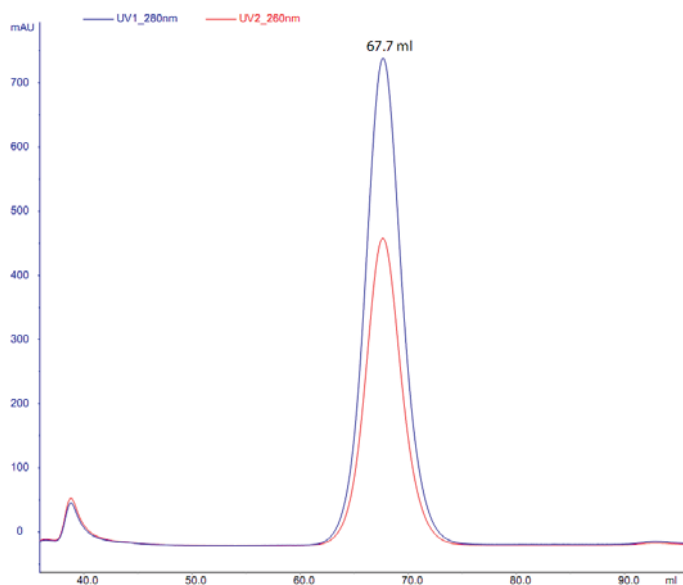
*The name of CinG has been changed to CinF due to changes in the annotation.



Supplementary Figure 3: ¹H-NMR spectrum of 2-octenoyl-CoA (blue) compared to the spectrum of commercially available crotonyl-CoA (red).



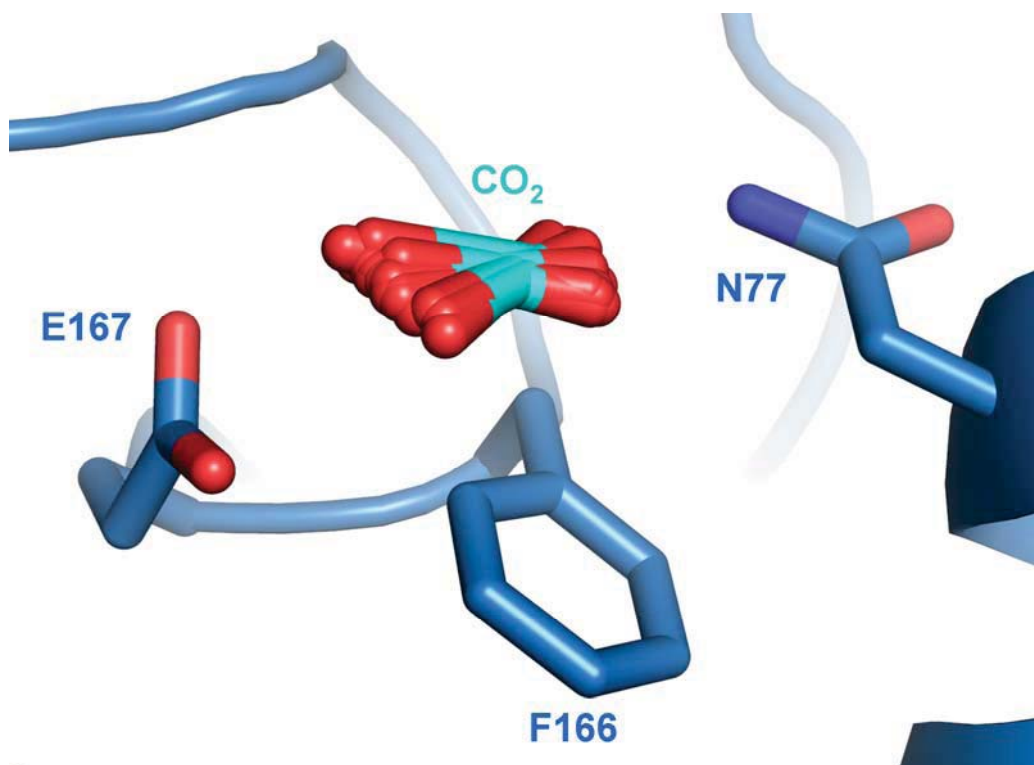
Supplementary Figure 4: ¹H-NMR spectrum of 2-octenoyl-SNAC ester.



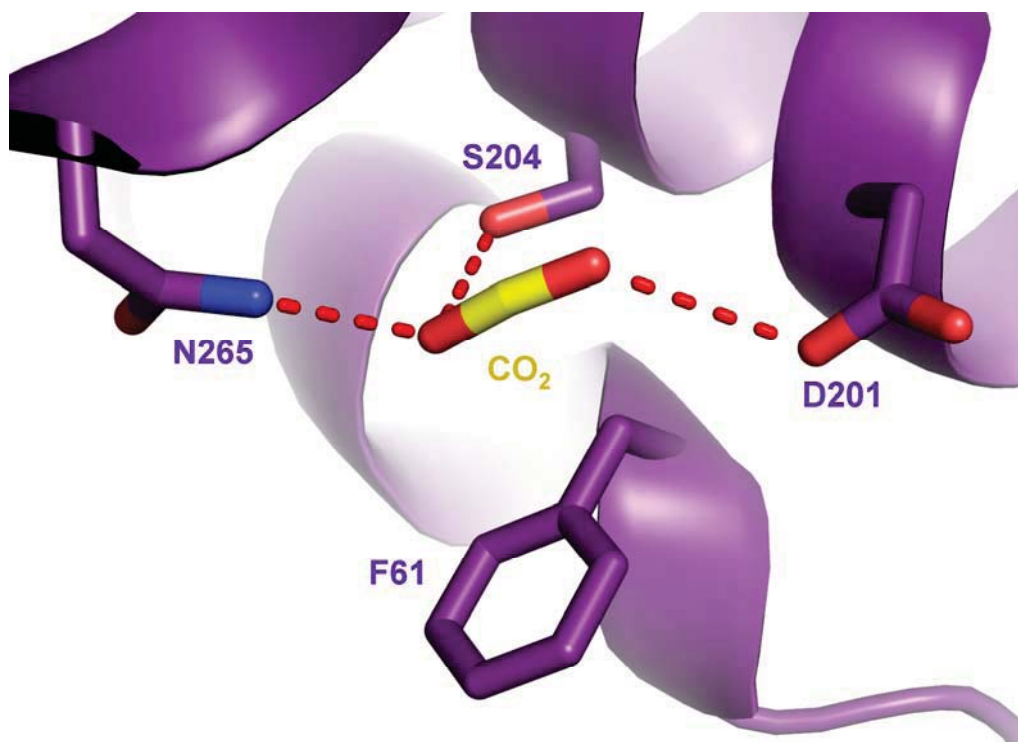
Supplementary Figure 5: Chromatogram of size exclusion chromatography of CinF using a Superdex 200 16/60 column (GE healthcare). Absorption at 280nm is shown in blue, absorption at 260nm in red.

#	Structure 1					x	Structure 2				interface	Δ^iG	Δ^iG	N _{H_B}	N _{S_B}	N _{B_S}	CSS	
Id	NN	$\langle r \rangle$	Range	iNat	iNres		Range	Symmetry op-n	Sym.ID	iNat	iNres	area, Å ²	kcal/mol					P-value
1	1		C	167	40	\diamond	B	x,y,z	1_555	167	41	1480.1	-16.6	0.079	25	0	0	1.000
	2		D	167	40	\diamond	A	x,y,z	1_555	165	39	1473.8	-18.8	0.050	25	0	0	1.000
	Average:										1477.0	-17.7	0.065	25	0	0	1.000	
2	3		B	134	33	\diamond	A	x,y,z	1_555	130	33	1255.5	-0.1	0.773	19	17	0	0.265
	4		D	126	33	\diamond	C	x,y,z	1_555	131	33	1243.8	-0.1	0.772	20	18	0	0.265
	Average:										1249.6	-0.1	0.772	20	18	0	0.265	
3	5		[NAP]F:1	46	1	\diamond	A	x,y,z	1_555	86	34	611.5	-2.2	-1.000	0	0	0	-0.000
	6		[NAP]J:1	48	1	\diamond	C	x,y,z	1_555	83	33	609.1	-1.3	-1.000	0	0	0	-0.000
	7		[NAP]G:1	47	1	\diamond	B	x,y,z	1_555	83	31	608.7	-1.9	-1.000	0	0	0	-0.000
	8		[NAP]I:1	46	1	\diamond	D	x,y,z	1_555	80	35	608.4	-0.6	-1.000	0	0	0	-0.000
	Average:										609.4	-1.5	-1.000	0	0	0	0.000	
4	9		D	62	20	\diamond	B	x,y,z	1_555	63	20	610.5	-1.2	0.616	4	0	0	0.066
	10		C	64	20	\diamond	A	x,y,z	1_555	62	20	602.4	-1.0	0.643	4	0	0	0.066
	Average:										606.5	-1.1	0.630	4	0	0	0.066	
5	11		[DRG]E:1	43	1	\diamond	A	x,y,z	1_555	62	27	509.3	-1.8	0.000	0	0	0	-0.000
	12		[DRG]M:1	41	1	\diamond	D	x,y,z	1_555	63	26	498.1	-2.1	0.000	0	0	0	-0.000
	13		[DRG]K:1	41	1	\diamond	B	x,y,z	1_555	57	25	491.6	-1.9	0.000	0	0	0	-0.000
	14		[DRG]L:1	40	1	\diamond	C	x,y,z	1_555	65	27	479.1	-2.2	0.000	0	0	0	-0.000
	Average:										494.5	-2.0	0.000	0	0	0	0.000	
6	15		B	49	16	\diamond	D	-x+1,y-1/2,-z+1	2_646	40	14	401.1	1.3	0.740	5	4	0	-0.000
	16		C	37	14	\diamond	A	-x,y-1/2,-z	2_545	43	16	351.7	-0.9	0.590	3	4	0	-0.000
	Average:										376.4	0.2	0.665	4	4	0	0.000	

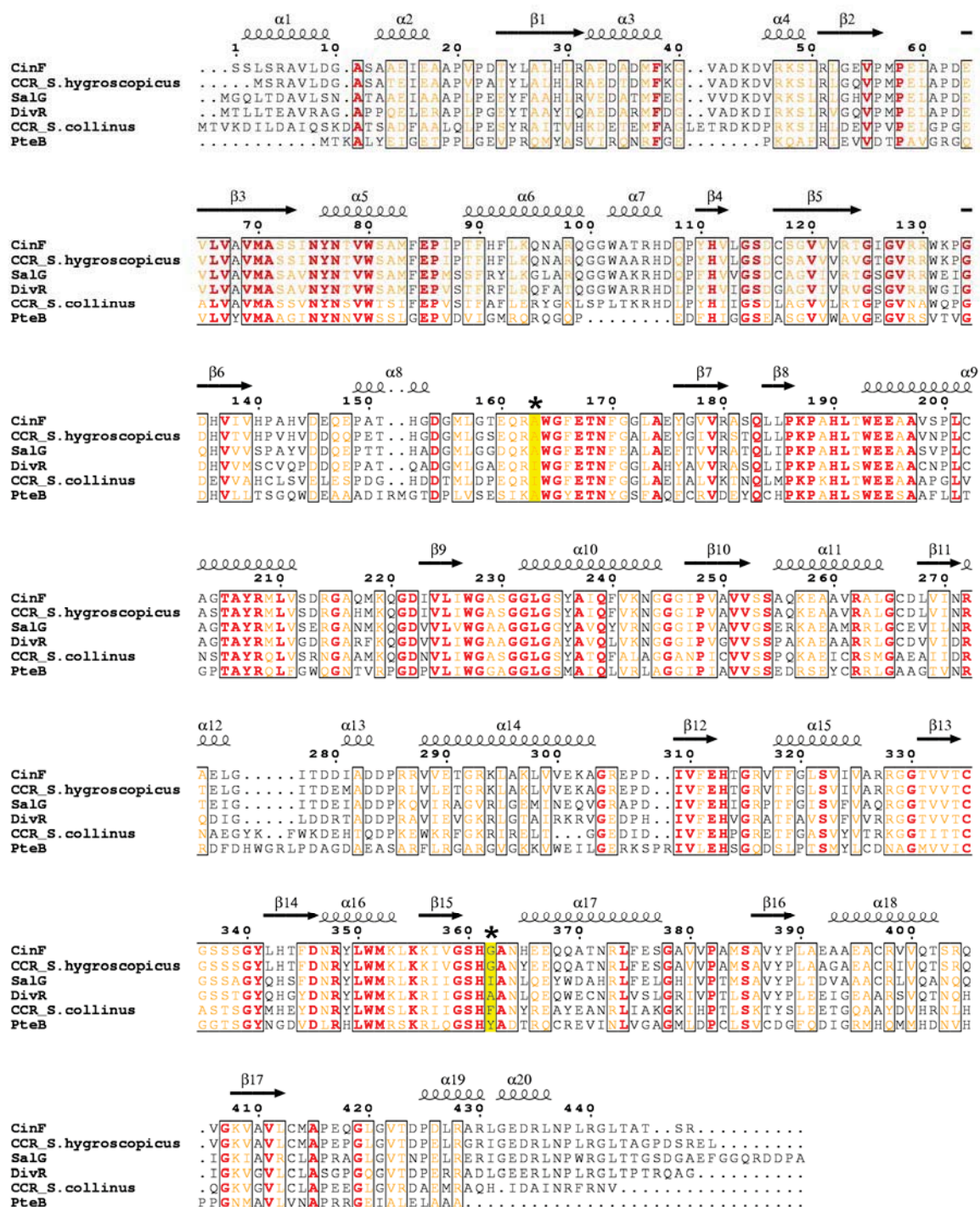
Supplementary Figure 6: Monomer interfaces of CinF as determined by the PISA server¹. Shown are only the most relevant interactions.



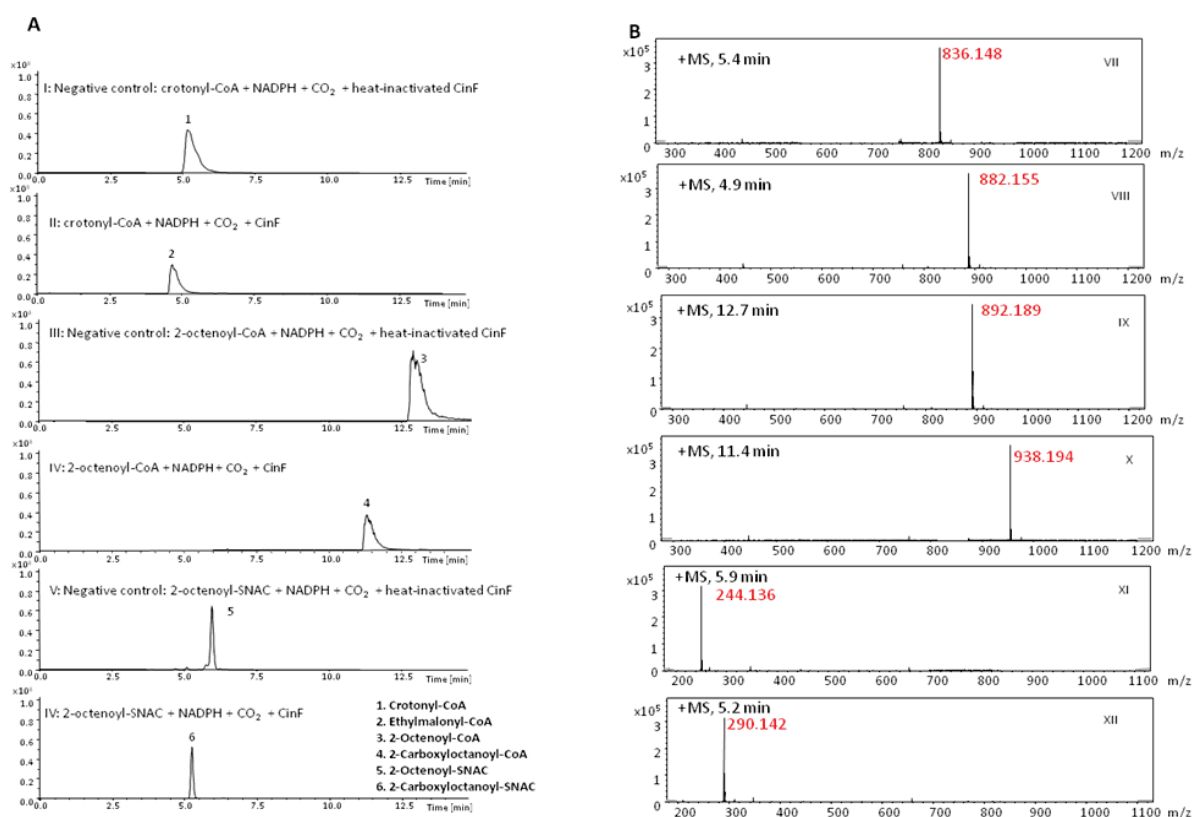
Supplementary Figure 7: *In silico* modelling results for CO₂ binding by CinF. CinF is shown in blue, the CO₂ in cyan.



Supplementary Figure 8: CO₂ binding by Rh protein (pdb code: 3B9Z, *Nitrosomonas europaea*)². Rh protein is shown in purple, CO₂ in yellow.



Supplementary Figure 9: Structure based sequence alignment of CinF with other carboxylase/reductase proteins. The sequences of CinF (*Streptomyces coelicolor*), a putative 2-octenoyl-CoA carboxylase/reductase (*Streptomyces hygroscopicus*), SalG (*Salinospira tropica*), DivR (*Streptomyces* sp. HKI0576) and a putative CCR (*Streptomyces collinus*, pdb code: 3KRT) and the 2-octenoyl-CoA carboxylase/reductase PteB (*Streptomyces avermitilis*) were aligned with ClustalW³ and displayed with ESPrpt2.2⁴. The positions of the residues determining the substrate specificity are marked in yellow and with an asterisk.

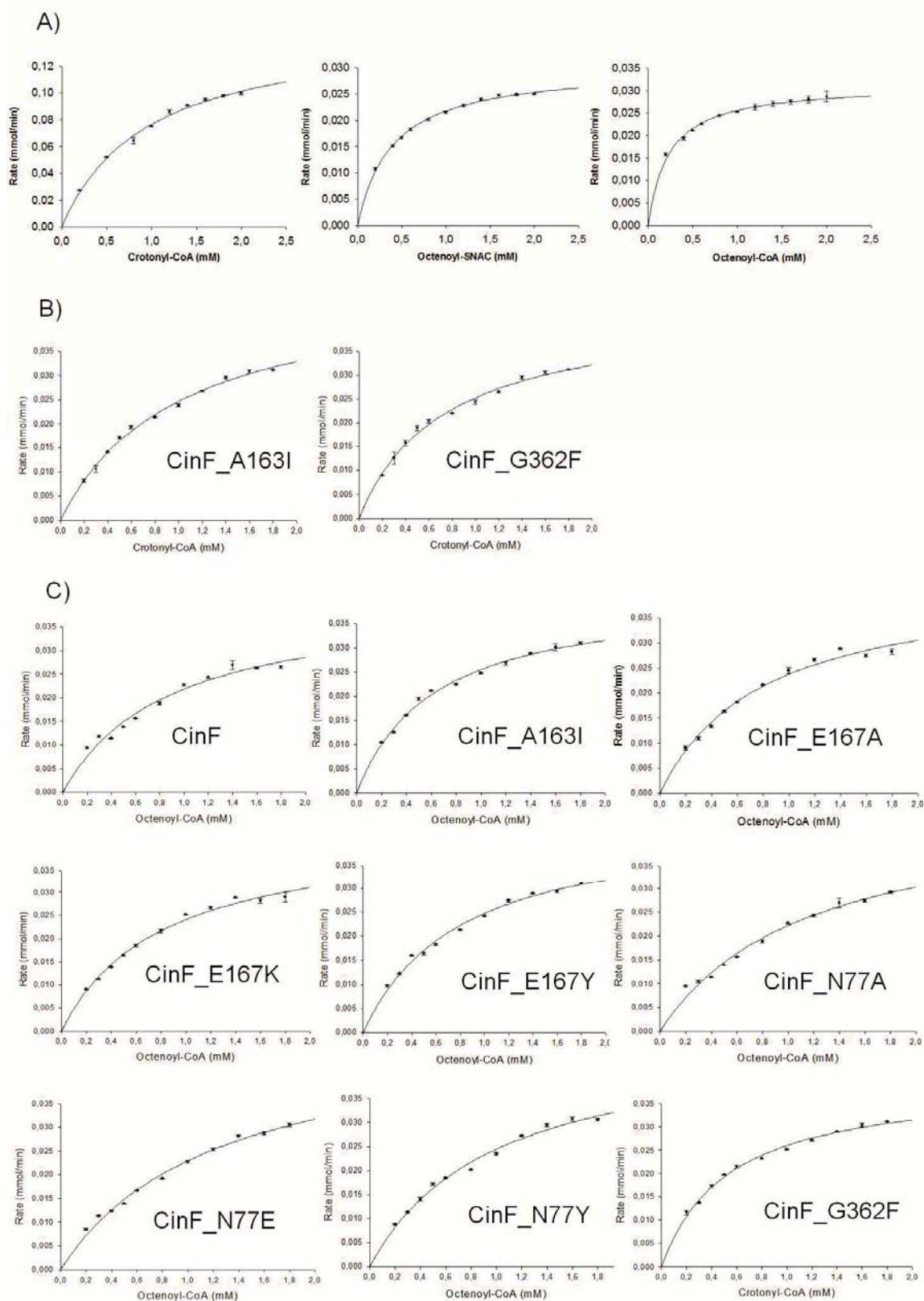


Supplementary Figure 10: High-resolution UPLC-MS analysis of CinF-catalyzed reductive carboxylation of different substrates.

A: I: Shown is the extracted ion chromatogram (EIC) of m/z $[M+H]^+ = 836$ and of m/z $[M+H]^+ = 882$. **II:** Shown is the EIC of m/z $[M+H]^+ = 836$ and of m/z $[M+H]^+ = 882$. **III:** Shown is the EIC of m/z $[M+H]^+ = 892.2$ and m/z $[M+H]^+ = 938.2$. **IV:** Shown is the EIC of m/z $[M+H]^+ = 892.2$ and m/z $[M+H]^+ = 938.2$. **V:** Shown is the EIC of m/z $[M+H]^+ = 244.2$ and m/z $[M+H]^+ = 290.2$. **VI:** Shown is the EIC of m/z $[M+H]^+ = 244.2$ and m/z $[M+H]^+ = 290.2$.

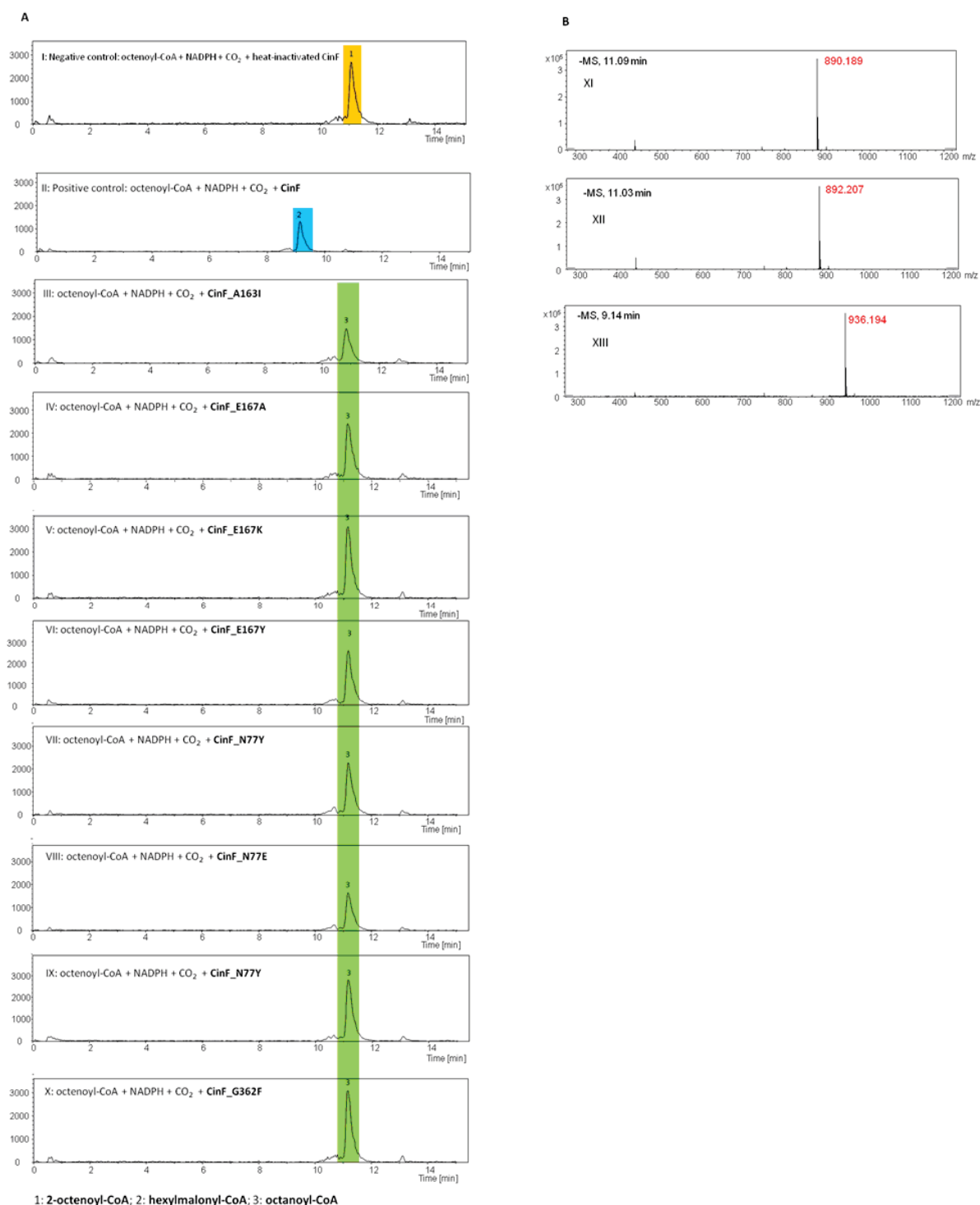
B: Shown is the high-resolution mass spectrum corresponding to substrate and products. **VII** shows the mass spectrum at r.t. 5.4 min of **I**, where the m/z $[M+H]^+ = 836.148$ corresponds to crotonyl-CoA. **VIII** shows the mass spectrum at r.t. 4.9 min of **II**, where the m/z $[M+H]^+ = 882.155$ corresponds to ethylmalonyl-CoA. **IX** shows the mass spectrum at r.t. 12.7 min of **III**, where the m/z $[M+H]^+ = 892.189$ corresponds to 2-octenoyl-CoA. **X** shows the mass spectrum at r.t. 11.4 min of **IV**, where the m/z $[M+H]^+ = 938.194$ corresponds to hexylmalonyl-CoA. **XI** shows the mass spectrum at r.t. 5.9 min of **V**, where the m/z $[M+H]^+ = 244.136$ corresponds to 2-octenoyl-SNAC. **XII** shows the mass spectrum at r.t. 5.2 min of **VI**, where the m/z $[M+H]^+ = 290.142$ corresponds to hexylmalonyl-SNAC.

All compounds were identified by known molecular weight through high-resolution UPLC-MS.



Supplementary Figure: 11 Results of *in vitro* biochemical characterisation of CinF and its variants.

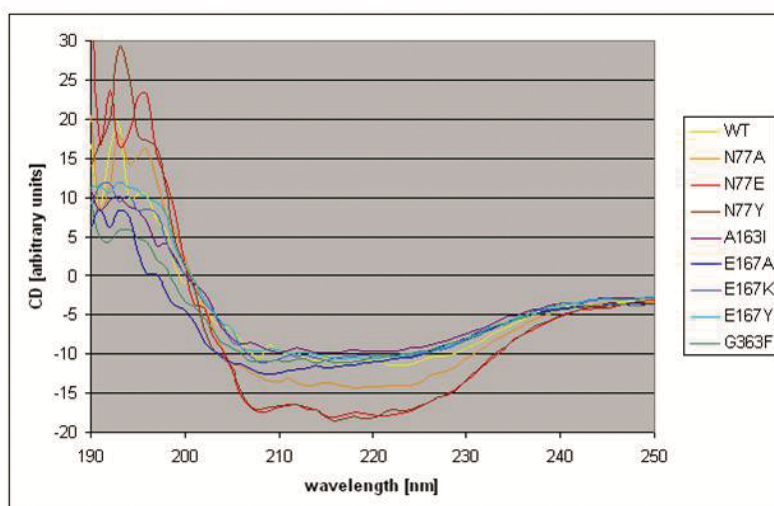
A): Determination of steady-state kinetic parameters of CinF for different thioester substrates in the presence of NADPH and NaHCO_3 . **B)** Determination of steady-state kinetic parameters of CinF_A163I and CinF_G362F variants for crotonyl-CoA in the presence of NADPH and NaHCO_3 . **C)** Determination of steady-state kinetic parameters of CinF and its different variants only for their reductive activities versus octenoyl-CoA in the presence of NADPH and in the absence of NaHCO_3 .



Supplementary Figure 12 High-resolution UPLC-MS analysis of CinF and its variants-catalysed reductive carboxylation of 2-octenoyl-CoA. **A:** Shown is the extracted ion chromatogram (EIC) of m/z $[M-H]^- = 890.2$, 892.2 and 936.2, respectively. In the negative control (**I**) only the deprotonated mass m/z $[M-H]^- = 890.1890$ corresponding to 2-octenoyl-CoA could be observed. In the positive control (**II**) only the deprotonated mass m/z $[M-H]^- = 936.194$ corresponding to the carboxylated product hexyl-malonyl-CoA could be observed. In assays incubated with all other variants (**III-X**) only the deprotonated mass m/z $[M-H]^- = 892.207$ corresponding to the reduced product octanoyl-CoA could be observed.

B: Shown is the high-resolution mass spectrum at different retention times corresponding to substrate and products. **XI** shows the mass spectrum at r.t. 11.09 min of **I**, where the deprotonated mass m/z $[M-H]^- = 890.189$ corresponds to 2-octenoyl-CoA. **XII** shows the mass spectrum at r.t. 11.03 min of **III-X**, where the deprotonated mass m/z $[M-H]^- = 892.207$ corresponds to reduced product octanoyl-CoA. **XIII** shows the mass spectrum at r.t. 9.14 min of **II**, where the deprotonated mass m/z $[M-H]^- = 936.194$ corresponds to carboxylated product hexylmalonyl-CoA.

All compounds were identified by known molecular weight through high-resolution UPLC-MS.



Supplementary Figure 13: Circular dichroism spectra of CinF and its variants.

Supplementary Methods

Cloning and heterologous expression of *cinF* and its variants.

Polymerase chain reaction was performed with primers CinF-up (5'-**TTTAAACGGGAGGTAGCAGCTTGTCCCGAGCC**-3') and CinF-dn (5'- **CTCGAGTCA** CCGGGAGGTGGCGGTGAG-3'; inserted restriction sites are highlighted in bold) to generate *cinF* PCR products with *DraI/XhoI* restriction site overhangs. To generate *cinF* variants' fragments with *DraI/XhoI* restriction site overhangs, a three-step PCR strategy was performed as following: firstly with primer CinF-up and CinF-part1 (5'-GACGATCTTCTTGAGCTTCATCCACA-3') to amplify fragment 1 containing partial sequence of *cinF*, concurrently using primers CinF-part2 (for Gly362Phe: 5'-*TGTGGATGAAGCTCAAGAAGATCGTCGGCAGCCACT***TCGCCAACCACGA**-3', for Gly362Ile: 5'-*TGTGGATGAAGCTCAAGAAGATCGTCGGCAGCCACAT***TCGCCAACCACGA**-3'; for A163I: *ACGGCGACGGCATGCTCGGCACCGAACAGCGGATCTGGGGCTTCGAGACCA*; for E167A: *AGGAGCCCGCCACCCACGGCGACGGCATGCTCGGCACCGCCCAGCGGGCCT*; for E167K: *AGGAGCCCGCCACCCACGGCGACGGCATGCTCGGCACCA***AGCAGCGGGCCT**; for E167Y: *AGGAGCCCGCCACCCACGGCGACGGCATGCTCGGCACCTACCAGCGGGCCT*; for N77A: *TCGTCGCGGTGATGGCCAGCTCCATCAACTACGCCACGGTCTGGTCGGCGA*; for N77E: *TCGTCGCGGTGATGGCCAGCTCCATCAACTAC***GAGACGGTCTGGTCGGCGA**; for N77Y: *TCGTCGCGGTGATGGCCAGCTCCATCAACTACTACACGGTCTGGTCGGCGA*; mutations are highlighted in bold) containing the corresponding mutation points for different *cinF* variants respectively, and CinF-dn to generate fragment 2, another partial sequence of *cinF* sharing a 26 bp homolog arm with fragment 1 (shown in italic in primers CinF-part2). Finally, using these two fragments mixture as template and CinF-up and CinF-dn as primers, a third PCR was performed to generate the different full-length *cinF* variants' fragments repectively. All PCR products were ligated into cloning vector pJET and transformed into *E. coli* HS966. Correct

construct of pJET_*cinF* and pJET_*cinF*_variants were confirmed by *DraI/XhoI* restriction digest and subsequent sequencing.

To generate the expression construct pSTW42_*cinF* and pSTW42_*cinF*_variants expression vector pSTW42 (derived from pGEX-6P-1, GST-tag sequences was replaced by SUMO-tag from pSUMO3 vector, resistance gene was exchanged from ampicillin to kanamycin) containing *XhoI/Eco47III* restriction sites was used. An N-terminal SUMO-tag containing poly-histidine was utilized for facilitating the subsequent purification. Ligation products were transformed into *E.coli* HS966. Correct constructs were confirmed by *XhoI/NdeI* restriction digest and sequencing.

For heterologous expression of CinF and its variants correct constructs were transformed into the expression strain *E. coli* BL21 (DE3) by electroporation. Correct clones were transferred into 5 ml LB-medium with 50 µg/ml kanamycin and grown overnight at 37 °C. On the next day the overnight culture was used for inoculation of LB-medium with 50 µg/ml kanamycin. The expression culture was incubated at 37 °C to an OD₆₀₀ of 0.6, and expression was induced by addition of 1mM IPTG. Subsequently, the culture was incubated at 16 °C overnight.

Verification of CinF using Matrix MALDI-TOF-MS

The desired protein band was excised from the SDS-PAGE gel and washed three times with water. 500 µl of a 50% acetonitrile (ACN) mixture in H₂O were added until the gel piece was destained. The gel piece was then air dried and 500 µl ACN were added for 10 min. The gel piece was air dried again and 20 µl trypsin were added to each sample. After complete rehydration of the gel piece, 20 µl 40 mM NH₄HCO₃ were added and the gel piece was incubated overnight at 30 °C. TFA was added to a final concentration of 0.1%. The sample was then stored at -20 °C prior to analysis by MALDI-MS. 0.7 µl desalted protein sample was mixed with 0.4 µl CCA matrix solution. The sample was co-crystallized with matrix on the metal plate and sent to the analysis in positive modus with MALDI-TOF-MS/MS.

Identification of the target protein was done by peptide mass-fingerprint (PMF) and peptide fragment-fingerprint (PFF) analysis.

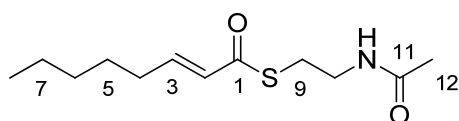
The results of the MALDI-TOF-MS were compared to the theoretical fragments after tyrosine digestion and uploaded to protein data bank to perform the evaluation. This analysis was able to identify 99.99% of the amino acid sequence and therefore confirmed correct CinF expression. Both fused protein (SUMO-CinF) and cleaved protein (CinF) were successfully identified (See **Supplementary Figure 2**).

Synthesis and purification of 2-octenoyl-CoA/-SNAC

Synthesis of 2-octenoyl-CoA/-SNAC was conducted by following procedures as described previously⁵. Crude, synthetic 2-octenoyl-CoA/-SNAC were purified using a Waters auto-purification system operating in positive ionization mode. Separation was achieved using an XBridgePrep C₁₈ column (Waters; 19 × 150 mm, 5 μm partical size, flow 25 mL min⁻¹) with a solvent system consisting of solvent A (20 mM ammonium formiate (pH 6.0)) and solvent B (1:1 solvent A: MeOH). The following gradient was applied: 40% B for 4 min, 40–60% B over 4 min, 60–100% B over 0.5 min, 100% B for 3 min, 100 - 40% B over 0.5 min, and 4 min 40% B. Compounds were detected by diode array and ESI-MS analysis. The purified product was subsequently desalted using a Sephadex column. The structure of the product was verified by NMR and purity was calculated according to HPLC measurement.

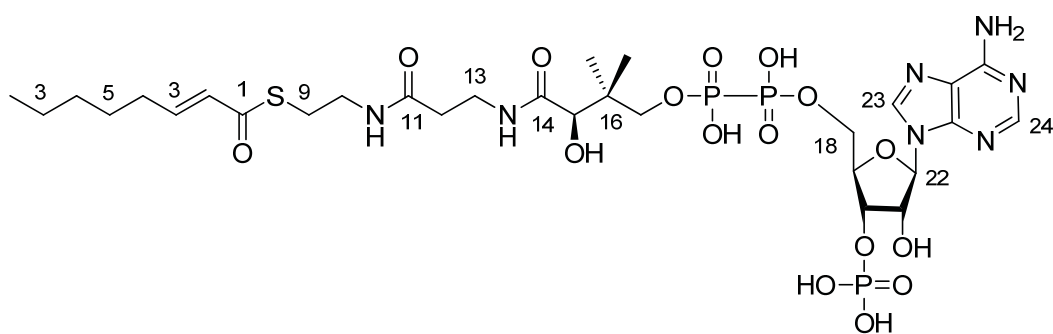
NMR spectroscopy

2-Octenoyl-SNAC-ester



^1H -NMR (700 MHz, MeOD): δ 6.92 (dt, $J = 15.5, 7.1$ Hz, 1H, 3-H), 6.16 (d, $J = 15.5$ Hz, 1H, 2-H), 3.34 (t, $J = 6.7$ Hz, 2H, 10-H), 3.06 (t, $J = 6.7$ Hz, 2H, 9-H), 2.24 - 2.20 (m, 2H, 4-H), 1.92 (s, 3H, 12-H), 1.51 - 1.46 (m, 2H, 5-H), 1.38 - 1.29 (m, 4H, 6-H and 7-H), 0.92 (t, $J = 7.1$ Hz, 8-H) ppm; ^{13}C -NMR (175 MHz, MeOD): δ 191.2 (1-C), 173.4 (11-C), 147.7 (3-C), 129.4 (2-C), 40.2 (10-C), 33.1 (4-C), 32.6 (6-C), 28.8 (9-C), 28.8 (5-C), 23.4 (7-C), 22.4 (12-C), 14.1 (8-C) ppm. HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_{12}\text{H}_{21}\text{NO}_2\text{S}$, 244.1367; found, 244.1363: analysis (calcd., found for $\text{C}_{12}\text{H}_{21}\text{NO}_2\text{S}$): C (59.22, 59.23), H (8.70, 8.68), N (5.76, 5.76), O (13.15, 13.16), S (13.18, 13.17).

2-Octenoyl-CoA



^1H -NMR (700 MHz, MeOD): δ 8.59 (bs, 1H, 23-H), 8.22 (s, 1H, 24-H), 6.91 (dt, $J = 15.4, 7.0$ Hz, 1H, 3-H), 6.14 (d, $J = 15.4$ Hz, 1H, 2-H), 6.12 (d, $J = 4.9$ Hz, 1H, 22-H), 4.99 (m, 1H, 20-H), 4.79 (m, 1H, 21-H), 4.48 (bs, 1H, 19-H), 4.29 (bs, 2H, 18-H), 4.08 (bs, 1H, 15-H), 4.01 (bs, 1H, 17-H_a), 3.63 (bs, 1H, 17-H_b), 3.50 - 3.44 (m, 2H, 13-H), 3.35 - 3.32 (m, 2H, 10-H), 3.05 (t, $J = 6.8$ Hz, 2H, 9-H), 2.43 (t, $J = 7.1$ Hz, 2H, 12-H), 2.23 - 2.18 (m, 2H, 4-H), 1.50 - 1.44 (m, 2H, 5-H), 1.36 - 1.28 (m, 4H, 6-H and 7-H), 1.06 (s, 3H, 16-Me), 0.91 (t, $J = 7.0$ Hz, 8-H), 0.85 (s, 3H, 16-Me) ppm. HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_{29}\text{H}_{49}\text{N}_7\text{O}_{17}\text{P}_3\text{S}$, 892.1907; found, 892.1890: analysis (calcd., found for $\text{C}_{29}\text{H}_{49}\text{N}_7\text{O}_{17}\text{P}_3\text{S}$): C (39.06, 39.08), H (5.43, 5.44), N (11.00, 10.98), O (30.50, 30.51), P (10.42, 10.42), S (3.60, 3.61).

^1H -NMR shifts of Coenzyme A moiety are according to those from commercially available Crotonyl-CoA ester.

Analysis of *in vitro* reaction assay

Measurements of CinF reaction assays with different substrates were accomplished via chromatography conducted using an Ultimate3000 RSLC system (Dionex, Idstein, Germany) equipped with a Waters BEH C18 100 x 2.1 mm, 1.7 μ m dp column (Waters, Milford, USA). The separation was carried out at 600 μ l/min and 45 °C. Solvent A was 0.1 % formic acid (v/v) and solvent B was acetonitrile with 0.1 % formic acid (v/v). Solvent B initially remained at 5 % B for 2 min before increasing to 20 % during a 7 min gradient. UV spectra were acquired at 260 nm. The HPLC flow was splitted to 450 nl/min to allow positive nanoESI using a TriVersa Nanomate (Advion Biosciences, Ithaca, USA). Line spectra ranging from 200 – 2000 m/z ($R = 30000$) were measured using a linear iontrap–orbitrap hybrid mass spectrometer (LTQ-Orbitrap, ThermoFisher Scientific, Bremen, Germany).

Measurements of CinF and its variants in assays with 2-octenoyl-CoA were conducted using an Ultimate3000 RSLC system (Dionex, Idstein, Germany) equipped with a Luna HILIC 100 x 2 mm, 3 μ m dp column (Phenomenex, Torrance, USA). The separation was carried out at 400 μ l/min and 40 °C. Solvent A was 20 mM ammonium formate adjusted to pH 8.5 by $\text{NH}_4\text{OH}_{\text{aq}}$. Solvent B was acetonitrile. Solvent B initially remained at 85 % B for 0.5 min before decreasing from 85 to 50 % during a 15 min gradient. UV spectra were acquired at 260 nm. The HPLC flow was splitted to 50 μ l/min before ionization in (-)-ESI mode using a Maxis 3G mass spectrometer (Bruker Daltonics, Bremen, Germany). Line spectra were acquired ranging from 300 – 1200 m/z at 2 Hz.

Data collection and refinement

The CinF crystals belong to the space group $P2_1$ with unit cell dimensions $a = 96.01 \text{ \AA}$, $b = 83.30 \text{ \AA}$ and $c = 122.74 \text{ \AA}$ and $\alpha = \gamma = 90^\circ$ and $\beta = 110.96^\circ$. The data processing was carried out using XDS⁶ and scaling was done with XSCALE. A set of 5% of randomly chosen reflections

were set aside for the calculation of the free R factor (R_{free}). The Matthews coefficient⁷ of $VM = 2.39 \text{ \AA}^3 \text{ Da}^{-1}$ indicated four monomers per asymmetric unit corresponding to a solvent content of 48%. To obtain initial phases molecular replacement by MOLREP⁸ was performed using the putative crotonyl-CoA carboxylase/reductase (3KRT) as a search model, finding four monomers in the asymmetric unit (initial FOM=0.34). One round of manual model building and repositioning drastically increased the FOM to 0.76. The structure was refined with REFMAC5⁸ and completed in Coot⁹. For the data set of CinF in complex with its substrates the model of the apo-structure was used for molecular replacement. Clear density for NADP was visible in all monomers and also for 2-octenoyl-CoA in two monomers, whereas in the other two monomers the density was much weaker. Structure and restraint files for 2-octenoyl-CoA were generated using the Dundee Prodrgr server¹⁰. The 2-octenoyl-CoA was build into the monomer with the best density (chain A) and, by applying the symmetry, placed in the other monomers as well. Water molecules were added and edited manually as well as using the water find tool in Coot with a cutoff of 1 sigma within hydrogen bonding distance to protein residues. A final TLS refinement¹¹ step using five TLS groups per monomer as determined by the TLS Motion Determination Server¹² was carried out with REFMAC5.

MolProbity^{13,14} validation indicated a good quality model. Nearly all residues fall within the allowed regions of the Ramachandran plot¹⁵ (>99%). Only in the complex structure H313, which is involved in substrate binding, is an outlier in all chains. The relevance of the crystal contacts was assessed using the PISA server¹.

References

1. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97 (2007).
2. Li, X., Jayachandran, S., Nguyen, H.H. & Chan, M.K. Structure of the *Nitrosomonas europaea* Rh protein. *Proc Natl Acad Sci U S A* **104**, 19279-84 (2007).

3. Larkin, M.A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-8 (2007).
4. Gouet, P., Courcelle, E., Stuart, D.I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-8 (1999).
5. Valenzano, C.R. et al. Stereospecificity of the Dehydratase Domain of the Erythromycin Polyketide Synthase. *Journal of the American Chemical Society* **132**, 14697-14699 (2010).
6. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).
7. Matthews, B.W. Some crystal forms of bovine chymotrypsinogen B and chymotrypsinogen A. *J Mol Biol* **33**, 499-501 (1968).
8. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55 (1997).
9. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
10. Schuttelkopf, A.W. & van Aalten, D.M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **60**, 1355-63 (2004).
11. Painter, J. & Merritt, E.A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-50 (2006).
12. Painter, J. & Merritt, E.A. TLSMD webserver for the generation of multi-group TLS models. *Journal of Applied Crystallography* **39**, 109-111 (2006).
13. Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
14. Davis, I.W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-83 (2007).
15. Ramachandran, G.N. & Sasisekharan, V. Conformation of polypeptides and proteins. *Adv Protein Chem* **23**, 283-438 (1968).

III

Stereoselective Synthesis of Deuterium-Labeled (2*S*)-Cyclohexenyl Alanines, Biosynthetic Intermediates of Cinnabaramides

Philipp Barbie,[†] Liujie Huo,[‡] Rolf Müller[‡] and Uli Kazmaier^{*†}

[†]Saarland University, Institute of Organic Chemistry, Campus, Bldg. C4.2, Email: u.kazmaier@mx.uni-saarland.de

[‡]Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, D-66123 Saarbrücken, Germany.

Stereoselective Synthesis of Deuterium-Labeled (2S)-Cyclohexenyl Alanines, Biosynthetic Intermediates of Cinnabaramide

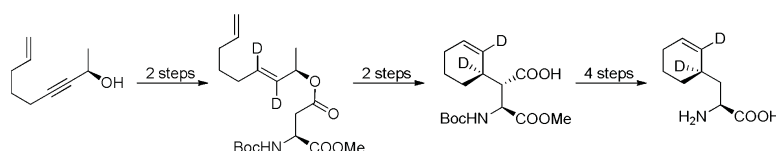
Philipp Barbie,[†] Liujie Huo,[‡] Rolf Müller,[‡] and Uli Kazmaier^{*,†}

Saarland University, Institute of Organic Chemistry, Campus, Bldg. C4.2, and Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, D-66123 Saarbrücken, Germany

u.kazmaier@mx.uni-saarland.de

Received October 26, 2012

ABSTRACT



Dideuterated β -cyclohexenylalanines, proposed biosynthetic intermediates of the cinnabaramides, can be obtained from chiral alkynols via a sequence of Ireland–Claisen rearrangement, ring closing metathesis, and radical decarboxylation. Feeding experiments indicate that both (2S)- β -cyclohexenylalanines can be incorporated into cinnabaramide, while the configuration at the cyclohexenyl ring does not restrict biosynthetic processing.

The proteasome is a key player in the regulation of a wide range of cellular processes by degradation of proteins, mediating processes such as amino acid recycling, cell differentiation, and apoptosis.¹ Therefore, proteasome inhibitors are interesting candidates as antitumor drugs.² Effective proteasome inhibitors are widespread in nature, produced by a wide range of microorganisms. Many of these inhibitors are peptide derived small molecules, such as the group of γ -lactam– β -lactones.³

In 1991, Omura et al. isolated a potent inhibitor from *Streptomyces lactacystineus*, called lactacystin.⁴ But actually, the active compound was not lactacystin itself, but its

clasto-form omuralide (Figure 1).⁵ A closely related family of compounds are the salinosporamides,⁶ isolated by Fenical from the marine actinomycete *Salinospora tropica*.⁷ Almost the same structural motif was found in the cinnabaramides, a secondary metabolite of the terrestrial strain *Streptomyces JS 360*.⁸

The cinnabaramides and salinosporamides only differ in the side chain of the γ -lactam ring. Cinnabaramides and salinosporamide A inhibit the 20S-proteasome in the low nanomolar range.^{8,9} The strained β -lactone ring of these

[†] Saarland University.

[‡] Helmholtz Institute for Pharmaceutical Research Saarland.

(1) (a) Goldberg, A. L. *Nature* **2003**, *426*, 895–899. (b) Goldberg, A. L. *Biochem. Soc. Trans.* **2007**, *35*, 12–17. (c) Murata, S.; Yashiroda, H.; Tanaka, K. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 104–115.

(2) Borissenko, L.; Groll, M. *Chem. Rev.* **2007**, *107*, 687–717.

(3) (a) Gulder, T. A. M.; Moore, B. S. *Angew. Chem.* **2010**, *122*, 9534–9556. *Angew. Chem., Int. Ed.* **2010**, *49*, 9346–9367. (b) Kale, A. J.; McGlinchey, R. P.; Lechner, A.; Moore, B. S. *ACS Chem. Biol.* **2011**, *6*, 1257–1264.

(4) Omura, S.; Fujimoto, T.; Matsuzaki, K.; Moriguchi, R.; Tanaka, H.; Sasaki, Y. *J. Antibiot.* **1991**, *44*, 113–116.

(5) Dick, L. R.; Cruikshank, A. A.; Grenier, L.; Melandri, F. D.; Nunes, S. L.; Stein, R. L. *J. Biol. Chem.* **1996**, *271*, 7273–7275.

(6) Reed, K. A.; Manam, R. R.; Mitchell, S. S.; Xu, J.; Teisan, S.; Chao, T. H.; Deyanat-Yazdi, G.; Neuteboom, S. T. C.; Lam, K. S.; Potts, B. C. M. *J. Nat. Prod.* **2007**, *70*, 269–276.

(7) Feling, R. H.; Buchanan, G. O.; Minzer, T. J.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. *Angew. Chem.* **2003**, *115*, 369–371. *Angew. Chem., Int. Ed.* **2003**, *42*, 355–357.

(8) Stadler, M.; Bitzer, J.; Mayer-Bartschmid, A.; Müller, H.; Benet-Buchholz, J.; Gantner, F.; Tichy, H.-V.; Reinemer, P.; Bacon, K. B. *J. Nat. Prod.* **2007**, *70*, 246–252.

(9) (a) Fenteany, G.; Standaert, R. F.; Lane, W. S.; Choi, S.; Corey, E. J.; Schreiber, S. L. *Science* **1995**, *268*, 726–731. (b) Groll, M.; Ditzel, L.; Löwe, J.; Stock, D.; Bochtler, M.; Bartunik, H. D.; Huber, R. *Nature* **1997**, *386*, 463–471.

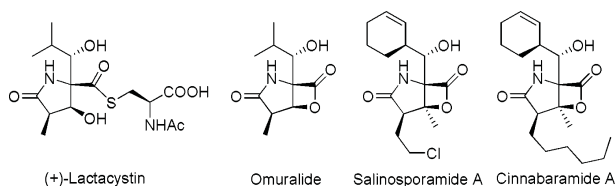


Figure 1. Naturally occurring γ -lactam- β -lactones.

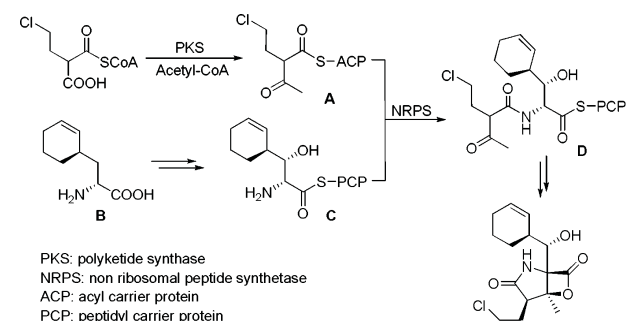
compounds interacts with a threonine in the active center of the proteasome, inactivating the enzyme by ring opening and covalent blocking of the active center.⁹ Salinosporamide is found to trigger apoptosis and is in phase I of clinical trials for the treatment of multiple myeloma.¹⁰ Not surprising, these natural products arouse interest in the community of synthetic organic chemists, and a wide range of interesting syntheses have been developed in recent years for omuralide,¹¹ cinnabaramide,¹² and especially salinosporamide¹³ and derivatives.¹⁴ The recent developments in this field are nicely covered in reviews by Moore^{3a} and Potts.¹⁵

Besides the development of straightforward protocols for the synthesis of these compounds, much effort has also been focused toward investigating their biosynthesis. While Moore et al. are involved in the biosynthesis of the

salinosporamides,^{3b,16} we are mainly interested in the cinnabaramides.¹⁷

In principle, the molecule can be divided into two major parts. In the case of the salinosporamides, the lower part of the molecule is the product of a polyketide synthase (PKS) leading to an activated β -ketothioester **A** (Scheme 1). The upper part, the unusual amino acid β -hydroxy cyclohexenylalanine, is formed *via* a shunt in the phenylalanine biosynthetic pathway, leading primarily to β -cyclohexenylalanine **B**.^{16a} After coupling to the peptidyl carrier protein and oxidation by a cytochrome P450 hydroxylase (C), the two building blocks are coupled on a nonribosomal peptide synthetase (NRPS) to give **D**. Subsequent cyclization gives rise to salinosporamide.¹⁶

Scheme 1. Biosynthesis of Salinosporamide A according to Moore^{16a}



In principle, one might expect a very similar biosynthetic pathway also for the cinnabaramides, which should differ mainly in the PKS-subunit. To study the cinnabaramide biosynthesis, we were interested in obtaining deuterium-labeled building blocks which can be used for feeding experiments. Recently, we described the stereoselective synthesis of dideuterated (2*R*,3*S*,4*S*)- β -cyclohexenylserine,¹⁸ a postulated intermediate in an early biosynthetic proposal.^{16a} Later, the configuration of the amino acid was determined to be (2*S*).^{16b} Therefore, it was not so surprising that no incorporation of this amino acid itself, or in the activated form, was observed in feeding experiments. Recent biosynthetic studies by the Müller group indicate that, in the case of the cinnabaramides, **B** is probably coupled to the corresponding β -ketoester and that the cyclochrome P-450 oxidation at the β -position proceeds later on in the biosynthesis.¹⁹

To prove this proposal, we developed a stereoselective synthesis of the two isomeric dideuterated (2*S*)- β -cyclohexenylalanines. So far, the absolute configuration of the cinnabaramide intermediates is not yet determined, but the (2*S*) configuration seems reasonable, based on the analogy to the salinosporamide biosynthesis. In the natural product, the (4*S*)-configuration is found, and therefore the (2*S*,4*R*)- β -cyclohexenylalanine **1** should be the correct

(10) (a) Chauhan, D.; Catley, L.; Li, G.; Podar, K.; Hideshima, T.; Velankar, M.; Mitsiades, C.; Mitsiades, N.; Yasui, H.; Letai, A.; Ovao, H.; Berkers, C.; Nicholson, B.; Chao, T.-H.; Neuteboom, S. T. C.; Richardson, P.; Palladino, M. A.; Anderson, K. C. *Cancer Cell* **2005**, *8*, 407–419. (b) Potts, B. C.; Albitar, M. X.; Anderson, K. C.; Baritaki, S.; Berkers, C.; Bonavida, B.; Chandra, J.; Chauhan, D.; Cusack, J. C., Jr.; Selby, M. D.; Prodger, J. C. *J. Org. Chem.* **2008**, *73*, 2041–2051. (c) Gu, W.; Silverman, R. B. *J. Org. Chem.* **2011**, *76*, 8287–8293.

(11) (a) Corey, E. J.; Li, W. D. *Z. Chem. Pharm. Bull.* **1999**, *47*, 1–10. (b) Crane, S. N.; Corey, E. J. *Org. Lett.* **2001**, *3*, 1395–1397. (c) Gilley, C. B.; Buller, M. J.; Kobayashi, Y. *Org. Lett.* **2007**, *9*, 3631–3634. (d) Hayes, C. J.; Sherlock, A. E.; Green, M. P.; Wilson, C.; Blake, A. J.; Selby, M. D.; Prodger, J. C. *J. Org. Chem.* **2008**, *73*, 2041–2051. (e) Gu, W.; Silverman, R. B. *J. Org. Chem.* **2011**, *76*, 8287–8293.

(12) Ma, G.; Nguyen, H.; Romo, D. *Org. Lett.* **2007**, *9*, 2143–2146.

(13) (a) Reddy, L. R.; Saravanan, P.; Corey, E. J. *J. Am. Chem. Soc.* **2004**, *126*, 6230–6231. (b) Endo, A.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2005**, *127*, 8298–8299. (c) Mulholland, N. P.; Pattenden, G.; Walters, I. A. S. *Org. Biomol. Chem.* **2006**, *4*, 2845–2846. (d) Ling, T.; Macherla, V. R.; Manam, R. R.; McArthur, K. A.; Potts, B. C. M. *Org. Lett.* **2007**, *9*, 2289–2292. (e) Mulholland, N. P.; Pattenden, G.; Walters, I. A. S. *Org. Biomol. Chem.* **2008**, *6*, 2782–2789. (f) Takahashi, K.; Midori, M.; Kawano, K.; Ishihara, J.; Hatakeyama, S. *Angew. Chem.* **2008**, *120*, 6340–6342. *Angew. Chem., Int. Ed.* **2008**, *47*, 6244–6246. (g) Fukuda, T.; Sugiyama, K.; Arima, S.; Harigaya, Y.; Nagamitsu, T.; Omura, S. *Org. Lett.* **2008**, *10*, 4239–4242. (h) Mosey, R. A.; Tepe, J. J. *Tetrahedron* **2009**, *50*, 295–297. (i) Nguyen, H.; Ma, G.; Romo, D. *Chem. Commun.* **2010**, *46*, 4803–4805. (j) Ling, M.; Potts, B. C.; Macherla, V. R. *J. Org. Chem.* **2010**, *75*, 3882–3885. (k) Sato, Y.; Fukuda, H.; Tomizawa, M.; Masaki, T.; Shibuya, M.; Kanoh, N.; Iwabuchi, Y. *Heterocycles* **2010**, *81*, 2239–2246. (l) Nguyen, H.; Ma, G.; Gladysheva, T.; Fremgen, T.; Romo, D. *J. Org. Chem.* **2011**, *76*, 2–12. (m) Kaiya, Y.; Hasegawa, J.-L.; Momose, T.; Sato, T.; Chida, N. *Chem.-Asian J.* **2011**, *6*, 209–219. (n) Satoh, N.; Yokoshima, S.; Fukuyama, T. *Org. Lett.* **2011**, *13*, 3028–3031.

(14) Chen, Z. H.; Wang, B. C.; Kale, A. J.; Moore, B. S.; Wang, R. W.; Qing, F. L. *J. Fluorine Chem.* **2012**, *136*, 12–19.

(15) Potts, B. C.; Lam, K. S. *Mar. Drugs* **2010**, *8*, 835–880.

(16) (a) Beer, L. L.; Moore, B. S. *Org. Lett.* **2007**, *9*, 845–848. (b) McGlinchey, R. P.; Nett, M.; Eustáquio, A. S.; Asolkar, R. N.; Fenical, W.; Moore, B. S. *J. Am. Chem. Soc.* **2008**, *130*, 7822–7823.

(17) Rachid, S.; Huo, L.; Hermann, J.; Stadler, M.; Köpcke, B.; Bitzer, J.; Müller, R. *ChemBioChem* **2011**, *12*, 922–931.

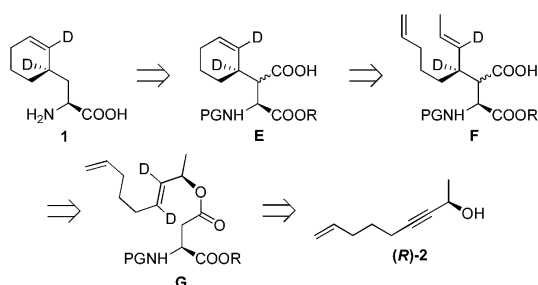
(18) Deska, J.; Hahn, S.; Kazmaier, U. *Org. Lett.* **2011**, *13*, 3210–3213.

(19) Huo, L.; Müller, R., unpublished results.

intermediate. Interestingly, the biosynthetic enzyme responsible for the activation of the β -cyclohexenylalanine seems to have a rather broad substrate spectrum, allowing also the incorporation of other amino acids.^{16b,20} Therefore, we were interested to see if the configuration of the 4-position does play any role.

Our retrosynthetic plan toward the (2*S*,4*R*)-**1** is shown in Scheme 2. D₂-labeled β -cyclohexenylalanine **1** should be accessible *via* radical decarboxylation of the corresponding aspartate derivative **E**, while the cyclohexenyl ring should be the result of a ring closing metathesis (RCM). The key step of the synthesis is a stereoselective Claisen rearrangement providing the diene **F** from the allyl ester **G**. In this step the stereogenic center at C-4 is incorporated. Depending on the configuration of the allyl alcohol used, both isomers can be obtained stereoselectively. The configuration at C-3 of **F** does not play any role, because this stereogenic center is removed later on. The required di-deuterated *trans* allyl alcohol should be available from the corresponding propargyl alcohol (**R**)-**2** *via* reduction using LiAlD₄.

Scheme 2. Retrosynthesis of (2*S*,4*R*)-Dideuterated Cyclohexenylalanine **1**



We began our synthesis with the coupling of 5-bromopentene with butyn-3-ol (Scheme 3). Although this reaction is described in the literature,²¹ we were able to improve the yield of **2** significantly by using LiNH₂ as a base (instead of BuLi). The racemic alkynol **2** was subsequently subjected to an enzymatic kinetic resolution using Novozym 435. The reaction stops after 50% conversion, and both enantiomeric compounds were obtained with high yield and *ee*. Unfortunately, we were not able to determine the *ee*-value of the unreacted alcohol directly by GC (no separation) and therefore, for analytical purposes, the remaining alcohol (*S*)-**2** was also converted into the acetate (*S*)-**3**. Upon saponification, both enantiomers of alkynol **2** were obtained.

For the synthesis of the deuterated (2*S*,4*R*)-cyclohexenylalanine **1**, (*R*)-**2** was reduced with LiAlD₄ (Scheme 3). The *in situ* formed vinylalumoxane was quenched by the addition of D₂O giving rise to the *trans*-configured

dideuterated allyl alcohol (*R*)-**4**. Coupling to semiprotected aspartate²² using the Steglich protocol provided the required allylester **5**, the substrate for the subsequent Claisen rearrangement. Our group was investigating Claisen rearrangements of amino acids for a long time.²³ While the best results are obtained for α -amino acid esters with chelated enolates,²⁴ this protocol is not really suitable for β -amino acid esters, such as **5**. Here the Ireland–Claisen rearrangement²⁵ is the method of choice,²⁶ providing **6** in high yield and diastereoselectivity. While the stereogenic center at C-2 remains unchanged, the configuration at C-4 is determined by the stereogenic center of the allyl alcohol (chirality transfer almost perfect). The configuration at C-3 is the result of the enolate formation in the deprotonation step. This center is formed as a 9:1 mixture, according to NMR. But this center is negligible, because it is removed later on. The diastereomeric rearrangement products could be separated by flash chromatography. Diastereomerically pure **6** was subjected to ring closing metathesis providing cyclohexenyl aspartate **7** in high yield. In case the diastereomeric mixture was subjected to RCM also at this stage, a separation of the isomers was possible.

The next step, the radical decarboxylation, was found to be the most critical one of the whole sequence. We decided to apply the Barton protocol using the corresponding *N*-hydroxythiopyridone esters.²⁷ Unfortunately, this ester was found to be very sensitive, undergoing fast decomposition. Attempts to isolate and purify it were unsuccessful. Therefore, we had to carry out the activation and decarboxylation in a one-pot protocol. We investigated a wide range of coupling reagents for the initial step, the formation of the active ester. No ester formation was observed by using chloroformates, carbonyldiimidazole, or T3P as coupling reagents. DCC at -20°C provided the required ester, but in impure form, and therefore, the decarboxylated product was very difficult to separate from the byproducts formed. The best results were obtained using PhPOCl₂/NEt₃ in the activation step, and the very mild BEt₃/O₂ protocol²⁸ for the radical formation. *t*-BuSH was used as a H-source. Under these optimized conditions, the required product **8** could be obtained in 49% yield. The final steps toward **1** proceeded quantitatively. The isomeric (2*S*,4*S*)-derivative was obtained from (*S*)-**2** in an analogous fashion and comparable yield.

(22) Sato, S.; Tetsuhashi, M.; Sekine, K.; Miyachi, H.; Naito, M.; Hashimoto, Y.; Aoyama, H. *Bioorg. Med. Chem.* **2008**, *16*, 4685–4698.

(23) (a) Kazmaier, U. *Amino Acids* **1996**, *11*, 283–299. (b) Kazmaier, U.; Maier, S. *Org. Lett.* **1999**, *1*, 1763–1766. (c) Kazmaier, U.; Mues, H.; Krebs, A. *Chem.—Eur. J.* **2002**, *8*, 1850–1855.

(24) (a) Kazmaier, U.; Maier, S. *Tetrahedron* **1996**, *52*, 941–954. (b) Kazmaier, U.; Schneider, C. *Synthesis* **1998**, 1321–1326. (c) Kazmaier, U.; Schneider, C. *Tetrahedron Lett.* **1998**, *39*, 817–818. (d) Kazmaier, U.; Maier, S. *J. Org. Chem.* **1999**, *64*, 4574–4575.

(25) (a) Ireland, R. E.; Mueller, R. H. *J. Am. Chem. Soc.* **1972**, *94*, 5897–5898. (b) Ireland, R. E.; Mueller, R. H.; Willard, A. K. *J. Am. Chem. Soc.* **1976**, *98*, 2868–2877.

(26) Becker, D.; Kazmaier, U. *J. Org. Chem.*, in press (DOI: 10.1021/jo301693d).

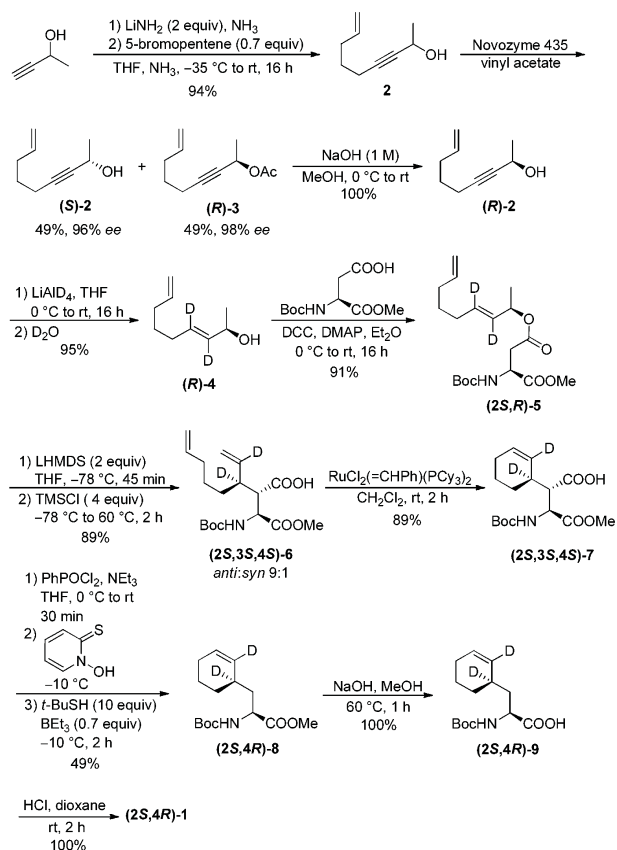
(27) Barton, D. H. R.; Crich, D.; Motherwell, W. B. *Tetrahedron* **1985**, *41*, 3901–3924.

(28) Ollivier, C.; Renaud, P. *Chem. Rev.* **2001**, *101*, 3415–3434.

(20) Nett, M.; Gulder, T. A. M.; Kale, A. J.; Hughes, C. C.; Moore, B. S. *J. Med. Chem.* **2009**, *52*, 6163–6167.

(21) Roush, W. R.; Spada, A. P. *Tetrahedron Lett.* **1982**, *23*, 3773–3776.

Scheme 3. Synthesis of d₂-(2*S*,4*R*)-1



Subsequent feeding experiments were conducted using compound **1**, which was stepwise administrated to the culture broth of *Streptomyces* sp. JS360 (see Supporting Information). Organic extracts of the culture supernatants were analyzed using HPLC-MS, which unambiguously demonstrated incorporation of **1** into cinnabaramide resulting in the production of dideuterated cinnabaramides in a yield comparable to that of native cinnabaramides.

Based on this efficient incorporation we can conclude that (2*S*,4*R*)- β -cyclohexenylalanine **1** most likely serves as the native substrate of the cinnabaramide peptidase. Interestingly, while the salinosporamide pathway shows a broad substrate tolerance allowing for the incorporation of alternative amino acids, no cinnabaramides derived from incorporation of proteinogenic amino acids have been detected in extracts of the wild type producer. But interestingly, the diastereomeric (2*S*,4*S*)-cyclohexenylalanine was incorporated as well. This may indicate that either (1) the free amino acids cannot be activated for further incorporation or (2) *in vivo* insufficient quantities of amino acids are available for secondary metabolite biosynthesis. As the latter scenario is unlikely we assume that it should be possible to incorporate alternative amino acids in the respective position of cinnabaramides. Ongoing work in our laboratories is addressing this question. After conducting a bioactivity screening, we expect to discover more cytotoxic novel cinnabaramides.

In conclusion, we have shown that the Ireland–Claisen rearrangement of β -amino acid allylic esters is a highly suitable tool for the stereoselective synthesis of β -substituted aspartates. Subsequent ring closing metatheses and decarboxylations give rise to deuterated β -cyclohexenylalanine derivatives. Feeding experiments indicate that (2*S*)- β -cyclohexenylalanine is an intermediate in the biosynthesis of cinnabaramide. The configuration at the cyclohexenyl ring does not play any role, as obviously both stereoisomers are accepted by the peptidyl carrier protein. Further biosynthetic studies are currently under investigation.

Acknowledgment. Financial support from the Deutsche Forschungsgemeinschaft was gratefully acknowledged.

Supporting Information Available. Experimental procedures as well as analytical and spectroscopic data of all new compounds, details of the feeding experiments, and copies of NMR spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

The authors declare no competing financial interest.

Supporting Information

Stereoselective synthesis of deuterium-labeled (2S)-cyclohexenyl alanines, biosynthetic intermediates of cinnabaramide.

Philipp Barbie, Liujie Huo, Rolf Müller and Uli Kazmaier*

*Saarland University, Institute of Organic Chemistry, Campus, Bldg. C4.2, and
Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection
Research and Department of Pharmaceutical Biotechnology, D-66123 Saarbruecken,
Germany.*

u.kazmaier@mx.uni-saarland.de

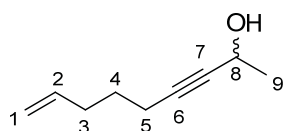
1) General remarks	S2
2) Synthesis of (2S,4R)-1	S2
3) Synthesis of (2S,4S)-1	S9
4) Supplemental information on feeding experiment	S14
5) Copies of NMR spectra	S16

1) General remarks

All air- or moisture-sensitive reactions were carried out in dried glassware (>100 °C) under an atmosphere of nitrogen. Dried solvents were distilled before use: THF was distilled from LiAlH₄, CH₂Cl₂ was dried with CaH₂ before distillation and DMF was purchased from Sigma-Aldrich. The products were purified by flash chromatography on silica gel columns (Macherey-Nagel 60, 0.063-0.2 mm). Mixtures of ethyl acetate and hexane were generally used as eluents. Analytical TLC was performed on pre-coated silica gel plates (Fluka). Visualization was accomplished with UV-light and KMnO₄ or Ninhydrin solution. Melting points were determined with a *Laboratory Devices MEL-TEMP II* melting point apparatus and are uncorrected. ¹H and ¹³C NMR spectra were recorded with a *Bruker AC-400* [400 MHz (¹H) and 100 MHz (¹³C)] spectrometer in CDCl₃, unless otherwise specified. Chemical shifts are reported in ppm relative to TMS, and CHCl₃ was used as the internal standard. Multiplicities are reported as (br) broad, (s) singlet, (d) doublet, (t) triplet, (q) quartet and (m) multiplet. The diastereomeric ratios were determined by NMR and GC. Enantiomeric ratios were determined by GC using chiral columns (*Chirasil-Dex-CB* or *Chirasil-Val-CB*). Mass spectra were recorded with a *Finnigan MAT 95 spectrometer* using the CI technique.

2) Synthesis of (2*S*,4*R*)-1

Non-8-en-3-yn-2-ol (**2**)¹



Lithium (700 mg, 101 mmol) was added in small portions to liquid ammonia (250 mL) at –40 °C. After addition was complete, a small amount of Fe(NO₃)₃ · 9H₂O was added. When the blue color had disappeared completely, a solution of 3-butyne-2-ol (3.50 g, 50.0 mmol) in dry THF (5 mL) was added dropwise at –40 °C and stirred for one hour. Then, a solution of 5-bromo-1-pentene (5.00 g, 33.5 mmol) in THF (5 mL) was added dropwise and the reaction mixture was allowed to warm up to rt slowly overnight. The resulting mixture was quenched with water (50 mL) and the aqueous phase was extracted with Et₂O (3 x 50 mL). The combined organic layers were dried over Na₂SO₄ and the solvent was evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (pentane/Et₂O, 8:2) to give racemic alcohol **2** (4.35 g, 31.5 mmol, 94%) as a colorless liquid. R_f 0.21 (pentane/Et₂O, 8:2)

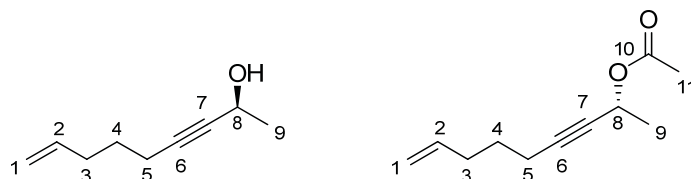
¹H NMR (400 MHz, CDCl₃): δ = 1.39 (d, ³J_{9,8} = 6.8 Hz, 3 H, 9-H), 1.57 (m, 2 H, 4-H), 1.96 (bs, 1 H, OH), 2.13 (tdd, ³J_{3,4} = 14.4 Hz, ³J_{3,2} = 7.1 Hz, ⁴J_{3,1cis} = 1.5 Hz, 2 H, 3-H), 2.20 (td, ³J_{5,4} = 7.2 Hz, ⁵J_{5,8} = 1.9 Hz, 2 H, 5-H), 4.50 (m, 1 H, 8-H), 5.00 (m, 1 H, 1-H_{trans}), 5.02 (ddt, ³J_{1cis,2} = 17.2 Hz, ²J_{1cis,1trans} = ⁴J_{1cis,3} = 1.5 Hz, 1 H, 1-H_{cis}),

¹ Roush, W. R.; Spada, A. P. *Tetrahedron Lett.* **1982**, 23, 3773-3776.

5.77 (ddt, $^3J_{2,1cis} = 17.2$ Hz, $^3J_{2,1trans} = 10.2$ Hz, $^3J_{2,3} = 7.1$ Hz, 1 H, 2-H). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 17.9$ (t, C-5), 24.6 (q, C-9), 27.7 (t, C-4), 32.6 (t, C-3), 58.4 (d, C-8), 82.6 (s, C-6), 84.0 (s, C-7), 115.1 (t, C-1), 137.7 (d, C-2).

GC: *Chirasil-Dex-CB*, T0 [5 min] = 75 °C, 0.5 °C/min - 120 °C, 8.0 °C/min - 180 °C, (**S**)-**2**: $t_R = 43.04$ min, (**R**)-**2**: $t_R = 45.86$ min.

(S)-Non-8-en-3-yn-2-ol [(S)-2]¹ and (R)-Non-8-en-3-yn-2-yl acetate [(R)-3]



To a solution of racemic alcohol **2** (2.75 g, 19.9 mmol) in vinyl acetate (30 mL) Novozym[®] 435 (150 mg) was added at rt. The mixture was shaken for 16 h. After the reaction was completed (NMR), the enzyme was filtered off and the solvent was evaporated *in vacuo*. The residue was purified by flash chromatography (pentane/Et₂O 8:2 to 6:4) to give alcohol (**S**)-**2** (1.30 g, 9.41 mmol, 47%, 96% ee) and acetate (**R**)-**3** (1.75 g, 9.68 mmol, 49%, 98% ee) as colorless liquids.

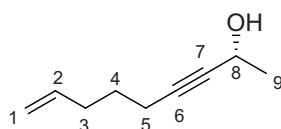
(**S**)-**2**: R_f 0.21 (pentane/Et₂O, 8:2); $[\alpha]_D^{20} = -19.5^\circ$ ($c = 1.9$, CHCl_3 , 96% ee).

(**R**)-**3**: R_f 0.46 (pentane/Et₂O, 8:2); $[\alpha]_D^{20} = +137.5^\circ$ ($c = 1.1$, CHCl_3 , 98% ee); ^1H NMR (400 MHz, CDCl_3): $\delta = 1.46$ (d, $^3J_{9,8} = 6.7$ Hz, 3 H, 9-H), 1.57 (m, 2 H, 4-H), 2.06 (s, 3 H, 11-H), 2.13 (td, $^3J_{3,4} = 14.4$ Hz, $^3J_{3,2} = 7.2$ Hz, $^4J_{3,1cis} = 1.7$ Hz, 2 H, 3-H), 2.21 (td, $^3J_{5,4} = 7.2$ Hz, $^5J_{5,8} = 1.9$ Hz, 2 H, 5-H), 4.99 (m, 1 H, 1-H_{cis}), 5.03 (ddt, $^3J_{1cis,2} = 17.2$ Hz, $^2J_{1cis,1trans} = ^4J_{1cis,3} = 1.7$ Hz, 1 H, 1-H_{cis}), 5.44 (qt, $^3J_{8,9} = 6.7$ Hz, $^5J_{8,5} = 1.9$ Hz, 1 H, 8-H), 5.78 (ddt, $^3J_{2,1cis} = 17.2$ Hz, $^3J_{2,1trans} = 10.2$ Hz, $^3J_{2,3} = 7.2$ Hz, 1 H, 2-H); ^{13}C NMR (100 MHz, CDCl_3): $\delta = 18.0$ (t, C-5), 21.2 (t, C-11), 21.8 (q, C-9), 27.6 (t, C-4), 32.7 (t, C-3), 60.8 (d, C-8), 78.9 (s, C-6), 85.1 (s, C-7), 115.2 (t, C-1), 137.8 (d, C-2), 170.0 (s, C-10).

GC: *Chirasil-Dex-CB*, T0 [5 min] = 60 °C, 1.0 °C/min - 120 °C, 6.0 °C/min - 180 °C, (**S**)-**3**: $t_R = 45.60$ min, (**R**)-**3**: $t_R = 45.97$ min.

HRMS ($\text{C}_{11}\text{H}_{19}\text{O}_3$ [$\text{M}+\text{H}_2\text{O}+\text{H}$]⁺): calculated 199.1329, found 199.1256.

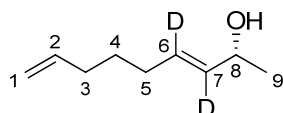
(R)-Non-8-en-3-yn-2-ol [(R)-2]¹



To a solution of acetate (**R**)-**3** (3.10 g, 17.2 mmol) in methanol (80 mL) 1 M aq. NaOH (19.0 mL, 19.0 mmol) was added at rt and stirred overnight. After removal of the methanol *in vacuo*, the residue was diluted with Et₂O (30 mL) and the aqueous phase

was extracted with Et₂O (3 x 30 mL). The combined organic layers were washed with brine (20 mL) and dried over Na₂SO₄. The solvent was evaporated in vacuo and the residue was purified by flash chromatography on silica gel (pentane/Et₂O 7:3) to give alcohol **(R)-2** (2.36 g, 17.2 mmol, 100%) as colorless liquid. *R*_f 0.21 (pentane/Et₂O, 8:2); [α]_D²⁰ = + 30.5 ° (c = 1.0, CHCl₃).

(R,E)-3,4-Dideuteronona-3,8-dien-2-ol [(R)-4]¹

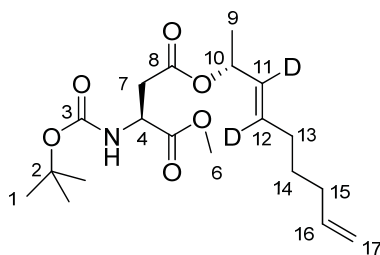


To a suspension of lithium aluminium deuteride (609 mg, 14.5 mmol) in dry THF (10 mL) a solution of alcohol **(R)-4** (880 mg, 6.37 mmol) in THF (8 mL) was added at 0 °C. The reaction mixture was allowed to warm to rt overnight and quenched with deuterium oxide (5 mL). A 2 M solution of Na₂CO₃ (20 mL) was added to the reaction mixture and stirring was continued for 10 min. The layers were separated and the aqueous phase was extracted with Et₂O (3 x 50 mL). The combined organic layers were dried over Na₂SO₄ and the solvent was removed *in vacuo*. After purification by column chromatography on silica gel (pentane/Et₂O, 6:4), the deuterated alcohol **(R)-4** (860 mg, 6.05 mmol, 95%) was obtained as colorless liquid. *R*_f 0.16 (pentane/Et₂O, 8:2); [α]_D²⁰ = + 12.7 ° (c=1.0, CHCl₃).

¹H NMR (400 MHz, CDCl₃): δ = 1.24 (d, ³*J*_{8,9} = 6.4 Hz, 1 H, 9-H), 1.46 (m, 2 H, 4-H), 1.55 (bs, 1 H, OH), 2.00–2.08 (m, 4 H, 3-H, 5-H), 4.24 (q, ³*J*_{9,8} = 6.4 Hz, 1 H, 8-H), 4.94 (ddt, ³*J*_{1trans,2} = 10.4 Hz, ²*J*_{1trans,1cis} = 2.0 Hz, ⁴*J*_{1trans,3} = 1.2 Hz, 1 H, 1-H_{trans}), 4.99 (dtd, ³*J*_{1cis,2} = 16.8 Hz, ⁴*J*_{1cis,3} = ²*J*_{1cis,1trans} = 2.0 Hz, 1 H, 1-H_{cis}), 5.79 (ddt, ³*J*_{2,1cis} = 16.8 Hz, ³*J*_{2,1trans} = 10.4 Hz, ³*J*_{2,3} = 6.7 Hz, 1 H, 2-H). ¹³C NMR (100 MHz, CDCl₃): δ = 23.4 (q, C-9), 28.3 (t, C-4), 31.3 (t, C-3 oder C-5), 33.1 (t, C-3 oder C-5), 68.7 (d, C-8), 114.5 (t, C-1), 130.1 (t, *J*_{7,D} = 22.8 Hz, C-7), 134.0 (t, *J*_{7,D} = 23.0 Hz, C-6), 138.6 (d, C-2).

HRMS (C₉H₁₅D₂O [M+H]⁺): calculated 143.1399, found 143.1398.

(S)-4-[(R,E)-3,4-Dideuteronona-3,8-dien-2-yl] 1-methyl 2-(tert-butoxycarbonyl-amino)succinate [(2S,R)-5]



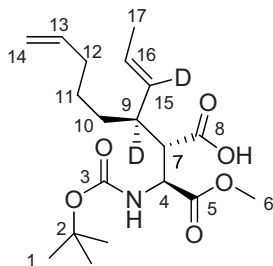
To a solution of (*L*)-1-methyl-*N*-boc-aspartate (1.22 g, 4.93 mmol), alcohol **(R)-4** (705 mg, 4.93 mmol) and 4-(dimethylamino)pyridine (49.0 mg, 0.40 mmol) in Et₂O (30 mL)

N,N'-dicyclohexylcarbodiimide (1.12 g, 5.42 mmol) was added at 0 °C. The reaction mixture was allowed to warm to rt overnight and filtered. The filtrate was washed with 1 M KHSO₄ (10 mL), sat. NaHCO₃ (10 mL) and brine (20 mL). After drying over Na₂SO₄, the solvent was removed *in vacuo* and the residue was purified by flash chromatography on silica gel (petroleum ether/EtOAc 8:2). The allylic ester **(2S,R)-5** (1.67 g, 4.50 mmol, 91%) was obtained as colorless oil. *R*_f 0.32 (petroleum ether/Et₂O, 8:2); [α]_D²⁰ = + 70.2 ° (c=1.0, CHCl₃).

¹H NMR (400 MHz, CDCl₃): δ = 1.26 (d, ³J_{9,10} = 6.5 Hz, 3 H, 9-H), 1.43 (s, 9 H, 1-H), 1.45 (m, 2 H, 14-H), 2.00–2.05 (m, 4 H, 13-H, 15-H), 2.77 (dd, ²J_{7a,7b} = 16.8 Hz, ³J_{7a,4} = 4.4 Hz, 1 H, 7-H_a), 2.94 (dd, ²J_{7b,7a} = 16.8 Hz, ³J_{7b,4} = 4.4 Hz, 1 H, 7-H_b), 3.72 (s, 3 H, 6-H), 4.54 (m, 1 H, 4-H), 4.94 (ddt, ³J_{17trans,16} = 10.0 Hz, ²J_{17trans,17cis} = 1.6 Hz, ⁴J_{17trans,15} = 1.2 Hz, 1 H, 17-H_{trans}), 4.99 (ddt, ³J_{17cis,16} = 17.2 Hz, ²J_{17cis,17trans} = ⁴J_{17cis,15} = 1.6 Hz, 1 H, 17-H_{cis}), 5.30 (q, ³J_{10,9} = 6.5 Hz, 1 H, 10-H), 5.45 (d, ³J_{NH,4} = 8.0 Hz, 1 H, N-H), 5.77 (ddt, ³J_{16,17cis} = 17.2 Hz, ³J_{16,17trans} = 10.0 Hz, ³J_{16,15} = 6.7 Hz, 1 H, 16-H). ¹³C NMR (100 MHz, CDCl₃): δ = 20.2 (q, C-9), 28.0 (t, C-14), 28.2 (s, C-1), 31.3 (t, C-13 oder C-15), 33.0 (t, C-13 oder C-15), 37.1 (t, C-7), 50.0 (d, C-4), 52.5 (q, C-6), 71.9 (d, C-10), 80.0 (s, C-2), 114.6 (t, C-17), 128.9 (t, *J*_{7,D} = 23.4 Hz, C-11), 132.9 (t, *J*_{7,D} = 22.8 Hz, C-12), 138.4 (d, C-16), 155.3 (s, C-3), 170.0 (s, C-8), 171.5 (s, C-5).

HRMS (C₁₄H₂₂D₂NO₄ [M–Boc+H]⁺): calculated 272.1825, found 272.1826.

(2S,3S)-2-[(S)-1-(tert-Butoxycarbonylamino)-2-methoxy-2-oxoethyl]-3-deutero-3-((Z)-1-deuterioprop-1-enyl)oct-7-enoic acid [(2S,3S,4S)-6]



Freshly distilled HMDS (1.67 mL, 8.08 mmol) was dissolved in dry THF (8 mL) in a Schlenk tube and cooled to –30 °C. *N*-Butyllithium (2.5 M in hexanes, 2.15 mL, 5.38 mmol) was added dropwise and stirring was continued for 5 min. The mixture was allowed to warm to rt and was cooled again to –78 °C. In a second Schlenk tube, ester **(2S,R)-5** (1.00 g, 2.69 mmol) was dissolved in THF (10 mL) and cooled to –78 °C. The LHMDS solution was transferred dropwise to the second flask *via* a double-ended needle and stirred for 30 min at –78 °C. TMSCl (1.72 mL, 13.5 mmol) was added and stirring was continued for 1 h. The reaction mixture was then quickly warmed up to 60 °C. This temperature was maintained for 2 h while stirring. The mixture was quenched with 1 M HCl (15 mL) and the organic layer was extracted with Et₂O. The combined organic layers were washed with 1 M KHSO₄ and brine. After drying over Na₂SO₄, the solvent was removed *in vacuo*. Column chromatography on

silica gel (petroleum ether/EtOAc 9:1, 1% HOAc) afforded **(2S,3S,4S)-6** as the major diastereomer (784 mg, 2.10 mmol, 78%) and **(2S,3R,4S)-6** as the minor diastereomer (109 mg, 0.29 mmol, 11%) als colorless solids; mp. 137-138 °C.

GC (methyl ester of crude reaction mixture) *Chirasil-Val-CB*, T_0 [5 min] = 145 °C, 1.0 °C/min - 180 °C, 6.0 °C/min - 180 °C **(2S,3S,4S)-6**: t_R = 34.66 min. **(2S,3R,4S)-6**: = 43.63 min.

(2S,3S,4S)-6: R_f 0.15 (petroleum ether/Et₂O, 8:2, 1% HOAc); $[\alpha]_D^{20} = + 8.6^\circ$ (c=1.0, CHCl₃). Rotameric ratio at rt: 3:1.

Major rotamer: ¹H NMR (400 MHz, CDCl₃): δ = 1.28–1.46 (m, 3 H, 11-H, 10-H_a), 1.43 (s, 9 H, 1-H), 1.57 (m, 1 H, 10-H_b), 1.66 (d, ³ $J_{17,16}$ = 6.2 Hz, 3 H, 17-H), 2.02 (m, 2 H, 12-H), 2.93 (d, ³ $J_{7,4}$ = 6.8 Hz, 1 H, 7-H), 3.73 (s, 3 H, 6-H), 4.55 (dd, ³ $J_{4,7}$ = ³ $J_{4,NH}$ = 7.4 Hz, 1 H, 4-H), 4.93 (ddt, ³ $J_{14trans,13}$ = 10.0 Hz, ² $J_{14trans,14cis}$ = 2.0 Hz, ³ $J_{14trans,12}$ = 1.2 Hz, 1 H, 14-H_{trans}), 4.98 (ddt, ³ $J_{14cis,13}$ = 17.2 Hz, ² $J_{14cis,14trans}$ = ⁵ $J_{14cis,12}$ = 2.0 Hz, 1 H, 14-H_{cis}), 5.23 (d, ³ $J_{NH,4}$ = 7.4 Hz, 1 H, NH), 5.47 (m, 1 H, 16-H), 5.78 (ddt, ³ $J_{13,14cis}$ = 17.2 Hz, ³ $J_{13,14trans}$ = 10.0 Hz, ³ $J_{13,12}$ = 6.7 Hz, 1 H, 13-H), 9.35 (bs, 1 H, OH). ¹³C NMR (100 MHz, CDCl₃): δ = 17.9 (q, C-17), 26.6 (t, C-11), 28.3 (q, C-1), 32.0 (t, C-10), 33.4 (t, C-12), 51.7 (d, C-7), 52.5 (q, C-6), 53.4 (d, C-4), 80.3 (s, C-2), 114.5 (t, C-14), 128.5 (d, C-16), 130.1 (t, $J_{15,D}$ = 16.7 Hz, C-15), 138.7 (d, C-13), 155.0 (s, C-3), 171.5 (s, C-5), 176.9 (s, C-8).

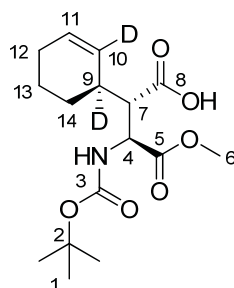
Minor rotamer (selected signals): ¹H NMR (400 MHz, CDCl₃): δ = 4.26 (m, 1 H, 4-H), 5.80 (bs, 1 H, NH).

(2S,3R,4S)-6: R_f 0.17 (petroleum ether/Et₂O, 8:2, 1% HOAc).

¹H NMR (selected signals) (400 MHz, CDCl₃): δ = 1.44 (s, 9 H, 1-H), 3.71 (s, 3 H, 6-H). ¹³C NMR (100 MHz, CDCl₃): δ = 18.0 (q, C-17), 28.3 (q, C-1), 32.1 (t, C-10), 33.5 (t, C-12), 41.9 (d, C-9), 50.5 (d, C-7), 52.7 (d, C-6), 79.9 (s, C-2), 138.6 (d, C-13), 155.4 (s, C-2), 172.4 (s, C-5).

HRMS (C₁₉H₃₀D₂NO₆ [M+H]⁺): calculated 372.2350, found 372.2311.

(2S,3S)-3-(tert-Butoxycarbonylamino)-2-[(S)-1,2-dideuterocyclohex-2-enyl]-4-methoxy-4-oxobutanoic acid [(2S,3S,4S)-7]



To a solution of acid **(2S,3S,4S)-6** (100 mg, 0.27 mmol) in dry dichloromethane (3 mL) benzylidene-bis(tricyclohexylphosphine)dichlororuthenium (2.7 mg, 0.003 mmol) was added at rt and stirring was continued for 2 h. The solvent was evaporated in

vacuo and the residue was purified by flash chromatography on silica gel (petroleum ether/EtOAc 9:1, 1% HOAc). The metathesis product **(2S,3S,4S)-7** (78 mg, 0.24 mmol, 89%) was obtained as colorless foam.

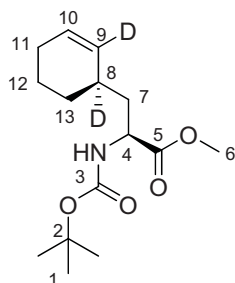
R_f 0.06 (petroleum ether/Et₂O, 9:1, 1% HOAc); [α]_D²⁰ = − 8.1 ° (c=1.2, CHCl₃); Rotameric ratio at rt: 4:1.

Major rotamer: ¹H NMR (400 MHz, CDCl₃): δ = 1.44 (s, 9 H, 1-H), 1.41–1.59 (m, 2 H, 14-H_a, 13-H_a), 1.78 (m, 1 H, 13-H_b), 1.88–2.01 (m, 3 H, 14-H_b, 12-H), 2.76 (d, ³J_{7,4} = 6.0 Hz, 1 H, 7-H), 3.75 (s, 3 H, 6-H), 4.71 (m, 1 H, 4-H), 5.38 (d, ³J_{NH,4} = 8.0 Hz, 1 H, NH), 5.75 (m, 1 H, 11-H), 10.13 (bs, 1 H, OH). ¹³C NMR (100 MHz, CDCl₃): δ = 21.5 (t, C-13), 24.9 (t, C-12), 26.6 (t, C-14), 28.3 (q, C-1), 33.9 (t, J_{9,D} = 19.3 Hz, C-9), 52.5 (d,q, C-4, C-6), 53.5 (d, C-7), 80.3 (s, C-2), 127.3 (t, J_{9,D} = 23.5 Hz, C-10), 129.0 (d, C-11), 155.1 (s, C-3), 171.4 (s, C-5), 177.4 (s, C-8).

Minor rotamer (selected signals): ¹H NMR (400 MHz, CDCl₃): δ = 4.53 (m, 1 H, 4-H), 6.13 (bs, 1 H, NH). ¹³C NMR (100 MHz, CDCl₃): δ = 26.7 (t, C-14), 52.8 (d,q, C-4, C-6), 53.7 (d, C-7), 81.8 (s, C-2), 128.5 (d, C-11), 155.6 (s, C-3), 171.2 (s, C-5), 170.9 (s, C-8), 176.7 (s, C-8).

HRMS (C₁₆H₂₄D₂NO₆ [M+H]⁺): calculated 330.1880, found 330.1853.

(S)-Methyl 2-(tert-butoxycarbonylamino)-3-[(R)-1,2-dideuterocyclohex-2-enyl]-propanoate [(2S,4R)-8]



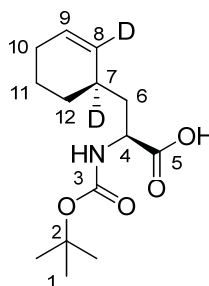
To a solution of acid **(2S,3S,4S)-7** (191 mg, 0.578 mmol) in dry THF (1 mL) triethylamine (500 μ L, 3.57 mmol) and phenylphosphonic acid dichloride (90%, 100 μ L, 643 μ mol) was added at 0 °C. After stirring for 15 min at 0 °C and 15 min at rt, the reaction mixture was cooled to −10 °C. Mercaptopyridine-*N*-oxide (189 mg, 1.49 mmol) was added and stirring was continued for 1.5 h in the dark. To the resulting mixture subsequently *tert*-butylmercaptane (670 μ L, 5.95 mmol) and triethylborane (1M in THF, 200 μ L, 200 μ mol) were added at −10 °C. After 3h, the reaction mixture was diluted with Et₂O (10 mL) and washed with 1 N KHSO₄ (5 mL), sat. NaHCO₃ (5 mL) and brine (5 mL). After drying over Na₂SO₄, the solvent was evaporated *in vacuo*, and the residue was purified by column chromatography on silica gel (petroleum ether/EtOAc, 95:5). The deuterated cyclohexenyl alanine **(2S,4R)-8** (80.2 mg, 281 μ mol, 49%) was obtained as colorless solid. R_f 0.35 (petroleum ether/Et₂O, 8:2). [α]_D²⁰ = − 31.3 ° (c=1.1, CHCl₃); rotameric ratio at rt: 85:15.

Major rotamer: ^1H NMR (400 MHz, CDCl_3): δ = 1.24 (m, 1 H, 13- H_a), 1.42 (s, 9 H, 1-H), 1.48–1.58 (m, 2 H, 7- H_a , 12- H_a), 1.68 (m, 1 H, 12- H_b), 1.73–1.81 (m, 2 H, 7- H_b , 13- H_b), 1.95 (m, 2 H, 11-H), 3.71 (s, 3 H, 6-H), 4.38 (m, 1 H, 4-H), 4.96 (d, $^3J_{\text{NH},4}$ = 7.8 Hz, 1 H, NH), 5.70 (m, 1 H, 10-H). ^{13}C NMR (100 MHz, CDCl_3): δ = 21.0 (t, C-12), 25.1 (t, C-11), 28.3 (q, C-1), 29.0 (s, C-13), 30.9 (t, $J_{8,\text{D}}$ = 19.0 Hz, C-8), 39.3 (t, C-7), 51.6 (d, C-4), 52.2 (q, C-6), 79.9 (s, C-2), 127.9 (d, C-10), 129.5 (t, $J_{9,\text{D}}$ = 23.5 Hz, C-9), 155.3 (s, C-3), 173.8 (s, C-5).

Minor rotamer (selected signals): ^1H NMR (400 MHz, CDCl_3): δ = 4.22 (m, 1 H, 4-H), 4.74 (bs, 1 H, NH).

HRMS ($\text{C}_{15}\text{H}_{24}\text{D}_2\text{NO}_4$ $[\text{M}+\text{H}]^+$): calculated 286.1982, found 286.2002

(S)-2-(tert-Butoxycarbonylamino)-3-[(R)-1,2-dideuterocyclohex-2-enyl]propanoic acid [(2S,4R)-9]



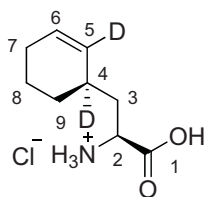
To a solution of ester **(2S,4R)-8** (65.0 mg, 228 μmol) in methanol (0.5 mL) was added 1 M aq. NaOH (250 μL , 250 μmol). The reaction mixture was heated to 60 $^\circ\text{C}$ for 1 h. The solvent was evaporated *in vacuo* and the residue was acidified with 1 M KHSO_4 . The mixture was extracted with EtOAc (3 x 10 mL) and dried over Na_2SO_4 . The solvent was removed again *in vacuo* and the resulting product (61.5 mg, 228 μmol , 100%) was used without further purification. $[\alpha]_{\text{D}}^{20}$ = -55.0° (c = 1.1, CHCl_3); rotameric ratio at rt: 2:1.

Major rotamer: ^1H NMR (400 MHz, CDCl_3): δ = 1.27 (m, 1 H, 12- H_a), 1.44 (s, 9 H, 1-H), 1.50–1.61 (m, 2 H, 6- H_a , 11- H_a), 1.70 (m, 1 H, 11- H_b), 1.77–1.87 (m, 2 H, 6- H_b , 12- H_b), 1.97 (m, 2 H, 10-H), 4.40 (m, 1 H, 4-H), 4.99 (d, $^3J_{\text{NH},4}$ = 8.0 Hz, 1 H, NH), 5.72 (t, $^3J_{9,10}$ = 3.4 Hz, 1 H, 9-H), 9.66 (bs, 1 H, OH). ^{13}C NMR (100 MHz, CDCl_3): δ = 21.0 (t, C-11), 25.1 (t, C-10), 28.3 (q, C-1), 29.0 (s, C-12), 30.9 (t, $J_{7,\text{D}}$ = 20.3 Hz, C-7), 38.9 (t, C-6), 51.6 (d, C-4), 80.2 (s, C-2), 128.0 (d, C-9), 129.4 (t, $J_{8,\text{D}}$ = 21.4 Hz, C-8), 155.6 (s, C-3), 178.1 (s, C-5).

Minor rotamer (selected signals): ^1H NMR (400 MHz, CDCl_3): δ = 4.23 (m, 1 H, 4-H), 6.37 (bs, 1 H, NH). ^{13}C NMR (100 MHz, CDCl_3): δ = 39.3 (t, C-6), 52.7 (d, C-4), 81.7 (s, C-2), 156.6 (s, C-3), 177.5 (s, C-5).

HMRS ($\text{C}_{14}\text{H}_{22}\text{D}_2\text{NO}_4$ $[\text{M}+\text{H}]^+$): calculated 272.1825, found 272.1847.

(S)-2-Amino-3-[(R)-1,2-dideuterocyclohex-2-enyl]propanoic acid hydrochloride [(2S,4R)-1]



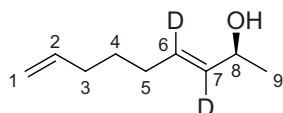
HCl (4 M in Dioxane, 380 μ L, 2.40 mmol) was added dropwise to the Boc-protected amino acid **(2S,4R)-9** (41.5 mg, 151 μ mol) at rt. After stirring for 1 h, the solvent was evaporated *in vacuo* to yield the HCl-salt **(2S,4R)-1** (31.3 mg, 151 μ mol, 100%) as white crystalline solid; mp. 200 $^{\circ}$ C (decomp.). $[\alpha]_{\text{D}}^{20} = -19.3^{\circ}$ (c = 1.1, MeOH).

^1H NMR (400 MHz, CD_3OD): δ = 1.31 (m, 1 H, 9- H_a), 1.58 (m, 1 H, 8- H_a), 1.73–1.79 (m, 2 H, 3- H_a , 8- H_b), 1.90–1.96 (m, 2 H, 3- H_b , 9- H_b), 2.01 (m, 2 H, 7-H), 4.03 (dd, $^3J_{2,3a} = ^3J_{2,3b} = 7.0$ Hz, 1 H, 2-H), 5.78 (t, $^3J_{6,7} = 3.4$ Hz, 1 H, 6-H). ^{13}C NMR (100 MHz, CD_3OD): δ = 21.9 (t, C-8), 26.1 (t, C-7), 29.6 (t, C-9), 31.9 (t, $J_{4,\text{D}} = 19.2$ Hz, C-4), 38.3 (t, C-3), 52.0 (d, C-2), 129.6 (d, C-6), 129.9 (t, $J_{5,\text{D}} = 23.7$ Hz, C-5), 172.2 (s, C-1).

HRMS ($\text{C}_9\text{H}_{15}\text{D}_2\text{NO}_2$ $[\text{M}+\text{H}]^+$): calculated 172.1301, found 172.1304.

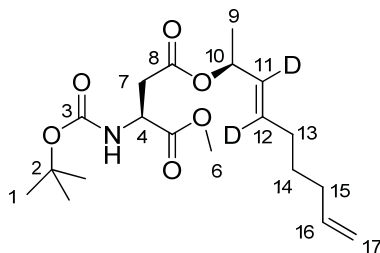
3) Synthesis of (2S,4S)-1

(S,E)-3,4-Dideuteronona-3,8-dien-2-ol [(S)-4]



(S)-Non-8-en-3-yn-2-ol **(S)-2** (1.00 g, 7.24 mmol) was reduced with LiAlD_4 (609 mg, 14.5 mmol) according to the procedure described for **(R)-4**, yielding alcohol **(S)-4** (808 mg, 5.69 mmol, 79%) as a colorless oil. R_f 0.16 (pentane/ Et_2O , 8:2); $[\alpha]_{\text{D}}^{20} = -9.7^{\circ}$ (c=1.0, CHCl_3); NMR according to **(R)-4**.

(S)-4-[(S,E)-3,4-dideuteronona-3,8-dien-2-yl] 1-methyl 2-(tert-butoxycarbonyl-amino)succinate [(2S,S)-5]



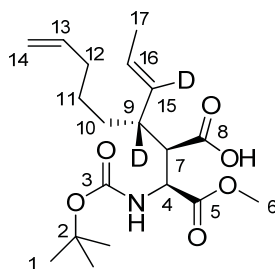
(S)-4 (687 mg, 4.83 mmol) was esterified with (*L*)-1-methyl-*N*-*boc* aspartate (1.22 g, 4.93 mmol), using 4-(dimethylamino)pyridine (59.0 mg, 0.84 mmol) and *N,N'*-dicyclohexylcarbodiimide (1.09 g, 5.28 mmol), according to the procedure described for **(2S,*R*)-5**. Ester **(2S,*S*)-5** (1.52g, 4.09 mmol, 85%) was obtained as colorless oil. *R*_f 0.27 (petroleum ether/EtOAc, 9:1). $[\alpha]_D^{20} = -17.6^\circ$ (*c*=1.0, CHCl₃); rotameric ratio at rt 9:1.

Major rotamer: ¹H NMR (400 MHz, CDCl₃): δ = 1.26 (d, ³*J*_{9,10} = 6.5 Hz, 1 H, 9-H), 1.42 (s, 9 H, 1-H), 1.44 (m, 2 H, 14-H), 1.99–2.05 (m, 4 H, 13-H, 15-H), 2.75 (dd, ²*J*_{7a,7b} = 16.8 Hz, ³*J*_{7a,4} = 4.8 Hz, 1 H, 7-H_a), 2.94 (dd, ²*J*_{7b,7a} = 16.8 Hz, ³*J*_{7b,4} = 4.5 Hz, 1 H, 7-H_b), 3.72 (s, 3 H, 6 H), 4.54 (m, 1 H, 4-H), 4.92 (ddt, ³*J*_{17trans,16} = 10.4 Hz, ²*J*_{17trans,17cis} = 2 Hz, ⁴*J*_{17trans,15} = 1.2 Hz, 1 H, 17-H_{trans}), 4.97 (ddt, ³*J*_{17cis,16} = 17.2 Hz, ²*J*_{17cis,17trans} = ⁴*J*_{17cis,15} = 2.0 Hz, 1 H, 17-H_{cis}), 5.29 (q, ³*J*_{10,9} = 6.4 Hz, 1 H, 10-H), 5.46 (d, ³*J*_{NH,4} = 8.5 Hz, 1 H, N-H), 5.76 (ddt, ³*J*_{16,17cis} = 17.2 Hz, ³*J*_{16,17trans} = 10.4 Hz, ³*J*_{16,15} = 6.7 Hz, 1 H, 16-H). ¹³C NMR (100 MHz, CDCl₃): δ = 20.2 (q, C-9), 27.9 (t, C-14), 28.2 (q, C-1), 31.2 (t, C-13 oder C-15), 33.0 (t, C-13 oder C-15), 37.1 (t, C-7), 49.9 (d, C-4), 52.5 (q, C-6), 71.9 (d, C-10), 80.0 (s, C-2), 114.6 (t, C-17), 128.8 (t, *J*_{7,D} = 23.3 Hz, C-11), 132.9 (t, *J*_{7,D} = 23.2 Hz, C-12), 138.4 (d, C-16), 155.3 (s, C-3), 170.1 (s, C-8), 171.5 (s, C-8).

Minor rotamer (selected signals): ¹H NMR (400 MHz, CDCl₃): δ = 4.40 (m, 1 H, 4-H), 5.13 (bs, 1 H, NH).

HRMS (C₁₄H₂₂D₂NO₄ [M–Boc+H]⁺): calculated 272.1825, found 272.1835.

(2*R*,3*R*)-2-((*S*)-1-(*tert*-Butoxycarbonylamino)-2-methoxy-2-oxoethyl)-3-deutero-3-((*Z*)-1-deuterioprop-1-enyl)oct-7-enoic acid [(2*S*,3*R*,4*R*)-6]



According to **(2S,3*S*,4*S*)-6**, **(2S,3*R*,4*R*)-6** was prepared from **(2S,*S*)-5** (1.43 g, 3.85 mmol), hexamethyldisilazide (2.4 mL, 11.5 mmol), *n*-butyllithium (2.5 M in hexanes, 3.08 mL, 7.69 mmol) and TMSCl (2.45 mL, 19.2 mmol), yielding the major diastereomer **(2S,3*R*,4*R*)-6** (1.20 g, 3.22 mmol, 85%) and the minor diastereomer **(2S,3*S*,4*R*)-6** (65.0 mg, 0.177 mmol, 4%) as colorless oils.

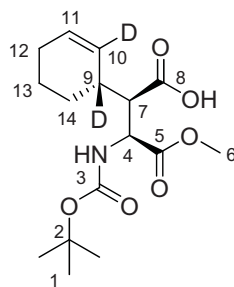
(2S,3*R*,4*R*)-6: *R*_f 0.18 (pentane/EtOAc, 8:2, 1% HOAc); $[\alpha]_D^{20} = +24.7^\circ$ (*c* = 1.0, CHCl₃); ¹H NMR (400 MHz, CDCl₃): δ = 1.21–1.32 (m, 2 H, 10-H_a, 11-H_a), 1.36–1.49 (m, 10 H, 11-H_b, 1-H), 1.56 (m, 1 H, 10-H_b), 1.65 (d, ³*J*_{17,16} = 6.5 Hz, 3 H, 17-H), 2.01 (m, 2 H, 12-H), 2.98 (d, ³*J*_{7,4} = 3.9 Hz, 1 H, 7-H), 3.72 (s, 3 H, 6-H), 4.65 (dd,

$^3J_{4,\text{NH}} = 10.1$ Hz, $^3J_{4,7} = 6.5$ Hz, 1 H, 4-H), 4.93 (m, 1 H, 14- H_{trans}), 4.98 (m, 1H, 14- H_{cis}), 5.47 (q, $^3J_{16,17} = 6.5$ Hz, 1 H, 16-H), 5.63 (d, $^3J_{\text{NH},4} = 10.1$ Hz, 1 H, NH), 5.77 (ddt, $^3J_{13,14\text{cis}} = 16.9$ Hz, $^3J_{13,14\text{trans}} = 10.2$ Hz, $^3J_{13,12} = 6.7$ Hz, 1 H, 13-H); ^{13}C NMR (100 MHz, CDCl_3): $\delta = 17.8$ (q, C-17), 26.4 (t, C-11), 28.3 (q, C-1), 31.2 (t, C-10), 33.4 (t, C-12), 50.8 (d, C-7), 52.3 (d, C-4), 52.6 (q, C-6), 80.1 (s, C-2), 114.4 (t, C-14), 128.0 (d, C-16), 138.6 (d, C-13), 155.7 (d, C-3), 172.1 (s, C-5), 178.1 (s, C-8).

(2S,3S,4R)-6: R_f 0.15 (pentane/EtOAc, 8:2, 1% HOAc); Selected Signals: ^1H NMR (400 MHz, CDCl_3): $\delta = 1.44$ (s, 9 H, 1-H), 1.60 (d, $^3J_{16,17} = 6.5$ Hz, 3 H, 17-H), 2.87 (d, $^3J_{7,4} = 3.3$ Hz, 1 H, 7-H), 3.74 (s, 3 H, 6-H), 4.50 (dd, $^3J_{4,\text{NH}} = 9.9$ Hz, $^3J_{4,7} = 3.3$ Hz, 1 H, 4-H), 5.30 (m, 1 H, 16-H), 5.58 (bs, 1 H, NH). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 17.9$ (q, 17-C), 26.3 (t, C-11), 33.5 (t, C-12), 52.5 (d, C-4), 52.6 (q, C-6), 114.5 (t, C-14), 138.5 (d, C-13), 155.6 (d, C-3), 172.1 (s, C-5).

HRMS ($\text{C}_{19}\text{H}_{30}\text{D}_2\text{NO}_6$ $[\text{M}+\text{H}]^+$): calculated 372.2350, found 372.2342.

(2R,3S)-3-(tert-butoxycarbonylamino)-2-[(R)-1,2-dideuterocyclohex-2-enyl]-4-methoxy-4-oxobutanoic acid [(2S,3R,4R)-7]



According to **(2S,3S,4S)-7**, **(2S,3R,4R)-7** was prepared from **(2S,3R,4R)-6** (790 mg, 2.13 mmol) and Grubbs I catalyst (21.0 mg, 25.5 μmol), yielding a colorless foam (550 mg, 1.55 mmol, 86%).

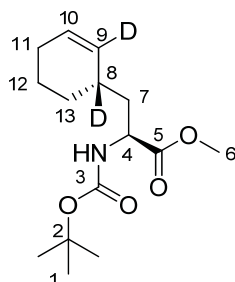
(2S,3R,4R)-7: R_f 0.18 (pentane/EtOAc, 8:2, 1% HOAc); $[\alpha]_{\text{D}}^{20} = + 53.3^\circ$ ($c = 1.0$, CHCl_3); rotameric ratio at rt 9:1.

Major rotamer ^1H NMR (400 MHz, CDCl_3): $\delta = 1.45$ (s, 9 H, 1-H), 1.40–1.61 (m, 2 H, 14- H_a , 13- H_a), 1.75 (m, 1 H, 13- H_b), 1.87 (m, 1 H, 14- H_b), 2.01 (m, 2 H, 12-H), 2.91 (d, $^3J_{7,4} = 3.6$ Hz, 1 H, 7-H), 3.73 (s, 3 H, 6-H), 4.66 (dd, $^3J_{4,\text{NH}} = 10.7$ Hz, $^3J_{4,7} = 3.7$ Hz, 1 H, 4-H), 5.67 (d, $^3J_{\text{NH},4} = 10.7$ Hz, 1 H, NH), 5.81 (t, $^3J_{11,12} = 3.4$ Hz, 1 H, 11-H), 7.61 (bs, 1 H, OH). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 20.4$ (t, C-13), 25.0 (t, C-12), 26.3 (t, C-14), 28.3 (q, C-1), 33.3 (t, $J_{9,\text{D}} = 22.3$ Hz, C-9), 50.9 (d, C-7), 52.6 (q, C-6), 52.7 (d, C-4), 80.4 (s, C-2), 129.0 (d, C-11), 155.2 (s, C-3), 171.6 (s, C-5), 177.2 (s, C-8).

Minor rotamer (selected signals): ^1H NMR (400 MHz, CDCl_3): $\delta = 4.46$ (m, 1 H, 4-H), 6.09 (bs, 1 H, NH).

HRMS ($\text{C}_{16}\text{H}_{24}\text{D}_2\text{NO}_6$ $[\text{M}+\text{H}]^+$): calculated 330.1880, found 330.1878.

(S)-Methyl 2-(tert-butoxycarbonylamino)-3-[(S)-1,2-dideuterocyclohex-2-enyl]-propanoate [(2S,4S)-8]



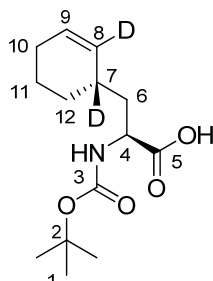
According to the procedure described for **(2S,4R)-8**, the **(2S,4S)**-diastereomer was prepared from **(2S,3R,4R)-7** (380 mg, 1.15 mmol), PhPOCl₂ (90%, 196 μ L, 1.24 mmol), Et₃N (970 μ L, 6.92 mmol), mercaptopyridine *N*-oxide (365 mg, 2.88 mmol), *tert*-butyl mercaptane (1.30 mL, 11.5 mmol) and triethyl borane (1M in THF, 400 μ L, 400 μ mol) yielding the cyclohexenyl alanine derivative (186 mg, 652 μ mol, 57%) as colorless resin. *R*_f 0.33 (pentane/EtOAc, 8:2,); [α]_D²⁰ = + 47.0 ° (c=0.9, CHCl₃); rotameric ratio at rt 9:1.

Major rotamer ¹H NMR (400 MHz, CDCl₃): δ = 1.24 (m, 1 H, 13-H_a), 1.43 (s, 9 H, 1-H), 1.47–1.60 (m, 2 H, 7-H_a, 12-H_a), 1.65–1.74 (m, 2 H, 12-H_b, 7-H_b), 1.85 (m, 1 H, 13-H_b), 1.95 (m, 2 H, 11-H), 3.71 (s, 3 H, 6-H), 4.36 (m, 1 H, 4-H), 4.92 (d, ³*J*_{NH,4} = 8.1 Hz 1 H, NH), 5.69 (t, ³*J*_{10,11} = 3.4 Hz, 1 H, 10-H). ¹³C NMR (100 MHz, CDCl₃): δ = 20.9 (t, C-12), 25.1 (t, C-11), 28.1 (s, C-13), 28.3 (q, C-1), 31.9 (t, *J*_{8,D} = 18.8 Hz, C-8), 39.3 (t, C-7), 51.5 (d, C-4), 52.2 (q, C-6), 79.9 (s, C-2), 128.0 (d, C-10), 130.3 (t, *J*_{9,D} = 23.6 Hz, C-9), 155.4 (s, C-3), 173.9 (s, C-5).

Minor rotamer (selected signals): ¹H NMR (400 MHz, CDCl₃): δ = 4.22 (m, 1 H, 4-H), 4.74 (m, 1 H, NH).

HRMS (C₁₅H₂₄D₂NO₄ [M+H]⁺): calculated 286.1982, found 286.1986.

(S)-2-(tert-Butoxycarbonylamino)-3-[(S)-1,2-dideuterocyclohex-2-enyl]propanoic acid [(2S,4S)-9]



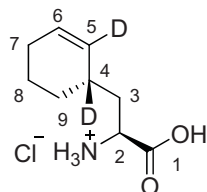
According to **(2S,4R)-9**, Ester **(2S,4S)-8** (150 mg, 526 μ mol) was saponified with NaOH (1.0 M, 580 μ L, 580 μ mol), yielding acid **(2S,4S)-9** (138 mg, 510 μ mol, 97%) as colorless resin. [α]_D²⁰ = + 77.9 ° (c=0.4, CHCl₃); rotameric ratio at rt 2:1.

Major rotamer ^1H NMR (400 MHz, CDCl_3): δ = 1.27 (m, 1 H, 12- H_a), 1.44 (s, 9 H, 1-H), 1.52 (m, 1 H, 11- H_a), 1.60–1.79 (m, 3 H, 6-H, 11- H_b), 1.86 (m, 1 H, 12- H_b), 1.97 (m, 2 H, 10-H), 4.38 (m, 1 H, 4-H), 4.94 (d, $^3J_{\text{NH},4}$ = 8.0 Hz 1 H, NH), 5.70 (t, $^3J_{9,10}$ = 3.5 Hz, 1 H, 10-H), 9.20 (bs, 1 H, OH). ^{13}C NMR (100 MHz, CDCl_3): δ = 20.9 (t, C-11), 25.1 (t, C-10), 28.0 (s, C-12), 28.3 (q, C-1), 31.0 (t, $J_{7,\text{D}}$ = 20.6 Hz, C-7), 38.8 (t, C-6), 51.5 (d, C-4), 80.2 (s, C-2), 128.1 (d, C-9), 130.2 (t, $J_{8,\text{D}}$ = 22.4 Hz, C-8), 155.7 (s, C-3), 178.1 (s, C-5).

Minor rotamer (selected signals): ^1H NMR (400 MHz, CDCl_3): δ = 4.20 (m, 1 H, 4-H), 6.22 (bs, 1 H, NH); ^{13}C NMR (100 MHz, CDCl_3): δ = 52.6 (d, C-4), 81.7 (s, C-2), 127.9 (d, C-9), 156.6 (s, C-3), 177.7 (s, C-5).

HRMS ($\text{C}_{14}\text{H}_{22}\text{D}_2\text{NO}_4$ $[\text{M}+\text{H}]^+$): calculated 272.1825, found 272.1875.

(S)-2-Amino-3-[(S)-1,2-dideuterocyclohex-2-enyl]propanoic acid hydrochloride (2S,4S)-1



According to **(2S,4R)-1**, **(2S,4S)-9** (97.0 mg, 309 μmol) was deprotected with HCl (4 M in dioxane, 750 μL , 3.00 mmol), giving the hydrochloride (64.1 mg, 309 μmol , 100%) as white crystalline solid, mp. 205 $^{\circ}\text{C}$ (decomp.). $[\alpha]_{\text{D}}^{20}$ = + 75.3 $^{\circ}$ (c =1.0, MeOH).

^1H NMR (400 MHz, CD_3OD): δ = 1.30 (m, 1 H, 9- H_a), 1.59 (m, 1 H, 8- H_a), 1.75 (m, 1 H, 8- H_b), 1.80–1.93 (m, 3 H, 3-H, 9- H_b), 2.01 (m, 2 H, 7-H), 4.01 (d, $^3J_{2,3}$ = 6.8 Hz, 1 H, 2-H), 5.76 (t, $^3J_{6,7}$ = 3.4 Hz, 1 H, 6-H). ^{13}C NMR (100 MHz, CD_3OD): δ = 21.7 (t, C-8), 26.1 (t, C-7), 29.3 (t, C-9), 31.8 (t, $J_{4,\text{D}}$ = 18.9 Hz, C-4), 38.2 (t, C-3), 52.0 (d, C-2), 129.5 (d, C-6), 130.2 (t, $J_{5,\text{D}}$ = 24.2 Hz, C-5), 172.3 (s, C-1).

HRMS ($\text{C}_{14}\text{H}_{22}\text{D}_2\text{NO}_4$ $[\text{M}+\text{H}]^+$): calculated 172.1301, found 172.1304.

4) Supplemental informations on feeding experiment

Materials

YMG medium¹ (D-glucose 20 g/L, meat extract 5 g/L, yeast extract 5 g/L, casein hydrolysate 3 g/L, sodium chloride 1.5 g/L, peptone 5 g/L) was used for *Streptomyces* sp. JS360 propagation. Q6 medium²¹ (D-glucose 0.5%, cotton seed meal 1%, glycerol 2%, calcium carbonate 0.1%, pH 7.2 adjusted with sodium hydroxide) was used for cinnabaramide production.

Methods

Feeding experiments:

Streptomyces sp. JS360 was grown in 20 mL V6 medium at 30 °C for 48 h and was transferred to 25 ml Q6 medium in dilution factor 1:50 supplemented variously with D-di-deuterated cyclohexenyl-alanine, L-di-deuterated cyclohexenyl-alanine in concentration of 0.1 mM. These substances were added in three portions to the bacterial culture after 0, 6, and 12 h, respectively, and afterwards cultures were further incubated at 30 °C for 24 h. To investigate the production level of cinnabaramides and their derivatives, culture supernatant was extracted by ethyl-acetate three times in ratio 1:1. After removal of the solvent in vacuo, the remainder was dissolved in methanol (1 ml) and 5 µL of the solution was injected for HPLC-MS analysis.

Analytical procedure:

High Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS) was performed with a HPLC-DAD system (Agilent 1100) coupled to an HCT ultra ESI-MS ion trap apparatus (Brucker Daltonics) operating in positive ionization mode. Separation was achieved by use of a Luna RP-C18 column with a solvent system consisting of water to acetonitrile gradient. For separation of cinnabaramides and their derivatives, a solvent system based on water (A)/acetonitrile (B) containing formic acid (0.1%) was applied. Gradient: 0-2 min 5% B; 2-22 min linear from 5% to 95% B; 22-25 min isocratic at 95% B; 25-27 min linear from 95-5% B.

High-resolution LC-MS measurements were performed with an Accela UPLC-system (Thermo-Fisher) coupled to an LTQ-Orbitrap (linear trap-FT-Orbitrap combination) operating in positive ionization mode with a Waters BEH-C18 column with a solvent system consisting of a water (A)/acetonitrile (B) gradient containing formic acid (0.1%).

² Stadler, M.; Bitzer, J.; Mayer-Bartschmid, A.; Müller, H.; Benet-Buchholz, J.; Gantner, F.; Tichy, H.-V.; Reinemer, P.; Bacon, K. B. *J. Nat. Prod.* **2007**, *70*, 246–252.

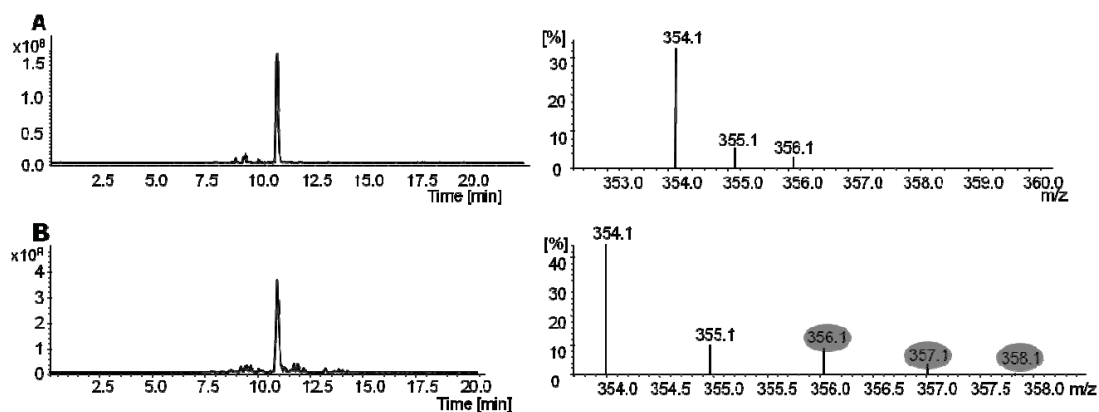
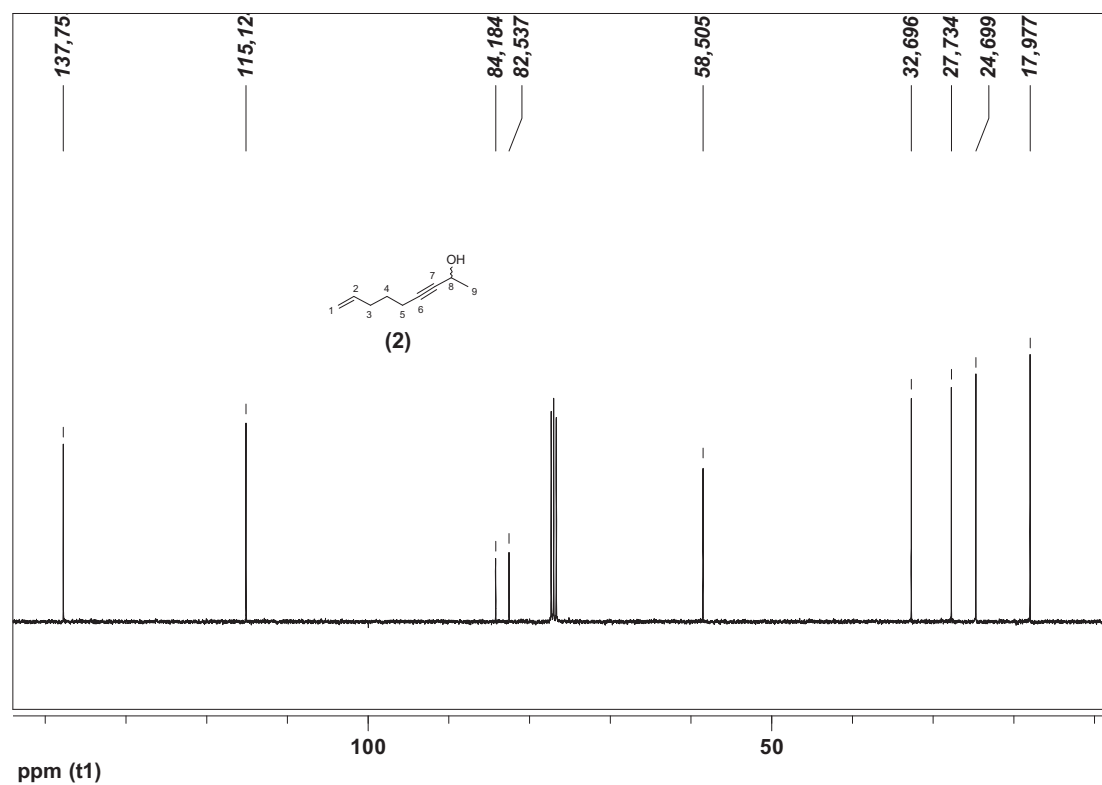
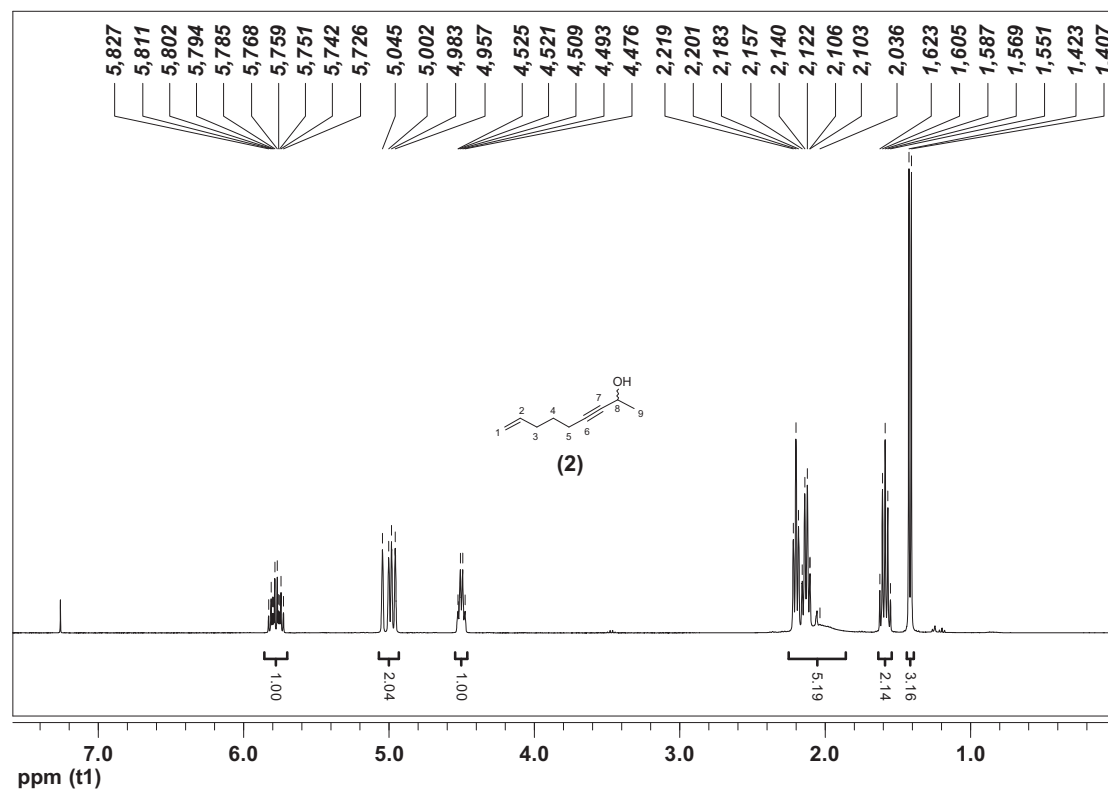


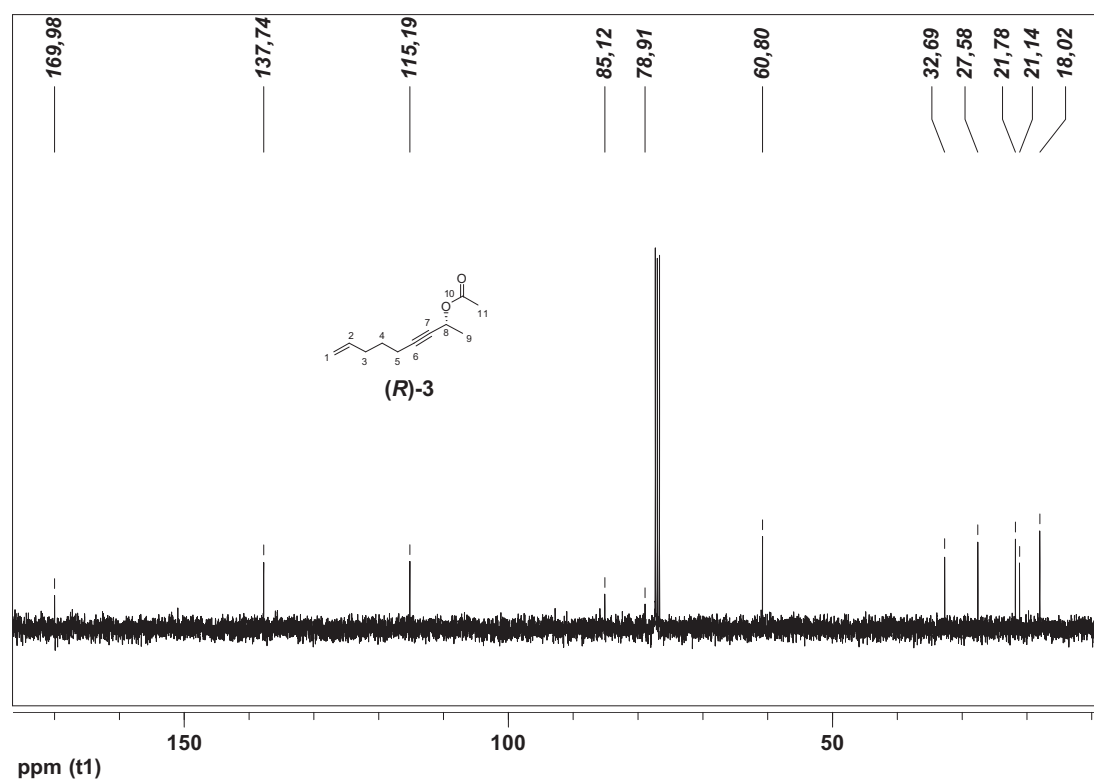
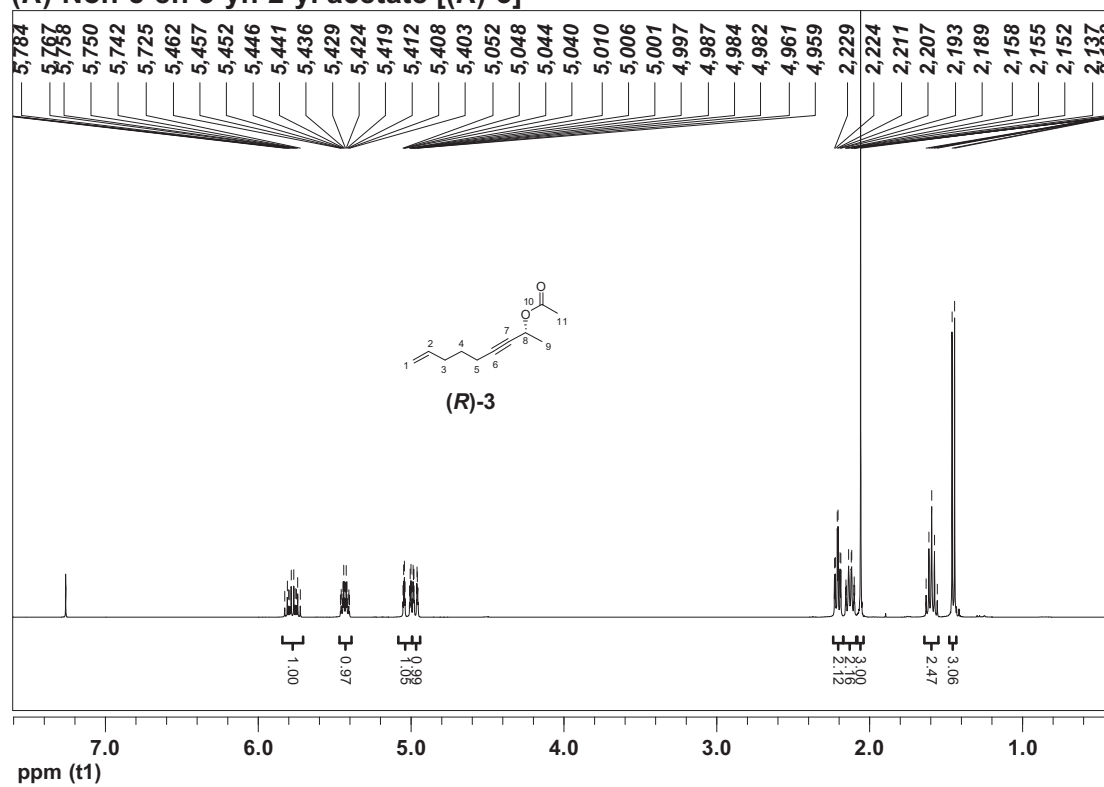
Figure S1: **A)** Left: HPLC-MS analysis of *S. sp* JS360 wild type extracts with extracted ion chromatogram (EIC) for $m/z = 354.2$ corresponding to cinnabaramide E ($t_R = 11.2$ min). Right: mass spectrum at $t_R = 11.2$ min showing isotopic pattern of cinnabaramide E. **B)** Left: HPLC-MS analysis of *S. sp* JS360 wild type extracts fed by D-di-deuterated cyclohexenyl-alanine with extracted ion chromatogram (EIC) for $m/z = 354.2$ corresponding to cinnabaramide E ($t_R = 11.2$ min). Right: mass spectrum at $t_R = 11.2$ min showing isotopic pattern of cinnabaramide E and di-deuterated cinnabaramide E analogue (highlighted in gray), showing that feeding of the wildtype strain with D-di-deuterated cyclohexenyl-alanine resulted in production of respective di-deuterated cinnabaramides.

5) Copies of NMR spectra

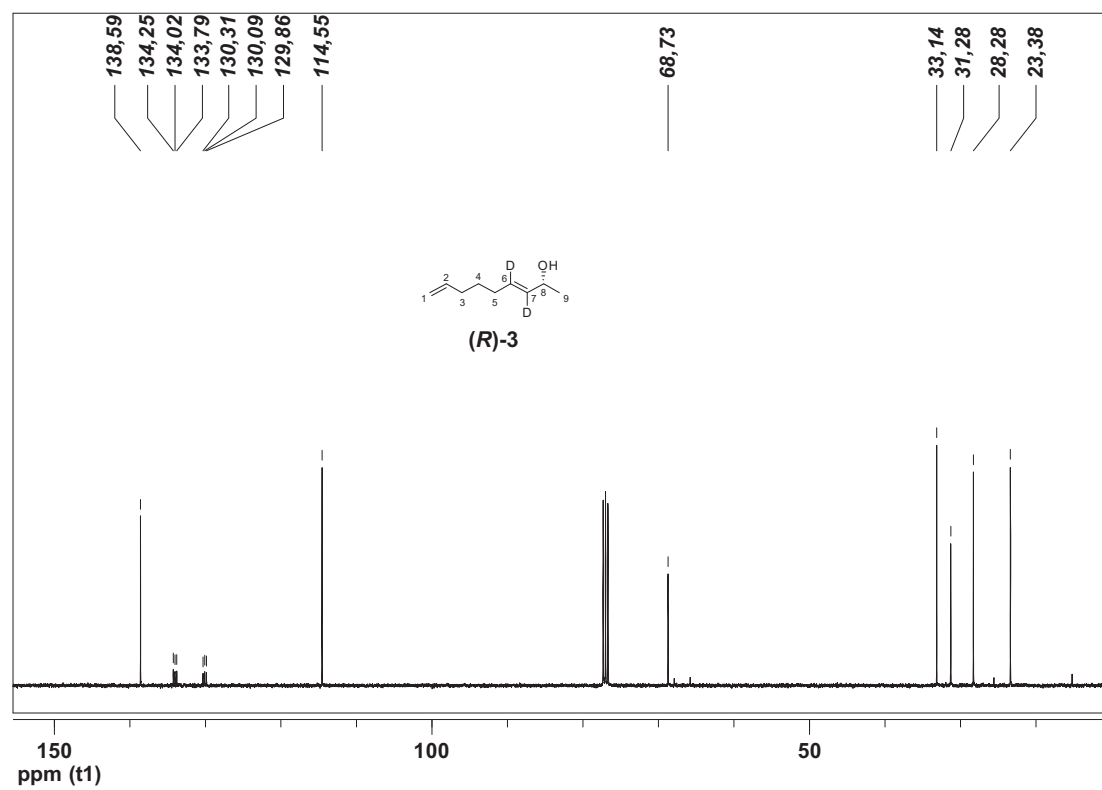
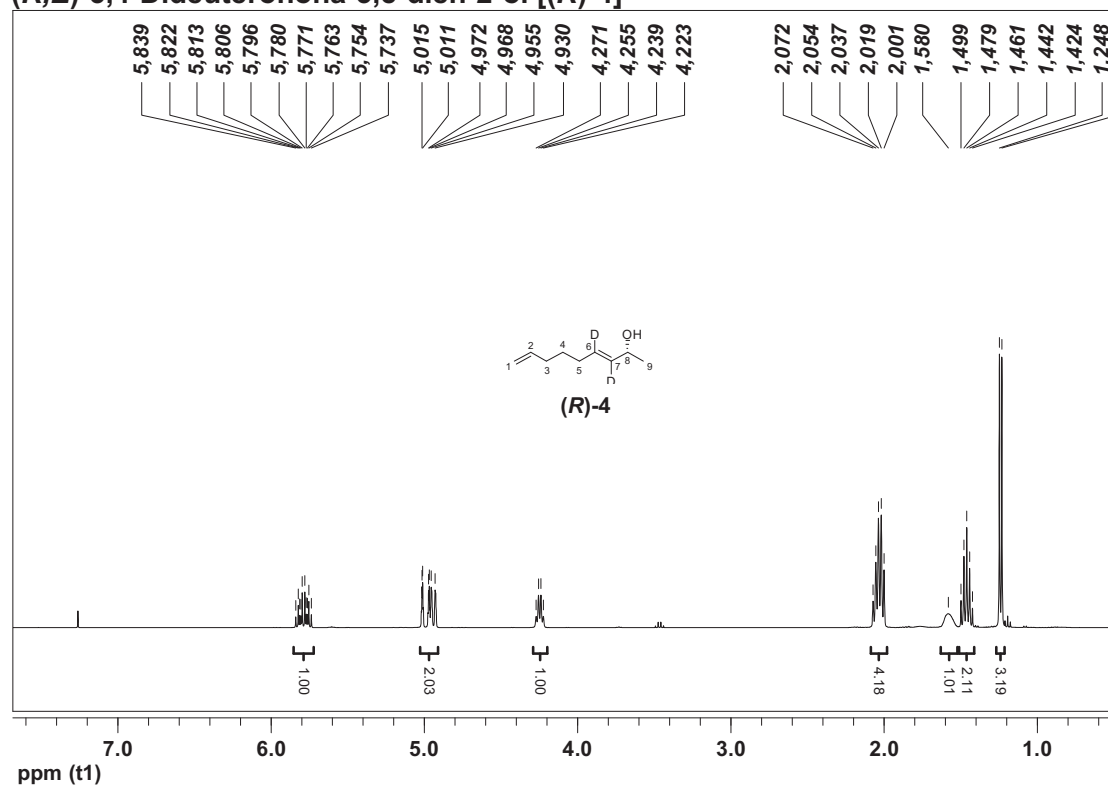
Non-8-en-3-yn-2-ol (2)



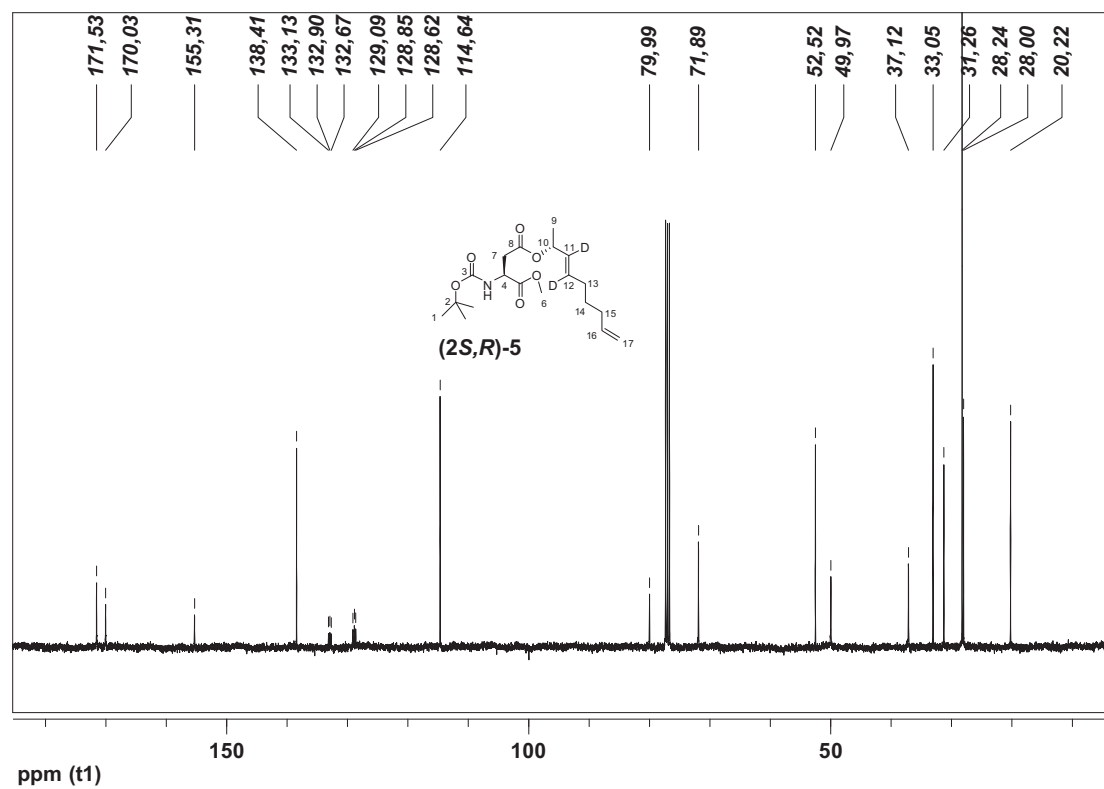
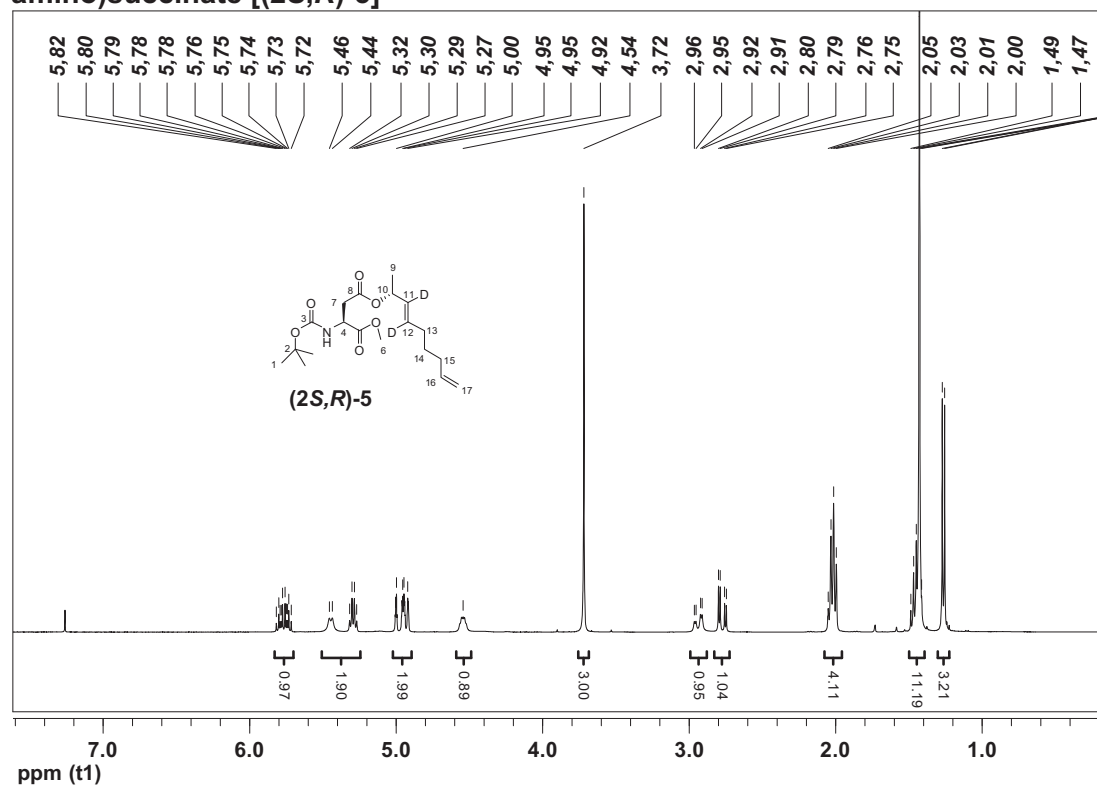
(R)-Non-8-en-3-yn-2-yl acetate [(R)-3]



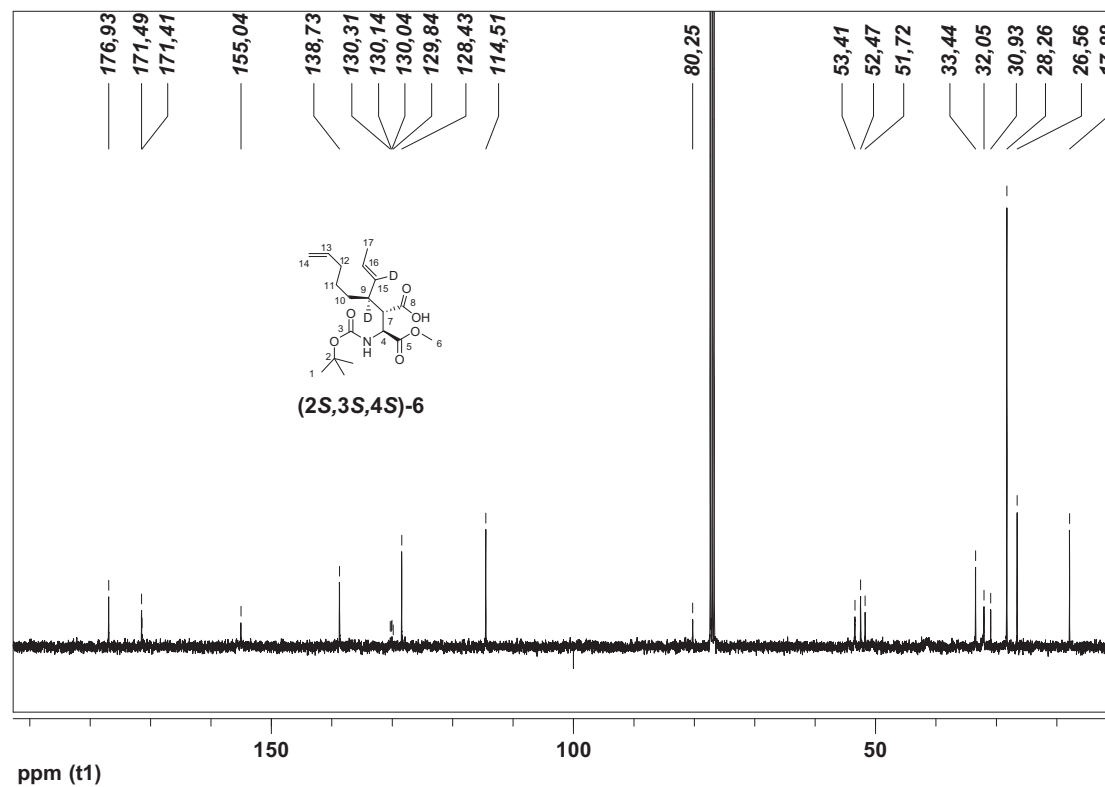
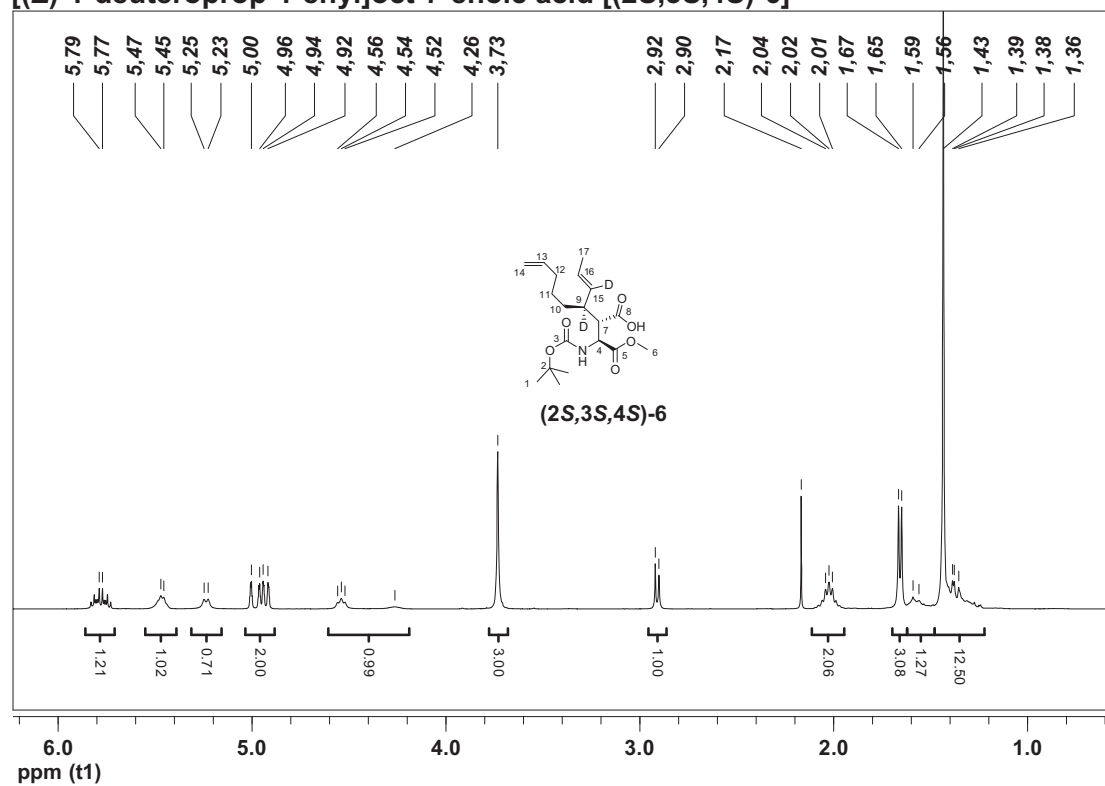
(R,E)-3,4-Dideuteronona-3,8-dien-2-ol [(R)-4]



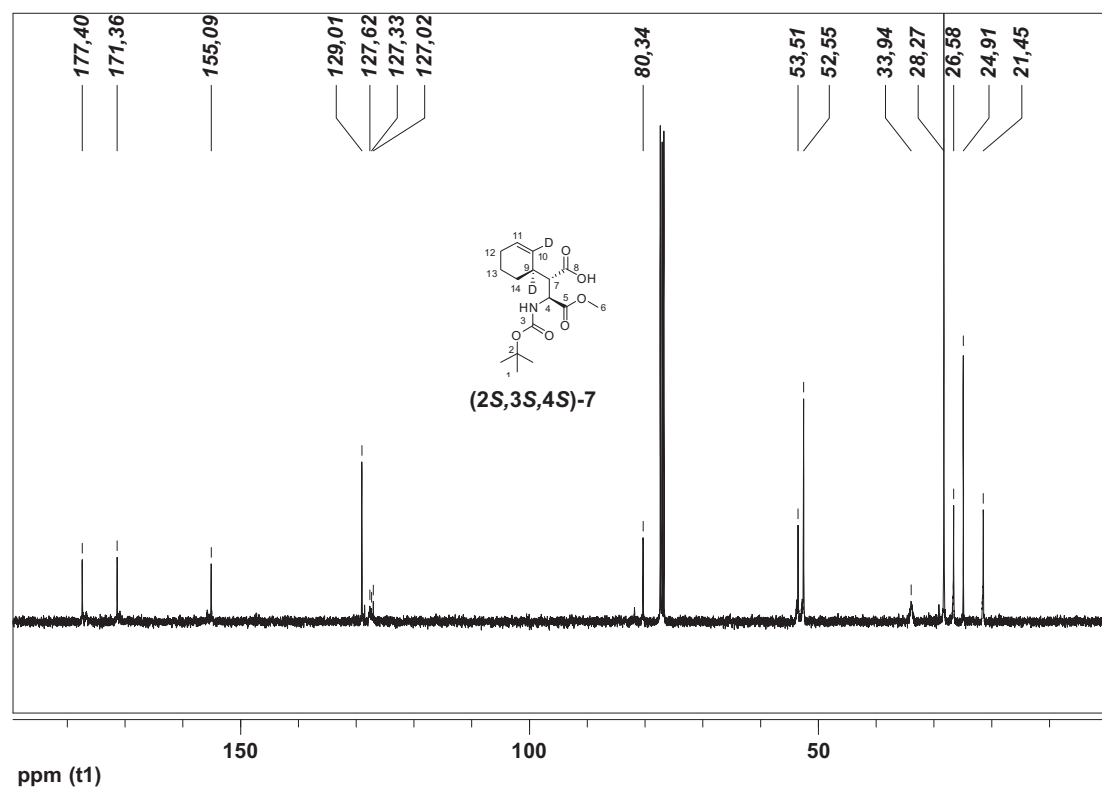
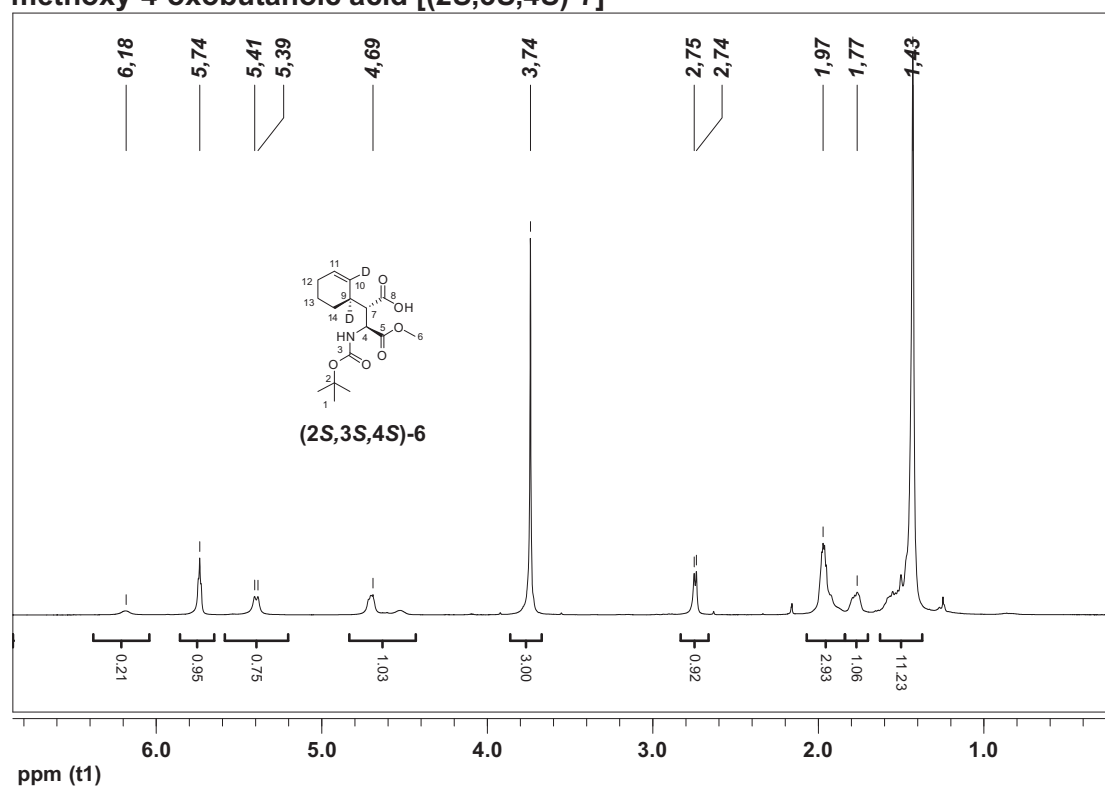
(S)-4-[(R,E)-3,4-Dideuteronona-3,8-dien-2-yl] 1-methyl 2-(tert-butoxycarbonyl-amino)succinate [(2S,R)-5]



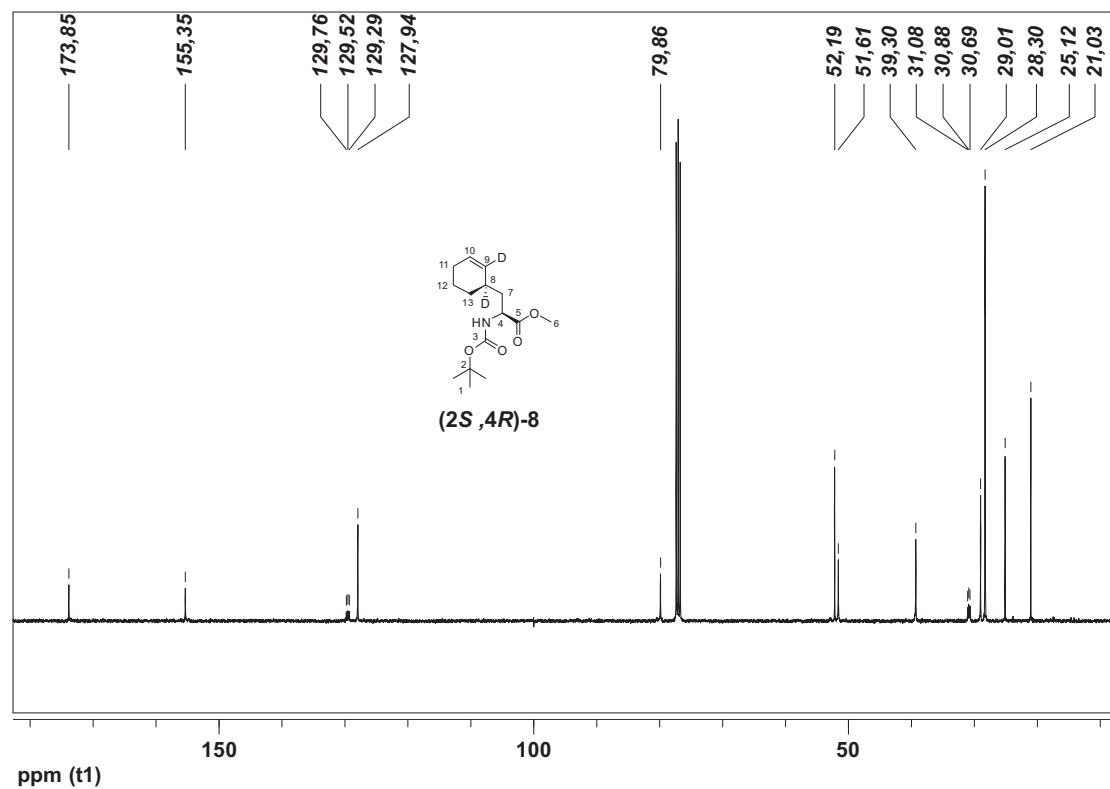
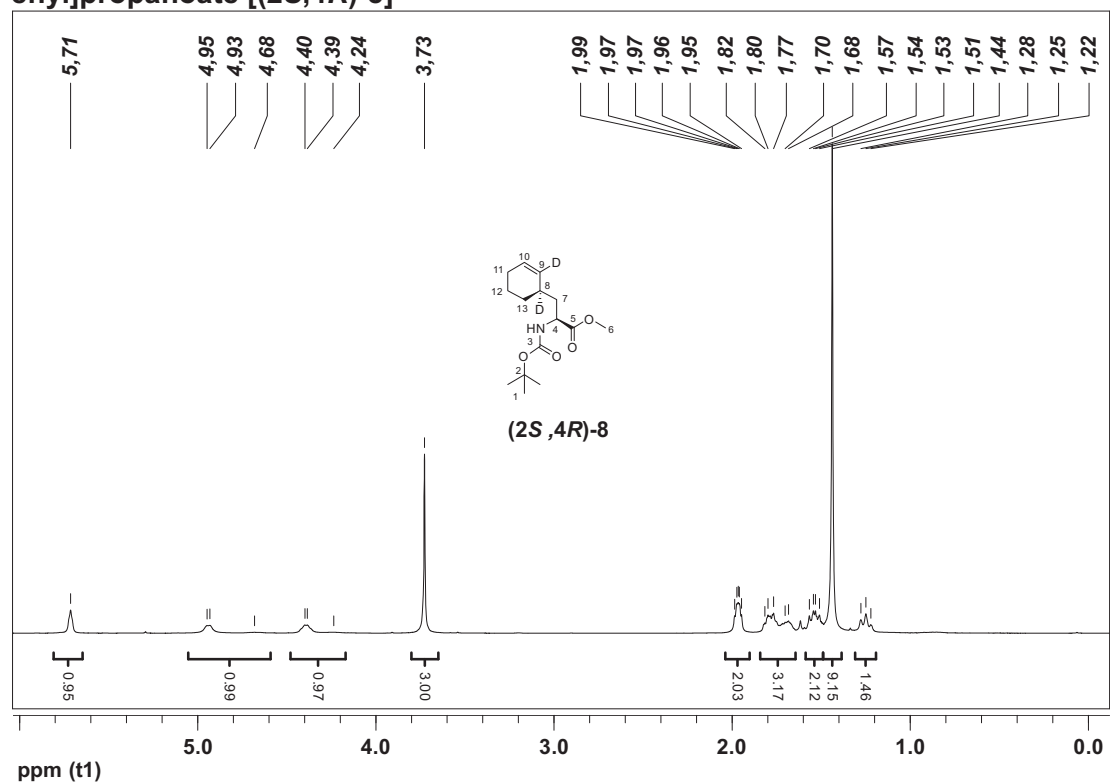
(2S,3S)-2-[(S)-1-(tert-Butoxycarbonylamino)-2-methoxy-2-oxoethyl]-3-deutero-3-[(Z)-1-deuterioprop-1-enyl]oct-7-enoic acid [(2S,3S,4S)-6]



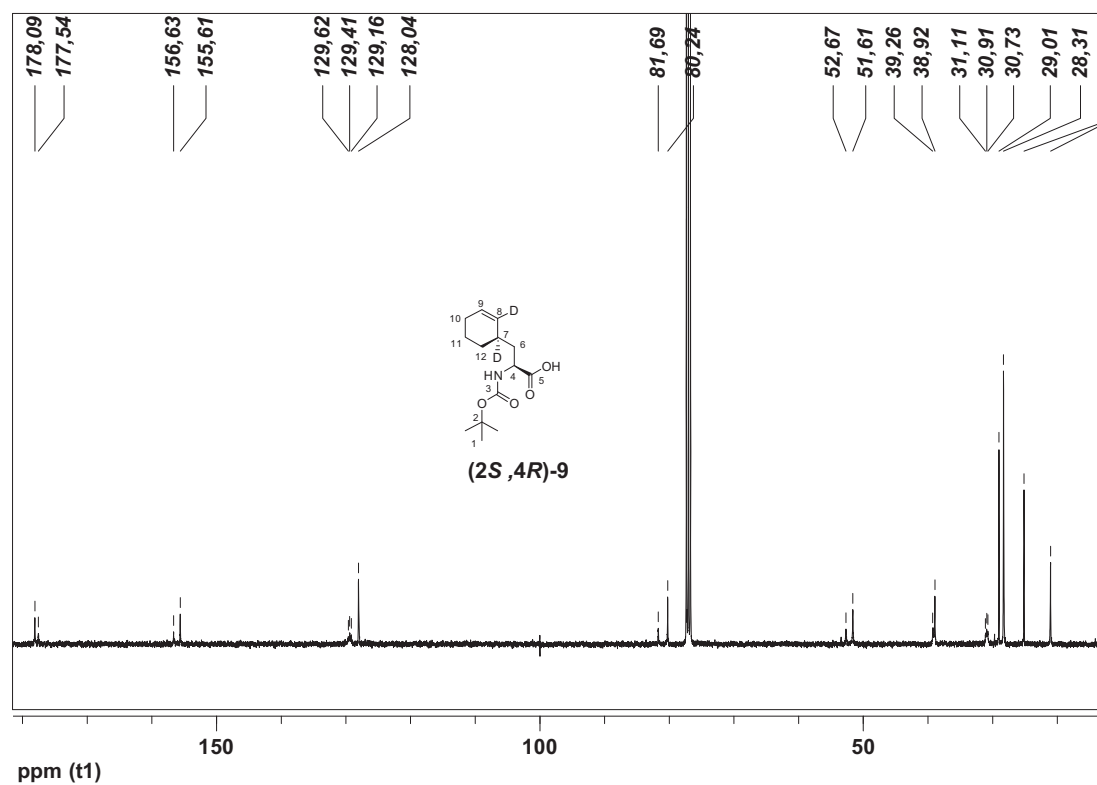
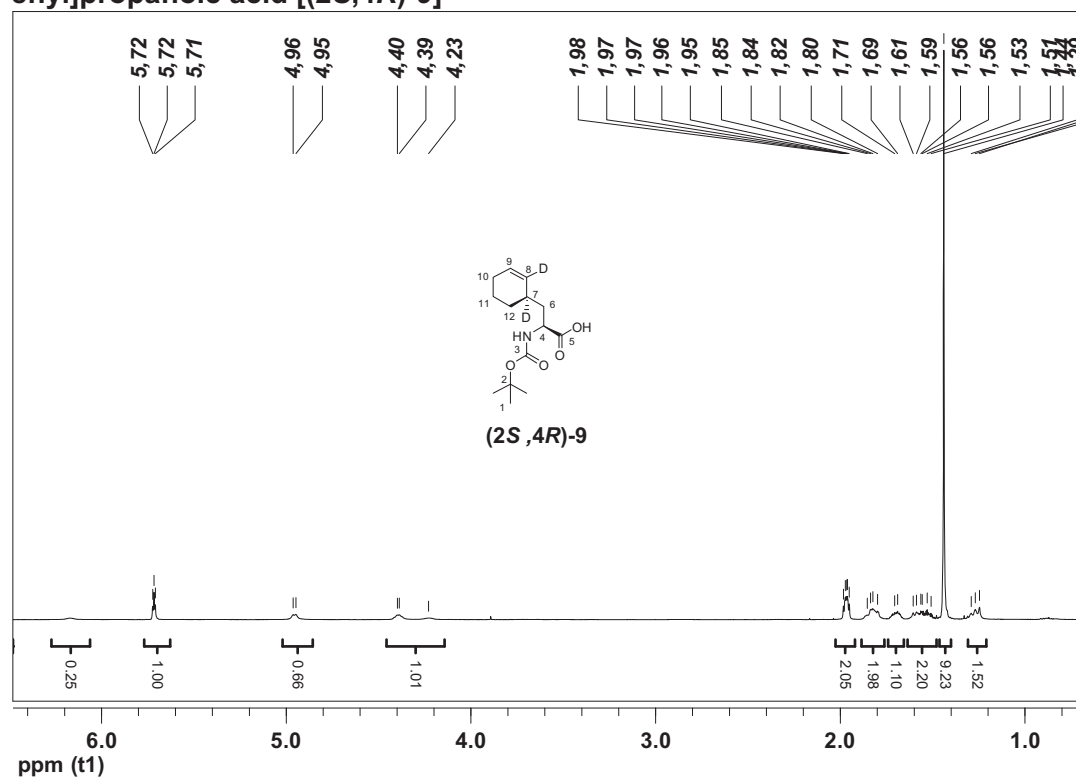
(2S,3S)-3-(tert-Butoxycarbonylamino)-2-[(S)-1,2-dideuterocyclohex-2-enyl]-4-methoxy-4-oxobutanoic acid [(2S,3S,4S)-7]



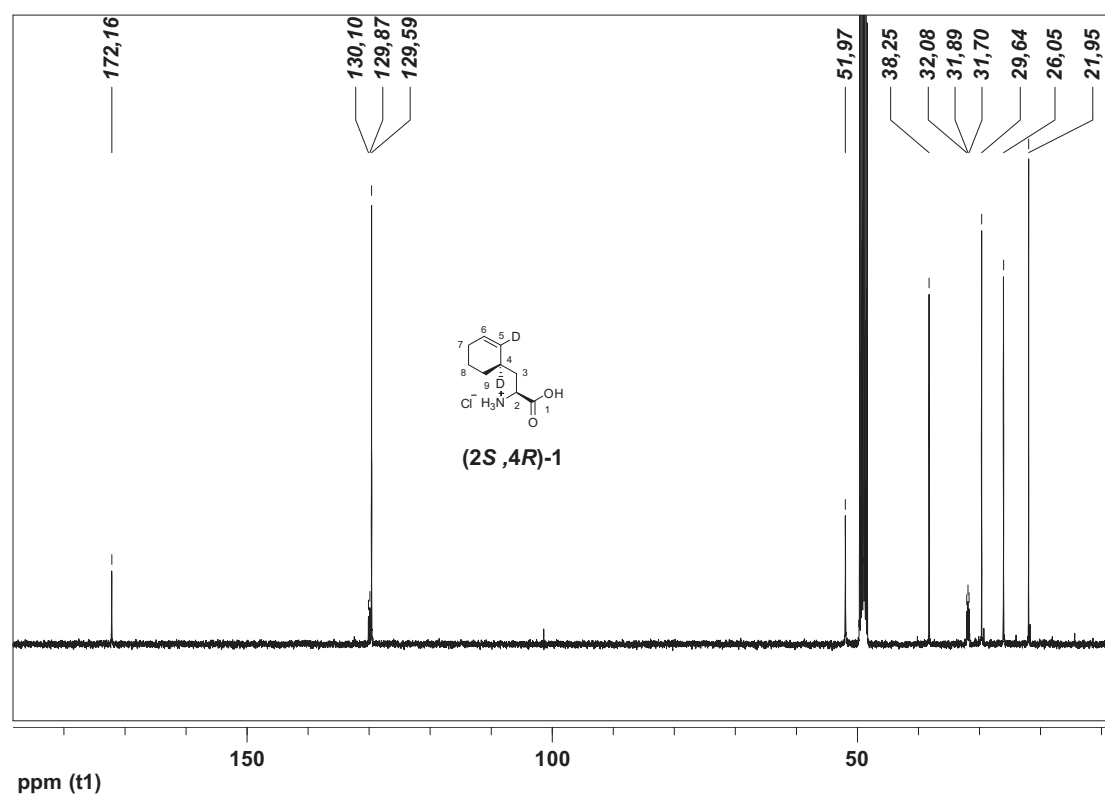
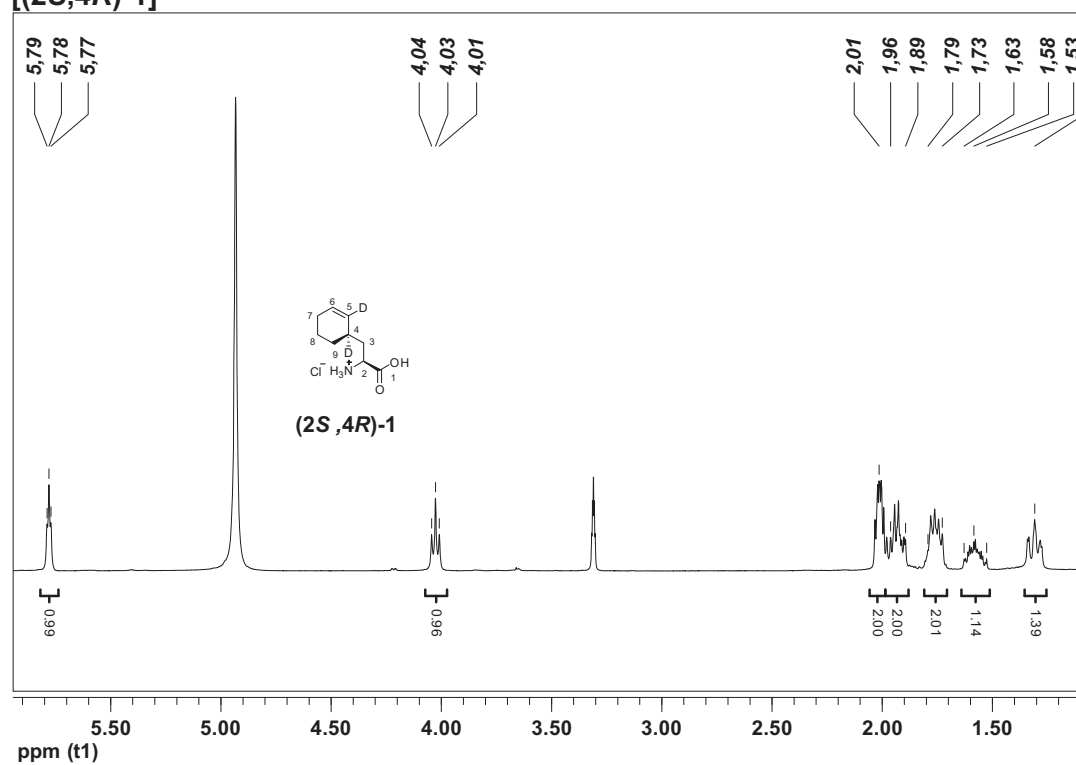
(S)-Methyl 2-(tert-butoxycarbonylamino)-3-[(R)-1,2-dideuteriocyclohex-2-enyl]propanoate [(2S,4R)-8]



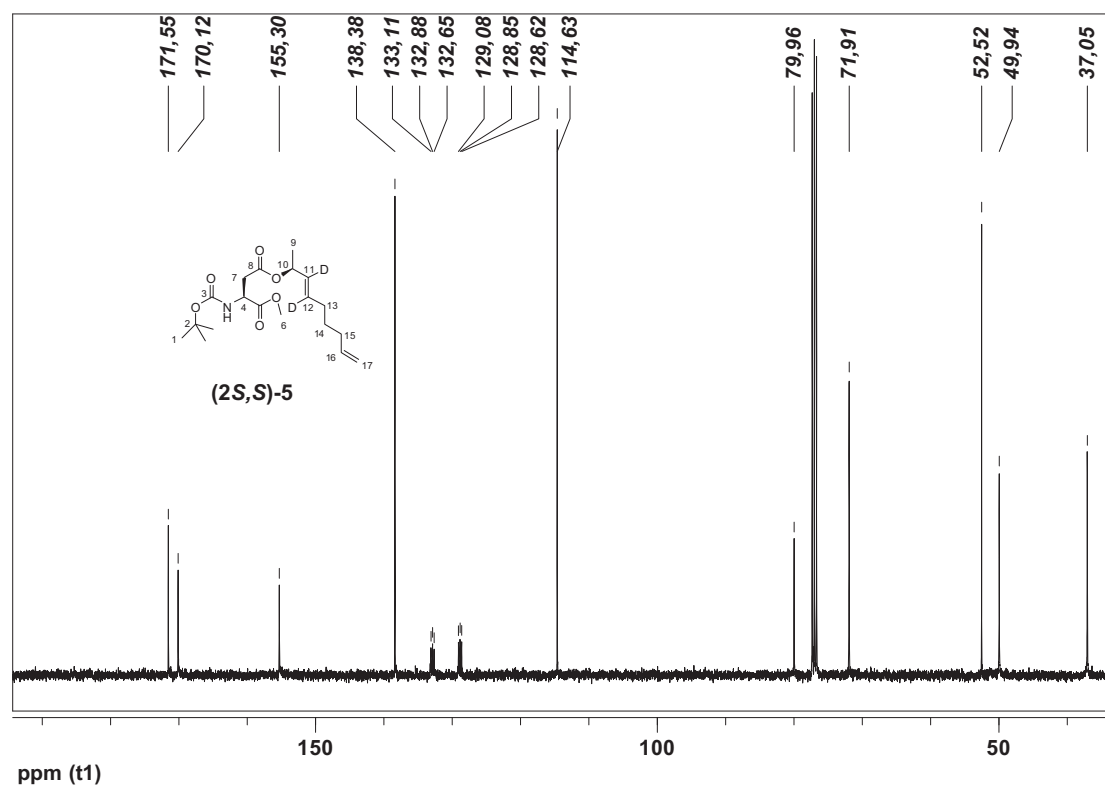
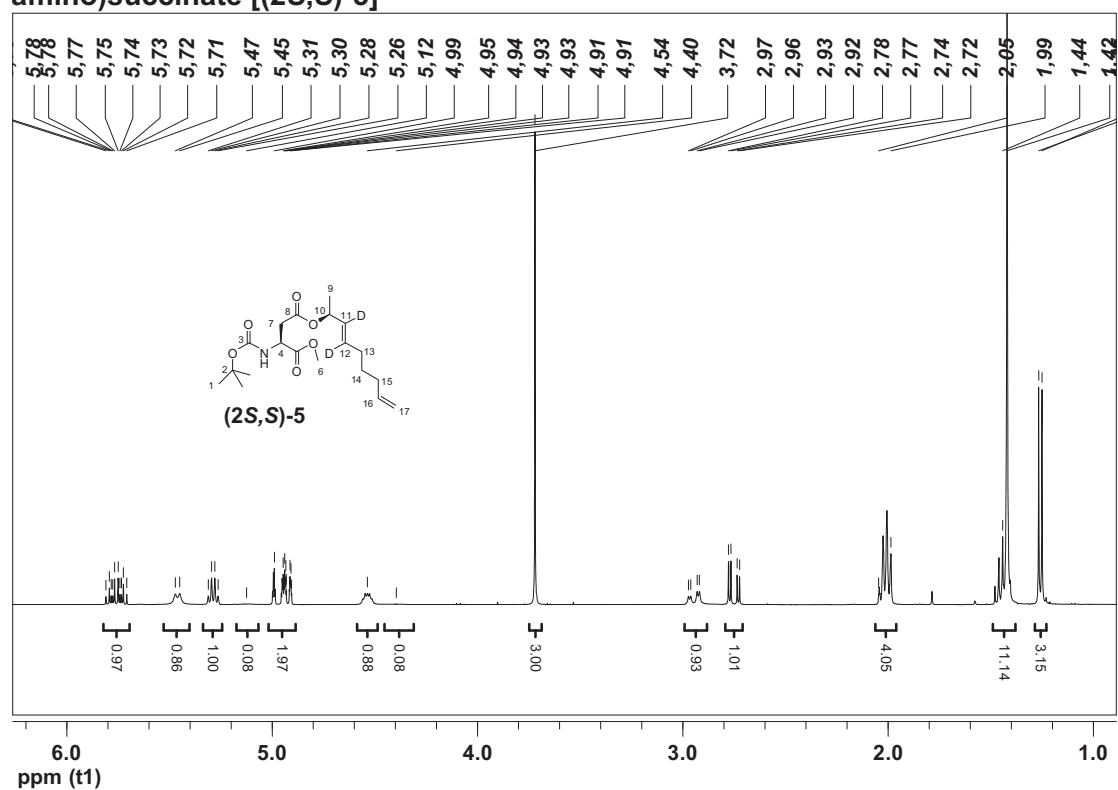
(S)-2-(tert-Butoxycarbonylamino)-3-[(R)-1,2-dideuteriocyclohex-2-enyl]propanoic acid [(2S,4R)-9]



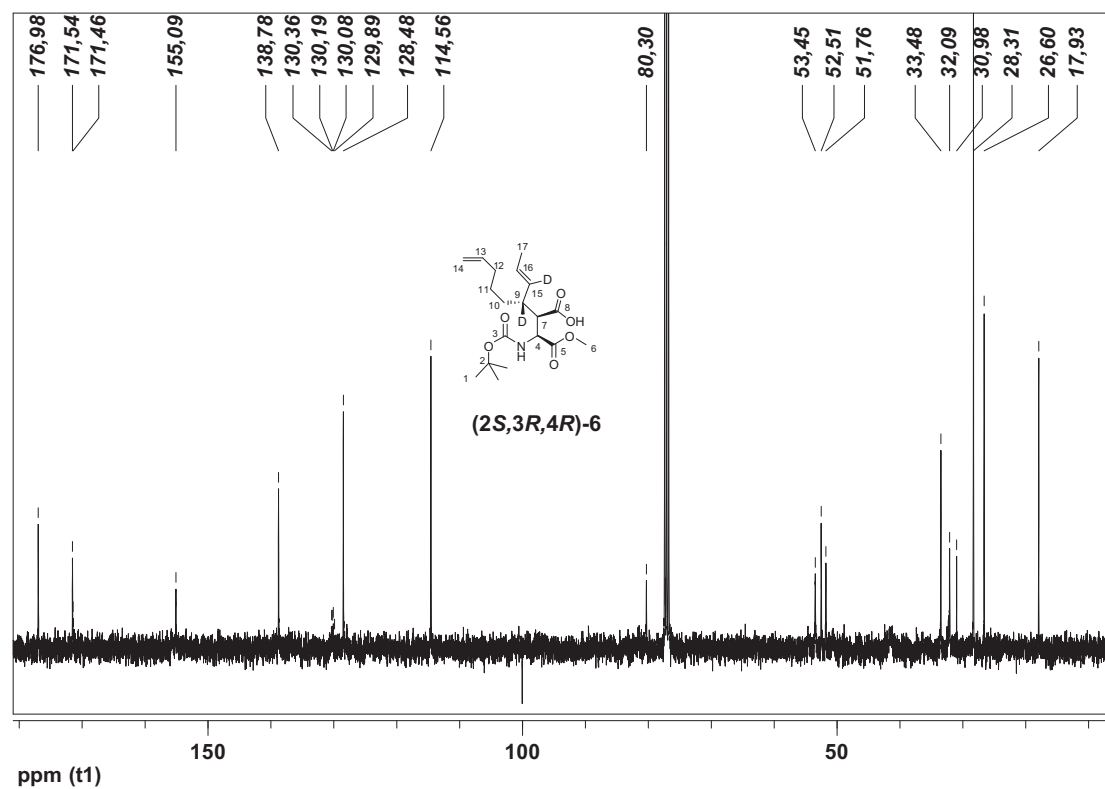
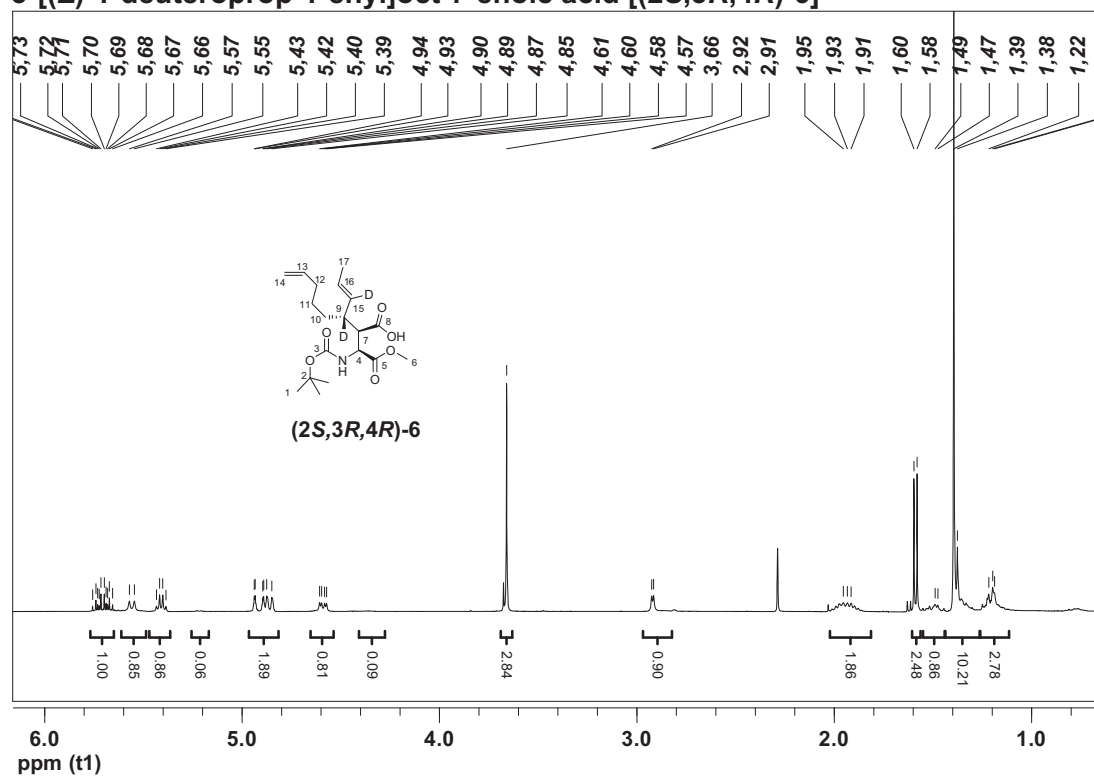
(S)-2-Amino-3-[(R)-1,2-dideuteriocyclohex-2-enyl]propanoic acid hydrochloride [(2S,4R)-1]



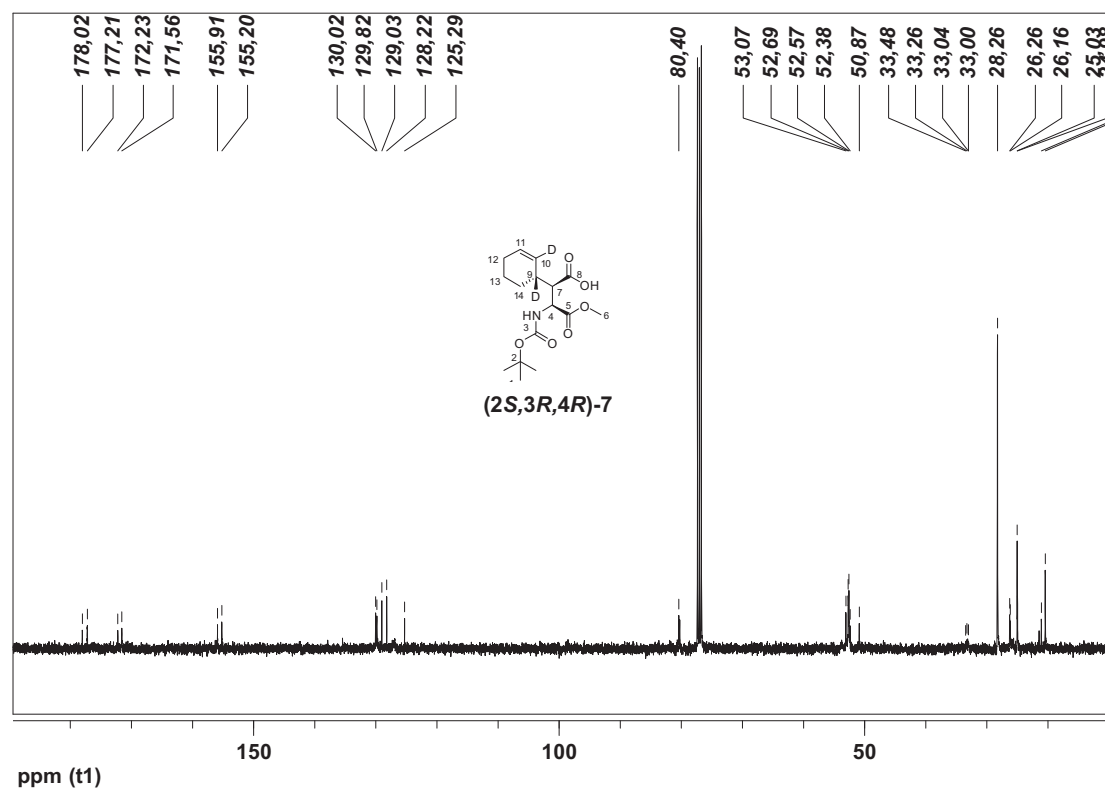
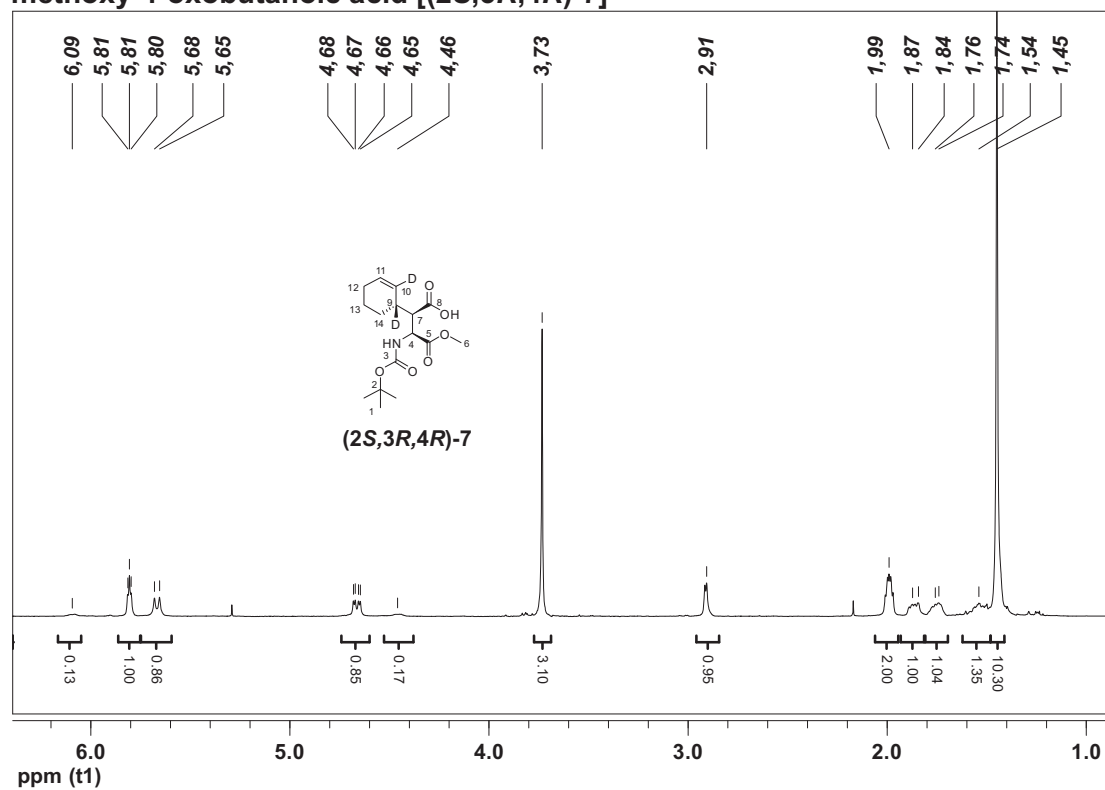
(S)-4-[(S,E)-3,4-Dideuteronona-3,8-dien-2-yl] 1-methyl 2-(tert-butoxycarbonyl-amino)succinate [(2S,S)-5]



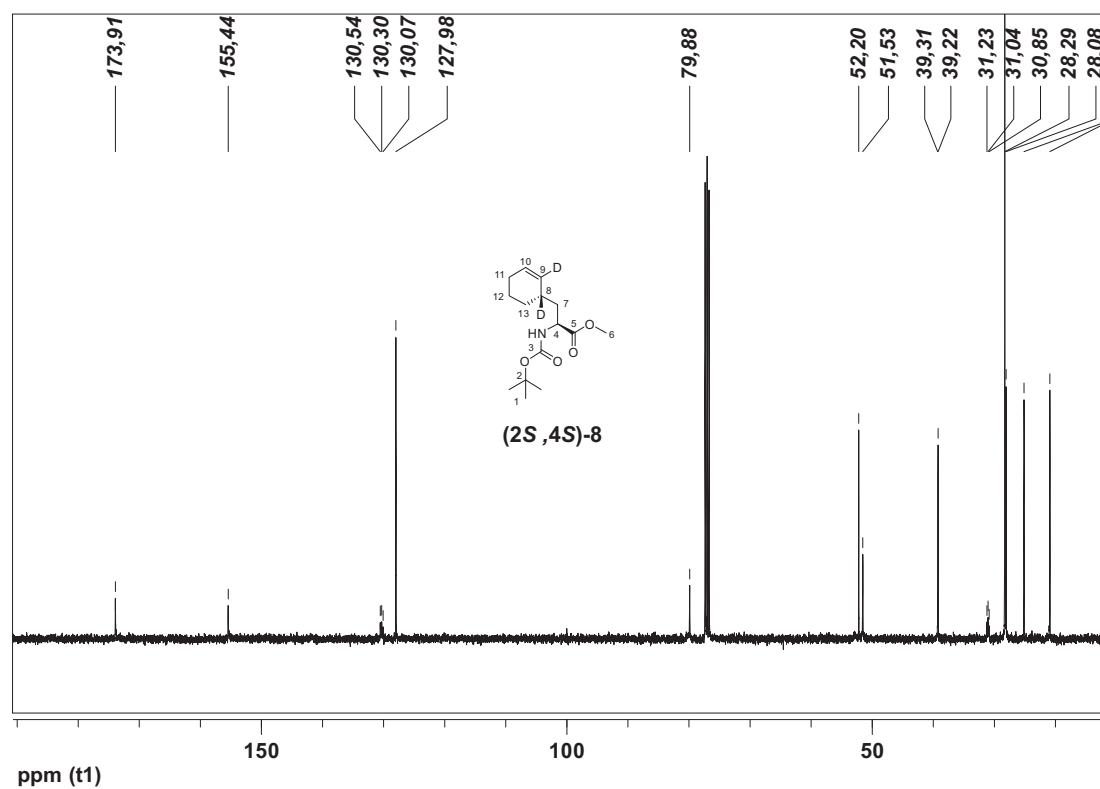
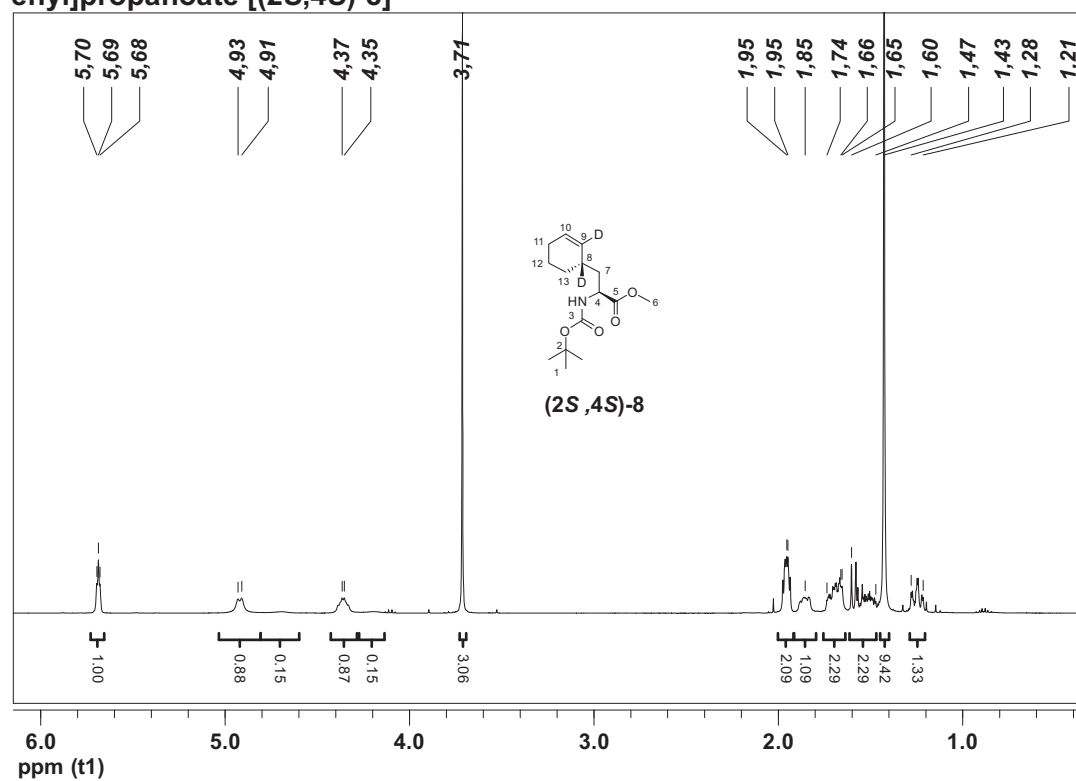
(2R,3R)-2-[(S)-1-(tert-Butoxycarbonylamino)-2-methoxy-2-oxoethyl]-3-deutero-3-[(Z)-1-deuterioprop-1-enyl]oct-7-enoic acid [(2S,3R,4R)-6]



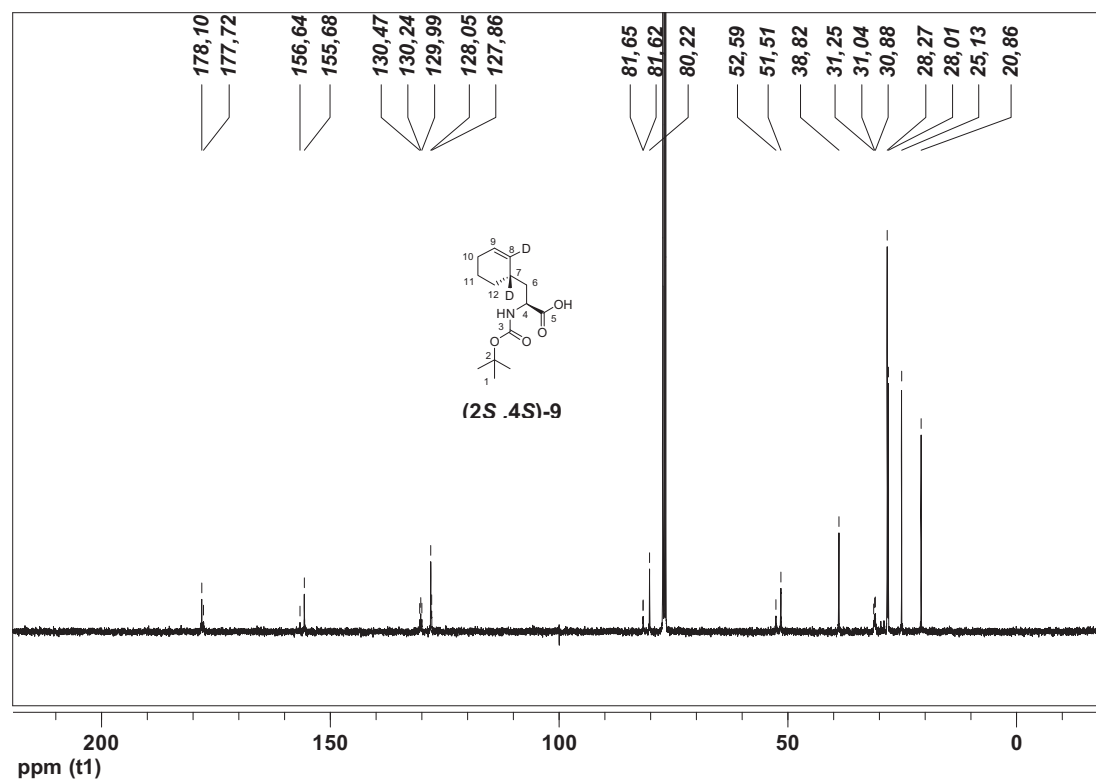
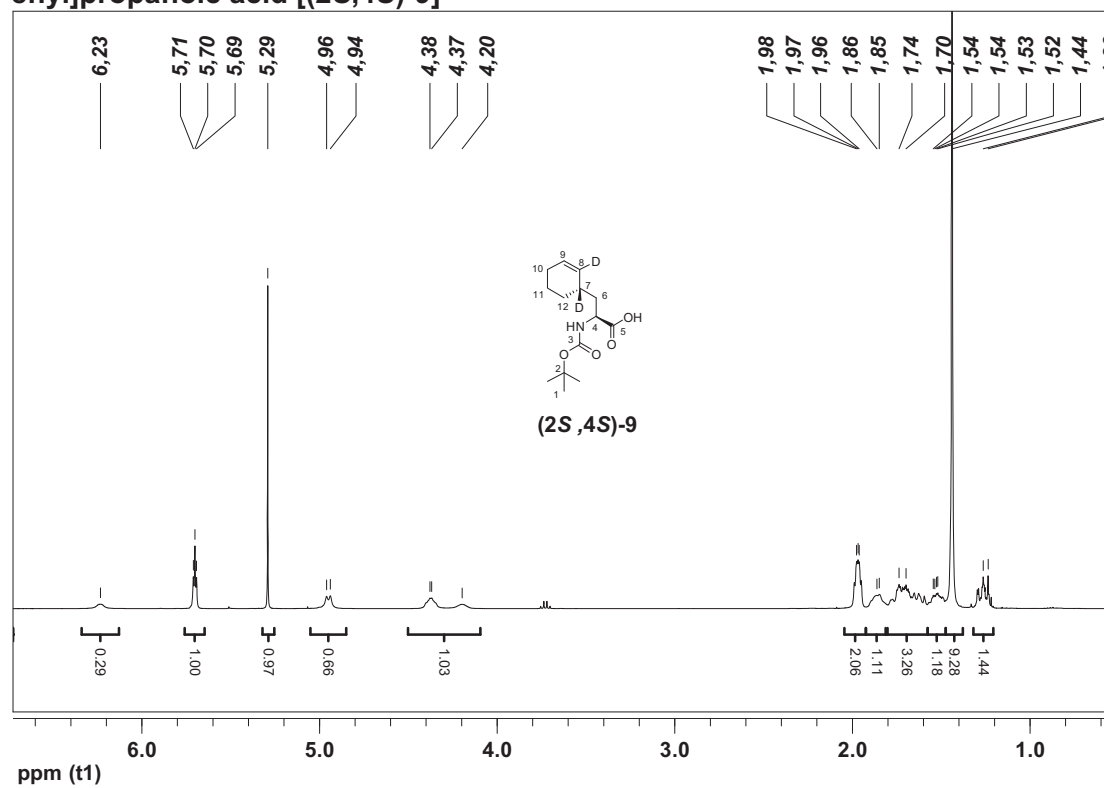
(2R,3S)-3-(tert-Butoxycarbonylamino)-2-[(R)-1,2-dideuterocyclohex-2-enyl]-4-methoxy-4-oxobutanoic acid [(2S,3R,4R)-7]



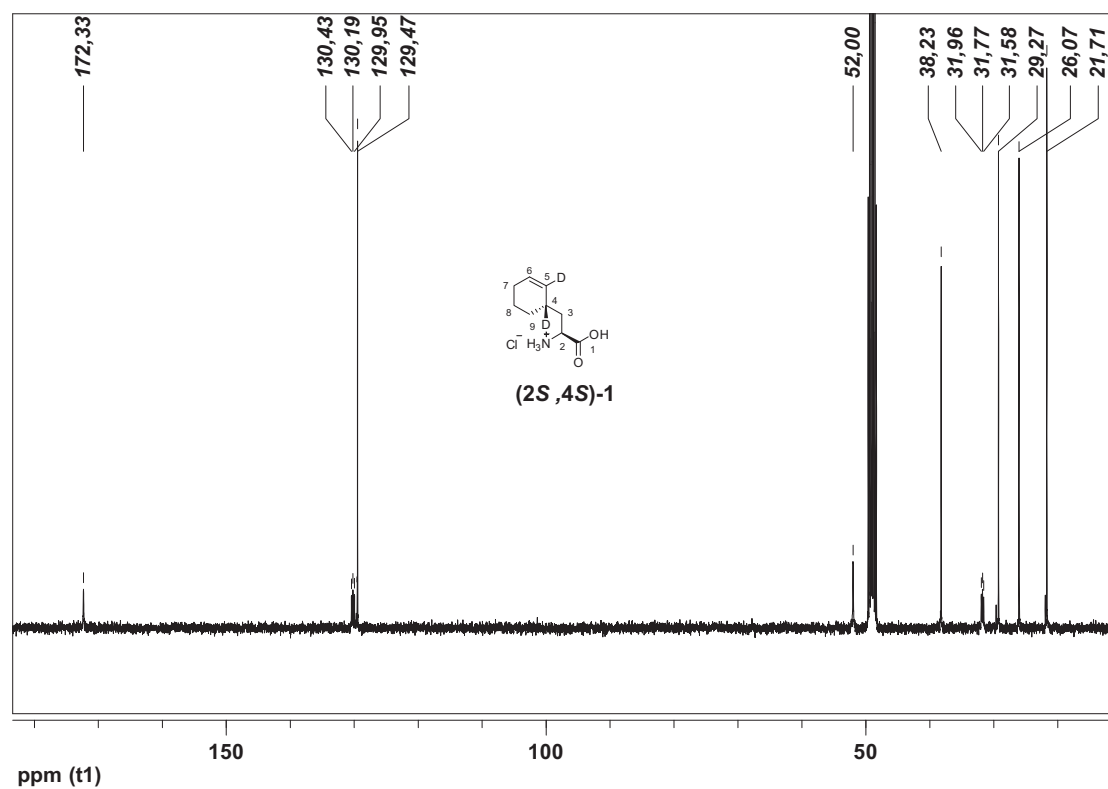
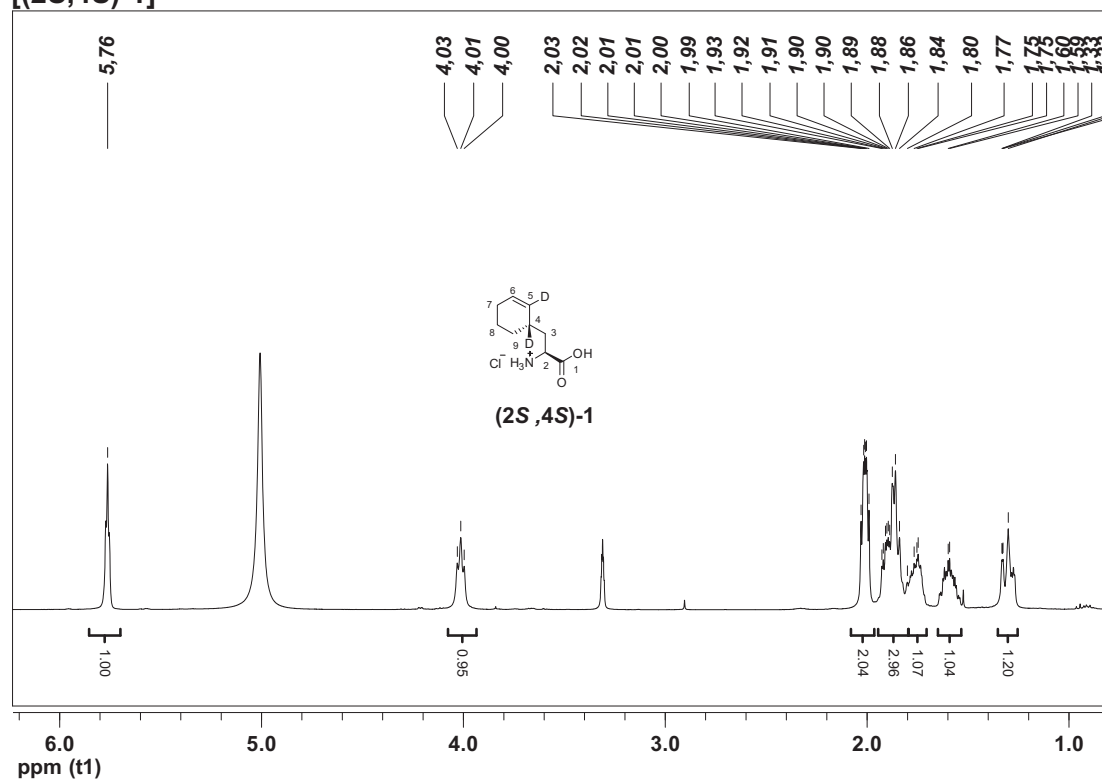
(S)-Methyl 2-(tert-butoxycarbonylamino)-3-[(S)-1,2-dideuteriocyclohex-2-enyl]propanoate [(2S,4S)-8]



(S)-2-(tert-Butoxycarbonylamino)-3-[(S)-1,2-dideuterocyclohex-2-enyl]propanoic acid [(2S,4S)-9]



(S)-2-Amino-3-[(S)-1,2-dideuterocyclohex-2-enyl]propanoic acid hydrochloride [(2S,4S)-1]



IV

Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis

Liujie Huo^{1,5}, Shwan Rachid^{1,2,5}, Marc Stadler^{3,4}, Silke C. Wenzel^{1*} and Rolf Müller^{1*}

¹ Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Saarland, 66123, Germany

² Faculty of Science and Health, University of Koya, Koya, Kurdistan, KO50 1001, Iraq

³ Intermed Discovery GmbH, Dortmund, Nordrhein-Westfalen, 44227, Germany

⁴ Department of Microbial Drugs, Helmholtz Centre for Infection Research, Braunschweig, Niedersachsen, 38124, Germany

⁵ These authors contributed equally to this work.

*e-mail: rom@helmholtz-hzi.de (+49-68130270201) or s.wenzel@mx.uni-saarland.de (+49-68130270204)

Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis

Lijie Huo,^{1,4} Shwan Rachid,^{1,2,4} Marc Stadler,³ Silke C. Wenzel,^{1,*} and Rolf Müller^{1,*}

¹Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Pharmaceutical Biotechnology, Saarland University, Saarbrücken, Saarland 66123, Germany

²Faculty of Science and Health, University of Koya, Koya, Kurdistan KO50 1001, Iraq

³Department of Microbial Drugs, Helmholtz Centre for Infection Research, Braunschweig, Niedersachsen 38124, Germany

⁴These authors contributed equally to this work

*Correspondence: s.wenzel@mx.uni-saarland.de (S.C.W.), rom@helmholtz-hzi.de (R.M.)

<http://dx.doi.org/10.1016/j.chembiol.2012.08.013>

SUMMARY

Bottromycins represent a promising class of antibiotics binding to the therapeutically unexploited A-site of the bacterial ribosome. By inhibiting translation they are active against clinically important pathogens, such as vancomycin-resistant *Enterococci*. Structurally, bottromycins are heavily modified peptides exhibiting various unusual biosynthetic features. To set the stage for compound modification and yield optimization, we identified the biosynthetic gene cluster, used synthetic biotechnology approaches to establish and improve heterologous production, and generated analogs by pathway genetic engineering. We unambiguously identified three radical SAM methyltransferase-encoding genes required for various methylations at unactivated carbons yielding *tert*-butyl valine, methyl-proline, and β -methyl-phenylalanine residues, plus a gene involved in aspartate methyl-ester formation. Evidence for the formation of the exo-thiazole unit and for a macrocyclodehydration mechanism leading to amidine ring formation is provided.

INTRODUCTION

One of the most prominent global public health threats is caused by antibiotic resistance in conjunction with new and reoccurring infectious diseases. In addition, antimicrobial research in pharmaceutical companies is challenged by a severe disproportion between the degree of investment and the expected profit in the course of drug development. Therefore, access to new hit-and-lead structures addressing novel targets and/or representing new chemical scaffolds exhibiting activity against multi-drug-resistant bacteria is of utmost importance (Fischbach and Walsh, 2009; Newman and Cragg, 2012; Li and Vederas, 2009).

Bottromycins were discovered as antibacterial peptides with promising activity against Gram-positive bacteria and mycoplasma from the fermentation broth of *Streptomyces bottropensis* (Waisvisz et al., 1957a, 1957b, 1957c; Waisvisz

and van der Hoeven, 1958; Tanaka et al., 1968). Later on it was shown that their antibacterial ability extends to methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE) (Shimamura et al., 2009). The structure elucidation process involved several chemical revisions (Nakamura et al., 1965a, 1965b, 1965c, 1966, 1967; Takahashi et al., 1976; Schipper, 1983) and ultimately led to the assignment of **1** (Figure 1), which was recently confirmed by total synthesis (Shimamura et al., 2009). Mode of action studies revealed the aminoacyl-tRNA binding site (A site) on the 50S ribosome as the target of bottromycins, ultimately leading to the inhibition of protein synthesis (Otaka and Kaji, 1976, 1981, 1983). As this site is currently not addressed by clinically used antibiotics, no cross-resistance was observed, and bottromycins are regarded as promising leads to be developed as novel anti-infectives, with renewed interest even in medicinal chemistry (Gouda et al., 2012).

Bottromycins represent octapeptides exhibiting an internal tetrapeptide cycle formed via a unique amidine linkage. The compound harbors an exo-thiazole and several unnatural amino acids, which carry methyl-groups at nonactivated carbons, posing additional challenges for total synthetic approaches aimed toward drug development.

As a valid alternative to total synthesis, biosynthetic engineering can be envisaged to improve structure and yield of any microbial natural product eventually resulting in fermentative production (Fischbach and Voigt, 2010). To achieve this goal the underlying principles of compound production need to be understood, and thus identification of the biosynthetic genes is mandatory. In principle, there are two natural ways for the production of highly modified and bioactive peptide scaffolds: biosynthesis can be achieved in a thiotemplated fashion on large multienzymatic systems termed nonribosomal peptide synthetases (NRPS), which were intensively studied in the past decades (Finking and Marahiel, 2004; Schwarzer et al., 2003). Alternatively, and only recently described as more common than originally anticipated, complex peptides can also be biosynthesized starting from simple, ribosomally made precursor peptides undergoing intriguing modification steps (McIntosh et al., 2009; Nolan and Walsh, 2009; Oman and van der Donk, 2010). Nonribosomally synthesized peptides (NRPs) and the latter compounds of ribosomal origin (RPs) differ mainly in the construction of the core-scaffold as RPs are limited in structural diversity by incorporation of the canonical proteinogenic amino acids only. However, they can be extensively posttranslationally

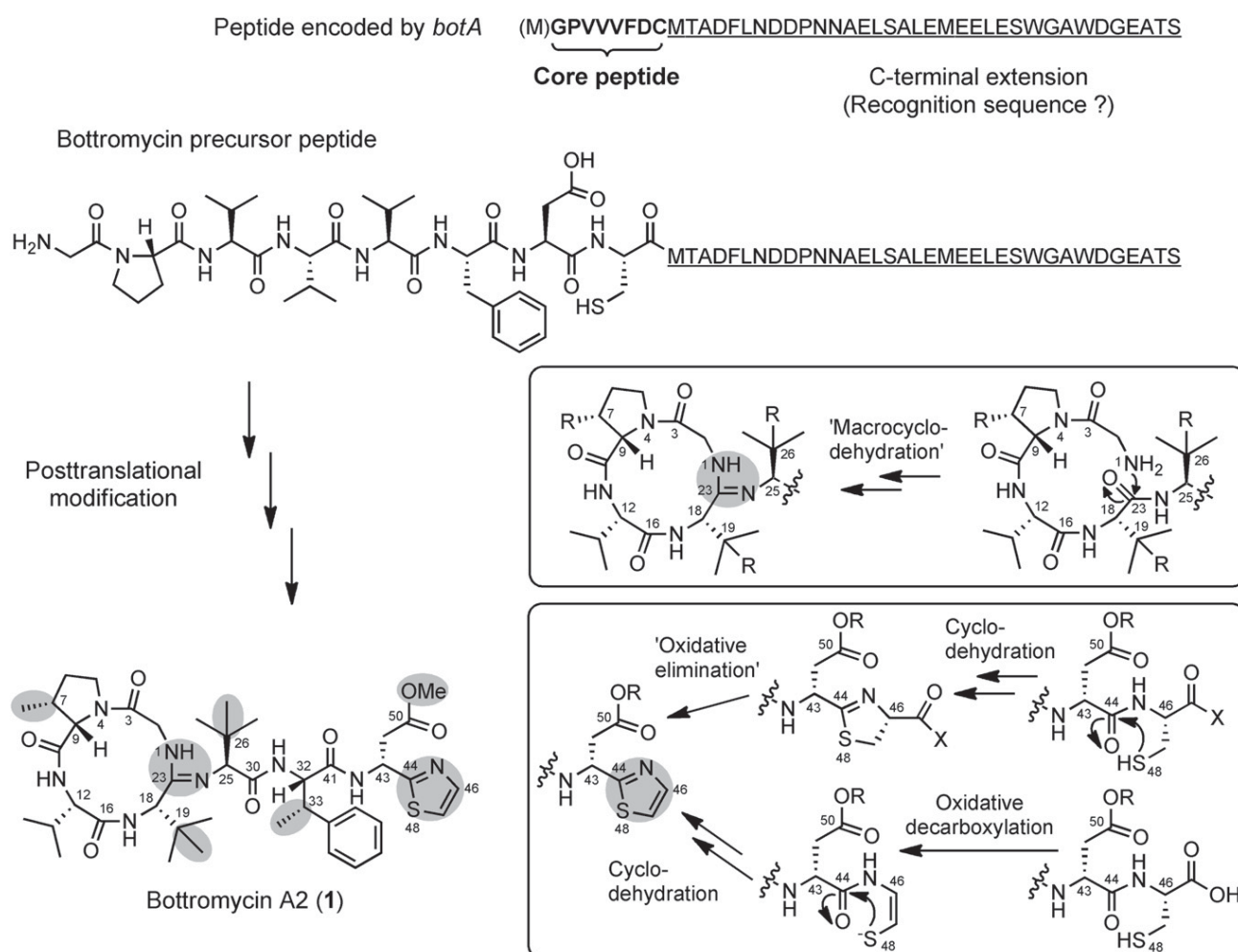


Figure 1. Maturation of the Ribosomal Bottromycin Peptide Scaffold

The bottromycin precursor peptide consists of the core peptide as well as a C-terminal extension (underlined) likely serving as recognition sequence. Maturation of the bottromycin peptide scaffold requires a number of posttranslational modifications (highlighted in gray), the timing of which is currently unknown. These include various methylations and proteolytic digestion as well as "macrocyclodehydration" and thiazole ring formation (proposed mechanisms for the latter two are boxed, for explanations see subsection "In Silico Analysis of the Bottromycin Biosynthetic Pathway"). R, H or Me; X, OH or MTADFLNDDPNNAELSALEMEELESWGAWDGEATS.

modified, which is mostly performed on the precursor peptide state (McIntosh et al., 2009; Oman and van der Donk, 2010). NRPS, in contrast, incorporate during chain assembly numerous "unnatural" amino acids and are thought to undergo less extensive post-NRPS modifications. However and despite their vastly different core biosynthetic routes, RPs and NRPs share many postassembly modifications, which are often required for biological activity. These include heterocyclizations, methylations, and glycosylations, just to name a few (Walsh et al., 2001; McIntosh et al., 2009; Oman and van der Donk, 2010). In NRPS the required enzymatic functionalities are optionally incorporated into the megasynthetases. As discussed above it is difficult to predict from the chemical structure whether modified peptides are based on NRP and RP assembly logic. Therefore, retrobiosynthetic analysis of bottromycin did not yield an assignment of the type of biosynthetic pathway to be expected.

Motivated by the promising activity and the chemical uniqueness of the bottromycin scaffold, we set out to identify the biosynthetic gene locus in the terrestrial actinomycete *Streptomyces* sp. BC16019, which was shown to produce bottromycin A2, B2, and C2 (Lerchen et al., 2006). This work identified a ribosomal locus required for bottromycin biosynthesis and set the stage for synthetic biotechnology approaches toward structure and yield improvement. We heterologously expressed the biosynthetic pathway and optimized initially low bottromycin production titers. Using pathway engineering we were able to assign each of the five unusual methylation reactions to specific genes, which in parallel enabled the production of three, to our knowledge, novel derivatives. In addition, work presented here provides evidence for a unique "macrocyclodehydration" reaction giving rise to the unique amidine linkage found in the bottromycin scaffold.

Table 1. Proteins Involved in Bottromycin Biosynthesis

Protein	Residues (aa)	Annotation	Proposed Function in Bottromycin Biosynthesis
BotRMT1	641	Radical SAM	C-Methylation (of phenylalanine)
BotA	44	Structural gene	Precursor peptide
BotC	434	Cyclodehydratase (YcaO-like family)	Macrocyclodehydration (?)
BotCD	400	Cyclodehydratase (YcaO-like family and docking_ocin superfamily)	Cyclodehydration during thiazole ring formation (?)
BotRMT2	660	Radical SAM	C-Methylation (of valines)
BotH	293	Putative α/β hydrolase	Macrocyclization or thiazole ring formation (?)
BotAH	465	Putative amidohydrolase	Macrocyclization or thiazole ring formation (?)
BotCYP	345	Cytochrome P450 enzyme	Macrocyclization or thiazole ring formation (?)
BotRMT3	681	Radical SAM	C-Methylation (of proline)
BotR	184	Transcriptional regulator	Pathway regulation
BotOMT	279	O-methyl transferase	O-Methylation
BotT	441	Multidrug transporter	Efflux pump, self-resistance
BotP	504	Leucyl-aminopeptidase	Proteolytic cleavage

See also Figure S1 and Table S1.

RESULTS AND DISCUSSION

Identification of the Bottromycin Biosynthetic Gene Cluster

Our initial attempts to identify the bottromycin biosynthetic pathway were based on the assumption that the compound is made via a nonribosomal peptide synthetase (NRPS) assembly line, similar to most other cyclic peptide pathways analyzed to date. Those systems usually exhibit a modular architecture, in which each module incorporates one amino acid building block to the growing peptide chain (Finking and Marahiel, 2004; Schwarzer et al., 2003). Retrobiosynthetic analysis revealed that the bottromycin peptide core almost certainly originates from eight amino acids (Gly-Pro-Val-Val-Val-Phe-Asp-Cys), which led us to predict an octamodular NRPS biosynthetic machinery. In accordance with textbook biosynthetic logic, bottromycin thiazole ring formation from a cysteine residue requires the catalytic activity of a heterocyclization (HC) and an oxidation (Ox) domain (Walsh et al., 2001; Schwarzer et al., 2003; Roy et al., 1999; Walsh and Nolan, 2008), which should therefore be included within the putative bottromycin assembly line. As the identification of heterocycle-forming peptide pathways using homologous and heterologous probes for HC and/or Ox domains is well established (Cheng et al., 2002), we considered this strategy to screen for the bottromycin biosynthetic pathway. A cosmid library of the bottromycin producer *Streptomyces* sp. BC16019 was constructed and analyzed for HC/Ox harboring NRPS pathways by hybridization experiments (S.R., L.H., and R.M., unpublished data). However, these efforts did not lead to the identification of a genetic locus possibly involved in bottromycin biosynthesis. To obtain comprehensive information about all putative secondary metabolite pathways of *S. sp. BC16019* whole-genome shotgun sequence data were generated using the 454 technology. Detailed in silico analysis of the resulting 319 sequence scaffolds revealed no putative NRPS-dependent bottromycin biosynthetic gene cluster, indicating that the compound might rather derive from a ribosomal pathway. In this case, one would expect a small precursor peptide-encoding

gene accompanied by a number of “modifying genes” for crafting the complex and bioactive bottromycin scaffold from a rather simple, linear peptide chain assembled from the canonical 20 proteinogenic amino acids. Indeed, we were able to identify a small open reading frame (orf; *botA*) in the sequence data encoding the bottromycin octapeptide sequence, plus some C-terminal extension. In silico analysis of the flanking regions led to the identification of a number of genes encoding enzymes that are most likely involved in the various posttranslational modification (PTM) steps predicted for the maturation of the bottromycin propeptide (see Figure 1 and Table 1). In order to confirm that the identified chromosomal region encodes the bottromycin biosynthetic pathway, we established a genetic modification procedure for the producer strain to perform targeted gene inactivation experiments. Using the suicide vector pKC1132-*botOMT*-KO, one of the putative-modifying genes, *botOMT*, was inactivated via homologous recombination, resulting in the mutant strain *S. sp. BC16019::pKC1132-botOMT-KO*. High-performance liquid chromatography (HPLC)-mass spectrometry (MS) analysis in comparison to wild-type extracts showed the abolishment of bottromycin production, confirming the involvement of *botOMT* in bottromycin biosynthesis. At this point, polar effects on genes located downstream could not be excluded using the applied insertion strategy. Further experiments (see subsection “Establishment and Optimization of Heterologous Bottromycin Production”) clearly established the *bot* locus as the bottromycin biosynthetic gene cluster.

In Silico Analysis of the Bottromycin Biosynthetic Pathway

The gene locus responsible for bottromycin biosynthesis was identified through in silico analysis of genome data coupled with targeted insertion mutagenesis as described above as well as via heterologous expression of the entire pathway (see subsection “Establishment and Optimization of Heterologous Bottromycin Production”). The pathway consists of (at least) 13 genes (*bot* genes) and spans about ~18 kb (see Figure 2). The *botA* gene encodes a 44-amino acid peptide and is flanked

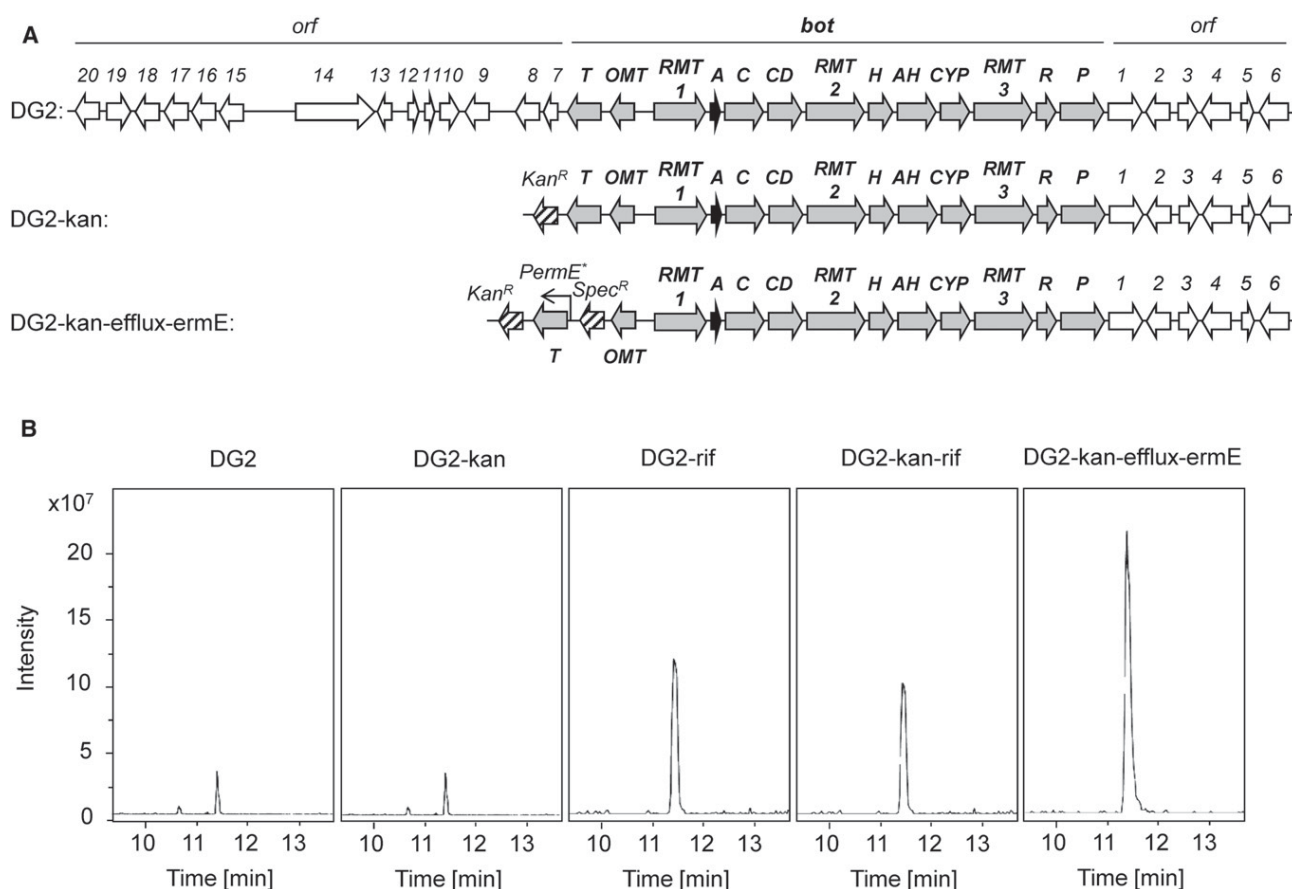


Figure 2. Heterologous Bottromycin Production

(A) Inserts of the three expression constructs (pOJ436 derivatives) used for heterologous bottromycin production. The original cosmid DG2 was modified via Red/ET recombineering to delete the 5'-flanking region (*orf7-orf20*) by insertion of *kan^R* and, in the next step, to insert the *PermE^{*}* promoter upstream of *botT* together with *spec^R*.

(B) Quantification of heterologous bottromycin production by HPLC-MS analysis of the culture extracts from different *S. coelicolor* mutant strains. Sections of extracted ion chromatograms at *m/z* = 823.45 corresponding to the [M+H]⁺ ion of bottromycin A2 (1) are illustrated as representative readout of productivity. The *S. coelicolor* A3(2) host strains contain one of the three expression constructs shown in (A). With *S. coelicolor*::DG2 and DG2-kan, "rifampicin-induced mutagenesis" was performed ("DG2-rif" and "DG2-kan-rif").

by a set of modifying genes required for maturation to an 8-residue highly modified macrocycle (see Figure 1). The BotA precursor peptide is bipartite: it contains an N-terminal core peptide as well as a C-terminal extension (the N-terminal methionine is most likely cleaved off by a methionine aminopeptidase during translation or by the putative peptidase BotP). This structure is unique among precursor peptides from ribosomal pathways described to date; in other precursors, the core peptides are typically flanked by N-terminal leader sequences, and in some cases additional C-terminal extensions are found (Oman and van der Donk, 2010). It is frequently postulated that the N-terminal leader sequences serve as a recognition motif for the PTM enzymes and that they are removed in the last step of the maturation process (Oman and van der Donk, 2010). The absence of an N-terminal leader sequence in BotA indicates that the C-terminal extension—or parts of it—functions as PTM enzyme recognition sequence. Furthermore, as discussed for the N-terminal leaders, some other functions can be attributed to the C-terminal BotA sequence, for example, serve as secre-

tion signal, act as chaperone (in *cis*) and assist with folding, stabilize the precursor against degradation, and/or keep the peptide inactive during biosynthesis inside the cell (Oman and van der Donk, 2010). (Final) Proteolytic cleavage of the C-terminal extension is most likely catalyzed either by the putative α/β -hydrolase BotH or by BotP, which shows homology to leucyl aminopeptidases (Matsui et al., 2006). For some members of this group, it could be shown that they also prefer methionine as N-terminal amino acid, which correlates well with the predicted cleavage site in BotA (see Figure 1) (Herrera-Camacho et al., 2007). The highly modified bottromycin scaffold results from various PTM reactions, and a number of candidate genes involved in these processes have been identified in the *bot* locus—although most of the encoded proteins exhibit only very low sequence identity to proteins found in the common databases. This finding and the associated difficulties in predicting enzyme function from low similarity scores on the protein level are currently a general challenge for the analysis of ribosomal peptide biosynthesis pathways, which complicates the *in silico* analysis and

functional gene annotation. However, a striking feature of the bottromycin scaffold is its extensive methylation pattern, and according to our expectations a number of putative S-adenosylmethionine (SAM)-dependent methyl transferases (MTs) could be identified (BotRMT1, BotRMT2, BotRMT3, and BotOMT). BotOMT shows homologies to O-MTs and was therefore assumed to be involved in O-methylation of the aspartate residue. The remaining three MTs (BotRMT1–BotRMT3) belong to the class of radical SAM-dependent MTs (“radical SAMs”; Frey et al., 2008; Atta et al., 2010) and were expected to catalyze the various C-methylations at the bottromycin scaffold. According to their didomain protein architecture (N-terminal cobalamin binding domain and C-terminal radical SAM domain), BotRMT1–BotRMT3 belong to class B of the recently classified radical SAM MTs (Zhang et al., 2012), which are indeed hypothesized to contain members methylating unactivated sp^3 carbons. This finding is also in accordance with the proposed methylation sites of BotRMT1–BotRMT3 at proline, phenylalanine, and the two valine residues. However, based on the *in silico* analysis, it is clearly impossible to assign the exact function of each protein. We therefore carried out further experiments to elucidate the pathway (see subsection “Functional Studies on the Bottromycin Methylation Pattern: A Suite of Hyperactive Methylases”). The thiazole ring is yet another structural feature of bottromycin, occurring widespread in ribosomally made and nonribosomally assembled peptide scaffolds (Walsh et al., 2001; Schwarzer et al., 2003; Roy et al., 1999; McIntosh et al., 2009; Walsh and Nolan, 2008; Nolan and Walsh, 2009). This heterocycle is typically installed by cyclodehydration starting from a cysteine residue, the thiol of which attacks the carbonyl carbon of the adjacent amino acid to form a thiazoline, which can subsequently undergo oxidation to form the thiazole. For the ribosomal peptide microcin B17, it was shown that a three-protein complex, consisting of a cyclodehydratase, a “docking protein,” and a flavin mononucleotide (FMN)-dependent dehydrogenase, is required for thiazole ring formation (Li et al., 1996). Whereas the function of the cyclodehydratase (cyclodehydration reaction) and the dehydrogenase (thiazoline oxidation) was clearly assigned, the precise function of the “docking protein” remained elusive until recently, and a direct role in regulating cyclodehydratase activity and assembly of an active complex was debated (Melby et al., 2011). However, a recent study demonstrates that the “docking protein,” which belongs to the YcaO-like protein family, itself catalyzes the cyclodehydration under consumption of ATP (Dunbar et al., 2012). This study not only provided first experimental evidence for the requirement of ATP-hydrolysis for cyclodehydration reactions but also showed the dispensability of the subunit formerly annotated as “cyclodehydratase” for the reaction. However, it could be shown that this protein can increase the reaction rate by approximately three orders of magnitude. When analyzing the bottromycin biosynthetic pathway for enzyme candidates possibly involved in thiazole ring formation, two putative “YcaO-like family proteins” (BotC and BotCD) were identified, whereas genes encoding homologs of the “cyclodehydratase” subunit could not be detected (see below). Based on the above findings, we assume that BotC and BotCD autonomously act as cyclodehydratases. Whereas one of these enzymes is most likely involved in thiazoline ring formation, the other candidate might play a role in the unique

“macrocyclodehydration” reaction as discussed below (see Figure 1). We speculate that BotCD is the enzyme involved in bottromycin thiazole formation, because it shows not only homology to YcaO-like family proteins but also to the “docking_ocin superfamily” of proteins (TIGR03604; members of this protein family include enzymes related to SagD; Lee et al., 2008), which includes “docking proteins” involved in heterocyclization reactions in other ribosomal pathways. BotC, on the other hand, would thus be a likely candidate for the unique catalyst performing the macrocyclodehydration discussed below. After cyclization conversion of thiazoline moieties into thiazoles usually occurs via dehydrogenation. However, a putative dehydrogenase activity could not be detected in the *bot* locus, suggesting that thiazole ring formation in bottromycin biosynthesis proceeds via a different mechanism. The absence of a dehydrogenase might be explained by the requirements to form an “exo-thiazole” moiety, which probably results from an oxidative elimination reaction (see Figure 1). One possible scenario is an oxidative decarboxylation after proteolysis of the C-terminal extension sequence. However, we cannot exclude other “elimination hypotheses,” which might rely on further premodifications of the thiazoline ring (e.g., hydroxylation or dehydrogenation). Alternatively, precursor peptide proteolysis and subsequent oxidative decarboxylation of the C-terminal cysteine might occur before cyclodehydration. Similar reactions have been described for the biosynthesis of aminovinyl-cysteine-containing ribosomal peptides, including the generation of reactive “thio-enol” intermediates via oxidative decarboxylation of C-terminal cysteine residues (Sit et al., 2011). The enzymes involved were characterized as homo-oligomeric flavin-containing cysteine decarboxylases, but no putative homolog of this enzyme family was found encoded within the bottromycin biosynthetic gene locus. Apparently, the underlying biochemistry for bottromycin “exo-thiazole” ring formation cannot be predicted based on the available *in silico* data and requires further experimental studies.

Another intriguing modification unique for ribosomal peptide biosynthesis is the formation of the bottromycin macrocycle established between the N-terminal glycine and the second valine residue. Similar to textbook cyclodehydration chemistry, we propose a nucleophilic attack of the glycine amino group to the valine carbonyl carbon, followed by dehydration. This transformation may be seen as “macrocyclodehydration” reaction, resulting in the formation of an amidine group, which is—to the best of our knowledge—unique among all ribosomal peptides biosynthetically characterized so far and also very rare among other natural products. Coformycin (Nakamura et al., 1974) and ectonine (Inbar and Lapidot, 1988), also described as pyrostatin B (Castellanos et al., 2006), are examples of natural products containing this structural feature. Studies on the genetics and biochemistry of ectoine metabolism revealed that the cyclic amidine moiety is formed via a “cyclodehydratase”-like reaction similar to the macrocyclization mechanism we postulate for the bottromycin biosynthesis. The enzyme involved is described as ectoine synthase (EctC) and belongs to the family of carbon-oxygen lyases. It is encoded in the ectoine biosynthetic gene locus, which is found widespread in microorganisms (Pastor et al., 2010), and could also be identified in the bottromycin producer as well as in *Streptomyces coelicolor*. However, no homolog of the ectoine synthase family could be identified in

the bottromycin biosynthetic gene cluster, and we assume that one of the putative cyclodehydratases (BotC and BotCD) might be involved in “macrocyclodehydration.” BotC is the favored candidate enzyme because of the arguments provided above. However, we cannot exclude the involvement of ectoine biosynthetic pathway enzymes in this reaction, as they are encoded in the chromosomes of both the native and the heterologous bottromycin producers. Additionally, there are some *bot* genes with yet unassigned biosynthetic function, like *botH*, *botAH*, and *botCYP*, encoding a putative α/β hydrolase, an amidohydrolase, and a P450 enzyme, respectively, which might be involved in this unusual “macrocyclodehydration” reaction or in the “exo-thiazole” ring formation discussed above. Besides the discussed enzyme candidates involved in the maturation of the bottromycin precursor peptide, the biosynthetic gene cluster encodes a putative transcriptional regulator (BotR) as well as a putative multidrug transporter (BotT). The latter might play an important role for the self-resistance of the bottromycin producer, which is underpinned by results from genetic engineering studies based on a heterologous expression system (see next subsection).

Establishment and Optimization of Heterologous Bottromycin Production

As genetic manipulation of the native bottromycin producer *S. sp.* BC16019 turned out to be rather difficult and was restricted to the established single-cross-over method with very low efficiency, we aimed to establish a heterologous expression system to further characterize and engineer bottromycin biosynthesis. To mobilize the bottromycin pathway for expression in suitable host strains, the entire ~18 kb biosynthetic gene cluster was required to be subcloned on one physical entity. We were able to identify a cosmid harboring the complete target region plus some flanking sequence when screening the gene library of the bottromycin producer with suitable probes (cosmid DG2, see Figure 2A). As the pOJ436 cosmid backbone already contains genetic elements for conjugation (origin of transfer, *oriT*) and site-specific integration into *Streptomyces* chromosomes (phage ϕ C31 integrase gene and attachment site; Bierman et al., 1992), the related host strains *Streptomyces coelicolor* A3(2) and *Streptomyces albus* J1074 were directly transformed with this construct. Transconjugants were verified by PCR and subsequently analyzed for heterologous bottromycin production. Parallel cultivations of host strains with and without the expression construct in comparison to the native bottromycin producer were carried out followed by HPLC-MS analysis of the culture extracts. Successful heterologous bottromycin production could be detected for both of the expression hosts, *S. coelicolor*::DG2 and *S. albus*::DG2, demonstrating that the putative bottromycin pathway was complete and active in the foreign circuits under control of its native regulatory elements. Because the heterologous production yields in both hosts were much lower compared to the native producer (~1 μ g/l in *S. albus* and ~4 μ g/l in *S. coelicolor*; ~100 times lower than in native host), optimization approaches focused on the *S. coelicolor* expression system were initiated. In a preliminary step, the bulky expression construct was simplified by deletion of a ~16 kb insert fragment at the 5' end of the *bot* gene cluster. This modification was performed via Red/ET recombineering

(Zhang et al., 1998, 2000), resulting in the integration of a kanamycin resistance gene to generate cosmid DG2-kan (see Figure 2A). As expected, subsequent heterologous expression in *S. coelicolor* showed that this deletion does not significantly affect the bottromycin production titer (see Figure 2B), and it also confirmed our assumption that the excised genetic region (*orf7*–*orf20*) is not required for bottromycin biosynthesis. Aiming at improvement of bottromycin production titers, we next applied an undirected approach based on the observation that certain drug-resistance mutations in the *rpoB* gene (encoding the RNA polymerase β -subunit), or the *rpsL* gene (encoding the ribosomal S12 protein), may effectively enhance secondary metabolite production in Gram-positive bacteria (Hu et al., 2002; Inaoka et al., 2004; Shima et al., 1996; Hu and Ochi, 2001; Tamehiro et al., 2003). Such mutations can be induced, for example, by challenging the bacteria with rifampicin (mutations in *rpoB*) or streptomycin (mutations in *rpsL*). After plating of the expression strains *S. coelicolor*::DG2 and *S. coelicolor*::DG2-Kan on rifampicin-containing medium, spontaneous isolates resistant to the antibiotic were obtained. Quantitative bottromycin production analysis of clones from each “rifampicin mutagenesis” experiment revealed mutant strains (*S. coelicolor*::DG2-rif and *S. coelicolor*::DG2-Kan-rif) with a considerably higher production titer (about ten times higher compared to the starting clones, see Figure 2B). In parallel to this successful undirected approach, we also considered targeted engineering strategies to improve bottromycin production yields; because of the antibacterial activity of bottromycin, we speculated that the self-resistance of the host strain might be a limiting factor for the production titer. Indeed we could show that the growth of *S. coelicolor* is already heavily impaired at bottromycin concentrations of 2 μ g/ml. In light of this result, the putative efflux pump (BotT) encoded in the bottromycin biosynthetic gene cluster represented a promising target for directed engineering approaches as we speculated it to play a crucial role for self-resistance. To overexpress the *botT* gene, its native promoter was replaced against the strong *Perme** promoter (Bibb et al., 1994) by insertion of a spectinomycin resistance cassette via Red/ET recombineering (see Figure 2A). *S. coelicolor* A3(2) wild-type was subsequently transformed with the resulting construct DG2-kan-efflux-ermE, and bottromycin production was quantified in comparison to *S. coelicolor*::DG2(-kan) and the corresponding “rif mutants.” The HPLC-MS analysis revealed that, once more, bottromycin production could be increased: the generated *S. coelicolor*::DG2-kan-efflux-ermE expression host shows a 20 times higher production titer compared to *S. coelicolor*::DG2(-kan) and an around two times higher production titer compared to *S. coelicolor*::DG2(-kan)-rif. The generated data on the successful enhancement of heterologous bottromycin production for future experiments suggest combining both strategies: the targeted approach (*Perme** promoter insertion) with the undirected approach (“rifampicin mutagenesis”). This could potentially lead to a further increase of bottromycin production yields.

Functional Studies on the Bottromycin Methylation Pattern: A Suite of Hyperactive Methylases

A remarkable characteristic of the bottromycin scaffold is its versatile methylation pattern, which can be attributed to a set

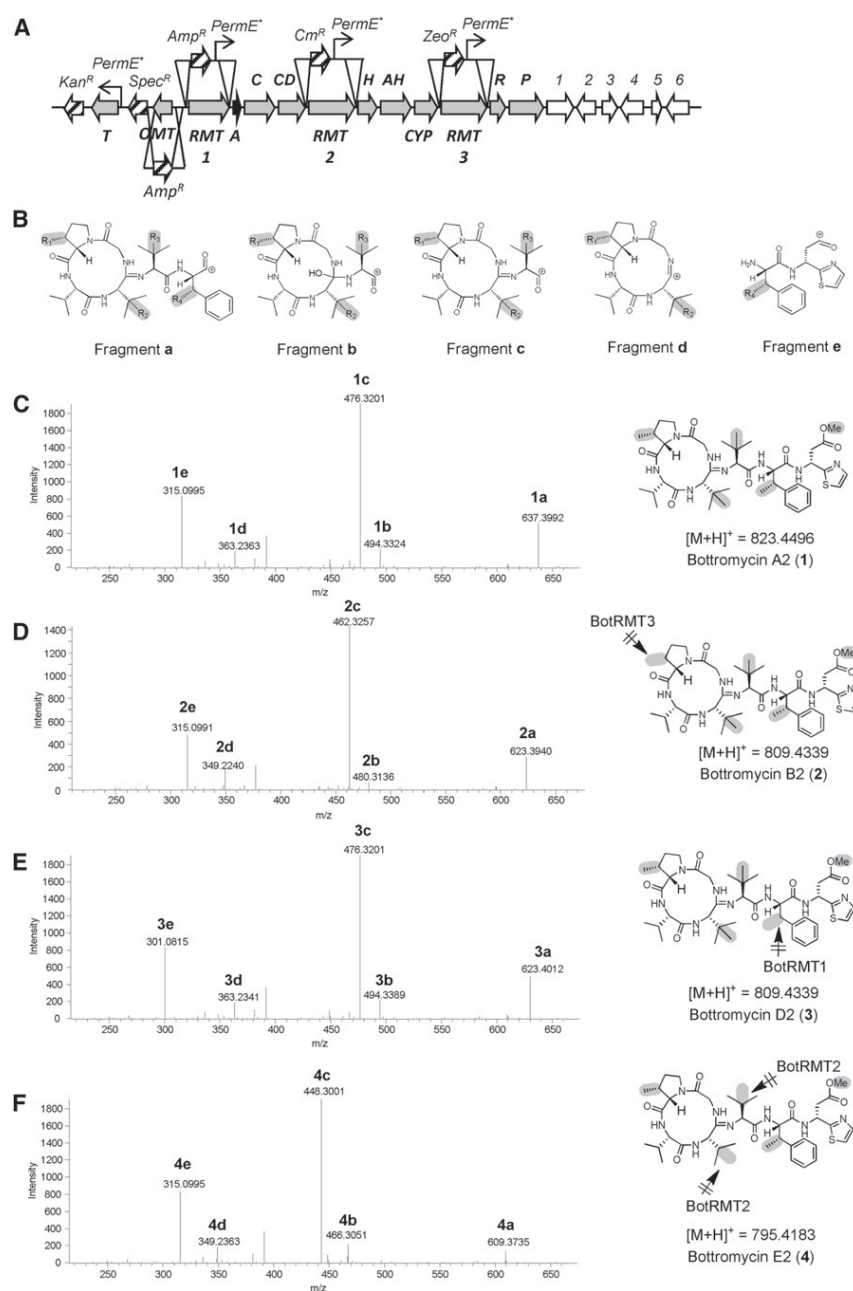


Figure 3. Engineering and Characterization of the Bottromycin "Methylation Machinery"

(A) Modification of the bottromycin biosynthetic gene cluster. The four methyl transferase encoding genes *botRMT1*, *botRMT2*, *botRMT3*, and *botOMT* were deleted from the expression construct DG2-kan-efflux-ermE in four orthogonal Red/ET recombineering steps by insertion of appropriate selection marker cassettes.

(B) Assigned fragment ions used for the comparison of the different bottromycin scaffolds detected in (C)–(F). R₁, R₂, R₃, and R₄ = H or Me.

(C)–(F) High-resolution MS/MS fragmentation analysis of bottromycin A2 reference substance (C) and of bottromycin scaffolds detected in extracts from the *S. coelicolor* A3(2) host strain expressing modified versions of the bottromycin biosynthetic gene cluster: *botRMT3* deletion (D), *botRMT1* deletion (E), and *botRMT2* deletion (F). For the fragment spectrum from the *botOMT* deletion experiment and a detailed comparison of calculated and detected high-resolution masses see Table S2 and Figure S1.

(Zhang et al., 1998, 2000). The respective MT-encoding genes were replaced by an antibiotic resistance gene (*amp^R*, *cm^R*, or *zeo^R*), and in case of *botRMT1*–*botRMT3* the *PermE** promoter was simultaneously introduced in front of the downstream genes to ensure proper transcription (see Figure 3A). After *S. coelicolor* A3(2) was transformed with the modified constructs, the heterologous production was studied by intensive HPLC-hrMS/MS analysis on the culture extracts to elucidate the methylation pattern of the produced bottromycin derivatives. By comparing the fragmentation pattern of the detected bottromycin scaffolds, the function of all four MTs could be unambiguously assigned (see Figure 3, Figure S1 [available online], and Table S2). Three bottromycin derivatives can be usually detected in extracts of the native bottromycin producer, which differ in the methylation pattern of the proline

of four putative MTs encoded in the *bot* gene cluster. Based on the in silico analysis, it was quite obvious that BotOMT might catalyze O-methylation of the aspartate residue, whereas the action sites of the three radical SAMs (BotRMT1–BotRMT3) could not be assigned (C-methylation at C-7, C-19, C-26, and/or C-33, see Figure 1). The established heterologous expression system provided an excellent opportunity to accomplish functional studies on the four MTs by performing targeted gene deletion experiments, followed by the heterologous expression of the modified pathways to elucidate the resulting bottromycin methylation pattern. In order to achieve this goal, four orthogonal gene deletions were performed on the expression construct DG2-kan-efflux-ermE using the Red/ET recombineering technology

residue: the monomethylated main product bottromycin A2 (1) plus unmethylated (bottromycin B2 (2)) and bimethylated (bottromycin C2) minor components (Nakamura et al., 1967). Deletion of *botRMT3* resulted in the (exclusive) production of a bottromycin scaffold exhibiting a molecular mass as well as a fragmentation pattern identical to authentic bottromycin B2 (2) reference substance. Consequently, we can draw the conclusion that *botRMT3* catalyzes radical C-methylation of the proline residue. After deleting *botRMT1*, again a bottromycin derivative with a molecular mass identical to bottromycin B2 could be detected, but numerous differences in the fragmentation pattern are obvious. The mass of all "phenylalanine-containing" fragments (3a and 3e) is reduced by 14 Dalton (Da) in comparison to the

corresponding fragments of bottromycin A2 and B2 (**1a** and **1e/2e**, **2a** is also reduced by 14 Da as it contains the nonmethylated proline residue). On the other hand, “non-phenylalanine-containing” fragments (**3b–3d**) show the same mass as corresponding fragments from bottromycin A2 (**1b–1d**), strongly indicating that the three other relevant positions (proline and two valine residues) are still methylated. Based on this, BotRMT1 was assigned to catalyze radical C-methylation of the phenylalanine residue. The deletion of the third radical SAM-encoding gene (*botRMT2*) resulted in the production of a bottromycin scaffold with a molecular mass 28 Da less than that of bottromycin A2 (**1**) and 14 Da less than that of bottromycin B2 (**2**), indicating the presence of two unmethylated positions. The mass of fragment **4e** verifies the presence of the methyl group at the phenylalanine residue, whereas fragments **4b–4d** indicate the “loss” of methyl groups at two valine sites. This suggests that BotRMT2 is involved in radical C-methylation of two of the three valine residues and therefore acts twice during bottromycin biosynthesis. A similar analysis was carried out for the *botOMT* deletion experiment (not shown in Figure 3) and provided experimental evidence for the assigned function: O-methylation of the aspartate residue. The produced “free acid” bottromycin scaffold cannot be detected in strains expressing the complete pathway, and the same holds true for compounds **3** and **4** identified after *botRMT1* and *botRMT2* deletion. Thus, in addition to the findings on the functional role of the bottromycin MTs, these experiments also led to the production of three, to our knowledge, novel bottromycin derivatives. It will be very interesting to analyze the bioactivity of the modified scaffolds, which is currently impeded by the very low production yields that also prevented us from characterizing all structures via NMR experiments. The reason(s) for the heavily reduced production yields in the deletion experiments currently remain elusive. Reduced conversion rates of desmethyl precursors by subsequent biosynthetic enzymes or breakdown of the immunity system in the deletion mutants are two of numerous possible explanations. As no cross-complementation reactions could be observed in the mutants analyzed, we conclude that the three radical SAMs act very specifically when crafting the bottromycin scaffold.

SIGNIFICANCE

Bottromycins represent an intriguing family of natural products exhibiting promising activity against methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococci*. As protein synthesis inhibitors, they are regarded as attractive lead compounds for the development of novel antibiotics because bottromycins target the unique aminoacyl-tRNA binding site of the 50S ribosomal subunit. Here, we discovered the bottromycin biosynthetic pathway from *Streptomyces* sp. BC16019, providing evidence that the heavily modified peptide scaffold is of ribosomal origin. Quite unusually, the bottromycin precursor peptide lacks an N-terminal leader sequence, which is regarded as a common structural feature of ribosomally made natural product peptide precursors. Our data also highlight the bottromycin biosynthetic pathway as an outstanding example for a variety of posttranslational modification reactions, which in general are typical for ribosomal peptide biosyntheses.

Maturation of the bottromycin precursor peptide includes macrocyclization via a unique amidine-forming “macrocycle-dehydration,” a second cyclodehydration most likely coupled to oxidative elimination building a thiazole ring, proteolytic cleavage, and various unusual methylation reactions. By use of synthetic biotechnology approaches, we functionally elucidated the methylation steps involving the action of an O-methyl transferase and three radical S-adenosylmethionine (SAM)-dependent methyl transferases required for alkylation of nonactivated carbons. The latter enzyme class is currently not well characterized and thus of special interest. These experiments also enabled the production of three bottromycin derivatives based on a heterologous expression system. Initial low yields were addressed by application of targeted and undirected approaches for production optimization. These resulted in significant increases of bottromycin production titers in the heterologous host now providing an important basis for advanced structure and yield improvements. In a broader sense, the discovery and characterization of the ribosomal bottromycin pathway revealed exciting insights into the biosynthesis of ribosomally made peptides—a biotechnologically hardly exploited class of natural products of increasing importance for drug discovery.

EXPERIMENTAL PROCEDURES

Strains and Culture Conditions

Streptomyces sp. BC16019, originally obtained from a soil sample collected in Germany (Lerchen et al., 2006), is maintained in the culture collection of InterMed Discovery GmbH, Dortmund, Germany, under liquid nitrogen. For the present study, it was revived and cultivated at 30°C and 180 rpm in liquid yeast and malt extract with glucose (YMG) medium (Stadler et al., 2007). *Streptomyces coelicolor* A3(2) (Bentley et al., 2002) and *Streptomyces albus* J1074 (Chater and Wilde, 1976) were grown at 30°C and 180 rpm in liquid Tryptone soya broth (TSB) medium (Kieser et al., 2000), and their mutants were cultivated in liquid TSB medium containing 60 µg/ml apramycin at 30°C and 180 rpm. *S. sp.* BC16019, *S. coelicolor* A3(2), and *Streptomyces albus* J1074 were grown on solid MS agar plates (Kieser et al., 2000). In addition to *Escherichia coli* HS996 (Invitrogen, Carlsbad, CA, USA) used as general host for molecular biology experiments, *E. coli* GB05-red (Fu et al., 2012) was employed in Red/ET recombineering experiments (Zhang et al., 2000), and *E. coli* ET12567/pUZ8002 (Kieser et al., 2000) was used for conjugation experiments. All *E. coli* strains were grown in Luria Broth (LB) medium at 37°C. If necessary, the medium was supplemented with the following antibiotics for screening appropriate mutants: chloramphenicol 30 µg/ml, spectinomycin 100 µg/ml, ampicillin 100 µg/ml, zeocin 25 µg/ml, kanamycin sulfate 50 µg/ml, and apramycin 60 µg/ml.

Sequencing and In Silico Analysis of the *Streptomyces* sp. BC16019 Genome

Total genomic DNA of *S. sp.* BC16019 was isolated in accordance with the Kirby mix procedure as described previously (Kieser et al., 2000) and sequenced by applying the 454 technology (Rothberg and Leamon, 2008), including 2 kb paired-end reads. The resulting sequence data (2,228 contigs and 319 scaffolds) were subsequently screened for putative open reading frames (*orfs*) encoding peptides that contain the amino acid sequence of the bottromycin scaffold (GPVVVFDC). For this, the EMBOSS “getorf” tool was applied to identify sequences of ORFs (defined as region between two stop codons), in accordance with the standard amino acid code. The output sequence list of all possible, translated *orfs* was subsequently screened for the GPVVVFDC-sequence using standard text editor software. A small *orf* (later named *botA*) was identified, encoding this sequence plus some C-terminal extension (see Figure 1).

Sequence Analysis of the Bottromycin Biosynthetic Gene Cluster

Based on the data from the library screening (see Supplemental Experimental Procedures), cosmid DG2 was chosen for further analysis, including whole cosmid shotgun sequencing as described previously (Silakowski et al., 1999). The obtained sequence of the ~41 kb chromosomal fragment (insert of cosmid DG2) was subsequently analyzed using bioinformatic tools. Prediction of open reading frames was performed with FramePlot 4.0 (<http://nocardia.nih.gov/jp4/>), and functional annotation was based on BlastP (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and Pfam (<http://pfam.sanger.ac.uk/search>) searches. The Geneious 5.6.2 software packages (Biomatters, New Zealand) was used for the analysis and annotation of DNA and protein sequences. In addition to the putative ~18 kb bottromycin biosynthetic gene cluster (*bot* genes; see Figure 2 and Table 1) a number of flanking *orf*s (*orf1*–*orf20*; see Figure 2 and Table S1) were identified.

ACCESSION NUMBERS

The GenBank accession number for the bottromycin biosynthetic gene cluster sequence from *Streptomyces* sp. BC16019 reported in this paper is JX235926.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two tables, one figure, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.chembiol.2012.08.013>.

ACKNOWLEDGMENTS

The authors would like to thank Eva Luxenburger and Thomas Hoffmann (Saarland University) for skillful help with various analytical technologies, Andrea Rademacher (InterMed Discovery GmbH) for expert technical assistance, and Jens Bitzer for reference compounds. This work was generously supported by a grant from the Bundesministerium für Bildung und Forschung (FKZ 0315385B) to Saarland University and InterMed Discovery GmbH.

Received: May 25, 2012

Revised: August 3, 2012

Accepted: August 9, 2012

Published online: September 27, 2012

REFERENCES

- Atta, M., Mulliez, E., Arragain, S., Forouhar, F., Hunt, J.F., and Fontecave, M. (2010). S-Adenosylmethionine-dependent radical-based modification of biological macromolecules. *Curr. Opin. Struct. Biol.* 20, 684–692.
- Bentley, S.D., Chater, K.F., Cerdeño-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147.
- Bibb, M.J., White, J., Ward, J.M., and Janssen, G.R. (1994). The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Mol. Microbiol.* 14, 533–545.
- Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., and Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* 116, 43–49.
- Castellanos, L., Duque, C., Zea, S., Espada, A., Rodríguez, J., and Jiménez, C. (2006). Isolation and synthesis of (–)-(5S)-2-imino-1-methylpyrrolidine-5-carboxylic acid from *Cliona tenuis*: structure revision of pyrostatins. *Org. Lett.* 8, 4967–4970.
- Chater, K.F., and Wilde, L.C. (1976). Restriction of a bacteriophage of *Streptomyces albus* G involving *endonuclease Sall*. *J. Bacteriol.* 128, 644–650.
- Cheng, Y.Q., Tang, G.L., and Shen, B. (2002). Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in *Streptomyces atroolivaceus* S-140. *J. Bacteriol.* 184, 7013–7024.
- Dunbar, K.L., Melby, J.O., and Mitchell, D.A. (2012). YcaO domains use ATP to activate amide backbones during peptide cyclodehydrations. *Nat. Chem. Biol.* 8, 569–575.
- Finking, R., and Marahiel, M.A. (2004). Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol.* 58, 453–488.
- Fischbach, M.A., and Walsh, C.T. (2009). Antibiotics for emerging pathogens. *Science* 325, 1089–1093.
- Fischbach, M., and Voigt, C.A. (2010). Prokaryotic gene clusters: a rich toolbox for synthetic biology. *Biotechnol. J.* 5, 1277–1296.
- Frey, P.A., Hegeman, A.D., and Ruzicka, F.J. (2008). The Radical SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol.* 43, 63–88.
- Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P.M., Plaza, A., Xia, L., Müller, R., Stewart, A.F., and Zhang, Y. (2012). Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bio-prospecting. *Nat. Biotechnol.* 30, 440–446.
- Gouda, H., Kobayashi, Y., Yamada, T., Ideguchi, T., Sugawara, A., Hirose, T., Omura, S., Sunazuka, T., and Hirono, S. (2012). Three-dimensional solution structure of bottromycin A2: a potent antibiotic active against methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococci*. *Chem. Pharm. Bull. (Tokyo)* 60, 169–171.
- Herrera-Camacho, I., Rosas-Murrieta, N.H., Rojo-Domínguez, A., Millán, L., Reyes-Leyva, J., Santos-López, G., and Suárez-Rendueles, P. (2007). Biochemical characterization and structural prediction of a novel cytosolic leucyl aminopeptidase of the M17 family from *Schizosaccharomyces pombe*. *FEBS J.* 274, 6228–6240.
- Hu, H., and Ochi, K. (2001). Novel approach for improving the productivity of antibiotic-producing strains by inducing combined resistant mutations. *Appl. Environ. Microbiol.* 67, 1885–1892.
- Hu, H., Zhang, Q., and Ochi, K. (2002). Activation of antibiotic biosynthesis by specified mutations in the *rpoB* gene (encoding the RNA polymerase beta subunit) of *Streptomyces lividans*. *J. Bacteriol.* 184, 3984–3991.
- Inaoka, T., Takahashi, K., Yada, H., Yoshida, M., and Ochi, K. (2004). RNA polymerase mutation activates the production of a dormant antibiotic 3,3'-neotrehalosadiamine via an autoinduction mechanism in *Bacillus subtilis*. *J. Biol. Chem.* 279, 3885–3892.
- Inbar, L., and Lapidot, A. (1988). The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ¹³C- and ¹⁵N-labeled L-glutamate and ¹³C and ¹⁵N NMR spectroscopy. *J. Biol. Chem.* 263, 16014–16022.
- Kieser, T., Bibb, M., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). *Practical Streptomyces Genetics*. 613 (Norwich, England: The John Innes Foundation).
- Lee, S.W., Mitchell, D.A., Markley, A.L., Hensler, M.E., Gonzalez, D., Wohlrab, A., Dorrestein, P.C., Nizet, V., and Dixon, J.E. (2008). Discovery of a widely distributed toxin biosynthetic gene cluster. *Proc. Natl. Acad. Sci. USA* 105, 5879–5884. Published online 2008.
- Lerchen, H.-G., Schiffer, G., Brötz-Österheld, H., Mayer-Bartschmid, A., Eckermann, S., Freiberg, C., Endermann, R., Schuhmann, J., Meier, H., Svenstrup, N., et al. (2006). Cyclic iminopeptide derivatives (Germany) [WO 2006/103010].
- Li, J.W.H., and Vederas, J.C. (2009). Drug discovery and natural products: end of an era or an endless frontier? *Science* 325, 161–165.
- Li, Y.M., Milne, J.C., Madison, L.L., Kolter, R., and Walsh, C.T. (1996). From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* 274, 1188–1193.
- Matsui, M., Fowler, J.H., and Walling, L.L. (2006). Leucine aminopeptidases: diversity in structure and function. *Biol. Chem.* 387, 1535–1544.
- McIntosh, J.A., Donia, M.S., and Schmidt, E.W. (2009). Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. *Nat. Prod. Rep.* 26, 537–559.
- Melby, J.O., Nard, N.J., and Mitchell, D.A. (2011). Thiazole/oxazole-modified microcins: complex natural products from ribosomal templates. *Curr. Opin. Chem. Biol.* 15, 369–378.

- Nakamura, H., Koyama, G., Iitaka, Y., Ono, M., and Yagiawa, N. (1974). Structure of coformycin, an unusual nucleoside of microbial origin. *J. Am. Chem. Soc.* **96**, 4327–4328.
- Nakamura, S., Chikaike, T., Yonehara, H., and Umezawa, H. (1965a). Structures of bottromycin A and B. *J. Antibiot. (Tokyo)* **18**, 60–61.
- Nakamura, S., Chikaike, T., Karasawa, K., Tanaka, N., Yonehara, H., and Umezawa, H. (1965b). Isolation and characterisation of bottromycin A and B. *J. Antibiot. (Tokyo)* **18**, 47–52.
- Nakamura, S., Chikaike, T., Yonehara, H., and Umezawa, H. (1965c). Isolation, characterization and structural elucidation of new amino acids from bottromycin A. *Chem. Pharm. Bull. (Tokyo)* **13**, 599–602.
- Nakamura, S., Tanaka, N., and Umezawa, H. (1966). Bottromycin A₁, A₂ and their structures. *J. Antibiot. (Tokyo)* **19**, 10–12.
- Nakamura, S., Yajima, T., Lin, Y.C., and Umezawa, H. (1967). Isolation and characterization of bottromycins A₂, B₂, C₂. *J. Antibiot. (Tokyo)* **20**, 1–5.
- Newman, D.J., and Cragg, G.M. (2012). Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **75**, 311–335.
- Nolan, E.M., and Walsh, C.T. (2009). How nature morphs peptide scaffolds into antibiotics. *ChemBioChem* **10**, 34–53.
- Oman, T.J., and van der Donk, W.A. (2010). Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **6**, 9–18.
- Otake, T., and Kaji, A. (1976). Mode of action of bottromycin A₂. Release of aminoacyl- or peptidyl-tRNA from ribosomes. *J. Biol. Chem.* **251**, 2299–2306.
- Otake, T., and Kaji, A. (1981). Mode of action of bottromycin A₂: effect on peptide bond formation. *FEBS Lett.* **123**, 173–176.
- Otake, T., and Kaji, A. (1983). Mode of action of bottromycin A₂: effect of bottromycin A₂ on polysomes. *FEBS Lett.* **153**, 53–59.
- Pastor, J.M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L.N., Iborra, J.L., Vargas, C., Nieto, J.J., and Canovas, M. (2010). Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol. Adv.* **28**, 782–801.
- Rothberg, J.M., and Leamon, J.H. (2008). The development and impact of 454 sequencing. *Nat. Biotechnol.* **26**, 1117–1124.
- Roy, R.S., Gehring, A.M., Milne, J.C., Belshaw, P.J., and Walsh, C.T. (1999). Thiazole and oxazole peptides: biosynthesis and molecular machinery. *Nat. Prod. Rep.* **16**, 249–263.
- Schipper, D. (1983). The revised structure of bottromycin A₂. *J. Antibiot. (Tokyo)* **36**, 1076–1077.
- Schwarzer, D., Finking, R., and Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* **20**, 275–287.
- Shima, J., Hesketh, A., Okamoto, S., Kawamoto, S., and Ochi, K. (1996). Induction of actinorhodin production by rpsL (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **178**, 7276–7284.
- Shimamura, H., Gouda, H., Nagai, K., Hirose, T., Ichioka, M., Furuya, Y., Kobayashi, Y., Hirono, S., Sunazuka, T., and Omura, S. (2009). Structure determination and total synthesis of bottromycin A₂: a potent antibiotic against MRSA and VRE. *Angew. Chem. Int. Ed. Engl.* **48**, 914–917.
- Silakowski, B., Schairer, H.U., Ehret, H., Kunze, B., Weinig, S., Nordsiek, G., Brandt, P., Blöcker, H., Höfle, G., Beyer, S., and Müller, R. (1999). New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. *J. Biol. Chem.* **274**, 37391–37399.
- Sit, C.S., Yoganathan, S., and Vederas, J.C. (2011). Biosynthesis of amino-vinyl-cysteine-containing peptides and its application in the production of potential drug candidates. *Acc. Chem. Res.* **44**, 261–268.
- Stadler, M., Bitzer, J., Mayer-Bartschmid, A., Müller, H., Benet-Buchholz, J., Gantner, F., Tichy, H.V., Reinemer, P., and Bacon, K.B. (2007). Cinnabaramides A-G: analogues of lactacystin and salinosporamide from a terrestrial streptomycete. *J. Nat. Prod.* **70**, 246–252.
- Takahashi, Y., Naganawa, H., Takita, T., Umezawa, H., and Nakamura, S. (1976). The revised structure of bottromycin A₂. *J. Antibiot. (Tokyo)* **29**, 1120–1123.
- Tamehiro, N., Hosaka, T., Xu, J., Hu, H.F., Otake, N., and Ochi, K. (2003). Innovative approach for improvement of an antibiotic-overproducing industrial strain of *Streptomyces albus*. *Appl. Environ. Microbiol.* **69**, 6412–6417.
- Tanaka, N., Nishimura, T., Nakamura, S., and Umezawa, H. (1968). Activity of bottromycin against *Mycoplasma gallisepticum*. *J. Antibiot. (Tokyo)* **21**, 75–76.
- Waisvisz, J.M., and van der Hoeven, M.G. (1958). The chemistry and partial structure of bottromycin. IV. *J. Am. Chem. Soc.* **80**, 383–385.
- Waisvisz, J.M., van der Hoeven, M.G., and te Nijnhuis, B. (1957a). The structure of the sulfur-containing moiety of bottromycin. *J. Am. Chem. Soc.* **79**, 4524–4527.
- Waisvisz, J.M., van der Hoeven, M.G., Hölscher, J.F., and te Nijnhuis, B. (1957b). Bottromycin. II. Preliminary degradation studies. *J. Am. Chem. Soc.* **79**, 4522–4524.
- Waisvisz, J.M., van der Hoeven, M.G., van Peppen, J., and Zwennis, W.C.M. (1957c). Bottromycin. I. A new sulfur-containing antibiotic. *J. Am. Chem. Soc.* **79**, 4520–4521.
- Walsh, C.T., and Nolan, E.M. (2008). Morphing peptide backbones into heterocycles. *Proc. Natl. Acad. Sci. USA* **105**, 5655–5656.
- Walsh, C.T., Chen, H.W., Keating, T.A., Hubbard, B.K., Losey, H.C., Luo, L.S., Marshall, C.G., Miller, D.A., and Patel, H.M. (2001). Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* **5**, 525–534.
- Zhang, Y., Buchholz, F., Muylers, J.P., and Stewart, A.F. (1998). A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123–128.
- Zhang, Y., Muylers, J.P.P., Testa, G., and Stewart, A.F. (2000). DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* **18**, 1314–1317.
- Zhang, Q., van der Donk, W.A., and Liu, W. (2012). Radical-mediated enzymatic methylation: a tale of two SAMS. *Acc. Chem. Res.* **45**, 555–564.

Supplemental Information

Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis

Liujie Huo, Shwan Rachid, Marc Stadler, Silke C. Wenzel, and Rolf Müller

Inventory of Supplemental Information

SUPPLEMENTAL DATA

Table S1, related to Table 1. Orfs encoded in the flanking regions of the proposed *bot* gene cluster.

Figure S1, related to Figure 3: Proposed chemical structures of detected fragments by high-resolution MS-MS analysis.

Table S2, related to Figure 3: Data for MS fragmentation analysis of bottromycin derivatives.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

SUPPLEMENTAL REFERENCES

SUPPLEMENTAL DATA

Table S1, related to Table 1. Orfs encoded in the flanking regions of the proposed *bot* gene cluster

Protein	Residues (aa)	Annotation ^[a]
Orf1	417	Putative sugar phosphotransferase system (PTS) IIC domain
Orf2	250	Putative beta-lactamase B superfamily domain
Orf3	174	Putative PemK-like protein
Orf4	307	Putative CBS-like (Cystathionine beta-synthase (CBS) and Cysteine synthase) protein
Orf5	92	Putative ThiS (thiaminS) family-like protein
Orf6	368	Putative binding domain of LacI transcriptional regulator family
Orf7	77	Putative sugar phosphotransferase system (PTS) IIB domain
Orf8	245	Putative ribonuclease PH-like 3'-5' exoribonuclease
Orf9	200	Putative deoxyribonucleotide triphosphate pyrophosphatase (HAM1-like protein)
Orf10	268	Putative HNH endonuclease domain-containing protein
Orf11	90	Hypothetical protein
Orf12	181	Putative peroxiredoxin (PRX) family-like protein
Orf13	113	Putative chaperonin 10kd subunit-like protein
Orf14	807	Putative penicillin binding proteins (PBPs)
Orf15	244	Putative ABC-2 family transporter
Orf16	293	Putative ABC-2 family transporter
Orf17	300	Putative ATP-binding domain of ABC family transporter
Orf18	284	Hypothetical protein
Orf19	503	Putative GDSL-link lipase/acylhydrolase
Orf20	403	Hypothetical protein

[a] Annotations was based on Protein Blast (Database: nonredundant protein sequences (nr)) and Pfam family search with E-value 1.0.

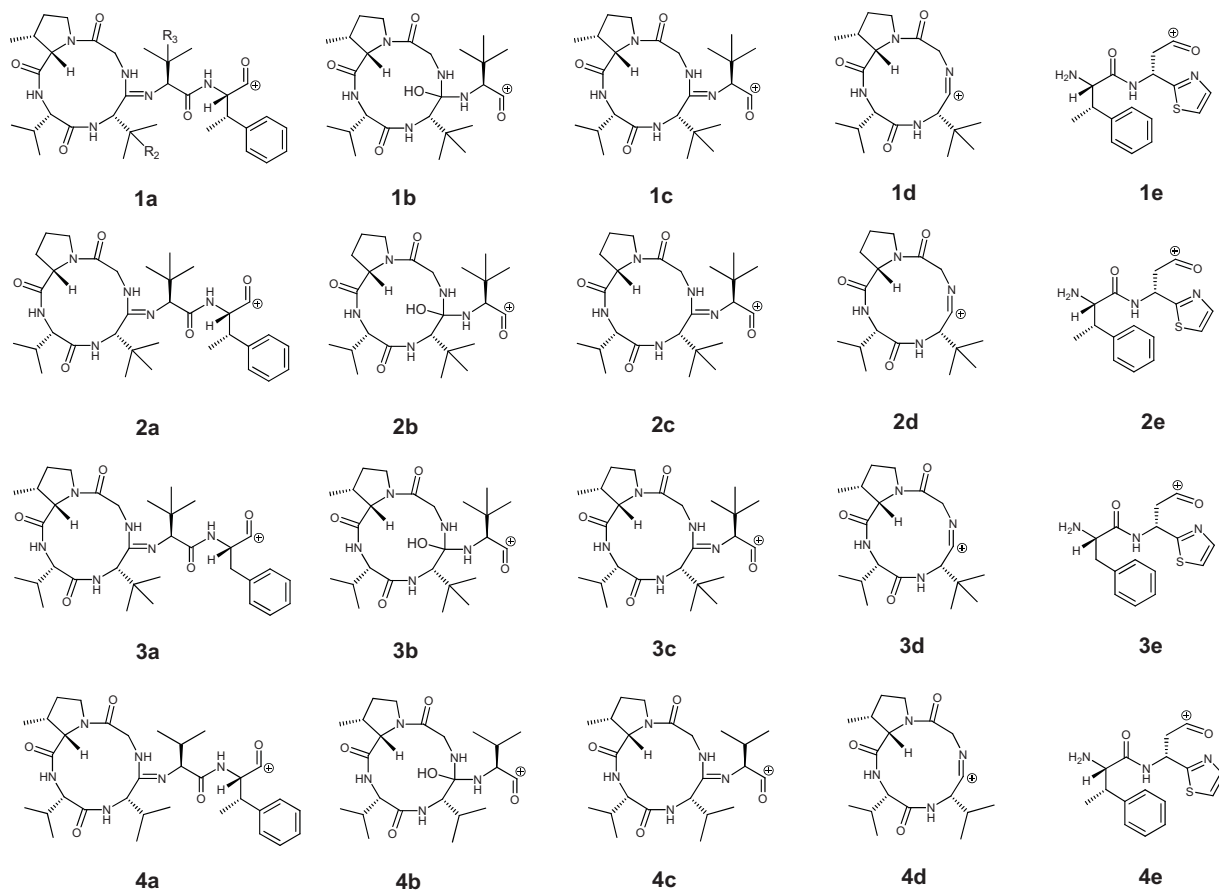


Figure S1, related to Figure 3. Proposed chemical structures of detected fragments by high-resolution MS-MS analysis.

Table S2, related to Figure 3. Data for MS fragmentation analysis of bottromycin derivatives.

Substance ($[M+H]^+$)	Chemical fomular	Calculated mass (Da)	Detected mass (Da)
Bottromycin A2 (1)	$C_{42}H_{63}N_8O_7S^+$	823.4496	823.4530
Fragment 1a	$C_{35}H_{53}N_6O_5^+$	637.4072	637.3992
Fragment 1b	$C_{25}H_{44}N_5O_5^+$	494.3337	494.3324
Fragment 1c	$C_{25}H_{42}N_5O_4^+$	476.3231	476.3201
Fragment 1d	$C_{19}H_{31}N_4O_3^+$	363.2391	363.2363
Fragment 1e	$C_{16}H_{17}N_3O_2S$	315.1041	315.0995
Bottromycin B2 (2)	$C_{41}H_{61}N_8O_7S^+$	809.4339	809.4385
Fragment 2a	$C_{34}H_{51}N_6O_5^+$	623.3915	623.3940
Fragment 2b	$C_{24}H_{42}N_5O_5^+$	480.3180	480.3136
Fragment 2c	$C_{24}H_{40}N_5O_4^+$	462.3075	462.3257
Fragment 2d	$C_{18}H_{29}N_4O_3^+$	349.2234	349.2240
Fragment 2e	$C_{16}H_{17}N_3O_2S$	315.1041	315.0995
Bottromycin D2 (3)	$C_{41}H_{61}N_8O_7S^+$	809.4339	809.4372
Fragment 3a	$C_{34}H_{51}N_6O_5^+$	623.3915	623.4012
Fragment 3b	$C_{25}H_{44}N_5O_5^+$	494.3337	494.3389
Fragment 3c	$C_{25}H_{42}N_5O_4^+$	476.3231	476.3201
Fragment 3d	$C_{19}H_{31}N_4O_3^+$	363.2391	363.2341
Fragment 3e	$C_{15}H_{15}N_3O_2S$	301.0885	301.0815
Bottromycin E2 (4)	$C_{40}H_{59}N_8O_7S^+$	795.4183	795.4121
Fragment 4a	$C_{33}H_{49}N_6O_5^+$	609.3759	609.3735
Fragment 4b	$C_{23}H_{40}N_5O_5^+$	466.3024	466.3051
Fragment 4c	$C_{23}H_{38}N_5O_4^+$	448.2918	448.3001
Fragment 4d	$C_{18}H_{29}N_4O_3^+$	349.2234	349.2363
Fragment 4e	$C_{16}H_{17}N_3O_2S$	315.1041	315.0995

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Construction and screening of a *Streptomyces* sp. BC16019 cosmid library

For the construction of a 35-45 kb fragment library, 100 µg of the chromosomal DNA from *S. sp.* BC16019 was partially digested with *Sau*3AI in 1 ml final volume. 150 µl aliquots taken at various time points from the digestion were analyzed by electrophoresis using a 0.4% (w/v) agarose gel. After dephosphorylation with SAP, 2.5 µg of the DNA was ligated with 1.0 µg of the pOJ436 cosmid vector hydrolyzed with *Pvu*II/SAP and *Bam*HI. Packaging of the resulting ligation mixture was performed with the Gigapack[®] III gold packaging kit (Stratagene) and the extract was subsequently transformed into *E. coli* HS966 cells. After incubation in LB medium at 37 °C for 60 min, aliquots were plated out on LB agar amended with apramycin 60 µg/ml. A cosmid library consisting of 2304 clones was generated and automatically transferred twice onto a 22.2 x 22.2 cm nylon membrane (Biodyne B, Pall) with a Qbot robot (Genetix) as described previously (Dunham, 1997).

For the identification of a cosmid carrying the putative bottromycin biosynthetic gene cluster including the putative precursor peptide gene (*botA*), which was identified in the genome sequence data (see above), a suitable probe was designed based on the available genome sequence information. The ~ 0.5 kb fragment was amplified from *S. sp.* BC16019 chromosomal DNA using the primers contig572-bot-for and contig572-bot-rev (see Table below). After DIG-labelling (DIG DNA labeling and detection kit, Roche, Germany) and purification of the probe, overnight hybridization experiments were carried out at 42 °C, followed by stringent washing at 68 °C and developing of the membrane using the anti-DIG polyclonal antibody and detection kit according to the manufacturer's recommendations (Roche, Germany). Hybridizing cosmids were isolated, pre-analyzed by *Bam*HI digestion in order to verify their size and to identify putative

duplicates and then subjected to end-sequencing using the primers pOJ436-T3 and pOJ436-T7 (see Table below).

Inactivation of *botOMT* in *Streptomyces* sp. BC16019

Inactivation of *botOMT* gene in *S. sp.* BC16019 was accomplished by homologous recombination in the encoding region. Primers botK-KO-up and botK-KO-dn (see Table below) were used to amplify a 0.6 kb fragment homologous to an internal region of the target gene *botOMT*. The amplified DNA fragment was cloned into the suicide vector pKC1132 (Kieser et al., 2000) to generate the inactivation construct pKC1132-botOMT-KO. The *botOMT* inactivation mutant *S. sp.* BC16019::pKC1132-botOMT-KO was obtained by conjugation as follows: pKC1132-botOMT-KO was electroporated into *E. coli* ET12567/pUZ8002. Transformants were selected on LB agar plates supplemented with apramycin (60 µg/ml). Spores of *S. sp.* BC16019 were prepared from 7 day-old MS agar plates according to published procedures (Kieser et al., 2000). Conjugation was performed between the *E. coli* ET12567/pUZ8002/pKC1132-botOMT-KO and spores of the *S. sp.* BC16019. Conjugants were selected after incubation for 7 days on MS agar plates supplemented with 25 µg/ml nalidixic acid and 60 µg/ml apramycin. Chromosomal DNA of selected conjugants was isolated according to the Kirby mix procedure (Kieser et al., 2000). To confirm integration of the plasmid into the chromosome, three different PCR reactions were conducted with oligonucleotides botK-KO-out-r, botK-KO-out-l, botK-KO-in-r and botK-KO-in-l. Using primer pair botK-KO-out-r and botK-KO-out-l, a 0.9 kb fragment could be generated for *S. sp.* BC16019 wild type strain and no PCR product could be obtained for correct *S. sp.* BC16019::pKC1132-botOMT-KO mutant. Using primer pair botK-KO-out-l and botK-KO-in-l, a 0.63 kb fragment could be generated for correct *S. sp.* BC16019::pKC1132-botOMT-KO mutant and no product for wild type strain. Using primers botK-KO-out-r and botK-KO-in-r

a 0.65 kb fragment could be generated for correct *S. sp.* BC16019::pKC1132-botOMT-KO mutant and no product for wild type strain.

Modification of cosmid DG2 by Red/ET Recombination

In all Red/ET recombination experiments linear DNA fragments ('modification cassettes') flanked by suitable homology arms were generated by PCR. The PCR reactions were carried out with Go-Taq[®] Flexi DNA polymerase (Promega, Germany) according to the manufacturer's protocol. PCR products were concentrated and purified by ethanol precipitation prior to further use. In general, *E. coli* GB05-red cells harboring the parent vector to be modified were inoculated from an overnight culture (1:40 dilution), and cultivated on a thermomixer (Eppendorf) at 37 °C and 1000 rpm. When the OD₆₀₀ value reached 0.2, 0.1% L-arabinose was added to the culture to induce expression of the recombinases. Cultivation was then continued until the OD₆₀₀ reached 0.4. The cells were then harvested by centrifugation, washed twice with ice-cold distilled H₂O (dH₂O), and resuspended in 40 µl dH₂O. PCR product (0.2-0.3 µg) was mixed with the electro-competent cells and transferred into an ice-cold electroporation cuvette (1 mm). The mixture was then electroporated at 1300 V (Eppendorf electroporator), followed by addition of 1 ml LB. The cells were next cultivated at 37 °C and 1000 rpm for 75 min, before the culture was concentrated and transferred to LB agar plates containing suitable antibiotics. The plates were incubated at 37 °C overnight. Correct transformants were verified by restriction analysis and/or sequencing of the isolated cosmid DNA.

Construction of DG2-kan: A kanamycin cassette was amplified from pRK-Pirate-km using the primers DG2-short-up and DG2-short-down (see Table below) containing 50 bp homologous sequences to the downstream region of *botT* and the T7 terminus of cosmid DG2. After Red/ET recombination, colonies were selected on LB agar supplemented with 60 µg/ml apramycin and

50 µg/ml kanamycin at 37 °C. Replacement of the ~16 kb upstream region by integration of the kanamycin cassette into cosmid DG2 resulted in construct DG2-kan, which was verified by restriction analysis.

Construction of DG2-kan-efflux-ermE: A spectinomycin cassette was amplified from pRK-spect (Gene Bridges, Germany) using the primers DG2-efflux-ermE-up and DG2-efflux-ermE-dn (see Table below) containing 50 bp homologous sequences to the 5' region of *botT* (including the ATG start codon) and to the 5'-upstream region of *botT*. In addition the *Perme** promoter sequence (Bibb et al., 1994) was integrated into the primer DG2-efflux-ermE-up. Transformation of the PCR product into *E. coli* GB05-red cells carrying cosmid DG2-kan resulted in apramycin/spectinomycin resistant colonies. Subsequent restriction analysis of the isolated cosmid DNA showed the correct integration of the spectinomycin *Perme** cassette upstream of *botT* to replace the native promoter structure. The sequence of the inserted *Perme** promoter as well as the homology region in the 5' part of *botT* was verified by sequencing using primers efflux-up and efflux-dn (see Table below).

Construction of DG2-kan-botRMT1, DG2-kan-botRMT2, DG2-kan-botRMT3 and DG2-kan-botOMT: To delete the four methyl transferase encoding genes *botRMT1*, *botRMT2*, *botRMT3* and *botOMT*, respectively, from the expression construct DG2-kan, different selection marker cassettes were introduced via Red/ET recombineering (see Figure 3 and Figure S1). For the deletion of *botRMT1*, the ampicillin resistance cassette was amplified from pJET1.2 (Fermentas) using the primer pair botA-KO-up and botA-KO-dn (see Table below) containing homology arms to the upstream and downstream region of *botRMT1* as well as the *Perme** promoter to assure transcription of downstream gene(s). Recombinant clones were selected on LB-agar containing ampicillin 100 µg/ml and apramycin 60 µg/ml. For the deletion of *botRMT2*, the chloramphenicol resistance cassette was amplified from pBACe3.6 (Frengen et al., 1999) using the primer pair

botE-KO-up and botE-KO-dn (see Table below) containing homology arms to the upstream and downstream region of *botRMT2* as well as the *Perme** promoter to assure transcription of downstream gene(s). Recombinant clones were selected on LB-agar containing chloramphenicol 25 µg/ml and apramycin 60 µg/ml. For the deletion of *botRMT3*, the zeocin resistance cassette was amplified from pKen35 (Wenzel et al., unpublished) using the primer pair botI-KO-up and botI-KO-dn (see Table below) containing homology arms to the upstream and downstream region of *botRMT3* as well as the *Perme** promoter to assure transcription of downstream gene(s). Recombinant clones were selected on LB-agar containing zeocin 25 µg/ml and apramycin 60 µg/ml. For the deletion of *botOMT*, the ampicillin resistance cassette was amplified from pJET1.2 (Fermentas) using the primer pair botK-KO-up and botK-KO-dn (see Table below) containing homology arms to the upstream and downstream region of *botOMT*. Recombinant clones were selected on LB-agar containing ampicillin 100 µg/ml and apramycin 60 µg/ml. The resulting modified expression constructs, DG2-kan-botRMT1, DG2-kan-botRMT2, DG2-kan-botRMT3 and DG2-kan-botOMT, were verified by detailed restriction analysis as well as sequencing using the primer pair botA-up/botA-dn for DG2-kan-botRMT1, botE-up/botE-dn for DG2-kan-botRMT2, botI-up/botI-dn for DG2-kan-botRMT3 and botK-up/botK-dn for DG2-kan-botOMT (see Table below).

Transformation of the expression constructs into heterologous host strains

The original cosmid DG2 as well as DG2-kan, DG2-kan-botRMT1, DG2-kan-botRMT2, DG2-kan-botRMT3 and DG2-kan-botOMT were transformed into standard *Streptomyces* host strains by bi-parental conjugation as described above. Exconjugants were selected on MS agar plates supplemented with 25 µg/ml nalidixic acid and 60 µg/ml apramycin. Chromosomal DNA of the generated mutant strains (*S. coelicolor*::DG2, *S. albus*::DG2, *S. coelicolor*::DG2-kan,

S. coelicolor::DG2-kan-botRMT1, *S. coelicolor*::DG2-kan-botRMT2, *S. coelicolor*::DG2-kan-botRMT3 and *S. coelicolor*::DG2-kan-botOMT) was isolated as described above. The correctness of integration of the expression constructs into the Φ C31 phage attachment site was verified by PCR analysis using primers attB-l, attB-r, attp-up and attp-dn (see Table below). Using primer pair attB-l and attB-r, a 0.47 kb fragment could be generated for *S. coelicolor* A3(2) wild type strain and *S. albus* J1047 wild type strain. No PCR product could be obtained for corresponding correct mutants. Using primer pair attB-l and attp-dn, a 0.6 kb fragment could be generated for corresponding correct mutants and no product for wild type strain. Using primers attB-r and attp-up a 0.46 kb fragment could be generated for correct mutants and no product was found for the wild type strain. For heterologous bottromycin production, the generated mutants were grown in liquid TSB medium containing 60 μ g/ml apramycin at 30 °C and 180 rpm for 3 days. Production analysis was carried out by HPLC-MS as described below.

Induction of rifampicin resistance mutations in *S. coelicolor*::DG2 and *S. coelicolor*::DG2-kan

Spore suspensions from *S. coelicolor*::DG2 and *S. coelicolor*::DG2-kan (10, 20, 30, 40, and 50 μ l from stocks having a concentration of 10^8 spores/ml) were plated out onto MS agar plates supplemented with 10, 50 and 100 μ g/ml rifampicin. After incubation at 30 °C for 5 days several colonies appeared predominantly on plates with the lowest rifampicin concentration. From each expression system (*S. coelicolor*::DG2 and *S. coelicolor*::DG2-kan) 30 colonies were then screened for enhanced bottromycin production in comparison to the starting clones. Cultivations were carried out in 24-well tissue culture plates (Sarstedt 24 Well Tissue Culture Plate, Flat Bottom, Polystyrene, Sterile With Lid) containing 1.5 ml TSB medium per well amended with 60 μ g/ml apramycin and 10 μ g/ml rifampicin. During cultivation for 7 days at 50 rpm sterilized

distilled water was added to the wells to compensate evaporation. After threefold extraction of the culture supernatant with equal amounts of ethylacetate, the organic phase was evaporated and the extracts were redissolved in 50 µl methanol. Quantitative bottromycin production analysis was carried out as described below (injection volume was 5 µl). From the 30 colonies analyzed from each mutagenesis experiment two to three mutants with around 20-fold higher bottromycin production yields were obtained, named as *S. coelicolor*::DG2-rif and *S. coelicolor*::DG2-kan-rif).

Analysis of bottromycin production and novel derivatives thereof

For bottromycin production analysis, the cultures were harvested by centrifugation and 25 ml of the supernatant was extracted with an equal volume of ethyl acetate three times. The organic phase extract was evaporated and re-dissolved in methanol for HPLC-MS analysis. Standard analysis of crude extracts was performed on an HPLC-DAD system from the Agilent 1100 series, coupled to a Bruker Daltonics HCTultra ESI-MS ion trap instrument operating in positive ionization mode. Compounds were separated on a Luna RP-C₁₈ column (125 × 2 mm; 2.5 µm particle diameter; flow rate 0.4 ml/min, Phenomenex), with a mobile phase of water/acetonitrile each containing 0.1% formic acid, using a gradient from 5% - 95% acetonitrile over 20 min. Detection was by both, diode array and ESI-MS. High-resolution mass spectrometry was performed on an Accela UPLC-system (Thermo-Fisher) coupled to a linear trap-FT-Orbitrap combination (LTQ-Orbitrap), operating in positive ionization mode. Separation was achieved using a BEH RP-C₁₈ column (50 × 2 mm; 1.7 µm particle diameter; flow rate 0.6 ml/min, Waters), with a mobile phase of water/acetonitrile each containing 0.1% formic acid over 9 min. The UPLC-system was coupled to the LTQ-Orbitrap by a Triversa Nanomate (Advion), a chip-based nano-ESI interface. The yield of bottromycin A2 was determined from the peak area

(extracted ion chromatogram at $m/z [M+H]^+ = 823.4528$) by reference to a standard curve generated with pure bottromycin A2. Determination of bottromycin B2 and other novel bottromycin derivatives bottromycin D2 and E2 produced by respective mutants were achieved via extracted ion chromatogram at $m/z [M+H]^+ = 809.4385$ (for bottromycin B2 and E2), at $[M+H]^+ = 795.4208$ (for bottromycin D2). For quantitative measurement of bottromycin A2 production, a standard curve was established by preparing a concentration gradient (from 0.2 mg/ml to 0.2 ng/ml in a dilution factor 1:10) of pure bottromycin A2 as reference substance. To determine the peak area, extracted ion chromatogram at $m/z [M+H]^+ = 637.3992$, $m/z [M+H]^+ = 476.3201$ and $m/z [M+H]^+ = 315.0995$ were selected to be observed.

Table. Oligonucleotides used in this study (related to Supplemental Experimental Procedures)

Primers	Sequence (5'-3')
botK-KO-up	GATATCACGCCCCGCTGACGCTCGACGCCCCGCT (underlined: introduced stop codon)
botK-KO-dn	AAGCTTCCAGTTCGCGTGGGTCGTTCA
botK-KO-out-l	GCTCCCGGTCGGTCGGAGGCTCC
botK-KO-out-r	TGTGCGCCAGAGAGGCGAGAA
botK-KO-in-l	TTCACACAGGAAACAGCTATGA
botK-KO-in-r	AGGGTTTTCCCAGTCACGACGTTG
contig572-bot-for	CCCCAACTACGACGG TTTC
contig572-bot-rev	ACCCGG TGAGGATGTTCCAG
pOJ436-T3	TTAATTGGGAGTGATTTCC
pOJ436-T7	CATCACCGGCGCCACAGGTG
DG2-short-up	GATCCCCCATCCCGGGCCAGGGCTCGGAAACAGCGGATACCCCGACGGCCGCT GCTCTAGTACACCTGAATAACT
DG2-short-dn	GATCACCAGGCCAGCACCACACCGAGGGCGAAGGTCAGCCGCCAGCCGAGT CGTTGCTGTATCTCCAGATAACT
DG2-efflux-ermE-up	GTGTTGCCCAGCAGGGACAGGCCGTGGGAGATGAGGTAGAGCGGCAGCATTT CCGCTCCCTTCTCTGACGCCGCTGGATCCTACCAACCGGCACGATTGTGCCCCAC AACAGCATCTCGATTTTCGTTTCGTGAATACATG, (underlined: <i>PerME*</i> promoter sequence)
DG2-efflux-ermE-dn	ACGCCGTCAACCATCCGACGGGCCATCGGCCACGACCAGAATGCCCCGTTA TAATTTTTTTAATCTGTTATTTA
efflux-up	ACGACGGGGAGTGCGGCGACGCAGG
efflux-dn	CCTCGGCCGCCGAGGGAGTTCGGAG
botA-KO-up	GTGCCTGCCGCCAAAAACGTTTCTTGTATCAATGCCTTGGAAGTGAACGAGACGA AAGGGCCTCGTGATACGCCT
botA-KO-dn	TGAGGAAATCCGCGGTCATGCAGTCGAATACGACTACGGGTCCCATTTCGCT CCCTTCTCTGACGCCGCTGGATCCTACCAACCGGCACGATTGTGCCCCACAACA GCATCGGTCTGACAGTTACCAATGC
BotE-KO-up	AGCTCGCGCTCGTGACCAACCCTGGAACAGCCTGTACCGCCCGTGATACCTG TGACGGAAGATCAC
BotE-KO-dn	ACACCTCGAAGACGTCCCTACGGCTCGTGCCGTTACCGTCCCGCACTTCCGCTC CCTTCTCTGACGCCGCTGGATCCTACCAACCGGCACGATTGTGCCCCACAACAG CATCGGGCACCAATAACTGCCTTA
BotI-KO-up	ATGACCGGAACGACCGAGAGGACCGAGGCGACCGGCCCGGCGGATTGAGGAT CTGATCAGCACGTGTT
BotI-KO-dn	CAGGTCCTCGGCGCGGGTCTGGTCTCGGTGGGGGCCGTTGGTCATTTCCGCTCC CTTCTCTGACGCCGCTGGATCCTACCAACCGGCACGATTGTGCCCCACAACAGC ATCTCAGTCCTGCTCCTCGGCCA
BotK-KO2-up	CCCGGTCGGTCGGAGGCTCCCGGTCGGTCGGAGGCTCCGGCGGGAGGTCTGAC AGTTACCAATGC
BotK-KO2-dn	GTGAAGATTTCCCTGACCGGCGCGGCGGAGACCCTGCTCGCACCCCTGAGACGA AAGGGCCTCGTGATACGCCT
botA-up	TGAATGAAGCCATACCAA
botA-dn	GAAATGTTGAATACTCATA
botE-up	GCAAATATTATACGCAAGGC
botE-dn	TAATATCCAGCTGAACGGT
botI-up	GGACAACACCCTGGCCTG
botI-dn	GTCGTCTCCACGAAGTC
botK-up	GAAATGTTGAATACTCATA
botK-dn	TGAATGAAGCCATACCAA
attB l	CAGGTTACCCACAGCTG
attB r	CTCAACTAAAGTGGGGCG
attp_up	AGAGATACAGGGTCATGGAA
attp_dn	AAATGCCCGACGAACCTGAA

SUPPLEMENTAL REFERENCES

Bibb, M.J., White, J., Ward, J.M., and Janssen, G.R. (1994). The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Molecular Microbiology*. 3, 533-545.

Dunham, I. Bacterial Cloning Systems. Birren, B. Genome Analysis. A Laboratory Manual. Cloning Systems. [3], 1-86. 1997. New York, Cold Spring Harbor Laboratory Press.

Frengen, E., Weichenhan, D., Zhao, B., Osoegawa, K., van, G.M., and de Jong, P.J. (1999). A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics*. 3, 250-253.

Kieser, T., Bibb, M., Buttner, M. J., Chater, K. F., and Hopwood, D. A. Practical Streptomyces Genetics. 613. 2000. Norwich, England, The John Innes Foundation.

C. Final Discussion

1. General scope of this work

The present thesis deals with the identification, elucidation and engineering of biosynthetic pathways of the natural products cinnabaramides from *Streptomyces* sp. JS360 and bottromycins from *Streptomyces* sp. BC16019. Synthetic biotechnology tools were employed to generate their novel natural product derivatives. Moreover, detailed biochemical and structural characterizations were performed on one key enzyme responsible for the formation of an unusual polyketide extender unit incorporated in the cinnabaramide biosynthetic assembly line which contributed to a better mechanistic understanding of this family of enzymes and revealed how the primary metabolic enzymes evolved to produce dicarboxylic acids incorporated in secondary metabolic pathways. Investigation of active site mutations were conducted to set the stage for generation of deliberately altered polyketides with improved bioactivity and bioavailability for clinical application.

2. Cinnabaramides biosynthesis in *Streptomyces* sp. JS360

Cinnabaramides are a class of compounds isolated from a terrestrial Actinomycete *Streptomyces* sp. JS360 and are a potent and selective inhibitor of the human 20S proteasome. The strain is known to produce cinnabaramides A-G, each containing a characteristic γ -lactam- β -lactone bicyclic ring structure attached to a cyclohexenyl unit together with a hexyl side chain (**Figure A9 and C1b**).^[123] Significant structural similarities were observed between cinnabaramides and salinosporamides, a family of type I PKS/NRPS hybrid antitumor metabolites produced by a marine Actinomycete *Salinispora tropica* CNB-440.^[140-142] Structures of cinnabaramides and salinosporamides only differ by the polyketide side chains, salinosporamides incorporate a chloroethylmalonyl-CoA into the biosynthetic pathway resulting in a chloroethyl side chain^[143;144] while we demonstrate in this work that cinnabaramides incorporate a hexylmalonyl-CoA leading to a hexyl side chain (**Figure C1b**).^[124;145] During the present thesis, the cinnabaramides biosynthetic pathway from *Streptomyces* sp. JS360 was elucidated, cloned and sequenced. The uncovered type I PKS/NRPS hybrid assembly line exhibited several unique features in utilization of unusual building blocks and their flexibility on substrate recognition.

2.1 Cinnabaramides biosynthesis is directed by a hybrid type I PKS/NRPS incorporating unusual building blocks

Based on the known hybrid PKS/NRPS machinery for salinosporamides^[143] and the high structural similarity to cinnabaramides, the biosynthetic origin for cinnabaramides was predicted to follow a PKS/NRPS hybrid architecture (**Figure C1a**). According to the co-linearity between the required biosynthetic steps and the complement of the existing modules, we presumed that the cinnabaramide

C. Final Discussion

biosynthetic pathway is initiated by selection and loading of acetyl-CoA by the first AT domain CinA-AT₁. This hypothesis is supported by the *in silico* prediction of substrate specificity of the CinA-AT₁ domain and by the comparison to the salinosporamide SalA-AT_L domain which was also proven to be specific for acetyl-CoA via stable isotope feeding experiments.^[146] *cinA* encodes a bi-modular PKS (**Figure C1b**) which shows uncommon domain organization harboring contiguous AT domains responsible for loading and extension distinctly different from the standard prototypic PKS organization of AT₁-ACP-KS-AT₂-ACP₂. Interestingly, this noncanonical type of architecture consisting of an intermixed loading and first extension module is found in several myxobacterial systems, such as those involved in stigmatellin, soraphen and aurafuranone biosyntheses.^[48;147;148] The substrate specificity analysis of the CinA-AT₂ domain revealed a distinct difference from its counterpart in the salinosporamide gene cluster SalA-AT₁ which is specific for an unusual PKS extender unit chloroethylmalonyl-CoA.^[142;146] Based on this analysis and the structures of cinnabaramide harboring a hexyl side chain, we inferred that the second PKS module catalyzes the chain elongation of the acetyl-S-ACP intermediates with the unusual extender unit hexylmalonyl-S-ACP resulting in the incorporation of the hexyl side chain. Similar example was reported in biosynthesis of thuggacins from myxobacterium *Sorangium cellulosum* So ce895, a family of macrolide antibiotics bearing hexyl side chains.^[35] However, CinA-AT₂ also strongly differs (similarity <20% on the protein level) from the hexylmalonyl-CoA specific TugC-AT₃ involved in the biosynthesis of thuggacins.^[35;149] This fact prompted us to verify this hypothesis by closer examination into the biosynthetic mechanism of the unusual PKS extender unit hexylmalonyl-CoA.

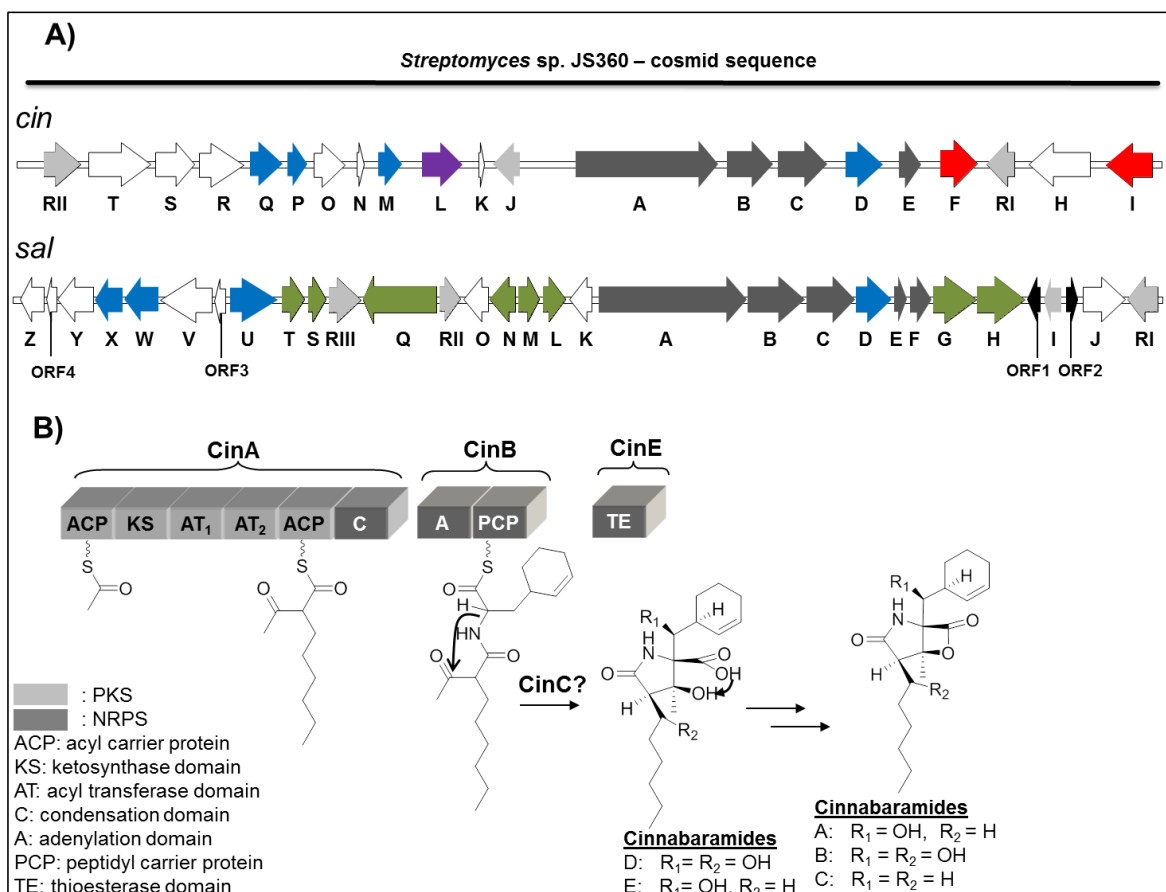


Figure C1. A) Schematics of the *cin* locus from *Streptomyces* sp. JS360 and the *sal* locus from *Salinospora tropica* CNB-440. Similar genes are involved in both salinosporamide and cinnabaramide biosyntheses: genes involved in construction of the core γ -lactam- β -lactone bicyclic ring (in dark gray), genes involved in assembly of the non proteinogenic amino acid β -hydroxyl-cyclohex-2'-enylalanine (in blue), and genes involved in regulation and resistance (in light gray). Specific genes existing only in cinnabaramide biosynthesis: genes involved in the hexylmalonyl-CoA pathway (in red) and genes involved in hydroxylation of C12 in cinnabaramide B and D (in purple). Specific genes existing only in salinosporamide biosynthesis: partial transposases (in black) and genes involved in the chloroethylmalonyl-CoA pathway (in green). Unknown functional genes involved in biosynthesis of cinnabaramides and salinosporamides (blank). **B) Proposed PKS/NRPS cinnabaramide biosynthetic assembly line.**

Based on previous observations in other natural product pathways involving unusual PKS extender units, e.g., as mentioned above, in the biosynthetic gene cluster of thuggacins, the unusual extender unit hexylmalonyl-CoA is thought to be generated from 2-octenoyl-CoA converted by a crotonyl-CoA carboxylase/reductase (CCR) analogue TgaD^[35], and later on, predominantly triggered by the work relating to the ethylmalonyl-CoA pathway in *Rhodobacter sphaeroides* in 2009^[150], whereas a purified CCR enzyme was *in vitro* characterized to catalyze the reductive carboxylation of (*E*)-crotonyl-CoA resulting in the production of ethylmalonyl-CoA, we predicted that the hexyl chain of cinnabaramides might originate from 2-octenoyl-CoA via a similar reductive carboxylation. Similar biochemical steps have also been described in production of PKS extender units such as ethylmalonyl-CoA and chloroethylmalonyl-CoA in the salinosporamide pathway.^[151] Indeed, within the *cin* gene cluster, we found a potential gene, *cinF*, encoding a protein showing 90% sequence identity on the protein level to a putative CCR from *Streptomyces hygroscopicus* (Accession No. AAR32675). Thus, CinF was

C. Final Discussion

thought to be the key enzyme for formation of the unusual extender unit hexylmalonyl-CoA in cinnabaramide pathway (**Figure C2**). However, the sequence similarity between CinF and TgaD is unexpectedly low (16% identity on the protein level) despite their assumed identical functionality.^[145] First of all, the disruption of *cinF* led to a complete abolishment of cinnabaramide production indicating that CinF is indispensable for the biosynthesis of cinnabaramides (L. Huo and R. Müller, unpublished data). We thus hypothesized that CinF catalyzes the reductive carboxylation of 2-octenoyl-CoA or -ACP, leading to hexylmalonyl-CoA or -ACP, respectively (**Figure C2**). This hypothesis has been verified by further investigation of CinF biochemically and structurally. These results have identified CinF as a hexylmalonyl-CoA synthase responsible for catalyzing the conversion of 2-octenoyl-CoA to hexylmalonyl-CoA in the presence of hydrogen carbonate and reduced nicotinamide adenine dinucleotide phosphate (NADPH).^[145] Moreover, further mutagenetic studies on the active site of CinF was also performed^[145] and is discussed below in **Discussion 2.2 Biosynthetic mechanism of the hexylmalonyl-CoA pathway**. In addition, a putative acyl-CoA dehydrogenase encoding gene, *cinI*, is presumably involved in generation of 2-octenoyl-CoA from octanoic acid that could originate from primary fatty acid degradation and/or synthesis (**Figure C2**).

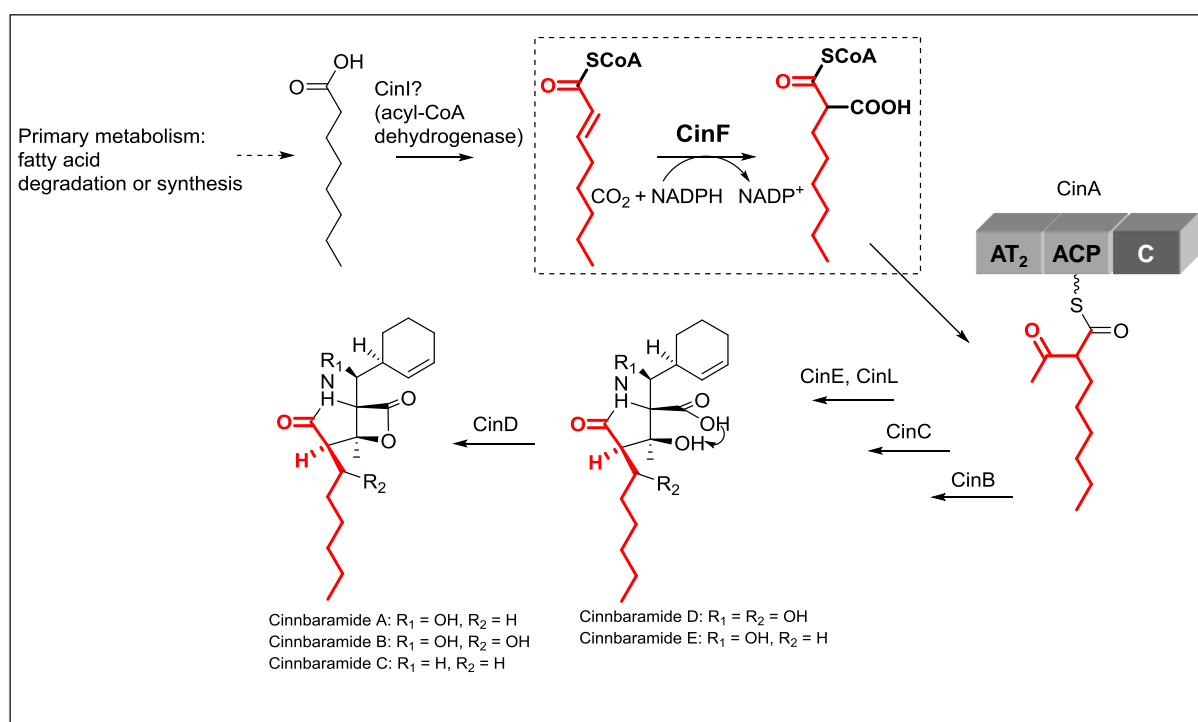


Figure C2. Proposed biosynthetic pathway of production of the hexyl side chain and its incorporation into cinnabaramides biosynthesis.

After we identified that the hexylmalonyl-CoA is the native extender unit recognized by CinA AT₂ domain, it was subsequently demonstrated CinA AT₂ has relaxed substrate specificities towards various chlorinated 2-octenoic acids.^[124] This is inspiring for us to use synthetic biotechnological tools to generate more polyketide-chain altered cinnabaramide derivatives of novel structures and/or

improved bioactivities. Detailed discussion is in **Discussion 2.3 Creation of molecular diversity based on the cinnabaramide scaffold via synthetic biotechnology.**

Subsequently, CinB and the C-terminal domain of CinA form the NRPS module of the assembly line, which is supposed to be responsible for incorporating the non-proteinogenic amino acid 2'-cyclohexenyl-alanine (**Figure C1**). The deduced substrate specificity code of the CinB-A domain is significantly similar to the SalB-A domain which has been proven to be specific for selection and activation of the exactly same substrate.^[140;152] We thus proposed that the elongation of the cinnabaramide intermediate proceeds with extension of the same amino acid residue. Indeed, this hypothesis was supported by the feeding of the synthesized (2*S*,4*R*)- β -cyclohexenylalanine and the diastereomeric (2*S*,4*S*)-cyclohexenylalanine to the growth culture of *S. sp.* JS360 resulting in the production of deuterated cinnabaramides. The work of Moore's group has proposed that in the *sal* locus a set of enzymes including 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase (SalU), prephenate dehydratase (SalX) and L-amino acid amino transferase (SalW) are involved in the non proteinogenic amino acid formation. This pathway employs phosphoenolpyruvate and D-erythrose-4-phosphate as starting points, followed by a branch of the phenylalanine pathway leading to the formation of dihydroprephenate. After decarboxylative dehydration to yield 2,3-dihydrophenylpyruvate that undergoes a 1,3-reduction and transamination, the 3-R-cyclohex-2-enyl-L-alanine is finally formed.^[120;143] Recently, another similar pathway has been reported in *Bacillus subtilis* in the biosynthetic pathway of tetrahydrotyrosine as part of anticapsin biosynthesis.^[153] Expectedly, almost all of the corresponding counterparts of these involved enzyme encoding genes could also be found in the *cin* locus which is another supportive evidence for the hypothesis that 2'-cyclohexenyl-alanine is the native building block for the cinnabaramide NRPS pathway. Thus, a biosynthetic pathway of 2'-cyclohexenyl-alanine could also be postulated for cinnabaramide biosynthesis (**Figure C3**). It is noteworthy to state that since SalB-A domain exhibited relaxed substrate specificities towards a number of aliphatic amino acids^[120] and the substrate specificity code of CinB-A domain is highly similar to SalB-A domain, we determined that the CinB-A domain is also tolerant to other proteinogenic or non-proteinogenic acids. This encouraged us to use synthetic biotechnological tools to generate more amino acids-altered cinnabaramide derivatives of novel structures and/or improved bioactivities. Detailed discussion is below in **Discussion 2.3 Creation of molecular diversity based on the cinnabaramide scaffold via synthetic biotechnology.**

C. Final Discussion

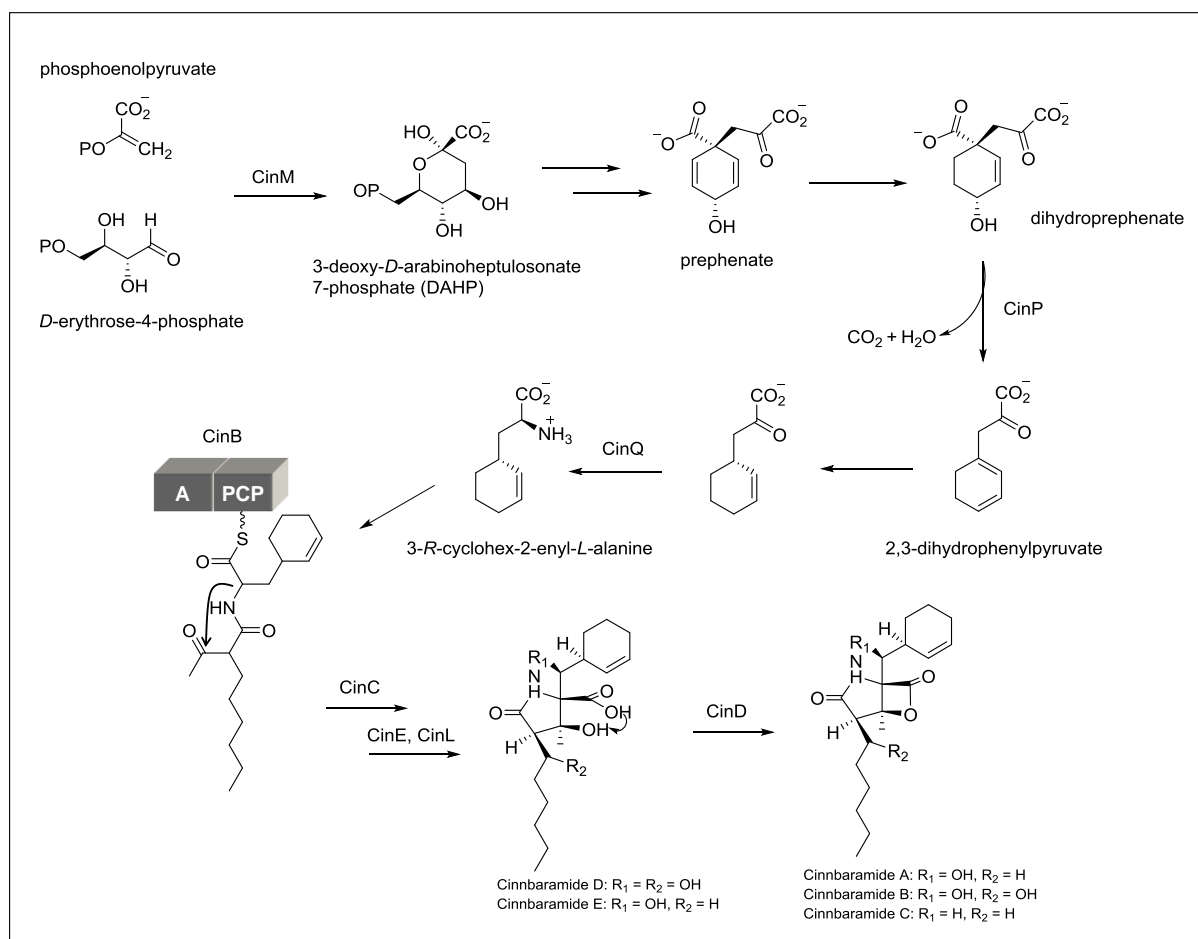


Figure C3. Proposed biosynthetic pathway of non proteinogenic amino acid building block 2'-cyclohexenyl-alanine for cinnabaramides biosynthesis. Scheme modified according to the reference.^[154]

After the NRPS module incorporates 2'-cyclohexenyl-alanine into the assembly line, the resultant PCP-tethered intermediate undergoes an intramolecular condensation presumably promoted by the ketosynthase (CinC) to generate the γ -lactam ring, which leads the CinB-bound substrate to give rise to the offloaded β -lactone (**Figure C1**). CinE, which shows a high similarity to SalF, is a putative type II thioesterase, catalyzes the release and cyclization of the intermediate to generate the β -lactone ring structure, and it is thought to be also responsible for removing mis-primed precursor molecules from PKS/NRPS modules.^[143]

To sum up from above, cinnabaramides biosynthetic gene cluster possesses a PKS/NRPS hybrid structure which employs unusual building blocks incorporated into the pathway. The analysis of such hybrid systems is of special interest for combinatorial biosynthesis because the combination of PKS and NRPS can lead to an immense variety of novel structures.^[155;156] Thus, future combinatorial biosynthesis work on cinnabaramide biosynthetic pathway could be planned to generate further derivatives of interesting chemical structures and/or bioactivities. Besides, the feature that both PKS and NRPS pathways exhibited relaxed substrate specificities could allow us to generate novel cinnabaramides derivatives via synthetic biotechnological tools, such as precursor-directed biosynthesis and/or mutasynthesis. The possible challenging obstacle would be the limited flexibility of substrate specificity of the megasynthetases, especially of the 'gate keeper' AT domain and A

C. Final Discussion

domain. Thus, further pathway engineering on CinA-AT₂ domain and CinB-A domain to alter their substrate specificity towards desired scaffolds could be the plan to unblock the obstacle. A number of successful examples have been reported using rational manipulations domains/modules to alter the selection of building blocks by site-directed mutagenesis, domain/module exchange, module insertion, fusion or deletion.^[157-167] Despite the reduced catalytic efficiency usually observed after the pathway engineering, this strategy is definitely worth trying in the future work.

Additionally, two putative cytochrome P450 enzyme encoding genes are found nearby, *cinD* inside the gene cluster and *cinL* outside of the gene cluster (**Figure C1**). We suggested that they are both involved in post-assembly line hydroxylation enabling the production of various cinnabaramide derivatives (cinnabaramides A, B, D, E). Compared to the salinosporamide gene cluster, no counterpart of *cinL* could be found suggesting that CinL possibly plays a role in hydroxylation of cinnabaramides B and D at C12 because no natural salinosporamide analogues are hydroxylated at the corresponding carbon positions.^[143] The generated *cinL*-deficient mutant was unable to produce cinnabaramide D but yielded distinctly more cinnabaramide E thereby strongly supporting this speculation (L. Huo and R. Müller, unpublished data). Accordingly, CinD is exclusively supposed to be involved in hydroxylation of the 2'-cyclohexenyl-alanine residue resulting in the β -hydroxy-2'-cyclohexenyl-alanine residue in cinnabaramides A, B, D and E. It is noteworthy that the salinosporamide megasynthetase was shown to tolerate both secondary and tertiary carbons in the γ -position of the amino acid precursors but substitution in the β -position was not accepted.^[140] Based on this observation we administrated deuterated cyclohexenyl-serine, in which the β -position is substituted with a hydroxyl group, into the culture of *Streptomyces* sp. JS360. No incorporation of the supplemented compound was observed indicating that the cinnabaramide biosynthetic megasynthetase cannot tolerate alterations in the β -position of the amino acid precursor.^[124] We suggested that the reason might be that the substitution directly affects the functionality of CinD in the pathway. These results not only give us very important guidance for precursor-directed biosynthesis and/or mutasynthesis, but also open the door for post-assembly line enzyme modifications. Since the *cinL*-deficient mutant indeed produced significantly more cinnabaramide E but no cinnabaramide E anymore, the deletion of CinD could also be considerable, which would lead to an selective production of novel cinnabaramides containing non-hydroxylated 2'-cyclohexenyl-alanine residue. Furthermore, the combined deletions of CinD and CinE could be optional to result in more novel cinnabaramides.

Last but not the least, the *cin* locus harbors genes responsible for regulation, self-resistance and unknown functions for which their encoding proteins are still to be investigated. Finally, the characterized biosynthetic gene cluster was cloned into one cosmid and heterologously expressed in *Streptomyces coelicolor* A3(2) and *Streptomyces albus* J1047 under the control of the native regulatory elements (L. Huo and R. Müller, unpublished data). This confirmed the involvement of all genes in cinnabaramide biosynthetic pathway and established a platform for further genetic engineering to optimize the heterologous production titer, such as the promoter exchange, deregulation

of the native regulation system and introduction of an inducible system into the heterologous host and enhancement of the self-resistance. Some of them was already successfully applied in yield optimization of heterologous production of bottromycin^[125], which gave us a very good model. In the other hand, this heterologous platform could also be utilized in combination of Red/ET technique to generate unnatural derivatives of cinnabaramides.

2.2 Biosynthetic mechanism of the hexylmalonyl-CoA pathway

Except for “usual” extender units as malonyl-CoA and methylmalonyl-CoA incorporated in PKS,^[24] several observations have been reported in polyketide biosynthetic pathways involving unusual extender units including methoxymalonyl-CoA,^[168;169] aminomalonyl-CoA,^[170] hydroxymalonyl-CoA,^[170] ethylmalonyl-CoA,^[171] hexylmalonyl-CoA,^[35;124] chloroethylmalonyl-CoA^[146] and 2-carboxyl-5-methylhexanoyl-CoA;^[172] Interestingly, the biosynthetic origins of most unusual extender units are poorly understood to date. According to the literature,^[168-170] a very limited number of biosynthetic pathways of these usual extender units have been characterized, such as the methoxymalonyl-CoA originates from a glycerol-derived precursor, which could be bound to a CP and further processed to form methoxymalonyl-CoA.^[53;173] Later report revealed that hydroxymalonyl-CoA and aminomalonyl-CoA can also be biosynthesized via similar fashion.^[174] However, none of those could retrobiosynthetically explain the formation of polyketides side chains (**Figure C4**). Until the recent reports on the biochemistry of the CCR family of enzymes appeared,^[150;175] the formation of many of those unusual extender units could possibly be explained retrobiosynthetically (**Figure C4**). As a new representative of CCRs, CinF was biochemically characterized *in vitro* and was found to convert 2-octenoyl-CoA and 2-octenoic acid *N*-acetylcysteamine (NAC) thioester into hexylmalonyl-CoA and hexylmalonic acid SNAC thioester in the presence of hydrogen carbonate and reduced nicotinamide adenine dinucleotide phosphate (NADPH).^[145] These results confirmed the hypothesis that CinF is responsible for providing hexylmalonyl-CoA for the synthesis of cinnabaramides.^[124]

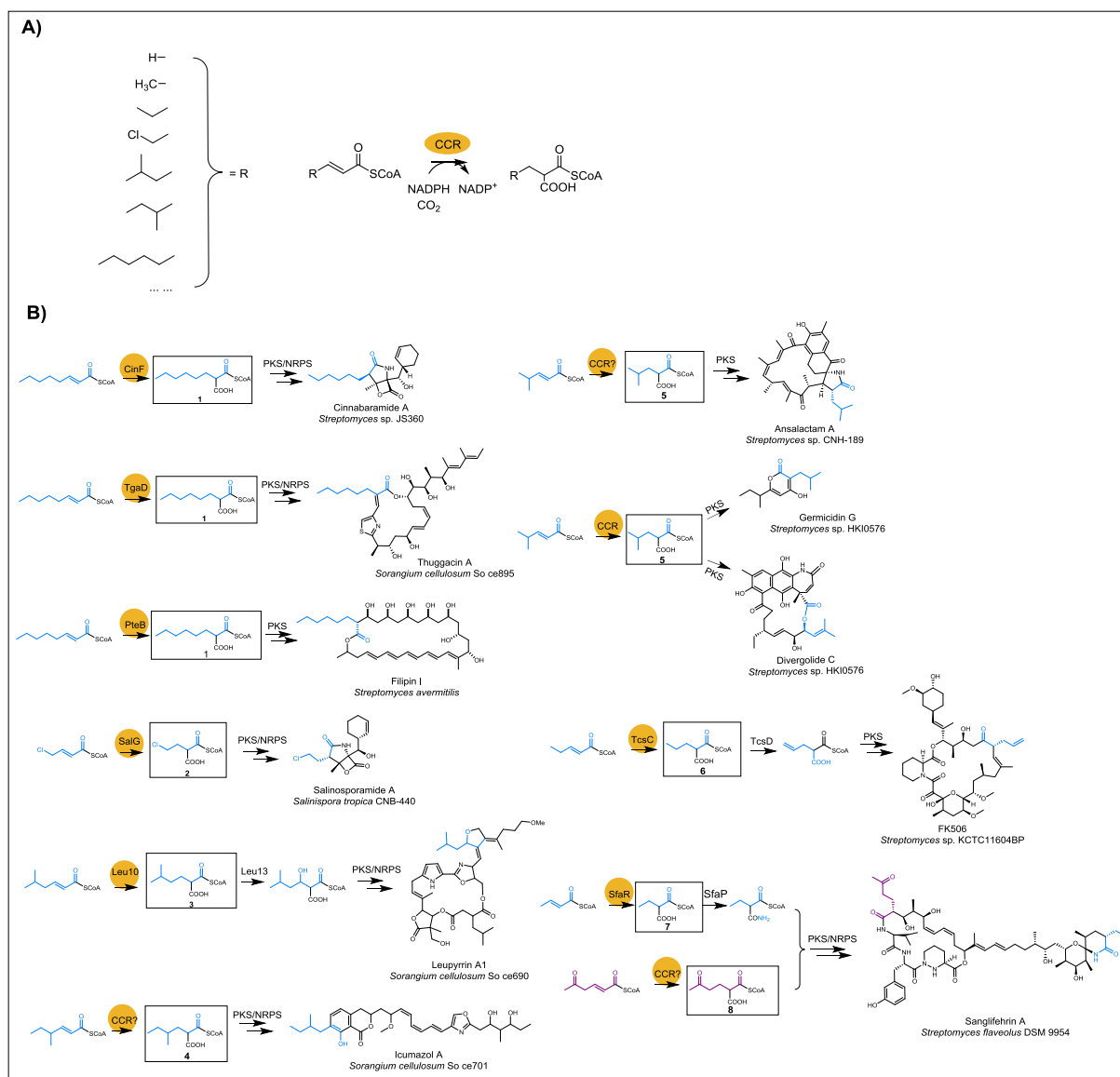


Figure C4: Reactions catalyzed by crotonyl-CoA carboxylase/reductases (CCRs). A) CCRs generate a number of different building blocks for polyketide biosynthesis while consuming NADPH and CO_2 . B) Detailed production of unusual building blocks for polyketide synthesis by means of reductive carboxylation and installation of bacterial polyketide metabolites: (1) hexylmalonyl-CoA; (2) chlorocrotonyl-CoA; (3) 2-carboxyl-5-methylhexanoyl-CoA; (4) 2-carboxyl-4-methylhexanoyl-CoA; (5) 2-carboxyl-4-methylpentanoyl-CoA; (6) propylmalonyl-CoA; (7) ethylmalonyl-CoA; and (8) 2-(2-oxobutyl)-malonyl-CoA.

Furthermore, we solved the structure of the enzyme in complex with NADP^+ and its substrate 2-octenoyl-CoA at high resolution in collaboration with the Structural Biology Department at Helmholtz Centre for Infection Research (HZI, Braunschweig, Germany). After several trials we failed to generate detectable CO_2 binding crystals with Na_2CO_3 or CS_2 as mimics of CO_2 . We then turned to docking CO_2 into the CinF structure which suggested that residues E167 and N77 were involved in binding and activation of CO_2 (**Figure C5a**). To verify this hypothesis we tested several mutants of these residues (CinF^{E167A}, CinF^{E167K}, CinF^{E167Y}, CinF^{N77A}, CinF^{N77E} and CinF^{N77Y}) for their ability to perform the reduction and carboxylation reactions. The activity assays showed that all 6 CinF mutants lost their reductive carboxylation activity but partially retained the reductive activity versus 2-

octenoyl-CoA.^[145] These data strongly support our hypothesis and our *in silico* analysis regarding the CO₂ binding site.

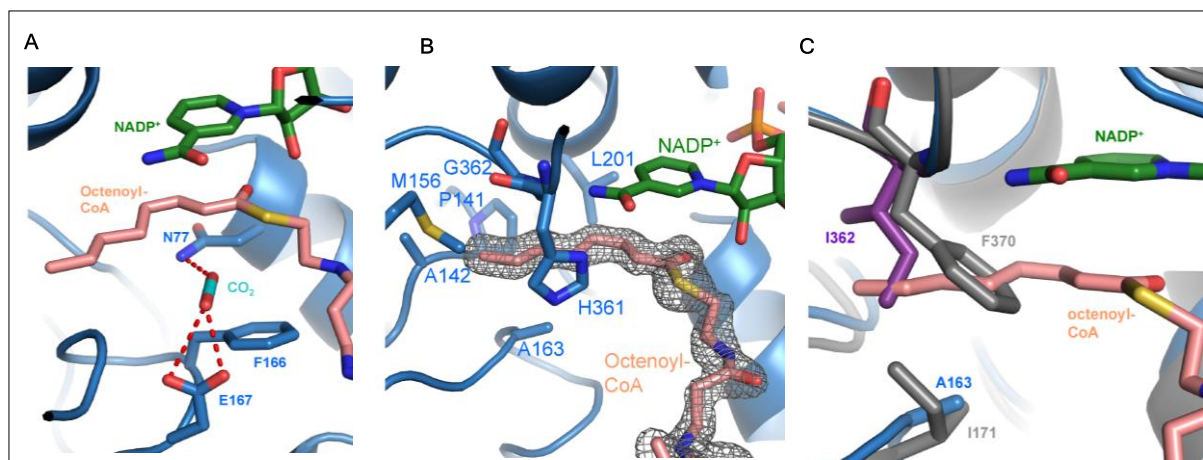


Figure C5. Ligand binding to CinF. **A)** Model of CO₂ docked into the structure of CinF, CO₂ is shown as cyan/red sticks, dashed lines indicate hydrogen bonds. **B)** Close-up of the hydrophobic substrate binding pocket for 2-octenoyl-CoA. **C)** Comparison of the substrate binding sites of CinF (in blue), CCR from *S. coelicolor* (in gray) and SalG (in purple).

CinF has a hydrophobic binding pocket for the long octenoyl chain (**Figure C5b**) while the other members of the carboxylase-reductase family have distinctly smaller ones for shorter chains. For example, the pocket is blocked by the large residues Phe370 and Ile171 in a crotonyl-CoA-specific CCR from *S. coelicolor* (PDB code: 3KRT) whereas in CinF the corresponding residues Gly362 and Ala163 are small to form a larger pocket. The chlorocrotonyl-CoA-specific SalG also has an Ala163 as in CinF, but instead of the Gly362 or Phe370 it has an Ile362 (**Figure C5c**).^[146] These observations prompted us to determine whether these two positions are crucial for CCR substrate specificity in terms of the length of the aliphatic chain. Mutagenetic studies on G362F, G362I and A163I render CinF unable to convert 2-octenoyl-CoA into product, indicating their relevance to the activity. Furthermore, the activity screening towards various substrates including crotonyl-CoA, 2-octenoyl-CoA and 2-octenoic acid SNAC thioester showed that CinF^{G362F} and CinF^{A163I} exhibited weak catalytic activity toward crotonyl-CoA, whereas CinF^{G362I} completely lost the activity towards all substrates. It was therefore possible to suppress the 2-octenoyl-CoA converting activity by one single mutation in the active site while retaining some of its crotonyl-CoA converting activity. The two crucial residues, Gly362 and Ala163, in substrate binding pocket explain why CinF is able to use the longer 2-octenoyl-CoA chain while the other CCRs are only able to accept substrates with shorter chains. Based on this, it becomes possible that various substrates harboring different chain lengths can be accommodated by varying the size of specific side chains within the substrate-binding pocket leading to different substrate specificities for these enzymes. Through this, engineering of CCRs to produce various PKs incorporating uncommon extender units is of high interest for drug development. For this purpose, we already further generated and biochemically characterized more CCRs, such as TgaD, Leu10 and CCR involved in icumazole biosynthetic pathway (**Figure C4**) (L. Huo, F. Yan and R. Müller, unpublished data). The results revealed that each CCR has its own favorite substrate despite they share a common

C. Final Discussion

reductive carboxylation mechanism. A detailed description of the CCR catalyzed reaction mechanism has been reported by Erb group very recently.^[176] Based on the determined various substrate specificities from generated CCRs and available crystal structure of CinF, especially the well-known active center and substrate recognition pocket, we can align the CinF sequence with other CCRs and make a series of rational site-directed mutagenesis on CinF substrate recognition pocket, expecting that modified CinF variants are able to catalyze the formation of shorter, longer, hydroxylated, halogenated etc. dicarboxylic acids. All of them could be served as extender units for polyketide biosynthesis, which could tremendously diversify the natural products. In conclusion, these findings provide a solid basis for attempts to modify the CCRs substrate specificity toward generating novel polyketide extender units and eventually the production of altered polyketide structures. However, generation of novel extender units is just the first step but not the last. Based on PKS enzymology, incorporation of those extender units into polyketides is dependent on acyl-transferase domains required for recognition and loading onto the corresponding carrier proteins of the PKS. Furthermore, the subsequent elongation modules also need to be capable of processing the altered intermediate and the corresponding unsaturated substrate for the CCR reaction must be available. Given these obstacles the success of biosynthetic engineering very challenging but definitely worth trying; nonetheless, first successful examples have been shown for the cinnabaramide biosynthetic pathway, specifically unnatural dicarboxylic acids were accepted by the PKS and subsequently processed to produce chlorinated analogues (discussed below).

2.3 Creation of molecular diversity based on the cinnabaramide scaffold via synthetic biotechnology

It has been reported that in salinosporamide biosynthesis the enzymes responsible for activation and incorporation of the native *sal*-specific amino acids exhibit relaxed substrate specificities toward a number of aliphatic amino acids. Moore and coworkers used this promiscuity to generate a series of salinosporamides analogues with altered bioactivities against human cancer cell lines;^[177] however, no cinnabaramides derived from proteinogenic amino acids have been detected in extracts of the wild type producer. Interestingly, the native substrate (2*S*,4*R*)- β -cyclohexenylalanine and the diastereomeric (2*S*,4*S*)-cyclohexenylalanine were both incorporated after feeding to the growth media of *S. sp.* JS360,^[178] indicating that either 1) the CinB A domain tolerantly activates the specific native substrates regardless of the chemical configurations but other free amino acids cannot be activated for further incorporation or 2) *in vivo* insufficient quantities of amino acids are available for secondary metabolite biosynthesis. Since the latter scenario is very unlikely we assume that it should be possible for the CinB A domain to directly transfer alternative amino acid *N*-acetylcysteamine (NAC) thioesters as mimics to the corresponding activated CoA thioesters to the CinB PCP domain and further incorporate them in the respective position of cinnabaramides. It has been reported that two salinosporamide derivatives with saturated cyclohexane or cyclopentane rings were successfully prepared by mutasynthesis through administration of derivatives and branched-chain amino acids to

C. Final Discussion

the culture of a *S. tropica salX*-deficient mutant.^[177] Similarly, the corresponding counterpart of *salX* in the *cin* gene cluster, *cinQ*, encoding a putative L-amino acid aminotransferase was inactivated leading to complete elimination of cinnabaramide production in *S. sp. JS360 cinQ*-deficient mutant. By feeding of cyclopentyl-DL-alanine to the growth medium of the *S. sp. JS360 cinQ*-deficient mutant a novel cyclopentyl-cinnabaramide analogue was generated via mutasynthesis. However, this successful incorporation of the free amino acid contrasted the hypothesis of specificity of the CinB A domain for the native substrates β -cyclohexenylalanine. A possible explanation for this would be the administrated cyclopentyl-DL-alanine exhibits a highly similar backbone to the native substrate which is tolerated by the active sites of CinB A domain. The isolated cyclopentyl cinnabaramide A retains the antifungal activity despite a significant loss in cytotoxicity (**Figure C6**).^[124] Furthermore, a series of amino acids and their corresponding synthetic amino acid *N*-acetylcysteamine (NAC) thioesters, e.g., cyclohexyl-DL-alanine, leucine, phenylalanine, and tyrosine were individually fed into the culture of the *S. sp. JS360 cinQ*-deficient mutant. High-performance liquid chromatography (HPLC) coupled mass spectroscopy (MS) data showed that only the corresponding amino acid SNAC esters were successfully incorporated into the cinnabaramide biosynthetic pathway resulting in various cinnabaramide analogues while feeding with free amino acids lead to no incorporation (**Figure C6**) (L. Huo, P. Barbie, U. Kazmaier and R. Müller, unpublished data). The successful incorporation of all thioesters confirms the assumption that the tolerant CinB A domain, as its counterpart in salinosporamide pathway, could transfer the alternative amino acid *N*-acetylcysteamine (NAC) thioesters mimicking their active CoA esters to the CinB PCP domain and the corresponding PCP bound intermediates could be recognized by the condensation domain so the biosynthesis could proceed further. This successful mutasynthesis studies on cinnabaramide NRPS pathway set a stage for us to utilize the relaxed substrate specificity of CinB A domain to generate the amino acids-altered cinnabaramide derivatives by feeding of *N*-acetylcysteamine (NAC) thioesters of proteinogenic and non-proteinogenic amino acids. Based on the fact that the generated cyclopentyl cinnabaramide A retains the antifungal activity despite a significant loss in cytotoxicity, a compound library containing various amino acids-derived cinnabaramides is supposed to be set up and be used for the structure-activity relationship (SAR) study in order to illustrate the mode of action of cinnabaramide antifungal activity which remains still unclear. The upcoming bottleneck could be the limited tolerance of CinB A domain towards desired amino acids. As discussed above, possible solutions would be rational engineering of modules/domains making them adapt to the corresponding fed building blocks.

Using precursor-directed biosynthesis (*E*)-6-chlorooct-2-enoic acid, (*E*)-8-chlorooct-2-enoic acid, and their corresponding *N*-acetylcysteamine (NAC) thioesters were unambiguously incorporated into the cinnabaramide biosynthetic pathway and resulted in production of chlorinated cinnabaramide A and other derivatives. These results demonstrate the promiscuity of the CinA AT₂ domain. The inability to generate derivatives of (*E*)-4-chlorooct-2-enoic acid or its -SNAC ester indicates that the presence of a chlorine atom at this position might cause the improper positioning of the substrate in the CinA AT₂

C. Final Discussion

active site. The alternative explanation would be that this substrate cannot be accepted by CinF for prerequisite reductive carboxylation leading to the formation of the corresponding chlorohexylmalonyl-CoA as the very substrate recognized by the CinA AT₂ domain. Further activity screening showed that the cytotoxic activities of the 15-chloro derivative were superior to those of cinnabaramide A while the 17-chloro derivative showed enhanced bioactivity only towards human myeloma RPMI-8226 cells. In general, the cytotoxic effects are still significantly weaker than those of salinosporamide A; however, the chlorinated derivatives mostly retain the cinnabaramides antifungal activity which is much stronger than salinosporamide A (**Figure C6**).

Since the key to the superior activity of salinosporamide lies in the chlorine atom attached to the side chain which is required for the irreversible binding to the target^[131], it is definitely worth further conducting the precursor-directed biosynthesis with other chlorinated enolates or corresponding –SNAC esters in order to generate various chlorinated cinnabaramides, which might give rise to a much more promising anti-cancer activity than salinosporamides, especially for feeding of the (*E*)-4-chlorooct-2-enoic acid or its -SNAC ester which could in principle lead to the production of 13-chloro cinnabaramide representing a most similar structure to salinosporamide. Given that the first attempt failed and that might be caused by intolerant CinA AT₂ domain, corresponding domain exchange or site-directed mutagenesis work is supposed to be done for adapting the AT specificity for the (*E*)-4-chlorooct-2-enoic acid. Another option would be the mutagenic study of CinF, in order to enable CinF to convert the administrated chlorinated carboxylic acids to corresponding chlorohexylmalonyl-CoAs by altering its substrate specificity. Eventually, the chlorohexylmalonyl-CoAs could be recognized and incorporated into the pathway resulting in production of novel chlorinated cinnabaramide derivatives.

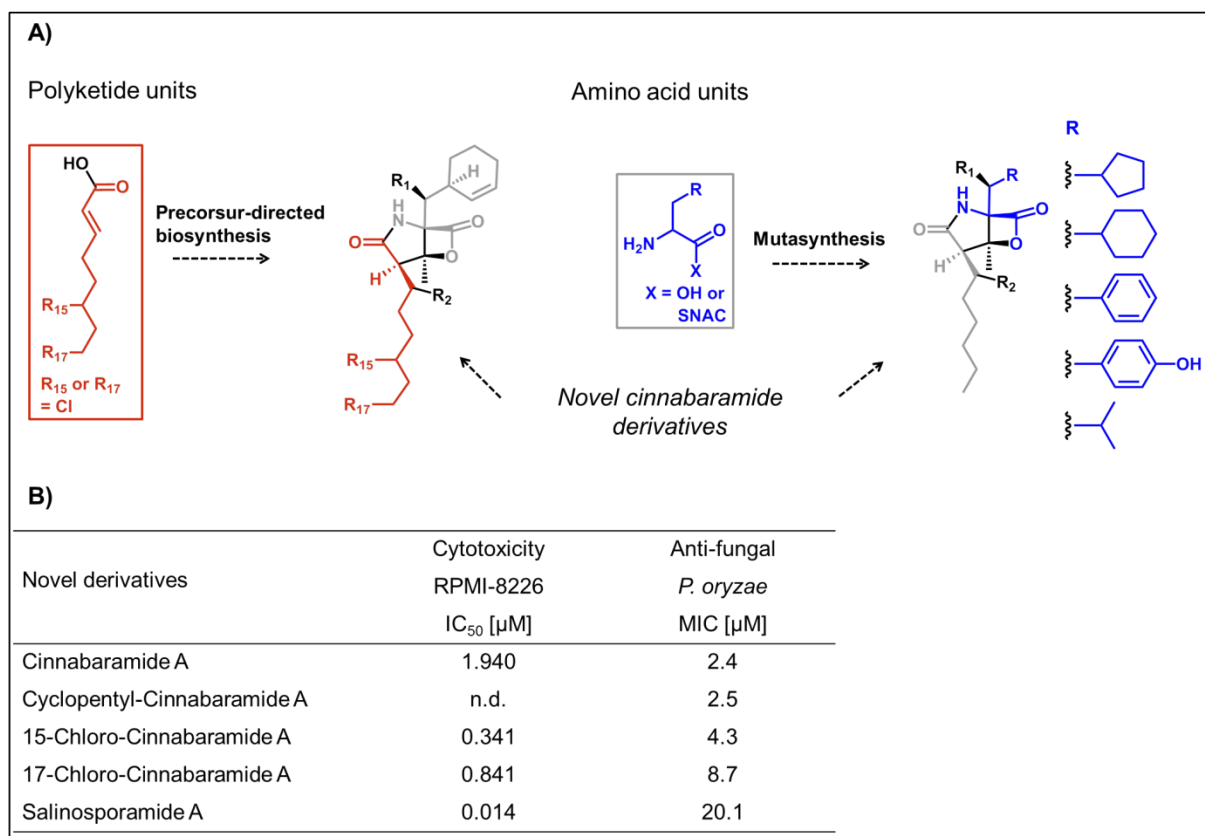


Figure C6. Rational engineering of cinnabaramide biosynthesis to generate structural diversity via synthetic biotechnology. A) Summary of generated cinnabaramide analogues. B) Biological activities of the isolated cinnabaramide derivatives.

3. Bottromycin biosynthesis from *Streptomyces* sp. BC16019

3.1 Bottromycin biosynthesis in *Streptomyces* sp. BC16019: Heavily modified peptides exhibiting unusual biosynthetic features

Bottromycins, initially isolated from the fermentation broth of *Streptomyces bottropensis* in 1957,^[126;132-135] represent a class of antibiotics exhibiting inhibitory activity on bacterial protein synthesis. Bottromycins were initially found to be active against Gram-positive bacteria and mycoplasma.^[132] Later on, bottromycins also showed their extended antibacterial abilities against methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE).^[136] Mode of mechanism studies revealed that the aminoacyl-*t*-RNA binding site (A-site) on the 50S ribosome is the target of bottromycins.^[137-139] This site is currently not targeted by clinically used antibiotics leading to no observed cross-resistance between bottromycins and other therapeutic antibiotics, which leave bottromycins as a promising class for development into novel antibiotics.^[179]

In this work we employed a terrestrial actinomycete, *Streptomyces* sp. BC16019, which is able to produce bottromycin A2 and B2 (**Figure A10**). Inspired by the promising structural features and biological activity, we first identified the biosynthetic locus of bottromycins in the genome of *Streptomyces* sp. BC16019. *In silico* analysis of the *bot* gene cluster revealed that the biosynthesis of bottromycins follows a ribosomal pathway. A 44-amino acid encoding structural gene *botA* was found flanked by a set of genes in charge of post-translational modifications. Unlike the typical RiPPs which

C. Final Discussion

harbor a N-terminal leader peptide appended to the core peptides,^[180] the BotA precursor peptide lacks the N-terminal leader sequence and only contains a single N-methionine residue instead which is experimentally verified to be cleaved off by a putative leucyl peptidase BotP (L. Huo and R. Müller, unpublished data). This unique organization for precursor peptides has not been previously described for ribosomal peptide pathways. As in some reported cases of RiPPs, BotA contains an additional C-terminal extension sequence which is supposed to be responsible for the recognition by the post-translational modification (PTM) enzymes in the absence of the N-terminal leader.^[180] Furthermore, additional tasks can be attributed to the C-terminal extension such as directing the export of mature compound, stabilizing the precursor peptide against degradation and keeping the peptide inactive during biosynthesis in the cell.^[180] A putative α/β hydrolase BotH most likely catalyzes the final proteolytic cleavage of the C-terminal extension from the precursor peptide and a number of tailoring enzymes are involved in the production of the highly modified bottromycin scaffold.

Like other RiPPs, such as thiopeptides, cyanobactins and linear azole-containing peptides,^[29] a thiazole ring also exist in bottromycins. This heterocycle is often produced by cyclodehydration of a cysteine when the thiol attacks the carbonyl carbon of the adjacent amino acid resulting in a thiazoline ring. This thiazoline ring can be subsequently oxidized to thiazole. This biosynthetic process was shown in the microcin B17 biosynthetic pathway where a three-protein complex consisting of a cyclodehydratase, a “docking protein” and a flavin mononucleotide (FMN)-dependent dehydrogenase are required for thiazole ring formation.^[181] The biochemical functionality of cyclodehydratase and flavin mononucleotide (FMN)-dependent dehydrogenase were clearly assigned as cyclodehydration of cysteine and thiazoline oxidation respectively. However, the precise function of the “docking protein” remained unclear. The hypothesis that it might play a role in regulation of the cyclodehydratase activity and initiation of an active complex was investigated by D. A. Mitchell and coworkers.^[182] A recent study on a “docking protein” belonging to the YcaO-like protein family revealed that it alone catalyzes the cyclodehydration with consumption of ATP.^[183] In the bottromycin biosynthetic pathway, two putative “YcaO-like proteins,” BotC and BotCD, were identified as possible cyclodehydratases. Because BotCD shows not only the homology to “YcaO like protein” but also to the “docking_ocin superfamily” of proteins (Accession number: TIGR03604; members of this protein family include enzymes related to SagD), which includes “docking proteins” involved in heterocycle formations in other RiPP pathways,^[184] we speculated BotCD is the enzyme responsible for the formation of thiazoline. Therefore, BotC is presumed to be involved in the macrocyclization process discussed below. After the thiazoline formation a subsequent oxidation to thiazole via a FMN-dependent dehydrogenase would be expected; however, no homologues to this family of enzymes could be found in the *bot* locus. Thus, we assumed two possible mechanisms of thiazole formation in bottromycin biosynthesis (**Figure C7**). One possibility would be an oxidative elimination occurring after the proteolysis of the C-terminal extension sequence and cyclodehydration. Another possibility would be the proteolysis of the C-terminal extension sequence and subsequent oxidative decarboxylation of the

C. Final Discussion

cysteine residue happens prior to cyclodehydration. Similar routes were reported in aminovinyl-cysteine-containing ribosomal peptide biosynthesis including formation of reactive “thio-enol” intermediates via oxidative decarboxylation of C-terminal cysteine residues^[185] which supports this hypothesis. However, these complex enzymes are characterized as homo-oligomeric flavin-containing cysteine decarboxylases but no homologous proteins could be found in the *bot* locus, leaving the underlying biochemistry of bottromycin “*exo*-thiazole” formation a mystery.

Another intriguing structural feature of bottromycins, especially for RiPPs, is the macrocycle formed between the N-terminal glycine residue and the second valine residue (**Figure A10 and C7**). According to the cyclodehydration chemistry described in textbooks the following pathway for this macrocycle formation was proposed: a nucleophile attack of the glycine amino group on the valine carbonyl carbon followed by dehydration. This macrocyclodehydration reaction mechanism results in an amidine moiety which is very unique amongst all the known RiPPs or even other natural products. To the best of our knowledge, this amidine moiety was only reported in coformycin^[186] and ectoine^[187] (also named as pyrostatin B^[188]). The studies on ectoine metabolism revealed the cyclic amidine was formed via a cyclodehydration-like reaction which supports our hypothesis. Interestingly, the ectoine synthase (EctC) responsible for amidine formation belongs to a family of carbon-oxygen lyases which is widespread in microorganisms.^[189] Not surprisingly, the EctC analogue encoding genes could also be identified in the genome of bottromycin producer *S. sp.* BC16019 as well as the heterologous host *S. coelicolor*.^[125] Thus, we cannot exclude the involvement of the ectoine biosynthetic enzyme(s) in the amidine moiety formation for bottromycin biosynthesis, although the locus of the EctC analogue encoding gene is very far away from the *bot* gene cluster. Alternatively, we assume that BotC or BotCD might catalyze the cyclodehydration reaction, whereas BotC is favorably supposed to be involved while BotCD is more likely to be responsible for the formation of thiazoline (as discussed above).

In addition to the previously discussed PTM enzymes involved in bottromycin precursor peptide maturation process, a putative amidohydrolase and putative cytochrome P450 encoding genes, *botAH* and *botCYP*, are also located in the *bot* gene cluster. Despite their unassigned biochemical functionalities, we speculated that they might be involved in either the macrocyclodehydration reaction or the *exo*-thiazole ring formation. Besides, a putative transcriptional regulator, BotR, and a putative multi-drug transporter, BotT, are also found in the *bot* locus. Additionally, the genetic engineering work based on the established heterologous expression system revealed that BotT plays a role in the self-resistance mechanism of the native producer.^[125]

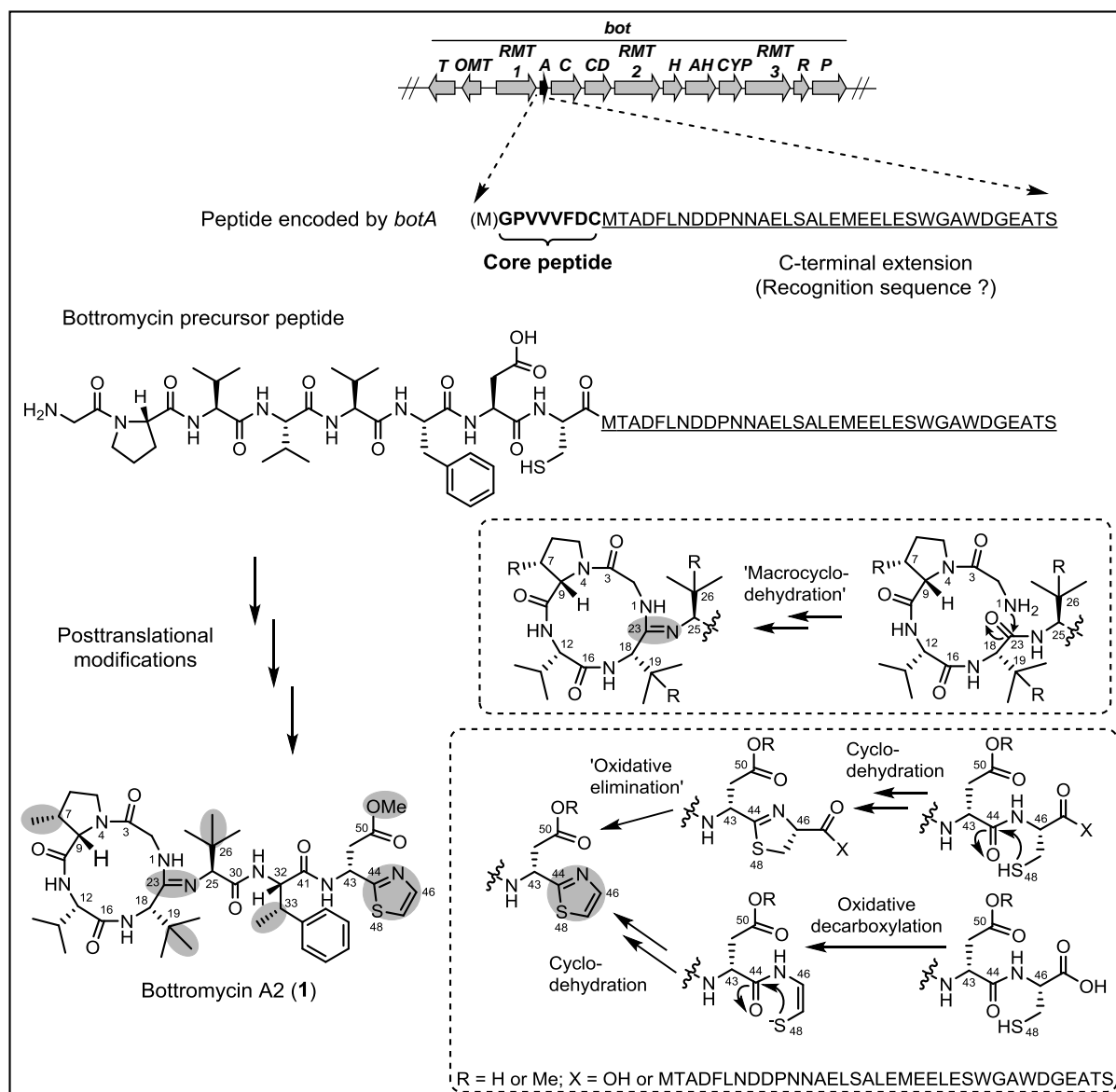


Figure C7. Maturation of the ribosomal bottromycin peptide scaffold. The bottromycin precursor peptide consists of the core peptide as well as a C-terminal extension (underlined) likely serving as a recognition sequence. Maturation of the bottromycin peptide scaffold requires a number of posttranslational modifications (highlighted in grey) the timing of which is currently unknown. These include various methylations, proteolytic digestion, 'macrocyclodehydration' and thiazole ring formation.

3.2 Engineering of the bottromycin biosynthetic pathway in a heterologous host

Faced with the tough genetic manipulation of the bottromycin native producer, we successfully engineered a heterologous expression system in order to characterize the bottromycin biosynthetic pathway. After mobilization of the proposed biosynthetic gene cluster localized on cosmid DG2 into *Streptomyces coelicolor* A3(2) and *Streptomyces albus* J1074, which were chosen as the suitable heterologous hosts because they are closely related to the native producer (see **Introduction 4.1 Heterologous expression: General aspects**), bottromycin A2 production could be detected via HPLC-MS in both *S. coelicolor*::DG2 and *S. albus*::DG2 mutants.^[125] These results indicate that the transformed biosynthetic pathway was complete and active in the heterologous hosts under the control of its native regulatory elements. However, similar to the commonly encountered bottleneck in

C. Final Discussion

heterologous expression,^[93] much lower (almost 100 times) production yields (*S.coelicolor*::DG2: ca. 4 µg/l; *S. albus*::DG2: ca. 1 µg/l) were observed compared to the native producer. Soon after, the bulky expression construct DG2 was simplified by deletion of ca. 16 kb insert fragment at the 5' end of the *bot* gene cluster using Red/ET recombineering.^[190;191] The resulting DG2-kan construct was transformed into *S. coelicolor* leading to a not significantly affected production titer compared to *S.coelicolor*::DG2. The result demonstrated that the deleted region is not involved in bottromycin biosynthesis (**Figure C8**).

Because of the high antibacterial activity of bottromycins and the fact that growth of *S. coelicolor* is impaired at bottromycin concentrations of 2 µg/l, we presumed that the limiting factor of the low production titer might lie in the self-resistance mechanism. Thus, further production optimization approaches were subsequently conducted with respect to undirected and directed strategies in order to improve the self-resistance of the heterologous hosts. The undirected approach was based on the observation that certain drug-resistant mutations in the *rpoB* gene (RNA-polymerase β -subunit encoding gene) or the *rpsL* gene (the ribosomal S12 protein encoding gene) might lead to a significant enhancement of the secondary metabolite's production in Gram-positive bacteria.^[192-196] Mutations could be triggered, e.g., by challenging the bacterial cells with rifampicin resulting in mutations in *rpoB* or by streptomycin resulting in mutations in *rpsL*. After using rifampicin mutagenesis some of the generated mutants exhibited roughly 10 times higher production titer than the starting clones (**Figure C8**).^[125] In parallel to this successful undirected approach, the putative efflux pump or multi-drug transporter (BotT) was targeted in the directed approach. In order to overexpress the *botT* gene we replaced its native promoter with the *PermE** promoter^[197] by insertion of a spectinomycin cassette via Red/ET recombineering resulting in the construct DG2-kan-efflux-ermE. The *S. coelicolor*::DG2-kan-efflux-ermE mutant produced ca. 20 times higher yield than *S. coelicolor*::DG2 (**Figure C8**).^[125] In summary, both undirected and directed mutagenetic approaches were employed to successfully improve the heterologous bottromycin production and the combination of these two strategies should lead to a further enhancement of bottromycin production titer. Additionally, the establishment of the bottromycin heterologous expression system allows for further characterization of the biosynthetic pathway in the heterologous host rather than the poorly manipulable native producer.

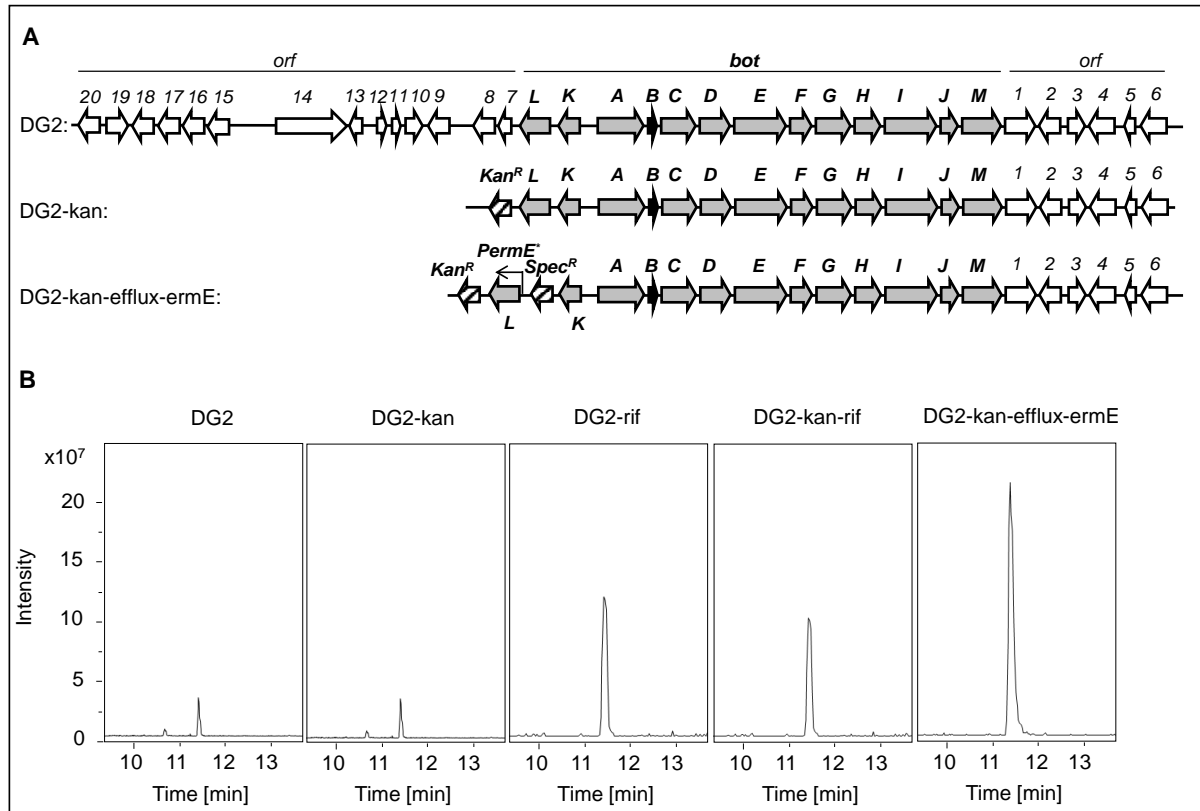


Figure C8. Heterologous bottromycin production. **A)** Inserts of the three expression constructs used for heterologous bottromycin production. The original cosmid DG2 was modified via Red/ET recombineering to delete the 5'-flanking region (*orf7-orf20*) by insertion of *kan^R* and, in the next step, to insert the *PerME** promoter upstream of *botL* together with *spec^R*. **B)** Quantification of heterologous bottromycin production by HPLC-MS analysis of the culture extracts from different *S. coelicolor* mutant strains. Sections of extracted ion chromatograms at *m/z* = 823.45 corresponding to the [M+H]⁺ ion of bottromycin A2 (1) are illustrated as representative readouts of productivity. The *S. coelicolor* A3(2) host strains contain one of the three expression constructs shown in (A). 'Rifampicin-induced mutagenesis' was performed with *S. coelicolor*::DG2 and DG2-kan ('DG2-rif' and 'DG2-kan-rif').

Using the established heterologous platform and synthetic biotechnology technique we functionally elucidated the methylation steps comprising one *O*-methyl transferase and three radical SAM-dependent methyl transferases required for alkylation of non-activated carbons. After the target deletions of the four methyl transferases were individually performed by replacement of different antibiotic resistant genes (ampicillin, chloramphenicol and zeocin) using Red/ET recombineering, the respective heterologous expression of the modified pathways were followed up. By comparing the fragmentation patterns of the detected bottromycin scaffolds the corresponding methyl transferases could be unambiguously assigned: BotRMT3 catalyzes the *C*-methylation of the proline residue, BotRMT1 is responsible for *C*-methylation of the phenylalanine residue and BotRMT2 turned out to catalyze the *C*-methylations of both valine residues. In addition, experimental evidence showed that BotOMT catalyzes the *O*-methylation of the aspartic acid residue although the corresponding non-methylated free acid product was not detected in the extracts of the corresponding mutant.^[125] It would be very intriguing if the non-methylated novel bottromycin derivatives could be isolated and screened for bioactivity which could not be accomplished in the present work due to their low production yields. However, based on the strategies mentioned above, we are absolutely convinced that the production

C. Final Discussion

titer could be easily improved and the isolation of these novel bottromycin derivatives could be achievable.

Future work would be still focused on generation of more bottromycin derivatives because the promising activity and the novel binding target making them to be potentially developed as new antibiotics. As a RiPP, despite the large biosynthetic gene cluster the structure backbone of bottromycin is actually determined by the 24 bp core peptide sequence involved in the 135 bp long *botA* gene. Such a short oligonucleotide can be easily synthesized in different versions. These *botA* variants may contain the altered core peptide sequences but the exactly same C-terminal extension sequence which can lead to various bottromycin scaffolds as well as guarantee the functionality of PTM enzymes. Combined with the established heterologous system with an improved production titer, this *in vivo* strategy could efficiently help us to generate a number of bottromycin derivatives.

Another route to generation of more bottromycin derivatives refers to the *in vitro* reconstitution of bottromycins. Based on this, different versions of BotA precursor peptides could also be easily synthesized whereas only the core peptide sequence would be altered. Until recently, the recombinant proteins of most PTM enzymes were expressed and purified in *E. coli* (L. Huo, J. Köhnke, S. C. Wenzel and R. Müller, unpublished data). The short BotA precursor peptide was also generated by chemical synthesis. According to the ambiguous annotations of these PTM enzymes, to find out how these enzymes finish the teamwork in turn, especially how the amidine-ring structure is enzymatically formed is challenging because this moiety is very rare among the RiPPs biochemically described so far. Even extended to the whole natural products, only coformycin and ectonine harbour this characteristic structural feature.^[198;199] However, several RiPPs were successfully reconstituted *in vitro*^[200-202], which is an optimistic support for this strategy. This work is definitely worth being continued in the future.

D. Reference

- [1.] G. R. Hamilton, T. F. Baskett, *Can.J.Anaesth.* **2000**, *47* 367-374.
- [2.] H. W. Florey, *Br.Med.J.* **1944**, 169-171.
- [3.] D. J. Newman, G. M. Cragg, K. M. Snader, *Nat.Prod.Rep.* **2000**, *17* 215-234.
- [4.] D. J. Newman, G. M. Cragg, *J.Nat.Prod.* **2012**, *75* 311-335.
- [5.] M. A. Fischbach, C. T. Walsh, *Science* **2009**, *325* 1089-1093.
- [6.] A. L. Demain, *J.Ind.Microbiol.Biotechnol.* **2006**, *33* 486-495.
- [7.] H. Gross, V. O. Stockwell, M. D. Henkels, B. Nowak-Thompson, J. E. Loper, W. H. Gerwick, *Chem.Biol.* **2007**, *14* 53-63.
- [8.] T. A. Gulder, B. S. Moore, *Angew.Chem.Int.Ed Engl.* **2010**, *49* 9346-9367.
- [9.] A. Harvey, *Drug Discov Today* **2000**, *5* 294-300.
- [10.] A. L. Demain, *J.Ind.Microbiol.Biotechnol.* **2006**, *33* 486-495.
- [11.] T. Kieser, M. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, The John Innes Foundation, Norwich, England **2000**, p. 613.
- [12.] J. C. Ensign, *Annu.Rev.Microbiol.* **1978**, *32* 185-219.
- [13.] M. Ramachandra, D. L. Crawford, A. L. Pometto, *Appl.Environ.Microbiol.* **1987**, *53* 2754-2760.
- [14.] D. L. Crawford, A. L. Pometto, III, *Methods Enzymol.* **1988**, *161* 249-258.
- [15.] A. J. McCarthy, S. T. Williams, *Gene* **1992**, *115* 189-192.
- [16.] T. Smith, *Nat.Struct.Biol.* **2000**, *7* 189-190.
- [17.] T. Smith, *Nat.Struct.Biol.* **2000**, *7* 189-190.
- [18.] M. G. Watve, R. Tickoo, M. M. Jog, B. D. Bhole, *Arch.Microbiol.* **2001**, *176* 386-390.
- [19.] *Pharmaceutical Biotechnology. drug discovery and clinical applications*, (Eds.: O. Kayser, R. H. Müller) WILEY-VCH, Weinheim **2008**.
- [20.] H. Reichenbach, G. Höfle, in *Drug Discovery from Nature* Eds.: S. Grabley, R. Thiericke), Springer, Berlin **1999**, p. pp. 149-179.
- [21.] G. Höfle, N. Bedorf, H. Steinmetz, D. Schomburg, K. Gerth, H. Reichenbach, *Angew.Chem.Int.Ed.* **1996**, *35* 1567-1569.
- [22.] M. N. Fornier, *Clinical Breast Cancer* **2007**, *7* 757-763.
- [23.] D. Schwarzer, M. A. Marahiel, *Naturwissenschaften* **2001**, *88* 93-101.
- [24.] C. T. Walsh, *Science* **2004**, *303* 1805-1810.
- [25.] M. A. Fischbach, C. T. Walsh, *Chem.Rev.* **2006**, *106* 3468-3496.
- [26.] R. Finking, M. A. Marahiel, *Annual Review of Microbiology* **2004**, *58* 453-488.
- [27.] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem.Biol.* **1996**, *3* 923-936.
- [28.] X. Yang, W. A. van der Donk, *Chemistry.* **2013**, *19* 7662-7677.
- [29.] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H. G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. E. Süßmuth, J. R. Tagg, G. L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, W. van der Donk, *Nat.Prod.Rep.* **2013**, *30* 108-160.
- [30.] F. Malpartida, D. A. Hopwood, *Nature* **1984**, *309* 462-464.
- [31.] J. Staunton, K. J. Weissman, *Nat.Prod.Rep.* **2001**, *18* 380-416.
- [32.] A. Sandmann, F. Sasse, R. Müller, *Chem.Biol.* **2004**, *11* 1071-1079.
- [33.] M. Kopp, H. Irschik, S. Pradella, R. Müller, *ChemBioChem* **2005**, *6* 1277-1286.
- [34.] O. Perlova, K. Gerth, A. Hans, O. Kaiser, R. Müller, *J.Biotechnol.* **2006**, *121* 174-191.
- [35.] K. Buntin, H. Irschik, K. J. Weissman, E. Luxenburger, H. Blöcker, R. Müller, *Chem.Biol.* **2010**, *17* 342-356.
- [36.] R. Dieckmann, Y. O. Lee, L. H. van, D. H. von, H. Kleinkauf, *FEBS Lett.* **1995**, *357* 212-216.
- [37.] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16* 4174-4183.
- [38.] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem.Biol.* **1999**, *6* 493-505.
- [39.] A. Schneider, M. A. Marahiel, *Arch.Microbiol.* **1998**, *169* 404-410.
- [40.] J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407* 215-218.

References

- [41.] A. Haese, M. Schubert, M. Herrmann, R. Zocher, *Mol.Microbiol.* **1993**, 7 905-914.
- [42.] A. Stindl, U. Keller, *Biochemistry* **1994**, 33 9358-9364.
- [43.] T. Stachelhaus, C. T. Walsh, *Biochemistry-US* **2000**, 39 5775-5787.
- [44.] L. E. Quadri, T. A. Keating, H. M. Patel, C. T. Walsh, *Biochemistry* **1999**, 38 14941-14954.
- [45.] Z. Suo, C. T. Walsh, D. A. Miller, *Biochemistry* **1999**, 38 17000.
- [46.] L. Du, M. Chen, C. Sanchez, B. Shen, *FEMS Microbiol.Lett.* **2000**, 189 171-175.
- [47.] C. Bisang, P. F. Long, J. Cortes, J. Westcott, J. Crosby, A. L. Matharu, R. J. Cox, T. J. Simpson, J. Staunton, P. F. Leadlay, *Nature* **1999**, 401 502-505.
- [48.] J. Ligon, S. Hill, J. Beck, R. Zirkle, I. Molnar, J. Zawodny, S. Money, T. Schupp, *Gene* **2002**, 285 257-267.
- [49.] B. Silakowski, G. Nordsiek, B. Kunze, H. Blocker, R. Muller, *Chem.Biol.* **2001**, 8 59-69.
- [50.] H. Ikeda, T. Nonomiya, S. Omura, *J.Ind.Microbiol.Biotechnol.* **2001**, 27 170-176.
- [51.] B. J. Carroll, S. J. Moss, L. Bai, Y. Kato, S. Toelzer, T. W. Yu, H. G. Floss, *J.Am.Chem.Soc.* **2002**, 124 4176-4177.
- [52.] Y. A. Chan, M. T. Boyne, A. M. Podevels, A. K. Klimowicz, J. Handelsman, N. L. Kelleher, M. G. Thomas, *Proc.Natl.Acad.Sci.U.S.A* **2006**, 103 14349-14354.
- [53.] S. C. Wenzel, R. M. Williamson, C. Grünanger, J. Xu, K. Gerth, R. A. Martinez, S. J. Moss, B. J. Carroll, S. Grond, C. J. Unkefer, R. Müller, H. G. Floss, *J.Am.Chem.Soc.* **2006**, 128 14325-14336.
- [54.] C. D. Reeves, L. M. Chung, Y. Liu, Q. Xue, J. R. Carney, W. P. Revill, L. Katz, *J.Biol.Chem.* **2002**, 277 9155-9159.
- [55.] A. S. Eustaquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore, *Proc.Natl.Acad.Sci.U.S.A* **2009**, 106 12295-12300.
- [56.] L. Katz, in *Chem RevEd.*: D. E. Cane), American Chemical Society, Columbus, OH **1997**, p. pp. 2557-2575.
- [57.] C. Hertweck, *Angew.Chem Int.Ed Engl.* **2009**, 48 4688-4716.
- [58.] M. B. Austin, J. P. Noel, *Nat.Prod.Rep.* **2003**, 20 79-110.
- [59.] J. Schröder, *Trends Plant Sci.* **1997**, 2 373-378.
- [60.] L. Du, Y. Q. Cheng, G. Ingenhorst, G. L. Tang, Y. Huang, B. Shen, *Genet.Eng (N.Y.)* **2003**, 25 227-267.
- [61.] L. Du, C. Sanchez, B. Shen, *Metab Eng* **2001**, 3 78-95.
- [62.] C. Bender, D. Palmer, A. Penaloza-Vazquez, V. Rangaswamy, M. Ullrich, *Arch.Microbiol.* **1996**, 166 71-75.
- [63.] R. J. Parry, S. V. Mhaskar, M. T. Lin, A. E. Walker, R. Mafoti, *Can J Chem* **1994**, 72 86-99.
- [64.] V. Rangaswamy, M. Ullrich, W. Jones, R. Mitchell, R. Parry, P. Reynolds, C. L. Bender, *FEMS Microbiol.Lett.* **1997**, 154 65-72.
- [65.] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blöcker, G. Höfle, S. Beyer, R. Müller, *J.Biol.Chem.* **1999**, 274 37391-37399.
- [66.] B. Silakowski, G. Nordsiek, B. Kunze, H. Blöcker, R. Müller, *Chemistry & Biology* **2001**, 8 59-69.
- [67.] S. Weinig, H. J. Hecht, T. Mahmud, R. Müller, *Chem.Biol.* **2003**, 10 939-952.
- [68.] Y. Paitan, G. Alon, E. Orr, E. Z. Ron, E. Rosenberg, *J.Mol.Biol.* **1999**, 286 465-474.
- [69.] N. Gaitatzis, A. Hans, R. Müller, S. Beyer, *J.Biochem.(Tokyo)* **2001**, 129 119-124.
- [70.] L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry-US* **1998**, 37 1585-1595.
- [71.] C. Sanchez, L. Du, D. J. Edwards, M. D. Toney, B. Shen, *Chem.Biol.* **2001**, 8 725-738.
- [72.] T. J. Oman, W. A. van der Donk, *Nat.Chem.Biol.* **2010**, 6 9-18.
- [73.] E. W. Schmidt, J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood, J. Ravel, *Proc.Natl.Acad.Sci.U.S.A* **2005**, 102 7315-7320.
- [74.] H. E. Hallen, H. Luo, J. S. Scott-Craig, J. D. Walton, *Proc.Natl.Acad.Sci.U.S.A* **2007**, 104 19097-19101.
- [75.] M. Trabi, J. S. Mylne, L. Sando, D. J. Craik, *Org.Biomol.Chem.* **2009**, 7 2378-2388.
- [76.] W. L. Kelly, L. Pan, C. Li, *J.Am.Chem Soc.* **2009**.
- [77.] M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, H. G. Sahl, S. Matsunaga, J. Piel, *Science* **2012**, 338 387-390.
- [78.] L. C. Wieland Brown, M. G. Acker, J. Clardy, C. T. Walsh, M. A. Fischbach, *Proc.Natl.Acad.Sci.U.S.A* **2009**, 106 2549-2553.
- [79.] R. Liao, L. Duan, C. Lei, H. Pan, Y. Ding, Q. Zhang, D. Chen, B. Shen, Y. Yu, W. Liu, *Chem.Biol.* **2009**, 16 141-147.
- [80.] R. P. Morris, J. A. Leeds, H. U. Naegeli, L. Oberer, K. Memmert, E. Weber, M. J. LaMarche, C. N. Parker, N. Burrer, S. Esterow, A. E. Hein, E. K. Schmitt, P. Krastel, *J.Am.Chem.Soc.* **2009**, 131 5946-5955.

References

- [81.] C. T. Walsh, S. J. Malcolmson, T. S. Young, *ACS Chem.Biol.* **2012**, 7 429-442.
- [82.] Y. M. Li, J. C. Milne, L. L. Madison, R. Kolter, C. T. Walsh, *Science* **1996**, 274 1188-1193.
- [83.] J. O. Melby, N. J. Nard, D. A. Mitchell, *Curr.Opin.Chem.Biol.* **2011**, 15 369-378.
- [84.] H. Onaka, M. Nakaho, K. Hayashi, Y. Igarashi, T. Furumai, *Microbiology* **2005**, 151 3923-3933.
- [85.] L. C. Wieland Brown, M. G. Acker, J. Clardy, C. T. Walsh, M. A. Fischbach, *Proc.Natl.Acad.Sci.U.S.A* **2009**, 106 2549-2553.
- [86.] J. Lee, J. McIntosh, B. J. Hathaway, E. W. Schmidt, *J.Am.Chem.Soc.* **2009**, 131 2122-2124.
- [87.] J. D. Walton, H. E. Hallen-Adams, H. Luo, *Biopolymers* **2010**, 94 659-664.
- [88.] M. J. van Belkum, L. A. Martin-Visscher, J. C. Vederas, *Trends Microbiol.* **2011**, 19 411-418.
- [89.] S. Gunasekera, N. L. Daly, M. A. Anderson, D. J. Craik, *IUBMB.Life* **2006**, 58 515-524.
- [90.] P. J. Knerr, W. A. van der Donk, *Annu.Rev.Biochem.* **2012**, 81 479-505.
- [91.] M. S. Butler, *J.Nat.Prod.* **2004**, 67 2141-2153.
- [92.] S. C. Wenzel, F. Gross, Y. Zhang, J. Fu, F. A. Stewart, R. Müller, *Chem.Biol.* **2005**, 12 349-356.
- [93.] S. Ongley, X. Bian, B. A. Neilan, R. Müller, *Nat.Prod.Rep.* **2013**, 30 1121-1138.
- [94.] H. B. Bode, R. Müller, *Angew.Chem.Int.Ed.* **2005**, 44 6828-6846.
- [95.] S. C. Wenzel, R. Müller, *Curr.Opin.Biotechnol.* **2005**, 16 594-606.
- [96.] J. P. P. Muylers, Y. M. Zhang, F. Buchholz, A. F. Stewart, *Genes Dev.* **2000**, 14 1971-1982.
- [97.] Y. Zhang, J. P. P. Muylers, G. Testa, A. F. Stewart, *Nat.Biotechnol.* **2000**, 18 1314-1317.
- [98.] J. P. P. Muylers, Y. Zhang, V. Benes, G. Testa, W. Ansorge, A. F. Stewart, *EMBO Rep.* **2000**, 1 239-243.
- [99.] J. Fu, X. Bian, S. Hu, H. Wang, F. Huang, P. M. Seibert, A. Plaza, L. Xia, R. Müller, A. F. Stewart, Y. Zhang, *Nat.Biotechnol.* **2012**, 30 440-446.
- [100.] J. P. P. Muylers, Y. Zhang, A. F. Stewart, *Trends Biochem.Sci.* **2001**, 26 325-331.
- [101.] O. Perlova, J. Fu, S. Kuhlmann, D. Krug, F. Stewart, Y. Zhang, R. Müller, *Appl.Environ.Microbiol.* **2006**, 72 7485-7494.
- [102.] F. Gross, M. W. Ring, O. Perlova, J. Fu, S. Schneider, K. Gerth, S. Kuhlmann, A. F. Stewart, Y. Zhang, R. Müller, *Chem.Biol.* **2006**, 13 1253-1264.
- [103.] Y. F. Hu, V. V. Phelan, C. M. Farnet, E. Zazopoulos, B. O. Bachmann, *ChemBioChem* **2008**, 9 1603-1608.
- [104.] J. Fu, S. C. Wenzel, O. Perlova, J. Wang, F. Gross, Z. Tang, Y. Yin, A. F. Stewart, R. Müller, Y. Zhang, *Nucleic Acids Res.* **2008**, 36 e113.
- [105.] X. Bian, F. Huang, F. A. Stewart, L. Xia, Y. Zhang, R. Müller, *ChemBioChem* **2012**, 13 1946-1952.
- [106.] Bian, X., **2012**.
- [107.] R. J. Goss, S. Shankar, A. A. Fayad, *Nat.Prod.Rep.* **2012**, 29 870-889.
- [108.] M. Nett, H. Ikeda, B. S. Moore, *Nat.Prod.Rep.* **2009**, 26 1362-1384.
- [109.] S. C. Wenzel, R. Müller, *Nat.Prod.Rep.* **2009**, 26 1385-1407.
- [110.] A. S. Ratnayake, J. E. Janso, X. Feng, G. Schlingmann, I. Goljer, G. T. Carter, *J.Nat.Prod.* **2009**, 72 496-499.
- [111.] W. T. Shier, K. L. Rinehart, Jr., D. Gottlieb, *Proc.Natl.Acad.Sci.U.S.A* **1969**, 63 198-204.
- [112.] R. G. Ankenbauer, A. L. Staley, K. L. Rinehart, C. D. Cox, *Proc.Natl.Acad.Sci.U.S.A* **1991**, 88 1878-1882.
- [113.] C. Bormann, A. Kalmanczhelyi, R. Sussmuth, G. Jung, *J.Antibiot.(Tokyo)* **1999**, 52 102-108.
- [114.] S. Weist, B. Bister, O. Puk, D. Bischoff, S. Pelzer, G. J. Nicholson, W. Wohlleben, G. Jung, R. D. Sussmuth, *Angew.Chem.Int.Ed Engl.* **2002**, 41 3383-3385.
- [115.] S. Weist, C. Kittel, D. Bischoff, B. Bister, V. Pfeifer, G. J. Nicholson, W. Wohlleben, R. D. Sussmuth, *J.Am.Chem.Soc.* **2004**, 126 5942-5943.
- [116.] M. A. Gregory, H. Petkovic, R. E. Lill, S. J. Moss, B. Wilkinson, S. Gaisser, P. F. Leadlay, R. M. Sheridan, *Angew.Chem.Int.Ed Engl.* **2005**, 44 4757-4760.
- [117.] K. J. Weissman, *Trends Biotechnol.* **2007**, 25 139-142.
- [118.] P. G. Williams, G. O. Buchanan, R. H. Feling, C. A. Kauffman, P. R. Jensen, W. Fenical, *J.Org.Chem.* **2005**, 70 6196-6203.
- [119.] K. A. Reed, R. R. Manam, S. S. Mitchell, J. Xu, S. Teisan, T. H. Chao, G. Yanat-Yazdi, S. T. Neuteboom, K. S. Lam, B. C. Potts, *J.Nat.Prod.* **2007**, 70 269-276.
- [120.] M. Nett, T. A. Gulder, A. J. Kale, C. C. Hughes, B. S. Moore, *J.Med.Chem.* **2009**, 52 6163-6167.
- [121.] A. S. Eustaquio, B. S. Moore, *Angew.Chem.Int.Ed.* **2008**, 47 3936-3938.
- [122.] Y. Liu, C. Hazzard, A. S. Eustaquio, K. A. Reynolds, B. S. Moore, *J.Am.Chem.Soc.* **2009**, 131 10376-10377.
- [123.] M. Stadler, J. Bitzer, A. Mayer-Bartschmid, H. Muller, J. et-Buchholz, F. Gantner, H. V. Tichy, P. Reinemer, K. B. Bacon, *J.Nat.Prod.* **2007**, 70 246-252.
- [124.] S. Rachid, L. Huo, J. Herrmann, M. Stadler, B. Köpcke, J. Bitzer, R. Müller, *ChemBioChem* **2011**, 12 922-931.

References

- [125.] L. Huo, S. Rachid, M. Stadler, S. C. Wenzel, R. Müller, *Chemistry & Biology* **2012**, *19* 1278-1287.
- [126.] J. M. Waisvisz, M. G. van der Hoeven, B. te Nijnhuis, *J.Am.Chem.Soc.* **1957**, *79* 4524-4527.
- [127.] K. S. Ahn, G. Sethi, T. H. Chao, S. T. Neuteboom, M. M. Chaturvedi, M. A. Palladino, A. Younes, B. B. Aggarwal, *Blood* **2007**, *110* 2286-2295.
- [128.] R. H. Feling, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen, W. Fenical, *Angew.Chem.Int.Ed.* **2003**, *42* 355-357.
- [129.] W. Fenical, P. R. Jensen, M. A. Palladino, K. S. Lam, G. K. Lloyd, B. C. Potts, *Bioorg.Med.Chem.* **2009**, *17* 2175-2180.
- [130.] B. S. Moore, A. S. Eustaquio, R. P. McGlinchey, *Curr.Opin.Chem.Biol.* **2008**, *12* 434-440.
- [131.] M. Groll, R. Huber, B. C. Potts, *J.Am.Chem.Soc.* **2006**, *128* 5136-5141.
- [132.] N. Tanaka, NISHIMUR.T, S. Nakamura, H. Umezawa, T. Hayami, *J.Antibiot.* **1968**, *21* 75-76.
- [133.] J. M. Waisvisz, M. G. van der Hoeven, J. van Peppen, W. C. M. Zwennis, *J.Am.Chem.Soc.* **1957**, *79* 4520-4521.
- [134.] J. M. Waisvisz, M. G. van der Hoeven, J. F. Hölscher, B. te Nijnhuis, *J.Am.Chem.Soc.* **1957**, *79* 4522-4524.
- [135.] J. M. Waisvisz, M. G. van der Hoeven, *J.Am.Chem.Soc.* **1958**, *80* 383-385.
- [136.] H. Shimamura, H. Gouda, K. Nagai, T. Hirose, M. Ichioka, Y. Furuya, Y. Kobayashi, S. Hirono, T. Sunazuka, S. Omura, *Angewandte Chemie-International Edition* **2009**, *48* 914-917.
- [137.] T. Otaka, A. Kaji, *J.Biol.Chem.* **1976**, *251* 2299-2306.
- [138.] T. Otaka, A. Kaji, *FEBS Lett.* **1981**, *123* 173-176.
- [139.] T. Otaka, A. Kaji, *FEBS Lett.* **1983**, *153* 53-59.
- [140.] R. P. McGlinchey, M. Nett, A. S. Eustaquio, R. N. Asolkar, W. Fenical, B. S. Moore, *J.Am.Chem.Soc.* **2008**, *130* 7822-7823.
- [141.] R. H. Feling, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen, W. Fenical, *Angew.Chem.Int.Ed Engl.* **2003**, *42* 355-357.
- [142.] L. L. Beer, B. S. Moore, *Org.Lett.* **2007**, *9* 845-848.
- [143.] T. A. M. Gulder, B. S. Moore, *Angewandte Chemie-International Edition* **2010**, *49* 9346-9367.
- [144.] A. J. Kale, R. P. McGlinchey, B. S. Moore, *J.Biol.Chem.* **2010**, *285* 33710-33717.
- [145.] N. Quade, L. Huo, S. Rachid, D. W. Heinz, R. Müller, *Nat Chem Biol Nat Chem Biol* **2011**, *8* 117-124.
- [146.] A. S. Eustaquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore, *Proc.Natl.Acad.Sci.U.S.A* **2009**, *106* 12295-12300.
- [147.] N. Gaitatzis, B. Silakowski, B. Kunze, G. Nordsiek, H. Blocker, G. Hofle, R. Muller, *J.Biol.Chem.* **2002**, *277* 13082-13090.
- [148.] B. Frank, S. C. Wenzel, H. B. Bode, M. Scharfe, H. Blocker, R. Muller, *J.Mol.Biol.* **2007**, *374* 24-38.
- [149.] H. Irschik, H. Reichenbach, G. Höfle, R. Jansen, *J.Antibiot.* **2007**, *60* 733-738.
- [150.] T. J. Erb, V. Brecht, G. Fuchs, M. Müller, B. E. Alber, *Proc.Natl.Acad.Sci.U.S.A* **2009**, *106* 8871-8876.
- [151.] Y. Liu, C. Hazzard, A. S. Eustaquio, K. A. Reynolds, B. S. Moore, *J.Am.Chem.Soc.* **2009**, *131* 10376-10377.
- [152.] D. W. Udvary, L. Zeigler, R. N. Asolkar, V. Singan, A. Lapidus, W. Fenical, P. R. Jensen, B. S. Moore, *Proc.Natl.Acad.Sci.U.S.A* **2007**, *104* 10376-10381.
- [153.] S. A. Mahlstedt, C. T. Walsh, *Biochemistry* **2010**, *49* 912-923.
- [154.] M. Nett, T. A. Gulder, A. J. Kale, C. C. Hughes, B. S. Moore, *J.Med.Chem.* **2009**, *52* 6163-6167.
- [155.] L. Du, B. Shen, *Curr Opin Drug Discov Devel* **2001**, *4* 215-228.
- [156.] L. H. Du, C. Sanchez, B. Shen, *Metab.Eng.* **2001**, *3* 78-95.
- [157.] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *Proc.Natl.Acad.Sci.U.S.A* **2000**, *97* 5848-5853.
- [158.] K. Patel, M. Piagentini, A. Rascher, Z. Q. Tian, G. O. Buchanan, R. Regentin, Z. Hu, C. R. Hutchinson, R. McDaniel, *Chem.Biol.* **2004**, *11* 1625-1633.
- [159.] H. Petkovic, R. E. Lill, R. M. Sheridan, B. Wilkinson, E. L. McCormick, H. A. McArthur, J. Staunton, P. F. Leadlay, S. G. Kendrew, *J.Antibiot.(Tokyo)* **2003**, *56* 543-551.
- [160.] T. Stachelhaus, A. Schneider, M. A. Marahiel, *Science* **1995**, *269* 69-72.
- [161.] K. T. Nguyen, D. Ritz, J. Q. Gu, D. Alexander, M. Chu, V. Miao, P. Brian, R. H. Baltz, *Proc.Natl.Acad.Sci.U.S.A* **2006**, *103* 17462-17467.
- [162.] K. T. Nguyen, X. He, D. C. Alexander, C. Li, J. Q. Gu, C. Mascio, P. A. Van, L. Mortin, M. Chu, J. A. Silverman, P. Brian, R. H. Baltz, *Antimicrob.Agents Chemother.* **2010**, *54* 1404-1413.
- [163.] V. Miao, M. F. Coeffet-Le Gal, K. Nguyen, P. Brian, J. Penn, A. Whiting, J. Steele, D. Kau, S. Martin, R. Ford, T. Gibson, M. Bouchard, S. K. Wrigley, R. H. Baltz, *Chem.Biol.* **2006**, *13* 269-276.
- [164.] A. Schneider, T. Stachelhaus, M. A. Marahiel, *Mol.Gen.Genet.* **1998**, *257* 308-318.

References

- [165.] C. J. Rowe, I. U. Bohm, I. P. Thomas, B. Wilkinson, B. A. Rudd, G. Foster, A. P. Blackaby, P. J. Sidebottom, Y. Roddis, A. D. Buss, J. Staunton, P. F. Leadlay, *Chem.Biol.* **2001**, *8* 475-485.
- [166.] H. D. Mootz, N. Kessler, U. Linne, K. Eppelmann, D. Schwarzer, M. A. Marahiel, *J.Am.Chem.Soc.* **2002**, *124* 10980-10981.
- [167.] C. J. Martin, M. C. Timoney, R. M. Sheridan, S. G. Kendrew, B. Wilkinson, J. C. Staunton, P. F. Leadlay, *Org.Biomol.Chem.* **2003**, *1* 4144-4147.
- [168.] S. C. Wenzel, R. M. Williamson, C. Grunanger, J. Xu, K. Gerth, R. A. Martinez, S. J. Moss, B. J. Carroll, S. Grond, C. J. Unkefer, R. Muller, H. G. Floss, *J.Am.Chem.Soc.* **2006**, *128* 14325-14336.
- [169.] P. C. Dorrestein, S. G. Van Lanen, W. Li, C. Zhao, Z. Deng, B. Shen, N. L. Kelleher, *J.Am.Chem.Soc.* **2006**, *128* 10386-10387.
- [170.] Y. A. Chan, M. T. Boyne, A. M. Podevels, A. K. Klimowicz, J. Handelsman, N. L. Kelleher, M. G. Thomas, *Proc.Natl.Acad.Sci.U.S.A* **2006**, *103* 14349-14354.
- [171.] L. L. Beer, B. S. Moore, *Organic Letters* **2007**, *9* 845-848.
- [172.] M. Kopp, H. Irschik, K. Gemperlein, K. Buntin, P. Meiser, K. J. Weissman, H. B. Bode, R. Müller, *Molecular Biosystems* **2011**, *7* 1549-1563.
- [173.] P. C. Dorrestein, S. G. Van Lanen, W. Li, C. Zhao, Z. Deng, B. Shen, N. L. Kelleher, *J.Am.Chem.Soc.* **2006**, *128* 10386-10387.
- [174.] Y. A. Chan, M. T. Boyne, A. M. Podevels, A. K. Klimowicz, J. Handelsman, N. L. Kelleher, M. G. Thomas, *P.Natl.Acad.Sci.USA* **2006**, *103* 14349-14354.
- [175.] T. J. Erb, I. A. Berg, V. Brecht, M. Muller, G. Fuchs, B. E. Alber, *P.Natl.Acad.Sci.USA* **2007**, *104* 10631-10636.
- [176.] R. G. Rosenthal, M. O. Ebert, P. Kiefer, D. M. Peter, J. A. Vorholt, T. J. Erb, *Nat.Chem.Biol.* **2014**, *10* 50-55.
- [177.] M. Nett, T. A. Gulder, A. J. Kale, C. C. Hughes, B. S. Moore, *J.Med.Chem.* **2009**, *52* 6163-6167.
- [178.] P. Barbie, L. Huo, R. Müller, U. Kazmaier, *Organic Letters* **2012**, *14* 6064-6067.
- [179.] H. Gouda, Y. Kobayashi, T. Yamada, T. Idegushi, A. Sugawara, T. Hirose, S. Omura, T. Sunazuka, S. Hirono, *Chem Pharm Bull Chem Pharm Bull* **2012**, *60* 169-171.
- [180.] T. J. Oman, W. A. van der Donk, *Nat.Chem.Biol.* **2010**, *6* 9-18.
- [181.] Y. M. Li, J. C. Milne, L. L. Madison, R. Kolter, C. T. Walsh, *Science* **1996**, *274* 1188-1193.
- [182.] J. O. Melby, N. J. Nard, D. A. Mitchell, *Curr.Opin.Chem.Biol.* **2011**, *15* 369-378.
- [183.] K. L. Dunbar, J. O. Melby, D. A. Mitchell, *Nat Chem Biol Nat Chem Biol* **2012**, *8* 569-575.
- [184.] S. W. Lee, D. A. Mitchell, A. L. Markley, M. E. Hensler, D. Gonzalez, A. Wohlrab, P. C. Dorrestein, V. Nizet, J. E. Dixon, *P.Natl.Acad.Sci.USA* **2008**, *105* 5879-5884.
- [185.] C. S. Sit, S. Yoganathan, J. C. Vederas, *Acc.Chem.Res.* **2011**, *44* 261-268.
- [186.] H. Nakamura, G. Koyama, Y. Iitaka, M. Ono, N. Yagiawa, *J.Am.Chem.Soc.* **1974**, *96* 4327-4328.
- [187.] L. Inbar, A. Lapidot, *J.Biol.Chem.* **1988**, *263* 16014-16022.
- [188.] L. Castellanos, C. Duque, S. Zea, A. Espada, J. Rodriguez, C. Jimenez, *Organic Letters* **2006**, *8* 4967-4970.
- [189.] J. M. Pastor, M. Salvador, M. Argandona, V. Bernal, M. Reina-Bueno, L. N. Csonka, J. L. Iborra, C. Vargas, J. J. Nieto, M. Canovas, *biotechnology advances* **2010**, *28* 782-801.
- [190.] Y. Zhang, F. Buchholz, J. P. Muyrers, A. F. Stewart, *Nat.Genet.* **1998**, *20* 123-128.
- [191.] Y. Zhang, J. P. Muyrers, G. Testa, A. F. Stewart, *Nat.Biotechnol.* **2000**, *18* 1314-1317.
- [192.] H. Hu, Q. Zhang, K. Ochi, *J.Bacteriol.* **2002**, *184* 3984-3991.
- [193.] T. Inaoka, K. Takahashi, H. Yada, M. Yoshida, K. Ochi, *J.Biol.Chem.* **2004**, *279* 3885-3892.
- [194.] J. Shima, A. Hesketh, S. Okamoto, S. Kawamoto, K. Ochi, *J.Bacteriol.* **1996**, *178* 7276-7284.
- [195.] H. F. Hu, K. Ochi, *Appl.Environ.Microbiol.* **2001**, *67* 1885-1892.
- [196.] N. Tamehiro, T. Hosaka, J. Xu, H. F. Hu, N. Otake, K. Ochi, *Appl.Environ.Microbiol.* **2003**, *69* 6412-6417.
- [197.] M. J. Bibb, J. White, J. M. Ward, G. R. Janssen, *Mol.Microbiol.* **1994**, *14* 533-545.
- [198.] H. Nakamura, G. Koyama, Y. Iitaka, M. Ono, N. Yagiawa, *J.Am.Chem.Soc.* **1974**, *96* 4327-4328.
- [199.] L. Inbar, A. Lapidot, *J.Biol.Chem.* **1988**, *263* 16014-16022.
- [200.] J. C. Milne, R. S. Roy, A. C. Eliot, N. L. Kelleher, A. Wokhlu, B. Nickels, C. T. Walsh, *Biochemistry* **1999**, *38* 4768-4781.
- [201.] L. A. Furgerson Ihnken, C. Chatterjee, W. A. van der Donk, *Biochemistry* **2008**, *47* 7352-7363.
- [202.] L. Xie, L. M. Miller, C. Chatterjee, O. Averin, N. L. Kelleher, W. A. van der Donk, *Science* **2004**, *303* 679-681.

E. Appendix

Author's efforts contributed to the publications presented in this work

Publication I:

Mining the cinnabaramide Biosynthetic pathway to Generate Novel Proteasome Inhibitors

S.R. and L.H. performed the cloning and characterization of the cinnabaramide biosynthetic gene cluster and *in silico* analysis of the pathway. S.R. and L.H. performed the feeding experiments. J.H. performed the bioactivity experiments. B.K. and J.B. isolated the novel cinnabaramide derivatives and elucidated the structures in M.S.'s lab. R.M. and M.S. designed the study. S.R., L.H., J.H., M.S. and R.M. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Publication II:

Unusual carbon fixation gives rise to diverse polyketide extender unit

N.Q. performed protein purification, crystallization, crystallographic data collection and structure determination. L.H. performed gene cloning, protein heterologous expression and purification experiments as well as *in vitro* characterization of proteins and generation of corresponding point mutations. S.R. contributed initial advice regarding cloning, protein expression and purification. R.M. and D.W.H. designed the study, and R.M. wrote the manuscript together with N.Q., L.H. and D.W.H. All authors discussed the results and commented on the manuscript.

Publication III:

Stereoselective Synthesis of Deuterium-Labeled (2S)-Cyclohexenyl Alanines, Biosynthetic Intermediates of Cinnabaramides

P.B. performed the stereoselective synthesis of dideuterated β -cyclohexenylalanines. L.H. performed the feeding experiment and corresponding HPLC-MS analyses. R.M. and U.K. designed the study. P.B., R.M. and U.K. wrote the manuscript with help of L.H.. All authors discussed the results and commented on the manuscript.

Publication IV:

Synthetic Biotechnology to Study and Engineer Ribosomal Bottromycin Biosynthesis

S.R. and L.H. identified the bottromycin biosynthetic origin, established and improved the heterologous production. L.H. and S.C.W. performed the *in silico* analysis of the bottromycin biosynthetic pathway. L.H. performed the generation of analogs by pathway genetic engineering and characterized three radical SAM methyltransferase-encoding genes. S.C.W. and R.M. designed the study. S.C.W., M.S. and R.M. wrote the manuscript with help of L.H.. All authors discussed the results and commented on the manuscript.

Lebenslauf

Name Liujie Huo

Geburtsdatum 29.08.1986

Geburtsort Chongqing, China.

Nationalität China

Schulbildung:

1992 - 1998 Grundschule in Chengdu, China.

1998 - 2001 Unterstufe der Mittelschule, Chengdu, China.

2001 - 2004 Oberstufe der Mittelschule, Chengdu, China.

Hochschulausbildung:

09. 2004 - 06. 2008 Bachelor of Engineering (B. Eng.), Tianjin University, China.

09. 2008 - 09. 2010 Master of Science (M.Sc.) in Biotechnology, Universität des Saarland, Germany.

01. 2010 - 09. 2010: Masterarbeit am Institut für pharmazeutische Biotechnologie unter Betreuung von Prof. Dr. Rolf Müller mit dem Titel: „Studies of cinnabaramide biosynthesis in *Streptomyces* sp. BC22709“.

10. 2010 - 03. 2014 Dissertation an dem Institut für pharmazeutische Biotechnologie, Universität des Saarlandes/Abteilung Mikrobielle Naturstoffe, Helmholtz-Instituts für Pharmazeutische Forschung Saarland (HIPS) unter der Leitung von Prof. Dr. Rolf Müller mit dem Thema: „Synthetic biotechnology to study and engineer natural product biosynthesis in actinomycetes“.

Beruflicher Werdegang:

04. 2009 - 09. 2009 studentischer Hilfskraft an der Arbeitsgruppe von Prof. Dipl. Ing. Dr. tech. Elmar Heinzle, Universität des Saarlandes, Germany.

06. 2009 - 12. 2009 studentischer Hilfskraft am Institut für pharmazeutische Biotechnologie bei Prof. Dr. Rolf Müller, Universität des Saarlandes, Germany.

Auszeichnungen:

09. 2005 - 02. 2006 First-class scholarship of the institute, Department of Chemical Engineering, Tianjin University.
09. 2007 Gold award of iGEM in 2006 as a team member of Tianjin University, International Genetically Engineered Machine (iGEM) competition, Massachusetts Institute of Technology, Massachusetts, USA.
09. 2007 - 02. 2008 Ansing Int. Group-Tianjin University scholarship, Tianjin University.
01. 2012 Paper of the Month, January, 2012. Awarded by Helmholtz Center for Infection Research.
09. 2012 Lecture award, 2012 the International Workshop of the VAAM on the "Biology and Chemistry of Antibiotic-Producing Bacteria and Fungi", Braunschweig, Germany.
05. 2013 Chinese Government Award for Outstanding Self-financed Students Abroad, China Scholarship Council.
09. 2013 Federation of European Microbiological Societies (FEMS) Meeting Attendance Grant.

Publikationen:

N. Quade⁺, **L. Huo**⁺, S. Rachid, D. W. Heinz, R. Müller, Unusual carbon fixation giving rise to diverse polyketide extender units. *Nature Chemical Biology*, 2012, 1, 117-124.

L. Huo⁺, S. Rachid⁺, M. Stadler, Silke C. Wenzel, R. Müller, Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chemistry&Biology*, 2012, 19, 1278-1287.

P. Barbie, **L. Huo**, R. Müller, U. Kazmaier, Stereoselective synthesis of deuterium-labeled (2S)-cyclohexenyl alanines, biosynthetic intermediates of cinnabaramide, *Organic Letters*, 2012, 14(23), 6064.

S. Rachid, **L. Huo**, J. Herrmann, M. Stadler, B. Köpcke, J. Bitzer, R. Müller, Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. *ChemBioChem*, 2011, 12, 922-931.

(⁺: Authors contributed equally to the work)