

**Biochemical Characterization of a
Non-Canonical Branching Module of the
Rhizoxin Polyketide Synthase**

Dissertation

To Fulfill the
Requirements for the Degree of
„doctor rerum naturalium“ (Dr. rer. nat.)

**Submitted to the Council of the Faculty
of Biology and Pharmacy
of the Friedrich Schiller University Jena**

by M.S. Srividhya Sundaram
born on 03.10.1985 in Coimbatore, India

The thesis entitled **Biochemical Characterization of a Non-Canonical Branching Module of the Rhizoxin Polyketide Synthase** was carried out at the Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute Jena in the department of Biomolecular Chemistry under the guidance of Prof. Dr. Christian Hertweck from July 2012 to January 2017.

Reviewers:

Prof. Dr. Christian Hertweck
Department of Biomolecular Chemistry
Leibniz Institute for Natural Product Research and Infection Biology
Hans Knöll Institute, Jena, Germany

Prof. Dr. Dirk Hoffmeister
Department of Pharmaceutical Microbiology
Leibniz Institute for Natural Product Research and Infection Biology
Hans Knöll Institute, Jena, Germany

Prof. Neil Oldham
Biomolecular Mass Spectrometry, Faculty of Science
School of Chemistry
University of Nottingham, UK

Date of the thesis defense: 11.08.2017

Table of contents

Table of contents

Abbreviations	1
1 Introduction	3
1.1 Natural products in drug discovery	3
1.2 Polyketide biosynthesis	5
1.2.1 Basic mechanisms	5
1.2.2 The ketosynthase domain.....	7
1.2.3 Classification of PKSs.....	8
1.2.4 <i>Trans</i> -AT PKSs.....	9
1.2.5 Structures of PKSs – opportunities for synthetic biology	10
1.3 Rhizoxin.....	12
1.3.1 Isolation, mode of action and biosynthesis	12
1.3.2 Polyketide-chain branching.....	13
2 Aims of the research	17
3 Manuscripts	19
3.1 List of manuscripts.....	19
3.2 Manuscript A.....	26
3.3 Manuscript B.....	76
3.4 Manuscript C.....	98
3.5 Manuscript D	120
3.6 Manuscript E	134
3.7 Manuscript F	158
4 Unpublished results	173
4.1 Crystallization attempts of the entire branching module	173
4.1.1 Introduction	173
4.1.2 Results and discussion	173
4.1.3 Methods.....	175
5 Discussion	179
5.1 Heterocycles formed by the branching module	179

5.2	Stereoselective dual polyketide branching	182
5.3	The structural role of the B domain.....	184
5.4	Investigating the mechanism of the branching KS	187
5.5	Small- to large-sized lactone formation	190
6	Summary	191
7	Zusammenfassung	195
8	References	199
9	Appendix	214
9.1	Manuscript E – Supplementary information	214
9.2	Manuscript F – Supplementary information	228
9.3	Manuscript G.....	249
	List of publications.....	285
	Conference presentations & awards	286
	Acknowledgements	287
	Curriculum vitae	288
	Selbstständigkeitserklärung	289

Abbreviations

ACP	Acyl carrier protein
AT	Acyl transferase
BLAST	Basic Local Alignment Search Tool
CoA	Coenzyme A
CD	Circular dichroism
Cryo-EM	Single-particle cryo-electron microscopy
DEBS	6-Deoxyerythronolide B synthase
DH	Dehydratase
E.	<i>Escherichia</i>
ECH	Enoyl-CoA-hydratase
ER	Enoyl reductase
FAS	Fatty acid synthase
FDA	Food and Drug Administration
GNAT	GCN5-related N-acetyltransferase
GTP	Guanosine triphosphate
HCS	Hydroxymethylglutaryl-ACP-synthase
HEPES	2-[4-(2-Hydroxyethyl) piperazin-1-yl] ethanesulfonic acid
HMG-CoA	3-Hydroxy-3-methylglutaryl-coenzyme A
HMGS	Hydroxymethylglutaryl synthase
HR	Highly-reducing
HRMS	High resolution mass spectrometry
KR	Ketoreductase
KS	Ketosynthase
LC-ESI-MS	Liquid chromatography-electrospray ionization-mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization
MCoA	Malonyl-CoA
MS	Mass spectrometry
MT	Methyl transferase
MWCO	Molecular weight cut-off
NMR	Nuclear Magnetic Resonance

NR	Non-reducing
NRPS	Nonribosomal peptide synthetase
PDB	Protein data bank
PKS	Polyketide synthase
PPant	Phosphopantetheine
PPTase	4'-Phosphopantetheinyl transferase
PR	Partially-reducing
R & D	Research and Development
SAH	<i>S</i> -Adenosyl-L-homocysteine
SAM	<i>S</i> -Adenosyl-L-methionine
SAXS	Small-angle X-ray scattering
SNAC	<i>N</i> -acetylcysteamine
TE	Thioesterase
TOF	Time of flight
TRIS	2-Amino-2-(hydroxymethyl) propane-1, 3-diol
WHO	World Health Organization

1 Introduction

1.1 Natural products in drug discovery

Nature has long been a definitive resource for basic human needs, especially for medicinal purposes. Natural products, with a history dating back to 2600 BC, serve as ingredients of traditional indigenous medicine in treating common ailments like colds, coughs and inflammation.¹ With their intrinsic pharmacophore conformations and biological activities, natural products also function as excellent starting points to tailor them into drugs, and hence hold high value in the drug discovery process.^{1,3} About one-third of the FDA-approved drugs in use today belong to natural products and their derivatives.⁴ These compounds include most of the known classes of antibiotics including the β -lactams, aminoglycosides, tetracyclines, macrolides, glycopeptides and lipopeptides.⁵ Natural products are mainly produced by plants, bacteria and fungi and broadly belong to the structural classes of polyketides, nonribosomal peptides, terpenes, ribosomally synthesized and posttranslationally modified peptides and alkaloids. They target a wide variety of proteins (enzymes) and function as inhibitors or modulators of protein-protein interactions.⁶ Consequently, they display diverse biological activities, acting as antibiotics, antiparasitics, antimalarials, statins, immunosuppressants, and anticancer agents, among others.⁷

The use of antibiotics puts selective pressure on microbes to develop multidrug resistance, which is compounded by their indiscriminate use. Therefore, there is an ongoing need for the discovery and development of new antibiotics. Despite this growing need, there exists marked gap in the discovery of new class of antimicrobial compounds over the past 30 years. This innovation gap is partly due to the decline of natural product research in many pharmaceutical companies. The diminished interest in natural product-based drug discovery could be attributed to several reasons.^{6, 8-10} First, the commonly practiced strategy of screening chemical libraries against defined targets in a high-throughput assay is not amenable to natural products of low natural abundance as they are often synthetically intractable on the required scale. Second, the growing interest in combinatorial chemistry resulted in synthetic combinations of compounds rather than natural product-based libraries and finally, the pressure for a short-time discovery of new leads precludes products from natural sources. However, with the emerging trends on genome mining and other R & D approaches, there is a renewed interest in using natural products as sources for lead identification and development.^{11, 12}

Natural products have been the target for synthetic chemists and synthetic biologists alike, with a focus on developing analogues with enhanced activity.¹³ Some of the examples of natural products and their derivatives are shown in **figure 1.1**.

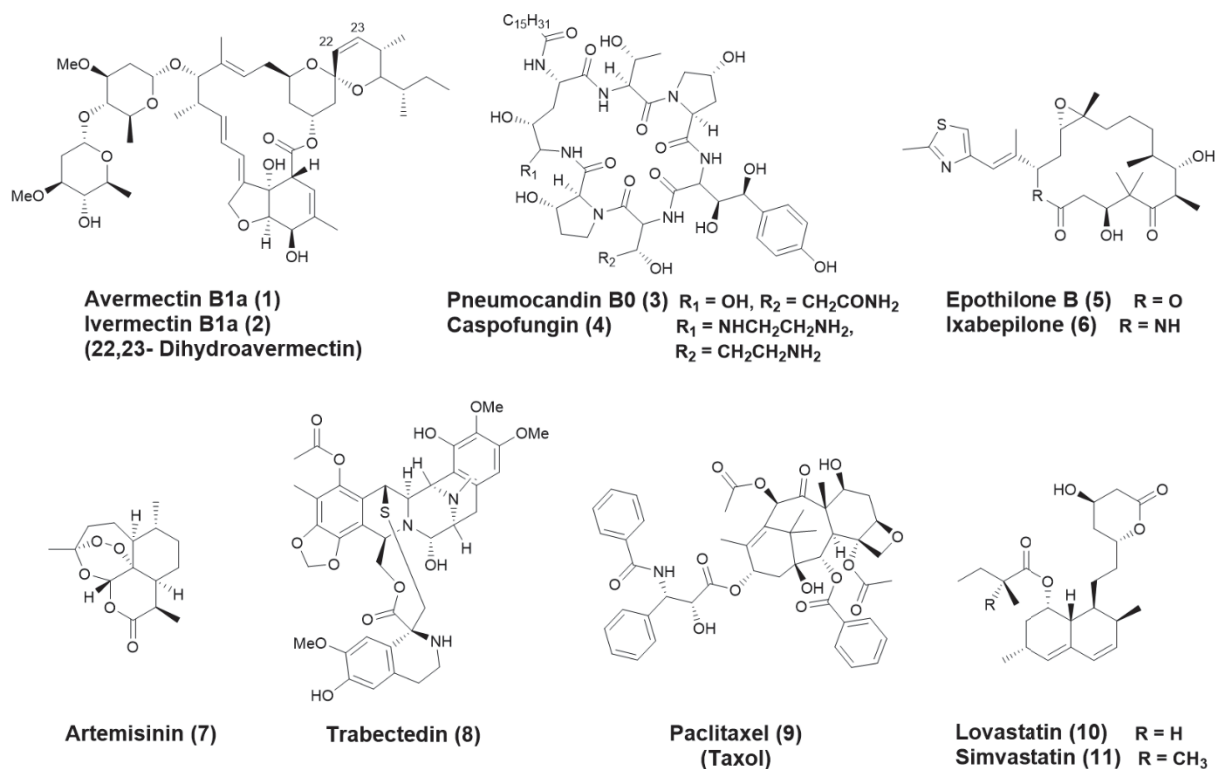


Figure 1.1 Natural products and their derivatives in clinical use

Avermectin B1a (**1**) belongs to the group of polyketides isolated from *Streptomyces avermitilis* that exhibit potent anthelmintic and insecticidal properties.¹⁴ Ivermectin (**2**), the synthetic dihydro derivative of avermectin B1a, has led to a significant reduction of river blindness in Africa and of scabies in India and Southeast Asia.¹⁵ Being a broad spectrum antiparasitic agent, ivermectin is also effective against human parasitic worm infestations including lymphatic filariasis.¹⁶ Hence it may come as no surprise that the discovery of ivermectin has been awarded a Nobel prize in 2015 and it is on the WHO's list of essential medicines.¹⁷ Pneumocandin Bo (**3**) is a lipohexapeptide isolated from the fungus *Glarea lozoyensis* and belongs to the class of echinocandins.¹⁸ Caspofungin (**4**), a semisynthetic derivative of pneumocandin Bo, is the first glucan-inhibiting antifungal drug approved by the FDA against the human pathogen *Candida albicans*.¹⁹ Selective chemical modifications of natural products have also resulted in reduced cytotoxicity of the anticancer agents such as the β -tubulin-binding epothilone B (**5**). The epothilones are effective microtubule disruptors initially isolated from the Myxobacterium *Sorangium cellulosum*.²⁰ Ixabepilone (**6**) is one of the synthetic analogues of epothilone B with substitution in its macrolide ring. This drug has been approved by the FDA to treat metastatic or advanced localized breast cancer.²¹

Semisynthetic strategies have also been employed towards addressing the problem of drug supply shortage. Artemisinin (**7**) is a sesquiterpene endoperoxide with a unique pharmacophore (1, 2, 4-trioxane peroxide), initially isolated from the plant *Artemisia annua*.²² Artemisinin and its derivatives have played an effective role in combating the deadly malaria caused by *Plasmodium falciparum* and *Plasmodium vivax* worldwide. A recent assessment shows that artemisinin combination therapies have averted nearly 22% of the 663 million clinical cases in the period 2000-2015.²³ As an alternative for extracting artemisinin from huge plant cultures, an engineered *Saccharomyces cerevisiae* strain produces its precursor,

artemisinic acid in very high yields.²⁴ Semi-synthetic artemisinin is now routinely derived from artemisinic acid in a two-step reaction.²⁵ The research on artemisinin has not only led to a Nobel prize, but also to a new direction for the synthesis of peroxide-based antimalarial drugs.²⁶ The alkaloid trabectedin (**8**) is an example for a product isolated from a symbiont of the sea squirt *Ecteinascidia turbinata*²⁷ which is in clinical use for soft tissue sarcoma and relapsed ovarian cancer.²⁸ The supply problem of this compound was met by a semisynthetic strategy commencing with a precursor, cyanosafracin B from *Pseudomonas fluorescens*.²⁹

Paclitaxel (**9**) is a complex diterpene isolated from the Pacific yew *Taxus brevifolia*.³⁰ It is a microtubule-stabilizing drug and is in clinical use for the treatment of a number of cancers including ovarian, breast, pancreatic and lung cancers.³¹ Synthetic approaches towards the production of paclitaxel led to the development of several analogues including cabazitaxel. This semi-synthetic drug is being used for the treatment of hormone-refractory metastatic prostate cancer in combination with prednisone.²⁶ The polyketide lovastatin (**10**) isolated from *Aspergillus terreus* was the first commercialized HMG-CoA reductase inhibitor of the cholesterol biosynthesis.³² Further biotransformation experiments with an esterase (LovD) to convert the 2-methylbutanoate side chain to 2, 2-dimethylbutanoate led to an improved drug, simvastatin (**11**).¹⁰ The statin family of drugs are easily one of the best-selling pharmaceutical drugs to date in treating hypercholesterolaemia.³³

Clearly, natural products are being extensively useful in drug discovery efforts. However, only a miniscule of the earth's soil surface has been screened for antibiotic production,⁹ leaving plenty of room for exploration, also pertaining to sources other than soil. In addition to the discovery of new compounds, combinatorial biosynthesis (i.e. reprogramming natural product biosynthetic pathways to produce novel compounds) holds great promises in lead development. The enzymes that constitute natural product biosynthetic pathways have emerged as excellent biocatalysts to create novel natural product analogues with increased bioactivity.³⁴ Therefore, studying the biosynthesis of natural products becomes highly imperative to generate compounds with defined stereochemistry that might be difficult to generate using synthetic chemistry.

The building blocks for polyketides, nonribosomal peptides and terpenes are most often their simple monomer constituents – acyl-CoA thioesters, amino acids and isoprenyl-pyrophosphates, respectively. In the following sections, the biosynthesis of natural products with an emphasis on polyketides is explained in detail.

1.2 Polyketide biosynthesis

1.2.1 Basic mechanisms

Polyketides belong to a structurally diverse class of natural products, encompassing macrolides, polyphenols, polyethers, enediynes and many more.³⁵ The structural diversity of polyketides mirrors the myriad biological activities including, but not limited to, antimicrobial, antiproliferative and immunosuppressive effects.^{10, 36, 37}

Equally astonishing as the diversity of polyketides are the enzymatic machineries that assemble them, with a stark similarity to the fatty acid synthases.³⁸ Typically, polyketides are synthesized by polyketide synthases (PKSs) via the successive coupling of esters through a Claisen condensation.³⁵ The starter units add to the chemical complexity of the polyketides which ranges from the common isobutyryl- or isovaleryl-CoA, to the unusual *p*-nitrobenzoate, glycolate, glycol moieties and so on.^{39, 40} Aside from malonyl-, the extender units can also be their 2-substituted methyl-, methoxy-, ethyl- or chloroethyl-derivatives.⁴¹ A typical reaction cycle in the biosynthesis of polyketides involves an acyl carrier protein (ACP), to which the growing intermediates are tethered to the thiol group of its 18 Å swinging phosphopantetheine (PPant) arm, an acyl transferase (AT) that selectively activates and loads simple or substituted malonyl-extender units on the ACP and a ketosynthase (KS) which then decarboxylates the extender unit and condenses with the acyl-intermediate (or a starter unit). The growing chain is therefore extended by two carbon atoms in each cycle. The nascent formed β -keto group can be processed further by optional enzymes such as ketoreductases (KRs), dehydratases (DHs) and enoylreductases (ERs) to afford β -hydroxyl, α,β double bond and fully-reduced methylene moieties, respectively (**Figure 1.2a**).³⁵

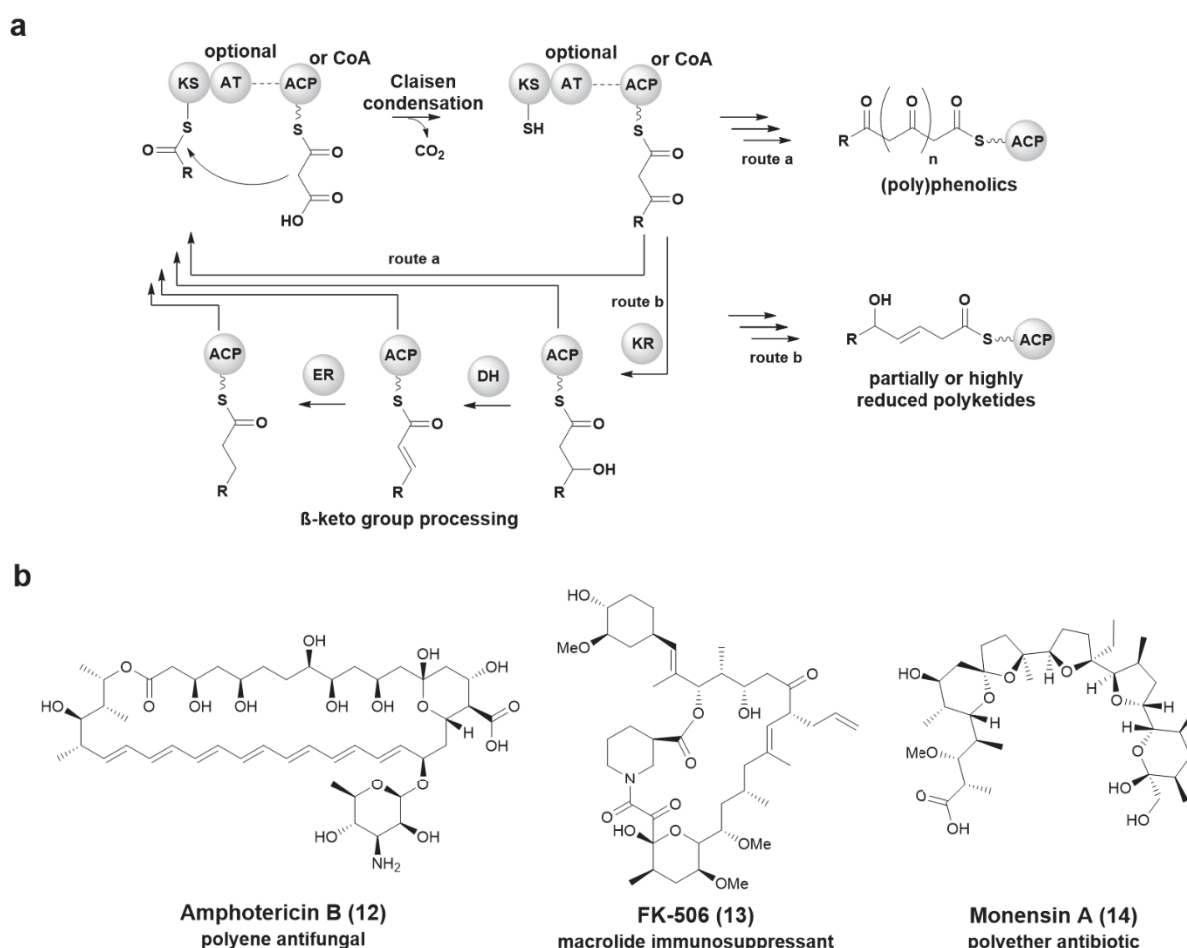


Figure 1.2 a) Schematic of polyketide biosynthesis.³⁵ Wavy bond in ACP refers to its phosphopantetheine arm. **b)** Examples of bioactive polyketides.

In most characterized PKSs, the thioester linkage attaching the polyketide intermediate to the terminal ACP is cleaved by a thioesterase (TE) by simple hydrolysis with bulk solvent or

intramolecular cyclization with a nucleophile on the polyketide;⁴²⁻⁴⁴ however, other offloading mechanisms are known.⁴⁵⁻⁴⁶ Several tailoring modifications might follow the offloading step including acylation, alkylation, hydroxylation, oxygenation, halogenation and glycosylation.⁴⁷ Together with the modifications introduced by the post-PKS tailoring enzymes, the structural diversity of polyketides can be attributed to the chain size, diversity of starter and extender units, alterations in the oxidation state, configuration of the β -keto groups and offloading mechanisms (**Figure 1.2b**).³⁵

1.2.2 The ketosynthase domain

The ~430 residue β -ketoacyl-S-ACP synthase, or the KS domain, catalyzes the core carbon-carbon bond formation that result in the chain elongation in polyketide and fatty acid biosynthesis.⁴⁸ The KS domain is highly conserved and possesses a thiolase fold with alternating layers of $2\alpha-5\beta-2\alpha-5\beta-2\alpha$ architecture (**Figure 1.3a**). The KS forms a dimer and its dimeric interface largely contributes to the dimerization of the entire PKS module. Typically, KS domains have a Cys-His-His catalytic triad found in the TACSSS, EAHGTG and KSNIGHT motifs.⁴⁸ Although the reaction mechanism catalyzed by the KS is not fully understood, studies of the KS domains from bacterial and mammalian fatty acid synthases have provided insights into the role of the conserved amino acids.⁴⁹⁻⁵¹ The cysteine nucleophile is located at the KS dimer interface and its nucleophilicity is enhanced by the helix dipole effect. Mutational experiments revealed that Cys is pivotal for the transthioesterification, where it attacks the thioester of an acyl-ACP upon its entry into the substrate channel. The negative charge on the thioester carbonyl is stabilized by the backbone amides of the Cys164 and Phe396 residues that constitute the oxyanion hole (**Figure 1.3b**).⁴⁹

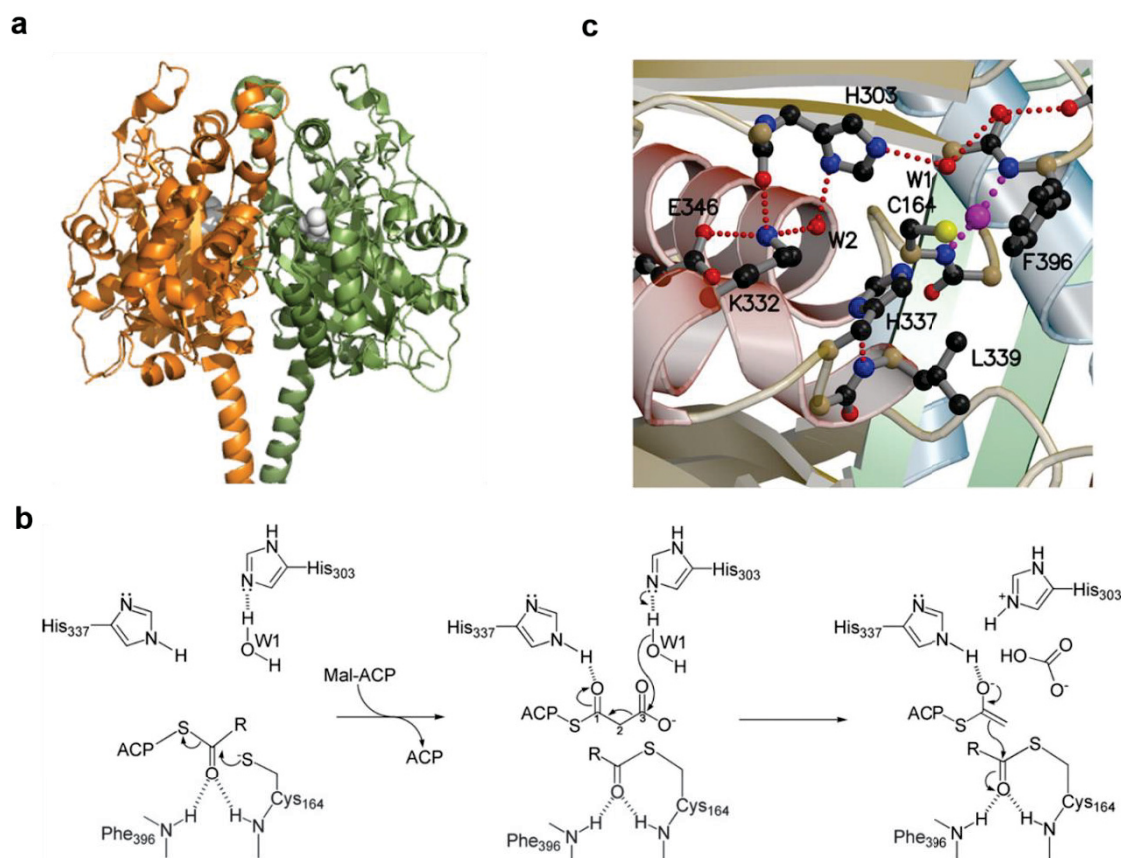


Figure 1.3 a) KS dimer of the DEBS KS module 5 (PDB code: 2hg4)⁵² with the active site cysteine shown as spheres. The figure was generated using Pymol.⁵³ **b)** Proposed mechanism of chain elongation catalyzed by a typical KS domain exemplified by the FabF KS⁴⁹ **c)** Active center of FabF of fatty acid biosynthesis from *Streptococcus pneumoniae*.⁴⁹ Pink ball refers to the oxyanion hole; pink dots represent the backbone amide interactions; and red dots refer to the hydrogen bonds.

Following acylation of the KS active site cysteine, the ACP is released and the extender unit (malonyl or substituted malonyl) is loaded to the upstream ACP. At this stage, the KS catalyzes the decarboxylation of the extender unit. One of the histidines activates a structured water molecule to attack the malonyl-ACP C₃ carboxylate, to release bicarbonate. The resulting negatively charged enolate oxygen is stabilized by the other histidine. Notably, this decarboxylation reaction is independent of Cys acylation as KS domains in which the Cys is mutated to Ala or Gln retained the ability to decarboxylate the extender units.^{54, 55} In the condensation step, the enolate attacks the acyl-thioester to afford a tetrahedral intermediate stabilized by the oxyanion hole (**Figure 1.3b**). After resolution of the tetrahedral intermediate, the β -ketoacyl-ACP product dissociates from the active site and associates with either the β -keto reductive domains (if they are present) or to the downstream module for further elongation/modification. Apart from the residues that comprise the critical catalytic triad, other residues with notable functions include Lys332 that contributes to the hydrogen bond network and Phe396 located on a hairpin turn that functions as a 'gatekeeper' that controls the order of the substrate addition, (residue numbers correspond to FabF KS from *S. pneumoniae*) (**Figure 1.3c**).⁴⁹ A proper orientation of the KS active site residues is essential for its substrate selectivity, and therefore the decarboxylative condensation reaction.

The mechanism described above corresponds to KS I and KS II-type of enzymes.^{49, 56} The KS III-type enzymes utilize distinct substrates and catalytic residues.⁵⁷ They are present across different types of PKSs, as described in the following section.

1.2.3 Classification of PKSs

Owing to the varied structure and mode of action, PKSs are classified into three distinct types –type I, type II and type III PKSs, analogous to the fatty acid synthases (FASs).^{58, 59} Although the reaction chemistry of PKSs is similar to FASs, PKSs have evolved to use a broader range of building blocks leading to the formation of diverse products.⁵⁹ Type I PKSs are multifunctional megaenzymes (up to ~10 MDa) with covalently fused catalytic domains arranged on a large polypeptide.⁶⁰ Type I PKSs are mainly found in prokaryotes and are subdivided into two major groups: modular and iterative (**Figure 1.4a**).⁶¹ The classical example for a bacterial modular type I PKS is the 6-deoxyerythronolide B synthase (DEBS) responsible for the synthesis of the aglycone core of the antibiotic erythromycin.⁶² In this case, each PKS 'module', where all the catalytic 'domains' are grouped, is involved in a single round of incorporation and reductive processing of a C₂ unit. Therefore, the number of modules can directly be translated to the number of precursors incorporated and therefore the expected product structure, a principle defined as 'colinearity'.^{63, 64} Iterative type I PKSs possess the same core catalytic domains as the modular PKSs, but the domains are arranged on a single polypeptide that functions in an iterative fashion to form the entire product.^{58, 65} Fungal iterative PKSs generate aromatic polyketides such as lovastatin, with varying degrees of reduction and are further classified into nonreducing (NR), partially reducing (PR) and highly reducing (HR) PKSs.⁶⁶ In recent years,

bacterial iterative PKSs have been characterized that produce highly complex molecules such as enediynes and tetramate macrolactams.⁶⁷

Type II PKSs are dissociable complexes of discrete enzymes typically with single functions. They are present only in bacterial systems and contain the same core catalytic domains found in type I PKSs. The exception is the presence of two KS domains, KS_{α} which catalyzes iterative, decarboxylative extension of the starter unit and KS_{β} , which controls the length of the polyketide intermediate (**Figure 1.4b**).⁶⁸ There is little or no reductive processing of the β -keto group but subsequent cyclizations generate aromatic molecules such as the anticancer drugs daunorubicin and doxorubicin.^{69, 70}

Both type I and type II PKSs rely on the thioester mechanism where the substrates and the growing intermediates are covalently attached to the PPant arm of the ACP domain. In contrast, type III PKSs utilize free CoA-linked thioester substrates and a single enzyme catalyzes acyl group transfers (**Figure 1.4c**).⁷¹ Type III PKSs are present mainly in plants but also in fungi and bacteria (for example, 3, 5-dihydroxyphenylglycine synthase).^{72, 73}

In addition to these PKSs, integrated NRPS-PKS hybrids are also responsible for the synthesis of many active compounds like epothilone and FK-506.^{74, 75}

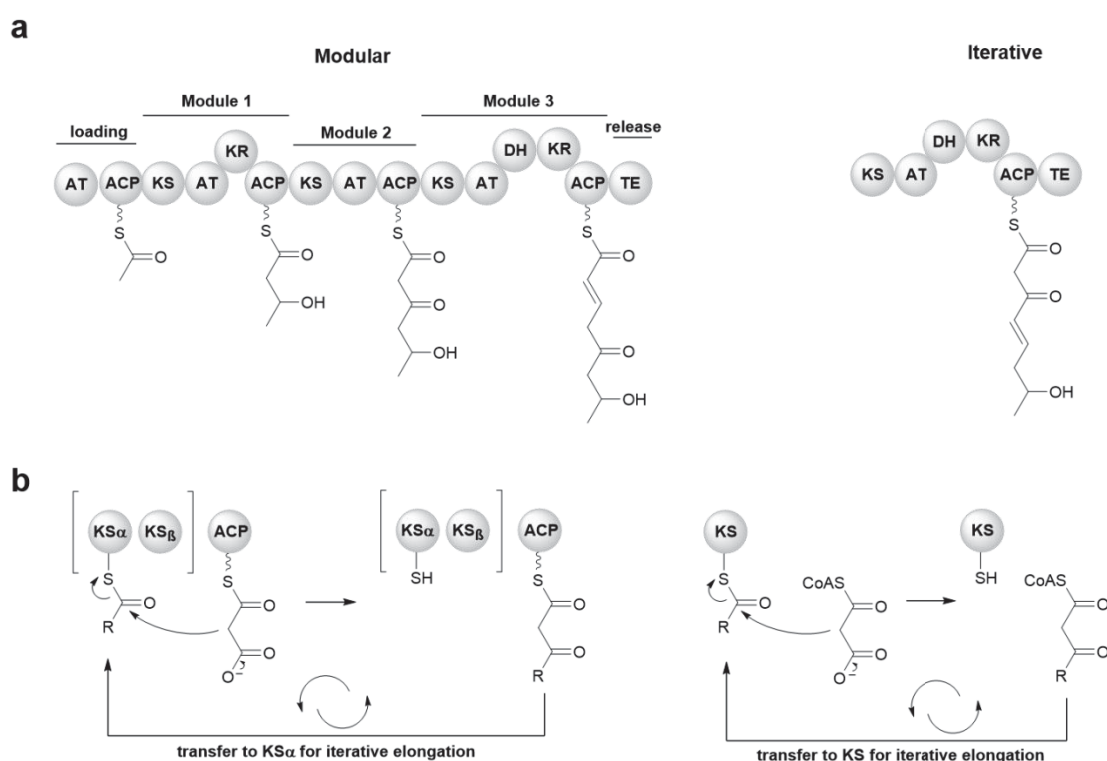


Figure 1.4 Representation of PKS systems from⁷⁶ **a**) type I PKS (modular and iterative) **b**) type II PKS and **c**) type III PKS

1.2.4 *Trans*-AT PKSs

The colinearity of type I PKSs described above is limited to the class of *cis*-AT PKSs, in which the AT domains are integral parts of the modules (**figure 1.4a**). Recently, the complexity of

polyketides has been further expanded by a second class of modular system termed as *trans*-AT PKSs.⁷⁷⁻⁸¹ A recent analysis of PKS-specific genes in bacterial genomes revealed that almost 38% of the modular PKSs belong to the *trans*-AT type, suggesting that *trans*-AT PKS-generated polyketides contribute to a major class of natural products.⁸² One prominent feature of *trans*-AT PKSs is the absence of integrated AT domains within their modules. Instead, stand-alone AT domains activate and load malonyl or substituted-malonyl units on the ACP.⁷⁸ The complex diversity of this class of polyketides can be attributed to the presence of inactive domains, module skipping, unusual domain orders, split modules (divided on two proteins) and the presence of other *trans*-acting enzymes such as ERs.^{83, 84} Phylogenetic analysis of *trans*-AT PKSs revealed that these enzymes have evolved by horizontal gene transfer that allowed recombination of PKS genes. In contrast to the phylogeny of *cis*-AT PKSs where the enzymes are grouped according to the polyketide types, *trans*-AT PKSs are grouped based on the substrate selectivity of KSs.⁸⁵ Several non-canonical modules are housed in the *trans*-AT PKSs such as GNAT loading modules,⁸⁶ non-elongating modules with inactive KS domains,⁸⁷ modules that perform double bond migration⁸⁸ and modules that generate a β -branch and/or rings. Due to their inherent mosaic architectures, the *trans*-AT PKSs are attractive targets for engineering and reprogramming strategies.⁸⁹ Among the *trans*-AT-derived polyketides, compounds with α - or β -branches represent an interesting class, as the alkyl branches serve as precursors for chemical diversification of functional groups in polyketides.

1.2.5 Structures of PKSs – opportunities for synthetic biology

The mechanisms of *cis*-AT and *trans*-AT PKSs are remarkably complex in that each catalytic domain must communicate with many others throughout the modular architecture. Understanding how these domains work in concert would greatly facilitate the engineering of the PKSs to generate novel metabolites. However, it is extremely challenging to generate a crystal structure of an intact PKS module mainly due to the inherent flexibility of the domains. Owing to the conserved homology between the FASs and PKSs, the first reported high resolution crystal structure of porcine FAS^{90, 91} propelled generation of several models of the *cis*-AT PKS modular architecture. These models were built on the mammalian FAS as well as based on the structures of individual domains and didomains.^{48, 92-94} Nevertheless, marked differences were observed in the oligomeric state of the domains and the domain interfaces between FASs and PKSs. In 2014, the first structures of an intact PKS module were reported for the pikromycin PKS module 5 (KS-AT-KR-ACP) by single-particle cryo-EM and for the DEBS module 3-TE (KS-AT-KR-ACP-TE) and the bimodules 5 and 6 by SAXS.^{95, 96} The structures revealed that the domains undergo extensive conformational changes adopting up to seven functional states in a single step of biosynthesis. The interdomain linkers play a massive role in allowing free movement of ACP to and from neighboring domains and also function as structural motifs.

However, understanding the structure-function relationships of linkers is incomplete without a high-resolution structure. Very recently, hybrid crystal structures have been reported for the condensing (KS-AT) and modifying (DH- Ψ KR-ER-KR) regions of the *Mycobacterium smegmatis* drug target mycerosic acid synthase (MAS), a fully reducing iterative PKS. The modifying region revealed a linker-based dimeric architecture that could be exploited for PKS engineering approaches.⁹⁷ Although the reported structures advance the basic understanding

of the dynamic *cis*-AT PKS architecture, we are only beginning to understand how the domain arrangement and intersubunit communication occurs in case of the complex *trans*-AT PKSs. In contrast to a number of high resolution structures of the *cis*-AT PKS domains,⁴⁸ very few structures of *trans*-AT PKSs are available. The only available structure of an intact module from the *trans*-AT family is the low resolution structure of module 5 (KS-ACP-ACP) from virginiamycin PKS.⁹⁸ In light of the very few high resolution structures available from the *trans*-AT family, the AT from disorazole synthase, the solution structure of a tandem ACP from the mupirocin β -branching cassette, the KS of module 2 from bacillaene PKS are some examples for structures from individual domains.⁹⁹⁻¹⁰¹

Knowledge on the structure of PKSs would propel the genetic engineering approaches further. The first example of combinatorial biosynthesis surfaced in 1989 where a carbomycin biosynthetic gene from *Streptomyces thermotolerans* produced isovalerylspiramycin in *S. ambofaciens*.¹⁰² Since then, several strategies have been implemented to selectively insert, exchange or delete domains and modules to generate hybrid compounds.^{103, 104} One of the well-established examples for combinatorial biosynthesis of complex polyketides is the synthesis of structural and functional variants of erythromycin.^{105, 106} In recent years, thanks to the advanced genetic and molecular biology techniques now available, it is possible to identify, express and manipulate diverse biosynthetic pathways in heterologous hosts. This 'synthetic biology' approach, geared towards the specific modification of individual PKS domains, is gaining attention either as an alternative or a parallel approach to combinatorial biosynthesis.¹⁰⁷

As more structures of individual domains or subunit domains are obtained, the boundaries of domains within the assembly line are coming into focus. Especially for the *trans*-AT PKSs, considerable progress has been made in understanding their function by studying individual domains using *in vitro* biochemical techniques.⁸⁹ Such *in vitro* analyses depend on the availability of suitable substrates, production of functional proteins and the usage of high resolution techniques such as mass spectrometry, NMR and X-ray crystallography. Only if the domain junctions and the linker regions of PKSs are better defined, can the true potential of PKS engineering to produce novel compounds be realized.

This thesis focusses on a *trans*-AT derived macrolide with a pharmacophoric moiety formed as a result of an unusual C-C bond forming mechanism. The biosynthesis of this macrolide has been mentioned in detail in the following section.

1.3 Rhizoxin

1.3.1 Isolation, mode of action and biosynthesis

Compounds that target the microtubules either to block their depolymerization or to inhibit their assembly belong to the most effective class of chemotherapeutics.¹⁰⁸ Rhizoxin (**15**), a 16-membered macrolide (**Figure 1.5a**), was first isolated from the pathogenic fungus *Rhizopus microsporus* in 1984, and shown to be the causative agent of rice seedling blight.¹⁰⁹ The structure of rhizoxin features 11 stereogenic centers, two epoxides, an oxazole and the pharmacophoric δ -lactone.¹¹⁰ Rhizoxin alone was shown to induce abnormal swelling and malformation of the rice seedling roots, by binding to the rice cell β -tubulin and thereby inhibiting cell division.¹¹¹ Rhizoxin and its derivatives inhibit tubulin-dependent GTP hydrolysis, microtubule assembly ($IC_{50} = 5\mu M$) and caused extensive depolymerisation of the microtubules in human and murine tumor cell lines.^{111, 112} Although rhizoxin entered clinical trials as an antitumor agent, poor *in vivo* efficacy prevented the compound and derivatives from progressing beyond phase II trials.^{113, 114} The unique structure of rhizoxin nevertheless triggered several total synthetic approaches with the aim to generate stable derivatives; however, these studies have not resulted in a clinically successful compound.^{115, 116} Stable isotope labeling experiments with ¹³C labeled acetate, methionine, serine and glycine demonstrated that the biosynthesis of rhizoxin involved a PKS.¹¹⁷ However, failure to identify fungal-specific PKS genes led to the unexpected finding that the actual producer of rhizoxin is not the rice-plant residing fungus *Rhizopus microsporus*, but a bacterial endosymbiont, *Burkholderia rhizoxinica* that lives in the fungal cytosol.¹¹⁸ The bacterium isolated from the fungus was amenable to genetic manipulation and produced several rhizoxin derivatives in pure culture, with enhanced antiproliferative activities.¹¹⁹ This example was the first of its kind to establish a metabolite connection in a plant-bacterial-fungal symbiotic association. The *Burkholderia-Rhizopus* symbiosis spreads across diverse ecological niches throughout the world.¹²⁰ The report parallels the endosymbiotic metabolites isolated from other eukaryotes like tunicates, sponges or insects.¹²¹

By cloning and gene inactivation experiments, the gene cluster responsible for rhizoxin biosynthesis was subsequently identified to be a type I *trans*-AT PKS-NRPS hybrid gene cluster. The proteins that comprise this PKS-NRPS hybrid are encoded by five open reading frames, *rhi* A-F and are arranged into modules that uses serine and malonate as the building blocks. The compound released from the assembly line is devoid of the C₂-C₃ and C₁₁-C₁₂ epoxide moieties and the C₁₇ O-methyl group (**Figure 1.5a**). The C₁₁-C₁₂ epoxidation and C₁₇ O-methylation are part of the post-PKS tailoring reactions catalysed by RhiH and RhiI, respectively.¹²² The second epoxidation is carried out by the fungal host by an unidentified enzyme. Although not required for bioactivity, the epoxidation at C₂-C₃ makes the molecule more potent. The Rhi PKS exhibits rare features typical of a *trans*-AT system, with a disrupted NRPS module 1 and PKS module 3, a silent KS-ACP divided between modules 8 and 9, an acetate-derived C₅ β -branching between modules 10 and 11.¹²² An identical but silent gene cluster was also found in *Pseudomonas fluorescens*, a biocontrol agent present in plant rhizosphere.¹²³ *P. fluorescens* produces several rhizoxin analogues that are toxic to fungal and oomycete plant pathogens.¹²⁴ This revealed the ecological role of rhizoxin to be dependent on the tolerance or sensitivity of the plant itself. In case of *Rhizopus*-sensitive rice plants, the

bacterium, in return for the nutrients from fungus, produces rhizoxin which the fungus then uses to kill the rice seedling.¹²⁵ However, for rhizoxin-tolerant plants, *P. fluorescens* can associate with the plant and produce rhizoxin that protects the plant against fungal pathogens.

Although the discovery of rhizoxin failed to make a clinical impact, it nevertheless led to the identification of a novel druggable site on β -tubulin. The crystal structure of tubulin and the 2,3-desepoxy rhizoxin derivative (rhizoxin F) shows that the compound shares a distinct binding site from that of the vinca alkaloid, vinblastine.¹²⁶ The key interactions of rhizoxin F and the β -tubulin were made by hydrophobic contacts between 8a methyl- and Asn 101, Asn 102, Val 182, Phe 404 and Tyr 408 and hydrogen bonds between the hydroxyl oxygen at position 13 and Val 181 and between the 5b carbonyl- and Asn 102 of the β -tubulin (**Figure 1.5b and 1.5c**).^{126, 127} Mutational analysis shows that the Asn 102 (or Asn 100 in some homologues¹²⁷) is the residue that determines sensitivity or resistance to rhizoxin.^{128, 129} From the tubulin binding experiments, the δ -lactone has been proven to be pivotal for the biological activity of rhizoxin. An open form of the δ -lactone had 4-fold decreased activity whereas the deletion of C5b carbonyl group resulted in complete loss of activity.^{114, 130} Modifications on the C5b carbonyl also resulted in a set of 16 amide and ester analogues, of which the glycerol ester exhibited the best antiproliferative activity.¹³¹

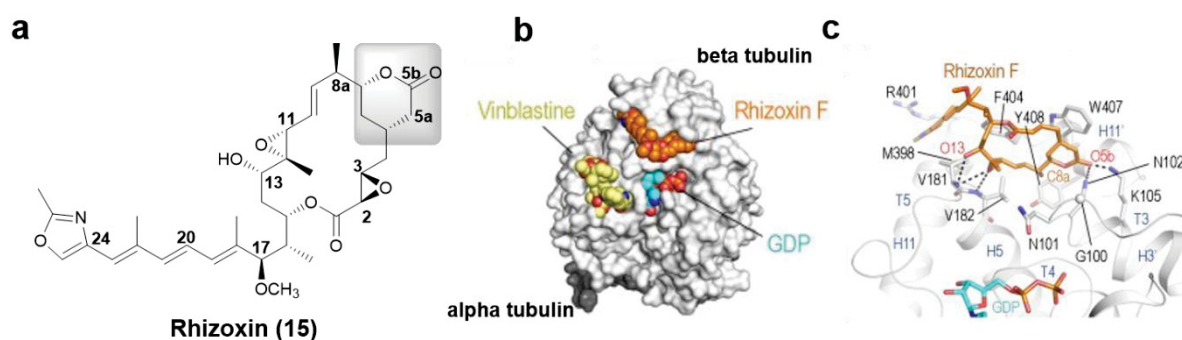


Figure 1.5 **a**) Structure of rhizoxin (**15**) **b**) Rhizoxin F (orange) bound to the β_2 subunit of the tubulin (grey) **c**) Interactions between the functional groups and side chains of rhizoxin F (orange sticks) and the β -tubulin (grey ribbon). Adapted from¹²⁶

1.3.2 Polyketide-chain branching

Installation of alkyl branches at the α -position is frequently observed in polyketides and is introduced by ATs that are specific for substituted-CoA extender units.⁴¹ An alternate way to generate α -methyl-branched intermediates is by the activity of C-methyl transferases (MTs) with *S*-adenosyl methionine (SAM) as a methyl donor (**Figure 1.6a**).¹³² These SAM-dependent methyl transferases are commonly seen embedded in the assembly line of *trans*-AT PKSs and function as a primary source for generating *gem*-dimethyl functionalities in polyketides.¹³³ Some examples of polyketides containing MT-generated branches are bacillaene,¹³⁴ mupirocin¹³⁵ and bongkreic acid¹³⁶ (**16**). Branches at the β -position are introduced by 3-hydroxymethylglutaryl synthase¹³⁷ (HMGS) cassettes that work in an analogous fashion to mevalonate biosynthetic enzymes.¹³⁸ Typically, the hydroxymethylglutaryl-ACP-synthase (HCS) catalyzes the aldol condensation of acyl donors to the β -keto polyketide intermediate. In select pathways, this intermediate is further processed by *cis* or *trans*-acting enoyl-CoA-

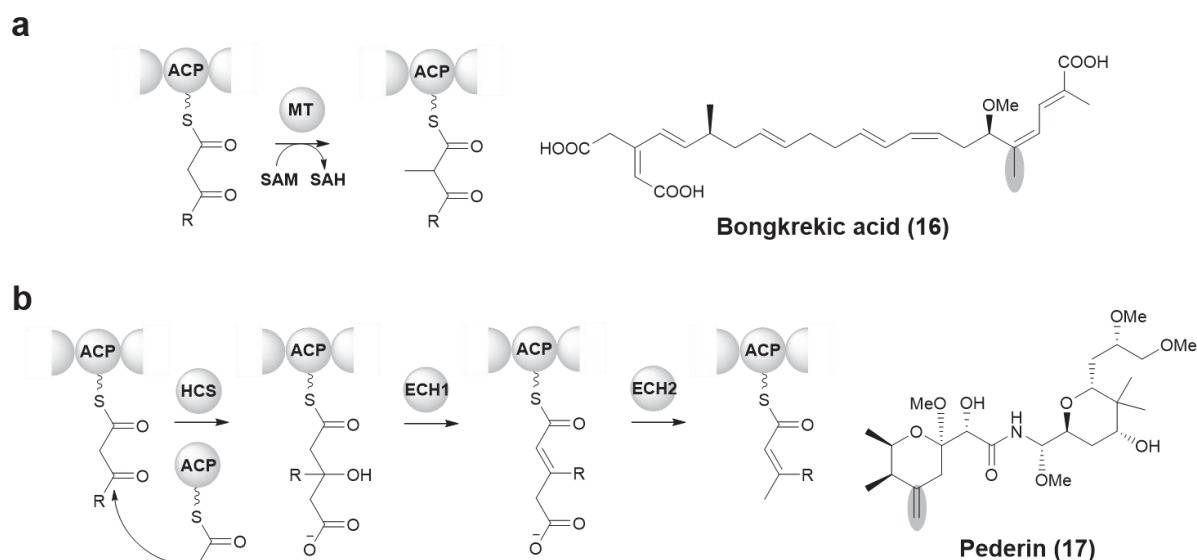


Figure 1.6 a) α -branching by SAM-dependent methyl transferases **b)** General mechanism of β -branching by HCS cassette enzymes. The branches are highlighted.

hydratases (ECH₁ and ECH₂) which catalyze the dehydration and decarboxylation of the growing chain, respectively (**Figure 1.6b**). The HMGS cassette has been shown to incorporate β -branches in curacin¹³⁹, bacillaene¹⁴⁰ and pederin⁷⁹ (**17**). The ACP of this cassette typically contains the motif DSxxxxxW that is specific for β -branching.¹⁰⁰

The absence of HMGS cassette-specific genes in the rhizoxin PKS¹²² suggested that an alternative β -branching reaction could be a starting point to generate its δ -lactone moiety. To understand the exact timing and mechanism of the pivotal C₅ branch in rhizoxin (**15**), the TE domain at the C-terminus of the megasynthase was deleted to prematurely release the pathway intermediates. Using rigorous MS and NMR analyses, it was found that the β -branch is introduced in the so called, 'branching module' (**Figure 1.7**) that is comprised of a KS, an ACP and a cryptic 530 amino acid branching (B) domain.¹⁴¹

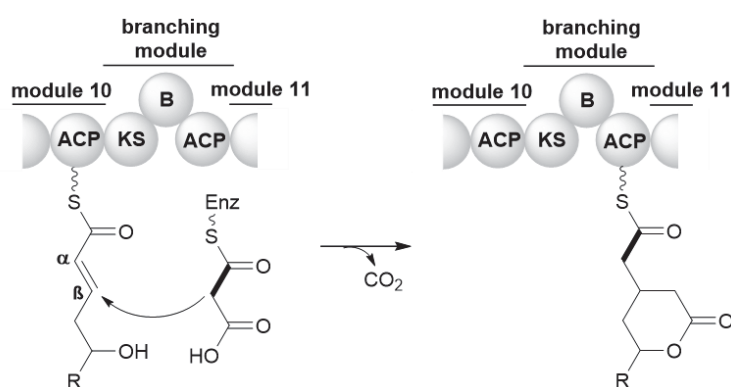


Figure 1.7 Mechanism of β -branching in rhizoxin PKS involves Michael addition of an acetate unit to the α,β -unsaturated thioester intermediate

Unlike the HMGS cassette, these enzymes are hosted within a single module. The substrate for the β -branch was a δ -hydroxy- α,β -unsaturated thioester rather than a β -keto thioester as seen for the mevalonate-like branching. Based on these results, a new mechanism of β -branching has been proposed. According to this proposal, the β -branch is formed by a novel Michael addition of a malonyl extender unit to the α,β -unsaturated thioester intermediate bound to the KS (**Figure 1.7**).¹⁴¹

Based on these results, a new mechanism of β -branching has been proposed. According to this proposal, the β -branch is formed by a novel Michael addition of a malonyl extender unit to the α,β -unsaturated thioester intermediate bound to the KS (**Figure 1.7**).¹⁴¹

A detailed biochemical insight into the β -branching reaction has been made possible by *in vitro* reconstitution of the entire branching module. Using purified proteins (KS-B didomain, ACP and AT) and N-acetylcysteamine (SNAC) thioester **18** as the substrate mimic, malonyl-acetate was added to the α,β -unsaturated thioester by a decarboxylative Michael addition. NMR analysis of the *in vitro* assay performed with ^{13}C -malonate revealed that the δ -hydroxyl group of the substrate further attacks the C_1 carbonyl of the KS-bound chain to form the δ -lactone (**figure 1.8a**). The ACP-bound lactone (**19**) and hydrolysed product (**20**) were observed by MALDI-MS and HRMS, respectively.¹⁴² The KS-B didomain of the rhizoxin branching module is the first reported structure of a KS domain from a *trans*-AT PKS (**figure 1.8b**).¹⁴² Point mutations on the catalytic residues of the KS domain completely abolished the branching activity proving that these residues are crucial for the β -branching. The B domain resembles a regular DH domain with a double-hotdog fold. However, the conserved DH catalytic dyad, His-Asp is altered to Tyr-Asp in the active site of the B domain. The B domain, when produced as a stand-alone protein, was not capable of catalyzing the branching reaction, suggesting that it might only have a structural role. Moreover, point mutation of the conserved Asp residue of the B domain also did not affect the branching activity. In addition to these observations, using a substrate mimic that lacks the δ -hydroxy group, an unprecedented covalent complex of the polyketide intermediate with the KS-B and ACP domains was observed. Taken together, the Rhi KS represents a novel β -carbon branching KS which is distinct from the regular linear chain-forming KSs, that could be exploited for generating specific branches in polyketides.

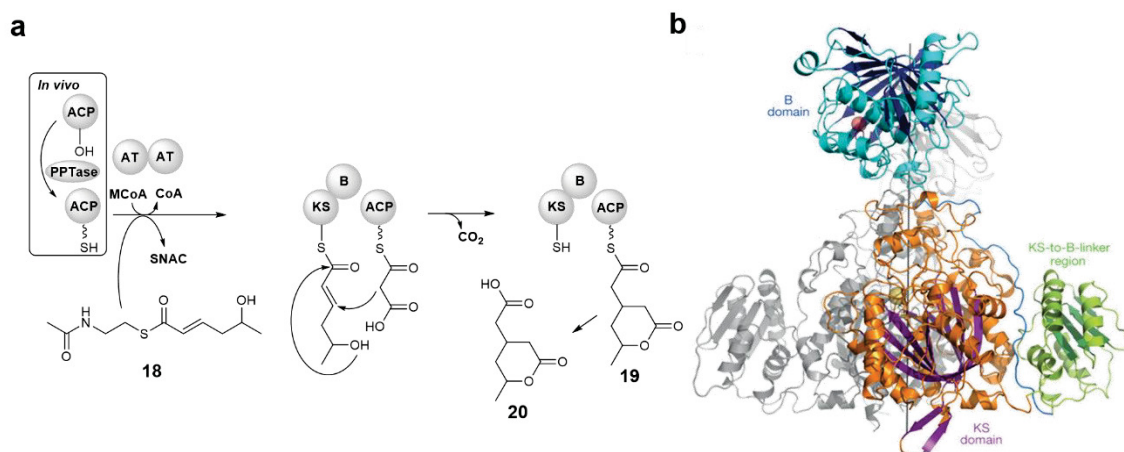


Figure 1.8 a) *In vitro* reconstitution setup of the branching reaction.¹⁴² b) Structure of the dimeric KS-B didomain of the rhizoxin branching module. PDB code: 4kc5

2 Aims of the research

Enzymes from natural product biosynthetic pathways are attractive biocatalysts for synthetic biology. Deeper understanding of the biochemistry of these enzymes would facilitate their use in the generation of 'non-natural' natural products. The non-canonical branching module (KS-B-ACP) of the rhizoxin PKS catalyzes δ -lactone formation via a two-step Michael addition and lactonization reaction. As this mechanism of δ -lactone formation is unprecedented, the rhizoxin branching module represents an excellent candidate for product diversification or pathway engineering.

The branching module generates the δ -lactone which is pivotal for the antimetabolic activity of rhizoxin. The first aim of this thesis was to determine the substrate tolerance of the branching module. As a first step to generate 6-membered heterocycles, a series of *N*-acetylcysteamine thioesters with varied δ -substituents should be synthesized and tested *in vitro*. This study will not only diversify the product-forming capabilities of the branching module but also provide a model for the biosynthesis of natural products with hitherto unknown β -branched architectures. In addition to the use of polyketide SNAC thioester mimics, the second aim was to find out the tolerance of the module towards α -substituted malonyl extender units. These results would address the feasibility of introducing more than one branch in a single reaction. Third, this research also aimed to generate lactones with varied ring sizes, in addition to the 6-membered heterocycles. This should be demonstrated with chemically-synthesized thioester mimics with shorter or longer methylene chains. Generating lactones with increased ring sizes will expand the scope of this complex enzymatic cyclization reaction, especially when such reactions are difficult or not accessible by synthetic chemistry.

Finally, the major aim of the thesis was to understand the mechanistic details of the branching-KS and the cryptic B domain. Although the B domain has been proposed to have a structural role in the branching reaction, to provide direct biochemical proof, chimeric modules should be generated in which the B domain is substituted with similar domains from other PKSs. This analysis would reveal whether the B domain is needed for the structural stability of the KS-B didomain or if it directly participates in the β -branching-lactonization reactions. To decipher the exact mechanism of KS domain in the branching reaction, the role of its conserved amino acid residues should be investigated. Site-specific mutations on selected residues should be performed to observe their impact on each step of the branching reaction. To complement the mutational analysis, KS domain substitution experiments should be performed to understand the role of the entire KS in catalysis. The study would also reveal the key determinants for the construction of a successful hybrid didomain. Taken together, the results from this research would establish the substrate promiscuity of the branching module with a detailed understanding of the mechanism of its key enzyme components.

3 Manuscripts

3.1 List of manuscripts

Manuscript A

Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles

Daniel Heine, Tom Bretschneider, Srividhya Sundaram, Christian Hertweck: Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles, *Angewandte Chemistry International Edition* 53 (43), 11645-11649, 2014.

Polyketides typically result from head-to-tail condensation of acyl thioesters to produce highly functionalized linear chains. The biosynthesis of the phytotoxin rhizoxin, however, involves a polyketide synthase (PKS) module that introduces a δ -lactone chain branch through Michael addition of a malonyl extender to an α,β -unsaturated intermediate unit. To evaluate the scope of the branching module, polyketide mimics were synthesized and their biotransformation by the reconstituted PKS module from the *Rhizopus* symbiont *Burkholderia rhizoxinica* was monitored in vitro. The impact of the type and configuration of the δ -substituents was probed and it was found that amino-substituted surrogates yield the corresponding lactams. A carboxamide analogue was transformed into a glutarimide unit, which can be found in many natural products. Our findings illuminate the biosynthesis of glutarimide-bearing polyketides and also demonstrate the utility of this branching module for synthetic biology.

Daniel Heine	Experiment design, Substrate and reference compounds synthesis, Data analysis, Manuscript writing
Tom Bretschneider	Experiment design, Genetic and biochemical experiments, Data analysis, Manuscript writing
Srividhya Sundaram	Genetic and biochemical experiments, Data analysis (20%)
Christian Hertweck	Experiment design, Manuscript writing

Manuscript B**Twofold polyketide branching by a stereoselective enzymatic Michael addition**

Daniel Heine, Srividhya Sundaram, Tom Bretschneider, Christian Hertweck: Twofold polyketide branching by a stereoselective enzymatic Michael addition, *Chemical Communications* 51 (48), 9872-9875, 2015.

The versatility of the branching module of the rhizoxin polyketide synthase was tested in an in vitro enzyme assay with a polyketide mimic and branched (di) methylmalonyl-CoA extender units. Comparison of the products with synthetic reference compounds revealed that the module is able to stereoselectively introduce two branches in one step by a Michael addition–lactonisation sequence, thus expanding the scope of previously studied PKS systems.

Daniel Heine	Experiment design, Substrate and reference compounds synthesis, Data analysis, Manuscript writing
Srividhya Sundaram	Genetic and biochemical experiments, Data analysis (20%)
Tom Bretschneider	Experiment design, Genetic and biochemical experiments, Data analysis, Manuscript writing
Christian Hertweck	Experiment design, Manuscript writing

Manuscript C**Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching**

Srividhya Sundaram, Daniel Heine, Christian Hertweck: Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching, *Nature Chemical Biology* 11 (12), 949-951, 2015.

Biosynthesis of rhizoxin in *Burkholderia rhizoxinica* affords an unusual polyketide synthase module with ketosynthase and branching domains that install the δ -lactone, conferring antimetabolic activity. To investigate their functions in chain branching, we designed chimeric modules with structurally similar domains from a glutarimide-forming module and a dehydratase. Biochemical, kinetic and mutational analyses reveal a structural role of the accessory domains and multifarious catalytic actions of the ketosynthase.

Srividhya Sundaram	Experiment design, Genetic and biochemical experiments, Data analysis, Manuscript writing (70%)
Daniel Heine	Substrate and reference compounds synthesis
Christian Hertweck	Experiment design, Manuscript writing

Manuscript D**On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems (Review)**

Srividhya Sundaram, Christian Hertweck: On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems, *Current Opinion in Chemical Biology* 31, 82-94, 2016.

Nonribosomal peptide synthetases (NRPS) and type I polyketide synthases (PKS) are versatile thiotemplate systems for the programmed assembly of biosynthetic building blocks. Typically, the post-PKS/NRPS enzymes tailor the resulting chains to yield the bioactive natural product scaffolds. However, more and more examples have surfaced showing that important structural modifications take place while the intermediates are still bound to the assembly line. A growing number of enzymatic domains and trans-acting enzymes as well as their recruiting areas in the modules have been identified and characterized. In addition to the widespread on-line alkylations, hydroxylations and heterocyclizations into oxazole/thiazole residues, on-line modifications lead to a variety of ring systems such as cycloethers, lactones, lactams, glutarimides, cyclopropanes, decalins and cyclic biaryls.

Srividhya Sundaram	Literature research, Manuscript writing (60%)
Christian Hertweck	Literature research, Manuscript writing

Manuscript E**Investigation of a non-canonical chain-branching ketosynthase by site-directed mutagenesis and domain swaps**

Srividhya Sundaram, Ruth Bauer, Hak Joong Kim, Tawatchai Thongkongkaew; Christian Hertweck, Investigation of a non-canonical chain-branching ketosynthase by site-directed mutagenesis and domain swaps, *in preparation*, 2017.

Ketosynthase (KS) domains of modular type I polyketide synthases (PKSs) typically catalyze Claisen condensation of simple acyl and malonyl units to form linear chains that eventually make bioactive polyketides. The KS domain of the branching module of rhizoxin PKS installs a β -branch in the otherwise linear polyketide intermediate which is then cyclized to form the pharmacophoric δ -lactone moiety. To know the exact role of KS domain in the branching-cyclization sequence, we performed targeted mutagenesis of 10 of the conserved amino acids. Biochemical and kinetic analysis revealed that the C3228A and T3290Y mutations completely abolished the β -branch formation. The H3364A, H3404A S3141D, I3478F and S3409A mutants exhibited low turnover activity with attenuated rate for lactonization. Complementing the mutational analyses, we generated chimeras of the branching-KS domain with elongating-KSs and show that a canonical linear chain is formed. These results reveal the mechanistic details of a complex branching reaction catalyzed solely by a KS domain.

Srividhya Sundaram	Experiment design, Genetic and biochemical experiments, Data analysis, Manuscript writing (50%)
Ruth Bauer	Genetic and biochemical experiments
Hak Joong Kim	Reference compounds synthesis
Tawatchai Thongkongkaew	Substrate synthesis
Christian Hertweck	Experiment design, Manuscript writing

Manuscript F**Versatile polyketide-chain lactonization by a non-canonical ketosynthase of the rhizoxin megasynthase**

Srividhya Sundaram, Tawatchai Thongkongkaew, Hak Joong Kim, Ruth Bauer, Daniel Heine, Christian Hertweck: Versatile polyketide-chain lactonization by a non-canonical ketosynthase of the rhizoxin megasynthase, *in preparation*, 2017.

In vitro reconstitution of a branching module from the rhizoxin PKS revealed a non-canonical β -branching mechanism that forms its pharmacophoric δ -lactone moiety. The branching module is tolerant towards substrates with varied δ substituents that yield 6-membered rings like lactam and glutarimide structures. Here, we expand the range of products and show that that the branching module lactonizes substrates with different chain lengths to yield an array of 5- to 10-membered lactones. The ability of KS to promote lactone formation complements the thioesterase- (TE) catalyzed cyclizations. It further sets the stage to generate ring-containing molecular architectures with potential bioactivities, especially when large-membered rings are challenging to produce by synthetic chemistry.

Srividhya Sundaram	Experiment design, Genetic and biochemical experiments, Data analysis, Manuscript writing (40%)
Tawatchai Thongkongkaew	Substrate synthesis
Hak Joong Kim	Reference compounds synthesis
Ruth Bauer	Substrate synthesis
Daniel Heine	Substrate synthesis
Christian Hertweck	Experiment design, Manuscript writing

Manuscript G**A widespread bacterial phenazine forms S-conjugates with biogenic thiols and crosslinks proteins**

Daniel Heine, Srividhya Sundaram, Matthias Beudert, Karin Martin, Christian Hertweck: A widespread bacterial phenazine forms S-conjugates with biogenic thiols and crosslinks proteins, *Chemical Science* 7, 4848-4855, 2016.

Phenazines are redox-active compounds produced by a range of bacteria, including many pathogens. Endowed with various biological activities, these ubiquitous N-heterocycles are well known for their ability to generate reactive oxygen species by redox cycling. Phenazines may lead to an irreversible depletion of glutathione, but a detailed mechanism has remained elusive. Furthermore, it is not understood why phenazines have so many protein targets and cause protein misfolding as well as their aggregation. Here we report the discovery of unprecedented conjugates (panphenazines A, B) of pantetheine and phenazine-1-carboxylic (PCA) acid from a *Kitasatospora* sp., which prompted us to investigate their biogenesis. We found that PCA reacts with diverse biogenic thiols under radical-forming conditions, which provides a plausible model for irreversible glutathione depletion. To evaluate the scope of the reaction in cells we designed biotin and rhodamine conjugates for protein labelling and examined their covalent fusion with model proteins (ketosynthase, carbonic anhydrase III, albumin). Our results reveal important, yet overlooked biological roles of phenazines and show for the first time their function in protein conjugation and crosslinking.

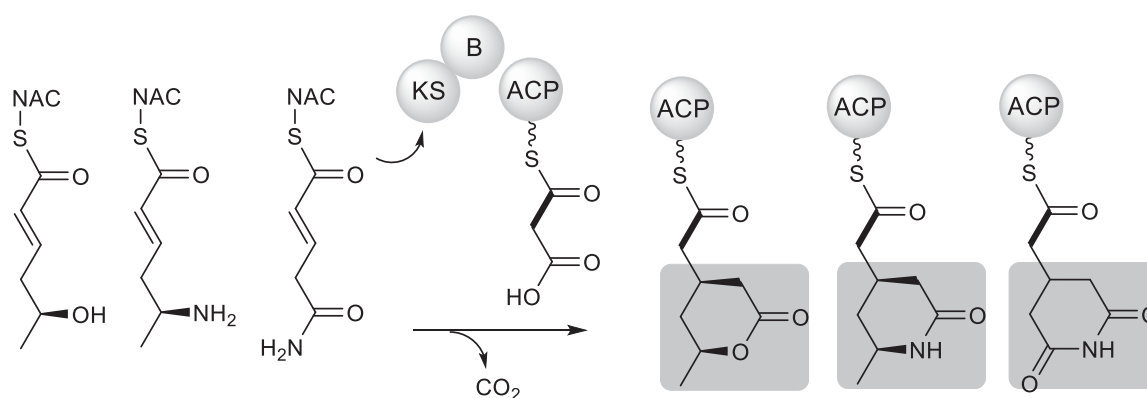
Daniel Heine	Experiment design, Compound isolation, Chemical synthesis, Data analysis, Manuscript writing
Srividhya Sundaram	Experiment design, Protein crosslinking assays, Data analysis (20%)
Matthias Beudert	Chemical synthesis
Karin Martin	Strain handling and fermentation
Christian Hertweck	Experiment design, Manuscript writing

This manuscript results from an unrelated side project undertaken. The complete manuscript is shown in the Appendix section (9.3).

3.2 Manuscript A

Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles

Daniel Heine, Tom Bretschneider, Srividhya Sundaram, Christian Hertweck: Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles, *Angewandte Chemistry International Edition* 53 (43), 11645-11649, 2014.



The branching module of the rhizoxin PKS is tolerant towards substrates with varied δ -substituents. An amine-substituted substrate yields a lactam moiety whereas a carboxamide surrogate results in a glutarimide moiety.

Polyketide Biosynthesis

Enzymatic Polyketide Chain Branching To Give Substituted Lactone, Lactam, and Glutarimide Heterocycles**

Daniel Heine, Tom Bretschneider, Srividhya Sundaram, and Christian Hertweck*

Dedicated to Professor Heinz G. Floss on the occasion of his 80th birthday

Abstract: Polyketides typically result from head-to-tail condensation of acyl thioesters to produce highly functionalized linear chains. The biosynthesis of the phytotoxin rhizoxin, however, involves a polyketide synthase (PKS) module that introduces a δ -lactone chain branch through Michael addition of a malonyl extender to an α,β -unsaturated intermediate unit. To evaluate the scope of the branching module, polyketide mimics were synthesized and their biotransformation by the reconstituted PKS module from the *Rhizopus* symbiont *Burkholderia rhizoxinica* was monitored in vitro. The impact of the type and configuration of the δ -substituents was probed and it was found that amino-substituted surrogates yield the corresponding lactams. A carboxamide analogue was transformed into a glutarimide unit, which can be found in many natural products. Our findings illuminate the biosynthesis of glutarimide-bearing polyketides and also demonstrate the utility of this branching module for synthetic biology.

Polyketides encompass an impressive range of natural products, many of which are in clinical use. Their structural diversity primarily emerges from simple head-to-tail condensations of activated acyl and malonyl building blocks. The resulting polyketide chains are enzymatically processed and may undergo enzymatic tailoring, for example, cyclization, oxygenation, halogenation, and glycosylation.^[1] With respect to the assembly of the carbon backbone, it is remarkable that only a few variations to the unidirectional 1,2-condensation scheme have been observed. However, noncanonical substructures can be crucial for the biological activity of the polyketide, as exemplified by the remarkably potent anti-mitotic agent rhizoxin (**1**, Figure 1).^[2] This phytotoxin is employed in a symbiosis of the plant-pathogenic fungus *Rhizopus microsporus* and its bacterial endosymbiont, *Bur-*

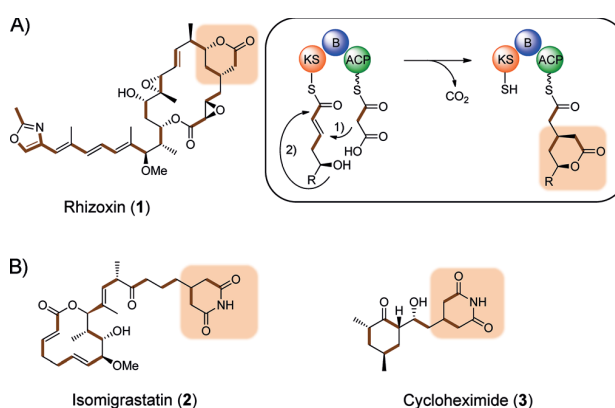


Figure 1. A) The structure of rhizoxin (**1**) and a model for vinylogous chain branching catalyzed by a specialized PKS module; 1) Michael addition, 2) lactonization. B) The structures of isomigrastatin (**2**) and cycloheximide (**3**) with the glutarimide motif highlighted. KS, ketosynthase; B, branching domain; ACP, acyl carrier domain; AT, acyl-transferase.

kholderia rhizoxinica.^[3] The structure of the macrolide is highly unusual because it features a δ -lactone ring that is tethered to the polyketide backbone. Notably, this chain branch is essential for the binding of rhizoxin to the β -tubulin subunit,^[4] through which it efficiently halts cell division in the picomolar range.^[3b] Isotope-labeling experiments have showed that the δ -lactone ring integrates an acetyl unit that branches off at a β -position relative to a former acetyl carbonyl in the polyketide chain (Figure 1).^[5]

Whereas alkyl substitutions at α -positions result from methylation or the incorporation of substituted malonyl units,^[6] branches at β -positions typically involve enzymes that resemble those of the isoprenoid enzymatic machinery.^[7] Surprisingly, genes for such components could not be found in the genome of the rhizoxin producer.^[8] Instead, the architecture of the rhizoxin (*rhi*) type I polyketide synthase (PKS)^[9] and the structures of isolated pathway intermediates^[10] have suggested that the chain branch is introduced through a Michael addition. To shed light on the enzyme mechanism, we reconstituted the designated “branching module” of the *rhi* PKS, which consists of ketosynthase (KS), branching (B), and acyl carrier (ACP) domains, and a trans-acting acyl transferase (AT). By means of an in vitro enzyme assay, we successfully transformed a synthetic surrogate of the polyketide intermediate into the corresponding δ -lactone derivative. We thus confirmed that the pharmacophoric side chain

[*] D. Heine, T. Bretschneider, S. Sundaram, Prof. Dr. C. Hertweck
Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI)
Beutenbergstr. 11a, 07745 Jena (Germany)
E-mail: christian.hertweck@hki-jena.de
Homepage: <http://www.hki-jena.de>

Prof. Dr. C. Hertweck
Chair for Natural Product Chemistry, Friedrich Schiller University Jena (Germany)

[**] We thank A. Perner and H. Heinecke for MS and NMR measurements. Financial support by the Studienstiftung des deutschen Volkes for a fellowship to D.H. is gratefully acknowledged.

Supporting information (including experimental details) for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201407282>.

is introduced through a vinylogous addition of a malonyl unit to an α,β -unsaturated thioester, followed by lactonization (Figure 1).^[11] Notably, similar PKS modules are encoded in the biosynthetic gene clusters for glutarimide-bearing natural products such as isomigrastatin (**2**),^[12] an important cell migration inhibitor (Figure 1). As in the structurally related antibiotic cycloheximide (**3**), the glutarimide residue is important for the biological activity of these compounds.^[13] Although it had been proposed that an enzymatic Michael addition could give rise to the glutarimide structure, the roles of these PKS modules have remained elusive and the proposed mechanism still needs to be experimentally confirmed. Herein, we show that the chain-branching *rhi* PKS module is more versatile than expected. Through an in vitro multienzyme assay with synthetic polyketide mimics, we have demonstrated that the unusual KS–B–ACP module is also capable of forming lactams and even glutarimides.

To test the substrate specificity of the rhizoxin PKS branching module and to explore the scope of potential applications, we synthesized various polyketide analogues for in vitro multienzyme biotransformation experiments. Specifically, a series of δ -functionalized, α,β -unsaturated *N*-acetyl cysteamine thioester (SNAC) derivatives that mimic the activated intermediates were prepared. First, we explored the impact of the configuration of the δ -hydroxy group on the course of the reaction. For this purpose, we synthesized the *R* and *S* enantiomers of the 5-hydroxyhexenoic acid SNAC thioester from the corresponding protected hydroxybutanoates through a reduction–olefination sequence. The stereoisomers **4** and **5** were individually subjected to the in vitro enzyme assay, and the products were analyzed by HPLC–HRMS, with the synthetic δ -lactone as reference (Figure 2). Surprisingly, in both cases formation of the branched product was detected in only one diastereoisomeric form. From retention-time comparisons and deductions from the natural biosynthetic pathway, we inferred that in both cases the *syn*-substituted lactone was formed.

To shed light on the preferences of the KS, we determined the kinetic parameters of the branching reaction. For the *R* enantiomer, we found a K_M value of $(2694 \pm 627) \mu\text{M}$ and a v_{max} value of $(20.9 \pm 2.4) \mu\text{M min}^{-1}$. By contrast, using the *S* enantiomer gave a K_M value of $(1348 \pm 315) \mu\text{M}$ and a v_{max} value of $(2.1 \pm 0.3) \mu\text{M min}^{-1}$. Binding of the non-natural *R* enantiomer to the KS is thus only lowered twofold, whereas the transformation rate of this substrate is ten times lower compared to the *S* enantiomer. This result indicates that the KS basically does not differentiate between the configurations of the δ -hydroxy groups, but lactone ring formation is drastically reduced when the non-natural isomer is applied. It appears that this sequence of Michael addition and lactonization exclusively yields *syn*-substituted lactones and that the configuration of the hydroxy group determines the stereochemical course of the reaction.

To verify the proposed course of the reaction and to corroborate the idea that non-covalent interactions between the substrate and the malonyl unit define the stereochemical course of the reaction, we synthesized and tested the δ -hydroxy-substituted dihydro surrogate **8**, which cannot undergo a branching reaction because of the missing double

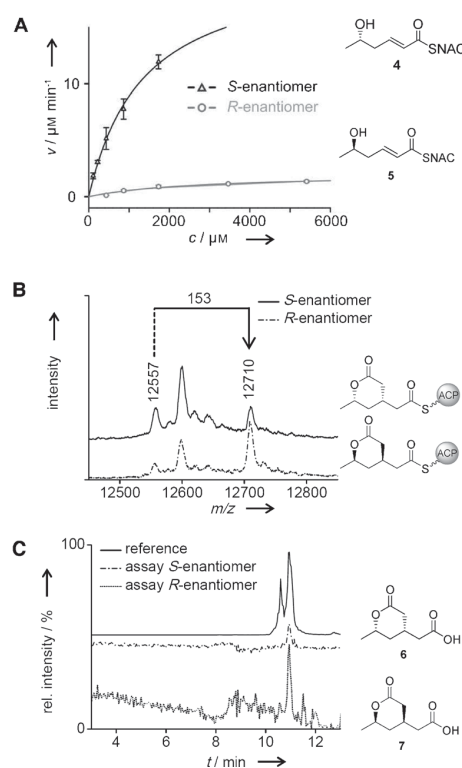


Figure 2. Vinylogous chain branching of pure *R* and *S* enantiomers of the *N*-acetyl cysteamine thioester of 5-hydroxyhexenoic acid, and subsequent lactone formation. A) Michaelis–Menten kinetics of the biotransformation. B) MALDI analysis of ACP-bound products (peak in middle relates to decarboxylated malonylated ACP species). C) SIM–LC–HRMS analysis of hydrolyzed products, and comparison with the racemic synthetic reference. SIM = selected ion monitoring.

bond. Furthermore, a phenol analogue (**10**) was prepared (Figure 3). The activated malonyl units and the rhizoxin polyketide chain mimics were individually added to the enzyme mixture that constitutes the functional PKS module in vitro. The reactions were monitored by MALDI analysis of the ACP with potentially bound products, as well as by high-resolution mass spectrometry (HRMS) of the hydrolyzed products. Unsurprisingly, the saturated alcohol **8** did not undergo any chain branching reaction, thus demonstrating once again that the Michael addition precedes ester bond formation. Furthermore, this experiment showed that the KS does not catalyze Claisen condensation with this substrate when a vinylogous attack is hampered. Changing the aliphatic alcohol to a phenol nucleophile did not give rise to the hypothetical product **11**. These results indicate a high degree of specificity for the KS–B–ACP module.

Next, we explored the possibility of replacing the nucleophilic δ -hydroxy substituent of the polyketide chain. It should be highlighted that this OH group is required for the release of the KS-bound intermediate by lactonization, which allows the propagation of the intermediate. Nonetheless, it is conceivable that other nucleophilic residues could also cleave the thioester bond, which would result in the formation of

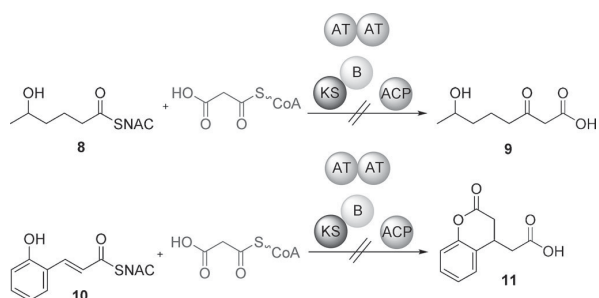


Figure 3. Probing the substrate tolerance of the branching module towards dihydro and phenol variants by using synthetic polyketide surrogates.

alternative heterocyclic systems. If the surrogate was accepted as a substrate for the Michael addition, aminolysis of the thioester would give a lactam ring. To evaluate the effect of an amino group in lieu of the hydroxy group, we synthesized the corresponding δ -amino-substituted surrogate of the natural polyketide precursor. To this end, *N*-Boc-protected ethyl *D,L*-3-aminobutanoate was reduced with DibalH. The amino-aldehyde was immediately subjected to a Horner–Wadsworth–Emmons (HWE) olefination, and subsequent deprotection gave the desired amine **12** (see the Supporting Information). This synthetic surrogate of the growing polyketide chain was employed in the in vitro assay and product formation was monitored by HRMS. The formation of a new compound with $m/z = 172.0966$ and a molecular formula of $C_8H_{14}O_3N$ [$M+H$]⁺ as deduced from LC–HRMS suggested that the desired lactam compound was produced. To confirm the identity of the lactam, we prepared a synthetic reference. In brief, a protected *D,L*- β -amino acid ester was subjected to a reductive Horner–Wadsworth–Emmons olefination to obtain the substituted α,β -unsaturated ester. A cesium carbonate catalyzed Michael addition of diethylmalonate followed by hydrolysis yielded the lactam as a diastereoisomeric mixture (see the Supporting Information). Comparison of the high-resolution mass spectra and HPLC retention times of the enzyme products and the synthetic reference confirmed the successful conversion of **12** into the δ -lactam **13** (Figure 4). Furthermore, only the (late-eluting) *syn*-diastereomer is formed, thus confirming that the enzymatic reaction is highly selective, yielding only one diastereoisomer from a mixture of enantiomeric δ -amino-substituted precursors, as has been observed for the δ -lactones.

The successful biotransformation of the amino analogue encouraged an expansion of this approach to carbonyl substituents such as carboxylic acid and amide residues. In principle, a carboxylate could give rise to an anhydride moiety or the respective dicarboxylic acid, whereas the carboxamide would lead to the glutarimide moiety, which is found in many natural products. Unexpectedly, the synthesis of carboxy- and carboxamide-substituted SNAC derivatives proved to be highly challenging. Despite many attempts, the functionalization of glutaric acid and HWE olefination were not feasible, possibly owing to the highly reactive α -proton at the allylic position. Finally, we succeeded in preparing the

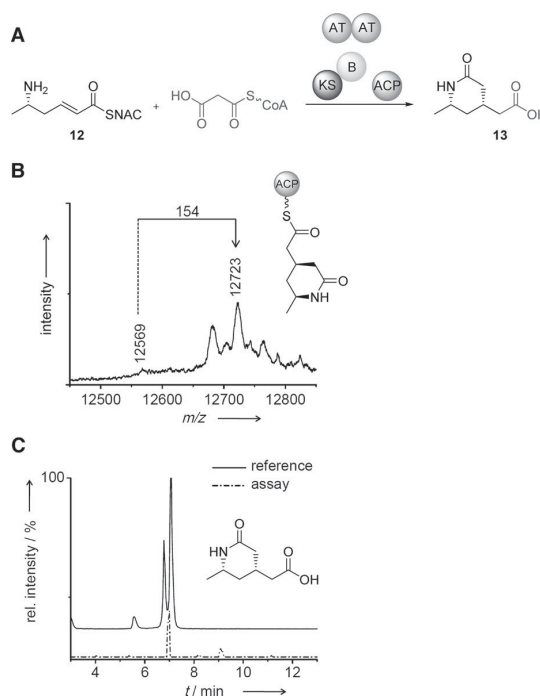


Figure 4. In vitro biotransformation of an amino-substituted polyketide surrogate into a lactam moiety. A) General Scheme of the enzyme assay. B) MALDI analysis of the ACP-bound product. C) SIM-LC–HRMS analysis of the hydrolyzed product and comparison with the synthetic reference (as a mixture of diastereoisomers); only one enantiomer is shown.

desired polyketide surrogates **14** and **16** (see the Supporting Information) through alkene metathesis of propenyl SNAC with but-3-enoic acid and but-3-enamide, respectively. Addition of the carboxy-substituted SNAC derivative **14** to the enzyme assay resulted in protein precipitation, and no biotransformation took place. To exclude the possibility of a pH effect, we repeated the experiment after accurately adjusting the pH value to 7.0 with Tris buffer. However, neither the respective anhydride **15** nor the hydrolyzed product could be detected.

By contrast, the carboxamide-substituted polyketide surrogate **16** was accepted as a substrate by the branching module. Through analysis of the enzyme mixture, we detected a MALDI signal corresponding to the ACP-bound glutarimide product (Figure 5). Moreover, acidic workup released the expected product with $m/z = 172.0610$ and a molecular formula of $C_7H_{10}O_4N$ (HRMS). The identity of the glutarimide **17** was rigorously confirmed through comparison of the HPLC retention times and high-resolution mass signals of the reaction product and a synthetic reference (Figure 5). A negative control experiment with a heat-inactivated KS–B didomain confirmed that the reaction is indeed enzyme-catalyzed.

Our results unequivocally show that the unusual KS–B–ACP module of the *rhi* PKS is not limited to lactone formation but is also capable of introducing branches that

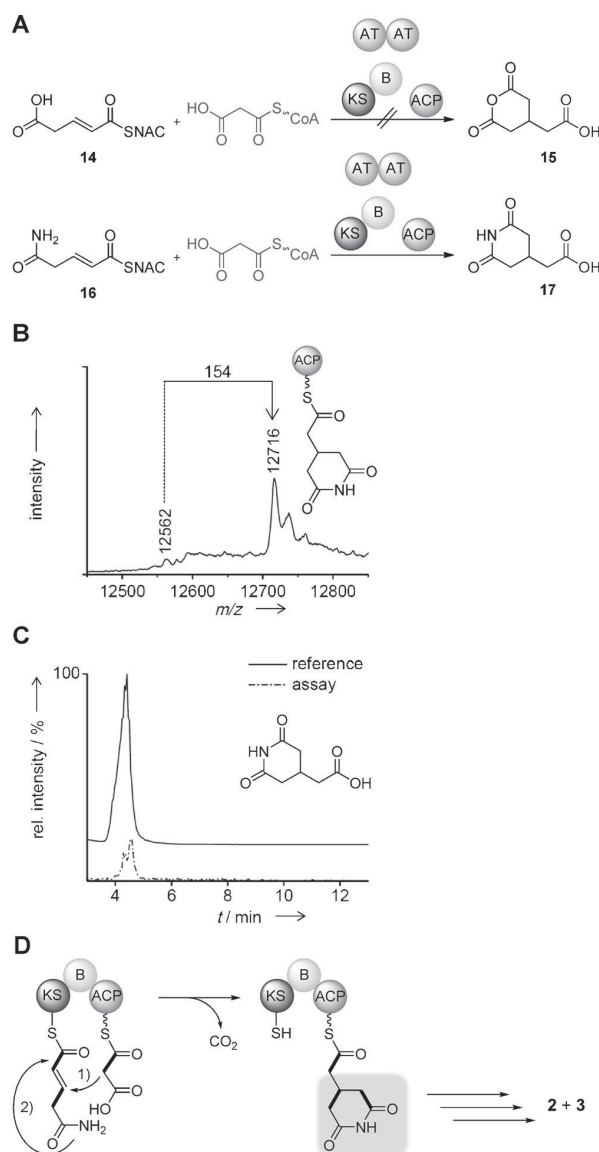


Figure 5. In vitro biotransformation of the amide-functionalized polyketide surrogate into a glutarimide moiety. A) General Scheme of the enzyme assay. B) MALDI analysis of the ACP-bound product. C) SIM-ILC-HRMS analysis of hydrolyzed product and comparison with the reference. D) Proposed model for the biosynthesis of glutarimide-bearing polyketides such as **2** and **3** (see Figure 1).

constitute lactam and glutarimide moieties. The latter are particularly noteworthy since related modules (KS-“X”-ACP, or KS-B-ACP) were found in PKS complexes involved in the biosynthesis of the glutarimide-bearing polyketides such as 9-methylstreptimidone,^[14] isomigrastatin^[12] and cycloheximide.^[15] The successful enzymatic synthesis of a glutarimide with the *rhi* PKS module for the first time confirms the proposed pathway involving Michael addition and heterocyclization.

In conclusion, we have shed light on the scope and versatility of a specialized PKS module that catalyzes the noncanonical vinylogous addition of a C2 extender unit to generate polyketide chain branches. Through the synthesis of a series of thioester surrogates with different δ -substituents, in vitro biotransformation experiments, and product analyses with synthetic references, we showed that the branching module generates lactone, lactam, and glutarimide heterocycles. We found that this set of domains exhibits a marked stereochemical substrate preference and that exclusively *syn*-substituted lactones and lactams are formed. However, there is a degree of flexibility that leaves room for various alternative conversions. Among these, the enzymatic formation of glutarimide provides first experimental support for postulated biosynthetic pathways to known glutarimide-bearing natural products. Although corresponding δ -lactam substructures have not yet been observed in natural polyketide structures, we postulate that the biosynthesis of polyketides with δ -lactam branches would be feasible and that such natural products await discovery. From a practical point of view, our findings are an important addition to our current understanding of non-terpenoid polyketide branching. Furthermore, the use of this branching module in synthetic biology or the engineering of enzymatic pathways could provide unique opportunities for compound diversification.

Received: July 16, 2014

Published online: September 11, 2014

Keywords: beta-branching · biotransformation · macrolides · Michael addition · polyketide synthases

- [1] C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716; *Angew. Chem.* **2009**, *121*, 4782–4811.
- [2] J. Hong, J. D. White, *Tetrahedron* **2004**, *60*, 5653–5681.
- [3] a) L. P. Partida-Martinez, C. Hertweck, *Nature* **2005**, *437*, 884–888; b) K. Scherlach, L. P. Partida-Martinez, H.-M. Dahse, C. Hertweck, *J. Am. Chem. Soc.* **2006**, *128*, 11529–11536; c) L. P. Partida-Martinez, I. Groth, M. Roth, I. Schmitt, K. Buder, C. Hertweck, *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 2583–2590; d) K. Scherlach, B. Busch, G. Lackner, U. Paszkowski, C. Hertweck, *Angew. Chem. Int. Ed.* **2012**, *51*, 9615–9618; *Angew. Chem.* **2012**, *124*, 9753–9756.
- [4] a) M. Takahashi, S. Iwasaki, H. Kobayashi, S. Okuda, T. Murai, Y. Sato, *Biochim. Biophys. Acta Gen. Subj.* **1987**, *926*, 215–223; b) I. Schmitt, L. P. Partida-Martinez, R. Winkler, K. Voigt, A. Einax, J. Wöstemeyer, C. Hertweck, *ISME J.* **2008**, *2*, 632–641; c) B. Kusebauch, K. Scherlach, H. Kirchner, H. M. Dahse, C. Hertweck, *ChemMedChem* **2011**, *6*, 1998–2001.
- [5] H. Kobayashi, S. Iwasaki, E. Yamada, S. Okuda, *J. Chem. Soc. Chem. Commun.* **1986**, 1702–1703.
- [6] a) Y. A. Chan, A. M. Podevels, B. M. Kevany, M. G. Thomas, *Nat. Prod. Rep.* **2009**, *26*, 90–114; b) M. C. Wilson, B. S. Moore, *Nat. Prod. Rep.* **2012**, *29*, 72–86.
- [7] a) C. T. Calderone, W. E. Kowtoniuk, N. L. Kelleher, C. T. Walsh, P. C. Dorrestein, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8977–8982; b) C. T. Calderone, *Nat. Prod. Rep.* **2008**, *25*, 845–853.
- [8] a) G. Lackner, N. Moebius, L. P. Partida-Martinez, C. Hertweck, *J. Bacteriol.* **2011**, *193*, 783–784; b) G. Lackner, N. Moebius, L. P.

- Partida-Martinez, S. Boland, C. Hertweck, *BMC Genomics* **2011**, *12*, 210.
- [9] a) L. P. Partida-Martinez, C. Hertweck, *ChemBioChem* **2007**, *8*, 41–45; b) T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck, J. Piel, *Nat. Biotechnol.* **2008**, *26*, 225–233.
- [10] a) B. Kusebauch, B. Busch, K. Scherlach, M. Roth, C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 5001–5004; *Angew. Chem.* **2009**, *121*, 5101–5104; b) B. Kusebauch, B. Busch, K. Scherlach, M. Roth, C. Hertweck, *Angew. Chem. Int. Ed.* **2010**, *49*, 1460–1464; *Angew. Chem.* **2010**, *122*, 1502–1506.
- [11] a) T. Bretschneider, J. B. Heim, D. Heine, R. Winkler, B. Busch, B. Kusebauch, T. Stehle, G. Zocher, C. Hertweck, *Nature* **2013**, *502*, 124–128; related highlight articles: b) C. A. Townsend, *Nature* **2013**, *502*, 44–45, c) R. J. Cox, *ChemBioChem* **2014**, *15*, 27–29.
- [12] S. K. Lim, J. Ju, E. Zazopoulos, H. Jiang, J. W. Seo, Y. Chen, Z. Feng, S. R. Rajsiki, C. M. Farnet, B. Shen, *J. Biol. Chem.* **2009**, *284*, 29746–29756.
- [13] S. R. Rajsiki, B. Shen, *ChemBioChem* **2010**, *11*, 1951–1954.
- [14] B. Wang, Y. Song, M. Luo, Q. Chen, J. Ma, H. Huang, J. Ju, *Org. Lett.* **2013**, *15*, 1278–1281.
- [15] M. Yin, Y. Yan, J. R. Lohman, S. X. Huang, M. Ma, G. R. Zhao, L. H. Xu, W. Xiang, B. Shen, *Org. Lett.* **2014**, *16*, 3072–3075.
-



Supporting Information

© Wiley-VCH 2014

69451 Weinheim, Germany

Enzymatic Polyketide Chain Branching to Give Substituted Lactone, Lactam, and Glutarimide Heterocycles**

*Daniel Heine, Tom Bretschneider, Srividhya Sundaram, and Christian Hertweck**

Supporting Information for

1. Instrumentation.....	2
2. General Methods.....	3
3. Protein Production.....	3
4. <i>In vitro</i> Enzymatic Biotransformations.....	4
5. Synthetic Protocols and Analytical Data.....	5
6. NMR Spectra.....	19
7. Original HRMS Data.....	32
8. Original MALDI Data.....	39
9. Supplemental References.....	43

1. Instrumentation

NMR measurements were performed on a Bruker AVANCE II 300, AVANCE III 500 or 600 MHz spectrometer, equipped with a Bruker Cryo platform. Chemical shifts are reported in parts per million (ppm) relative to the solvent residual peak of chloroform- d_1 (^1H : 7.24 ppm, singlet; ^{13}C : 77.00 ppm, triplet), methanol- d_4 (^1H : 3.30 ppm, quintet; ^{13}C : 49.00 ppm, septet) or DMSO- d_6 (^1H : 2.50 ppm, quintet; ^{13}C : 39.52 ppm, septet). MALDI measurements were performed with a Bruker UltrafleXtreme machine. HRESI-MS was carried out on an Accela UPLC-system (Thermo Scientific) combined with an Exactive mass spectrometer (Thermo Scientific) equipped with an electrospray ion source. Solid phase extraction was carried out using Chromabond C_{18} ec Cartridges filled with 2,000 mg of octadecyl-modified silica gel (Macherey-Nagel). HPLC-MS measurements were performed on a HPLC 1100 System connected to a 1100 Series LC/MSD Trap using a Zorbax Eclipse XDB-C8, 4.6 x 150 mm column, particle size 5 μm (Agilent Technologies). Semi-preparative HPLC was performed on a 1260 Infinity System (Agilent Technologies) using the following columns: Zorbax Eclipse XDB-C8, 9.5 x 250 mm, particle size 5 μm (Agilent Technologies), Zorbax Eclipse XDB-C18, 9.5 x 250 mm, particle size 5 μm (Agilent Technologies) and EC 250/4 Nucleodur Sphinx RP, particle size 5 μm (Macherey-Nagel). All solvents for analytical and preparative HPLC measurements were obtained commercially at least in gradient grade and were filtered prior to use. To avoid microbial growing 0.1 % formic acid was added to the water, used for analytical and preparative HPLC. IR-spectra were recorded on an FT/IR-4100 ATR spectrometer (Jasco).

2. General Methods

All reagents and glutarimide reference **17** were obtained from commercial suppliers (Sigma Aldrich and Acros Organics) and were used without further purification except where reported. All solvents were at least obtained in HPLC grade. Reactions were performed under inert atmosphere (Argon) using the common Schlenk technique. Solvents were dried prior to use by the following procedures: *N,N*-dimethylformamide (DMF) and dichloromethane (DCM) were distilled from a calcium hydride suspension. Tetrahydrofuran (THF) was pre-dried over potassium hydroxide and then freshly distilled from lithium aluminium hydride. Methanol, dichloromethane, chloroform and ethyl acetate were distilled prior to use. For thin-layer chromatography TLC aluminum sheets silica gel 60 F₂₅₄ (MERCK) were used. Open column chromatography was performed on silica gel 60; particle size 0.015 - 0.04 mm (Macherey-Nagel) and on Sephadex® LH-20 (Sigma-Aldrich).

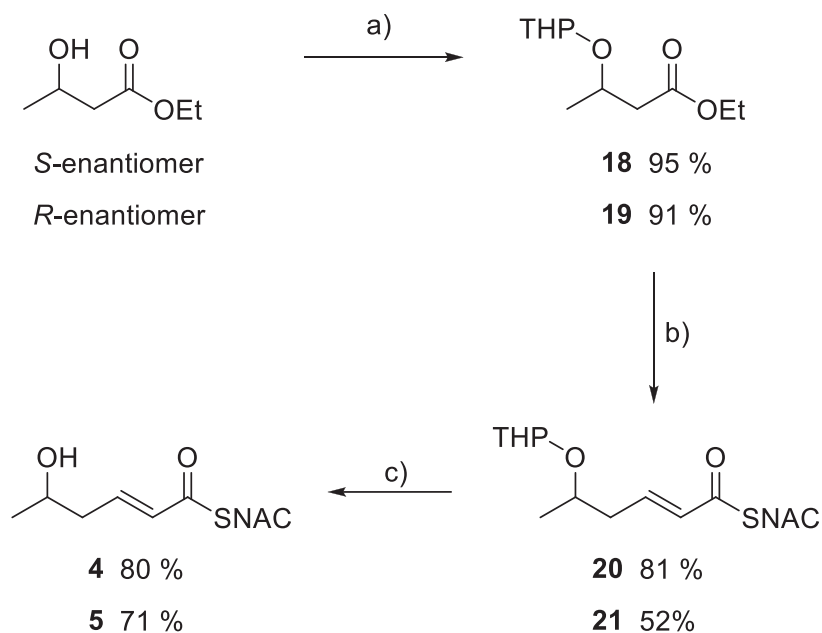
3. Protein Production

Recombinant RhiE*, consisting of the KS-B-bidomain, was produced and purified as described previously.^[1] Briefly, 5 mg of *Escherichia coli* BL21 cells containing the expression plasmids (pTB57, pTB41 and pTB39 respectively) were dissolved in 50 mL of 20 mM Tris (pH 7.5) supplemented with 200 mM NaCl and 50 mM imidazole. After disrupting the cells by sonication (sonotrode, 20% power), and removing the cell debris by centrifugation at 10,000 rpm for 20 min, the supernatant was loaded onto a 5 mL Ni-NTA matrix (Machery Nagel) connected to a FPLC machine (Äkta Explorer, Amersham Biosciences). The recombinant protein was eluted using the same buffer with 500 mM imidazole. The protein was diluted twice with 20 mM Tris (pH 7.5) and further purified using an ion-exchange column (5 ml HiTrap Q-HP, GE Healthcare). Elution was carried out with 30 column volumes of 20 mM Tris (pH 7.5) supplemented with 1 M NaCl. The fractions containing the target protein were concentrated using a vivaspin column (MWCO 5 & 10 kDa - Sartorius) and stored at 4 °C until use.

4. *In vitro* Enzymatic Biotransformations

In vitro biotransformations were conducted in analogy to the method established for the functional analysis of the branching module.^[1] At a 40 μL scale, 167 μM ACP, 3 μM RhiE*, 0.2 μM RhiG, 733 μM malonyl-CoA and 2,000 μM of the *N*-acetylcysteamine congeners were incubated in 20 mM Tris buffer (pH 7.0) at 23 °C with shaking at 400 rpm. After 2 h, 2 μL of the sample was mixed with 2 μL of 100 mM 2,5-dihydroxyacetophenone (containing 2.5 mM diammonium hydrogen citrate) and spotted onto an anchor chip 800/384 T F target (Bruker) to analyze products. For mass spectrometric measurements, the enzyme reaction was stopped by either adding 200 μL of ethyl acetate or an equal volume of methanol (in case of substrate **12**). The organic phase was recovered and concentrated to dryness under a stream of nitrogen gas. The residue was dissolved in 60 μL methanol and analyzed by HRMS (Exactive).

5. Synthetic Protocols and Analytical Data



Scheme S1. a) 3,4-dihydro-2*H*-pyran, pyridinium *p*-toluenesulfonate, CH₂Cl₂, 0 °C then RT; b) 1. DIBAL-H, toluene, -78 °C; 2. LiCl, *i*Pr₂NEt, *S*-(2-Acetamidoethyl)2-diethoxyphosphoryl)ethanethioate, acetonitrile, 0 °C then RT; c) *p*-toluenesulfonic acid, CH₃OH, RT. DIBAL-H = diisobutyl aluminium hydride.

(*S*)-Ethyl 3-((tetrahydro-2*H*-pyran-2-yl)oxy)butanoate **18^[1]**

To a solution of 3,4-dihydro-2*H*-pyran (1.9 g, 22.7 mmol) and 3-hydroxybutyric acid (2 g, 15.2 mmol) was slowly added pyridinium *p*-toluenesulfonate (380 mg, 1.52 mmol) in dichloromethane (5 mL). The solution was allowed to warm room temperature and stirred for additional 24 h. A saturated solution of sodium hydrogen carbonate (5 mL) was added and the aqueous phase was extracted with dichloromethane (3×10 mL). The combined organic phases were dried with sodium sulfate. After evaporation of the solvent and bulb to bulb distillation at 125 °C and 13 mbar, a colorless liquid (3.12 g) was obtained.

Yield: 95 %. ¹H NMR (300 MHz, CDCl₃): δ 4.71 (1H, m), 4.66 (1H, m), 4.25-4.05 (2H, m), 4.17 (4H, q, ³*J* = 6.3 Hz), 3.86 (2H, m), 3.48 (2H, m), 2.71-2.34 (4H, m), 1.81-1.47 (12 H, m), 1.27 (6H, d, ³*J* = 7.1 Hz), 1.22 ppm (6H, t, ³*J* = 6.3 HZ); [α]_D²⁴ = +14.29 (7.7 mg mL⁻¹, MeOH).

(R)-Ethyl 3-((tetrahydro-2H-pyran-2-yl)oxy)butanoate 19

The synthesis of **19** was performed by the same procedure as for compound **18** by using (*R*)-(+)-ethyl-3-hydroxybutyrate (2 g, 15.2 mmol). After workup and bulb to bulb distillation, 2.99 g of a colorless liquid could be obtained.

Yield: 91 %. NMR data are identical to those of compound **18**; $[\alpha]_D^{24} = -15.48$ (5.8 mg mL⁻¹, MeOH).

(S)-(E)-S-(2-Acetamidoethyl)-5-((tetrahydro-2H-pyran-2-yl)oxy)hex-2-enethioate 20^[1]

A stirred solution of **18** (60 mg, 0.28 mmol) in dry toluene (1 mL) was cooled to 78 °C and diisobutylaluminium hydride (DIBAL-H) (1 M in *n*-hexane, 7.5 mL) was added slowly. A mixture of methanol/1 M HCl (2:1; 4 mL) was added after 3 h, and the solution was filtered through celite. The filtrate was extracted with ethyl acetate (3×10 mL) and the combined organic phases were dried with sodium sulfate and concentrated under reduced pressure. The aldehyde could be used for the next step without further purification steps.

LiCl (13 mg, 0.3 mmol) was suspended together with **2** (50 mg, 0.16 mmol) in dry acetonitrile (2 mL). After addition of *N,N*-diisopropylethylamine (50 µl, 0.3 mmol) the mixture was cooled to 0 °C and the aldehyde, dissolved in acetonitrile (2 mL), was added. After warming to room temperature, the solution was stirred for 17 h. Saturated ammonium chloride solution (20 mL), water (20 mL) and diethyl ether (20 mL) were added, and the aqueous phase was extracted with diethyl ether (3×20 mL). The combined organic phases were dried with sodium sulfate and concentrated under reduced pressure. After column chromatography (chloroform/methanol 100:0 to 95:5) a colorless oil (72 mg) was obtained.

Yield: 81 %. ¹H NMR (300 MHz, CDCl₃): δ 6.97 (1H, dt, ³*J* = 15.6 Hz, ⁴*J* = 7.3), 6.88 (1H, dt, ³*J* = 15.6 Hz, ⁴*J* = 7.3), 6.18 (1H, dt, ³*J* = 15.6 Hz, ⁴*J* = 1.4 Hz), 6.14 (1H, dt, ³*J* = 15.6 Hz, ⁴*J* = 1.4 Hz), 5.90 (2H, br s), 4.69 (1H, dd, ³*J*₁ = 4.4 Hz, ³*J*₂ = 2.9 Hz), 4.61 (1H, dd, ³*J*₁ = 4.4 Hz, ³*J*₂ = 2.9 Hz), 4.00-3.80 (4H, m), 3.49-3.45 (2H, m), 3.44 (4H, m), 3.07 (4H, t, ³*J* = 6.4 Hz), 2.49-2.31 (4H, m), 1.94 (3H, s), 1.94 (3H, s), 1.71-1.48 (12H, m), 1.23 (3H, d, ³*J* = 6.3 Hz), 1.12 ppm (3H, d, ³*J* = 6.3 Hz); $[\alpha]_D^{22} = +7.05$ (c = 4.2 mg mL⁻¹, MeOH).

(R)-(E)-S-(2-Acetamidoethyl)-5-((tetrahydro-2H-pyran-2-yl)oxy)hex-2-enethioate 21

Compound **21** was prepared by the same procedure used for the synthesis of compound **20** by using **19** (60 mg, 0.28 mmol). Purification by column chromatography (chloroform/methanol

100/0 to 95/5) gave 46 mg of a colorless oil. Yield: 52 %. NMR data are identical to those of compound **20**; Yield: 52 %. $[\alpha]_D^{24} = -8.48$ ($c = 4.7 \text{ mg mL}^{-1}$, MeOH).

(S)-(E)-S-(2-Acetamidoethyl) 5-hydroxyhex-2-enethioate 4^[1]

The protected ester **20** (63 mg, 0.2 mmol) was dissolved in methanol (2 mL) and *p*-toluenesulfonic acid monohydrate (4 mg, 0.02 mmol) in methanol (1 mL) was added dropwise. After 16 h saturated sodium hydrogen carbonate solution (3 mL) and water (3 mL) were added. The aqueous phase was extracted with ethyl acetate (3×5 mL) and the combined organic fractions were dried with sodium sulfate and concentrated under reduced pressure. Column chromatography over silica gel (chloroform/methanol 95:5) gave a colorless oil (37 mg).

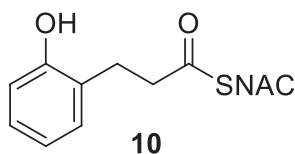
Yield: 80 %. ¹H NMR (600 MHz, CDCl₃): δ 6.92 (1H, dt, ³*J* = 15.6 Hz, ⁴*J* = 7.5 Hz), 6.20 (1H, dt, ³*J* = 15.6 Hz, ⁴*J* = 1.4 Hz), 3.97 (1H, m), 3.45 (2H, q, ³*J* = 6.3), 3.08 (2H, t, ³*J* = 6.4 Hz), 2.36-2.33 (2H, m), 1.97 (3H, s), 1.23 ppm (3H, d, ³*J* = 6.3 Hz); $[\alpha]_D^{22} = +8.66$ ($c = 1 \text{ mg mL}^{-1}$, MeOH).

(R)-(E)-S-(2-Acetamidoethyl) 5-hydroxyhex-2-enethioate 5

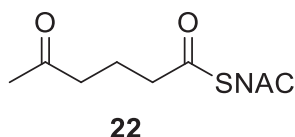
Compound **5** was synthesized from **22** (4.7 mg, 0.015 mmol) by the same procedure as described for the *S*-enantiomer **4**. Isolation of the product gave a colorless oil (2.4 mg). Yield: 71 %. NMR data are identical to those of compound **4**; $[\alpha]_D^{22} = -6.77$ ($c = 1 \text{ mg mL}^{-1}$, MeOH).

General Procedure for the Synthesis of SNAC-Thioesters

The free carboxylic acid (2 mmol) was dissolved or suspended in 15 mL of dry dichloromethane, and the mixture was cooled to 0 °C. Subsequently *N*-acetylcysteamine (2 mmol), *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride (2.5 mmol) and 4-dimethylaminopyridine (0.2 mmol) were added. The solution was slowly allowed to warm to room temperature and stirred for 12 hours. Water (20 mL) and dichloromethane (20 mL) were added to the solution, and the aqueous phase was extracted with dichloromethane (2×20 mL). The combined organic phases were dried over sodium sulfate and concentrated under reduced pressure. The product was subsequently purified by column chromatography on silica gel (gradient elution: chloroform/methanol: 98:2 to 90:10).

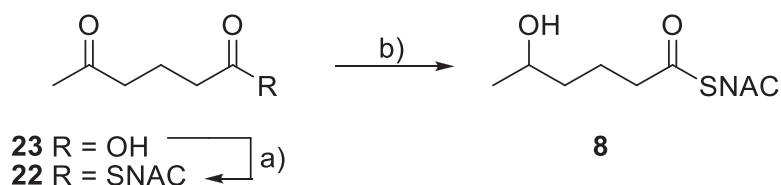
(*E*)-*S*-(2-Acetamidoethyl) 3-(2-hydroxyphenyl)prop-2-enethioate 10

Yield: 84 %. Yellow solid; m.p. 150.4 – 151.2 °C; ¹H NMR (500 MHz, MeOD): δ 7.88 (1H, d; ³*J* = 3.1 Hz), 7.49 (1H, dd, ³*J*₁ = 8.0 Hz, ³*J*₂ = 1.6 Hz), 7.22 (1H, m), 6.93 (1H, d, ³*J*_{trans} = 15.9 Hz), 6.84 (2H, m), 3.37 (2H, t, ³*J* = 6.7 Hz), 3.11 (2H, t, ³*J* = 6.7 Hz), 1.93 ppm (3H, s); ¹³C NMR (125 MHz, MeOD): δ 191.6, 173.5, 158.9, 138.4, 133.0, 130.5, 125.5, 122.2, 120.8, 117.1, 40.4, 29.1, 22.5 ppm; IR (film): 3363, 1641, 1600, 1446, 1263, 1020, 973, 743 cm⁻¹; HRMS: (ESI+): *m/z* calculated for C₁₃H₁₆O₃NS: 266.0845, found 266.0845 [M+H]⁺.

***S*-(2-Acetamidoethyl) 5-oxohexanethioate 22**

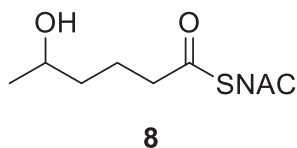
Yield: 90 %. Colorless liquid; ¹H NMR (500 MHz, CDCl₃): δ 6.26 (1H, bs), 3.40 (2H, q, ³*J* = 6.4 Hz), 3.00 (1H, t, ³*J* = 6.4 Hz), 2.57 (2H, t, ³*J* = 7.1 Hz), 2.48 (2H, t, ³*J* = 7.1 Hz), 2.11 (3H, s), 1.98 (3H, s), 1.90 ppm (2H, quint, ³*J* = 7.1 Hz); ¹³C NMR (125 MHz, CDCl₃): δ

207.8, 199.4, 170.8, 42.9, 42.1, 39.6, 29.9, 28.4, 22.9, 19.4 ppm; IR (film): 3289, 3097, 2951, 1705, 1679, 1636, 1557, 1441, 1362, 1295, 987 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_{10}\text{H}_{18}\text{O}_3\text{NS}$: 232.1002, found 232.0999 $[\text{M}+\text{H}]^+$.



Scheme S2. a) *N*-acetylcysteamine, DMAP, EDCI, CH_2Cl_2 , 0 °C then RT, 90 %; b) NaBH_4 , CeCl_3 , THF, 0 °C, 66 %. DMAP = 4-dimethylaminopyridine, EDCI = *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride.

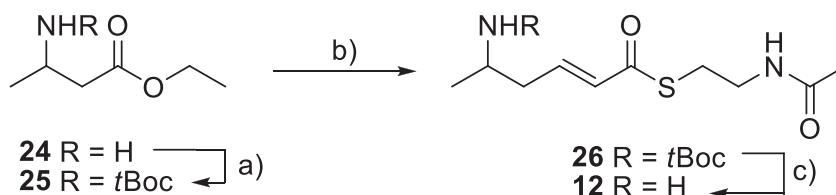
S-(2-Acetamidoethyl) 5-hydroxyhexanethioate **8**



To an ice-cooled, stirred suspension of *S*-(2-acetamidoethyl) 5-oxohexanethioate **22** (35 mg, 0.15 mmol) and CeCl_3 (100 mg, 0.4 mmol, 1.5 equivalents) in dry THF (0.5 mL) was given NaBH_4 (10 mg, 0.27 mmol). After 2 hours, the reaction mixture was cooled to -78 °C and methanol (0.2 mL) was added. After 30 minutes water (1 mL) and ethyl acetate (1 mL) were added. The aqueous layer was extracted with ethyl acetate (2×1 mL). The combined organic fractions were dried with Na_2SO_4 and the solvent evaporated under reduced pressure. After silica gel column chromatography ($\text{CHCl}_3/\text{MeOH}$ 98:2) a colorless liquid oil (23 mg) was obtained.

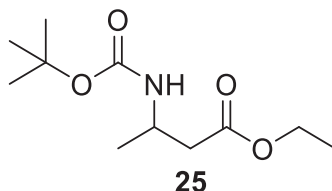
Yield: 66 %. Colorless oil; ^1H NMR (500 MHz, MeOD): δ 7.87 (1H, s), 3.64-3.75 (1H, m), 3.29 (3H, m), 2.98 (2H, t, $^3J = 6.7$ Hz), 2.58 (2H, t, $^3J = 7.5$ Hz), 1.89 (3H, s), 1.69 (2H, m), 1.38-1.45 (2H, m), 1.12 ppm (3H, d, $^3J = 6.1$ Hz); ^{13}C NMR (125 MHz, MeOD): δ 200.5, 173.4, 68.1, 44.7, 40.1, 39.1, 29.1, 23.5, 23.0, 22.5 ppm; IR (film): 3304, 2944, 2832, 1658,

1558, 1448, 1023 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_{10}\text{H}_{20}\text{O}_3\text{NS}$: 234.1158, found 234.1159 $[\text{M}+\text{H}]^+$.



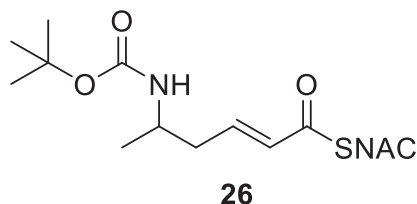
Scheme S3. a) $(\text{Boc})_2\text{O}$, THF, RT, 73 %; b) 1. DIBAL-H, toluene, $-78\text{ }^\circ\text{C}$; 2. S-(2-acetamidoethyl) 2-(diethoxyphosphoryl) ethanethioate, LiCl, Et_3N , acetonitrile, $0\text{ }^\circ\text{C}$ then RT, 2 steps 90 %; c) TFA, THF, RT, 96 %. Boc = *tert*butoxycarbonyl, TFA = trifluoroacetic acid.

Ethyl 3-((*tert*-butoxycarbonyl)amino)butanoate **25**



To a stirred solution of (*R,S*)-ethyl 3-aminobutanoate (131 mg, 1 mmol) in dry THF di-*t*-butyl dicarbonate (240 mg, 1.1 mmol) was added in small portions. After 12 h, the solvent was removed under reduced pressure, and the product was purified by column chromatography on silica gel ($\text{CHCl}_3/\text{MeOH}$ 98:2) to give 170 mg of a colorless oil.

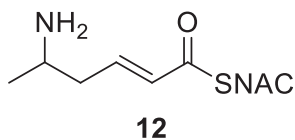
Yield: 73 %. Colorless oil; ^1H NMR (500 MHz, CDCl_3): δ 4.90 (1H, bs), 4.12 (2H, q, $^3J = 7.1$ Hz), 4.01 (1H, bs), 2.40-2.52 (2H, m), 1.41 (9H, s), 1.24 (3H, t, $^3J = 7.1$ Hz), 1.19 ppm (3H, d, $^3J = 6.8$ Hz).

S-(2-acetamidoethyl) (E)-5-((tert-butoxycarbonyl)amino)hex-2-enethioate 26

A solution of (*R,S*)-ethyl 3-((*tert*-butoxycarbonyl)amino)butanoate **25** (50 mg, 0.2 mmol) in dry toluene (1.5 mL) was cooled to $-78\text{ }^{\circ}\text{C}$, and DibalH solution (1 mL, 1 M in cyclohexane) was added dropwise. After 2 h the reaction was stopped by adding a mixture of methanol and HCl (1 M) (2:1, 4 mL). After warming to room temperature, the solution was filtered through Celite and washed thoroughly with ethyl acetate. The aqueous phase was extracted with ethyl acetate (2 \times 25 mL), and the combined organic fractions were dried over sodium sulfate. After removal of the solvent under reduced pressure, the crude product was directly used for the next step.

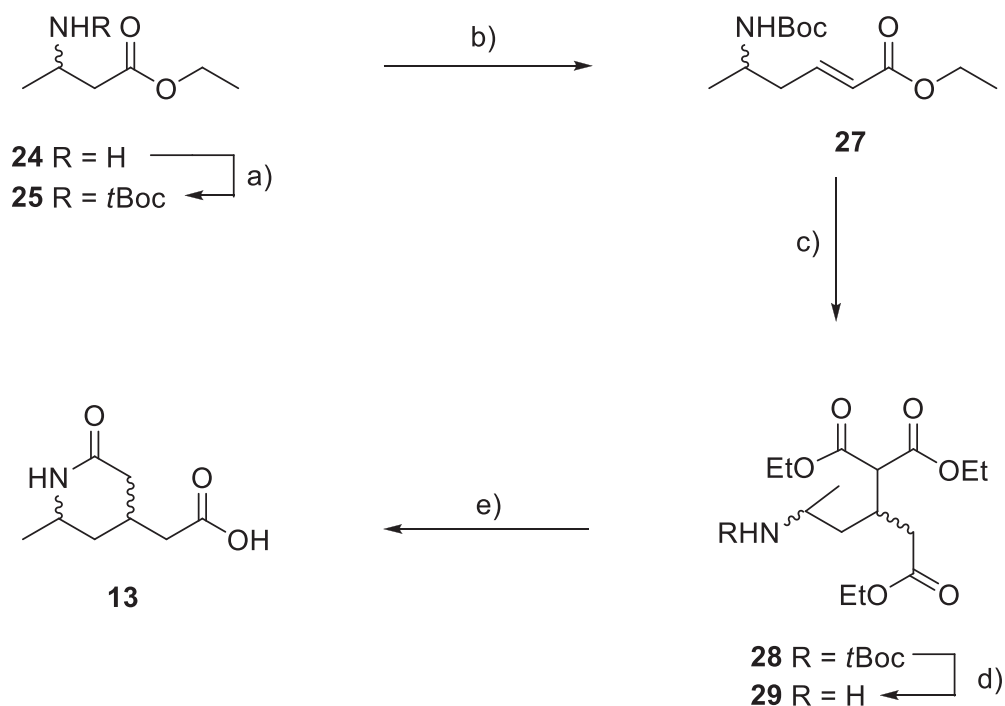
Phosphonate **1** (50 mg, 0.17 mmol) was dissolved in dry acetonitrile (1 mL) and given to a suspension of LiCl (8 mg, 0.2 mmol) in dry acetonitrile (5 mL). The mixture was cooled to $0\text{ }^{\circ}\text{C}$, and *N,N*-diisopropylethylamine (80 μL) was added dropwise. After the crude product of the first previous step was given to the mixture, the stirred solution was allowed to warm to room temperature. The reaction was quenched by adding a saturated ammonium chloride solution (5 mL), and the crude product was extracted with ethyl acetate (3 \times 20 mL). The combined organic fractions were washed with water (50 mL) and dried over sodium sulfate. The product (60 mg) was obtained after column chromatography over silica gel ($\text{CHCl}_3/\text{MeOH}$ 95:5).

Yield: 90 %. Colorless liquid; ^1H NMR (300 MHz, MeOD): δ 6.83 (1H, dt, $^3J_1 = 15.5\text{ Hz}$, $^3J_2 = 7.7\text{ Hz}$), 6.30 (1H, bs), 6.12 (1H, d, $^3J = 15.5\text{ Hz}$), 4.41 (1H, bs), 4.14 (1H, m), 3.42 (2H, m), 3.06 (2H, d, $^3J = 6.3\text{ Hz}$), 2.33 (2H, t, $^3J = 6.8\text{ Hz}$), 1.96 (3H, s), 1.40 (9H, s), 1.14 ppm (3H, d, $^3J = 7.2\text{ Hz}$); ^{13}C NMR (75 MHz, MeOD): δ 190.0, 170.4, 155.7, 141.9, 130.5, 79.5, 58.9, 39.7, 28.2, 23.1, 21.0, 20.9, 18.9 ppm; IR (film): 3305, 2938, 2828, 1661, 1538, 1367, 1168, 1025 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_{15}\text{H}_{27}\text{O}_4\text{N}_2\text{S}$: 331.17, found 331.17 $[\text{M}+\text{H}]^+$.

S-(2-Acetamidoethyl) (E)-5-aminohex-2-enethioate 12

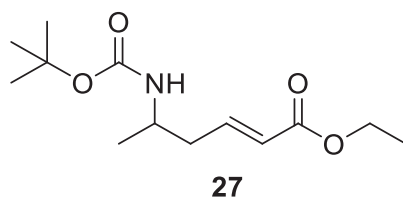
To a stirred solution of (*R,S*)-S-(2-acetamidoethyl) (*E*)-5-((*tert*-butoxycarbonyl)amino)hex-2-enethioate **26** (17 mg, 0.05 mmol) in dry THF trifluoroacetic acid (24 μ L, 0.3 mmol) was added at room temperature. After stirring for 48 h the solvent was removed under reduced pressure, and the product was purified by column chromatography (CHCl₃/MeOH 95:5) in a Pasteur pipette to give 11 mg of colorless oil.

Yield: 96 %. Colorless liquid; ¹H NMR (300 MHz, MeOD): δ 6.83 (1H, dt; ³*J* = 15.5 Hz; ³*J* = 7.4 Hz), 6.33 (1H, d, ³*J*₁ = 15.5 Hz), 3.62-3.82 (1H, m), 3.35 (2H, m), 3.07 (2H, t, ³*J* = 6.70 Hz), 2.40-2.65 (2H, m), 1.90 (3H, s), 1.30 ppm (3H, d, ³*J* = 6.70 Hz); ¹³C NMR (75 MHz, MeOD): δ 190.4, 173.5, 139.5, 133.1, 59.6, 40.0, 38.0, 29.1, 22.5, 18.6 ppm; IR (film): 3315, 2945, 2834, 1679, 1427, 1203, 1141, 1021 cm⁻¹; HRMS: (ESI+): *m/z* calculated for C₁₀H₁₉O₂N₂S: 231.1162, found 231.1163 [M+H]⁺.



Scheme S4. a) $(\text{Boc})_2\text{O}$, THF, RT, 73 %; b) 1. DIBAL-H, toluene, 78 °C; 2. LiCl, triethyl phosphonoacetate, $i\text{Pr}_2\text{NEt}$, acetonitrile, 0 °C then RT, 2 steps, 80 %; c) diethyl malonate, Cs_2CO_3 , acetonitrile, 80 °C, 91 %; d) TFA, CH_3OH , RT, 88 %; e) 2 M HCl, 100 °C, 74 %.

***tert*-Butyl (*E*)-(6-ethoxy-6-oxohex-4-en-2-yl)-12-azanecarboxylate (27)**

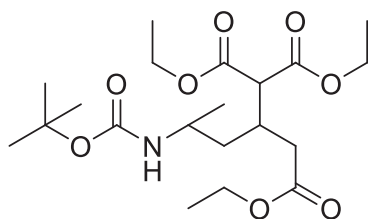


Ethyl 3-((*tert*-butoxycarbonyl)amino)butanoate (941 mg, 4 mmol) was dissolved in toluene (5 mL) and cooled to -78 °C. Dibal (1 M in cyclohexane, 6 mL) was added dropwise, and the mixture was stirred for additional 2 h. After adding a solution of methanol and HCl (1 M) (2:1, 6 mL) the solution was warmed to room temperature and filtered through Celite and washed with diethyl ether (3×40 mL). The combined organic phases were dried over sodium sulfate and concentrated under reduced pressure. LiCl (168 mg, 4 mmol) and triethyl

phosphonoacetate (896 mg, 4 mmol) were suspended in dry acetonitrile (10 mL) and *N,N*-diisopropylethylamine (680 μ L, 4 mmol) was added to the mixture. After cooling to 0 °C, a solution of the freshly prepared aldehyde in acetonitrile (1 mL) was added. The mixture was allowed to warm to room temperature and stirred for 12 h. Saturated ammonium chloride solution (10 mL), water (10 mL) and diethyl ether (30 mL) were added, and the aqueous phase was extracted with diethyl ether (3 \times 40 mL). The combined organic phases were dried over sodium sulfate and concentrated under reduced pressure. Column chromatography (chloroform/methanol 100:0 to 95:5) gave 822 mg of viscous oil.

Yield: 80 %. Colorless oil; ^1H NMR (300 MHz, MeOD): δ 6.86-6.96 (1H, dt, $^3J = 15.6$ Hz, $^3J = 7.5$ Hz), 5.85 (1H, d, $^3J = 15.6$ Hz), 4.15 (2H, q, $^3J = 7.1$ Hz), 3.64-3.75 (1H, m), 3.29-3.31 (1H, m), 2.32 (2H, q, $^3J = 7.5$ Hz), 1.41 (9H, s), 1.26 (3H, t, $^3J = 7.1$ Hz), 1.12 ppm (3H, d, $^3J = 6.7$ Hz); ^{13}C NMR (75 MHz, CDCl_3): δ 166.6, 156.4, 145.8, 122.8, 78.5, 60.0, 45.5, 39.1, 27.3, 19.6, 13.1 ppm; IR (film): 3333, 2942, 2832, 1688, 1528, 1448, 1368, 1272, 1119 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_{13}\text{H}_{24}\text{O}_4\text{N}$: 258.1700, found 258.1697 $[\text{M}+\text{H}]^+$.

Triethyl 2-(2-((*tert*-butoxycarbonyl)amino)propyl)propane-1,1,3-tricarboxylate **28**



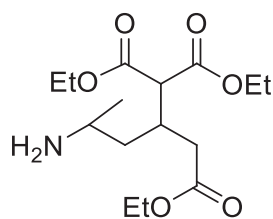
28

Tert-butyl (*E*)-(6-ethoxy-6-oxohex-4-en-2-yl)-12-azanecarboxylate (**27**) (26 mg, 0.1 mmol) was dissolved in dry acetonitrile (0.5 mL). Diethyl malonate (32 mg, 0.2 mmol) and cesium carbonate (16 mg, 0.05 mmol) were added, and the mixture was stirred for 10 h at 80 °C. Non-dissolved components were filtered off, and the filtrate was concentrated under reduced pressure. Column chromatography (*n*-hexane/ethyl acetate 90:10) gave 38.5 mg of clear oil.

Yield: 91 %. Colorless oil; ^1H NMR (500 MHz, CDCl_3): δ 4.2-4.08 (8H, m), δ 3.67 (1H, d, $^3J = 6.0$ Hz), 2.66-2.51 (2H, m), 1.62 (1H, bs), 1.55-1.45 (2H, m), 1.42 (9H, s), 1.26-1.22 (9H, m), 1.11 ppm (3H, d, $^3J = 6.4$ Hz); ^{13}C NMR (125 MHz, CDCl_3): δ 172.0, 168.5, 168.4, 155.5, 62.6, 61.2, 60.3, 53.8, 39.4, 35.9, 31.6, 28.3, 21.6, 16.2, 14.1, 14.0 ppm (data given for the

main stereoisomer); IR (film): 2978, 1729, 1520, 1448, 1367, 1247, 1158 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_{20}\text{H}_{36}\text{O}_8\text{N}$: 418.2435, found 418.2437 $[\text{M}+\text{H}]^+$.

Triethyl 2-(2-aminopropyl)propane-1,1,3-tricarboxylate **29**

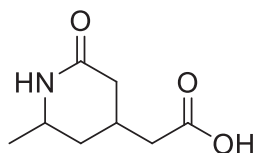


29

Compound **28** (12 mg, 0.029 mmol) was dissolved in dichloromethane (0.5 mL), and trifluoroacetic acid (20 μL) was added cautiously. After 5 h the solvent was removed under a stream of nitrogen gas, and the product was purified by small-scale column chromatography ($\text{CHCl}_3/\text{MeOH}$ 98:2 to 95:5) to give 8 mg of a viscous liquid.

Yield: 88 %. Colorless oil; ^1H NMR (500 MHz, MeOD): δ 4.24-4.12 (6H, m), 3.66 (1H, d, $^3J = 5.6$ Hz), 3.36 (1H, m), 2.67 (1H, m), 2.65-2.56 (1H, m), 2.51-2.45 (1H, m), 1.83-1.74 (1H, m), 1.71-1.60 (1H, m), 1.31 (3H, d, $^3J = 6.5$ Hz) 1.28-1.24 ppm (9H, m); ^{13}C NMR (125 MHz, MeOD): δ 174.0, 173.6, 169.9, 169.8, 169.8, 169.7, 62.8, 62.7, 62.0, 61.9, 55.7, 55.3, 47.2, 47.0, 38.7, 38.3, 37.2, 37.0, 37.0, 32.3, 32.3, 19.3, 18.5, 14.5, 14.4, 14.4 ppm; IR (film): 3335, 2949, 2836, 1672, 1402, 1203, 1142 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_{15}\text{H}_{28}\text{O}_6\text{N}$: 318.1911, found 318.1905 $[\text{M}+\text{H}]^+$.

2-(2-Methyl-6-oxopiperidin-4-yl)acetic acid **13**

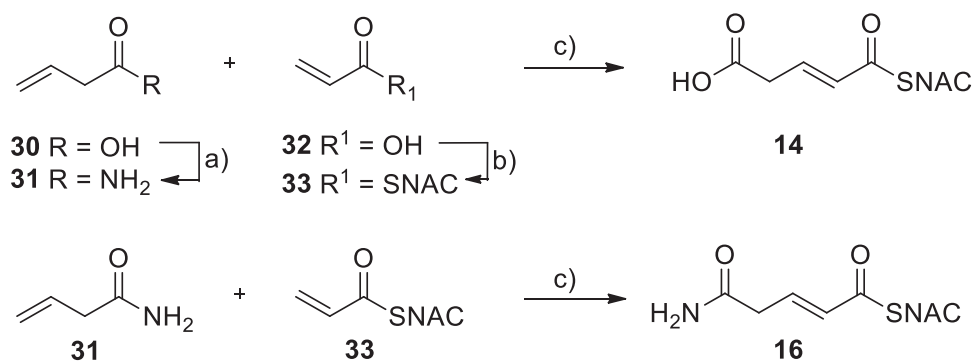


13

Compound **29** (7.8 mg, 0.025 mmol) was suspended in HCl (2 M, 400 μL) and heated to 100 $^\circ\text{C}$ in a clear glass vial (1.5 mL) closed with a screw cap. After 28 hours the solution was cooled to room temperature and water (1 mL) was added. After washing with ethyl acetate

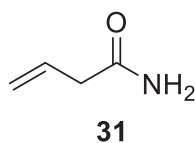
(3×1 mL) the aqueous phase was concentrated in an argon flow to give 3.1 mg (0.018 mmol) of clear oil.

Yield: 74 %. Colorless liquid; mixture of diastereoisomers, NMR signals are given for the main diastereoisomer, ^1H NMR (600 MHz, D_2O): δ 3.32-3.26 (1H, m), 2.39-2.33 (3H, m), 2.32-2.23 (2H, m), 1.61-1.44 (2H, m), 1.17 ppm (3H, d, $^3J = 6.6$ Hz); ^{13}C NMR (150 MHz, D_2O): δ 174.5, 173.2, 47.0, 43.7, 36.7, 36.4, 35.7, 26.3, 15.8 ppm; IR (film): 3320, 2945, 2832, 1447, 1415, 1022 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_8\text{H}_{14}\text{O}_3\text{N}$: 172.0968, found 172.0969 $[\text{M}+\text{H}]^+$.



Scheme S5. a) 1. SOCl_2 , CH_2Cl_2 , RT; 2. NH_3 in MeOH, -78 °C, 2 steps, 84 %; b) *N*-acetylcysteamine, DMAP, EDC, CH_2Cl_2 , 0 °C then RT, 66 %, c) Grubbs II, CH_2Cl_2 , RT.

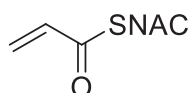
But-3-enamide 31



But-3-enoic acid (516 mg, 6 mmol) was suspended in dry dichloromethane (10 mL), and thionyl chloride was added dropwise. The solution was stirred for 4 h after the addition of DMF (50 μL). The solution was cooled to -78 °C, and a solution of ammonia in MeOH (7 N, 6 mL) was added. The solution was allowed to warm to room temperature. After 15 hours it was concentrated under reduced pressure. Column chromatography over silica gel ($\text{CHCl}_3/\text{MeOH}$ 98:2 to 7:1) gave pure amide **31** (426 mg). All analytical data were identical to those reported earlier.^[2]

Yield: 84 %. Colorless liquid, ^1H NMR (300 MHz, CDCl_3): δ 5.92 (1H, m), 5.89 (2H, bs), 5.25 (1H, m), 5.21 (1H, m), 3.02 ppm (2H, dt, $^3J_1 = 7.1$ Hz, $^4J_2 = 1.2$ Hz); HRMS: (ESI+): m/z calculated for $\text{C}_8\text{H}_{15}\text{O}_2\text{N}_2$: 171.1128, found 171.1127 $[\text{2M}+\text{H}]^+$.

S-(2-Acetamidoethyl) prop-2-enethioate 33

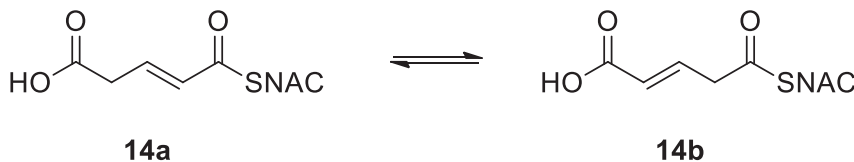


33

Acrylic acid (288 mg, 4 mmol), *N*-acetylcysteamine (476 mg, 4 mmol) and DMAP (48 mg, 0.4 mmol) were dissolved in dry dichloromethane under argon. The solution was cooled to 0 °C, and *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide hydrochloride (840 mg, 4.4 mmol) was added. The solution was stirred for 15 h at room temperature. The white slurry was concentrated under reduced pressure, and the residue was subjected to column chromatography using silica gel ($\text{CHCl}_3/\text{MeOH}$ 95:5) to give 453 mg of a viscous liquid.

Yield: 66 %. Colorless liquid, ^1H NMR (600 MHz, CDCl_3): δ 6.38 (1H, dd, $^3J_1 = 17.2$ Hz, $^3J_2 = 10.0$ Hz), 6.31 (1H, dd, $^3J = 17.2$ Hz, $^2J = 1.0$ Hz), 5.88 (1H, bs), 5.72 (1H, dd, $^3J = 10.0$ Hz, $^2J = 1.0$ Hz), 3.46 (2H, q, $^3J = 6.20$ Hz), 3.10 (2H, t, $^3J = 6.20$ Hz), 1.95 ppm (3H, s); ^{13}C NMR (75 MHz, MeOD): δ 190.8, 170.6, 134.7, 127.2, 39.7, 28.3, 23.1 ppm; IR (film): 3304, 2946, 2835, 1658, 1559, 1399, 1021 cm^{-1} ; HRMS: (ESI+): m/z calculated for $\text{C}_7\text{H}_{12}\text{O}_2\text{NS}$: 174.0583, found 174.0583 $[\text{M}+\text{H}]^+$.

(*E*)-5-((2-acetamidoethyl)thio)-5-oxopent-3-enoic acid 14

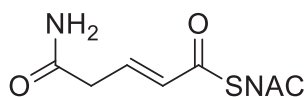


But-3-enoic acid (2.5 mg, 0.029 mmol) and thioester **33** (5 mg, 0.029 mmol) were dissolved in dry dichloromethane (0.5 mL) under argon. After addition of Hoveyda-Grubbs catalyst 2nd generation (7 mg, 0.01 mmol) the solution was stirred for 3 h. The solvent was evaporated

under reduced pressure, and the residue subjected to column chromatography on silica gel (CHCl₃/MeOH 100:0 to 90:10) to give a white solid.

Yield: 48 % (based on recovered starting material **33**). Compound **16a** appears as a mixture with its isomer **16b**. All attempts to obtain the pure compound **16a** failed due to a fast equilibrium at room temperature. White solid, ¹H NMR (500 MHz, MeOD): δ 6.91-6.70 (1H, m), 6.27 (1H, dt, ³J = 15.6 Hz, ³J = 1.6 Hz), 3.31-3.35 (2H, m), 3.30 (2H, m), 3.07 (2H, t, ³J = 6.6 Hz), 1.91 ppm (3H, s); ¹³C NMR (125 MHz, MeOD): δ 196.8, 190.6, 173.5, 140.8, 138.8, 40.0, 29.4, 29.0, 22.5 ppm; IR (film): 3325, 2948, 2835, 1660, 1449, 1412, 1112 cm⁻¹; HRMS: (ESI+): *m/z* calculated for C₉H₁₄O₄NS: 232.0644, found 232.0631 [M+H]⁺.

S-(2-Acetamidoethyl) (E)-5-amino-5-oxopent-2-enethioate 16

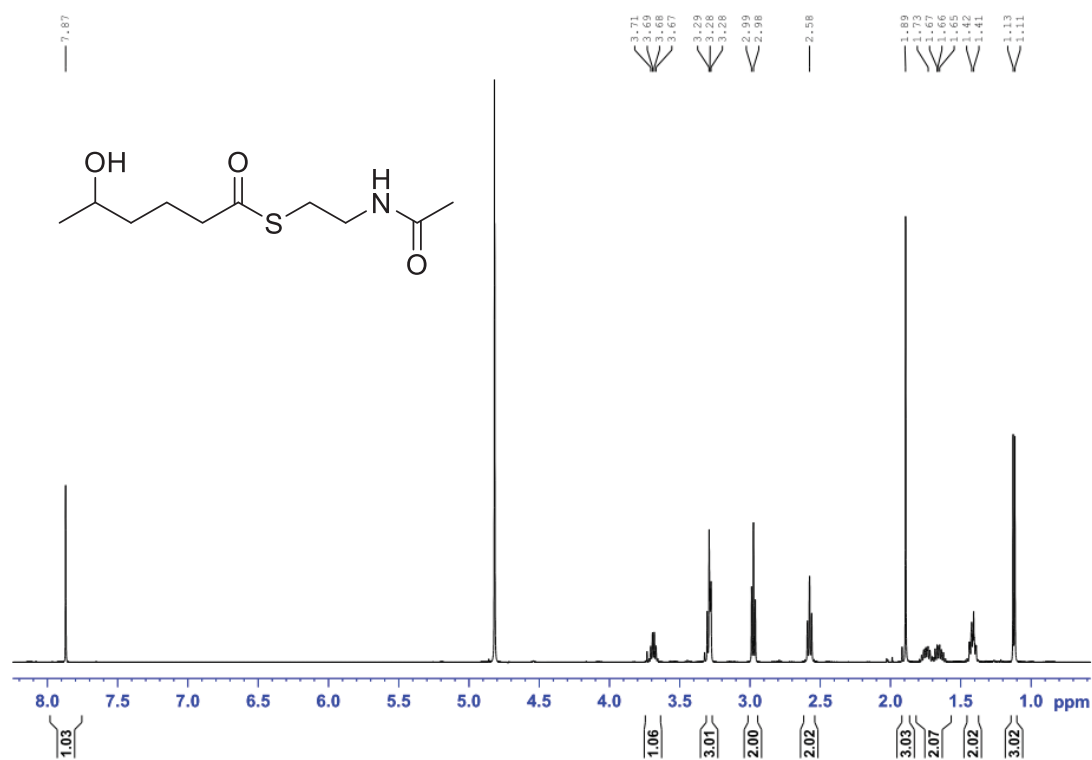
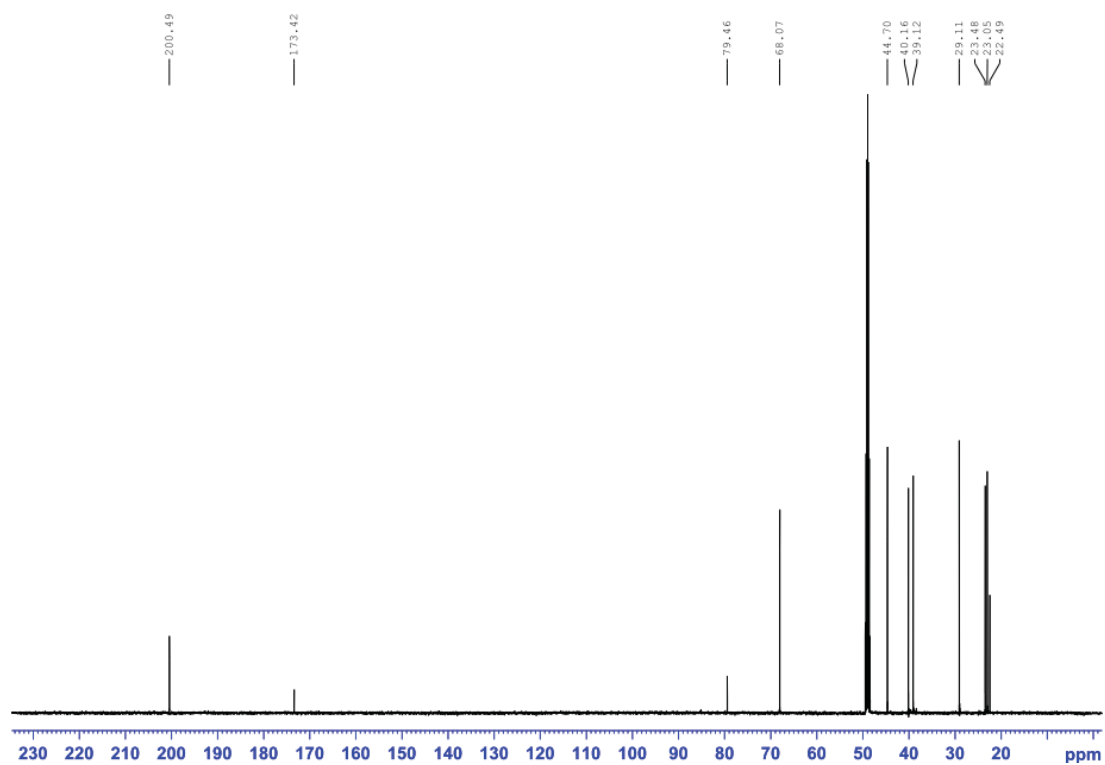


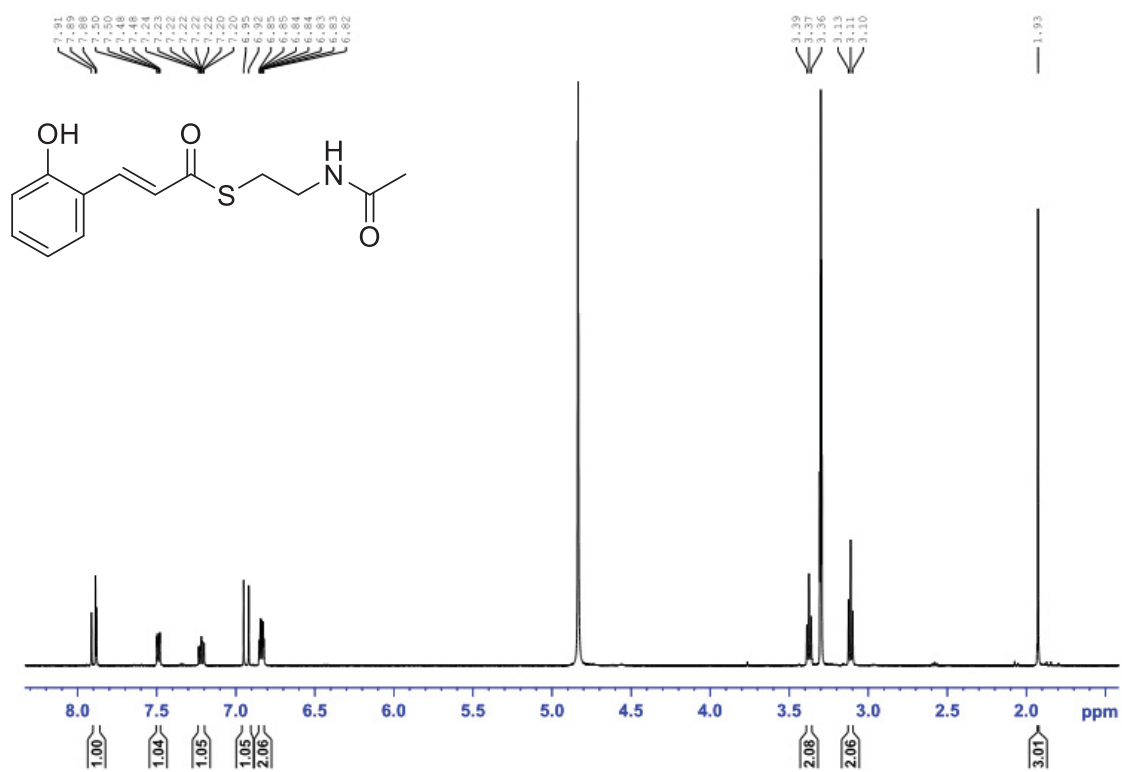
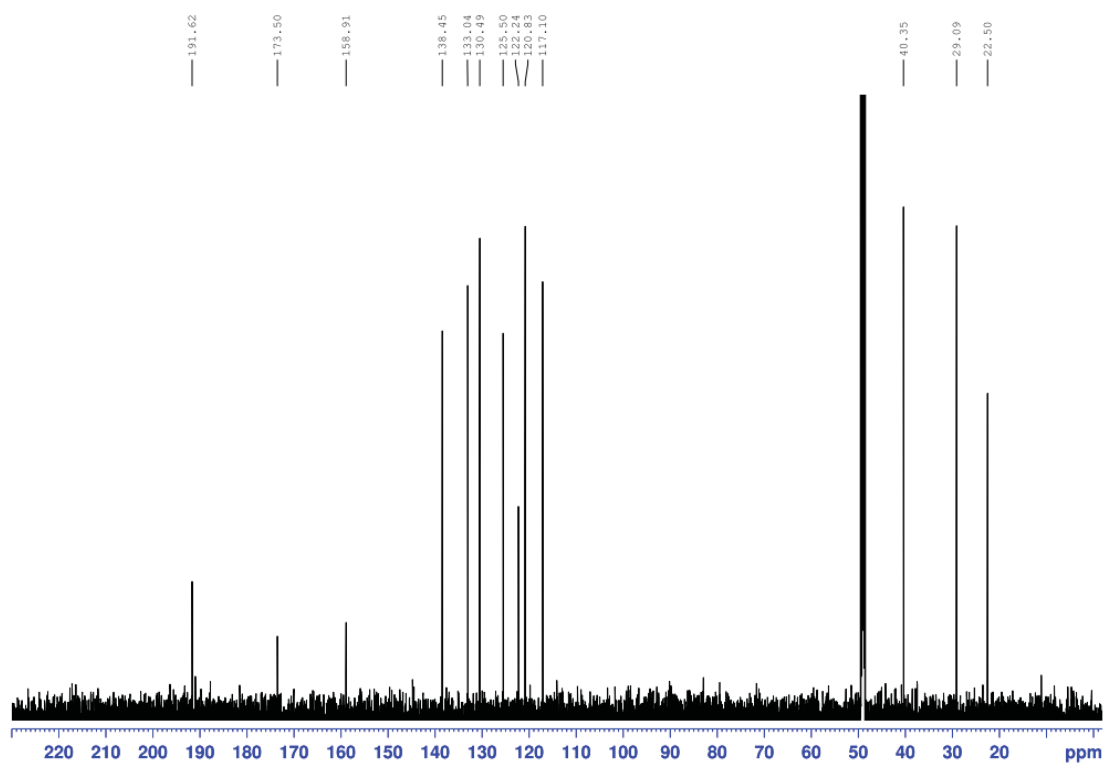
16

But-3-enamide (8.5 mg, 0.1 mmol) and thioester **33** (17 mg, 0.1 mmol) were dissolved in dry dichloromethane (0.5 mL) under argon. After addition of Hoveyda-Grubbs catalyst 2nd generation (24 mg, 0.04 mmol) the solution was stirred for 3 h. The solvent was evaporated under reduced pressure, and the residue subjected to column chromatography on silica gel (CHCl₃/MeOH 100:0 to 80:20) to give a white solid.

Yield: 70 % (based on recovered starting material **33**). ¹H NMR (600 MHz, CDCl₃): δ 6.96 (1H, dt, ³J = 15.5 Hz, ³J = 7.2 Hz), 6.28 (1H, dt, ³J = 15.5 Hz, ⁴J = 1.5 Hz), 3.34 (2H, t, ³J = 6.7 Hz), 3.15 (2H, dd, ³J = 7.2 Hz, ⁴J = 1.5 Hz), 3.07 (2H, t, ³J = 6.7 Hz), 1.91 ppm (3H, s); ¹³C NMR (150 MHz, CDCl₃): δ 190.6, 174.4, 173.5, 139.5, 131.8, 40.1, 39.3, 29.0, 22.5 ppm; IR (film): 3310, 2947, 2831, 1663, 1407, 1111, 1020 cm⁻¹; *m/z* calculated for C₉H₁₅O₃N₂S: 231.0798, found 231.0793 [M+H]⁺.

6. NMR Spectra

Figure S1. ¹H NMR spectrum (MeOD, 500 MHz) of **8**.Figure S2. ¹³C NMR spectrum (MeOD, 125 MHz) of **8**.

Figure S3. ¹H NMR spectrum (MeOD, 500 MHz) of **10**.Figure S4. ¹³C NMR spectrum (MeOD, 125 MHz) of **10**.

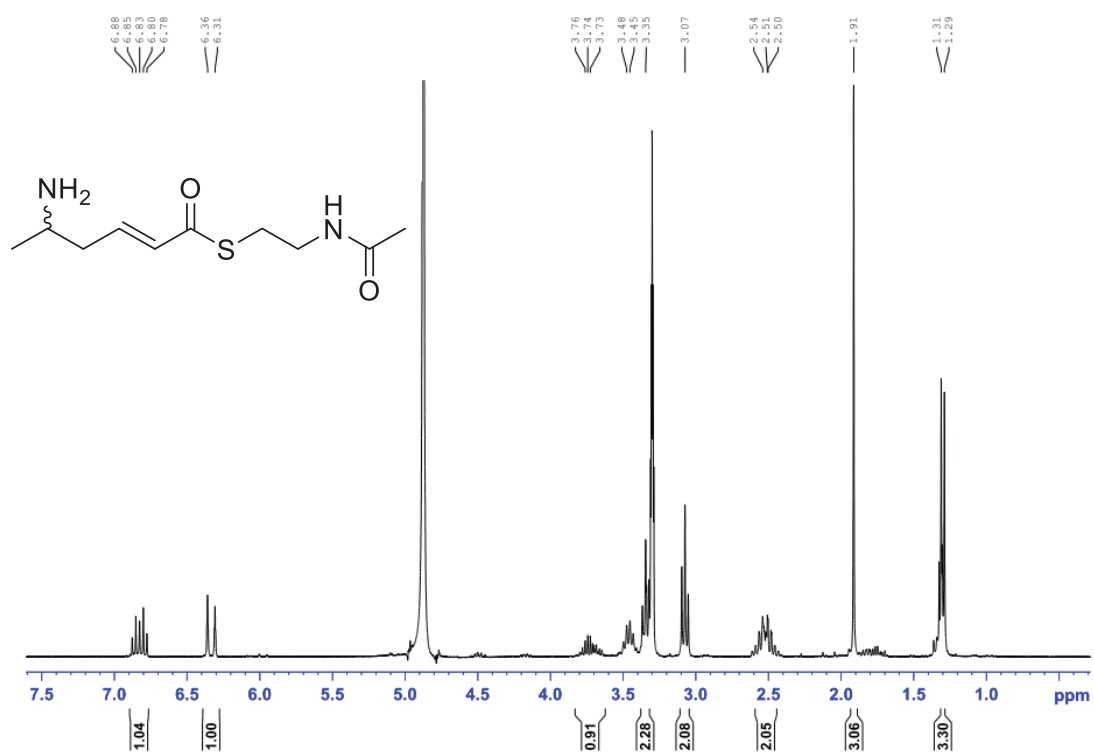


Figure S5. ¹H NMR spectrum (MeOD, 300 MHz) of 12.

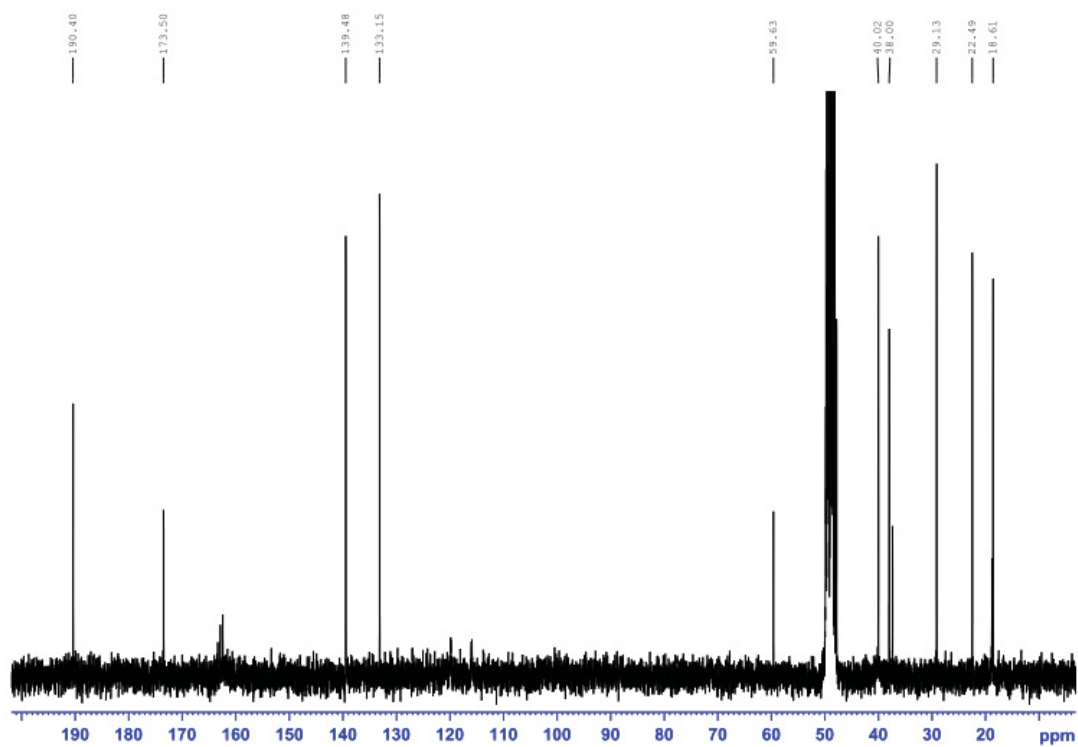
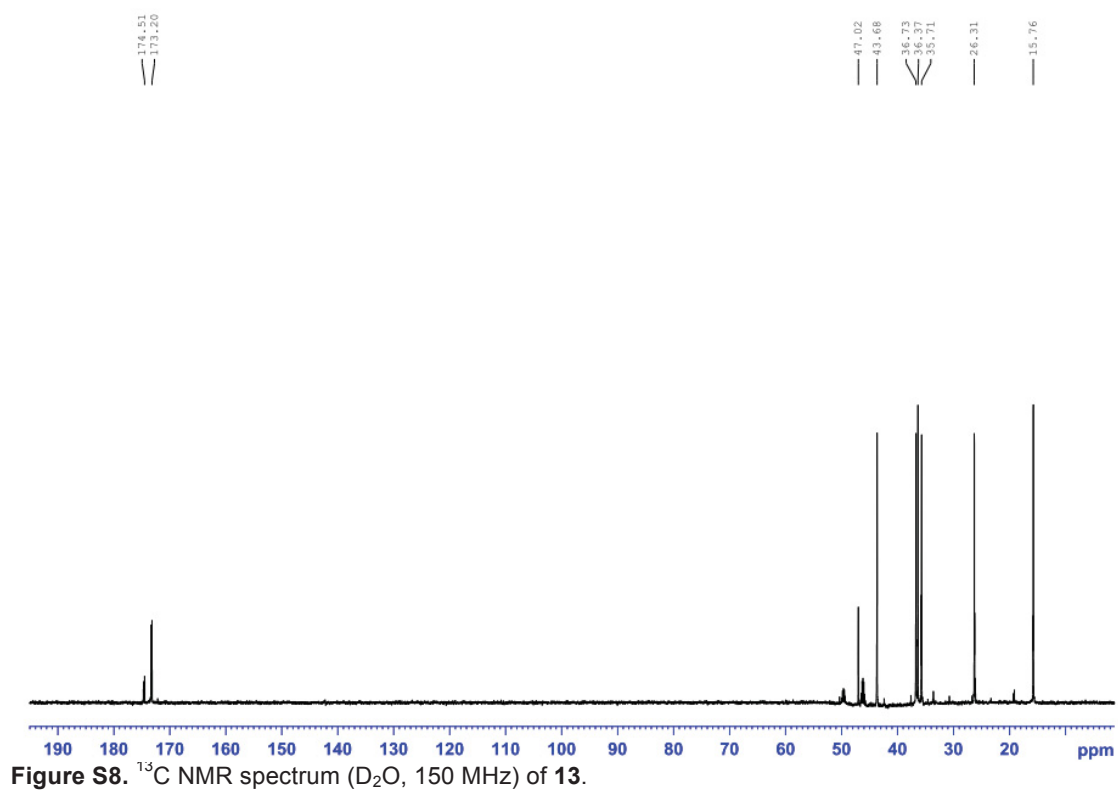
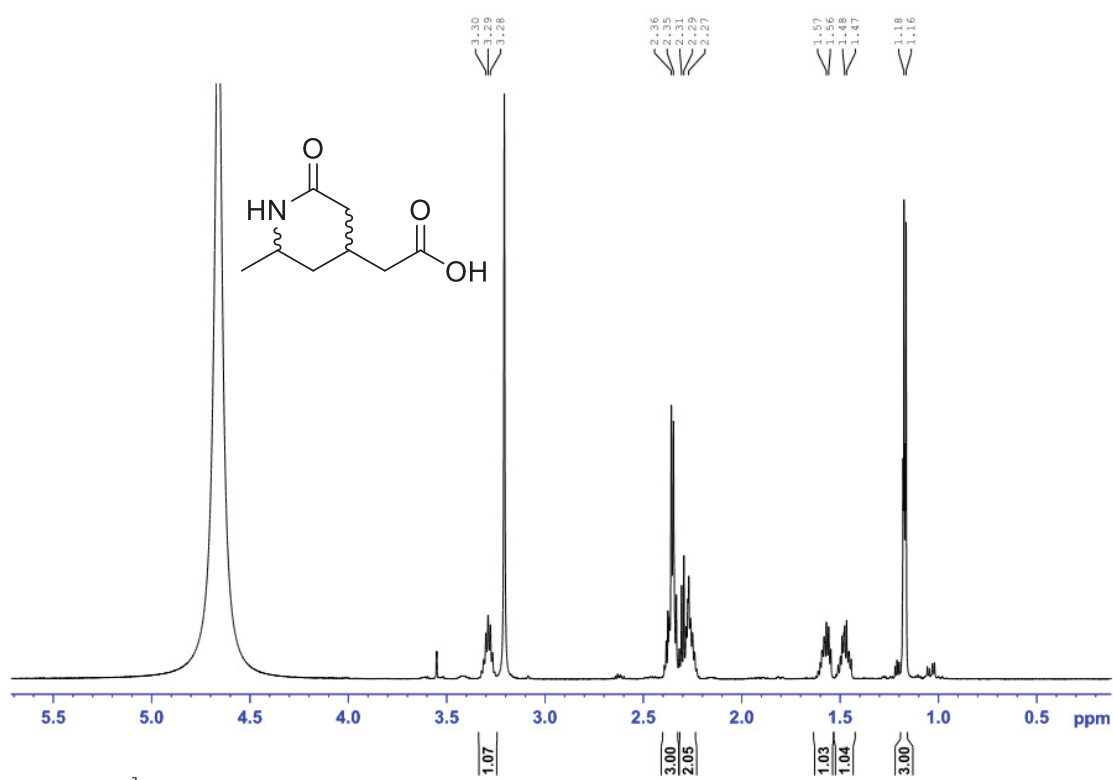
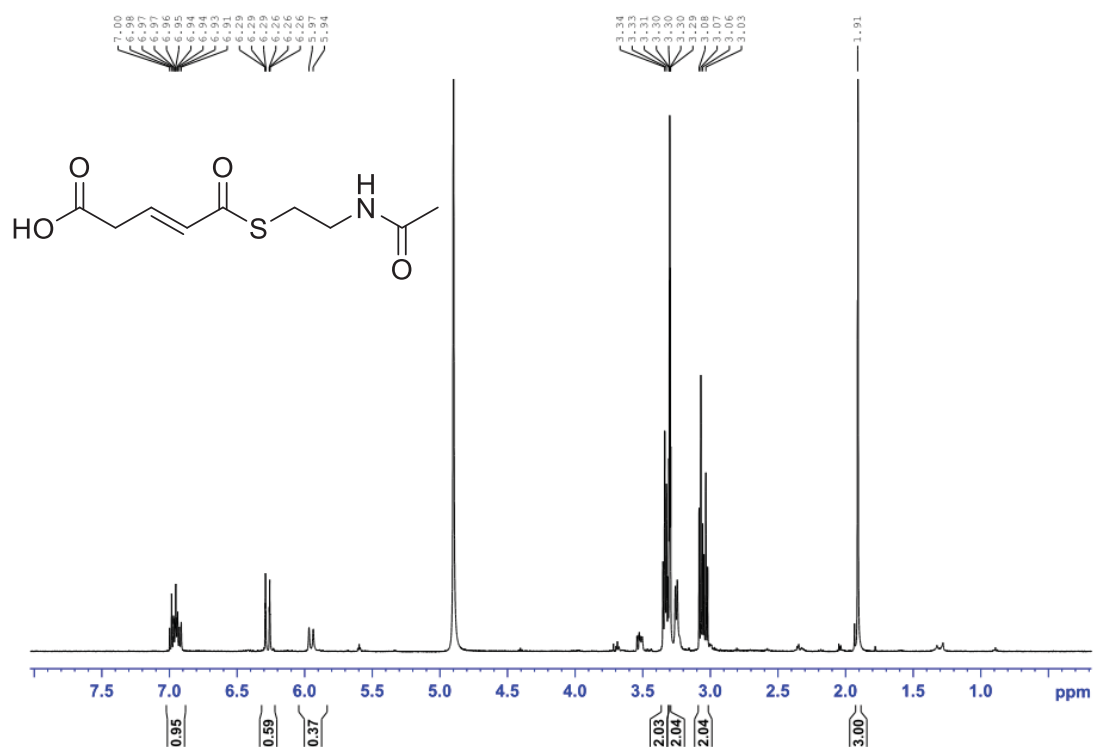
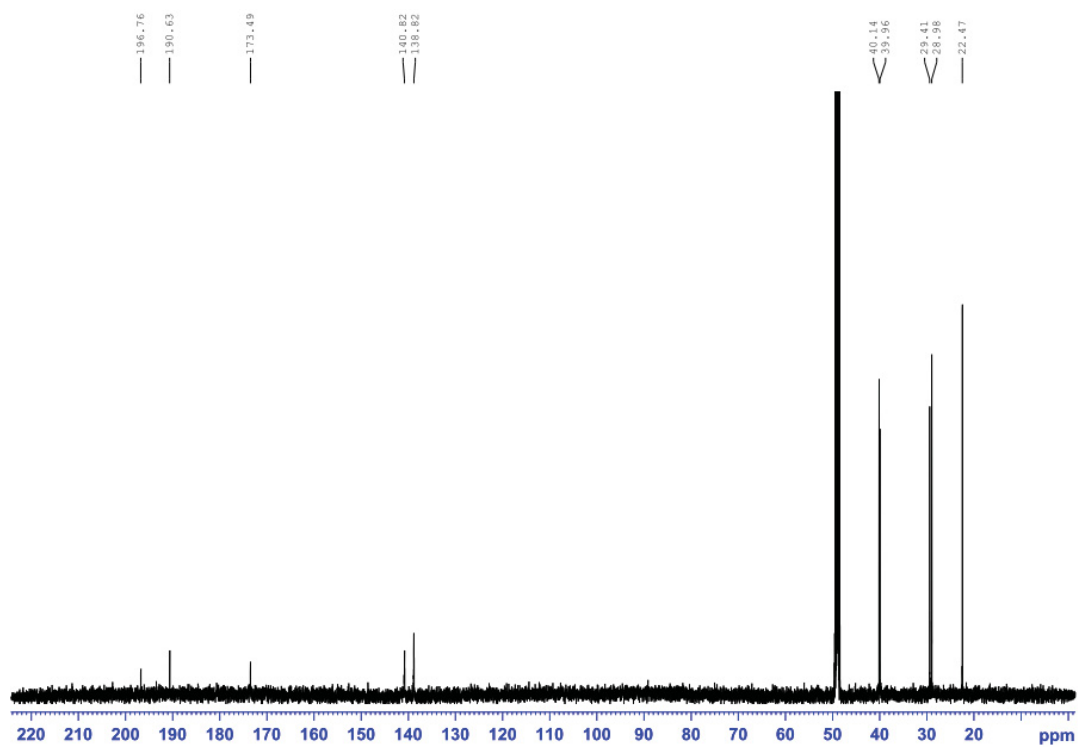


Figure S6. ¹³C NMR spectrum (MeOD, 75 MHz) of 12.



Figure S9. ¹H NMR spectrum (CDCl₃, 500 MHz) of 14.Figure S10. ¹³C NMR spectrum (CDCl₃, 125 MHz) of 14.

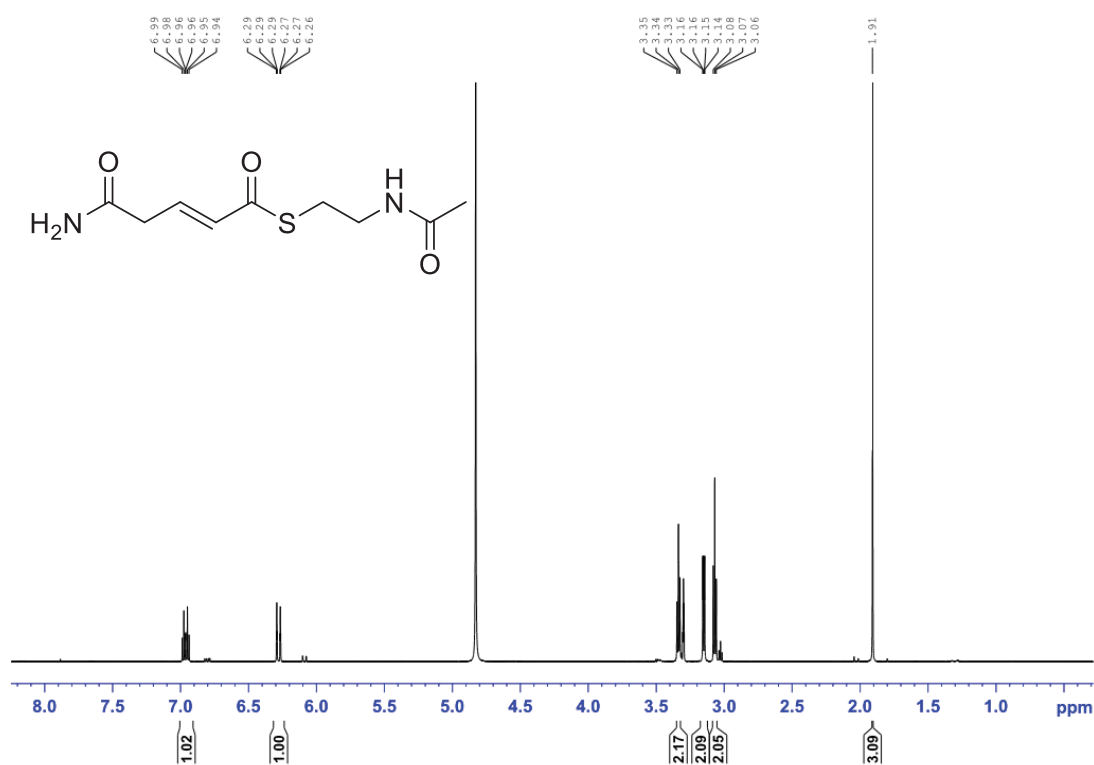


Figure S11. ^1H NMR spectrum (CDCl_3 , 600 MHz) of **16**.

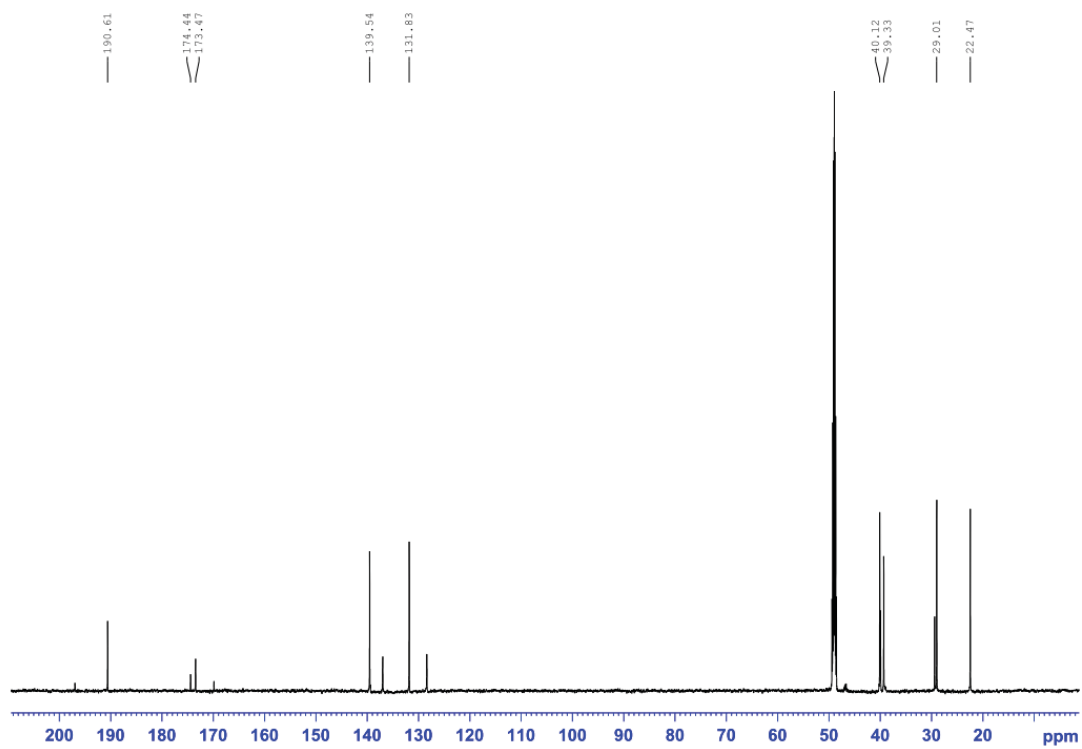


Figure S12. ^{13}C NMR spectrum (CDCl_3 , 150 MHz) of **16**.

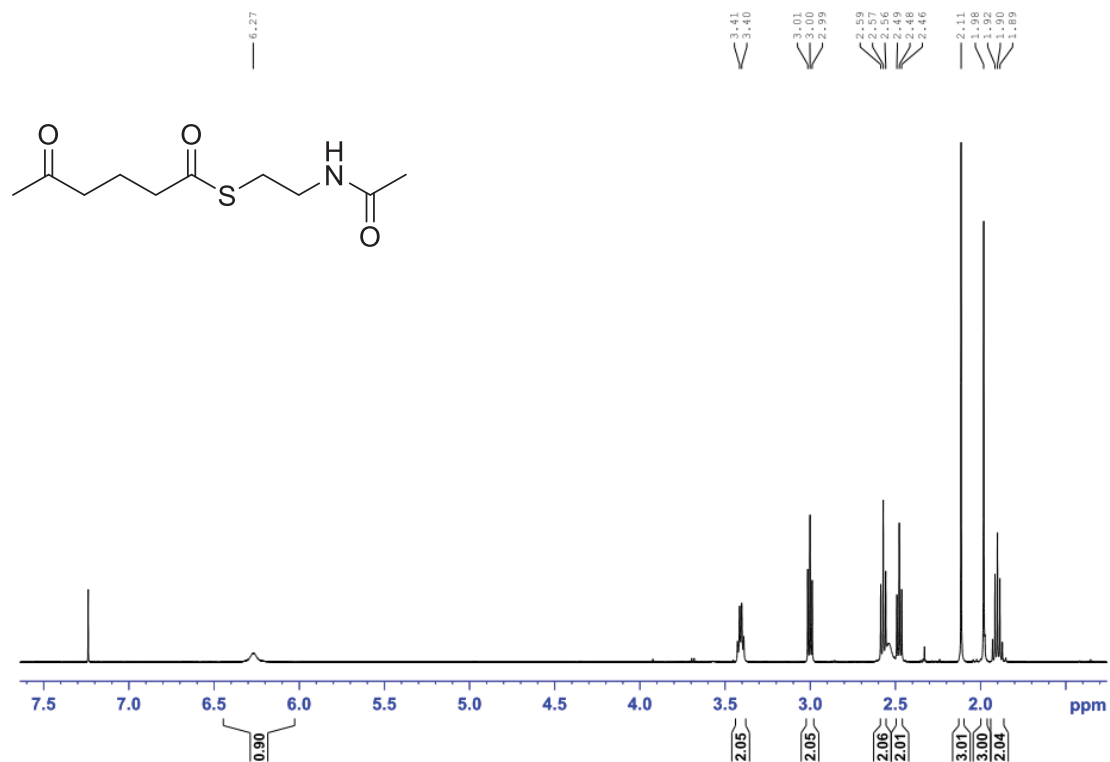


Figure S13. ¹H NMR spectrum (CDCl₃, 600 MHz) of 22.

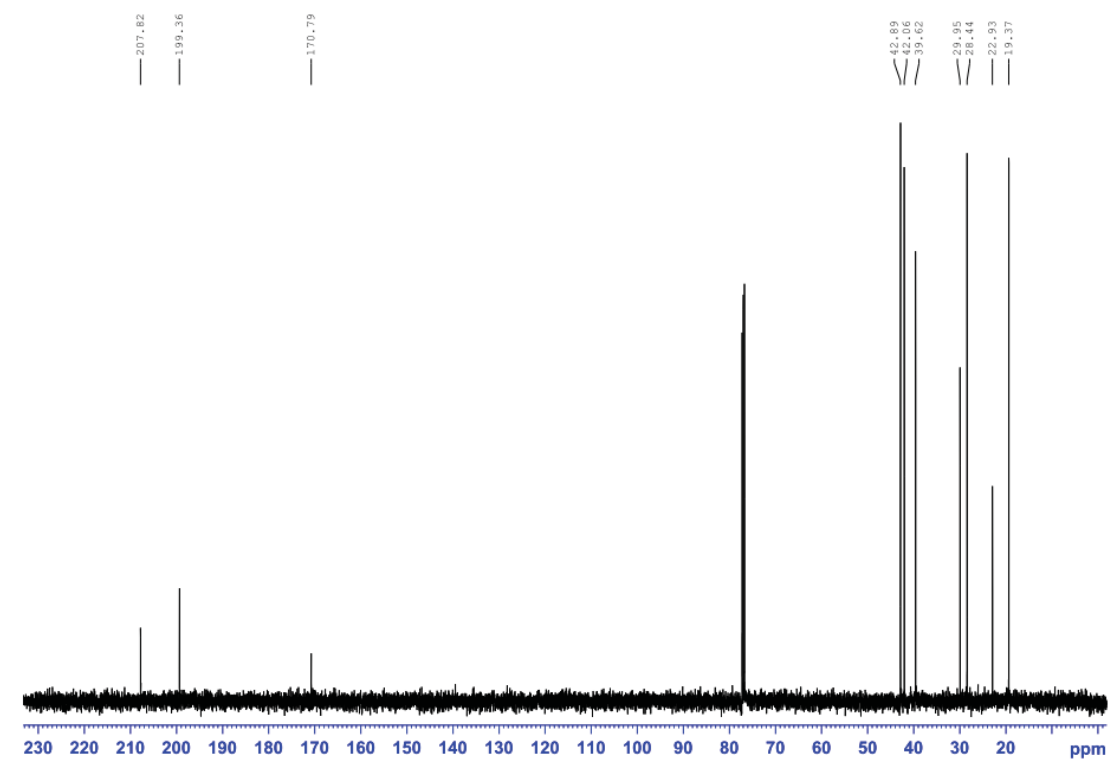


Figure S14. ¹³C NMR spectrum (CDCl₃, 125 MHz) of 22.

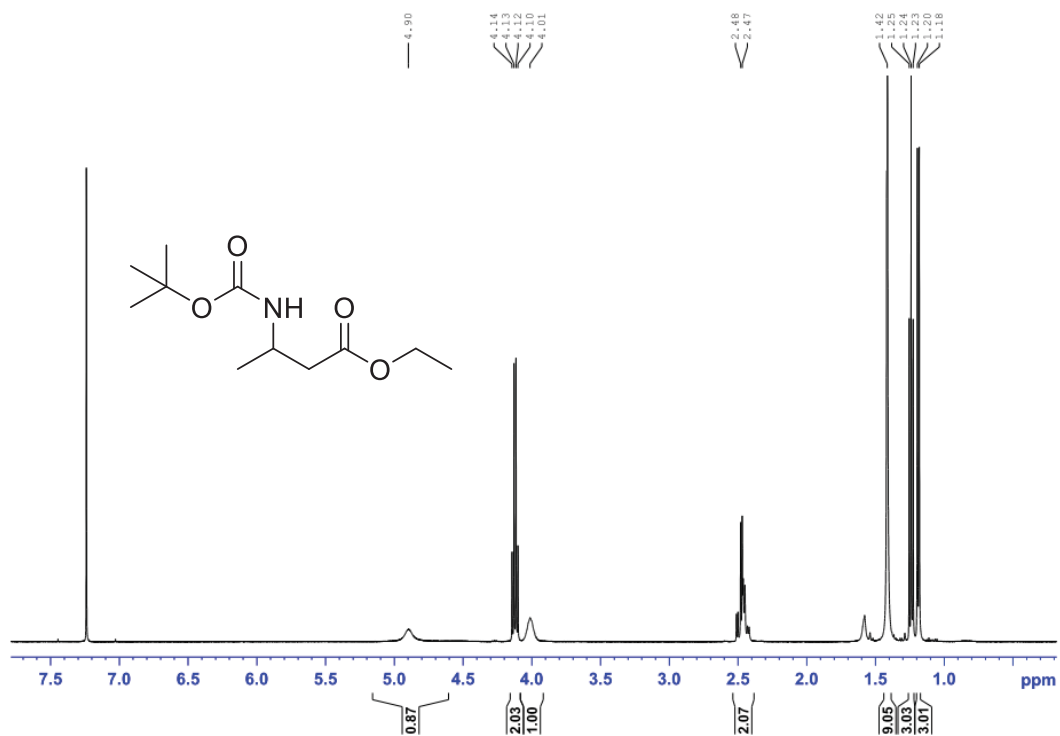


Figure S15. ¹H NMR spectrum (CDCl₃, 500 MHz) of 25.

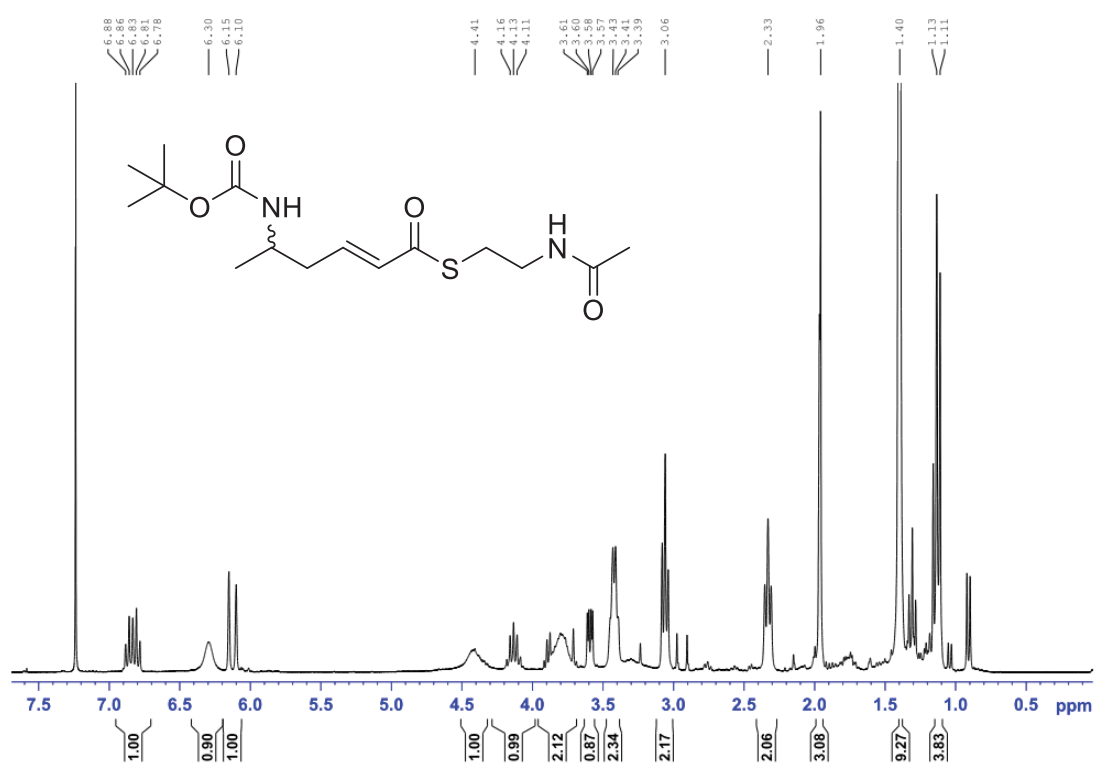


Figure S16. ¹H NMR spectrum (CDCl₃, 300 MHz) of 26.

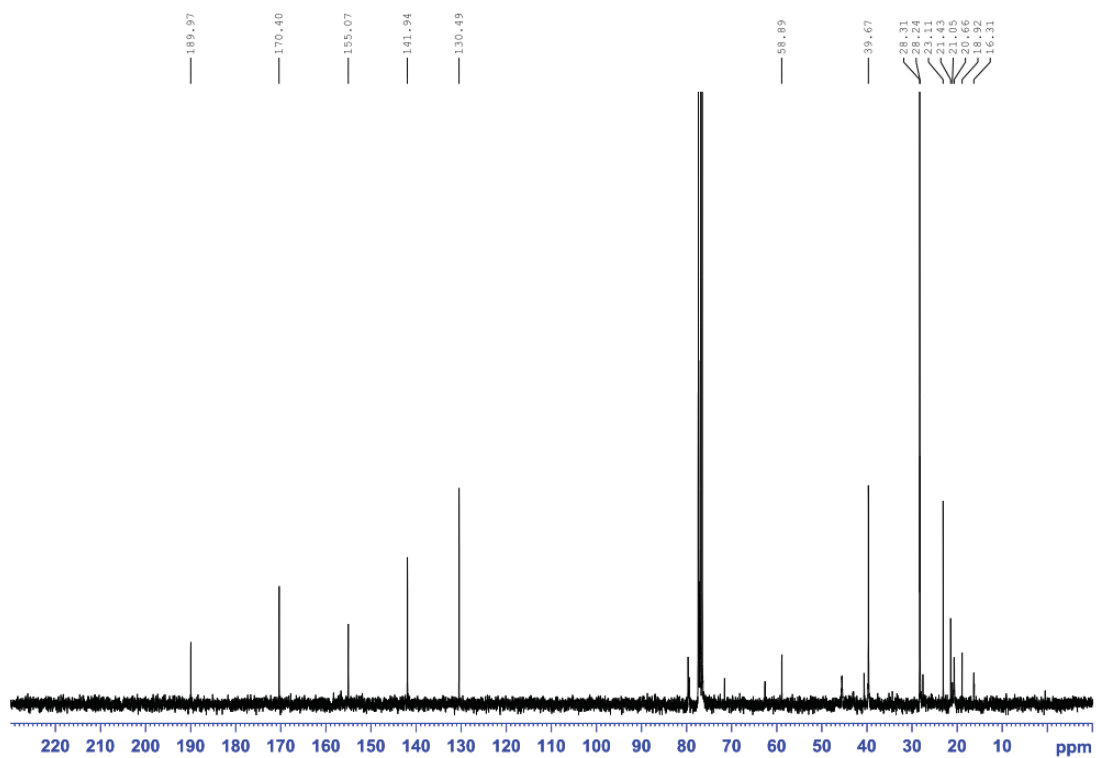


Figure S17. ¹³C NMR spectrum (CDCl₃, 75 MHz) of 26.

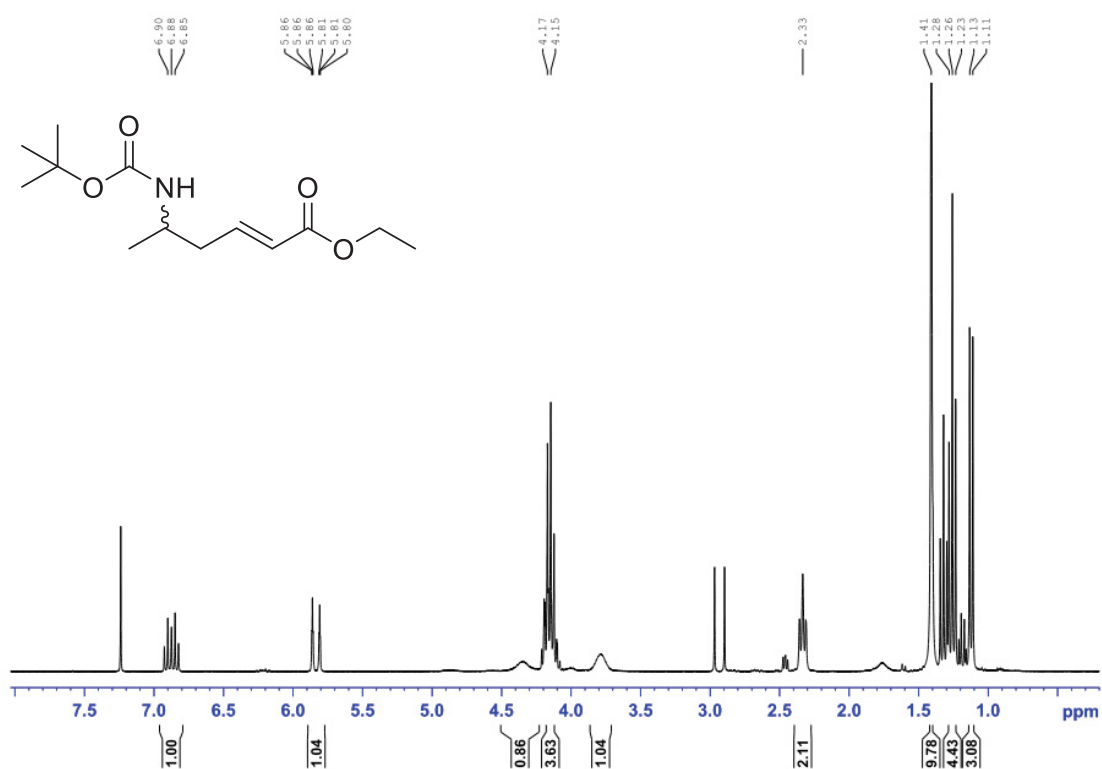


Figure S18. ^1H NMR spectrum (CDCl_3 , 300 MHz) of 27.

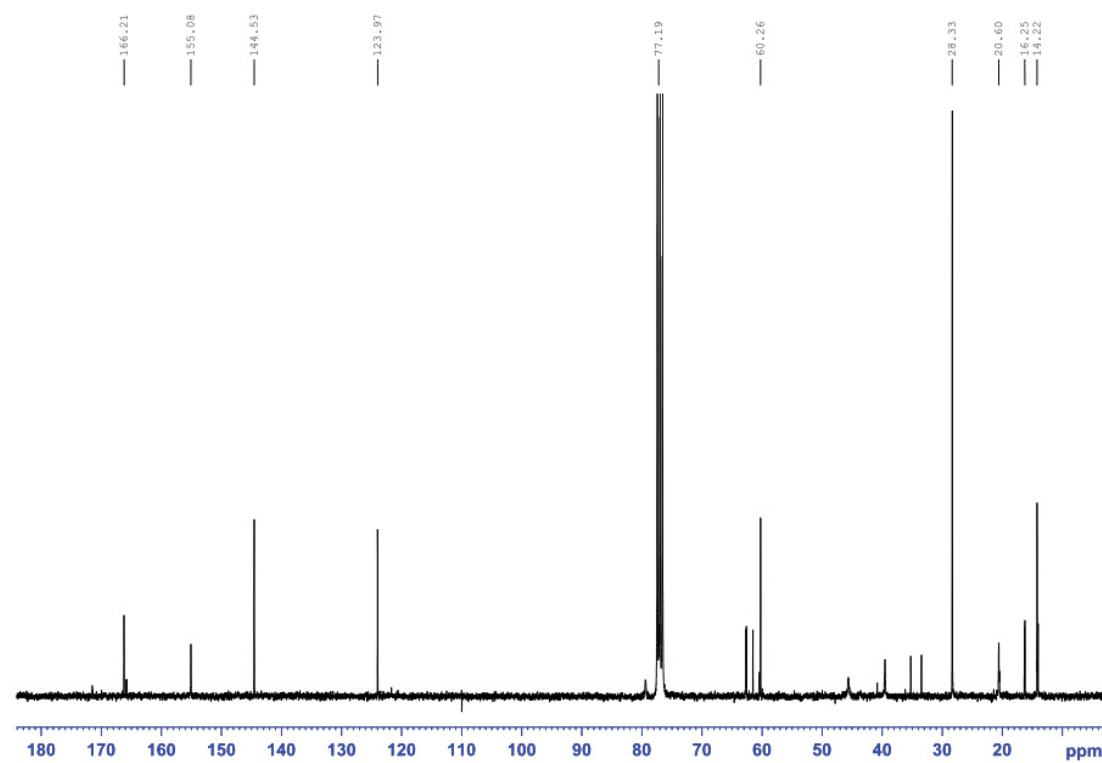


Figure S19. ^{13}C NMR spectrum (CDCl_3 , 75 MHz) of 27.

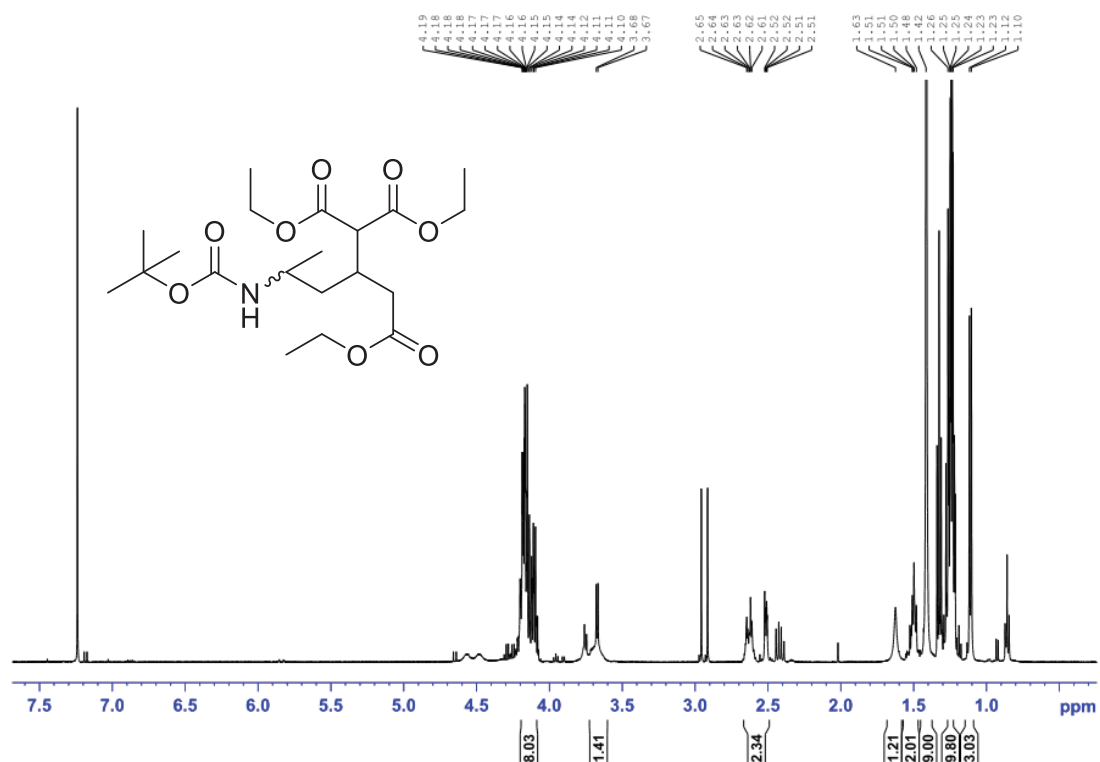


Figure S20. ¹H NMR spectrum (CDCl₃, 500 MHz) of 28.

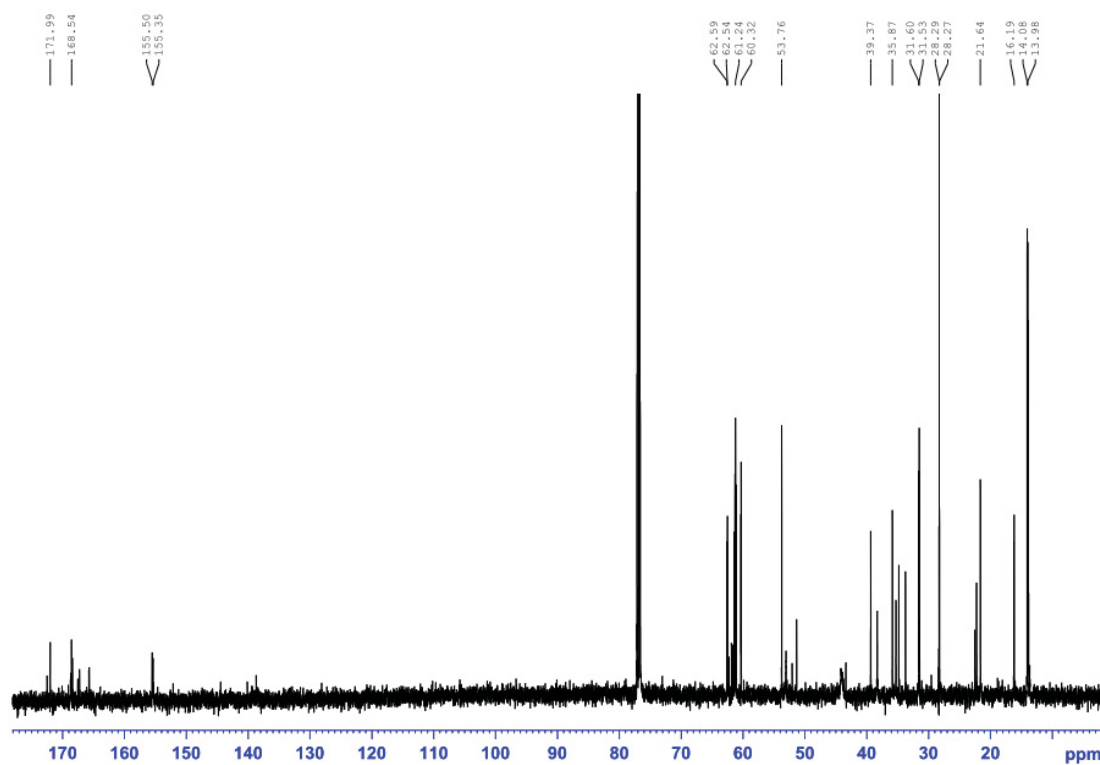
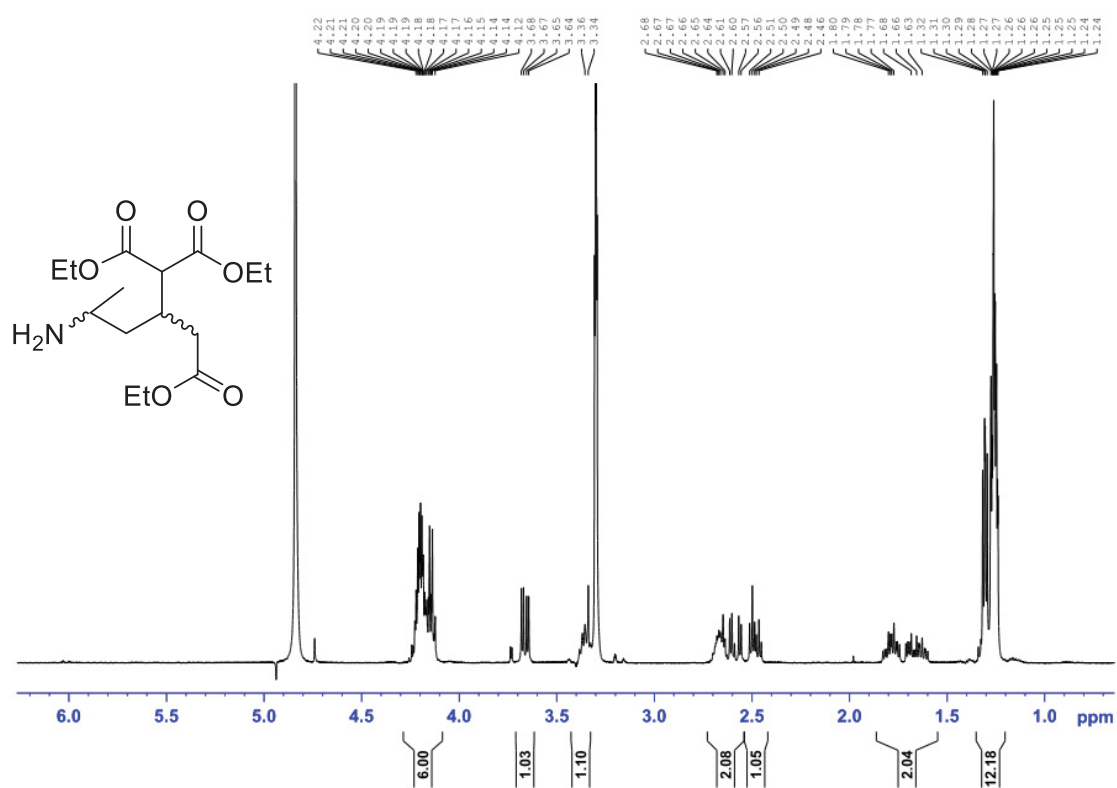
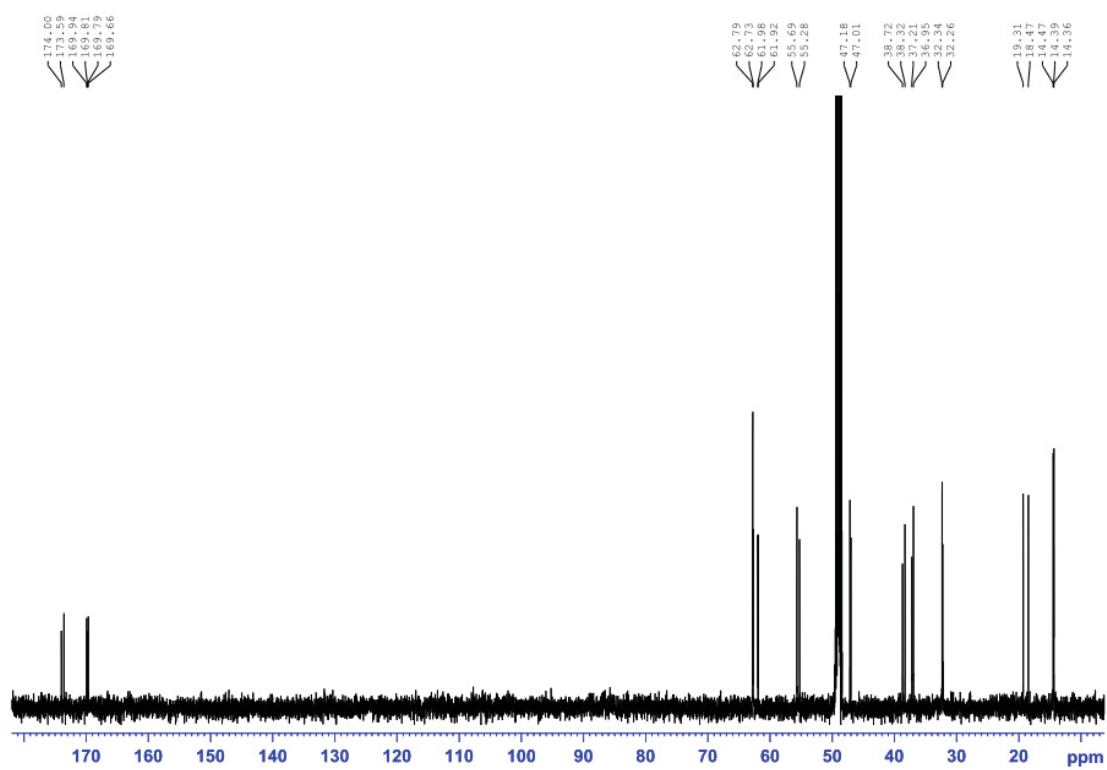
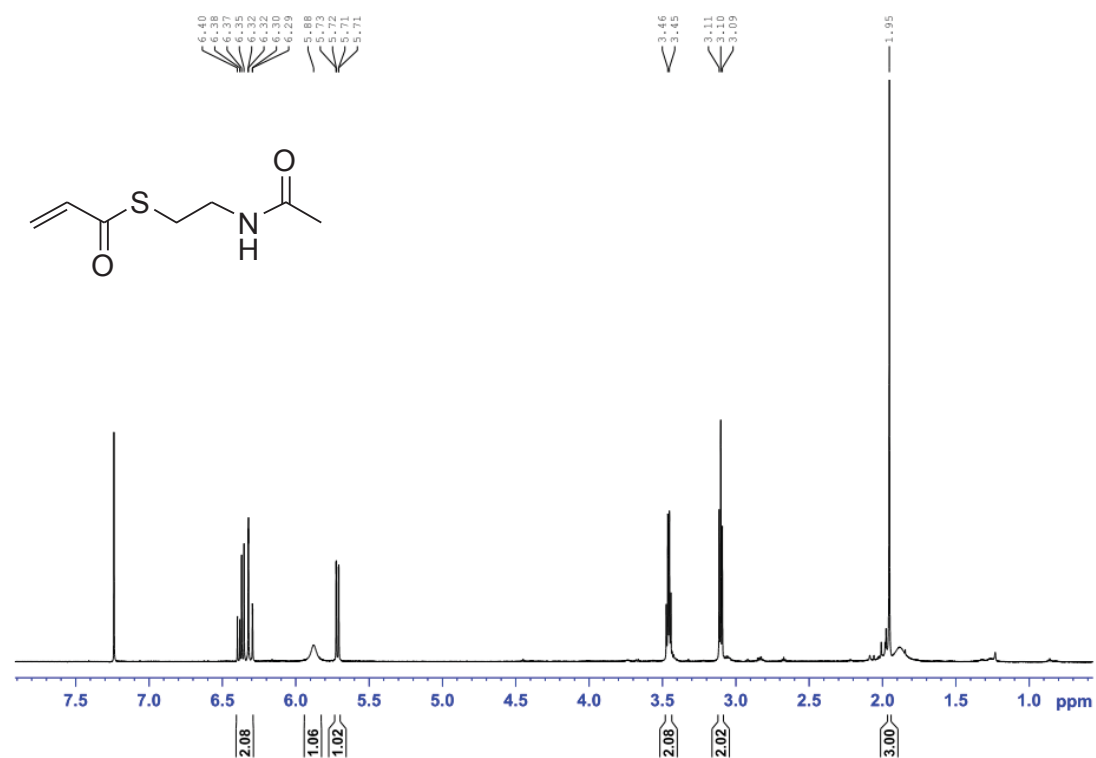
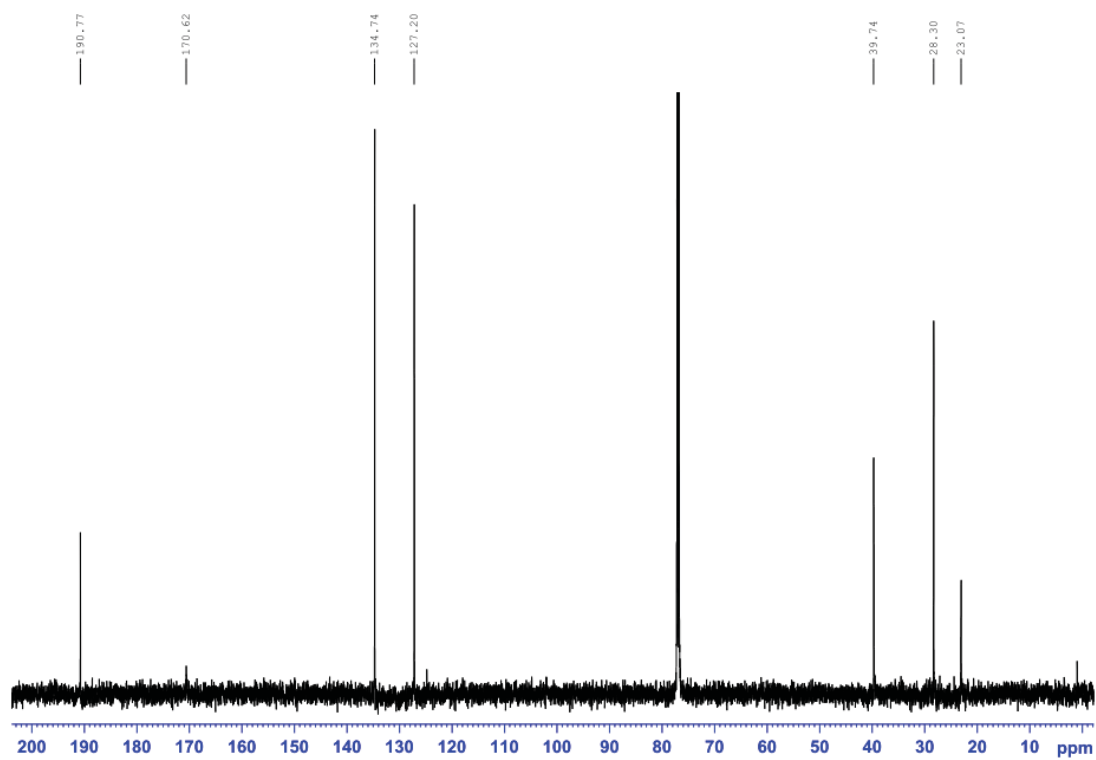


Figure S21. ¹³C NMR spectrum (CDCl₃, 125 MHz) of 28.

Figure S22. ¹H NMR spectrum (MeOD, 500 MHz) of 29.Figure S23. ¹³C NMR spectrum (MeOD, 125 MHz) of 29.

Figure S24. ¹H NMR spectrum (CDCl₃, 600 MHz) of 33.Figure S25. ¹³C NMR spectrum (CDCl₃, 150 MHz) of 33.

7. Original HRMS Data

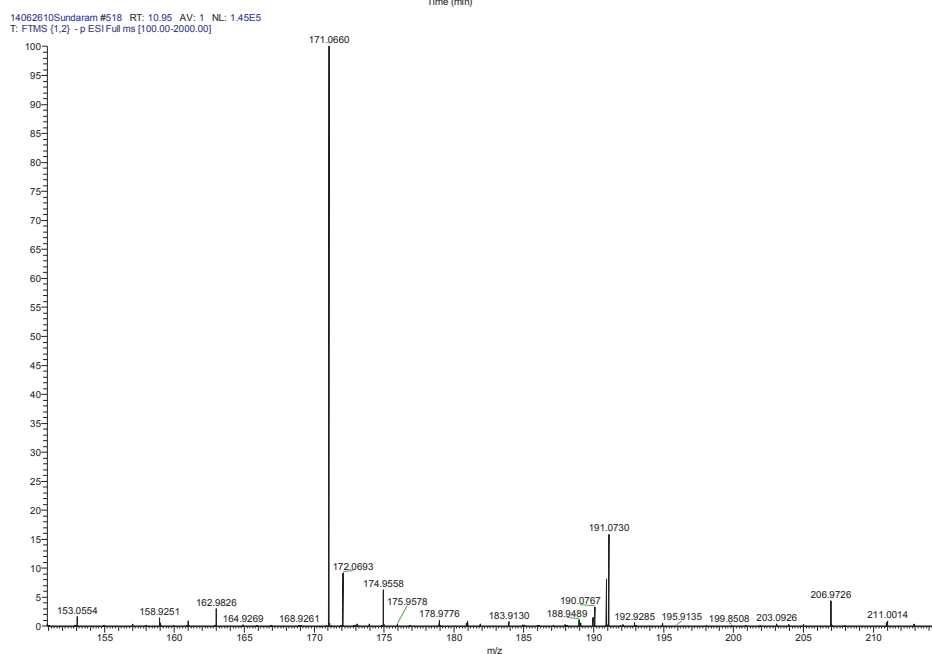
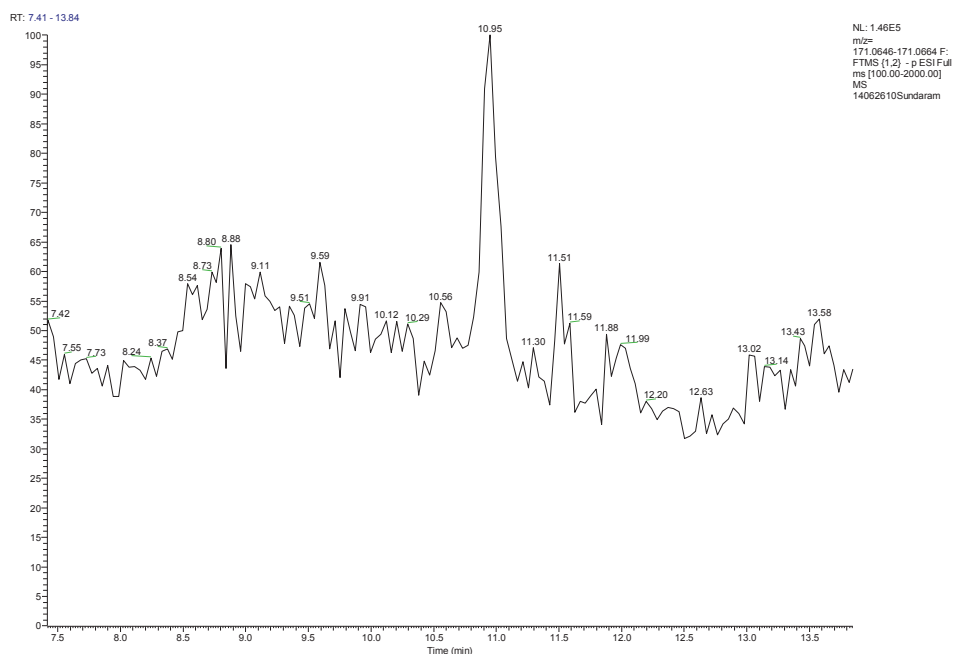


Figure S26. EIC ($m/z = 171.0646-171.0664$, negative mode) and mass spectrum after in vitro enzymatic biotransformation of **4**.

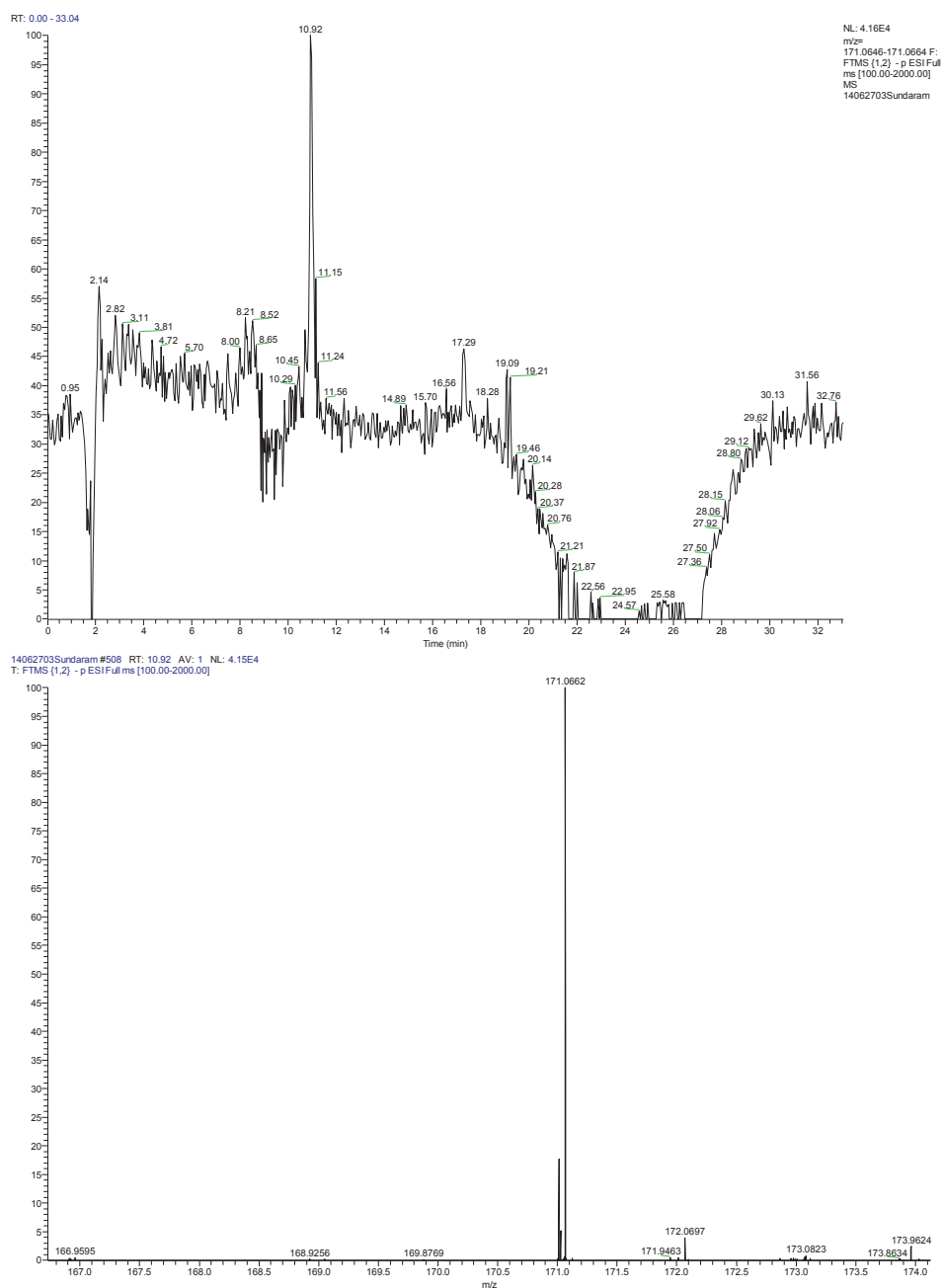


Figure S27. EIC ($m/z = 171.0646-171.0664$, negative mode) and mass spectrum after in vitro enzymatic biotransformation of **5**.

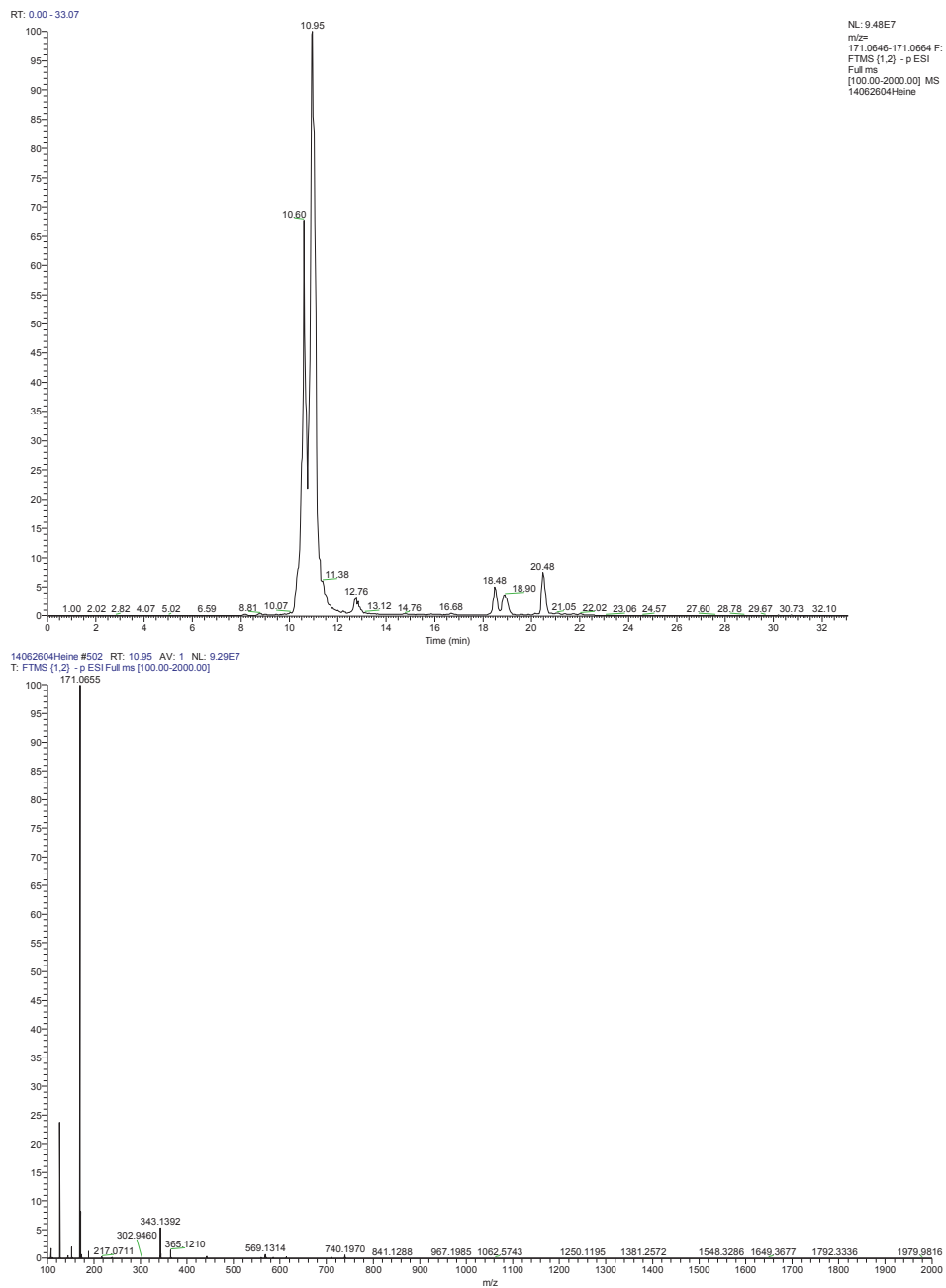


Figure S28. EIC ($m/z = 171.0646-171.0664$, negative mode) mass spectrum of reference compounds **6** and **7**.

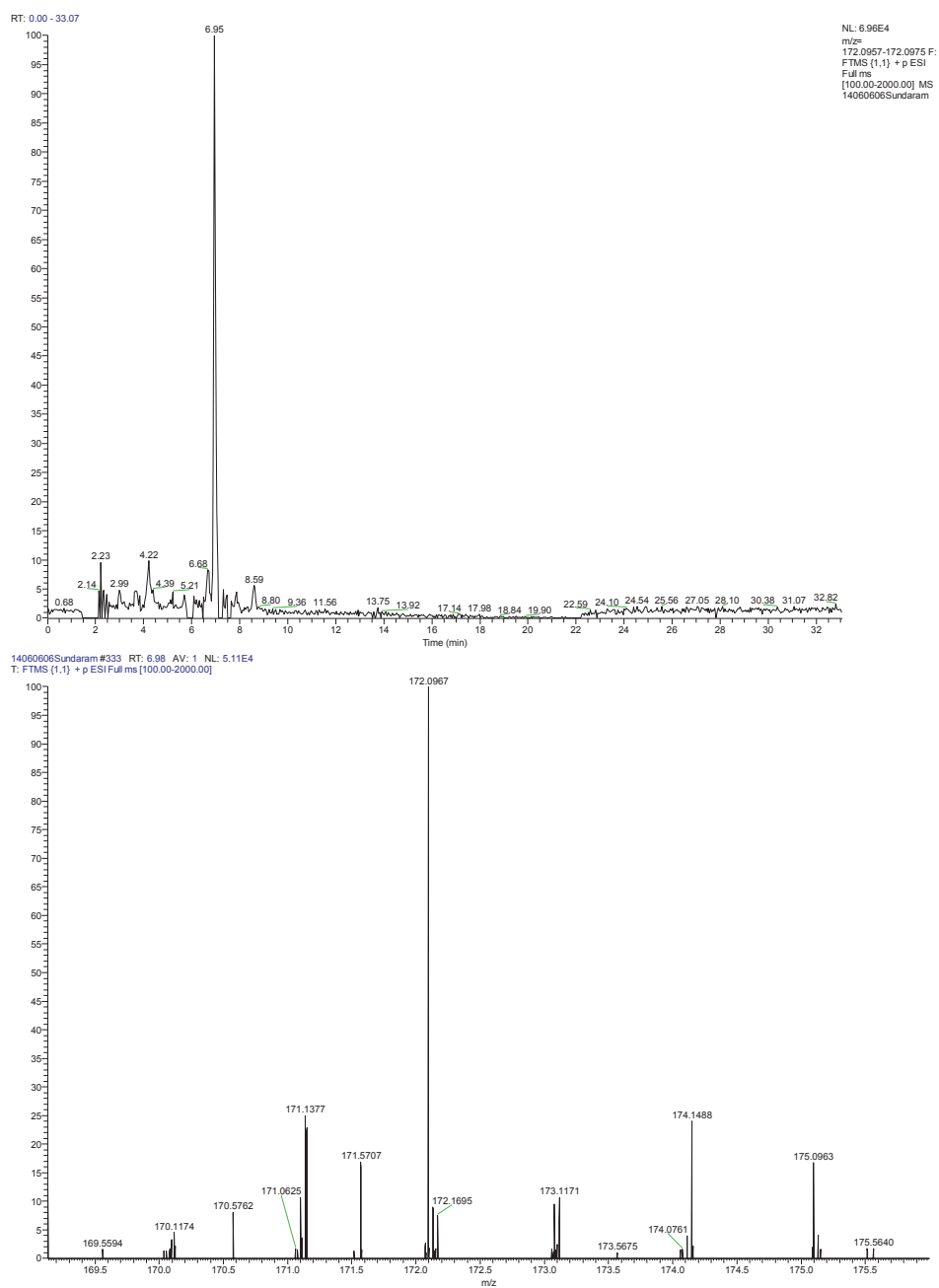


Figure S29. EIC ($m/z = 172.0957-172.0975$, positive mode) and mass spectrum after in vitro enzymatic biotransformation of **12**.

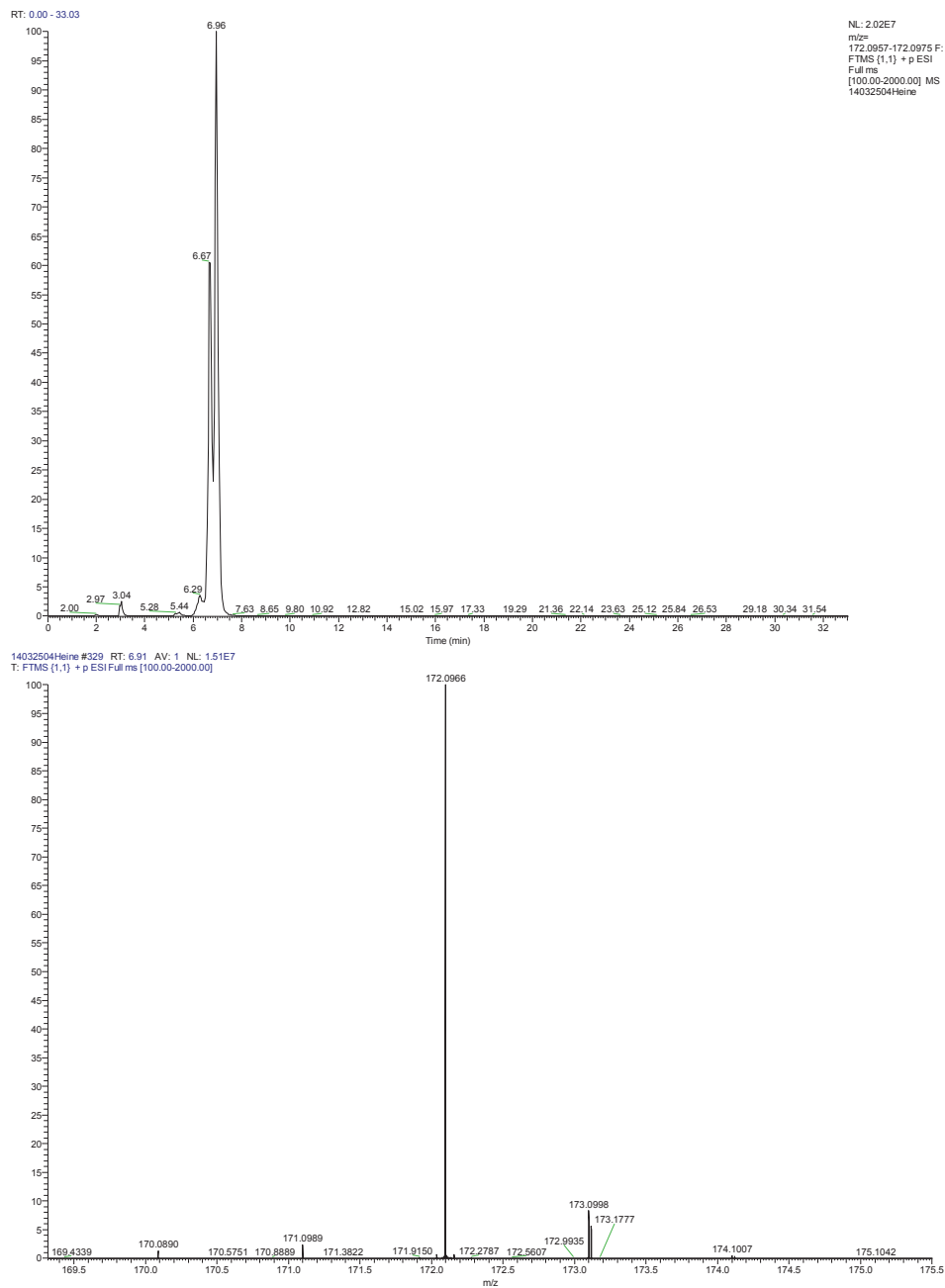


Figure S30. EIC ($m/z = 172.0957-172.0975$, positive mode) and mass spectrum of reference compound 13.

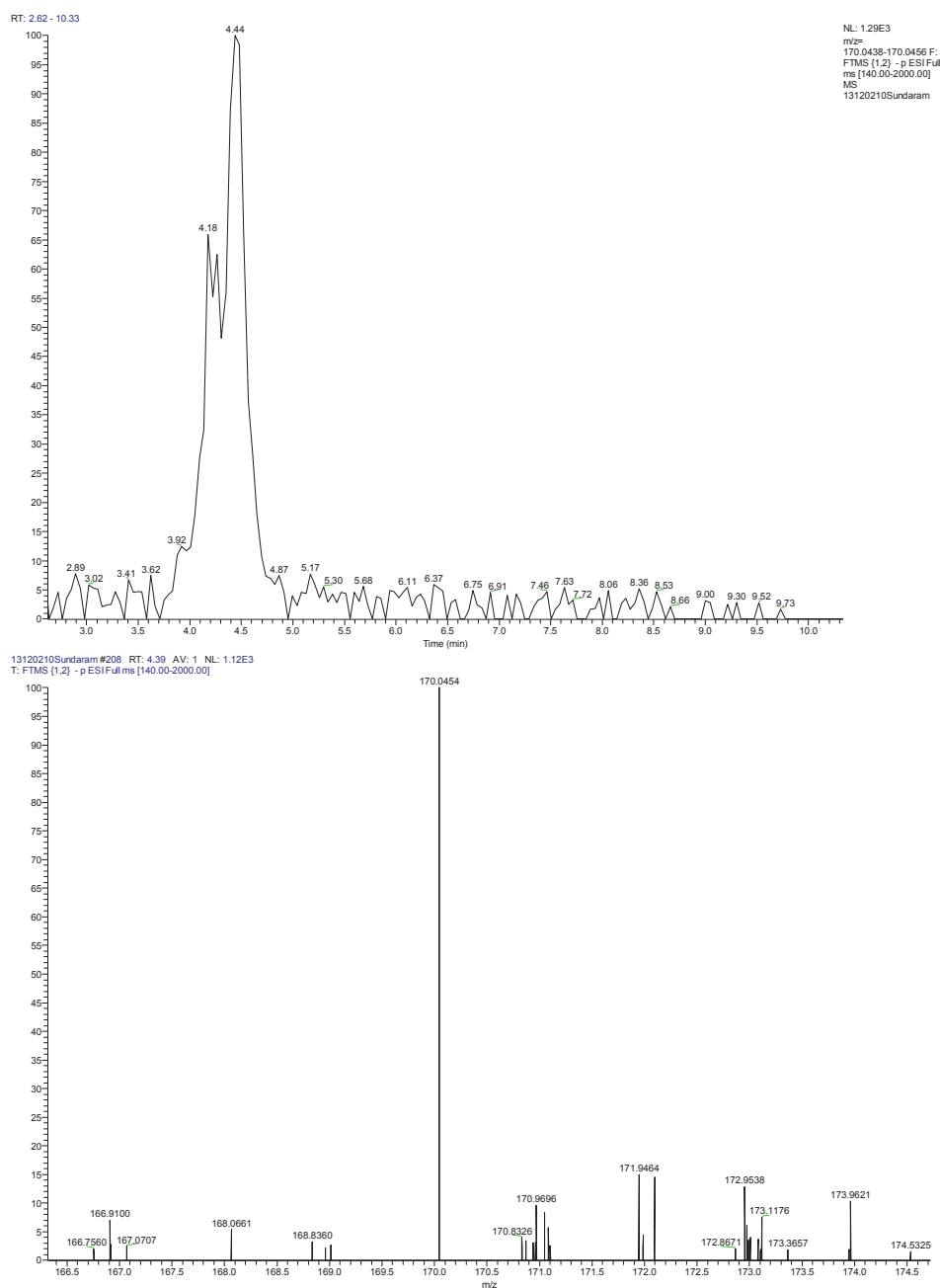


Figure S31. EIC ($m/z = 170.0438-170.0456$, negative mode) and mass spectrum after in vitro enzymatic biotransformation of **16**.

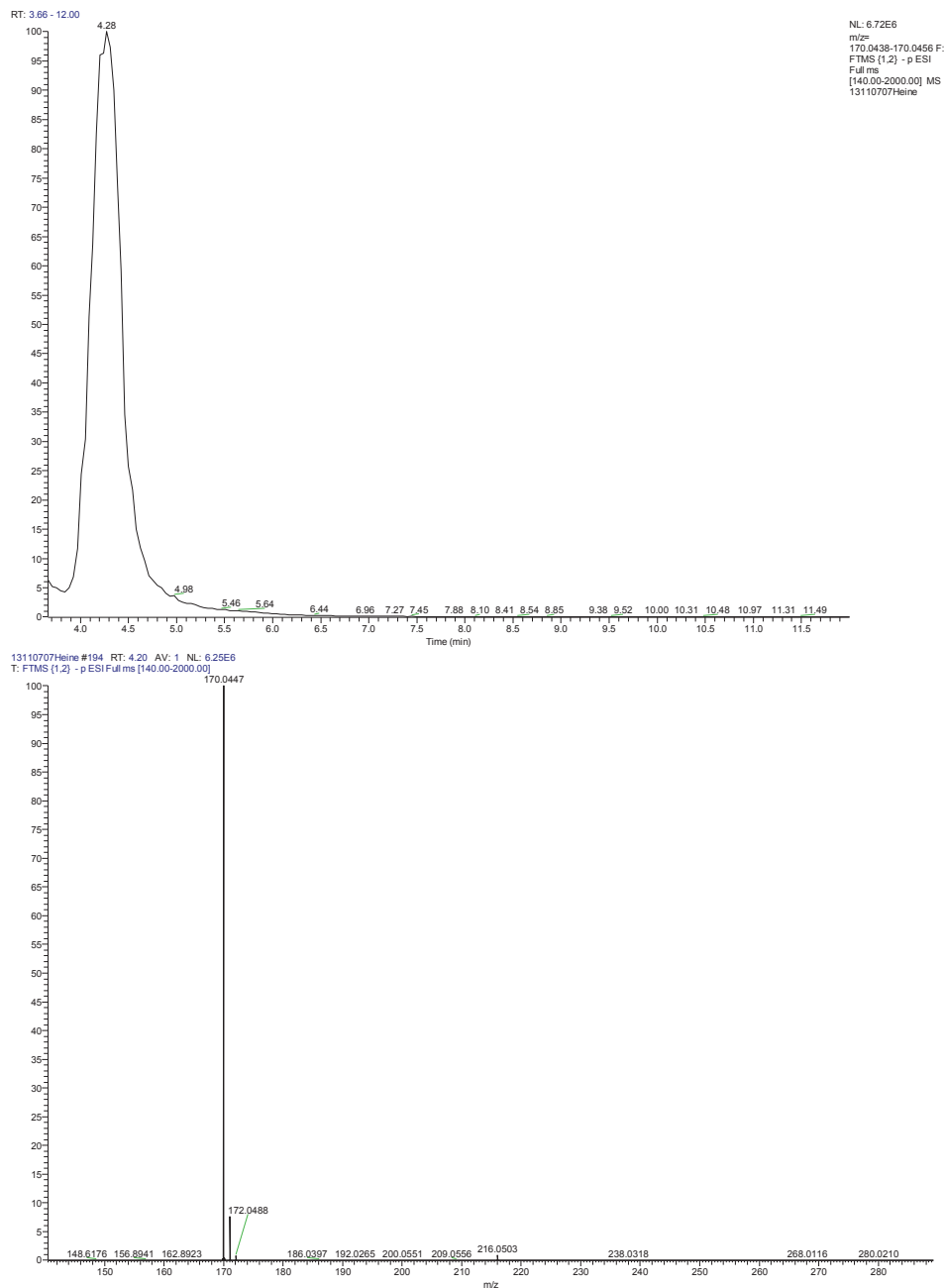


Figure S32. EIC ($m/z = 170.0438-170.0456$, negative mode) and mass spectrum of reference compound 17.

8. Original MALDI Data

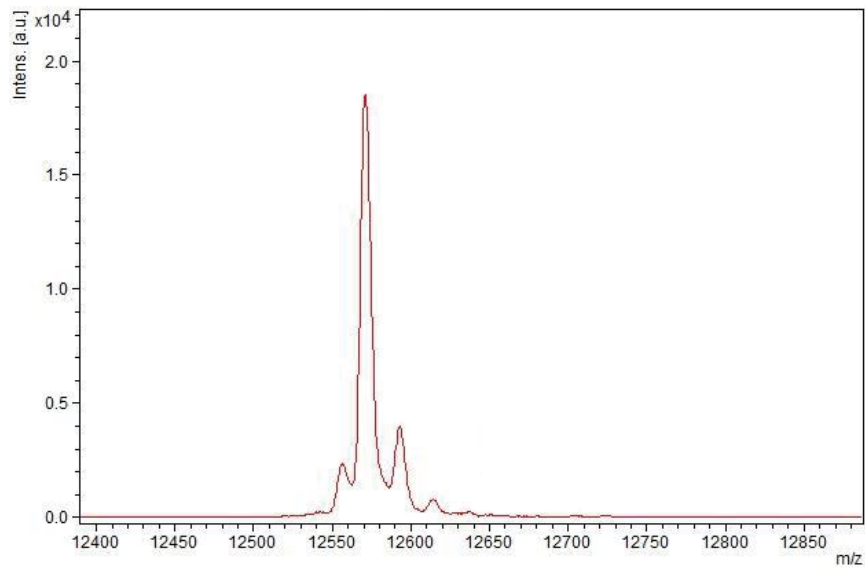


Figure S33. MALDI-analysis of holo-ACP.

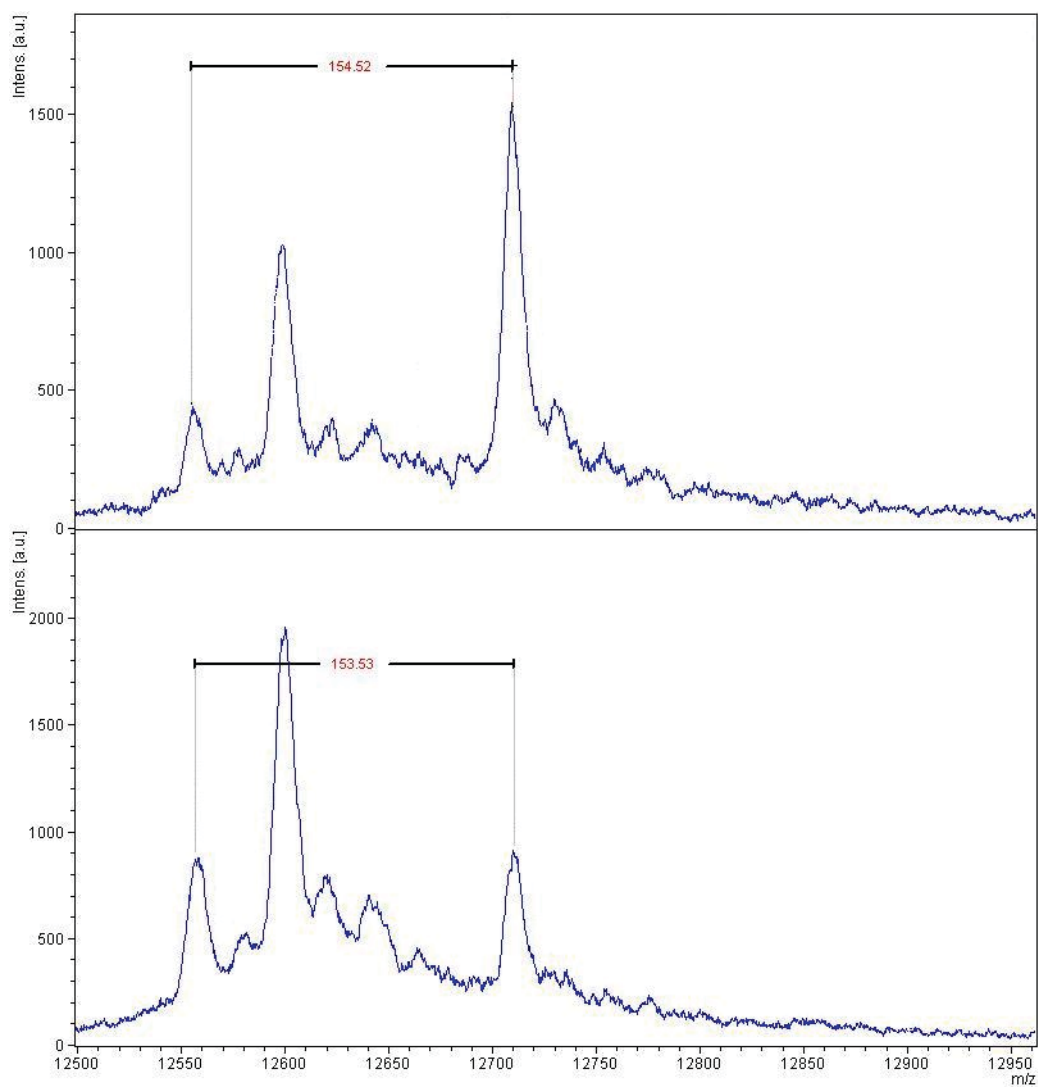


Figure S34. MALDI-analyses of ACP-bound products of in vitro biotransformation of **5** (top) and **4** (bottom).

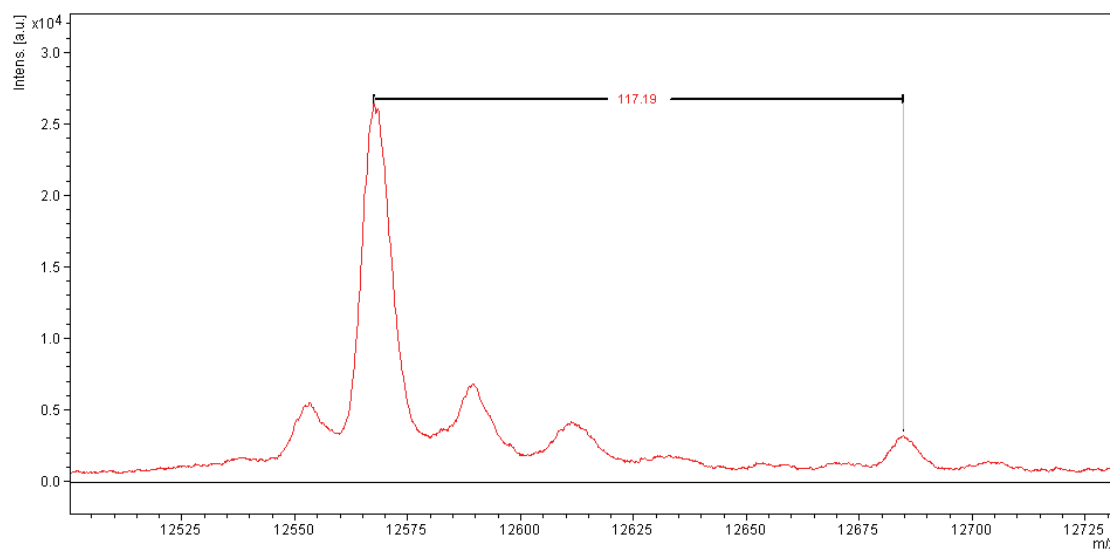


Figure S35. MALDI-analysis after attempted in vitro biotransformation of **8**.

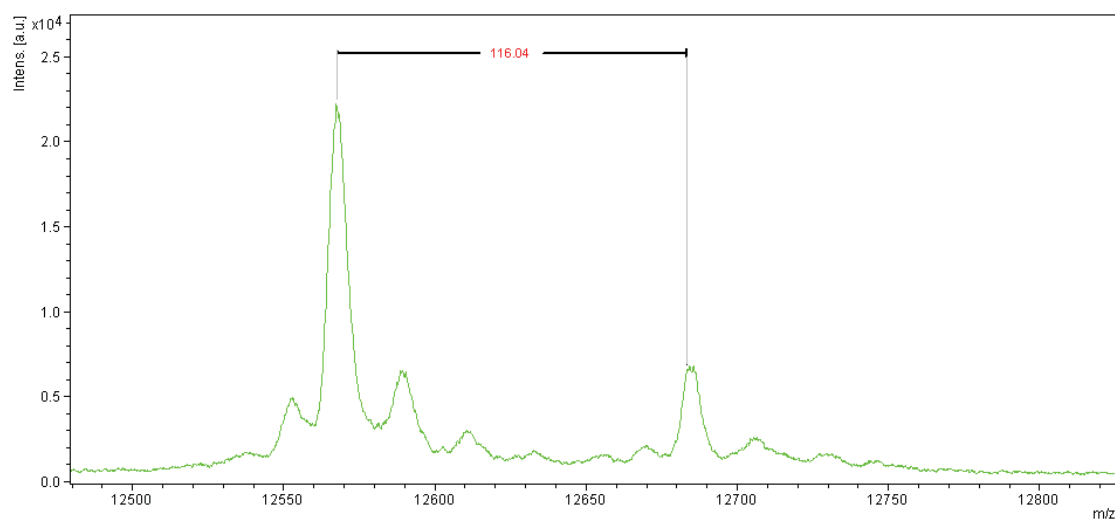


Figure S36. MALDI-analysis after attempted in vitro biotransformation of **10**.

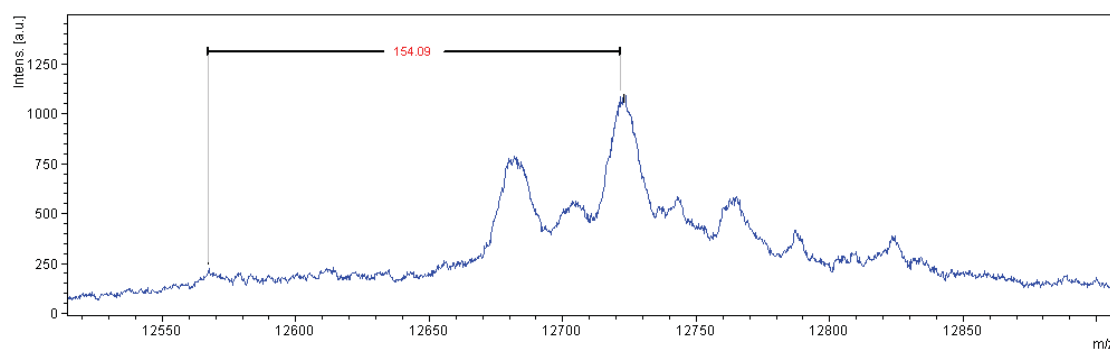


Figure S37. MALDI-analysis of ACP-bound product of in vitro biotransformation of **12**.

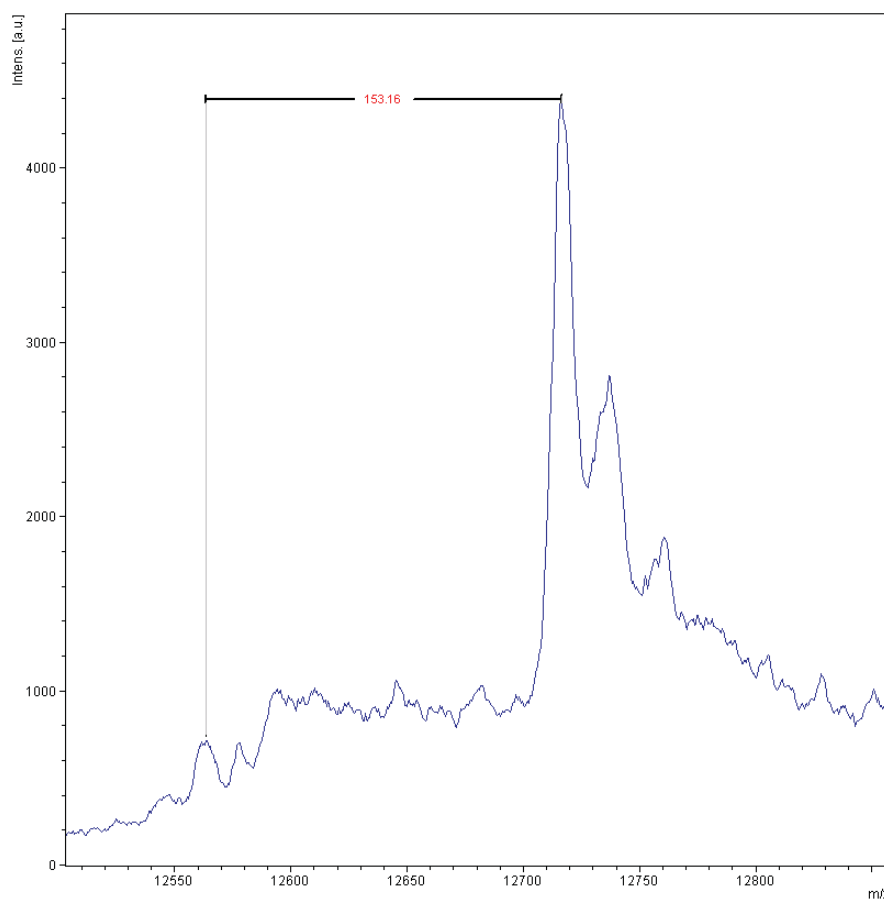


Figure S38. MALDI-analysis of ACP-bound product of in vitro biotransformation of **16**.

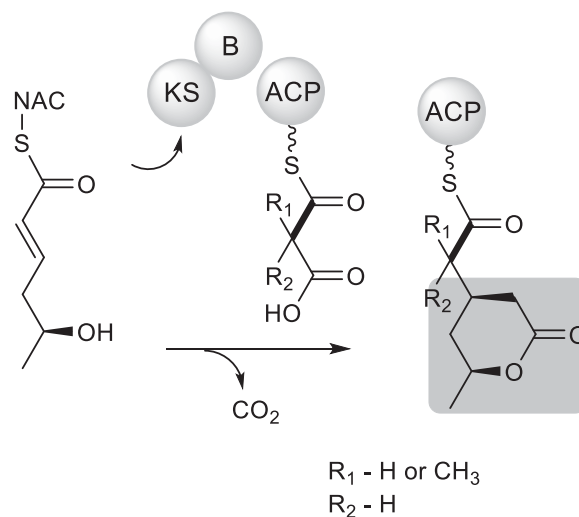
9. Supplemental References

- [1] T. Bretschneider, J. B. Heim, D. Heine, R. Winkler, B. Busch, B. Kusebauch, T. Stehle, G. Zocher, C. Hertweck, *Nature* **2013**, *502*, 124-128.
- [2] V. Theodorou, M. Gogou, M. Philippidou, V. Ragoussis, G. Paraskevopoulos, K. Skobridis, *Tetrahedron* **2011**, *67*, 5630-5634.

3.3 Manuscript B

Twofold polyketide branching by a stereoselective enzymatic Michael addition

Daniel Heine, Srividhya Sundaram, Tom Bretschneider, Christian Hertweck: Twofold polyketide branching by a stereoselective enzymatic Michael addition, *Chemical Communications* 51 (48), 9872-9875, 2015.



The branching module of the rhizoxin PKS accepts methylmalonyl-CoA in addition to the regular malonyl-CoA extender unit and introduce two branches in a single step by Michael addition and a subsequent lactonization.

ChemComm



COMMUNICATION

View Article Online
View Journal | View IssueCite this: *Chem. Commun.*, 2015, 51, 9872Received 14th April 2015,
Accepted 12th May 2015

DOI: 10.1039/c5cc03085d

www.rsc.org/chemcomm

Twofold polyketide branching by a stereoselective enzymatic Michael addition†

Daniel Heine,^a Srividhya Sundaram,^a Tom Bretschneider^a and Christian Hertweck^{*ab}

The versatility of the branching module of the rhizoxin polyketide synthase was tested in an *in vitro* enzyme assay with a polyketide mimic and branched (di)methylmalonyl-CoA extender units. Comparison of the products with synthetic reference compounds revealed that the module is able to stereoselectively introduce two branches in one step by a Michael addition–lactonisation sequence, thus expanding the scope of previously studied PKS systems.

Modular polyketide synthases represent miniaturized assembly lines that produce many pharmaceutically relevant natural products with highly diverse structures and biological activities.¹ The biosynthesis of the complex polyketides typically involves simple head-to-tail fusions of activated acyl and malonyl units. Specifically, a ketosynthase (KS) is primed with an acyl unit, a malonyl unit is loaded onto the adjacent acyl carrier protein (ACP) by an acyl transferase (AT), and the KS catalyzes the Claisen condensation of the bound thioesters to give a β -keto thioester.² The variety of polyketide structures result from several levels of polyketide diversification, including the size of the chain, the type of building blocks incorporated, the degree of β -keto processing, and diverse enzymatic tailoring reactions during or downstream of chain assembly.³ Yet, the range of realized metabolite structures is limited by the canonical KS-catalyzed Claisen condensation, which exclusively yields linear polyketide backbones. Deviations from this scheme are extremely rare, but can have a major impact on the polyketide structures and their binding to biological targets. One of these examples is the phytotoxic rhizoxin complex, which comprises some of the strongest antimitotic agents known to date.⁴ Rhizoxin is produced in an unusual symbiosis

of the bacterium *Burkholderia rhizoxinica* that lives in the cytosol of the plant-pathogenic fungus *Rhizopus microsporus*.⁵ The bacterial endosymbionts assemble the macrolactone core,⁴ and the fungus increases phytotoxicity of the macrolide by oxidative tailoring.^{5b} Most importantly, the structure of this highly active antiproliferative agent features an internal δ -lactone residue that proved to be crucial for the effective interaction of the polyketide with its biological target, β -tubulin.⁶ The small lactone ring branches off the polyketide backbone in the β -position of a former acetate unit.⁷ In general, branches in polyketide backbones may be introduced by the utilization of branched starter units,⁸ chain elongation by substituted malonyl groups,⁹ or α -methylation by *S*-adenosylmethionine (SAM)-dependent methyltransferases, in particular in modular *trans*-AT PKS.¹⁰ The polyketide can also be alkylated at the β -keto group by a mechanism known from mevalonate biosynthesis.¹¹ However, none of these mechanisms could rationalize the formation of the δ -lactone in rhizoxin, in particular since the rhizoxin-producing symbiont lacks genes for isoprenoid-type branching enzymes.¹² Dissection of the modular rhizoxin (rhi) PKS¹³ and structural analysis of all pathway intermediates¹⁴ suggested that the branch resulted from the vinylogous addition of a malonyl-unit to an α,β -unsaturated thioester. The architecture of the modular PKS, which is colinear with the rhizoxin structure,^{13a,15} suggested that an unprecedented module consisting of KS, B (branching) and ACP domains is involved in the chain branching reaction (Fig. 1).

The *in vitro* reconstitution of the 'branching module' was succeeded by mixing heterologously produced, soluble KS-B didomain, ACP and *trans*-AT.¹⁶ By using synthetic pathway surrogates, mutational analyses and elucidation of the KS-B protein structure we gained first insight into the course of the reaction.¹⁶ Accordingly, the C–C bond is formed by Michael addition of the C-2 of the ACP-bound malonyl, and nucleophilic attack of the δ -hydroxyl group at the thioester-bound polyketide yields the lactone ring.¹⁶ We also found that the branching module tolerates alternative substrates with amine and amide substituents in lieu of the alcohol, giving rise to δ -lactam and

^a Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Beutenbergstr. 11a, 07745 Jena, Germany. E-mail: christian.hertweck@leibniz-hki.de

^b Chair for Natural Product Chemistry, Friedrich Schiller University, 07743 Jena, Germany

† Electronic supplementary information (ESI) available: Experimental procedures, NMR spectra and analytical data of all synthetic compounds. See DOI: 10.1039/c5cc03085d

View Article Online

ChemComm

Communication

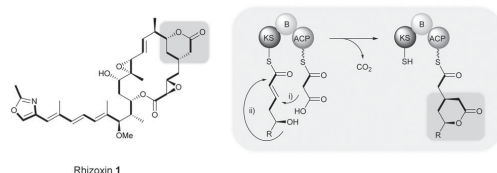


Fig. 1 Structure of rhizoxin (**1**) and model for Michael-type polyketide branching to form the δ -lactone side chain; (i) Michael addition and (ii) hydrolysis of the thioester bond at the KS and lactonisation. KS, ketosynthase; B, branching domain; ACP, acyl carrier protein.

glutarimide residues, respectively.¹⁷ An open question remained whether or not the branching module would also allow the incorporation of a substituted extender unit, which would lead to two chain branches and two new stereocenters in a single biotransformation. Here, we report the finding that the KS-B module is indeed capable of introducing a branched extender unit and elucidate the mechanistic and stereochemical implications of this reaction. To test the scope of the vinylogous chain branching reaction *in vitro* we interrogated the module with a surrogate of the growing polyketide chain and substituted malonyl-CoA thioesters. The functional module was successfully reconstituted using His-tagged KS-B didomain, *holo*-ACP and *trans*-AT (Fig. 2). As surrogate, an α,β -unsaturated thioester was designed that contains an *S*-configured hydroxyl group in δ -position, as in the natural substrate produced upstream of the β -branching module.^{16,17} The *N*-acetyl cysteamine (SNAC) thioester mimics the phosphopantetheinyl linker in the natural ACP, which allows for the transfer of the acyl onto the KS by transesterification. To monitor product formation in the *in vitro* assay the ACP-lactone adduct was measured by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) analysis, and the detached lactone was analyzed by high-performance liquid chromatography-high resolution mass spectrometry (HPLC-HRMS) using synthetic references. The successful transfer of methylmalonyl-CoA (MMCoA) onto the ACP is a prerequisite for the biotransformation. In this context it should be noted that the loading of MMCoA onto ACP by a regular *trans*-AT is unusual as *trans*-AT PKSs typically employ malonyl-CoA (MCoA), but a degree of flexibility towards MMCoA *in vitro* has been observed previously in some systems.¹⁸

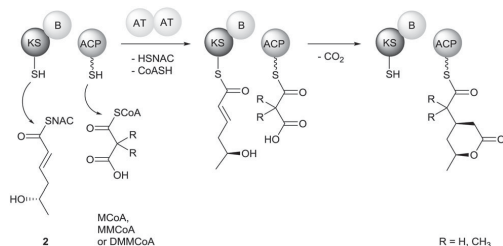


Fig. 2 Experimental setup for the *in vitro* β -branching assay. SNAC, *N*-acetylcysteamine; AT, acyltransferase; MCoA, malonyl-CoA, MMCoA, methylmalonyl-CoA, DMCoA, dimethylmalonyl-CoA.

The successful formation of the desired adduct was verified by MALDI-TOF. When performing the assay with MMCoA in place of MCoA, we detected a new signal in the MALDI trace at $m/z = 12738$ (Fig. 3). Comparison with the signal of the *holo*-ACP indicated a mass shift of 168 Da, which points towards the effective transfer of the substrate and the incorporation of a C_3H_5O unit. Notably, performing the assay with a heat-inactivated version of the KS or without substrate (substrate blank) did not give any comparable signals, thus proving that the reaction is enzyme catalyzed. After workup and extraction of the new product with methanol, HPLC-HRMS revealed the formation of two new compounds with $m/z = 187.0964$ $[M + H]^+$, which correspond to a molecular formula of $C_9H_{15}O_4$. The deduced molecular mass differs from the product obtained with MCoA by 14 Da, thus indicating that MMCoA was successfully incorporated. Although MSⁿ fragmentation data already strongly suggested that δ -lactones with methyl branches are produced, a synthetic reference was required to (a) verify its constitution and to elucidate (b) the exact position of the methyl group (at the side chain or attached to the lactone ring) and (c) the configuration of the substituents. Synthesis of the reference was initiated by a *Z*-selective Still-Gennari olefination variant of the Horner-Wadsworth-Emmons reaction.¹⁹ Deprotection of the THP ether **6** by *p*-toluenesulfonic acid gave *p*-sorbic acid (**7**) in excellent yield. Michael-addition of diethyl methylmalonate at elevated temperature proceeded smoothly, but resulted in a mixture of *syn*-*anti*-diastereoisomers of the product of the Michael addition.

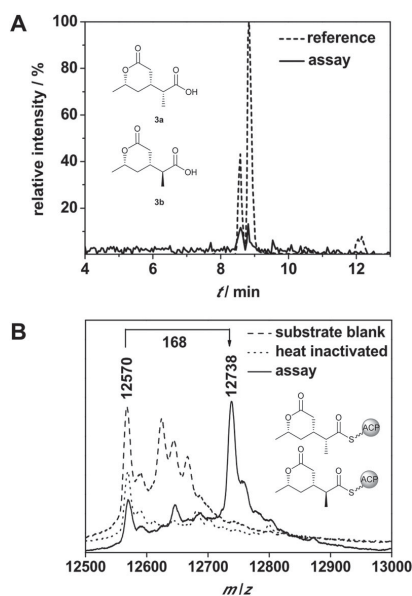


Fig. 3 (A) SIM-LC-HRMS ($m/z = 185.0814$, negative mode) analysis of hydrolyzed product and comparison with the synthetic reference (mixture of diastereoisomers). (B) MALDI analysis of ACP-bound product. SIM, selected ion monitoring.

View Article Online

Communication

ChemComm

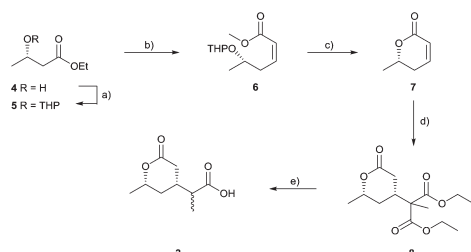


Fig. 4 Chemical synthesis of product reference **3**. (a) 3,4-Dihydro-2H-pyran, PPTS, CH_2Cl_2 , 0 °C then RT, 93% (b) (1) DIBAL, toluene, -78 °C, (2) methyl *P,P*-bis(2,2,2-trifluoroethyl)phosphonoacetate, KHMDS, 18-crown-6, THF, -78 °C, 2 steps 55% (95% *cis*) (c) TosOH, CH_2Cl_2 , 94% (d) diethyl methyl-malonate, Cs_2CO_3 , acetonitrile, 0 °C then RT, 69% (e) 2 M HCl, 100 °C, 78%. PPTS = pyridinium *p*-toluenesulfonate, DIBAL = diisobutylaluminum hydride, KHMDS = potassium bis(trimethylsilyl)amide, THF = tetrahydrofuran, TosOH = *p*-toluenesulfonic acid.

To improve diastereoselectivity, the reaction conditions were altered, eventually leading to a single product of the vinylogous addition at room temperature. The *syn*-configuration of **8** was proven by NOE spectroscopy. Acidic hydrolysis resulted in decarboxylation, and yielded the desired lactone reference **3** in 78% yield (Fig. 4). Analysis of 2D NMR data including NOESY data shed light on the relative configuration of product reference **3**. Substituents at C-3 and C-5 are *syn*, and since C-5 (*S*) is predefined in the δ -hydroxy ester **4**, position C-3 has an *R* configuration. As position C-2 is not defined, reference **3** exists as a 1:1 mixture of the two diastereoisomers, **3a** and **3b** (Fig. S1, ESI[†]). HPLC-HRMS comparison of the synthetic reference with the product of the enzyme assay showed that the synthetic diastereoisomers **3a/b** are identical with the two compounds formed in the *in vitro* assay. Thus, we unequivocally verified the structures of the reaction products. Since both *R*- and *S*-MMCoA are accepted as substrates for the branching module, the active site obviously provides sufficient space for both enantiomers. Thus, it appeared reasonable to examine whether dimethylmalonyl-CoA (DMMCoA) would also be accepted as a branching/elongation unit. Recently, we have shown that a KS-like enzyme loads DMMCoA onto a sugar residue of the antibiotic cervimycin.²⁰ Furthermore, DMMCoA is accepted as an elongation unit by the yersiniabactin and epothilone PKSs.²¹ To test whether this substrate is also incorporated into the rhizoxin substructure, we synthetically prepared DMMCoA and loaded the dimethylmalonyl moiety onto the *holo*-ACP by means of the *trans*-AT. Since the *trans*-AT-catalyzed malonyl transfer could be a potential bottleneck, we verified the successful formation of DMM-ACP by MALDI-TOF analysis. It is remarkable that the *trans*-AT is also capable of transferring this unusual building block. We then applied the ACP adduct to the *in vitro* enzyme assay. Monitoring the fate of DMMCoA-ACP in the *in vitro* branching reaction by MALDI and HPLC-HRMS did not show any formation of a dimethylated product. However, we observed the decarboxylated DMM-ACP, isobutyryl-ACP by MALDI analyses (Fig. S2, ESI[†]). Thus, one can conclude that

DMM-ACP is not accepted as substrate for the branching module. A plausible explanation could be that the space for the elongation unit is limited and DMM might be orientated differently in the KS, hampering C-C-bond formation. Our findings, in particular the results obtained with MMCoA, have several implications for the reaction mechanism of the rare PKS branching module. First, the finding that the methyl branch is exclusively located in the side chain supports the proposed reaction mechanism in which the δ -hydroxy group of the Michael adduct attacks the thioester carbonyl at the KS. As a consequence, the initial chain is terminated and the newly introduced malonyl-derived carbon chain is passed on to the downstream module for further elongation. Second, we found that the reaction is highly specific with respect to diastereocontrol of lactone ring formation. MMCoA attacks from the *Re* face, resulting in the *R*-configuration of the β -carbon atom. In contrast, the configuration of the α -position appears as a mixture of both configurations. It is well conceivable that an isomerisation occurs after decarboxylation and formation of the carbon nucleophile. The proposed mechanism also implies that the vinylogous attack of an *E*-enolate would yield a product with the opposite configuration at the β -carbon (Fig. 5).

In conclusion, we have shown that the KS-B didomain is capable of performing a twofold chain branching by Michael addition of a substituted extender unit. Besides the incorporation of MCoA extender units, the KS-B didomain allows for the incorporation of MMCoA resulting in a double-branch of the polyketide backbone. While attack of the methyl malonyl unit at the α,β -unsaturated thioester and the subsequent lactonisation take place with marked *syn*-diastereoselectivity, an *R/S*-mixture was observed for the third stereocenter that results from the incorporation of the substituted malonyl unit. Our results grant insight into the scope of the unusual PKS module and strongly support the proposed reaction mechanism. From a synthetic biology perspective, these findings are important since the use of the branching module could greatly enhance the current options for polyketide chain diversification. In particular, the alteration of complex polyketides that are not easily accessible by

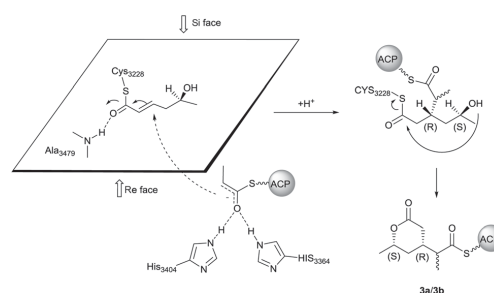


Fig. 5 Proposed mechanism of the diastereoselective Michael addition of the decarboxylated, branched extender unit MMCoA at the polyketide surrogate **2**. The nucleophilic attack exclusively takes place at the *Re* face, resulting in an *R* configured new chiral center. No rigorous stereocontrol over the second stereocenter formed in this reaction was observed.

View Article Online

Communication

ChemComm

synthetic methods and their sustainable production might come within reach.

The authors would like to thank H. Heinecke for recording NMR spectra, A. Perner for HPLC-MS measurements and C. Karkowski for assistance in general lab work. We are grateful to the "Studienstiftung des deutschen Volkes" for providing a scholarship to D.H.

Notes and references

- (a) C. Hertweck, *Angew. Chem., Int. Ed.*, 2009, **48**, 4688–4716; (b) M. A. Fischbach and C. T. Walsh, *Chem. Rev.*, 2006, **106**, 3468–3496.
- C. Khosla, D. Herschlag, D. E. Cane and C. T. Walsh, *Biochemistry*, 2014, **53**, 2875–2883.
- C. Hertweck, *Trends Biochem. Sci.*, 2015, **40**, 189–199.
- K. Scherlach, L. P. Partida-Martinez, H.-M. Dahse and C. Hertweck, *J. Am. Chem. Soc.*, 2006, **128**, 11529–11536.
- (a) L. P. Partida-Martinez and C. Hertweck, *Nature*, 2005, **437**, 884–888; (b) K. Scherlach, B. Busch, G. Lackner, U. Paszkowski and C. Hertweck, *Angew. Chem., Int. Ed.*, 2012, **51**, 9615–9618; (c) G. Lackner, L. P. Partida-Martinez and C. Hertweck, *Trends Microbiol.*, 2009, **17**, 570–576; (d) G. Lackner and C. Hertweck, *PLoS Pathog.*, 2011, **7**, e1002096.
- (a) M. Takahashi, S. Iwasaki, H. Kobayashi, S. Okuda, T. Murai and Y. Sato, *Biochim. Biophys. Acta*, 1987, **926**, 215–223; (b) I. Schmitt, L. P. Partida-Martinez, R. Winkler, K. Voigt, A. Einax, J. Wöstemeyer and C. Hertweck, *ISME J.*, 2008, **2**, 632–641; (c) B. Kusebauch, K. Scherlach, H. Kirchner, H. M. Dahse and C. Hertweck, *ChemMedChem*, 2011, **6**, 1998–2001.
- H. Kobayashi, S. Iwasaki, E. Yamada and S. Okuda, *J. Chem. Soc., Chem. Commun.*, 1986, 1702–1703.
- B. S. Moore and C. Hertweck, *Nat. Prod. Rep.*, 2002, **19**, 70–99.
- Y. A. Chan, A. M. Podevels, B. M. Kevany and M. G. Thomas, *Nat. Prod. Rep.*, 2009, **26**, 90–114.
- J. Piel, *Nat. Prod. Rep.*, 2010, **27**, 996–1047.
- C. T. Calderone, *Nat. Prod. Rep.*, 2008, **25**, 845–853.
- (a) G. Lackner, N. Moebius, L. P. Partida-Martinez and C. Hertweck, *J. Bacteriol.*, 2011, **193**, 783–784; (b) G. Lackner, N. Moebius, L. P. Partida-Martinez, S. Boland and C. Hertweck, *BMC Genomics*, 2011, **12**, 210.
- (a) L. P. Partida-Martinez and C. Hertweck, *ChemBioChem*, 2007, **8**, 41–45; (b) T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck and J. Piel, *Nat. Biotechnol.*, 2008, **26**, 225–233.
- (a) B. Kusebauch, B. Busch, K. Scherlach, M. Roth and C. Hertweck, *Angew. Chem., Int. Ed.*, 2010, **49**, 1460–1464; (b) B. Kusebauch, B. Busch, K. Scherlach, M. Roth and C. Hertweck, *Angew. Chem., Int. Ed.*, 2009, **48**, 5001–5004.
- N. Brendel, L. P. Partida-Martinez, K. Scherlach and C. Hertweck, *Org. Biomol. Chem.*, 2007, **5**, 2211–2213.
- T. Bretschneider, J. B. Heim, D. Heine, R. Winkler, B. Busch, B. Kusebauch, T. Stehle, G. Zocher and C. Hertweck, *Nature*, 2013, **502**, 124–128.
- D. Heine, T. Bretschneider, S. Sundaram and C. Hertweck, *Angew. Chem., Int. Ed.*, 2014, **53**, 11645–11649.
- (a) N. B. Lopanik, J. A. Shields, T. J. Buchholz, C. M. Rath, J. Hotherhall, M. G. Haygood, K. Hakansson, C. M. Thomas and D. H. Sherman, *Chem. Biol.*, 2008, **15**, 1175–1186; (b) B. J. Dunn, K. R. Watts, T. Robbins, D. E. Cane and C. Khosla, *Biochemistry*, 2014, **53**, 3796–3806.
- W. C. Still and C. Gennari, *Tetrahedron Lett.*, 1983, **24**, 4405–4408.
- T. Bretschneider, G. Zocher, M. Unger, K. Scherlach, T. Stehle and C. Hertweck, *Nat. Chem. Biol.*, 2012, **8**, 154–161.
- S. Poust, R. M. Phelan, K. Deng, L. Katz, C. J. Petzold and J. D. Keasling, *Angew. Chem., Int. Ed.*, 2015, **127**, 2400–2403.

Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2015

-Supporting Information for-

Twofold Polyketide Branching by a Stereoselective Enzymatic Michael Addition

Daniel Heine,^a Srividhya Sundaram,^a Tom Bretschneider^a and Christian Hertweck^{a,b*}

Instrumentation

NMR spectra were recorded on Bruker AVANCE II 300, AVANCE III 500 or 600 MHz spectrometer, equipped with a Bruker Cryo platform. Chemical shifts are reported in parts per million (ppm) relative to the solvent residual peak of chloroform-d₁ (¹H: 7.24 ppm, singlet; ¹³C: 77.00 ppm, triplet) or methanol-d₄ (¹H: 3.30 ppm, quintet; ¹³C: 49.00 ppm, septet). MALDI measurements were performed with a Bruker UltrafleXtreme machine. HRESI-MS was carried out with an Accela UPLC-system (Thermo Scientific) combined with an Exactive mass spectrometer (Thermo Scientific) equipped with an electrospray ion source. HPLC-MS measurements were performed on a HPLC 1100 System connected to a 1100 Series LC/MSD Trap using a Zorbax Eclipse XDB-C8, 4.6 x 150 mm column, particle size 5 μm (Agilent Technologies). Semi-preparative HPLC was performed on a 1260 Infinity System (Agilent Technologies) using the following columns: Zorbax Eclipse XDB-C8, 9.5 x 250 mm, particle size 5 μm (Agilent Technologies), Zorbax Eclipse XDB-C18, 9.5 x 250 mm, particle size 5 μm (Agilent Technologies) and EC 250/4 Nucleodur Sphinx RP, particle size 5 μm (Macherey-Nagel). All solvents for analytical and preparative HPLC measurements were obtained commercially at least in gradient grade and were filtered prior to use. To avoid microbial growth 0.1 % formic acid was added to the water, used for analytical and preparative HPLC. IR-spectra were recorded on an FT/IR-4100 ATR spectrometer (Jasco).

General experimental procedures

All reagents were obtained from commercial suppliers (Sigma Aldrich and Acros Organics) and were used without further purification. Chemical solvents were obtained at least in HPLC grade and dried prior to use by the following procedures: dichloromethane (DCM) was distilled from a calcium hydride suspension. Tetrahydrofurane (THF) was pre-dried over potassium hydroxide and then freshly distilled from lithium aluminum hydride. Acetonitrile was distilled from P₂O₅. Toluene was dried over molecular sieve 4A. Methanol, chloroform and ethyl acetate were distilled prior to use. Reactions were performed under inert atmosphere (Argon) using the Schlenk technique. For thin-layer chromatography TLC aluminum sheets silica gel 60 F₂₅₄ (MERCK) were used. Open column chromatography was performed on silica gel 60; particle size 0.015 - 0.04 mm (Macherey-Nagel) and on Sephadex® LH-20 (Sigma-Aldrich).

Protein purification

The proteins used in the assay were purified as described previously.^[17] Briefly, 5 mg of *Escherichia coli* BL21 cells containing the expression plasmids (pTB57, pTB41 and pTB39 for KS-B, ACP-holo and RhiG respectively) were suspended in 50 mL of equilibration buffer (20 mM Tris buffer (pH 7.5) supplemented with 200 mM NaCl and 50 mM imidazole). Cell disruption was carried out by sonication (Sonotrode, 20% power) and the cell debris was removed by centrifugation at 9,000 g for 20 min at 4 °C. The supernatant was loaded onto a 5 mL Ni-NTA matrix (Macherey Nagel) connected to a FPLC machine (Akta Explorer, Amersham Biosciences). The recombinant protein was eluted using the same buffer with 500 mM imidazole. The protein was diluted twice with 20 mM Tris (pH 7.5) and purified further by an ion-exchange column (5 mL HiTrap Q HP, GE Healthcare). Elution of the protein was carried out over 30 column volumes of 20 mM Tris (pH 7.5) and 1 M NaCl and the fractions containing target protein were concentrated using a vivaspin column (MWCO 5 and 10 kDa – Sartorius). All proteins were stored in 4 °C for *in vitro* assays.

***In vitro* biotransformation**

In vitro reconstitution of the branching reaction was performed as reported previously.^[16] In a total of 40 μL reaction, 167 μM ACP, 3 μM KS-B, 0.2 μM RhiG, 750 μM MMCoA and 1,000 μM (*R,S*)-(*E*)-*S*-(2-acetamidoethyl) 5-hydroxyhex-2-enethioate were incubated in 20 mM Tris buffer (pH 7.0) at 23 $^{\circ}\text{C}$ with shaking at 400 rpm. After 2 h, 2 μL of the sample was mixed with 2 μL of 100 mM 2',5'-dihydroxyacetophenone (containing 2.5 mM diammonium hydrogen citrate) and spotted onto an anchor chip 800/384 T F target (Bruker) to analyse the product attached to ACP-holo. For mass measurements, the enzyme reaction was stopped by adding an equal volume of methanol. After a brief spin to precipitate the proteins, the supernatant was directly analyzed by HRMS (Exactive).

Relative configuration of 3

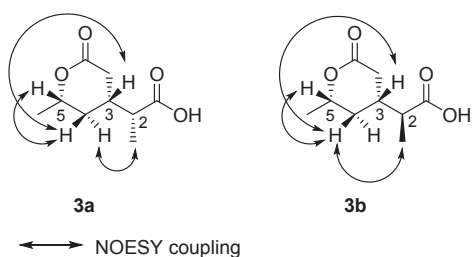


Figure S1. NMR-based analysis of the product reference and selected NOESY correlations of compounds **3a** and **3b**.

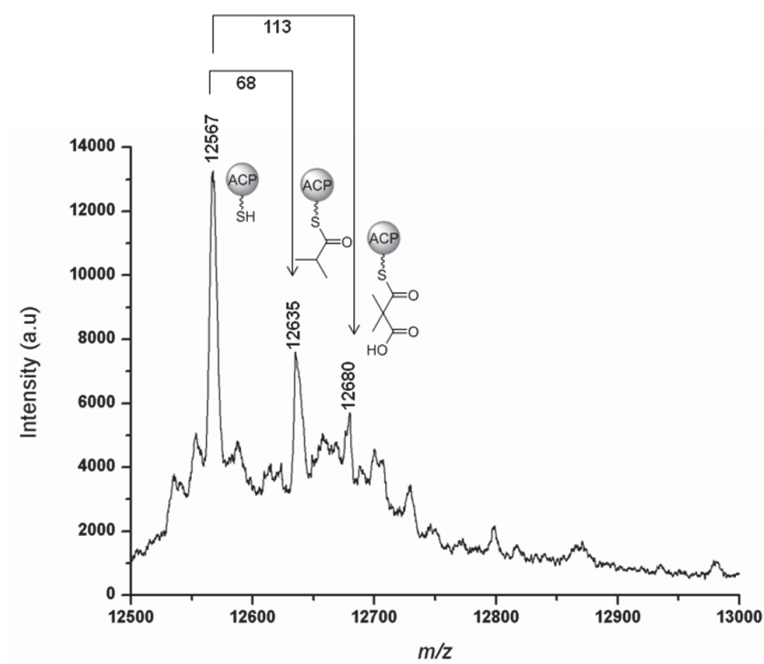
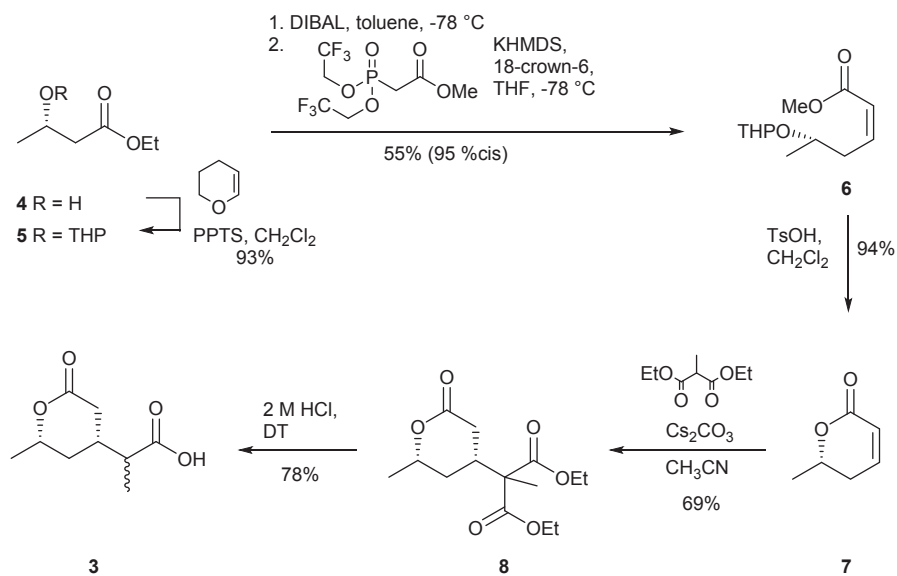
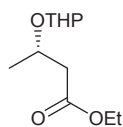


Figure S2. Analysis of DMM-ACP by MALDI-TOF.

Chemical Syntheses



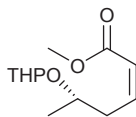
Ethyl (3S)-3-((tetrahydro-2H-pyran-2-yl)oxy)butanoate (5)



5

Compound 5 was synthesized in analogy to a literature procedure.^[1]

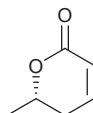
^1H NMR (300 MHz, CDCl_3): δ 1.17 (3H, d, $^3J = 6.1$ Hz), 1.23 (3H, t, $^3J = 7.2$ Hz), 1.22 (3H, t, $^3J = 7.2$ Hz), 1.27 (3H, d, $^3J = 6.2$ Hz), 1.42-1.86 (12 H, m), 2.35 (1H, dd, $^2J = 11.6$ Hz, $^3J = 5.6$ Hz), 2.40 (1H, dd, $^2J = 11.6$ Hz, $^3J = 6.3$ Hz), 2.54 (1H, dd, $^2J = 15.0$, $^3J = 7.5$), 2.65 (1H, dd, $^2J = 15.0$ Hz, $^3J = 7.2$ Hz), 3.41-3.52 (2H, m), 3.77-3.93 (2H, m), 4.06-4.16 (4H, m), 4.16-4.28 (2H, m), 4.67 (1H, m), 4.72 (1H, m) ppm; HRMS (ESI⁺): m/z calculated for $\text{C}_{11}\text{H}_{20}\text{O}_4\text{Na}$: 239.1254, found: 239.1250 $[\text{M}+\text{Na}]^+$.

Z-Methyl (5S)-5-((tetrahydro-2H-pyran-2-yl)oxy)hex-2-enoate 6

6

THP ether **5** (216 mg, 1 mmol) was dissolved in dry toluene (10 mL) and cooled to $-78\text{ }^{\circ}\text{C}$ under argon. DIBAL (1 M in cyclohexane, 2.5 mL) was added dropwise and the solution stirred for two hours. A solution of hydrochloric acid and methanol (1:2, 5 mL) was added and the mixture allowed to reach room temperature. After filtration over Celite and sodium sulfate, the solution was concentrated under reduced pressure and used without further purification. Methyl *P,P*-bis(2,2,2-trifluoroethyl)phosphonoacetate (318 mg, 1 mmol) and 18-crown-6 (1.32 g, 5 mmol) were dissolved in dry tetrahydrofuran (20 mL) and cooled to $-78\text{ }^{\circ}\text{C}$. KHMDS (1 M in toluene, 1 mL) and the crude aldehyde dissolved in dry THF (1 mL) were added slowly. The resulting mixture was stirred for 30 min at $-78\text{ }^{\circ}\text{C}$. Saturated NH_4Cl was added and the aqueous phase was extracted with diethyl ether. The combined organic phase was dried over sodium sulfate and concentrated under reduced pressure. The product was purified by column chromatography (*n*-hexane:ethyl acetate, 4:1) using silica gel to give 125 mg of a clear oil.

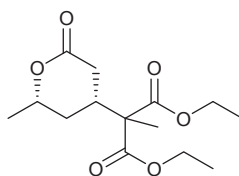
Yield: 55% (95% *cis*). ^1H NMR (300 MHz, CDCl_3): δ 1.13 (3H, d, $^3J=6.2$ Hz), 1.23 (3H, d, $^3J=6.3$ Hz), 1.46-1.86 (12H, m), 2.76-2.97 (4H, m), 3.43-3.53 (2H, m), 3.67 (6H, s), 3.82-3.99 (4 H, m), 4.65-4.69 (2H, m), 5.84 (2H, dt, $^3J=11.6$ Hz, $^4J=1.7$ Hz), 6.29 (1H, dt, $^3J_1=11.6$ Hz, $^3J_2=7.4$ Hz), 6.43 (1H, dt, $^3J_1=11.7$ Hz, $^3J_2=7.2$ Hz). ^{13}C NMR (75 MHz, CDCl_3): δ 19.7, 19.9, 22.7, 23.5, 25.4, 25.5, 30.7, 31.1, 35.2, 36.6, 51.0, 51.0, 62.7, 62.9, 70.8, 72.4, 96.1, 98.1, 120.3, 120.8, 146.3, 147.2, 166.7, 166.8 ppm, IR (film): 813, 869, 996, 1020, 1075, 1131, 1173, 1438, 1650, 1721, 2869, 2942 cm^{-1} , HRMS (ESI⁺): *m/z* calculated for $\text{C}_{13}\text{H}_{21}\text{O}_4$: 241.1445, found: 241.1442 [M+H]⁺.

(S)-6-methyl-5,6-dihydro-2H-pyran-2-one (7)**7**

THP-Ether **6** (326 mg, 1.4 mmol) was dissolved in dichloromethane (10 mL) and *p*-toluenesulfonic acid monohydrate (13.3 mg, 0.07 mmol) dissolved in dichloromethane (1 mL) was added. After 3 h, the solution was concentrated under reduced pressure and the crude product purified by column chromatography over silica gel (dichloromethane) to give 148 mg of a clear oil.

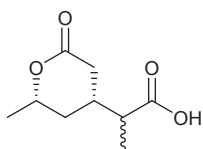
Yield: 94%. $^1\text{H NMR}$ (300 MHz, CDCl_3): δ 1.43 (3H, d, $^3J = 6.3$ Hz), 2.22-2.41 (1H, m), 4.50-4.62 (1H, m), 6.00 (1H, ddd, $^3J_1 = 9.8$ Hz, $^3J_2 = 5.5$ Hz, $^3J_3 = 3.8$ Hz), 6.85 (1H, ddd, $^3J_1 = 9.8$ Hz, $^3J_2 = 2.5$ Hz, $^3J_3 = 1.2$), $^{13}\text{C NMR}$ (75 MHz, CDCl_3): δ 20.7, 31.0, 74.3, 121.4, 144.9, 164.5 ppm; HRMS (ESI⁺): m/z calculated for $\text{C}_6\text{H}_9\text{O}_2$: 113.0597, found: 113.0607 [M+H]⁺.

Analytical data are in full agreement with the reported data.^[2]

Diethyl 2-methyl-2-((2S)-2-methyl-6-oxotetrahydro-2H-pyran-4-yl) malonate 8**8**

Pyrane **7** (134 mg, 1.2 mmol) and diethyl methylmalonate (417 mg, 2.4 mmol, 2 eq.) were dissolved in dry acetonitrile at 0 °C. Cs₂CO₃ (195 mg, 0.6 mmol, 0.5 eq.) was added, and the solution slowly allowed to reach room temperature. After 72 h, the solution was concentrated under reduced pressure and the crude product purified by column chromatography over silica gel (*n*-hexane:ethyl acetate 9:1) to give 236 mg of a clear oil.

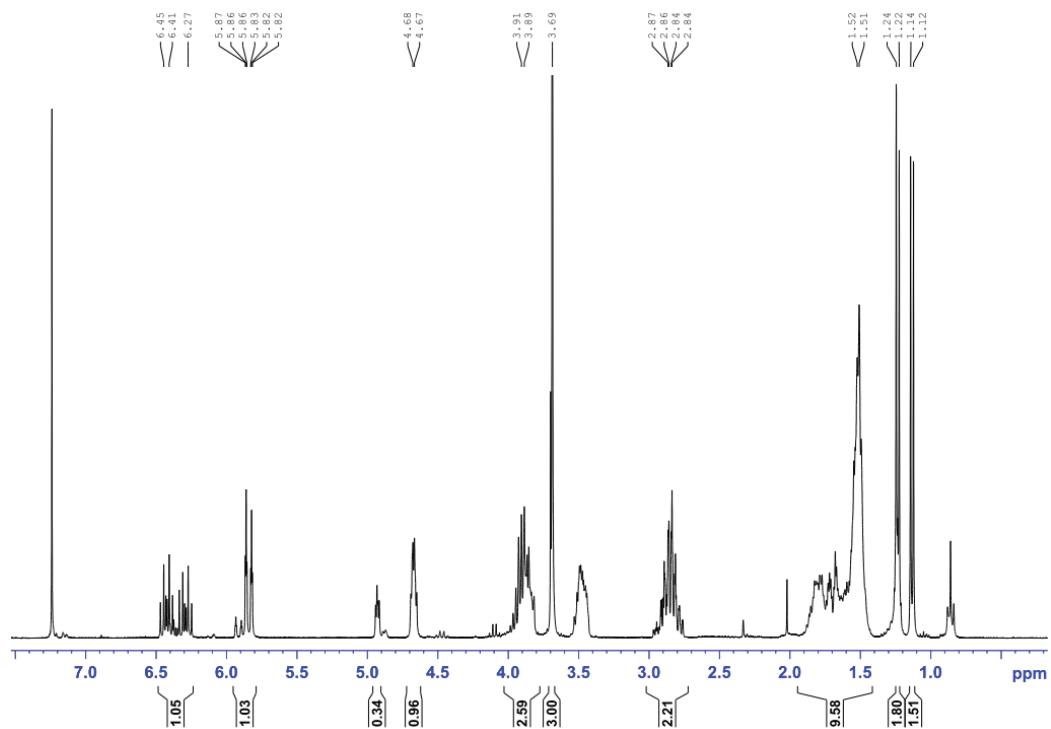
Yield: 69%. ¹H NMR (500 MHz, CDCl₃): δ 1.25 (6H, t, ³J = 7.1 Hz), 1.35 (3H, d, ³J = 6.3 Hz), 1.42 (3H, s), 1.78-1.83 (2H, m) 2.50 (1H, d, ³J = 4.3 Hz), 2.52 (1H, s), 2.65-2.73 (1H, m), 4.16-4.22 (4H, m), 4.49 (1H, m) ¹³C NMR (125 MHz, CDCl₃): δ 14.0, 17.9, 20.6, 31.1, 31.6, 33.8, 56.7, 61.7, 61.7, 73.8, 170.8, 172.4 ppm, IR (film): 860, 1021, 1095, 1154, 1329, 1369, 1456, 1731, 2945, 2983 cm⁻¹, HRMS (ESI⁺): *m/z* calculated for C₁₄H₂₃O₆: 287.1489, found: 287.1487 [M+H]⁺.

(*R,S*)-2-((2*S,4R*)-2-methyl-6-oxotetrahydro-2*H*-pyran-4-yl) propanoic acid 3**3**

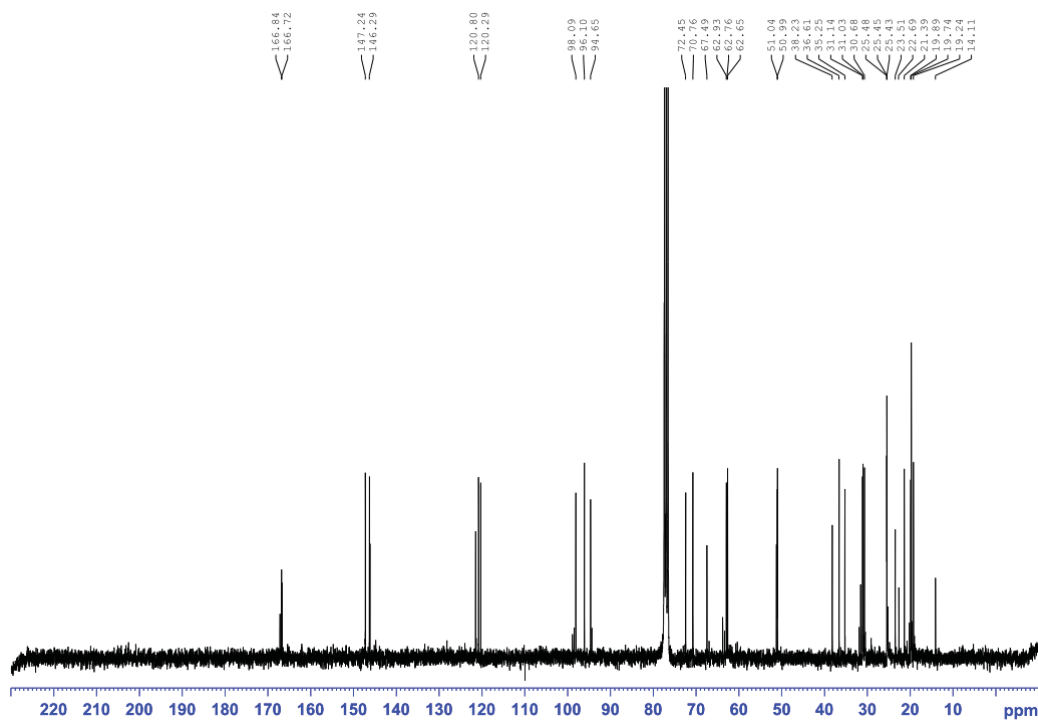
Compound **8** (55 mg, 0.19 mmol) was dissolved in HCl (2 M, 0.5 mL) and heated in a vial with closed-top screw cap to 120 °C for 18 h. Water (1 mL) was added and the aqueous phase was extracted with ethyl acetate (3×1 mL) to remove impurities. The aqueous phase was then concentrated under reduced pressure and subjected to column chromatography over silica gel (chloroform:methanol 98:2) to give a clear oil.

Yield: 78 % ¹H NMR (600 MHz, MeOD): δ 1.13 (3H, d, ³J = 7.1 Hz), 1.17 (3H, d, ³J = 7.0 Hz), 1.32 (3H, d, ³J = 6.2 Hz), 1.34 (3H, d, ³J = 6.3 Hz), 1.35-1.39 (1H, m), 1.77 (1H, ddd, ²J = 14.4 Hz, ³J₁ = 8.2 Hz, ³J₂ = 8.1 Hz), 1.81 (1H, ddd, ²J = 14.4 Hz, ³J₁ = 6.6 Hz, ³J₂ = 4.5 Hz), 2.01 (1H, dd, ³J = 15.5 Hz, ²J = 9.6 Hz), 2.10- 2.15 (1H, m), 2.21-2.27 (1H, m), 2.36-2.39 (1H, m), 2.40 (1H, dd, ²J = 9.6 Hz, ³J = 5.4 Hz), 2.59-2.66 (1H, m), 2.92 (1H, qd, ³J₁ = 7.1 Hz, ³J₂ = 7.0 Hz), 4.49-4.55 (1H, m), 4.55-4.61 (1H, m) ¹³C NMR (150 MHz, MeOD): δ 12.6, 15.0, 21.2, 21.5, 33.0, 33.1, 33.1, 34.2, 35.7, 37.9, 38.6, 45.9, 75.6, 76.3, 174.1, 175.6, 178.1, 178.4 ppm; IR (film): 1021, 1112, 1416, 1449, 1653, 2341, 2832, 2945, 3324 cm⁻¹, HRMS (ESI⁺): *m/z* calculated for C₉H₁₅O₄: 187.0965, found: 187.0964 [M+H]⁺.

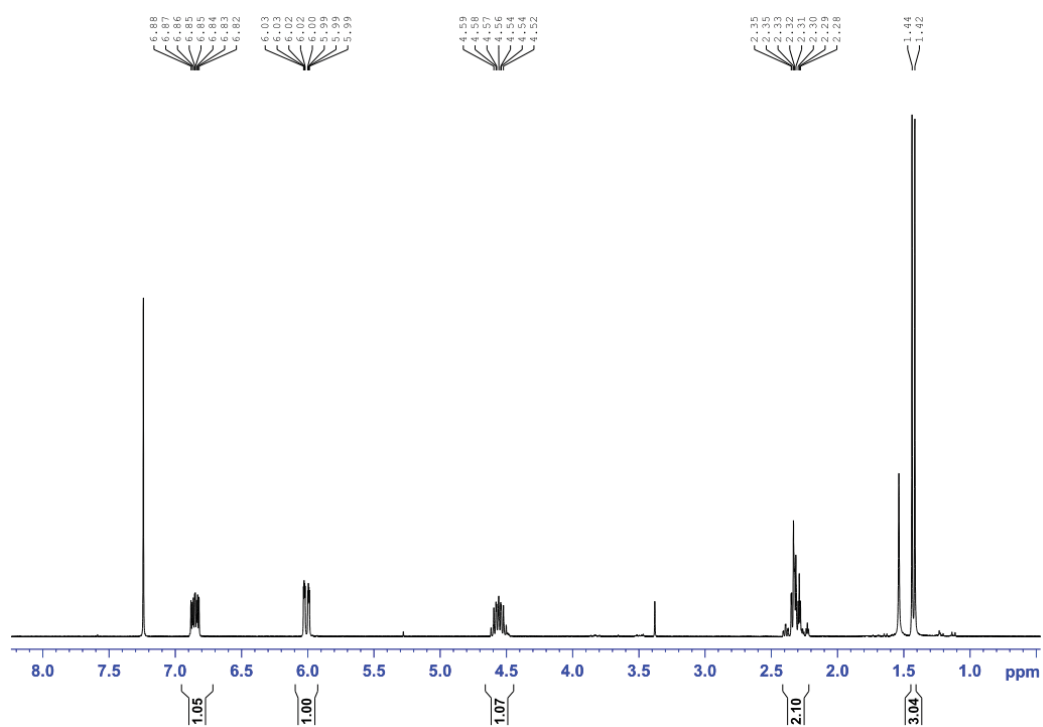
3. NMR spectra



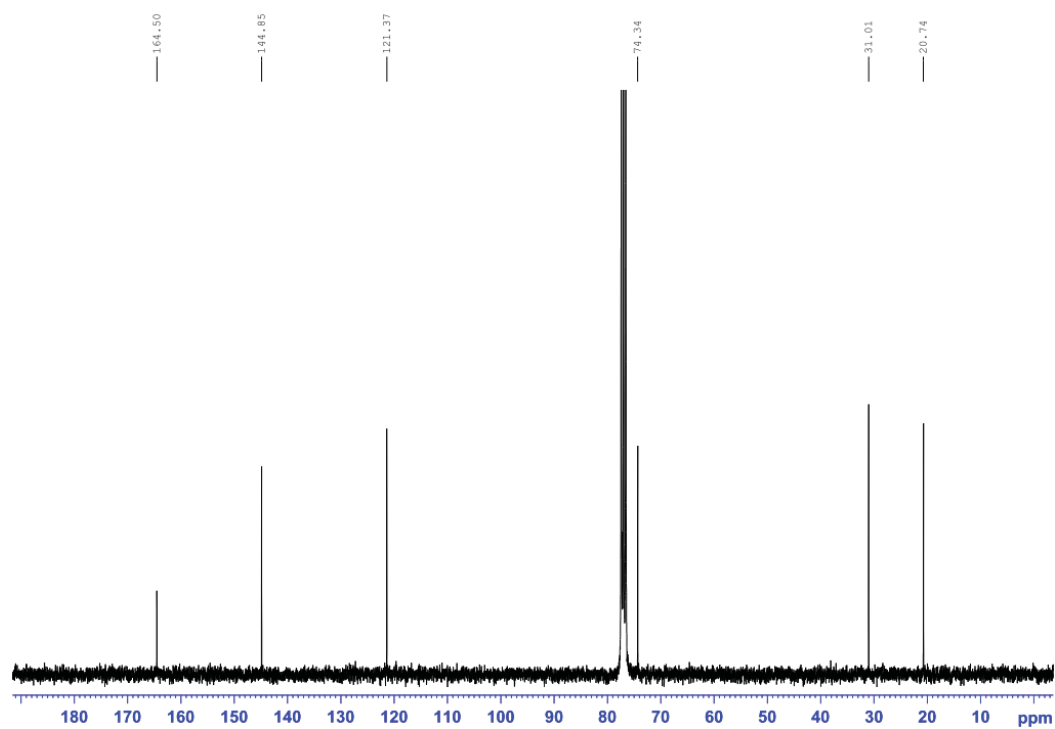
¹H NMR Spectrum (300 MHz, CDCl₃) of compound **6**.



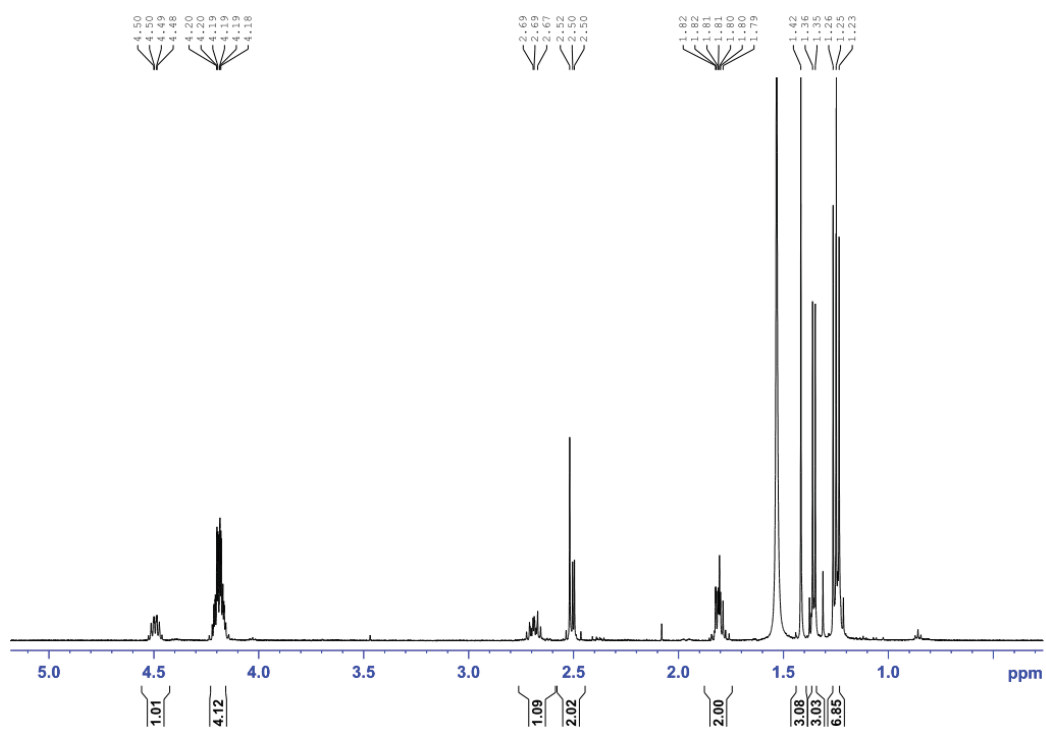
¹³C NMR Spectrum (300 MHz, CDCl₃) of compound **6**.



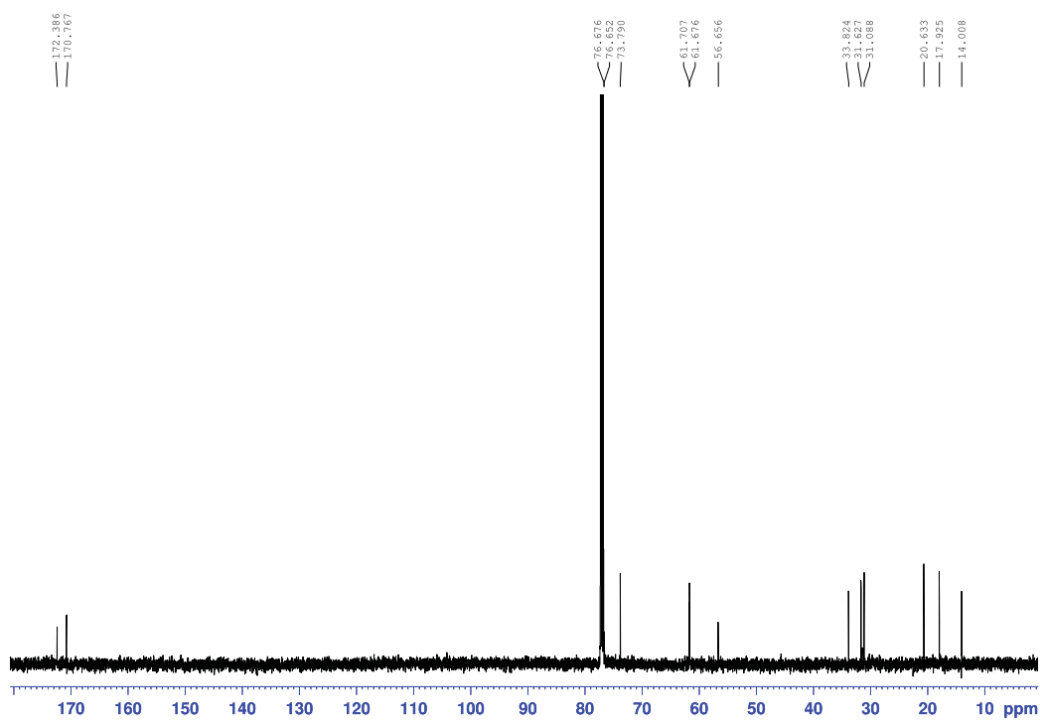
¹H NMR Spectrum (300 MHz, CDCl₃) of compound 7.

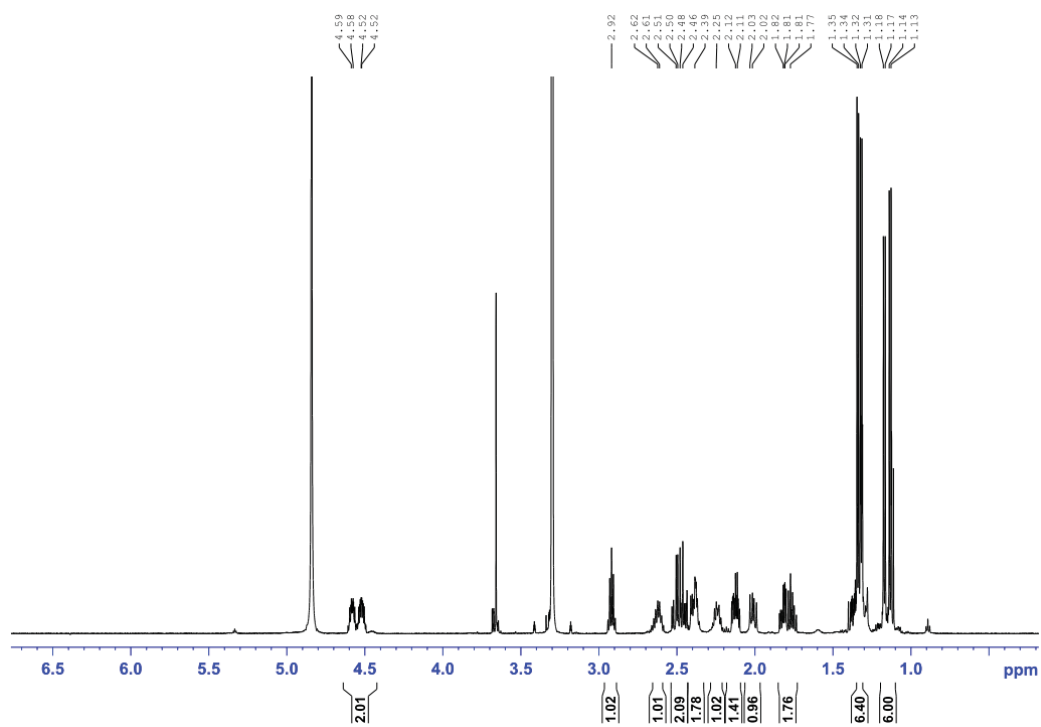


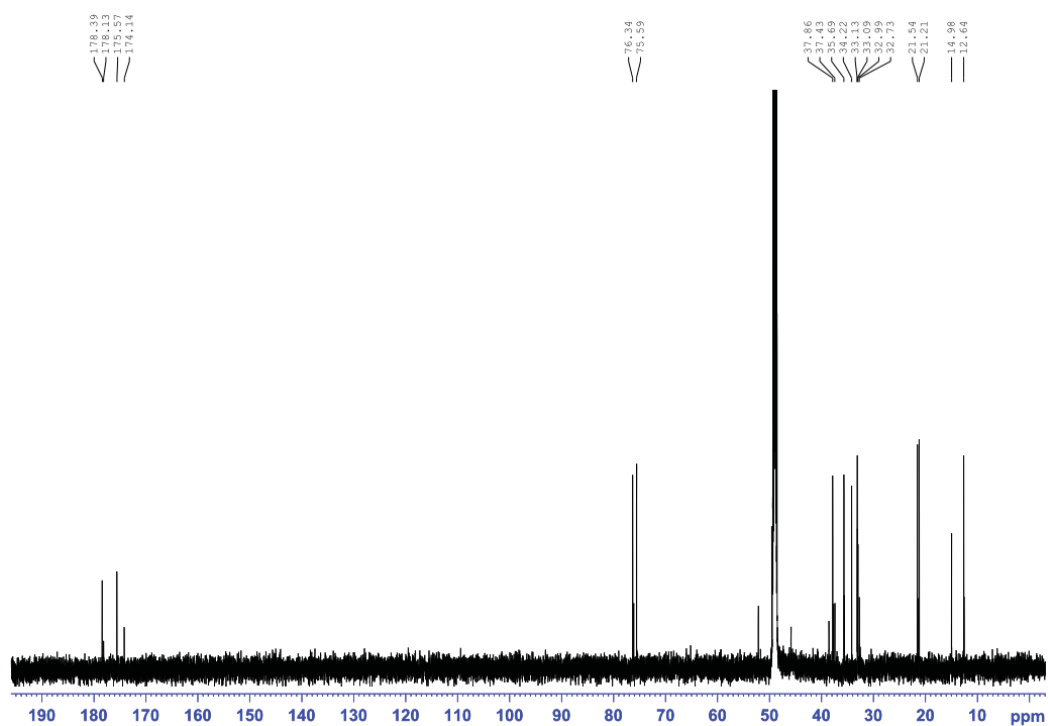
¹³C NMR Spectrum (75 MHz, CDCl₃) of compound 7.



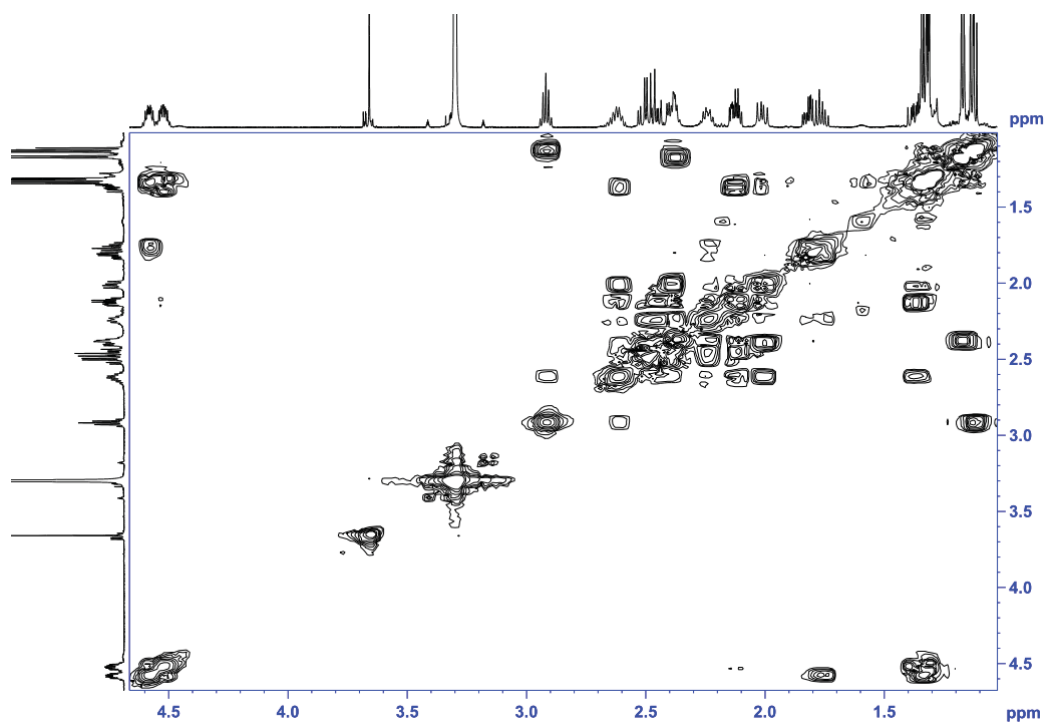
^1H NMR Spectrum (500 MHz, CDCl_3) of compound **8**.



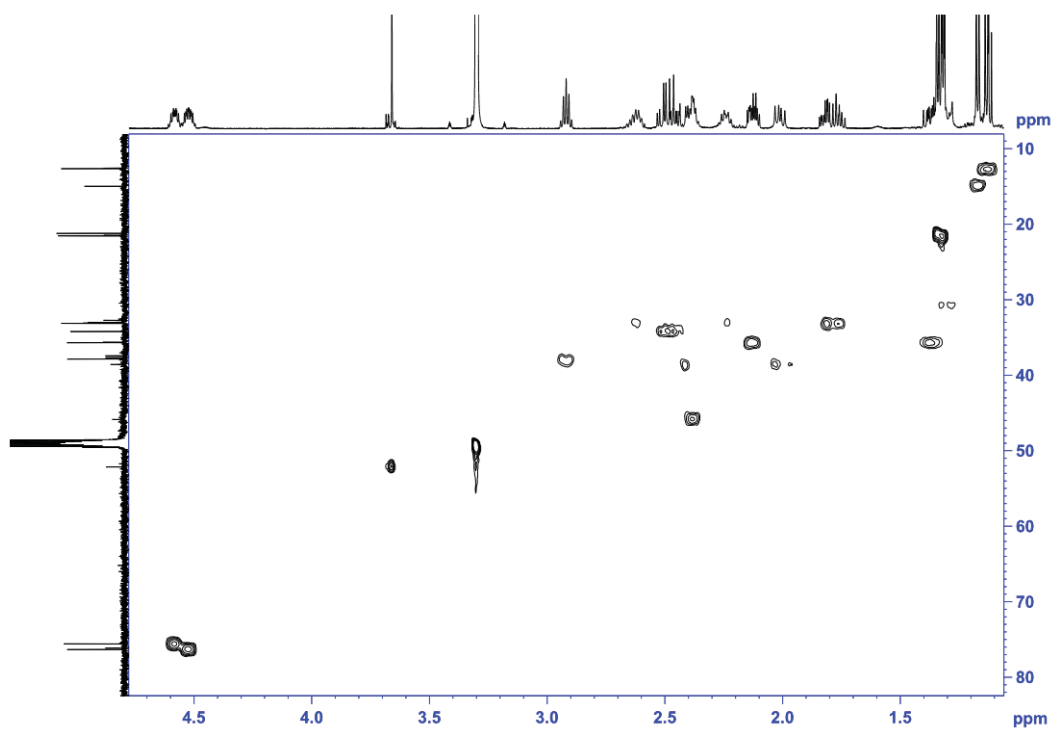
^{13}C NMR Spectrum (125 MHz, CDCl_3) of compound **8**. ^1H NMR Spectrum (600 MHz, MeOD) of compound **3**.



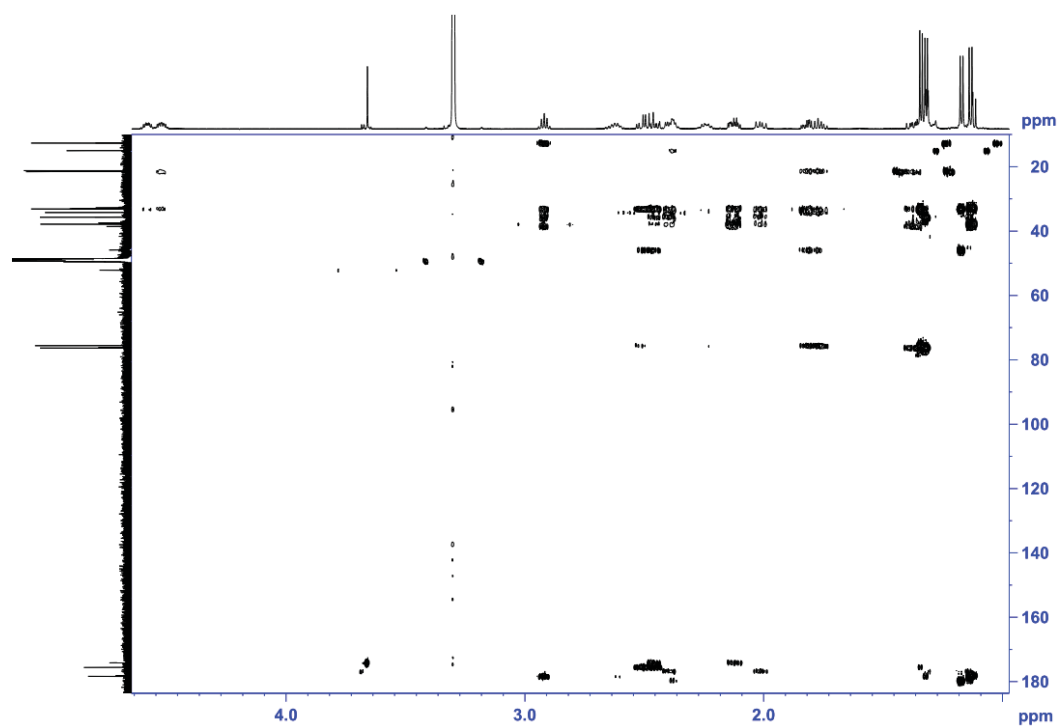
^{13}C NMR Spectrum (150 MHz, MeOD) of compound **3**.



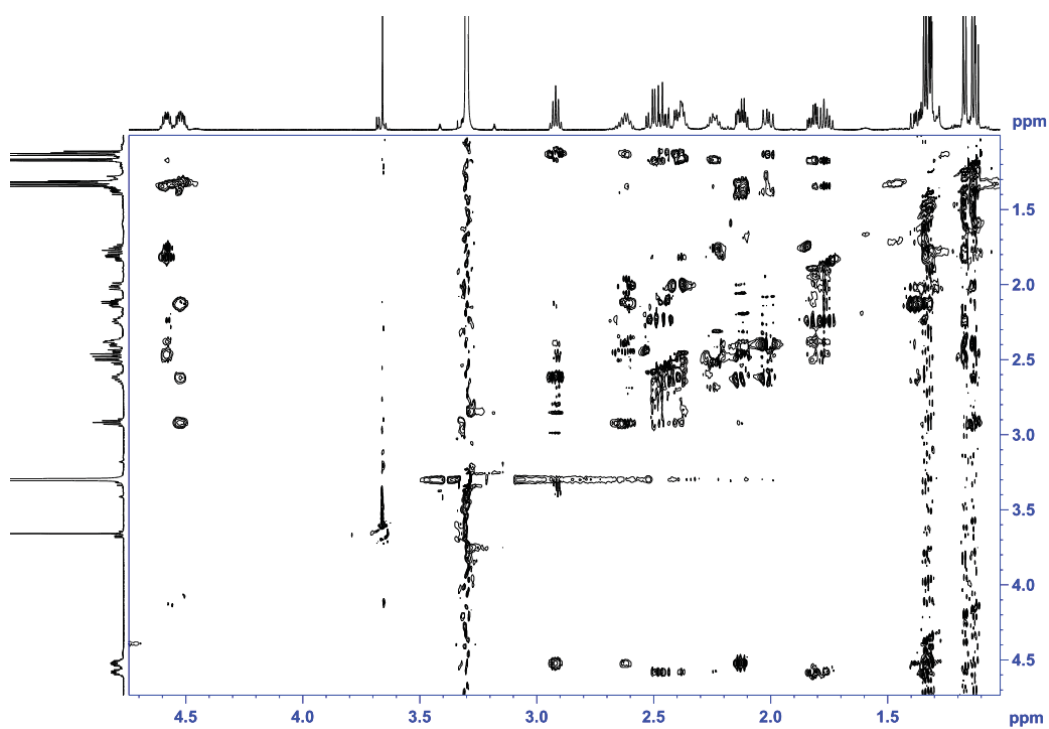
H,H COSY NMR Spectrum (600 MHz, MeOD) of compound **3**.



HSQC NMR Spectrum (600 MHz, MeOD) of compound **3**.



HMBC NMR Spectrum (600 MHz, MeOD) of compound 3.



NOESY NMR Spectrum (600 MHz, MeOD) of compound **3**.

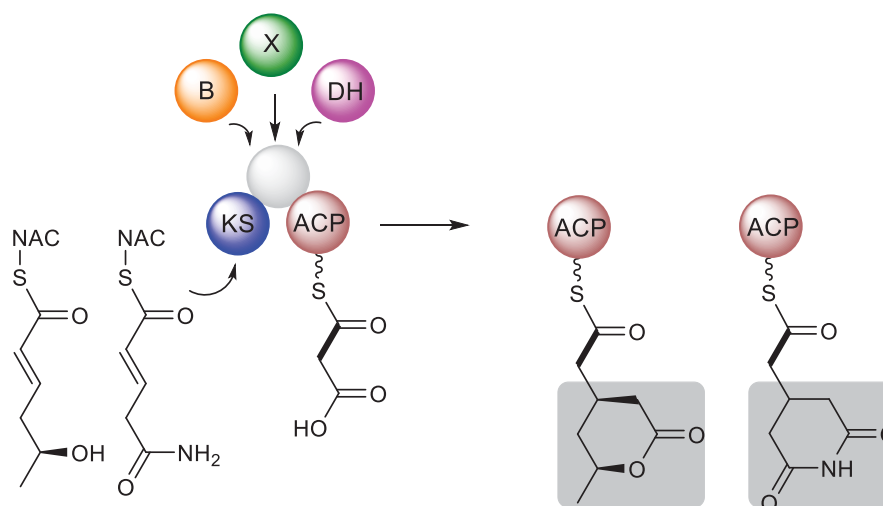
References

- [1] a) D. Heine, T. Bretschneider, S. Sundaram, C. Hertweck, *Angew. Chem. Int. Ed.* **2014**, *53*, 11645-11649; b) D. Heine, T. Bretschneider, S. Sundaram, C. Hertweck, *Angew. Chem.* **2014**, *126*, 11829-11833.
- [2] S. Bartlett, D. Böse, D. Ghorji, B. Mechsner, J. Pietruszka, *Synthesis* **2013**, *45*, 1106-1114.

3.4 Manuscript C

Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching

Srividhya Sundaram, Daniel Heine, Christian Hertweck: Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching, *Nature Chemical Biology* 11 (12), 949-951, 2015.



Some polyketide synthase pathways include branching modules that insert branched monomers into polyketide products. *In vitro* reconstitution using swapped domains shows that the mysterious branching (B) domain and the homologous X domain in these modules have structural rather than catalytic roles (from *Nature Chemical Biology*).

Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching

Srividhya Sundaram¹, Daniel Heine¹ & Christian Hertweck^{1,2*}

Biosynthesis of rhizoxin in *Burkholderia rhizoxinica* affords an unusual polyketide synthase module with ketosynthase and branching domains that install the δ -lactone, conferring antimetabolic activity. To investigate their functions in chain branching, we designed chimeric modules with structurally similar domains from a glutarimide-forming module and a dehydratase. Biochemical, kinetic and mutational analyses reveal a structural role of the accessory domains and multifarious catalytic actions of the ketosynthase.

Polyketides constitute a highly heterogeneous family of natural products, many of which are important leads for drug development. Given their high structural diversity, it is remarkable that the basis of their biosynthesis is the simple head-to-tail Claisen condensation of acyl and malonyl units, which exclusively yields linear carbon backbones¹. Polyketide chain assembly minimally requires a ketosynthase (KS) that mediates the attack of the malonyl-derived carbanion at the acyl thioester to form the C-C bond. In addition, an acyl carrier protein (ACP) is needed as an anchor for the growing polyketide chain; a malonyl transacylase-acyltransferase (AT) is needed to load the extender units onto the ACP; and the degree of β -keto group processing is controlled by the presence of ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains. Considering the uniform KS-mediated Claisen condensation biochemistry, it is surprising that a specialized polyketide synthase (PKS) module mediates a radically different reaction in rhizoxin biosynthesis. Rhizoxin (**1**; **Fig. 1a**) is a highly potent phytotoxin that is produced in a symbiosis of the rice seedling blight fungus *Rhizopus microsporus* and its bacterial endosymbiont, *B. rhizoxinica*^{2,3}. The bacteria assemble the scaffold of the macrolide, and the fungal host introduces an additional epoxide moiety to increase the phytotoxicity⁴. Notably, the rhizoxin congeners assembled by the rhizoxin (rhi) PKS⁵ represent some of the strongest antimetabolic agents known to date⁶. Binding to the β -tubulin subunit critically depends on an unusual chain branch^{7,8} that consists of a malonyl-derived C2 unit. Genetic and chemical analyses of the rhi PKS⁹ indicate that the rare pharmacophoric unit is introduced by a 'branching module' consisting of a KS domain, an enigmatic 'branching domain' (B) and an ACP¹⁰. The successful *in vitro* reconstitution of the entire branching module shows that the chain branch is introduced by the vinylogous addition of acetate to a KS-bound α,β -unsaturated thioester (**Fig. 1a**)¹¹. As a result, the branched intermediate is bound to both the KS and the ACP, and only through nucleophilic attack of the thioester at the KS can the side chain be cyclized and the polyketide chain be propagated. In the genuine rhi pathway, a δ -lactone is formed by cyclization of the δ -hydroxy-substituted intermediate. However, the branching module shows a degree of substrate tolerance and also accepts C3 extenders¹² as well as amino and carboxamide nucleophiles in lieu of the δ -hydroxy group, thus yielding lactam and glutarimide moieties, respectively¹³. This is a key finding because the glutarimide residue is known to be an important pharmacophoric group in many

antibiotics, such as cycloheximide¹⁴. In fact, genetic analysis of the biosynthesis of several glutarimide-containing polyketides—including 9-methylstreptimidone¹⁵, migrastatin¹⁶ and cycloheximide¹⁷—reveals that the PKS genes encode a branching module that is highly similar to the one found in the rhi PKS, consisting of an unknown (denoted as X) domain flanked by KS and ACP domains (**Fig. 1b**). Despite the elucidation of its crystal structure and numerous attempts to characterize the B domain biochemically¹¹, the role of the mysterious B or X domains has remained a riddle. Here we report the successful interrogation of the B and X domains from lactone- and glutarimide-forming PKS modules *in vitro*. Through engineered functional hybrid modules, kinetic analyses and mutational studies, we provide compelling evidence that the linker region and an adjacent domain with a double-hotdog fold are essential for the branching module but do not have catalytic roles for branching and cyclization. Thus, very unexpectedly, the branching KS domain alone mediates the entire complex reaction sequence.

To elucidate the exact role of the B and X domains in polyketide chain branching, we exchanged the B domain of the rhi branching module with the functional equivalent of a glutarimide-forming PKS, the X domain from the 9-methylstreptimidone PKS of *Streptomyces himastatinicus* (**Fig. 2a**)¹⁵. To test its ability to mediate glutarimide formation, we subjected the heterologously produced hybrid didomain (**Supplementary Results, Supplementary Note 1**) to an *in vitro* biotransformation experiment in the presence of phosphopantetheinylated ACP and the *trans*-acting AT. We used malonyl CoA and carboxamide-substituted *N*-acetylcysteamine (SNAC) thioester **2** (ref. 13), a mimic of the natural polyketide intermediate for glutarimide biosynthesis, as substrates. MALDI-TOF analysis of the ACP-bound product showed a new peak, indicating the expected mass of the glutarimide-substituted product (**Supplementary Fig. 7a**). Furthermore, extraction of the assay mix with methanol released the free acid **3** (**Fig. 2b**). Its structural identity was unequivocally confirmed by high-resolution MS (HRMS) comparing the retention time and mass to that of a synthetic reference¹³. This finding proved that the hybrid module is functional and enabled a direct comparison of the B and X domains. In addition to the natural role as a branching δ -lactone synthase, the KS-B didomain also mediated glutarimide formation in the presence of a suitable carboxamide nucleophile. With the working counterpart from a genuine glutarimide pathway at hand, we tested whether the KS-X didomain could also produce a δ -lactone when using the synthetic *S*-hydroxy-substituted SNAC thioester **4**, a substrate that mimics the polyketide intermediate produced upstream in the rhi pathway¹¹. MALDI-TOF measurements indicated that the ACP-bound lactone was formed (**Supplementary Fig. 7b**), and we confirmed the identity of the hydrolyzed product (**5**) by HRMS analyses using an authentic reference (**Fig. 2c**). Thus, both KS-B and KS-X catalyze the Michael addition-cyclization sequence to yield δ -lactone- and glutarimide-substituted polyketides.

¹Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Jena, Germany. ²Chair of Natural Product Chemistry, Friedrich Schiller University, Jena, Germany. *e-mail: christian.hertweck@leibniz-hki.de

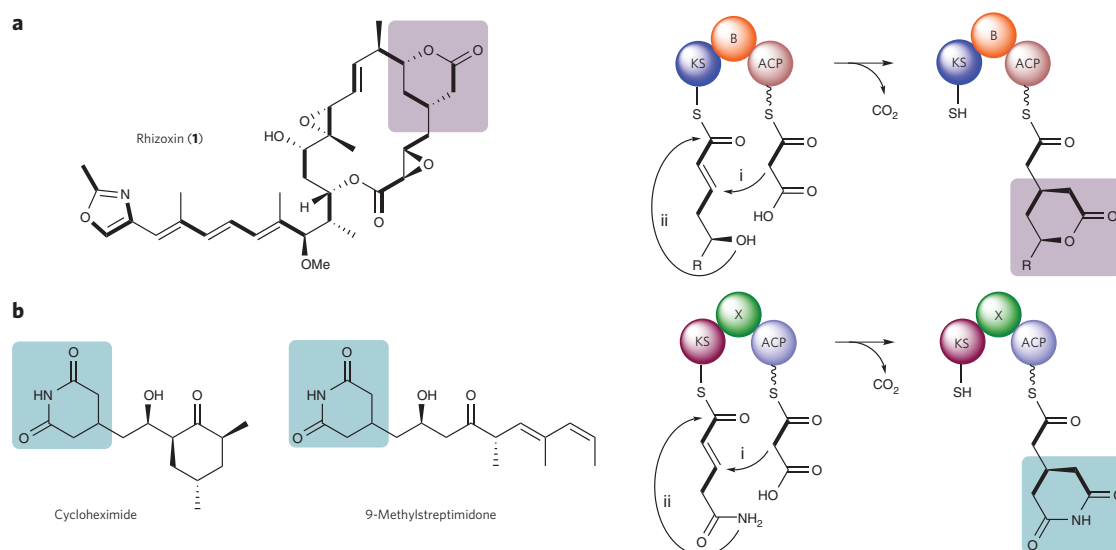


Figure 1 | Polyketide chain branches introduced by polyketide synthase modules. (a) Structure of rhizoxin (**1**) and a model for vinylogous chain branching by the PKS branching module for lactone formation. (b) Structures of cycloheximide and 9-methylstreptimidone and model for chain branching for glutarimide formation. Lactone and glutarimide moieties are highlighted. Michael addition (i) and lactonization and glutarimide formation (ii) are shown.

In addition to the qualitative analysis, which demonstrated the functionally equivalent substrate tolerance of the KS-B and KS-X didomains, we performed kinetic analyses to gain further insight into a potential prevalence for a particular branching reaction. To determine the kinetic determinants, we performed the *in vitro* assay with either KS-B or KS-X, with increasing concentrations of **2** or **4**. In the assay for δ -lactone formation, we obtained data indicating fairly similar substrate affinities of both constructs. Likewise, the rate of lactone formation proved to be similar (Table 1)¹³. Using **2** as the substrate, the K_M and V_{max} values were, again, almost identical

for KS-B and KS-X. Both branching didomains showed a three- to fourfold reduced reaction rate and a threefold lower affinity to **2** compared to **4** (Table 1 and Supplementary Fig. 8a). We thus concluded that the B or X domains do not have a major role in the selection and turnover of the different substrates, irrespective of the biosynthetic context (lactone- or glutarimide-producing PKS).

To interrogate the impact of the B and X domains, we generated another chimera in which the branching domain is substituted with a domain that has a similar fold. As the double-hotdog fold of the B domain is a hallmark of DH domains, we explored the catalytic

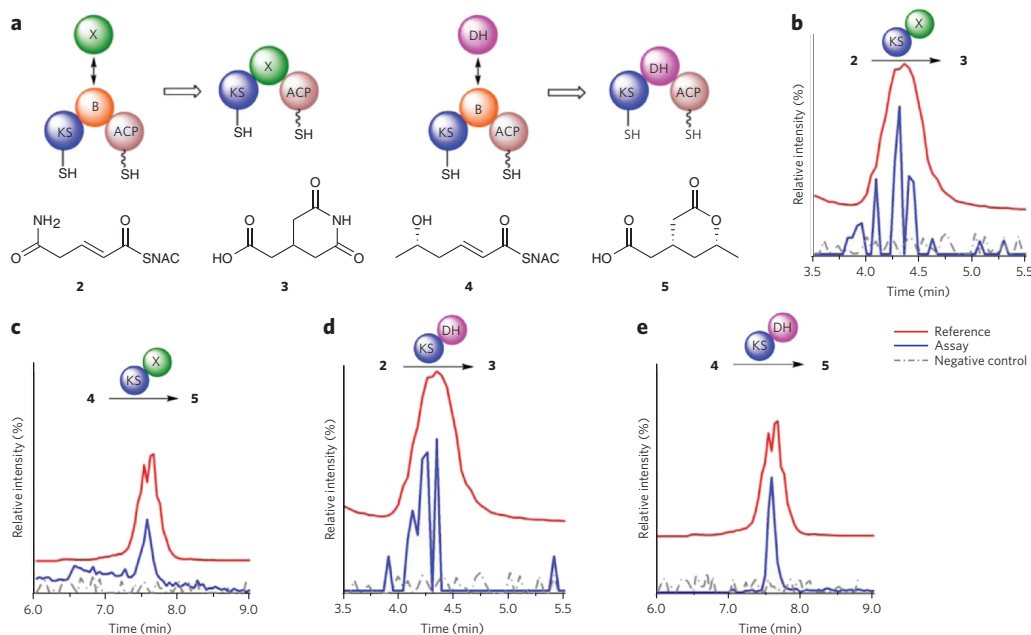


Figure 2 | *In vitro* lactone and glutarimide formation by chimeric modules. (a) Chimeric modules and the structures of substrates and products. (b, c) Biotransformation of amide-substituted (**2**; b) and hydroxy-substituted SNAC thioester (**4**; c) using the KS-X didomain. (d, e) Biotransformation of amide-substituted (**2**; d) and hydroxy-substituted SNAC thioester (**4**; e) using KS-DH didomain. HRMS analysis of the assay mix revealed glutarimide (**3**) and lactone (**5**) formation. Negative control refers to the branching reaction without **2** or **4**. All tests were performed in duplicate.

Table 1 | Kinetic determinants of wild-type and chimeric proteins for the production of 5 and 3

	KS-B		KS-X		KS-DH	
	K_M	V_{max}	K_M	V_{max}	K_M	V_{max}
5	1,348 ± 315	20.9 ± 2.4	1,871 ± 728	18.4 ± 5.7	1,428 ± 461	15.2 ± 3.7
3	4,082 ± 519	4.7 ± 0.4	3,707 ± 255	4.4 ± 1.1	4,878 ± 724	5.4 ± 0.8

Data are from three independent experiments and represent mean ± s.d.

abilities of an engineered KS-DH didomain (Fig. 2a). Specifically, we cloned the gene fragment coding for a DH domain from the downstream module 10 of the rhi PKS and fused it with the gene coding for the branching KS. In addition to the KS, we found that the KS-to-B linker domain was also crucial for protein solubility (Supplementary Note 2). We repeated the biotransformation experiments with the KS-DH hybrid and found that both the glutarimide (Fig. 2d) and the δ -lactone (Fig. 2e) were formed in the enzyme assay (Supplementary Fig. 7c,d). Furthermore, in all of the kinetic studies using KS-DH and 2 or 4, we determined K_M and V_{max} values equivalent to those obtained for KS-B and KS-X (Table 1 and Supplementary Fig. 8b,c). Thus, in the branching module, a regular DH domain can replace the B and X domains without loss of function.

Our findings could be rationalized in two ways: (i) the B and X domains support the branching-cyclization sequence by a mechanism that can also be catalyzed by a typical DH domain or (ii) the B, X and DH domains are catalytically inactive and have a solely structural role. To rule out a potential catalytic function of the additional domain in the genuine and hybrid branching modules, we performed mutational analyses. It is well known that DH catalysis depends on a highly conserved catalytic histidine-aspartate dyad. A remarkable feature of the B domain is the presence of Tyr3697 in place of the active site histidine normally present in DH domains (Supplementary Fig. 9a). This mutation clearly ruled out the involvement of a histidine in the chain branching-cyclization sequence. The aspartate residue is conserved among the B (Asp3876), X (Asp2415) and DH (Asp2488) domains (Supplementary Fig. 9a-c). To exclude the possibility that the conserved aspartate residues in the dyads of the B, X and DH domains are essential for the module function, we replaced them with alanine to yield B^o, X^o and DH^o domains, respectively (Supplementary Fig. 9d). For all aspartate-to-alanine mutants (KS-B^o-ACP, KS-X^o-ACP and KS-DH^o-ACP), we still observed branching activity in the *in vitro* enzyme assay (Supplementary Fig. 10b,d). Because the point mutations at the active sites did not affect the enzyme activity, the dyad is clearly not involved in the branching reaction. Thus, the additional domain with the double-hotdog fold seems to have a structural role. To support this conclusion, we designed constructs in which the KS and the requisite linker region are fused to the ACP (Supplementary Fig. 9e). However, we could not detect any soluble protein (data not shown), again demonstrating the pivotal role of the B domain for structural stability.

Taken together, we provide multiple lines of evidence indicating that B, X and DH domains solely have a structural role in the native and chimeric branching modules. Given their structural similarity, it is conceivable that the B and X domains from branching modules have evolved from DH domains yet lost their catalytic function. In this context, it is of interest to note that inactive or intact but nonfunctional DH domains are sometimes found in PKS modules^{18–20}. Domains with a double-hotdog fold could be important as spacers between the KS and the ACP. It is also noteworthy that the KS-to-B linker region is crucial for the solubility of the KS-B didomain. However, our results unequivocally demonstrate that the KS domain alone is in charge of mediating the entire catalytic sequence, the stereoselective vinylogous addition and the subsequent heterocyclization, which leads to lactone or glutarimide rings. This finding has profound implications for elucidating the branching process. The branching KS domains that catalyze Michael additions and heterocyclizations represent a functionally

unique family of ketosynthases that clearly differ from the canonical KSs of modular assembly lines that exclusively yield linear chains. KS domains promoting Michael additions complement the related enzymes that catalyze terpenoid-like alkylations of β -carbonyl groups²¹, which are also found in some type I PKS systems. This functional divergence of the KS domains is notable from an evolutionary perspective. In the case of the rhi PKS, the evolutionary advantage is clear, because the chain branch is pivotal to the antimetabolic activity of the macrolide. Likewise, the glutarimide residues represent pharmacophoric groups in various polyketide antibiotics^{14,22}. Thus, the branching KS is an important addition to the currently available ketosynthase-like biocatalysts^{23–25}. As the KS-mediated branching reaction enhances the structural diversity of polyketides to yield potent pharmacophoric groups, insights gained from our functional analysis are also valuable from a synthetic biology perspective.

Received 7 July 2015; accepted 28 August 2015; published online 19 October 2015

Methods

Methods and any associated references are available in the online version of the paper.

References

- Hertweck, C. *Angew. Chem. Int. Ed.* **48**, 4688–4716 (2009).
- Partida-Martinez, L.P. & Hertweck, C. *Nature* **437**, 884–888 (2005).
- Lackner, G., Partida-Martinez, L.P. & Hertweck, C. *Trends Microbiol.* **17**, 570–576 (2009).
- Scherlach, K., Busch, B., Lackner, G., Paszkowski, U. & Hertweck, C. *Angew. Chem. Int. Ed.* **51**, 9615–9618 (2012).
- Partida-Martinez, L.P. & Hertweck, C. *ChemBioChem* **8**, 41–45 (2007).
- Scherlach, K., Partida-Martinez, L.P., Dahse, H.-M. & Hertweck, C. *J. Am. Chem. Soc.* **128**, 11529–11536 (2006).
- Schmitt, I. *et al. ISME J.* **2**, 632–641 (2008).
- Kusebauch, B., Scherlach, K., Kirchner, H., Dahse, H.M. & Hertweck, C. *ChemMedChem* **6**, 1998–2001 (2011).
- Nguyen, T. *et al. Nat. Biotechnol.* **26**, 225–233 (2008).
- Kusebauch, B., Busch, B., Scherlach, K., Roth, M. & Hertweck, C. *Angew. Chem. Int. Ed.* **48**, 5001–5004 (2009).
- Bretschneider, T. *et al. Nature* **502**, 124–128 (2013).
- Heine, D., Sundaram, S., Bretschneider, T. & Hertweck, C. *Chem. Commun. (Camb.)* **51**, 9872–9875 (2015).
- Heine, D., Bretschneider, T., Sundaram, S. & Hertweck, C. *Angew. Chem. Int. Ed.* **53**, 11645–11649 (2014).
- Rajski, S.R. & Shen, B. *ChemBioChem* **11**, 1951–1954 (2010).
- Wang, B. *et al. Org. Lett.* **15**, 1278–1281 (2013).
- Lim, S.K. *et al. J. Biol. Chem.* **284**, 29746–29756 (2009).
- Yin, M. *et al. Org. Lett.* **16**, 3072–3075 (2014).
- Chen, S. *et al. Chem. Biol.* **10**, 1065–1076 (2003).
- Hothersall, J. *et al. J. Biol. Chem.* **282**, 15451–15461 (2007).
- Xu, Z. *et al. ChemBioChem* **15**, 1274–1279 (2014).
- Calderone, C.T. *Nat. Prod. Rep.* **25**, 845–853 (2008).
- Schneider-Poetsch, T. *et al. Nat. Chem. Biol.* **6**, 209–217 (2010).
- Bretschneider, T. *et al. Nat. Chem. Biol.* **8**, 154–161 (2012).
- Fuchs, S.W. *et al. Angew. Chem. Int. Ed.* **52**, 4108–4112 (2013).
- Zocher, G. *et al. Chem. Sci.* doi:10.1039/C5SC02488A (6 August 2015).

Acknowledgments

We thank A. Perner for MS analyses and M. Poetsch for MALDI measurements. We are grateful for financial support by the International Leibniz Research School (to S.S.) and the Studienstiftung des Deutschen Volkes (to D.H.).

Author contributions

S.S. and C.H. designed experiments; S.S. performed genetic and biochemical experiments and analyzed data; D.H. synthesized substrates and reference compounds; and S.S. and C.H. wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Correspondence and requests for materials should be addressed to C.H.



ONLINE METHODS

General. Chemicals were purchased from Sigma-Aldrich and Roth. MALDI measurements were performed using an UltrafleXtreme MALDI TOF/TOF (Bruker). HRMS measurements were performed using an Exactive Orbitrap High-Performance Benchtop HRMS with an electrospray ion source and an Accela HPLC system (Thermo Scientific). All solvents for analytical HPLC measurements were obtained commercially at least in gradient grade and were filtered before use. To avoid microbial growth, 0.1% formic acid was added to the water. CD measurements were performed using a J815 spectropolarimeter (Jasco). Site-directed mutagenesis was performed using a Quikchange II XL kit (Stratagene). The identity of all recombinant proteins was confirmed using SDS-PAGE.

Alignment and threading of KS-X. The structure of KS-X was predicted by aligning amino acids corresponding to residues 1614–2525 (GenBank [CCC21123.1](#)) of KS-B (Protein Data Bank code [4KC5](#)). The prediction was done using the Phyre2 server²⁶, and the structure was visualized in Pymol (Version 1.5.0.4; 2010).

Cloning experiments. To facilitate cloning of the genes for X and DH domains to the C-terminal of KS, two separate constructs were generated: (i) the region corresponding to KS spanning residues 3152–3661 (GenBank [CBW75249.1](#)) was cloned in pHis8-3 by introducing a C-terminal KpnI site, using the primer pair KS-BamHI-fw (5'-GGA TCC TCC GGT GAG CGC GTC GAG-3') and KS-KpnI-rv (5'-GGG GTA CCT TCC TCG ACG ATC ATA TGG-3'), resulting in pSS31. (ii) For the second construct, the region spanning residues 3152–3821 (GenBank [CBW75249.1](#)) that corresponds to KS and the KS-B linker region was cloned with the insertion of the KpnI site using primers KS-BamHI-fw (5'-GGA TCC TCC GGT GAG CGC GTC GAG-3') and KS-KpnI-rv (5'-GGG GTA CCC GAC GAC TCG TAA CCT TCC TC-3'), resulting in pSS33. The X-encoding gene from region 2053–2525 (GenBank [CCC21123.1](#)) and the DH-encoding gene corresponding to the region 2286–2576 (GenBank [CBW75249.1](#)) were amplified from the genomic DNA of *Streptomyces himastatinicus* and *Burkholderia rhizoxinica*, respectively, using the following primer pairs: X-KpnI-fw (5'-GGG GTA CCC ACA TGA TTC TGG AGG AGT ACG CGG GA-3'), X-EcoRI-rv (5'-CGG AAT TCC TAC AGT TCG TCC ACG ACG AAG G-3'); DH-KpnI-fw (5'-GGG GTA CCC TGC ACC CGC TGG TGC ATC-3') and DH-EcoRI-rv (5'-GAA TTC CTA GGG TTC AAC GCC AGT C-3'). The corresponding PCR products were cloned in pJET 1.2 (Thermo Scientific) for sequence confirmation and subcloned into pHis 8-3 to generate pSS32, pSS34 and pSS42 for KS-DH^o (without linker), KS-DH (with linker) and KS-X, respectively. All of the recombinant plasmids were heterologously expressed in *E. coli* BL21 (New England BioLabs).

Protein production and purification. The procedures for the production and purification of ACP (pTB41), KS-B (pTB57) and RhiG (pTB39) were as described previously¹¹. *E. coli* BL21 transformants containing pSS32, pSS34 and pSS42 were cultivated in LB medium at 37 °C. After A_{600nm} reached ~0.5, the cells were induced with 0.1 mM IPTG at 15 °C for 18 h. The cell pellet was dissolved (10 ml buffer per g pellet) in 20 mM Tris buffer, pH 7.5, containing 0.2 M NaCl and 50 mM imidazole. After disrupting the cells by sonication after four 3-min cycles at 20% power (Sonotrode), the cells were centrifuged at 9,000g at 4 °C for 30 min. The lysed supernatant was then loaded onto a Ni-NTA column (Macherey Nagel) connected to an FPLC machine. Proteins were eluted using the same buffer with 0.5 M imidazole. The protein was diluted twice with 20 mM Tris (pH 7.5) and purified further by an ion-exchange column (5 ml HiTrap Q HP, GE Healthcare). Proteins were eluted over 30 column volumes of 20 mM Tris (pH 7.5) and 1 M NaCl, and the fractions containing target proteins were concentrated using Amicon columns (MWCO 5 and 10 kDa; Millipore). All of the proteins were stored at 4 °C for *in vitro* assays. Long-term storage of the proteins was possible at –20 °C supplemented with 50% glycerol.

***In vitro* reconstitution of the branching reaction.** The *in vitro* assay for the reconstitution of the branching reaction was performed as described

previously¹¹ with minor modifications. Briefly, in a 40- μ l reaction, 167 μ M ACP; 3 μ M KS-B, KS-X or KS-DH; 0.2 μ M RhiG; 750 μ M malonyl-CoA; and 1,000 μ M of SNAC thioesters were incubated in 20 mM Tris buffer (pH 7.0) at 23 °C with shaking at 400 r.p.m. After 2 h, 2 μ l of the sample was mixed with 2 μ l of 100 mM 2',5'-dihydroxyacetophenone (containing 2.5 mM diammonium hydrogen citrate) and spotted onto an anchor chip 800/384 T F target (Bruker) to analyze the product. The MALDI-TOF was operated in positive linear mode. Data acquisition was performed using flexControl 3.3 in the range of 5–35 kDa with 1,000-Hz laser frequency, 60% laser power and 500 shots at each spot. Analysis of the data was executed in flexanalysis 3.3. For mass measurements, the enzyme reaction was stopped by adding an equal volume of methanol. After a brief spin to precipitate the proteins, the supernatant was directly analyzed by HRMS (Exactive). HPLC conditions: C18 column (Betasil C18 3 μ m, 150 \times 2.1 mm) and gradient elution (MeCN/0.1% (v/v) HCOOH (H₂O) 5/95 for 1 min, gradient up to 98/2 in 15 min and 98/2 for 3 min; flow rate 0.2 ml min⁻¹; injection volume 3–5 μ l).

Kinetic studies. To determine the kinetic parameters for the KS-X and KS-DH didomains, we varied the concentration of the substrate **4** in the range of 108 μ M to 2,164 μ M, and the reaction was performed for 75 min. Higher concentrations ranging from 435 μ M to 6,519 μ M were used for **2**, and the reaction was performed for 250 min. Samples were taken at regular intervals, and the peak height/area of the product attached to the ACP was analyzed by MALDI-TOF. The mean kinetic parameters were calculated from three independent experiments using a Lineweaver-Burk plot. The substrates and the reference compounds were synthesized as below.

S-(2-Acetamidoethyl) (E)-5-amino-5-oxopent-2-enethioate (2). Compound **2** was synthesized according to a synthesis procedure reported before¹³. All analytical data are in good agreement with the reported values. ¹H NMR (600 MHz, CDCl₃): δ 6.96 (dt, ³J₁ = 15.5 Hz, ³J₂ = 7.2 Hz, 1H), 6.28 (dt, ³J = 15.5 Hz, ⁴J = 1.5 Hz, 1H), 3.34 (t, ³J = 6.7 Hz, 2H), 3.15 (dd, ³J = 7.2 Hz, ⁴J = 1.5 Hz, 2H), 3.07 (t, ³J = 6.7 Hz, 2H), 1.91 p.p.m. (s, 3H); HRMS (*m/z*): [M+H]⁺ calculated for C₉H₁₅O₃N₂S, 231.0798; found, 231.0793.

2-(2,6-Dioxopiperidin-4-yl) acetic acid (3). Compound **3** was obtained from Sigma-Aldrich. Analytical data are in full agreement with the literature data²⁷. ¹H NMR (600 MHz, DMSO-d₆): δ 12.3 (s, 1H), 10.7 (s, 1H), 2.55 (d, ³J = 3.9 Hz, 1H) 2.53 (d, ³J = 3.9 Hz, 1H) 2.36–2.44, (m, 1H), 2.26–2.36 (m, 4H), HRMS (*m/z*): [M+H]⁺ calculated for C₇H₁₀O₄N, 172.0610; found, 172.0610.

S-(2-Acetamidoethyl) (S,E)-5-hydroxyhex-2-enethioate (4). Compound **4** was synthesized according to a procedure reported before¹³. All analytical data are in good agreement with the reported values¹³. ¹H NMR (600 MHz, CDCl₃): δ 6.92 (dt, ³J = 15.6 Hz, ⁴J = 7.5 Hz, 1H), 6.20 (dt, ³J = 15.6 Hz, ⁴J = 1.4 Hz, 1H), 3.97 (m, 1H), 3.45 (q, ³J = 6.3, 2H), 3.08 (t, ³J = 6.4 Hz, 2H), 2.36–2.33 (m, 2H), 1.97 (s, 3H), 1.23 p.p.m. (d, ³J = 6.3 Hz, 3H); HRMS (*m/z*): [M+H]⁺ calculated for C₁₀H₁₈O₃NS, 232.1002; found, 232.0998; [α]_D²⁵ = +8.66 (c = 1 mg ml⁻¹, MeOH).

2-((2R,4R)-2-Methyl-6-oxotetrahydro-2H-pyran-4-yl) acetic acid (5): Compound **5** was synthesized according to a known procedure¹¹. Analytical data are in good agreement with the literature data. ¹H NMR (300 MHz, MeOD): δ 4.59–4.68 (m, 1H), 4.43–4.56 (m, 1H), 2.75 (ddd, ³J₁ = 17.5 Hz, ³J₂ = 5.9 Hz, ⁴J = 1.9 Hz, 1H), 2.61 (dd, ³J₁ = 15.5 Hz; ³J₂ = 5 Hz, 1H), 2.29–2.50 (m, 8H), 2.17 (dd, ³J₁ = 17.5 Hz, ³J₂ = 10.2 Hz, 1H), 2.00–2.09 (m, 1H), 1.73–1.87 (m, 2H), 1.34 p.p.m. (d, ³J = 6.3 Hz, 6H); HRMS (*m/z*): [M+H]⁺ calculated for C₇H₁₃O₄, 173.0808; found, 173.0810.

Secondary structure determination by CD. The secondary structures of KS-B, KS-X and KS-DH were determined at 23 °C. Wavelength scans were recorded for the far-UV region between 250 nm and 190 nm at a speed of 10 nm min⁻¹, a bandwidth of 1 nm and with accumulations of 3–5 spectra using a 1-mm quartz cuvette (Hellma). For all measurements, a concentration of 0.1 mg ml⁻¹ protein was used.

Mutational analysis of KS-X and KS-DH. Point mutations on X and DH were made using the following primers on plasmids pSS34 and pSS42 respectively: X-D2415A-fw (5'-CTC GCG GCC CTC GCT CAG GGC GCG-3'), X-D2415A-rv (5'-CGC GCC CTG AGC GAG GGC CGC GAG-3'); DH-D2488A-fw (5'-CCG CCG GGC GTA CTC GCT AGT GCG CTG CAT-3'), DH-D2488A-rv (5'-ATG CAG CGC ACT AGC GAG TAC GCC CGG CCG-3'). All mutations were confirmed by sequencing, and the production

of protein was carried out using the same procedure described above. For enzyme assays, the same concentrations as the wild type were used. CD (Jasco) measurements confirmed the correct folding of the mutated proteins (**Supplementary Fig. 10a,c**).

26. Kelley, L.A. & Sternberg, M.J. *Nat. Protoc.* **4**, 363–371 (2009).

27. Ji, X.-Y. *et al. Chem. Pharm. Bull. (Tokyo)* **58**, 1436–1441 (2010).

SUPPLEMENTARY INFORMATION

**Polyketide Synthase Chimeras Reveal
Key Role of Ketosynthase Domain in Chain Branching**

Srividhya Sundaram,¹ Daniel Heine,¹ Christian Hertweck^{1,2*}

¹Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Jena, Germany;

²Chair for Natural Product Chemistry, Friedrich Schiller University, Jena, Germany.

SUPPLEMENTARY RESULTS

Supplementary Notes

Supplementary Note 1

Rational design and functional analysis of hybrid branching modules

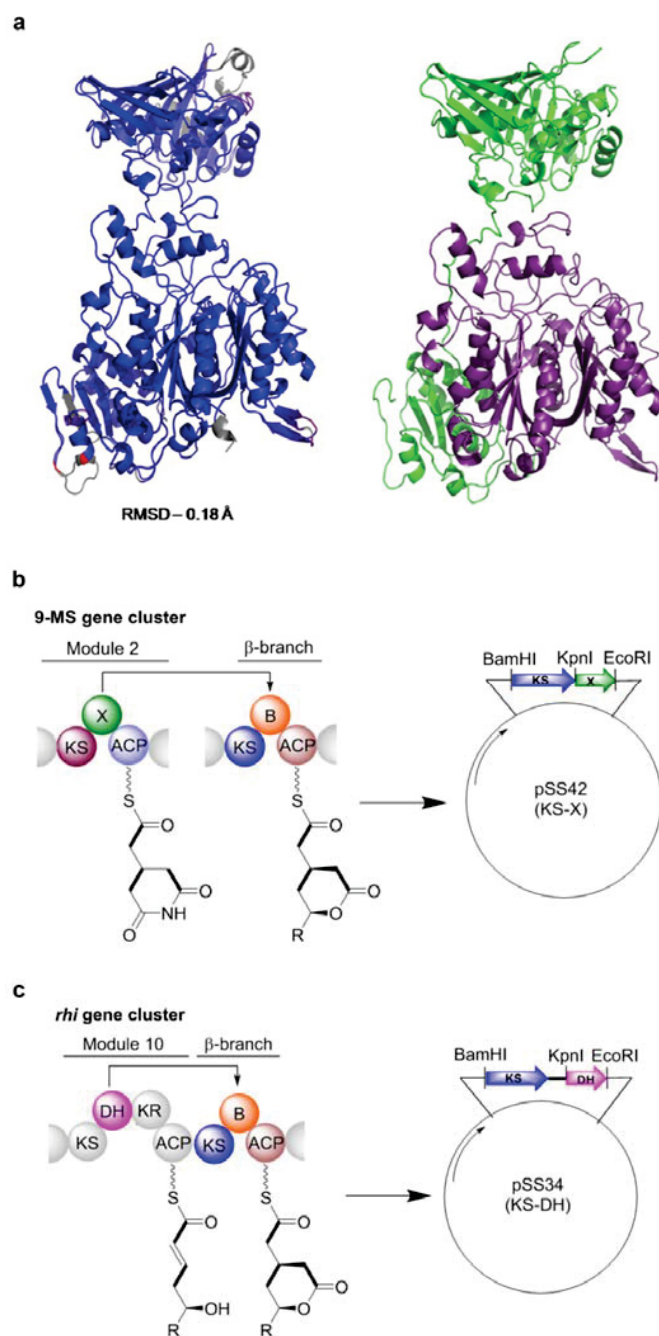
To elucidate the exact role of the enzyme domains (B and X) in polyketide chain branching we first sought to cross-complement the B domain of the functional *rhi* branching module with the functional equivalent of a glutarimide-forming PKS. On the basis of crystal data for KS-B (PDB ID – 4KC5) we performed a structural threading of proteins related to the B domain. Alignment and structure prediction showed that the X domain from a *trans*-AT PKS module SmdI¹ shares the highest similarity in folding (**Supplementary Fig. 1a**). Therefore, an intergenic exchange of B with a domain designated as unknown (X) from the 9-methylstreptimidone producer, *Streptomyces himastatinicus*,¹ was performed. Specifically, the gene for the X domain (including the linker region) was amplified from the genome of *Streptomyces himastatinicus* and fused to the C-terminus of the *rhi* KS gene (**Supplementary Fig. 1b and 2a**). The hybrid protein in which the linker region and B domain of the *rhi* PKS were replaced with the X domain from the 9-methylstreptimidone pathway, was heterologously produced in *E. coli* BL21 and purified by affinity and ion-exchange chromatography. Initially, the solubility of the hybrid KS-X didomain was substantially reduced compared to the native protein. However, after changing the induction temperature to 15 °C we obtained sufficient amounts of soluble protein. The successful construction of the desired chimera was proven by SDS-PAGE (**Supplementary Fig. 3a**), tryptic digestion and MALDI analyses (**Supplementary Fig. 4**), and the correct folding of the protein was verified by comparison of circular dichroism (CD) spectra of KS-B and KS-X (**Supplementary Fig. 3b**).

Supplementary Note 2

Construction of a functional branching module lacking B and X domains

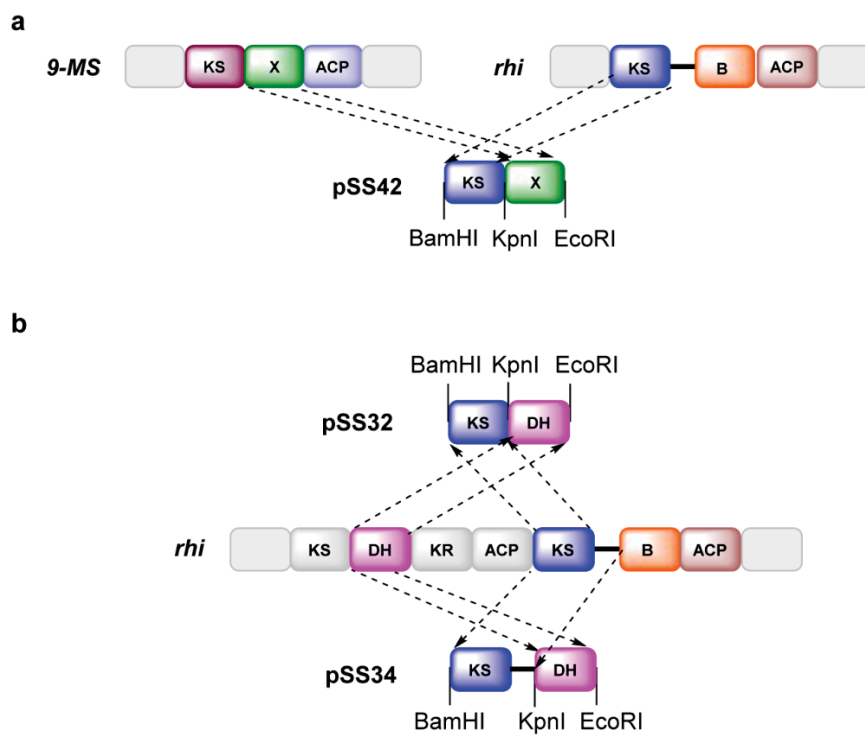
To construct a module without B or X, as a first step, the KS-to-B linker region and the B domain were replaced with the DH domain (**Supplementary Fig. 1c and 2b**). The resulting hybrid protein (KS-DH°), behaved differently compared to wild-type KS-B and showed a drastically reduced solubility, which resulted in inclusion bodies when produced at 28 °C (**Supplementary Fig. 5**). Since solubility could not be improved, we considered the possibility that the linker region is crucial for protein solubility. Therefore, we constructed another chimera that comprised the KS-to-B linker domain, and eventually succeeded in the production of soluble protein at reduced temperature (15 °C) (**Supplementary Fig. 3a**). The identity of the KS-DH didomain was confirmed by tryptic digest and MALDI-TOF analysis (**Supplementary Fig. 6**), and comparison of the CD spectra showed that its secondary structure is similar to KS-B and KS-X (**Supplementary Fig. 3b**).

Supplementary figures

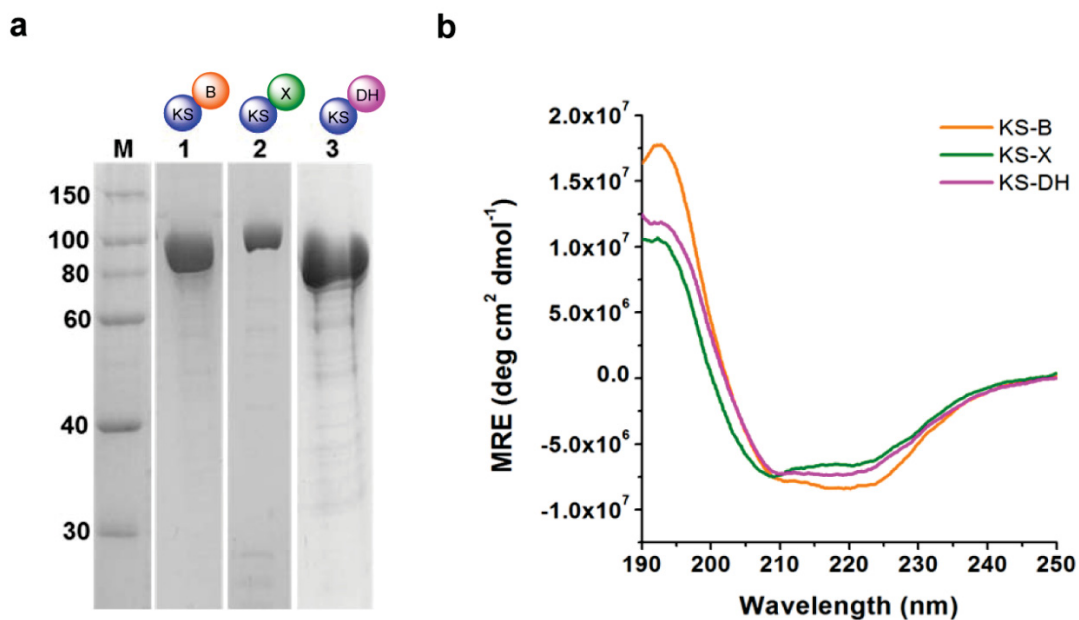


Supplementary Figure 1. Modeling of KS-X and chimera constructs. (a) Modeled KS-X didomain from module 2 (*Sdm1* gene) of 9-methylstreptimidone cluster showing high similarity to KS-B (left). Modeled KS-X with color codes for the domains is also shown (right). **(b)** Cloning strategy for KS-X chimera encoded on the plasmid pSS42. The gene

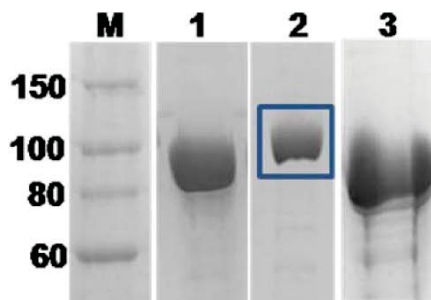
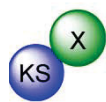
corresponding to the unknown X domain from 9-methyl streptimidone (9-MS) cluster is cloned to the C-terminus of KS domain of rhizoxin (*rhi*) cluster to generate pSS42. (c) Cloning strategy for KS-DH chimera encoded on the plasmid pSS34. The gene for DH domain from module 10 of *rhi* gene cluster is fused to the C-terminus of KS and linker domain using KpnI and EcoRI restriction sites. Black filled line refers to the KS-B linker region.



Supplementary Figure 2. Cloning strategy of the chimeras. (a) Cloning strategy for generation of chimeras using gene fragments coding for domains of the rhizoxin (*rhi*) and 9-methyl streptimidone (9-MS) PKSs. (b) Gene fragments corresponding to DH domain of the rhizoxin (*rhi*) was exchanged for B domain. Black filled line refers to the KS-B linker region.



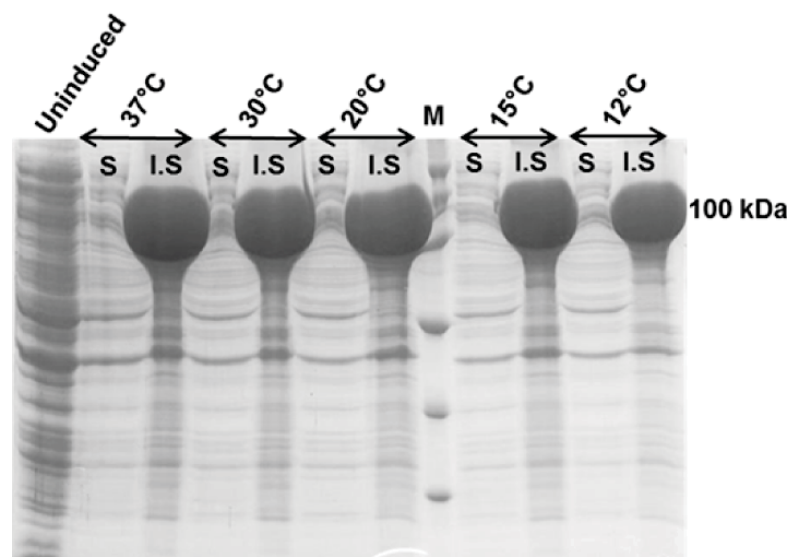
Supplementary Figure 3. Protein production with the hybrid didomains. (a) SDS-PAGE (12%) of purified KS-B (1), KS-X (2), KS-DH (3). The molecular weights of the proteins are 104.7, 100 and 99 kDa respectively. M – protein marker. (b) Secondary structure fingerprint of the chimeras by far-UV CD.



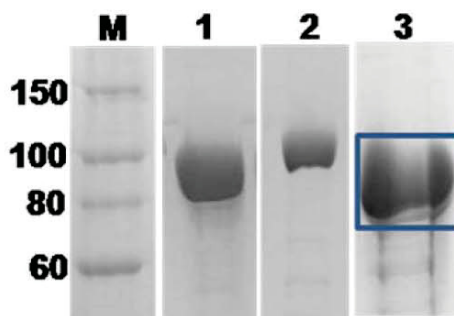
Protein: His-KS-X										Peak threshold: 0.0
Intensity coverage: 80.1 % (64114 cnts)		Sequence coverage MS: 45.8 %		Sequence coverage MS/MS: 8.6 %		pI: 5.7		MW (kDa): 104.0		
10	20	30	40	50	60	70	80	90		
MKHHHHHHH	GGLVPRGSHG	GSSGERVEDN	ELANYIAVIG	LGGYYPGADS	IDELWQNLAN	GVDCMSDFPA	DRWDHSKIYY	KNRKVLGKTT		
100	110	120	130	140	150	160	170	180		
CINGSFIKDV	DKFDYSYFKM	PKVYADHNSP	EVRLFQVAV	HTFEDAGYSK	ETLLSRYNGD	VGVLGTMNSN	DYHYGFESN	VFRGSMASGS		
190	200	210	220	230	240	250	260	270		
GMATIPMTVS	YFYGLTGPSL	FIDTMCSSSS	TCIHTACQML	KHDETKMFLA	GGLNLMYHPY	TTVNTSQGNF	TSITSESVNS	YGVGADGTVI		
280	290	300	310	320	330	340	350	360		
GEGIGAVLLK	RLDRAIADR	QIYGVKGS	MTNAGERNGF	NVPNPDQLT	AIRQAMDQAK	VHPSSISYIE	GHSGGTLGD	PIEVLGLNNA		
370	380	390	400	410	420	430	440	450		
FRWATDDRQF	CYLGSIKSN	GHLAASGIA	GLTKTLQFK	HKQIAPSIHS	SQLNQDIDFA	DTPFVVPQQL	IEWQPERII	NGRKQVFPFR		
460	470	480	490	500	510	520	530	540		
AGLTSIAAGG	MNAHMIVEEG	THMILEEYAG	RSQDAHDVDR	GRDELVFVSA	HTDAALDTSL	ARFRAYLSEA	DESDLPSIAF	TLRVGKNELP		
550	560	570	580	590	600	610	620	630		
RRWAFFASDT	RGAIEAIDRY	LAGDRDLTTA	LSSDDPRAVA	ARDLGTWVAG	GKSVDWARLG	DPRRLRVSL	PAYPFDRVSC	WVTPPEGAPS		
640	650	660	670	680	690	700	710	720		
VLAPLALRDK	LHPFLGRNAS	DVDGLRYILD	VHLDLDDY	YQGRERNIV	PAYALDVALA	TARISGLAAD	PATARVRDLR	VLDPVENTAT		
730	740	750	760	770	780	790	800	810		
ERLITLDTT	HREAARGAVF	AEDASGTRTR	VVEFEVGADG	TPGPANPRAS	LRFDVVRTLD	QPQTLAELAD	GGLDYEPLNC	GVDSAAWLRD		
820	830	840	850	860	870	880	890	900		
GRLVLTLTAP	EVQRDHTKPI	WTLAPHVLA	LDQGAHLAAK	SAGLPHWERQ	TWRHTADVRL	HTGAPDVAYA	VLDLAAEGGA	LACTITLLGP		
910	920	930	940	950	960					
GGEVAVELTG	VRYGDHDTPG	QDAPPQGVPR	DATPRHDTPR	DGGIVSFVVD	EL					

Protein: RhiE-8b										Peak threshold: 0.0
Intensity coverage: 0.4 % (339 cnts)		Sequence coverage MS: 9.9 %		Sequence coverage MS/MS: 0.0 %		pI: 5.2		MW (kDa): 35.0		
10	20	30	40	50	60	70	80	90		
IVNPLMFKNK	LHPLVAKNCS	TPQGAIFRT	DFVEDELDY	VYSGRGRRL	SAFNFADVAL	AMPALASRFD	GRTLSVSCAF	EHYIADWTTV		
100	110	120	130	140	150	160	170	180		
TGLEVRLFEI	DSEQLELEFD	FRRSGEQPTH	LGFVINPLT	SDEPPLPQQM	LDDARELLNR	QALQAGRQLS	AAEVSORLAQ	ACYDFAPYLD		
190	200	210	220	230	240	250	260	270		
HDGELTIGRS	GLVLRGRPPV	NRHNHYADNV	QLSPYLATTI	DKALYLLLDE	LGLPQGRVIV	RNIERLCCYH	TPAGGFSVVL	SGIGLNDNEL		
280	290	300	310	320						
SLSLVLDER	EQICVKLDKV	SLYLKQEVA	SVDRKHSLLT	GT						

Supplementary Figure 4. KS-X evaluation by MALDI fingerprinting. Tryptic digestion and MALDI analysis of the marked band corresponding to histidine-tagged KS-X protein. Peptides from KS and X domain were detected. Peptides corresponding to the B domain were not present. Lanes: 1 – KS-B; 2 – KS-X; 3 – KS-DH; M – protein marker.



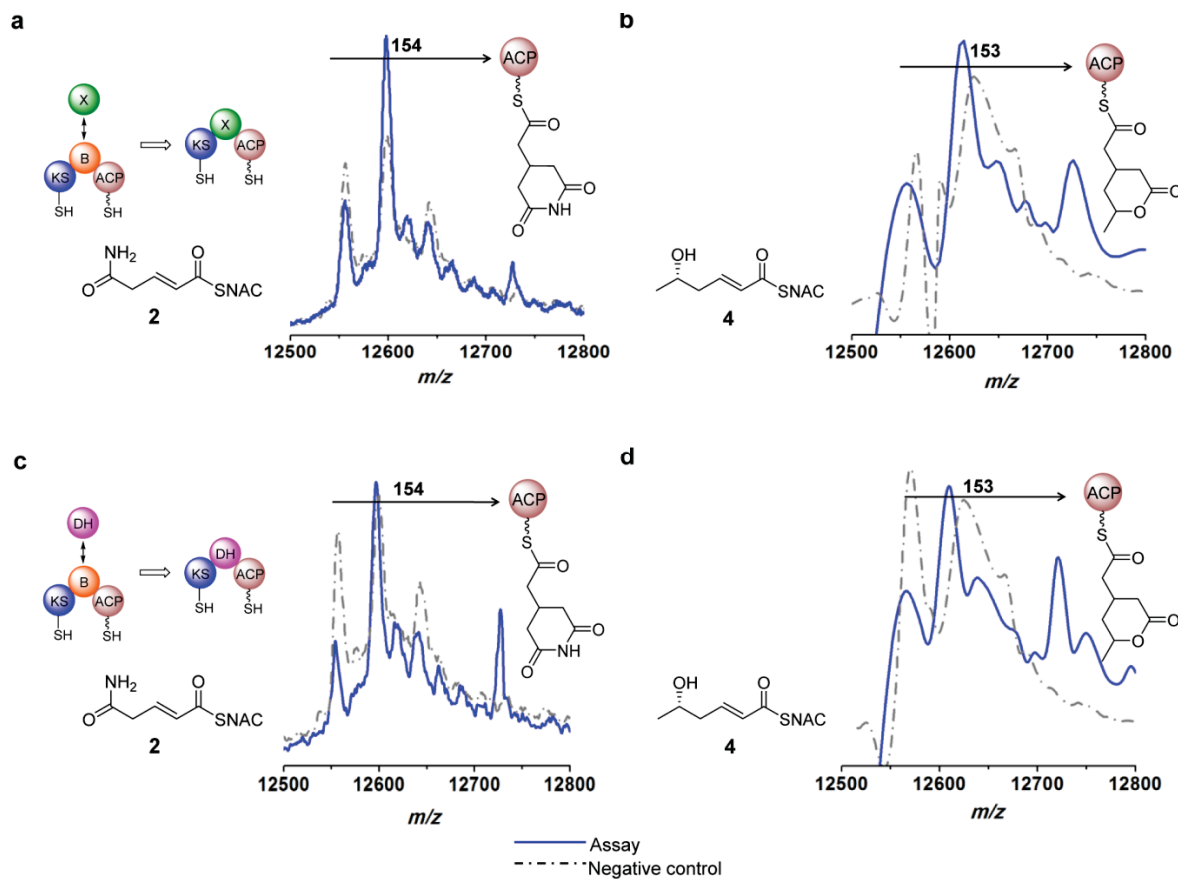
Supplementary Figure 5. SDS-PAGE of KS-DH^o. In the absence of KS-to-B linker domain, soluble protein could not be obtained even at an induction temperature of 12 °C. S – soluble; I.S – insoluble.



Protein:		His-KS-DH		Peak threshold:		0.0		
Intensity coverage:		68.4 % (56394 cnts)		Sequence coverage MS:		40.8 %		
		Sequence coverage MS/MS:		9.2 %		pl: 6.1		
		MW (kDa):		102.3				
10	20	30	40	50	60	70	80	90
MKHHHHHHH	GGLVPRGSHG	GSSGERVEDN	ELANYIAVIG	LGGYYPGADS	IDELWQNLAN	GVDCHSDFFA	DRWDHSKIYY	KMRKVLGKTT
100	110	120	130	140	150	160	170	180
CINGSFIKDV	DKFDYSYFKM	PKVYADHMSK	EVRLFLQVAV	HTFEDAGYSK	ETLLSRVNGD	VGVLGTMEN	DYHYGFESN	VFRGSMASGS
190	200	210	220	230	240	250	260	270
GHATIPMTVS	YFYGLTGPSL	FIDTMCSSSS	TCIHTACQML	KHDETKMFLA	GGLNLHYHPY	TTVNTSQGNF	TSITSESVNS	YGVGADGTVI
280	290	300	310	320	330	340	350	360
GEGIGAVLLK	RLDRAIADR	QIYGVIRGSA	MTNAGERNGF	NVPNPDQLT	AIRQAMDQAK	VHPSSISYIE	GHGSGTRLGD	PIEVLGLNNA
370	380	390	400	410	420	430	440	450
FRWATDDKQF	CYLGSIKSN	GHLAASGIA	GLTKTLQFK	HKQIAPSIHS	SQLNQDIDFA	DTPFVVPQQL	IENRQPERII	NGRKQVPRR
460	470	480	490	500	510	520	530	540
AGLTSIAAGG	MNAHMIVEEY	PEPADSAGQI	SEDQLVVFVS	VHKLALLAQ	LTSFRDWLAS	SEAPLAQIAY	TLOVGKNNLR	NRLAIRCTR
550	560	570	580	590	600	610	620	630
QALSRLNAC	IDGHYQSSAD	SKIFYRFQES	DAVQPLESDL	NDPLAPLLTQ	WLNQDSQVDW	ASLYAQPPVR	ISLPAYRFEK	TRCWYTEEGY
640	650	660	670	680	690	700	710	720
ESSGTLHPLV	HNISTLATQ	RFVSRFSGRE	TFLADHRIGD	QLIFFGAAYL	ELAREAIVRA	SELPPQQVTL	QNVVLRPLA	FAVGGDRDV
730	740	750	760	770	780	790	800	810
ELYCELAVNG	NDSVSFRIYS	LDGGTQQLHA	QGDAQLAAVA	APLPQVDLA	QWAARCNVSE	RSHDVLYHDF	DTMGVHYGVT	FRAAGQIRIG
820	830	840	850	860	870	880	890	900
RDRVIGRLIH	EALDPQLEGV	LLPPGVLDSA	LHIAASLESR	RQGAEPMLP	FALESRLYG	SVEQALRAGP	LTVTAAYSEA	RGGGAIEKID
910	920	930						
LQICNDAGQC	LVQLKGFSSR	KTGVEP						

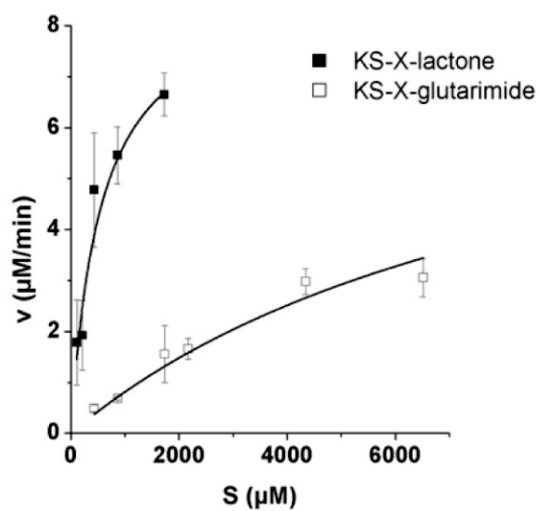
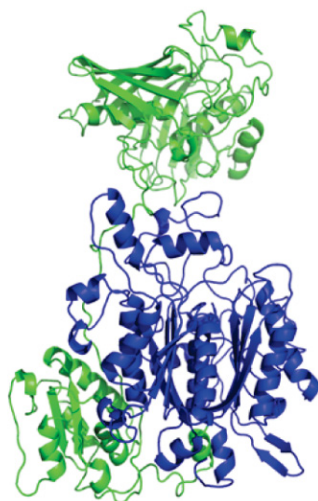
Protein:		RhE-Bb		Peak threshold:		0.0		
Intensity coverage:		19.0 % (15628 cnts)		Sequence coverage MS:		3.2 %		
		Sequence coverage MS/MS:		0.0 %		pl: 5.2		
		MW (kDa):		35.0				
10	20	30	40	50	60	70	80	90
IVNPLMFKNK	LHPLVAKNCS	TPQGAIFRT	DFVEDELLD	VYSGRGRRR	SAFNFDVAL	AMPALASRFD	GRTLVSVCFA	EHYIADWTTV
100	110	120	130	140	150	160	170	180
TGLEYLFEI	DSEQLELEFD	FRRSGEQPTH	LGFVINPLT	SDEPPLPQW	LDDARELLNR	QALQAGRQLS	AAEVSQRLAQ	AGYDFAPYLD
190	200	210	220	230	240	250	260	270
HDGELTIGRS	GLVLKGRPPV	NRNHYADNV	QLSPYLATTI	DKALYLLDE	LGLPQGRVIV	RNIERLCCYH	TPAGGFSVVL	SGIGLNDNEL
280	290	300	310	320				
SLSLLVLDER	EQICVKLDKV	SLYLKQEVA	SVDRKHSLLT	GT				

Supplementary Figure 6. KS-DH evaluation by MALDI fingerprinting. Tryptic digestion and MALDI analysis of the marked band corresponding to histidine-tagged KS-DH protein. Peptides from KS, KS-B-linker and DH domains were detected. Peptides corresponding to the B domain were not present. Lanes: 1 – KS-B; 2 – KS-X; 3 – KS-DH; M – protein marker.

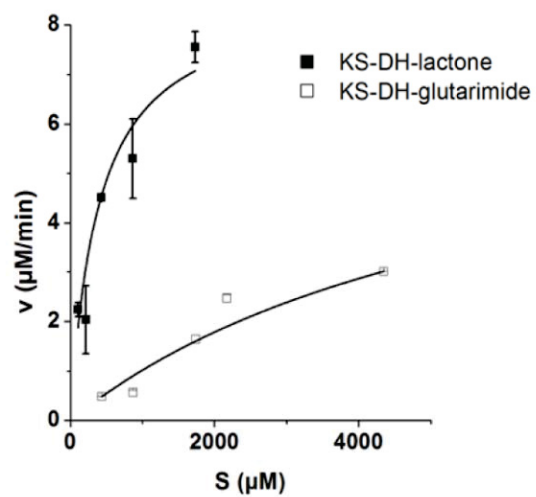


Supplementary Figure 7. *In vitro* branching assay analysis by MALDI-TOF. (a) Biotransformation of amide- (**2**) and (b) hydroxy-substituted SNAC thioester (**4**) using KS-X didomain. (c) Biotransformation of amide- (**2**) and (d) hydroxy-substituted SNAC thioester (**4**) using KS-DH didomain. MALDI mass spectrum revealed ACP-bound products. Negative control refers to the branching reaction without **2** or **4**. All examinations were performed in duplicates.

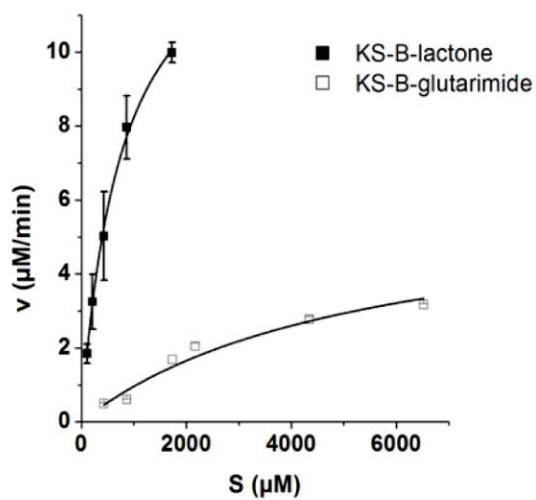
a



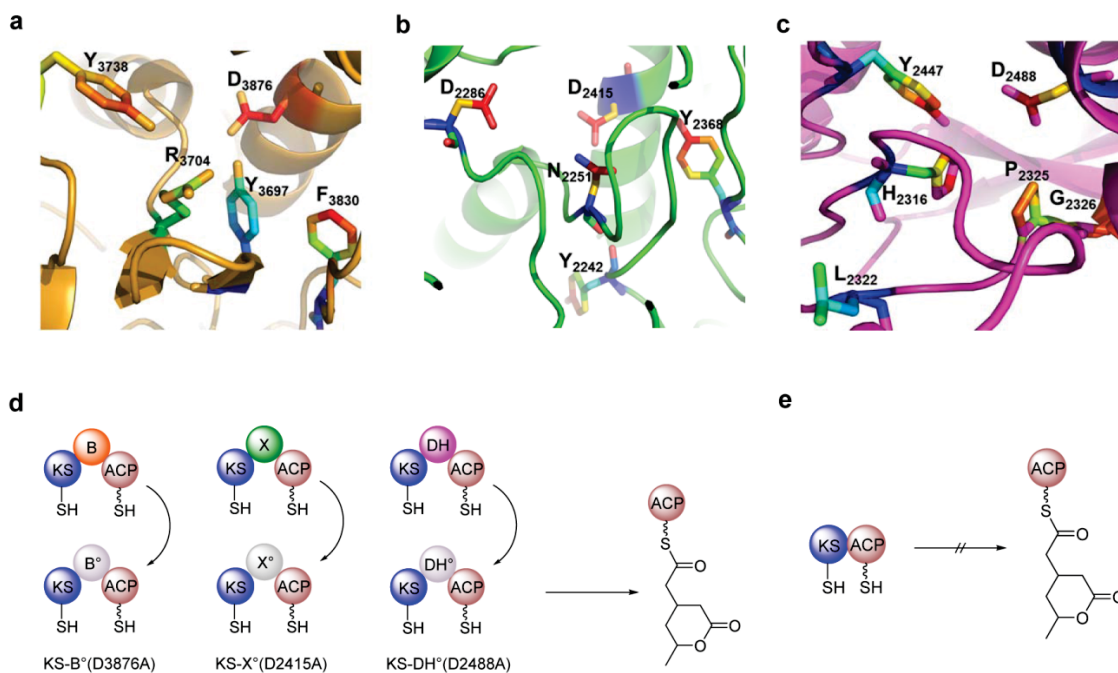
b



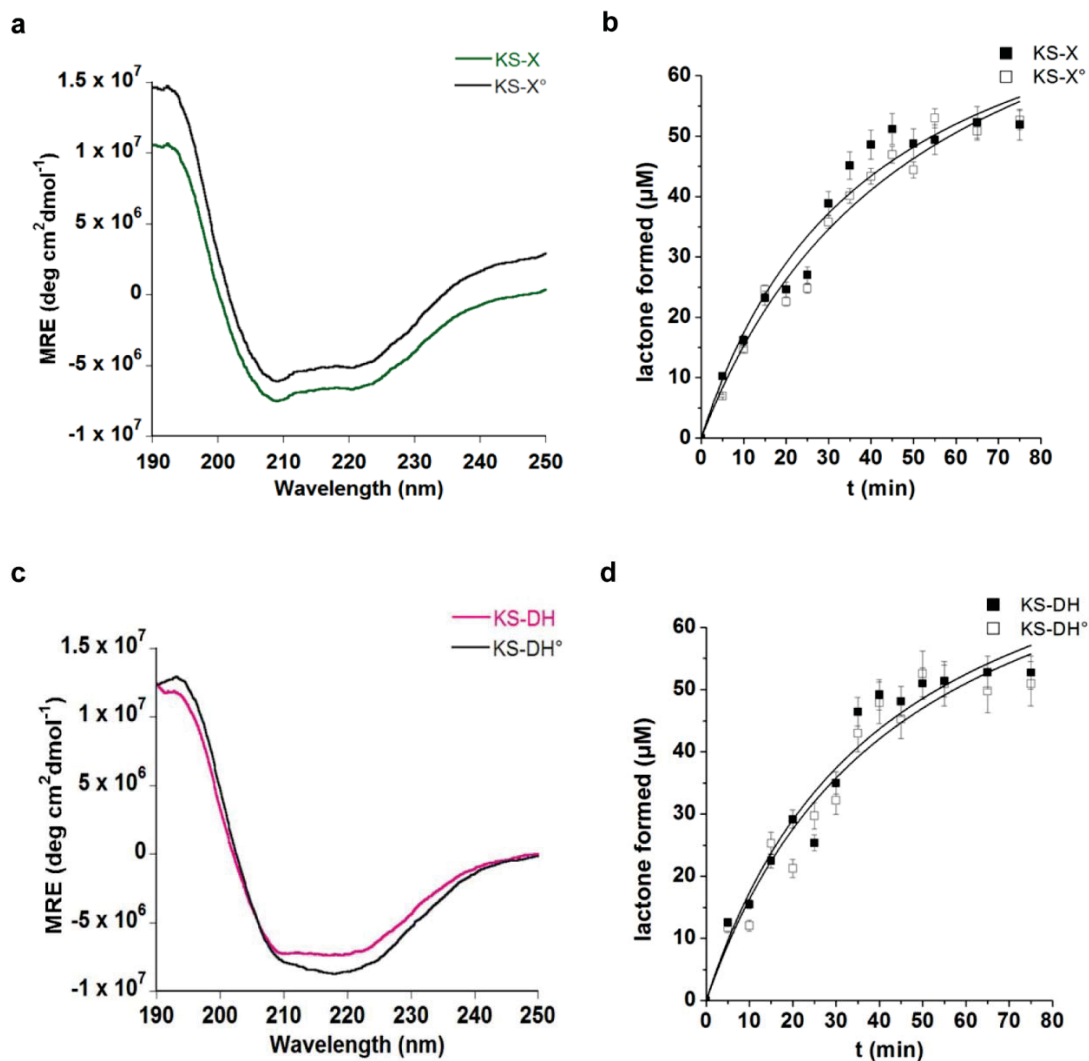
c



Supplementary Figure 8. Michaelis-Menten kinetics of lactone and glutarimide formation. Biocatalysis by (a) KS-X, (b) KS-DH and (c) KS-B. Lactone product turnover is at least 3–5 fold higher than that of glutarimide. The data represent mean \pm standard deviation resulting from 3 independent experiments. 3D structures of (a) KS-X (modeled), (b) KS-DH (modeled) and (c) KS-B are shown. Color codes are consistent with the corresponding domains. Modeling of chimeric proteins was done using Phyre 2 server and the structures were generated using Pymol.



Supplementary Figure 9. Mutational analysis of B, X and DH domains. Comparison of active site residues of (a) B, (b) X and (c) DH domains. (d) Mutants with point mutations of conserved Asp residue retained the branching activity. (e) KS-ACP fusion resulted in insoluble protein production hampering assay conditions.



Supplementary Figure 10. CD spectra and reaction rate of mutants. (a) CD analysis of KS-X in comparison with the mutant KS-X° (D2415A). (b) Kinetic data (mean \pm standard deviation from 3 independent experiments) of KS-X° in comparison with KS-X. (c) CD analysis of KS-DH in comparison with the mutant KS-DH° (D2488A). (d) Kinetic data (mean \pm standard deviation from 3 independent experiments) of KS-DH° in comparison with KS-DH.

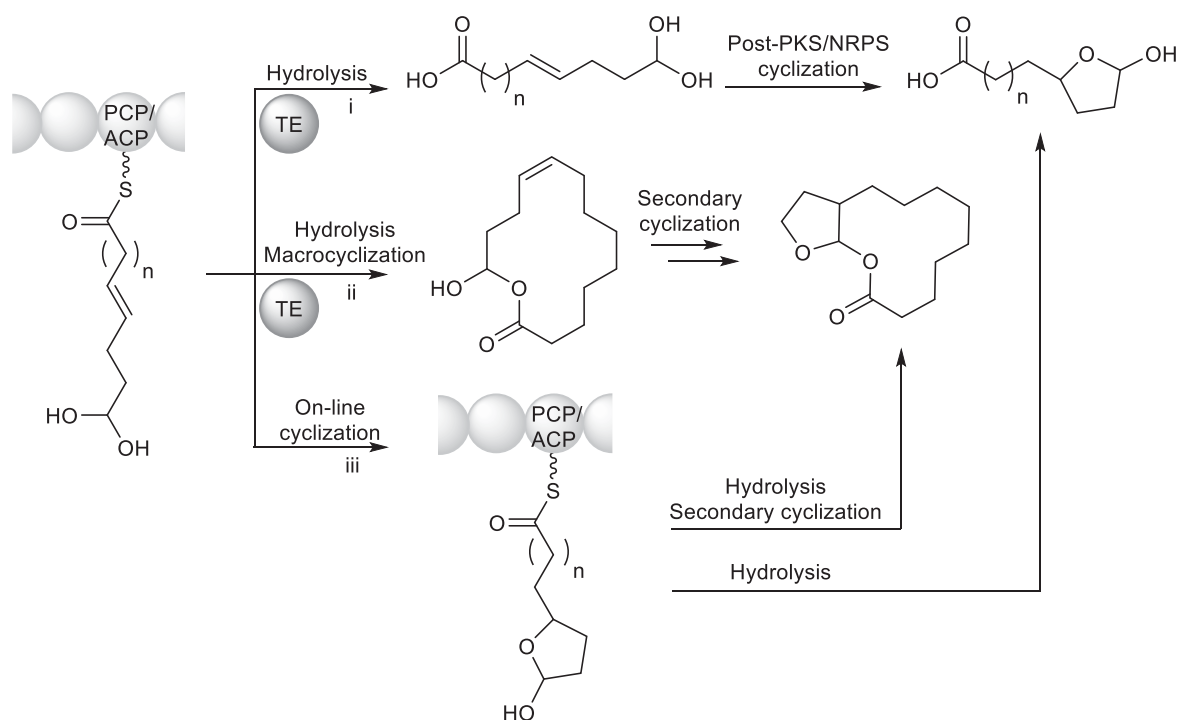
References

- 1 Wang, B. *et al.* Biosynthesis of 9-methylstreptimidone involves a new decarboxylative step for polyketide terminal diene formation. *Org. Lett.* **15**, 1278-1281 (2013).

3.5 Manuscript D

On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems (Review)

Srividhya Sundaram, Christian Hertweck: On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems, *Current Opinion in Chemical Biology* 31, 82-94, 2016.



Polyketides or nonribosomal peptide intermediates bound to carrier protein can be (i) released from the assembly line as a hydrolytic product by TE domain and cyclized further (ii) released by hydrolysis as a cyclized product by TE domain or (iii) cyclized while bound to the thiotemplate (on-line) by assembly line domains or *trans*-acting enzymes. The review discusses rings formed by on-line enzymatic tailoring mechanisms.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Chemical Biology

On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems

Srividhya Sundaram¹ and Christian Hertweck²

Non-ribosomal peptide synthetases (NRPS) and type I polyketide synthases (PKS) are versatile thiotemplate systems for the programmed assembly of biosynthetic building blocks. Typically, the post-PKS/NRPS enzymes tailor the resulting chains to yield the bioactive natural product scaffolds. However, more and more examples have surfaced showing that important structural modifications take place while the intermediates are still bound to the assembly line. A growing number of enzymatic domains and trans-acting enzymes as well as their recruiting areas in the modules have been identified and characterized. In addition to the widespread on-line alkylations, hydroxylations and heterocyclizations into oxazole/thiazole residues, on-line modifications lead to a variety of ring systems such as cycloethers, lactones, lactams, glutarimides, cyclopropanes, decalins and cyclic biaryls.

Addresses

¹ Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Beutenbergstrasse 11a, 07745 Jena, Germany

² Friedrich Schiller University, 07743 Jena, Germany

Corresponding author: Hertweck, Christian
(christian.hertweck@leibniz-hki.de)

Current Opinion in Chemical Biology 2016, 31:82–94

This review comes from a themed issue on **Biotransformation and biocatalysis**

Edited by **Dan S Tawfik** and **Wilfred A van der Donk**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 8th February 2016

<http://dx.doi.org/10.1016/j.cbpa.2016.01.012>

1367-5931/© 2016 Elsevier Ltd. All rights reserved.

Introduction

Thiotemplates are fascinating microbial multienzyme systems that produce a wealth of polyketide and peptide natural products, many of which are medically important as antibiotics, anticancer drugs and immunosuppressants. These versatile molecular assembly lines have a modular architecture with specialized domains for substrate loading, assembly, modification and chain release. Depending on the type of building blocks and mechanisms employed for the chain assembly, thiotemplates fall into two groups: type I polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS), yet there are many hybrids that generate mixed structures [1–4].

By contrast to biosynthetic pathways where several free-standing enzymes catalyze the stepwise transformation of free-floating, non-covalently bound substrates, in thiotemplates the growing polyketide or peptide chains are tethered to the phosphopantetheinyl arms of acyl-carrier protein (ACP) or peptidyl-carrier protein (PCP) domains until full maturation of the polyketide or peptide backbones. Thus, the thiotemplate systems prevents diffusion of the substrates and ensures spatial and temporal control over the many fine-tuned biochemical processes, defining the choice of building blocks, order of incorporation, degree of processing and the size of the chains. In addition, the terminal modules or domains determine the fate of the product (Figure 1). Typically, thioesterase (TE) domains liberate the products by either simple hydrolysis or (macro)cyclizations [5,6], including dimerizations to create macrodiolides [7,8]. Alternatively, reductive chain release, aminolysis as well as various more sophisticated off-loading mechanisms are known, which may lead to tetronates and several other ring systems (Figure 1b) [9,10]. Despite these variations governed by chain-termination processes, according to text-book knowledge thiotemplates are mainly in charge of assembling the basic peptide and polyketide scaffolds, whereas so-called post-PKS and post-NRPS enzymes are responsible for tailoring reactions such as oxygenations, glycosylations and cyclizations, which are often crucial for the biological activities of the final products. As the names post-PKS and post-NRPS imply, chain assembly and tailoring reactions would be time-resolved events. However, with the growing number of studied biosynthetic pathways involving thiotemplates, more and more examples surfaced where *trans*-acting enzymes or novel PKS/NRPS domains catalyze peptide/polyketide transformations during chain assembly.

Common on-line modifications (alkylations, hydroxylations, and amino acid heterocyclizations)

In addition to the canonical NRPS and PKS domains for β -keto processing and amino acid epimerization, domains for on-line hydroxylation and methyl transfer are widespread. In particular, integral SAM-dependent methyltransferase (MT) domains for α -methylation are frequently found in *trans*-AT PKSs, a subgroup of bacterial modular PKSs in which the malonyl units are loaded by a *trans*-acting acyl transferases (AT) [11]. For the same type of PKSs, another on-line alkylation mechanism is prevalent that leads to terpenoid-like alkyl branches at the β -carbon relative to the thioester. A set of enzymes encoded by the hydroxymethylglutaryl-CoA synthase

(HCS) gene cassette is in charge of mediating the β -branching reactions in *trans* [12]. Specialized motifs in the PKS warrant the formation of the HCS-ACP complex in the correct module [13].

Perhaps the best-known on-line cyclization reactions are the cyclodehydrations of serine and cysteine residues of peptides into oxazoline and thiazoline rings, respectively, which are often found in siderophores (Figure 1c). For this common modification, NRPS modules harbor cyclization (CY) domains, which may be paired with oxygenase (OXY) domains for the formation of oxazole and thiazole heterocycles [6*]. Such backbone cyclizations are often important for the overall architecture of the natural products and rigidify the scaffold. The following sections highlight recent discoveries of on-line tailoring reactions with emphasis on cyclizations (Figure 1d).

Cycloether formation during polyketide biosynthesis

Cycloethers are often important moieties that confer biological activity to polyketide metabolites. Prominent examples are polyether ionophores like monensin and lasolocid. Most often, the oxacycles are incorporated by epoxidation followed by intramolecular nucleophilic epoxide ring opening catalyzed by oxygenases and epoxide hydrolases/cyclases [14–16]. Tetrahydrofuran and tetrahydropyran rings may also be formed by vinylogous addition reactions and cytochrome-mediated C–O-bond formation, yet until recently all of these reactions were known as typical post-PKS modifications.

Recently, specific PKS domains have been discovered that mediate polyketide heterocyclization during chain assembly. One such example is the dehydratase (DH) domain involved in the biosynthesis of a potent haterulamide, oocydin A (5) produced by plant-associated enterobacteria of the *Serratia* and *Dickeya* genera. One of the two DH domains mediates the attack of a hydroxyl group to the β -position of an α , β -unsaturated thioester in a Michael addition reaction, resulting in a furan ring (Figure 2a) [17]. A similar reaction is shown to be catalyzed by a pyran synthase-like (PS) domain in the biosynthesis of phormidolide A (6), a brominated macrocyclic polyketide isolated from a marine cyanobacterium (*Leptolyngbya* sp.) (Figure 2b) [18]. In both cases, the catalytic domains are part of a type I *trans*-AT PKS. The PS domain is also identified in the biosynthesis of pederin (7), an antimetabolic agent produced by a symbiont (*Pseudomonas* sp.) of rove beetles (*Paederus* spp.) [19]. The catalytic domain, which is part of a type I *trans*-AT PKS, exhibits moderate homology to DH domains with mutated active site residues. The PS domain was shown to mediate the oxa-conjugate addition in a stereocontrolled manner. As a stand-alone enzyme the PS domain is capable of forming 5-membered and 6-membered cycloethers, thus representing a versatile biocatalyst

(Figure 2c) [20]. Notably, till date no PS-like domains have been identified in *cis*-AT PKS systems. However, a functionally related DH domain with both dehydratase and cyclase activities plays a key role in the biosynthesis of ambruticin (8), an antifungal agent produced by the myxobacterium *Sorangium cellulosum*. Analysis of the ambruticin modular (*cis*-AT) PKS and *in vitro* assays using synthetic substrate analogs showed that the DH3 domain catalyzes a dehydration-cyclization cascade to yield the pyran ring (Figure 2d) [21].

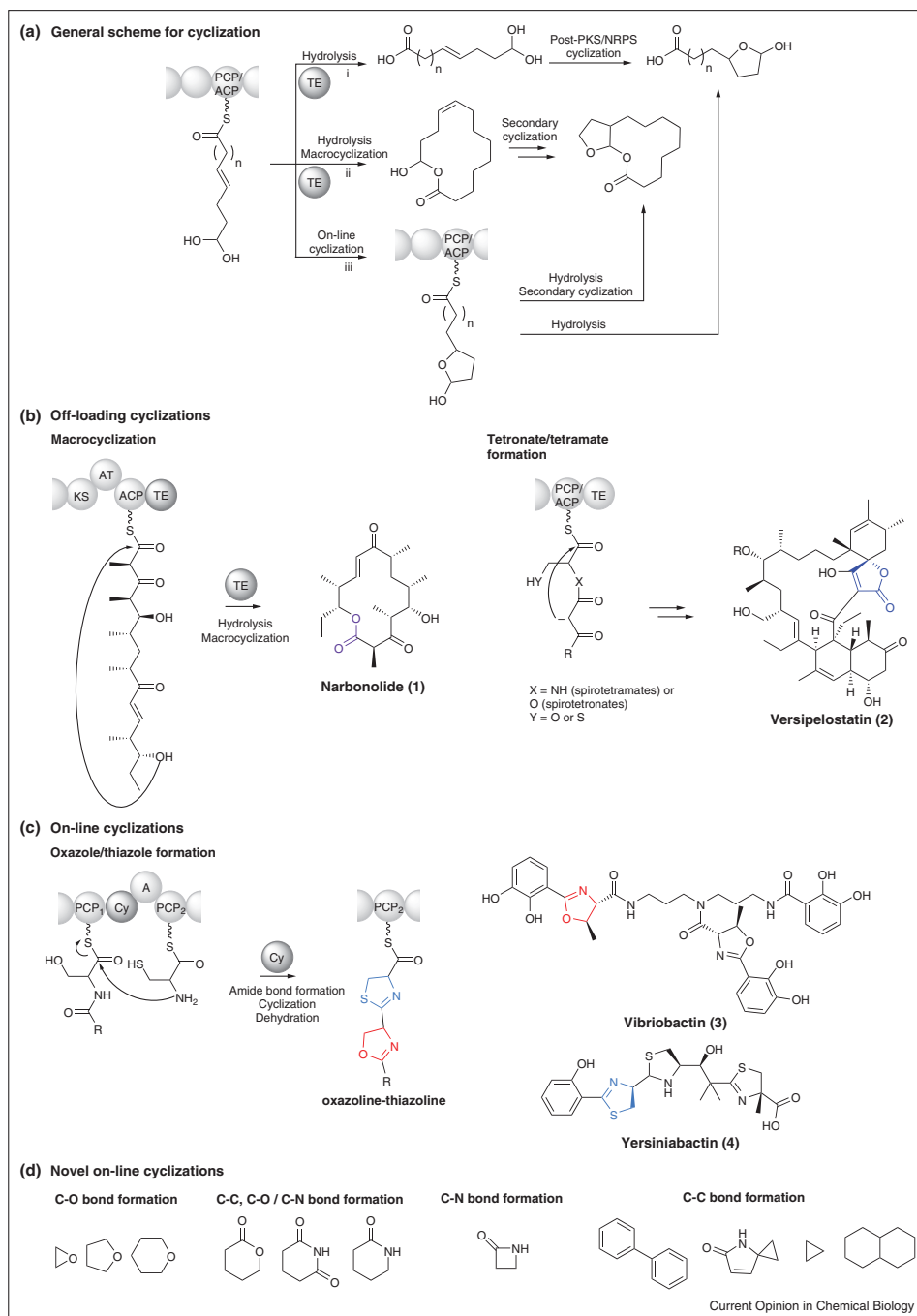
A tetrahydropyran ring is also found in the structure of salinomycin (9), a polyether ionophore from *Streptomyces albus*, yet in contrast to the above-mentioned pathways, no 'PS' or 'DH'-like domains can be found in the modular salinomycin PKS. It was shown that pyran ring formation requires a novel cyclase, SalBIII, which is similar to the MonB-type of epoxide hydrolases. By contrast to the conventional epoxide ring opening reactions typically mediated by such enzymes, SalBIII catalyzes an E1cb-like β -elimination reaction (Figure 2e) [22]. It was proposed that SalBIII acts in *trans* to transform the PKS-bound intermediate into the oxacycle. The crystal structure of the enzyme has been solved, which revealed two aspartate residues (Asp38 and Asp104) in the active site, which are implicated in the elimination-addition reaction. On the basis of gene cluster analyses it is conceivable that an integral Cyc11 domain of the indanomycin PKS mediates pyran ring formation by a similar mechanism [23]. Taken together, the PS and DH domains of the pederin and ambruticin PKSs and the novel *trans*-acting SalBIII cyclase broaden the range of online polyketide heterocyclizations.

Polyketide branching with lactonization, glutarimide formation and lactamization

In addition to cycloether biosynthesis, on-line Michael additions are also involved in the formation of various other pharmacophoric heterocycles such as δ -lactone and glutarimide moieties that branch off the polyketide backbone. Contrary to O-heterocyclization, these additional ring systems are installed by the vinylogous attack of an α , β -unsaturated thioester intermediate by a C-nucleophile, produced by decarboxylation of the ACP-bound malonyl extender unit, followed by cyclization. First insights into this unusual polyketide branching mechanism were gleaned from analysis of the modular *trans*-AT PKS for the antimetabolic phytotoxin rhizoxin (10) (Figure 3a) that is produced in symbiosis of the fungus *Rhizopus microsporus* and its endobacterium *Burkholderia rhizoxinica* [24,25]. Elucidation of the pathway intermediates and the assembly line program [26] suggested that a designated branching module (KS (Ketosynthase)-B (Branching)-ACP) plays a key role in chain branching. By the successful *in vitro* reconstitution of the module it was found that this set of domains is sufficient to catalyze the entire branching/cyclization sequence. According to results obtained from stable-isotope labeling experiments and trapping

84 Biocatalysis and biotransformation

Figure 1



the Michael adduct, the exact course of the reaction could be deduced [27^{**}]. The Michael addition leads to a bridged KS-ACP intermediate, which is cleaved by attack of the δ -hydroxy group on the KS thioester. In addition to the regular C2 extender unit, the KS-B didomain also accepts a methylmalonyl-derived building block to give a two-fold branched product [28]. The mechanistic study of the rhizoxin branching module also provided a model for the formation of glutarimide-substituted polyketides such as migrastatin (**11**), 9-methyl streptimidone (**12**) and cycloheximide (**13**) (Figure 3b). On the basis of PKS gene analyses it was proposed that, in analogy to rhizoxin δ -lactone formation, a KS-X-ACP module (X for mystery domain) would be involved in glutarimide formation [29–31]. Using the rhizoxin branching module it was shown that changing the OH group to a carboxamide substituent yields the glutarimide residue [32]. In light of the above-mentioned, PS/DH-catalyzed Michael additions it is remarkable that the B domain has a double hotdog fold that is a hallmark of DH domains [33]. To clarify the role of the B and X domains in lactone and glutarimide biosynthesis, chimeras of branching modules were constructed. Mutational and kinetic studies and the substitution with functional and mutated DH domains indicated, however, that these domains do not play a catalytic role [34]. Lastly, the branching/cyclization sequence works also for amine-substituted polyketides, thus providing a new biocatalytic avenue to δ -lactams (Figure 3c). Although this route has not yet been observed for any natural product, it is plausible that such pathways exist in nature. Intriguingly, only recently new on-line modification reactions leading to β -lactams and γ -lactams have been revealed, as outlined in the following sections.

β -Lactamization during peptide assembly

β -Lactams are typical motifs of a broad range of antibiotics including penicillins and cephalosporins, and are essential for their biological activity. The strained rings are typically formed as a post-NRPS reaction by oxidative bicyclization (for isopenicillin N) or by an additional β -lactam synthetase (for clavulanic acid) [35]. Recently, a third route for β -lactam formation has been elucidated in the course of studying nocardicin G (**14**) biosynthesis in *Nocardia uniformis* [36^{**}]. Nocardicin belongs to the family of monocyclic β -lactam antibiotics that are not inactivated by β -lactamase, and acts against Gram-negative bacteria. Nocardicin is synthesized by a non-ribosomal peptide synthetase (composed of NocA and NocB) that posed several riddles. Domain analysis suggested that a pentapeptide would be the product of the pentamodular

NRPS. However, in lieu of the predicted L-pHPG-L-Arg-D-pHPG-L-Ser-L-pHPG pentapeptide backbone (pHPG, para-hydroxyphenylglycine), naturally occurring nocardicins show only a tripeptide backbone with a β -lactam core and an epimerized C-terminus. By a series of studies at the genetic and biochemical levels it was found that the immediate product of the NRPS is indeed a pentapeptide, and that the L-pHPG-L-Arg terminus is important for the processing of the peptide. With help of synthetic surrogates of pathway intermediates it was found that the TE domain plays an unusual role. It not only functions as an epimerase to invert the configuration of the C-terminal pHPG [37], but also acts as a gate keeper to hold the peptide backbone until the β -lactam ring had formed. Thus, compelling evidence existed that β -lactam formation takes place during, not after peptide assembly. According to a plausible model, the new route to β -lactams involves β -elimination at the Ser residue to give an intermediary dehydroalanyl; β -addition of the incoming C-terminal pHPG residue and aminolysis of the thioester yields the β -lactam ring [36^{**}]. Using synthetic L-pHPG-L-Arg-D-pHPG-L-Ser-S-PCP₄ and purified holomodule 5 (Figure 3d), the reaction sequence was reconstituted *in vitro*. The condensation domain in module 5 (C₅) was found to be essential for this unprecedented sequence. In addition to the typical conserved residues, C₅ has a third His residue (His₇₉₀) in the catalytic motif, H₇₉₀HH₇₉₂XXXDG that plays a central role in the lactam formation. Since a His790A mutant proved to be non-functional, it was concluded that this histidine residue promotes both the dehydration of the Ser residue and the conjugate addition [36^{**}].

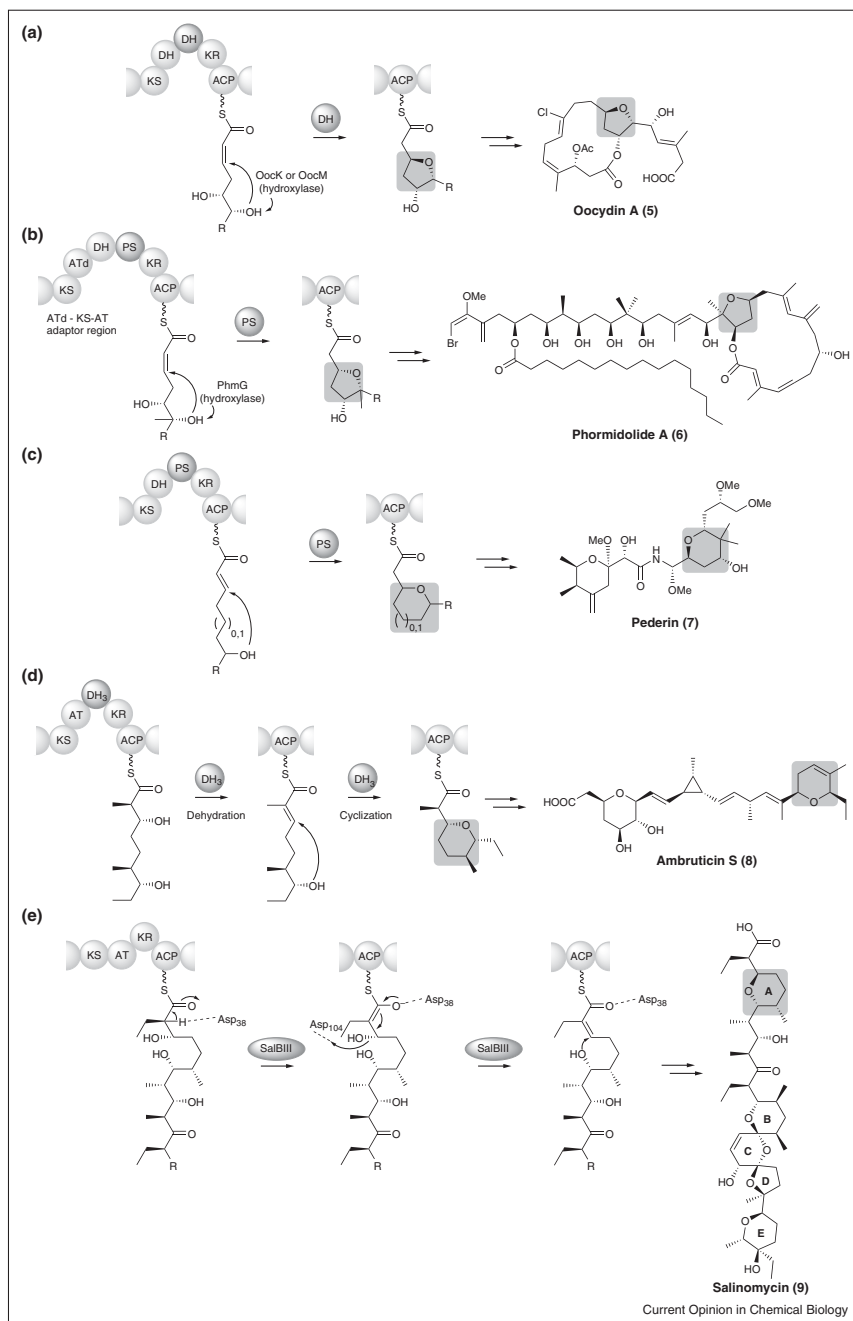
Aryl cross-coupling of NRPS-bound intermediates

As for β -lactams, intramolecular cyclizations of peptide backbones are indispensable for the antibiotic action of actinomycete-derived glycopeptide antibiotics, such as vancomycin, balhimycin and teicoplanin. Specifically, aryl cross-couplings of various aromatic amino acid residues lead to interconnected lactam rings that stabilize the three-dimensional heptapeptide structures of these valuable antibacterial agents [38,39]. For a long time it has been known that the cross-coupling is catalyzed by *trans*-acting cytochrome monooxygenases (CYPs) in a highly specific order [40]. Furthermore, these specialized CYPs fuse their substrates only when they are bound to the NRPS. For example, in teicoplanin (**15**) biosynthesis by *Actinoplanes teichomyeticus*, OxyB_{te} installs the first crosslink between the side chains of aromatic residues [41,42]. *In vitro* assays

(Figure 1 Legend) (a) General cyclization patterns. Polyketide or peptide intermediates bound to carrier protein can be (i) released from the assembly line as linear compounds and cyclized by tailoring enzymes, or (ii) released as a macrocyclized product that may further be cyclized or (iii) cyclized on-line by assembly line domains or *trans*-acting enzymes. (b) Examples of off-loading mechanisms in the formation of a macrocyclized narbonolide **1** and the spiro-tetronate versipelostatin **2**. (c) Examples of well-known on-line cyclizations that involves the Cy domain for the formation of oxazole/thiazole moieties (vibriobactin **3** and yersiniabactin **4**). (d) Ring structures formed by on-line tailoring as discussed in this review.

86 Biocatalysis and biotransformation

Figure 2

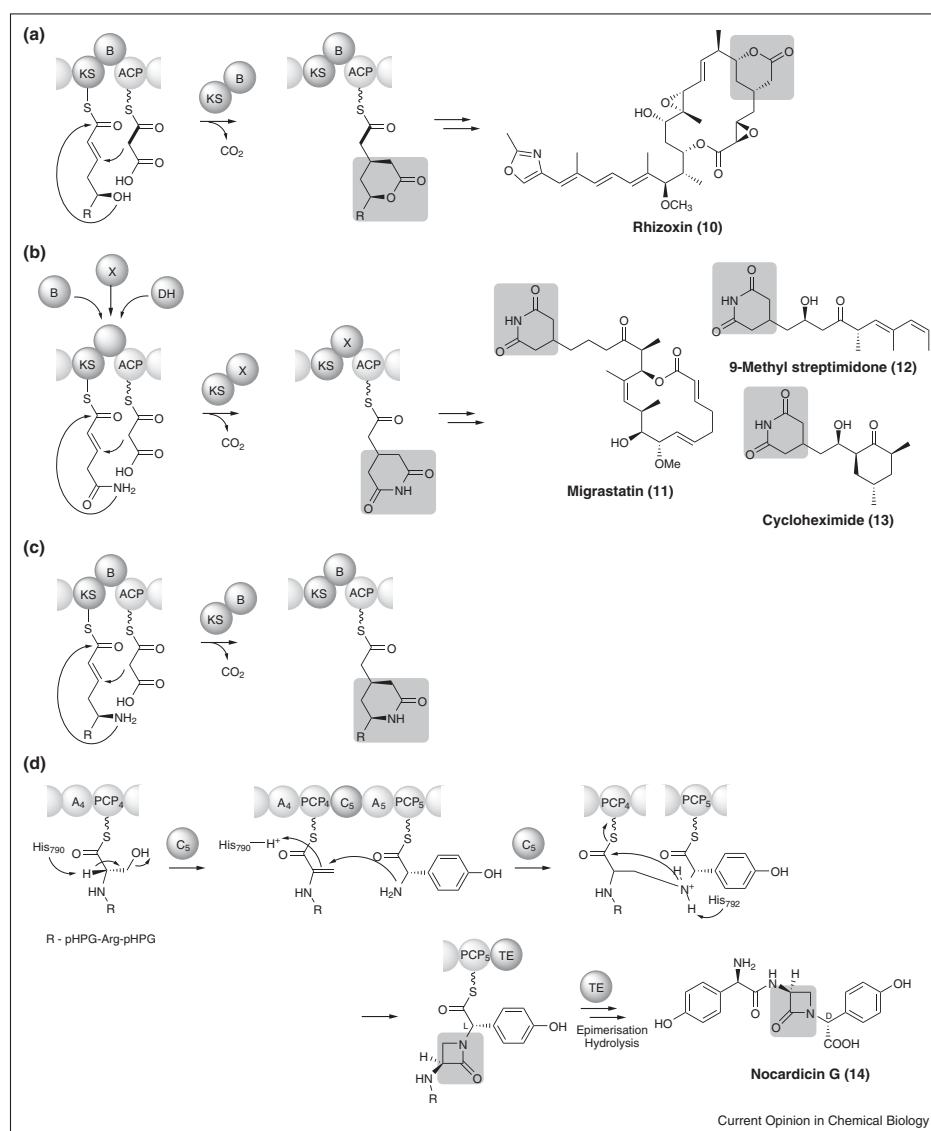


Formation of furan rings by (a) DH domain (oocidin A 5) and (b) PS domain (phormidolide A 6). Formation of pyran rings by (c) PS domain (pederin 7), (d) DH₃ domain (ambruticin S 8) and (e) SalBIll, a tailoring enzyme (salinomycin 9).

showed that for this protein, the peptidyl-PCP alone is not sufficient to mediate the entire crosslinking. Although the other oxygenase enzymes (OxyA, OxyC) have been identified, the mechanism of recruiting these enzymes to the PCP-bound peptide has remained elusive.

Only recently, it was shown that all glycopeptide-type NRPS assembly lines feature a specialized domain (named X) in the last module, which resembles condensation/epimerization domains but has a mutated active site. By generating constructs of the entire module or specific

Figure 3



Lactone, glutarimide and lactam formation by non-canonical domains. (a) Vinylogous chain branching in rhizoxin **10** biosynthesis catalyzed by KS-B didomain. (b) Similar branching mechanisms are observed in the synthesis of glutarimide residue seen in migrastatin **11**, 9-methyl streptimidone **12** and cycloheximide **13**, catalyzed by KS-X didomain; a regular DH domain can replace B or X domains in the module. (c) δ -Lactam formation by the branching module. (d) Mechanism of β -lactam ring formation in nocardicin G **14** by a novel condensation domain (C_5).

88 Biocatalysis and biotransformation

domains, it has been demonstrated that the minimal requirement for OxyBtei to interact with the NRPS-bound peptide is the presence of the X domain, and that the interaction happens in the last module (7) of the teicoplanin or vancomycin NRPS (Figure 4) [43**]. This is different from the other P450 recruitment systems that rely only on the carrier protein [44]. In addition to OxyB, the X domain strongly interacts with OxyA and, to a lesser degree, with OxyC. This role of a domain to act as a recruiter of oxygenase enzymes has been unprecedented and appears to be a hallmark of glycopeptide biosynthesis. A range of heptapeptide and hexapeptide substrates, when presented on the PCP-X didomain *in vitro*, was also shown to be sequentially cyclized by OxyB and OxyA, indicating the impressive substrate tolerance of these enzymes [45]. However, one may also conceive that similarly controlled mechanisms take place in other biosynthetic pathways leading to aryl-coupled natural products.

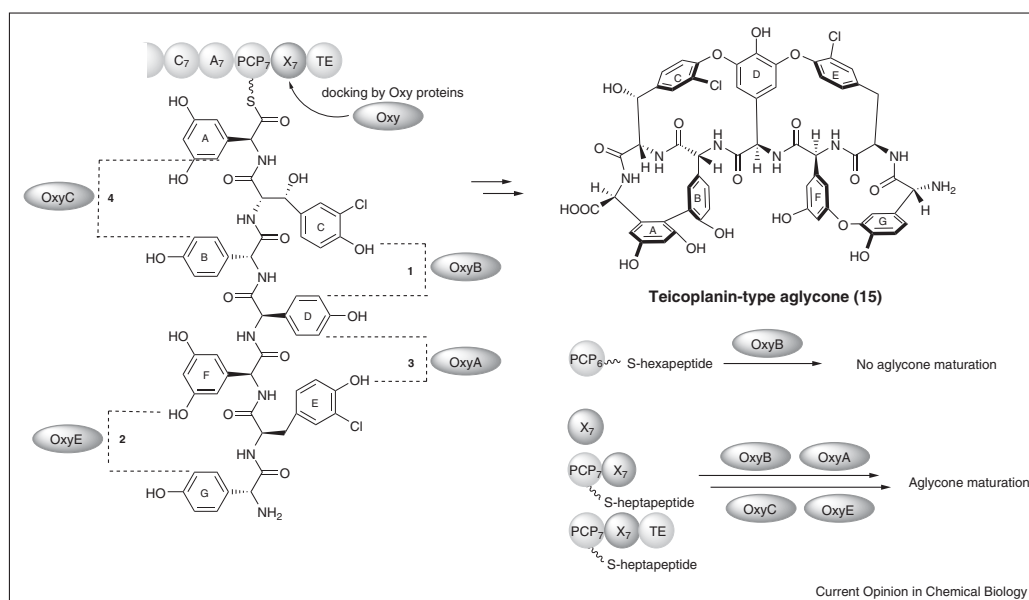
Knoevenagel cyclocondensation

In addition to the above-mentioned avenues, yet another pathway involving an on-line lactam formation has recently been proposed in the biosynthesis of colibactin, a not yet fully characterized metabolite of the human commensal *Escherichia coli*. Through genetic analyses it has been known that colibactin is produced by a hybrid PKS-NRPS pathway known as the '*pks* island'. The enigmatic compound leads to DNA damage, genomic

instability in mammalian cells and thus correlates with the induction of colorectal cancers [46]. Although the gene cluster responsible for colibactin has been known for over 10 years (*clb*), its structure has been one of the challenging elements to address because of the low yield, extreme instability of the intermediates and the contact-dependent cytopathic activity. Moreover, it appears that colibactin affords a prodrug strategy as seen in xenocoumacin biosynthesis [47] with the N-terminal part of the 'precolibactin' removed at a later stage by an encoded peptidase, ClbP. Over the recent years, many groups have tried to fit the puzzle with the pathway intermediates to propose the structure of colibactin. A NMR-based structural network analysis approach targeting the pathway intermediates had revealed structures of shunt products, precolibactin. One of the isolated compounds (16) features an azaspiro bicyclic ring structure, which probably functions as a warhead of colibactin. Upon release by the peptidase ClbP (with the exposed N-terminal primary amine and a thiazoline-thiazole tail) the electrophilic moiety can crosslink DNA by alkylation, probably by a homo-Michael addition [48*].

The intriguing cyclic structure, which incorporates a Met-derived 1-aminocyclopropane-1-carboxylic acid (ACC) extender unit, is probably formed by an intramolecular Knoevenagel cyclocondensation reaction while the peptide-polyketide intermediate is still tethered to the PKS-NRPS,

Figure 4



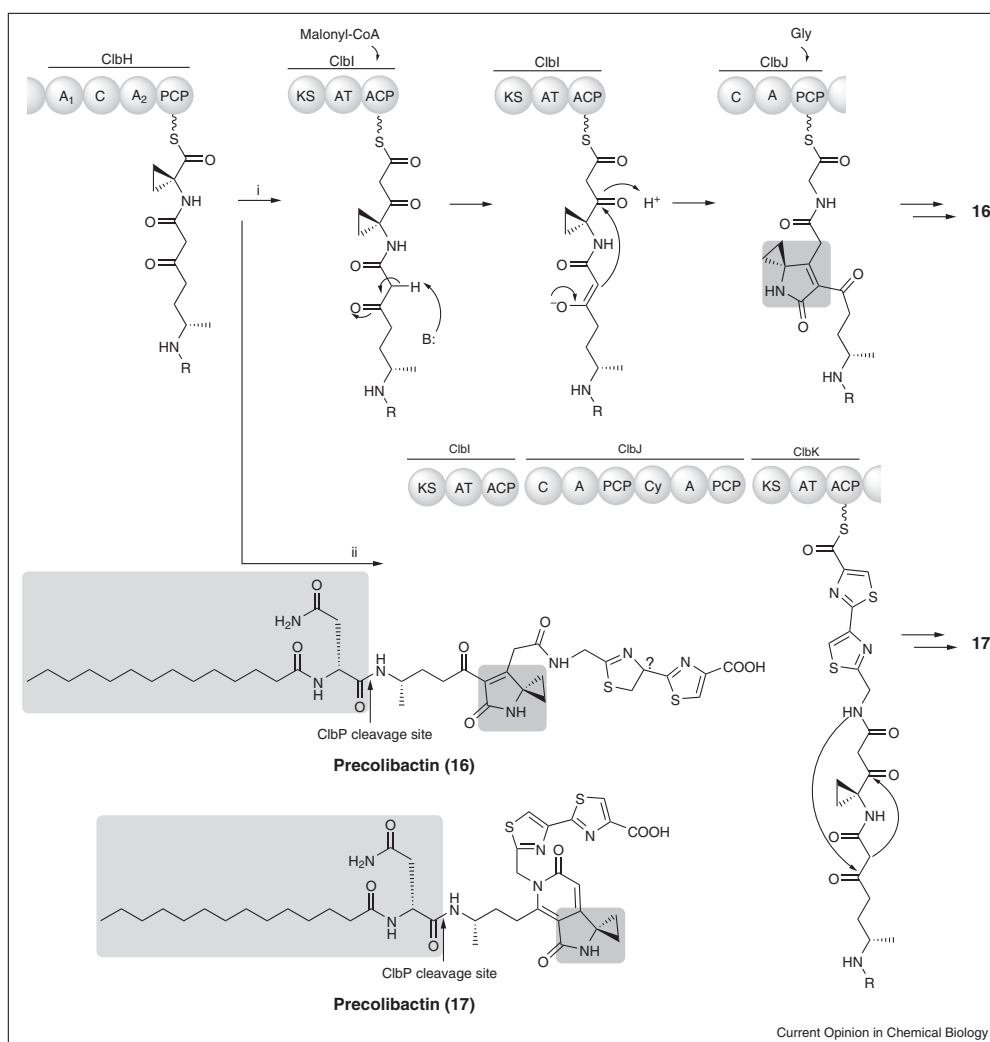
A conserved X domain in glycopeptide biosynthesis recruits oxygenases to form mature aglycone as shown for teicoplanin 15 biosynthesis.

followed by thioester cleavage and decarboxylation (Figure 5, Route i) [48*,49]. However, the biochemical basis of and the timing of the reaction need to be clarified. It cannot be ruled out that the cyclization occurs spontaneously after chain release [50]. Furthermore, a slightly different structure was elucidated for another isolated precolibactin candidate (17), which suggested that decarboxylation could also precede cyclization to form the precolibactin warhead (Figure 5, Route ii) [51].

On-line cyclopropanations

Whereas the cyclopropyl group of colibactin obviously derives from an ACC building block, small carbocycles may form during polyketide assembly. An unexpected variation of the terpenoid-like polyketide β -branching reaction [12] leads to the characteristic cyclopropyl residue of curacin A (18), a PKS-NRPS hybrid produced by the cyanobacterium *Lyngbya majuscula*. Dissection of the pathway and biochemical assays revealed that

Figure 5



Proposed intramolecular cyclocondensation reaction in colibactin biosynthesis and structures of full-length precolibactin derivatives (16 and 17) with the spirobicyclic warhead.

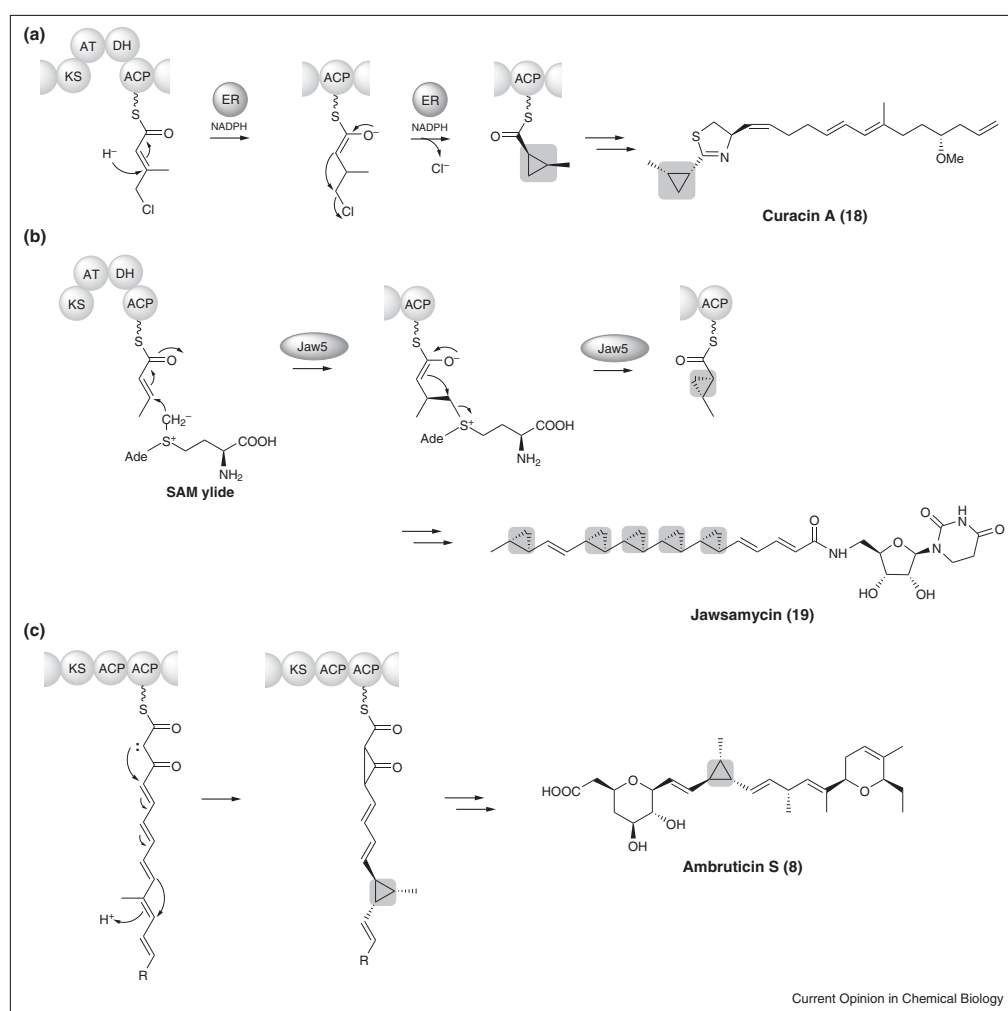
90 Biocatalysis and biotransformation

the (*S*)-3-hydroxy-3-methylglutaryl intermediate undergoes a chlorination by a ketoglutarate-dependent non-heme Fe(II) halogenase [52,53]. The halogenation reaction is followed by dehydration and decarboxylation, catalyzed by ECH1 and ECH2, respectively, to yield the allylic chloride. Finally, the enoyl reductase domain of CurF promotes the subsequent cyclization to afford the strained three-membered ring [52] (Figure 6a). It should be noted that the chlorination of the early pathway intermediate does not lead to a chloro-substituted final product, but sets the stage for the cyclization, thus

representing another rare case of a cryptic halogenation en route to cyclopropyl rings [54].

A different biosynthetic strategy for on-line cyclopropanations has evolved for the structurally intriguing polycyclopropyl natural product jawsamycin (**19**) (initially termed FR-900848) [55]. Heterologous expression and mutational experiments indicated that the small rings are introduced stepwise into the growing polyunsaturated polyketide chain. An enzyme with homology to radical *S*-adenosylmethionine (SAM) enzymes (Jaw5) plays a key

Figure 6



Formation of cyclopropyl rings by (a) NADPH-mediated ER domain (curacin A **18**), (b) SAM-dependent cyclopropanase Jaw5 (jawsamycin **19**) and (c) a putative Favorskiiase (ambruticin S **8**).

role in the polycyclopropanation (Figure 6b). In the current biosynthetic model, first a SAM ylide is formed by homolytic C–H cleavage of a SAM methyl and subsequent electron transfer from the Fe–S cluster, which reacts with the enoyl group. Alternatively, a reaction with a SAM methyl radical would also be possible. The exact mechanism and the biosynthetic program for these highly selective reactions remain to be elucidated. Nonetheless, the Jaw5-mediated reactions are an important addition to the known on-line alkylation reactions, which may open new possibilities for pathway engineering.

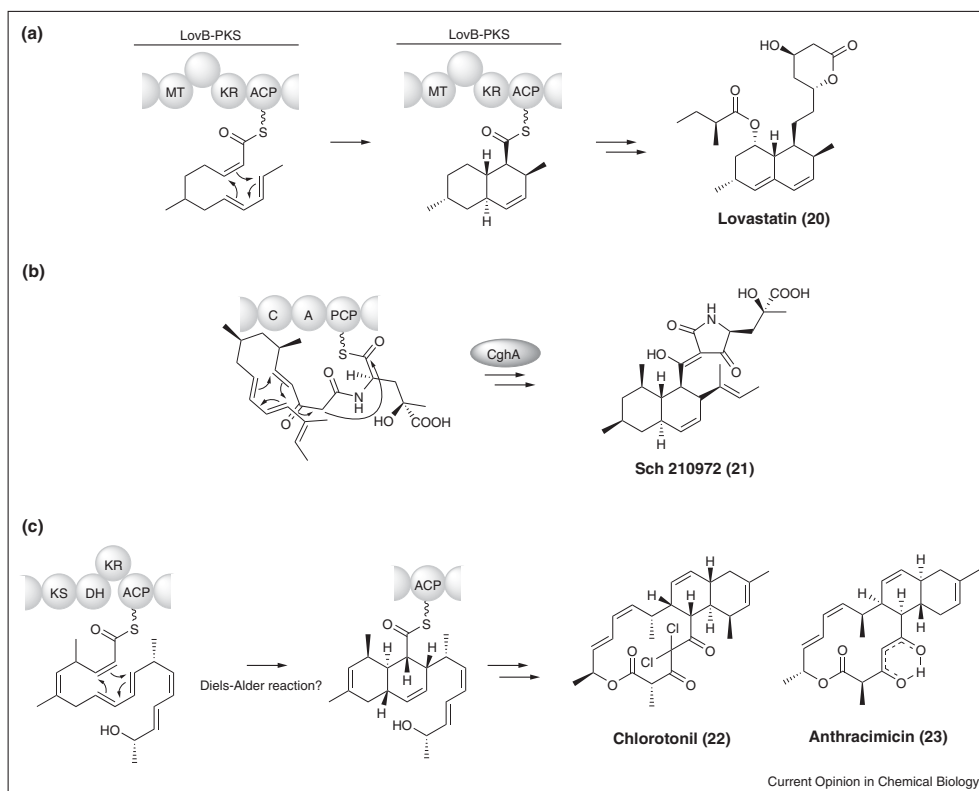
A route to cyclopropanes that is probably independent of both halogenation and alkylation has been realized in the biosynthetic pathway to ambruticin (8), an antifungal agent from *S. cellulosum* [56]. On the basis of the deduced biosynthetic program of the multimodular PKS it was proposed that a Favorskii-like rearrangement takes place on an ACP-bound polyene intermediate (Figure 6c). This

reaction would afford the loss of one carbon of an intermediary cyclopropanone residue. Still the putative Favorskiiase component is elusive and the exact reaction mechanism needs to be clarified.

Intramolecular [4 + 2] cycloadditions

By contrast to the numerous well-defined domains for intermediate modification or recruitment of *trans*-acting enzymes, on-line modification processes involving cycloadditions are rather cryptic. This applies in particular to the large number of natural products that derive from intramolecular [4 + 2] cycloadditions. A growing, heterogeneous family of enzymes have been implicated as mediators of Diels–Alder reactions, including lovastatin synthase (LovB) [57], SpnF in spinosyn biosynthesis [58], and VstJ in the biosynthesis of the spirotetronate versipelostatin [59]. Similarly, in the biosynthesis of a spirotetramate-containing pyrroindomycins, two distinct [4 + 2] cycloadditions have been shown to form the spiroconjugate [60]. The

Figure 7



(a) LovB-catalyzed Diels–Alder reaction in lovastatin 20 biosynthesis. (b) Lipocalin-like CghA catalyzed Diels–Alder reaction in the formation of tetramic acid-containing Sch 210972 21. (c) Intramolecular [4 + 2] cycloaddition occurs spontaneously in case of polyketides chlorotonil 22 and its analog, anthracimicin 23.

Diels–Alder reaction in lovastatin (**20**) biosynthesis has long been the text-book example for an on-line cycloaddition (Figure 7a), yet no distinct domain or enzymatic feature could be attributed to the [4 + 2] cycloaddition. Recently, in the biosynthesis of a potential anti-HIV agent Sch 210972 (**21**), produced by the fungus *Chaetomium globosum*, the decalin core has been proposed to form through a similar Diels–Alder reaction, but in the presence of a lipocalin-like enzyme, CghA (Figure 7b) [61]. This tetramic-acid containing natural product is by a PKS-NRPS hybrid pathway. It has been proposed that CghA-like proteins might help bind to specific conformation of the polyketide intermediate to perform stereoselective cycloaddition reaction.

For modular PKS systems, cases are known where intermediates seem to spontaneously undergo Diels–Alder reactions. Chlorotoniol A (**22**) from the myxobacterium *S. cellulosum* So ce 1525 [62] and anthracimycin (**23**) produced by *Streptomyces* sp. CNH365 [63] are related complex polyketides that differ in inverted stereocenters and in the geminal dichloro moiety at C₄. In both deduced pathways, the ACP-tethered elongated α,β -unsaturated nonaketides have *trans*-dienophiles that favor an *endo*-transition state, thus leading to the *trans*-decalin geometry observed in the molecules (Figure 7c). This Diels–Alder-like [4 + 2] cycloaddition probably occurs spontaneously as the diene and dienophile moieties are in close proximities to promote ring closure. The inverted stereocenters could be rationalized by the stereoinduction of the C₁₆ methyl group and the *R*-configured hydroxyl group in anthracimycin [62]. Even in cases where intermediates cyclize spontaneously, it is the thiotemplate program that generates structures with inherent reactivity, such as diene and dienophile moieties.

Conclusions

This brief overview illustrates that in addition to the manifold post-PKS and post-NRPS modifications catalyzed by diverse enzymes, more and more examples of on-line reactions have emerged. Although the number of these thiotemplate-related processes is still comparably low, it should be highlighted that the mentioned reactions are of eminent importance for the bioactivity of the final products. On-line cyclizations not only lead to pharmacophoric groups that are crucial for the binding to the targets by introducing critical residues or arresting an otherwise flexible molecule in a preferred conformation, but may also yield residues that function as warhead. The novel PKS and NRPS domains as well as the *trans*-acting enzymes and their recruiting domains are important additions to the biosynthetic toolbox. Knowledge of the assembly line logic will open new possibilities to engineer biosynthetic pathway to create structural diversity. In addition, elucidation of the biosynthetic code will aid in genome mining approaches for the discovery of hitherto unknown enzymes and natural products.

Acknowledgements

In the author's laboratory original research on the subject was financially supported by the International Leibniz Research School (ILRS, to S.S).

We thank Dr. Keishi Ishida for critically reading the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Walsh CT: **The chemical versatility of natural-product assembly lines.** *Acc Chem Res* 2008, **41**:4-10.
2. Meier JL, Burkart MD: **The chemical biology of modular biosynthetic enzymes.** *Chem Soc Rev* 2009, **38**:2012-2045.
3. Hertweck C: **The biosynthetic logic of polyketide diversity.** *Angew Chem Int Ed* 2009, **48**:4688-4716.
4. Fischbach MA, Walsh CT: **Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms.** *Chem Rev* 2006, **106**:3468-3496.
5. Hari T, Labana P, Boileau M, Boddy CN: **An evolutionary model encompassing substrate specificity and reactivity of type I polyketide synthase thioesterases.** *ChemBioChem* 2014, **15**:2656-2661.
6. Walsh CT: **Insights into the chemical logic and enzymatic machinery of NRPS assembly lines.** *Nat Prod Rep* 2015. This review discusses the function of NRPS modular enzymes in assembling the structurally complex macrocyclic and heterocyclic scaffolds
7. Zhou Y, Prediger P, Dias LC, Murphy AC, Leadlay PF: **Macrolide formation by the thioesterase of a modular polyketide synthase.** *Angew Chem Int Ed* 2015, **127**:5321-5324.
8. Zhou Y, Murphy AC, Samborsky M, Prediger P, Dias LC, Leadlay PF: **Iterative mechanism of macrolide formation in the anticancer compound conglobatin.** *Chem Biol* 2015, **22**:745-754.
9. Kopp M, Marahel MA: **Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis.** *Nat. Prod. Rep.* 2007, **24**:735-749.
10. Hertweck C: **Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology.** *Trends Biochem Sci* 2015, **40**:189-199. This review discusses the importance of biosynthetic reactions to generate pharmaceutically important polyketide scaffolds and the role of enzymes in generating structural diversity
11. Piel J: **Biosynthesis of polyketides by trans-AT polyketide synthases.** *Nat Prod Rep* 2010, **27**:996-1047.
12. Calderone CT: **Isoprenoid-like alkylations in polyketide biosynthesis.** *Nat Prod Rep* 2008, **25**:845-853.
13. Haines AS, Dong X, Song Z, Farmer RCW, Hothersall J, Płoskoří E, Wattana-amorn P, Stephens ER, Yamada E *et al.*: **A conserved motif flags acyl carrier proteins for beta-branching in polyketide synthesis.** *Nat Chem Biol* 2013, **9**:685-692.
14. Hotta K, Chen X, Paton RS, Minami A, Li H, Swaminathan K, Mathews II, Watanabe K, Oikawa H, Houk KN *et al.*: **Enzymatic catalysis of anti-Baldwin ring closure in polyether biosynthesis.** *Nature* 2012, **483**:355-358.
15. Wong FT, Hotta K, Chen X, Fang M, Watanabe K, Kim C-Y: **Epoxide hydrolase-lasalocid A structure provides mechanistic insight into polyether natural product biosynthesis.** *J Am Chem Soc* 2014, **137**:86-89.
16. Gallimore AR, Stark CB, Bhatt A, Harvey BM, Demydchuk Y, Bolanos-Garcia V, Fowler DJ, Staunton J, Leadlay PF, Spencer JB: **Evidence for the role of the monB genes in polyether ring formation during monensin biosynthesis.** *Chem Biol* 2006, **13**:453-460.

17. Matilla MA, Stöckmann H, Leeper FJ, Salmund GP: **Bacterial biosynthetic gene clusters encoding the anti-cancer haterumalide class of molecules: biogenesis of the broad spectrum antifungal and anti-oomycete compound, oocydin A.** *J Biol Chem* 2012, **287**:39125-39138.
18. Bertin M, Vulpanovici A, Monroe EA, Korobeynikov A, Sherman DH, Gerwick L, Gerwick W: **The phormidolide biosynthetic gene cluster: a trans AT PKS pathway encoding a toxic macrocyclic polyketide.** *ChemBioChem* 2015, **17**:164-173.
19. Piel J, Butzke D, Fusetani N, Hui D, Platzer M, Wen G, Matsunaga S: **Exploring the chemistry of uncultivated bacterial symbionts: antitumor polyketides of the pederin family.** *J Nat Prod* 2005, **68**:472-479.
20. Pöplau P, Frank S, Morinaka BI, Piel J: **An enzymatic domain for the formation of cyclic ethers in complex polyketides.** *Angew Chem Int Ed* 2013, **52**:13215-13218.
21. Berkhan G, Hahn F: **A dehydratase domain in ambruticin biosynthesis displays additional activity as a pyran-forming cyclase.** *Angew Chem Int Ed* 2014, **53**:14240-14244.
22. Luhavaya H, Dias MV, Williams SR, Hong H, de Oliveira LG, Leadlay PF: **Enzymology of pyran ring A formation in salinomycin biosynthesis.** *Angew Chem Int Ed* 2015, **54**:13622-13625.
23. Li C, Roegel KE, Kelly WL: **Analysis of the indanomycin biosynthetic gene cluster from *Streptomyces antibioticus* NRRL 8167.** *ChemBioChem* 2009, **10**:1064-1072.
24. Partida-Martinez LP, Hertweck C: **Pathogenic fungus harbours endosymbiotic bacteria for toxin production.** *Nature* 2005, **437**:884-888.
25. Scherlach K, Busch B, Lackner G, Paszkowski U, Hertweck C: **Symbiotic cooperation in the biosynthesis of a phytotoxin.** *Angew Chem Int Ed* 2012, **51**:9615-9618.
26. Kusebauch B, Busch B, Scherlach K, Roth M, Hertweck C: **Polyketide-chain branching by an enzymatic michael addition.** *Angew Chem Int Ed* 2009, **48**:5001-5004.
27. Bretschneider T, Heim JB, Heine D, Winkler R, Busch B, Kusebauch B, Stehle T, Zocher G, Hertweck C: **Vinylogous chain branching catalysed by a dedicated polyketide synthase module.** *Nature* 2013, **502**:124-128.
- A new type of PKS branching module (KS-B-ACP) has been shown to mediate a complex Michael addition-lactonization reaction in the biosynthesis of the anticancer agent, rhizoxin.
28. Heine D, Sundaram S, Bretschneider T, Hertweck C: **Twofold polyketide branching by a stereoselective enzymatic Michael addition.** *Chem Commun* 2015, **51**:9872-9875.
29. Lim SK, Ju J, Zazopoulos E, Jiang H, Seo JW, Chen Y, Feng Z, Rajski SR, Farnet CM, Shen B: **Iso-Migrastatin, migrastatin, and dorrigocin production in *Streptomyces platensis* NRRL 18993 is governed by a single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase.** *J Biol Chem* 2009, **284**:29746-29756.
30. Wang B, Song Y, Luo M, Chen Q, Ma J, Huang H, Ju J: **Biosynthesis of 9-methylstreptimidone involves a new decarboxylative step for polyketide terminal diene formation.** *Org Lett* 2013, **15**:1278-1281.
31. Yin M, Yan Y, Lohman JR, Huang S-X, Ma M, Zhao G-R, Xu L-H, Xiang W, Shen B: **Cycloheximide and actiphenol production in *Streptomyces* sp YIM56141 governed by single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase.** *Org Lett* 2014, **16**:3072-3075.
32. Heine D, Bretschneider T, Sundaram S, Hertweck C: **Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles.** *Angew Chem Int Ed* 2014, **53**:11645-11649.
33. Maier T, Leibundgut M, Ban N: **The crystal structure of a mammalian fatty acid synthase.** *Science* 2008, **321**:1315-1322.
34. Sundaram S, Heine D, Hertweck C: **Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching.** *Nat Chem Biol* 2015, **11**:949-951.
35. Tahlan K, Jensen SE: **Origins of the β -lactam rings in natural products.** *J Antibiot* 2013, **66**:401-410.
36. Gaudelli NM, Long DH, Townsend CA: **β -Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis.** *Nature* 2015, **520**:383-387.
- The synthesis of lactam ring in nocardicin G has been shown to involve a condensation (C) domain to perform a novel mechanism in contrast to the oxidative strategy seen in beta-lactam antibiotics.
37. Gaudelli NM, Townsend CA: **Epimerization and substrate gating by a TE domain in β -lactam antibiotic biosynthesis.** *Nat Chem Biol* 2014, **10**:251-258.
38. Stegmann E, Frasch HJ, Wohlleben W: **Glycopeptide biosynthesis in the context of basic cellular functions.** *Curr Opin Microbiol* 2010, **13**:595-602.
39. Yim G, Thaker MN, Koteva K, Wright G: **Glycopeptide antibiotic biosynthesis.** *J Antibiot* 2014, **67**:31-41.
40. Süßmuth RD, Wohlleben W: **The biosynthesis of glycopeptide antibiotics — a model for complex, non-ribosomally synthesized, peptidic secondary metabolites.** *Appl Microbiol Biotechnol* 2004, **63**:344-350.
41. Hadatsch B, Butz D, Schmiederer T, Steudle J, Wohlleben W, Süßmuth R, Stegmann E: **The biosynthesis of teicoplanin-type glycopeptide antibiotics: assignment of P450 mono-oxygenases to side chain cyclizations of glycopeptide a47934.** *Chem Biol* 2007, **14**:1078-1089.
42. Haslinger K, Maximowitsch E, Brieke C, Koch A, Cryle MJ: **Cytochrome P450 OxyBte1 catalyzes the first phenolic coupling step in teicoplanin biosynthesis.** *ChemBioChem* 2014, **15**:2719-2728.
43. Haslinger K, Peschke M, Brieke C, Maximowitsch E, Cryle MJ: **X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis.** *Nature* 2015, **521**:105-109.
- A conserved X domain has been shown to recruit oxygenases to the NRPS-bound peptide, to perform side-chain crosslinking in the context of glycopeptide antibiotic synthesis. X domain resembles condensation domains and is essential to ensure interaction with oxygenases and conversion of the bound peptide to a mature aglycone
44. Haslinger K, Brieke C, Uhlmann S, Sieverling L, Süßmuth RD, Cryle MJ: **The structure of a transient complex of a nonribosomal peptide synthetase and a cytochrome P450 monooxygenase.** *Angew Chem Int Ed* 2014, **53**:8518-8522.
45. Brieke C, Peschke M, Haslinger K, Cryle MJ: **Sequential in vitro cyclization by cytochrome P450 enzymes of glycopeptide antibiotic precursors bearing the X-domain from nonribosomal peptide biosynthesis.** *Angew Chem Int Ed* 2015, **127**:15941-15945.
46. Arthur JC, Gharaibeh RZ, Mühlbauer M, Perez-Chanona E, Uronis JM, McCafferty J, Fodor AA, Jobin C: **Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer.** *Nat Commun* 2014, **5**:4724.
47. Reimer D, Pos KM, Thines M, Grün P, Bode HB: **A natural prodrug activation mechanism in nonribosomal peptide synthesis.** *Nat Chem Biol* 2011, **7**:888-890.
48. Vizcaino MI, Crawford JM: **The colibactin warhead crosslinks DNA.** *Nat Chem* 2015, **7**:411-417.
- The structure of precolibactin, an unstable metabolite from *E. coli* has been predicted and is shown to interact with DNA to induce damage by alkylation. Precolibactin has a spirobicyclic warhead that is synthesized when the intermediate is tethered to the assembly line.
49. Brotherton CA, Wilson M, Byrd G, Balskus EP: **Isolation of a metabolite from the pks island provides insights into colibactin biosynthesis and activity.** *Org Lett* 2015, **17**:1545-1548.
50. Bian X, Plaza A, Zhang Y, Müller R: **Two more pieces of the colibactin genotoxin puzzle from *Escherichia coli* show incorporation of an unusual 1-aminocyclopropanecarboxylic acid moiety.** *Chem Sci* 2015, **6**:3154-3160.
51. Li ZR, Li Y, Lai JY, Tang J, Wang B, Lu L, Zhu G, Wu X, Xu Y, Qian PY: **Critical intermediates reveal new biosynthetic events in the enigmatic colibactin pathway.** *ChemBioChem* 2015, **16**:1715-1719.

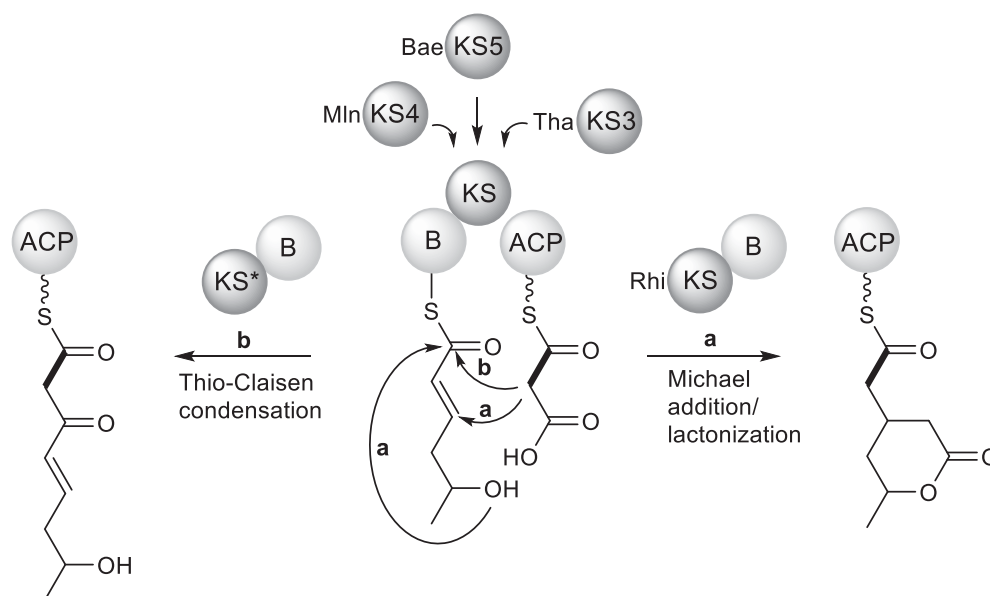
94 Biocatalysis and biotransformation

52. Gu L, Wang B, Kulkarni A, Geders TW, Grindberg RV, Gerwick L, Håkansson K, Wipf P, Smith JL, Gerwick WH *et al.*: **Metamorphic enzyme assembly in polyketide diversification**. *Nature* 2009, **459**:731-735.
53. Khare D, Wang B, Gu L, Razelun J, Sherman DH, Gerwick WH, Håkansson K, Smith JL: **Conformational switch triggered by alpha-ketoglutarate in a halogenase of curacin A biosynthesis**. *Proc Natl Acad Sci USA* 2010, **107**:14099-14104.
54. Vaillancourt FH, Yeh E, Vosburg DA, O'Connor SE, Walsh CT: **Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis**. *Nature* 2005, **436**: 1191-1194.
55. Hiratsuka T, Suzuki H, Kariya R, Seo T, Minami A, Oikawa H: **Biosynthesis of the structurally unique polycyclopropanated polyketide-nucleoside hybrid jawsamycin (FR-900848)**. *Angew Chem Int Ed* 2014, **53**:5423-5426.
56. Julien B, Tian Z-Q, Reid R, Reeves CD: **Analysis of the ambruticin and jerangolid gene clusters of *Sorangium cellulosum* reveals unusual mechanisms of polyketide biosynthesis**. *Chem Biol* 2006, **13**:1277-1286.
57. Ma SM, Li JW-H, Choi JW, Zhou H, Lee KM, Moorthie VA, Xie X, Kealey JT, Da Silva NA, Vederas JC *et al.*: **Complete reconstitution of a highly reducing iterative polyketide synthase**. *Science* 2009, **326**:589-592.
58. Fage CD, Isiorho EA, Liu Y, Wagner DT, Liu H-W, Keatinge-Clay AT: **The structure of SpnF, a standalone enzyme that catalyzes [4 + 2] cycloaddition**. *Nat Chem Biol* 2015, **11**: 256-258.
59. Hashimoto T, Hashimoto J, Teruya K, Hirano T, Shin-Ya K, Ikeda H, Liu H-W, Nishiyama M, Kuzuyama T: **Biosynthesis of versipelostatin: identification of an enzyme-catalyzed [4 + 2]-cycloaddition required for macrocyclization of spirotetronate-containing polyketides**. *J Am Chem Soc* 2015, **137**:572-575.
60. Tian Z, Sun P, Yan Y, Wu Z, Zheng Q, Zhou S, Zhang H, Yu F, Jia X, Chen D *et al.*: **An enzymatic [4 + 2] cyclization cascade creates the pentacyclic core of pyrroindomycins**. *Nat Chem Biol* 2015, **11**:259-265.
61. Sato M, Yagishita F, Mino T, Uchiyama N, Patel A, Chooi YH, Goda Y, Xu W, Noguchi H, Yamamoto T *et al.*: **Involvement of lipocalin-like CghA in decalin-forming stereoselective intramolecular [4 + 2] cycloaddition**. *ChemBioChem* 2015, **16**:2294-2298.
62. Jungmann K, Jansen R, Gerth K, Huch V, Krug D, Fenical W, Müller R: **Two of a kind – the biosynthetic pathways of chlorotoniil and anthracimycin**. *ACS Chem Biol* 2015, **10**: 2480-2490.
63. Alt S, Wilkinson B: **Biosynthesis of the novel macrolide antibiotic anthracimycin**. *ACS Chem Biol* 2015, **10**:2468-2479.

3.6 Manuscript E

Investigation of a non-canonical chain-branching ketosynthase by site-directed mutagenesis and domain swaps

Srividhya Sundaram, Ruth Bauer, Hak Joong Kim, Tawatchai Thongkongkaew; Christian Hertweck: Investigation of a non-canonical chain-branching ketosynthase by site-directed mutagenesis and domain swaps, *in preparation*, 2017.



Both the β -branching and lactonization reactions are catalyzed solely by the KS domain of the rhizoxin PKS. Site-directed mutagenesis revealed complex interplay among the critical conserved residues of the KS domain. The branching-KS substituted with a regular elongating-KS resulted in a linear polyketide product confirming the non-canonical catalytic role played by the KS.

Investigation of a non-canonical chain-branching ketosynthase by site-directed mutagenesis and domain swaps

Srividhya Sundaram¹, Ruth Bauer¹, Hak Joong Kim¹, Tawatchai Thongkongkaew¹, Christian Hertweck^{1,2,3*}

¹ Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), 07745 Jena, Germany

² Chair for Natural Product Chemistry, Friedrich Schiller University, 07743 Jena, Germany

³Lead Contact

*Correspondence: christian.hertweck@leibniz-hki.de

SUMMARY

Ketosynthase (KS) domains of modular type I polyketide synthases (PKSs) typically catalyze Claisen condensation of simple acyl and malonyl units to form linear chains that eventually make bioactive polyketides. The KS domain of the branching module of rhizoxin PKS installs a β -branch in the otherwise linear polyketide intermediate which is then cyclized to form the pharmacophoric δ -lactone moiety. To know the exact role of KS domain in the branching-cyclization sequence, we performed targeted mutagenesis of 10 of the conserved amino acids. Biochemical and kinetic analysis revealed that the C3228A and T3290Y mutations completely abolished the β -branch formation. The H3364A, H3404A S3141D, I3478F and S3409A mutants exhibited low turnover activity with attenuated rate for lactonization. Complementing the mutational analyses, we generated chimeras of the branching-KS domain with elongating-KSs and show that a canonical linear chain is formed. These results reveal the mechanistic details of a complex branching reaction catalyzed solely by a KS domain.

Keywords

Trans-AT PKSs, β -branching, KS domain, domain exchange, site-directed mutagenesis, biotransformation, enzyme kinetics

Highlights

- Polyketide β -branching as well as lactonization is catalyzed by the rhi KS domain.
- Site-specific mutagenesis of the KS reveals roles of critical conserved residues.
- A regular linear polyketide chain results from swapping branching-KS with elongating KSs.
- The results reveal complex interplay of the active site residues in the branching-cyclization reactions.

In Brief

Sundaram et al., investigates the mechanism of a branching-KS domain of the rhizoxin PKS and reports the role of critical conserved residues. The catalytic role of KS is also studied by whole domain exchange experiments.

INTRODUCTION

The extremely complex bioactive polyketides are assembled from simple acyl-CoA building blocks by giant multienzyme polyketide synthases (PKSs). Bacterial type I modular PKSs recruit domains to perform chain elongation followed by ketoreduction, dehydration, and enoylreduction that yield structurally diverse polyketides (Chan et al., 2009; Khosla et al., 2009). Typically, each domain within a module performs a single reaction. Therefore the number of modules can directly be translated to the expected product, a principle defined as colinearity (Hertweck, 2009). *Trans*-AT PKSs differ from this regular paradigm. These PKSs possess *trans*-acting acyltransferase (AT) and may contain skipped modules, irregular domain orders, inactive as well as dual functional domains and novel domains generating branches and rings (Helfrich and Piel, 2016; Piel, 2010). A single *trans*-AT PKS module minimally contains an acyl carrier protein (ACP) domain that shuttles starter/extender acyl-CoA units or intermediates and a ketosynthase (KS) domain that decarboxylates the extender unit and condenses with the intermediate. The KS domain is the most highly conserved of all the domains and contains a thiolase fold with $\alpha/\beta/\alpha/\beta/\alpha$ design. The active center of the KS domain comprises of a well conserved Cys-His-His catalytic triad. Crystal structures of KSs from *cis*- as well as *trans*-AT PKSs revealed clues about the reaction mechanism (Gay et al., 2014; Gay et al., 2016; Keatinge-Clay, 2012). Typically, the reactive Cys is the main nucleophile involved in the reaction and is first acylated by attack of the upstream acyl-ACP thioester intermediate. The nucleophilicity of Cys is enhanced by one of the histidines which might be important in activating the carboxylate of the extender unit attached to a downstream ACP. The other histidine recruits a well-placed water molecule to attack the activated carboxylate. This yields an enolate that attacks the acyl-KS thioester to elongate the polyketide chain. The reaction thus happens in two steps: decarboxylation of extender unit attached to ACP and carbon-carbon bond formation – both catalyzed by the KS domain (Keatinge-Clay, 2012). The mechanism of a typical KS domain is exemplified by the module 10 of thailandamide PKS (**Figure 1A**) (Ishida et al., 2012).

Some of the KS domains from *trans*-AT PKSs perform novel activities such as β -branching (Hertweck, 2015). One example for such a non-canonical KS domain from the phytotoxic rhizoxin *trans*-AT PKS is part of a branching module that also contains B and ACP domains. This module installs the pharmacophoric δ -lactone moiety of the rhizoxin that is vital for its antimetabolic activity (Kusebauch et al., 2009; Schmitt et al., 2008). The δ -lactone is formed by the addition of an ACP-bound malonyl unit to the α,β -unsaturated, unbranched thioester polyketide intermediate. As a result of this Michael addition, a β -branch is generated, which is then cyclized (**Figure 1B**) (Bretschneider et al., 2013). Analogous to this, the glutarimide moiety of 9-methyl streptimidone, isomigrastatin and cycloheximide (Ju et al., 2005; Wang et al., 2013; Yin et al., 2014) is also proposed to be generated by a similar KS domain.

Although the cryptic B domain shares structural homology with dehydratase (DH) domains, it has a distorted active site with a His to Tyr mutation. Mutational analysis of the conserved Asp of the B domain revealed that this domain might only have a structural role in the branching reaction (Bretschneider et al., 2013). Retaining of the branching activity in the B domain substitution experiments with structurally related X (Wang et al., 2013) and DH

domains confirmed that B and X domains maintains the structural integrity while the KS domain solely catalyzes the branching-cyclization sequence (Sundaram et al., 2015). Point mutations of the active site Cys-His-His residues of KS domain resulted in no branching activity confirming the direct role played by these residues (Bretschneider et al., 2013). Earlier mutational analyses on the canonical KS domains demonstrated that additional residues might be important for the substrate selectivity and specificity (Brautaset et al., 2003; Gay et al., 2014). It remains obscure how the branching KS have evolved to catalyze the novel β -branching despite the presence of a well conserved catalytic triad.

In this study, we expand the knowledge of the active site residues of the branching-KS as well as sought for other residues that might be important for either Michael addition or lactonization or both. We also generated chimeric KS domains and show that elongating KSs appended to B domain form a regular linear product.

RESULTS

Site-specific mutational analysis of KS domain

To decipher the exact role and mechanism of KS domain in chain branching, we probed into the primary sequence of the KS domain. The KS domain in the rhizoxin PKS branching module (Rhi-KS14) falls in clade IX category in the cladogram that corresponds to olefinic substrates (Nguyen et al., 2008). We compared the amino acid sequence of Rhi-KS14 with other KSs (both branching and elongating) from clade IX. 10 positions in the branching KSs differ considerably when compared against other KSs that mediate Claisen condensation (**Figure 2A and S1**). Site-directed mutagenesis was performed on these residues where they were changed to those that correspond in the other KSs. In case of S3232, it was changed to Leu as well as Ala. **Figure 2B** shows the conserved residues of the Rhi-KS14 (PDB code: 4kc5) superimposed on the bacillaene KS5 structure (PDB code: 5erb) (Gay et al., 2016). In total, 11 single mutations were introduced in the plasmid pTB57 (Bretschneider et al., 2013). To know the cumulative effect of all the selected amino acids, a synthetic gene (*KS'-B*) (MWG Eurofins) where all the 9 residues were mutated (except T3290). All the proteins were heterologously produced in *E. coli* BL21 (DE3). **Figure 2C** shows the purified proteins together with the active site mutants (C3228A, H3364A, H3404A) on a 10% SDS gel. The secondary structure of all the recombinant proteins when compared against the KS-B showed that there are no significant differences (except for *KS'-B*) (**Figure 2D and 2E**).

In vitro biotransformation with KS mutants

In the previous study, to decipher the course of the reaction, a SNAC thioester that lacks the δ -hydroxy group **1** was used to trap the KSB and ACP domains in place (**Figure 3A**). This resulted in a covalent intermediate bound to both the domains (**2**) which was then visualized on the SDS gel (Bretschneider et al., 2013). To identify any additional residues that might be involved in Michael addition, the *in vitro* assay was initially performed with **1**. The assays were performed at 23 °C for 16 h. All the mutants except C3228A, T3290Y and the *KS'-B*, retained the vinylogous activity (**Figure 3B**). It is intriguing that both the histidine residues are still capable of undergoing the vinylogous addition, contrary to the previous report (Bretschneider et al., 2013). In the previous study, the standard assay timing was only 60 min.

It is plausible that the C-C bond formation takes place at a much slower rate when histidines are mutated.

The *in vitro* reconstitution of the branching reaction was then performed with the native substrate mimic **3** to monitor the formation of the lactone **4** (**Figure 4A**). H3097A, S3117G, S3141D, S3232L, S3232A, C3234A, T3267L, L3406E, S3409A and I3478F (**Figure 4B, 4C, 4D, 4F, 4G, 4H, 4I, 4M, 4N and 4O**) mutations did not significantly affect the branching activity. In all these cases, the δ -lactone **4** was observed in comparable amounts to KS-B (**Figure 4Q**). Corroborating the SDS gel-shift experiments, the mutants C3228A, T3290Y and KS'-B did not result in **4** even when the corresponding enzymes were tested at 10 times the normal concentration for prolonged timing (16 h) (**Figure 4E, 4J and 4P**). Both C3228A and T3290Y mutants however still retained malonyl-CoA decarboxylation activity (data not shown). For the mutants H3364A and H3404A, increasing the enzyme concentration and the duration of the assay also increased the amount of product formed (**Figure 4K and 4L**). It is possible that both these residues complement each other in the branching reaction. The KS'B with 9 mutations had attenuated malonyl CoA decarboxylation activity and the product was not observed. During protein production and purification, unlike other mutants, KS'-B resulted mostly in the formation of inclusion bodies. Even after several rounds of optimization during protein purification, it was only possible to get very low amount of soluble protein. The CD data of KS-B' also exhibited slight differences in the near-UV region (**Figure 2D**). It is possible that the introduction of several mutations rendered the protein inactive. Therefore, it is difficult to conclude whether the inactivity of KS-B' is due to structural or functional differences.

LC-ESI-MS of intact proteins

To complement the LC-MS analysis, LC-ESI-MS of the intact ACP-bound proteins was also performed for KS-B, H3364A and H3404A. For the ease of comparison, the assays were performed at 10x enzyme concentration for 16 h. The PPant-ejected product was analyzed in the first quadrupole. The lactone-bound PPant arm was observed for all the samples tested. However, compared to the wild type KS-B (**Figure S2**), the intensities of the mass fragments were approximately 10-fold lower in H3364A and H3404A (**Figure S3 and S4**).

Kinetic analysis of lactone formation

The preference of the mutants for the native substrate mimic **3** was further analyzed by performing kinetic experiments. The mutants C3228A and T3290Y did not exhibit any turnover activity in the *in vitro* biotransformation and therefore were excluded from the kinetic analysis. The assay was performed at saturating concentration of **3** (1 mM) and increased enzyme concentrations. Increasing the concentrations of KSB and ACP to 4.5 μ M and 300 μ M respectively aided in comparing the activities of the mutants. **Figure 5** shows the cumulative data from three independent experiments. The mutants colored in black all have comparable activities and reached product saturation from approximately 45 min. S3141D, I3478F and S3409A resulted in the accumulation of decarboxylated malonyl-ACP (data not shown) whereas the lactonization was less pronounced. The H3364A, H3404A and KS'-B mutants exhibited negligible activities within the time limit specified (75 min). It is worth

mentioning that the decarboxylation activity of H3404 is still affected even after 16 h assay, suggesting that this residue is very important for the decarboxylation of malonyl-ACP. A double mutant (H3364A and H3404A) did not result in **4** proving that they are indispensable for decarboxylation.

Construction of KS-B chimeric modules

Even though the mutational studies shed some information on the kinetics of the branching reaction, it was difficult to assign specific roles to individual residues. It is possible that two or more residues work together in the branching reaction. To complement this study, KS whole domain exchange experiments were performed. The branching KS when exchanged with a non-branching KS should ideally lead to a linear product in place of the branched one. The KSs, KS5 from bacillaene (Bae) PKS (Butcher et al., 2007), KS4, KS5 and KS9 from macrolactin (Mln) PKS (Chen et al., 2007), and KS3, KS12 from thailandamide (Tha) PKS (Ishida et al., 2012) belong to clade IX similar to the Rhi-KS14. These KSs accept α , β -unbranched, unsaturated substrates (**Figure 6A**) (Nguyen et al., 2008) and yield linear polyketide intermediates. Of the 6 KSs, the Bae-KS5 has been reported to show tolerance towards shorter acyl-thioesters (Jenner et al., 2013). The genes corresponding to these KSs were amplified from the genomic DNA of *Bacillus amyloliquefacians* (for BaeKS5, MlnKS4, MlnKS5 and MlnKS9) and *Burkholderia thailandensis* (for ThaKS3 and ThaKS12). The linker region between the Rhi-KS14 and the B domain in the native KS-B didomain was kept intact since this region is also important for structural stability (Sundaram et al., 2015). After DNA sequence confirmation, the KS genes were fused to the N-terminus of the linker and the B domain and expressed in *E. coli* BL21 (DE3). All the chimeric proteins were produced as KS-B didomains (**Figure 6B**). Owing to the high GC content of *B. amyloliquefacians*, except for MlnKS4-B, the solubility of the chimeric proteins drastically reduced. To avoid further sample loss, only a single step of affinity purification was done for all the proteins (**Figure 6C**). However, except MlnKS4-B, none of the enzymes were stable when stored at 4 °C. Domain swapping usually results in insoluble or inactive proteins as the interdomain communication is compromised (Weissman, 2014).

A recent study reported the presence of a three-helix LINKS (Laterally-Interacting Ketosynthase Sequence) motif located C-terminal to some of the *trans*-AT KSs. This ~40 residue region contains a mixture of charged and hydrophobic residues that zip together additional copies of KSs within the assembly line (Gay et al., 2016). Some of the KSs do not contain LINKS region (for example, Rhi-KS14) whereas in some, the conservation of the LINKS region is poor. In the BaeKS5, the LINKS region is well established whereas in ThaKS3, ThaKS12, structurally uncharacterized mode of interactions have been proposed to occur (Gay et al., 2016). We reasoned that including the LINKS in the chimeric constructs would improve the KS dimer stability. Therefore, a new cloning strategy was planned to substitute this LINKS region for the linker region that is originally present between Rhi-KS14 and the B domain (**Figure 6B**). The resulting hybrids are shown as BaeKS5*-B, ThaKS3*-B and ThaKS12*-B. Including the LINKS region specific to each of the above mentioned KSs greatly improved the solubility of the protein (**Figure 6C**). The constructs were verified by

MALDI-TOF followed by tryptic digestion (**Figure S5**). Owing to poor storage stability, the CD spectra were recorded only for MlnKS4-B, BaeKS5*-B and ThaKS3*-B (**Figure 6D**).

***In vitro* biotransformation with KS chimeras**

The activity of all the chimeric KS-B was checked by *in vitro* reconstitution of the branching reaction with **3**. **Figure 7A** shows the possible reactions from the assay with the hybrids. If the incoming malonyl unit is added to the first (route b) rather than the third carbon atom (route a) of the polyketide intermediate bound to KS, a regular Claisen condensation would occur. This will lead to a linear product **5** in place of the branched one (**4**). Of all the 9 hybrids tested, only 3 (MlnKS4-B, BaeKS5*-B and ThaKS3*-B) resulted in a new compound which has the same mass as **4** with a retention time of ~7.2 min. The rest of the hybrids resulted in no turnover of **3**. **Figure 7B, 7C and 7D** shows the extracted ion chromatograms of $[M+H]^+$ 173.0808 for these three hybrids. **Figure 7E** shows the difference in retention times of the lactone **4** and the probable linear product **5**. MS/MS analysis of the new product (with MlnKS4-B as an example) was also performed along with the wild type KS-B as the positive control (**Figure S6**). As shown, the mass fragments were different for **4** and **5**.

DISCUSSION

The modular structure of the multienzyme polyketide synthases (PKSs) is a very efficient platform to modify the domains to produce analogues of desired structure as well as to decipher the function of enigmatic domains (Weissman, 2016). Domain modifications have been successfully performed with specific reductive, KR (Kellenberger et al., 2008) and AT domains of DEBS (Walker et al., 2013). The exchanging and engineering experiments are not only limited to individual domains but also to the whole modules as in the case of avermectin, aureothin, neo-aureothin and luteoreticulin PKSs (Marsden et al., 1998; Sugimoto et al., 2014; Sugimoto et al., 2015). Modifications on KS domain, which performs the core carbon-carbon bond formation is a very challenging feat, mainly because of the inability of the chimeric KS to maintain domain-domain interfaces. Many of the silent interfaces have been shown to activate by rational replacement of the KS domain of DEBS PKS (Chandran et al., 2006). Most of the studies on KS domain were directed towards specific mutations of individual residues to broaden the substrate specificity (Gay et al., 2014; Jenner et al., 2013; Murphy et al., 2016) as well as to understand the role of active site residues (in case of *cis*-AT PKS) (Robbins et al., 2016).

This study analyses the impact of individual amino acid residues of a non-canonical branching KS on the decarboxylation of malonyl-ACP, Michael addition and the subsequent lactonization. Although the versatility of the KS cannot be assigned to any specific residue, this study revealed insights into the conserved residues. The inactivity of C3228A mutation is not surprising (**Figure 4E**), which is expected by the absence of exceptionally nucleophilic C3228 residue needed for binding the polyketide intermediate. The H3364 enhances decarboxylation of malonyl-ACP by deprotonating a water molecule to attack the malonate. The H3364A mutant exhibited slow turnover in terms of the formation of a decarboxylated malonate as well as the subsequent lactone (**4**) formation (**Figure 4K**). A previous report on

FabF also suggested that this histidine functions to enhance the overall rate of the reaction (Zhang et al., 2006) and thus may not be an absolute necessity for the chain branching. The H3404 activates the malonyl carboxylate by stabilising the enolate. Reduction in the carbon-carbon bond formation as well as the lactone (**4**) (**Figure 3B and 4L**) supports that H3404 plays a major role in all the steps of the reaction. Neither of the histidines has marked influence on the ionization of C3228 residue. Similar to C3228A, the completely inactive T3290Y mutant (**Figure 4J**) suggests that this residue might interact with one of the histidines to promote C-C bond formation. The analogous Phe residue in FabB and FabF enzymes is reported to promote decarboxylation and the formation of C2 carbanion (Huang et al., 1998; Olsen et al., 1999; Price et al., 2003). We propose that the presence of T3290 promotes attack of the malonyl extender unit to the β -carbon of the polyketide intermediate and provides space for the formation of a lactone moiety. However, both C3228A and T3290Y mutants still functioned as an effective malonyl-CoA decarboxylase (data not shown). The reduction in the activity of S3409A mutant is intriguing (**Figure 5**) and we propose that this residue might promote the final lactonization step.

In this study, we also analyze the role of the entire KS domain by constructing chimeric KS*-B didomains. Insolubility of chimeric domains has been a major challenge in domain or module swapping experiments mainly due to the disruption of the interdomain linkers (Sun et al., 2015; Weissman, 2016). Interdomain contacts are critical for maintaining the functional state of the module as revealed by the cryo-electron microscopy of module 5 of pikromycin PKS (Dutta et al., 2014). Here, although the LINKS region is not conserved among the KSs (Gay et al., 2016), we exploited this region to generate soluble and active KS-B chimeras (**Figure 6C**). Although the turnover rate is slow, production of a canonical linear product **5** with the KS-B chimeras establishes the complex role of the Rhi KS in the Michael addition/cyclization sequence.

To summarize, from a series of functional analyses, we provide multiple evidences that the KS domain alone catalyzes the Michael addition and cyclization. We also verified the residues that play critical role at each step of the branching reaction. This KS-mediated branching reaction is very significant from a synthetic biology perspective.

SIGNIFICANCE

Novel domains/modules from polyketide synthases are emerging as effective biocatalysts to produce non-natural variants. The KS domain of the rhizoxin PKS is highly versatile for it produces the pharmacophoric δ -lactone moiety by a non-canonical polyketide β -branching followed by lactonization. Similar β -branching-KS domains are observed to take part in the formation of glutarimide group in cycloheximide-type of antibiotics. Understanding the mechanism of such KS domains is imperative towards extended synthetic biology approaches. In this study, the minimally invasive site-specific mutational analyses of the residues conserved in branching KSs but different in elongating-KSs reveal that the branching-lactonization reaction is more complex than expected. Insights gained from the mutational

analyses and KS domain exchange experiments might enlighten the mechanistic details of the branching reaction.

EXPERIMENTAL PROCEDURES

General

Chemicals were purchased from Roth and Sigma-Aldrich. MALDI measurements were done using an UltrafleXtreme MALDI TOF/TOF (Bruker). High resolution mass measurements were performed using an Exactive Orbitrap High Performance Benchtop HRMS with an electrospray ion source and an Accela HPLC system (Thermo Scientific). LC-MS/MS experiments were performed using a QExactive equipment with Accela HPLC system (Thermo Scientific). LC-ESI-MS analysis of intact proteins was performed using ZORBAX 300SB-CN column in a ThermoFinnigan Surveyor LC equipment with a LTQ velos MS attached running on an electrospray ion source. The spectra from LC-ESI-MS measurements were deconvoluted using ProMass software (Thermo Scientific). NMR measurements were performed on a Bruker AVANCE II 300, AVANCE III 500 or 600 MHz spectrometer, with a Bruker cryo platform. Circular dichroism (CD) measurements were performed using a J815 spectropolarimeter (Jasco). Site-directed mutagenesis was performed using Quikchange II XL kit (Stratagene). Identity of all recombinant proteins was confirmed using SDS-PAGE and tryptic fingerprinting followed by MALDI. Crystal structures were visualized using Pymol.

Cloning experiments

Mutational procedures

Primers for point mutations were designed based on the Quikchange II XL mutagenesis kit manual (Stratagene) and a previous report (Zheng et al., 2004). The plasmid pTB57 that expresses RhiE KS-B under the control of T7 promoter (Bretschneider et al., 2013) was used as template. For GC-rich primers, up to 6% DMSO was used to prevent non-specific primer annealing. The primers used for generating all the point mutations are listed in **table 1**. After PCR amplification and inactivation of any template DNA by DpnI, 1/10th the sample volume of 3 M sodium acetate (pH 5.2) was added to precipitate the amplified product. The precipitate was washed with 2.5 volumes of 100% ice-cold ethanol. After a brief centrifugation, the DNA pellet was further washed with 70% ethanol. The enriched plasmid was then directly used for transformation of *E. coli* Top10 cells by electroporation. After confirming the mutants by DNA sequencing (MWG Eurofins), the plasmids were introduced into *E. coli* BL21 (DE3) (New England Biolabs). The double mutant (H3364A and H3404A) was generated using the primers already reported (Bretschneider et al., 2013). The synthetic gene with 9 mutations (H3097A, S3117G, S3141D, S3232L, C3234A, T3267L, L3406E, S3409A and I3478F) was purchased from MWG Eurofins. The gene was directly cloned in the vector pHis8-3 and expressed in *E. coli* BL21 (DE3) (New England Biolabs).

Primers	Sequence
H3097A_fw	GCGACTTTCCGGCAGACCGGTGGGACGCCAGCAAAATCTATTAC
H3097A_rv	GTAATAGATTTTGTCTGGCGTCCACCGGTCTGCCGGAAAGTTCGC
S3117G_fw	GTGCTAGGCAAGACCACCTGCATCAATGGCGGGTTCATCAAGGAC
S3117G_rv	GTCCTTGATGAACCCGCCATTGATGCAGGTGGTCTTGCCCTAGCAC
S3141D_fw	GCCCAAGGTCTATGCCGACCATATGGACCAGAGGTTCAGACTG
S3141D_rv	CAGTCTGACCTCTGGGTCCATATGGTCGGCATAGACCTTGGGC
S3232L_fw	CGATACCATGTGCTCGTCCCTCACCTGTATCCACACCGCTTG
S3232L_rv	CAAGCGGTGTGGATACAGGTGAGGGACGACGAGCACATGGTATCG
S3232A_fw	CGATACCATGTGCTCGTCCGCTACCTGTATCCACACCGCTTG
S3232A_rv	CAAGCGGTGTGGATACAGGTAGCGGACGACGAGCACATGGTATCG
C3234A_fw	GCTCGTCGTCCTCGACCGCCATCCACACCGCTTGTC
C3234A_rv	GACAAGCGGTGTGGATGGCGGTTCGAGGACGACGAGC
T3267L_fw	CACCCGTACACCACGGTGAATCTCTCCCAAGGCAATTCACATCG
T3267L_rv	CGATGTGAAATTGCCTTGGGAGAGATTACCGTGGTGTACGGGTG
T3290Y_fw	CTACGGGGTGGGTGCTGATGGTTACGTGATTGGTGAAGGCATCGG
T3290Y_rv	CCGATGCCTTCACCAATCACGTAACCATCAGCACCCACCCCGTAG
L3406E_fw	TCGAATATCGGCCACCTAGAGGCGGCTTCGGGCATTGGC
L3406E_rv	GCCAATGCCCGAAGCCGCCCTCTAGGTGGCCGATATTCGA
S3409A_fw	GCGGCTGCCGGCATTGCCGGCTTGACTAAAACCTTGCTG
S3409A_rv	CAAGCCGGCAATGCCGGCAGCCGCAATAGGTGGC
I3478F_fw	GGGGCTTACCTCGTTCGCTGCCGGCGGCATGAAC
I3478F_rv	GTTTCATGCCCGGCAGCGAACGAGGTAAGCCCC

Table 1 Primers for generating point mutations in pTB57.

Cloning of KS-B chimeras

To facilitate cloning of the genes for KSs, either KpnI or NdeI site was introduced in the N-terminus by PCR amplification of region corresponding to linker-to-B and B domains. This region was cloned in pHis 8-3. Separately, for cloning KSs with its corresponding LINKS motif, a standalone B domain (without the KS to B linker region) was cloned in pHis 8-3 vector. The genes for corresponding KSs were amplified from the genomic DNA of *Bacillus amyloliquefacians* (DSMZ) (for BaeKS5, BaeKS5*, MlnKS4, MlnKS5, MlnKS9) or *Burkholderia thailandensis* (for ThaKS3, ThaKS3*, ThaKS12, ThaKS12*). The primers and the region used for amplification are listed in **table 2**. The corresponding BamHI - KpnI, BamHI - NdeI or NcoI - NdeI digested KS fragments were cloned to the N-terminus of linker-to-B (without LINKS) or B domain (with LINKS). After sequence confirmation (MWG Eurofins), all recombinant plasmids were heterologously expressed in *E. coli* BL21 (DE3).

Primers	Sequence	Accession code	Protein region
Linker B_KpnI_fw	GGGGTACCTACCCAGAGCCAGCGGAC	CBW75249.1	3658 - 4133
Linker B_NdeI_fw	GGAATTCCATATGTACCCAGAGCCAGCGGAC		
Linker B_EcoRI_rv	CGGAATTCCTATGTCCCAGTCAGTAAGG		
B_KpnI_fw	GGGGTACCATCGTCAATCCGCTGATG	CBW75249.1	3822 - 4133
Bae KS5_BamHI_fw	CGGGATCCGAAGATATCGCGATAATCGG	YP_001421293.1	1893 - 2333
Bae KS5_KpnI_rv	GGGGTACCGGAATATTCTTCGATGACG		
Mln KS4_BamHI_fw	CGGGATCCGAGGATATTGCCATTATCGG	YP_001421029.1	660 - 1086
Mln KS4_NdeI_rv	GGAATTCCATATGTTGATATTCTTCCAC AATGACATGG		
Mln KS5_BamHI_fw	CGGGATCCAATGATATCGCGATTATCGG	YP_008420824	665 - 1084
Mln KS5_KpnI_rv	GGGGTACCCGCAAATTCTTCTAAAATAA CATGGGC		
Mln KS9_NcoI_fw	CATGCCATGGGAGGACATTGCCATTATCGG	YP_001421032	1018 - 1441
Mln KS9_NdeI_rv	GGAATTCCATATGCTCATAACGCTTCCAATA TAAGATGCGC		
Tha KS3_BamHI_fw	CGGGATCCGCCGCGACGGCGGACGCGATC	WP_011401356.1	3003 - 3441
Tha KS3_KpnI_rv	GGGGTACCGTCGCGGCCGTCGCGCTCGTC		
Tha KS12_BamHI_fw	CGGGATCCGCGGCCGCGTTCGCGGCG	WP_009900597.1	15 - 464
Tha KS12_KpnI_rv	GGGGTACCCGCGCCCCGCTCCGCGC		
Bae KS5*_BamHI_fw	CGGGATCCGCGCAAGTCCGTCCAGACACGCG	YP_001421293.1	1877 - 2513
Bae KS5*_KpnI_fw	GGGGTACCGGCCGGTCTTTTGACAGCT GCATCAGC		
Tha KS3*_BamHI_fw	CGGGATCCGCCGCGACGGCGGACGCGATC	WP_011401356.1	3003 - 3631
Tha KS3*_KpnI_rv	GGGGTACCGAATCCGAGCGGTTTCG CATCCAGCAT		
Tha KS12*_BamHI_fw	CGGGATCCGCGGCCGCGTTCGCGGCG	WP_009900597.1	15 - 635
Tha KS12*_KpnI_rv	GGGGTACCCACACCGGCCGCTTCGCCAGTTG		

Table 2 Primers for generating KS-B chimeras. ORF – open reading frame. The genes corresponding to the ORFs were amplified.

Protein production and purification

E. coli BL21 transformants bearing the genes for KS-B mutants and chimeras were cultivated in LB medium at 37 °C or 27 °C. After A_{600nm} reached ~0.4 - 0.5, the cells were induced with 0.1 mM IPTG at 15 °C for 16 - 20 h. The cell pellet was dissolved (10 mL buffer/g pellet) in 20 mM Tris buffer pH 7.5 containing 0.2 M NaCl and 50 mM imidazole. After disrupting the cells by sonication for 4 cycles for 3 min at 20% power (Sonotrode), the cells were centrifuged at 9,000 g at 4 °C for 30 min. The lysed supernatant was then loaded onto a Ni-NTA column (Macherey Nagel) connected to an FPLC machine. Proteins were eluted using the same buffer with 0.5 M imidazole. For KS-B mutants, the resulting protein was diluted twice with 20 mM Tris (pH 7.5) and purified further by an ion-exchange column (5 mL HiTrap Q HP, GE Healthcare). Proteins were eluted over 30 column volumes of 20 mM Tris

(pH 7.5) and 1 M NaCl, and the fractions containing target proteins were concentrated using Amicon columns (MWCO 10 kDa – Millipore). For KS-B chimeras, the ion-exchange chromatography was not performed to prevent further sample loss. All the purified proteins were stored at 4 °C for *in vitro* assays. The mutants were stable when stored at –20 °C supplemented with 50% glycerol while the chimeras regularly needed fresh preparation. The procedures for the production of ACP, wild type KS-B and AT were as described previously (Bretschneider et al., 2013).

***In vitro* reconstitution of the branching reaction**

The *in vitro* assay for the reconstitution of branching reaction was performed as described previously (Bretschneider et al., 2013) with minor modifications (specified separately in the results section). Briefly, in a 40 µL reaction, 167 µM ACP, 3 µM or ~30 µM KS-B, mutants and chimeras, 0.2 µM RhiG, 750 µM malonyl-CoA and 1,000 µM of **1** or **3** were incubated in 20 mM Tris buffer (pH 7.0) at 23 °C with shaking at 400 rpm for 2 h or 16 h. Analysis of the reaction by MALDI (Bruker) was done by spotting 2.5 µL of the assay mix with 1 µL of 100 µM of 2', 5' -dihydroxyacetophenone matrix. The MALDI-TOF was operated in positive linear mode. Data acquisition was performed using flexControl 3.3 in the range of 5–35 kDa with 1000 Hz laser frequency, 60% laser power and 500 shots at each spot. Analysis of the data was executed in flexAnalysis 3.3. For HRMS analysis, the reaction was stopped by adding equal volume of methanol. After a brief spin to remove precipitated proteins, the supernatant was dried using an upstream N₂ flow and the residue suspended in methanol was used for LC-HRMS analysis (Exactive). HPLC conditions: C18 column (Betasil C18 3 µm, 150 x 2.1 mm) and gradient elution (MeCN/0.1% (v/v) HCOOH (H₂O) 5/95 for 1 min, gradient up to 98/2 in 15 min and 98/2 for 3 min; flow rate – 0.2 mL/min; injection volume – 3–5 µL).

LC-ESI-MS and LC-MS/MS measurements

The branching reaction was also analyzed by LC-ESI-MS using a ZORBAX 300SB-CN column in positive mode. For this, in a 120 µL reaction, 167 µM ACP, ~30 µM KS-B, H3364A or H3404A, 0.2 µM RhiG, 750 µM malonyl-CoA and 1,000 µM of **3** were incubated in 20 mM Tris buffer (pH 7.0) at 23 °C with shaking at 400 rpm for 16 h. The assay mix was then desalted using ZebaTM spin desalting columns (7K MWCO, Thermo Scientific) followed by elution of the desalted sample in water. The components were separated using a gradient of 30 – 98% B for 26 min with a flow rate of 0.6 mL min⁻¹. MS/MS fragmentation of the selected ion fragments was also performed with the same instrument. The spectra from LC-ESI-MS measurements were deconvoluted using ProMass software (Thermo Scientific). For MS/MS analysis, Accucore C18 column (100 x 2.1 mm) was used in a QExactive operating in positive mode.

Kinetic studies

For the determination of kinetic parameters for KS-B mutants, concentration of the substrate **3** was maintained at 1,000 µM and the reaction was performed from 0–75 min. To warrant comparison of the rate of the reaction, higher concentrations of ACP (300 µM) and KS-B mutants (4.5 µM) were used. Samples were taken at regular intervals and the peak intensity of

the product attached to the ACP was analyzed by MALDI-TOF. The molarity of product formed was calculated from three independent experiments.

Secondary structure determination by CD

The secondary structures of all the proteins were determined at 20 °C. Wavelength scans were recorded for the far-UV region between 250 and 190 nm at a speed of 10 nm min⁻¹, a band width of 1 nm and accumulations of 5 spectra using a 1 mm quartz cuvette (Hellma). For all measurements, a concentration of 0.1 mg mL⁻¹ protein was used.

SUPPLEMENTAL INFORMATION

Supplemental information includes supplemental experimental procedures, figures and all relevant spectra.

AUTHOR CONTRIBUTIONS

S.S. and C.H. designed experiments. S.S. and R.B. performed genetic and biochemical experiments. S.S. analysed data. H.J.K. synthesized reference compound. T.T. synthesized substrates for biotransformation experiments. S.S. and C.H. wrote the manuscript.

ACKNOWLEDGEMENTS

We thank M.Poetsch and T.Kindell for MALDI-MS measurements and A.Perner for HRMS and ESI-MS measurements. We are grateful for financial support by the International Leibniz Research School (ILRS to S.S.) and the Deutscher Akademischer Austauschdienst (DAAD to T.T.).

The authors declare no conflicts of interest.

REFERENCES

- Brautaset, T., Borgos, S.E., Sletta, H., Ellingsen, T.E., and Zotchev, S.B. (2003). Site-specific mutagenesis and domain substitutions in the loading module of the nystatin polyketide synthase, and their effects on nystatin biosynthesis in *Streptomyces noursei*. *J. Biol. Chem.* 278, 14913-14919.
- Bretschneider, T., Heim, J.B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013). Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* 502, 124-128.
- Butcher, R.A., Schroeder, F.C., Fischbach, M.A., Straight, P.D., Kolter, R., Walsh, C.T., and Clardy, J. (2007). The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1506-1509.

- Chan, Y.A., Podevels, A.M., Kevany, B.M., and Thomas, M.G. (2009). Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep.* 26, 90-114.
- Chandran, S.S., Menzella, H.G., Carney, J.R., and Santi, D.V. (2006). Activating hybrid modular interfaces in synthetic polyketide synthases by cassette replacement of ketosynthase domains. *Chem. Biol.* 13, 469-474.
- Chen, X.H., Koumoutsis, A., Scholz, R., Eisenreich, A., Schneider, K., Heinemeyer, I., Morgenstern, B., Voss, B., Hess, W.R., and Reva, O. (2007). Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Nat. Biotechnol.* 25, 1007-1014.
- Dutta, S., Whicher, J.R., Hansen, D.A., Hale, W.A., Chemler, J.A., Congdon, G.R., Narayan, A.R., Håkansson, K., Sherman, D.H., and Smith, J.L. (2014). Structure of a modular polyketide synthase. *Nature* 510, 512-517.
- Gay, D.C., Gay, G., Axelrod, A.J., Jenner, M., Kohlhaas, C., Kampa, A., Oldham, N.J., Piel, J., and Keatinge-Clay, A.T. (2014). A close look at a ketosynthase from a *trans*-acyltransferase modular polyketide synthase. *Structure* 22, 444-451.
- Gay, D.C., Wagner, D.T., Meinke, J.L., Zogzas, C.E., Gay, G.R., and Keatinge-Clay, A.T. (2016). The LINKS motif zippers *trans*-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *J. Struct. Biol.* 193, 196-205.
- Helfrich, E.J., and Piel, J. (2016). Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* 33, 231-316.
- Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* 48, 4688-4716.
- Hertweck, C. (2015). Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem. Sci.* 40, 189-199.
- Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G., and Lindqvist, Y. (1998). Crystal structure of β -ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes. *EMBO J.* 17, 1183-1191.
- Ishida, K., Lincke, T., and Hertweck, C. (2012). Assembly and absolute configuration of short-lived polyketides from *Burkholderia thailandensis*. *Angew. Chem. Int. Ed.* 51, 5470-5474.
- Jenner, M., Frank, S., Kampa, A., Kohlhaas, C., Pöplau, P., Briggs, G.S., Piel, J., and Oldham, N.J. (2013). Substrate specificity in ketosynthase domains from *trans*-AT polyketide synthases. *Angew. Chem. Int. Ed.* 125, 1181-1185.
- Ju, J., Lim, S.-K., Jiang, H., Seo, J.-W., and Shen, B. (2005). Iso-Migrastatin congeners from *Streptomyces platensis* and generation of a glutarimide polyketide library featuring the dorrigocin, lactimidomycin, migrastatin, and NK30424 scaffolds. *J. Am. Chem. Soc.* 127, 11930-11931.
- Keatinge-Clay, A.T. (2012). The structures of type I polyketide synthases. *Nat. Prod. Rep.* 29, 1050-1073.

- Kellenberger, L., Galloway, I.S., Sauter, G., Böhm, G., Hanefeld, U., Cortés, J., Staunton, J., and Leadlay, P.F. (2008). A polylinker approach to reductive loop swaps in modular polyketide synthases. *Chembiochem* 9, 2740-2749.
- Khosla, C., Kapur, S., and Cane, D.E. (2009). Revisiting the modularity of modular polyketide synthases. *Curr. Opin. Chem. Biol.* 13, 135-143.
- Kusebauch, B., Busch, B., Scherlach, K., Roth, M., and Hertweck, C. (2009). Polyketide-chain branching by an enzymatic Michael addition. *Angew. Chem. Int. Ed.* 48, 5001-5004.
- Marsden, A.F., Wilkinson, B., Cortés, J., Dunster, N.J., Staunton, J., and Leadlay, P.F. (1998). Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* 279, 199-202.
- Murphy, A.C., Hong, H., Vance, S., Broadhurst, R.W., and Leadlay, P.F. (2016). Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chem. Commun.* 52, 8373-8376.
- Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., and Piel, J. (2008). Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat. Biotechnol.* 26, 225-233.
- Olsen, J.G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., Lindquist, Y., and Larsen, S. (1999). The X-ray crystal structure of β -ketoacyl [acyl carrier protein] synthase I. *FEBS Lett.* 460, 46-52.
- Piel, J. (2010). Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* 27, 996-1047.
- Price, A.C., Rock, C.O., and White, S.W. (2003). The 1.3-Angstrom-resolution crystal structure of β -ketoacyl-acyl carrier protein synthase II from *Streptococcus pneumoniae*. *J. Bacteriol.* 185, 4136-4143.
- Robbins, T., Kapilivsky, J., Cane, D.E., and Khosla, C. (2016). Roles of conserved active site residues in the ketosynthase domain of an assembly line polyketide synthase. *Biochemistry* 55, 4476-4484.
- Schmitt, I., Partida-Martinez, L.P., Winkler, R., Voigt, K., Einax, E., Dölz, F., Telle, S., Wöstemeyer, J., and Hertweck, C. (2008). Evolution of host resistance in a toxin-producing bacterial–fungal alliance. *ISME J.* 2, 632-641.
- Sugimoto, Y., Ding, L., Ishida, K., and Hertweck, C. (2014). Rational design of modular polyketide synthases: morphing the aureothin pathway into a luteoreticulin assembly line. *Angew. Chem. Int. Ed.* 53, 1560-1564.
- Sugimoto, Y., Ishida, K., Traitcheva, N., Busch, B., Dahse, H.-M., and Hertweck, C. (2015). Freedom and constraint in engineered noncolinear polyketide assembly lines. *Chem. Biol.* 22, 229-240.
- Sun, H., Liu, Z., Zhao, H., and Ang, E.L. (2015). Recent advances in combinatorial biosynthesis for drug discovery. *Drug Des. Devel. Ther.* 9, 823-833.

- Sundaram, S., Heine, D., and Hertweck, C. (2015). Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching. *Nat. Chem. Biol.* 11, 949-951.
- Walker, M.C., Thuronyi, B.W., Charkoudian, L.K., Lowry, B., Khosla, C., and Chang, M.C. (2013). Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* 341, 1089-1094.
- Wang, B., Song, Y., Luo, M., Chen, Q., Ma, J., Huang, H., and Ju, J. (2013). Biosynthesis of 9-methylstreptimidone involves a new decarboxylative step for polyketide terminal diene formation. *Org. Lett.* 15, 1278-1281.
- Weissman, K.J. (2016). Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.* 33, 203-230.
- Weissman, K.J. (2014). The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* 11, 660-670.
- Yin, M., Yan, Y., Lohman, J.R., Huang, S.-X., Ma, M., Zhao, G.-R., Xu, L.-H., Xiang, W., and Shen, B. (2014). Cycloheximide and actiphenol production in *Streptomyces* sp. YIM56141 governed by single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase. *Org. Lett.* 16, 3072-3075.
- Zhang, Y.-M., Hurlbert, J., White, S.W., and Rock, C.O. (2006). Roles of the active site water, histidine 303, and phenylalanine 396 in the catalytic mechanism of the elongation condensing enzyme of *Streptococcus pneumoniae*. *J. Biol. Chem.* 281, 17390-17399.
- Zheng, L., Baumann, U., and Reymond, J.-L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* 32, e115.

FIGURES LEGENDS

Figure 1 (A) Mechanism of a typical KS domain exemplified by module 10 of the thailandamide A PKS. Acyl-ACP intermediate acylates the active site Cys of the KS domain. The KS then decarboxylates malonyl extender units to form enolate which is then attacked by the acyl-KS intermediate to form a C2-elongated polyketide chain. **(B)** Non-canonical β -branching reaction catalyzed by the branching module of rhizoxin PKS. The malonyl extender unit is added to the α , β unsaturated thioester of the KS-bound polyketide intermediate. The δ -hydroxyl group then attacks the KS carbonyl to form the δ -lactone moiety of rhizoxin. KS – ketosynthase; B – branching domain; ACP – acyl carrier protein.

Figure 2 (A) Alignment of the conserved amino acids of branching- and elongating KSs. The numbers on top correspond to the residues of Rhi-KS14 (PDB code 4kc5). The entries in bold is part of the branching modules. **(B)** Comparison of the active site residues of KS-B (orange) (PDB code – 4kc5) with the bacillaene KS5 (green) (PDB code: 5erb). Residues chosen for mutational analysis are colored in black. Residues colored in blue forms the catalytic triad. **(C)** Recombinant proteins purified from *E. coli* on a 10% SDS gel. M – protein molecular weight marker. **(D)** CD spectrum of mutants used in this study compared with the wild type KS-B. **(E)** CD spectrum of the active site mutants.

Figure 3 (A) Chain branching assay setup with the deoxy variant of the substrate **1**. (B) SDS-PAGE analysis of the assay mix with **1**. The band corresponding to 100 kDa corresponds to the KS-B didomain whereas the band at the top corresponds to the trapped KS-B and ACP domains with the intermediate (**2**). All the proteins were verified by tryptic mass fingerprinting followed by MALDI. M – protein molecular weight marker.

Figure 4 (A) *In vitro* biotransformation setup of the substrate mimic **3** to generate **4**. LC-HRMS analysis of mutants showing extracted ion chromatograms of **4** ($m/z = 173.0808$, $[M+H]^+$). The EIC in cases of **4e**, **4j**, **4k**, **4l** and **4p** correspond to assays with increased enzyme concentration (10x). All other assays were performed under standard conditions (see Methods).

Figure 5 Rate of lactone (**4**) formation by the mutants and wild type KS-B. Data represent mean \pm s.d from three independent experiments.

Figure 6 (A) Partial bacillaene, macrolactin and thailandamide PKS. KSs highlighted were chosen for domain exchange experiments. (B) Cloning strategy for the generation of KS-B chimeras. (C) KS-B chimeras on a 10% SDS gel. All the proteins were verified by tryptic mass fingerprinting followed by MALDI. M – protein molecular weight marker. Rhi – Rhizoxin; Bae – bacillaene; Mln – macrolactin; Tha – thailandamide. (D) CD spectrum of KS-B chimeras.

Figure 7 (A) Possible routes for product formation with the KS-B chimeras. (B-D) LC-HRMS spectra of KS-B chimeras. Extracted ion chromatograms for $[M+H]^+$ 173.0808 is shown. (E) Shift in the retention time of **4** and **5**. Negative control refers to the assays performed with heat-inactivated KS-B chimeras.

Figure 1

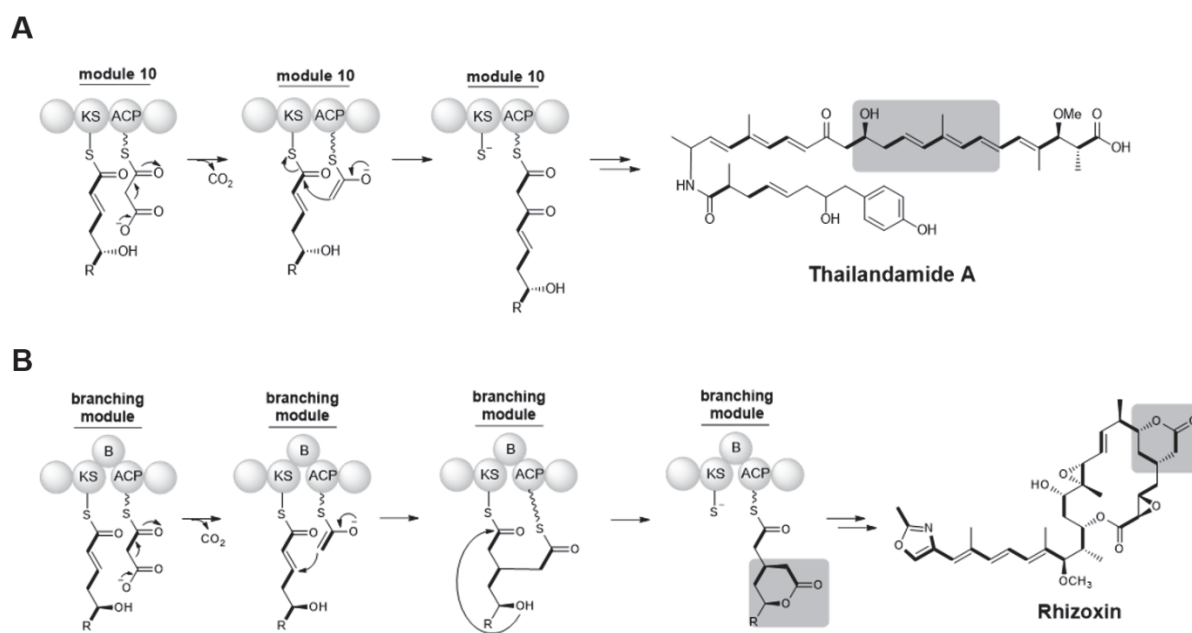


Figure 2

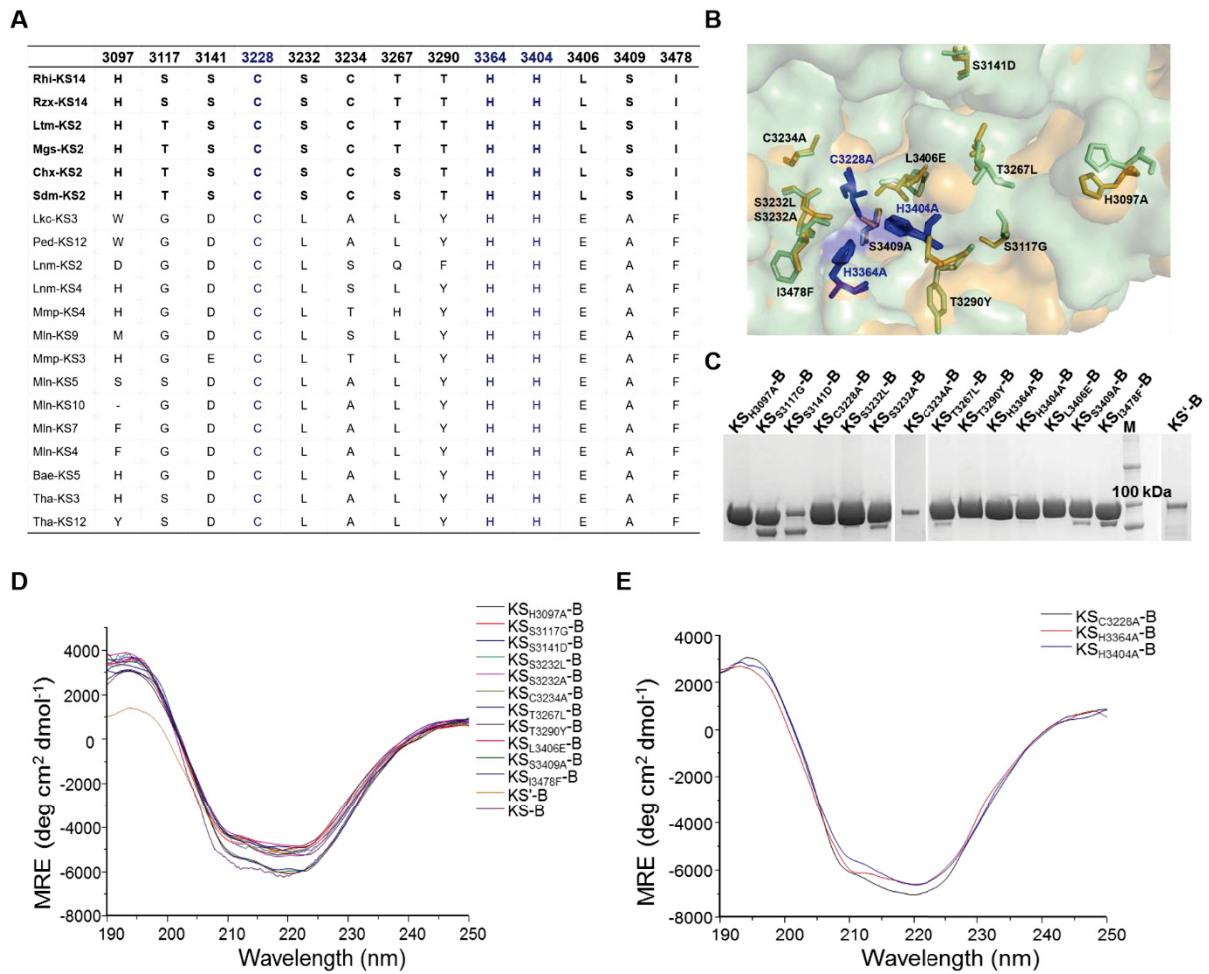


Figure 3

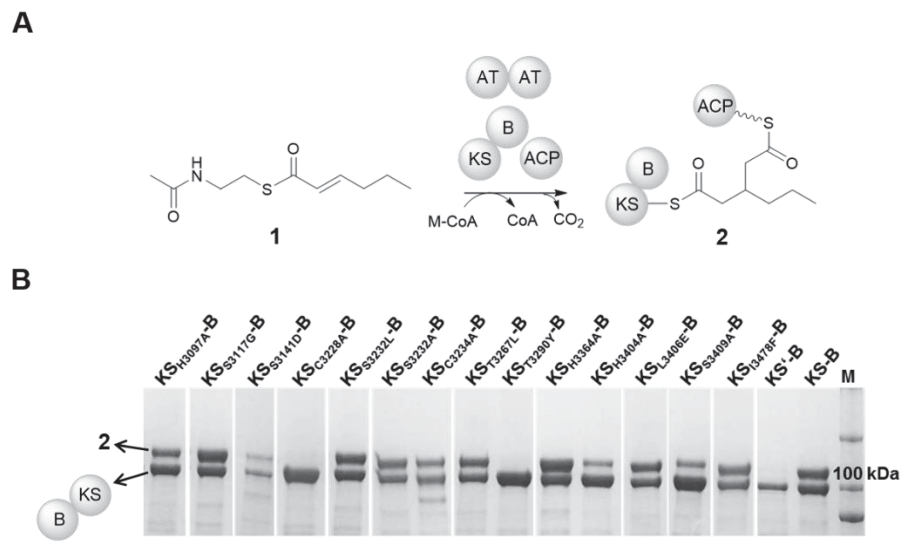


Figure 4

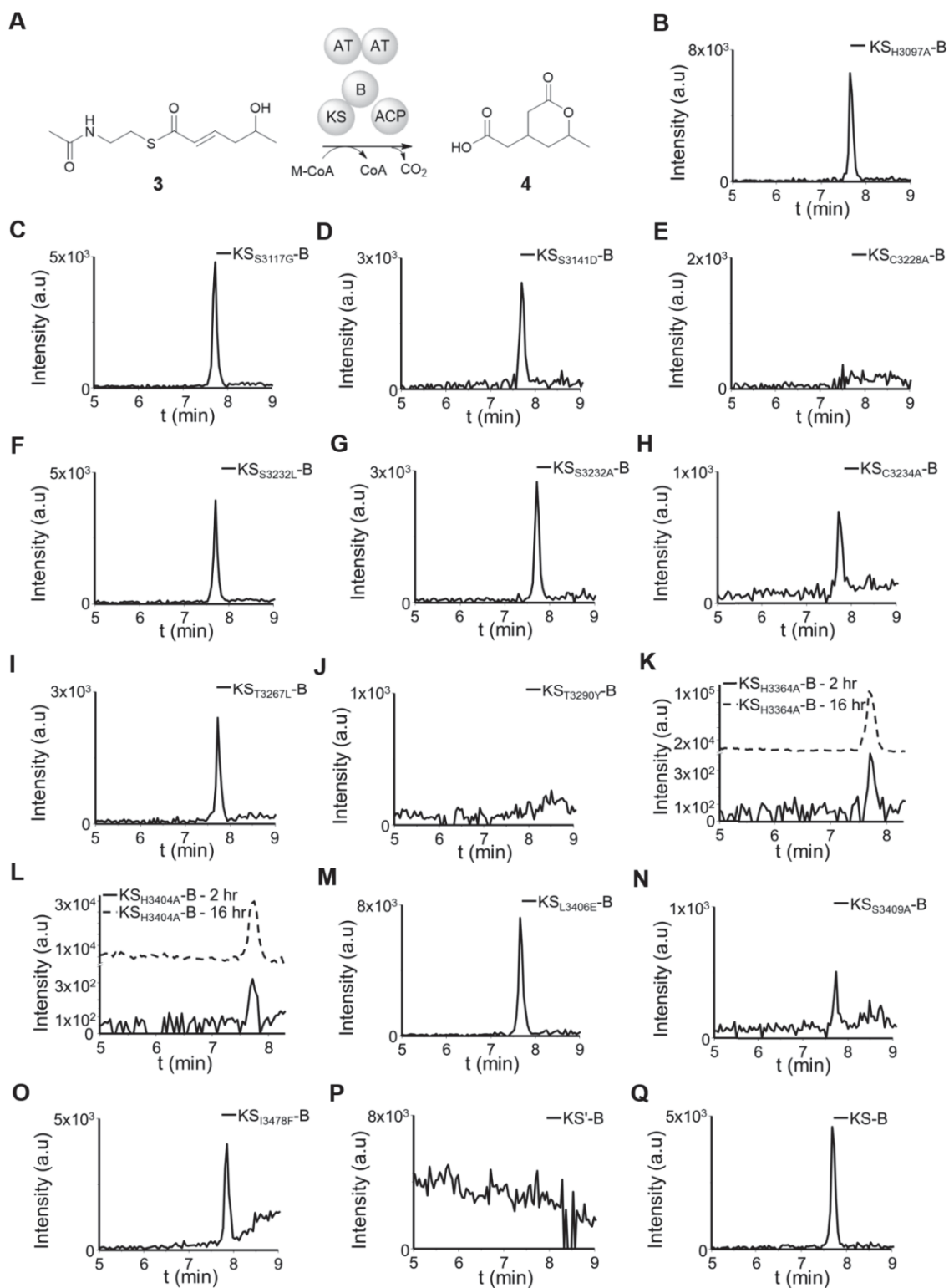


Figure 5

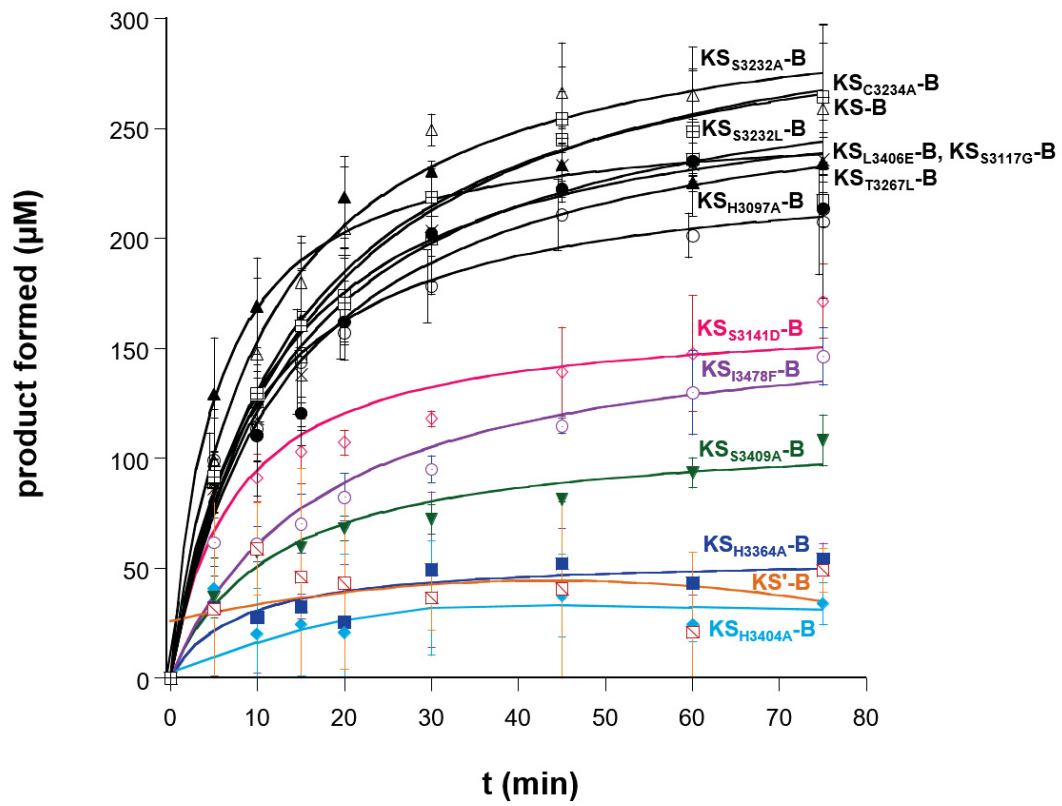


Figure 6

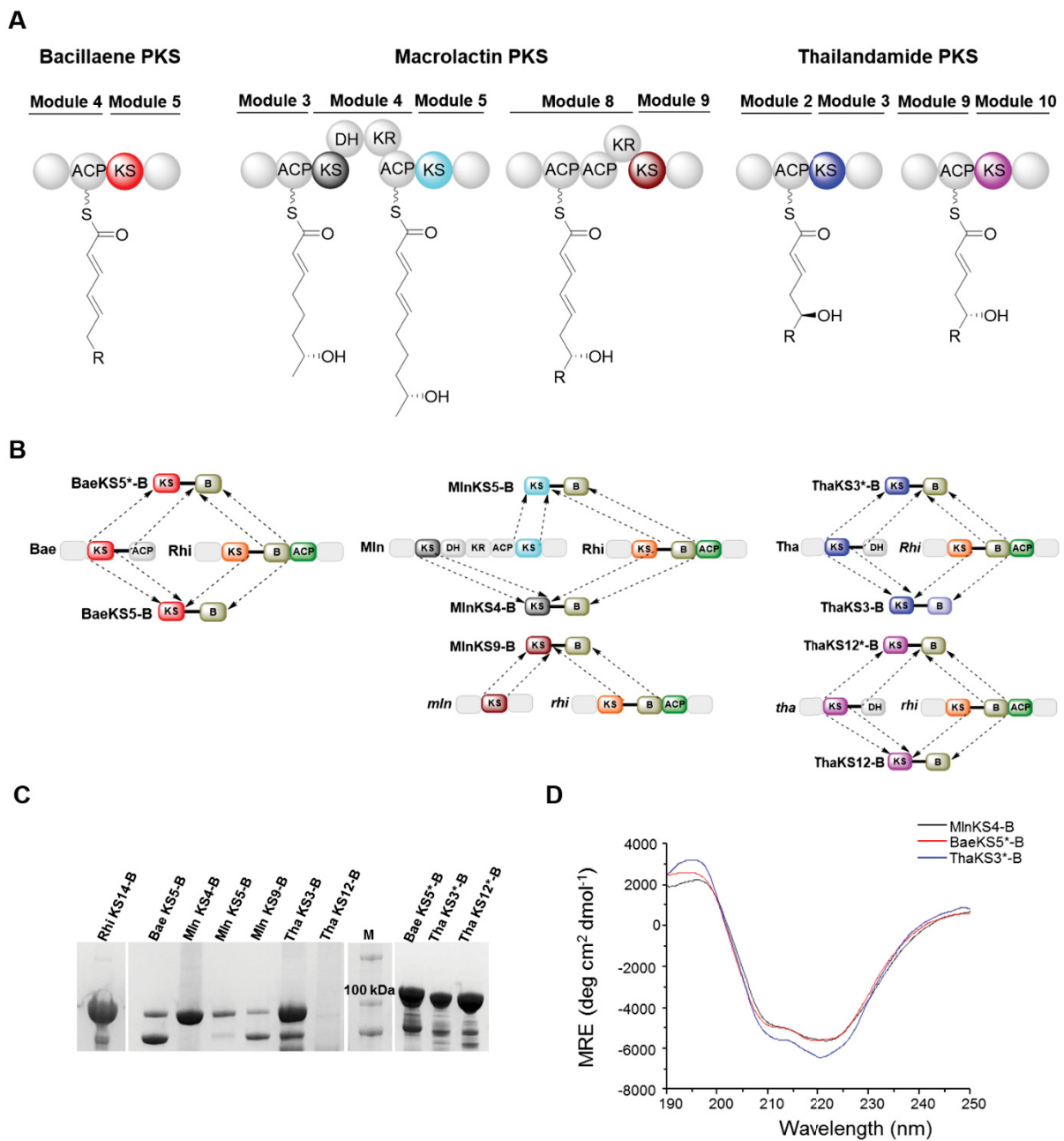
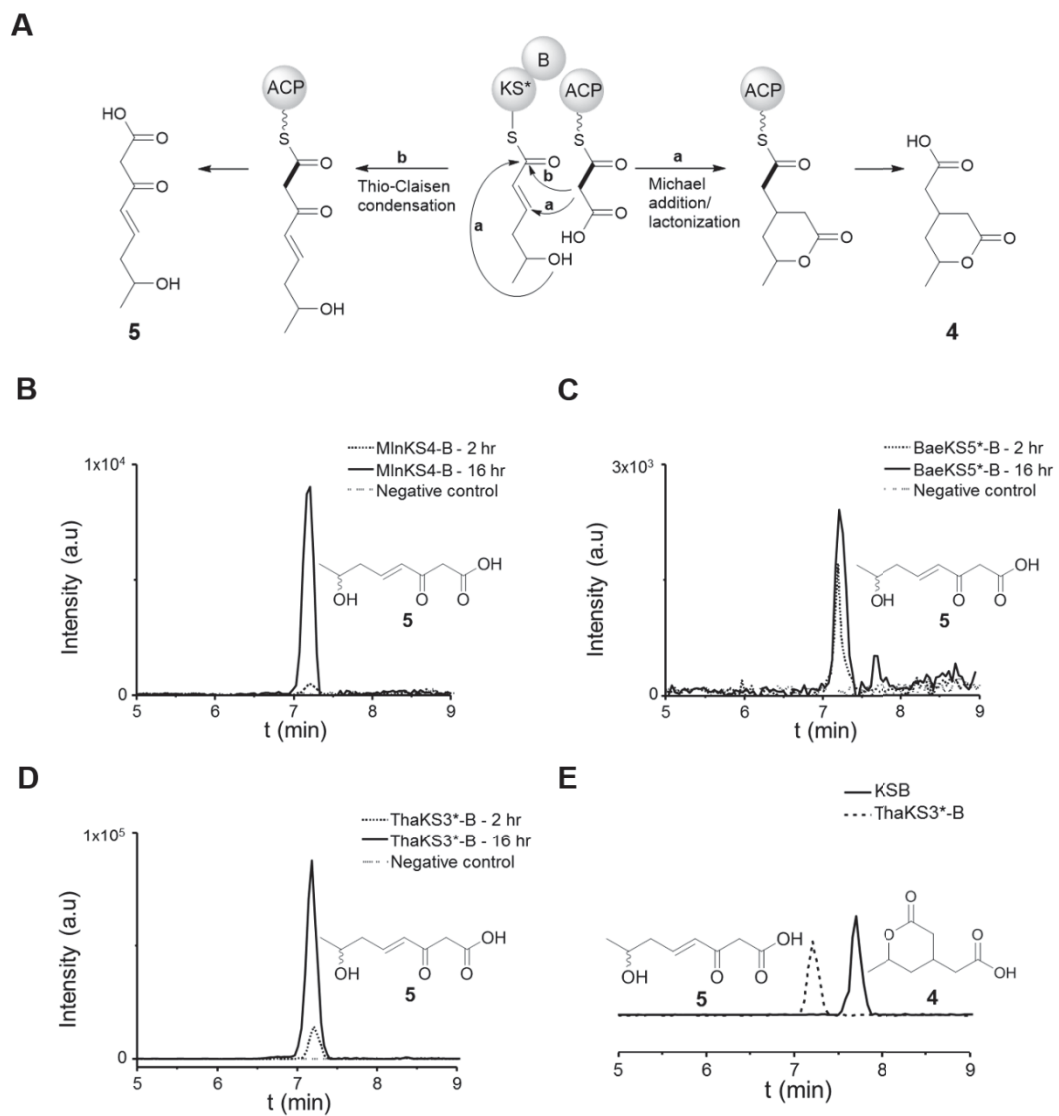


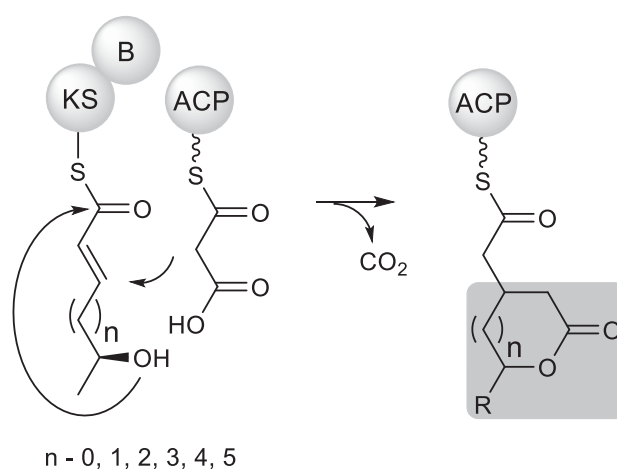
Figure 7



3.7 Manuscript F

Versatile polyketide-chain lactonization by a non-canonical ketosynthase of the rhizoxin megasynthase

Srividhya Sundaram, Tawatchai Thongkongkaew, Hak Joong Kim, Ruth Bauer, Daniel Heine, Christian Hertweck: Versatile polyketide-chain lactonization by a non-canonical ketosynthase of the rhizoxin megasynthase, *in preparation*, 2017.



The branching module of the rhizoxin PKS is tolerant towards substrates with varied acyl chain lengths. An array of 5- to 10-membered lactones is formed *in vitro* when tested with SNAC thioesters, independent of a TE domain.

Versatile polyketide-chain lactonization by a non-canonical ketosynthase of the rhizoxin megasynthase

Srividhya Sundaram¹, Tawatchai Thongkongkaew¹, Hak Joong Kim¹, Ruth Bauer¹, Daniel Heine¹, Christian Hertweck^{1, 2, 3*}

¹ Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology (HKI), 07745 Jena, Germany

² Chair for Natural Product Chemistry, Friedrich Schiller University, 07743 Jena, Germany

³Lead Contact

*Correspondence: christian.hertweck@leibniz-hki.de

SUMMARY

In vitro reconstitution of a branching module from the rhizoxin PKS revealed a non-canonical β -branching mechanism that forms its pharmacophoric δ -lactone moiety. The branching module is tolerant towards substrates with varied δ substituents that yield 6-membered rings like lactam and glutarimide structures. Here, we expand the range of products and show that the branching module lactonizes substrates with different chain lengths to yield an array of 5- to 10-membered lactones. The ability of KS to promote lactone formation complements the thioesterase- (TE) catalyzed cyclizations. It further sets the stage to generate ring-containing molecular architectures with potential bioactivities, especially when large-membered rings are challenging to produce by synthetic chemistry.

Keywords

Polyketide synthase, branching module, beta branching, lactones, ketosynthase, SNAC thioesters

Highlights

- Branching module of the rhizoxin PKS is tolerant towards substrates with varied chain lengths.
- Lactones with 5- to 10-membered rings are produced independent of a TE domain.
- The branching module is an excellent candidate for generating novel polyketide products.

In Brief

Sundaram et al., demonstrates the ability of branching module of the rhizoxin polyketide synthase to form 5- to 10-membered lactones by an atypical Michael addition-cyclization sequence.

INTRODUCTION

Polyketide synthases (PKSs) are efficient biosynthetic machineries that generate highly diverse natural products with varied chemical structures and biological activities (Staunton and Weissman, 2001). These enzymes are evolutionarily related to fatty acid biosynthetic enzymes and can be modular in nature (Tsai, 2016). Individual proteins (domains) within each module catalyze a single reaction in a sequential manner to generate polyketide intermediates that are tethered to the assembly line (Hertweck, 2009). The resulting polyketide can be tailored by post-PKS enzymes to yield the bioactive scaffolds (Olano et al., 2010). However, structural modifications can also occur while the intermediates are tethered to the assembly line (Pang et al., 2016; Sundaram and Hertweck, 2016). One of the on-line modifications is the generation of cyclic products or rings. Such ring systems add rigidity to the molecule and can be used as intermediates to generate active compounds (Townsend, 2013). Natural products containing small-sized lactone motifs are widely distributed in nature and most of these compounds are associated with effective antibiotic, antifungal and anti-inflammatory activities (Hertweck, 2009). Typically, cyclizations are carried out to offload a polyketide or peptide intermediate by the chain terminating thioesterase (TE) domains (Pinto et al., 2012; Wang et al., 2009; Xu et al., 2013; Zhou et al., 2015). **Figure 1A** shows the typical chain release by intramolecular condensation of the thioester and alcohol to generate macrolides like the 6-deoxyerythronolide B (**1**) core of the erythromycin (Weissman and Leadlay, 2005). Some examples of lactone formation independent of a TE domain have been reported, for example, the β -lactonization of the polyketide ebelactone A (**2**) (Wyatt et al., 2013) and the pH dependent δ -lactonization of lovastatin (**3**) (Campbell and Vederas, 2010) (**Figure 1B**). Although stand-alone TE domains are effective catalysts for *in vitro* cyclization (Boddy et al., 2003; Sharma and Boddy, 2007), unwanted hydrolysis or product inhibition hampers the cyclization. In light of the classical cyclization reactions, a non-canonical module of the branching module of the rhizoxin (**4**) PKS catalyzes β -branching and subsequent cyclization reactions. This module comprises of KS (ketosynthase), B (branching) and ACP (acyl carrier protein) domains. In an *in vitro* reconstitution experiment, the branching module has been shown to catalyze the transfer of ACP-bound malonyl unit to the α,β -unsaturated thioester mimic bound to the KS, by Michael addition (Bretschneider et al., 2013). Following lactonization, a branched product (δ -lactone) is generated (**Figure 1C**), in place of a canonical linear intermediate. This lactone group is pivotal for the antimitotic activity of rhizoxin (Partida-Martinez and Hertweck, 2007).

The ability of the branching module to form 6-membered rings have been well investigated (Heine et al., 2014). The branching module accepts SNAC thioester analogues with varied δ -substituents. Amine-substituted substrate resulted in lactam formation whereas the δ -carboxamide group resulted in glutarimide formation (Heine et al., 2014). While natural products with a δ -lactam group are not yet reported, glutarimide-containing compounds are reported in case of antibiotics such as cycloheximide (Yin et al., 2014) and 9-methylstreptimidone (Wang et al., 2013) in addition to the anticancer agents, iso-migrastatin and migrastatin (Ju et al., 2005). In addition to the malonyl extender unit, the KS-B didomain

also recognizes methyl-malonyl ACP to give a dual-branched δ -lactone product (Heine et al., 2015).

Natural compounds containing lactone motifs are usually associated with biological activities. Compounds with a gamma-lactone motif are widespread in nature that exhibits antibiotic, antifungal, cytotoxic and anti-inflammatory activities (Adames et al., 2015; Chianese et al., 2015; Liu et al., 2013; Seitz and Reiser, 2005). Novel caprolactone alkaloids isolated from a marine *Streptomyces* sp. show moderate phytotoxicity and activity against cancer cells (Stritzke et al., 2004). Higher members like penicillide (lactone 8), griseoviridin (lactone 9) have also been found in nature (Ferraz et al., 2008).

However, polyketide-derived large-membered lactones (>8) are not very abundant in nature. Synthetic and semi-synthetic approaches to generate cyclic products are challenging and the complexity increases with an increase in the ring size. In this study, we exploited the substrate scope of the branching module to produce medium- and large-membered lactone motifs. We show that the branching module is very versatile for it produces 5- to 10-membered lactone structures, independent of a TE domain, by β -branching and subsequent cyclization reactions.

RESULTS & DISCUSSION

In vitro biotransformation to lactones

To test the extent of the ring systems that can be formed by the branching module, thioester analogues that have shorter or longer methylene chains (**7**, **9**, **11**, **13** and **15**) were synthesized (**Figure 2**). Initially, the *in vitro* assay was performed only for 2 h, the condition that is standard for the formation of the 6-membered lactone. The reaction was analyzed by MALDI for the detection of the ACP-bound products and by HRMS, for the detection of the hydrolyzed product. However, analysis of the assay mix did not show any product formation. The assay was then performed for 16 h at 23 °C. After workup of the assay mix with methanol, high resolution MS analysis of the hydrolyzed products revealed the formation of new compounds in all the cases with the desired mass (**Figure 3A-3F**). The compounds have $[M+H]^+$ of 159.0652 (C₇H₁₁O₄), 173.0808 (C₈H₁₃O₄) (Bretschneider et al., 2013), 187.0965 (C₉H₁₅O₄), 201.1121 (C₁₀H₁₇O₄), 215.1278 (C₁₁H₁₉O₄) and 229.1434 (C₁₂H₂₁O₄) for 5-, 6-, 7-, 8-, 9-, and 10-membered rings respectively. When the assay was performed with heat-inactivated KS-B (negative control), no products were observed thus proving that the reaction is enzyme catalyzed. However, the ACP-bound lactones could not be detected by MALDI. **Figure S1** shows that the MS/MS fragments strongly suggest the formation of lactones. While small-sized polyketide-derived lactones are widespread in nature, medium- and large-sized lactones are scarce but are associated with effective biological activities (**17-20**) (**Figure 3G**) (Adames et al., 2015; Dräger et al., 1996; Ferraz et al., 2008; Ishigami, 2009; Seipke and Hutchings, 2013).

In addition to the thioester analogues mentioned above, substrates that yield 5-membered ethyl-substituted- (lactone 5*) or 7-membered lactones (lactone 7*) which has the same *m/z* as the native lactone-6 were also tested (**Figure 4A**). MALDI analyses of the assay mix resulted

in a new peak with the m/z of ~ 12719.7 which corresponds to the ACP-bound 5- and 7-membered lactones* (**Figure 4B**). In contrast to the lactones mentioned before, the products were observed after performing the assay for 2 h. However, the intensity of the ACP-bound 7-membered lactone was lower than that of the 5-membered lactone. However, the hydrolyzed products could not be detected by high resolution mass (result not shown), even when the assay was done for 16 h. In order to verify the formation of the corresponding lactones, LC-ESI-MS was also performed directly with the desalted-assay mix to detect the product bound and thereafter ejected from the PPant arm of ACP by further mass fragmentations. As shown in **figure 5**, the mass spectra were recorded for both the products. The spectra were deconvoluted to yield the observed mass of 12726.0 ± 5.0 Da and 12728.2 Da ± 3.2 Da for the formation of ACP-bound 5*-membered and 7*-membered lactones, respectively (**Figure 5A and 5B**). The ACP species generated ions with m/z 849.28 and m/z 849.41 for lactone 5* (**Figure 5A**) and lactone 7* (**Figure 5B**), respectively, which was then selected for MS/MS fragmentation experiments. The MS/MS analysis ejected the PPant-bound lactone fragments with m/z of ~ 415 for both the samples (**Figure 5A and 5B**). For all the lactones resulted, synthesis of reference compounds to validate the compounds are in progress.

In conclusion, the branching module is shown to accept SNAC thioesters with varied length to give 5- to 10-membered lactones. Although the ACP-bound proteins could not be detected by MALDI, evidences based on HRMS/MS and LC-ESI-MS clearly corroborate the results. Cyclization by branching module comes as an alternative approach for generating structural diversity in polyketide natural products and opens up strong possibilities for combinatorial biosynthesis. As the branching module also accepts amino-substituted surrogates, generation of different membered lactam rings could also be tested in the future.

SIGNIFICANCE

Modular PKSs play a major role in the production of therapeutically relevant natural products. Understanding the mechanism of the key enzyme components of these megasynthases would help in the production of non-natural and potentially bioactive compounds. The branching module of the rhizoxin PKS generates a δ -lactone moiety by a non-canonical β -branch formation followed by lactonization. The δ -lactone moiety is critical for the antimetabolic activity of the rhizoxin. The branching module has been utilized for its ability to produce existing as well as novel 6-membered rings *in vitro*. Since the lactone rings represent well-established pharmacophores, the branching module has been tested for rational production of small- to large-sized lactones. The results reveal that the branching module is highly versatile and produce lactones of up to 10-membered ring size. Although small-sized lactones are observed in nature, it is extremely challenging to synthesize stereospecific medium- and large-sized lactones by chemical approaches. Therefore, the branching module comes up as an excellent biocatalyst for efficient production of rings. These observations are relevant for expanding the novel branching-cyclization reactions to a variety of other ring systems.

EXPERIMENTAL PROCEDURES

General

Chemicals were purchased from Roth and Sigma-Aldrich. MALDI measurements were done using an UltrafleXtreme MALDI TOF/TOF (Bruker). High resolution mass measurements were performed using an Exactive Orbitrap High Performance Benchtop HRMS with an electrospray ion source and an Accela HPLC system (Thermo Scientific). LC-MS/MS experiments were performed using a QExactive equipment with Accela HPLC system (Thermo Scientific). LC-ESI-MS analysis of intact proteins was performed using ZORBAX 300SB-CN column in a ThermoFinnigan Surveyor LC equipment with a LTQ velos MS attached running on an electrospray ion source. The spectra from LC-ESI-MS measurements were deconvoluted using ProMass software (Thermo Scientific). NMR measurements were performed on a Bruker AVANCE II 300, AVANCE III 500 or 600 MHz spectrometer, with a Bruker cryo platform.

Protein production and purification

The procedures for the production and purification of ACP (pTB41), KS-B (pTB57) and RhiG (pTB39) were as described previously (Sundaram et al., 2015).

In vitro reconstitution of the branching reaction

The *in vitro* assay for the reconstitution of branching module was performed as described previously (Bretschneider et al., 2013) with minor modifications. Briefly, in a 40 μ L reaction, 167 μ M ACP, 3 μ M KS-B, 0.2 μ M RhiG, 750 μ M malonyl-CoA and 1,000 μ M of SNAC thioesters (**5**, **7**, **9**, **11**, **13** and **15**) were incubated in 20 mM Tris buffer (pH 7.0) at 23 °C with shaking at 400 rpm for 2 h or 16 h. Analysis of the reaction by MALDI (Bruker) was done by spotting 2.5 μ L of the assay mix with 1 μ L of 100 μ M of 2', 5' -dihydroxyacetophenone matrix. The MALDI-TOF was operated in positive linear mode. Data acquisition was performed using flexControl 3.3 in the range of 5–35 kDa with 1000 Hz laser frequency, 60% laser power and 500 shots at each spot. Analysis of the data was executed in flexAnalysis 3.3. For HRMS analysis, the reaction was stopped by adding equal volume of methanol. After a brief spin to remove precipitated proteins, the supernatant was dried using an upstream N₂ flow and the residue suspended in methanol was used for LC-HRMS analysis (Exactive). HPLC conditions: C18 column (Betasil C18 3 μ m, 150 x 2.1 mm) and gradient elution (MeCN/0.1% (v/v) HCOOH (H₂O) 5/95 for 1 min, gradient up to 98/2 in 15 min and 98/2 for 3 min; flow rate – 0.2 mL/min; injection volume – 3–5 μ L).

LC-ESI-MS and LC-MS/MS measurements

The branching reaction was also analyzed by LC-ESI-MS using a ZORBAX 300SB-CN column in positive mode. For this, in a 120 μ L reaction, 167 μ M ACP, 3 μ M KS-B, 0.2 μ M RhiG, 750 μ M malonyl-CoA and 1,000 μ M of **21** or **22** were incubated in 20 mM Tris buffer (pH 7.0) at 23 °C with shaking at 400 rpm for 16 h. The assay mix was then desalted using Zeba™ spin desalting columns (7K MWCO, Thermo Scientific) followed by elution of the desalted sample in water. The components were separated using a gradient of 30 – 98% B for 26 min with a flow rate of 0.6 mL min⁻¹. MS/MS fragmentation of the selected ion fragments

was also performed with the same instrument. The spectra from LC-ESI-MS measurements were deconvoluted using ProMass software (Thermo Scientific). For MS/MS analysis, Accucore C18 column (100 x 2.1 mm) was used in a QExactive operating in positive mode.

SUPPLEMENTAL INFORMATION

Supplemental information includes supplemental experimental procedures, synthetic data, figures and all relevant spectra.

AUTHOR CONTRIBUTIONS

S.S. and C.H. designed experiments. S.S. and R.B. performed biochemical experiments. S.S. analysed data. T.T. synthesized substrates for biotransformation experiments. H.J.K. synthesized reference compounds. S.S. and C.H. wrote the manuscript.

ACKNOWLEDGEMENTS

We thank M.Poetsch and T.Kindell for MALDI-MS measurements and A. Perner for HRMS and ESI-MS measurements. We are grateful for financial support by the International Leibniz Research School (ILRS to S.S.) and the Deutscher Akademischer Austauschdienst (DAAD to T.T.).

The authors declare no conflicts of interest.

REFERENCES

- Adames, I., Ortega, H.E., Asai, Y., Kato, M., Nagaoka, K., TenDyke, K., Shen, Y.Y., and Cubilla-Ríos, L. (2015). 3-epi-Waol A and Waol C: polyketide-derived γ -lactones isolated from the endophytic fungus *Libertella blepharis* F2644. *Tetrahedron Lett.* 56, 252-255.
- Boddy, C.N., Schneider, T.L., Hotta, K., Walsh, C.T., and Khosla, C. (2003). Epothilone C macrolactonization and hydrolysis are catalyzed by the isolated thioesterase domain of epothilone polyketide synthase. *J. Am. Chem. Soc.* 125, 3428-3429.
- Bretschneider, T., Heim, J.B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013). Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* 502, 124-128.
- Campbell, C.D., and Vederas, J.C. (2010). Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* 93, 755-763.
- Chianese, G., Gu, B.-B., Yang, F., Jiao, W.-H., Guo, Y.-W., Lin, H.-W., and Tagliatalata-Scafati, O. (2015). Spiroplakortone, an unprecedented spiroketal lactone from the Chinese sponge *Plakortis simplex*. *RSC Adv.* 5, 63372-63376.

- Dräger, G., Kirschning, A., Thiericke, R., and Zerlin, M. (1996). Decanolides, 10-membered lactones of natural origin. *Nat. Prod. Rep.* 13, 365-375.
- Ferraz, H., Bombonato, F.I., Sano, M.K., and Longo Jr, L.S. (2008). Natural occurrence, biological activities and synthesis of eight-, nine-, and eleven-membered ring lactones. *Quím. Nova* 31, 885-900.
- Heine, D., Bretschneider, T., Sundaram, S., and Hertweck, C. (2014). Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles. *Angew. Chem. Int. Ed.* 53, 11645-11649.
- Heine, D., Sundaram, S., Bretschneider, T., and Hertweck, C. (2015). Twofold polyketide branching by a stereoselective enzymatic Michael addition. *Chem. Commun.* 51, 9872-9875.
- Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* 48, 4688-4716.
- Ishigami, K. (2009). Synthetic studies of natural 10-membered lactones, mueggelone, microcarpalide, and Sch 642305, which have interesting bioactivities. *Biosci. Biotechnol. Biochem.* 73, 971-979.
- Ju, J., Lim, S.-K., Jiang, H., Seo, J.-W., and Shen, B. (2005). Iso-Migrastatin congeners from *Streptomyces platensis* and generation of a glutarimide polyketide library featuring the dorrigocin, lactimidomycin, migrastatin, and NK30424 scaffolds. *J. Am. Chem. Soc.* 127, 11930-11931.
- Liu, Y., Hu, Z., Lin, X., Lu, C., and Shen, Y. (2013). A new polyketide from *Diaporthe* sp. SXZ-19, an endophytic fungal strain of *Camptotheca acuminata*. *Nat. Prod. Res.* 27, 2100-2104.
- Olano, C., Méndez, C., and Salas, J.A. (2010). Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis. *Nat. Prod. Rep.* 27, 571-616.
- Pang, B., Wang, M., and Liu, W. (2016). Cyclization of polyketides and non-ribosomal peptides on and off their assembly lines. *Nat. Prod. Rep.* 33, 162-173.
- Partida-Martinez, L.P., and Hertweck, C. (2007). A gene cluster encoding rhizoxin biosynthesis in "*Burkholderia rhizoxina*", the bacterial endosymbiont of the fungus *Rhizopus microsporus*. *Chembiochem* 8, 41-45.
- Pinto, A., Wang, M., Horsman, M., and Boddy, C.N. (2012). 6-Deoxyerythronolide B synthase thioesterase-catalyzed macrocyclization is highly stereoselective. *Org. Lett.* 14, 2278-2281.
- Seipke, R.F., and Hutchings, M.I. (2013). The regulation and biosynthesis of antimycins. *Beilstein J. Org. Chem.* 9, 2556-2563.
- Seitz, M., and Reiser, O. (2005). Synthetic approaches towards structurally diverse γ -butyrolactone natural-product-like compounds. *Curr. Opin. Chem. Biol.* 9, 285-292.

- Sharma, K.K., and Boddy, C.N. (2007). The thioesterase domain from the pimarinic and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates. *Bioorg. Med. Chem. Lett.* 17, 3034-3037.
- Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 18, 380-416.
- Stritzke, K., Schulz, S., Laatsch, H., Helmke, E., and Beil, W. (2004). Novel caprolactones from a marine streptomycete. *J. Nat. Prod.* 67, 395-401.
- Sundaram, S., Heine, D., and Hertweck, C. (2015). Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching. *Nat. Chem. Biol.* 11, 949-951.
- Sundaram, S., and Hertweck, C. (2016). On-line enzymatic tailoring of polyketides and peptides in thio-template systems. *Curr. Opin. Chem. Biol.* 31, 82-94.
- Townsend, C.A. (2013). Enzymology: Modular biosynthesis branches out. *Nature* 502, 44-45.
- Tsai, S.-C.S. (2016). The substrate-protein and protein-protein interactions of fatty Acid and polyketide mega-synthases. *FASEB J.* 30, 390.392.
- Wang, B., Song, Y., Luo, M., Chen, Q., Ma, J., Huang, H., and Ju, J. (2013). Biosynthesis of 9-methylstreptimidone involves a new decarboxylative step for polyketide terminal diene formation. *Org. Lett.* 15, 1278-1281.
- Wang, M., Zhou, H., Wirz, M., Tang, Y., and Boddy, C.N. (2009). A thioesterase from an iterative fungal polyketide synthase shows macrocyclization and cross coupling activity and may play a role in controlling iterative cycling through product offloading. *Biochemistry* 48, 6288-6290.
- Weissman, K.J., and Leadlay, P.F. (2005). Combinatorial biosynthesis of reduced polyketides. *Nature Rev. Microbiol.* 3, 925-936.
- Wyatt, M.A., Ahilan, Y., Argyropoulos, P., Boddy, C.N., Magarvey, N.A., and Harrison, P.H. (2013). Biosynthesis of ebelactone A: isotopic tracer, advanced precursor and genetic studies reveal a thioesterase-independent cyclization to give a polyketide β -lactone. *J. Antibiot.* 66, 421-430.
- Xu, Y., Zhou, T., Zhang, S., Xuan, L.-J., Zhan, J., and Molnár, I.n. (2013). Thioesterase domains of fungal nonreducing polyketide synthases act as decision gates during combinatorial biosynthesis. *J. Am. Chem. Soc.* 135, 10783-10791.
- Yin, M., Yan, Y., Lohman, J.R., Huang, S.-X., Ma, M., Zhao, G.-R., Xu, L.-H., Xiang, W., and Shen, B. (2014). Cycloheximide and actiphenol production in *Streptomyces* sp. YIM56141 governed by single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase. *Org. Lett.* 16, 3072-3075.
- Zhou, Y., Prediger, P., Dias, L.C., Murphy, A.C., and Leadlay, P.F. (2015). Macrodiolide formation by the thioesterase of a modular polyketide synthase. *Angew. Chem. Int. Ed.* 127, 5321-5324.

FIGURES LEGENDS

Figure 1 (A) Classical cyclization reaction catalyzed by a TE domain. (B) Examples of bioactive lactone polyketides. (C) Non-canonical Michael addition and lactonization reactions catalyzed by KS-B didomain to generate the δ -lactone moiety of rhizoxin (4). TE – thioesterase; AT – acyl transferase; KR – ketoreductase; KS – ketosynthase; B – branching domain; ACP – acyl carrier protein.

Figure 2 *In vitro* branching assay set-up. SNAC thioesters (5, 7, 9, 11, 13 and 15) used in the study. 5 (boxed) is the substrate mimic that generates the native 6-membered lactone.

Figure 3 (A-F) LC-HRMS analysis of 5-, 6- 7-, 8-, 9- and 10-membered lactones. Negative control refers to the assays performed with heat-inactivated KS-B enzyme. (G) Examples of less-abundant natural 5-, 8-, 9- and 10-membered lactones.

Figure 4 (A) General scheme of the branching assay with 21 and 22. (B) MALDI analysis of the ACP-bound products for the formation of 5- and 7-membered lactones*. Negative control – enzyme assay without 21 or 22. * indicates a different ring system compared to **figure 3**.

Figure 5 Analysis of ACP species by LC-ESI-MS of (A) lactone 5* and ejected PPant arm with product attached. (B) Spectrum corresponding to lactone 7* and the ejected fragment.

Figure 1

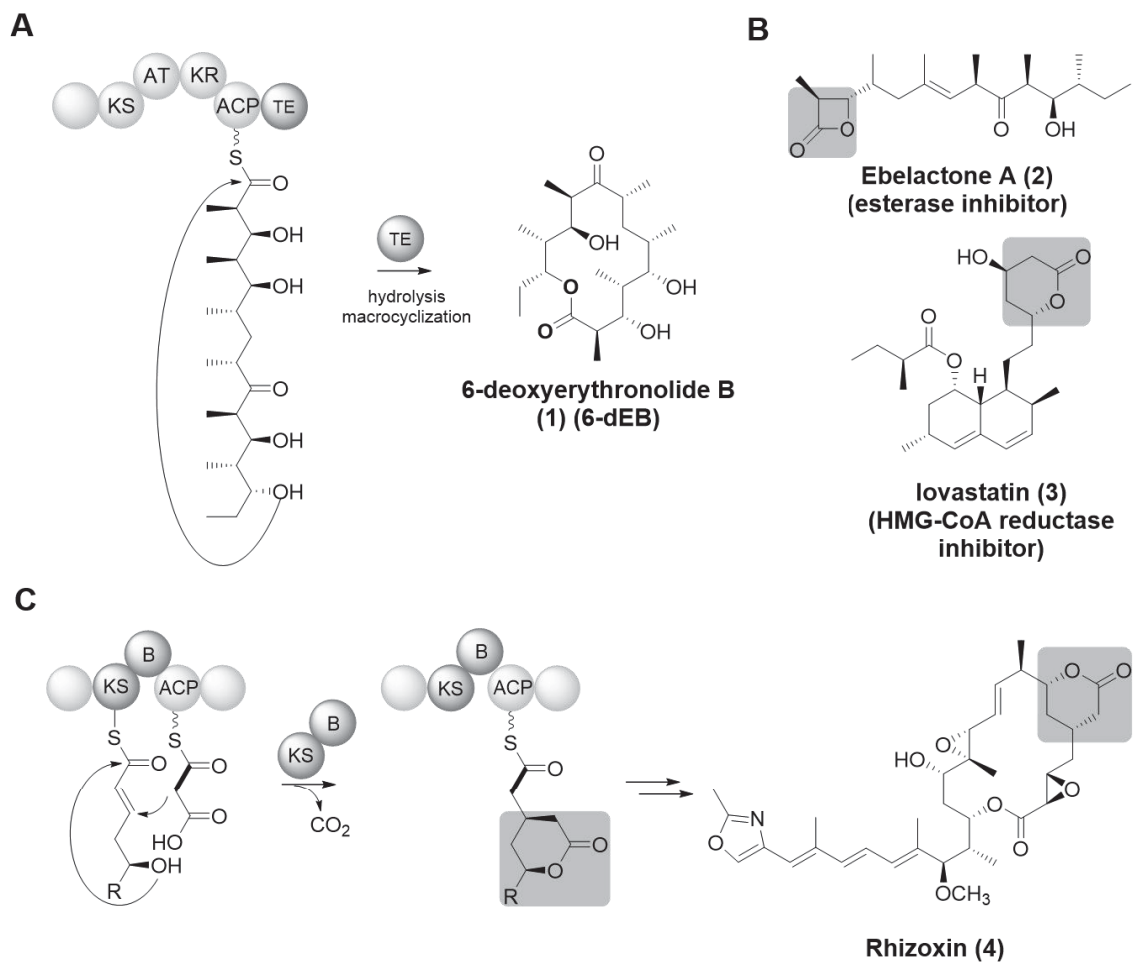


Figure 2

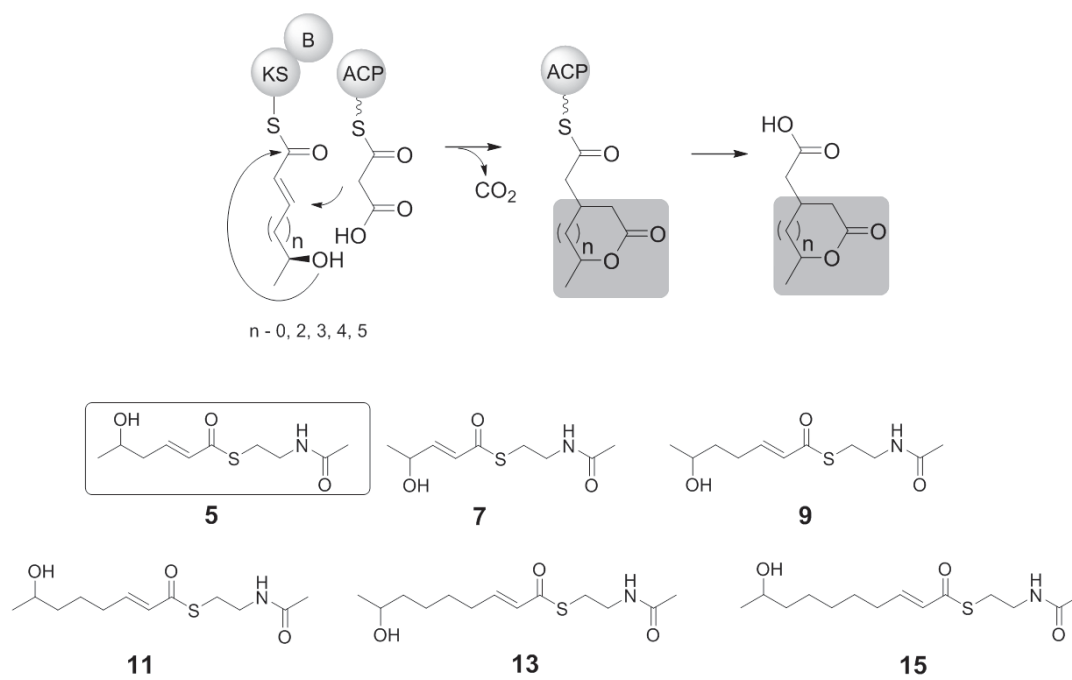


Figure 3

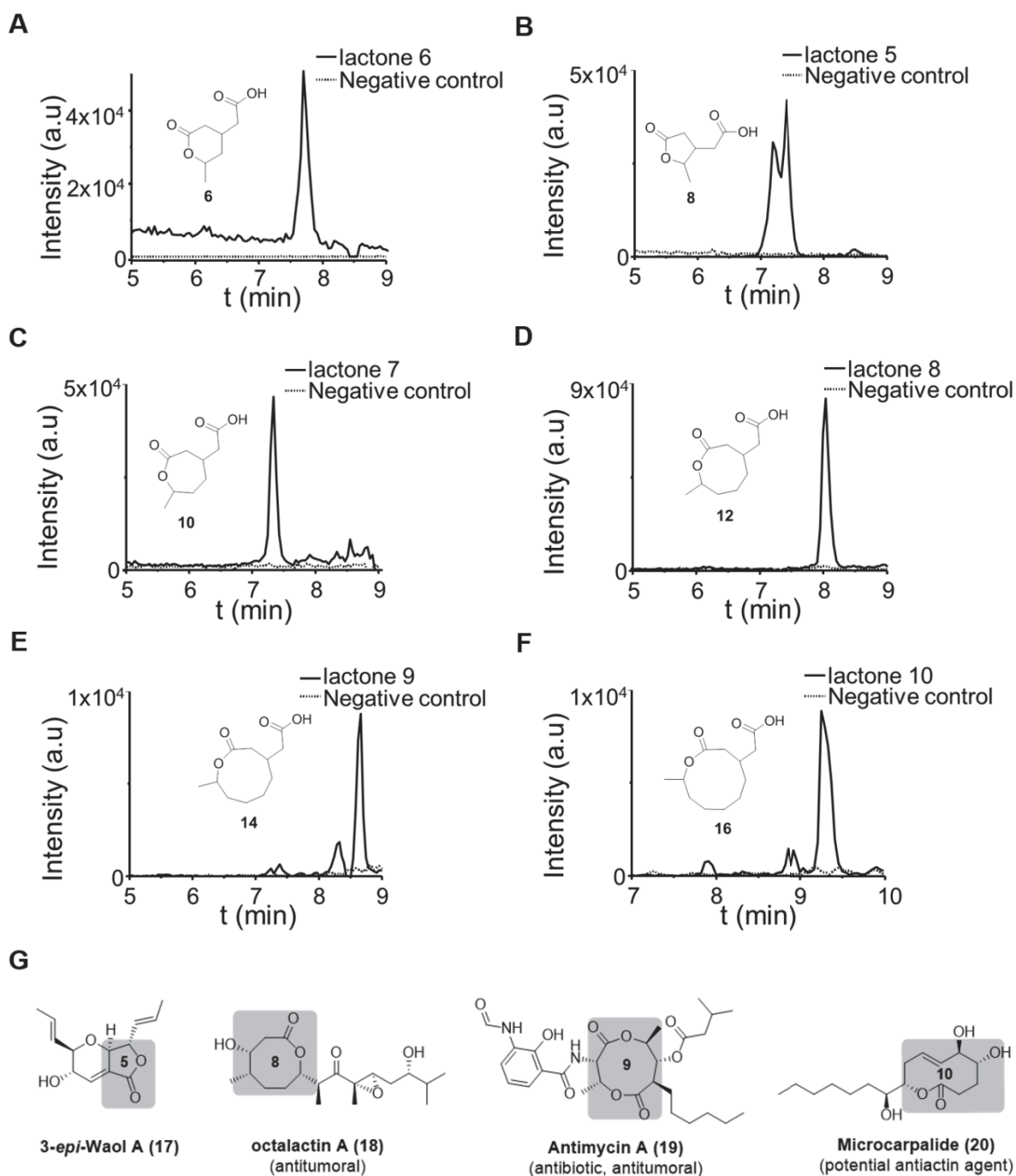


Figure 4

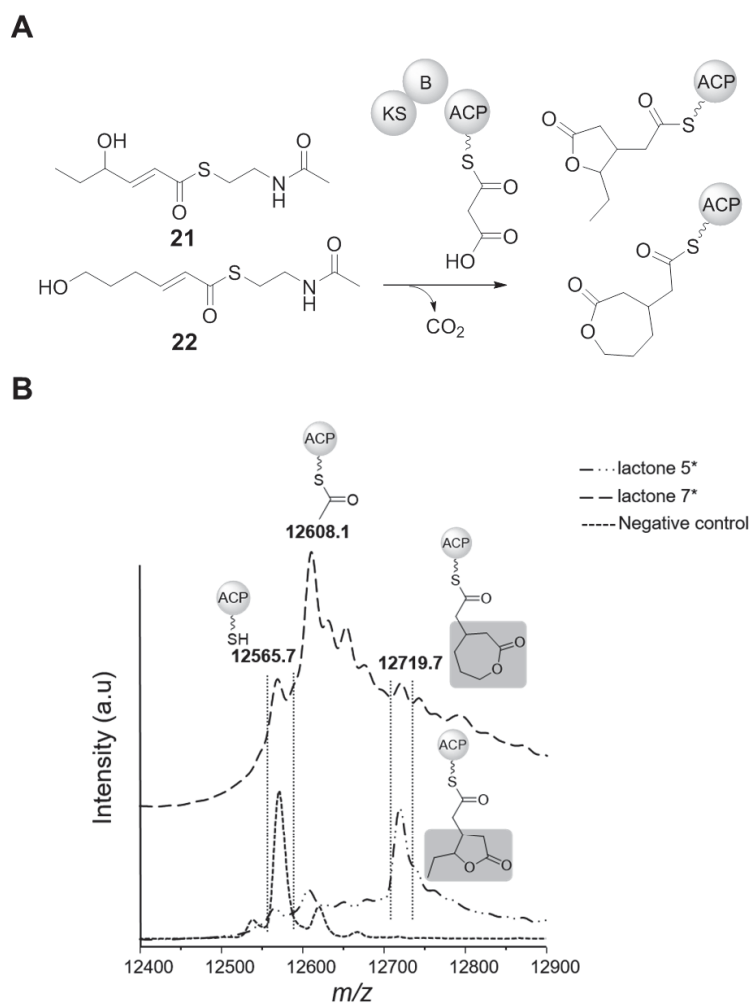
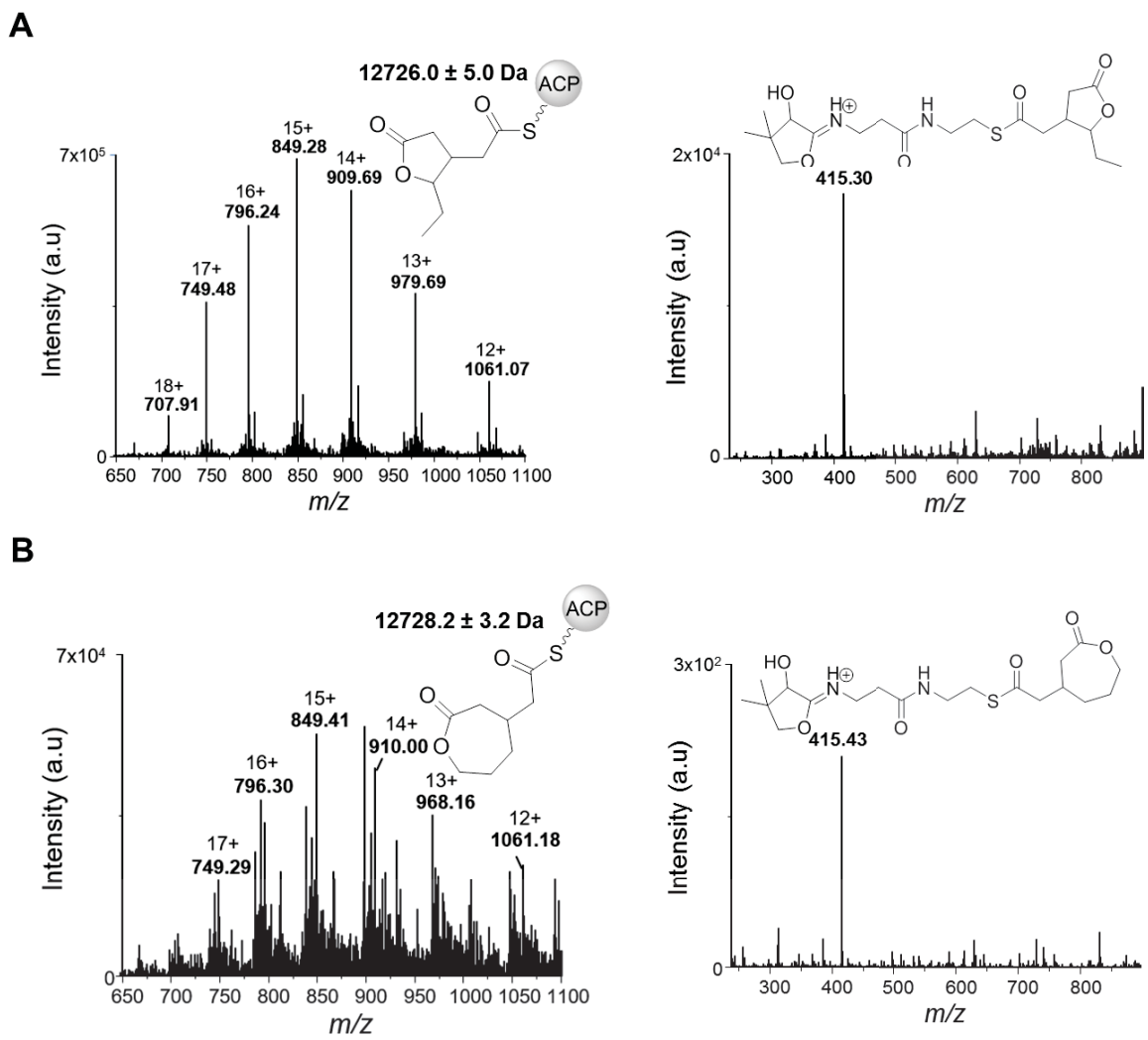


Figure 5



4 Unpublished results

4.1 Crystallization attempts of the entire branching module

4.1.1 Introduction

Since the report of the high resolution crystal structure of the porcine FAS in 2008,^{90, 91} many structural biologists endeavored to obtain the structure of an intact module to understand the exceptional dynamics of the individual catalytic domains between their chemical states, their

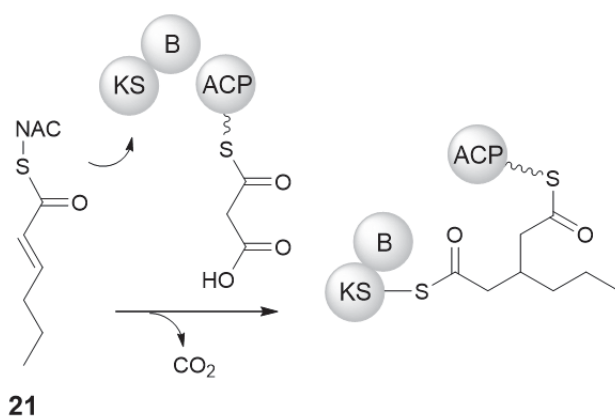


Figure 4.1 Deoxy variant of the SNAC thioester (**21**) promotes Michael addition to give a covalent intermediate bound to KS and ACP.

communication among the *cis*- and *trans*-acting enzymes and to facilitate the production of novel compounds.¹⁴³ Inspired by the FAS structure, several models for the PKS modular structure have been proposed.¹⁴³ Although a high resolution structure of a complete module has not yet been obtained, complementary techniques such as single-particle cryo-EM and SAXS have been used to generate models of *cis*-AT PKSs.⁹⁵⁻⁹⁸ However, not a single high resolution crystal structure of a complete module from a *trans*-AT PKS is available. The KSB didomain of the rhizoxin PKS branching module is the first solved X-ray structure from a *trans*-AT PKS family (PDB code: 4kc5),¹⁴² closely followed by the bacillaene PKS KS2.¹⁰¹ For a branching module that hosts versatile domains with extended substrate flexibility, little is known about how the domains interact with the ACP during the complex branching reaction. Trapping the highly flexible ACP in a crystal is very challenging. The crystal structure of the *E. coli* AcpP and the fatty acid 3-hydroxyacyl-ACP dehydratase FabA complex was obtained only using sulphonyl-3-alkynyl pantetheinamide as a mechanism-based crosslinking probe.¹⁴⁴ In the case of the rhizoxin PKS branching module, the SNAC thioester that lacks the δ -hydroxy group **21** resulted in a natural intermediate where both the KSB and ACP domains are covalently crosslinked (**Figure 4.1**).¹⁴² Therefore, we sought to crystallize the entire branching module-crosslinked-intermediate complex to visualize the interdomain contacts.

4.1.2 Results and discussion

Construction and functional analysis of a complete branching module

Attempts to maintain a stable KSB-ACP covalent intermediate has been very challenging. As ACP is very dynamic, it was constantly lost either during purification of the intermediate complex or during crystallization trials. Therefore, we planned to produce the KSB-ACP as a full-length protein. We reasoned that the additional covalent link between KSB and ACP would stabilize the crystal formation. **Figure 4.2a** shows the ACP of the branching module modeled into the KSB didomain (PDB code: 4kc5). Earlier attempts to crystallize the full-

length branching module (pRW₁₂ – construct from Robert Winkler) were not successful partly due to the flexible linker region between the B domain and ACP (colored red in **figure 4.2a**). The crystal structure of the KS-B didomain is therefore based on a new construct which is devoid of the unstable N-terminal part of KS (not shown) and the B-ACP long linker region (red).¹⁴² In this study, cloning strategies were planned to express KSB-ACP as a full length protein, in which the B-ACP linker (red) is excluded and the length of C-terminal part of the linker (dark blue) is optimized.

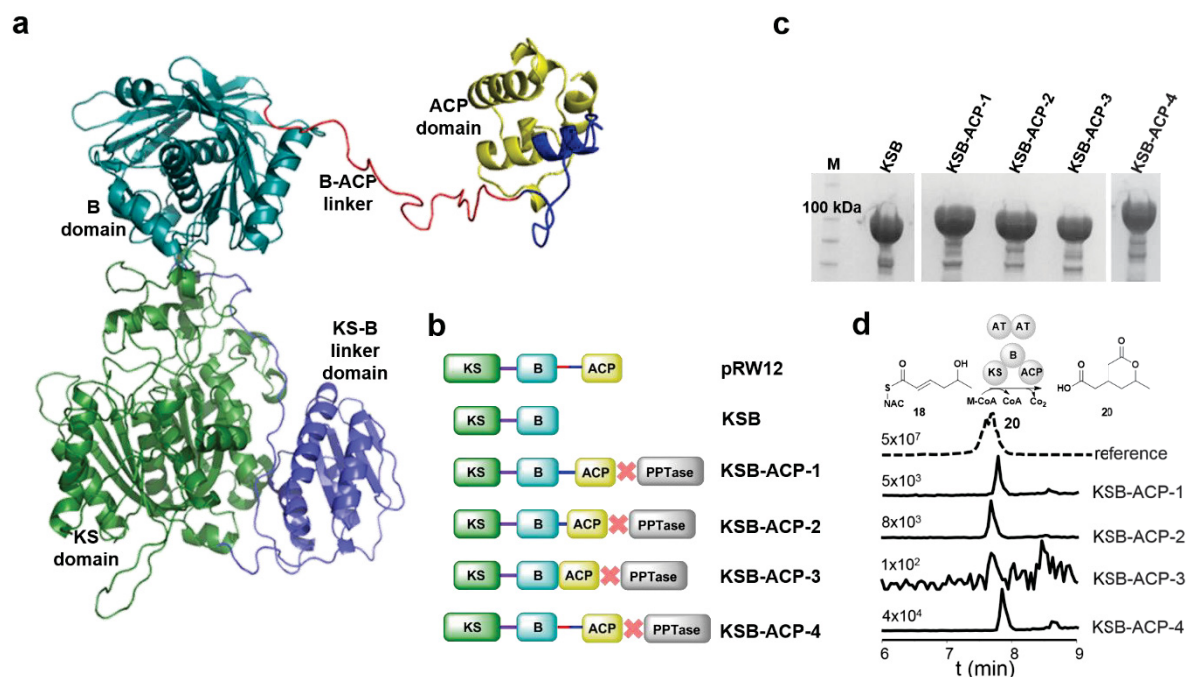


Figure 4.2 (a) ACP of the branching module modeled to the KS-B crystal structure (PDB code: 4kc5). (b) Domain information of the existing (pRW₁₂ and KSB) and new constructs (KSB-ACP-1, KSB-ACP-2, KSB-ACP-3 and KSB-ACP-4). The symbol X in red refers to the insertion of the stop codon. The constructs KSB-ACP-4 contains both the N- and C-terminal flanking regions of the KS domain. The color codes correspond to the respective regions in a. (c) SDS-PAGE of recombinant proteins. (d) LC-HRMS analysis of branching reaction showing extracted *m/z* chromatograms of 173.0808 [M+ H]⁺.

Three constructs were generated which differ in the length of the linker region between the B and ACP domains (dark blue). Although the construct pRW₁₂ was challenging to crystallize, it is possible that with the substrate **21**, the crosslinking might be established. For all the constructs, the region coding for PPTase from *B. rhizoxinica* pTB41¹⁴² was cloned to the downstream of the region for ACP, for *in vivo* phosphopantetheinylation of apo- to holo-ACP. These constructs are named as KSB-ACP₁, KSB-ACP₂, KSB-ACP₃ and KSB-ACP₄ (**figure 4.2b**). The protein KSB-ACP-3 lacks the B-ACP linker region. **Figure 4.2c** shows the SDS-PAGE analysis of all the recombinant proteins. *In vitro* biotransformation of the native substrate mimic **18** revealed that KSB-ACP-1, KSB-ACP-2 and KSB-ACP-4 resulted in the lactone product **20** whereas the turnover activity of KSB-ACP-3 was negligible (**figure 4.2d**).

The branching reaction was then performed with the deoxy substrate **21**. For the initial experiments, only KSB-ACP-1 was considered. To confirm the covalent linkage formation, LC-ESI-MS was performed directly with the assay mix after desalting. The mass spectra were

recorded for KSB (**figure 4.3a**), KSB-ACP-1 (**figure 4.3b**) and KSB-ACP-1 with **21** (**figure 4.3c**). After several attempts to improve the yield of a homogenous sample, an assay condition with controlled addition of KSB-ACP-1 (10 μ L/min) was selected for obtaining a crosslinked sample for crystallization trials. The initial crystals grew over a period of \sim 8 weeks (**figure 4.3 d**) but the diffraction of the crystals is not yet suitable for obtaining structural data. The crystallization work is being performed under collaboration with Dr. Georg Zocher (Universität Tübingen). If the crystallization of the protein KSB-ACP-1 is not successful, the assay and the subsequent steps will be optimized with other active proteins. Although X-ray crystal structure will not depict the entire mechanism, trapping the branching module in a specific state would help understand the conformational changes of the domains.

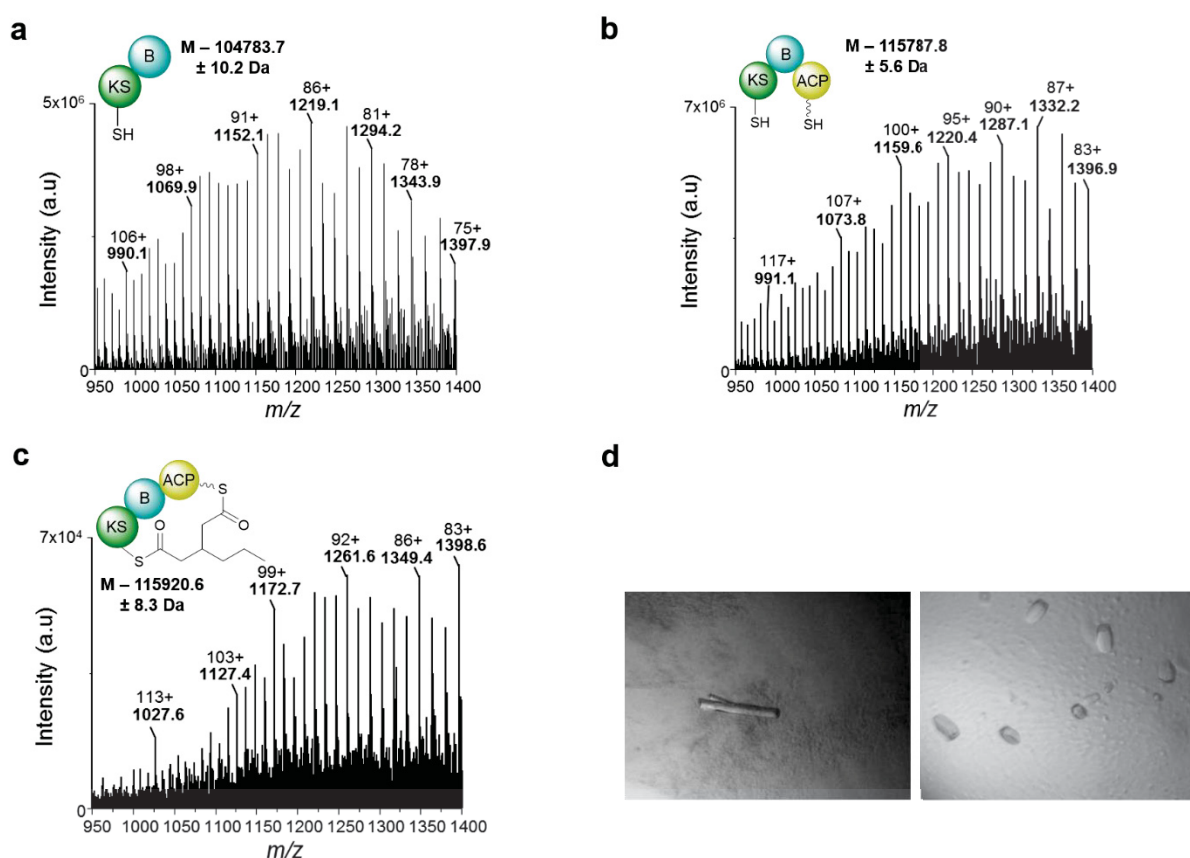


Figure 4.3 Analysis of KSB-ACP species by LC-ESI-MS. (a) KSB (b) KSB-ACP-1 (c) KSB-ACP-1 with substrate **21**. (d) Initial crystals formed

4.1.3 Methods

Cloning experiments

The genes corresponding to the B-ACP linker region, ACP and PPTase were amplified from pTB41¹⁴² and cloned to the C-terminus of KSB in pTB57.¹⁴² The primers used for KSB-ACP constructs are listed in **table 4.4**. KSB-ACP-1, KSB-ACP-2 and KSB-ACP-3 proteins contain the truncated KS domain, KS-B linker region, ACP but differ in the length of B-ACP linker region (**Figure 4.2b**).

Primers	Sequence	Accession no.	Protein region
PPTase_SacI_fw	CGAGCTCAGTGAATTGAAGGAGAT ATAATTCATG	WP_013428913.1	1-253
PPTase_XhoI_rv	CTCGAGTCAGCCAGAAGTGGGG GCTT		
PPTase_HindIII_rv	CCCAAGCTTTCAGCCAGAAGT GGGGGCTT		
KSB_EcoRI_rv	CGGAATTCGGCCATTTCTCTGT CCCAGTC	CBW75249.1	
ACP_EcoRI_fw1	CGGAATTCAAGCAGGAGGCTAA GCCGCAG		4162-4251
ACP_EcoRI_fw2	ATGAATTCGGCGAAAAAATTCTTG CTTTTATCCAACAG		4177-4251
ACP_EcoRI_fw3	ATGAATTCATCCAACAGGAGTTAC A GGACAAGC		4184-4251

Table 4.4 Primers used in the study.

Protein production and purification

The constructs were heterologously expressed in *E. coli* BL21 (DE3). Transformants containing the target genes were induced with 0.1 mM IPTG for 16 – 20 h at 15 °C. The proteins were purified by affinity and ion-exchange chromatography as described previously.¹⁴⁵ As a final purification step, size exclusion chromatography was performed in 20 mM HEPES, 150 mM NaCl (pH 7.5). At this stage, the aggregates were separated from the native dimer KSB-ACPs. The purity and identity of the proteins was checked by SDS-PAGE and tryptic digestion followed by MALDI analysis.

In vitro branching assay

The *in vitro* reconstitution of the branching reaction was done with 10 µM of the KSB-ACP, 750 µM malonyl-CoA, 0.2 µM RhiG and 1,000 µM of the SNAC thioester mimic **18** in 20 mM Tris buffer (pH 7.0) at 23 °C at 400 rpm for 2 h. For HRMS analysis, the reaction was stopped by adding equal volume of methanol. After a brief spin to remove precipitated proteins, the organic phase was dried using an upstream N₂ flow and the residue suspended in methanol was used for analysis (Exactive).

Crosslinking assay

To crystallize the KSB-ACP complex together with **21**, the *in vitro* assay was scaled up. The first trial was performed in 2 mL scale containing 87 µM KSB-ACP-1, 2 µM RhiG, 1.5 mM malonyl-CoA and 2.5 mM **21** in 20 mM Tris buffer (pH 7.0) at 23 °C, 400 rpm for 4 h. As the protein precipitated when RhiG was used in the assay, two different conditions were performed in the next trials. (i) In a 5 mL scale, 25 µM KSB-ACP-1, 1.2 mM malonyl-CoA and 4 mM **21** was used in 20 mM Tris buffer (pH 7.0) at 23 °C at 400 rpm for 3 h or (ii) In a 5 mL scale, 1.2 mM malonyl-CoA and 4 mM **21** was used in 20 mM Tris buffer (pH 7.0) and KSB-ACP-1 was added to the assay mix set at 23 °C, 400 rpm, at the rate of 10 µL/min over a period of 80 min, using a syringe pump, to a final concentration of 25 µM. The sample was further incubated for 90 min at 23 °C. Slow addition of KSB-ACP-1 improved the yield of homogenous sample. After

filtration to remove any residual precipitated proteins, the sample was further purified by size exclusion chromatography and used for crystallization trials.

LC-ESI-MS measurements and data analysis

To check the formation of covalent linkage with **21**, the branching reaction was analyzed by LC-ESI-MS using a ZORBAX 300SB-CN column in positive mode. For this, the assay mix was desalted using Zeba™ spin desalting columns (7K MWCO, Thermo Scientific) followed by elution of the desalted sample in water. The components were separated using a gradient of 30 – 98% B for 26 min with a flow rate of 0.6 mL min⁻¹. The spectra from LC-ESI-MS measurements were deconvoluted using ProMass software (Thermo Scientific).

Crystallization condition

The covalent intermediate was crystallized at 4 °C and 20 °C by mixing 6.5 mg/mL or 4.2 mg/mL of the protein solution with either 0.2 M magnesium chloride hexahydrate, 0.1 M HEPES (pH 7.5), 30% PEG₄₀₀ or 0.1 M sodium acetate trihydrate (pH 7.0), 12% PEG₃₃₅₀. For all the screening experiments, sitting drop vapor diffusion as well as micro seeding techniques have been tried. The diffraction experiments were performed at beamline Xo6DA at the Swiss Light Source in Villigen, Switzerland.

5 Discussion

5.1 Heterocycles formed by the branching module

Canonical ketosynthases catalyze the formation of linear chains in fatty acid and polyketide biosynthesis in a head-to-tail fashion.¹⁴⁶ In contrast to these canonical ketosynthase domains, rhizoxin biosynthesis utilizes a specialized branching module that catalyzes the nucleophilic addition of a malonyl building block to the α,β -unsaturated thioester, a mechanism which is distinct from the terpenoid-like β -branching known so far.¹⁴¹ The entire branching reaction has been reconstituted *in vitro* and has been shown to catalyze the vinylogous Michael addition and the subsequent lactonization.¹⁴² Expanding the substrate scope of this branching module has great impact in the synthesis of novel polyketide products.

The plasmids to express the genes for KS-B, ACP and AT were constructed by Dr. Tom Bretschneider. The KS-B could only be produced as a soluble and functional didomain. The construct for the production of ACP also contains a downstream PPTase gene from *B. rhizoxinica* to allow the *in vivo* activation of apo to holo-ACP.¹⁴² The *in vitro* assay involves the addition of purified, functional proteins (KS-B, ACP and AT), malonyl-CoA and 5-hydroxyhexenoic acid SNAC thioester mimic **18**. One of the strategies to demonstrate the substrate promiscuity of the branching module is by feeding non-cognate synthetic variants of the substrate in an *in vitro* reaction. Such thioesters mimic the phosphopantetheine arm of the ACP and are highly suitable for *in vitro* substrate tolerance analyses.¹⁴⁷ Phylogenetic analysis of KSs from *trans*-AT PKSs revealed that the substrate specificity is determined by the first four carbons of the intermediate (e.g. branched or unbranched).⁸⁵ The branching KS belongs to the clade that accepts α,β -unsaturated, unbranched polyketide intermediates. We tested the effect of the α,β -dihydro derivative and several δ -substituted substrates on the branching reaction. The latter is particularly important as the δ -hydroxy group of **18** is pivotal for the lactonization. A series of biotransformation experiments with SNAC thioester mimics (synthesized by Dr. Daniel Heine) revealed that the branching module is highly versatile in the formation of different heterocycles (**Manuscript A**).¹⁴⁸

Figure 5.1a shows the SNAC thioesters used for *in vitro* biotransformation. The substrates **18**, **22** and **23** were accepted by the branching module whereas **24-26** were not. The result that the KS is incapable of accepting **24** reassures that the enzyme recognizes the α,β double bond to perform the Michael addition. The KS5 from bacillaene (Bae) PKS, which shares the same phylogenetic clade as the branching KS, however, is tolerant towards both α,β -saturated and unsaturated substrates.¹⁴⁹ The BaeKS5 catalyzes the typical Claisen condensation and produces linear product. Therefore, the result shows that the branching KS is highly specialized and cannot be tuned to make linear chains with a saturated substrate. It is difficult to explain why the biotransformation does not take place with **25** and **26**. It is possible that the use of **25** resulted in steric clashes in the KS binding pocket. However, the decarboxylated form of malonyl-CoA bound to the ACP was observed by MALDI measurement, proving that the enzymes used in the assay were functional and that the cyclized product was not observed only

because of the δ -substituent. When the assay was performed with **26**, due to its acidic nature, the assay mix precipitated hampering further analysis. Adjusting the pH conditions of the assay also did not result in any product formation. Whether the KS was acylated by these substrates remains unclear. It is possible that the KS acylation is reversible that leads to unloading of nonproductive substrates.

One of the significant findings from these biotransformation experiments is the ability of the branching module to process a δ -carboxamide-substituted SNAC thioester to form a glutarimide moiety. The first report of a natural product with a glutarimide group, cycloheximide (**27**), came in 1949.¹⁵⁰ Several glutarimide-containing polyketides followed this initial discovery including iso-migrastatin (**28**),¹⁵¹ migrastatin,¹⁵² dorrigocins,¹⁵³ lactimidomycin¹⁵⁴ and 9-methylstreptimidone¹⁵⁵ (**29**). These compounds have diverse functions as antiviral, antifungal, antitumoral agents and function as protein synthesis inhibitors by binding to the 60S ribosome unit.^{156, 157} Owing to the pharmacological role of these compounds, studying their biosynthesis has gained attention over the recent years. The enzymatic machineries that govern the biosynthesis of these compounds turned out to be *trans*-AT PKSs that have been suggested to employ similar branching module to install the glutarimide moieties.¹⁵⁸⁻¹⁶¹ However, the enzymes have not been investigated *in vivo* or *in vitro*. The current study (**Manuscript A**) provides the first biochemical proof that the glutarimide is indeed a result of a Michael addition-cyclization reaction by the branching module.

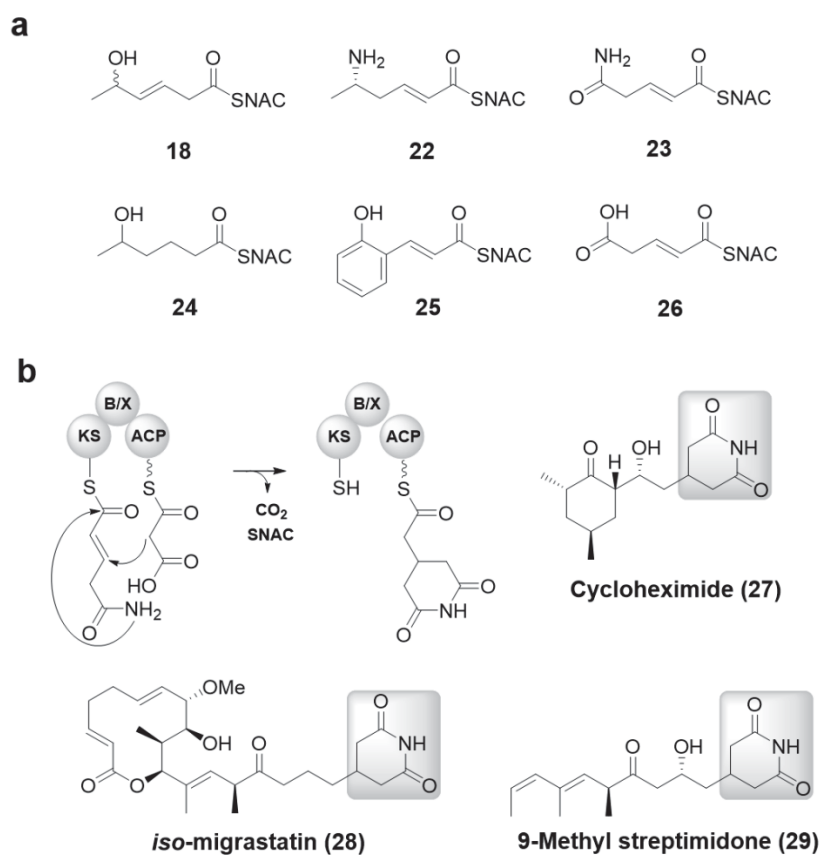


Figure 5.1 (a) Substrates used in the study (b) Proposed mechanism of β -branching in the biosynthesis of glutarimide unit and the natural products with glutarimide groups critical in bioactivity

Apart from the glutarimide ring, δ -lactam moieties result when the assay is performed with the δ -amine-substituted SNAC thioester. δ -lactam-containing terpenes and alkaloids have been reported from marine and soil fungi that exhibit anti-allergic and cytotoxic properties.¹⁶²⁻¹⁶⁴ Although no δ -lactam-containing polyketides have been reported so far, it is conceivable that such molecules are yet to be discovered.

Kinetic analyses demonstrated that the KS could accept both *S* and *R* enantiomers of **18**, albeit at different rates. The turnover rate of the reaction was greater for the *S* enantiomer ($v_{\max} = 20.9 \pm 2.4 \mu\text{M min}^{-1}$) compared to the non-natural *R* enantiomer ($v_{\max} = 2.1 \pm 0.3 \mu\text{M min}^{-1}$). Although the KS is flexible towards the configuration of the δ -hydroxy or the δ -amine group, the branching occurs by a *syn*-selective nucleophilic addition of the malonyl extender unit. HRMS analysis shows that exclusively *syn*-substituted lactone and lactam are formed. It appears that the branching KSs have steric requirements to specifically catalyze *syn*-specific β -branching and further yield 6-membered heterocycles. A recent report on the substrate specificity of KSs from *trans*-AT system revealed that the presence of a bulky Met residue immediately preceding the active site Cys precludes acceptance of β -branched substrates.¹⁴⁹ It has also been shown that mutating the Met to a smaller Ala residue provided the necessary space in the KS binding pocket to bind a β -branched SNAC derivative. Due to the presence of Met residue in the branching KS, the added β -branch likely points in the direction opposite to that of Met, providing a possible explanation for the *syn*-selectivity. Non-canonical enzymes that perform dual reactions are scarce from *trans*-AT PKSs. A dehydratase in the *cis*-AT-mediated ambruticin biosynthesis (AmbDH₃) has been reported to catalyze stereoselective dehydration and pyran ring formation.^{165, 166} Such heterocycles in natural products are important as they represent structural motifs and pharmacophoric units as in the case of rhizoxin or cycloheximide. Taken together, the non-canonical Rhi KS is promiscuous in processing non-native substrates and is a potential candidate for biosynthetic engineering of polyketide structures.

5.2 Stereoselective dual polyketide branching

Select *cis*-acting ATs can load unusual extender units including allylmalonyl-CoA, ethylmalonyl-CoA and chloroethylmalonyl-CoA.⁴⁰ However, characterized ATs from *trans*-AT PKSs have limited selectivity over α -substituted extender units. Very few ATs have been reported to incorporate non-malonyl extender units to separate modules (**figure 5.2**).¹⁶⁷⁻¹⁶⁹

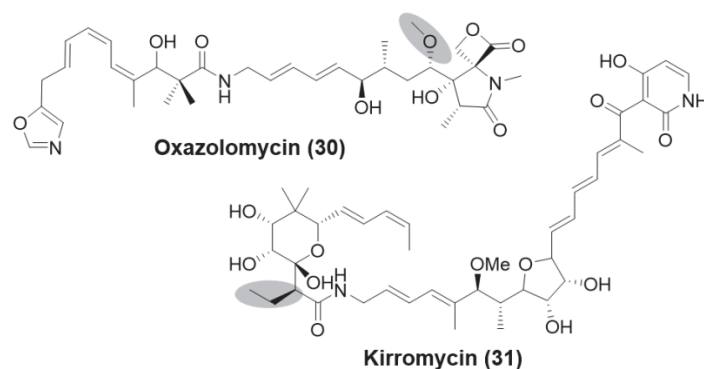


Figure 5.2 Extender unit variability in *trans*-AT derived polyketides: Oxazolomycin (methoxymalonyl-); Kirromycin (ethylmalonyl-). The extender units are highlighted.

In this context, the ability of the tandem-AT (RhiG) from the rhizoxin PKS to transfer non-natural methylmalonyl and dimethylmalonyl units on the ACP is an important step forward (**Manuscript B**).¹⁷⁰ The finding is significant especially for the β -branching module, where more than one branch could be introduced at the same time *in vitro*. To detect the corresponding products by HRMS, the extraction condition of the assay mix needed to be changed slightly. After performing the *in vitro* assay, the reaction mix was extracted with methanol instead of ethyl acetate. The intensity of the released product greatly improved with methanol extraction. Reactions conducted with methylmalonyl extender units resulted in the production of the expected dually branched product. Notably the methyl-group of the extender unit is found in the side chain, proving that the lactonization involves the attack of the δ -hydroxy group to the KS-bound thioester. It is important to note that both *R* and *S* enantiomers of methylmalonyl-CoA are accepted by the branching module, resulting in both the diastereoisomers in the final lactone product. Hence, in a single step, two (α and β) branches are produced in a vinylogous stereoselective chain branching event (**Manuscript B**). It is conceivable that the binding cleft of KS can accommodate both the *R* and *S* configurations of the newly added methyl branch in addition to the *syn*-lactone. A similar example of a malonyl-specific AT tolerant towards methylmalonyl-CoA has been observed in an *in vitro* assay with the bryostatin AT.¹⁷¹

The structures of epothilone and yersiniabactin contain gem-dimethyl groups derived from dimethylmalonyl-CoA extender unit.¹⁷² However, it is loaded on the ACP by a methyltransferase-dependent mechanism. To the best of our knowledge, the *trans*-AT described in this study is the first of its kind to load a dimethylmalonyl extender unit on the ACP. The fact that KS could decarboxylate the dimethylmalonyl unit to yield isobutryl-ACP shows that only the Michael addition and the subsequent lactonization are hampered. The wrong orientation of the additional methyl branch in the KS binding cleft could likely be the

reason for this. The extender unit specificity of AT could be determined based on signature fingerprint motifs in the sequence.^{173, 174} According to this, GHSXG and HAFH indicate malonyl-CoA specificity whereas GHSQG and YASH motifs indicate methylmalonyl-CoA specificity (**figure 5.3a**). The Phe residue preceding the catalytic His has been predicted to sterically clash with α -substituted extender units including methylmalonyl-CoA.⁵⁹ Considering that RhiG contains the unfavored Phe-His sequence, it remains unclear how the enzyme could still bind to methylmalonyl- and dimethylmalonyl-CoA. A similar result was obtained for the *trans*-acting disorazol AT, which showed some tolerance towards methylmalonyl-CoA despite containing the Phe-His motif.⁹⁹ One possibility for this unexpected promiscuity is that the fingerprint residues are not well established for non-canonical extender units and *trans*-acting ATs. A modeled RhiG based on the available structures of disorazol AT⁹⁹ and FabD¹⁷⁵ shows that some of the residues could be oriented differently in the active site (**figure 5.3b**). A crystal structure of extender unit-bound RhiG would reveal insights into the substrate specificity. In parallel to this, the AT specificity could also be engineered to incorporate unusual extender units via domain swapping or site-specific mutations.^{54, 176} However, in the current β -branching setting, the orientation of these ACP-bound extender units within the KS binding pocket should be taken into consideration. It is challenging to incorporate regiospecific methyl groups to natural products using synthetic chemistry approaches. In this context, the branching module and the *trans*-AT from the rhizoxin system are excellent candidates for generating novel non-natural polyketide derivatives.

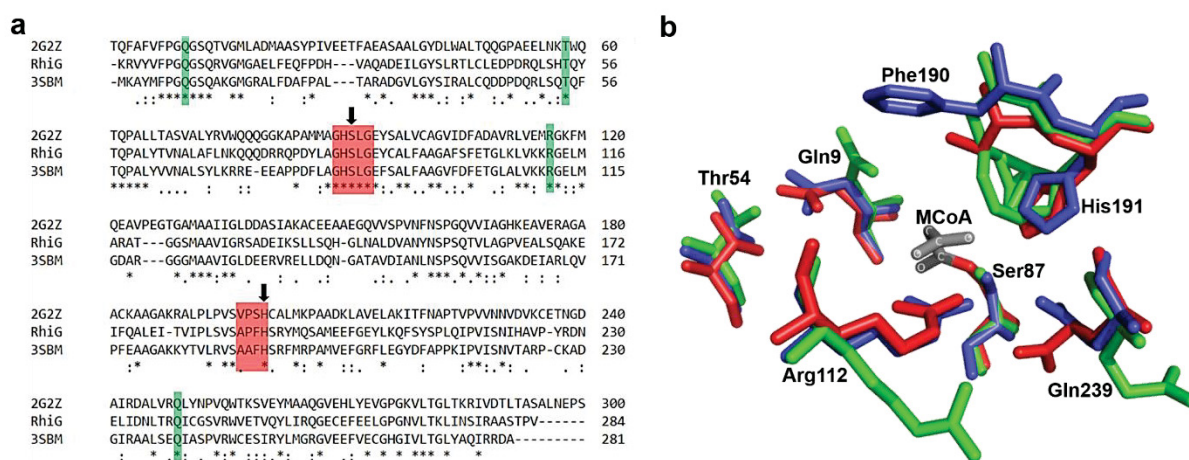


Figure 5.3 a) Sequence alignment of the malonyl-transferring ATs. 2GzZ – FabD from *E. coli*; 3SBM – AT from *trans*-AT disorazol synthase (DSZS). Fingerprint motifs are highlighted in red. Other active site residues are colored green. Arrows indicate the catalytic residues **b)** Alignment of the active site residues from FabD bound to malonyl-CoA (red), DSZS AT (blue) and RhiG (green). The structure of RhiG was modelled using Phyrez server and the figure was generated in Pymol.^{53, 177} The numbers correspond to the malonyl-transferring part of RhiG.

5.3 The structural role of the B domain

Although the inherent mosaic architecture of *trans*-AT PKSs render them suitable for engineering approaches to understand the role of novel domains, such enzymes have not been significantly targeted for genetic manipulation.¹⁷⁸ This holds particularly true for the cryptic B domain in the branching module of the Rhi PKS. Structural analysis of the B domain suggested that it has a double hotdog fold reminiscent of DH domains but has a distorted active site.¹⁴² In DH domains, an active site Asp residue has been proposed to donate a proton to the β -hydroxy unit and the catalytic His extracts an α -proton with the concomitant release of water, to yield an α,β -double bond.^{179, 180} Although a non-canonical DH has been reported to catalyze both dehydration and cyclization to yield a pyran ring,¹⁶⁵ all DH-catalyzed reactions are mediated by an invariant active site histidine. The catalytic His is mutated to Tyr in the B domain, which precludes the possibility that it might be involved in the branching reaction.¹⁴² Although the B domain shares a similar fold to DH domains, it lacks conserved motifs that are specific for a DH domain. In fact, alignment of the primary sequence of the B domain and the RhiE_DH domain of module 10 revealed an identity of only 23%. A BLAST search of the B domain against the PDB database resulted in no reliable hits. Domains homologous to the B domain have been speculated to be part of similar branching modules in the biosynthesis of glutarimide-containing antibiotics (9-methylstreptimidone, migrastatin, etc.).¹⁵⁸⁻¹⁶¹ A multiple sequence alignment of these homologues showed that the catalytic His mutation to Tyr was also observed in the 'X' (unknown) domains of 9-methylstreptimidone (SmdI_X) and migrastatin (Mgs_X) PKSs (**figure 5.4**). While the catalytic His is observed in the X domains of dorrigocin (Dgn_X) and lactimidomycin (Ltm_X) PKSs, these domains lack the catalytic Asp. The catalytic Asp was only conserved in SmdI_X apart from the rhizoxin PKS from *P. fluorescens* (RzxE_B). None of the Asp residues close to the active site of Mgs_X, Dgn_X, Ltm_X or ChxE_X could be rationalized for catalytic activity. In addition, other residues in the active site also exhibited significant variations.

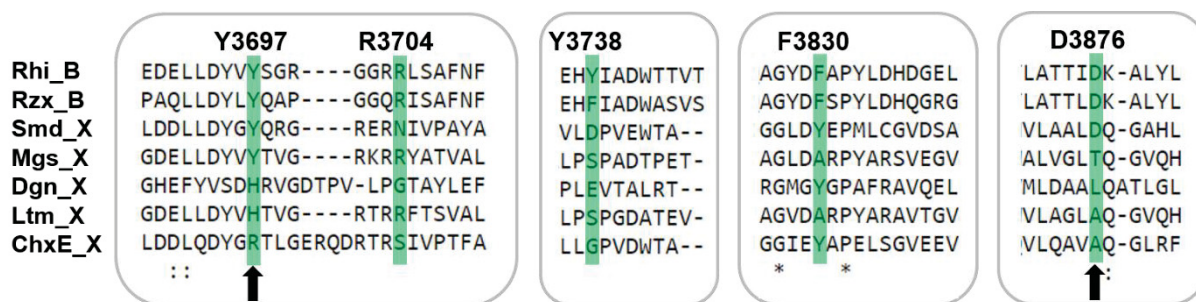


Figure 5.4 Multiple sequence alignment of B domain (CAL69893.1) with its homologous B/X domains. Arrows indicate the probable position of the catalytic residues. Active site residues are highlighted in green. The numbering on top corresponds to the B domain. Rzx – Rhizoxin pathway from *Pseudomonas fluorescens* (AAY92265.1); Smd – 9-Methylstreptimidone pathway from *Streptomyces himastatinicus* (CCC21123); Mgs – Migrastatin pathway from *S. platensis* subsp. *rosaceus* (ACY01390.1); Dgn – Dorrigocin pathway from *S. platensis* subsp. *rosaceus* (ACY01392.1); lactimidomycin pathway from *S. amphibiosporus* (ACY01400.1) and cycloheximide pathway from *Streptomyces* sp. YIM 56141 (AFO59866.1).

In addition to these discrepancies in the amino acid sequence, the crystal structure of the KS-B didomain showed that the active centers of KS and the B domain are separated by $> 40 \text{ \AA}$,¹⁴² a distance that might be too far to shuttle the intermediates. Also, point mutation of the conserved Asp3876 to Ala did not affect the branching activity. Based on these results, it was suggested that the B domain has a structural rather than a functional role in the branching reaction.¹⁴² The current study provides a direct biochemical proof that the B domain can be substituted by the SmdI_X or a regular DH domain (RhiE_DH10) with no influence on the substrate selectivity or turnover (**Manuscript C**).¹⁴⁵

Based on the sequence comparison results (**figure 5.4**), it seemed logical that the B domain substitution experiments should be based on the overall fold of the protein (**figure 5.5a**). The structural alignment of the KS-B domain against the KS-X counterparts revealed that although B and SmdI_X (or simply X) domains share only 25.8% sequence identity, the overall deviation based on the fold was only 1.8 \AA . The deviation was mainly observed in the KS-to-B linker region and the C-terminal part of the B domain. In assembly line PKSs, the domains are connected by linker regions which are not just covalent couplers. These linkers are very important for maintaining protein-protein interactions between neighbouring domains. Therefore domain substitution experiments should be performed without disrupting this interdomain communication.¹⁸¹ The linker region connecting KS-B and KS-X aligned well except for an additional α -helix and a β -sheet in KS-to-B linker, which might not be present in the KS-to-X linker (**figure 5.5b**). Therefore, for the B domain substitution experiments with X domain, the constructed chimera has the Rhi KS fused with the region spanning both the linker and X regions of the SmdI.¹⁵⁹ The second chimera was constructed by fusing the Rhi KS and KS-to-B linker region with a regular DH domain. For this purpose, the DH domain upstream of the branching module (RhiE_DH10) was used.¹²² When this construct was initially constructed without the KS-to-B linker, the protein could not be produced in a soluble form, proving that the linker region is also needed to obtain a soluble functional protein.

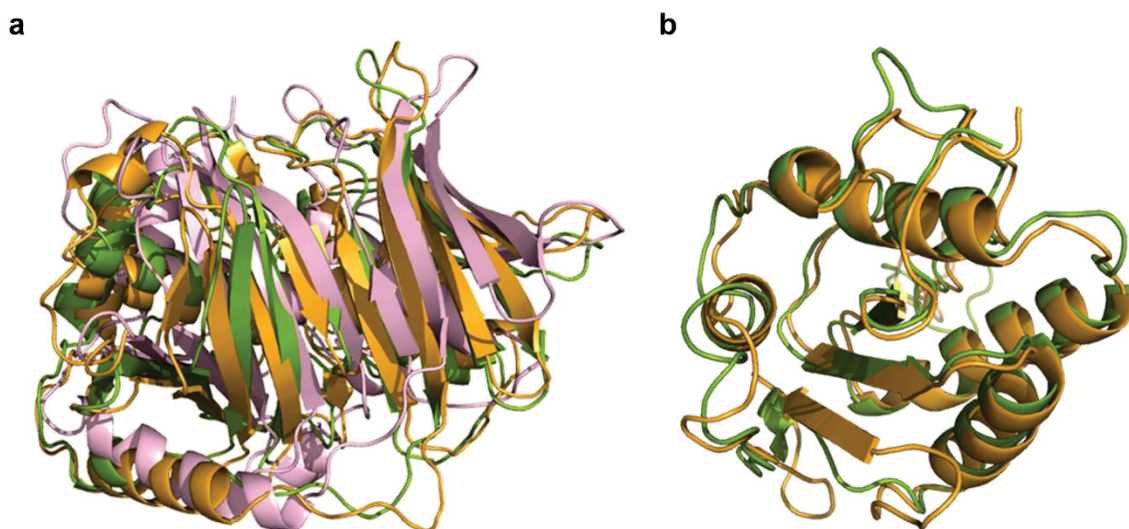


Figure 5.5 a) B domain (orange) aligned to X (SmdI_X – green; modelled) and DH (RhiE_DH10 – pink; modelled) shows the typical double hotdog fold where the central helix is wrapped around by six-stranded β -sheets.¹⁸² RMSD – 11.398 \AA . **b)** Alignment of the linker regions between the KS and the B/X domains from RhiE and SmdI. RMSD – 3.362 \AA . The figures are generated using Pymol⁵³

Even after careful consideration of the specific domain boundaries in the cloning experiments, the solubility of the chimeric proteins drastically reduced compared to the wild type KS-B. Therefore, to maximise the production of a soluble KS-X and KS-B hybrids, the fermentation condition had to be optimized. Therefore, gene induction experiments were performed at low temperature (15 °C) for 16 – 18 h. The proteins were also not suitable for long term storage and demanded frequent fresh preparations.

In the *in vitro* biotransformation experiments, the fact that both the chimeras could produce the lactone and glutarimide structures proves that, irrespective of B, X or DH domains, the branching reaction was not affected. These results suggest that the KS domain exclusively catalyzes the stereoselective β -branching and lactonization reactions. Since KS-ACP architectures are present in other PKS machineries,¹⁸³ attempts to produce a soluble Rhi KS-ACP didomain were made. However, owing to poor solubility, the protein could not be tested further. Taken together, our data indicate that the B domain as well as the adjacent KS-to-B linker region is needed to maintain the structural integrity of the KS-B didomain. Based on bioinformatic analysis, the linker regions C-terminal to the KS of trans-AT PKSs were hypothesized to be AT-docking domains (ATDs) that might be structurally similar to the KS-AT adapter regions of *cis*-AT PKSs.^{174, 184} However, because of the lack of functional characterization, this region was renamed to 'flanking subdomain'.¹⁰¹ Although the current result suggests that the linker region is needed to produce a soluble protein, whether it binds to AT remains to be determined. In addition to the mutation of the key catalytic histidine, the structure of the B domain suggests that the Tyr³⁷³⁸ residue is positioned such that it might block the substrate tunnel entrance found in characterized DH domains. This would imply that the substrates are unable to even enter the active site of the B domain. This observation is consistent with the ability of the KS domain to catalyse both steps of the branching reaction by itself. Considering all these results, the B domain appears to be a relic of a DH domain that has lost its activity over time. The B domain can possibly control the conformational changes that are needed during the interactions of the substrates with the KS domain. Inactive or non-catalytic domains are often found among PKSs, for example, a DH domain in the FK228 PKS,¹⁸⁵ a DH and a TE domain in the mycolactone PKS^{186, 187} and a DH, KR and ER domain in the rapamycin PKS.¹⁸⁸ Such domains might function as mediators to aid in communication between the upstream and downstream domains.

To summarise, the results from the B domain substitution experiments confirm the vital role played by the cryptic B and X domains to aid in structural stability and possibly interdomain communication. The valuable information gained on the role of B domains and the linker regions could prove useful in synthetic biology, helping to generate a functional PKS hybrid with the entire branching module in order to synthesize β -branch-containing polyketides.

5.4 Investigating the mechanism of the branching KS

As the B domain substitution experiments reveal that the KS alone is in charge of the branching reaction, the sequence of KS was examined to identify potential residues that might catalyze Michael addition and lactonization. A detailed phylogenetic analysis revealed that the KS of the branching module clustered among KSs (Clade IX) that accept olefinic substrates.⁸⁵ While the other KSs in the clade perform a regular Claisen condensation, the Rhi KS alone catalyzes a conjugate addition to form the non-canonical β -branch. Phylogenetic analysis of the KS domains from *cis*-AT PKSs, *trans*-AT PKSs and FASs showed that the known branching reliably KSs group together (bootstrap with 1000 replications) (**figure 5.6a**). Comparison of these sequences revealed that there are significant differences in the amino acids that are located close to the Cys-His-His catalytic triad. Site-specific mutation of residues in the Rhi branching KS to the corresponding residue of an elongating KS was performed in order to understand how the β -branching activity evolved. Furthermore, these results might also help assign specific roles for the amino acids in the Michael addition and lactonization reactions. Such an approach is common to study the PKS enzyme components and has been employed to decipher the reaction mechanism of various PKS domains as well as to alter substrate specificity, while still maintaining the protein-protein interactions.¹⁷⁸ Based on the mutational analysis of the Rhi KS, some of the additional residues that might involve in the reaction have been proposed (**Manuscript E**).

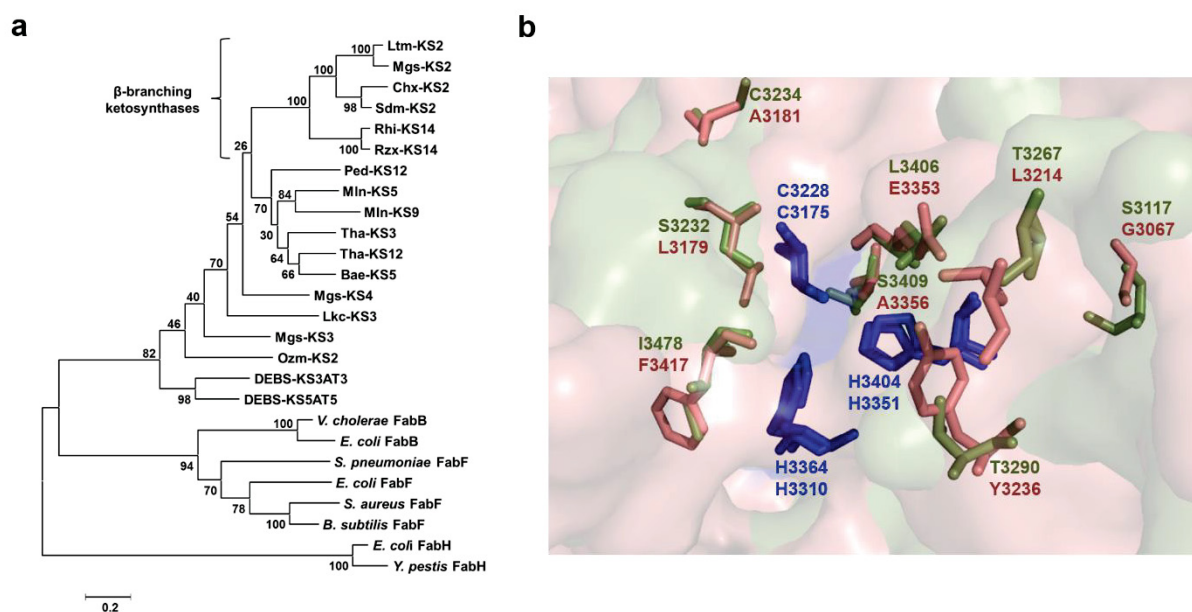


Figure 5.6 a) Phylogeny of KS domains. *Ltm-KS2* (ACY01400.1) from *Streptomyces amphibiosporus*, *Mgs-KS2* (ACY01390.1); *Mgs-KS3* and *Mgs-KS4* from *Streptomyces platensis*, *ChxE-KS2* (AFO59866.1) from *Streptomyces* sp., *Smd-KS2* (CCC21123) from *Streptomyces himastatinicus*, *Rhi-KS14* (CAL69893.1) from *Burkholderia rhizoxinica*, *Rzx-KS2* (AAY92265.1) from *Pseudomonas fluorescens*, *Ped-KS12* (AA547562.1) from *Pseudomonas* sp., *Mln-KS5* (YP_008420824); *Mln-KS9* (YP_001421032.1) and *Bae-KS5* (YP_001421293.1) from *Bacillus amyloliquefacians*, *Tha-KS3* (WP_011401356.1) and *Tha-KS12* (WP_009900597.1) from *Burkholderia thailandensis*, *Lkc-KS3* (ADN64228.1) from *Streptomyces rochei*, *Ozm-KS2* (4WKY) from *Streptomyces albus*, *DEBS* KSs (2QO3, 2HG4) from *Saccharopolyspora erythraea*, *FabB* (4XOX) from *Vibrio cholerae*, *FabB* (2VB8) from *E. coli*, *FabF* (1OXo) from *Streptococcus*

pneumoniae, *FabF* (2GFW) from *E. coli*, *FabF* (2GQD) from *Staphylococcus aureus*, *FabF* (4LS5) from *Bacillus subtilis*, *FabH* (5BNM) from *E. coli* and *FabH* (4R8E) from *Yersinia pestis*. The sequences were aligned using Muscle and the phylogenetic tree was constructed using Mega6 with 100 bootstrap replications.^{189, 190} **b**) Comparison of active site residues of Rhi-KS14 (green) and Tha-KS3 (pink). The structure of Tha-KS3 was modelled using Phyre2 server and the figure was generated in Pymol.^{53, 177}

Mutations of most of the active site residues did not affect the branching reaction (**Manuscript E**). However, for some mutants, the overall rate of the reaction was attenuated (Ser314Asp, Ile3478Phe and Ser3409Ala) while other mutants were completely inactive (Cys3228Ala and Thr3290Tyr). The complete inactivity of Thr3290Tyr mutant is intriguing. In β -ketoacyl synthase I (*FabB*) and II (*FabF*), this position is represented by Phe229 (Phe157 in *FabH*⁵⁷). Phe229 is one of the highly-conserved residues of the Fab-family of enzymes. This residue is located close to the carboxyl group of malonate and has been proposed to promote electron movement for decarboxylation and the formation of C2 carbanion.¹⁹¹⁻¹⁹⁵ Although the Thr3290Tyr mutant still retained the malonyl-ACP decarboxylation activity, it was not able to catalyze β -branch formation. One explanation for the inability to catalyze the conjugate addition is that Thr3290 facilitates the attack of the enolate intermediate to the β -position of the acyl thioester. In other words, Thr3290 might promote the C-C bond formation.

The crystal structure of KS-B bound the SNAC thioester substrate analog (**18**) showed positive electron density in the catalytic cysteine but the exact conformation of the substrate could not be determined. The bulkier Phe residues that lines the active site tunnel of *FabH* (Phe213, Phe157⁵⁷) are substituted with smaller Gly or Ala residues in the Rhi KS. Hence it is likely that there is more flexibility for processing an α,β -unsaturated substrate. Biotransformation and kinetic experiments demonstrated that His3364 and His3404 are critical for substrate turnover. While the His3364Ala and His3404Ala mutants retained measurable turnover activity after 16 hours, the rate of the reaction was slower compared to the wild type. Analogous mutations of the catalytic His residues of the KS of DEBS module 1 and *E. coli* *FabB* also resulted in moderate to very low turnover of the substrate.^{49, 196} However, a double mutant (His3364Ala and His3404Ala) was totally inactive, confirming the role of histidines in catalysis. From the mutational analyses, it is not possible to specify a particular residue that might induce lactonization by the attack of the δ -hydroxy group at this stage. However, the data provide a defined starting point for greater understanding of this unusual KS domain. A high resolution co-crystal structure of additional β -branching domains bound to reaction intermediates would help in the elucidation of the finer details of substrate binding and catalysis. To this end, attempts to obtain a crystal structure of a KS-B-ACP complex together with a crosslinking substrate (**21**) are in progress (**Unpublished results**).

To complement the mutational analysis and to strengthen the fact that the KS domain is needed for branching, the KS domain was also substituted with 6 other KSs that mediate the canonical Claisen condensation. However, as with the B domain substitution experiments, soluble protein production was not obtained for most of the KS chimeras. When the construction of the chimeras was repeated with KS-specific C-terminal flanking regions named LINKS,¹⁹⁷ soluble and functional proteins could be produced. **Figure 5.6b** shows the comparison of active site residues of the RhiE beta-branch KS and the canonical ThaKS3. Biotransformation assays revealed that the chimeras MlnKS4-B, BaeKS5*-B and ThaKS3*-B

produced a product that had the same mass and molecular formula as the native lactone (**20** in **figure 1.8a**), albeit with low turnover activity. However, the product eluted at a different time (~7.2 min) to that of the lactone **20** (~7.7 min). MS/MS experiments showed that the product has a distinct fragmentation pattern to that of the lactone (**20**) suggesting that it might be the linear or a cyclized product different from the δ -lactone. These results show that the native KS domain is needed to generate a β -branch followed by lactone formation. It also represents a successful example of manipulating the substrate specificity of the KS domains by careful domain substitution.

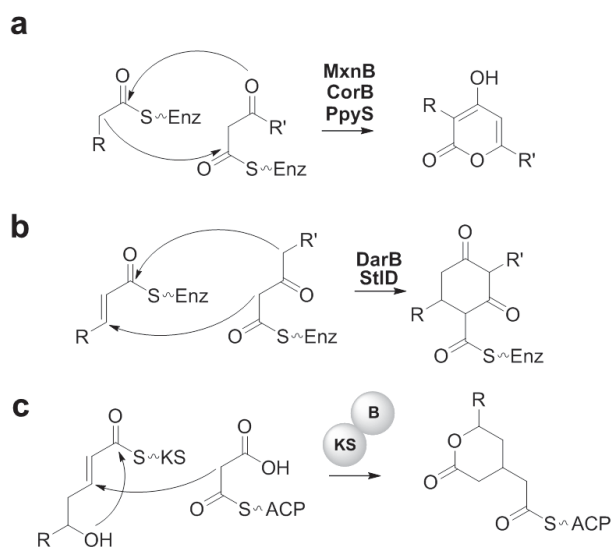


Figure 5.7 Schemes for novel KS-catalyzed condensation and cyclization

In light of these novel reactions, the Rhi KS holds a unique place and represents a new family of β -branching cyclase (**figure 5.7c**).

Over the recent years, ketosynthases that perform novel condensation and cyclization reactions have been reported. MxnB from *Myxococcus fulvus* Mxf50, CorB from *Coralloccoccus coralloides*, and PpyS from *Photorhabdus luminiscens* all catalyze the Claisen condensation reaction in head-to-head fusion of two polyketide chains followed by lactonization to yield rigid pyrone rings (**figure 5.7a**).¹⁹⁸⁻²⁰² In contrast, KSIII-type DarB from *Chitinophaga pinensis* and StlD from *Photorhabdus luminiscens* catalyze two distinct C-C bond formations – both a Claisen condensation and a Michael addition – of a β -ketoacyl substrate and an α,β -unsaturated precursor to produce cyclohexanediones (**figure 5.7b**).^{203, 204} In

5.5 Small- to large-sized lactone formation

The lactone motifs or related ring structures represent reactive centers for nucleophilic or electrophilic reactions in the target biomolecules. The presence of medium-sized rings might also provide structural rigidity to natural products which is crucial for their binding to biological targets. Typically, TE domains catalyze the lactonization reaction as a means to off-load a polyketide or a nonribosomal peptide intermediate, to generate medium- to large-sized macrocycles.²⁰⁵ Recently, several non-canonical cyclization reactions have been reported to occur either on- or off- the PKSs and NRPSs assembly lines, by specialized *cis*- or *trans*-acting domains and tailoring enzymes (**Manuscript D**).^{166, 206, 207} As the branching module of the rhizoxin PKS could form 6-membered lactam and glutarimide heterocycles, the ability to produce smaller or larger than the 6-membered lactones were tested with SNAC thioester mimics of varied acyl chain lengths. HRMS and MS/MS analysis revealed that the branching module might form up to 10-membered lactone macrocycles *in vitro* (**Manuscript F**). It is remarkable that the active site of KS can accommodate large-sized ring lactones. The crystal structure of an acyl-enzyme intermediate with the FabF-cerulinin complex revealed extensive movement of aromatic residues to aid in directing the acyl chain into the active site pocket. The Phe396 residue of the FabF moves away from the active site to leave space for substrate binding.⁴⁹ The substitution of Phe396 to Ala in the Rhi KS probably creates additional space for substrate binding as well as the formation of up to 10-membered cycles. Recently, Bae-KS5 has been shown to accept and catalyze elongation of shorter-to longer acyl chains. Mutation of a bulkier Met residue to Ala led to the elongation of β -methyl branched SNAC thioesters.²⁰⁸ The Rhi KS is inherently suitable to accept a wide range of non-natural substrates. Therefore, it might also be possible to generate large-sized lactams or other complex products *in vitro*. Medium- to large-sized lactones which show potential bioactivity do occur in nature, but 9- and 10-membered lactones are rare.²⁰⁹ The chemical synthesis of stereospecific lactones are also extremely challenging. Therefore, the production of a series of lactones by the branching module holds vital importance to use its enzyme components to generate complex products. This study provides insights into the extended substrate flexibility of the KS domain and opens possibilities to expand the range of products formed, for example, by site-specific mutagenesis.

6 Summary

The diversity of functional groups and stereochemical complexity of natural products continues to inspire synthetic chemists and biologists to develop pharmacophore mimics with improved bioactivity. Studying polyketide natural product biosynthetic enzymes has opened opportunities to engineer the assembly lines or individual modules/domains to generate non-natural variants. The biosynthesis of the macrolide rhizoxin varies from the archetypal head-to-tail condensation of acyl building blocks that yield exclusively linear chains. The pharmacophoric δ -lactone moiety of rhizoxin is generated by a rare β -branch formed in the branching module (KS-B-ACP) of a *trans*-AT type I PKS. The β -branch is formed as a result of a nucleophilic vinylogous addition of a C₂ unit to the α,β -unsaturated polyketide intermediate. This reaction is further followed by attack of the δ -hydroxy group of the intermediate to the KS-bound thioester to yield the δ -lactone. This thesis explains how the branching module of the rhizoxin biosynthetic pathway has been utilized for product diversification. It further provides a detailed analysis into the individual domains within the module. The results obtained from this thesis are summarized as follows.

- To extend the scope of the branching module, chemically synthesized SNAC thioester mimics with different δ -substituents have been tested *in vitro*. An amine-substituted substrate resulted in the formation of a δ -lactam whereas the carboxamide-substituted substrate yielded a glutarimide ring. These results, together with kinetic assays, showed that exclusively *syn*-substituted products are formed. This study is the first to verify that the biosynthesis of the glutarimide in cycloheximide-type of antibiotics follows a Michael addition-cyclization sequence catalyzed by the non-canonical KS domain. In addition to the formation of 6-membered lactone, lactam and glutarimide rings, this thesis also shows that the branching module could form an array of 5- to 10-membered lactones. It is intriguing that the active site of KS domain could accommodate rings of bigger sizes. It is highly possible that in future, varied ring sizes with substituted lactones or lactams could be formed by the branching module. Not only does the branching module form lactam, glutarimide and higher order lactones, but, provided an appropriate extender unit, can also stereoselectively form a dual branched product. These results (**Manuscript A**, **Manuscript B** and **Manuscript F**) provide proof that the branching module stereospecifically generates branched rings and is of particular interest in the field of synthetic biology to generate potential therapeutics. The work was done in collaboration with Dr. Tom Bretschneider, Dr. Daniel Heine, Tawatchai Thongkongkaew and Dr. Hak Joong Kim. As part of this study, a detailed review discussing novel on-line cyclization reactions of polyketide or nonribosomal peptide intermediates catalyzed by assembly line domains and tailoring enzymes has been published (**Manuscript D**).
- The PKSs for the biosynthesis of the glutarimide moiety of the cycloheximide-family of antibiotics possess a branching module homologous to the rhizoxin PKS. Prior to the

Summary

work described in this thesis, the role of the B domain and the homologous X domain in the branching modules remained obscure. Since the catalytic Asp to Ala mutant of the B domain retained the branching activity, it was proposed that these enigmatic domains do not participate directly in catalyzing the branching reaction. To shed light on the function of these domains, two different chimeras were constructed. The first chimeric didomain constitutes the KS from the branching module fused to the X domain from a glutarimide-forming module from 9-methylstreptimidone PKS. Biochemical characterization of the chimeric didomain showed that both the KS-B-ACP and KS-X-ACP modules are functionally equivalent. In addition, kinetic studies of the chimeras revealed that the B and X domains do not have any significant impact on substrate specificity or the product turnover. Irrespective of their origin, we found that the modules containing B or X domains show similar reaction rates and substrate affinities. The second chimeric didomain was generated by fusing the KS of the branching module with the structurally similar DH domain from the downstream module of the rhizoxin PKS. Most surprisingly, the KS-DH-ACP engineered module in which the B/X domains were exchanged for the DH domain also catalyzed the Michael addition - cyclization sequence with the same efficiency and specificity. The conserved Asp residues of the DH and B/X domains were mutated to investigate their potential catalytic role. The loss of the typical DH active site residues did not have any impact on the chain branching activity, proving that the aspartate-histidine dyad is not required for the branching reaction. In an attempt to observe the effect of a complete removal of the B domain, the KS and ACP were fused without the linker. However, the KS-ACP hybrid could not be produced as a soluble protein, hampering further analysis. The results from **Manuscript C** reveal that the B/X domains and the KS-to-B linker are critical for maintenance of the structural integrity of the KS-B didomain. The double hotdog fold of the B/X domains could be important for facilitating conformational changes between the adjacent domains. Surprisingly, the KS domain alone is capable of stereoselectively adding the C₂ unit to the β -carbon as well as cyclizing the polyketide intermediate to form the δ -lactone. The KS-mediated branching therefore represents a unique system to incorporate β -branches in polyketides to yield novel and potentially bioactive cyclic moieties.

- The KS domain of the branching module resembles a classical elongating KS with the well conserved Cys-His-His catalytic triad and a thiolase fold. Yet, it is solely capable of catalyzing the unprecedented β -branching and cyclization reactions. To know if there are other potential active site residues that might explain this novel reaction, site-specific mutagenesis of 10 conserved residues surrounding the active site was performed. The residues were chosen based on sequence comparison of KSs that catalyze β -branching or elongation. This is the first study reporting an extensive mechanistic investigation into the branching KS from a *trans*-AT PKS. As expected, based on previous studies on canonical KS domains, the catalytic Cys is essential for transacylation with the polyketide intermediate. Mutation of the catalytic histidines decreased the malonyl-CoA decarboxylation activity as well as product turnover. Kinetic analysis revealed that mutation of the residues Ser₃₁₄₁, Ser₃₄₀₉ and Ile₃₄₇₈ also decreased the rate of product formation. Like the catalytic Cys to Ala mutant, the

Summary

Thr3290Tyr mutation also rendered the protein inactive. Based on the analogous FAS structures, the Thr3290 might be positioned close to the malonyl carboxylate and thus aid in β -branch formation. The mutational analysis also indicated that Michael addition and the subsequent lactonization could be mediated by two or more residues. While elucidation of the mechanism requires further investigation, this study identified some key residues for catalysis and created a well-defined starting point to study the mechanistic details of the complex branching reaction. In addition to the mutational analysis, the KS domain was also substituted with elongating KSs from macrolactin (KS4), bacillaene (KS5) and thailandamide (KS3) PKSs. In all the chimeras constructed, the B domain and the linker region C-terminal to the KS were kept intact so as not to compromise structural integrity. The KS-swapped chimeras were functionally similar to a classical KS and catalyzed decarboxylative β -carbon elongation, leading to a proposed linear product. These results once again confirm that the active site architecture of the branching Rhi KS is very important to perform the Michael addition and lactonization (**Manuscript E**). A crystal structure of the entire branching module would help to elucidate the conformational changes occurring within the active site upon substrate binding.

To summarize, the results from this thesis validate that the branching module exhibits tolerance towards a variety of non-natural substrates to produce novel ring systems. It further provides a detailed functional analysis of its key components (KS and B) that could be leveraged for synthetic biology. The KS-B-ACP architecture serves as a hallmark to identify such non-canonical branching modules in natural product biosynthetic gene clusters.

7 Zusammenfassung

Die Vielfalt an funktionellen Gruppen und stereochemischer Komplexität von Naturstoffen inspiriert synthetische Chemiker und Biologen fortwährend dazu, Wirkstoffderivate mit verbesserter Bioaktivität zu entwickeln. Die Erforschung und das daraus folgende Verständnis über Polyketidsynthesen eröffnete Möglichkeiten, um Assemblierungslinien oder individuelle Module/Domänen zu entwickeln, welche dazu genutzt werden können, nicht natürlich vorkommende Polyketid-Varianten zu generieren. Die Biosynthese von Rhizoxin weicht von der archetypischen Kopf-Schwanz-Kondensation der Acyl-bildenden Einheiten ab, welche ausschließlich zur Produktion von linearen Ketten führt. Die pharmakophore δ -Lacton-Gruppe von Rhizoxin wird mittels einer seltenen β -Verzweigung im (KS-B-ACP) einer *trans*-AT Typ I PKS gebildet. Es konnte gezeigt werden, dass die β -Verzweigung aus einer nukleophilen vinylogenen Addition einer C₂-Einheit an einem α,β -ungesättigten Polyketid-Intermediat hervorgeht. An diese Reaktion schließt der Angriff der δ -Hydroxylgruppe des Intermediaten an den KS-gebundenem Thioester an, wodurch das δ -Lacton gebildet wird. Die vorliegende Arbeit beschreibt die Nutzung des β -Verzweigungs-Moduls in der Rhizoxin-PKS für die Herstellung variabler Produkte. Weiterhin wird eine detaillierte Analyse der einzelnen Domänen innerhalb des Moduls aufgezeigt. Im Folgenden werden die Ergebnisse der vorliegenden Arbeit zusammenfassend aufgeführt.

- Um die Substratflexibilität des β -Verzweigungs-Moduls zu analysieren wurden chemisch synthetisierte SNAC-Thioester-Analoga mit verschiedenen δ -Substituenten *in vitro* getestet. Während aus einem Amin-substituierten Substrat ein δ -Lactam hervorging, resultierte ein Carboxamid-substituiertes Substrat in einem Glutarimid-Ring. Diese Ergebnisse, zusammen mit den kinetischen Betrachtungen belegen, dass ausschließlich *syn*-substituierte Produkte gebildet werden. Diese Studie verifiziert zum ersten Mal, dass die Biosynthese von Glutarimid im Cycloheximid-Typ der Antibiotika einer Michael-Addition-Zyklisations-Sequenz folgt, welche von einer nicht-kanonischen KS-Domäne katalysiert wird. In dieser Arbeit konnte zusätzlich zu der Formation eines 6-gliedrigen Lactons, Lactams und Glutarimid-Rings gezeigt werden, dass der β -Verzweigungs-Modul weiterhin die Bildung einer Reihe von 5- bis 10-gliedrigen Lactonen katalysiert. Es ist erstaunlich, dass das aktive Zentrum der KS-Domäne auch Reaktionen mit größeren Ringen unterstützt. Es ist anzunehmen, dass in Zukunft verschiedene Ringgrößen mit substituierten Lactonen oder Laktamen durch das Verzweigungs-Modul gebildet werden können. Neben der Herstellung von Laktamen, Glutarimiden und Lactonen höherer Ordnung ist das Verzweigungs-Modul weiterhin in der Lage, stereoselektiv zweifach verzweigte Produkte bei entsprechend bereitgestellten Verlängerungseinheiten zu bilden. Diese Ergebnisse (**Manuskript A**, **Manuskript B** und **Manuskript F**) liefern den Beweis, dass das Verzweigungs-Modul stereospezifisch verzweigte Ringe bilden kann, was von besonderem Interesse für die synthetische Biologie ist, da dies für die Generierung von potentiellen Therapeutika genutzt werden könnte. Die Arbeit erfolgte in Kollaboration mit Dr. Tom Bretschneider, Dr. Daniel Heine, Tawatchai Thongkongkaew und Dr. Hak Joong

Kim. Als Teil dieser Studie wurde ein detaillierter Übersichtsartikel geliefert (**Manuskript D**), welcher neuartige *on-line* Zyklisierungsreaktionen der Polyketide oder nichtribosomalen Peptid-Intermediate beschreibt, die durch Domänen von Assemblierungslinien und modifizierende Enzyme katalysiert werden.

- PKS, welche die Biosynthese der Glutarimid-Einheit der Cycloheximid-Familie unter den Antibiotika katalysieren, enthalten ein zu den Rhizoxin-PKS homologes Verzweigungs-Modul. Im Vorfeld dieser Arbeit war die Rolle der B-Domäne des Verzweigungs-Modulen, und der homologen X-Domäne unklar. Da die Asp zu Ala-Mutante der B-Domäne die katalytische Aktivität der Verzweigungsreaktion bewahrte, wurde angenommen, dass diese enigmatische Domäne nicht direkt an der Verzweigungsreaktion teilnimmt. Um die Funktion dieser Domänen besser verstehen zu können, wurden zwei verschiedene Chimären generiert. Die erste chimäre Bidomäne besteht aus der KS-Domäne der Verzweigungs-Module welche mit der X-Domäne des Glutarimid-bildenden Moduls der 9-Methylstreptimidon-PKS fusioniert wurde. Die biochemische Charakterisierung der chimären Bidomänen zeigte auf, dass die Module KS-B-ACP- und KS-X-ACP- funktionell äquivalent sind. Kinetische Studien der Chimären zeigten außerdem, dass die B- und X-Domänen keinen signifikanten Einfluss auf die Substratspezifität oder die Produktbildungsrate haben. Ungeachtet derer Herkunft wurde gezeigt, dass die B- oder X-Domänen enthaltenen Module ähnliche Umsatzraten und Substrataffinitäten aufweisen. Die zweite chimäre Bidomäne wurde generiert, indem die KS-Domäne der Verzweigungs-Module mit der strukturell ähnlichen DH-Domäne aus dem darauffolgenden Modul der Rhizoxin-PKS fusioniert wurde. Überraschenderweise konnte gezeigt werden, dass das generierte KS-DH-ACP-Modul, indem die B/X-Domänen mit der DH-Domäne ausgetauscht wurden, ebenso die Michaelis Addition-Zyklisierungs-Sequenz in selber Effizienz und Spezifität katalysiert. Die konservierten Asp-Reste der DH- und B/X-Domänen wurden mutiert um deren potentiell katalytische Rolle zu untersuchen. Der Verlust der typischen aktiven Reste der DH-Domäne hatte keinen Einfluss auf die Kettenverzweigungs-Aktivität. Hiermit konnte bewiesen werden, dass die Aspartat-Histidin-Dyade nicht für die Verzweigungsreaktion benötigt wird. Um den Effekt einer kompletten Entfernung der B-Domäne zu untersuchen, wurden KS- und ACP-Domäne ohne einen Linker fusioniert. Jedoch konnte der KS-ACP-Hybrid nicht als lösliches Protein hergestellt werden, wodurch weitere Analysen nicht möglich waren. Die Ergebnisse des **Manuskripts C** zeigen auf, dass die B/X-Domänen und der Linker zwischen den Domänen KS und B essentiell für die Erhaltung der strukturellen Integrität der KS-B-Bidomäne sind. Der *double hotdog fold* der B/X-Domänen könnte bedeutend für eine erleichterte Konformationsänderung zwischen den benachbarten Domänen sein. Interessanterweise ist die KS-Domäne eigenständig zur stereoselektiven Addition der C₂-Einheit an den β -Kohlenstoff als auch zur Zyklisierung des Polyketid-Intermediates für die Ausbildung des δ -Lactons in der Lage. Die KS-vermittelte Verzweigung repräsentiert daher ein einzigartiges System für die Inkorporation von β -Verzweigungen in Polyketiden, welche zu neuartigen und potentiell bioaktiven zyklischen Einheiten führt.

- Die KS-Domäne der Verzweigungs-Module ähnelt mit ihrer konservierten katalytischen Triade aus Cystein und zwei Histidinen sowie der Thiolase-Faltung einer klassischen Elongations-KS. Dennoch ist sie eigenständig dazu in der Lage, die neuartige β -Verzweigung und eine Zyklisierungsreaktion zu katalysieren. Um herauszufinden, ob weitere potentiell aktive Seitenketten in dieser neuartigen Reaktion involviert sind, wurden Punktmutationen bei zehn konservierten Aminosäuren durchgeführt, welche das aktive Zentrum umgeben. Die Aminosäuren wurden anhand eines Sequenzvergleiches von KS-Domänen ausgewählt, welche β -Verzweigungen und β -Elongationen katalysieren. Dies ist die erste Studie über eine umfangreiche mechanistische Untersuchung von Verzweigungs-KS-Domänen einer *trans*-AT PKS. Entsprechend vorangegangener Studien an kanonischen KS-Domänen wurde bestätigt, dass das katalytisch aktive Cys essentiell für die Transacylierung mit Polyketid-Intermediaten ist. Die Mutation des katalytischen His verringerte sowohl die Malonyl-CoA Decarboxylierungs-Aktivität als auch die Produktbildungsrate. Kinetische Analysen zeigten, dass Mutationen an Ser₃₁₄₁, Ser₃₄₀₉ und Ile₃₄₇₈ ebenso zu einer reduzierten Produktbildungsrate führten. Neben der Mutation des katalytischen Cys zu Ala führte auch eine Thr₃₂₉₀Tyr-Mutation zu einem inaktiven Protein. Basierend auf der analogen FAS-Struktur, könnte Thr₃₂₉₀ in Nähe zu dem Malonylcarboxylat positioniert sein und daher die Formation der β -Verzweigung unterstützen. Mutationsanalysen weisen außerdem darauf hin, dass die Michael-Addition und die nachfolgende Lactonisierung durch zwei oder mehr Aminosäurereste vermittelt werden könnte. Währenddessen die Aufklärung des Mechanismus weiterer Untersuchungen bedarf, liefert diese Studie bereits die Identifikation einiger essentieller Aminosäurereste für die Katalyse und kreierte einen gut-definierten Startpunkt um mechanistische Details der komplexen Verzweigungsreaktion zu untersuchen. Zusätzlich zur Mutationsanalyse wurde die KS-Domäne außerdem durch Elongations-Ketosynthasen von Macrolactin (KS₄)-, Bacillaene (KS₅)- und Thailandamid (KS₃)-PKS ersetzt. In allen konstruierten Chimären wurden die B-Domäne und die Linker-Region C-terminal zur KS intakt gehalten, um die strukturelle Integrität nicht zu beeinträchtigen. Diese Chimären erwiesen sich funktionell als ähnlich zu einer klassischen KS und katalysierten die decarboxylative β -Kohlenstoff-Verlängerung, welche zu dem erwarteten linearen Produkt führt. Diese Ergebnisse bestätigen erneut, dass die Architektur des aktiven Zentrums der Verzweigungs-Rhi-KS-Domäne von hoher Bedeutung für die Ausführung der Michael-Addition und Lactonisierung ist (**Manuskript E**). Die Kristallstruktur der gesamten Verzweigungs-Module würde dazu beitragen, die konformationellen Veränderungen innerhalb des aktiven Zentrums während der Substratbindung aufzuklären.

Zusammenfassend bestätigen die Ergebnisse der vorliegenden Arbeit, dass das Verzweigungs-Modul eine Toleranz hinsichtlich einer Variabilität von nicht-natürlichen Substraten aufweist und somit die Produktion neuartiger Ringsysteme unterstützt. Die Arbeit liefert weiterhin eine detaillierte funktionelle Analyse derer Schlüsselkomponente (KS und B), was für die synthetische Biologie von Nutzen sein könnte. Die Architektur von KS-B-ACP dient als Kennzeichen für die Identifikation derartiger nicht-kanonischen Verzweigungs-Module in biosynthetischen Genclustern von Naturstoffen.

8 References

1. Newman, D.J., Cragg, G.M., and Snader, K.M., The influence of natural products upon drug discovery. *Nat Prod Rep*, 2000. **17**(3): p. 215-234.
2. Cragg, G.M., Newman, D.J., and Snader, K.M., Natural products in drug discovery and development. *J Nat Prod*, 1997. **60**(1): p. 52-60.
3. Clardy, J. and Walsh, C., Lessons from natural molecules. *Nature*, 2004. **432**(7019): p. 829-837.
4. Patridge, E., Gareiss, P., Kinch, M.S., and Hoyer, D., An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov Today*, 2016. **21**(2): p. 204-207.
5. Fischbach, M.A. and Walsh, C.T., Antibiotics for emerging pathogens. *Science*, 2009. **325**(5944): p. 1089-1093.
6. Koehn, F.E. and Carter, G.T., The evolving role of natural products in drug discovery. *Nat Rev Drug Discov*, 2005. **4**(3): p. 206-220.
7. Katz, L. and Baltz, R.H., Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol*, 2016. **43**(2-3): p. 155-176.
8. Clardy, J., Fischbach, M.A., and Walsh, C.T., New antibiotics from bacterial natural products. *Nat Biotechnol*, 2006. **24**(12): p. 1541-1550.
9. Baltz, R.H., Antimicrobials from actinomycetes: back to the future. *Microbe*, 2007. **2**(3): p. 125.
10. Li, J.W.-H. and Vederas, J.C., Drug discovery and natural products: end of an era or an endless frontier? *Science*, 2009. **325**(5937): p. 161-165.
11. Rodrigues, T., Reker, D., Schneider, P., and Schneider, G., Counting on natural products for drug design. *Nat Chem*, 2016. **8**(6): p. 531-541.
12. Schmitt, E.K., Hoepfner, D., and Krastel, P., Natural products as probes in pharmaceutical research. *J Ind Microbiol Biotechnol*, 2016. **43**(2-3): p. 249-260.
13. Harvey, A.L., Natural products in drug discovery. *Drug Discov Today*, 2008. **13**(19): p. 894-901.
14. Burg, R.W., Miller, B.M., Baker, E.E., Birnbaum, J., Currie, S.A., Hartman, R., Kong, Y.-L., Monaghan, R.L., Olson, G., and Putter, I., Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. *Antimicrob Agents Chemother*, 1979. **15**(3): p. 361-367.
15. World Health Organization, Programme for the elimination of neglected diseases in Africa (PENDA), 2016-2025. 2013.

References

16. Hoerauf, A., Filariasis: new drugs and new opportunities for lymphatic filariasis and onchocerciasis. *Curr Opin Infect Dis*, 2008. **21**(6): p. 673-681.
17. World Health Organization, WHO model list of essential medicines. 2013.
18. Masurekar, P.S., Fountoulakis, J.M., Hallada, T.C., Sosa, M.S., and Kaplan, L., Pneumocandins from *Zalerion arboricola*. *J Antibiot*, 1992. **45**(12): p. 1867-1874.
19. Cragg, G.M. and Newman, D.J., Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta*, 2013. **1830**(6): p. 3670-3695.
20. Gerth, K., Bedorf, N., Höfle, G., Irschik, H., and Reichenbach, H., Epothilons A and B: antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties. *J Antibiot*, 1996. **49**(6): p. 560-563.
21. Alvarez, R.H., Valero, V., and Hortobagyi, G.N., Ixabepilone for the treatment of breast cancer. *Ann Med*, 2011. **43**(6): p. 477-486.
22. Klayman, D.L., Qinghaosu (artemisinin): an antimalarial drug from China. *Science*, 1985. **228**(4703): p. 1049-1055.
23. Bhatt, S., Weiss, D., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K., Moyes, C., Henry, A., and Eckhoff, P., The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature*, 2015. **526**(7572): p. 207-211.
24. Paddon, C.J. and Keasling, J.D., Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol*, 2014. **12**(5): p. 355-367.
25. Roth, R.J. and Acton, N., A simple conversion of artemisinic acid into artemisinin. *J Nat Prod*, 1989. **52**(5): p. 1183-1185.
26. Butler, M.S., Robertson, A.A., and Cooper, M.A., Natural product and natural product derived drugs in clinical trials. *Nat Prod Rep*, 2014. **31**(11): p. 1612-1661.
27. Adrian, T.E., Novel marine-derived anti-cancer agents. *Curr Pharm Des*, 2007. **13**(33): p. 3417-3426.
28. D'Incalci, M. and Galmarini, C.M., A review of trabectedin (ET-743): a unique mechanism of action. *Mol Cancer Ther*, 2010. **9**(8): p. 2157-2163.
29. Cuevas, C. and Francesch, A., Development of Yondelis®(trabectedin, ET-743). A semisynthetic process solves the supply problem. *Nat Prod Rep*, 2009. **26**(3): p. 322-337.
30. Wani, M.C., Taylor, H.L., Wall, M.E., Coggon, P., and McPhail, A.T., Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc*, 1971. **93**(9): p. 2325-2327.
31. Weaver, B.A., How Taxol/paclitaxel kills cancer cells. *Mol Biol Cell*, 2014. **25**(18): p. 2677-2681.
32. Zhou, H., Xie, X., and Tang, Y., Engineering natural products using combinatorial biosynthesis and biocatalysis. *Curr Opin Biotechnol*, 2008. **19**(6): p. 590-596.

References

33. Tulp, M. and Bohlin, L., Functional versus chemical diversity: is biodiversity important for drug discovery? *Trends Pharmacol Sci*, 2002. **23**(5): p. 225-231.
34. Tibrewal, N. and Tang, Y., Biocatalysts for natural product biosynthesis. *Annu Rev Chem Biomol Eng*, 2014. **5**: p. 347-366.
35. Hertweck, C., The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed*, 2009. **48**(26): p. 4688-4716.
36. Staunton, J. and Weissman, K.J., Polyketide biosynthesis: a millennium review. *Nat Prod Rep*, 2001. **18**(4): p. 380-416.
37. Weissman, K.J., Introduction to polyketide biosynthesis. *Methods Enzymol*, 2009. **459**: p. 3-16.
38. Hopwood, D.A. and Sherman, D.H., Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu Rev Genet*, 1990. **24**(1): p. 37-62.
39. Moore, B.S. and Hertweck, C., Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat Prod Rep*, 2002. **19**(1): p. 70-99.
40. Ray, L. and Moore, B.S., Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Nat Prod Rep*, 2016. **33**(2): p. 150-161.
41. Chan, Y.A., Podevels, A.M., Kevany, B.M., and Thomas, M.G., Biosynthesis of polyketide synthase extender units. *Nat Prod Rep*, 2009. **26**(1): p. 90-114.
42. Khosla, C., Kapur, S., and Cane, D.E., Revisiting the modularity of modular polyketide synthases. *Curr Opin Chem Biol*, 2009. **13**(2): p. 135-143.
43. Meier, J.L. and Burkart, M.D., The chemical biology of modular biosynthetic enzymes. *Chem Soc Rev*, 2009. **38**(7): p. 2012-2045.
44. Horsman, M.E., Hari, T.P., and Boddy, C.N., Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat Prod Rep*, 2016. **33**(2): p. 183-202.
45. Keating, T.A. and Walsh, C.T., Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr Opin Chem Biol*, 1999. **3**(5): p. 598-606.
46. Du, L. and Lou, L., PKS and NRPS release mechanisms. *Nat Prod Rep*, 2010. **27**(2): p. 255-278.
47. Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E., and Flugel, R.S., Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr Opin Chem Biol*, 1997. **1**(3): p. 309-315.
48. Keatinge-Clay, A.T., The structures of type I polyketide synthases. *Nat Prod Rep*, 2012. **29**(10): p. 1050-1073.
49. Zhang, Y.-M., Hurlbert, J., White, S.W., and Rock, C.O., Roles of the active site water, histidine 303, and phenylalanine 396 in the catalytic mechanism of the elongation

References

- condensing enzyme of *Streptococcus pneumoniae*. *J Biol Chem*, 2006. **281**(25): p. 17390-17399.
50. Wang, J., Soisson, S.M., Young, K., Shoop, W., Kodali, S., Galgoci, A., Painter, R., Parthasarathy, G., Tang, Y.S., and Cummings, R., Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature*, 2006. **441**(7091): p. 358-361.
 51. Tang, Y., Chen, A.Y., Kim, C.-Y., Cane, D.E., and Khosla, C., Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem Biol*, 2007. **14**(8): p. 931-943.
 52. Tang, Y., Kim, C.-Y., Mathews, I.I., Cane, D.E., and Khosla, C., The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc Natl Acad Sci U S A*, 2006. **103**(30): p. 11124-11129.
 53. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.
 54. Witkowski, A., Joshi, A.K., and Smith, S., Mechanism of the β -ketoacyl synthase reaction catalyzed by the animal fatty acid synthase. *Biochemistry*, 2002. **41**(35): p. 10877-10887.
 55. Witkowski, A., Joshi, A.K., Lindqvist, Y., and Smith, S., Conversion of a β -ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine. *Biochemistry*, 1999. **38**(36): p. 11643-11650.
 56. White, S.W., Zheng, J., Zhang, Y.-M., and Rock, C.O., The structural biology of type II fatty acid biosynthesis. *Annu Rev Biochem*, 2005. **74**: p. 791-831.
 57. Davies, C., Heath, R.J., White, S.W., and Rock, C.O., The 1.8 Å crystal structure and active-site architecture of β -ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure*, 2000. **8**(2): p. 185-195.
 58. Shen, B., Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol*, 2003. **7**(2): p. 285-295.
 59. Smith, S. and Tsai, S.-C., The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep*, 2007. **24**(5): p. 1041-1072.
 60. Sherman, D.H., The Lego-ization of polyketide biosynthesis. *Nat Biotechnol*, 2005. **23**(9): p. 1083-1084.
 61. Rawlings, B.J., Type I polyketide biosynthesis in bacteria (part B). *Nat Prod Rep*, 2001. **18**(3): p. 231-281.
 62. Cortes, J., Haydock, S.F., Roberts, G.A., Bevitt, D.J., and Leadlay, P.F., An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature*, 1990. **348**(6297): p. 176-178.
 63. Weissman, K.J., Polyketide biosynthesis: understanding and exploiting modularity. *Philos Trans A Math Phys Eng Sci*, 2004. **362**(1825): p. 2671-2690.
 64. Fischbach, M.A. and Walsh, C.T., Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev*, 2006. **106**(8): p. 3468-3496.

References

65. Müller, R., Don't classify polyketide synthases. *Chem Biol*, 2004. **11**(1): p. 4-6.
66. Boettger, D., Bergmann, H., Kuehn, B., Shelest, E., and Hertweck, C., Evolutionary imprint of catalytic domains in fungal PKS–NRPS hybrids. *Chembiochem*, 2012. **13**(16): p. 2363-2373.
67. Chen, H. and Du, L., Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl Microbiol Biotechnol*, 2016. **100**(2): p. 541-557.
68. Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A., Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep*, 2007. **24**(1): p. 162-190.
69. Meurer, G., Gerlitz, M., Wendt-Pienkowski, E., Vining, L.C., Rohr, J., and Hutchinson, C.R., Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem Biol*, 1997. **4**(6): p. 433-443.
70. Strohl, W.R., Biotechnology of antibiotics. 2 ed. 1997: *CRC Press*.
71. Moore, B.S. and Hopke, J.N., Discovery of a new bacterial polyketide biosynthetic pathway. *Chembiochem*, 2001. **2**(1): p. 35-38.
72. Chen, H., Tseng, C.C., Hubbard, B.K., and Walsh, C.T., Glycopeptide antibiotic biosynthesis: enzymatic assembly of the dedicated amino acid monomer (*S*)-3, 5-dihydroxyphenylglycine. *Proc Natl Acad Sci U S A*, 2001. **98**(26): p. 14901-14906.
73. Gross, F., Luniak, N., Perlova, O., Gaitatzis, N., Jenke-Kodama, H., Gerth, K., Gottschalk, D., Dittmann, E., and Müller, R., Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. *Arch Microbiol*, 2006. **185**(1): p. 28-38.
74. Motamedi, H. and Shafiee, A., The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. *Eur J Biochem*, 1998. **256**(3): p. 528-534.
75. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B., Cloning and heterologous expression of the epothilone gene cluster. *Science*, 2000. **287**(5453): p. 640-642.
76. Zografos, A.L., From biosynthesis to total synthesis: Strategies and tactics for natural products. 1 ed. 2016: *Wiley*.
77. Albertini, A.M., Caramori, T., Scoffone, F., Scotti, C., and Galizzi, A., Sequence around the 159° region of the *Bacillus subtilis* genome: the pksX locus spans 33· 6 kb. *Microbiology*, 1995. **141**(2): p. 299-309.
78. Cheng, Y.-Q., Tang, G.-L., and Shen, B., Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc Natl Acad Sci U S A*, 2003. **100**(6): p. 3149-3154.
79. Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N., and Matsunaga, S., Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc Natl Acad Sci U S A*, 2004. **101**(46): p. 16222-16227.

References

80. Wenzel, S.C. and Müller, R., Formation of novel secondary metabolites by bacterial multimodular assembly lines: deviations from textbook biosynthetic logic. *Curr Opin Chem Biol*, 2005. **9**(5): p. 447-458.
81. Moldenhauer, J., Chen, X.H., Borriss, R., and Piel, J., Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the *trans*-AT family. *Angew Chem Int Ed*, 2007. **119**(43): p. 8343-8345.
82. O'brien, R.V., Davis, R.W., Khosla, C., and Hillenmeyer, M.E., Computational identification and analysis of orphan assembly-line polyketide synthases. *J Antibiot*, 2014. **67**(1): p. 89-97.
83. Piel, J., Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep*, 2010. **27**(7): p. 996-1047.
84. Helfrich, E.J. and Piel, J., Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep*, 2016. **33**(2): p. 231-316.
85. Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., and Piel, J., Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat Biotechnol*, 2008. **26**(2): p. 225-233.
86. Piel, J., Wen, G., Platzer, M., and Hui, D., Unprecedented diversity of catalytic domains in the first four modules of the putative pederin polyketide synthase. *ChemBiochem*, 2004. **5**(1): p. 93-98.
87. Chen, X.-H., Vater, J., Piel, J., Franke, P., Scholz, R., Schneider, K., Koumoutsis, A., Hitzeroth, G., Grammel, N., and Strittmatter, A.W., Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42. *J Bacteriol*, 2006. **188**(11): p. 4024-4036.
88. Moldenhauer, J., Götz, D.C., Albert, C.R., Bischof, S.K., Schneider, K., Süssmuth, R.D., Engeser, M., Gross, H., Bringmann, G., and Piel, J., The final steps of bacillaene biosynthesis in *Bacillus amyloliquefaciens* FZB42: Direct evidence for β , γ dehydration by a *trans*-acyltransferase polyketide synthase. *Angew Chem Int Ed*, 2010. **49**(8): p. 1465-1467.
89. Till, M. and Race, P., Progress challenges and opportunities for the re-engineering of *trans*-AT polyketide synthases. *Biotechnol Lett*, 2014. **36**(5): p. 877-888.
90. Maier, T., Jenni, S., and Ban, N., Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science*, 2006. **311**(5765): p. 1258-1262.
91. Maier, T., Leibundgut, M., and Ban, N., The crystal structure of a mammalian fatty acid synthase. *Science*, 2008. **321**(5894): p. 1315-1322.
92. Zheng, J., Gay, D.C., Demeler, B., White, M.A., and Keatinge-Clay, A.T., Divergence of multimodular polyketide synthases revealed by a didomain structure. *Nat Chem Biol*, 2012. **8**(7): p. 615-621.
93. Zheng, J., Fage, C.D., Demeler, B., Hoffman, D.W., and Keatinge-Clay, A.T., The missing linker: a dimerization motif located within polyketide synthase modules. *ACS Chem Biol*, 2013. **8**(6): p. 1263-1270.

References

94. Gay, D., You, Y.-O., Keatinge-Clay, A., and Cane, D.E., Structure and stereospecificity of the dehydratase domain from the terminal module of the rifamycin polyketide synthase. *Biochemistry*, 2013. **52**(49): p. 8916-8928.
95. Dutta, S., Whicher, J.R., Hansen, D.A., Hale, W.A., Chemler, J.A., Congdon, G.R., Narayan, A.R., Håkansson, K., Sherman, D.H., and Smith, J.L., Structure of a modular polyketide synthase. *Nature*, 2014. **510**(7506): p. 512-517.
96. Edwards, A.L., Matsui, T., Weiss, T.M., and Khosla, C., Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J Mol Biol*, 2014. **426**(11): p. 2229-2245.
97. Herbst, D.A., Jakob, R.P., Zähringer, F., and Maier, T., Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature*, 2016. **531**(7595): p. 533-537.
98. Davison, J., Dorival, J., Rabeharindranto, H., Mazon, H., Chagot, B., Gruez, A., and Weissman, K.J., Insights into the function of *trans*-acyl transferase polyketide synthases from the SAXS structure of a complete module. *Chem Sci*, 2014. **5**(8): p. 3081-3095.
99. Wong, F.T., Jin, X., Mathews, I.I., Cane, D.E., and Khosla, C., Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. *Biochemistry*, 2011. **50**(30): p. 6539-6548.
100. Haines, A.S., Dong, X., Song, Z., Farmer, R., Williams, C., Hothersall, J., Płoskoń, E., Wattana-amorn, P., Stephens, E.R., and Yamada, E., A conserved motif flags acyl carrier proteins for β -branching in polyketide synthesis. *Nat Chem Biol*, 2013. **9**(11): p. 685-692.
101. Gay, D.C., Gay, G., Axelrod, A.J., Jenner, M., Kohlhaas, C., Kampa, A., Oldham, N.J., Piel, J., and Keatinge-Clay, A.T., A close look at a ketosynthase from a *trans*-acyltransferase modular polyketide synthase. *Structure*, 2014. **22**(3): p. 444-451.
102. Epp, J.K., Huber, M., Turner, J., Goodson, T., and Schoner, B.E., Production of a hybrid macrolide antibiotic in *Streptomyces ambofaciens* and *Streptomyces lividans* by introduction of a cloned carbomycin biosynthetic gene from *Streptomyces thermotolerans*. *Gene*, 1989. **85**(2): p. 293-301.
103. Wong, F.T. and Khosla, C., Combinatorial biosynthesis of polyketides—a perspective. *Curr Opin Chem Biol*, 2012. **16**(1): p. 117-123.
104. Hertweck, C., Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem Sci*, 2015. **40**(4): p. 189-199.
105. McDaniel, R., Kao, C.M., Hwang, S.J., and Khosla, C., Engineered intermodular and intramodular polyketide synthase fusions. *Chem Biol*, 1997. **4**(9): p. 667-674.
106. Staunton, J., Combinatorial biosynthesis of erythromycin and complex polyketides. *Curr Opin Chem Biol*, 1998. **2**(3): p. 339-345.
107. Zhang, W. and Liu, J., Recent advances in understanding and engineering polyketide synthesis. *F1000Research*, 2016. **5**: F1000 Faculty Rev-208.
108. Gigant, B., Cormier, A., Dorléans, A., Ravelli, R., and Knossow, M., Microtubule-destabilizing agents: structural and mechanistic insights from the interaction of

References

- colchicine and vinblastine with tubulin, in *Tubulin-binding agents*. 2008, *Springer*. p. 259-278.
109. Iwasaki, S., Kobayashi, H., Furukawa, J., Namikoshi, M., Okuda, S., Sato, Z., Matsuda, I., and Noda, T., Studies on macrocyclic lactone antibiotics. VII. Structure of a phytotoxin" rhizoxin" produced by *Rhizopus chinensis*. *J Antibiot*, 1984. **37**(4): p. 354-362.
110. Iwasaki, S., Namikoshi, M., Kobayashi, H., Furukawa, J., Okuda, S., Itai, A., Kasuya, A., Iitaka, Y., and Sato, Z., Studies on macrocyclic lactone antibiotics. VIII. Absolute structures of rhizoxin and a related compound. *J Antibiot*, 1986. **39**(3): p. 424-429.
111. Takahashi, M., Iwasaki, S., Kobayashi, H., Okuda, S., Murai, T., Sato, Y., Haraguchi-Hiraoka, T., and Nagano, H., Studies on macrocyclic lactone antibiotics. XI. Antimitotic and anti-tubulin activity of new antitumor antibiotics, rhizoxin and its homologues. *J Antibiot*, 1987. **40**(1): p. 66-72.
112. Bai, R., Pettit, G.R., and Hamel, E., Binding of dolastatin 10 to tubulin at a distinct site for peptide antimitotic agents near the exchangeable nucleotide and vinca alkaloid sites. *J Biol Chem*, 1990. **265**(28): p. 17141-17149.
113. Tsuruo, T., Oh-hara, T., Iida, H., Tsukagoshi, S., Sato, Z., Matsuda, I., Iwasaki, S., Okuda, S., Shimizu, F., and Sasagawa, K., Rhizoxin, a macrocyclic lactone antibiotic, as a new antitumor agent against human and murine tumor cells and their vincristine-resistant sublines. *Cancer Res*, 1986. **46**(1): p. 381-385.
114. Hong, J. and White, J.D., The chemistry and biology of rhizoxins, novel antitumor macrolides from *Rhizopus chinensis*. *Tetrahedron*, 2004. **60**(27): p. 5653-5681.
115. Nakada, M., Kobayashi, S., Iwasaki, S., and Ohno, M., The first total synthesis of the antitumor macrolide rhizoxin: synthesis of the key building blocks. *Tetrahedron Lett*, 1993. **34**(6): p. 1035-1038.
116. Keck, G.E., Wager, C.A., Wager, T.T., Savin, K.A., Covel, J.A., McLaws, M.D., Krishnamurthy, D., and Cee, V.J., Asymmetric total synthesis of rhizoxin D. *Angew Chem Int Ed*, 2001. **40**(1): p. 231-234.
117. Kobayashi, H., Iwasaki, S., Yamada, E., and Okuda, S., Biosynthesis of the antimitotic antitumor antibiotic rhizoxin by *Rhizopus chinensis*; Origins of the carbon atoms. *J Chem Soc Chem Commun*, 1986: p. 1702-1703.
118. Partida-Martinez, L.P. and Hertweck, C., Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature*, 2005. **437**(7060): p. 884-888.
119. Scherlach, K., Partida-Martinez, L.P., Dahse, H.-M., and Hertweck, C., Antimitotic rhizoxin derivatives from a cultured bacterial endosymbiont of the rice pathogenic fungus *Rhizopus microsporus*. *J Am Chem Soc*, 2006. **128**(35): p. 11529-11536.
120. Lackner, G., Möbius, N., Scherlach, K., Partida-Martinez, L.P., Winkler, R., Schmitt, I., and Hertweck, C., Global distribution and evolution of a toxinogenic *Burkholderia-Rhizopus* symbiosis. *Appl Environ Microbiol*, 2009. **75**(9): p. 2982-2986.
121. Piel, J., Metabolites from symbiotic bacteria. *Nat Prod Rep*, 2009. **26**(3): p. 338-362.

References

122. Partida-Martinez, L.P. and Hertweck, C., A gene cluster encoding rhizoxin biosynthesis in "*Burkholderia rhizoxina*", the bacterial endosymbiont of the fungus *Rhizopus microsporus*. *Chembiochem*, 2007. **8**(1): p. 41-45.
123. Brendel, N., Partida-Martinez, L.P., Scherlach, K., and Hertweck, C., A cryptic PKS-NRPS gene locus in the plant commensal *Pseudomonas fluorescens* Pf-5 codes for the biosynthesis of an antimetabolic rhizoxin complex. *Org Biomol Chem*, 2007. **5**(14): p. 2211-2213.
124. Loper, J.E., Henkels, M.D., Shaffer, B.T., Valeriote, F.A., and Gross, H., Isolation and identification of rhizoxin analogs from *Pseudomonas fluorescens* Pf-5 by using a genomic mining strategy. *Appl Environ Microbiol*, 2008. **74**(10): p. 3085-3093.
125. Scherlach, K., Busch, B., Lackner, G., Paszkowski, U., and Hertweck, C., Symbiotic cooperation in the biosynthesis of a phytotoxin. *Angew Chem Int Ed*, 2012. **124**(38): p. 9753-9756.
126. Prota, A.E., Bargsten, K., Diaz, J.F., Marsh, M., Cuevas, C., Liniger, M., Neuhaus, C., Andreu, J.M., Altmann, K.-H., and Steinmetz, M.O., A new tubulin-binding site and pharmacophore for microtubule-destabilizing anticancer drugs. *Proc Natl Acad Sci U S A*, 2014. **111**(38): p. 13817-13821.
127. Schmitt, I., Partida-Martinez, L.P., Winkler, R., Voigt, K., Einax, E., Dölz, F., Telle, S., Wöstemeyer, J., and Hertweck, C., Evolution of host resistance in a toxin-producing bacterial-fungal alliance. *ISME J*, 2008. **2**(6): p. 632-641.
128. Takahashi, M., Kobayashi, M., and Iwasaki, S., Rhizoxin resistant mutants with an altered β -tubulin gene in *Aspergillus nidulans*. *Mol Gen Genet*, 1989. **220**(1): p. 53-59.
129. Takahashi, M., Matsumoto, S., Iwasaki, S., and Yahara, I., Molecular basis for determining the sensitivity of eucaryotes to the antimetabolic drug rhizoxin. *Mol Gen Genet*, 1990. **222**(2-3): p. 169-175.
130. Hamel, E., Natural products which interact with tubulin in the vinca domain: maytansine, rhizoxin, phomopsin A, dolastatins 10 and 15 and halichondrin B. *Pharmacol Ther*, 1992. **55**(1): p. 31-51.
131. Kusebauch, B., Scherlach, K., Kirchner, H., Dahse, H.M., and Hertweck, C., Antiproliferative effects of ester- and amide-functionalized rhizoxin derivatives. *ChemMedChem*, 2011. **6**(11): p. 1998-2001.
132. Liscombe, D.K., Louie, G.V., and Noel, J.P., Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Nat Prod Rep*, 2012. **29**(10): p. 1238-1250.
133. Miller, D.A., Luo, L., Hillson, N., Keating, T.A., and Walsh, C.T., Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem Biol*, 2002. **9**(3): p. 333-344.
134. Butcher, R.A., Schroeder, F.C., Fischbach, M.A., Straight, P.D., Kolter, R., Walsh, C.T., and Clardy, J., The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proc Natl Acad Sci U S A*, 2007. **104**(5): p. 1506-1509.

References

135. Gurney, R. and Thomas, C.M., Mupirocin: biosynthesis, special features and applications of an antibiotic from a Gram-negative bacterium. *Appl Microbiol Biotechnol*, 2011. **90**(1): p. 11-21.
136. Moebius, N., Ross, C., Scherlach, K., Rohm, B., Roth, M., and Hertweck, C., Biosynthesis of the respiratory toxin bongkrekic acid in the pathogenic bacterium *Burkholderia gladioli*. *Chem Biol*, 2012. **19**(9): p. 1164-1174.
137. Maloney, F.P., Gerwick, L., Gerwick, W.H., Sherman, D.H., and Smith, J.L., Anatomy of the β -branching enzyme of polyketide biosynthesis and its interaction with an acyl-ACP substrate. *Proc Natl Acad Sci U S A*, 2016. **113**(37): p. 10316-10321.
138. Miziorko, H.M., Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch Biochem Biophys*, 2011. **505**(2): p. 131-143.
139. Geders, T.W., Gu, L., Mowers, J.C., Liu, H., Gerwick, W.H., Håkansson, K., Sherman, D.H., and Smith, J.L., Crystal structure of the ECH₂ catalytic domain of CurF from *Lyngbya majuscula* - Insights into a decarboxylase involved in polyketide chain β -branching. *J Biol Chem*, 2007. **282**(49): p. 35954-35963.
140. Calderone, C.T., Kowtoniuk, W.E., Kelleher, N.L., Walsh, C.T., and Dorrestein, P.C., Convergence of isoprene and polyketide biosynthetic machinery: isoprenyl-S-carrier proteins in the pksX pathway of *Bacillus subtilis*. *Proc Natl Acad Sci U S A*, 2006. **103**(24): p. 8977-8982.
141. Kusebauch, B., Busch, B., Scherlach, K., Roth, M., and Hertweck, C., Polyketide-chain branching by an enzymatic Michael addition. *Angew Chem Int Ed*, 2009. **48**(27): p. 5001-5004.
142. Bretschneider, T., Heim, J.B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C., Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature*, 2013. **502**(7469): p. 124-128.
143. Weissman, K.J., Uncovering the structures of modular polyketide synthases. *Nat Prod Rep*, 2015. **32**(3): p. 436-453.
144. Nguyen, C., Haushalter, R.W., Lee, D.J., Markwick, P.R., Bruegger, J., Caldara-Festin, G., Finzel, K., Jackson, D.R., Ishikawa, F., and O'Dowd, B., Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature*, 2014. **505**(7483): p. 427-431.
145. Sundaram, S., Heine, D., and Hertweck, C., Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching. *Nat Chem Biol*, 2015. **11**: p. 949-951.
146. Smith, J.L. and Sherman, D.H., An enzyme assembly line. *Science*, 2008. **321**(5894): p. 1304-1305.
147. Franke, J. and Hertweck, C., Biomimetic thioesters as probes for enzymatic assembly lines: synthesis, applications, and challenges. *Cell Chem Biol*, 2016. **23**(10): p. 1179-1192.
148. Heine, D., Bretschneider, T., Sundaram, S., and Hertweck, C., Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles. *Angew Chem Int Ed*, 2014. **53**(43): p. 11645-11649.

References

149. Jenner, M., Frank, S., Kampa, A., Kohlhaas, C., Pöplau, P., Briggs, G.S., Piel, J., and Oldham, N.J., Substrate specificity in ketosynthase domains from *trans*-AT polyketide synthases. *Angew Chem Int Ed*, 2013. **52**(4): p. 1143-1147.
150. Kornfeld, E.C., Jones, R.G., and Parke, T.V., The structure and chemistry of actidione, an antibiotic from *Streptomyces griseus*. *J Am Chem Soc*, 1949. **71**(1): p. 150-159.
151. Woo, E.J., Starks, C.M., Carney, J.R., Arslanian, R., Cadapan, L., Zavala, S., and Licari, P., Migrastatin and a new compound, isomigrastatin, from *Streptomyces platensis*. *J Antibiot*, 2002. **55**(2): p. 141-146.
152. Nakae, K., Yoshimoto, Y., Sawa, T., Homma, Y., Hamada, M., Takeuch, T., and Imoto, M., Migrastatin, a new inhibitor of tumor cell migration from *Streptomyces* sp. MK929-43F1. Taxonomy, fermentation, isolation and biological activities. *J Antibiot*, 2000. **53**(10): p. 1130-1136.
153. Kadam, S. and Mcalpine, J.B., Dorrigocins: novel antifungal antibiotics that change the morphology of ras-transformed NIH/3T3 cells to that of normal cells. III. Biological properties and mechanism of action. *J Antibiot*, 1994. **47**(8): p. 875-880.
154. Sugawara, K., Nishiyama, Y., Toda, S., Komiyama, N., Hatori, M., Moriyama, T., Sawada, Y., Kamei, H., Konishi, M., and Oki, T., Lactimidomycin, a new glutarimide group antibiotic. Production, isolation, structure and biological activity. *J Antibiot*, 1992. **45**(9): p. 1433-1441.
155. Saito, N., Kitame, F., Kikuchi, M., and Ishida, N., Studies on a new antiviral antibiotic, 9-methyl streptimidone. Physicochemical and biological properties. *J Antibiot*, 1974. **27**(3): p. 206-214.
156. Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O., Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol*, 2010. **6**(3): p. 209-217.
157. Obrig, T.G., Culp, W.J., McKeehan, W.L., and Hardesty, B., The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *J Biol Chem*, 1971. **246**(1): p. 174-181.
158. Lim, S.-K., Ju, J., Zazopoulos, E., Jiang, H., Seo, J.-W., Chen, Y., Feng, Z., Rajski, S.R., Farnet, C.M., and Shen, B., iso-Migrastatin, migrastatin, and dorrigocin production in *Streptomyces platensis* NRRL 18993 is governed by a single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase. *J Biol Chem*, 2009. **284**(43): p. 29746-29756.
159. Wang, B., Song, Y., Luo, M., Chen, Q., Ma, J., Huang, H., and Ju, J., Biosynthesis of 9-methylstreptimidone involves a new decarboxylative step for polyketide terminal diene formation. *Org Lett*, 2013. **15**(6): p. 1278-1281.
160. Yin, M., Yan, Y., Lohman, J.R., Huang, S.-X., Ma, M., Zhao, G.-R., Xu, L.-H., Xiang, W., and Shen, B., Cycloheximide and actiphenol production in *Streptomyces* sp. YIM56141 governed by single biosynthetic machinery featuring an acyltransferase-less type I polyketide synthase. *Org Lett*, 2014. **16**(11): p. 3072-3075.
161. Seo, J.-W., Ma, M., Kwong, T., Ju, J., Lim, S.-K., Jiang, H., Lohman, J.R., Yang, C., Cleveland, J., and Zazopoulos, E., Comparative characterization of the lactimidomycin

References

- and iso-migrastatin biosynthetic machineries revealing unusual features for acyltransferase-less type I polyketide synthases and providing an opportunity to engineer new analogues. *Biochemistry*, 2014. **53**(49): p. 7854-7865.
162. Andrioli, W.J., Santos, M., Silva, V., Oliveira, R., Chagas-Paula, D., Jorge, J., Furtado, N.A.J.C., Pupo, M.T., Silva, C.H.T.d.P.d., and Naal, R.M.Z.G., δ -Lactam derivative from thermophilic soil fungus exhibits in vitro anti-allergic activity. *Nat Prod Res*, 2012. **26**(23): p. 2168-2175.
163. Gopichand, Y. and Schmitz, F.J., Two novel lactams from the marine sponge *Halichondria melanodocia*. *J Org Chem*, 1979. **44**(26): p. 4995-4997.
164. Jang, J.-H., Kanoh, K., Adachi, K., and Shizuri, Y., Awajanomycin, a cytotoxic γ -lactone- δ -lactam metabolite from marine-derived *Acremonium* sp. AWA16-1. *J Nat Prod*, 2006. **69**(9): p. 1358-1360.
165. Berkhan, G. and Hahn, F., A dehydratase domain in ambruticin biosynthesis displays additional activity as a pyran-forming cyclase. *Angew Chem Int Ed*, 2014. **53**(51): p. 14240-14244.
166. Sundaram, S. and Hertweck, C., On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems. *Curr Opin Chem Biol*, 2016. **31**: p. 82-94.
167. Zhao, C., Coughlin, J.M., Ju, J., Zhu, D., Wendt-Pienkowski, E., Zhou, X., Wang, Z., Shen, B., and Deng, Z., Oxazolomycin biosynthesis in *Streptomyces albus* JA3453 featuring an "acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. *J Biol Chem*, 2010. **285**(26): p. 20097-20108.
168. Musiol, E.M., Härtner, T., Kulik, A., Moldenhauer, J., Piel, J., Wohlleben, W., and Weber, T., Supramolecular templating in kirromycin biosynthesis: the acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the *trans*-AT PKS. *Chem Biol*, 2011. **18**(4): p. 438-444.
169. Menche, D., Arikan, F., Perlova, O., Horstmann, N., Ahlbrecht, W., Wenzel, S.C., Jansen, R., Irschik, H., and Müller, R., Stereochemical determination and complex biosynthetic assembly of etnangien, a highly potent RNA polymerase inhibitor from the myxobacterium *Sorangium cellulosum*. *J Am Chem Soc*, 2008. **130**(43): p. 14234-14243.
170. Heine, D., Sundaram, S., Bretschneider, T., and Hertweck, C., Twofold polyketide branching by a stereoselective enzymatic Michael addition. *Chem Commun*, 2015. **51**(48): p. 9872-9875.
171. Lopanik, N.B., Shields, J.A., Buchholz, T.J., Rath, C.M., Hothersall, J., Haygood, M.G., Håkansson, K., Thomas, C.M., and Sherman, D.H., *In vivo* and *in vitro* trans-acylation by BryP, the putative bryostatin pathway acyltransferase derived from an uncultured marine symbiont. *Chem Biol*, 2008. **15**(11): p. 1175-1186.
172. Poust, S., Phelan, R.M., Deng, K., Katz, L., Petzold, C.J., and Keasling, J.D., Divergent mechanistic routes for the formation of gem-dimethyl groups in the biosynthesis of complex polyketides. *Angew Chem Int Ed*, 2015. **127**(8): p. 2400-2403.
173. Starcevic, A., Zucko, J., Simunkovic, J., Long, P.F., Cullum, J., and Hranueli, D., ClustScan: an integrated program package for the semi-automatic annotation of

References

- modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res*, 2008. **36**(21): p. 6882-6892.
174. Yadav, G., Gokhale, R.S., and Mohanty, D., Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J Mol Biol*, 2003. **328**(2): p. 335-363.
175. Oefner, C., Schulz, H., D'Arcy, A., and Dale, G.E., Mapping the active site of *Escherichia coli* malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography. *Acta Crystallogr D Biol Crystallogr*, 2006. **62**(6): p. 613-618.
176. Dunn, B.J. and Khosla, C., Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *J R Soc Interface*, 2013. **10**(85): p. 20130297.
177. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J., The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, 2015. **10**(6): p. 845-858.
178. Weissman, K.J., Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Nat Prod Rep*, 2016. **33**(2): p. 203-230.
179. Keatinge-Clay, A., Crystal structure of the erythromycin polyketide synthase dehydratase. *J Mol Biol*, 2008. **384**(4): p. 941-953.
180. Akey, D.L., Razelun, J.R., Tehranisa, J., Sherman, D.H., Gerwick, W.H., and Smith, J.L., Crystal structures of dehydratase domains from the curacin polyketide biosynthetic pathway. *Structure*, 2010. **18**(1): p. 94-105.
181. Gokhale, R.S. and Khosla, C., Role of linkers in communication between protein modules. *Curr Opin Chem Biol*, 2000. **4**(1): p. 22-27.
182. Pidugu, L.S., Maity, K., Ramaswamy, K., Surolia, N., and Suguna, K., Analysis of proteins with the 'hot dog' fold: Prediction of function and identification of catalytic residues of hypothetical proteins. *BMC Struct Biol*, 2009. **9**(1): p. 1.
183. Huang, Y., Tang, G.-L., Pan, G., Chang, C.-Y., and Shen, B., Characterization of the ketosynthase and acyl carrier protein domains at the LnmI nonribosomal peptide synthetase-polyketide synthase interface for leinamycin biosynthesis. *Org Lett*, 2016. **18**(17): p. 4288-4291.
184. Tang, G.-L., Cheng, Y.-Q., and Shen, B., Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. *Chem Biol*, 2004. **11**(1): p. 33-45.
185. Wesener, S.R., Potharla, V.Y., and Cheng, Y.-Q., Reconstitution of the FK228 biosynthetic pathway reveals cross talk between modular polyketide synthases and fatty acid synthase. *Appl Environ Microbiol*, 2011. **77**(4): p. 1501-1507.
186. Stinear, T.P., Mve-Obiang, A., Small, P.L., Frigui, W., Pryor, M.J., Brosch, R., Jenkin, G.A., Johnson, P.D., Davies, J.K., and Lee, R.E., Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc Natl Acad Sci U S A*, 2004. **101**(5): p. 1345-1349.

References

187. Bali, S. and Weissman, K.J., Ketoreduction in mycolactone biosynthesis: insight into substrate specificity and stereocontrol from studies of discrete ketoreductase domains *in vitro*. *Chembiochem*, 2006. **7**(12): p. 1935-1942.
188. Schwecke, T., Aparicio, J.F., Molnar, I., König, A., Khaw, L.E., Haydock, S.F., Oliynyk, M., Caffrey, P., Cortes, J., and Lester, J.B., The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc Natl Acad Sci U S A*, 1995. **92**(17): p. 7839-7843.
189. Edgar, R.C., MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, 2004. **32**(5): p. 1792-1797.
190. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S., MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*, 2013. **30**(12): p. 2725-2729.
191. Price, A.C., Choi, K.-H., Heath, R.J., Li, Z., White, S.W., and Rock, C.O., Inhibition of β -ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin structure and mechanism. *J Biol Chem*, 2001. **276**(9): p. 6551-6559.
192. Waller, R.F., Ralph, S.A., Reed, M.B., Su, V., Douglas, J.D., Minnikin, D.E., Cowman, A.F., Besra, G.S., and McFadden, G.I., A type II pathway for fatty acid biosynthesis presents drug targets in *Plasmodium falciparum*. *Antimicrob Agents Chemother*, 2003. **47**(1): p. 297-301.
193. Price, A.C., Rock, C.O., and White, S.W., The 1.3-Angstrom-resolution crystal structure of β -ketoacyl-acyl carrier protein synthase II from *Streptococcus pneumoniae*. *J Bacteriol*, 2003. **185**(14): p. 4136-4143.
194. Olsen, J.G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., Lindquist, Y., and Larsen, S., The X-ray crystal structure of β -ketoacyl [acyl carrier protein] synthase I. *FEBS Lett*, 1999. **460**(1): p. 46-52.
195. Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G., and Lindqvist, Y., Crystal structure of β -ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes. *EMBO J*, 1998. **17**(5): p. 1183-1191.
196. Robbins, T., Kapilivsky, J., Cane, D.E., and Khosla, C., Roles of conserved active site residues in the ketosynthase domain of an assembly line polyketide synthase. *Biochemistry*, 2016. **55**(32): p. 4476-4484.
197. Gay, D.C., Wagner, D.T., Meinke, J.L., Zogzas, C.E., Gay, G.R., and Keatinge-Clay, A.T., The LINKS motif zippers *trans*-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *J Struct Biol*, 2016. **193**(3): p. 196-205.
198. Kresovic, D., Schempp, F., Cheikh-Ali, Z., and Bode, H.B., A novel and widespread class of ketosynthase is responsible for the head-to-head condensation of two acyl moieties in bacterial pyrone biosynthesis. *Beilstein J Org Chem*, 2015. **11**(1): p. 1412-1417.
199. Zocher, G., Vilstrup, J., Heine, D., Hallab, A., Goralski, E., Hertweck, C., Stahl, M., Schäberle, T.F., and Stehle, T., Structural basis of head to head polyketide fusion by CorB. *Chem Sci*, 2015. **6**(11): p. 6525-6536.

References

200. Erol, Ö., Schäberle, T.F., Schmitz, A., Rachid, S., Gurgui, C., El Omari, M., Lohr, F., Kehraus, S., Piel, J., and Müller, R., Biosynthesis of the myxobacterial antibiotic coralopyronin A. *Chembiochem*, 2010. **11**(9): p. 1253-1265.
201. Sucipto, H., Wenzel, S.C., and Müller, R., Exploring chemical diversity of α -pyrone antibiotics: Molecular basis of myxopyronin biosynthesis. *Chembiochem*, 2013. **14**(13): p. 1581-1589.
202. Sucipto, H., Sahner, J., Prusov, E., Wenzel, S., Hartmann, R., Koehnke, J., and Müller, R., *In vitro* reconstitution of α -pyrone ring formation in myxopyronin biosynthesis. *Chem Sci*, 2015. **6**(8): p. 5076-5085.
203. Fuchs, S.W., Bozhüyük, K.A., Kresovic, D., Grundmann, F., Dill, V., Brachmann, A.O., Waterfield, N.R., and Bode, H.B., Formation of 1, 3-cyclohexanediones and resorcinols catalyzed by a widely occurring ketosynthase. *Angew Chem Int Ed*, 2013. **52**(15): p. 4108-4112.
204. Mori, T., Awakawa, T., Shimomura, K., Saito, Y., Yang, D., Morita, H., and Abe, I., Structural insight into the enzymatic formation of bacterial stilbene. *Cell Chem Biol*, 2016. **23**(12): p. 1468-1479.
205. Hari, T., Labana, P., Boileau, M., and Boddy, C.N., An evolutionary model encompassing substrate specificity and reactivity of type I polyketide synthase thioesterases. *Chembiochem*, 2014. **15**(18): p. 2656-2661.
206. Hemmerling, F. and Hahn, F., Biosynthesis of oxygen and nitrogen-containing heterocycles in polyketides. *Beilstein J Org Chem*, 2016. **12**(1): p. 1512-1550.
207. Pang, B., Wang, M., and Liu, W., Cyclization of polyketides and non-ribosomal peptides on and off their assembly lines. *Nat Prod Rep*, 2016. **33**(2): p. 162-173.
208. Jenner, M., Afonso, J.P., Bailey, H.R., Frank, S., Kampa, A., Piel, J., and Oldham, N.J., Acyl-chain elongation drives ketosynthase substrate selectivity in *trans*-acyltransferase polyketide synthases. *Angew Chem Int Ed*, 2015. **54**(6): p. 1817-1821.
209. Ferraz, H., Bombonato, F.I., Sano, M.K., and Longo Jr, L.S., Natural occurrence, biological activities and synthesis of eight-, nine-, and eleven-membered ring lactones. *Quim Nova*, 2008. **31**(4): p. 885-900.

9 Appendix

9.1 Manuscript E – Supplementary information

Investigation of a non-canonical chain-branching ketosynthase by site-directed mutagenesis and domain swaps

Srividhya Sundaram¹, Ruth Bauer¹, Hak Joong Kim¹, Tawatchai Thongkongkaew¹, Christian Hertweck^{1, 2, 3*}

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

General

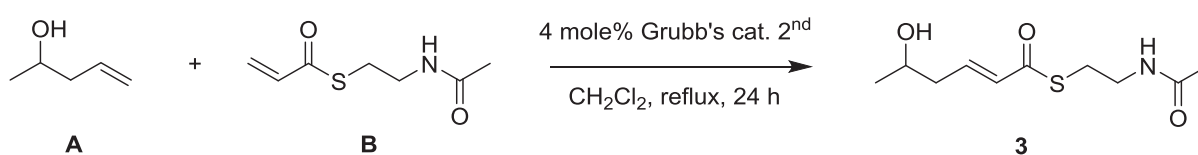
All reagents and 4-penten-2-ol were commercially purchased from Sigma-Aldrich and Roth. All the solvents used for preparative and analytical measurements were obtained in HPLC grade. The chemical synthesis of (E)-(S)-(2-acetamidoethyl) hex-2-enethioate and 2-(2-methyl-6-oxotetrahydro-2H-pyran-4-yl) acetic acid were performed as described previously (Bretschneider, et al., 2013).

Chemical synthesis of (R,S)-(E)-S-(2-acetamidoethyl) 5-hydroxyhex-2-enethioate (**3**)

According to Bretschneider et al. (2013), (R,S)-(E)-S-(2-acetamidoethyl) 5-hydroxyhex-2-enethioate (**3**) was synthesized by a reduction-olefination sequence in 6 steps. This method has three disadvantages. The reaction sequence is long and time consuming. It also requires a protecting group. Moreover, the synthetic route is not flexible for longer and shorter substrates because the starting materials are not commercially available. Thus, we generated the substrates using a new synthetic route. (R,S)-(E)-S-(2-acetamidoethyl) 5-hydroxyhex-2-enethioate (**3**) can be synthesized by a cross metathesis reaction between the S-(2-acetamidoethyl) prop-2-enethiolate and 4-penten-2-ol. The S-(2-acetamidoethyl) prop-2-enethiolate was synthesized by the method described by (Heine, et al., 2014). Its spectroscopic data agrees with the previous report.

General procedure for a cross metathesis reaction

0.18 mmole of 4-penten-2-ol (**A**) and 0.06 mole of S-(2-acetamidoethyl) prop-2-enethiolate (**B**) were added into a heat dry 4 ml vial filled with 500 μ L of dry CH_2Cl_2 . 1 mg (2 mole %) of Grubb's catalyst second generation was added and another portion 1 mg (2 mole %) of catalyst was added after 16 hours. The reaction was heated at refluxing for 24 hours. To work-up, the reaction was dried under N_2 gas and the product was purified by silica gel column chromatography using 5% MeOH: CHCl_3 as eluent.



Molecular formula: $\text{C}_{10}\text{H}_{17}\text{NO}_3\text{S}$; **HRESIMS:** m/z 232.1001 $[\text{M}+\text{H}]^+$ (calculated 232.1007, Δm -2.6 ppm); **IR (film) ν (cm^{-1}):** 3289 (O-H and N-H stretching), 3093 ($\text{C}_{\text{sp}^2}\text{-H}$)

stretching), 1652 and 1629 (C=O and C=C stretching) ; ^1H (MeOH- d_4 , 500 MHz): δ 6.94 (1H, dt, $J=15.5, 7.6$ Hz), 6.22 (1H, dt, $J=15.5, 1.3$ Hz), 3.80 (1H, sextet, $J = 6.2$ Hz), 3.33 (2H, t, $J = 6.8$ Hz), 3.06 (2H, t, $J = 6.5$ Hz), 2.33 (2H, m), 1.91 (3H, s), 1.17 (3H, d, $J = 6.0$ Hz) ppm; ^{13}C (MeOH- d_4 , 125 MHz): δ 190.9, 173.4, 143.9, 131.3, 67.4, 42.5, 40.2, 28.9, 23.3, 22.5 ppm

Rhi-KS14	SGERVEDNELANYIAVIGLGGYYPGADSIDELWQNLANGVDCMSDFPADRWDH	SKIYYKN	60
Rhi-RzxE	-----VAVIGLGGYYPGADSIDQLWQNLANGVDCMSEFPAERWDH	RKLYYKN	47
Ltm-KSX	-----GDIAVVMAGRYPGADTLEEFWELLSEGRHSFEPVPSRWRH	GDLYFDE	49
Mgs-KSX	-----HDIAVIGMAGRYPGADTLEEFWELLSEGRHSFEPVPSRWRH	GDIYFDE	49
ChxE-KSX	-----HDAIIGVSGSYPGADTLEEFWELLSEGRHAFREVPREWRH	DAIYSSD	49
Ms-KSX	-----HDAVIGISGTYPGADTMDLWLSLLEEGRHGFEEVPRDRWH	DAIYSRD	49
Lkc-KS3	-----DDIAVIGMAGEFPPGADTEAFWQALLRGDDAVRVIPGDRWH	RENHSFA	49
Ped-KS12	-----IAVIGISGRYPMANDLDEFWLNLRREGKDCVSEVPSQRWNR	RDHYSEE	47
LnM-KS2	-----EPIAVIGIAGRYPGAGDLETFWNLAEVDSVGLPAERARD	-----	42
LnM-KS4	-----DDIAIIGVAGRYPEAEDLEAFWRNLAEGRDCVGEVPADRWH	HAAYDPE	49
Mmp-KS4	-----IAIIGLAGRYPQAEINIEELWENLKLGRDCITTVPSQRWH	DAIYDPS	47
Mln-KS9	-----EDIAIIGLSGKYPMAEDLPEFWENLKSJKDCITEIPPDRWN	DDFFSED	49
Mmp-KS3	-----VAIIGLAGRYPQARTLEEFWQVLSQGRDCISEIPTERWH	SRYYSD	47
Mln-KS5	-----NDIAIIGISGRYPMANDLDTLWENLLNGKDCITEIPEDRWH	SLGHYQPA	49
Mln-KS10	-----ADIAIIGVSGRYPGAENIREFQNVLMNGRDCVSEIPDDRKM	-----	42
Mln-KS7	-----EDIAIIGIAGRYPMAKDVEQFWENISSGKDCITEIPKERWH	RKYYHTG	49
Mln-KS4	-----EDIAIIGVSGRYPNSDTEEFWENIAGGKNCIEIPEERWH	RSNYSPD	49
Bae-KS5	-----EDIAIIGISGRYPQAEINLQEFWKNLSEGTDCITEIPNDRWH	SLYYDAD	49
Tha-KS3	-----DAIAVIGLAGRYPQARDLDAFWRNLRDGRDCITEIPAERWH	GDFFDPQ	49
Tha-KS12	-----DIAIVGIAGRYPQADDLAQFWRNLARGVDSVTEIPADRWDY	RRFYDPQ	48
	:*::*:* * * . : : * . . . * . *		

Rhi-KS14	RKVLGKTTTCING	SFIKDVDFDYSYFKMPKVYADHMS	PEVRLFLQVAVHTFEDAGYSKET	120
Rhi-RzxE	REVLGKSTCLNG	SFIKDVDFDYSYFKMPKVYADHMS	PEVRLFLQTAVHTFEDAGYSRET	107
Ltm-KSX	RDVLGKTTVRTGT	FLRDVDAFDPYFYSISQRDAELLS	PEVRLFLQAGVTALEDAGYSKET	109
Mgs-KSX	RDVDGKTVVKTGT	FLRDVEAFDPYFNISQRDAELLS	PEVRLFLQAGVEALEDAGYSRET	109
ChxE-KSX	RGVLGKSTIRTGT	FLSDIDTFDPYFRISKREAENMS	PEVRLFLRAGVEALEDAGYSRET	109
Ms-KSX	RAVLGKSSIRTGT	FLRGIDAFDPYFRISKRAEHMS	PEVRLFLQTVGEALEDAGYSRET	109
Lkc-KS3	PDEPGTAYGRHGG	FLDHALDFDPVFFNIAPLEARFMD	PQERRFLQSVYHALEDAGYFARP	109
Ped-KS12	HSRAGGHFCKWGG	FIDIDKDFDPLFFNISPSAAEYMD	PQERLFLFHAWMAMEDAGYRRED	107
LnM-KS2	---GWPTQMWWGG	FLDGVDRFDALFFGIAPRDAQLMD	PQERQLFQVWVETLEDAGCTRAR	99
LnM-KS4	RGKEGRTYGRRGG	FLDGVDRFDAASFGISRREAELMD	PQERLFLTVGRQAVENAGYRPEE	109
Mmp-KS4	KGVSFKTYSKWGG	FLRGVDEFDPRFFNISPREAEIMD	PQERLFLQCAHYVLEDAGYTRQS	107
Mln-KS9	RDMKGSYSKWGG	FISGADQFDPLFFHISPREAEIMD	PQERLFLFLETAKHCFEDSGYPRNM	109
Mmp-KS3	EDAPGKTYARWGG	FIDGVAEFDPAFFGISPREAMAME	PQERLFLQTAHEAIEDAGYTRQG	107
Mln-KS5	AGDEVKNQSKWGG	FIDDVYHFDPLFFNISPAEAEIID	PQERLFLFLETVWHTIEDAGYAKKT	109
Mln-KS10	---LEGSCYKWGG	FLKNPDRFDPLFFRISPKEAAFLD	PQERLFLFLETVYNTIEDAGYKPSD	99
Mln-KS7	--DNARHDSKWGG	FIKDADKDFDPLFFSISPRDAELID	PQERLFLFLETAWHTFEDAGYTKKK	107
Mln-KS4	RMEKGKINSKWGG	FISGVDQFDPLFFHISHKEAELMD	PQERIFLETSWHTFEDSGYTKEK	109
Bae-KS5	KDKEGKTYGKWGG	FLKDVDFDQFFSISPRDAKLMD	PQERLFLQCVYETMEDAGYTRKK	109
Tha-KS3	KGVAGKTYSKWGG	FIEGVDEFDAFFNIAPRDAERMD	PQERLFLQASYQAIEDAGYARAS	109
Tha-KS12	KGRLGKSYSKWGG	FLSDVARFDAFFNISAREAQIMD	PQERLFLFLECIVYHTLEDAGYTRRN	108
	* * : * * * * : * : . * : * * * . * : * *			

Rhi-KS14	LL-----SRYNGDVGVLLGTMSNDYHYYGFESNVFRGSMASGSGMATI	PMTVSYFYG	172	
Rhi-RzxE	LQ-----SRYGGNVGVLLGAMSNDYHYYGFENHLSGGSIASGSGMATI	PMTVSYFYG	159	
Ltm-KSX	LR-----RRYDGDVGVLVGSMNNSYAYYGFENMLMRGTAMSGSEVGVMANM	LSYYYG	161	
Mgs-KSX	LR-----RRYDGDVGVLVGSMNNSYSLYGFQNLMLMRGTATSGSELGVMANM	LSYHYG	161	
ChxE-KSX	LR-----RKYDGDVGVLAGSMSNHYGLYGFENGLTRGSAASGSYTGTL	PNMLSYFYG	161	
Ms-KSX	IQ-----RQYGGDVGVLVGAMSNDYHYYGLYGFQNSLTRGSAASGSYTGTL	PNMLSYFYG	161	
Lkc-KS3	-----TGDMGVFVAAMFGHYQDLTAADR-----	VIGSSFAAIANRVSYAFD	150	
Ped-KS12	LRKLARGSAAEDLP	GGVYAGVMYSEYQLL	GIEAARQGGKATVANFHASVANRVSYVLD	167

Lnm-KS2	IR-----EQLGSDVGVFVGTMYNEYPPFGVERSLAGESADTGS AVAGIANRVSYFLD	151
Lnm-KS4	LA-----RTRVGVFAGVMWNHYQLCTDGS A---EPVAPTALHCSVANRLSYCLD	155
Mmp-KS4	LSA-----KGRVGVYVGVQYTEYTAFTAQT-----LVAALPASIANRVSFCD	151
Mln-KS9	LE-----KMKVGVFVGAMWGQYQLFEGEEN--GVAFSPASVYSSIANRVSYFH	156
Mmp-KS3	LAA---SARHEDAEGMVG VFGVTYEYQLYGAQQTA EGRPLVLSMSPSSIANRVSFVNG	164
Mln-KS5	LE-----RKNVGVFAGVMYQYQLYSVG-----HEAAVSSSYASIANRVSYFFN	153
Mln-KS10	LT-----RKKVGVFAGV TYGSYQFFGVEESLKGNELAVGSPFSAIANRISYFFN	148
Mln-KS7	LK-----RHKTGVFVGV TYSHYQLFAAGASDDEERQALLSSYSSIANRVSYFFD	156
Mln-KS4	LD-----GMKAGV FVGAMYQYQLFGAEETARGNP IALSSFFSSIANRVSYFFN	158
Bae-KS5	LTE---KSG-DLLGANVGVYVGM YEEYQLYGAEEQARGKSLALTGNPSSIANRASVFG	165
Tha-KS3	LG-----AGRVGVFAGVMYSEYQLYGIEESAAGRPAALSGSAASIANRVSYHLD	158
Tha-KS12	VSR-----SRRVGVFVGV MYEEYQLYGVERMLEGTPVALAGNPAAIANRVSYFCD	158

* * . * : *

Rhi-KS14	LTGPSLFIDTMCSSSSSTCIHTACQMLKHDETKMVLAGGLNLMYHPYTTVNTSQGNFTSIT	232
Rhi-Rzx E	LTGPSLFIDTMCSSSSSTCIHSAFQMLRHDETMVLAGGLNLMYHPYTTVNTSQGNFTSIT	219
Ltm-KSX	FTGPSMFVDTMCSSSSSACVHQALSMLRGGECRMV VVGGINLMLHPYDLIATSQAHFTTKS	221
Mgs-KSX	FTGPSVFLDTMCSSSASACVHQAVRMLRSGECRMTV VVGGINLMLHPFDLIATSQAHFTTKS	221
ChxE-KSX	FTGPSLFVDTMCSSSSTCVHQAVQMLRARECSMAVAGGVNLLLHPYNLISSSQEHFTTAT	221
Ms-KSX	LTGPSIFLDTMCSSSSSACI HQAVQMLRARECPMVVAGGVNLLLHPYNLISSSQEHFTTAT	221
Lkc-KS3	LHGPSIALDTMCSSGGLTALHLAVRSIRAGDCALAVAGGVNLMTHPGKYRLLSEGRFLSPT	210
Ped-KS12	LHGPSMTVDTMCSSSLLTALHLACQDLKTGR TDMALAGGVNLSVHPNKYSVLSLNEFISSQ	227
Lnm-KS2	LHGPSLAVDTMCSSSLLTALHLAVESLRRGEC AA AVAGGVNLSLHPHKFRQTRLKMS SSD	211
Lnm-KS4	LSGPSMAVDTACSSSLLTSLHLAVESIRRGECALAVAGGVNVA AHPQKYLQLAQGRFLSSD	215
Mmp-KS4	FRGPSLTLDSCSSSLLTTHLACQSLRSGESEYAIAGGVNVS IHPNKYLLHAQGRFASSV	211
Mln-KS9	LKGPSLALDTMCSSSLLTSIHYACQSIHCGESDMALAGGVNLSIHPNKYKFLSMGQFLSSD	216
Mmp-KS3	FHGPSMAIDAMCASSLTTIHLACQSLHSGECQVALAGGVNVS IHPAKFLMLGAGKFASRR	224
Mln-KS5	FSGPSIALDTMCSSSLLTAVHFACESLIRNECEIAIAGGVNLSIHPHKYELLTQGGFLSSD	213
Mln-KS10	FKGPSMAVDTMCSSSLLNAIHLACESIRRGESTLAVAGGVNLT LHPNKYTLNLSRFFASD	208
Mln-KS7	LHGPSLAIDTMCSSSLLSAVHFACESLKRGESTLAIAGGVNLSLHPDKYTLNLSQNHVSSD	216
Mln-KS4	LSGPSIALDTMCSSSLLTAIHLACESIKRGESEIALAGGVNVS IHPSKYLWLSQGNFVSTE	218
Bae-KS5	FNGPSMALDTMCSSSLLTAIHLACQSLRNGECEAA FAGGVNVS VHPNKYMLGQNRFLSSK	225
Tha-KS3	LHGPSMAVDTMCSSSLLTALHLACRSLQRGECELALAGGVNVS VHPNKYMLADNRFVSSA	218
Tha-KS12	FHGPSMAIDTMCSSSLLSAIHLACQSLMLGECEVAVAGGVNVS IHPNKYQMLSQGRFASSN	218

: ***: *: *... : * * : ...**:* : ** : :

Rhi-KS14	SESVNSYGVGADGTVIGEGIGAVLLKRLDRAIADR DQIYGVIKGSAMTNAGERNGFNVPN	292
Rhi-Rzx E	SQSVNSYGVGADGTVIGEGVAVLLKRLDRAIADR DQIYGVIKGSAITNAGERNGFNVPN	279
Ltm-KSX	AEVVRSYGLGADGTVILGEGVGLV LKPLAEAVADGDH VYGVIKGSGMTNAGVRNGFTVPS	281
Mgs-KSX	AEVVRSYGLGADGTVILGEGVGLV LKPLAEAVADGDH VYGVIKGSGMTNAGVRNGFTVPS	281
ChxE-KSX	SDVIRSFGLGADGTVILGEGVGLV LKPLVEAERDGDH VYAVIKGTALANAGVRNGFTVPS	281
Ms-KSX	SDVIRSFGLGADGTVILGEGVGLV LKPLVEAERDGDQ VYAVIKGTALTNAGVRNGFTVPS	281
Lkc-KS3	-GHCQAFGVEADGYVPGEETA AVVLLKPLSDALR DGDGRVHAVIRATAVNSGGRTAGFTVPS	269
Ped-KS12	-GRCTSFEGGDDGYVPSEGVGVLLKRLVDAERDR DHIHAVIKSSVLNHHGGKTHGFSVPN	286
Lnm-KS2	-HRCRSFGAGGDGFVPAEGVGAVLLKPLSAAEADGDRI HAVIRGTAVNHHGGKTNGYMVPN	270
Lnm-KS4	-GRCRAFGADGDGYVPGEVGVAVLLKPLADALADGDH VHAVIKGSFLNHSGRTS GFFTVPS	274
Mmp-KS4	-GCCKTFGQGGDGYVPAEGVGAVLLKPLPQAIADGDRI HAVIKGSSINHGGRTAGFTVPK	270
Mln-KS9	-GRCRSFGKDGSGYVPGEVGVAVLLKPLKKAIEDQDQI YGVIKGS AVNHHGGKTS GYSIPN	275
Mmp-KS3	-GRCESFGASGSGYVPSEGVGVAVLLKPLHQAQADGDRI YGVIRGSAINHGGRTNGYAVPN	283
Mln-KS5	-GRCRSFGSGDDGYVPGEGTGAVLLKPVKKA IADGDQIYSVIKATAINHGAKTNGYTVPN	272
Mln-KS10	-GRCRTFEGGDDGYVPGEVGVSVLLKPAVEAVRDGDHI YALIKATAVNHGGRTNGFTVPN	267
Mln-KS7	-GRCRSFGAGGDGYVPGEVGVAVLLKSLSAAKADGDRI YGVIKGTSVNHGGKTNGYTVPN	275

Mln-KS4	-GLCRSFGEGGDGYVPGEVGSVAVLLKPLEKAIADQDHIYGVKIGTSVNHGGKTNFTVPN	277
Bae-KS5	-GRCESFGEGGDGYVPGEVGVAVLLKPLSKAKADGDHIYGLIKGTAVNHDGKTNVSVPN	284
Tha-KS3	-GRCESFGAGGDGYVPAEGVGVAVLKLPLRAAIADGDIAIHGVICATALNHGGKNNGYTVPN	277
Tha-KS12	-GRCESFGAGGDGYVPSEVGVAVLLKPLARAIADGDRIHAVIKATALNHGGKTNVSVPN	277
	:*: ..* : .** . :** * * * :*: * . : . * : :*	
Rhi-KS14	PDLQTLAIRQAMDQAKVHPSSISYIEGHGSGTKLGDPIEVLGLNNAFRWATDDK-QFCYL	351
Rhi-RzxE	PDLQTLAIRQAMEQAKVHPASISYVEGHGSGTRLGDPIEVLGLKNAFAVDTDEK-QFCYL	338
Ltm-KSX	PQQQARAIALDDAAVDARTVSYLEGHGSATSLGDPIEIKGASLAFGRDTRDV-GYCAI	340
Mgs-KSX	PQQQARAIEKALDDAAVDARTISYLEGHGSATSLGDPIEIKGAALAFGRDTRDQDL-GFCAL	340
ChxE-KSX	PRMQARAIERAIDDAGVDPRTISYVEAHGSGTSLGDPIEVKALTTAYRKYTPET-GFCAL	340
Ms-KSX	PHAQARAIEKAIDDAGVDPRTISYVEGHGSGTSLGDPIEVKALTDAFRKHTEDT-QFCAL	340
Lkc-KS3	ERAQHRVIAAALAEAEIEPGAVTYVEAHGTGTRLGDPIEVRALSRAYGGPEQGR---AYL	326
Ped-KS12	PKAQQHLLISRALREAEVDPRAITYVEAHGTGTPLDPIEVTALSKAFAQYSLGG-QPYWI	345
Lnm-KS2	PVAQGDVLRRAALRRAGADPATIGYVEAHGTGTQLGDPIEINGLNRAFAGASVAP-ASRAI	329
Lnm-KS4	PAAQATLIADALDRSGVAADSVGYIEAHGTGTALGDPIEIEGLRQAFADAGLAP-GSCAI	333
Mmp-KS4	SSAQAVAVRSALHQAGLASSDLSYIEAHGTGTALGDPIEIEGLRTVFEADGFEP-QTCAI	329
Mln-KS9	PDAQAEVISEALRKADIDPNTVNYIEAHGTGTSLGDPIEIQGLKDSYQL-----NHPCAI	330
Mmp-KS3	PAAQAAVISRALRQANVQPRQIGYVEAHGTGTSLGDPIEVAGLSRAWRAYTPDR-QFCAL	342
Mln-KS5	PKAQSEVITRALKKGFIDPKSISYIEAHGTGTALGDPIEITGLTRAFSDAERDK-KTCAI	331
Mln-KS10	PNGQSELIQDALQKACIDPRTISYIEAHGTGTSLGDPIEITGLTKAFREFTEHT-QYCSI	326
Mln-KS7	PGEQANVISETLDKAGIDADSISYIEAHGTGTSLGDPIEISGLTKALHSSKSD--YRCAI	333
Mln-KS4	PNAQAELIAENFKKSGISPRSVSYIEAHGTGTSLGDPIEITGLTKAFKFTDDK-QFCAL	336
Bae-KS5	PNAQAAVIKQALKDAGTDPRAVSYIEAHGTGTSLGDPIEITGLTKAFSEQTQDK-QFCAL	343
Tha-KS3	PAWQAAVIEAALGEAGVAPGDVSYVEAHGTGTSLGDPIEITGLTKAFGEFPGARRGAPCAI	337
Tha-KS12	PNAQADVIGDALARAGIDARSIGYVEAHGTGTSLGDPIEITGLTKAFGEFPGARRGAPCAI	336
	* : : . : *:*.**:* ** * :* : . :	
Rhi-KS14	GSIKSNIGHLLAASGIAGLTKTLLQFKHKQIAPSI-HSSQLNQDIDFADTPFVVPQQQLIE	410
Rhi-RzxE	GSIKSNIGHLLAASGVAGLTKTLLQFKHGHIAPSI-HSGLLNQDIDFADTPFVVPQQQLTQ	397
Ltm-KSX	GSVKSNVAHLLSGSGLVGLTKVLLQLRHRRLAPSL-HSETLSPAIDFGSTPFVVPQRERAE	399
Mgs-KSX	GSVKSNVAHLLSGSGMAGLTKVLLQLKHRRLAPSL-HAGTLSSAIDFEETPFVVPQRHRDT	399
ChxE-KSX	GSVKSNLGHLLHAAGMIGIAKVLLQFQFQGLVPSL-HSSVINPDIDFSRTPFRVQDELTA	399
Ms-KSX	GSVKSNLGHLLHAAGMIGFAKVILQMRKGTLPAPSL-HSDELNPDIDFGGTPFRVQRALAP	399
Lkc-KS3	GSVKSNIGHLESAAGMAGLVKVVQLAHRRLAPTL-HCTLENPDIDLHLSRFAVVKETLP	385
Ped-KS12	GSVKSNIGHTESTAGIAGLSKVILQMRREGQLAPSL-HSQTLPNPNIDFASPPFQVNRQLRE	404
Lnm-KS2	GSVKANIGHAEAAAGIAGLTKVLLQLRHRRLVPSL-HTEELNDAVDWASSPFVVPVREGRP	388
Lnm-KS4	GSVKSGIGHLESAAGIAAVTKVLLQMRHRELVPAPSL-HSEQPNPHIDFAATPFVVPQRTRAP	392
Mmp-KS4	GSIKSNIGHCESAAGIAGLTKVLLQLKHRQLVPSL-HSTQLNRNIDFAGSPFHVQQTLEP	388
Mln-KS9	GSVKSNIGHLESAAGIAAVTKVLLQMKNKQLVPSL-HSDIVNPFIDFDQVPPFVQQRKAAE	389
Mmp-KS3	GSVKSNIGHCESAAGIAGLTKVLLQMKHGLLVSSL-HAETLNPINIDFSETPFVVPQRTRAP	401
Mln-KS5	GSIKSNIGHLEAAAGIAGITKVILQMKHKTIVPSLTHSAELNENIPFQKTPFVVPDSEAEK	391
Mln-KS10	GSVKSNIGHLEAAAGIAGLTKVLLQMKLKLFPPI-HAETLNPINIDFKASPFVVPQRSQD	385
Mln-KS7	GSVKSNIGHLEAAAGIASITKVLMQMKYKQLAPSL-HSGNLNPNIDFEETPFVVPQSELTD	392
Mln-KS4	GSVKSNIGHLESAAGIAALTKVLLQFKHQQLAPSI-HSEQLNQNIREFEKTFFVVPQKEEE	395
Bae-KS5	GSAKSNIGHCESAAGIAGLTKVLLQMKHKQLAPSL-HSRTLNPINIDFLATPFVVPQRTLEE	402
Tha-KS3	GSVKSNIGHAESAAAGIAGLTKVLLQMRHRMLVPSL-HADTLNPNIAFETTPFCVQRALER	396
Tha-KS12	GSVKSNIGHGESAAAGMAGLTKIVLQMRHRRLVPSL-HADTPNPNIDFADTPFVVPQTTLAP	395
	** * : : * : : : * : : * : . : * :	

Rhi-KS14	WRQPE-RIINGRKQVFP RRAGLTS IAAGGMNAHMIVEEYP	449
Rhi-RzxE	WHRPE-RMIQGELQRF RRAGLTS IAAGGMNAHLIVEEYP	436
Ltm-KSX	WRRPVVHG-----AEVRRAGVTS IGAGGINVHLIVEEFD	434
Mgs-KSX	WRRPVVGG-----EEAPRRAGVTS IGAGGINVHIVVEEYD	434
ChxE-KSX	WKPAV-TMVDGREITHPRRAGITS IGAGGMNSHMILEEYP	438
Ms-KSX	WDPVTVRAADGRVVTHPRRAGITS IGAGGMNSHMILEEYA	439
Lkc-KS3	WPKG-----SGPTRFAGLSS FGAGGANHAIIVQEFV	416
Ped-KS12	WPRPVL D-----GRLQPRVASLSS FGAGGSNAHLVISEYI	439
LnM-KS2	WAPLTGAD----GAPLRRAGLSA FGAGGANAHVVVEEYV	424
LnM-KS4	WVPR-----PGSTVLRAGVSA FGAGGSNAHVLLSEAP	424
Mmp-KS4	WLPKGAA-----DAVQRRAGISS FGAGGSNAHLVVEEYS	423
Mln-KS9	WNIPQNS-----NAVHPLRAGVSA FGAGGTNAHLILEAYE	424
Mmp-KS3	WERPLGSD----GVPGLRMAAISS FGAGGANAH LIVSEAP	437
Mln-KS5	WES-----SVRRRAGISS FGAGGANAHVILEEFA	420
Mln-KS10	WKKPKLQ-ENGKTVTVRRRAGISS FGAGGTNVHIIIEEYE	424
Mln-KS7	WET-----PQNQARRSGISS FGAGGSNAHIIIEEYK	423
Mln-KS4	WRKE-----SGASNRRAAISS FGAGGANAHVIVEEYQ	427
Bae-KS5	WKRPVIN-ENGVNKELPRTAGLSS FGAGGVNAHIVIEEYS	441
Tha-KS3	WER-----PGDGERVAGVSS FGAGGANAHVIVREYR	427
Tha-KS12	WRGAILPDEAGRPAELPLRAGLSS FGAGGANAHVVVEAYD	435
	* : : : : : . * * * * * * * * : :	

Figure S1, related to figure 2A: Multiple sequence alignment of selected clade IX full-length KSs. Residues highlighted in red were chosen for mutational analysis in the current study. Residues colored in blue constitute the active site.

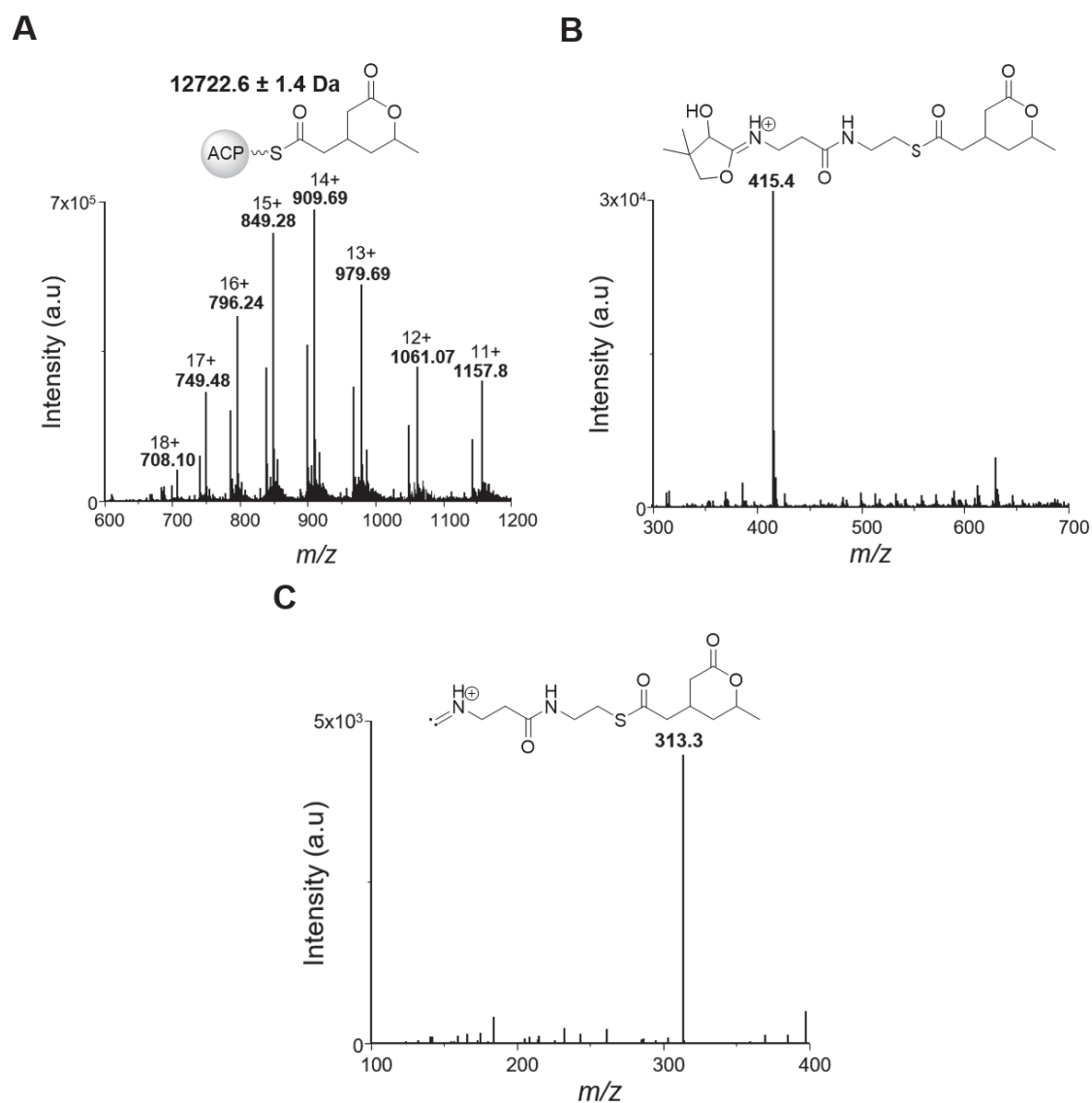


Figure S2, related to “LC-ESI-MS and LC-MS/MS measurements” in the Experimental procedures section: ESI-MS of KS-B. (A) Mass spectrum corresponding to ACP-bound lactone product. (B) MS² of m/z 849.28 ejected the PPant arm of ACP with the product attached. (C) MS³ fragmentation

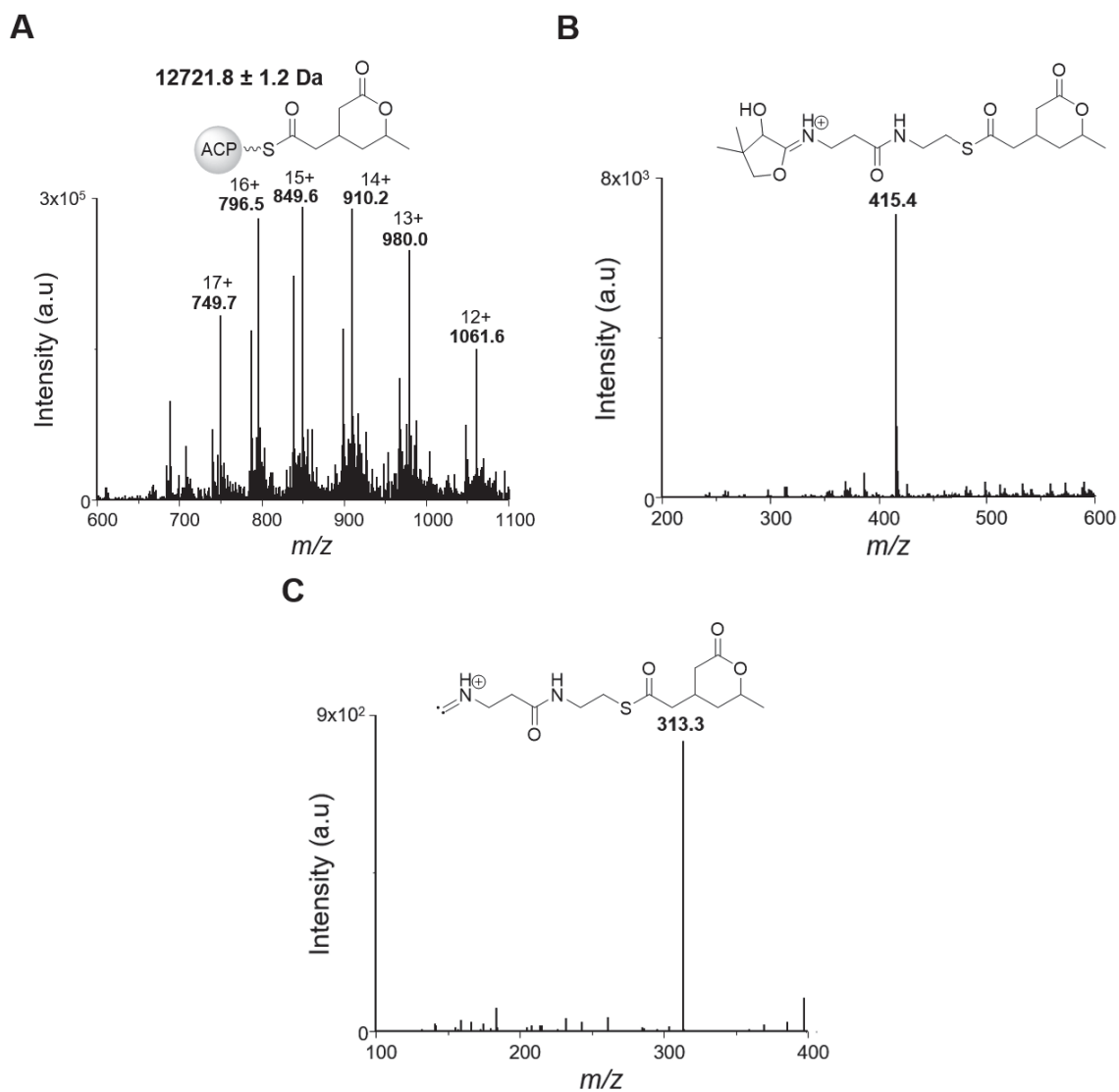


Figure S3, related to “LC-ESI-MS and LC-MS/MS measurements” in the Experimental procedures section: ESI-MS of KS_{H3364A}-B. (A) Mass spectrum corresponding to ACP-bound lactone product. (B) MS² of *m/z* 849.60 ejected the PPant arm of ACP with the product attached. (C) MS³ fragmentation

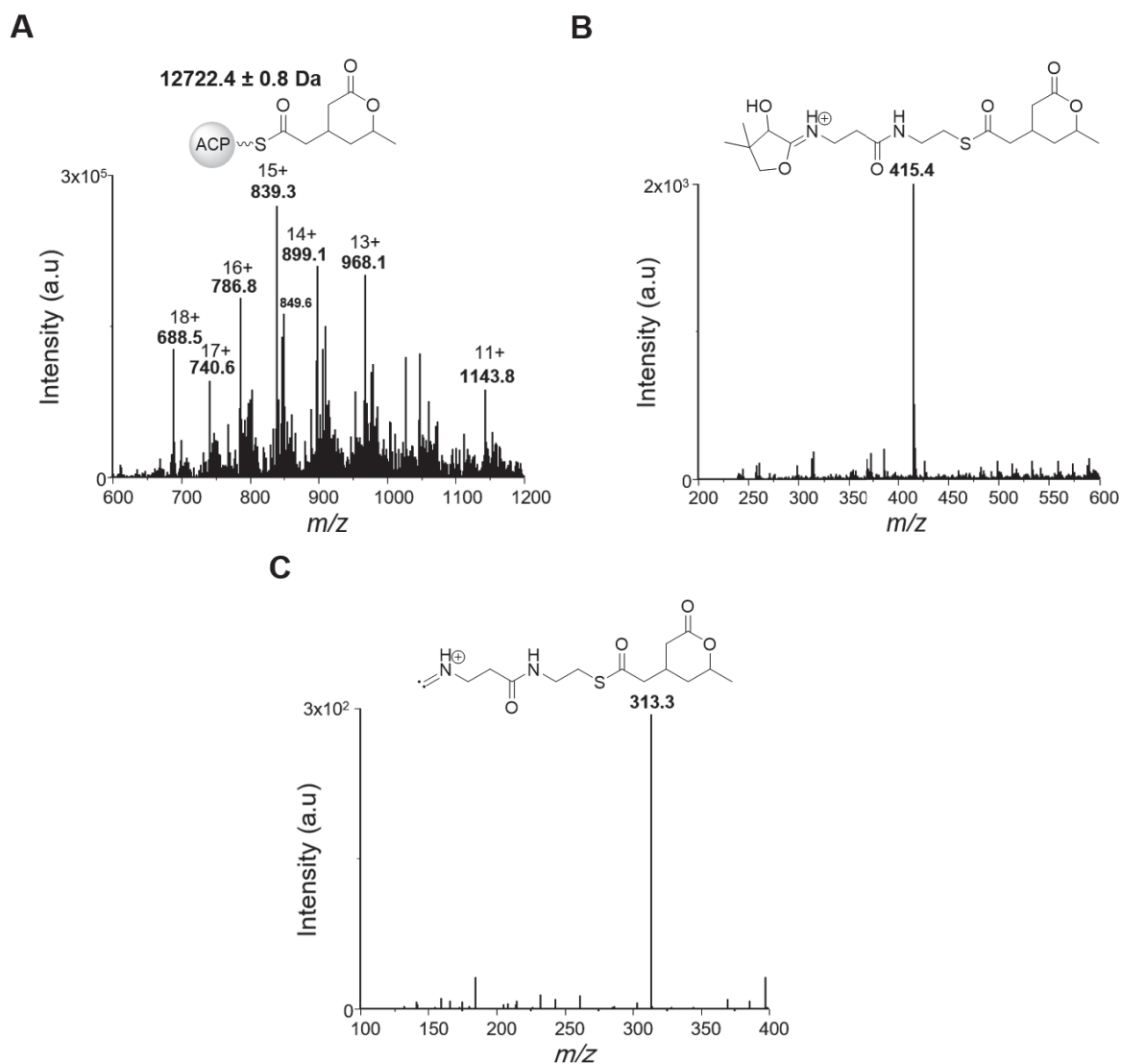


Figure S4, related to “LC-ESI-MS and LC-MS/MS measurements” in the Experimental procedures section ESI-MS of $KS_{H3404A-B}$. (A) Mass spectrum corresponding to ACP-bound lactone product. (B) MS^2 of m/z 849.60 ejected the PPant arm of ACP with the product attached. (C) MS^3 fragmentation

A

Protein:		MlnKS4-B		Peak threshold:	0.0			
Intensity coverage:	86.5 % (154250 crnt)	Sequence coverage MS:	54.3 %	Sequence coverage MS/MS:	0.0 %			
				pt:	5.7			
				Mw (kDa):	103.3			
10	20	30	40	50	60	70	80	90
MKHHHHHHH	GGLVPRGSHG	GSEDAIITGV	SGRYPNSDTL	EEFMENIAGG	KNCIIKIPPE	RUDFRN	DMMEKGGKINS	KUGGFISQVD
100	110	120	130	140	150	160	170	180
QFDPFLFHIS	HKEAELMDPQ	ERIFLETSHH	TFEDSGYTK	KLDGMRAGVF	VGAMVQQQL	FGAETARGN	PIALSFFSS	IANRVSVFFN
190	200	210	220	230	240	250	260	270
LSGPIALDT	NCSSSLTAIH	LACESIKRGE	SEIALAGGVV	WTHPSKYLW	LSQGNFVSTE	GLCRSFGEGG	DGVVFGESG	AVLLKPLEKA
280	290	300	310	320	330	340	350	360
IADQDHIYGV	IKGTSVNHGG	KTNQFTVNPV	NAQAEI AEN	FKKSGISPRR	VSYLEARGTO	TSLODPFIEA	GLTKAFKKFT	DDKQFCAIQS
370	380	390	400	410	420	430	440	450
VKSNIGHLES	AAGIAALIKV	LLQFKHQALA	PSIHSEQLNQ	NIRFEKTFY	VQKKEEMRK	ESGASNRRAA	ISSFGAGGAN	AHVIVEEYQH
460	470	480	490	500	510	520	530	540
NYPEPADSAG	QISEDQLVIV	FSVHKALLA	CMLTFRDWL	ASSEAPLAQI	AYTLQVQKN	LRNRLAIRC	TRQALSRALN	ACIDGHYSS
550	560	570	580	590	600	610	620	630
ADSKFVRFQ	ESDAVQPLES	DLNDPLAPL	TQULNGDSQV	DWASLYAQP	VRISLPAYRF	EKTRCVYTE	GVESSIVNPL	HFNKHLV
640	650	660	670	680	690	700	710	720
AKKCS	TPOPG	AIFRDFVED	ELLDTVYSGR	GGRRLSAFNF	ADVALAMPAL	ASRFDRGLS	VSCAFEHYA	DWTVTGLEV
730	740	750	760	770	780	790	800	810
LEFFDFRSG	EQPHLGFV	INPLTSDEPP	LPQWLDDAR	ELLNQALQA	GROLAAEVS	ORLAQAQYDF	APYLDHDEL	TIGRGLVLK
820	830	840	850	860	870	880	890	900
GRPPVNRNH	YADNVQLSPY	LATTIDKALY	LLDELGLPQ	GRVIVRNIR	LCYHTPAGG	FSVLSGIGL	DNELSLSLL	VLDEREIQCV
910	920	930						
KLDKVSLYL	KQEVASVDRK	HSLLTGT						

B

Protein:		BaeKS5*-B		Peak threshold:	0.0			
Intensity coverage:	73.4 % (75260 crnt)	Sequence coverage MS:	47.9 %	Sequence coverage MS/MS:	0.0 %			
				pt:	6.0			
				Mw (kDa):	108.2			
10	20	30	40	50	60	70	80	90
MKHHHHHHH	GGLVPRGSHG	GSAQVRPDR	ENQLTERED	IATIGISGV	PQAEMLQEFV	KMLSEGTDCI	TEIPNDRDH	SLYYADKDK
100	110	120	130	140	150	160	170	180
EGKTYGKGG	FLKDVDFDF	OFFSISPRDA	KLMDPQERLF	LOCVYETHED	AGYTKKLTTE	KSGDLLGANV	GUYVGWYEE	YQLYGAEOA
190	200	210	220	230	240	250	260	270
FGKSIALTON	PSSIANRASY	VFGFNGPSMA	LDTMCSSSLT	AHLACQSLR	NGECEAAFG	GVNVSHPNK	YLMQLGNFL	SSKGRCSFG
280	290	300	310	320	330	340	350	360
EGDGVTPGE	GVGAVLLKPL	SKAKADGHI	YGLIKGTAVN	HDGKINGYV	PNPNAQAIVI	KQALKDAGT	PRAVSTIEAH	GTGSLGDFI
370	380	390	400	410	420	430	440	450
EITGLTKAFS	EQTQDKQFCA	IGSARSNIGH	CESAAGIAGL	TKVLLQMKHK	QLAPSLHSRT	LNPNIIDFLAT	PFKVQQTLEE	UKRPVINENG
460	470	480	490	500	510	520	530	540
VNKLFPRTAG	LSSFAGGVN	AHIVIEEYSA	DEKRETAFAA	PHPSMIVLSA	KNEQLQKRA	KRLDALFSG	RYREADLSRI	AYTLQVGRFP
550	560	570	580	590	600	610	620	630
NEERLGHIVS	NLELEEKLD	EFTGGKESID	QLTRGQVQKN	KDTHALFID	EDHEKTEAW	LEKGAARV	ELWVKGFLP	WDKLYQGRP
640	650	660	670	680	690	700	710	720
QKISLPAYFF	AKDRYVIDTS	ADAAVKRFAP	TIWNPLMK	KLHPLVARK	STPQGAIFR	TDFVEDELD	VYVSGRGRR	LSAFNFADVA
730	740	750	760	770	780	790	800	810
LAMPALASRF	DGRTLSVSCA	FEHYIADWTT	VTGLEYRLF	IDSEQLELEF	DFRRSGEPT	HLGFVAVINPL	TSDEPFLPQ	ULDDARELLN
820	830	840	850	860	870	880	890	900
RQALQAQRQL	SAAEVSQRLA	QAGYDFAPYL	DHDELGTIGR	SGLVKGRFP	VNRNHNADN	VQLSPYLATT	IDKALYLLD	ELGLPQGRVI
910	920	930	940	950	960	970	980	
VNRIERLCY	HTPAGGFVSV	LSGIGLNDNE	LSLSLLVLDE	REQICKVLDK	VSLYLKQEV	ASVDRKHSLL	TGT	

C

Protein:		ThaKS3*-B		Peak threshold:	0.0			
Intensity coverage:	88.7 % (210775 crnt)	Sequence coverage MS:	60.7 %	Sequence coverage MS/MS:	0.0 %			
				pt:	6.2			
				Mw (kDa):	104.8			
10	20	30	40	50	60	70	80	90
MKHHHHHHH	GGLVPRGSHG	GSAATAADAIA	VIGLAGRYPO	ARDLDAFUN	LRDGRDCITE	IPAERUNHGD	FFDPQKGVAG	KTYSKGGFI
100	110	120	130	140	150	160	170	180
EGVDFDAAF	FNIAPRDAER	MDPQERFLQ	ASYQAIEDAG	YARASLGAGR	VGVFAGVMS	EYQLYIEES	AAGRPAALSG	SAASTANRVS
190	200	210	220	230	240	250	260	270
YHLDLGPHSM	AVDTKCSSSL	TALHLACESL	QRGECALALA	GGVNSVHPN	KYLMADNRF	VSSAGRCESF	GAGDGVVPA	EGVGVAVLKP
280	290	300	310	320	330	340	350	360
LRAAIDGDA	INGVICATAL	NHGKNGNYT	VPNPANQAAV	IEAALGEAGV	APGDVSVVEA	HGTGTTLGD	IEIAGLARAF	GEPGARBPAG
370	380	390	400	410	420	430	440	450
CAIGSVKSN	GHAESAAGIA	GLTKVLLQNR	HRNLVPSLHA	DTLNPNIAFE	TIPFCVQRAL	ERUERPGDE	RVAGVSSFGA	GGANAHYIVR
460	470	480	490	500	510	520	530	540
EYRSDDERDG	RDEPASTAPA	RARPAWILS	ARNEDGLRAR	AAQLRELAAG	CEGDADLHAI	AYTLQGRDA	HDERLAFEAT	SIADLIASLD
550	560	570	580	590	600	610	620	630
AFARGEPEGK	QLRNGRAR	GDAPPAGVAD	APLARGEHRA	ALDDVVRGAS	IDWRRAYGPG	APFGPAPRRM	HLPVYPFART	RHULPAPLDA
640	650	660	670	680	690	700	710	720
FRRAARAGGG	LHPMLDANS	EFGTIVNPLM	FRNKLHPLVA	KKCS	TPOPGA	IFRTDFVEDE	LLDTVYSGR	GRRLSAFNFA
730	740	750	760	770	780	790	800	810
SRFDGRTLSV	SCAFEHYAD	WTVTGLEVYR	LFEIDSEOLE	LEFDFFRSGE	OPHLLGFVAVI	NPLTSDEPFL	PQWLDDARE	LLNRQALQAG
820	830	840	850	860	870	880	890	900
RQLSAAEVSQ	RLQAQYDFA	PYLDHDELGT	IGRSGVLVKG	RPPVNRNHHY	ADNVQLSPYL	ATTIDKALYL	LLDELGLPQ	EVIVRNIRL
910	920	930	940	950	960	970		
CCYHTPAGGF	SVVLSGIGLN	DNELSLSLLV	LDEREIQCVK	LDKVSYLKQ	QEVASVDRKH	SLLTGT		

Figure S5, related to figure 6C: MALDI-TOF analysis following tryptic digests of recombinant proteins. (A) MlnKS4-B (B) BaeKS5*-B and (C) ThaKS3*-B.

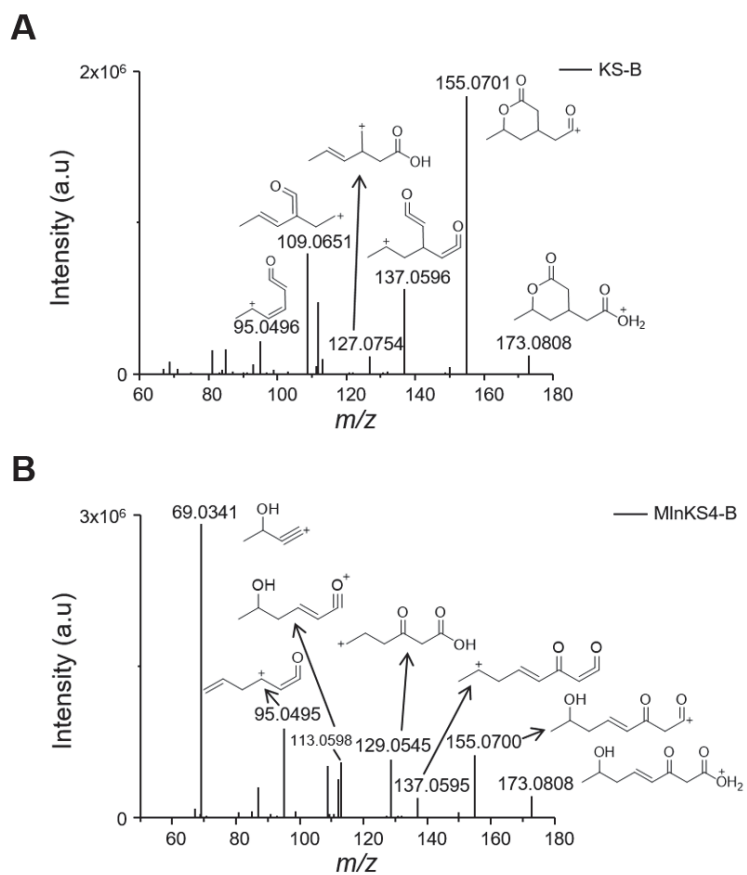


Figure S6, related to figure 7: LC-MS/MS of (A) **4** and (B) **5** tested with RhiE KS-B and MlnKS4-B respectively.

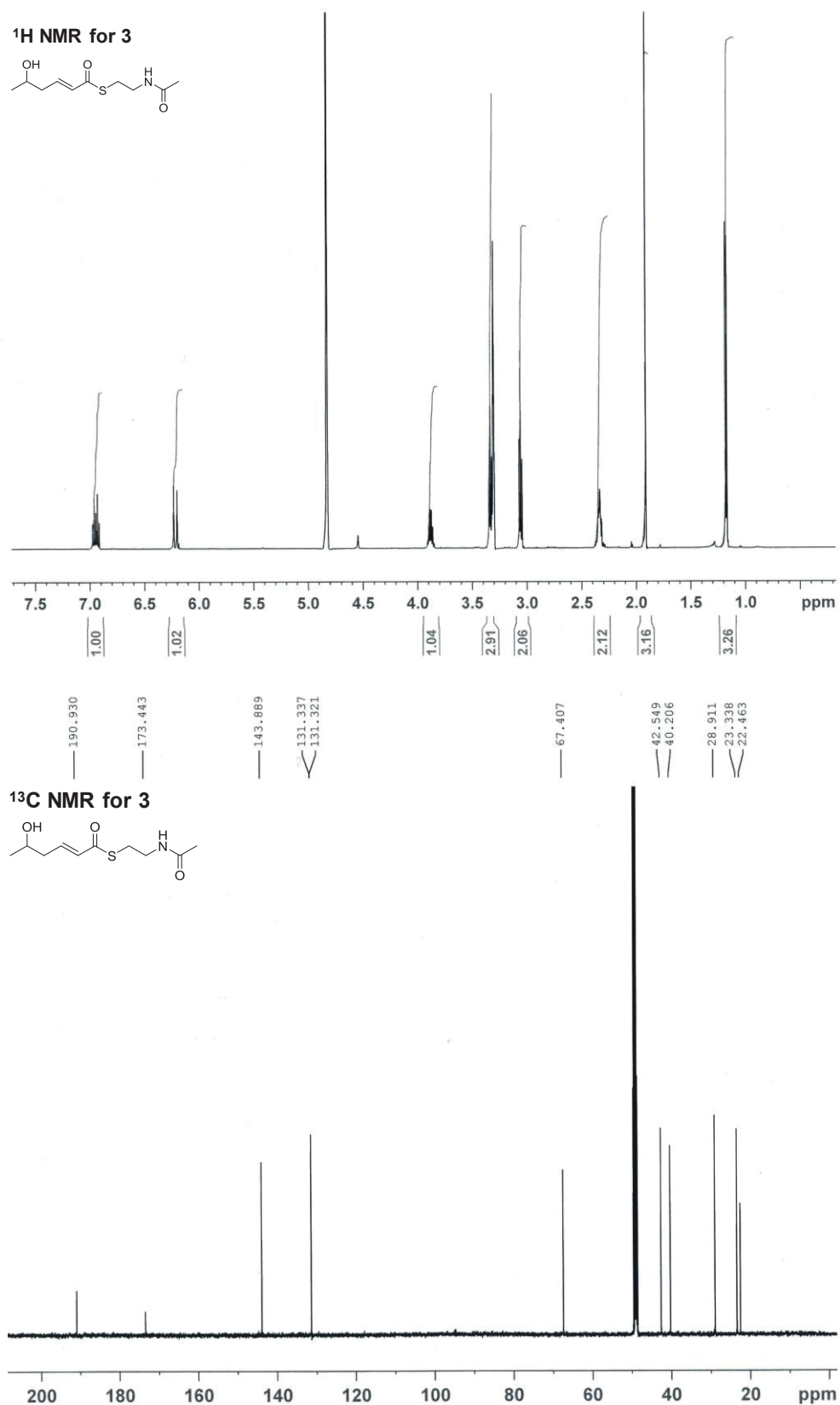


Figure S7, related to figure 4 and figure 7: ¹H and ¹³C NMR spectrum of 3.

REFERENCES

Bretschneider, T., Heim, J.B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013). Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* 502, 124-128.

Heine, D., Bretschneider, T., Sundaram, S., and Hertweck, C. (2014). Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles. *Angew. Chem. Int. Ed.* 53, 11645-11649.

9.2 Manuscript F – Supplementary information

Versatile polyketide-chain lactonization by a non-canonical ketosynthase of the rhizoxin megasynthase

Srividhya Sundaram¹, Tawatchai Thongkongkaew¹, Hak Joong Kim¹, Ruth Bauer¹, Daniel Heine¹, Christian Hertweck^{1, 2, 3*}

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

General

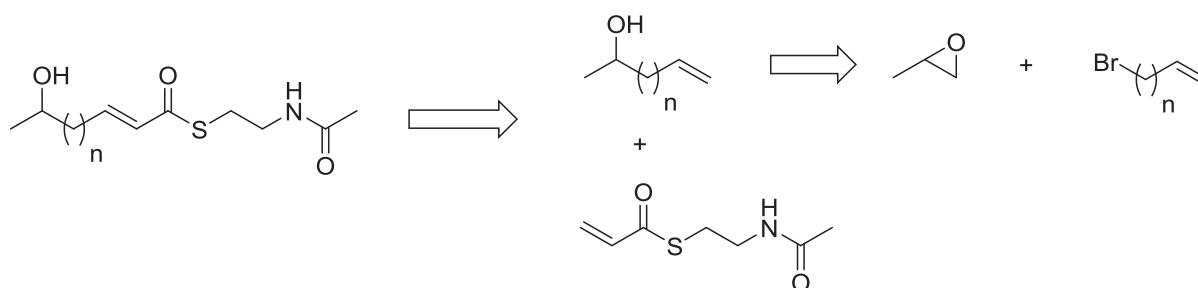
All reagents, 3-buten-2-ol and 4-penten-2-ol were commercially purchased from Sigma-Aldrich and Roth. All the solvents used for preparative and analytical measurements were obtained in HPLC grade.

Chemical synthesis of (*R,S*)-(*E*)-*S*-(2-acetamidoethyl) thioate derivatives

According to (Bretschneider et al., 2013), an (*R,S*)-(*E*)-*S*-(2-acetamidoethyl) 5-hydroxyhex-2-enethioate was synthesized by a reduction-olefination sequence in 6 steps. This method has three disadvantages. The reaction sequence is long and time consuming. It also requires a protecting group. Moreover, the synthetic route is not flexible for longer and shorter substrates because the starting materials are not commercially available. Thus, we generated the substrates using a new synthetic route.

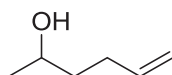
Retrosynthetic analysis

The (*R,S*)-(*E*)-*S*-(2-acetamidoethyl) thioate derivatives can be synthesized by a cross metathesis reaction between the *S*-(2-acetamidoethyl) prop-2-enethiolate and alken-2-ol analogues.

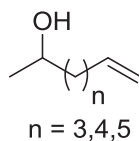


Chemical synthesis of alken-2-ol analogues

5-hexen-2-ol

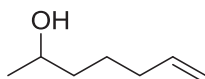


The 5-hexen-2-ol was synthesized by reduction of 5-hexen-2-one according to a method described by (Warner and Bäckvall, 2015). Its spectroscopic data is in agreement with the previous report.

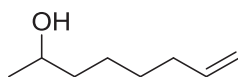
General procedure for the synthesis of 6-hepten-2-ol, 7-octen-2-ol, and 8-nonen-2-ol

6-hepten-2-ol, 7-octen-2-ol, and 8-nonen-2-ol were synthesized by epoxide opening with suitable organo cuprates (Leijondahl et al., 2009). Specifically, 7.7 mmole of Mg(s) was added into a heat dry schlank flask under Ar. Next, 4 mL of dry THF was added and the mixture was stirred overnight. Then 3 drops of 1,2-dibromoethane was added. 1 mL of bromoalkene was added slowly and the reaction was stirred at room temperature for 1 hour.

In another schlank flask, 0.5 ml of epoxypropane and 143 mg of CuI, and 4 mL of dry THF were added. The flask was cooled at -78°C . Next, the Grignard reagent was added dropwise. The reaction was stirred at -78°C for 1.5 hours and at room temperature for another 1 hour. To work-up, 10 mL of saturated aqueous NH_4Cl was added. The reaction was extracted 3 times with 10 mL Et_2O . The organic layer was combined, dried with Na_2SO_4 , and evaporated to dryness. The crude reaction was purified by silica gel column chromatography using 20% Et_2O : hexane as eluent to yield 395.4 mg (46%) of 6-hepten-2-ol, 758.8 mg (78%) of 7-octen-2-ol, 632.8 mg (60%) of 8-nonen-2-ol.

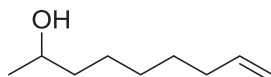
6-Hepten-2-ol

Molecular formula: $\text{C}_7\text{H}_{14}\text{O}$; **HRESIMS:** m/z 115.1125 $[\text{M}+\text{H}]^+$ (calculated 115.1123, Δm +1.7 ppm); **IR (film) ν (cm^{-1}):** 3337 (O-H stretching), 3076 ($\text{C}_{\text{sp}^2}\text{-H}$ stretching), 1641 (C=C stretching); **^1H (CDCl_3 , 500 MHz):** δ 5.79 (1H, m), 4.99 (1H, dq, $J = 17.0, 1.6$ Hz), 4.93 (1H, dd, $J = 9.5, 1.5$ Hz), 3.78 (1H, m), 2.05 (2H, m), 1.36-1.54 (4H, m), 1.17 (3H, d, $J = 6.5$ Hz) ppm; **^{13}C (CDCl_3 , 125 MHz):** δ 138.7, 114.6, 68.0, 38.7, 33.7, 25.0, 23.5 ppm.

7-Octen-2-ol

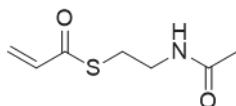
Molecular formula: C₈H₁₆O; **HRESIMS:** m/z 129.1281 [M+H]⁺ (calculated 129.1279, Δm +1.5 ppm); **IR (film) ν (cm⁻¹):** 3356 (O-H stretching), 3077 (C_{sp2}-H stretching), 1641 (C=C stretching); **¹H (CDCl₃, 500 MHz):** δ 5.78 (1H, m), 5.00 (1H, m), 4.92 (1H, m), 3.77 (1H, m), 2.04 (2H, m), 1.20-1.50 (6H, m), 1.16 (3H, d, J = 6.0 Hz) ppm; **¹³C (CDCl₃, 125 MHz):** δ 138.8, 114.4, 68.1, 39.1, 33.7, 28.8, 25.2, 23.5 ppm.

8-Nonen-2-ol



Molecular formula: C₉H₁₈O; **HRESIMS:** m/z 143.1432 [M+H]⁺ (calculated 143.1436, Δm -2.8 ppm); **IR (film) ν (cm⁻¹):** 3356 (O-H stretching), 3076 (C_{sp2}-H stretching), 1640 (C=C stretching); **¹H (CDCl₃, 500 MHz):** δ 5.78 (1H, m), 5.00 (1H, m), 4.91 (1H, m), 3.77 (1H, m), 2.02 (2H, m), 1.22-1.49 (8H, m), 1.16 (3H, d, J = 6.0 Hz) ppm; **¹³C (CDCl₃, 125 MHz):** δ 139.0, 114.2, 68.1, 39.2, 33.7, 29.1, 28.8, 25.6, 23.4 ppm

Chemical synthesis of S-(2-acetamidoethyl) prop-2-enethiolate

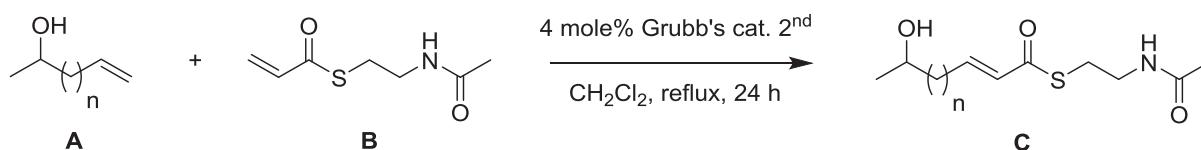


The S-(2-acetamidoethyl) prop-2-enethiolate was synthesized according to a method described by (Heine et al., 2014). Its spectroscopic data is in agreement with the previous report.

General procedure for a cross metathesis reaction

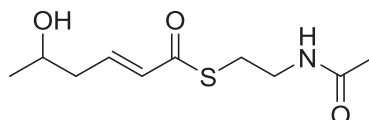
0.18 mmole of alken-2-ol (**A**) and 0.06 mole of S-(2-Acetamidoethyl) prop-2-enethiolate (**B**) were added into a heat dry 4 ml vial filled with 500 μL of dry CH₂Cl₂. 1 mg (2 mole %) of Grubb's catalyst second generation was added and another portion 1 mg (2 mole %) of catalyst was added after 16 hours. The reaction was heated at refluxing for 24 hours. To work-up, the reaction was dried under N₂ gas and a product was purified by silica gel column chromatography using 5% MeOH: CHCl₃ as eluent.

A general reaction scheme:

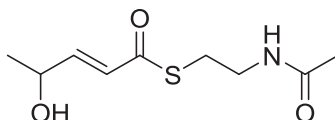


No	n	mmole		%yield	recovery of B (mmole)
		A	B		
1	0	0.18	0.06	43*	0.03
2	1	0.18	0.07	36*	0.03
3	2	0.18	0.06	24*	0.02
4	3	0.18	0.06	91*	0.02
5	4	0.18	0.06	58	-
5	5	0.18	0.07	72	-

*The percent yield was calculated based on a recovery of starting material.

(R,S)-(E)-S-(2-acetamidoethyl) 5-hydroxyhex-2-enethioate (5)

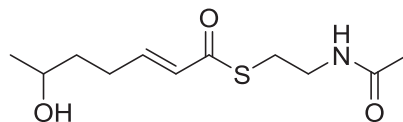
Molecular formula: C₁₀H₁₇NO₃S; **HRESIMS:** m/z 232.1001 [M+H]⁺ (calculated 232.1007, Δm -2.6 ppm); **IR (film) ν (cm⁻¹):** 3289 (O-H and N-H stretching), 3093 (C_{sp2}-H stretching), 1652 and 1629 (C=O and C=C stretching); **¹H (MeOH-d₄, 500 MHz):** δ 6.94 (1H, dt, J=15.5, 7.6 Hz), 6.22 (1H, dt, J=15.5, 1.3 Hz), 3.80 (1H, sextet, J = 6.2 Hz), 3.33 (2H, t, J = 6.8 Hz), 3.06 (2H, t, J = 6.5 Hz), 2.33 (2H, m), 1.91 (3H, s), 1.17 (3H, d, J = 6.0 Hz) ppm; **¹³C (MeOH-d₄, 125 MHz):** δ 190.9, 173.4, 143.9, 131.3, 67.4, 42.5, 40.2, 28.9, 23.3, 22.5 ppm

(R,S)-(E)-S-(2-acetamidoethyl) 4-hydroxypent-2-enethioate (7)

Molecular formula: C₉H₁₅NO₃S; **HRESIMS:** m/z 218.0847 [M+H]⁺ (calculated 218.0851, Δm -1.8 ppm); **IR (film) ν (cm⁻¹):** 3284 (O-H and N-H stretching), 3097 (C_{sp2}-H stretching), 1652 and 1633 (C=O and C=C stretching); **¹H (MeOH-d₄, 300 MHz):** δ 6.90 (1H, dd,

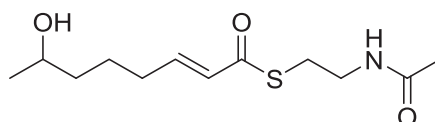
$J=15.3, 4.4$ Hz), 6.31 (1H, dd, $J=15.3, 6.6$ Hz), 4.40 (1H, m), 3.34 (2H, t, $J = 6.9$ Hz), 3.06 (2H, t, $J = 6.6$ Hz), 1.91 (3H, s), 1.27 (3H, d, $J = 6.6$ Hz) ppm; ^{13}C (MeOH- d_4 , 125 MHz): δ 191.2, 173.4, 149.6, 126.9, 67.5, 40.2, 29.0, 22.7, 22.5 ppm

(*R,S*)-(E)-S-(2-acetamidoethyl) 6-hydroxyhept-2-enethioate (9)



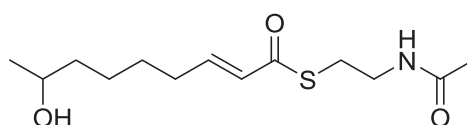
Molecular formula: $\text{C}_{11}\text{H}_{19}\text{NO}_3\text{S}$; **HRESIMS:** m/z 246.1159 $[\text{M}+\text{H}]^+$ (calculated 246.1164, Δm -2.0 ppm); **IR (film) ν (cm^{-1}):** 3289 (O-H and N-H stretching), 3091 ($\text{C}_{\text{sp}2}$ -H stretching), 1652 and 1627 (C=O and C=C stretching); **^1H (MeOH- d_4 , 500 MHz):** δ 6.94 (1H, dt, $J=15.5, 7.0$ Hz), 6.18 (1H, dt, $J=15.5, 1.4$ Hz), 3.72 (1H, sextet, $J = 6.2$ Hz), 3.34 (2H, t, $J = 6.7$ Hz), 3.05 (2H, t, $J = 6.7$ Hz), 2.30 (2H, m), 1.90 (3H, s), 1.56 (2H, m), 1.16 (3H, d, $J = 6.2$ Hz) ppm; **^{13}C (MeOH- d_4 , 75 MHz):** δ 190.9, 173.4, 147.2, 129.5, 67.7, 40.2, 38.2, 29.5, 28.8, 23.4, 22.4 ppm

(*R,S*)-(E)-S-(2-acetamidoethyl) 7-hydroxyoct-2-enethioate (11)



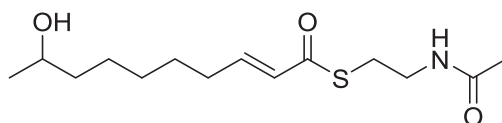
Molecular formula: $\text{C}_{12}\text{H}_{21}\text{NO}_3\text{S}$; **HRESIMS:** m/z 260.1314 $[\text{M}+\text{H}]^+$ (calculated 260.1320, Δm -2.3 ppm); **IR (film) ν (cm^{-1}):** 3494 (O-H and N-H stretching), 3087 ($\text{C}_{\text{sp}2}$ -H stretching), 1653 and 1631 (C=O and C=C stretching) ; **^1H (MeOH- d_4 , 300 MHz):** δ 6.92 (1H, dt, $J=15.6, 7.4$ Hz), 6.17 (1H, dt, $J=15.6$ Hz), 3.71 (1H, sextet, $J = 6.1$ Hz), 3.32 (2H, t, $J = 6.6$ Hz), 3.04 (2H, t, $J = 6.8$ Hz), 2.22 (2H, m), 1.90 (3H, s), 1.27-1.69 (4H, m), 1.14 (3H, d, $J = 6.0$ Hz) ppm; **^{13}C (MeOH- d_4 , 75 MHz):** δ 191.0, 173.5, 147.3, 129.6, 68.2, 40.2, 39.6, 33.1, 28.9, 25.4, 23.5, 22.5 ppm

(*R,S*)-(E)-S-(2-acetamidoethyl) 8-hydroxynon-2-enethioate (13)



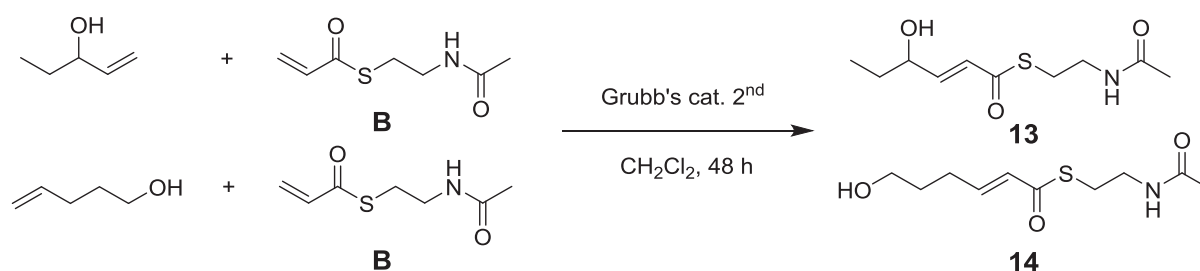
Molecular formula: C₁₃H₂₃NO₃S; **HRESIMS:** m/z 274.1472 [M+H]⁺ (calculated 274.1477, Δm -1.8 ppm); **IR (film) ν (cm⁻¹):** 3286 (O-H and N-H stretching), 1652 and 1629 (C=O and C=C stretching); **¹H (MeOH-d₄, 300 MHz):** δ 6.92 (1H, dt, J=15.6, 7.0 Hz), 6.17 (1H, dt, J=15.3, 1.4 Hz), 3.70 (1H, sextet, J = 5.8 Hz), 3.32 (2H, t, J = 6.6 Hz), 3.04 (2H, t, J = 6.8 Hz), 2.22 (2H, m), 1.91 (3H, s), 1.28-1.57 (6H, m), 1.13 (3H, d, J = 6.3 Hz) ppm; **¹³C (MeOH-d₄, 75 MHz):** δ 191.0, 173.4, 147.4, 129.6, 68.4, 40.2, 39.8, 33.1, 29.2, 28.9, 26.5, 23.5, 22.6 ppm

(R,S)-(E)-S-(2-acetamidoethyl) 9-hydroxydec-2-enethioate (15)



Molecular formula: C₁₄H₂₅NO₃S; **HRESIMS:** m/z 288.1628 [M+H]⁺ (calculated 288.1633, Δm -1.7 ppm); **IR (film) ν (cm⁻¹):** 3289 (O-H and N-H stretching), 1652 and 1632 (C=O and C=C stretching); **¹H (MeOH-d₄, 300 MHz):** δ 6.92 (1H, dt, J=15.6, 7.1 Hz), 6.16 (1H, dt, J=15.3, 1.5 Hz), 3.70 (1H, sextet, J = 5.9 Hz), 3.33 (2H, t, J = 6.6 Hz), 3.05 (2H, t, J = 6.8 Hz), 2.22 (2H, m), 1.91 (3H, s), 1.28-1.55 (8H, m), 1.13 (3H, d, J = 6.0 Hz) ppm; **¹³C (MeOH-d₄, 75 MHz):** δ 191.0, 173.4, 147.5, 129.6, 68.5, 40.2, 40.1, 33.1, 30.3, 29.1, 28.9, 26.6, 23.5, 22.5 ppm

Chemical synthesis of (R,S)-E-S-(2-acetamidoethyl)-4-hydroxyhex-2-enethioate (21) and (R,S)-E-S-(2-acetamidoethyl)-6-hydroxyhex-2-enethioate (22)



For synthesis of (R,S)-E-S-(2-acetamidoethyl)-4-hydroxyhex-2-enethioate **21**, 1-penten-3-ol (0.22 mmol, 17.7 μl) was added to SNAC ester **B** (0.06 mmol, 10.0 mg). The compound (R,S)-E-S-(2-acetamidoethyl)-6-hydroxyhex-2-enethioate **22**, was synthesized analogously using 4-penten-1-ol (0.22 mmol, 17.7 μl) and SNAC ester **B** (0.06 mmol, 11.1 mg). The mixture was dissolved in dry dichloromethane (0.5 ml) and the mixture was stirred for two

days at room temperature. Hoveyda-Grubbs 2nd generation catalyst (10 mol %, **21**: 0.01 mmol, 13 mg; **22**: 0.02 mmol, 14 mg) was added in two portions at day one and day two. The solvent was removed under a nitrogen stream and the product was purified by open column chromatography. A gradient of chloroform:methanol (98:2) and chloroform:methanol (95:5) was used for elution to give 4.2 mg of **21** and 7.8 mg of **22**.

(*R,S*)-*E-S*-(2-acetamidoethyl)-4-hydroxyhex-2-enethioate **21** (4.2 mg, yield: 42%). Colorless oil; ¹H NMR (500 MHz, MeOD): δ 6.88 (1H, dd, ³*J*₁ = 15.5 Hz; ³*J*₂ = 4.7 Hz), 6.32 (1H, dd, ³*J*₁ = 15.5 Hz; ⁴*J*₂ = 1.6 Hz), 4.17 (1H, m), 3.34 (2H, t, *J* = 6.7 Hz), 3.07 (2H, t, *J* = 6.7 Hz), 1.91 (3H, s), 1.58 (2H, m), 0.95 (3H, t, *J* = 7.5 Hz); ¹³C NMR (600 MHz, MeOD): δ 191.26 (COS), 173.55 (CON), 148.64 (CH), 127.65 (CHCO), 72.72 (COH), 40.14 (NCH₂), 30.40 (SCH₂), 29.05 (CH₂), 22.48 (CH₃CO), 9.98 (CH₃) ppm; HRMS: (ESI+): *m/z* calculated for C₁₀H₁₈O₃NS: 232.1005, found 232.1002 [M+H]⁺.

(*R,S*)-*E-S*-(2-acetamidoethyl)-6-hydroxyhex-2-enethioate **22** (7.8 mg, yield: 70%). Colorless oil; ¹H NMR (500 MHz, MeOD): δ 6.94 (1H, dt, ³*J*₁ = 15.6 Hz; ³*J*₂ = 7.0 Hz), 6.19 (1H, dt, ³*J*₁ = 15.5 Hz; ⁴*J*₂ = 1.5 Hz), 3.56 (2H, t, *J* = 6.4 Hz), 3.33 (2H, t, *J* = 6.7 Hz), 3.05 (2H, t, *J* = 6.6 Hz), 2.29 (2H, m), 1.91 (3H, s), 1.68 (2H, m); ¹³C NMR (600 MHz, MeOD): δ 191.23 (COS), 173.56 (CON), 147.04 (CH), 129.71 (CHCO), 62.00 (COH), 40.21 (NCH₂), 31.90 (SCH₂), 29.56 (CH₂), 28.91 (CH₂), 22.50 (CH₃O) ppm; HRMS: (ESI+): *m/z* calculated for C₁₀H₁₈O₃NS: 232.1003, found 232.1002 [M+H]⁺.

To achieve sufficient purity of the substrate analogues for the *vitro* branching assay studies, the fractions containing the SNAC esters **21** and **22** were purified again by semi-preparative HPLC on a 1260 Infinity system using the column Eclipse XDB-C18, 250 x 4.6 mm, particle size 5 μ m (Agilent Technologies). SNAC esters **21** and **22** were characterized by HRMS, ¹H and ¹³C NMR measurements.

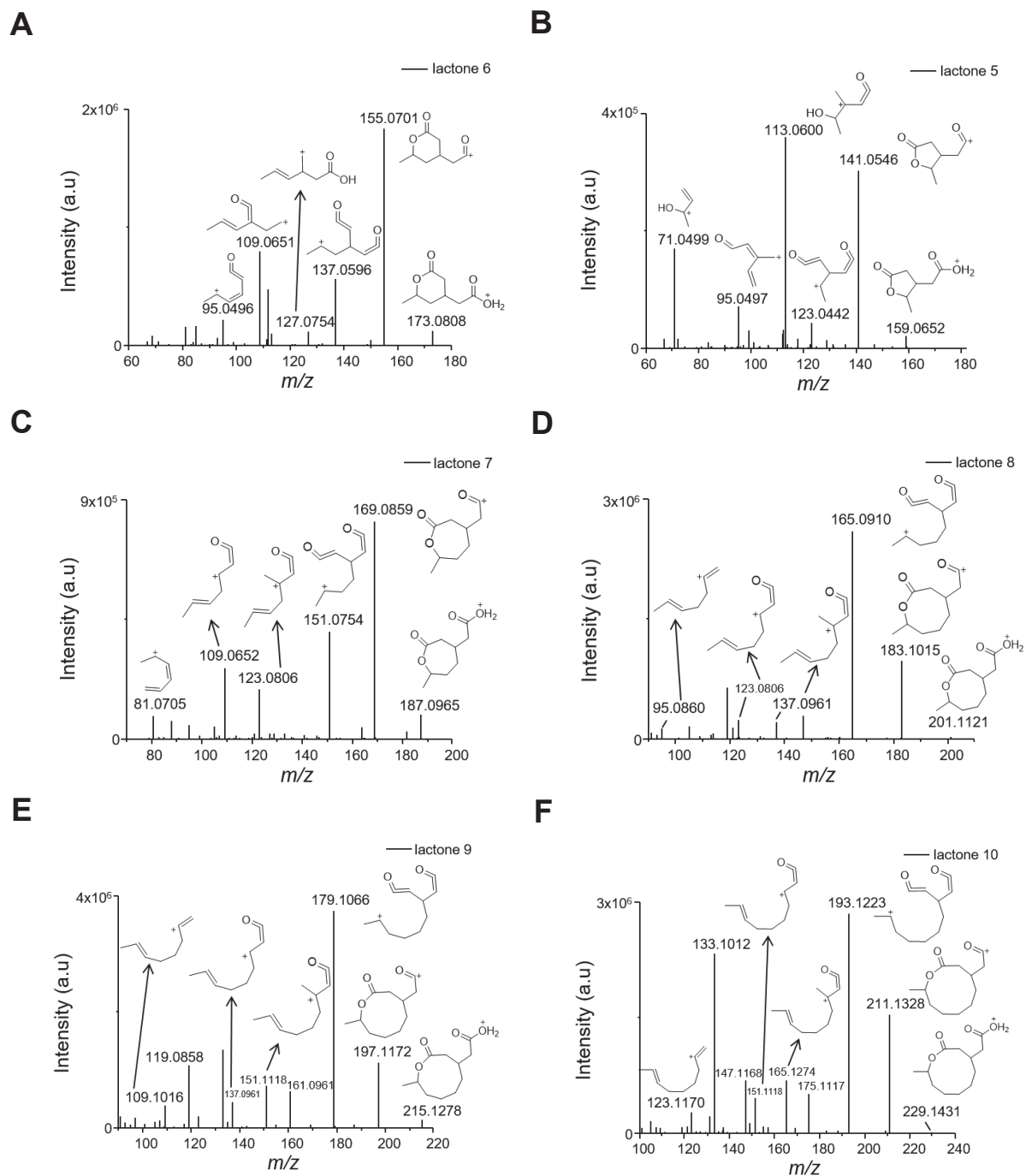


Figure S1 (A-F), related to figure 3: LC-MS/MS of 6, 8, 10, 12, 14 and 16 respectively.

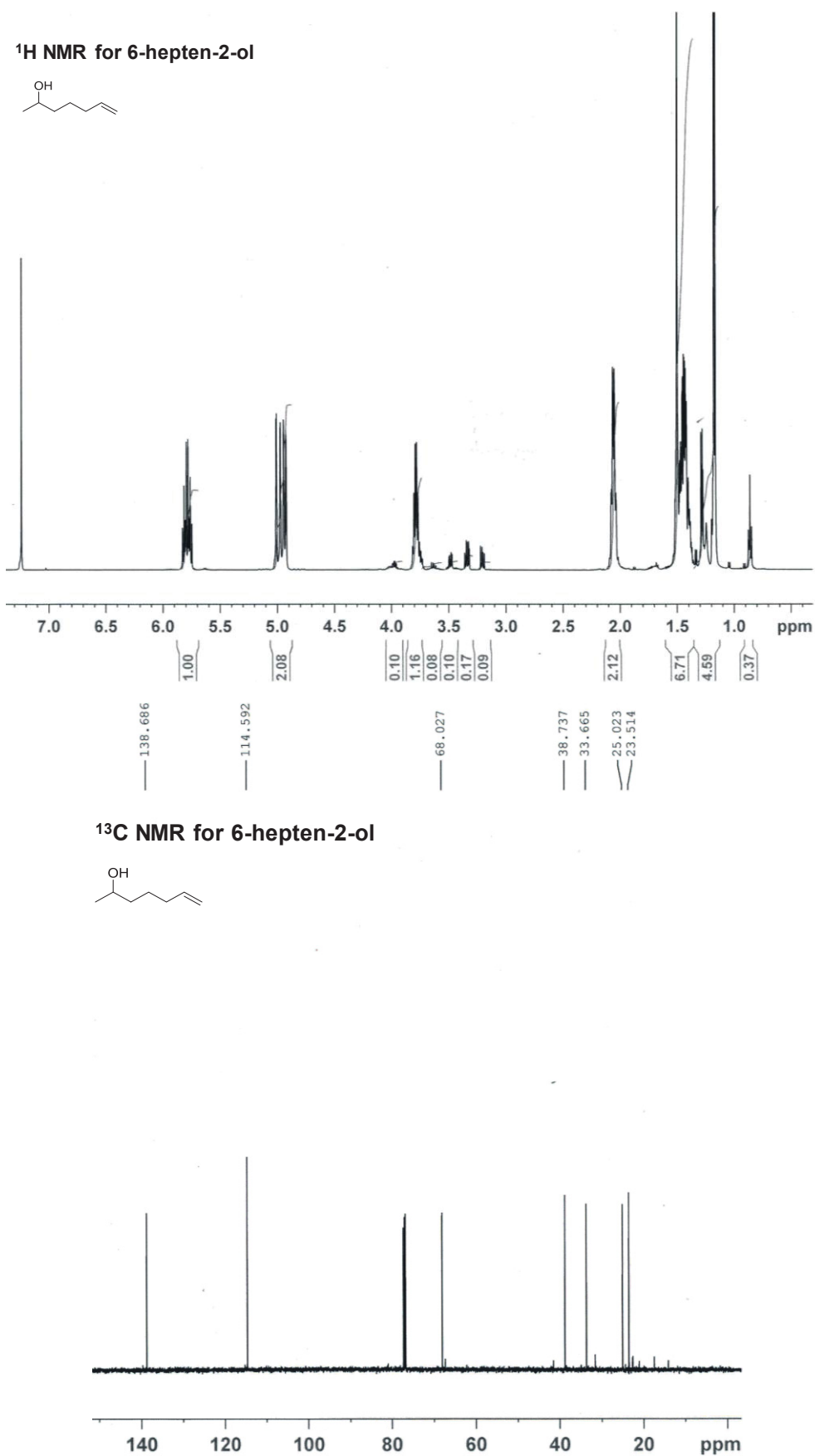


Figure S2 ¹H and ¹³C NMR spectra of 6-hepten-2-ol.

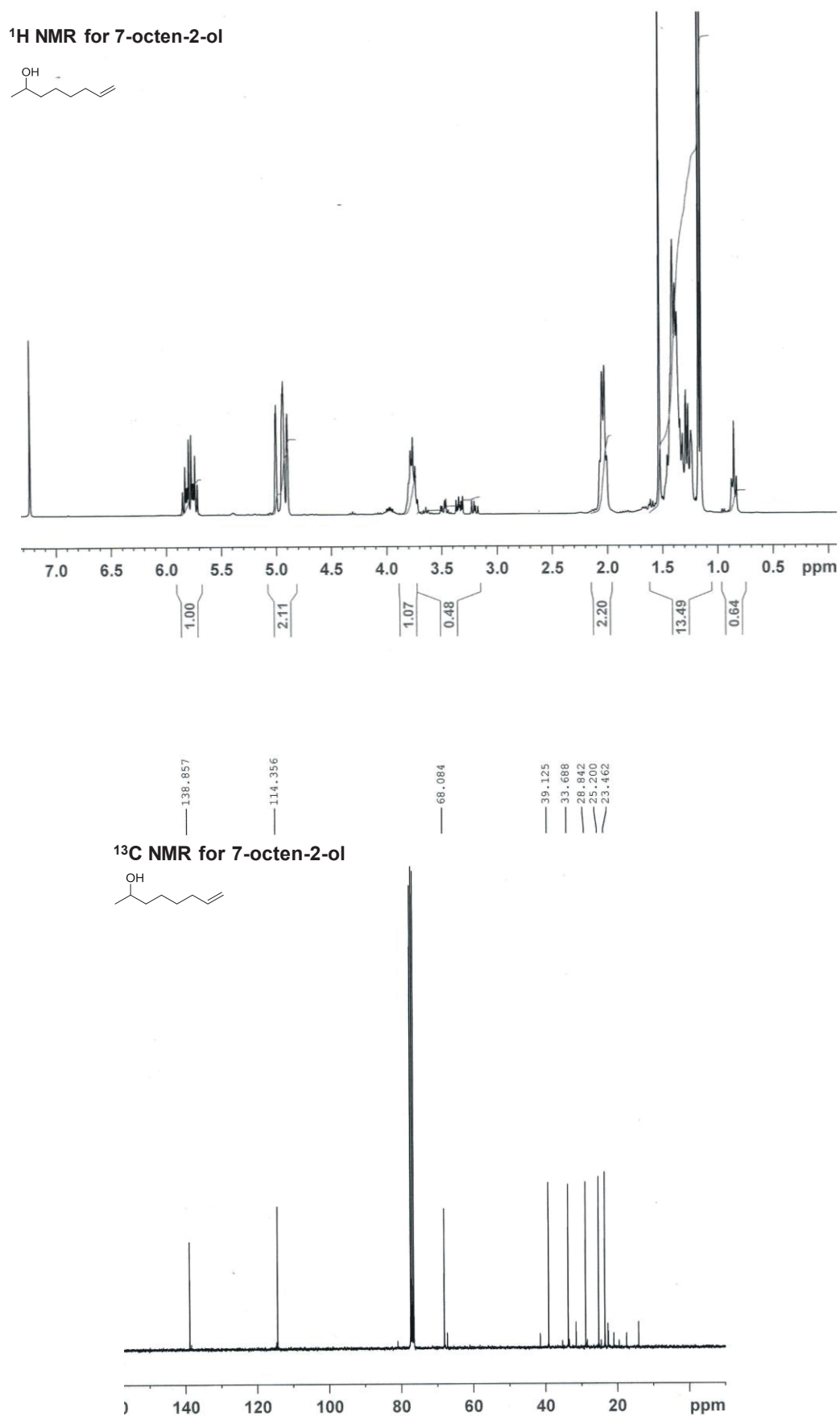


Figure S3 ¹H and ¹³C NMR spectra of 7-octen-2-ol.

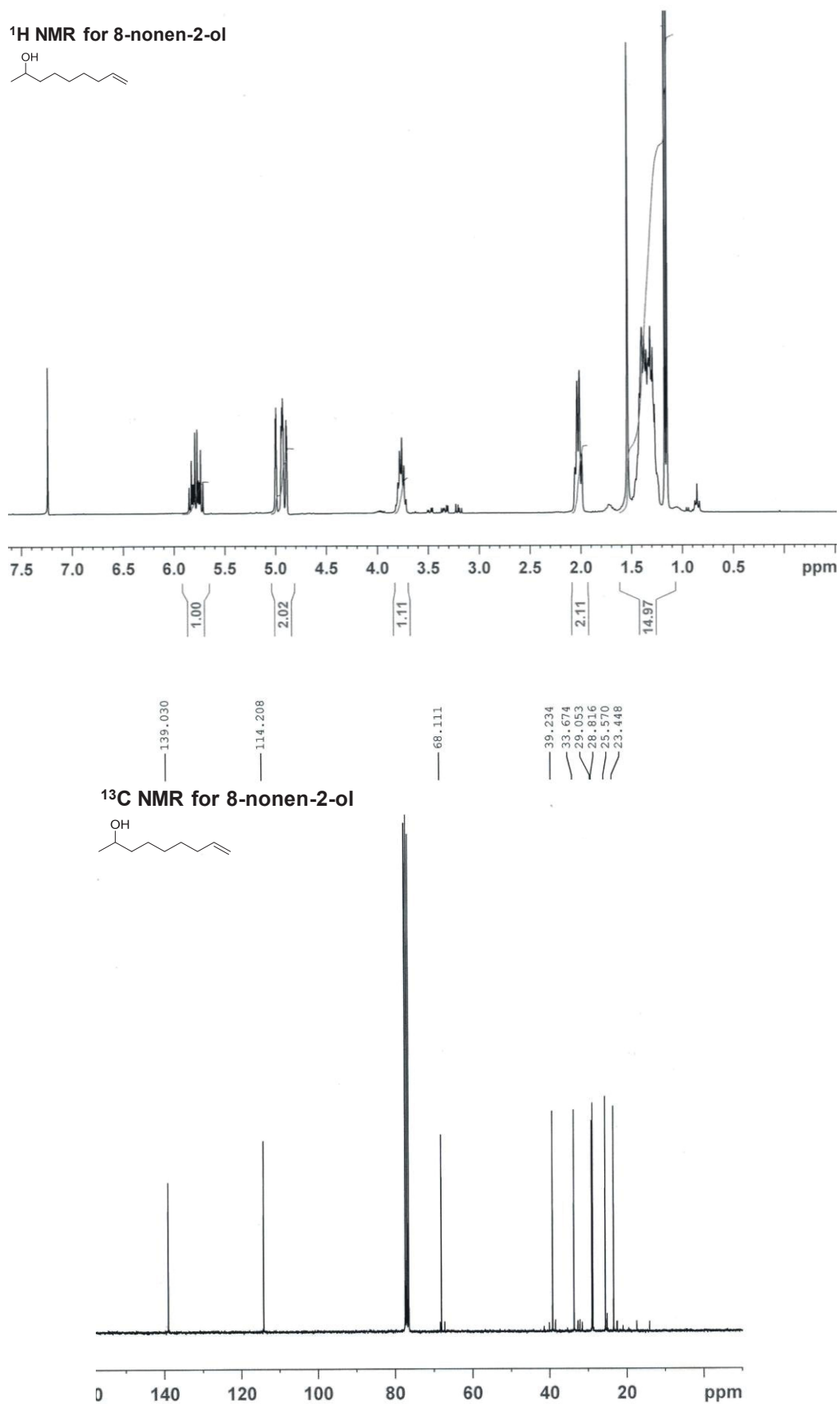
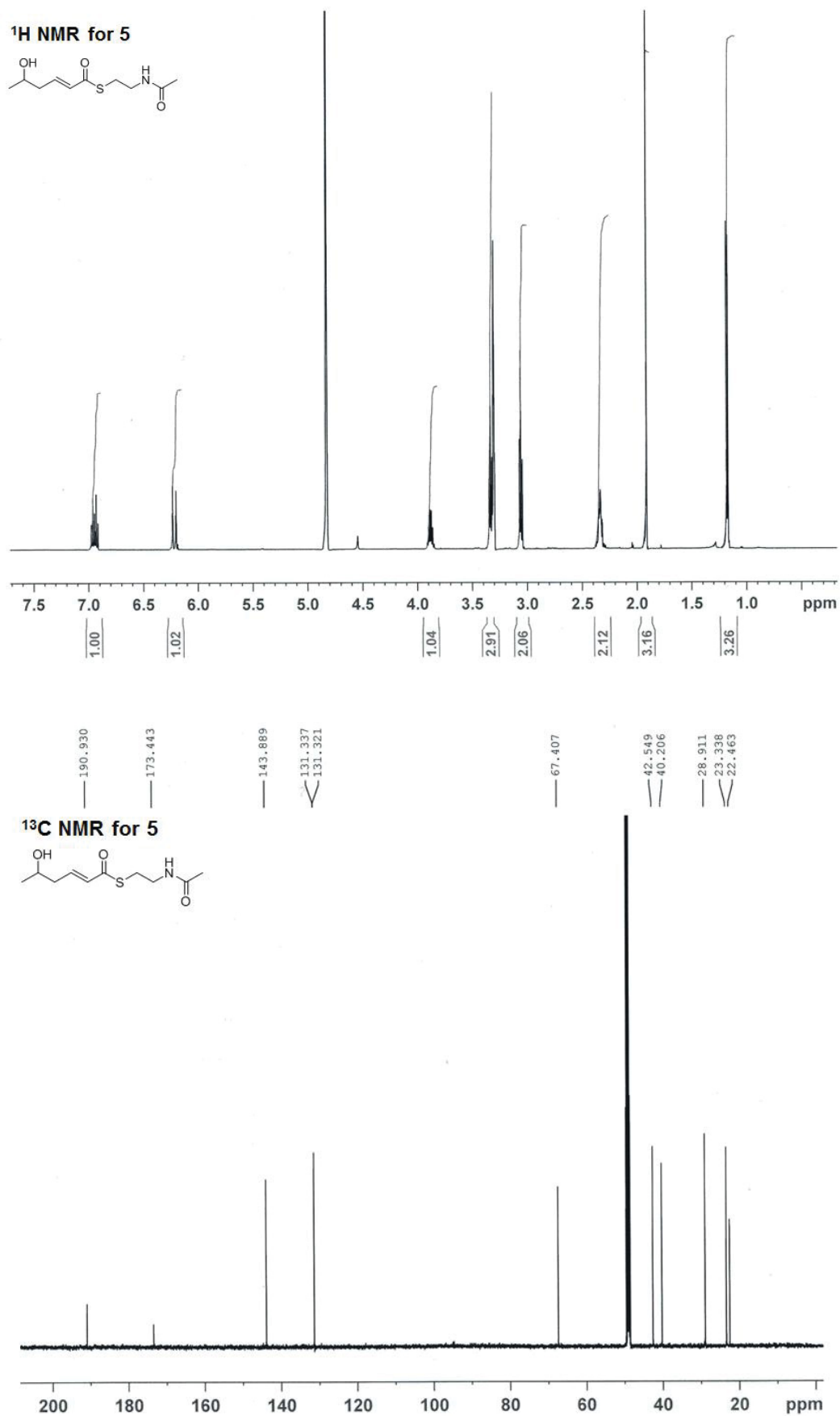


Figure S4 ¹H and ¹³C NMR spectra of 8-nonen-2-ol.

Figure S5 ¹H and ¹³C NMR spectra of 5.

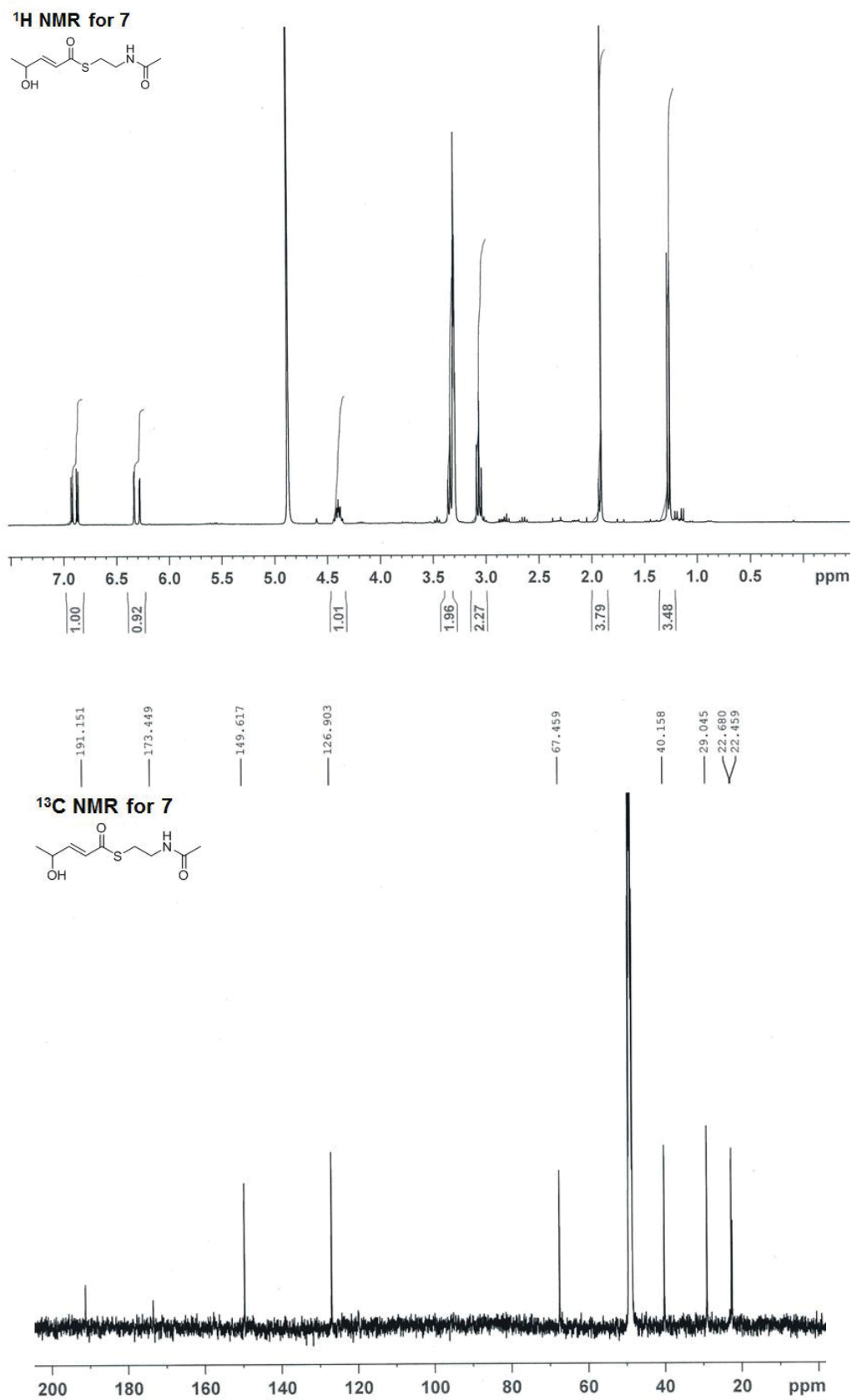


Figure S6 ¹H and ¹³C NMR spectra of 7.

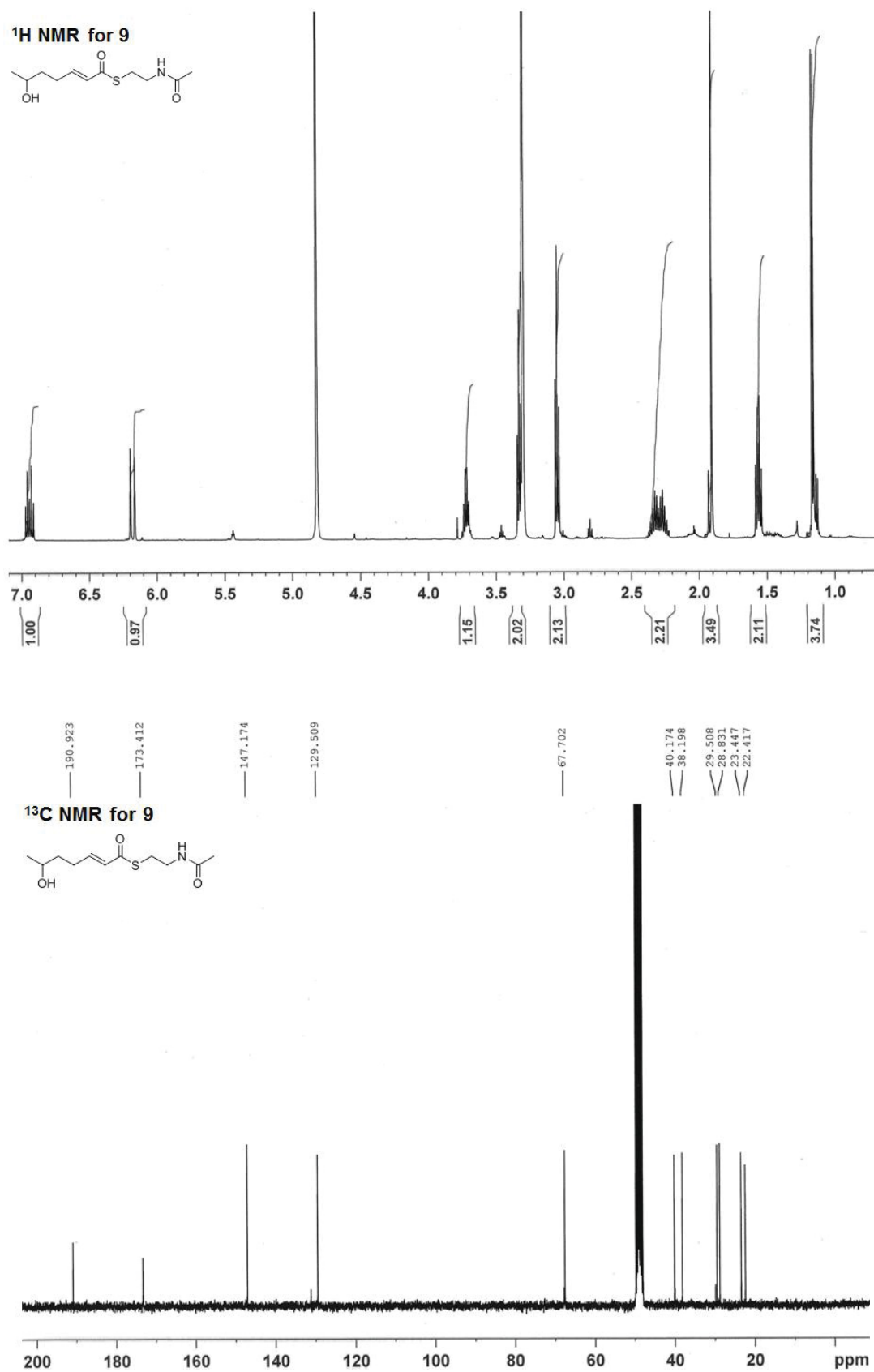


Figure S7 ¹H and ¹³C NMR spectra of 9.

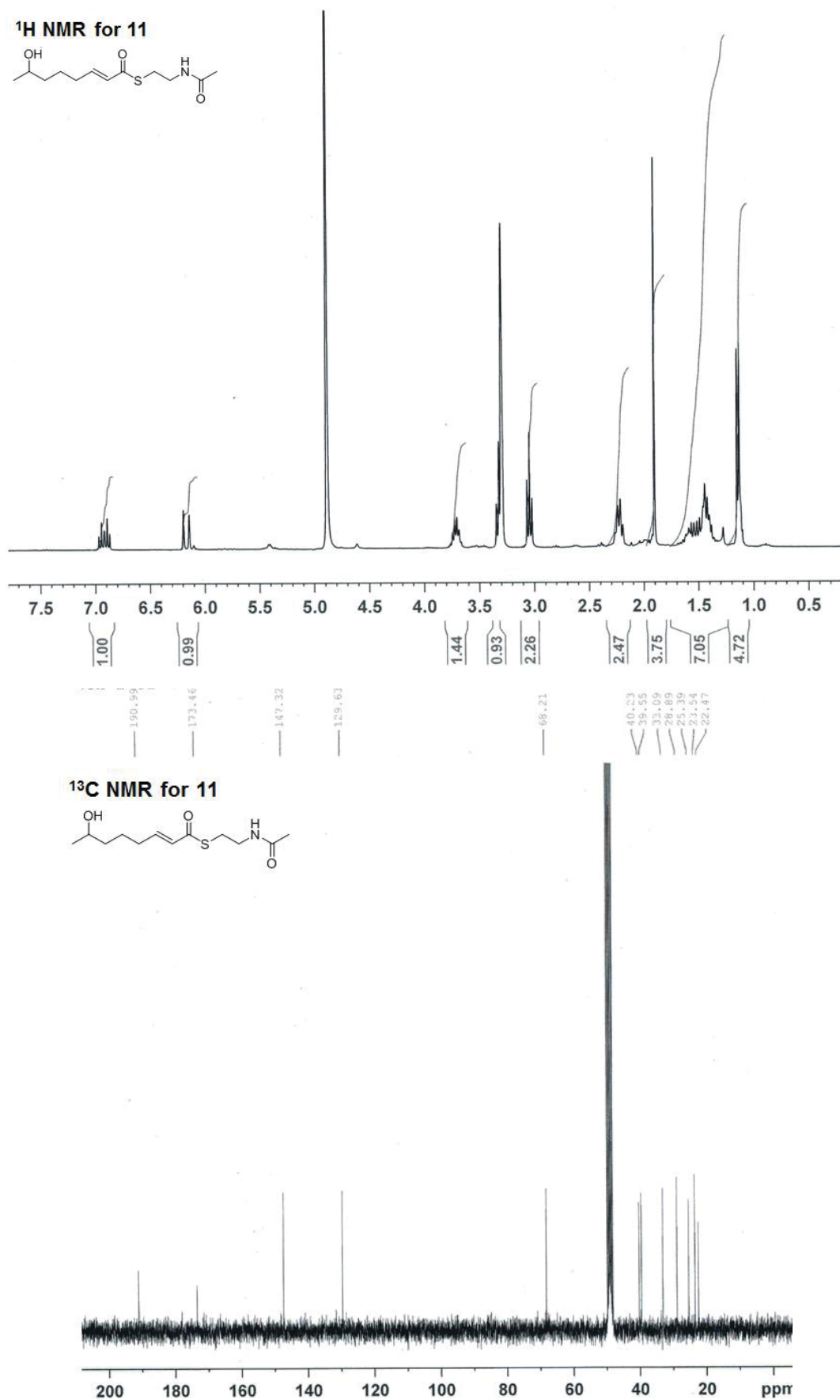
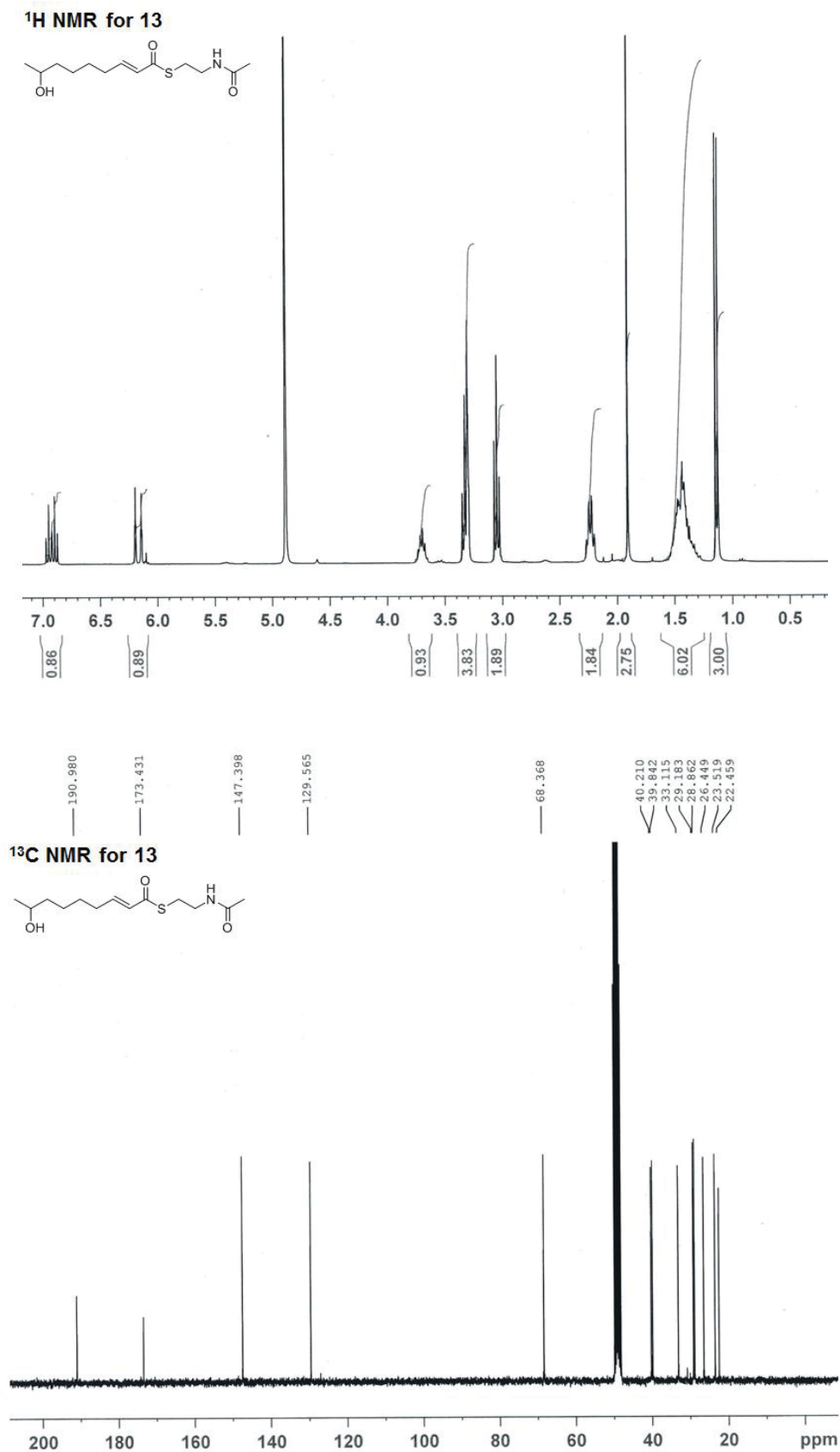


Figure S8 ¹H and ¹³C NMR spectra of 11.



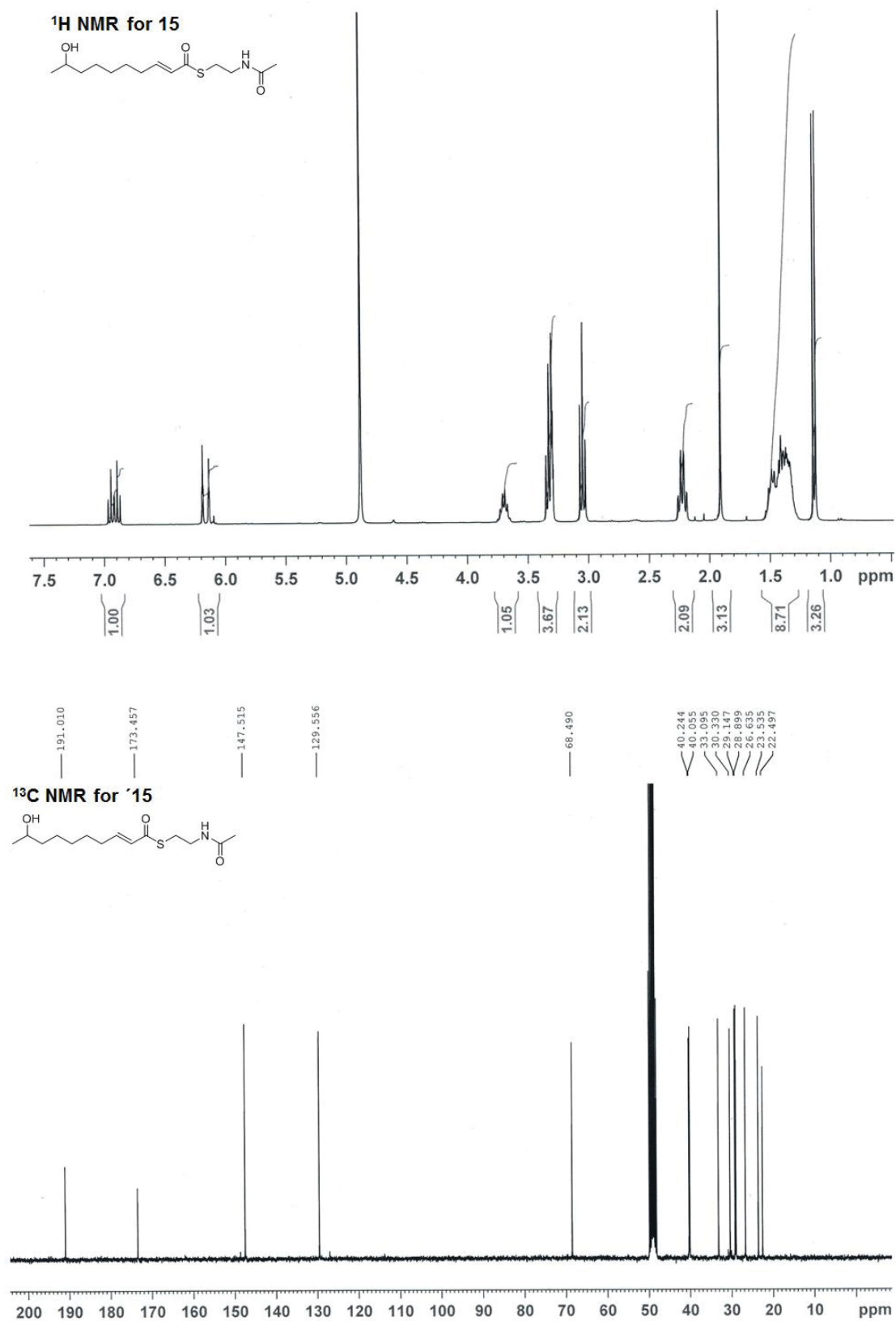
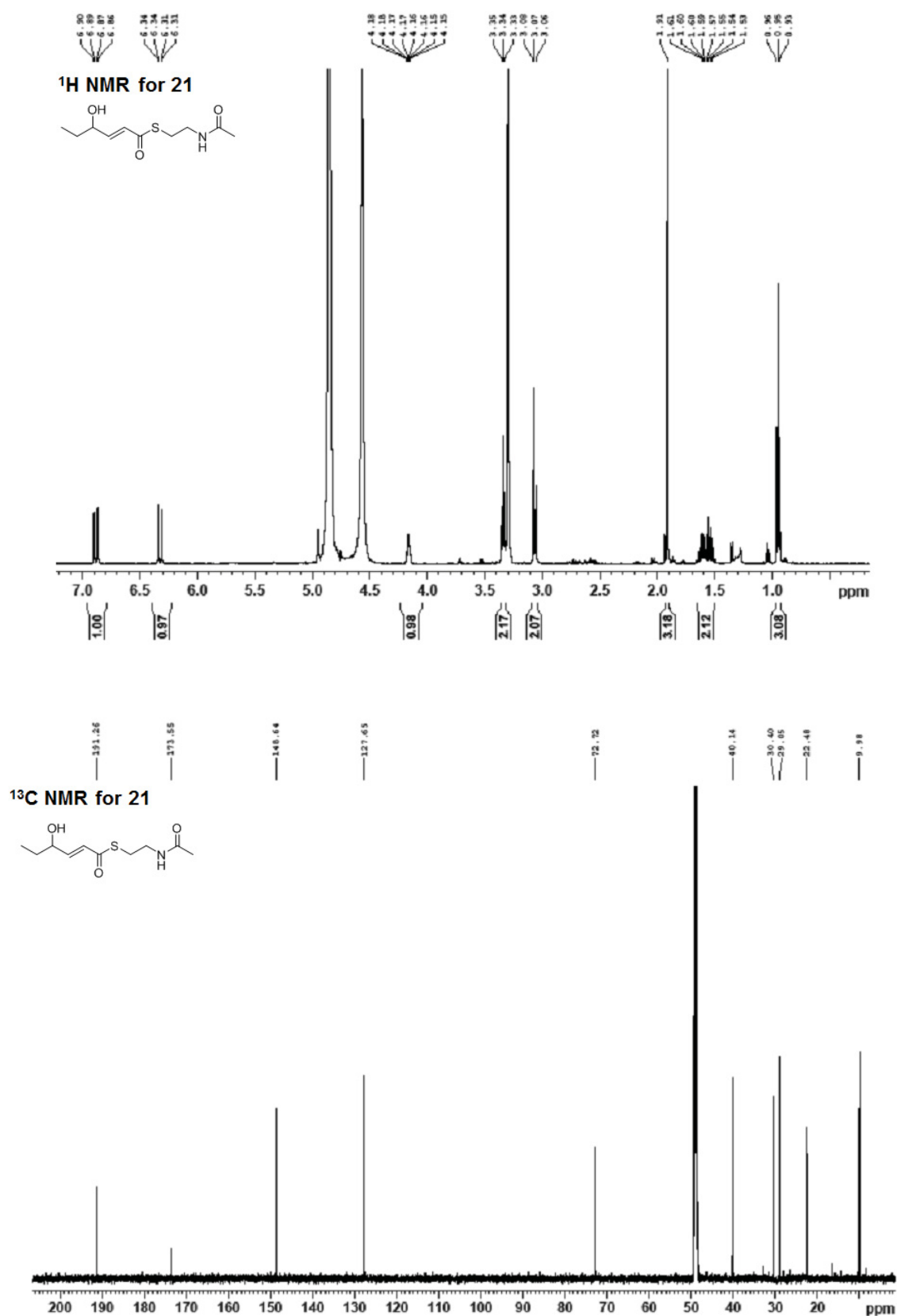


Figure S10 ¹H and ¹³C NMR spectra of 15.

Figure S11 ¹H and ¹³C NMR spectra of 21.

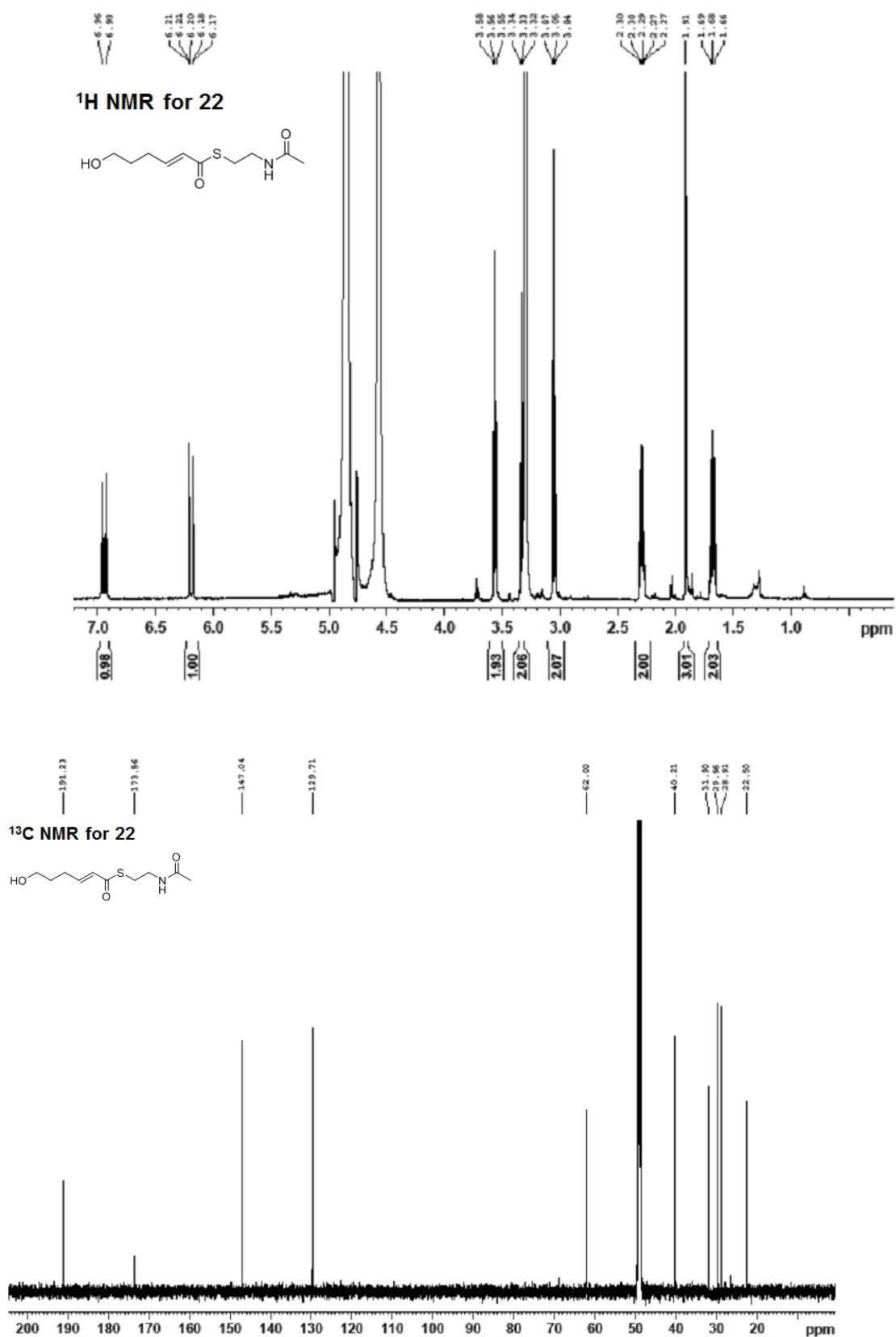


Figure S12 ¹H and ¹³C NMR spectra of 22.

REFERENCES

Bretschneider, T., Heim, J.B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013). Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* 502, 124-128.

Heine, D., Bretschneider, T., Sundaram, S., and Hertweck, C. (2014). Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles. *Angew. Chem. Int. Ed.* 53, 11645-11649.

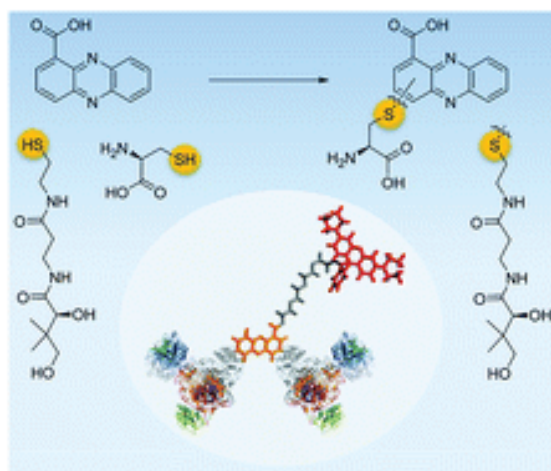
Leijondahl, K., Borén, L., Braun, R., and Bäckvall, J.-E. (2009). Enzyme- and ruthenium-catalyzed dynamic kinetic asymmetric transformation of 1, 5-diols. application to the synthesis of (+)-solenopsin A. *J. Org. Chem.* 74, 1988-1993.

Warner, M.C., and Bäckvall, J.E. (2015). Racemization of olefinic alcohols by a carbonyl (cyclopentadienyl) ruthenium complex: Inhibition by the carbon-carbon double bond. *European J. Org. Chem.* 2015, 2388-2393.

9.3 Manuscript G

A widespread bacterial phenazine forms S-conjugates with biogenic thiols and crosslinks proteins

Daniel Heine, Srividhya Sundaram, Matthias Beudert, Karin Martin, Christian Hertweck: A widespread bacterial phenazine forms S-conjugates with biogenic thiols and crosslinks proteins, *Chemical Science* 7, 4848-4855, 2016.



The discovery of unprecedented pantetheine conjugates (panphenazines A and B) and phenazine-1-carboxylic acid (PCA) from a *Kitasatospora* sp. shed light on the novel mechanism of the phenazines to form radical-based thiol conjugation and crosslinking of proteins.

CrossMark
click for updatesCite this: *Chem. Sci.*, 2016, 7, 4848

A widespread bacterial phenazine forms S-conjugates with biogenic thiols and crosslinks proteins†

D. Heine,^a S. Sundaram,^a Matthias Beudert,^a K. Martin^a and C. Hertweck^{*ab}

Phenazines are redox-active compounds produced by a range of bacteria, including many pathogens. Endowed with various biological activities, these ubiquitous N-heterocycles are well known for their ability to generate reactive oxygen species by redox cycling. Phenazines may lead to an irreversible depletion of glutathione, but a detailed mechanism has remained elusive. Furthermore, it is not understood why phenazines have so many protein targets and cause protein misfolding as well as their aggregation. Here we report the discovery of unprecedented conjugates (panphenazines A, B) of panthetheine and phenazine-1-carboxylic (PCA) acid from a *Kitasatospora* sp., which prompted us to investigate their biogenesis. We found that PCA reacts with diverse biogenic thiols under radical-forming conditions, which provides a plausible model for irreversible glutathione depletion. To evaluate the scope of the reaction in cells we designed biotin and rhodamine conjugates for protein labelling and examined their covalent fusion with model proteins (ketosynthase, carbonic anhydrase III, albumin). Our results reveal important, yet overlooked biological roles of phenazines and show for the first time their function in protein conjugation and crosslinking.

Received 1st February 2016
Accepted 13th April 2016

DOI: 10.1039/c6sc00503a

www.rsc.org/chemicalscience

Introduction

From an evolutionary perspective, it is most remarkable how small biomolecules have been shaped into privileged structures that fulfil multiple biological functions. This is particularly true for the large family of phenazines, ancient N-heterocycles that are produced by a large variety of bacteria.¹ Initially discovered as colourful bacterial pigments over 150 years ago, to date over one hundred natural phenazines have been characterised, most of which are endowed with highly diverse bioactivities.² These findings inspired the chemical synthesis of thousands of phenazine derivatives and propelled the development of antibiotics and anticancer agents.^{1–3} However, phenazines are a remarkable example of antibiotic compounds that are actually crucial for microbial interaction and competitive fitness.^{4–6} They not only act as signalling molecules in microbial communication,^{7–9} mediating biofilm formation and morphogenesis of colonies,¹⁰ but also control gene expression¹¹ and influence the intracellular redox state¹⁰ as well as exocellular electron transfer.¹² Owing to their redox properties, phenazines play major roles in infection processes of pathogens such as *Pseudomonas*

aeruginosa, especially in cystic fibrosis patients.¹³ In particular, pyocyanin (PYO, **1**) and phenazine-1-carboxylic acid (PCA, **2**), the key intermediate in phenazine biosynthesis,¹⁴ represent important virulence factors of *P. aeruginosa* that damage infected tissue (Fig. 1).^{15,16} The primary reason for the deleterious effects is phenazine-mediated redox cycling,¹ which leads to high concentrations of reactive oxygen species (ROS). Naturally, ROS is effectively counteracted by reduced glutathione (GSH) as redox buffer, with formation of the oxidised disulfide form (GSSG).^{17,18} Despite an impressive body of data on phenazine functions, there are numerous open questions on their mode of action. Specifically, it was observed that phenazines

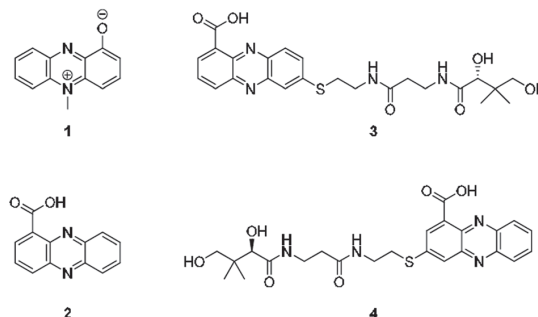


Fig. 1 Structures of pyocyanin (**1**), PCA (**2**), and panphenazines A (**3**) and B (**4**) discovered in this study.

^aLeibniz Institute for Natural Product Research and Infection Biology, Hans Knoell Institute, Beutenbergstrasse 11a, 07745 Jena, Germany. E-mail: christian.hertweck@leibniz-hki.de

^bFriedrich Schiller University, 07737 Jena, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6sc00503a



may lead to a decrease in GSH levels by conjugation, but the reaction mechanism has remained elusive.^{19–22} Moreover, it was found that ROS formation alone is neither responsible for the observed cytotoxic effects nor for misfolding and aggregation of proteins in the model organism *C. elegans* when exposed to phenazines.²³ Notably, addition of antioxidative agents failed to rescue the proteotoxic and protein aggregation phenotypes.²³ In this context it is remarkable that phenazine-mediated toxicity may be attributed to autophagy, which can result from high titres of misfolded proteins.²⁴ Yet, a link between non-functional proteins and phenazines has been missing. Furthermore, it is not yet sufficiently understood why phenazines have such broad activity profiles and highly diverse biological targets.^{2,25} Here, we report the discovery of unprecedented phenazine–pantetheine conjugates and the elucidation of their biogenesis, which revealed an overlooked mode of action of phenazines. Using two different synthetic probes we analysed the ability of phenazines to form thiol conjugates and crosslink proteins by a radical mechanism, which may clarify previously inexplicable disease symptoms and enigmatic phenazine functions.

Results and discussion

Discovery of panphenazines

Through genome mining²⁶ of the rare actinomycete *Kitasatospora* sp. HKI 714 we found that this soil-derived actinobacterium has the potential to produce prenylated phenazines. The genomics-guided search led to the discovery of a series of new endophenazines along with acylated phenazine derivatives.²⁷ As the strain produced large amounts of PCA (2), the universal precursor of phenazine derivatives, we expected to discover more congeners. In fact, metabolic profiling of concentrated culture extracts indicated the presence of unusual sulfur-containing phenazines. HRMS data ($m/z = 501.1807$ ($[M + H]^+$ for 3), $m/z = 501.1808$ ($[M + H]^+$ for 4)) and sulfur isotope pattern analysis suggested a molecular formula of $C_{24}H_{29}O_6N_4S$ for both compounds. Because of the extremely low titres of the new compounds ($\approx 5 \mu\text{g L}^{-1}$) the strain needed to be grown on a large scale (200 L). By semipreparative HPLC we eventually succeeded in isolating both compounds, named panphenazine A (3, 1.4 mg) and B (4, 0.9 mg), in sufficient amounts for NMR analyses. NMR data confirmed the phenazine core, and the downfield-shift (145.2 ppm) of the substituted quaternary C-7 of 3 suggested the presence of a thioether-bound side chain. The adjacent ethylene unit belongs to a cysteamine residue, and HMBC data pointed towards a peptide bond to the carboxy terminus of β -alanine. The remaining part of the molecule represents pantoic acid, which is connected to the β -alanine by another amide bond. The side chain is therefore identical to pantetheine, the acyl group carrier of coenzyme A or acyl carrier proteins during fatty acid and polyketide biosynthesis. The position of the side chain was confirmed by COSY and HMBC experiments. In an analogous way we elucidated the structure of compound 4, which differs from 3 only in the substitution site (C-3 *in lieu* of C-7) (Fig. S1 and Table S1†). Thus, panphenazines 3 and 4 are highly unusual pantetheine adducts of PCA (2).

Biomimetic synthesis of PCA thiol conjugates

Considering the hundreds of known natural phenazines, reports on sulfur-containing derivatives are scarce. The few reported examples are two *N*-acetylcysteine conjugates, dermacozine J²⁸ and SB 212305,²⁹ and a partially characterised pyocyanin glutathione adduct.^{21,22} To date, practically nothing is known about their formation. Various enzymes have been described that mediate C–S bond formation in natural product biosynthetic pathways.^{30–36} However, inspection of the endophenazine (*epa*) biosynthesis gene cluster in the *Kitasatospora* sp. HKI 714 genome²⁷ did not reveal any candidate genes that could be involved in the formation of the thioconjugates. Considering the substitution pattern, it was conceivable that 3 and 4 have resulted from a non-enzymatic reaction. Possible scenarios are (a) an oxidative aromatic coupling, (b) a Michael addition followed by reoxidation of the aromatic system, and (c) a coupling reaction involving radicals.

To elucidate the mechanism of phenazine thioether formation, we performed a series of experiments using PCA (2) and *N*-acetylcysteine as a simplified surrogate of pantetheine. Since addition of mild oxidizing agents or bases or incubation of PCA with disulfides (cysteine, pantetheine) did not lead to *S*-conjugates, mechanisms involving oxidation or nucleophilic attack of a deprotonated thiol were ruled out. Instead, we found that thiol conjugation proceeded smoothly under the influence of sunlight. PCA has UV absorption maxima at 250 nm and 370 nm, which could serve for photo-excitation. Irradiation of the reaction mixture with an UV lamp (370 nm) yielded various regioisomeric conjugates, 5 (10%), 6 (8%), 7 (5%) and 8 (6%) (Fig. 2). HPLC–HRMS analysis indicated that, in addition to the monosubstituted conjugates, various other derivatives containing two or three *N*-acetyl cysteamine moieties were formed. Using these reaction conditions, we mimicked the biogenesis of the panphenazines from pantetheine and PCA. In contrast to the experiment with *N*-acetylcysteine, irradiation of the solution at 370 nm yielded exclusively two regioisomers. HPLC–HRMS comparison unequivocally proved that the products are

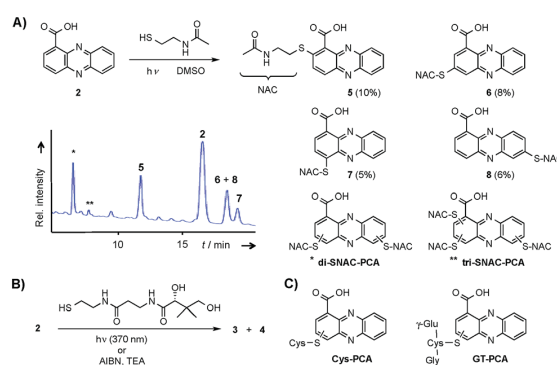


Fig. 2 Formation of PCA-thiol conjugates. (A) Photo-induced conjugate formation of PCA and *N*-acetylcysteine (NAC) at 370 nm (HPLC profile after 24 h); (B) biomimetic synthesis of 3 and 4; (C) general structure of cysteine–PCA and glutathione–PCA adducts resulting from irradiation.



identical with the conjugates **3** and **4** isolated from *Kitasatospora* sp. HKI 714 (Fig. S2†). Furthermore, the photoreaction proved to be a viable route for coupling cysteine and glutathione with PCA, as evidenced by HRMS and MSⁿ analyses (Fig. S3 and S4†).

The successful photoreactions and the substitution pattern are highly suggestive for the involvement of resonance-stabilised radicals in thiol conjugation. To interrogate the reaction mechanism, we conducted the coupling reaction of PCA and *N*-acetylcysteamine in the dark, but added azobisisobutyronitrile (AIBN) as a radical initiator. HPLC-HRMS analyses clearly indicated the formation of the corresponding adducts. Likewise, **3** and **4** were formed after incubation of PCA with pantetheine and AIBN (Fig. S5†).

Mechanistic considerations

The radical mechanism is supported by two other observations: first, addition of an electron source such as triethylamine (TEA) substantially increases product formation (up to tenfold), which is in agreement with the reported mechanism of the formation of the phenazine radical anion.³⁷ Second, addition of the radical scavenger *t*-butylhydroxytoluene (TBHT) hampered the photoreaction; only traces of conjugates could be detected. Thus, in addition to the observed substitution pattern of the thioconjugates, three lines of experimental evidence support the radical-based reaction mechanism (Fig. 3).

However, a direct light-induced formation of the thiyl radical can be ruled out. The homolytic cleavage of the SH bond occurs at wavelengths below 300 nm,³⁸ yet the PCA-thioconjugates are readily produced at 370 nm. Alternatively, it is more plausible that radical formation is initiated by photoexcitation of PCA into the triplet state. Transferring an electron from TEA (or an alternative electron source) leads to the phenazine radical anion (Fig. 3).³⁷ Subsequent protonation generates a neutral radical, which can react with biogenic thiols (GSH, pantetheine, cysteamine, cysteine) to form a thiyl radical while regenerating PCA. By analogy to the aromatic substitution reaction with thiyl

radicals,³⁹ one may infer that the thiyl radicals readily attack the electron-rich phenazine core. After H abstraction and re-aromatization, the radical chain reaction continues, which may also lead to multiple substitutions as observed for *N*-acetylcysteamine. On the contrary, conjugate formation with pantetheine seems to appear more selectively, possibly due to an increased steric demand of the thiol. Even though various conjugates could be detected in the HPLC-HRMS trace, only two isomers (**3** and **4**) appear in higher amounts. It should be highlighted that light-independent avenues to the formation of phenazine radicals have been reported in the biological context. It is well known that PYO is chemically reduced through radical intermediates in the cell, and phenazine redox-cycling (intracellular as well as extracellular) may generate PCA radicals,¹ which can react with thiols to form conjugates.

Phenazine protein adduct formation and crosslinking

Single-electron redox processes yield phenazine radicals in host cells; consequently, these reactive species may not only react with biogenic thiols, but also with SH residues of proteins, an important target of many small molecules.^{40,41} Such a process would be reminiscent of non-enzymatic reactions of sugars with proteins, yielding advanced glycation end-products (AGEs), which play important roles in aging and degenerative diseases.⁴² Considering the non-selective adduct formation of phenazines and the multitude of thiols in living cells, it is obvious that the corresponding conjugates would barely be detectable *in vivo* as individual entities. Thus we investigated whether PCA radicals form conjugates with individual proteins *in vitro*. For this, we synthesised a chemical probe (**9**) (Fig. S6†) in which the phenazine is fused to biotin *via* a tetraethylene glycol (TEG) linker (Fig. 4A).

First, we incubated **9** and AIBN with a cysteine-rich didomain (ketosynthase-branching, KS-B, 14 cysteines) of a modular polyketide synthase (PKS) module.^{43–45} The reaction mixture was loaded onto a streptavidin column, and the biotinylated protein was eluted at 60 °C in buffer with 3 mM biotin after multiple washing steps. Subsequent sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis indicated that the protein was indeed linked to the biotin-labelled probe, unequivocally proving that the phenazine was bound to the PKS component (Fig. 4B). To gain insight into the fate of proteins in the presence of the phenazine and UV light, we also prepared a fluorescent probe (**10**). Although PCA shows fluorescence itself, the respective signal proved to be too weak for the detection of protein conjugates. Therefore we attached PCA to rhodamine B, a fluorophore with an exceptionally high fluorescence quantum yield. The KS-B didomain was subjected to the rhodamine-based probe **10** with irradiation at 370 nm, and the reaction mixture was analysed by SDS-PAGE on a 12% gel. Surprisingly, the fluorescent bands were substantially larger than expected. Instead of the predicted band at 100 kDa, we observed a fluorescent signal that corresponds to a molecular weight of over 200 kDa (Fig. S7†). One possible explanation for this finding is that the probe might have cross-linked two or more equivalents of the

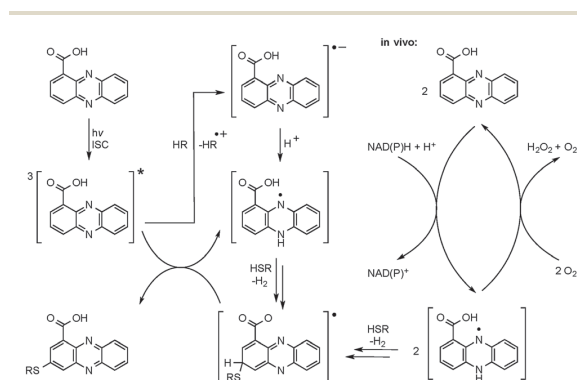


Fig. 3 Proposed reaction mechanism for the radical-induced conjugate addition. Radical formation can proceed *via* a photochemical reaction (left) or by redox cycling inside the cell (right). Only one of the possible intermediates (resonance form) is shown, leading to the derivative substituted at position 7. ISC = intersystem crossing; HR, electron donor (triethylamine, thiol or other); HSR, thiol.



Edge Article

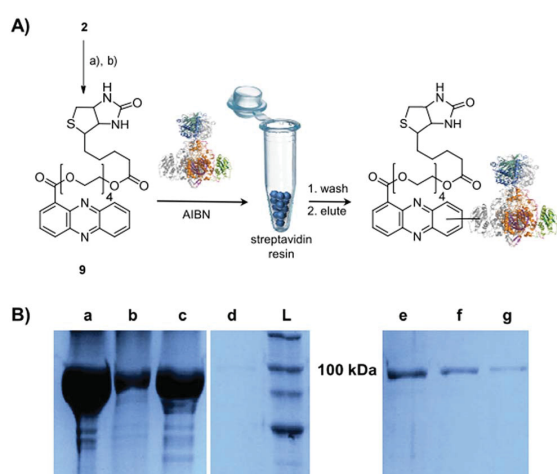


Fig. 4 Evidence for radical-mediated protein conjugation by PCA. (A) Synthesis of biotin-based probe **9**, thio-adduct formation with keto-synthase (KS-B) and purification with streptavidin resin; (a) tetraethylene glycol, EDC, DMAP, DCM, 86% (b) biotin, EDC, DMAP, DCM, 96%. (B) SDS-PAGE (10%) of KS-B after the incubation with the probe **9** for 24 h in the presence of AIBN. (a) Load, (b) flow-through, (c) first wash, (d) second wash, (e) first elution, (f) second elution, (g) third elution, (L) ladder.

protein. When analysing the products on a lower concentrated gel (10%) for better separation, we detected several protein aggregates of various sizes (Fig. 5B).

By tryptic digestion and MALDI analyses we corroborated that all protein complexes resulting from the incubation of KS-B with **9** and **10** are indeed derived from the KS (Fig. S8.1–S8.6†). In control experiments where the protein was irradiated in the presence of rhodamine B alone no aggregation could be observed (Fig. 5B). MS/MS analysis of the protein (lane a) revealed one of the peptides where **10** was bound to Cys3559 (Fig. 5C). A model was generated to show the proposed binding mode of **10** to this cysteine using Pymol (Fig. 5D). These findings unequivocally showed that the crosslinking activity correlates with the presence of the phenazine moiety and that multiple linkages are formed. To test the scope of the reaction we employed two proteins that are completely unrelated to a PKS, carbonic anhydrase III (5 cysteines) and albumin (35 cysteines). After incubation of carbonic anhydrase III with probe **10** we monitored the time course of its aggregation. Bands corresponding to the native protein slowly disappeared with concomitant formation of extensive protein aggregates bound to the fluorescent probe (Fig. S9†). Using the same conditions, albumin was functionalised with the rhodamine-based probe **10**, yet formation of higher aggregates was negligible (Fig. S10†). Nonetheless, the monoadduct was ideally suited to confirm the binding of the probe to a cysteine side chain by MALDI analysis (Fig. S11 and S12†). Taken together, all selected proteins are derivatised with the phenazine probe, yet they differ in their tendency to form higher aggregates, which may be controlled by steric factors.

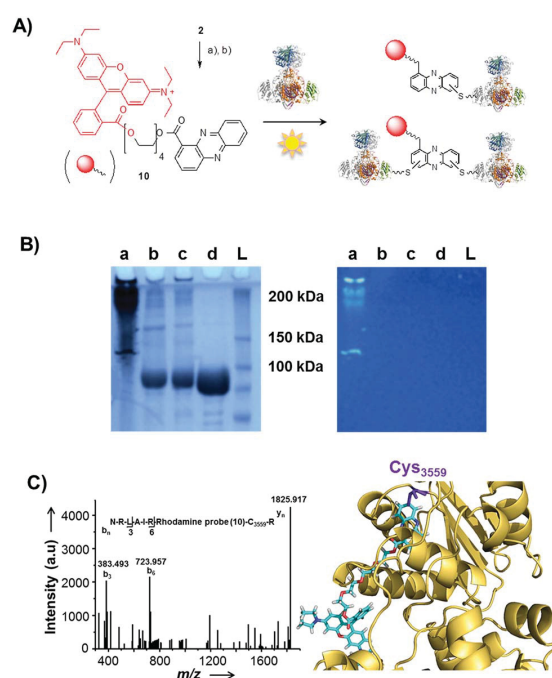


Fig. 5 Photo-induced protein crosslinking by PCA. (A) Synthesis of PCA-rhodamine B conjugate **10** and thio-conjugate formation with KS-B, (a) tetraethylene glycol, EDC, DMAP, DCM, 86%, (b) rhodamine B, PyBOP, Hünig's base, DCM, 63%. (B) SDS-PAGE (10%) of KS-B after the incubation with probe **10** or rhodamine B in water (left picture: under white light; right picture at 370 nm); (contrast adjusted) (a) **10**, 370 nm, 18 h, (b) rhodamine B, 370 nm, 18 h, (c) H₂O, 370 nm, 18 h, (d) KS-B control; (L) ladder. (C) MALDI-MS/MS spectrum of the rhodamine-based probe (**10**) linked to a peptide fragment of KS-B and model of the rhodamine-based probe (**10**) linked to Cys3559 of KS-B (PDB code 4KC5).

Conclusions

Considering the large number of studies on phenazines and their biological activities, it is surprising that their ability to capture biogenic thiols and crosslink proteins has so far been overlooked. The discovery of the panphenazines and the mechanistic investigations revealed that phenazine radicals form conjugates with a range of thiols. Thus, we provide a plausible model for irreversible GSH reduction in human endothelial cells.¹⁹ Furthermore, we report for the first time that phenazines form S-conjugates with proteins. Protein modification and aggregation mediated by small molecules are the focus of diverse research on biomaterials,^{46,47} and these processes are also relevant for protein aging and degenerative diseases.⁴² Our findings are significant because PCA is an important biosynthetic building block that is produced in large amounts by many pathogenic pseudomonads during infection processes.^{2,48} Since phenazine radicals are readily formed in cells by redox cycling, PCA and congeners may contribute to protein degeneration processes. Specifically, our findings may explain the broad activity profiles of phenazines including protein aggregation in



C. elegans.²³ Radical-based thiol conjugation and protein crosslinking by phenazines represent novel modes of action that are an important addition to the body of knowledge on this widespread group of virulence factors.⁶

Experimental

General analytical procedures

All NMR experiments were performed on Bruker AVANCE II 300, AVANCE III 500 or a 600 MHz spectrometer, equipped with a Bruker Cryo Platform. The chemical shifts are reported in parts per million (ppm) relative to the solvent residual peak of chloroform-*d*₁ (¹H: 7.24 ppm, singlet; ¹³C: 77.00 ppm, triplet), methanol-*d*₄ (¹H: 3.30 ppm, quintet; ¹³C: 49.00 ppm, septet) or DMSO-*D*₆ (¹H: 2.50 ppm, quintet; ¹³C: 39.50 ppm, septet). HR ESI-MS was carried out on an Accela UPLC-system (Thermo Scientific) combined with an Exactive mass spectrometer (Thermo Scientific) equipped with an electrospray ion source. Solid phase extraction was carried out using Chromabond C₁₈ ec Cartridges filled with 2000 mg of octadecyl-modified silica gel (Macherey-Nagel). HPLC-MS measurements were performed on a HPLC 1100 System connected to a 1100 Series LC/MSD Trap using an Zorbax Eclipse XDB-C8, 4.6 × 150 mm column, particle size 5 μm (Agilent Technologies). For thin layer chromatography TLC aluminium sheets silica gel 60 F₂₅₄ (MERCK) were used. Open column chromatography was performed using silica gel 60; particle size 0.015–0.04 mm (Macherey-Nagel) and on Sephadex® LH-20 (Sigma-Aldrich). Semi preparative HPLC was performed on an 1260 Infinity System (Agilent Technologies) using the following columns: Zorbax Eclipse XDB-C8, 9.5 × 250 mm, particle size 5 μm (Agilent Technologies), Zorbax Eclipse XDB-C18, 9.5 × 250 mm, particle size 5 μm (Agilent Technologies) and EC 250/4 Nucleodur Sphinx RP, 10 × 250 mm, particle size 5 μm (Macherey-Nagel). All solvents for analytical and preparative HPLC were obtained commercially in gradient grade and were filtered prior to use. To avoid microbial growth 0.1% formic acid was added to the water, used for analytical and preparative HPLC. UV-spectra were obtained using a UV-spectrometer UV-1800 (Shimadzu). IR-spectra were recorded on an FT/IR-4100 ATR spectrometer (Jasco). For irradiation experiments an UV-lamp TL MINI 6 Watt 33-640 (Philips) and an UV-lamp F 6W T5 G5 129 3000K 280 lm (Sylvania) were used.

Strain cultivation and isolation of panphenazines

The cultivation of *Kitasatospora* sp. HKI 714 and the fractionation of the crude extract were performed in exactly the same way as reported before.²⁷ Compounds **1** and **2** were isolated by semipreparative HPLC (Zorbax Eclipse XDB-C18, 9.5 × 250 mm, 5 μm, flow rate 4 mL min⁻¹, gradient starting from H₂O/MeOH 90/10, in 2.5 min to 50/50, in 15.5 min to 46/54, in 7 min to 42/58, in 5 min to 0/100).

Compound 3. Yellow solid, for a full list of NMR signals see Table S1† in the SI; IR (film): 3308, 2923, 2854, 1700, 1652, 1421, 1265, 1053, 759 cm⁻¹; UV/vis (MeOH): λ_{max} nm (log ε): 236 (6.22), 274 (6.36), 381 (5.87), 427 (5.78) nm; HR ESI-MS: *m/z* = 501.1807 [M + H]⁺ (calculated 501.1809 for C₂₄H₂₉O₆N₄S).

Compound 4. Yellow solid, for a full list of NMR signals see Table S1† in the SI; IR (film): 3297, 2929, 1651, 1541, 1410, 1045, 757 cm⁻¹; UV/vis (MeOH): λ_{max} nm (log ε): 261 (6.05), 372 (5.52), 408 (5.41) nm; HR ESI-MS: *m/z* = 501.1808 [M + H]⁺ (calculated 501.1809 for C₂₄H₂₉O₆N₄S).

General synthetic procedures

All reagents were obtained from commercial suppliers (Sigma Aldrich and Acros Organics) and were used without further purification. Carbonic anhydrase III (29.5 kDa) was obtained from GE Healthcare UK Limited and albumin (66 kDa) was obtained from Carl Roth GmbH. Chemical solvents were obtained at least in HPLC grade. Methanol and chloroform were distilled prior to use. Reactions were performed under inert atmosphere (Argon) using the Schlenk technique.

Pantetheine

A solution of *D*-pantetheine (200 mg, 0.36 mmol) and dithiothreitol (100 mg, 0.65 mmol) in a methanolic solution (4 mL; methanol : water; 1 : 1; v/v) was stirred for 2 days in an inert atmosphere (argon). The solvent was removed under reduced pressure and the product purified by column chromatography using silica gel (chloroform/methanol: 98/2: v/v) to give 195 mg of a colourless liquid. Yield 97%. ¹H NMR (300 MHz, MeOD): δ 0.9 (6H, s), 2.4 (2H, t), 2.6 (2H, t), 3.3–3.6 (6H, m), 3.9 (1H, s) ppm. ¹³C NMR (75.5 MHz, MeOD): δ 20.9, 21.3, 24.5, 36.4, 40.4, 43.9, 70.3, 77.2, 79.5, 173.9, 176.1 ppm.

Synthesis of panphenazines (3 and 4)

Phenazine-1-carboxylic acid (2, 3 mg, 0.013 mmol) was dissolved in DMSO (300 μL) and pantetheine (4 mg, 0.013 mmol) was added. The stirred mixture was incubated at 370 nm for 24 h. The solvent was removed under reduced pressure and the crude product dissolved in MeOH. HPLC-HRMS proved the identity of one product with compound **4**. Compound **3** was only produced in minor amounts under these conditions. The yield of **4** was quantified by integration of its peak area at a detector wavelength of 430 nm. Under these conditions yield was determined to be 2%. Repeating the reaction with the addition of TEA (1.9 μL, 0.014 mmol) increased the yield to 24%. In contrast, addition of TBHT instead of TEA hampered the reaction and no corresponding signal for **4** could be detected in the UV trace.

Synthesis of phenazine *N*-acetylcysteamine conjugates (5–8)

Phenazine-1-carboxylic acid (2, 10 mg, 0.045 mmol) and *N*-acetylcysteamine (5 mg, 0.042 mmol) were dissolved in DMSO and incubated at 370 nm for 24 h. The solvent was removed under reduced pressure and the residue was subjected to semipreparative HPLC (Zorbax Eclipse XDB-C8, 9.5 × 250 mm, 5 μm, flow rate 4 mL min⁻¹, gradient starting from H₂O/ACN 90/10, in 3 min to 80/20, in 25 min to 60/40, in 2 min to 0/100). Yields of monosubstituted phenazines: **5** (10%), **6** (8%), **7** (5%), **8** (6%) (yellow solids).

Compound 5. ¹H NMR (500 MHz, DMSO-*D*₆): δ 1.58 (3H, s), 3.15 (2H, t, ³*J* = 6.7 Hz), 3.22–3.27 (2H, m), 7.92–7.98 (2H, m),



8.03 (1H, d, $^3J = 9.2$ Hz), 8.12 (1H, d, $^3J = 9.2$ Hz), 8.16–8.21 (1H, m), 8.21–8.25 (1H, m) ppm. ^{13}C NMR (125 MHz, DMSO- D_6): δ 22.3, 32.7, 38.8, 128.2, 129.3, 131.0, 131.3, 132.4, 139.8, 141.3, 142.3, 142.5, 163.1[#], 168.2, 169.3 ppm.

Compound 6. ^1H NMR (600 MHz, DMSO- D_6): δ 1.82 (3H, s), 3.28 (2H, t, $^3J = 6.8$ Hz), 3.36–3.44 (2H, m)*, 7.58 (1H, m), 7.87–7.94 (3H, m), 8.16 (1H, dd, $^3J = 8.4$ Hz, $^4J = 1.4$ Hz), 8.19 (1H, d, $^3J = 8.4$ Hz), 8.23 (1H, t, $^3J_1 = 5.5$ Hz, NH) ppm. ^{13}C NMR (150.9 MHz, DMSO- D_6): δ 22.6, 30.4, 37.7, 120.7, 127.0, 128.9, 129.4, 130.2, 131.2, 139.2, 141.2, 141.3, 142.6, 143.2, 164.3, 168.5, 169.6 ppm.

Compound 7. ^1H NMR (600 MHz, DMSO- D_6): δ 1.83 (3H, s), 3.22 (2H, t, $^3J = 6.9$ Hz), 3.40–3.44 (2H, m), 7.75 (1H, d, $^3J = 7.2$ Hz), 7.86 (1H, d, $^3J = 7.2$ Hz), 7.95–8.01 (2H, m), 8.20–8.29 (3H, m) ppm. ^{13}C NMR (150.9 MHz, DMSO- D_6): δ 22.6, 29.6, 37.5, 124.0, 128.6, 129.2, 129.6, 131.2, 131.4, 133.6, 140.1, 140.6, 140.7, 142.1, 143.8, 165.4, 169.5 ppm.

Compound 8. ^1H NMR (600 MHz, DMSO- D_6): δ 1.82 (3H, s), 3.30 (2H, t, $^3J = 6.9$ Hz), 3.39–3.43 (2H, m), 7.73 (1H, d, $^3J = 6.6$ Hz), 7.77 (1H, dd, $^3J = 9.2$ Hz, $^4J = 1.5$ Hz), 7.80 (1H, dd, $^3J_1 = 8.6$ Hz, $^3J_2 = 6.6$ Hz), 7.99 (1H, s), 8.04 (1H, d, $^3J = 8.6$ Hz), 8.09 (1H, d, $^3J = 9.2$ Hz), 8.28–8.33 (1H, m, NH) ppm. ^{13}C NMR (150.9 MHz, DMSO- D_6): δ 22.5, 30.5, 37.6, 121.9, 125.4, 127.0, 127.5, 129.3, 130.3, 130.9, 140.4, 141.1, 142.0, 142.5, 165.5, 169.6, 169.8 ppm. *Signal partly obscured, [#]signal only detectable from HMBC data.

Synthesis of phenazine-biotin and -rhodamine probes

Tetraethyleneglycol phenazine-1-carboxylate (11)

Phenazine-1-carboxylic acid **2** (10 mg, 0.044 mmol) was dissolved in dichloromethane (5 mL) and cooled to 0 °C. Tetraethylene glycol (26 mg, 0.13 mmol, 3 eq.), DMAP (5 mg, 0.04 mmol) and EDC (8.4 mg, 0.05 mmol) were added. After 24 h, water was added and the aqueous phase extracted with dichloromethane (2 × 10 mL). The combined organic fractions were dried over sodium sulfate and concentrated under reduced pressure. Column chromatography over silica gel (chloroform/methanol: 98/2, v/v) gave the product (15.3 mg) as yellow oil. Yield: 86%. ^1H NMR (600 MHz, MeOD): δ 3.47 (2H, t, $^3J = 4.9$ Hz), 3.54 (2H, m), 3.58 (2H, t, $^3J = 4.9$ Hz), 3.61 (2H, m), 3.66 (2H, m), 3.73 (2H, m), 3.93 (2H, t, $^3J = 4.7$ Hz), 4.65 (2H, t, $^3J = 4.7$ Hz), 7.96–7.99 (3H, m), 8.24–8.27 (1H, m), 8.31 (1H, dd, $^3J = 6.8$ Hz, $^4J = 1.3$ Hz), 8.33 (1H, m), 8.40 (1H, dd, $^3J = 8.8$ Hz, $^4J = 1.3$ Hz) ppm. ^{13}C NMR (150.9 MHz, MeOD): δ 62.2, 66.0, 70.2, 71.4, 71.6, 71.7, 73.6, 79.5, 130.3, 130.9, 131.0, 132.7, 132.8, 132.9, 133.4, 134.0, 142.0, 143.8, 144.5, 144.9, 168.1 ppm; IR (film): 1119, 1214, 1237, 1274, 1350, 1421, 1523, 1727, 2874, 3018 cm^{-1} , HRMS: (ESI⁺): m/z calculated for $\text{C}_{21}\text{H}_{25}\text{O}_6\text{N}_2$ [M + H]⁺ 401.1707, found: 401.1710.

Biotin-labelled phenazine-1-carboxylic acid (9)

Tetraethylene-glycol phenazine-1-carboxylate (7 mg, 0.017 mmol) **11** and biotin (8.5 mg, 0.035 mmol, 2 eq.) were dissolved in dry dichloromethane (2 mL) and cooled to 0 °C. DMAP (2 mg, 0.017 mmol) and EDC (8 mg, 0.044 mmol, 2.5 eq.) were added. Water (5 mL) was added after 20 h and the aqueous phase

extracted with dichloromethane (2 × 5 mL). The combined organic fractions were dried over sodium sulfate and concentrated under reduced pressure. Column chromatography over silica gel (chloroform/methanol: 95/5, v/v) gave 10.5 mg of a yellow solid. Yield: 96%. ^1H NMR (500 MHz, CDCl_3): δ 1.56–1.70 (6H, m), 2.28–2.32 (2H, m), 2.67–2.73 (1H, m), 2.81–2.88 (1H, m), 3.06–3.13 (1H, m), 3.57–3.65 (10H, m), 3.72 (2H, dd, $^3J_1 = 5.4$ Hz, $^3J_2 = 4.1$ Hz), 3.91 (2H, t, $^3J_1 = 4.8$ Hz), 4.14–4.18 (2H, m), 4.23–4.28 (1H, m), 4.43–4.48 (1H, m), 4.65 (2H, t, $^3J_1 = 4.8$ Hz), 7.82–7.86 (3H, m), 8.20–8.23 (1H, m), 8.23 (1H, dd, $^3J_1 = 7.1$ Hz, $^4J_2 = 1.3$ Hz), 8.27–8.30 (1H, m), 8.37 (1H, dd, $^3J_1 = 8.8$ Hz, $^4J_2 = 1.3$ Hz) ppm. ^{13}C NMR (125 MHz, MeOD): δ ppm. ^{13}C NMR (150.9 MHz, CDCl_3): δ 24.6, 28.1, 28.2, 33.7, 40.4, 55.4, 60.1, 61.9, 63.3, 64.6, 69.0, 69.1, 70.4, 70.5, 70.6, 70.6, 129.1, 129.3, 130.3, 131.0, 131.2, 131.3, 132.1, 133.2, 141.0, 142.5, 143.1, 143.6, 166.4, 173.5, 174.1 ppm; IR (film): 747, 1038, 1119, 1179, 1236, 1264, 1455, 1522, 1700, 2867, 2925, 3236 cm^{-1} , HRMS: (ESI⁺): m/z calculated for $\text{C}_{31}\text{H}_{39}\text{O}_8\text{N}_4\text{S}$ [M + H]⁺ 627.2483, found: 627.2477.

Rhodamine-labelled phenazine-1-carboxylic acid (10)

To a solution of rhodamine B (1.9 mg, 0.0043 mmol), tetraethyleneglycol phenazine-1-carboxylate **11** (1.7 mg, 0.0043 mmol) and PyBop (2.7 mg, 0.0052 mmol, 1.2 eq.) in anhydrous dichloromethane (1 mL) was added Hünigs base (2.9 μL , 4 eq.) at 0 °C. The reaction was allowed to warm to room temperature and stirred for 24 h. The solution was concentrated under reduced pressure and the residue was subjected to column chromatography over silica gel (gradient chloroform/MeOH: 100/0–95/5, v/v). The crude product was subjected to semi-preparative HPLC using a Zorbax Eclipse XDB-C8 (acetonitrile- H_2O , 0.1% formic acid; starting conditions: 20% acetonitrile, 40% in 4 min, 60% in 12.5 min, 100% in 0.5 min, 100% for 5 min, 20% in 1 min) to give 2.2 mg of a white solid. Yield: 63%. ^1H NMR (CDCl_3 , 600 MHz) δ = 1.26–1.37 (12H, m), 3.45–3.49 (3H, m), 3.50–3.59 (9H, m), 3.61–3.64 (2H, m), 3.64–3.68 (1H, m), 3.69–3.72 (2H, m), 3.73–3.79 (1H, m), 3.90 (2H, t, $^3J = 4.8$ Hz), 4.12 (2H, t, $^3J = 4.7$ Hz), 4.63 (2H, t, $^3J = 4.8$ Hz), 6.78–6.81 (3H, m), 7.02 (1H, s), 7.04 (1H, s), 7.28 (1H, dd, $^3J_1 = 7.6$ Hz, $^4J_2 = 0.7$ Hz), 7.61–7.82 (3H, m), 7.85–7.91 (3H, m), 8.25 (1H, dd, $^3J_1 = 6.9$ Hz, $^4J_2 = 1.1$ Hz), 8.26 (1H, dd, $^3J_1 = 8.0$ Hz, $^4J_2 = 1.0$ Hz), 8.28–8.32 (2H, m), 8.28 (1H, d, $^3J_1 = 8.8$ Hz) ppm. ^{13}C NMR (CDCl_3 , 600 MHz) δ = 12.6, 12.6, 12.6, 13.7, 46.0, 46.0, 46.0, 47.4, 64.7, 64.8, 68.6, 69.2, 70.4, 70.5, 70.6, 70.7, 96.4, 113.6, 114.0, 119.0, 128.7, 129.0, 129.6, 129.8, 130.2, 130.3, 131.2, 131.3, 131.3, 131.4, 131.4, 131.7, 132.3, 132.8, 132.9, 133.0, 133.6, 141.1, 141.9, 142.5, 143.8, 153.8, 155.5, 157.8, 158.8, 165.0, 166.3 ppm. IR (film): 1077, 1126, 1183, 1342, 1410, 1462, 1590, 1651, 1722, 2353, 2860, 2923, 3318 cm^{-1} , HRMS: (ESI⁺): m/z calculated for $\text{C}_{49}\text{H}_{53}\text{O}_8\text{N}_4$ [M]⁺ 825.3858, found: 825.3868.

Phenazine-protein binding assays

The biotinylated probe **9** was directly mixed with the protein (KS-B – 104.7 kDa) in 20 mM Tris-HCl (pH 8.0). Final concentrations of the probe and protein were 0.5 mM and 20 μM respectively. The reaction mixture was incubated for 24 h under



the following conditions: (a) at 254 nm (b) at 370 nm or (c) at 22 °C in the presence of 0.5 mM azobisisobutyronitrile (AIBN). For purification of the biotinylated KS-B, the reaction mixture was diluted twice with 50 mM Tris, 0.2 M NaCl, 0.1% SDS (pH 8.0) and incubated with 1 mL of streptavidin beads (Thermo Scientific) at 22 °C for 1 h with shaking at 400 rpm. The beads were washed twice with 50 mM Tris, 0.2 M NaCl, 2% SDS (pH 8.0) after incubation at 60 °C for 10 min. Elution of the bound protein was achieved by heating the sample thrice at 60 °C for 30 min in 50 mM Tris, 0.2 M NaCl, 2% SDS, 3 mM D-biotin (pH 8.0).

The rhodamine-based probe **10** was incubated with the protein (KS-B, carbonic anhydrase III, albumin) in water or in DMSO. Final concentrations of the probe and the protein were 10 μM and 20 μM respectively. The reaction mixture was incubated for varying times between 15 min and 20 h under UV-light at 254 nm, at 370 nm, at 22 °C in the presence of 0.5 mM AIBN in the dark or at sunlight (specific times and conditions for each experiment are indicated in the respective figure legend). After the assay the protein samples were analysed by PAGE, on a 10% or 12% SDS-gel for KS-B. The structures of KS-B and albumin were visualised using Pymol. The probe **10** was generated as a fragment and added specifically to the Cys residue of the proteins by replacing the hydrogen atom in Pymol.

Acknowledgements

We thank H. Heinecke and A. Perner for NMR and HRMS measurements, M. Poetsch for MALDI experiments, K. Martin and K. Perlet for strain handling, and K.-D. Menzel and M. Steinacker for support in large-scale fermentation and downstream processing. We gratefully acknowledge the Studienstiftung des Deutschen Volkes (to D. H.) and the International Leibniz Research School for Microbial and Biomolecular Interactions (ILRS, to S. S.) for financial support.

References

- S. Chincholkar and L. Thomashow, *Microbial Phenazines*, Springer, Berlin, 2013.
- J. B. Laursen and J. Nielsen, *Chem. Rev.*, 2004, **104**, 1663–1686.
- A. Cimmino, A. Evidente, V. Mathieu, A. Andolfi, F. Lefranc, A. Kornienko and R. Kiss, *Nat. Prod. Rep.*, 2012, **29**, 487–501.
- J. Davies and K. S. Ryan, *ACS Chem. Biol.*, 2012, **7**, 252–259.
- A. Price-Whelan, L. E. P. Dietrich and D. K. Newman, *Nat. Chem. Biol.*, 2006, **2**, 71–78.
- L. S. I. Pierson and E. A. Pierson, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 1659–1670.
- D. Romero, M. F. Traxler, D. Lopez and R. Kolter, *Chem. Rev.*, 2011, **111**, 5492–5505.
- D. V. Mavrodi, J. A. Parejko, O. V. Mavrodi, Y. S. Kwak, D. M. Weller, W. Blankenfeldt and L. S. Thomashow, *Environ. Microbiol.*, 2013, **15**, 1462–2920.
- L. E. Dietrich, A. Price-Whelan, A. Petersen, M. Whiteley and D. K. Newman, *Mol. Microbiol.*, 2006, **61**, 1308–1321.
- L. E. Dietrich, C. Okegbe, A. Price-Whelan, H. Sakhtah, R. C. Hunter and D. K. Newman, *J. Bacteriol.*, 2013, **195**, 1371–1380.
- L. E. Dietrich, T. K. Teal, A. Price-Whelan and D. K. Newman, *Science*, 2008, **321**, 1203–1206.
- A. J. Stams, F. A. de Bok, C. M. Plugge, M. H. van Eekert, J. Dolfin and G. Schraa, *Environ. Microbiol.*, 2006, **8**, 371–382.
- C. C. Caldwell, Y. Chen, H. S. Goetzmann, Y. Hao, M. T. Borchers, D. J. Hassett, L. R. Young, D. Mavrodi, L. Thomashow and G. W. Lau, *Am. J. Pathol.*, 2009, **175**, 2473–2488.
- M. Mentel, E. G. Ahuja, D. V. Mavrodi, R. Breinbauer, L. S. Thomashow and W. Blankenfeldt, *ChemBioChem*, 2009, **10**, 2295–2304.
- R. C. Hunter, V. Klepac-Ceraj, M. M. Lorenzi, H. Grotzinger, T. R. Martin and D. K. Newman, *Am. J. Respir. Cell Mol. Biol.*, 2012, **47**, 738–745.
- H. Ran, D. J. Hassett and G. W. Lau, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 14315–14320.
- K. Apel and H. Hirt, *Annu. Rev. Plant Biol.*, 2004, **55**, 373–399.
- Y. Q. O'Malley, K. J. Reszka, D. R. Spitz, G. M. Denning and B. E. Britigan, *Am. J. Physiol.: Lung Cell. Mol. Physiol.*, 2004, **287**, L94–L103.
- M. Muller, *Free Radical Biol. Med.*, 2002, **33**, 1527–1533.
- M. Muller, *Free Radical Biol. Med.*, 2011, **50**, 971–977.
- R. Cheluvappa, R. Shimmon, M. Dawson, S. N. Hilmer and D. G. Le Couteur, *Acta Biochim. Pol.*, 2008, **55**, 571–580.
- M. Muller and N. D. Merrett, *Chem. Biol. Interact.*, 2015, **232**, 30–37.
- A. Ray, C. Rentas, G. A. Caldwell and K. A. Caldwell, *Neurosci. Lett.*, 2015, **584**, 23–27.
- A. J. McFarland, S. Anoopkumar-Dukie, A. V. Perkins, A. K. Davey and G. D. Grant, *Arch. Toxicol.*, 2012, **86**, 275–284.
- B. Rada and T. L. Leto, *Trends Microbiol.*, 2013, **21**, 73–81.
- J. M. Winter, S. Behnken and C. Hertweck, *Curr. Opin. Chem. Biol.*, 2011, **15**, 22–31.
- D. Heine, K. Martin and C. Hertweck, *J. Nat. Prod.*, 2014, **77**, 1083–1087.
- M. Wagner, W. M. Abdel-Mageed, R. Ebel, A. T. Bull, M. Goodfellow, H.-P. Fiedler and M. Jaspars, *J. Nat. Prod.*, 2014, **77**, 416–420.
- M. L. Gilpin, M. Fulston, D. Payne, R. Cramp and I. Hood, *J. Antibiot.*, 1995, **48**, 1081–1085.
- A. Braunshausen and F. P. Seebeck, *J. Am. Chem. Soc.*, 2011, **133**, 1757–1759.
- D. H. Scharf, P. Chankhamjon, K. Scherlach, T. Heinekamp, K. Willing, A. A. Brakhage and C. Hertweck, *Angew. Chem., Int. Ed.*, 2013, **52**, 11092–11095.
- D. H. Scharf, N. Remme, A. Habel, P. Chankhamjon, K. Scherlach, T. Heinekamp, P. Hortschansky, A. A. Brakhage and C. Hertweck, *J. Am. Chem. Soc.*, 2011, **133**, 12322–12325.
- S. Coyne, C. Chizzali, M. N. Khalil, A. Litomska, K. Richter, L. Beerhues and C. Hertweck, *Angew. Chem., Int. Ed.*, 2013, **52**, 10564–10568.



[View Article Online](#)

Edge Article

Chemical Science

- 34 E. Sasaki, X. Zhang, H. G. Sun, M. Y. Lu, T. L. Liu, A. Ou, J. Y. Li, Y. H. Chen, S. E. Ealick and H. W. Liu, *Nature*, 2014, **510**, 427–431.
- 35 K. V. Goncharenko, A. Vit, W. Blankenfeldt and F. P. Seebeck, *Angew. Chem., Int. Ed.*, 2015, **54**, 2821–2824.
- 36 B. Li, J. P. Yu, J. S. Brunzelle, G. N. Moll, W. A. van der Donk and S. K. Nair, *Science*, 2006, **311**, 1464–1467.
- 37 T. Ogata, Y. Yamamoto, Y. Wada, K. Murakoshi, M. Kusaba, N. Nakashima, A. Ishida, S. Takamuku and S. Yanagida, *J. Phys. Chem.*, 1995, **99**, 11916–11922.
- 38 G. A. Amaral, F. Ausfelder, J. G. Izquierdo, L. Rubio-Lago and L. Banares, *J. Chem. Phys.*, 2007, **126**, 024301.
- 39 S. Naumov and C. Schoneich, *J. Phys. Chem. A*, 2009, **113**, 3560–3565.
- 40 M. Gersch, J. Kreuzer and S. A. Sieber, *Nat. Prod. Rep.*, 2012, **29**, 659–682.
- 41 T. Böttcher, M. Pitscheider and S. A. Sieber, *Angew. Chem., Int. Ed.*, 2010, **49**, 2680–2698.
- 42 P. Ulrich and A. Cerami, *Recent Prog. Horm. Res.*, 2001, **56**, 1–22.
- 43 T. Bretschneider, J. B. Heim, D. Heine, R. Winkler, B. Busch, B. Kusebauch, T. Stehle, G. Zocher and C. Hertweck, *Nature*, 2013, **502**, 124–128.
- 44 D. Heine, T. Bretschneider, S. Sundaram and C. Hertweck, *Angew. Chem., Int. Ed.*, 2014, **53**, 11645–11649.
- 45 D. Heine, S. Sundaram, T. Bretschneider and C. Hertweck, *Chem. Commun.*, 2015, **51**, 9872–9875.
- 46 S. Adolph, V. Jung, J. Rattke and G. Pohnert, *Angew. Chem., Int. Ed.*, 2005, **44**, 2806–2808.
- 47 N. Reddy, R. Reddy and Q. Jiang, *Trends Biotechnol.*, 2015, **33**, 362–369.
- 48 G. V. Bloembergen and B. J. J. Lugtenberg, *New Phytol.*, 2003, **157**, 503–523.

Open Access Article. Published on 14 April 2016. Downloaded on 24/01/2017 13:51:30.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.



Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2016

Supplemental Material for

**A Widespread Bacterial Phenazine
Forms Conjugates with Biogenic Thiols and Crosslinks Proteins**

Daniel Heine, Srividhya Sundaram, Matthias Beudert, Karin Martin and Christian Hertweck

Analytical data of compounds 3 and 4	2
Binding of phenazine to biogenic thiols.....	3
Synthesis of probes.....	5
Phenazine-protein binding assays with KS-B.....	5
Tryptic mass fingerprinting of labelled KS-B.....	6
Phenazine-protein binding assays with carbonic anhydrase III and albumin.....	10
Tryptic mass fingerprinting of labelled albumin.....	11
Antimicrobial properties of isolated phenazines 3 and 4	13
NMR spectra.....	14

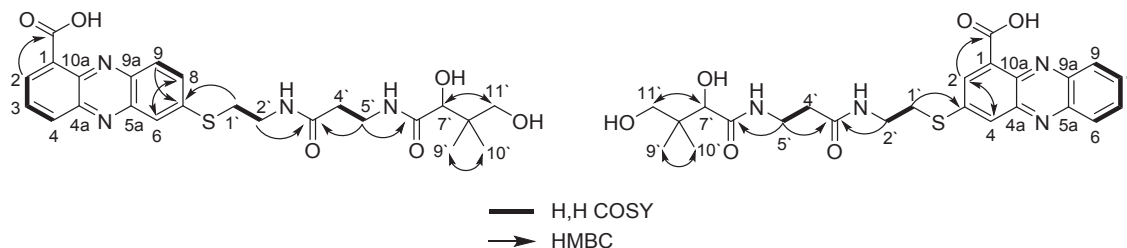
Analytical data of compounds **3** and **4**

Figure S1: Selected ^1H , ^{13}C HMBC and ^1H , ^1H COSY correlations of panphenazines **3** and **4**.

Table S1: NMR spectroscopic data for Compounds **3** and **4** in methanol- d_4 .

	3		4	
Position	$\delta_{\text{H}}^{\text{a}}$ (J in Hz)	$\delta_{\text{C}}^{\text{b}}$	$\delta_{\text{H}}^{\text{a}}$ (J in Hz)	$\delta_{\text{C}}^{\text{b}}$
1	-	- ^c	-	- ^c
2	8.62, d (7.0)	135.5	8.32, m	134.7
3	8.03, dd (8.7, 7.0)	131.7	-	143.3
4	8.40, dd (8.7, 1.3)	134.8	8.17, d (2.2)	126.6
4a	-	144.6	-	144.8
5a	-	145.3	-	145.1
6	8.05, d (2.1)	123.2	8.22, dd (7.6, 2.2)	130.2
7	-	145.2	7.96, m	133.0
8	7.87, dd (9.1, 2.1)	134.3	7.96, m	132.8
9	8.18, d (9.1)	129.4	8.28, dd (7.6, 2.2)	129.9
9a	-	141.0	-	142.0
10a	-	140.8	-	140.4
1'	3.38, m	32.0	3.39, m	32.3
2'	3.59, td (6.90, 1.5)	39.3	3.59, t (6.8)	39.2
3'	-	174.1	-	174.1
4'	2.43, t (6.7)	36.5	2.42, t (6.6)	36.5
5'	3.48, m	36.4	3.49, m	36.3
6'	-	176.1	-	176.1
7'	3.87, s	77.3	3.87, s	77.3
8'	-	40.4	-	40.4
9'/10'	0.89, s	21.3/20.9	0.89, s	21.3/20.9
11'	3.39, m	70.4	3.40, m	70.4
COOH	-	169.7	-	170.2

^a recorded at 600.3 MHz, ^b recorded at 150.9 MHz; ^c signals not detectable.

Binding of phenazine to biogenic thiols

Conjugate formation of PCA with pantetheine following irradiation at 370 nm

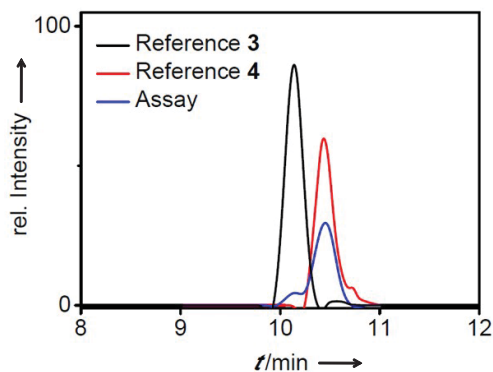


Figure S2. SIM-HRMS ($m/z = 501.1788\text{--}501.1828$) analysis of the obtained product after incubation of PCA and pantetheine with irradiation at 370 nm; comparison with the isolated reference compounds **3** and **4**. SIM: selected ion monitoring.

Coupling of PCA to cysteine and glutathione

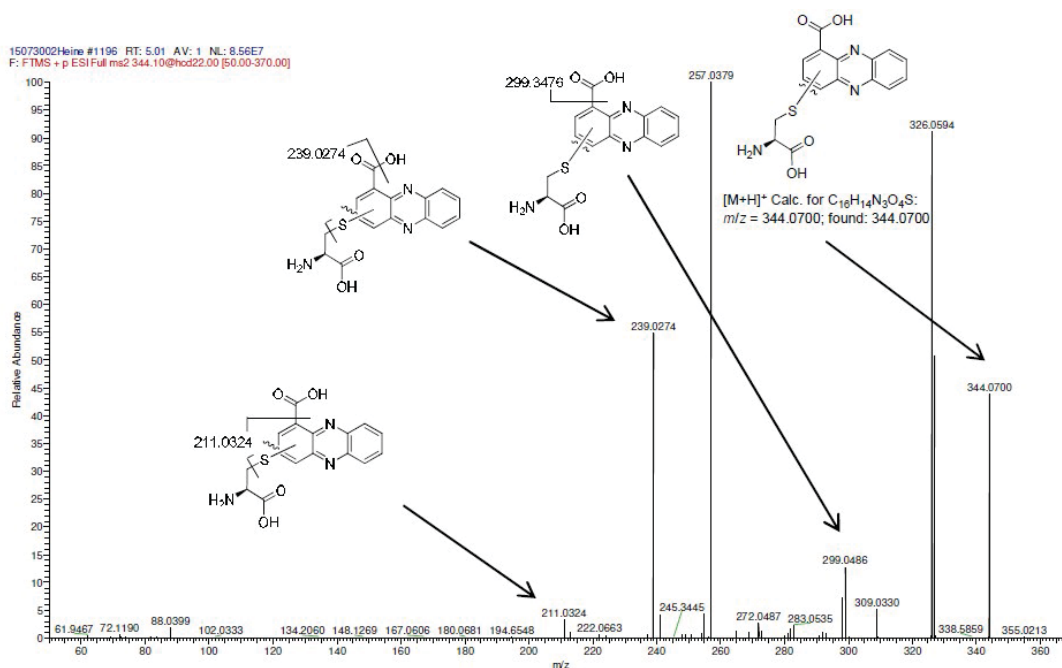


Figure S3. HRMS(/MS) spectrum of PCA-cysteine adduct ($[M+H]^+$) after incubation of PCA with cysteine under sunlight.

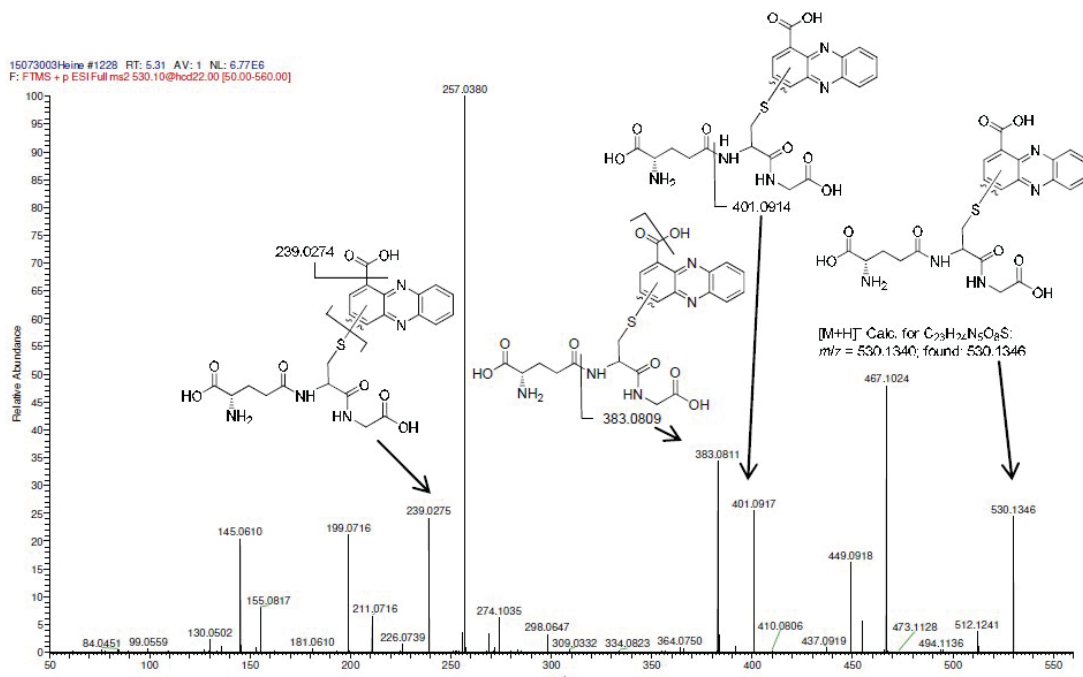


Figure S4. HRMS(/MS) spectrum of PCA-glutathione-adduct ($[M+H]^+$) after incubation of PCA with glutathione under sunlight.

Conjugate formation of PCA with pantetheine in presence of AIBN

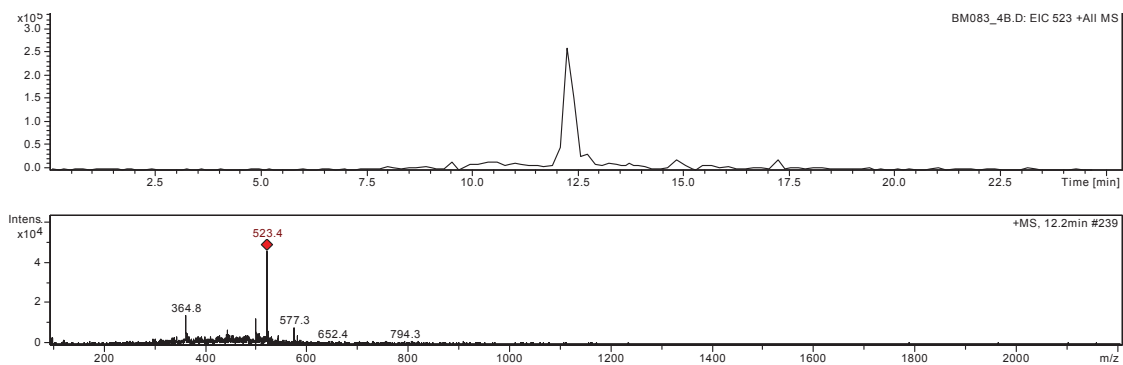


Figure S5. SIM-MS spectrum of the product after incubation of PCA with pantetheine in presence of AIBN in the dark (top). MS spectrum of PCA-pantetheine-adduct ($[M+Na]^+$) (bottom).

Synthesis of phenazine probes **9** (with biotin) and **10** (with rhodamine).

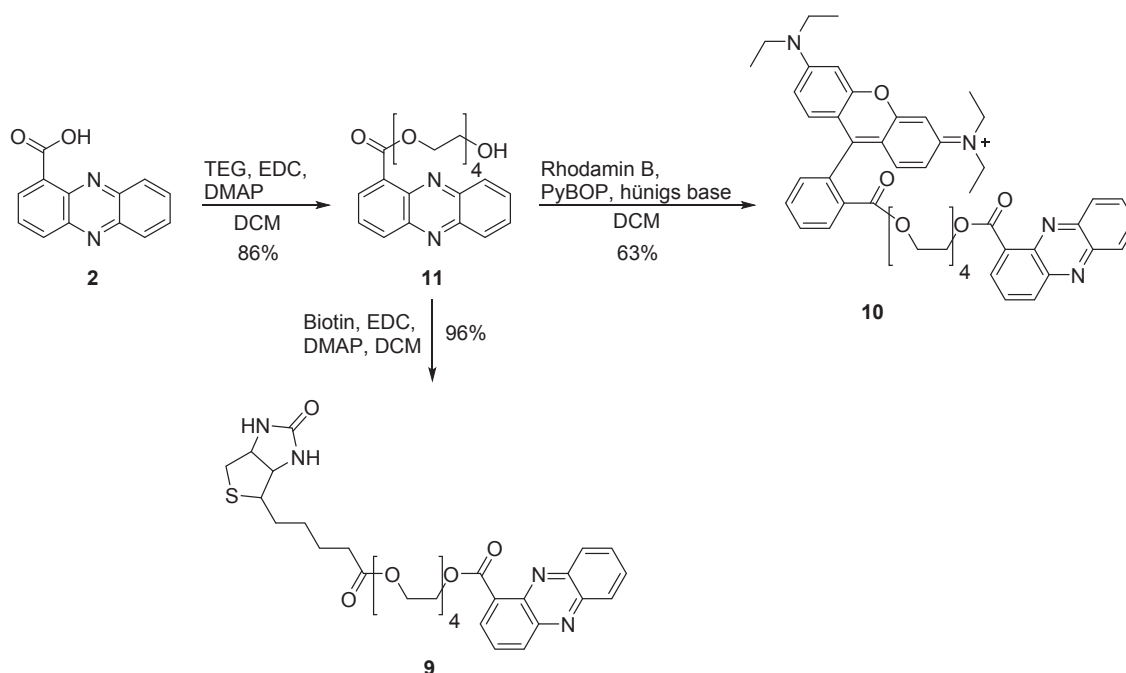


Figure S6: Synthesis of biotin- and rhodamine-B-tagged phenazine probes.

Phenazine-protein binding assays with KS-B

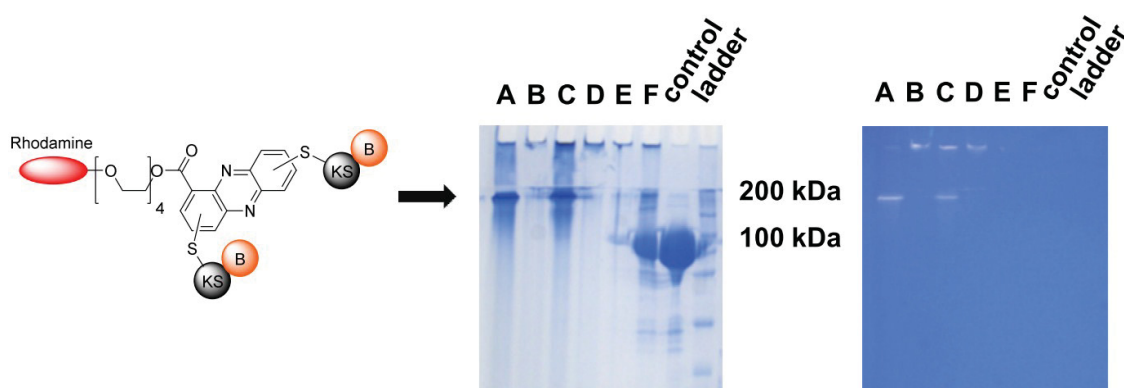
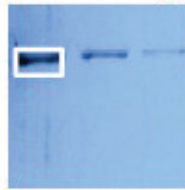


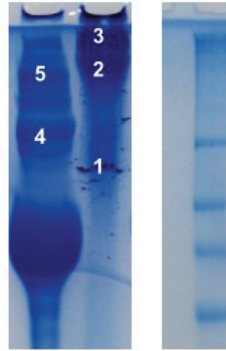
Figure S7. SDS-PAGE (12% gel) after incubation of KS-B with rhodamine-B-labelled phenazine probe **10** under the following conditions: A) $\lambda = 370$ nm, 20 h, in H₂O, B) $\lambda = 370$ nm, 20 h, DMSO, C) sunlight, 8 h, H₂O, D) D: sunlight, 8 h, DMSO, E) E: AIBN, 20 h, DMSO, F) F: AIBN, 20 h, H₂O.

Tryptic mass fingerprinting of labelled KS-B



Protein:		RhlE KsBb					
Intensity coverage:	54.2% (250997 cnts)	Sequence coverage MS:	42.5%				
		Sequence coverage MS/MS:	5.5%				
		pl:	5.9				
10	20	30	40	50	60	70	80
MKHHHHHHHH	GGLVPRGSHG	GSSGERVEDN	ELANYIAVIG	LGGYYPGADS	IDELWQNLAN	GVDCMSDFPA	DRWDHSKIYY
90	100	110	120	130	140	150	160
KNRKVLGKTT	CINGSFIDV	DKFDYSYFKM	PKVYADHMSP	EVRLFQAV	HTFEDAGYSK	ETLLSRVNGD	VGVLGTMSN
170	180	190	200	210	220	230	240
DYHYGFESN	VFRGSMASGS	GMATIPMTVS	YFYGLTGPSL	FIDTMCSSSS	TCIHTACQML	KHDETKMVL	GGLNLMYHPY
250	260	270	280	290	300	310	320
TTVNTSQGNF	TSITSESVNS	YVGADGTVI	GEGIGAVLLK	RLDRAIADRD	QIYGVIKGSA	MTNAGERNGF	NVPNPDQLTL
330	340	350	360	370	380	390	400
AIRQANDQAK	VHPSSISYIE	GHGSGTKLGD	PIEVLGLNNA	FRWATDDKQF	CYLGSIKSN	GHLAASGIA	GLTKTLLOFK
410	420	430	440	450	460	470	480
HKQIAPSIHS	SQLNQDIDFA	DTFFVVPQQL	IEWRQPERII	NGRKQVFPFR	AGLTSIAAGG	MNAHMIVEEY	PEPADSAGQI
490	500	510	520	530	540	550	560
SEDQLVFFVS	VHKLALLAQN	LTSFRDWLAS	SEAPLAQIAY	TLQVGKNNLR	NRLAIRCRT	QALSRALNAC	IDGHYQSSAD
570	580	590	600	610	620	630	640
SKIFRYFQES	DAVQPLESDL	NDPLAPLLTQ	ULNGDSQVDW	ASLYAQPVR	ISLPAYRFK	TRCWYTEEGY	ESSIVNPLMF
650	660	670	680	690	700	710	720
KNKLHPLVAK	NCSTPOPGAI	FRTDFVEDEL	LDYVYSGRGG	RRLSAFNFAD	VALAMPALAS	RFDGRTLSVS	CAFHEYIADW
730	740	750	760	770	780	790	800
TTVTGLETRL	FEIDSEQLEL	EFDFFRSGEQ	PTHLGFAVIN	PLTSDEPLP	QQWLDAREL	LMRQALQAGR	QLSAAEVSQR
810	820	830	840	850	860	870	880
LAQAGYDFAP	YLDHDGELTI	GRSGLVLKGR	PPVNRHNYA	DNVQLSPYLA	TTIDKALYLL	LDELGLPQGR	VIVRNIERLC
890	900	910	920	930	940	950	
CYHTPAGGFS	VVLSGIGLND	NELSLSLVL	DEREQICVKL	DKVSLYLKQ	EVASVDRKHS	LLTGT	

Figure S8.1. Tryptic mass fingerprinting of KS-B after incubation with **9** and AIBN followed by affinity chromatography.



1

Protein:		Rhf KsRh					
Intensity coverage:	49.7% (2283686 cnts)	Sequence coverage MS:	16.3%				
		Sequence coverage MS/MS:	83%				
		pI:	5.9				
1U	2U	3U	4U	5U	6U	7U	8U
MKHHHHHHH	GGLVPRGSHG	GSSGERVEDN	ELANYIAVIG	LGGYYPGADG	IDELCQNLAD	GVDCMSDFPA	DRWDHSKIYY
90	100	110	120	130	140	150	160
KNRKVLGKTT	CING ^S FIKDV	DKFD ^S YSYFKM	PKVYADHMSP	EVRLFLOQAV	ETFE ^S AG ^S SE	ETLLSRVNGD	VGVLGTMSN
170	180	190	200	210	220	230	240
DYHYGFESN	VFRGSMASGS	GMATIPMTVS	YFYGLTGPSL	FIDTMCSSSS	TCIED ^S CC ^S DL	SDDETRM ^S VLA	GGLNLMYHPY
250	260	270	280	290	300	310	320
TTVNTS ^S QGN ^S F	TSITSES ^S VNS	YGVGADGTVI	GEGIGAVLLK	RLDRAIADRE	QIYGV ^S EG ^S SA	SDNAGERN ^S GF	NVNPDLQTL
330	340	350	360	370	380	390	400
AI ^S RQAMDQAK	VHPSSISYIE	GHGSGTKLGD	PIEVLGLNNA	FRWATDDKQF	CHLSE ^S ES ^S CI	GHL ^S LR ^S AG ^S GIA	GLTKTLLOPK
410	420	430	440	450	460	470	480
HKQIAPSIHS	SQLNQDIDFA	DTPFVVPQQL	IEWRQPERII	NGRKQVFPRE	AGL ^S TR ^S IA ^S GG	SDN ^S AM ^S IV ^S E ^S EY	PEPADSAGQI
490	500	510	520	530	540	550	560
SEDQLVVFVS	VHKLALLAQN	LTSFRD ^S WLAS	SEAPLAQIAY	TLQVGKNNLE	NRL ^S AE ^S CR ^S TR	QALS ^S RAL ^S NAC	IDGHYQSSAD
570	580	590	600	610	620	630	640
SKIFYRFQES	DAVQPLESDL	NDPLAPLLTQ	WLNQDSQVDW	ASLYAQQPVF	ISL ^S FA ^S SE ^S FER	TRCQVTE ^S EGY	ESSIVNPLMF
650	660	670	680	690	700	710	720
KNKLHPLVAK	NCS ^S TPQPGAI	FRTDFVEDEL	LDYVYSGRGG	RRLSAFN ^S FAD	V ^S AL ^S AE ^S FL ^S AS	RFDGRTLSVS	CAFEHYIADW
730	740	750	760	770	780	790	800
TTVTGLE ^S YRL	FEIDSEOLEL	EFDFRRSGEQ	PTHLGF ^S AVIN	PLTSD ^S EPPLF	QQK ^S LD ^S AE ^S REL	LNR ^S CAL ^S Q ^S GR	QLSAAEVSQR
810	820	830	840	850	860	870	880
LAQACVDFAP	YLDHDC ^S EIT ^S T	GRSGIVL ^S RGR	DPVNR ^S R ^S NHYA	DNVQL ^S SPVIA	TTT ^S CF ^S AI ^S RL	LD ^S DF ^S LD ^S Q ^S GR	VTVR ^S NT ^S ERL ^S C
890	900	910	920	930	940	950	
CYHTPACGFS	VVLSGICLND	NELSL ^S LLV ^S L	DERE ^S IC ^S VKL	DKVSL ^S YL ^S LCK	EVA ^S EV ^S CR ^S BS	LLTGT	

Figure S8.2. Tryptic mass fingerprinting of KS-B after crosslinking with **10** under UV light (1).

2 Protein: RHE KsBb

Intensity coverage: 54.9% (1173396 ions) Sequence coverage MS: 14.5% Sequence coverage MS/MS: 6.6% pI: 5.9

10	20	30	40	50	60	70	80
MKHHHHHHHH	GGLVPRGSHG	GSSGERVEDN	ELANYIAVIG	LGYYYPGADS	IDELWQNLAN	GVDCMSDFPA	DRWDHSKIYY
90	100	110	120	130	140	150	160
KNRKVLGKTT	CINGSFIRKDV	DKFDYGYFKM	PKVYADHMSP	EVRLFLQVAV	HTFEDAGYSK	ETLLSRVNGD	VGULLGTMSN
170	180	190	200	210	220	230	240
DYHYHGFESN	VFRGSMASGS	GMATIPMTVS	YFYGLTGPPL	FIDTMCSSSS	TCIHTACQML	KHDETKMVL	GGLNLNMYHPY
250	260	270	280	290	300	310	320
TTVNTSOGNF	TSITSESVNS	YGVGADGIVI	GEGIGAVLLK	RLDRAIADRD	QIYGVIKGSA	MTNAGERNGF	NVPNPDLOTL
330	340	350	360	370	380	390	400
AIRQAMDQAK	VHPSSISYIE	GHGSGTKLGD	PIEVLGLNNA	FRUATDDKQF	CYLGSIKSN	GHLAASGIA	GLTKTLLOFK
410	420	430	440	450	460	470	480
HKQIAPSIHS	SQLNQDIDFA	DTPFVVPQQL	IEWRQPERII	NGRKQVFFRR	AGLTSIAAGG	MNAHMIVEEY	PEPADSAGQI
490	500	510	520	530	540	550	560
SEDQLVVFVS	VHKLALLAQN	LTSFRDWLAS	SEAPLAQIAY	TLQVGNLNR	NRLAIRCRR	QALSRALNAC	IDGHYQSSAD
570	580	590	600	610	620	630	640
SKIFRYFQES	DAVQPLESDL	NDPLAPLLTQ	WLNQDSQVDW	ASLYAQPPVR	ISLPAYRFEK	TRCWYTEEGY	ESSIVNPLMF
650	660	670	680	690	700	710	720
KNKLHPLVAK	NCSTPQGA	FRTDFVEDEL	LDYVYSRGG	RRLSAFNFAD	VALAMPALAS	RFDGRRLSVS	CAFEHYIADW
730	740	750	760	770	780	790	800
TTVTGLEAYL	FEIDSEOLEL	EFDFRSAGEQ	PTHLGFAVIN	PLTSDPEPLP	QOQLDDAREL	LNROALOAGR	OLSAAEVSOR
810	820	830	840	850	860	870	880
LAQAGYDFAP	YLDHDCGELTI	GRSGLVLKGR	PPVNRHMYA	DNVQLSPYLA	TTIDKALYLL	LDELGLPQGR	VIVNRIERLC
890	900	910	920	930	940	950	
CYHTPAGGFS	VVLSGIGLND	NELSLSLVL	DEREQICVKL	DKVSLYLGRQ	EVASVDRKHS	LLTGT	

Figure S8.3. Tryptic mass fingerprinting of KS-B after crosslinking with **10** under UV light (1).

3 Protein: RHE KsBb

Intensity coverage: 47.8% (976679 ions) Sequence coverage MS: 14.4% Sequence coverage MS/MS: 6.6% pI: 5.9

10	20	30	40	50	60	70	80
MKHHHHHHHH	GGLVPRGSHG	GSSGERVEDN	ELANYIAVIG	LGYYYPGADS	IDELWQNLAN	GVDCMSDFPA	DRWDHSKIYY
90	100	110	120	130	140	150	160
KNRKVLGKTT	CINGSFIRKDV	DKFDYGYFKM	PKVYADHMSP	EVRLFLQVAV	HTFEDAGYSK	ETLLSRVNGD	VGULLGTMSN
170	180	190	200	210	220	230	240
DYHYHGFESN	VFRGSMASGS	GMATIPMTVS	YFYGLTGPPL	FIDTMCSSSS	TCIHTACQML	KHDETKMVL	GGLNLNMYHPY
250	260	270	280	290	300	310	320
TTVNTSOGNF	TSITSESVNS	YGVGADGIVI	GEGIGAVLLK	RLDRAIADRD	QIYGVIKGSA	MTNAGERNGF	NVPNPDLOTL
330	340	350	360	370	380	390	400
AIRQAMDQAK	VHPSSISYIE	GHGSGTKLGD	PIEVLGLNNA	FRUATDDKQF	CYLGSIKSN	GHLAASGIA	GLTKTLLOFK
410	420	430	440	450	460	470	480
HKQIAPSIHS	SQLNQDIDFA	DTPFVVPQQL	IEWRQPERII	NGRKQVFFRR	AGLTSIAAGG	MNAHMIVEEY	PEPADSAGQI
490	500	510	520	530	540	550	560
SEDQLVVFVS	VHKLALLAQN	LTSFRDWLAS	SEAPLAQIAY	TLQVGNLNR	NRLAIRCRR	QALSRALNAC	IDGHYQSSAD
570	580	590	600	610	620	630	640
SKIFRYFQES	DAVQPLESDL	NDPLAPLLTQ	WLNQDSQVDW	ASLYAQPPVR	ISLPAYRFEK	TRCWYTEEGY	ESSIVNPLMF
650	660	670	680	690	700	710	720
KNKLHPLVAK	NCSTPQGA	FRTDFVEDEL	LDYVYSRGG	RRLSAFNFAD	VALAMPALAS	RFDGRRLSVS	CAFEHYIADW
730	740	750	760	770	780	790	800
TTVTGLEAYL	FEIDSEOLEL	EFDFRSAGEQ	PTHLGFAVIN	PLTSDPEPLP	QOQLDDAREL	LNROALOAGR	OLSAAEVSOR
810	820	830	840	850	860	870	880
LAQAGYDFAP	YLDHDCGELTI	GRSGLVLKGR	PPVNRHMYA	DNVQLSPYLA	TTIDKALYLL	LDELGLPQGR	VIVNRIERLC
890	900	910	920	930	940	950	
CYHTPAGGFS	VVLSGIGLND	NELSLSLVL	DEREQICVKL	DKVSLYLGRQ	EVASVDRKHS	LLTGT	

Figure S8.4. Tryptic mass fingerprinting of KS-B after crosslinking with **10** under UV light (1).

4

Protein:		PihC KsDb	
Intensity coverage:	47.8 % (1543905 cnts)	Sequence coverage MS:	35.7 %
		Sequence coverage MS/MS:	8.3 %
		pI:	5.9
10	20	30	40
50	100	110	120
130	140	150	160
170	180	190	200
210	220	230	240
250	260	270	280
290	300	310	320
330	340	350	360
370	380	390	400
410	420	430	440
450	460	470	480
490	500	510	520
530	540	550	560
570	580	590	600
610	620	630	640
650	660	670	680
690	700	710	720
730	740	750	760
770	780	790	800
810	820	830	840
850	860	870	880
890	900	910	920
930	940	950	

Protein: PihC KsDb
Intensity coverage: 47.8 % (1543905 cnts) Sequence coverage MS: 35.7 % Sequence coverage MS/MS: 8.3 % pI: 5.9

MRHHHHHHHHH GGLVPRGSHG GSSGERVEDN ELANYIAVIG LGGYYPGADS IDELWQNLAN GVDCHSDFFPA DRWDHSKIYY
KNRKVLGKTT CINGSFIDKDV DKFDYSYFKM PKVYADHNSP EVRLFLQVAV HTFEDAGYSK ETLLSRVNGD VGVLLGTMEN
DYHYGYFESN VFRGSMASGS GMATIPMTVS YFYGLTGPSL FIDTMCSSSS TCIHTACQML KHDETKMLA GGLNLVYHPY
TTVNTSQCNF TSITSESVNS YCVGADCTVI CEGICAVLLK RLDRAIADRD QIYCVIKCSA MTNACERNCF NVNPNDLQTL
AIRQANDQAK VHPSSISYIE GHGSGTKLGD PIEVLGLNNA FRWATDDKQF CYLGSIKSNI GHLLAASGIA GLTKILLQFK
HRQIAPSIHS SQLNQDIDFA DTPFVVQQQL IEWRQPERII NGRKQVPPRR AGLTSAIAGG MNAHMIVEEY PEPADSAGOI
SEDQLVVFVS VHKLALLAQN LTFSRDULAS SEAPLAQIAY TLQVGNLNR NRLAIRCRT RALSRALNAC IDGHYQSSAD
SKIFYRFQES DAVQPLESDL NDPLAPLLTQ WLNCDSQVDM ASLYAQPPVR ISLPAYRFEK TRCWYTEECY ESSIVNPLMF
KNKLHPLVAK NCS TPQPGAI FRDFVEDEL LDVYVSGRGG RRLSAFNFAD VALAMPALAS RFDGRTLSVS CAFEHYIADW
TTVTGLEVRL FEIDGQLEL EFDPRSGEQ PTHLGFAVIN PLTSDPEPLP QQQLDDAREL LNRQALQAGR QLSAAEVQSR
LAQAGYDFAP YLDHIDGELTI GRGGLVLRGR PPNRHNHYA DNVQLSPYLA TTIDKALYLL LDELGLPQGR VIVRNIERLC
CYHTPAGGFS VVLSGIGLND NELSLSLVL DEREQICVKL DKVSLYLGKQ EVASVDRKHS LLTGT

Figure S8.5. Tryptic mass fingerprinting of KS-B after crosslinking with 10 under UV light (1).

5

Protein:		RihE KsDb	
Intensity coverage:	40.7 % (1732434 cnts)	Sequence coverage MS:	35.9 %
		Sequence coverage MS/MS:	6.6 %
		pI:	5.9
10	20	30	40
50	100	110	120
130	140	150	160
170	180	190	200
210	220	230	240
250	260	270	280
290	300	310	320
330	340	350	360
370	380	390	400
410	420	430	440
450	460	470	480
490	500	510	520
530	540	550	560
570	580	590	600
610	620	630	640
650	660	670	680
690	700	710	720
730	740	750	760
770	780	790	800
810	820	830	840
850	860	870	880
890	900	910	920
930	940	950	

Protein: RihE KsDb
Intensity coverage: 40.7 % (1732434 cnts) Sequence coverage MS: 35.9 % Sequence coverage MS/MS: 6.6 % pI: 5.9

MRHHHHHHHHH GGLVPRGSHG GSSGERVEDN ELANYIAVIG LGGYYPGADS IDELWQNLAN GVDCHSDFFPA DRWDHSKIYY
KNRKVLGKTT CINGSFIDKDV DKFDYSYFKM PKVYADHNSP EVRLFLQVAV HTFEDAGYSK ETLLSRVNGD VGVLLGTMEN
DYHYGYFESN VFRGSMASGS GMATIPMTVS YFYGLTGPSL FIDTMCSSSS TCIHTACQML KHDETKMLA GGLNLVYHPY
TTVNTSQCNF TSITSESVNS YCVGADCTVI CEGICAVLLK RLDRAIADRD QIYCVIKCSA MTNACERNCF NVNPNDLQTL
AIRQANDQAK VHPSSISYIE GHGSGTKLGD PIEVLGLNNA FRWATDDKQF CYLGSIKSNI GHLLAASGIA GLTKILLQFK
HRQIAPSIHS SQLNQDIDFA DTPFVVQQQL IEWRQPERII NGRKQVPPRR AGLTSAIAGG MNAHMIVEEY PEPADSAGOI
SEDQLVVFVS VHKLALLAQN LTFSRDULAS SEAPLAQIAY TLQVGNLNR NRLAIRCRT RALSRALNAC IDGHYQSSAD
SKIFYRFQES DAVQPLESDL NDPLAPLLTQ WLNCDSQVDM ASLYAQPPVR ISLPAYRFEK TRCWYTEECY ESSIVNPLMF
KNKLHPLVAK NCS TPQPGAI FRDFVEDEL LDVYVSGRGG RRLSAFNFAD VALAMPALAS RFDGRTLSVS CAFEHYIADW
TTVTGLEVRL FEIDGQLEL EFDPRSGEQ PTHLGFAVIN PLTSDPEPLP QQQLDDAREL LNRQALQAGR QLSAAEVQSR
LAQAGYDFAP YLDHIDGELTI GRGGLVLRGR PPNRHNHYA DNVQLSPYLA TTIDKALYLL LDELGLPQGR VIVRNIERLC
CYHTPAGGFS VVLSGIGLND NELSLSLVL DEREQICVKL DKVSLYLGKQ EVASVDRKHS LLTGT

Figure S8.6. Tryptic mass fingerprinting of KS-B after crosslinking with 10 under UV light (1).

Phenazine-protein binding assays with carbonic anhydrase III and albumin

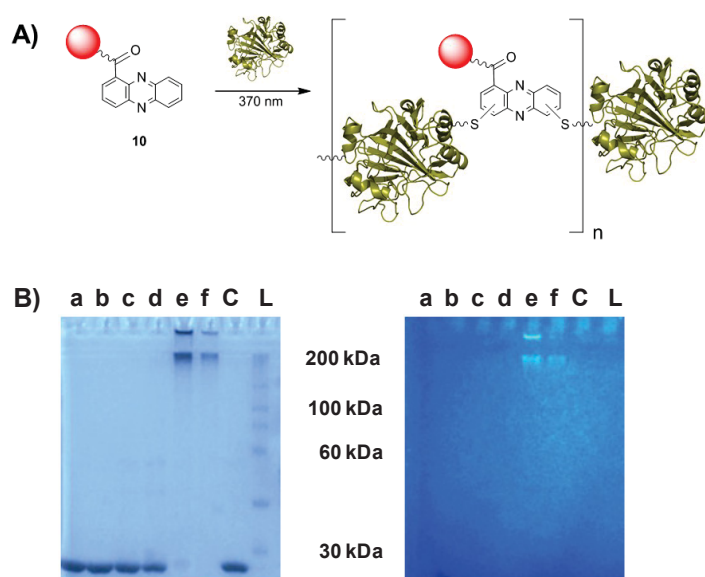


Figure S9. A) Scheme for the functionalization of carbonic anhydrase III by probe **10** after irradiation with UV-light. B) SDS-PAGE (10%) of carbonic anhydrase III after the incubation with probe **10** or rhodamine B in water; a) rhodamine B, 370 nm, 10 h; b) **10**, 370 nm, 15 min; c) **10**, 370 nm, 30 min; d) **10**, 370 nm, 1 h; e) **10**, 370 nm, 5 h; f) **10**, 370 nm, 10 h, C) carbonic anhydrase III control, L: Ladder. The fluorescent bands would correspond to potential conjugates of carbonic anhydrase III and **9** with $n > 5$.

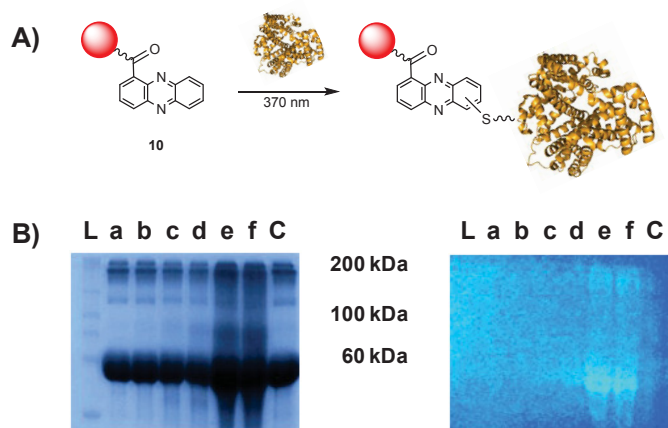
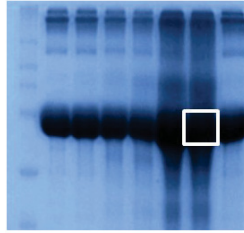


Figure S10. SDS-PAGE (10%) of albumin after the incubation with probe **10** or rhodamine B in water (left picture: under white light; right picture at 370 nm); L, Ladder, a) rhodamine B, 370 nm, 10 h; b) **10**, 370 nm, 15 min; c) **10**, 370 nm, 30 min; d) **10**, 370 nm, 1 h; e) **10**,

370 nm, 5 h; f) **10**, 370 nm, 10 h, C) albumin control. The fluorescent band potentially corresponds to a 70 kDa monoadduct of **10** and albumin.

Tryptic mass fingerprinting of labelled albumin



Protein: BSA							
Intensity coverage: 51.2% (1746939 cnts)	Sequence coverage MS: 52.2%						
Sequence coverage MS/MS: 5.1%	pI: 5.8						
10	20	30	40	50	60	70	80
MKWVTFISLL	LLFSSAYSRG	VFRRDTHRSE	IAHRFKDLGE	EHFKGLVLIA	FSQYLQOCFF	DEHVKLVNEL	TEFAKTCVAD
90	100	110	120	130	140	150	160
ESHAGCEKSL	HTLFGDELCK	VASLRETYGD	MADCCEKQEP	ERNECFLSHK	DDSPDLPLK	PDPNTLCDEF	KADEKFWGK
170	180	190	200	210	220	230	240
YLYEIARRHP	YFYAPELLEY	ANKYNGVFQE	CCQAEDKGAC	LLPKIETMRE	KVLASSARQR	LRCASIQKFG	ERALKAWSVA
250	260	270	280	290	300	310	320
RLSQKFPKAE	FVEVTKLVTD	LTKVHKECCH	GDLLECADDR	ADLAKYICDN	QDTISSKLE	CCDKPLEKS	HCIAEVEKDA
330	340	350	360	370	380	390	400
IPENLPPLTA	DFAEDKDVCK	NYQEAQDAFL	GSFLYEYSRR	HPEYAVSVLL	RLAKEYEATL	EECCAKDDPH	ACYSTVFDKL
410	420	430	440	450	460	470	480
KHLVDEPQNL	IKQNCQDFEK	LGEYGFQNAL	IVRYTRKVPQ	VSTPTLVEVS	RSLGKVGTRC	CTRPESERMP	CTEDYLSLIL
490	500	510	520	530	540	550	560
NRLCVLHEKT	PVSEKVTKCC	TESLVNRRPC	FSALTPDETY	VPKAFDEKLF	TFHADICTLP	DTEKQIKKQT	ALVELLKHKP
570	580	590	600	610			
KATEEQKTV	MENFVAFVDK	CCAADDKEAC	FAVEGPKLVV	STQTALA			

Figure S11. Tryptic mass fingerprinting of albumin after functionalization with **10** under UV light (1).

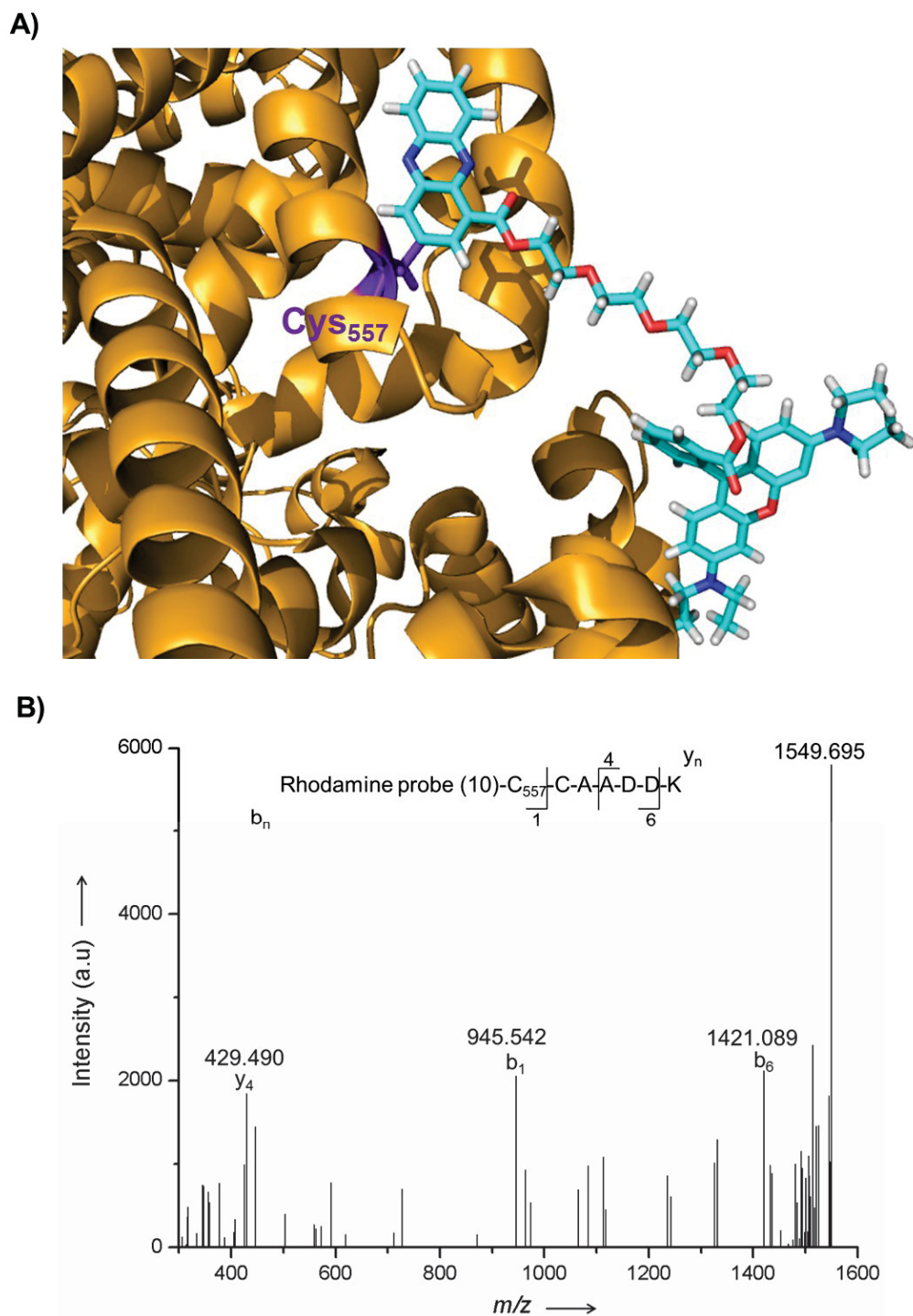


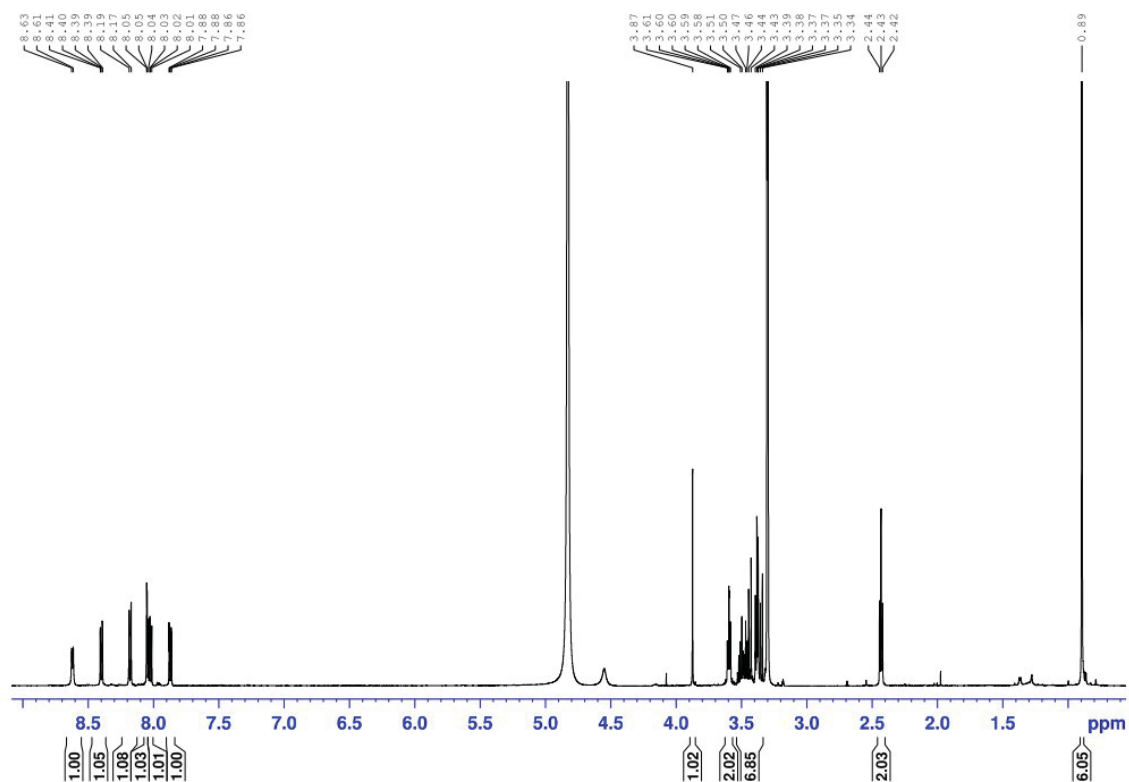
Figure S12. A) Model of the rhodamine-based probe (**10**) crosslinked to the Cys₅₅₇ of albumin under UV light (PDB code 4F5S) B) MALDI-MS/MS spectrum of the rhodamine-based probe (**10**)-linked peptide of albumin.

Antimicrobial properties of isolated phenazines **3** and **4**

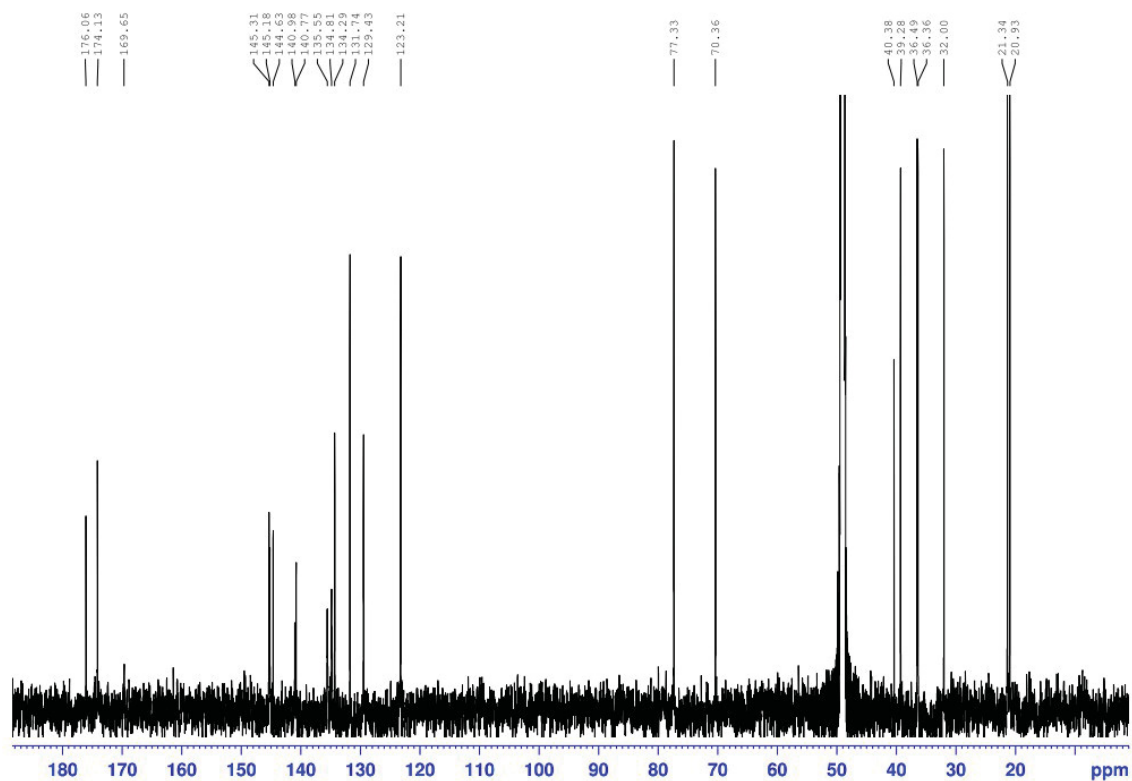
Cpd.	Inhibition zone (mm)									
	<i>B. sub</i>	<i>S. aur</i>	<i>E. col</i>	<i>P. aer</i>	MRSA	VRE	<i>M. vac</i>	<i>S. sal</i>	<i>C. alb</i>	<i>P. not</i>
3	0	0	17	0	0	15	13	20	0	0
4	10	0	0	15	0	0	13	0	0	0
PCA (2)	31	25	32	0	20	0	27	20	35	16

Table S2. Antimicrobial properties of the phenazine-pantetheine conjugates **3** and **4** in comparison with PCA (**2**). *B. sub*: *Bacillus subtilis* ATCC 6633; *S. aur*: *Staphylococcus aureus* SG 511; *E. col*: *Escherichia coli* SG 458; *P. aer*: *Pseudomonas aeruginosa* SG 137; MRSA: Methicillin- and quinolone-resistant *Staphylococcus aureus* 134/93; VRE: Vancomycin-resistant *Enterococcus faecalis* 1528; *M. vac*: *Mycobacterium vaccae* IMET 10670; *S. sal*: *Sporobolomyces salmonicolor* SBUG 549; *C. alb*: *Candida albicans* BSMY 212; *P. not*: *Penicillium notatum* JP36; cpd, compound; PCA: phenazine-1-carboxylic acid (**2**).

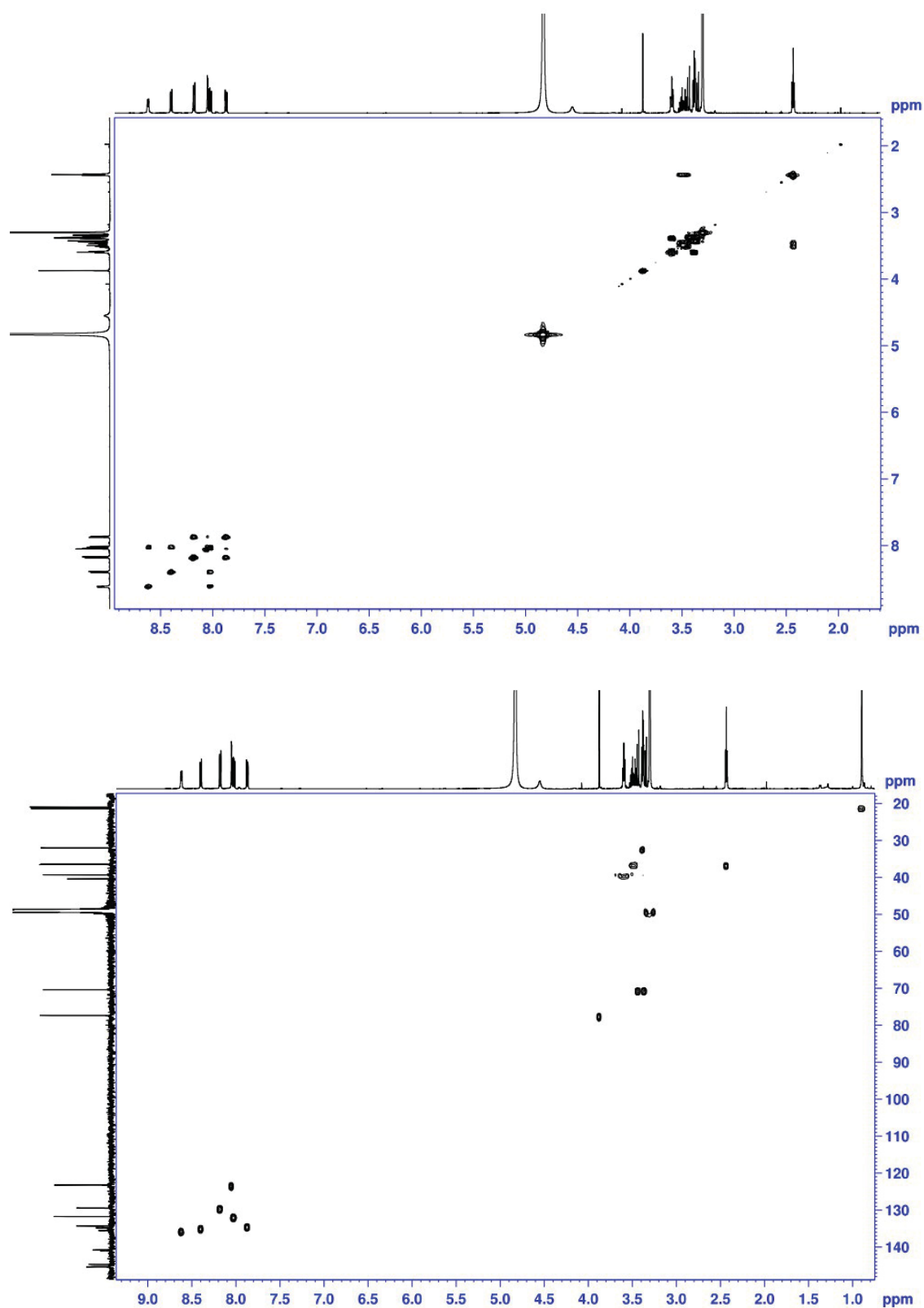
NMR Spectra



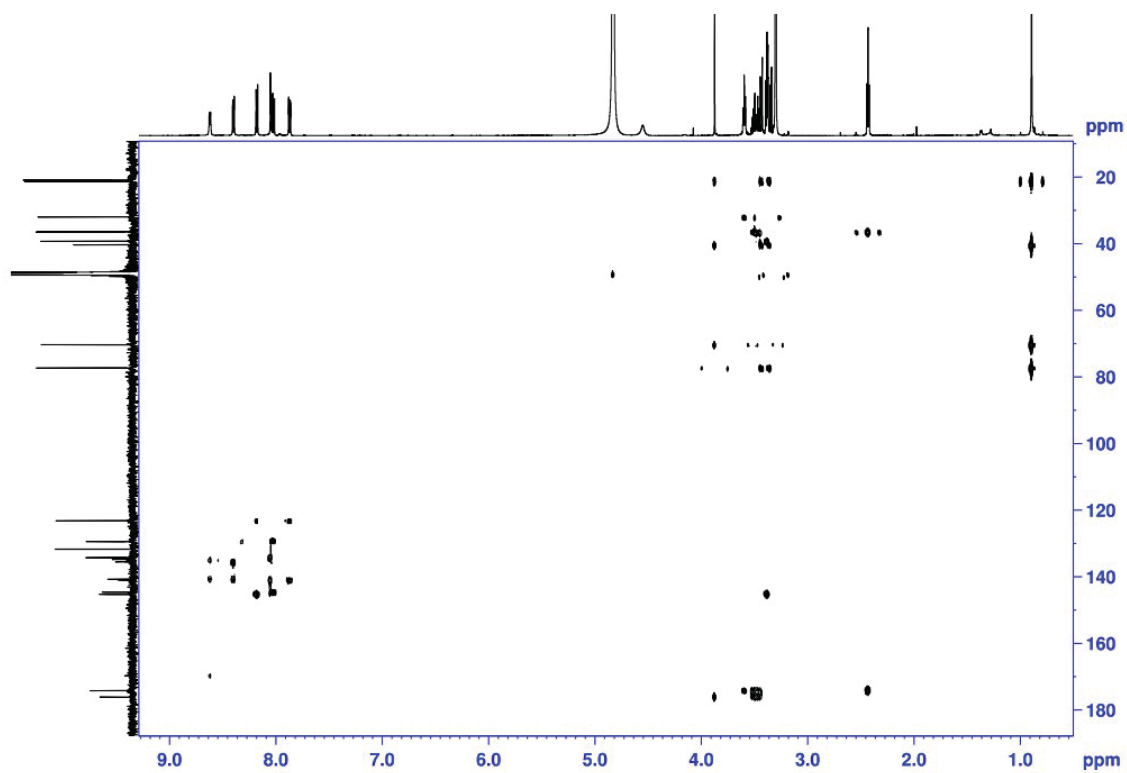
¹H NMR spectrum (600 MHz, MeOD) of **3**.



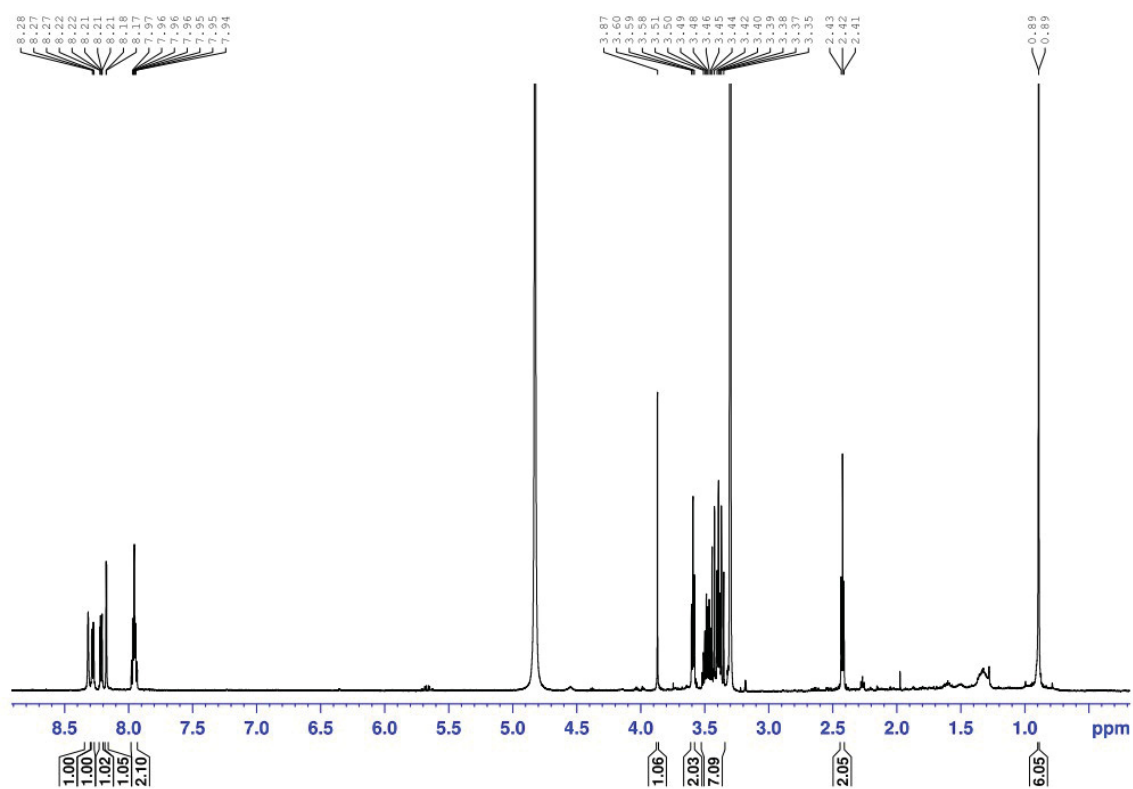
¹³C NMR spectrum (150 MHz, MeOD) of **3**.



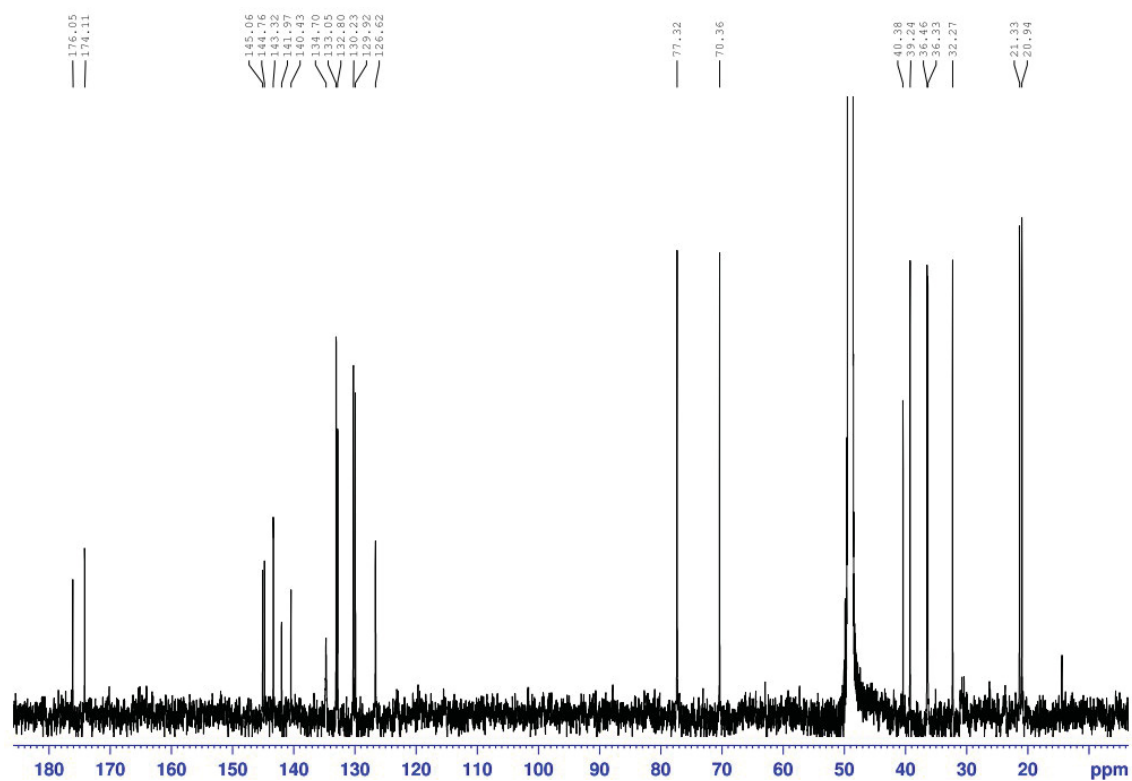
COSY and HSQC spectra of compound 3.



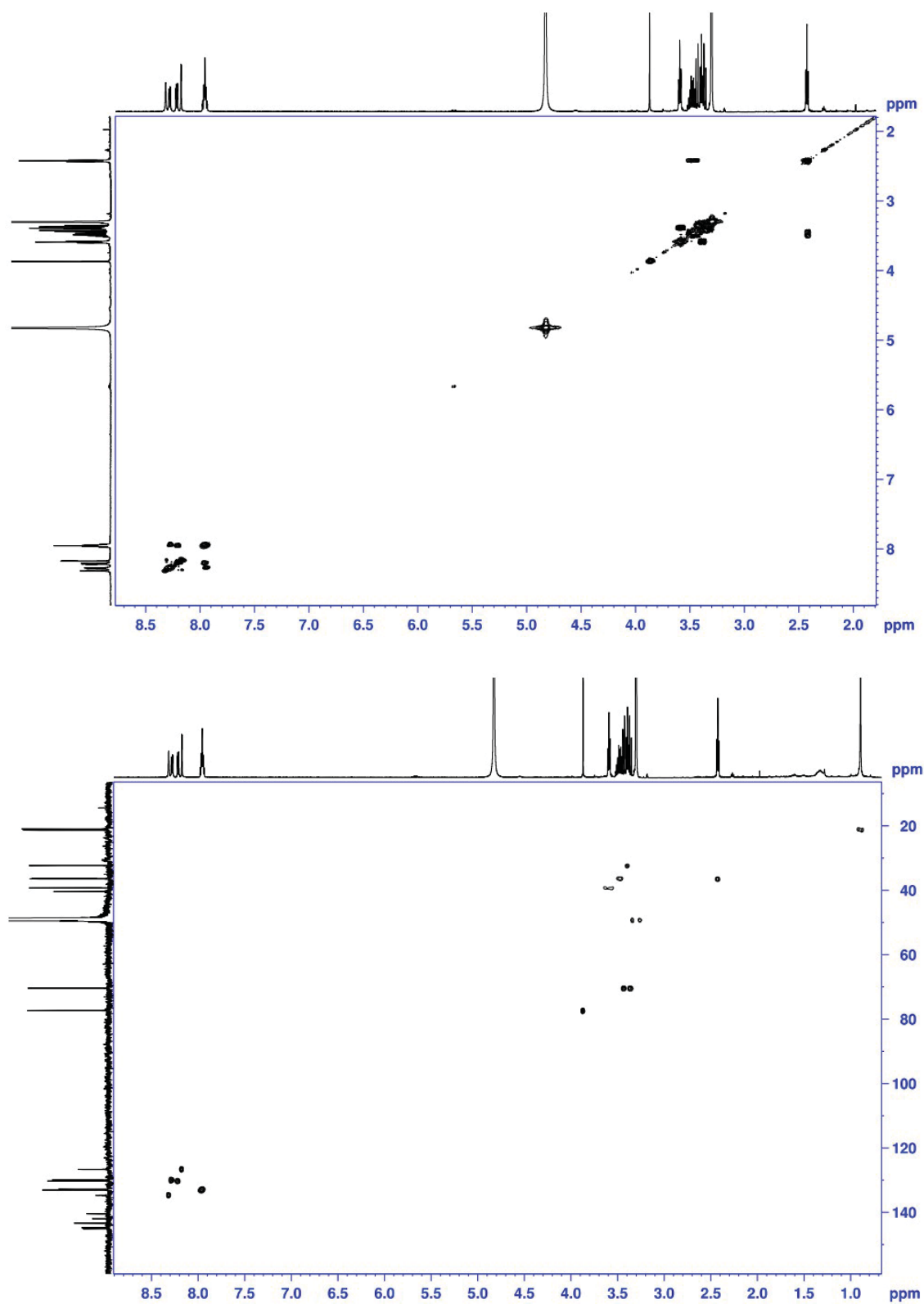
HMBC spectrum of compound **3**.



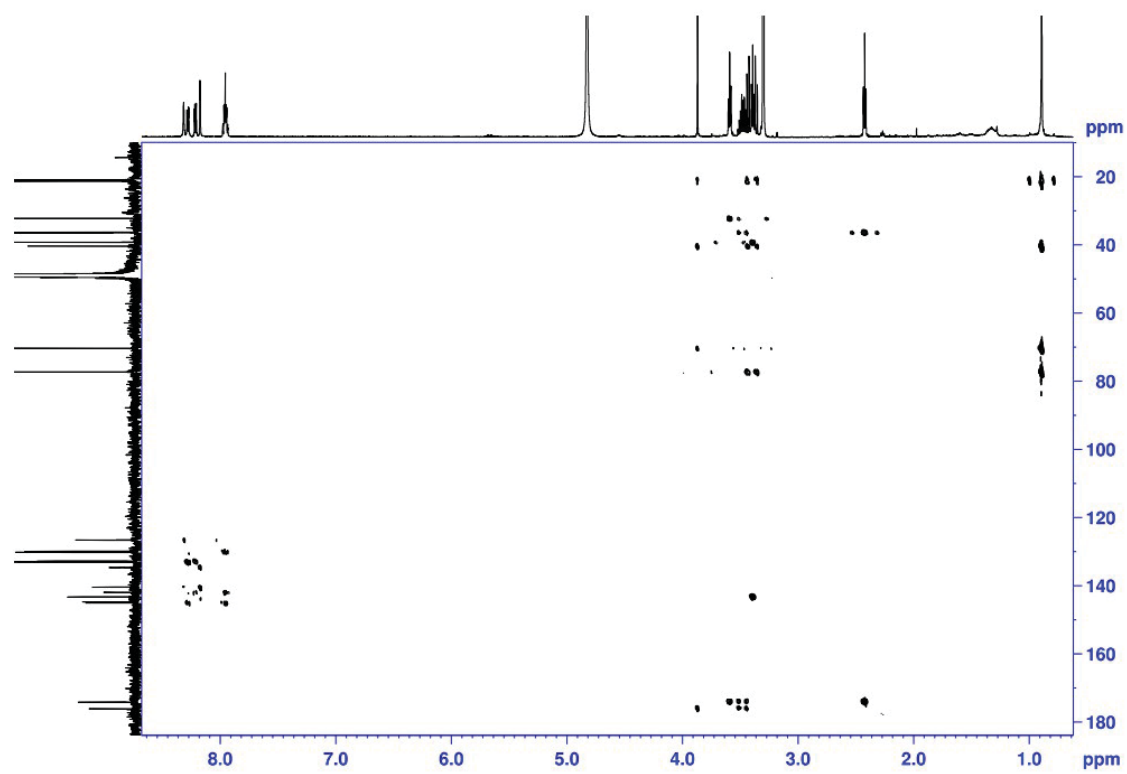
¹H NMR spectrum (600 MHz, MeOD) of 4.



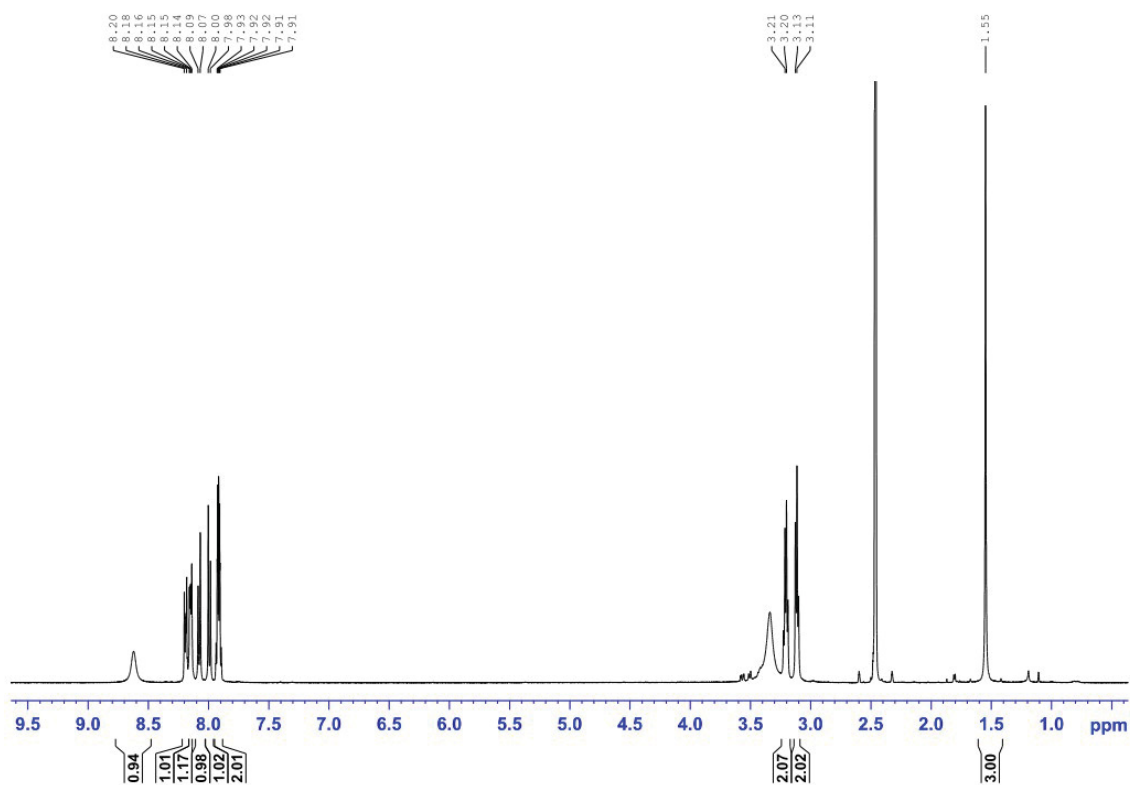
¹³C NMR spectrum (150 MHz, MeOD) of 4.



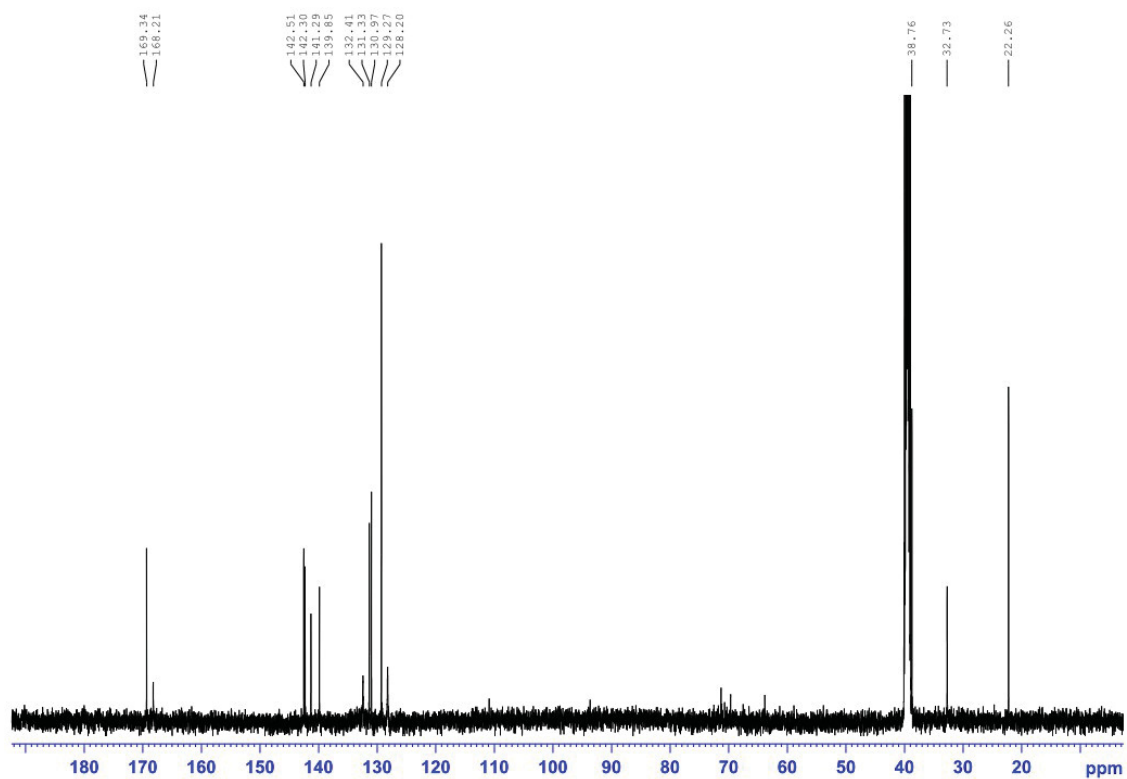
COSY and HSQC spectra of compound 4.



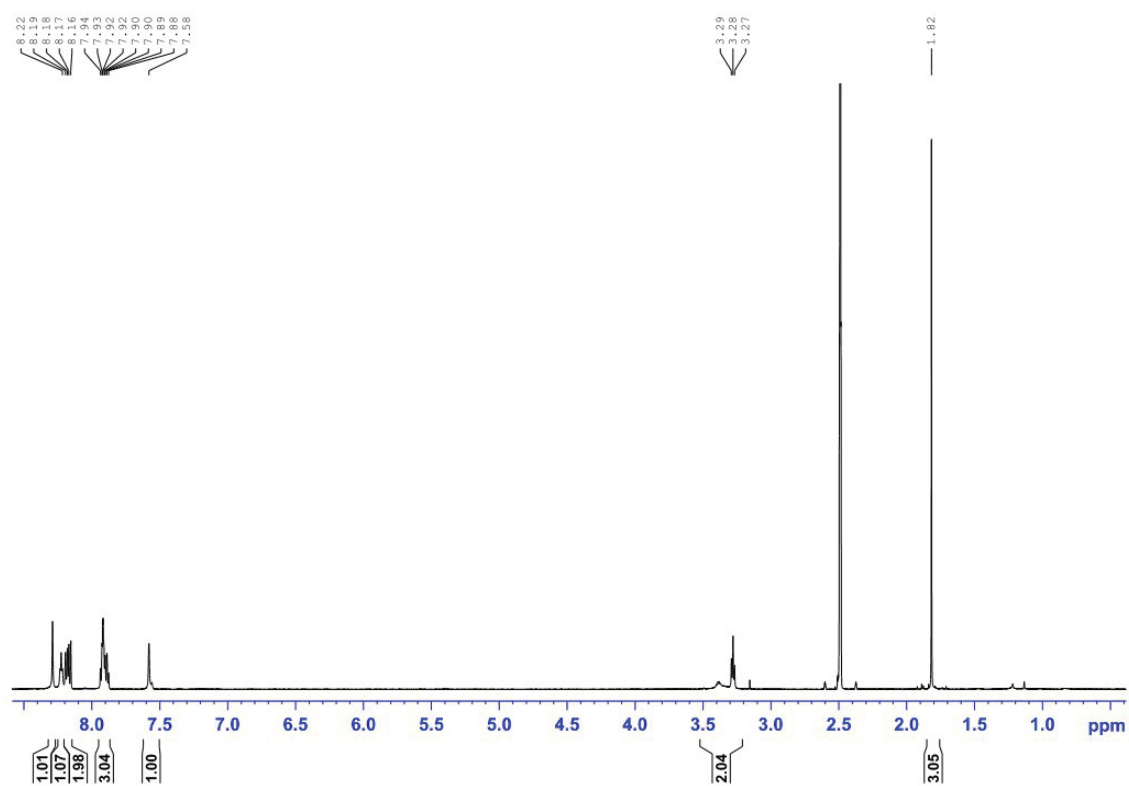
HMBC spectrum of compound **4**.



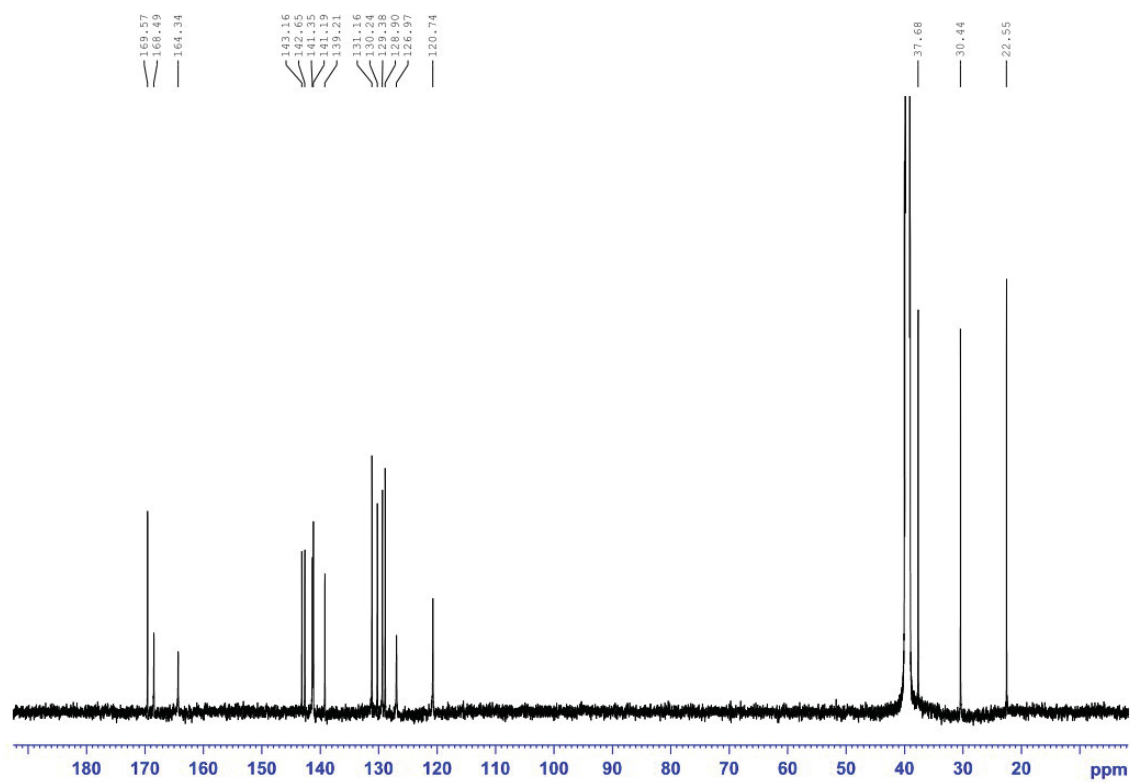
¹H NMR spectrum (500 MHz, DMSO-D₆) of **5**.



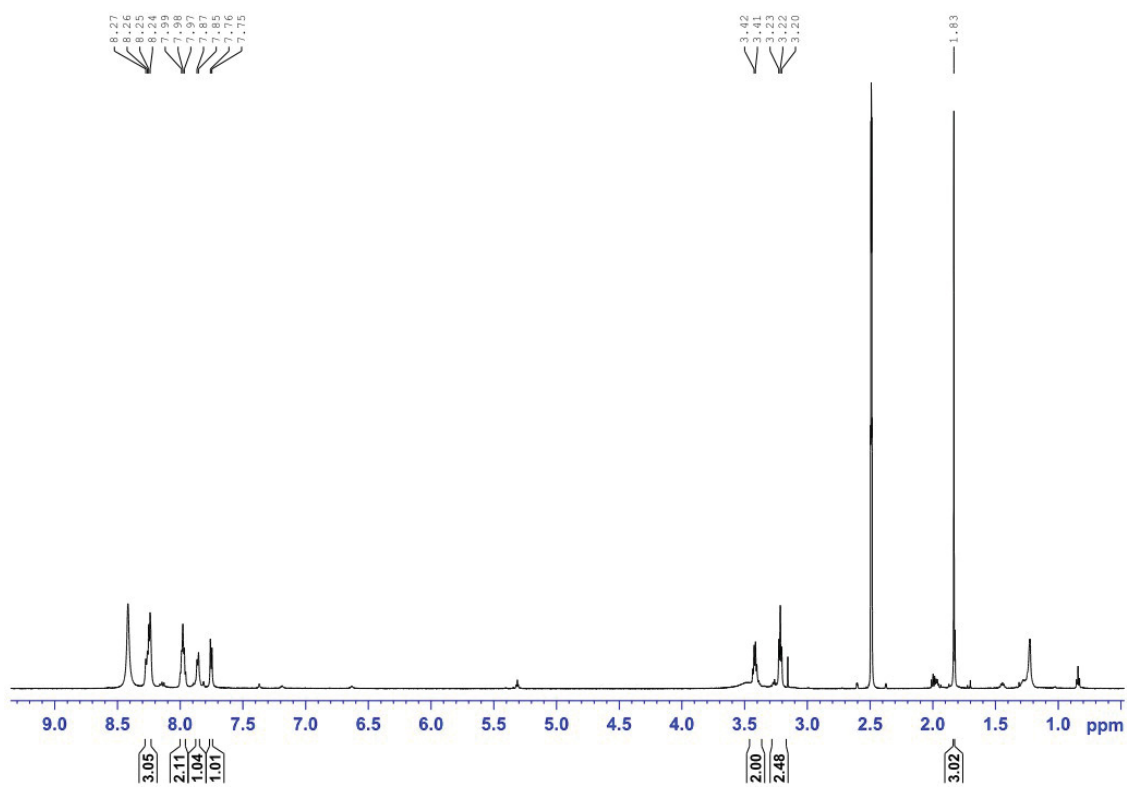
^{13}C NMR spectrum (125 MHz, DMSO-D_6) of **5**.



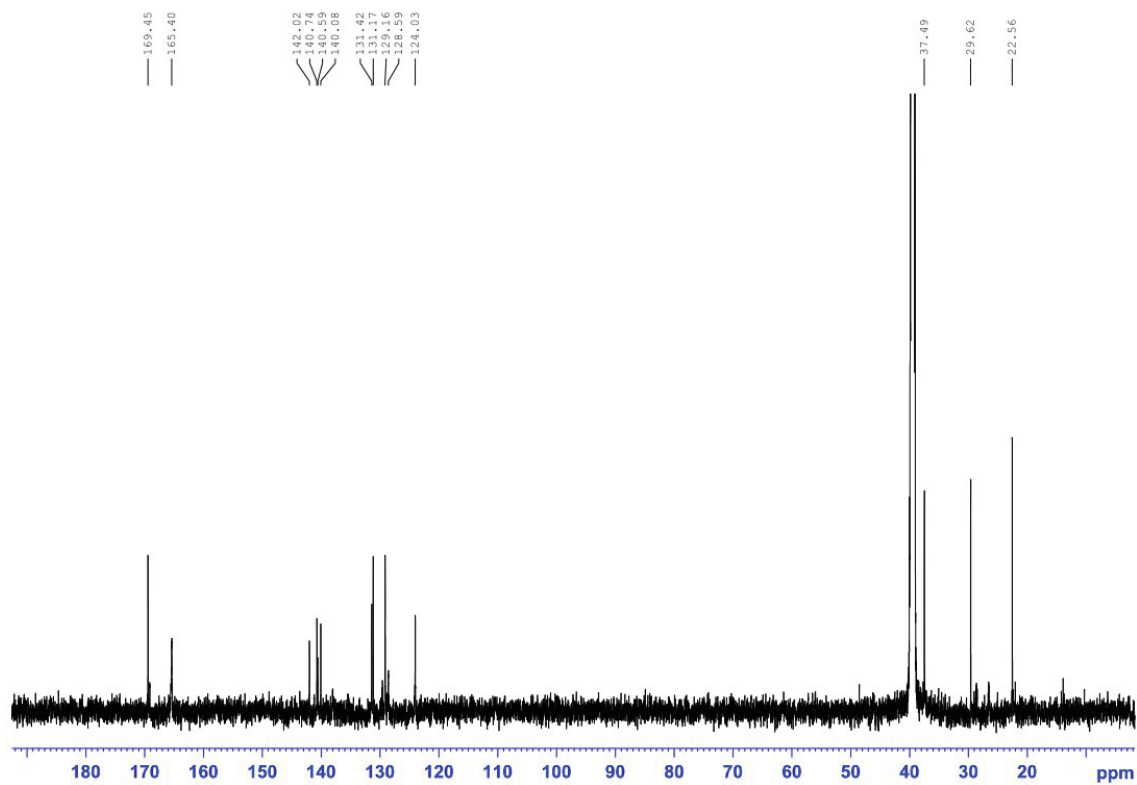
^1H NMR spectrum (600 MHz, DMSO-D_6) of **6**.



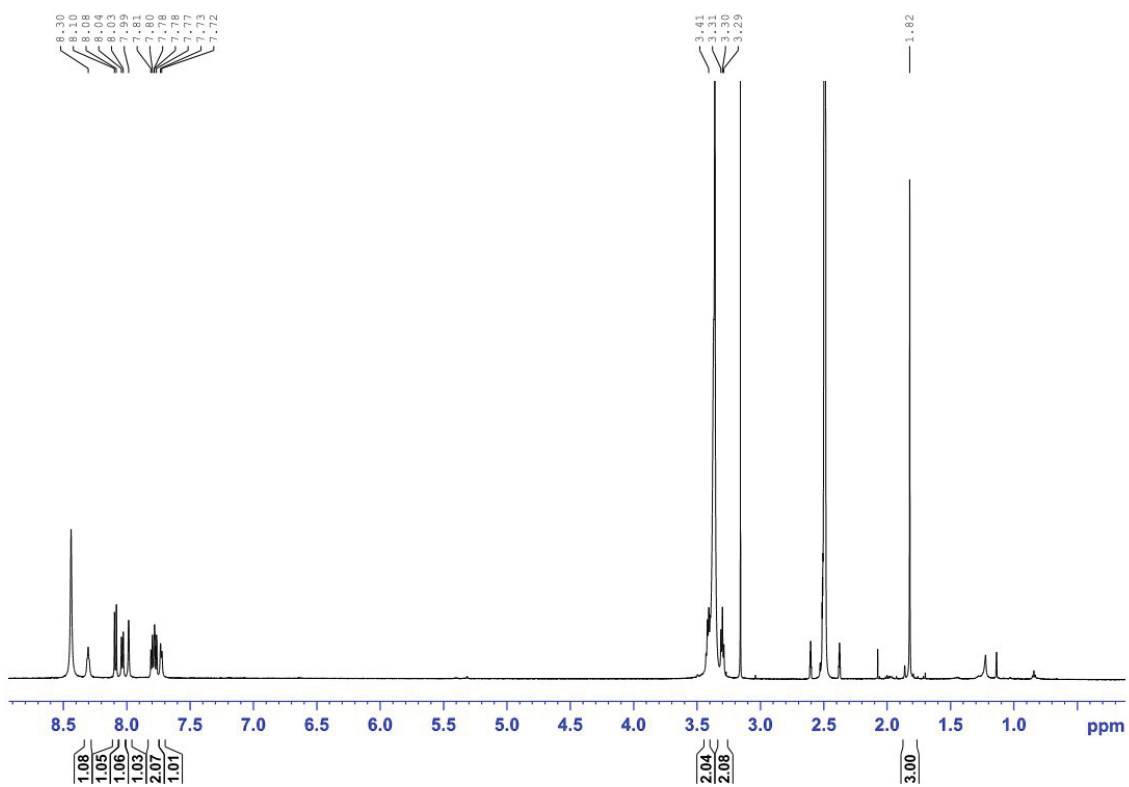
¹³C NMR spectrum (150 MHz, DMSO-D₆) of 6.



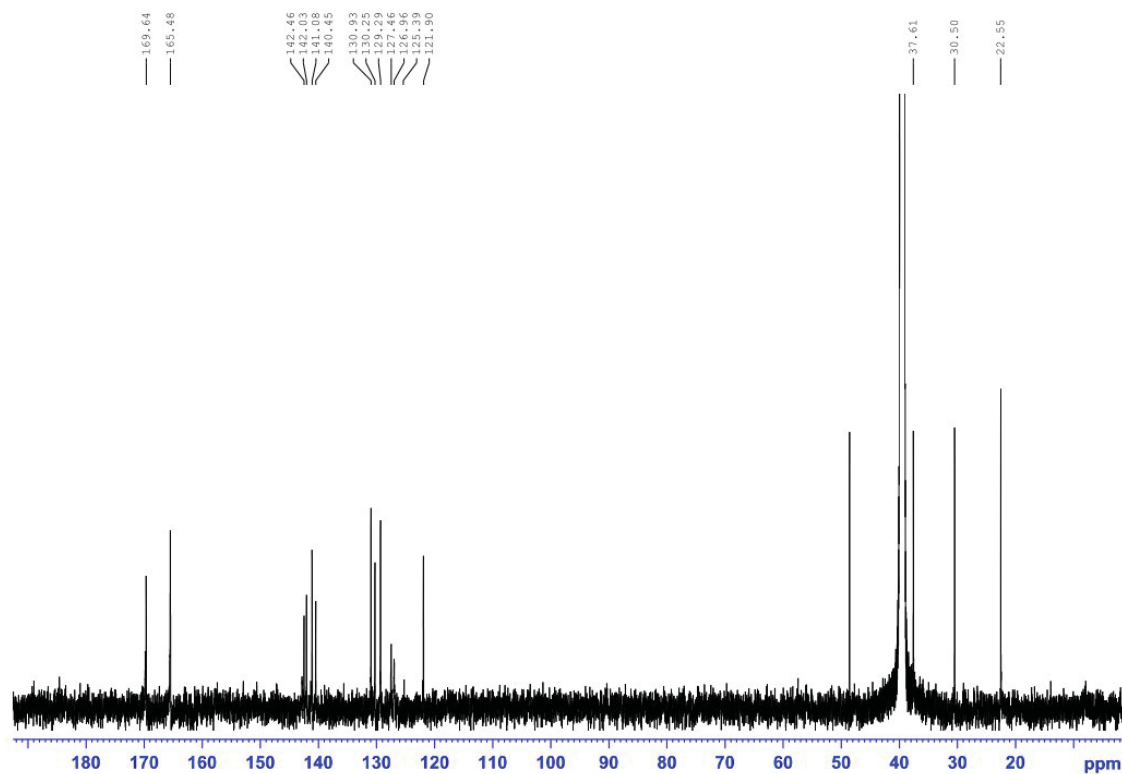
¹H NMR spectrum (500 MHz, DMSO-D₆) of 7.



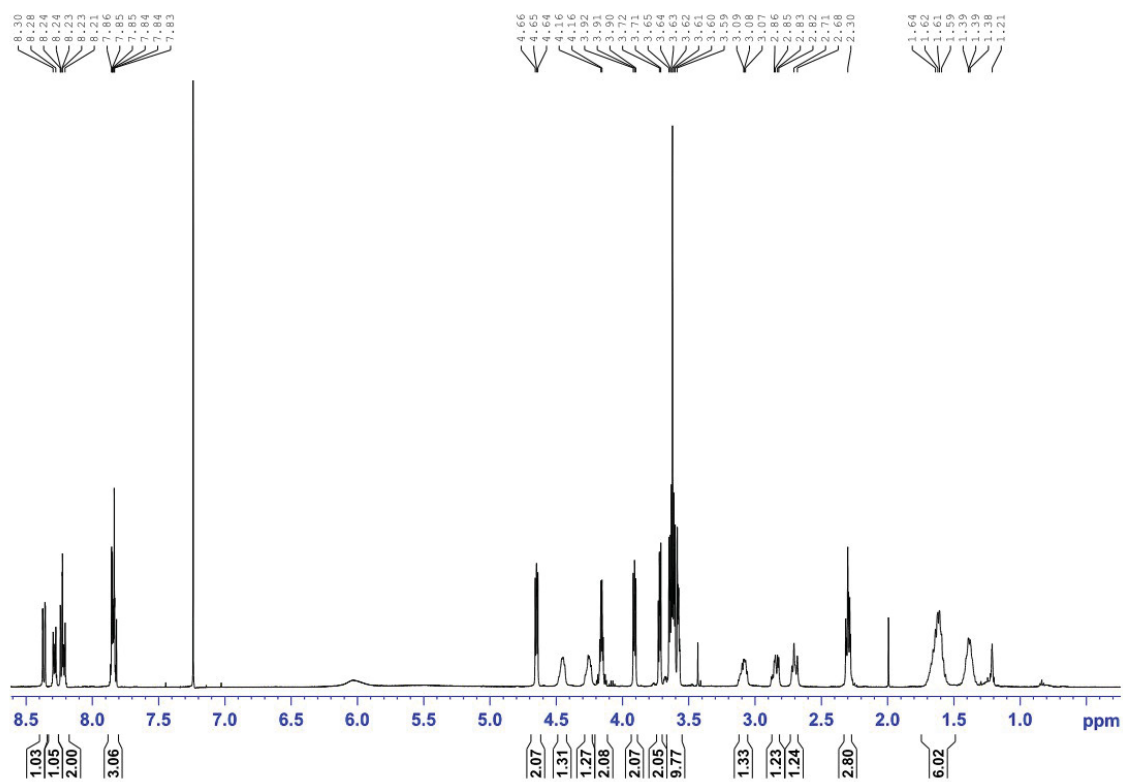
^{13}C NMR spectrum (125 MHz, DMSO-D_6) of **7**.



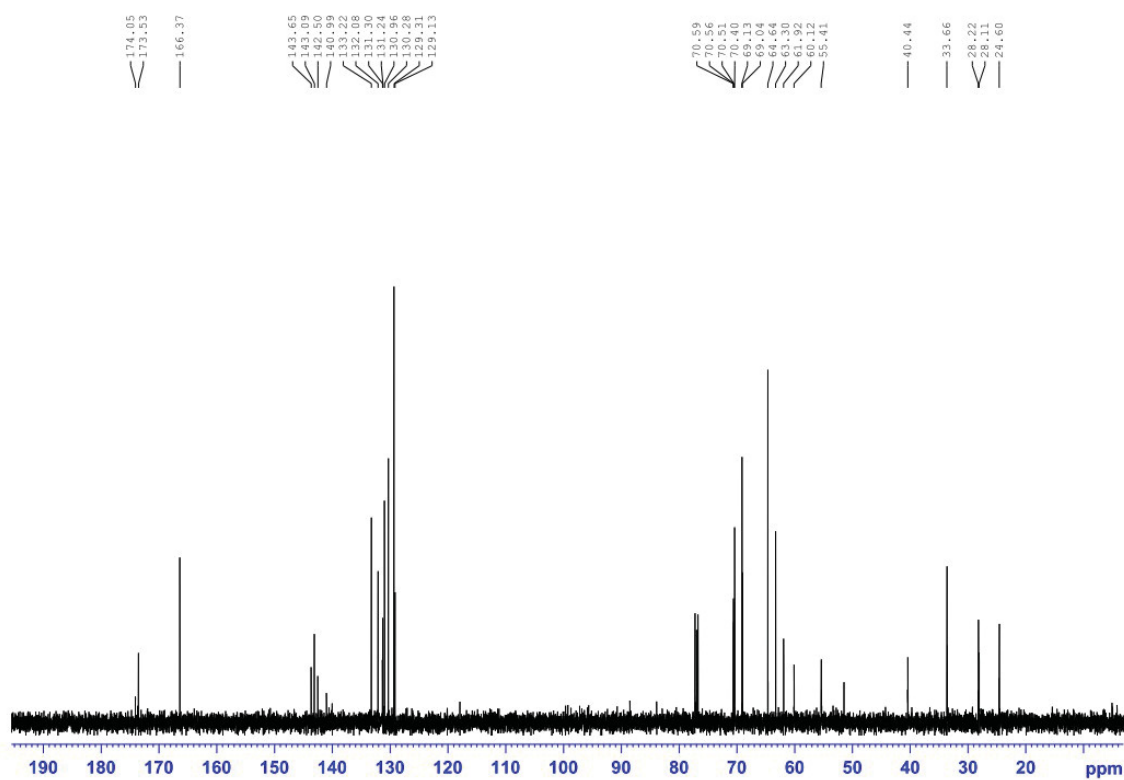
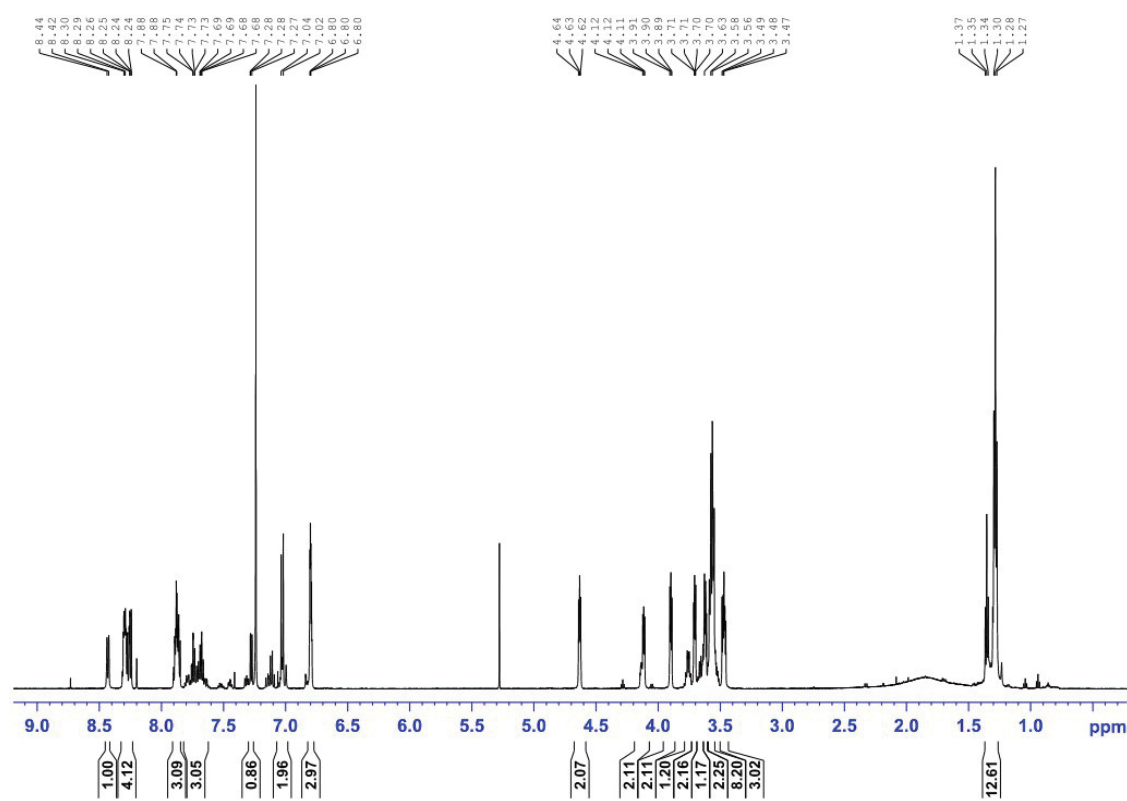
^1H NMR spectrum (600 MHz, DMSO-D_6) of **8**.

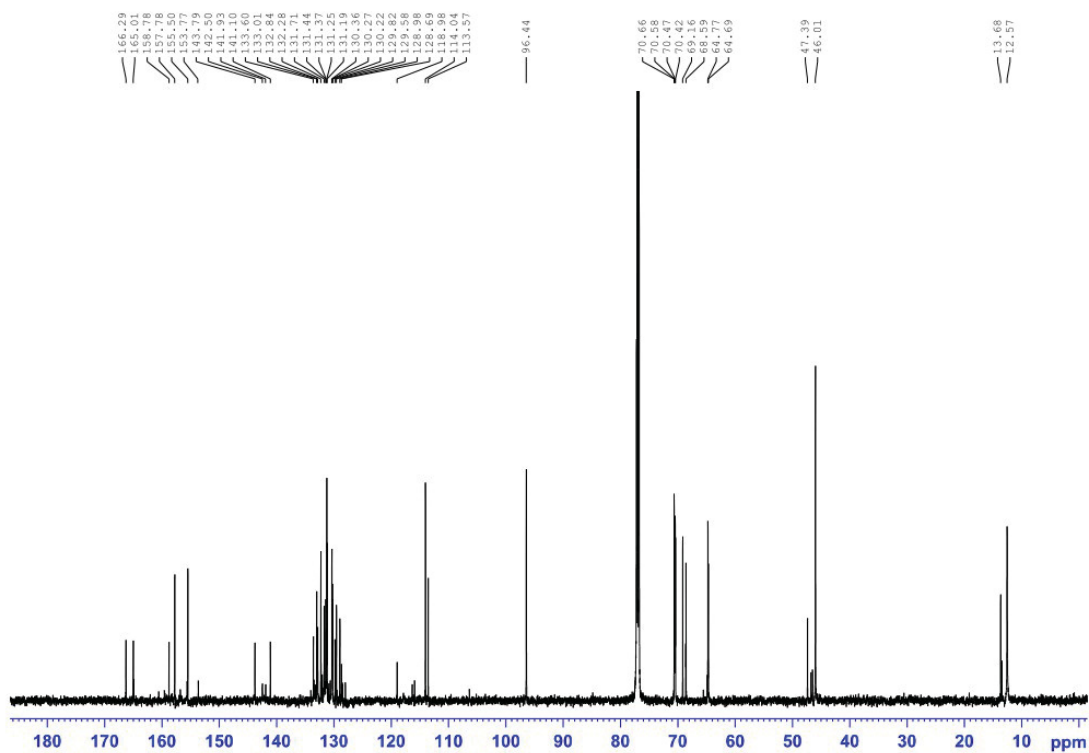


^{13}C NMR spectrum (150 MHz, DMSO-D_6) of **8**.

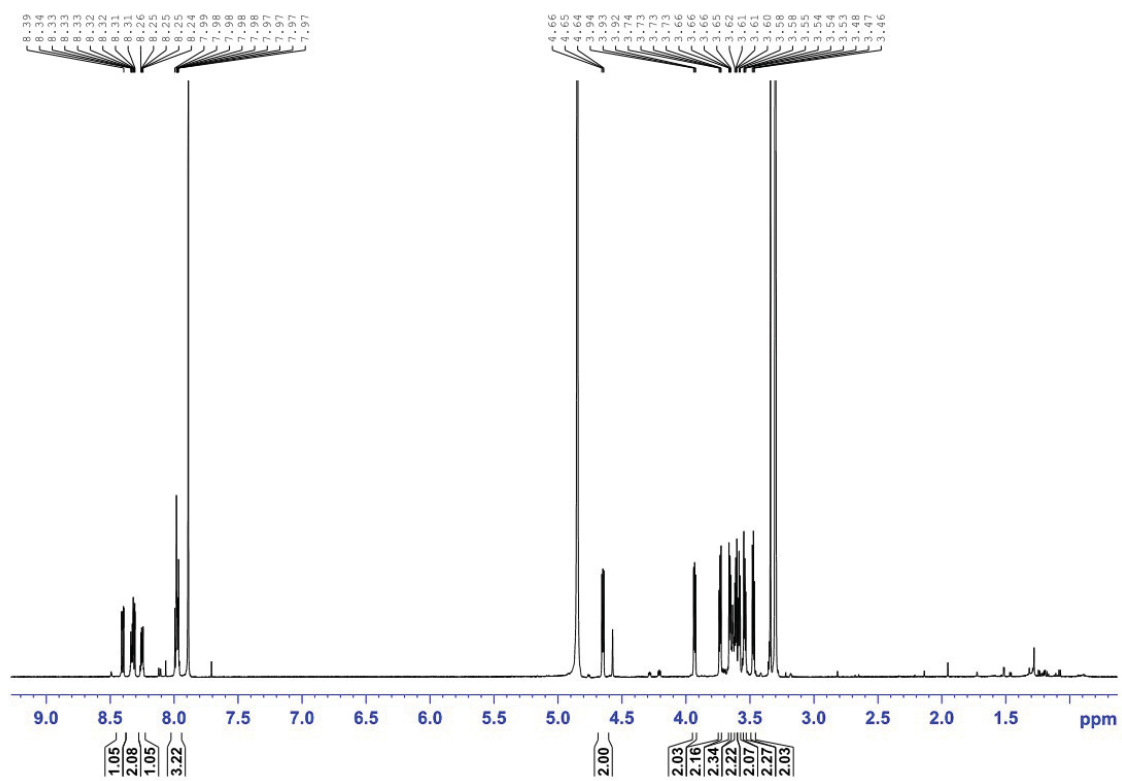


^1H NMR (500 MHz, CDCl_3) spectrum of **9**.

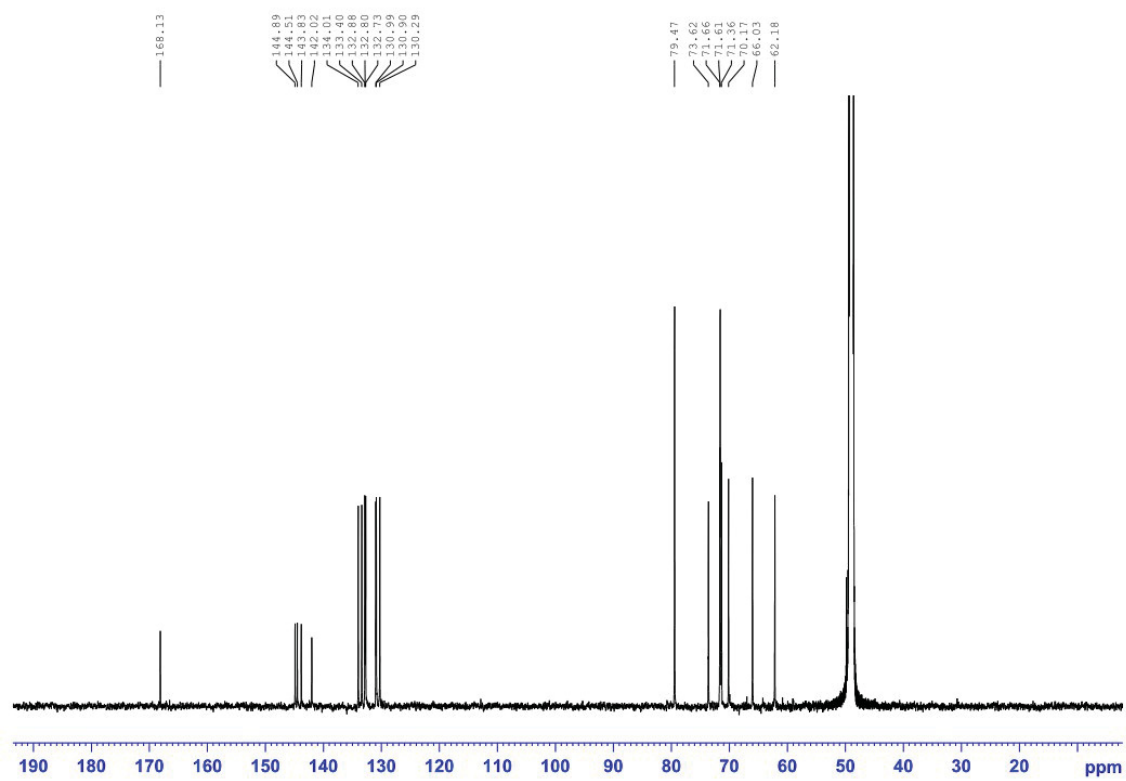
 ^{13}C NMR (150 MHz, CDCl_3) spectrum of 9. ^1H NMR spectrum (600 MHz, CDCl_3) of 10.



¹³C NMR spectrum (150 MHz, CDCl₃) of **10**.



¹H NMR spectrum (600 MHz, MeOD) of **11**.



^{13}C NMR spectrum (150 MHz, MeOD) of **11**.

List of publications

Publications from main project

Heine, Daniel; Bretschneider, Tom; Sundaram, Srividhya; Hertweck, Christian: Enzymatic polyketide chain branching to give substituted lactone, lactam, and glutarimide heterocycles, *Angewandte Chemistry International Edition* 53 (43), 11645-11649, **2014**.

Heine, Daniel; Sundaram, Srividhya; Bretschneider, Tom; Hertweck, Christian: Twofold polyketide branching by a stereoselective enzymatic Michael addition, *Chemical Communications* 51 (48), 9872-9875, **2015**.

Sundaram, Srividhya; Heine, Daniel; Hertweck, Christian: Polyketide synthase chimeras reveal key role of ketosynthase domain in chain branching, *Nature Chemical Biology* 11 (12), 949-951, **2015**.

Sundaram, Srividhya; Hertweck, Christian: On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems (Review), *Current Opinion in Chemical Biology* 31, 82-94, **2016**.

Publications from side projects

Heine, Daniel; Sundaram, Srividhya; Beudert, Matthias; Martin, Karin; Hertweck, Christian: A widespread bacterial phenazine forms conjugates with biogenic thiols and crosslinks proteins, *Chemical Science* 7, 4848-4855, **2016**.

Kugel, Susann; Baunach, Martin; Baer, Philipp; Ishida-Ito, Mie; Sundaram, Srividhya; Xu, Zhongli; Hertweck, Christian: Cryptic indole hydroxylation by a non-canonical terpenoid cyclase parallels bacterial xenobiotic detoxification, *Nature Communications*, in press, **2017**.

Conference presentations & awards

Poster presentations

'Enzymatic Polyketide Chain Branching into Lactone, Lactam and Glutarimide Heterocycles', Directing Biosynthesis IV, 03/2015, The John Innes Centre, Norwich, UK.

'Enzymatic Polyketide Chain Branching into Lactone, Lactam and Glutarimide Heterocycles', International Leibniz Research School Symposium, 05/2015, Hans Knöll Institute, Jena, Germany.

'Enzymatic Polyketide Chain Branching into Lactone, Lactam and Glutarimide Heterocycles', Biology of Bacteria Producing Natural Products (VAAM), 09/2015, Goethe Universität, Frankfurt.

'Polyketide Synthase Chimeras Reveal Key Role of Ketosynthase Domain in Chain Branching', 5th International Conference on Microbial Communication for Young Scientists (MICOM), 03/2016, Jena, Germany.

Awards

Best poster award for *'Enzymatic Polyketide Chain Branching into Lactone, Lactam and Glutarimide Heterocycles'*, International Leibniz Research School Symposium, 05/2015, Hans Knöll Institute, Jena, Germany.

Best poster award for *'Enzymatic Polyketide Chain Branching into Lactone, Lactam and Glutarimide Heterocycles'*, Biology of Bacteria Producing Natural Products (VAAM), 09/2015, Goethe Universität, Frankfurt.

Acknowledgements

I would like to thank several people who have supported me during the course of this thesis. I extend my sincere gratitude to my supervisor, Prof. Dr. Christian Hertweck for accepting me as a doctoral student and providing an exciting topic to work. It has been a wonderful learning experience and I am forever thankful to you.

I thank Dr. Tom Bretschneider for handing me over this interesting project and for being an inspiration all these years. I thank Dr. Daniel Heine, Tawatchai Thongkongkaew and Dr. Hak Joong Kim for supporting the work with their synthetic chemistry skills. I am thankful to Dr. Georg Zocher for his excellent collaboration and support over these years. Special thanks to Dr. Keishi Ishida and Dr. Kirstin Scherlach for always welcoming questions regardless of the field. I also thank Dr. Mie Ishida for initially supporting me to get comfortable in the protein lab. Thank you, Susann Kugel, for being the 'protein partner' to discuss and troubleshoot any issues we have encountered. A special thanks to you for translating and for Katharina Dornblut for proofreading the summary of this thesis. I thank Ruth Bauer for her very committed Bachelors work on the project. It has been an absolute pleasure to having worked with you.

Many thanks to Dr. Kyle Dunbar for providing time to critically read this thesis. I really appreciate your invaluable suggestions that helped improve this work.

I am very thankful to Andrea Perner, Maria Poetsch and Till Kindell for making innumerable MS and MALDI measurements. I sincerely thank Barbara Urbansky and Caroline Rückert for managing the lab logistics. You all have been a great support for me.

Special thanks to Agnieszka Litomska and my ex-colleague Dr. Emma Barnes for backing me up both in personal and professional ways. I would cherish the moments we shared together, forever. I extend my sincere thanks to all the past and present members of BMC who have supported me over these years. I will not forget all the lovely chats and the delicious food we shared. It has been the best working atmosphere one could ever ask for.

I am thankful to Uwe Knüpfer for his help in fermentation experiments. I thank the International Leibniz Research School for funding my PhD project. I sincerely thank Dr. Christine Vogler for her administrative support, right from the beginning.

I am very much grateful to my parents and in-laws for their encouragement. My most heartfelt thanks go to my husband, Vijay. You have never stopped feeling proud of me and have always been very supportive, encouraging, kind and lovable throughout my life and especially, during the course of my PhD.

Curriculum vitae

Name Srividhya Sundaram
Date of Birth 03 Oct 1985
Place of Birth Coimbatore, India
Email srividhya.sundaram@leibniz-hki.de
Address 512, Schlegelstrasse 4, 07747 Jena, Germany



EDUCATION

07/2012 – 01/2017 Leibniz Institute for Natural Product Research and Infection Biology, Hans-Knöll Institute, Jena, Germany
PhD thesis
Mentor: Christian Hertweck
Biochemical characterization of natural product biosynthetic enzymes

01/2007 – 09/2010 Centre for Biotechnology, Anna University, India
Masters (By Research) thesis
Biochemical characterization of bacterial post-translational lipid modification enzymes

02/2006 – 06/2006 Vittal Mallya Scientific Research Foundation, India
Bachelors thesis
Characterization of home-made Taq polymerase

08/2002 – 06/2006 K.S.R. College of Technology, India
Bachelors of Technology

SKILLS & INTERESTS

Languages English (Bilingual), German (Limited), Tamil (Native speaker)

Communication Effective in discussing and presenting (writing & oral) information on general as well as science-related topics
Proven ability to work at ease in an international team

Teaching Experience in teaching students from school to Masters level
Assisted in teaching basic biochemistry experiments to school teachers

Organizational Assisted in organizing exhibits for the promotion of science and technology to the general public

Other interests Reading, Gardening, Yoga, Cooking and Singing

Jena, 10.10.17

Srividhya Sundaram

Selbstständigkeitserklärung

Hiermit erkläre ich ehrenwörtlich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die aus fremden Quellen oder eigenen Prüfungsarbeiten direkt oder indirekt übernommenen Gedanken sind entsprechend gekennzeichnet. Ich habe keine unerlaubte fremde Hilfe in Anspruch genommen. Die Hilfe eines Promotionsberaters wurde nicht in Anspruch genommen, Dritte haben weder unmittelbar noch mittelbar geldwerte Leistungen für Tätigkeiten im Zusammenhang mit dem Inhalt dieser Arbeit erhalten. Diese Arbeit wurde nicht bereits zuvor als Prüfungsarbeit eingereicht. Eine gleiche, ähnliche oder andere Abhandlung wurde bei keiner anderen Hochschule als Dissertation eingereicht.

Die Promotionsordnung der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität ist mir bekannt.

Jena, 10.10.2017

Srividhya Sundaram