

Characterization of PKS II systems from entomopathogenic bacteria

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich für Biowissenschaften (15)
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Gina Luisa Carina Grammbitter

Frankfurt am Main 2019
(D 30)

vom Fachbereich für Biowissenschaften (15) der Johann Wolfgang Goethe-Universität als
Dissertation angenommen

Dekan: Professor Dr. Sven Klimpel

1. Gutachter: Professor Dr. Helge B. Bode

2. Gutachter: Professor Dr. Martin Grininger

Datum der Disputation:

1	Table of Contents	
2	LIST OF ABBREVIATIONS	III
3	ZUSAMMENFASSUNG	V
4	SUMMARY	X
5	NATURAL PRODUCT RESEARCH	1
6	THE LOGIC OF THE PKS AND FAS THIOTEMPLATE REACTIONS	3
6.1	Biosynthetic Mechanism of Polyketides and Fatty Acids	3
6.2	Classification of PKS and FAS	6
6.2.1	PKS	6
6.2.1.1.	PKS Type I.....	7
6.2.1.2.	PKS Type II.....	10
6.2.1.3.	PKS Type III.....	12
6.2.2	FAS	14
6.2.2.1.	FAS Type I.....	14
6.2.2.2.	FAS Type II.....	15
6.2.3	Comparison of Non-Acetate Priming Reactions in Type II Systems	17
6.3	The Protein Architecture of PKS and FAS Systems	20
6.3.1	Type I Systems – Different Synthases.....	20
6.3.2	Type II systems – Protein-Protein Interactions	22
6.3.2.1.	Small but Powerful – The ACP.....	23
6.3.2.2.	Protein-Protein Interactions in PKS II Systems	24
6.3.2.3.	Protein-Protein Interactions in FAS II Systems	27
6.4	PKS II Systems – Rare in Gram-Negative Bacteria	31
6.4.1	Anthraquinone (AQ) Biosynthesis.....	31
6.4.2	Isopropylstilbene (IPS) Biosynthesis	33
6.4.3	Aryl polyene (APE) Biosynthesis	35
6.5	Aim of this Thesis	38
7	PUBLICATIONS AND MANUSCRIPTS	41
7.1	An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments	41
7.2	Biosynthesis of the Multifunctional Isopropylstilbene in <i>Photo-rhabdus laumondii</i> Involves Cross-Talk between Specialized and Primary Metabolism.	42
7.3	Anthraquinone Production is Influenced by Cinnamic Acid	43
7.4	Structural Snapshots of the Minimal PKS System Responsible for Octaketide Biosynthesis	44
7.5	Molecular Mechanism of Polyketide Shortening in Anthraquinone Biosynthesis of <i>Photorhabdus laumondii</i>	45
8	ADDITIONAL RESULTS: TRACING THE FULL-LENGTH APE	47
8.1	Switching On the APE-Production in <i>X. doucetiae</i>	48
8.2	Identification of Unusual Conjugated-FAs	49
8.3	Identification of APE-Lipids	53

8.4	Materials and Methods	64
9	DISCUSSION AND FUTURE WORK	75
9.1	APEs – More than Small Molecules	75
9.1.1	APE _{Xd} -Methyl Esters from <i>X. doucetiae</i>	75
9.1.2	In Small Steps from APE-Methyl Esters to APE-Lipids	77
9.2	Multitasking APE-Synthase	84
9.3	Physiological function of APEs – Really an Antioxidant Agent?	88
9.4	The APE-Lipid Synthase – Forming a Multienzyme Complex?	91
9.5	Special Features in IPS Biosynthesis	96
9.5.1	Unraveling the Missing Steps in Biosynthesis of IPS	96
9.5.2	Gatekeeper Enzymes StIB and pIFabH	97
9.6	Comparing ACPs from APE, IPS and AQ biosynthesis	100
10	CONCLUSION	107
11	LITERATURE	111
12	ATTACHMENTS	119
12.1	Publication: An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments	119
12.2	Manuscript: Biosynthesis of the Multifunctional Isopropylstilbene in <i>Photorhabdus laumondii</i> Involves Cross-Talk between Specialized and Primary Metabolism	181
12.3	Manuscript: Anthraquinone Production is Influenced by Cinnamic Acid	215
12.4	Manuscript: Structural Snapshots of the Minimal PKS System Responsible for Octaketide Biosynthesis	227
12.5	Publication: Molecular Mechanism of Polyketide Shortening in Anthraquinone Biosynthesis of <i>Photorhabdus laumondii</i>	283
12.6	Supplemental information: Tracing the Full-length APE	325
13	LIST OF PUBLICATIONS	333
14	RECORD OF CONFERENCES	334
15	ERKLÄRUNG UND VERSICHERUNG	335

2 List of Abbreviations

aa	amino acid
AasS	acyl-acyl-carrier protein synthetase
ACC	acetyl-CoA:carboxylase
ACN	acetonitrile
ACP	acyl-carrier protein
APE	aryl polyene
AQ	anthraquinone
AR	adapter region
ARO	aromatase
AT	acyl transferase
BGC	biosynthetic gene cluster
BKD	branched-chain α -keto acid dehydrogenase
bp	base pairs
CHS	chalcone synthase
CLF	chain-length factor
CoA	Coenzyme A
CX-MS	chemical cross-linking coupled with mass spectrometry
CYC	cyclase
Da/MDa	dalton/megadalton
DAR	2,5-dialkylresorcinol
DEBS	6-deoxyerythronolide B synthase
DH	β -hydroxyacyl-ACP dehydratase
DNA	deoxyribonucleic acid
EN	double-bond in APE compound
ER	enoyl-ACP reductase
ESI	electrospray ionization
FAS	fatty acid synthase
G3P	glycerol-3-phosphate
GC	gas chromatography
GlcNAc	glucosamine
GNAT	GCN5-related N-acetyltransferase
HBA	hydroxybenzoic acid
HPLC	high performance liquid chromatography
HR	high resolution
IPS	3,5-Dihydroxy-4-isopropylstilbene
IPTG	isopropyl- β -D-thiogalactopyranosid
KAS	β -ketoacyl-ACP synthase
kb	kilobase
KR	β -ketoacyl-ACP reductase
KS	ketosynthase
LPS	lipopolysaccharide
MALDI	matrix assisted laser desorption/ionization
MAT	malonyl-acetyl-transferase
MCAT	malonyl-CoA:acyltransferase
Met	methylation of aryl-ring in APE compound
MmpL	Mycobacterial membrane protein Large
MPT	malonyl-palmitoyl-transferase
MS	mass spectrometry
MS ²	tandem mass spectrometry
MT	methyltransferase
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance

List of Abbreviations

NRPS	nonribosomal peptide synthetase
PCP	peptidyl-carrierprotein
PGL	phenolic glycolipids
PKS	polyketide synthase
Ppant	4'-phosphopantetheine
PPTase	4'-phosphopantetheinyl transferase
RND	Resistance-Nodulation-Cell Division
ROS	reactive oxygen species
STS	stilbene synthase
TE	thioesterase
TOF	time of flight
UDP	uridine diphosphate glucose
UV-Vis	Ultraviolet-visible

3 Zusammenfassung

Polyketide stellen eine wichtige Substanzklasse von Naturstoffen dar, deren Synthese aus chemischer Sicht der Synthese von Fettsäuren ähnelt. Über eine decarboxylierende Claisen-Kondensationsreaktion erfolgt die Verknüpfung von Acetateinheiten unter der Bildung einer C-C-Bindung. Die Biosynthese erfolgt dabei proteingebunden, mit sogenannten Acyl-Carrier-Proteinen (ACPs) unter der Einbeziehung einer komplexen Biosynthesemaschinerie aus Enzymen. Dabei sorgt die posttranslationale Modifikation der ACPs mit dem hoch flexiblen Phosphopantetheinyl-Arm (Ppant arm) dafür, dass Starter-, Verlängerungseinheiten und die entstehenden Intermediate, kovalent gekoppelt werden können, um für die Prozessierung in die räumliche Nähe der Enzyme gebracht zu werden.

Im Gegensatz zu Fettsäuren weisen Polyketide eine enorme strukturelle Vielfalt auf. Die erzeugte β -Keto-Verbindung, die nach der Verknüpfungsreaktion von zwei Acetateinheiten entsteht und eine C-C-Bindung erzeugt, wird während der Fettsäurebiosynthese meist zu einer vollständig gesättigten Verbindung prozessiert. Dies geschieht durch die Ketoreduktion mit einer Ketoreduktase (KR), gefolgt von einer Wassereliminierung durch eine Dehydratase (DH) und der erneuten Reduktion mit Hilfe einer Enoylreduktase (ER). Diese Schritte werden in der Polyketidsynthese teilweise nicht vollständig durchgeführt und es gibt zudem eine Vielzahl an weiteren Prozessierungsreaktionen, wie beispielsweise Zyklisierungen, Aromatisierungen, Methylierungen, Oxidationen und Glykosylierungen usw., die für die strukturelle Vielfalt der Polyketidprodukte sorgen.

Gesteuert wird die Biosynthese unter anderem von der Biosynthesemaschinerie, die in insgesamt drei unterschiedliche Typen eingeteilt werden kann. Typ I PKS-Systeme stellen multifunktionale Proteinkomplexe dar, bestehend aus Proteinen, welche die jeweiligen Enzymdomänen tragen. Hier entstehen in der Regel reduzierte lineare oder zyklische Polyketide. Typ II PKS-Systeme beinhalten monofunktionale Enzyme, die miteinander interagieren. Dabei wird ein Polyketongrundgerüst mit Hilfe einer minimal-PKS hergestellt, das aus ACP, KS und einem sogenannten *chain-length factor* (CLF) besteht. Die Malonyl-CoA:Transacyltransferase (MCAT) aus dem Primärmetabolismus, sorgt für die Bereitstellung von Malonat. Das Malonat wird für die decarboxylierende Claisen-Kondensationsreaktion zur Verlängerung genutzt. KS und CLF bilden ein Heterodimer und schützen das hochreaktive Polyketonintermediat während der Synthese, das typischerweise, durch die weitere Prozessierung mit einer KR, einer Zyklase und einer Aromatase, in eine aromatische Verbindung umgesetzt wird.

PKS III-Systeme bestehen ebenso aus monofunktionalen Enzymen, mit dem Unterschied, statt ACP-gebundene Intermediate, Coenzym A-Intermediate für die Biosynthese zu nutzen. Außerdem werden in der Regel statt Acetat-Startereinheiten, aromatische Starter genutzt und es entstehen kleine aromatische Verbindungen.

In der vorliegenden Arbeit wird die ausführliche Charakterisierung von drei unterschiedlichen Typ II PKS-Systemen aus den Gram-negativen Bakterien *Xenorhabdus* und *Photorhabdus* beschrieben. Die jeweiligen PKS II-Systeme, für die Produktion von Arylpolyenen, Isopropylstilbenen und Anthrachinonen, wurden sowohl auf Protein- als auch Produktebene untersucht. PKS vom Typ II werden sonst, mit nur wenigen Ausnahmen, in Gram-positiven Bakterien gefunden.

Ein großer Teil dieser Arbeit beschäftigt sich dabei mit der Analyse der biochemisch wenig charakterisierten Substanzklasse der Arylpolyen (APE)-Pigmente. Vor fünf Jahren waren APEs lediglich unter dem Namen Xanthomonadin und Flexirubin bekannt, zwei Naturstoffe, die bereits um 1970 entdeckt und in Folgearbeiten strukturell charakterisiert wurden. Zu dieser Zeit war nicht bekannt, dass beide Substanzen zu den APEs zählen. Es wurde zunächst angenommen, dass Xanthomonadin, gattungsspezifisch, nur in Xanthomonaden und Flexirubin in Flavobakterien vorkommen, welche durch die Produktion der Pigmente einen charakteristischen, gelben Phänotyp aufweisen.

Im Jahr 2014 brachte die Publikation von Cimermancic *et al.* durch eine ausführliche bioinformatische Analyse, das Gencluster von APE-Pigmenten zum Vorschein, das bis dahin keiner Substanzklasse zugeordnet werden konnte. Anhand der Genausstattung der entsprechenden Biosynthesegencluster (BGCs) wurden APEs in insgesamt drei Familien eingeteilt, wobei Xanthomonadin und Flexirubin der Familie drei zugeordnet werden konnten. Diese Familie wurde noch einmal in die entsprechenden Subfamilien zu Xanthomonadin- oder Flexirubin-ähnlichen BGCs eingeteilt. Die Produkte der beiden anderen APE-Familien waren nicht bekannt und die Forscher charakterisierten jeweils einen repräsentativen Kandidaten. Die Struktur eines APEs aus *E. coli* CFT073, dem APE_{Ec}, aus Familie eins und eines APEs aus *Vibrio fischeri*, dem APE_{Vf}, aus Familie zwei, wurden aufgeklärt. Diese ausführliche Analyse zeigte ferner, dass das APE-BGC eines der am weit verbreitetsten Gencluster in Gram-negativen Proteobakterien darstellt.

Anhand der Einteilung von Cimermancic *et al.* stellte sich heraus, dass in Spezies der Gattung *Xenorhabdus*, das APE-BGC der Familie eins, ebenfalls weit verbreitet ist.

Neben der Identifikation der APE-Pigmente aus *Xenorhabdus*, stand vor allem die Aufklärung der bisher ungeklärten Biosynthese im Vordergrund. Das APE-BGC weist zwar einige Charakteristika von Typ II PKS auf, doch sind im entsprechenden BGC einige Enzyme codiert, die eher der Fettsäurebiosynthese zuzuordnen wären. Aufgrund dieser Gene und der eher linearen Struktur des APEs wurde postuliert, dass die Biosynthese in einem Fettsäurebiosynthese-ähnlichen Mechanismus stattfindet.

Die biochemische Analyse von Enzymen, die an der Biosynthese von APEs beteiligt sind und die *in vitro*-Rekonstruktion der ACP-gebundenen APE-Biosynthese konnte bestätigen, dass APEs durch einen fettsäureähnlichen Verlängerungsmechanismus hergestellt werden und demnach kein Polyketonintermediat bilden. Darüber hinaus lieferte die Entdeckung unerwarteter Protein-Protein-Interaktionen neue Erkenntnisse über die Multienzymkomplexbildung dieses außergewöhnlichen PKS II-Systems. Durch die Zusammenarbeit mit den Gruppen von Prof. Michael Groll und jr. Prof. Nina Morgner wurden zwei hetero-Proteinkomplexe strukturell gelöst und mehrere native Multimerisierungs-Ereignisse der Proteine identifiziert, die es ermöglichten, ein potentielles Protein-Interaktionsnetzwerk vorzuschlagen. Die Ergebnisse zur *in vitro*-Rekonstitution von APEs sind in der Veröffentlichung '*An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments*' (Publikation *J. Am. Chem. Soc.*, Erstautor) zusammengefasst.

Neben der *in vitro*-Analyse, wurden über *in vivo*-Studien, die von *Xenorhabdus doucetiae* hergestellten APE-Verbindungen, näher untersucht. Die Aktivierung des unter Laborbedingungen stillen BGCs führte zum Nachweis der APE-Verbindungen im homologen Wirt. Durch die Kombination aus homologer Überproduktion und gezielter Gendeletionen im APE BGC, wurde eine APE-lipidähnliche Struktur identifiziert. Mit Hilfe von massenspektrometrischen Analysen und der Isolation von Intermediaten, wurden die strukturellen Bausteine dieses APE-Lipids identifiziert. Es stellte sich heraus, dass das APE-Lipid, neben dem APE-Teil, aus einem Glucosaminrest, einer langkettigen Fettsäure mit ungewöhnlich konjugierten Doppelbindungen und einer Phosphoethanolamin-Kopfgruppe, wahrscheinlich über ein Glycerin verknüpft, besteht. In Kombination mit den oben genannten *in vitro*-Daten wurde ein Biosynthese-Mechanismus des APE-Lipids postuliert, der die Erkenntnisse von Cimermancic *et al.* über die Substanzklasse der APEs erweitert. Diese Ergebnisse sind im Abschnitt '*Additional Results: Tracing the Full-length APE*' dargestellt.

Ein weiteres, eher untypisches Typ II PKS-System stellt die Biosynthesemaschinerie zur Produktion von Isopropylstilben (IPS) dar. Große Teile der IPS-Biosynthese wurden bereits im Bode-Labor und der Arbeitsgruppe von Prof. Ikuro Abe untersucht. Studien mit *Photorhabdus laumondii* in der Arbeitsgruppe Bode zeigten, dass sich die Gene der Biosynthese nicht in

einem BGC anordnen, sondern im Genom verteilt vorliegen. Eine Funktionsanalyse dieser Gene zeigte weiter, dass die IPS-Vorläufer-Moleküle aus zwei unterschiedlichen Biosyntheserouten stammen. Die Bode-Gruppe entdeckte hierbei, dass eine ungewöhnliche KS/Cyclase (StlD) die Kondensation beider Vorläufer, dem verlängerten und prozessierten Zimtsäure- und dem verlängerten Isovalerat-Vorläufer, über eine Michael-Addition katalysiert. Ein solcher Weg zur Stilbenbildung unterscheidet sich von dem Biosyntheseweg, der in Pflanzen weit verbreitet ist und beispielsweise nur aus einem sich verlängernden Zimtsäure-Vorläufermolekül mit Hilfe einer PKS III-Maschinerie aufbaut. Die Abe-Gruppe löste die Proteinstruktur und den biochemischen Mechanismus der KS/Cyclase StlD und untersuchte weiter die Aromatisierungsreaktion der Aromatase StlC. Wie jedoch Zimtsäure auf das ACP geladen und weiter zu dem benötigten Michael-Akzeptor prozessiert wird, der für die Zyklisierungsreaktion erforderlich ist, blieb bisher ungeklärt.

In dieser Arbeit wurde die Biosynthese des Michael-Akzeptors erfolgreich *in vitro* rekonstruiert. Unter anderem wurde gezeigt, dass dabei Enzyme aus der Fettsäurebiosynthese involviert werden, wodurch sich ein erforderliches Zusammenspiel von Enzymen aus dem Primär- mit solchen aus dem Sekundärstoffwechsel nachweisen lässt. Anhand dieser Erkenntnisse war es möglich die minimale Proteinausstattung zur Produktion von Stilbenen in einem heterologen Wirt festzulegen. Es wurde gezeigt, dass StlB nicht wie bisher postuliert als Coenzym A (CoA)-Ligase fungiert und Zimtsäure zu Zimtsäure-CoA umwandelt, sondern die Acyl-Gruppe von adenylierter Zimtsäure direkt auf das entsprechende ACP belädt, um dadurch Zimtsäure-ACP herzustellen. Neben der Mitwirkung weiterer Enzyme aus der Fettsäurebiosynthese, spielt die KS FabH aus *P. laumondii* (*plFabH*), eine Schlüsselrolle als *gatekeeper*-Enzym. Es konnte gezeigt werden, dass das *E. coli*-Homolog nicht in der Lage ist Zimtsäure-ACP um eine C2-Einheit zu verlängern, im Gegensatz zu *plFabH*. Mit Hilfe dieser entscheidenden Informationen gelang die rekombinante IPS-Produktion in *E. coli*. Die Ergebnisse hierzu werden im Manuskript '*Biosynthesis of the Multifunctional Isopropylstilbene in Photorhabdus laumondii Involves Cross-Talk between Specialized and Primary Metabolism*' (Manuskript, geteilter Erstautor) vorgestellt.

Anhand der bislang durchgeführten Studien, stellt das PKS II-System der orange-rot pigmentierten Anthrachinone (AQs) ein bereits gut untersuchtes PKS II-System dar. Bereits 2007 entdeckten Brachmann *et al.* das AQ-BGC und dessen Produkte, darunter das Hauptprodukt AQ-256 und die daraus hervorgehenden, methylierten Derivate. Quiqin Zhou (Arbeitsgruppe Bode) führte weitere biochemische *in vitro*-Analysen durch. Kombiniert mit der heterologen Expression des AQ-BGCs führte dies zur Identifizierung von Nebenprodukten, sogenannten *shunt*-Produkten, die auf ein AQ-Gerüst hindeuteten, das von einem

Oktaketidzwischenprodukt abgeleitet werden kann. Da mit Markierungsexperimenten herausgefunden wurde, dass das Hauptanthrachinon AQ-256 nur aus sieben Acetat-Einheiten besteht, wurde postuliert, dass ein Oktaketidintermediat von der Hydrolase AntI zu einem Heptaketid verkürzt wird. Dieser PKS-Kürzungsmechanismus wurde durch die Proteinkristallstruktur von AntI durch die Arbeitsgruppe Groll strukturell bestätigt und brachte den molekularen Mechanismus zum Vorschein (*'Molecular Mechanism of Polyketide Shortening in Anthraquinone Biosynthesis of Photorhabdus laumondii'*, Publikation *Chem. Sci.*, Co-Autor).

Weitere Substratanalysen des *Photorhabdus laumondii* AQ-Produzenten und einer Mutante zeigten eine mögliche Hemmung von Zimtsäure gegenüber der Hydrolase AntI (*'Anthraquinone Production is Influenced by Cinnamic Acid'*, Manuskript, Erstautor).

Die biochemische Analyse während der Doktorarbeit von Quiqin Zhou mit dem minimalen-PKS der AQ-Synthase ergab weiterhin die exklusive Aktivierung des AQ-ACPs durch die Phosphopantetheinyl-Transferase (PPTase) AntB. PPTasen katalysieren die Übertragung des Ppant-Arms zur Aktivierung der ACPs. AntB allein ist unlöslich, wird aber durch eine CoA-Ligase stabilisiert, die höchstwahrscheinlich inaktiv ist und als Chaperon arbeitet. So beinhaltet die minimale PKS-Ausstattung zur Herstellung des Oktaketidgerüsts neben dem ACP, dem KS:CLF-Heterodimer und der MCAT, die weitere Co-Existenz der PPTase AntB und der CoA-Ligase AntG. Durch die strukturellen Analysen der Arbeitsgruppe Groll war es erstmals möglich das minimale-PKS der AQ-Biosynthese in Aktion darzustellen. Die Strukturdaten von Komplexen aus einem ACP:KS:CLF, dem KS:CLF allein und einem ACP:MCAT in ihrer nicht-aktiven und aktiven Form aufgeklärt wurden. Es war möglich, ein KS-gebundenes Hexaketid zu detektieren, das nach der heterologen Proteinproduktion von KS:CLF, am konservierten Cystein-Rest der KS, gebunden vorlag.

Zudem brachte die Mutagenese von Aminosäuren, die anhand der Strukturen an den Protein-Protein-Interaktionen im ACP:KS:CLF-Komplex beteiligt zu sein schienen, einige interessante Interaktionsstellen zum Vorschein. Anhand der Proteinkomplexstruktur wurde ferner ein *induced-fit*-Mechanismus der MCAT mit dem ACP postuliert, durch den sich eine monodirektionale Transferreaktion während der Malonylierung erklären lässt (*'Structural Snapshots of the Minimal PKS System Responsible for Octaketide Biosynthesis'*, Manuskript, Co-Autor).

4 Summary

This work deals with the characterization of three different type II polyketide synthase systems (PKS II) from the Gram-negative bacteria *Xenorhabdus* and *Photorhabdus*.

Particular attention was paid to a biochemically underexplored class of aryl polyene (APE) pigments. Bioinformatic analysis of enzymes involved in the biosynthesis and the *in vitro* reconstruction proved that the synthesis of APEs involves an unusual fatty acid-like elongation mechanism. Furthermore, the discovery of unexpected protein-protein interactions provided new insights into the multienzyme complex formation of this unusual PKS II system. Through collaboration with the groups from Prof. Michael Groll and junior Prof. Nina Morgner, two protein complexes were structurally solved and several native protein multimerization events were identified and allowed us to suggest a possible protein-interaction network. The results are summarized in publication '*An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments*' (first author; *J. Am. Chem. Soc.*).

In addition to *in vitro*-analysis, *in vivo*-studies were used to investigate the APE compound produced by *X. doucetiae* in more detail. The activation of the silent biosynthetic gene cluster (BGC) led to the detection of the APE compound in the homologous host. Further combination of homologous expression and targeted deletions of the APE BGC revealed an APE-lipid-like structure. MS-based analyses and purification of intermediates allowed us to deduce structural building blocks of the APE-lipid, which is composed of an APE structural core, a glucosamine residue and an unusual long-chain fatty acid with unusual conjugated double bonds and a phosphoethanolamine head group. In combination with the above stated *in vitro*-data, we assumed a plausible biosynthetic mechanism of the APE-lipid. The results are summarized in the section '*Additional Results: Tracing the Full-length APE*'.

The biosynthesis of isopropylstilbene (IPS) has already been well-studied by the Bode laboratory and the group of Prof. Ikuro Abe. Studies with *Photorhabdus laumondii* TT01 by the Bode group revealed the distributed locations and functions of the genes involved in biosynthesis, which originate from two pathways. Particularly, the Bode group first demonstrated that an unusual ketosynthase/cyclase (StlD) catalyzes the condensation of 5-phenyl-2,4-pentadienoyl-ACP and isovaleryl-beta-ketoacyl-ACP via a Michael addition. Such a pathway for stilbene formation is distinct from those widespread in plants. The Abe group solved the structure and biochemical mechanism of StlD and further investigated the aromatization reaction of the aromatase StlC. However, the generation of the required cinnamoyl-precursor 5-phenyl-2,4-pentadienoyl-ACP as a Michael acceptor for this cyclization reaction remained elusive. In this work, we were able to reconstitute the synthesis of the

Michael acceptor *in vitro*, by the action of enzymes from the fatty acid biosynthesis. With the knowledge about the crucial cross-talk from primary and specialized metabolism, we further determined the minimal endowment for stilbene production in a heterologous host. Here, the discovered AasS enzyme StlB is responsible for the generation of cinnamoyl-ACP and among others, *p*/FabH plays a key role as gatekeeper enzyme for further processing. With this information in hand, we were able to obtain IPS production in *E. coli*. These results are presented in the manuscript '*Biosynthesis of the Multifunctional Isopropylstilbene in Photorhabdus laumondii Involves Cross-talk Between Specialized and Primary Metabolism*' (co-first author, manuscript).

The biosynthesis of the orange-to-red-pigmented anthraquinones (AQs) is the best-studied type II PKS system according to preliminary results. While several investigations by Brachmann *et al.* discovered the BGC and the overall product spectrum of the main AQ-256 and its methylated derivatives, data of Quiqin Zhou (Bode group) performed biochemical *in vitro* analysis paired with *in vivo* heterologous expression of the *ant*-genes *antA-I*. This led to the identification of shunt products that indicated an AQ-scaffold derived from an octaketide intermediate that gets shortened to a heptaketide by the hydrolase AntI, resulting in the main anthraquinone AQ-256. This PKS-shortening mechanism was further confirmed by the protein crystal structure of AntI by the Groll group (publication, minor contributions, co-author, *Chem Sci*. 'Molecular Mechanism of Polyketide Shortening in Anthraquinone Biosynthesis of *Photorhabdus luminescens*'). Further substrate analysis of the *P. luminescens* AQ-producer and mutants revealed an inhibitory effect of cinnamic acid against the hydrolase AntI. Cinnamic acid might therefore be involved in regulation of AQ biosynthesis ('*Anthraquinone Production is Influenced by Cinnamic Acid*', first author, manuscript).

Biochemical analysis from Quiqin Zhou with the minimal PKS of the AQ-synthase further revealed the exclusive activation of the AQ-ACP by the PPTase AntB. The PPTase is insoluble alone but gets stabilized by the CoA-ligase, most likely inactive, working as a chaperone. Thus, the minimal PKS endowment to produce the octaketide scaffold comprises, besides the ACP, the KS:CLF heterodimer and the MCAT, the co-occurrence of the PPTase AntB and the CoA-ligase AntG. For the first time, X-ray crystallography depicted a minimal PKS in action, by obtaining the structural data of native complexes from an ACP:KS:CLF, the KS:CLF alone and an ACP:MCAT in their non-active and active forms. It was possible to confirm a KS-bound hexaketide, which was built upon heterologous expression of the KS:CLF. Mutagenesis with amino-acids proposed to be involved in protein-protein interactions in the ACP:KS:CLF complex revealed some interesting protein-interaction sites. Additionally, an induced-fit mechanism of the MCAT with the ACP during the malonylation reaction confirmed a

monodirectional transfer reaction (*Structural Snapshots of the Minimal PKS System Responsible for Octaketide Biosynthesis* co-author, manuscript under review).

5 Natural Product Research

In the course of evolution ranging over millions of years, microorganisms developed a vast diversity of fascinating and highly complex molecules, awaiting their identification by natural product research. While their physiological role is often unknown, they reveal a great range of useful bioactivities.^[1-5]

In a broad sense, the term 'natural product' covers all molecules produced by a living organism. It encompasses different compound classes crucial for life, including carbohydrates, proteins, fats, and nucleic acids.^[6] These metabolites are assigned to primary metabolism.

Natural product research deals most likely with molecules from the secondary metabolism, also referred to as specialized metabolism, with a molecular weight of <1500 Da.^[7] In contrast to the essential primary metabolites, secondary metabolites, are only found in specific organisms or groups of organisms.^[6] Secondary metabolites are known to provide benefit to the organism that results in an evolutionary advantage, e.g. by acting as toxic agents to defend the organisms against predators or acting as volatile attractants towards the same or other species.^[6] Although the physiological function of the produced secondary metabolites is not always known, they exhibit useful biological activities that result in their pharmacological use.^[1-5]

The main representatives of secondary metabolite classes in microorganisms used for natural product research include polyketides, non-ribosomal peptides, alkaloids, and terpenes.^[7] They all can be assigned to different metabolic pathways, consisting of a combination of genes within a contig, termed biosynthetic gene clusters (BGCs).^[8] These BGCs usually encode enzymes responsible for synthesis and modification, proteins for transport, regulation, and detoxification. The paradigm of natural product research began with the discovery of penicillin by Alexander Fleming in 1928.^[9,10] Due to contamination with a fungus on his bacterial agar plate, Fleming observed an inhibition zone killing the bacteria around the fungal colony.^[10] The natural product responsible for this effect was termed penicillin, after the producing fungus *Penicillium notatum*. Based on its structure, it was classified as a β -lactam antibiotic and emerged as one of the most important antibiotics to treat bacterial infections. The discovery of penicillin is a prelude to the 'Golden Age of Antibiotic Discovery' era, bringing on a tremendous number of antibiotics discovered from a plethora of different host organisms, such as bacteria, fungi or plants.^[5,7] Besides antibacterial or antifungal activities, further pharmaceutical applications of natural products from microorganisms were discovered, such as in cancer therapy, as immunosuppressive drugs, as cholesterol-lowering agents or anesthetics (Figure 1).^[1-5,11]

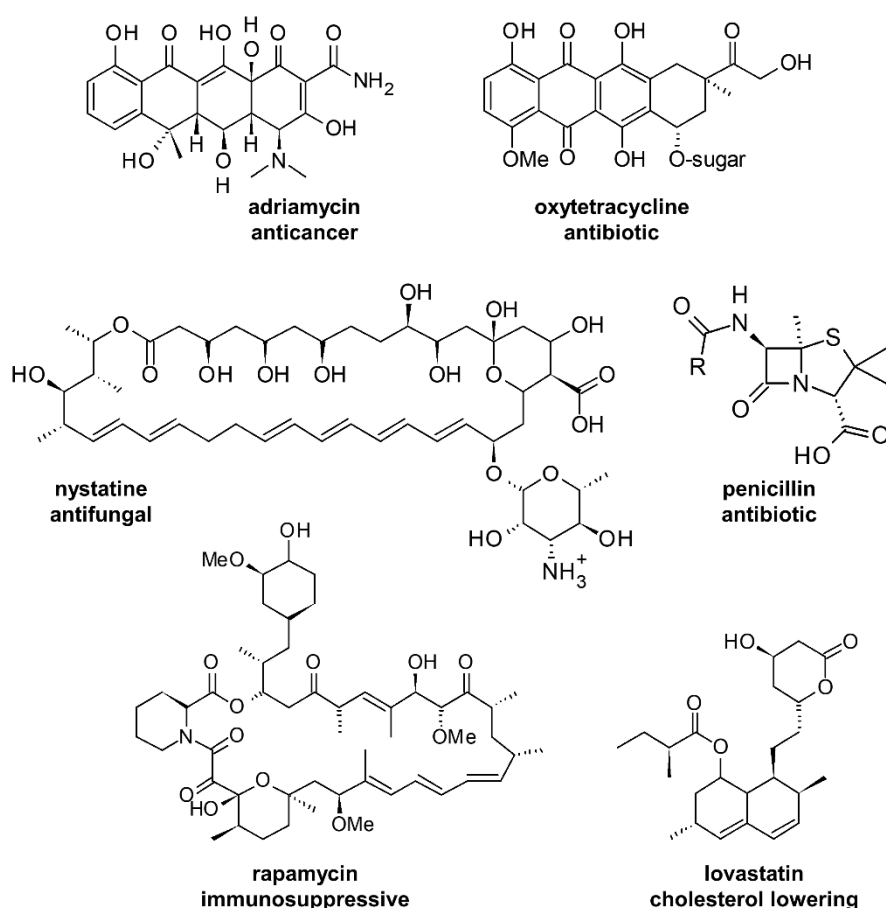


Figure 1. Penicillin and representative polyketide natural products with their pharmaceutical applications.^[7]

In recent years, the progress in bioinformatics enabled the efficient detection and prediction of distinct BGCs to their corresponding natural product classes, a process termed genome mining.^[12,13] Genome mining can be useful in terms of the vast number of genomic data sets that are accessible in the post-genomic era.^[13]

In times of multi-resistant organisms, there is a high interest to identify new active compounds and to avoid studies with already known ones. Computer-assisted techniques^[8,13,14] combined with advances in high-throughput cultivations have enabled networking with mass spectrometry (MS)-based data, to generate huge metabolic and proteomic data sets.^[15–19] Only recently, such data sets have been linked to genomics.^[20] The overall aim is to create publicly accessible data sets and suitable methods for simplifying natural product research.^[15]

In the above-stated approaches, among others, genetics plays an indispensable role in revealing the programming rules of the different classes of natural products. However, a thorough understanding of the catalytic mechanisms of the involved enzymes still requires the combination of *in vivo* and *in vitro* analysis, to characterize the biosynthesis of the desired compounds.^[21,22]

Once we understand the mechanisms and enzyme interaction network in detail, this will facilitate attempts to control and re-task the biosynthetic systems to produce novel compounds.

6 The Logic of the PKS and FAS Thiotemplate Reactions

Polyketides and non-ribosomal peptides are built upon simple acetate or amino acid building blocks.^[7] Their biosynthesis involves an enzyme machinery to which the different intermediates are covalently attached and shuttled to the active sites of the corresponding enzymatic domains.^[23] Such a reaction line is also known for the biosynthesis of fatty acids.

The natural products investigated in this study originate from entomopathogenic bacteria of the species *Xenorhabdus doucetiae* DSM 17909^T (*X. doucetiae*)^[24] and *Photorhabdus laumondii* subsp. *laumondii* TT01 (*P. laumondii* or TT01, renamed from previous *P. luminescens* subsp. *laumondii* TT01)^[25] and are produced by polyketide synthases (PKS) from specialized metabolism. Due to the similarity of PKS and fatty acid synthases (FAS) involved in primary metabolism, both systems will be presented in more detail in the following sections.

6.1 Biosynthetic Mechanism of Polyketides and Fatty Acids

PKS and FAS machineries use a thiotemplate mechanism to shuttle their substrates in an effective manner, with the help of so-called acyl-carrier proteins (ACPs).^[26] ACPs are small molecular weight proteins of about 8-10 kDa.^[7] They possess no catalytic activity but are post-translationally modified with the 4'-phosphopantetheine (Ppant) arm derived from Coenzyme A (CoA) (Figure 2).^[27] This modification results in a terminal thiol which forms the attachment site for substrates and intermediates and consequently builds a thiotemplate.

The reaction for obtaining the Ppant arm to form an active *holo*-ACP is catalyzed by a 4'-phosphopantetheinyl transferase (PPTase) (Figure 2).

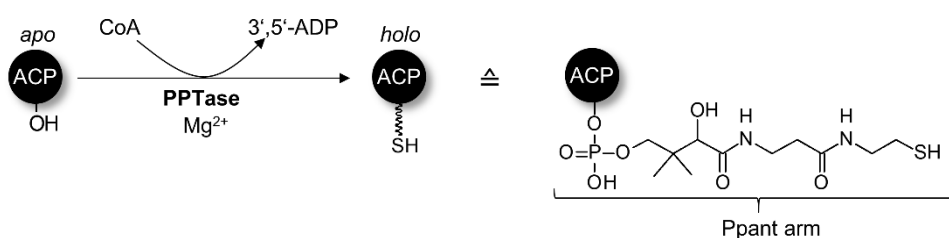


Figure 2: Posttranslational activation of ACPs. The inactive *apo*-ACP is activated to its *holo*-form by transferring the Ppant moiety of Coenzyme A to a conserved serine residue by a $MgCl_2$ -dependent phosphopantetheinyl transferase (PPTase) reaction.

Although fatty acids and polyketides are structurally different, from a biochemical perspective, the underlying biosynthetic mechanisms for the formation of their backbone are very similar (Figure 3).^[23,27,28]

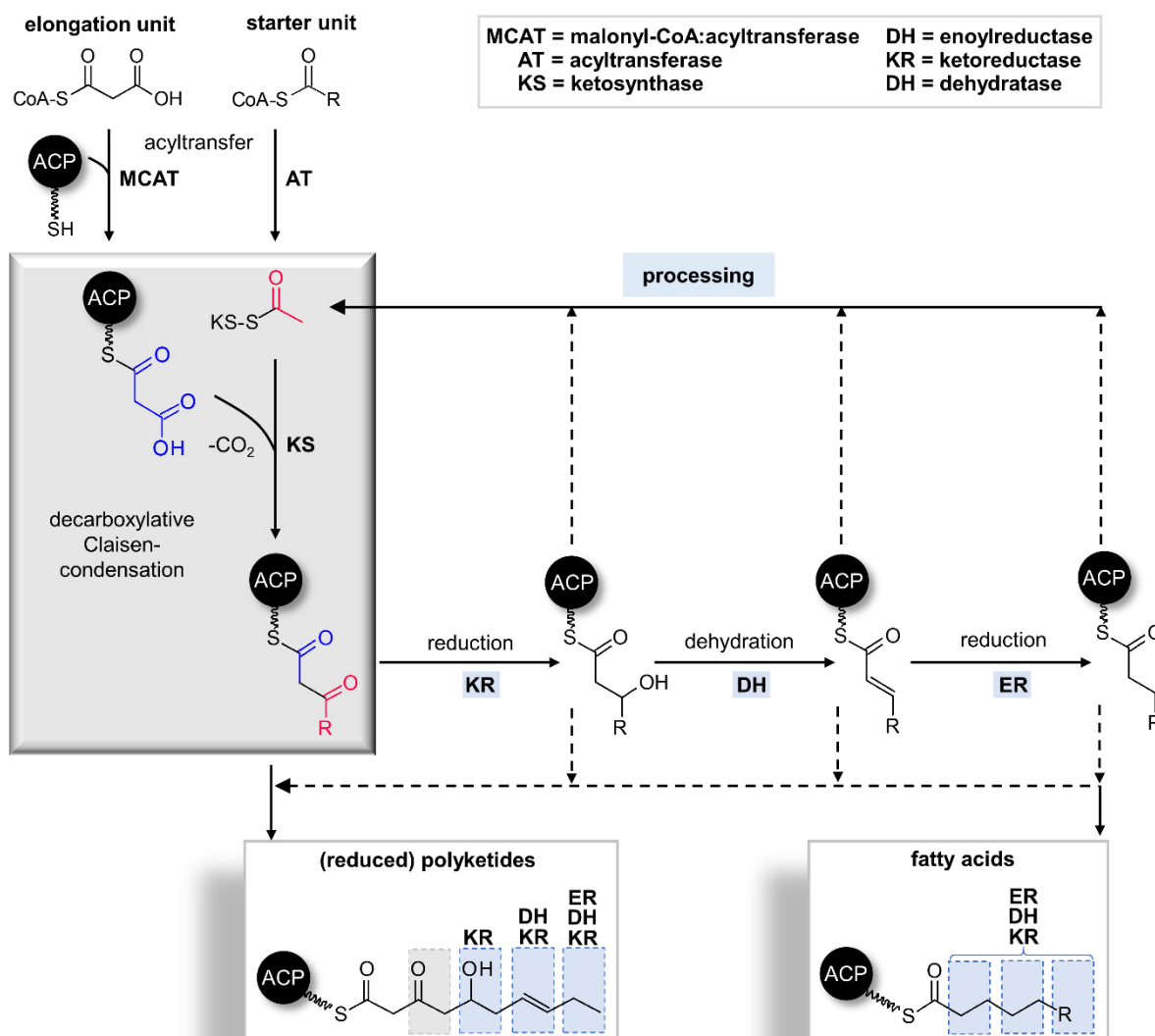


Figure 3. Polyketide and fatty acid mechanism during biosynthesis. Malonyl-CoA and acetyl- or acyl-CoA (represented with R) from primary metabolism are connected by a ketosynthase (KS) in a decarboxylative Claisen-condensation reaction to result in a β -ketoacyl ACP. This product gets further processed by a ketoreductase (KR), a dehydratase (DH) and an enoylreductase (ER) to result in a fully saturated acyl product during fatty acid biosynthesis. These processing steps can be partially omitted in the case of polyketide biosynthesis.^[28]

The biosynthesis of fatty acids and polyketides is typically conducted with acetate and malonate building blocks derived from primary metabolism.^[23,27,28] Those building blocks are loaded onto the ACP for acyl-chain elongation and further processing through the action of different enzymatic domains (Figure 3).

The biosynthesis starts with the loading of the starter ac(et)yl-group. This starter unit gets coupled to the *holo*-ACP by an acyltransferase (AT) and is further transferred to a conserved

cysteine residue of the adjacent ketosynthase (KS). This intermediate gets connected with the ACP-tethered malonate extender unit. Malonyl-ACP is derived by the malonyl-CoA:ACP transacylase (MCAT) with the use of *holo*-ACP and malonyl-CoA.^[27,29] The elongation takes place in a decarboxylative Claisen-condensation reaction to form the ACP-bound β -ketoacyl product, while decarboxylation is mediated by the KS (Figure 3, grey block).^[30] This product is further shuttled to the active centers of an NADPH dependent ketoreductase (KR), resulting in a β -hydroxy acyl product, which gets passed over to a dehydratase (DH) to form the corresponding enoyl product. Finally, this intermediate is transported to an NADH-dependent enoyl-reductase (ER) to result in a saturated acyl product. In FAS systems, this cycle is fully executed and iterated until the acyl product reaches its final length.^[23]

Regarding PKS systems, the reductive steps (KR, DH, ER) are optional and are sometimes partially omitted, before the next elongation step is conducted. This leads to a much more complex product pattern for polyketides.^[23] As a final step, the ACP-tethered acyl products are typically released by a thioesterase (TE) to form free acids or macrolactones (PKS only).^[23] Throughout the biosynthesis, the substrates and intermediates are ACP-tethered or covalently linked to the AT (conserved active site serine) and the KS (conserved active site cysteine). In contrast to fatty acids, polyketides show a huge structural diversity.^[23] As mentioned above, this is, on the one hand, explained by the partially incomplete elongation cycles, leading to different reduction states, on the other hand, by a selection of different starter and extender units (see 6.2.1). Further variations include different chain-lengths, folding and termination reactions and are dependent on the different protein architectures of PKS but also FAS systems.^[23] This leads to different classifications of the synthases. This classification will be discussed in more detail in the following chapters. Additionally, there are many post-assembly line tailoring events such as glycosylation, acylation, alkylation and oxidation, altogether resulting in an enormous structural diversity of PKS-derived products.^[23,31]

6.2 Classification of PKS and FAS

Although PKS and FAS systems rely on the same principal reaction mechanism as shown in Figure 3, huge differences in the arrangement of enzymes and product spectrum can be found (Figure 5-7). These differences include, among others, the overall protein structure and an iterative use of catalytic domains. Accordingly, PKS systems can be divided into three main classes, while FAS can be divided into two.^[23,32] In this section, the main differences of the classes of both systems are summarized, chapter 6.3 deals with insights in the overall protein structure and protein-protein interactions within those systems.

6.2.1 PKS

The overall classification of PKS types is based on their protein-arrangement, the resulting product and the host organism (Figure 4).^[23,28,33] Type I PKS are multi-modular and due to dimerization, they form a multifunctional enzyme complex that produces predominantly reduced linear or cyclic polyketides. This type is widespread in bacteria and fungi.^[34] In contrast, type II PKS consists of monofunctional enzymes, postulated to form a multienzyme complex, resulting in aromatic polyketides. Type II systems are only found in bacteria.^[35] Type III systems also consist of monofunctional enzymes but without ACPs. The reaction products derived from type III PKS are small aromatic compounds usually found in plants.^[36]

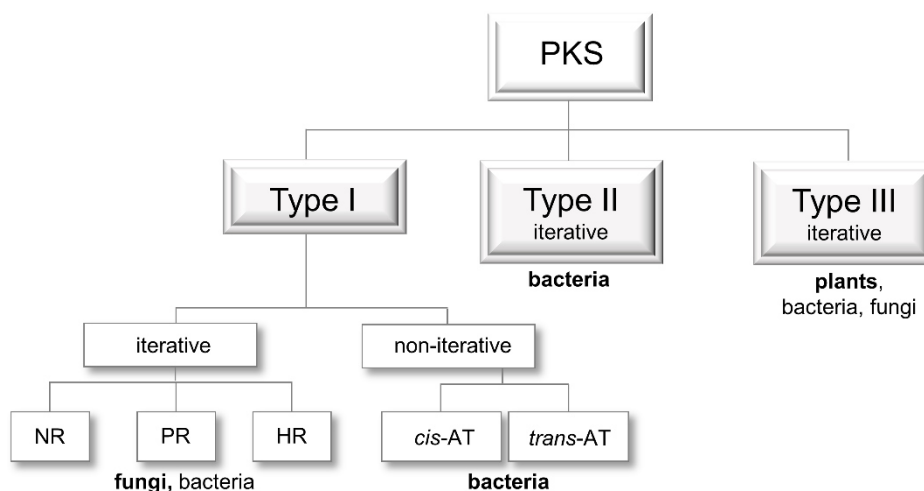


Figure 4. Different types and subtypes of PKS systems. Organisms in bold are the main representatives of the corresponding PKS system. NR = non-reducing, PR = partly reducing, HR = highly reducing.^[23]

6.2.1.1. PKS Type I

PKS I systems consist of modules. A module is (part of) one polypeptide chain, which harbors the different enzymatic domains for the processing of one polyketide building block. The vast majority of modular PKS I systems are known to act in a collinear fashion, in a non-iterative way and are consequently well-known as assembly lines.^[37–39] Thus, it might be possible to predict PKS I-derived products according to their modular organization equipped with the enzymatic domains. These modules can be divided into the loading module that enables the covalent linkage of the ac(et)yl-starter unit and several extension modules that are responsible for loading the malonate, or derivative thereof, in each elongation cycle (Figure 5b).

The final module (termination module) is typically terminated by a TE-domain to release the final polyketide product via hydrolysis or macro-cyclization.^[40,41]

To date, four different loading modules are well-known for PKS I systems, which shuttle acetate, propionate, or in rare cases, aromatic starter units (Figure 5a).^[42]

For instance, the 6-deoxyerythronolide B synthase (DEBS), responsible for the biosynthesis of erythromycin, contains a loading module composed of an acyltransferase (AT) and an ACP (Figure 5a,c). Here, a propionyl-CoA is loaded by the so-called AT_L (the loading AT) to the adjacent ACP (loading ACP, ACP_L, loading module: AT_L-ACP_L) followed by a further transfer to the conserved cysteine, located on a KS of the next extension module for further processing (Figure 5a, first loading mechanism).

However, the vast majority of other modular PKS possess an additional KS-domain in their loading module.^[42] This KS-domain has a mutation of the active site cysteine, which is replaced by a glutamine and therefore known as KS_Q (loading module: KS_Q-AT_L-ACP_L).^[43] This KS_Q-domain is unable to fulfill a condensation reaction but still owns its decarboxylation activity, as it is well known for the production of monensin molecules (Figure 5a, second loading mechanism).^[43] Hence, the acyl moiety of the malonyl-CoA or methylmalonyl-CoA starter unit is transferred to the ACP_L by the AT_L and decarboxylated by the KS_Q, before moving to the KS-domain of the extension module.

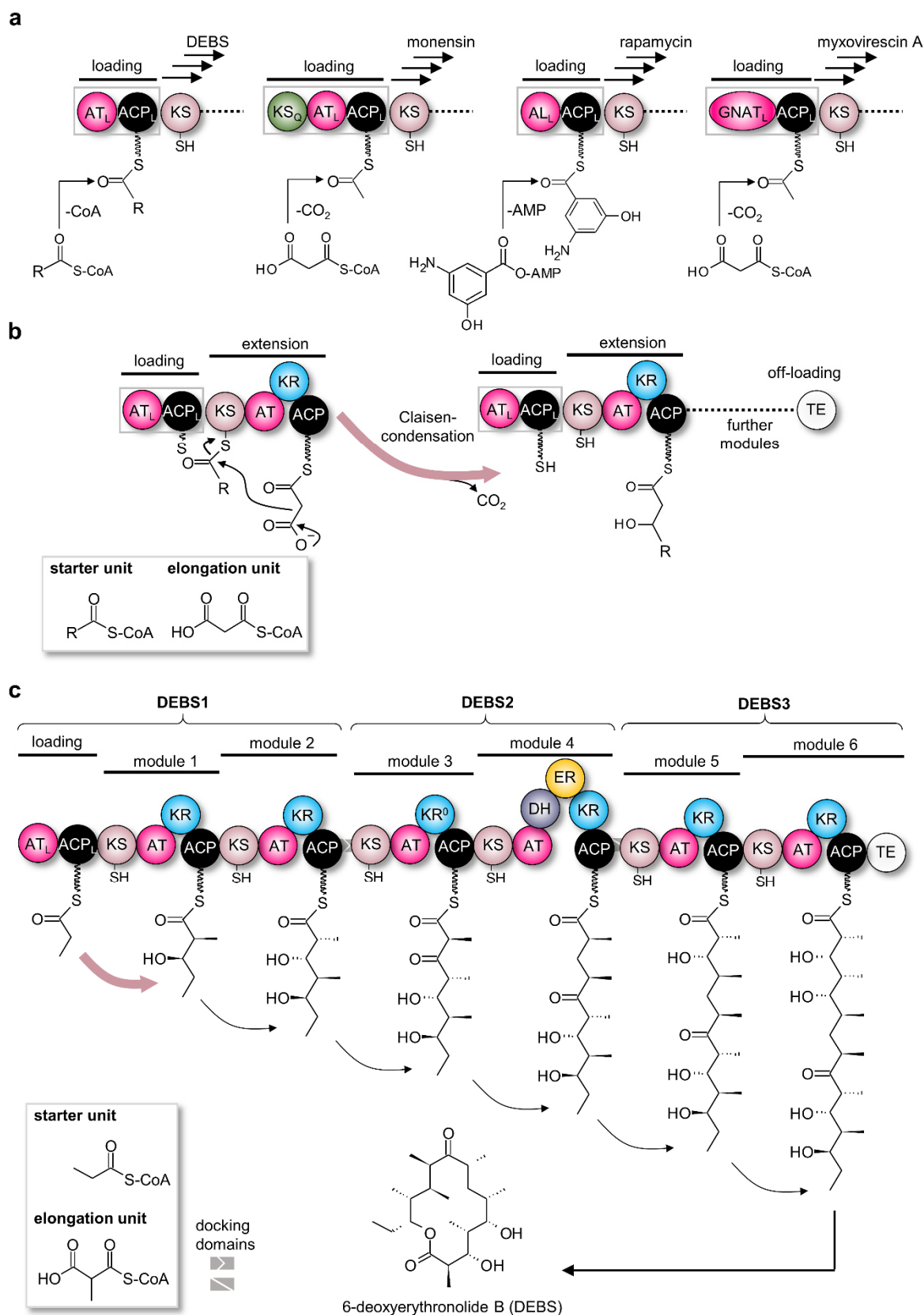


Figure 5. Type I PKS. Different loading mechanisms in PKS I systems (**a**). KS-domain catalyzed Claisen-condensation reaction (**b**). Logic of PKS I exemplified by the domain organization of the 6-deoxyerythronolide B synthase (DEBS, three proteins and seven modules). The proteins are connected by docking domains; catalytic domains are indicated by their corresponding label. The KR⁰-domain of DEBS module 3 has lost its reductive activity (**c**).^[44]

The third loading mechanism (Figure 5a) was well-described for the biosynthesis of rapamycin and uses a mechanism where free acids are adenylated and loaded subsequently by the AT-domain.^[45] This mechanism is analogous to the activation mechanism of adenylation-domains, the counterparts of PKS AT-domains in the biosynthesis of non-ribosomal peptides.^[38] This adenylation domain is therefore referred to as AD and acts in combination with a neighboring ACP, giving rise to the loading module AD_L-ACP_L. Generally, synthases using this loading module incorporate partially oxidized cyclohexane-carboxylate or aromatic units, derived from the shikimate pathway (see also 6.2.3).^[46,47] These precursors are further modified with hydroxyl or amino side chains, for example, as demonstrated in rifamycin synthesis.^[48]

The loading module of the myxobacterial polyketide myxovirescin A uses an additional, so-called GNAT-domain^[49], giving rise to GNAT_L-ACP_L-loading modules. GNAT represents an *N*-acyltransferase, belonging to enzymes that are typically responsible for an acyl transfer to a primary amine. To this date, GNAT_L-domains in PKS loading modules have been shown to decarboxylate malonyl-CoA to acetyl-CoA prior to ACP-loading, without the use of an AT_L-domain. This reaction mechanism differs from the decarboxylation and transfer mechanism in KS_Q-AT_L-ACP_L-loading modules. (Figure 5a, fourth loading mechanism).

An extension module of PKS I typically consist of KS-AT and the optional reduction domains KR, DH and ER.^[33] Additionally, some PKS extension modules consist of a C-methyltransferase-domain (MT) for methylation of the β -keto polyketide intermediate in α -position.^[50] Extension modules in type I systems typically load malonates but compared to the other PKS types, they are more promiscuous to malonate derivative extender units such as methyl-, ethyl- or chloromethyl-malonate.^[42] The acyl moieties can be directly loaded from their corresponding CoA ester onto the ACP by the AT of the extension module.^[51]

Other examples, such as methoxy-, hydroxy- or amino-malonate, are synthesized by modification enzymes on an ACP-tethered acyl unit.^[51] Genes encoding modification enzymes are often found within or flanking the associated BGC, while the CoA-linked extender units are used from primary metabolism.^[51]

Non-iterative PKS I can be further subdivided into *cis*- and *trans*-AT-PKS (Figure 4). *Trans*-AT-PKS I systems describe all systems with a *trans*-acting AT-domain, while *cis*-PKS I systems describe the conventional arrangement. Interestingly, the *trans*-AT-PKS has a flanking subdomain (FSD) appended to the C-terminus of the KS-domain. It is argued if that helical motif might be either important for protein-protein interaction of the *trans*-AT with the corresponding module or to serve as 'spacer' to provide room for AT-interaction.^[52] To date, *trans*-AT PKS I are exclusively found in bacteria.^[53]

In contrast to non-iterative PKS I, iterative ones are a fungi hallmark,^[54] which also do not follow the principle of collinearity. Here, the modules are used repeatedly with the degree of the reduction catalyzed by the KR-, DH- and ER-domains varying in the extension cycles. This mechanism is not yet fully understood, hampering the prediction of the resulting polyketide product.^[23]

According to the presence of the reduction domains, fungal PKSs are further sub-classified into nonreducing (NR), partially reducing (PR), or highly reducing (HR) PKS (Figure 4).^[23] The bacterial counterparts of non-iterative PKS I systems are rare. Usually they consist of monomodular PKS I systems, responsible for producing small aromatic compounds that result from linear polyene intermediates, such as enediyne.^[55,56]

6.2.1.2. PKS Type II

Type II PKS systems do not use modules equipped with the enzymatic domains, but rather consist of free-standing, monofunctional enzymes that act in an iterative manner (Figure 6). Currently, type II PKS systems are only found in Gram-positive bacteria, such as *Streptomyces*^[35] and only a few examples exist in Gram-negative bacteria (see section 6.4). The products of a type II PKS system are mostly aromatic.^[23] Typically, a malonyl-CoA-derived malonyl-ACP serves as a starter unit, whereas the MCAT is shared with FAS II systems^[57] and transfers the acyl moiety from malonyl-CoA to the *holo*-ACP to create the malonyl-ACP. The exclusive use of malonyl-CoA as an extender unit is controlled by the stringent substrate specificity of MCAT.^[58] The malonate is further transferred to the conserved cysteine of the elongation KS_{α} . KS_{α} forms a heterodimer with KS_{β} , also known as chain-length factor (CLF).^[59] The CLF acts similar to the KS_Q in the loading module of type I PKS (see Figure 5) to generate the acetyl-starter unit. This starter unit is elongated with a newly loaded malonyl-ACP to result in the first β -ketoacyl product, which is again transferred to KS_{α} , for repeated elongation cycles (Figure 6 or Figure 17).^[59]

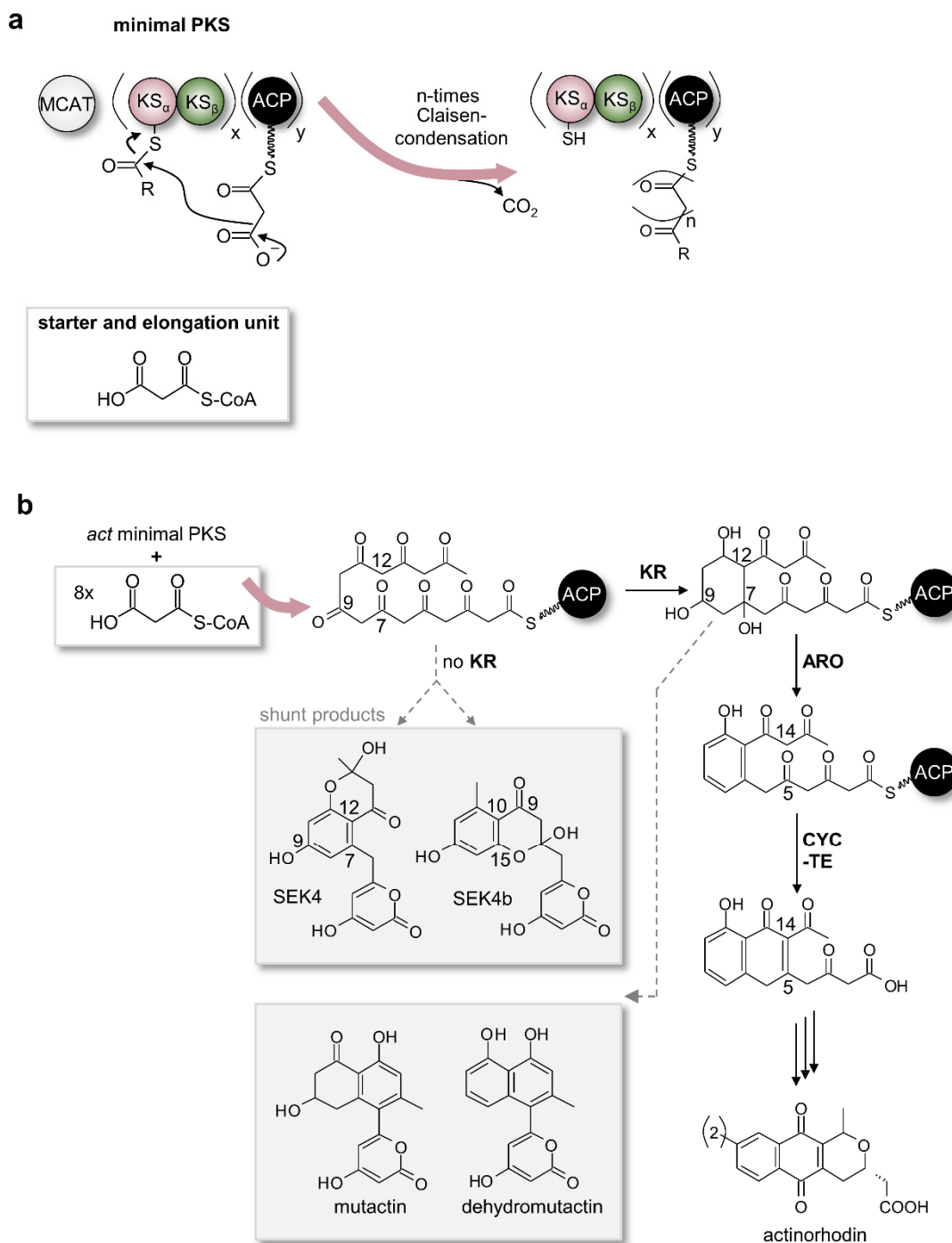


Figure 6. Type II PKS. Chain elongation reaction is conducted with malonate starter- and elongation units in a decarboxylative Claisen-condensation reaction mediated by the KS, which in turn forms a heterodimeric complex with the CLF. Together with the ACP and the MCAT, from primary metabolism, the minimal PKS is formed (**a**). The aromatic scaffold is formed by further interplay with the KR, ARO and CYC, which is exemplified by the actinorhodin-synthase. Several shunt products (SEK) can be built upon missing processing enzymes (KR, ARO, CYC) (**b**).^[60]

KS_α catalyzes malonyl decarboxylation of the elongation unit by the aid of two conserved histidines.^[61] The lack of any further reduction enzyme at this point in time makes products of type II PKS the quintessential polyketide, as the resulting ACP-attached intermediate forms a highly reactive polyketone product.^[7]

This intermediate is stabilized in the amphiphilic cavity of the heterodimeric $\text{KS}_\alpha\text{:KS}_\beta$ (further KS:CLF).^[61] The CLF is postulated to be further responsible for chain-length control, resulting in typical chain-lengths of 16 (octaketide), 20 (decaaketide) or 24 (dodecaacetide) carbon atoms.^[23,61] This corresponds to 8, 10 or 12 extension cycles for acetate, which is typically used as a starter unit. During *in vitro* investigations with non-acetate starter units, it was shown that the synthase adapts the number of carbons, corresponding to the original length, concluding that the KS:CLF -cavity ‘measures’ the product length instead of ‘counting’ the cycles.^[57] Several non-acetate priming mechanisms are described under 6.2.3.

KS:CLF and ACP (plus MCAT) are well-known as ‘minimal PKS’ and the encoding genes are typically found next to each other in one open reading frame (compare Figure 17 and Figure 21).^[35] In the following step, the polyketone is cyclized and further reduced in C_9 -position by a cyclase (CYC) and a KR, respectively, prior to be aromatized by an aromatase (ARO) (Figure 6). As a result, the polyketide is cleaved off and the overall aromatic structure is formed.^[35] The first cyclization reaction usually takes place between $\text{C}_7\text{-C}_{12}$ (actinorhodin) or $\text{C}_9\text{-C}_{14}$ ^[60] (see Figure 6). Despite the widely long-held assumption that reduction takes place prior to ring cyclization, recent studies on docking experiments with cyclic substrates show that the first ring is necessary to position the intermediate for the regioselective C_9 -reduction.^[62–65]

In the absence of the KR (and other tailoring enzymes), that exclusively reduces the C_9 position, the spontaneous reaction products (shunt products) SEK4 and SEK4b are synthesized. These products are the result of mis-cyclization events of the ACP-bound octaketide in actinorhodin biosynthesis,^[60] which is based on eight elongation cycles (Figure 6). Two other shunt product pairs are known to be expected, depending on the used minimal PKS (10 or 12 cycles).^[57] Upon heterologous production of the actinorhodin gene cluster, two additional shunt products, mutactin and dehydromutactin, are synthesized after cyclization and might occur due to the non-effective interplay of the enzymes within a heterologous host.^[66] However, such control mechanisms are still under investigation.

In most of the cases, if not all, the aromatic core structure is further processed by tailoring enzymes, resulting in methylation, oxygenation or glycosylation events.^[31]

Due to this immense diversity, product prediction of PKS II systems can be challenging and sometimes impossible, without preliminary analysis of the product spectrum derived from the ‘minimal PKS’.

6.2.1.3. PKS Type III

Type III PKS systems are commonly found in plants, with fewer occurrences in bacteria and fungi. In plants, they are named after their small aromatic compounds (typically mono- or bicyclic aromatic products) as stilbene (STS) or chalcone synthases (CHS) (Figure 7).^[23,36,67,68]

Although PKS III are monofunctional enzymes, similar to PKS II systems, they lack an ACP.^[67] Hence, the biosynthesis works throughout the interplay of CoA-derivatives and the homodimeric KS, which catalyzes the priming, extension, and cyclization reactions to form the polyketide intermediates. Typical starter unit representatives are *p*-coumaroyl-CoA and *trans*-cinnamoyl-CoA, which are derived from tyrosine or phenylalanine by an ammonia lyase and a CoA-ligase. The aromatic acyl moiety of the starter unit is attached to the KS and exclusively elongated with malonyl-CoA extender units. This step is sequentially repeated and results in a polyketide chain intermediate, which is formed and cyclized via intramolecular aldol cyclization within the active site cavity of the KS (Figure 7).

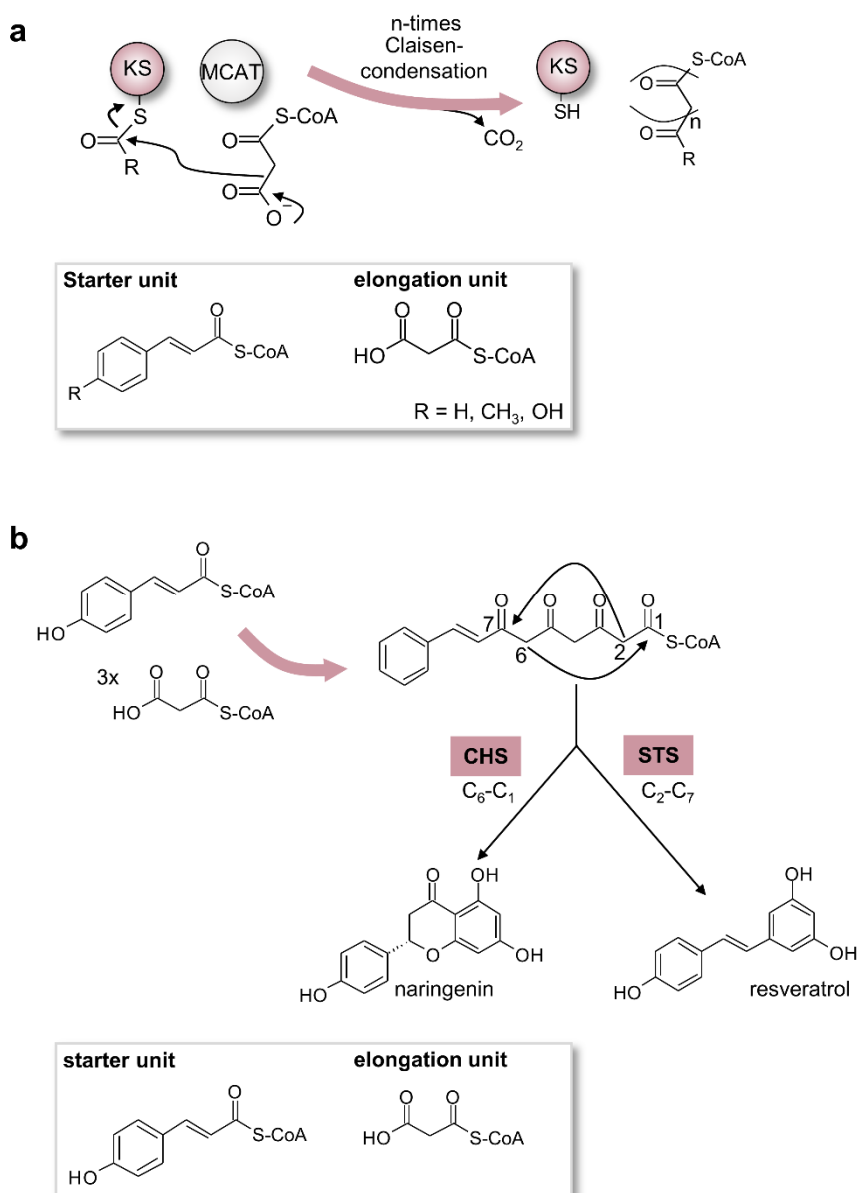


Figure 7. Type III PKS. Decarboxylative chain-elongation reaction takes place with aromatic starter- and malonate extender units by a homodimeric KS (**a**). Synthase-specific different products are built. Exemplified, a C₆-C₁ cyclization is performed by a chalcone-synthase (CHS) and leads to naringenin, a C₂-C₇ cyclization is catalyzed by a stilbene-synthase (STS) and leads to resveratrol, both produced in plants (**b**).^[36]

Depending on the PKS, the cyclization reaction results in either naringenin (CHS, C₆/C₁-cyclization) or resveratrol (STS, C₂/C₇-cyclization), the latter one when *p*-coumaroyl-CoA is used as a starter, instead of cinnamoyl-CoA (Figure 7).^[36]

6.2.2 FAS

Comparable to PKS systems, FAS systems can be distinguished into two types. Type I FAS are large megaenzyme complexes consisting of multiple catalytic domains, comparable to PKS I systems. All FAS I systems act in an iterative manner. This kind of system is found in mammals, fungi and corynebacterial.^[69] In contrast, type II FAS consists of monofunctional enzymes^[65,70,71], comparable to PKS II systems, and have been found in bacteria, plants, parasites and eukaryotic mitochondria.^[23,70,72]

Typically, a microorganism is equipped with only one type of FAS, which is located in the cytoplasm, while an exception is found in mycobacteria. Mycobacteria do have both systems, type I and type II FAS, enabling the production of very-long chained mycolic acids that are used to form their characteristic cell envelope (see also Figure 38).^[73]

In general, a type I FAS produces C₁₄-C₁₆ saturated fatty acids.^[74] Type II systems are more flexible and produce different chain-length, unsaturated, iso- and anteiso-branched and hydroxy fatty acids.^[70]

6.2.2.1 FAS Type I

FAS I are well-known to form reaction chambers, which can be subdivided in loading/unloading (MAT/MPT, ACP, TE), condensing (KS) and modification (KR, DH, ER) compartments upon domain distribution (Figure 8) (for more structural details, see section 6.3). FAS I systems are primed with acetyl-starter units, loaded by a malonyl-acetyl transferase (MAT, loads malonyl- and acetyl-starter units)-domain or in the case of fungal FAS a malonyl-palmitoyl transferase (MPT)-domain. The ACP-tethered cargo is elongated with malonate extender units by the KS-domain and reduced, dehydrated and reduced again with the processing domains (KR, DH, ER) as it was described in more detail under 6.1.^[75] Corresponding to the host, enzymatic release of the ACP-bound acyl moiety is mediated by an MPT/MAT- or a TE-domain.^[75,76]

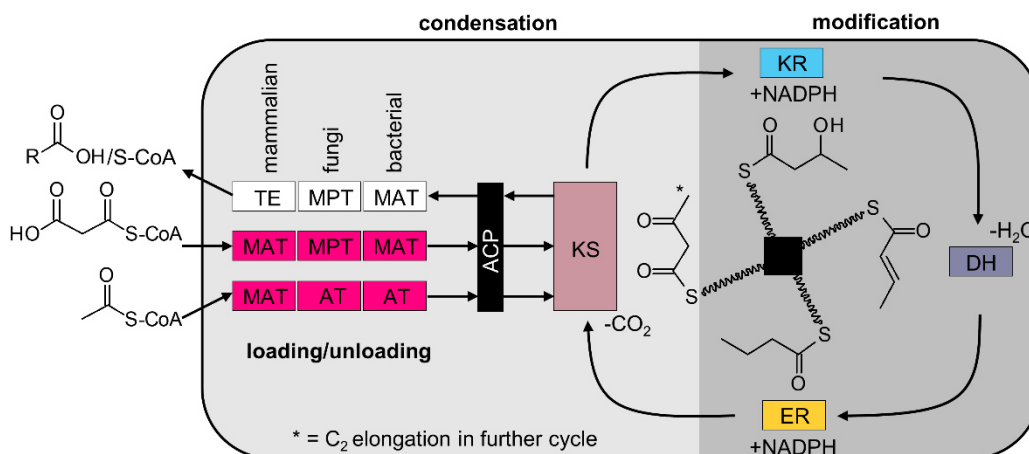


Figure 8. Type I FAS reaction chamber. The synthase is grouped in a condensation and modification compartment, equipped with the enzymatic domains, including the mobile ACP-domain. The reaction order is explained in the main text. The figure was modified according to reference [75].

6.2.2.2. FAS Type II

In contrast to type I systems, the reaction in FAS II involves monofunctional enzymes. Most of them are present in multimeric forms.^[65,70] This includes, among others, the three homodimeric KS FabH, FabB and FabF.

FabH is responsible for priming the *holo*-ACP with the ac(et)yl-CoA starter units. Depending on these starter units, different fatty acids are produced (compare 6.2.3). FabB and FabF are elongating KSs, which can iteratively condense acyl-ACP intermediates with malonyl-ACP, generated from malonyl-CoA by the MCAT (FabD). Malonyl-CoA, in turn, is synthesized from acetyl-CoA, which is carboxylated by an acetyl-CoA:carboxylase (ACC) (Figure 9). FabB elongates short-chain fatty acids from C_4 up to C_{16} and FabF typically from C_{16} to C_{18} .^[77] NADPH-dependent ketoreduction of the β -ketoacyl product is conducted by the tetrameric KR FabG^[78], prior to dehydration, which is either performed with the homodimeric DH FabA or the hexameric DH FabZ to result in *trans*-enoyl-ACP. FabA eliminates water from β -hydroxy decanoyl-ACP to enable the isomerization of *trans*-2-decenoyl-ACP for generating the *cis*-form.^[65] FabZ is not able to perform the isomerization reaction and prefers short ($\sim\text{C}_4$ - C_{12}) and longer chain substrates ($\sim\text{C}_{14}$ - C_{16}).^[79]

Mono-unsaturated fatty acids are produced by the interplay of FabA, FabB and FabF. Here, FabA produces *trans*-2-decenoyl-ACP and FabB further elongates it. To maintain the unsaturation, the ER step is omitted. FabF further elongates this 'long-chain'-substrate due to its C_{14} - C_{16} specificity.^[65] The fatty acid cycle is completed by the NADH-dependent enoyl reduction, catalyzed by the tetrameric ER FabI or FabK (Figure 9).

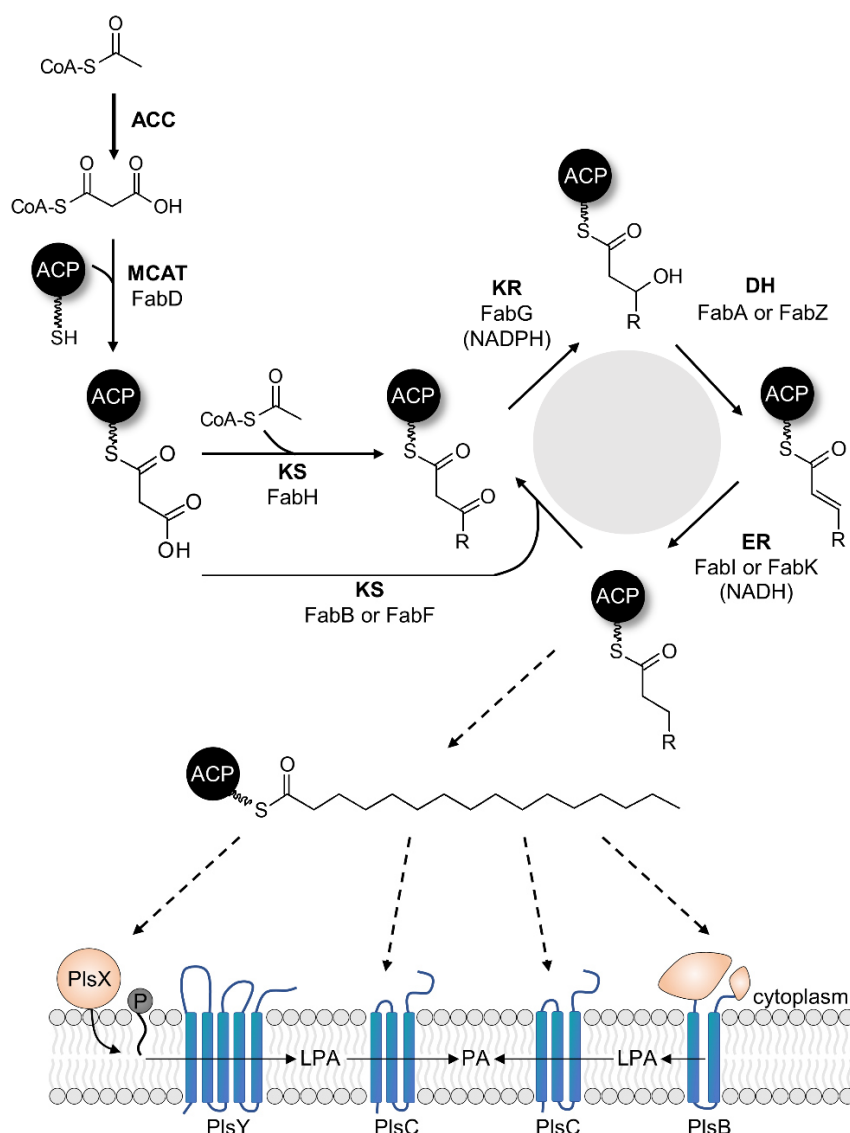


Figure 9. Type II FAS. Corresponding reactions are explained in the main text. LPA = lysophosphatidic acid PlsX = acyl-ACP phosphate acyltransferase, PlsY/PlsB = glycerol-3-phosphate (G3P) acyltransferase.

If not released by an MPT/MAT- or TE-domain, acyl-ACPs are also used for the biosynthesis of phospholipids (Figure 9). Hereby, the acyl-ACP serves as a starter to form 1-acyl-glycerol-3-phosphate by the use of different acyltransferases from the PlsX-Y pathway. This pathway is the most ubiquitous system in bacteria.^[80] The acyl-group from acyl-ACP can be either directly transferred to glycerol-3-phosphate (G3P) by PlsB, or first activated to an acyl phosphate by PlsX and then transferred to PlsY. In both cases, 1-acyl-glycerol-3-phosphate, also known as lysophosphatidic acid (LPA), is formed. The second position of the LPA is further acylated by PlsC to form phosphatidic acid (PA), which in turn forms the central building block in phospholipid biosynthesis.^[81] To this central building block, different polar head groups are attached,^[80] thus, resulting in cardiolipids, phosphatidylethanolamides and phosphatidylglycerides, which are the major lipids found in Gram-negative *E. coli*.^[82]

6.2.3 Comparison of Non-Acetate Priming Reactions in Type II Systems

The vast majority of polyketides derived from PKS II systems are malonyl-derived acetate starter units, as described in chapter 6.2.1.2.^[46] Nevertheless, an array of non-acetate examples have been described over the last few years, including straight- or branched-chain fatty acids, ω -alicyclic, aromatic and amino acids, with all of them playing a considerable role in product diversity.^[46,83]

To date, three main strategies for non-acetate priming in PKS II systems are known (Figure 11). The first strategy is based on an additional interplay of so-called KAS III-enzymes to load fatty acid precursors. The second strategy gives rise to an alternative route for loading propionate starter units. The third strategy includes the loading of acyl moieties by an acyl-CoA-ligase, in combination with an AT, to load aromatic or amino acid starter units (Figure 11).

In the previous section, the KAS III-enzyme FabH was presented to be in charge of the initial elongation reaction in FAS II systems. Here, the FabH is not only responsible for the starter unit selection but also for condensing of acetyl-CoA and malonyl-ACP units. Thus, this reaction includes transacylation, decarboxylation and condensation activity.^[84] Depending on its substrate specificity, FabHs can use different non-acetate starter units for priming. Besides acetyl starter units, also the use of propionyl- and butyryl-starter leads to straight-chain fatty acids. The use of isopropyl- isobutyryl and methylbutyryl-CoA, in contrast, results in branched-chain fatty acids (Figure 10).^[84,85] Iso-branched precursors are most-likely derived from the branched-chain α -keto acid dehydrogenase (BKD)-pathway, which creates a CoA-activated iso-branched precursor generated from valine, leucine and isoleucine (Figure 10).^[86,87] ω -alicyclic starter units that can be used by a KAS III are derived from multiple reactions from shikimic acid and products thereof.^[47]

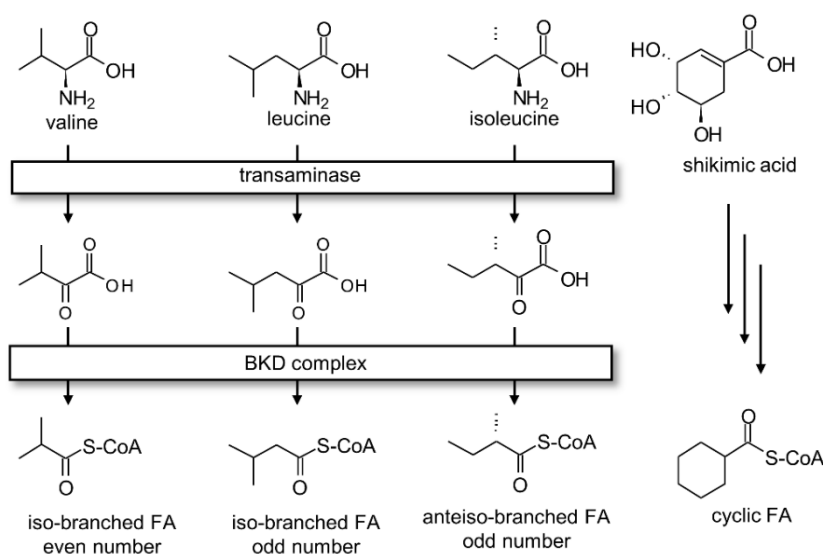


Figure 10. Different precursor molecules derived from primary metabolism.

In most organisms, only one FabH enzyme is present. However, the existence of two FabHs has been described from exceptional cases. The two FabHs from *Vibrio cholerae* for example, have different substrate specificities; one accepts short chain-length and the other one medium chain-length substrates.^[30] In the case of *Bacillus subtilis*, an additional FabH accepts iso- and anteiso-branched acyl-CoAs as substrates, besides the linear ones.^[85]

The KAS III priming strategy (Figure 11a) generates either diketides, which can be directly loaded to the ACP of the minimal PKS or used as ACP-processed intermediates. The first case was described for the propionyl starter in daunorubicin biosynthesis using a KAS III for the generation of the elongated propionyl starter, used by the minimal PKS.

The second process includes most likely an interaction of FAS II enzymes to generate the processed saturated products and involves the use of a second ACP and an additional AT, as it was described for the fatty acid precursors in frenolicin and R1128 biosynthesis (Figure 11a).^[23,46] Interestingly, KAS III enzymes involved in priming reactions with a second ACP show orthogonal ACP specificity, with the ACP carrying a conserved tyrosine that is important for KAS III recognition.^[88] Alternatively, in hedamycin biosynthesis, these reactions are replaced by an iterative PKS I module, which is assisted by the KAS III and an AT for loading the minimal PKS ACP.

In lomaiviticin biosynthesis, another strategy to generate propionate starter units was recently demonstrated (Figure 11b).^[89] Despite its seeming resemblance to the known mechanism of CLF(KS_β)-mediated acetate generation by decarboxylation of a malonate unit, a methylmalonate is used instead. A bifunctional enzyme catalyzes the decarboxylation of methylmalonyl-CoA and simultaneously and the transfer reaction of the acyl moiety to generate propionyl-ACP, which is further elongated with the minimal PKS (Figure 11b).

The third priming strategy involves a CoA-ligase and an AT (Figure 11c). To date, this strategy has only been described for a few examples. Well studied systems are the enterocin, tetracycline and aurachin pathways using benzoate, malonamate and anthranilate starter units, respectively.^[46,90] The transfer strategy is similar to the one from modular PKS I systems, where an acyl-CoA-ligase is responsible for activation of the acid substrate and the AT-domain for the transfer reaction of the non-acetate acyl moiety onto the ACP, prior to their attachment to the corresponding KS for further elongation (compare, Figure 5).^[83] Benzoate and anthranilate starter units are derived from phenylalanine and the shikimate pathway, respectively.^[47]

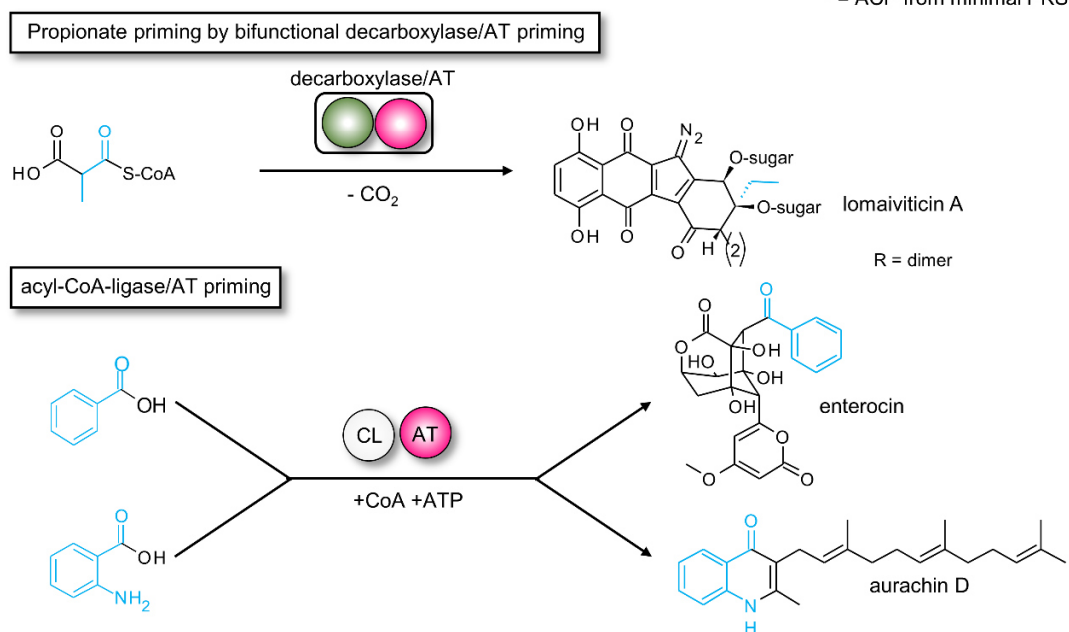
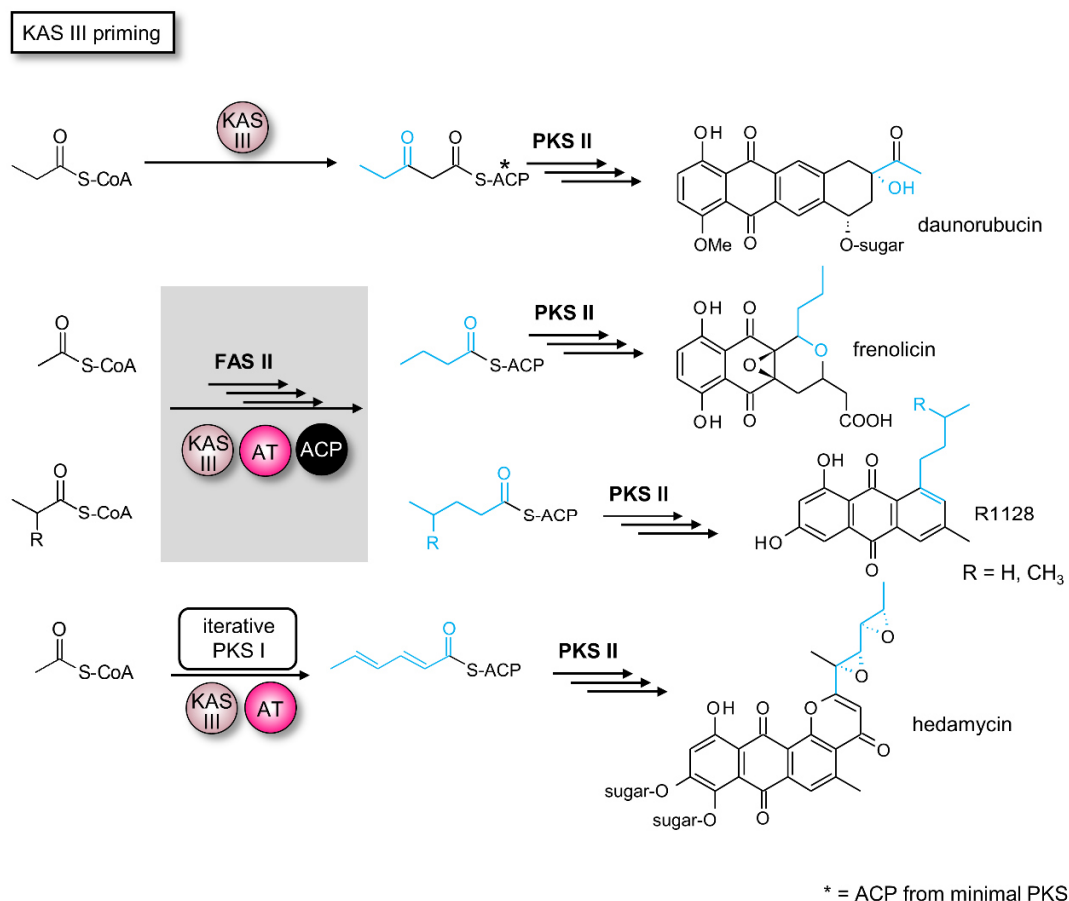


Figure 11. Non-acetate priming in PKS II systems. For more details of the different reactions, see main text.

Taken together, the co-occurrence of genes encoding for KAS III enzymes, an additional ACP, a CoA-ligase or an AT, are strong indicators of non-acetate starter units used for biosynthesis of PKS II-derived products.^[83]

6.3 The Protein Architecture of PKS and FAS Systems

The objective of this chapter is to provide an overview of the overall protein structure of FAS and PKS systems, emphasizing protein-protein interactions in type II systems. Elucidating the interactions of type II systems remains a challenging task, due to the non-covalent interactions of the involved enzymes, which are postulated to form multienzyme complexes.^[65] Type I systems have evolved from type II by a gene fusion event.^[91] Thus, an overview of the different shapes of the highly effective multifunctional enzyme complexes of type I systems will also be addressed.

6.3.1 Type I Systems – Different Synthases

As already introduced in section 6.2, FAS I and PKS I systems form multifunctional enzyme complexes. Comparative genomic analyses indicate that the megaenzymes of both systems share a common ancestor.^[76] Although Type I systems have in common their modular architecture, some remarkable and characteristic differences between FAS I as well as PKS I systems have evolved (Figure 12).^[76]

The dimerization of FAS and PKS modules is essential for their function. This dimerization is formed by dimeric KS-, DH-, TE- and ER- (not always) domains with the additional help of dimerization elements.^[92] Hereby, the KS is the only enzymatic domain that requires a dimeric state for activity.^[93] Each domain connection is formed by a linker region. Non-iterative modules, in the case of PKS assembly lines, are additionally connected in a head-to-tail fashion with unstructured linker regions or docking domains to form the effective assembly lines.^[44]

Because of the large size of these megaenzymes, ranging from several hundred kilo- to megadalton, it is challenging to get structural data, in particular, if there are ongoing dynamic processes. These dynamic processes include conformational changes, not only described for the flexible Ppant arm or the mobile ACP but also for the overall system.^[44,76]

Catching these dynamics within a whole module was faced in recent years with the combination of cryo-electron microscopy, small-angle-X-ray scattering, and crystal structure-based visualization methods.^[94–101] Fortunately, due to the latest structural data, synthases from each type can be compared (selected representatives, see Figure 12).

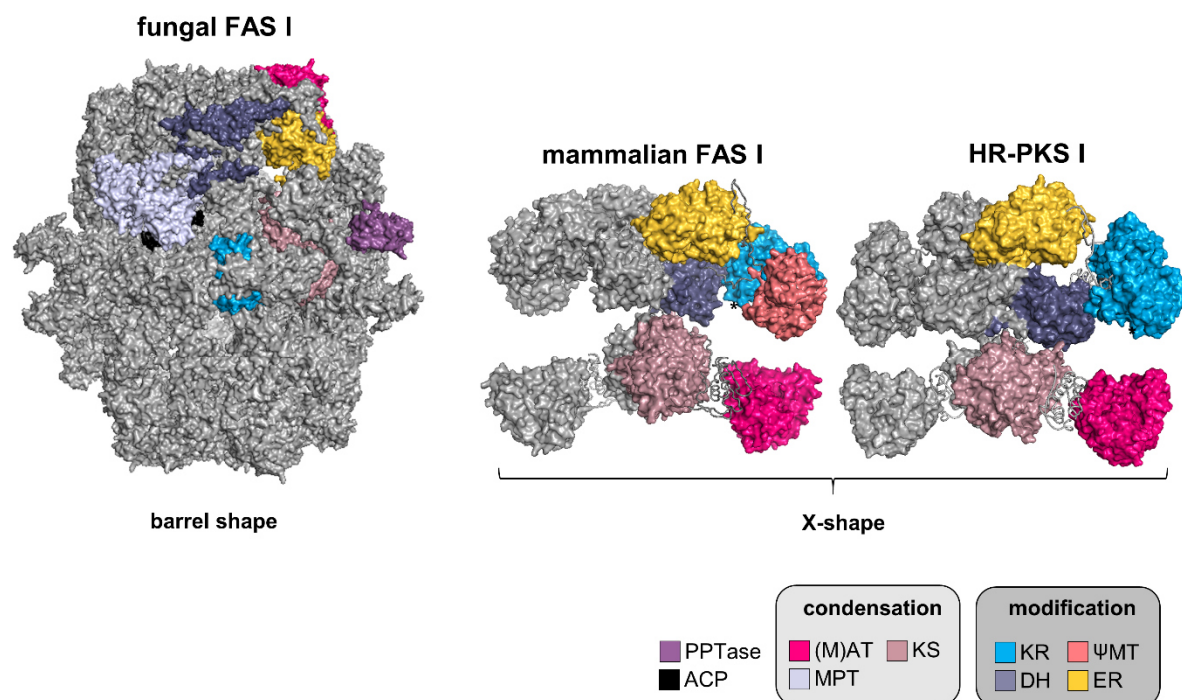


Figure 12. Comparison of protein architectures from type I synthases. Barrel-shaped fungal FAS from *Saccharomyces cerevisiae* (heterododecameric: 2.6 MDa, PDB: 3HMJ). Mammalian FAS, porcine (dimeric: 540 kDa, PDB: 2VZ9). Iterative HR-PKS, *Mycobacterium tuberculosis*, Mas-like (dimeric: 446 Da, PDB: 5BP1, 5BP4). The ACP-interaction site is marked with an asterisk in mammalian and HR-PKS I. Domain duplications and scaffolding elements are depicted in grey.

The largest and most static synthase from type I systems represents the fungal barrel-shaped megasynthase of *Saccharomyces cerevisiae*, also known as yeast FAS (Figure 12).^[100] This megasynthase consists of a 2.6 MDa heterododecameric complex (composed of six α and β -sheets, $\alpha_6\beta_6$) with an embedded PPTase-domain. Bacteria, typically use FAS II systems and the only known bacterial counterpart of a FAS I is assembled in the *Mycobacterium tuberculosis* and forms a smaller but still 1.9 MDa complex (β_6), without a PPTase-domain.^[99] The bacterial FAS I is described as a mini-version of the fungal FAS I. Both systems consist of reaction chambers, which are stabilized by a high proportion of scaffolding elements (35-50% of the overall structure).^[75] Less scaffolding domains in the bacterial FAS I are believed to have an impact on the structure by providing more flexibility.^[99] Within the reaction chambers, the mobile ACP-element shuttles the intermediates between the active sites.^[99]

The mammalian FAS I is structurally related to a PKS I module and in contrast to the static fungal and bacterial FAS I, harbors only minor amounts of scaffolding elements, resulting in an increased flexibility (Figure 12).^[76,102,103] The closest relative to the mammalian FAS I, according to the presence of processing domains (including KR-, DH- and ER-domain), is formed by an iterative HR-PKS (see Figure 4). The so far only solved structure for an HR-PKS is reported for the MAS-like synthase from *Mycobacterium tuberculosis* (Figure 12). In general,

the mammalian FAS I and PKS I synthases occupy a more open X-shaped structure, consisting of a homodimer (α_2). Here again, the ACP-bound acyl chains are sequestered in reaction chambers. According to their structure, FAS I and PKS I systems are divided into a condensing (including the loading modules) and a modifying part.^[104] The condensing part remains constant regarding its domain endowment and the modifying part changes upon addition or lack of the processing domains in PKS I systems (KR-, DH-, ER-domain).^{[102,105][75]} The flexibility of FAS I was demonstrated to take place around the linkage of the condensing and modifying parts, plus some flexibility in waving motions towards each other (Figure 13).

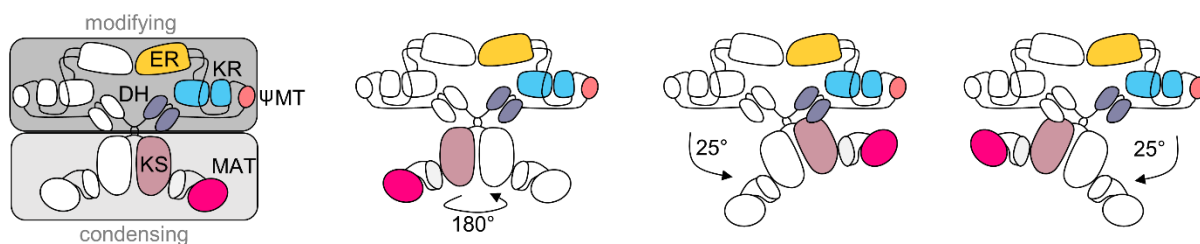


Figure 13. Flexibility in mammalian FAS in swinging and waving motions. The figure was modified according to reference ^[75].

Recently, cryo-EM structural data reported also a high flexibility in a *trans*-AT PKS I upon different acylation states.^[52] Although there are striking differences between the systems, there is still a large intersection that has opened the way for targeted protein engineering.^[44,106–109]

6.3.2 Type II systems – Protein-Protein Interactions

In contrast to type I systems, which evolved in effective reaction chambers^[106], type II systems require a different concept to control substrate selectivity and the order and timing of the multiple reactions.^[65] The corresponding protein-protein interactions are either more transient or result in tight complexes and will be presented in the following. Prior to addressing the differences between protein-protein interactions of PKS II and FAS II systems, the special features (additionally, to the prosthetic Ppant arm (see Figure 2)) of the ACP are explained in more detail.

6.3.2.1. Small but Powerful – The ACP

ACPs are responsible for the interaction with all the other proteins involved in polyketide and fatty acid biosynthesis.^[65] The structure of an ACP is conserved, consisting of a helical bundle fold with four helices, three major ones (helix I, II, IV) and one minor one (helix III) (Figure 14). Helix I and II are joined by a flexible loop. The conserved serine (within the motif DSX, X being typically L or R) for the attachment of the Ppant arm is located at the end of this loop and at the beginning of helix II.^[27]

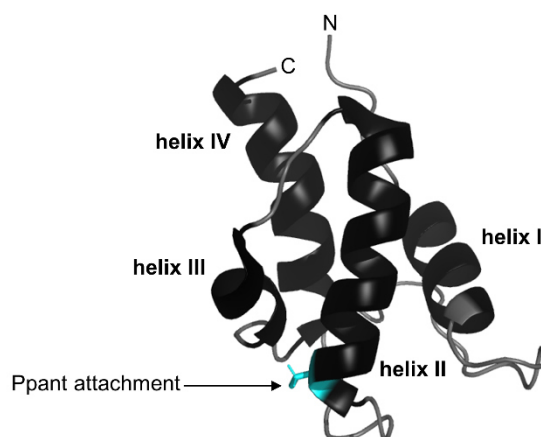


Figure 14. NMR solution structure of the typical helical bundle fold of the actinorhodin ACP from a type II PKS (PDB: 2K0X). Highlighted in light blue is the conserved cysteine side chain for Ppant attachment. The figure was generated using PyMOL 2.3.2.

Helix II is reported as the ‘recognition helix’, involved in initial protein-protein interactions. These interactions originate from the ACP by polar and negatively charged residues, followed by hydrophobic interactions. This kind of non-covalent interaction motif was recently reported to have similarities with the ‘ubiquitin interacting motif’.^[110]

Additionally, the ACP-coupled Ppant arm can switch between two different conformation states, bound in a hydrophobic pocket or on the protein surface (Figure 15). During the sequestering in the hydrophobic pocket, the intermediates are protected against (auto)catalytic reactions.^[111] Upon protein-protein interactions, the Ppant arm can be translocated to the active site of the interaction partner, a process which is also known as chain-flipping mechanism.^[111,112] This mechanism is exclusively found in type II systems since type I systems exclusively catalyze the corresponding reactions in their reaction chambers.^[60,113] Furthermore, in type II FAS, the Ppant-coupled intermediates are completely buried in the ACP, while in PKS II systems, the chain is only partially sequestered, representing a hydrophobic cave rather than a pocket. The chain sequestering takes place between helix II and III (Figure 15).^[60]

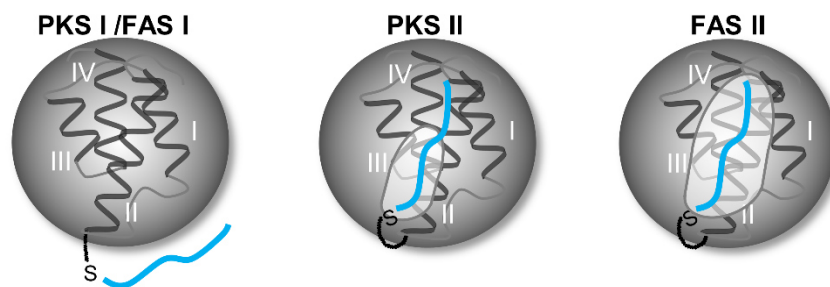


Figure 15. Acyl-chain (blue) sequestering in PKS and FAS systems. PKS I/FAS I: acyl-chain sequestering in reaction chambers, PKS II: partially sequestered acyl-chain in hydrophobic cave, FAS II fully sequestered acyl-chain in hydrophobic pocket. The figure was modified according to reference [60].

Consequently, the ACP has two different roles during the biosynthesis: Sequestering and transferring the intermediates to the active center of the involved enzymes and enabling the specific interactions with all interaction partners.[65] Although the general role of the carrier proteins in both, PKS II and FAS II systems, remains the same, the single reactions and interactions with involved proteins are different. Representative interaction mechanisms will be presented in the following.

6.3.2.2. Protein-Protein Interactions in PKS II Systems

The enzymes involved in protein-protein interactions of PKS type II systems are presented in Figure 16. Recent findings of the overall interaction network, which is still little known, will be discussed in this section.

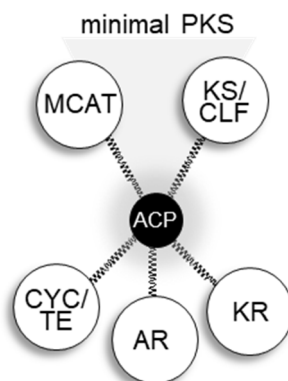


Figure 16. Enzymes involved in PKS II systems. Depicted are proposed protein-protein interactions being responsible for producing the aromatic polyketide scaffold. Not included are enzymes involved in post-tailoring processes. The figure was modified according to reference [65].

The groundbreaking and only X-ray structure of the heterodimeric KS:CLF complex of the actinorhodin synthase from *Streptomyces coelicolor*, gave striking insights in the action of the minimal PKS part (PDB 1TQY) (Figure 17).[61] The tube-like cavity, responsible for chain-length determination, is localized at the interface of the two enzymes. The acylation with the

polyketide takes place at the so-called nucleophilic elbow.^[61] The decarboxylation of the malonate extender unit during chain elongation, is mediated by two histidines in the active site.

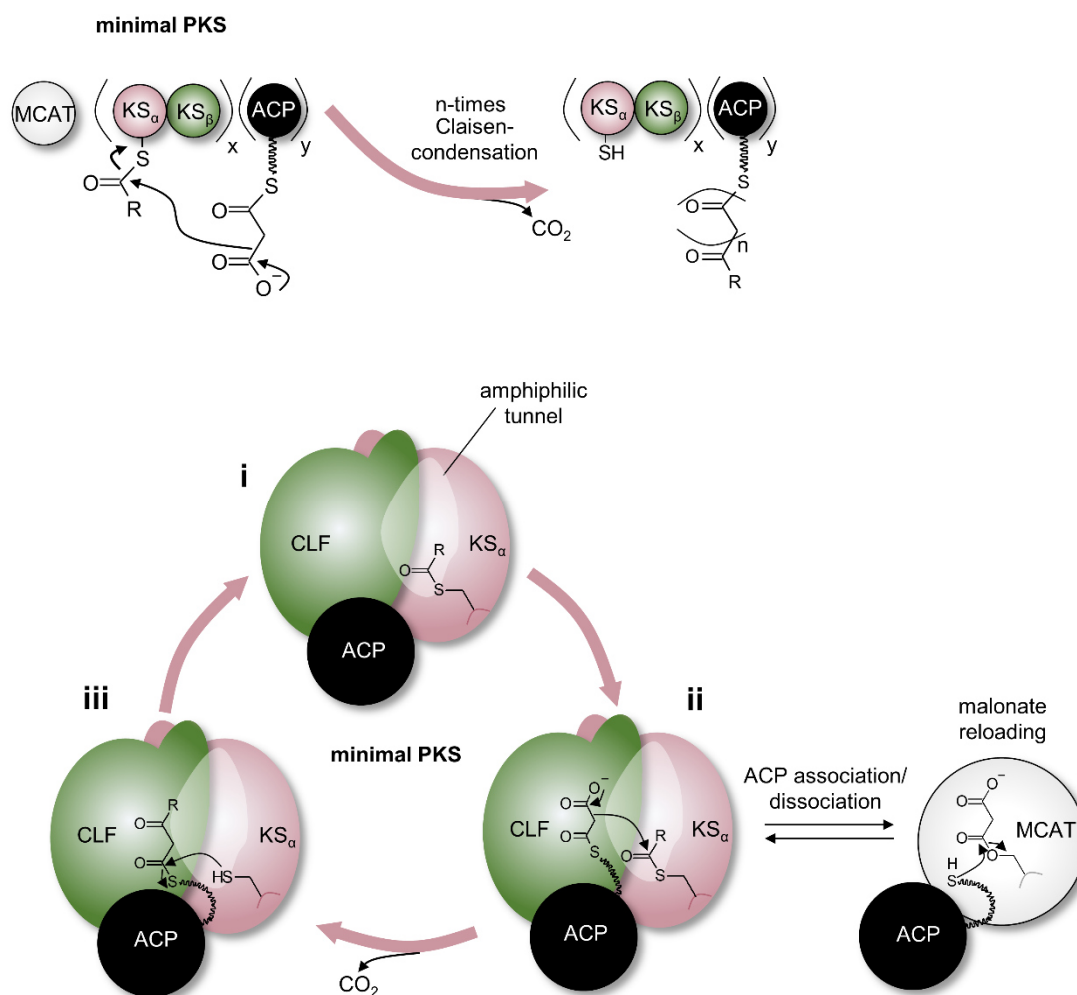


Figure 17. Minimal PKS in action. **i** The conserved cysteine of the KS is either acetylated (from decarboxylated malonate) or acylated with the growing acyl chain **ii** Decarboxylative Claisen-condensation takes place with a malonate extender unit, loaded onto the ACP by the MCAT (see also above). For reloading the malonate units, repeated ACP association and dissociation from the KS:CLF to the MCAT is needed. **iii** After elongation, the acyl-chain is back-transferred to the cysteine residue of the KS to start with a new elongation cycle. The figure was modified according to reference ^[61].

In contrast to the repeated interactions with the processing partner proteins (KR, DH and ER) of FAS II systems, the ACP from PKS II interacts during the elongation process exclusively with two enzymes. The MCAT, for reloading malonate elongation units as well as with the KS:CLF complex (Figure 17).^[61] In this process, the elongated polyketone is sequestered within the amphiphilic cavity of the KS:CLF heterodimer.^[61] In this way, the polyketide intermediate remains protected during the elongation process, while the ACP switches between MCAT and KS:CLF to reload the malonate units (Figure 17, ii).^[61] However, this repeated assembly and disassembly is poorly understood. During *in vitro* biosynthesis, there is even no need for an MCAT, since a self-malonylation of the ACP takes place.^[114] It is further

supposed that the polyketide intermediate gets stabilized by some ketides, which are enolized and interact with side chains of the KS:CLF-amphiphilic tunnel.^[60]

After reaching the final length, which is determined by some gate-keeper residues of the CLF (CLF: Phe109, Thr112 and Phe116),^[61] the chain is sequestered in the hydrophobic cavity of the ACP instead, between helix II and III.^[60,65,115]

It is described that the acyl cargo is not protected in total length in this cavity (compare Figure 15). In contrast, it was shown that it is partly solvent-exposed, which enables a further ACP-KR reaction and leads to the first cyclization event (actinorhodin C₇/C₁₂, see Figure 6). This is evident, if considered, that a missing KR leads to miss-cyclized products, due to the areas of the polyketide not being protected.^[60] The transferred linear polyketide intermediate gets stabilized by the ACP for the regiospecific C₉ reduction by the KR after dissociation of the KS:CLF complex.^[60] Nevertheless, the overall interplay of the first cyclization event remains elusive. The initial involvement of KS:CLF mediated cyclization cannot be excluded, as the reaction takes place by spontaneous aldol chemistry.^[35,116]

The binding of the first cyclized reduced intermediate is proposed to facilitate the dissociation of the tetrameric KR from the ACP, followed by association of the ARO/CYC (compare Figure 6).^[65] While CYCs are quite heterogeneous in sequence and structure (also regarding multimerization), they are supposed to function in a chaperon-like manner.^[117] Furthermore, they mediate the specific cyclization reactions via aldol condensations in defined reaction channels.^[35] Here, the highly reactive polyketide intermediates are protected prior to getting aromatized by the mono- or dimeric AROs.^[35,117,118] CYCs are not only supposed to be involved in stabilizing the overall complex, as their presence prevents the production of shorter polyketides but also to stimulate the turnover rate of the complex.^[117]

However, the exact interplay between the minimal PKS and the KR with the ARO/CYC remains obscure.

6.3.2.3. Protein-Protein Interactions in FAS II Systems

In contrast to type II PKS systems, where the polyketide is sequestered within the minimal PKS for elongation and is only pulled out for further tailoring reactions, the fatty acid intermediate in FAS II is buried in the hydrophobic pocket of the ACP instead, and is flipped out upon protein-protein interactions with the processing enzymes (Figure 18).^[65] Because of the multiple reactions with several enzymes, these interactions with the ACP must be tight enough, so that the chain-flipping of the intermediate is mediated, but loose enough to switch in an efficient manner between the different reaction partners.

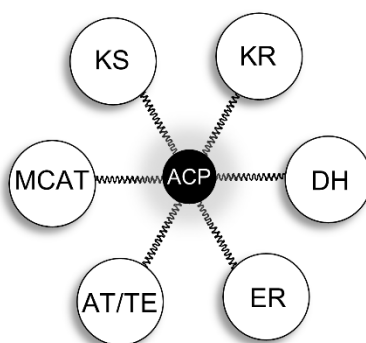


Figure 18. Enzymes involved in protein-protein interactions of FAS II systems. The figure was modified according to reference ^[65].

As already indicated, this chain-flipping mechanism is initiated with the recognition helix II of the ACP.^[60] Structural data of all stand-alone enzymes involved in FAS II systems exist.^[70] However, to understand the distinct interaction-mechanisms and interaction sites, it is necessary to get structural data of enzyme complexes. The transient interactions during FAS II biosynthesis are challenging to trap. To overcome this, several cross-linking (=) strategies have been developed and recently resulted in the X-ray structure of ACPs with the two elongation KS FabB and FabF (ACP=FabB, ACP=FabF) and of the ACP in complex with the DH FabA (ACP=FabA).^[65,119] Both KSs share a gating mechanism upon protein-protein interaction with the ACPs (Figure 19). Different to the minimal PKS (ACP plus KS:CLF, MCAT), each subunit of the homodimeric KS, forms a cross-linked complex with one ACP, leading to a hetero-tetrameric complex.^[120]

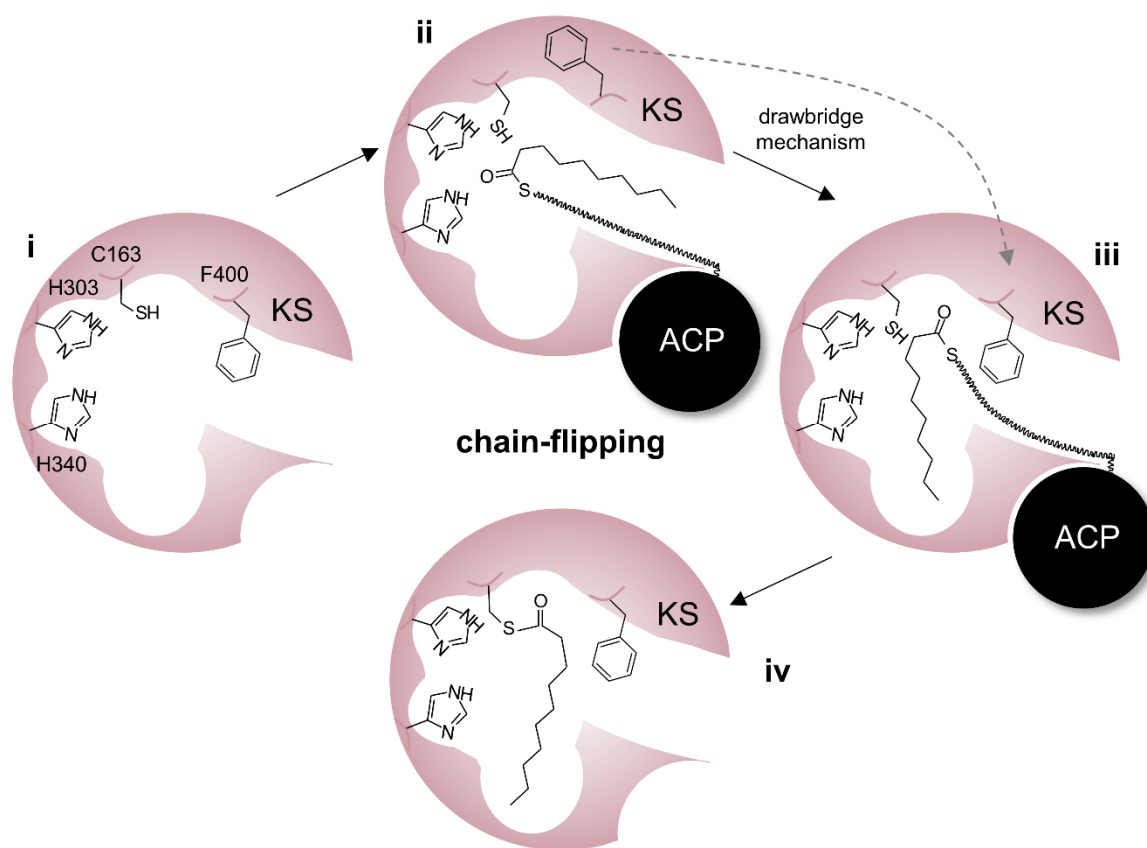


Figure 19. Chain-flipping mechanism during ACP:KS interaction. **i** Binding pocket of KS with catalytic triad C163, H303, H340 and gatekeeper residue F400. **ii** Association of ACP through recognition helix II, 'gate-open'. **iii** Conformational changes in loop 1 and 2 (not depicted) allows chain flipping and enclosing of reaction site by gating residue F400, 'gate-close'. **iv** Transacylated KS. The figure was modified according to reference ^[120].

The gating mechanism during ACP=KS interaction describes a conformational change in two loops (loop 1 and loop 2) of the KS, after protein-protein interactions with recognition helix II of the ACP.^[120] This conformational change leads to chain-flipping and transfer of the intermediate to the conserved cysteine in the active site of the KS (Figure 19). A phenylalanine residue is described to be the gatekeeper residue to prevent the solvent from entering the KS active site and to avoid unwanted side reactions. This gating mechanism was described as comparable to a drawbridge-like mechanism (Figure 19).^[120]

Striking interactions were revealed by solving the native crystal structure of hexameric FabZ (three arranged homodimers), with each homodimeric subunit binding one ACP (Figure 20).^[119] Here, a different mechanism to induce the chain-flipping takes place. This mechanism is supposed to function in a seesaw-like manner to regulate the substrate shuttling. The seesaw-like mechanism is mediated due to the large structural changes of FabZ. During this process, the ACP recognition helix II binds near the active tunnel of FabZ (catalytic diade: conserved His-Asp-motif), while a conserved arginine residue is supposed to play an important role in this

protein-specific recognition, prior to hydrophobic interaction takes place.^[65,119] The binding of the ACP through recognition helix II induces a β -layer within the DH to move. The L-shaped active site tunnel of the DH FabZ is changed to a Y-form conformation and induces chain-flipping of the long fatty acid substrate. Additionally, a gatekeeper tyrosine side chain allows entrance.^[121] This Y-shaped tunnel of FabZ has more steric hindrance compared with the L-shaped tunnel of FabA, which may be an explanation for the loss of isomerase activity of FabZ.^[121] The partner FabZ undergoes similar conformational changes but here, the β -layer is moved in a seesaw-like manner in the opposite direction (Figure 20, compare ii and v). At the same time, a specific loop is extruded at the ACP-binding interface of helix II and prevents the binding of a second ACP. This hindrance-process is further supported by the rotation of the tyrosine gatekeeper, which additionally ensures a more closed conformation.^[121]

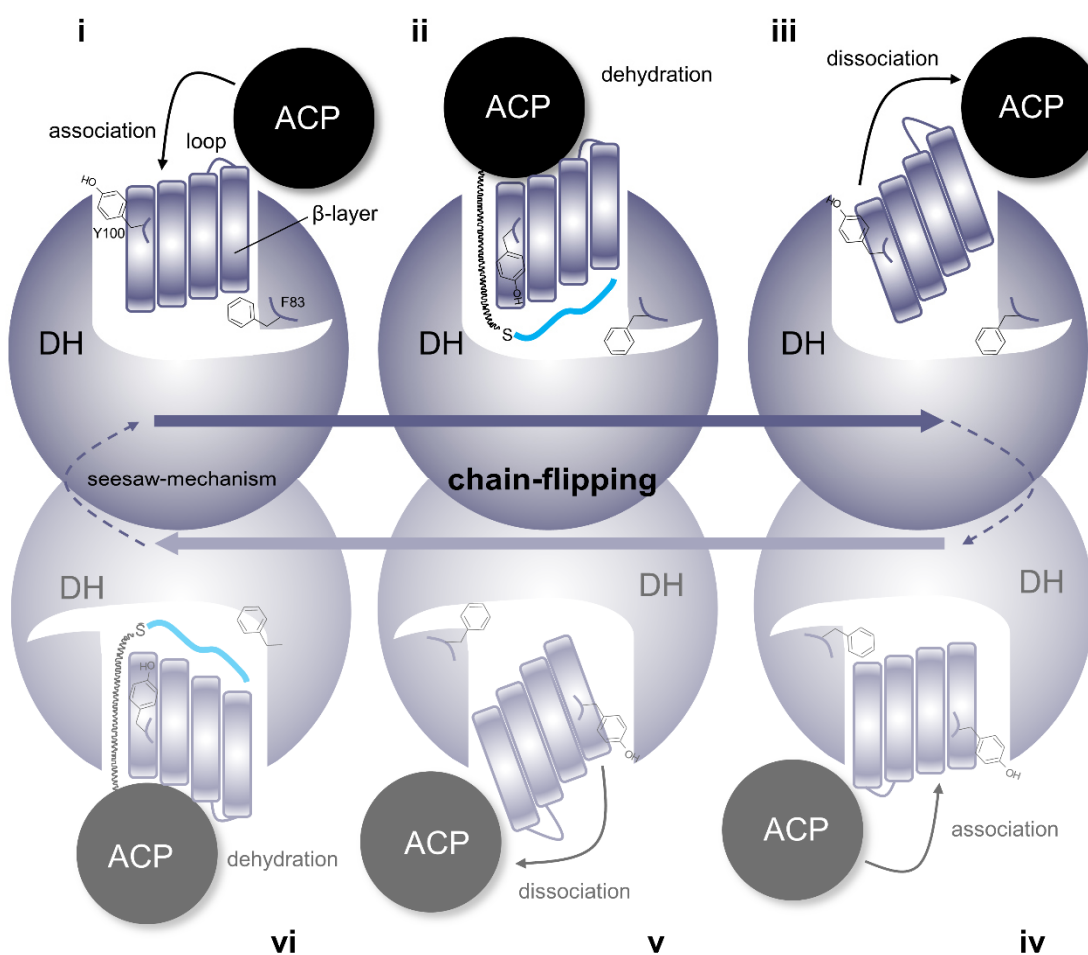


Figure 20. Chain-flipping mechanism during ACP:DH interaction. **i** association of the ACP through recognition helix II leads to structural changes within the DH **ii**. The tyrosine gatekeeper moves to allow chain-flipping and dehydration reaction (The active site-diade is not displayed). At the same time, the β -layer of the opposite DH moves and extrudes a specific loop that prevents the binding of a second ACP. **iii** dissociation of the ACP to allow association of a new ACP at the opposite DH. **iv-vi** repetition of steps i-iii. The figure was modified according to ^[121].

Different from FabZ, an ACP cross-linked FabA-homodimer binds two ACPs, one at each active site to result in a heterotetrameric complex.^[122] Interestingly, the positions of the ACP vary and are supposed to represent different reaction states, which might be artificially trapped by the cross-linkers.^[65] It is likely that due to the overall structural similarity of FabA and FabZ, both DHs undergo a conserved reaction mechanism.^[121] Two phenylalanine gating residues at the entrance of the active site are further supposed to mediate substrate specificity.^[122]

Within the last five years these structural insights of FAS II systems during protein-protein interaction were revealed, ten years after the discovery of the *actKS:CLF*.^[61] Due to the immense progress in structural biology, it is only a matter of time to get access to further complexes, being part of the postulated multienzyme complex of type II systems.

6.4 PKS II Systems – Rare in Gram-Negative Bacteria

As already introduced (6.2.1.2), type II PKS systems are almost exclusively found in Gram-positive actinomycetes^[23] and until today only five examples have been described from Gram-negative bacteria, while three of them are part of this work.

The first example was reported from the myxobacterium *Stigmatella aurantiaca* producing the aromatic aurachins, with anthranilate as a non-acetate starter unit (priming mechanism, see 6.2.3) and an unusual farnesyl-group, attached by the activity of a prenyltransferase (Figure 11).^[123,124] Only recently, the aromatic pyxidicyclines are described as another example from myxobacteria and revealed a nitrogen-containing quinone-like four ring structure.^[125]

The remaining three PKS II systems are derived from the soil-living entomopathogenic γ -proteobacteria *Photorhabdus* and *Xenorhabdus* that live in mutualistic symbiosis with *Heterorhabditis* and *Steinernema* nematodes, respectively.^[126–129] The PKS II-derived compounds are produced strain-specific: Exclusively found in *Photorhabdus* are anthraquinones (AQ) and isopropylstilbenes (IPS), while almost every *Xenorhabdus* strain harbors an aryl polyene (APE) BGC, mentioned that APE BGCs are widespread around Gram-negative bacteria (see 6.4.3). In the following, the state of the art of the underlying biosynthetic mechanisms is described. The systems are introduced in the order from typical PKS II systems (AQ) to more untypical ones (IPS, APE).

6.4.1 Anthraquinone (AQ) Biosynthesis

Photorhabdus strains possess a characteristic orange-to-reddish pigmentation upon their exponential growth phase. This pigmentation is caused by the production of different anthraquinone (AQ) molecules. In 2007 Brachmann *et al.*^[130] revealed the corresponding gene cluster and presented at this timepoint, besides the aurachins,^[123,124] only the second example for a PKS II-derived product from a Gram-negative bacterium. Besides the minimal PKS, the gene cluster encodes for a CoA-ligase, a KR, an ARO, a CYC, a hydrolase and an Sfp-type PPTase (Figure 21). Isotope labeled acetate feeding experiments revealed the incorporation of seven acetate derived malonate units.^[130] Homologous gene disruption of the cyclase AnthH further leads to the detection of the shunt products mutactin and dehydromutactin. These shunt products result typically by the production of the minimal PKS plus KR (compare actinorhodin biosynthesis, Figure 6) from an octaketide scaffold. This indicates a biosynthesis which resembles the PKS II-derived actinorhodin. This result was in contradiction to the formation of the main product, AQ-256, that was thought to originate from a heptaketide, as it was supported by the acetate-feeding experiments. A possible shortening mechanism from an octaketide to a heptaketide was postulated through the action of the hydrolase AntI but remained obscure at this point. Since the detected AQ-products are primed with acetate, the

function for the CoA-ligase remained elusive, too. In addition to AQ-256, three post-assembly line methylated anthraquinone derivatives were detected. The following studies revealed five methyltransferases being involved in methylation, as well as the transcription activator AntJ, involved in AQ-regulation. A topic that is well presented in the PhD thesis of Antje Heinrich.^[131] Taken together, the biosynthesis of AQs was postulated, based on the biosynthesis of actinorhodin (Figure 21).

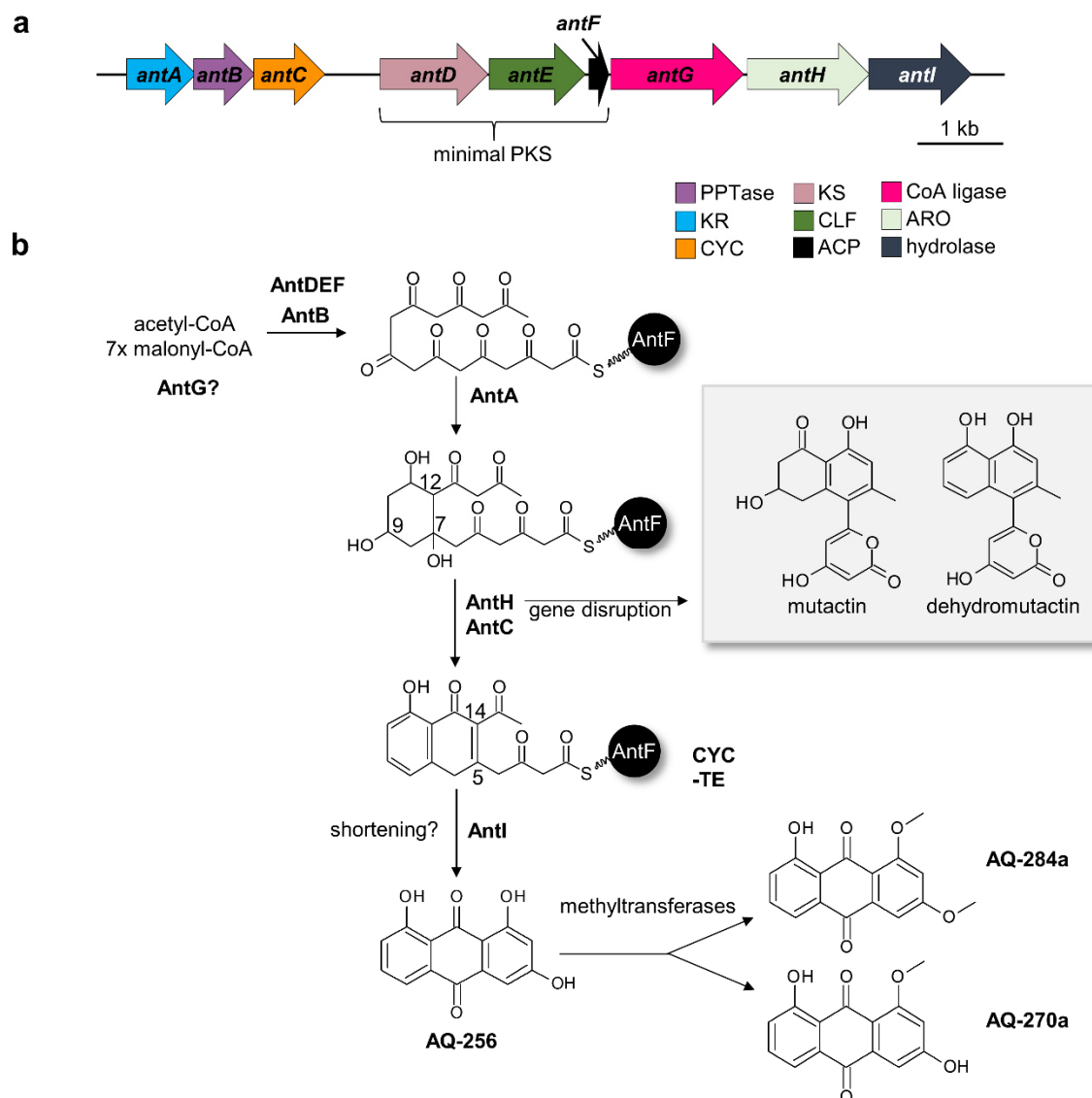


Figure 21. AQ biosynthesis. AQ BGC from *P. laumondii* with *antA-I* (*plu4195-4186*) (a). Proposed AQ biosynthesis (b). For more details, see main text.

6.4.2 Isopropylstilbene (IPS) Biosynthesis

Typically, stilbenes are produced by plant stilbene synthases (STS) corresponding to a type III PKS system.^[67,68] Yet, the only non-plant-derived stilbene product is isopropylstilbene (IPS) and some derivatives thereof, produced by *Photorhabdus* strains.^[132–134] As IPS is found in all *Photorhabdus* strains so far, it can be regarded as a strain-specific biomarker.^[128] *Photorhabdus* lives in mutualistic symbiosis with *Heterorhabditis* nematodes, acting together as an insect pathogen.^[126,127] Due to its antimicrobial activity, it was suggested that IPS is involved in protecting the insect cadaver against food competitors, such as other bacteria.^[132] Furthermore, IPS is essential for nematode development, inhibits protozoa and even the mammalian soluble epoxide hydrolase, which is involved in the arachidonic acid cascade and consequently, in several inflammatory processes.^[132] Since IPS selectively modulates the cytokine cascade, when applied as a topical cream, it is currently developed as treatment against psoriasis under the name tapinarof^[135] or benvitimod.^[136]

Interestingly, the representative synthase from *P. laumondii* uses a non-plant-like mechanism to produce IPS (Figure 22), which relies on the condensation of two intermediates instead of elongating one cinnamoyl moiety as it is known from plants (compare Figure 7). This unusual mechanism allows the generation of the characteristic iso-branch. The corresponding genes are not located within one BGC but distributed within the genome (*stIA* and *stIB*, *stCDE* and *bkdABC*) (Figure 22).

IPS biosynthesis starts with the generation of cinnamic acid from phenylalanine by a phenylalanine ammonium lyase (PAL, *StIA*). *StIB* is proposed to act as a CoA-ligase, activating cinnamic acid to cinnamoyl-CoA. It is further suggested that the acyl moiety of activated cinnamoyl-CoA is further loaded to an ACP, elongated, reduced and dehydrated to 5-phenyl-2,4-pentadienoyl-ACP, which serves as a Michael acceptor for further reactions. Although there is no proof yet, cinnamoyl-ACP extension might involve enzymes from the fatty acid biosynthesis. The Michael acceptor (5-phenyl-2,4-pentadienoyl-ACP) is required for the downstream reaction with the β -ketoacyl isovalerate moiety, derived from the BKD complex. Thus, cyclization (*StID*) of both precursor molecules and followed aromatization (*StIC*) finally leads to IPS, with the characteristic iso-branched side chain (Figure 22).

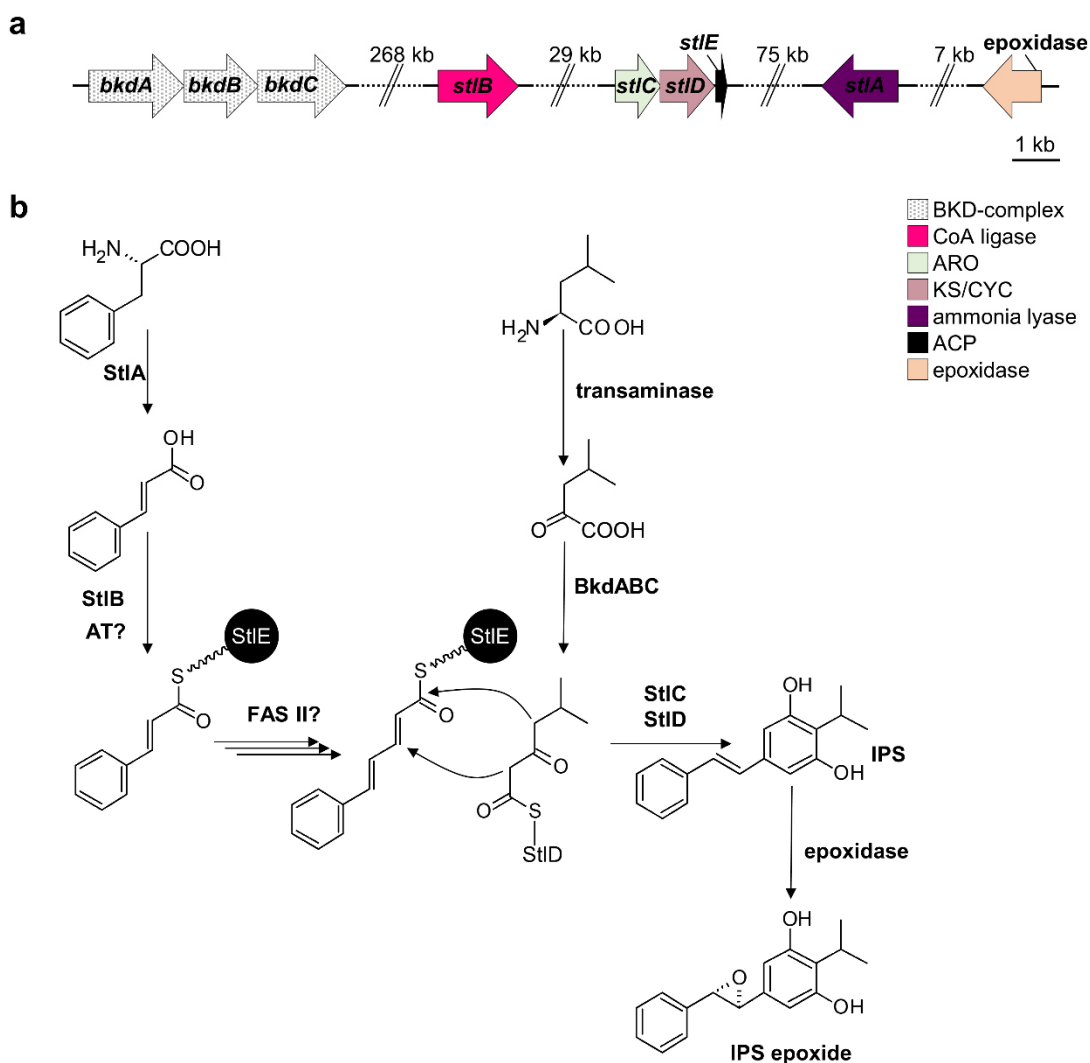


Figure 22. IPS biosynthesis. Distributed IPS-genes involved in biosynthesis from *P. laumondii*. *bkdABC*, BKD complex (*plu1883-1885*); *stlB*, CoA-ligase (*plu2134*); *stlCDE*, ARO-KS/CYC-ACP (*plu2163-65*); *stlA*, phenylalanine ammonia lyase (*plu2234*), epoxidase (*plu2236*) (a). Proposed IPS biosynthesis (b). For more details regarding the reaction mechanisms, see main text.

An X-ray protein crystal structure of the key cyclization enzyme StlD and *in vitro* analysis of the cyclization and aromatization (StlC) reactions, resulted in interesting molecular insights in the biochemistry of the unusual reaction mechanisms, used by the bacterial IPS synthase.^[137] Although StlD adopts the classic homodimeric thiolase-like fold, which is well-known for KS, the crystal structure revealed two acyl-binding pockets. In-between those two binding pockets, an elongation/cyclization cavity is located. Additionally, StlD has a large hidden hydrophobic pocket that is thought to accommodate longer acyl chain substrates. Furthermore, compared to other elongating KS, subtle changes around the catalytic triad allow for more open space around the reactive Cys126. Furthermore, the intramolecular Michael addition of the two substrates (5-phenyl-2,4-pentadienoyl- and isovaleryl- β -ketoacyl moiety), yields in a carboxy-cyclohexanedione product that is used by StlC to generate the aromatic stilbene. Hereby, the carboxylic acid moiety of the ring is essential for StlC recognition.^[137]

6.4.3 Aryl polyene (APE) Biosynthesis

APE compounds became of broad interest by the work of Cimermancic *et al.*^[138] Here, in an extensive bioinformatic analysis, the APE BGC was shown as one of the most widespread gene clusters among Gram-negative proteobacteria. Two APE products from *Vibrio fischeri* (*V. fischeri*) and *Escherichia coli* (*E. coli*) CFT073, respectively, were isolated and structurally confirmed (Figure 23). Intriguingly, those APE compounds were set in context with the well-known pigments xanthomonadin^[139,140] and flexirubin.^[141,142]

Both compounds were initially thought to be strain-specific. Only with the help of the global alignment, the assignment as APE compound was possible.^[138] Thus, by analyzing over 1000 BGCs, the authors sub-grouped the APE BGC further into three subfamilies: APE subfamily 1 with the BGC producing the APE_{Ec}, APE subfamily 2 with the BGC producing the APE_{Vf} and APE subfamily 3 with the BGCs producing xanthomonadin or flexirubin compounds (Figure 23). Due to the genetic differences within APE subfamily 3, this group is further divided into the BGCs producing xanthomonadin-like and the BGCs producing flexirubin-like compounds. Apart from flexirubin, the described APE compounds are presented as APE alkyl ester. Flexirubins are instead esterified to a dialkylresorcinol (DAR) moiety (see Figure 23, arcuflavin). The genes, responsible for the production of the DAR moiety are not found within the APE BGCs (Figure 23) and did not account for the classification.

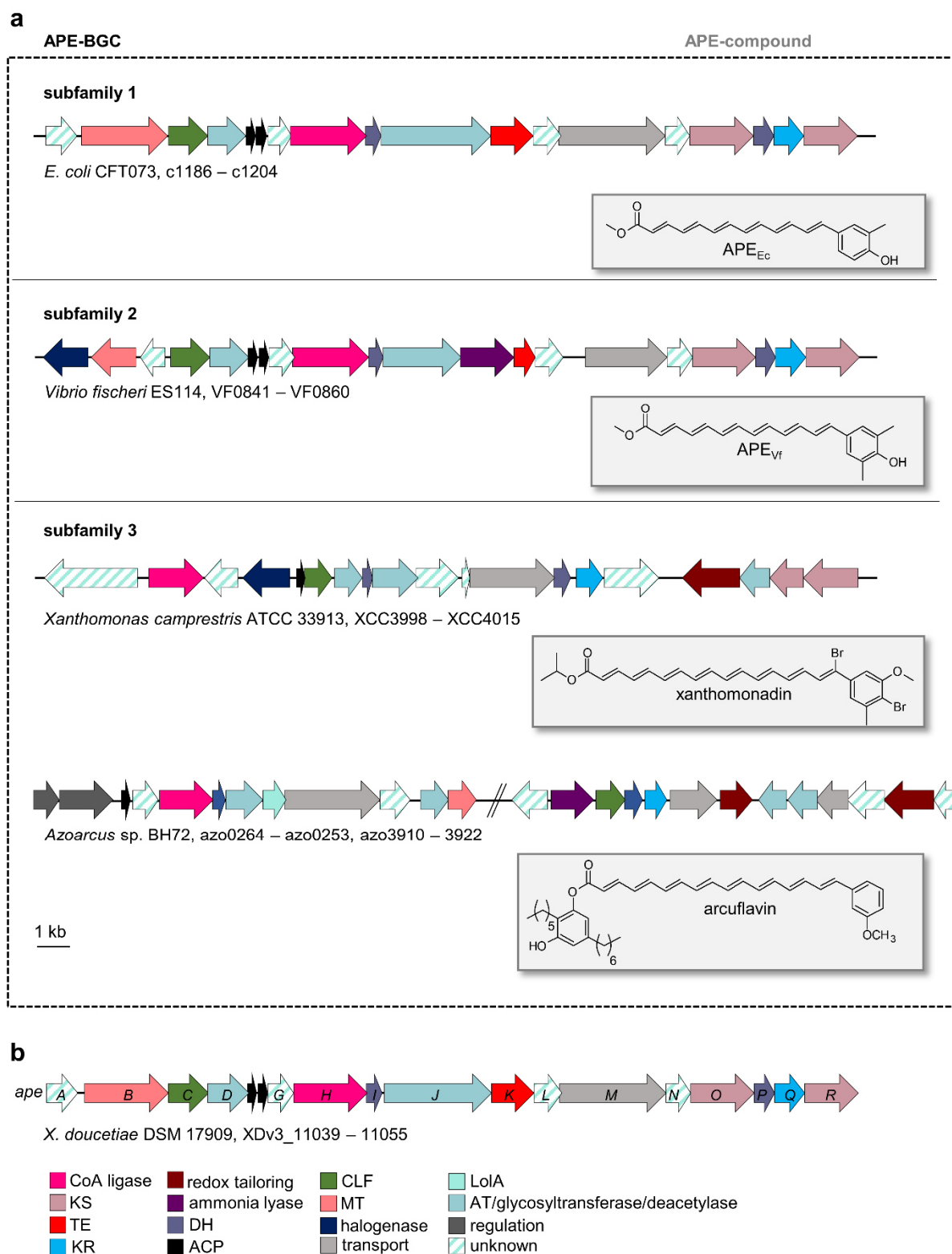


Figure 23. Different types of APE BGC representatives from subtype 1,2 and 3 with their corresponding APE compounds (**a**) and APE BGC found in *X. doucetiae* (**b**).

The characteristic feature of APE compounds is the aryl head group that is connected to a polyene core and leads to a bright yellow-to-orange coloration. Consequently, APE compounds are structurally related to carotenoids, which are, in contrast, synthesized by the

isoprenoid-pathway.^[143] It was shown that APE compounds possess similar antioxidative abilities,^[144] they are supposed to be located in the cell membrane, responsible to protect the organism from environmental stress, i.e. reactive oxygen species.^[145]

Typically, polyene compounds arise from (iterative) type I PKS, such as described for the polyene core structure in enediyne biosynthesis or for the biosynthesis of the polyene macrolactones amphotericin B or nystatin, produced by *Streptomyces* strains.^[146,147] It is therefore quite surprising that the discovered APE compounds are produced from monofunctional enzymes. The APE BGC contains genes, which can be involved in the starter unit selection, e.g. the CoA-ligases and ammonium lyases. Additionally, genes proposed to be involved in the fatty acid-like elongation mechanisms are found, namely, genes encoding for KS, KR and DH. According to the gene arrangement, there exists no classical 'minimal PKS' in one operon. However, genes encoding for an ACP, a KS and a C-terminally truncated version of a CLF can be identified as part of the BGC. The role and the effect of gene duplications of the key enzymes (ACP, KS, DH) is not known. Genetically, this PKS II system thus encompasses components shared from both, PKS II and FAS II systems (compare 6.2) and due to the more linear structure of APEs it is postulated that the polyene compound is synthesized in a fatty acid-like elongation mechanism.^[139,148]

Additional genes in the APE BGCs indicate that APE molecules might be transferred to a glycerol to result in a lipid-like structure. These genes, encoding for acyltransferases and a membrane transporter (Figure 23), are conserved among the APE subtypes. In entomopathogenic bacteria from *Xenorhabdus*, the APE BGC from subfamily 1 is widespread (Figure 23b).

6.5 Aim of this Thesis

This thesis is focusing on three PKS II systems found in the Gram-negative bacteria *Xenorhabdus* and *Photorhabdus*. Here, predominantly, the underexplored class of APE compounds were of special interest. Preliminary *in vitro*-results showed the first evidence of a fatty acid-like elongation mechanism by analyzing the chain-initiation reaction of an APE subfamily 3 system from *Azoarcus* sp. BH72, producing the pigment arcuflavin.^[148,149] However, further *in vitro*-investigations were needed to fully reconstitute the biosynthesis of these special molecules. Thus, within this thesis, the subfamily 1 APE-PKS II system from *Xenorhabdus doucetiae* DSM 17909^T (*X. doucetiae*) should be investigated. *In vitro*-studies with all enzymes proposed to be involved in APE biosynthesis are planned in order to reveal the mechanisms for the initiation and termination reactions and the role of gene duplications, encoding for the key proteins ACP, KS and DH.

At the same time *in vivo*-studies aim to further investigate the APE compound produced by *X. doucetiae*. The gene cluster is silent under laboratory conditions, but its activation is attended to be obtained by an inducible promotor-exchange system.^[150] Further homologous expression of the APE BGC is ought to study the function of the 'lipid-genes' with the help of targeted deletion studies. Hereby, MS-based analysis and purification are thought to provide solid structure evidence of the APE compounds.

The IPS biosynthesis was already well-studied in the past by the Bode group^[132,151] as well as by the Abe group.^[152] The bode group revealed the genes distributed within the genome of *Photorhabdus laumondii* TT01 (*P. laumondii* or TT01), responsible for precursor generation, ring formation and aromatization of this special, two-branched biosynthesis. Furthermore, the cyclization mechanism of the non-canonical KAS III-type KS/CYC, which is based on a Michael addition of two precursor molecules, was discovered. The Abe group supported this knowledge with the X-ray structure of the KS/CYC StID and biochemical studies with StID and the ARO StIC, giving rise to the molecular details of the cyclization reaction (compare Figure 22) However, the generation of the cinnamoyl-precursor, used as a Michael acceptor for the cyclization reaction remained elusive. There is not only a lack of a missing AT, which is proposed to load the cinnamoyl moiety to the ACP but also for the enzymes that generate the cinnamoyl-Michael acceptor. These reactions might be complemented by FAS II enzymes, as it was supposed for the PKS II derived frenolicin or R1228^[23,46] (see Figure 11). Thus, our investigations should not only deal with the starter unit selection but also with a possible cross-talk of specialized and primary metabolism using the purified enzymes.

The PKS II system in *Photorhabdus* producing AQs is the best-studied system according to preliminary results. While several investigations of Brachmann *et al.* discovered the BGC and the overall product spectrum of the main AQ-256 and its methylated derivatives, unpublished data of Quiquin Zhou engaged biochemical *in vitro* analysis paired with *in vivo* heterologous expression of the *ant*-genes *antA-I*.^[153] This leads to the identification of several more shunt products, supporting the hypothesis that the AQ-scaffold is derived by an octaketide which gets shortened to a heptaketide by the hydrolase AntI resulting in AQ-256. AntI was further investigated through X-ray analyses conducted by our cooperation partner from the Groll group. This reveals molecular insights in a polyketide shortening mechanism performed by the hydrolase AntI. Knowing these mechanistic insights, we were further interested in finding substrates for AntI-inhibition.

Biochemical analysis of Quiquin Zhou with the minimal PKS revealed the exclusive activation of the AQ-ACP by the PPTase AntB. The PPTase was insoluble alone but gets stabilized by the probably inactive CoA-ligase working as a chaperone. Again, X-ray data of the Groll group supported the biochemical analysis and furthermore enlightened the minimal PKS caught in addition, compromising ACP:KS:CLF and ACP:MCAT complexes. The work within this thesis is planned to confirm a polyketide, which was bound to the KS:CLF complex after heterologous production. Furthermore, point mutations of received mutants should confirm possible protein-protein interaction sites of the structurally solved minimal PKS. Additionally, guided by the protein crystal structure of the ACP:MCAT-complex, a possible induced-fit-theory should be investigated biochemically to confirm a monodirectional transfer reaction by the MCAT.^[154]

7 Publications and Manuscripts

7.1 An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments

Authors:

Gina L. C. Grammbitter,[†] Maximilian Schmalhofer,[‡] Kudratullah Karimi,[§] Yi-Ming Shi,[†] Tim A. Schöner,[†] Nicholas J. Tobias,[†] Nina Morgner,[§] Michael Groll,[‡] Helge B. Bode^{*†}

[†] Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe-Universität Frankfurt am Main and Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universität Frankfurt, Max-von-Laue-Straße 9 and 15, 60438 Frankfurt am Main, Germany.

[‡] Center for Integrated Protein Science at the Department Chemie, Lehrstuhl für Biochemie, Technische Universität München, Lichtenbergstr. 4, 85748 Garching, Germany.

[§] Institut für Physikalische und Theoretische Chemie, Goethe-Universität Frankfurt, Max-von-Laue-Straße 7, 60438 Frankfurt am Main, Germany

* corresponding author

Published in:

J. Am. Chem. Soc. 2019 141 (42), 16615-16623

Reprinted with permission from:

Grammbitter, G. L.-C., Schmalhofer, M., Karimi, K., Schöner, T.-A., Tobias, N.-J., Morgner, N., Groll, M., and Bode, H.-B. (2019) An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments, *J. Am. Chem. Soc.*

DOI: 10.1021/jacs.8b10776. Copyright © 2019, American Chemical Society

Online access:

<https://doi.org/10.1021/jacs.8b10776>

Attachments:

Declaration on the contribution of the authors and the publication.

7.2 Biosynthesis of the Multifunctional Isopropylstilbene in *Photorhabdus laumondii* Involves Cross-Talk between Specialized and Primary Metabolism

Authors:

Gina L. C. Grammbitter,^{1‡} Siyar Kavakli,^{1‡} Helge B. Bode^{1,2*}

¹ Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe-Universität Frankfurt, Frankfurt am Main, Germany

² Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universität Frankfurt, Frankfurt am Main, Germany

‡ authors contributed equally

* corresponding author

Status:

Submitted

Attachments:

Declaration on the contribution of the authors and the manuscript.

7.3 Anthraquinone Production is Influenced by Cinnamic Acid

Authors:

Gina L. C. Grammbitter¹, Helena Vural¹, Nicholas J. Tobias² and Helge B. Bode^{1,2,3*}

¹ Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe-Universität Frankfurt, Frankfurt am Main, Germany

² LOEWE Center for Translational Biodiversity in Genomics (TBG), Germany

³ Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universität Frankfurt, Frankfurt am Main, Germany

* corresponding author

Status:

Submitted

Attachments:

Declaration on the contribution of the authors and the manuscript.

7.4 Structural Snapshots of the Minimal PKS System Responsible for Octaketide Biosynthesis

Authors:

Alois Bräuer,¹ Qiuqin Zhou,² Gina L. C. Grammbitter,² Maximilian Schmalhofer,¹ Ville R. I. Kaila,^{1,3} Helge B. Bode,² Michael Groll^{1*}

¹ Center for Integrated Protein Science Munich (CIPSM) at the Department Chemie, Technische Universität München, Lichtenbergstraße 4, 85748 Garching, Germany.

² Merck Stiftungsprofessur für Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, Max-von-Laue-Strasse 9, Frankfurt am Main, 60438, Germany.

³ Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden.

* corresponding author

Status:

Under revision

Attachments:

Declaration on the contribution of the authors and the manuscript.

7.5 Molecular Mechanism of Polyketide Shortening in Anthraquinone Biosynthesis of *Photorhabdus laumondii*

Authors:

Qiuqin Zhou,^a Alois Bräuer,^b H el ene Adihou,^a Maximilian Schmalhofer,^b Patricia Saura, Ville R. I. Kaila,^b Gina L. C. Grammbitter,^a Michael Groll*^b and Helge B. Bode*^a

a Molekulare Biotechnologie Fachbereich Biowissenschaften, Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universit at Frankfurt, Max-von-Laue-Str. 9, Max-von-Laue-Str. 15, 60438 Frankfurt a. M., Germany

b Center for Integrated Protein Science Munich (CIPSM), Department of Chemistry, Technische Universit at M unchen, Lichtenbergstra e 4, 85748 Garching (Germany)

* corresponding author

Published in:

Chem. Sci., 2019, 10 (25), 6341-6349

Reprinted with permission from:

Q. Zhou, A. Br auer, H. Adihou, M. Schmalhofer, P. Saura, G. L. C. Grammbitter, V. R. I. Kaila, M. Groll and H. B. Bode, *Chem. Sci.*, 2019, 10, 6341

DOI: 10.1039/C9SC00749K

Attachments:

Declaration on the contribution of the authors and the publication.

8 Additional Results: Tracing the Full-length APE

In the publication ‘An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments’ (7.1) the structural core of APE compounds produced from *X. doucetiae* (APE_{Xd}) was obtained by hydrolysis and after isolation of the main compound, the structure was elucidated by NMR (see Figure 24).

This hydrolysis procedure includes a mild transesterification after cell pellet extraction, as described by Cimermancic *et al.* for the isolation of APE compounds.^[138] The authors used this extraction method for further purification and structure elucidation of APEs from *E. coli* CFT073 (APE_{Ec}) and from *Vibrio fischeri* (APE_{Vf}), which were classified respectively as members from APE 1 and APE 2 subfamily according to their BGC structure (compare Figure 23).

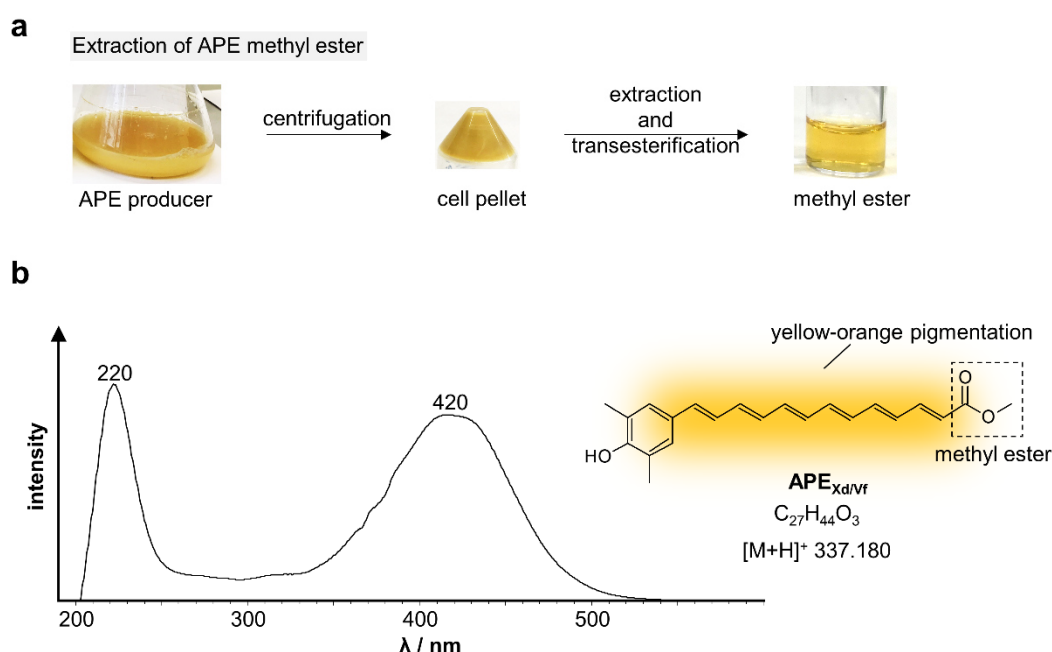


Figure 24. Generation of APE_{Xd}-methyl ester. General extraction procedure from the cell pellet followed by transesterification reaction (**a**). Characteristic absorbance spectrum of APE_{Xd}-methyl esters with a maximum at 420 nm (left). Structure of main APE_{Xd}-methyl ester isolated from *X. doucetiae* (right) (7.1) (**b**).

The authors further made some suggestions about the lipid behavior of APE molecules by comparing them to xanthomonadins, which were most likely found in the bacterial membrane. Xanthomonadins are a well-studied products from the APE subfamily 3.^[145] However, APE compounds were described as small molecules. According to some ‘lipid-genes’ within the APE BGC, such as genes encoding for a membrane transporter, (acyl)transferases and several genes encoding hypothetical proteins (compare turquoise genes in Figure 25)) and the results from our biochemical studies (see 7.1), allowed us to speculate that the APE moiety might be further processed to result in a lipid-like structure. Thus, this chapter deals with investigations all around the presence of an APE-lipid structure in *X. doucetiae*.

8.1 Switching On the APE-Production in *X. doucetiae*

Markerless promotor exchange in front of *apeB* induces the silent APE BGC in a secondary metabolite deficient *hfq*-background.^[155] Hfq is an RNA-chaperon and causes the absence of secondary metabolite production upon deletion.^[156] A yellow-pigmented phenotype results from an induced *ape*-promotor exchange mutant strain in a Δhfq -background (APE-producer) (Figure 25).

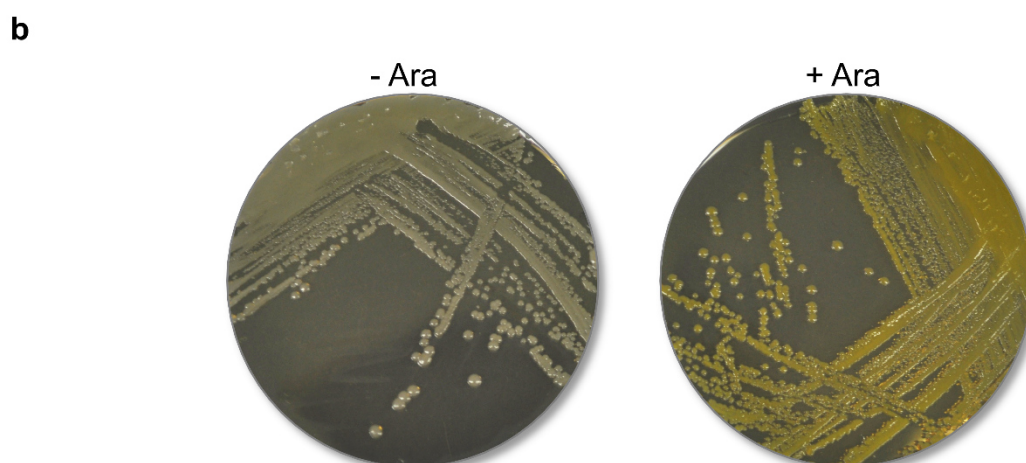
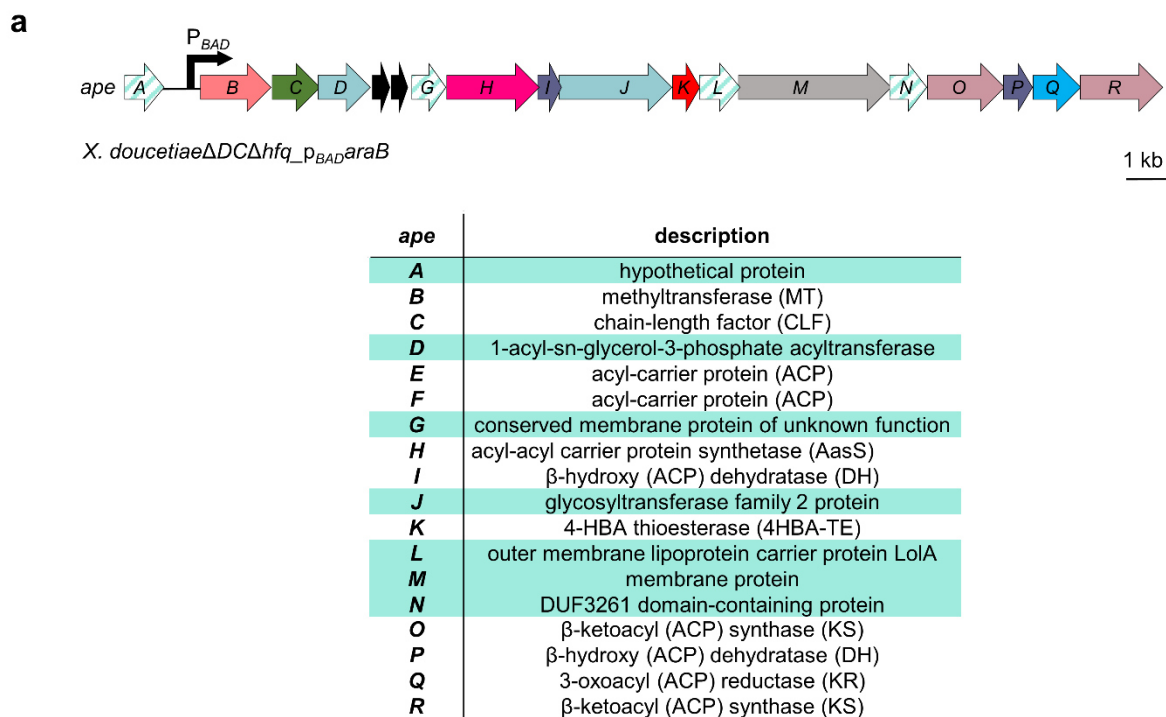


Figure 25. Non-induced vs. induced *ape*-promotor exchange mutant strain. Promotor exchange (P_{BAD} -promotor system) in front of *apeB* in *X. doucetiae*ΔDCΔhfq resulted in the 'APE-producer'. Gene functions are assigned according to blastp and biochemical studies from 7.1. Based on this, we postulated 'lipid-genes' involved in post-processing of the APE-core to a lipid-like structure (see table, highlighted in turquoise) (a). Induction of the APE-producer with 0.2% arabinose results in yellow pigmentation, observable on LB-agar plates if compared with the non-induced mutant (b).

8.2 Identification of Unusual Conjugated-FAs

Adapting the cell pellet extraction and transesterification procedure of Cimermanic *et al.*, some unknown hydrophobic compounds with specific absorbance in the lower absorbance range (> 400 nm) appeared during HPLC-analysis of the esterified APE_{Xd} compounds (further APE_{Xd}-methyl ester). Separation of the APE_{Xd}-methyl ester and those hydrophobic compounds was achieved using C3 reversed-phase chromatography (see Figure 26). Due to the absence of the hydrophobic compounds in the non-induced APE-producer strain, it was proposed that they might be (part of) a possible structural element of the postulated APE-lipid.

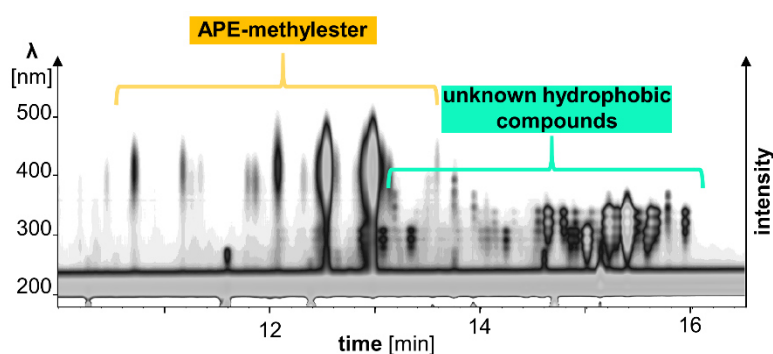


Figure 26. APE vs. hydrophobic compounds. Esterification of the APE-producer extract resulted in the detection of APE_{Xd}-methyl ester (UV_{max} 420 nm) and revealed unknown hydrophobic compounds with a specific absorbance in the Vis-range. Shown is the survey-view in the range 190–600 nm.

The masses of the hydrophobic compounds are distributed around m/z 400 and the detection was achieved in positive ionization mode. Because the corresponding MS^1 and MS^2 -spectra of the hydrophobic compounds gave no hint in any database, the main compound with an m/z of 399 was purified. The purification-strategy included the cell pellet extraction followed by a transesterification.^[138] The resulting crude extract was subjected to Sephadex LH-20 size exclusion and several rounds of reversed-phase chromatography.

The fractionating during size exclusion chromatography resulted in an initial separation of the different APE_{Xd}-methyl esters vs. hydrophobic compounds via HPLC-MS analysis (Figure 27). During size exclusion, the hydrophobic compounds are eluting first, followed by the APE_{Xd} methyl esters, indicating that the hydrophobic molecules share a higher molecular weight than APE_{Xd} methyl esters, which is already represented by the major APE_{Xd} methyl ester with a mass of m/z 337 and the main hydrophobic compound with a mass of m/z 399.

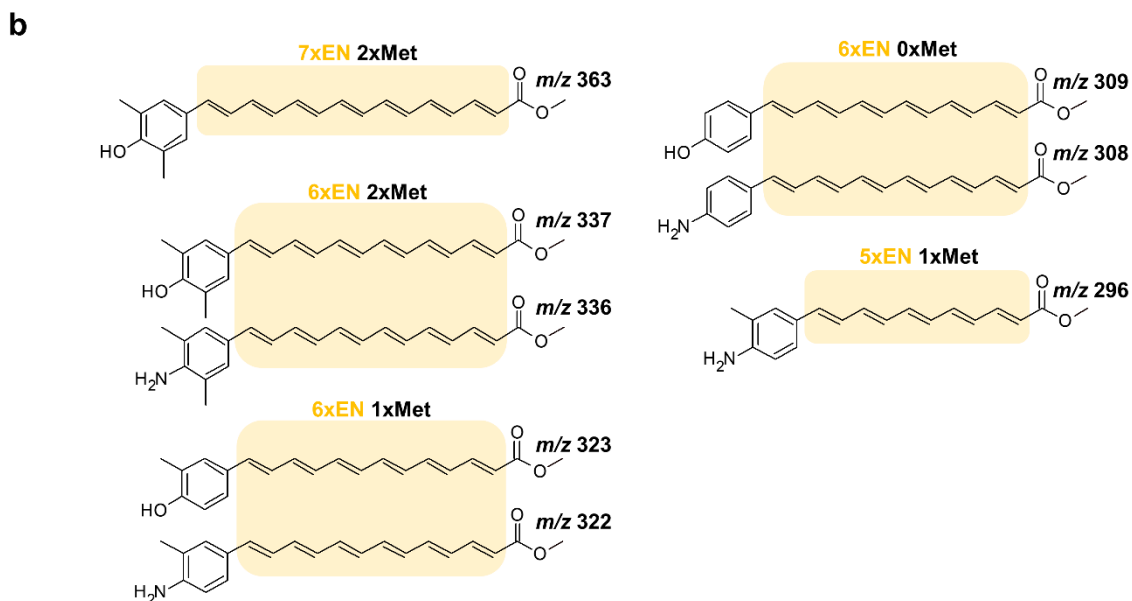
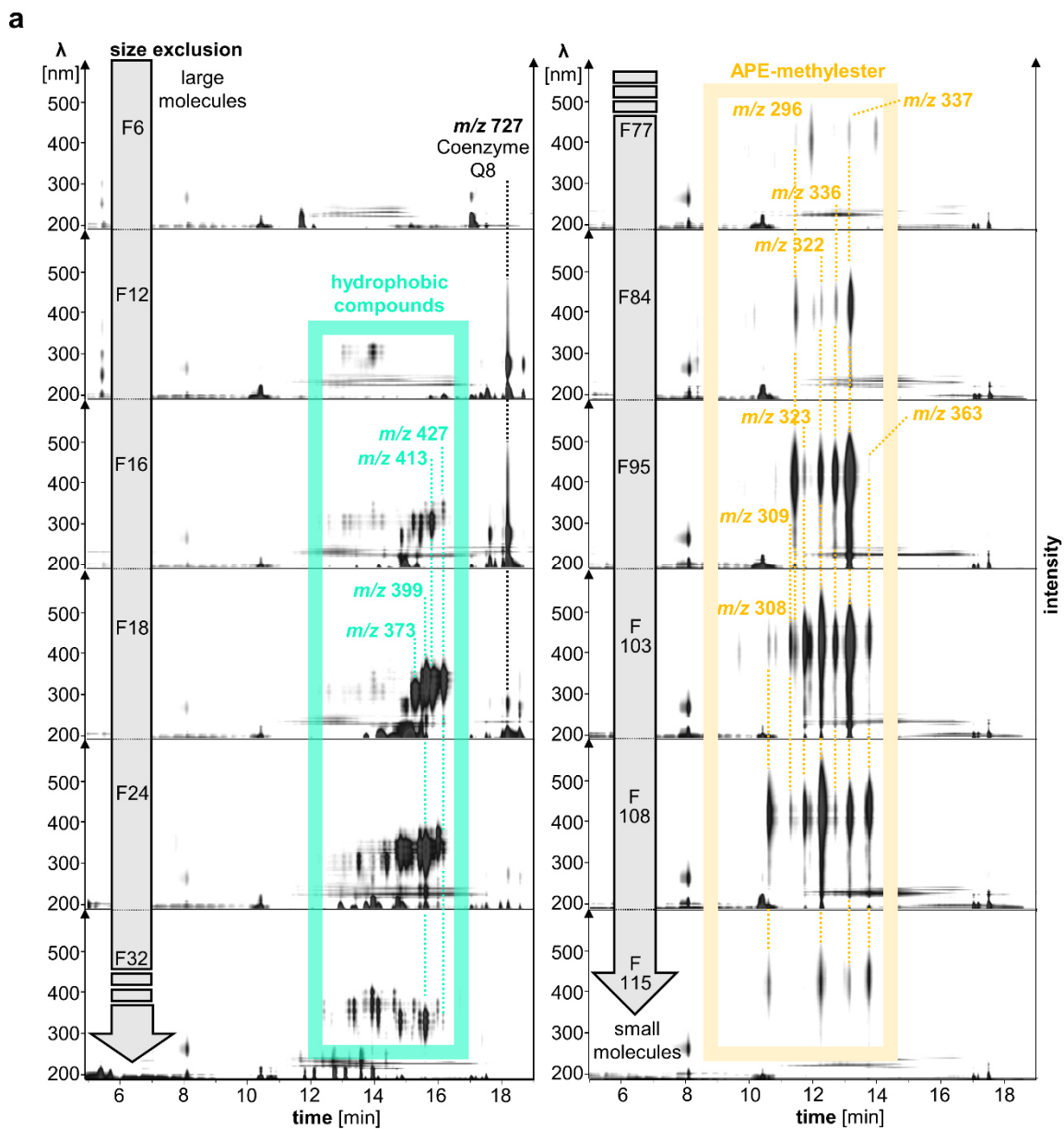


Figure 27. HPLC-UV/MS-analysis of esterified APE-producer extract derived from size exclusion chromatography. Depicted are samples from sephadex fractions F6, F12, F18, F24, F32, F77, F84, F95, F103, F108, F115 (a). Assigned APE_{x_d}-methyl ester with different chain-length (polyene core, EN) and methylations (Met), corresponding network-analysis, see Figure S1 (b). Shown is the survey-view with 190-600 nm.

By performing MS-based network analysis with the obtained HPLC-MS data, the APE_{x_d}-methyl esters and the hydrophobic compounds were grouped according to their specific MS²-spectra (see Figure S1).

Like that, the classification of APE_{x_d}-methyl esters from the publication in 7.1 was complemented with the obtained data (see Figure S1 and Table S1). Interestingly, not only *para*-substituted hydroxy-APEs were assigned but also aminated-APE_{x_d}-methyl ester (Figure 27b). Both precursors, *p*-hydroxybenzoic acid (HBA) and *p*-aminobenzoic acid, can be obtained from the shikimate pathway.^[47] A complete assignment of the APE_{x_d}-methyl esters was not possible but the mass shifts of ± 14 Da presented in the network analysis indicate differences in the methylation of the phenyl ring of the APE-methyl ester, a ± 26 Da indicates a loss or gain of a CH=CH-group in the polyene core and a ± 28 Da shift indicates a loss or gain of a CH₂-CH₂-group in a fully saturated alkyl chain (compare Figure S1). This is also true for the unknown hydrophobic compounds. According to the generated data, it was demonstrated again that APE_{x_d}-methyl esters with seven double bonds are the maximum length being synthesized from the APE-PKS II machinery, as it was already shown in section 7.1. A molecule, always being co-extracted, was Coenzyme Q₈ (727 [M+H]⁺, see Figure 27a), which was assigned using METLIN online database^[157,158] according to its specific MS².

Sephadex fractions containing the hydrophobic compounds were further separated by three reversed-phase chromatographic separations resulting in pure compound '399', which was further structure elucidated using NMR-techniques (Figure 28 and Figure S2-6 and Table S2).

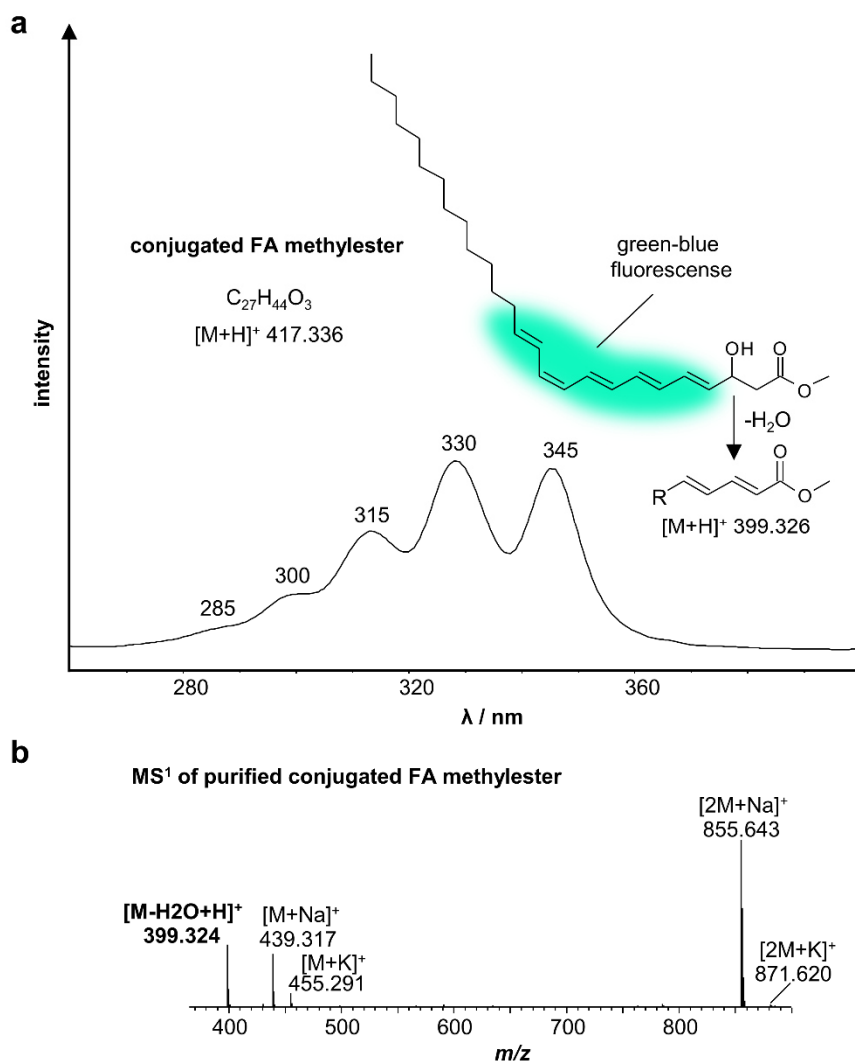


Figure 28. Characteristics of compound 399. Compound 399 was structure elucidated through NMR. The compound with a sum formula of $C_{27}H_{44}O_3$ shows a specific absorbance spectrum with maxima in (285 nm,) 305 nm, 315 nm, 330 nm and 345 nm. If the compound gets excited with 337 nm, it shows a natural green-to-blue fluorescence (a). MS¹-analysis of isolated compound '399' resulted in a several adducts: 399 [M-H₂O+H]⁺, 439.217 [M+Na]⁺, 455.291 [M+K]⁺, 855.643 [2M+Na]⁺, 871.620 [2M+K]⁺.

The purified compound 399 gives rise to a molecular formula of $C_{27}H_{44}O_3$ with a mass of 417 [M+H]⁺ and represents a fatty acid (FA) with an unusual C₁₀-C₁₁ *cis*-double bond in a large conjugated system (further conjugated-FA, see Figure 28a). The *cis*- Δ^{11} was hereby confirmed by the coupling constants of H-11/H-12 (Table S2). The difference of 18 Da from 417 [M+H]⁺ to the detected 399 [M+H]⁺ is attributed to a loss of water. 417 [M+H]⁺ was not detectable during MS-analysis, but instead the sodium- and potassium-adduct with a mass of 439 [M+Na]⁺ and 455 [M+K]⁺ were detected, respectively. A strong signal was observed for the corresponding [2M]⁺-ion of the sodium adduct with a mass of 855.643 [2M+Na]⁺ and lower signals for the potassium adduct 881.659 [2M+Ka]⁺ (Figure 28b). Due to the conjugated system, the FA shows a green-to-blue fluorescence.

8.3 Identification of APE-Lipids

A possible linkage of both, the APE_{Xd}-methyl ester and the conjugated-FA resulting in an APE-lipid molecule was further investigated by measuring the molecules without performing the transesterification. Due to the yellow-orange pigmentation of the APE_{Xd} moiety, it was easy to track the presence of the whole APE-lipid molecule. Unfortunately, the APE-lipid got stuck at the reversed-phase column using C18 and C3 material. The use of a solvent system typically used for phospholipids in combination with a C3-column resulted in the elution and separation of the postulated APE-lipids. A comparison of APE_{Xd}-methyl ester versus APE-lipids is represented in Figure 29.

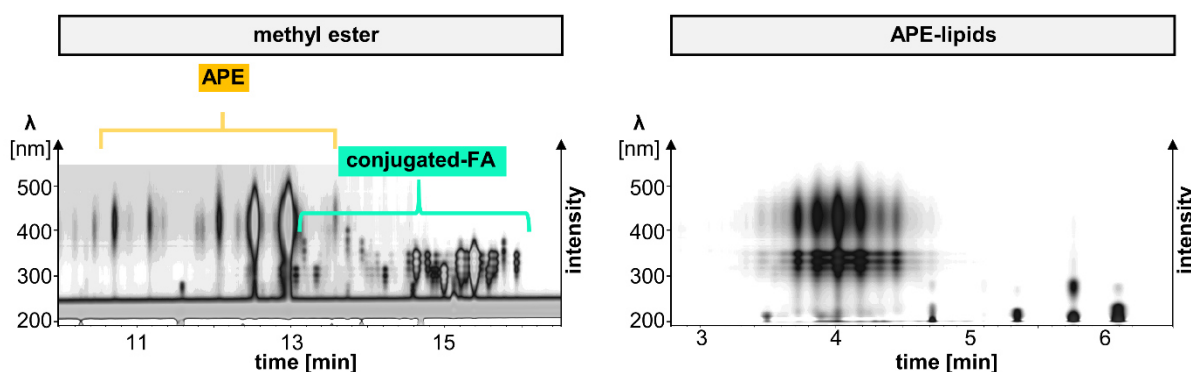


Figure 29. Separation of APE-lipid compounds. Comparison of HPLC-UV/MS data of generated APE_{Xd}-methyl ester and APE-lipids from the *X. doucetiae* APE-producer (survey view).

Indeed, the HPLC-UV-MS analysis supported the assumption of a probable linkage of the APE_{Xd}- and conjugated-FA moiety to result in an APE-lipid; characteristic UV- signals for the APE_{Xd}- and conjugated-FA moieties were detected at the same retention time, indicating that the corresponding molecules are connected to each other (Figure 29). Unfortunately, no new masses showed up during HPLC-MS analysis of the induced promotor exchange mutant in comparison with the non-induced control. Several attempts to purify one of the APE-lipids failed due to the high amount of impurities of phosphatidylethanolamines and probable other lipids (Figure S7) and only low amounts of the target compounds. The use of a combination of additives (formic acid and ammonium formate) ^[159] and changes in the ionization mode (see 8.4.1.2) enlightens for the first time a major mass of m/z 1285, which was not present in the uninduced sample.

Simple *in silico*-addition of the masses from the APE plus the conjugated-FA part does not reach an overall mass of 1285 Da. Thus, it was suggested that an unknown structural element is additionally involved in the formation of an APE-lipid. The absence of an additional UV signal of the APE-lipids (Figure 29) indicates that this unknown moiety is most likely UV-inactive.

Further deletion studies of the *X. doucetiae* APE-producer with targeted deletions in *apeB-R* were supposed to accumulate intermediates and provide information about the biosynthetic order of the APE-lipids. In Table 1 the results of the deletion studies regarding the two known building blocks, including the APE and conjugated-FA moieties are summarized (HPLC-UV spectra, see Figure 30).

Table 1. Summarized HPLC-MS/UV analysis of APE-producer mutants subjected to transesterification. Protein functions are assigned according to blastp. Signals corresponding to the APE_{Xd}-methyl ester are depicted as 'signal 420 nm'; signals corresponding to the conjugated-FA methyl ester are depicted as 'signal 305-315-330 nm'. For the corresponding survey-view of respective APE-producer mutants, see Figure 30.

mutant Δ	Function	signal 420 nm	signal 305-315-330 nm
<i>apeB</i>	MT	+	+
<i>apeC</i>	CLF	-	+
<i>apeD</i>	1-acyl-sn-glycerol-3-phosphate AT	-	-
<i>apeE</i>	ACP	-	+
<i>apeF</i>	ACP	-	-
<i>apeG</i>	membrane protein	-	+
<i>apeH</i>	AasS	-	-
<i>apeI</i>	DH	-	-
<i>apeJ</i>	glycosyltransferase	-	+
<i>apeK</i>	TE	-	+
<i>apeL</i>	LolA-like	+	+
<i>apeM</i>	membrane protein	+	+
<i>apeN</i>	unknown	-	-
<i>apeO</i>	KS	-	-
<i>apeP</i>	DH	-	-
<i>apeQ</i>	KR	-	-
<i>apeR</i>	KS	-	-

*less derivatives are produced **other APE_{Xd}-methyl ester derivatives are produced

The deletion of genes, which were already tested to be involved in the biosynthesis of the APE moiety (in 7.1) *apeE*, *apeF*, *apeH*, *apeI*, *apeO*, *apeP*, *apeQ* and *apeR*, resulted in the phenotypic loss of pigmentation, which is caused by the APE moiety and the absence of the *m/z* 1285 in MS¹ (Figure 31 and Figure 30)

Additionally, the deletion of *apeD*, *apeG* and *apeN*, encoding for a glycerol-3-phosphate AT, a membrane protein and a protein of unknown function, resulted in a phenotypic loss of pigmentation, too. The deletion of the MT encoding gene *apeB*, postulated to be responsible for the methylation of the aryl moiety in the APE-molecule, resulted in the production of less APE-lipid derivatives and a slightly pigmented phenotype, with the exclusive production of unmethylated APE-lipid, resulting in a mass of m/z 1256 (Figure 31). By analyzing the APE_{Xd}-methyl esters of the induced APE-producer over time, it was concluded that methylation does not occur post-PKS biosynthesis, as the non-methylated product was absent in early cultivation states (Figure S8).

Interestingly, some mutants with an observed non-pigmented phenotype showed obviously an UV-loss of the APE moieties but the presence of the signals of the conjugated-FA moieties. Noteworthy, this phenomenon was not observed *vice versa*. Due to a stronger interaction of the accumulated intermediates with the reversed phase material, compared to the conjugated-FA methyl esters, we suggested that these compounds represent intermediates, which indeed harbor the conjugated-FA but are further linked to a most probably UV-Vis-inactive, unknown molecule. Mutants, in which this intermediate with the Vis-profile of the conjugated-FA appeared are deletions in *apeC*, *apeE*, *apeJ* and *apeK* and *apeG*, encoding for a CLF, an ACP, a glycosyltransferase, a 4-HBA-TE and a membrane protein, respectively. In comparison to the *apeE*-, *apeK* and *apeG*-deletions, the production of the conjugated-FA plus the unknown part (further conjugated-FA+x) is increased in the *apeJ*- and *apeC*-deletions (Figure 30).

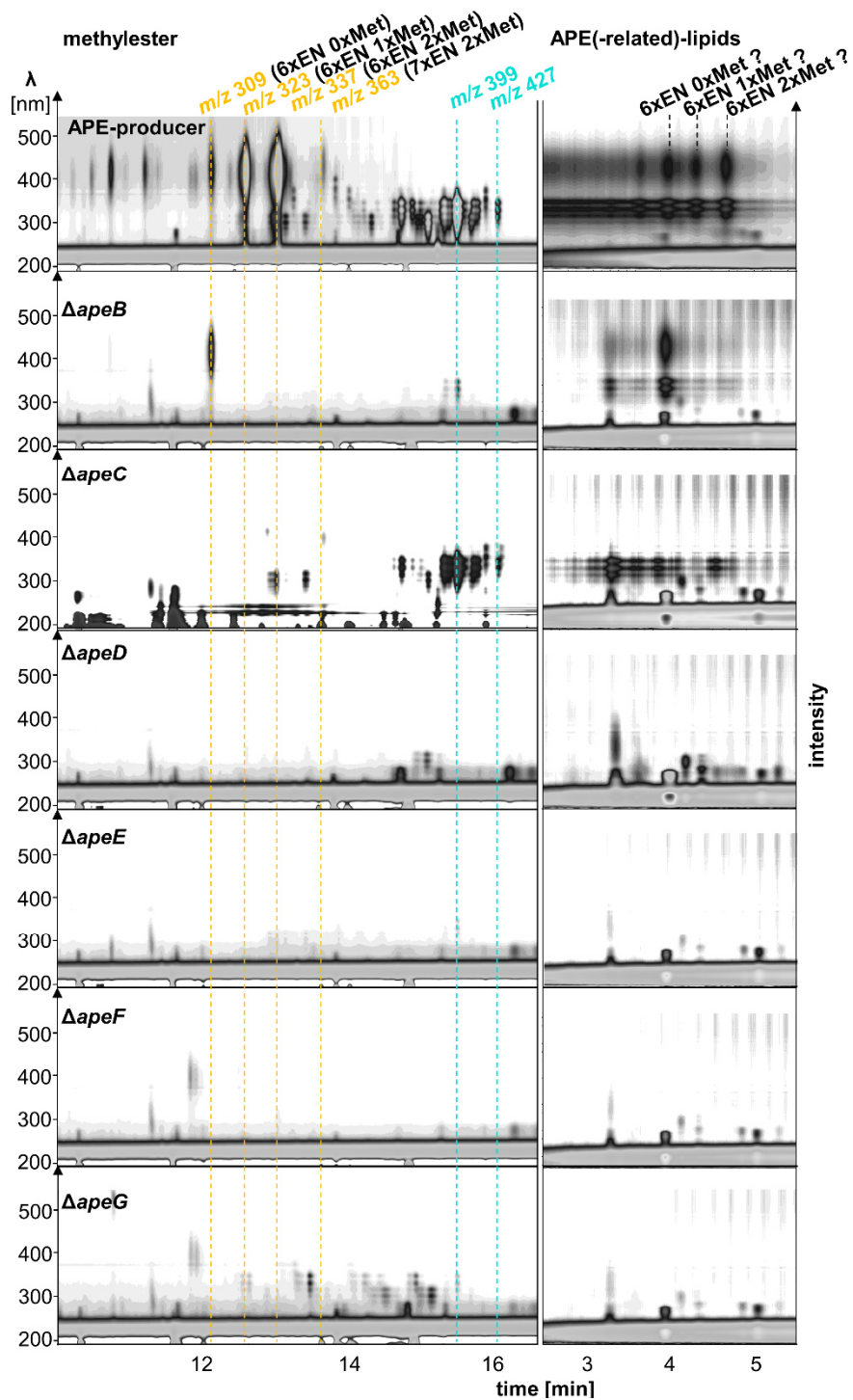


Figure 30. HPLC-UV/MS analysis of APE-producer mutants showing deletions of *apeB-G* genes. Shown in survey-view are the APE-lipids (right) and the corresponding APE_x-methyl ester after hydrolysis (left).

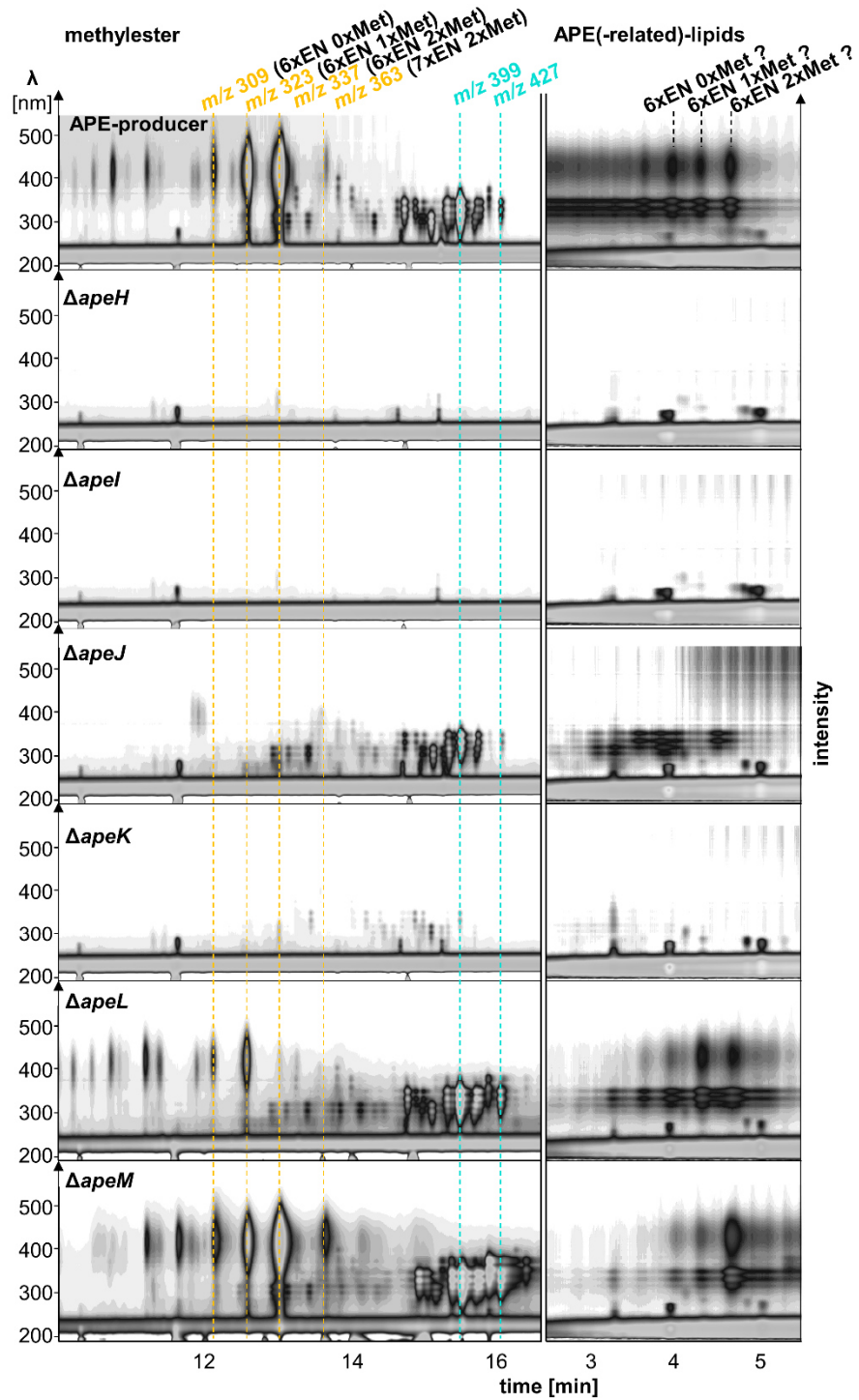


Figure 30 ff. HPLC-UV/MS analysis of APE-producer mutants showing deletions of *apeB*, *apeH-M* genes. Shown in survey-view are the APE-lipids (right) and the corresponding APE_x-methyl ester after hydrolysis (left).

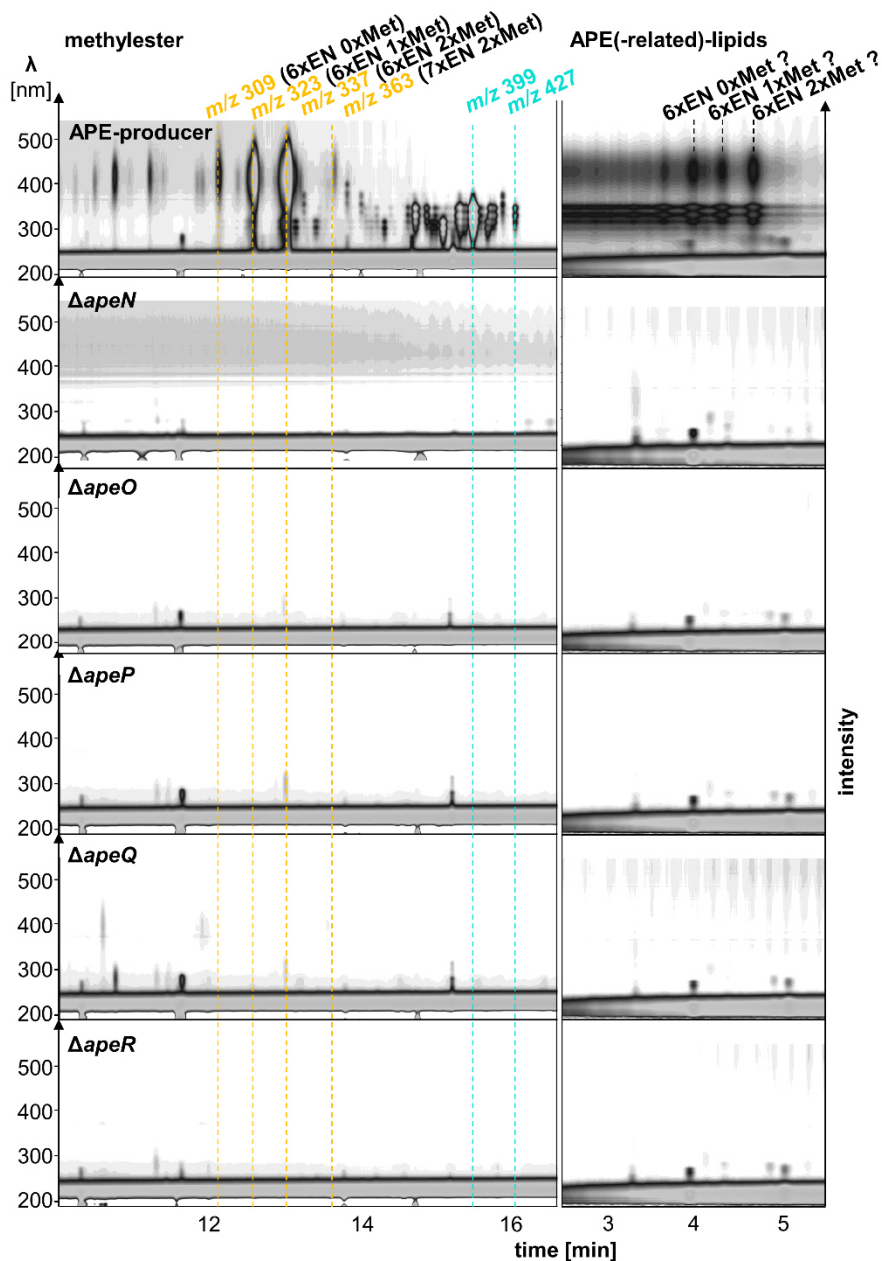


Figure 30 ff. HPLC-UV/MS analysis of APE-producer mutants showing deletions of *apeB*, *apeN-R* genes. Shown in survey-view are the APE-lipids (right) and the corresponding APE_{x_d}-methyl ester after hydrolysis (left).

During comparison of the APE-lipid extracts with their corresponding APE_{x_d}-methyl ester, it was possible to assign the differences in the APE-profile of the mutants of *ΔapeL* and *ΔapeM*. No significant difference in the distribution of the conjugated-FAs was detected, but variations in the APE moiety were noticed (see Figure 30). The deletion of *apeM* leads to an overall increase of the 6xEN-2xMet-APE_{x_d} methyl ester, while the deletion of *apeL* rather produced the 6xEN-1xMet-APE_{x_d} methyl ester (compare Figure 27b).

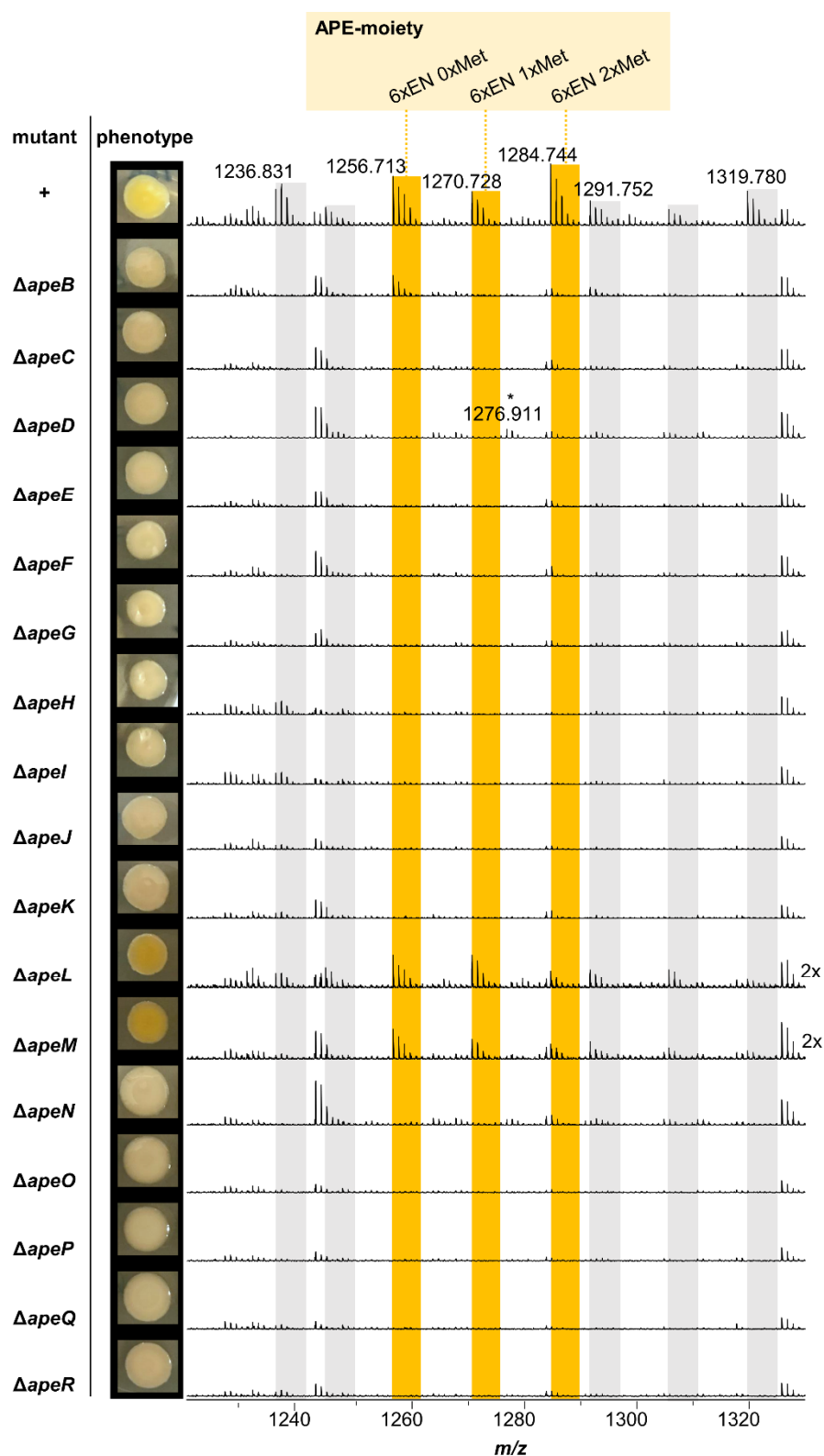


Figure 31. Detection of APE-lipids in the APE-producer mutants Δ *apeA-R*. MS¹ analysis of APE-producer mutants (right) with the corresponding phenotype on LB agar plates (left). Signals corresponding to molecules, probably harboring the APE_x_d moieties, are highlighted in yellow (6xEN 2xMet, 6xEN 1xMet, 6xEN 0xMet), while highlighted in grey are unknown signals, only arising in the yellow-pigmented phenotypes. A further unknown signal, present in the *apeD*-mutant extract, is marked with an asterisk. For comparison of the corresponding APE_x_d- and conjugated-FA methyl esters of the APE-producer mutants, see also Table 1 and the respective survey-view data in Figure 30.

Curiously, the yellow-pigmented APE-lipids do not show any fluorescence. In contrast, during thin-layer chromatography, a quenching effect from the reported fluorescence of the conjugated-FA moiety was observed (Figure 32a). This quenching ability was further demonstrated with an extract of the APE-producer, compared to an extract of the APE-producer $\Delta apeJ$ -mutant, with the latter one producing the conjugated-FA+X-molecule (Figure 32).

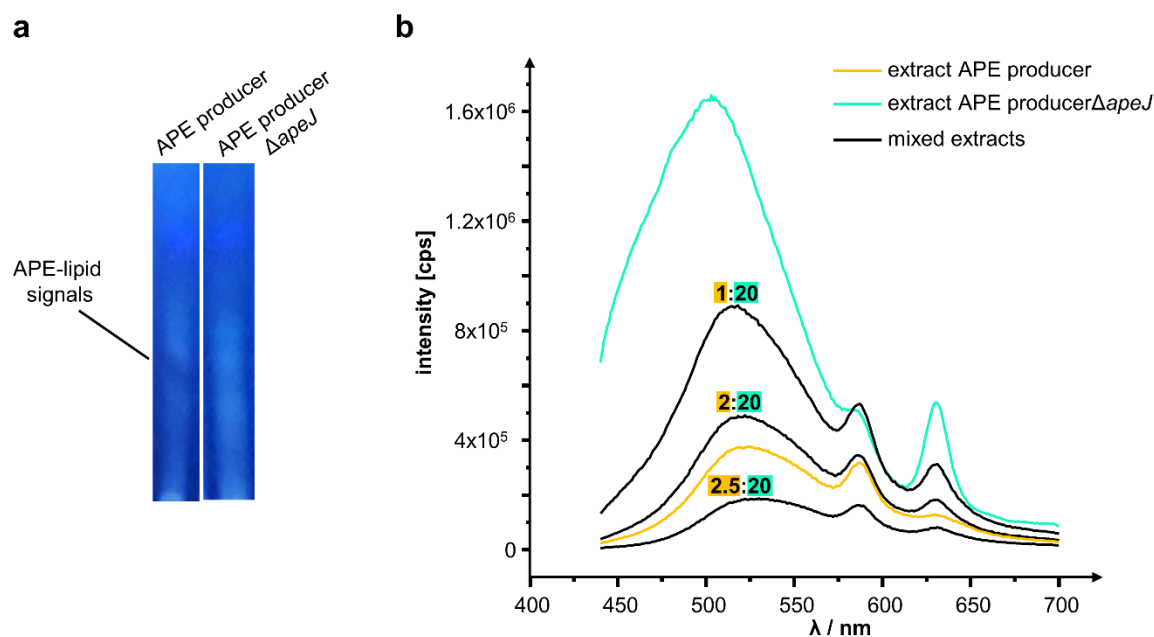


Figure 32. Quenching effect of APE-lipids. The extract of the APE-producer $\Delta apeJ$ shows fluorescence activity due to the accumulation of conjugated-FA. During TLC analysis of the extract from the APE-mutant (APE-producer) compared to the extract from the APE-mutant with a deletion of *apeJ*, a quenching effect of this fluorescence was detectable under excitation of 337 nm (a). The quenching effect, probably caused by APE-lipids, was further measurable with titrating minor amounts of extract from the APE-producer extract to extract from the APE-producer $\Delta apeJ$ (b).

Titration of the extract from the APE-producer in different ratios to the extract of the $\Delta apeJ$ -mutant indeed resulted in a quenching effect of the fluorescence of the *apeJ*-mutant extract (Figure 32b).

Unfortunately, no mass of the conjugated-FA+x-compound was detectable. Due to low amounts of this intermediate, a cell pellet of a 40 L fermentation of the *apeJ*-mutant was used for further isolation attempts. An enrichment of the conjugated-FA+x-compound was achieved using the combination of size exclusion and normal phase chromatography and resulted in the detection of a conjugated-FA+x-compound with a mass of m/z 886. Further purification attempts with reversed-phase failed, as the fluorophoric compound degraded during the purification process, resulting in the loss of fluorescence, and furthermore, due to remaining lipid impurities, in untraceable compound pieces.

Nevertheless, the MS²-spectrum of the APE-lipid (m/z 1285) can be compared with the MS-spectra of the different structural elements of the main APE_{Xd}-methyl ester (m/z 377), the conjugated-FA-methyl ester (m/z 399) and the conjugated-FA+x-compound (m/z 886) (Figure 33 and Figure 34).

Besides the characteristic MS²-signals for the APE moiety and the conjugated-FA during MS²-analysis of m/z 1285 (Figure 33, APE-signals, yellow; conjugated-FA signals, turquoise), signals corresponding to a glucosamine (GlcNAc) sugar moiety, detectable as monosaccharide^[160] were assigned (Figure 34, green), representing one of the UV inactive moieties of the APE-lipid. Additionally, a neutral loss of 142.019 Da, corresponding to a phosphoethanolamine, was detected (Figure 34).

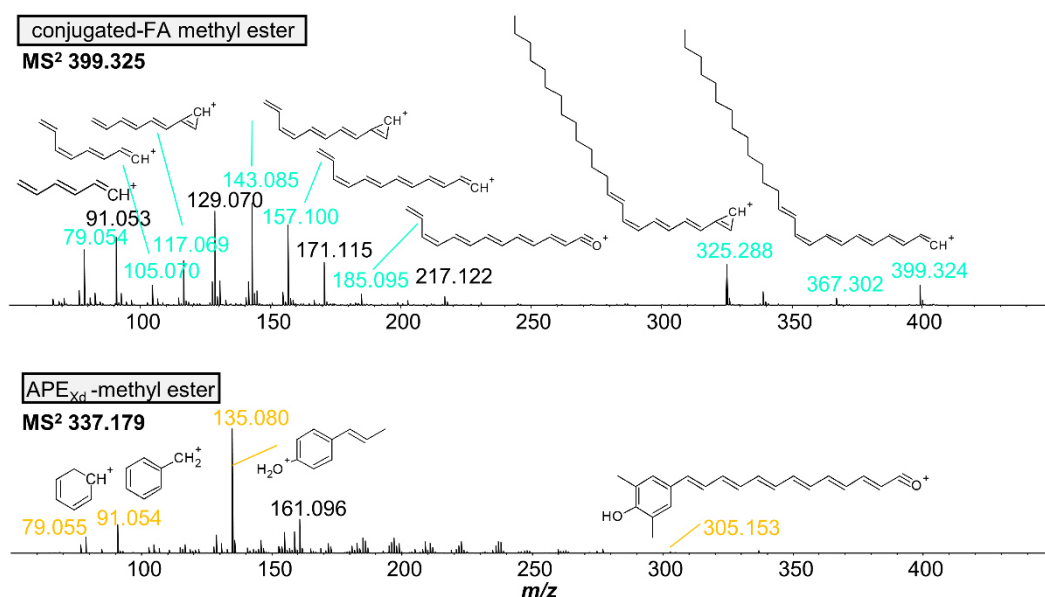


Figure 33. MS² data structural elements: purified APE_{Xd}-methyl ester (337.179 [M+H]⁺) and conjugated-FA (399.325 [M+H]⁺) in comparison to the APE-lipid with an m/z of 1284.744 and partially purified conjugated-FA+x with an m/z of 866.606 (see Figure 34).

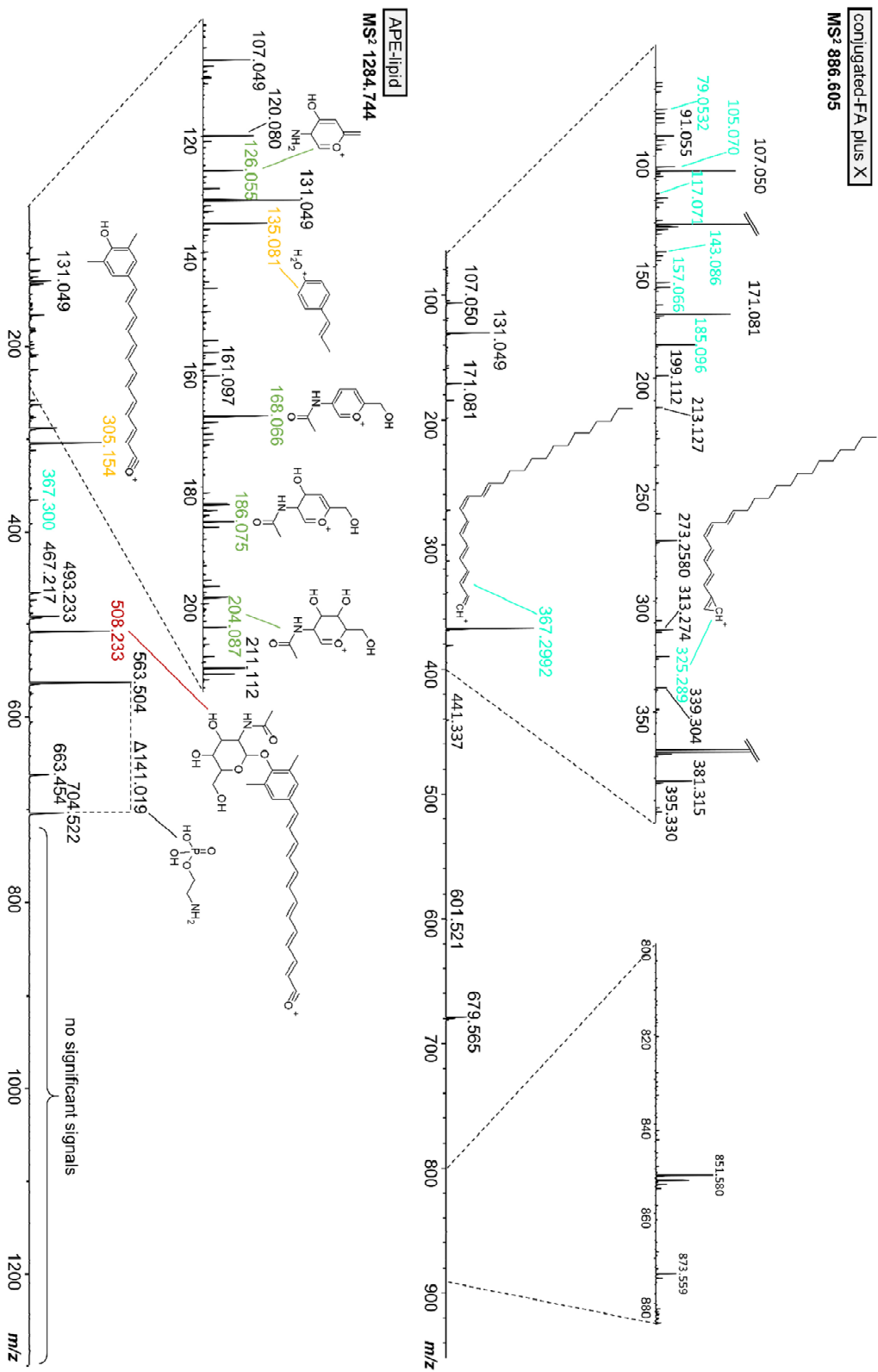


Figure 34. MS² data of the APE-lipid with an *m/z* of 1284.744 in comparison to structural elements: purified APE_x-methyl ester (337.179 [M+H]⁺) and conjugated-FA (399.325 [M+H]⁺) (Figure 33) or partially purified conjugated-FA+x with an *m/z* of 866.606.

The HR-data of the assigned signals from Figure 33 and Figure 34, corresponding to the different structural elements, are further summarized in Table 2.

Table 2. MS²-signals of structural elements of APE-lipids, corresponding spectra, see Figure 33 and 34.

structural element	sum formula [M] ⁺	theoretical mass [m/z]	detected mass [m/z]	Δppm
APE-lipid				
sugar (GlcN)	C ₆ H ₈ NO ₂	126.0550	126.0548	1
	C ₈ H ₁₀ NO ₃	168.0655	168.0659	2
	C ₈ H ₁₂ NO ₄	186.0761	186.0745	8
	C ₈ H ₁₄ NO ₅	204.0880	204.0866	7
APE	C ₉ H ₁₁ O	135.0804	135.0808	3
	C ₂₁ H ₂₁ O ₂	305.1536	305.1535	0
conjugated-FA moiety	C ₂₆ H ₃₉ O	367.2995	367.2996	0
APE and sugar moiety	C ₂₉ H ₃₄ NO ₇	508.2328	508.2303	0
Phosphoethanolamine	C ₂ H ₈ NO ₄ P [*]	141.0191	141.0186	4
conjugated-FA+X				
conjugated-FA	C ₆ H ₇	79.0543	79.0532	13
	C ₈ H ₉	105.0699	105.0704	5
	C ₉ H ₉	117.0699	117.0705	5
	C ₁₁ H ₁₁	143.0855	143.0858	2
	C ₁₁ H ₉ O	157.0648	157.0656	6
	C ₁₃ H ₁₃ O	185.0961	185.0962	1
	C ₂₄ H ₃₇	325.2890	325.2893	1
C ₂₆ H ₃₉ O	367.2995	367.2992	1	
conjugated-FA				
conjugated-FA	C ₆ H ₇	79.0543	79.0555	16
	C ₈ H ₉	105.0699	105.0699	0
	C ₉ H ₉	117.0699	117.0701	2
	C ₁₁ H ₁₁	143.0855	143.0857	2
	C ₁₁ H ₉ O	157.0648	157.1013	1
	C ₁₃ H ₁₃ O	185.0961	185.0944	9
	C ₂₄ H ₃₇	325.2890	325.2892	1
C ₂₆ H ₃₉ O	367.2995	367.2997	1	
APE _{Xd} -methyl ester				
APE	C ₆ H ₇	79.0543	79.0543	5
	C ₇ H ₇	91.0542	91.0541	1
	C ₉ H ₁₁ O	135.0804	135.0805	0
	C ₂₁ H ₂₁ O ₂	305.1536	305.1534	1

*detected as neutral loss

8.4 Materials and Methods

Table 3. Strains used in this study.

Strain	Genotype	Reference
<i>E. coli</i> ST18- λ pir	Tp ^r Sm ^r , <i>recA thi hsdR⁺</i> RP4-2-Tc::Mu-Km::Tn7, <i>λpir</i> phage lysogen, Δ <i>hemA</i>	[161]
<i>X. doucetiae</i> DSM 17909 ^T	Wild type, amp ^r	DSMZ
<i>X. doucetiae</i> DSM 17909 ^T Δ <i>xrdE</i>	DSM 17909 ^T wild type with a deletion in <i>xrdE</i>	[162]
<i>X. doucetiae</i> DSM 17909 ^T Δ <i>xrdE</i> _P _{BAD} <i>apeA</i>	<i>X. doucetiae</i> DSM 17909 ^T Δ <i>xrdE</i> with a promoter exchange in front of <i>apeA</i> ; amp ^r , kan ^r	see 7.1
<i>X. doucetiae</i> DSM 17909 ^T Δ <i>xrdE</i> _P _{BAD} <i>apeB</i>	<i>X. doucetiae</i> DSM 17909 ^T Δ <i>xrdE</i> with a promoter exchange in front of <i>apeB</i> ; amp ^r , kan ^r	see 7.1
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i>	<i>X. doucetiae</i> DSM 17909 ^T wild type with a deletion in XDD1_RS09835 (decarboxylase) and <i>hfq</i> ; amp ^r ,	[162]
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> (APE-producer)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> , amp ^r ,	see 7.1
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeB</i> (APE-producer Δ <i>apeB</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeB</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeC</i> (APE-producer Δ <i>apeC</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeC</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeD</i> (APE-producer Δ <i>apeD</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeD</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeE</i> (APE-producer Δ <i>apeE</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeE</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeF</i> (APE-producer Δ <i>apeF</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeF</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeG</i> (APE-producer Δ <i>apeG</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeG</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeH</i> (APE-producer Δ <i>apeH</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeH</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeI</i> (APE-producer Δ <i>apeI</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeI</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeJ</i> (APE-producer Δ <i>apeJ</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeJ</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeK</i> (APE-producer Δ <i>apeK</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeK</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeL</i> (APE-producer Δ <i>apeL</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeL</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeM</i> (APE-producer Δ <i>apeM</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeM</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeN</i> (APE-producer Δ <i>apeN</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeN</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeO</i> (APE-producer Δ <i>apeO</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeO</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeP</i> (APE-producer Δ <i>apeP</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeR</i> and with a deletion of <i>apeP</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeQ</i> (APE-producer Δ <i>apeQ</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeQ</i> , amp ^r	this work
<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i> Δ <i>apeR</i> (APE-producer Δ <i>apeR</i>)	<i>X. doucetiae</i> DSM 17909 ^T Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> and with a deletion of <i>apeR</i> , amp ^r	this work

Table 4. Oligonucleotides used for plasmid construction and verification. Overhangs are underlined.

Plasmid		Oligonucleotide 5' to 3'	Template
pEB17_ΔapeB	GG143	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTTAATCG</u> <u>CCACTTGAATCCTC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG144	<u>CCAATGTCAATTTTCATGAAATACCGTTCGATATCCA</u> <u>TATATCCGATAGGTAGCATATAG</u>	
	GG145	<u>CTATATGCTACCTATCGGATATATGGATATCGAACG</u> <u>GTATTTTCATGAAATTGAC</u>	
	GG146	<u>CAATTTGTGGAATTCGCGGGAGAGCTCGGTCATGA</u> <u>GCAGACGCCAG</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17
pEB17_ΔapeC	GG241	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTGGATTT</u> <u>CACCTCAAGACGTTTG</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG242	<u>GTAATATTGGCTGAAATCCTTTTCATCATCGAAATA</u> <u>CCGTTTCGTTACTGCTTC</u>	
	GG243	<u>CTGCAAGAAGCAGTAACGAACGGTATTTTCGATGAT</u> <u>GAAAAGGATTTGACG</u>	
	GG244	<u>CAATTTGTGGAATTCGCGGGAGAGCTCCCCGTTTT</u> <u>CTGTATCAACC</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17
pEB17_ΔapeD	GG167	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTGTTTT</u> <u>ACAGCATTGATGATTTTACC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG168	<u>AACTCTTGCATCATCGTATCCGATATCTACTTATCG</u> <u>GCTCCAATTCAC</u>	
	GG169	<u>GATCATTGTCAGTGGAAATGGAGCCGATAAGTAGA</u> <u>TATCGGATACGATGATGCAAGAG</u>	
	GG170	<u>CAATTTGTGGAATTCGCGGGAGAGCTCGATAGTAC</u> <u>ATAAGAAGTTGATGATTGCG</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17
pEB17_ΔapeE	GG147	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTGTTGCC</u> <u>ATCTTGCACGTC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG148	<u>CTTTTGTTTTTCCATGATGGCACCTGCTCCATCGTA</u> <u>TCCGATATCTATTAGTTTG</u>	
	GG149	<u>CCGTCAAACATAAGATATCGGATACGATGGAGCA</u> <u>GGTGCCATCATGG</u>	
	GG150	<u>CAATTTGTGGAATTCGCGGGAGAGCTCATCCTTTCA</u> <u>ATGACAGGTGGC</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17
pEB17_ΔapeF	GG151	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTGATGCA</u> <u>GAAACTTTGCTCCC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG152	<u>CTGTTTGTTTTGCTGATGGTTTATGGCTGAGATGGC</u> <u>ACCTGCTCTTATG</u>	
	GG153	<u>CAGAAAGCATAAGAGCAGGTGCCATCTCAGCCATA</u> <u>AACCATCAGC</u>	
	GG154	<u>CAATTTGTGGAATTCGCGGGAGAGCTCTCTTCCGA</u> <u>TAGCCATTGG</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17
pEB17_ΔapeG	GG187	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTTAGATA</u> <u>TCGGATACGATGATGCAAGAG</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG188	<u>GATAGCCATTGGGAGATTGATCGTGCTTTTAAAC</u> <u>TACCCAAATGTTTGCTGTTG</u>	
	GG189	<u>CAAAACAAACAGCAAACATTTGGGTAGTTAATAAAA</u> <u>GCACGATCAATCTCCC</u>	
	GG190	<u>CAATTTGTGGAATTCGCGGGAGAGCTCTGCGGAAA</u> <u>GTCAAACCATAC</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17
pEB17_ΔapeH	GG211	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTGCAGTA</u> <u>AAAAATTGCTATGGC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG212	<u>CATGTTGAGCAGGTTGCTCGCGGTGAATTTCAATG</u> <u>GAAATGAGCAACC</u>	
	GG213	<u>CTATTAATTGGGTTGCTCATTTCATTGAAATTCACC</u> <u>GCGAGCAAC</u>	
	GG214	<u>CAATTTGTGGAATTCGCGGGAGAGCTCATATCAAAA</u> <u>TCCATGCCAGC</u>	
	GG23	<u>GAGCTCTCCCGGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	pEB17

Additional Results: Tracing the Full-length APE

ff. Table 4. Oligonucleotides used for plasmid construction and verification. Overhangs are underlined.

pEB17_ΔapeI	GG219	<u>ATATGTGATGGGTAAAAAGGATCGATCCTTCCTGC</u> GTGATCAGATCC	X. doucetiae DSM 17909 ^T
	GG220	CCACGCAGGGCGTCATCGTATTCACGCTAAATAGC <u>TCCTGTAATGCAGGC</u>	
	GG221	<u>CTCATGGCCTGCATTACAGGAGCTATTTAGCGTGA</u> ATACGATGACGC	
	GG222	<u>CAATTTGTGGAATCCCGGGAGAGCTCCTTGCGTT</u> CTTGAATATCTGAC	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17
pEB17_ΔapeJ	GG155	<u>ATATGTGATGGGTAAAAAGGATCGATCCTGGCCTT</u> TGCATTTTGCCTATG	X. doucetiae DSM 17909 ^T
	GG156	G CAGTAAAACGAGGATCAGTCAGCATTTTTAACTTT ATCTTTCCCTGGCTGG	
	GG157	<u>CTGCCAGCCAAGGAAAGATAAAGTTAAAAATGCTG</u> ACTGATCCTCG	
	GG158	<u>CAATTTGTGGAATCCCGGGAGAGCTCGCATCAAG</u> GTTAGCCTGACC	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17
pEB17_ΔapeK	GG215	<u>ATATGTGATGGGTAAAAAGGATCGATCCTGTTGTT</u> GGCTATTTCTGGG	X. doucetiae DSM 17909 ^T
	GG216	CAGAAACAGCAGTATCCCGCGCCATTTTTTTTTTA TTCCTTATTTGTCGG	
	GG217	<u>CATCTGCCCGGACAAATAAGGAATAAAAAAAAAAAT</u> GGCGCGGGATAC	
	GG218	<u>CAATTTGTGGAATCCCGGGAGAGCTCGGACTCAC</u> CATCCACACCAG	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17
pEB17_ΔapeL	GG191	<u>ATATGTGATGGGTAAAAAGGATCGATCCTTATTGC</u> CGGTAACGGATG	X. doucetiae DSM 17909 ^T
	GG192	<u>CCAGAATCGTGCCAGCAGACGTGGCGGCTTGACC</u> CCCATGC	
	GG193	<u>GATATTCTGTTTGAACGCATGGGGTCAAGCCGCC</u> ACGTCTGCTGGCAC	
	GG194	<u>CAATTTGTGGAATCCCGGGAGAGCTCAATACTGAT</u> GGATAGCACACACAGC	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17
pEB17_ΔapeM	GG195	<u>ATATGTGATGGGTAAAAAGGATCGATCCTCGTGAA</u> TCCGGCTATGTTTG	X. doucetiae DSM 17909 ^T
	GG196	<u>CAGGTCAGGGACAGGATGAATAACCGTAGGCTCAG</u> CGTCTCTGGTG	
	GG197	<u>GACACCAATTAACACCAGAGACGCTGAGCCTACGG</u> TTATTCATCCTGTCCC	
	GG198	<u>CAATTTGTGGAATCCCGGGAGAGCTCGCTCCCGA</u> CCATGACG	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17
pEB17_ΔapeN	GG199	<u>ATATGTGATGGGTAAAAAGGATCGATCCTACGCA</u> GAAAAACGTTGCAAC	X. doucetiae DSM 17909 ^T
	GG200	<u>CATACCTACGGCAGAAATATAAATCATGTTGATTGC</u> TTTTTCCTTTTTGGC	
	GG201	<u>GCCAAAAAGGAAAAAGCAATCAACATGATTTATAT</u> TTCTGCCG	
	GG202	<u>CAATTTGTGGAATCCCGGGAGAGCTCGCAGGTTG</u> ATATAACCTATCTCTG	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17
pEB17_ΔapeO	GG159	<u>ATATGTGATGGGTAAAAAGGATCGATCCTGTGGTC</u> ATAGCCTGAACTATTTTC	X. doucetiae DSM 17909 ^T
	GG160	<u>ATAGCGATCCACGGGTAATAGTCAGGCAGTTTTA</u> CTTTCCAAATGCTGAATG	
	GG161	<u>CGCCATTCAGCATTGGAAGAGTAAACTGCCTGA</u> CTATTTACCCGTGG	
	GG162	<u>CAATTTGTGGAATCCCGGGAGAGCTCTTGGTATG</u> GATGACACTATCCCAG	
	GG23	GAGCTCTCCCGGAATTCC	
	GG138	AGGATCGATCCTTTTAACCCATC	pEB17

ff. Table 4. Oligonucleotides used for plasmid construction and verification. Overhangs are underlined.

pEB17_ΔapeP	GG223	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTACCAGC</u> <u>AATCACATGGGTC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG224	<u>GGCACCGGTCACGAGA</u> <u>ACTGAACGCATTAAGCCAC</u> <u>TCCAATATCAGG</u>	
	GG225	<u>CCAGCCTGATATTGGGAGTGGCTTATGCGTTCAGT</u> <u>TCTCGTGAC</u>	
	GG226	<u>CAATTTGTGGAATTCCCGGGAGAGCTCCGATAGCC</u> <u>AATCCTGTCCC</u>	
	GG23	<u>GAGCTCTCCCGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	
pEB17_ΔapeQ	GG227	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTCCATTG</u> <u>CAGCGATCAATATG</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG228	<u>CCTTTCCTGATGGGATGTCGCATTATGGCTATGCTA</u> <u>CTCCATCCTGTTATTTATG</u>	
	GG229	<u>CTCATAAATAAACAGGATGGAGTAGCATAGCCATAA</u> <u>TGCGACATCCC</u>	
	GG230	<u>CAATTTGTGGAATTCCCGGGAGAGCTCCCGTGGGC</u> <u>ACTGATGTAAC</u>	
	GG23	<u>GAGCTCTCCCGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	
pEB17_ΔapeR	GG163	<u>ATATGTGATGGGTTAAAAAGGATCGATCCTTTTGGC</u> <u>AGTTTTGAAGGTGAAATC</u>	<i>X. doucetiae</i> DSM 17909 ^T
	GG164	<u>GCTTTGTGACGGTCTTAACGCACTAAAAAGCATCC</u> <u>CTCCGTTGATTG</u>	
	GG165	<u>GGTCATCTCAATCAACGGAGGGATGCTTTTTAGTG</u> <u>CGTTAAGACCGCTG</u>	
	GG166	<u>CAATTTGTGGAATTCCCGGGAGAGCTCTAAAAGGG</u> <u>ATCATCCCCTGAGT</u>	
	GG23	<u>GAGCTCTCCCGGAATTCC</u>	
	GG138	<u>AGGATCGATCCTTTTTAACCCATC</u>	
Oligonucleotides used for verification of APE-producer mutants			
Ape-producer mutant	Oligonucleotide 5' to 3'		
ΔapeB	GG171	CTGCTGGCTAATCAATAAACATCCATGTTAAAGC	
	GG172	AAATCAGCAATAAAATTAACCAAATCAGGGAGAGTA	
ΔapeC	GG245	CATTACTGATCTGCAAGAAGCAGTAACGAACGG	
	GG246	GACGGTGTTAGTAATTGAGTAAGCTGGCGATTGAGTTGG	
ΔapeD	GG183	GAAACAGAACCCAGCCTTCCTCAAAG	
	GG184	CTTTTCAATGACAGGTGGCCATAAC	
ΔapeE	GG173	CAATCGCCAGCTTACTCAATTAACACCG	
	GG174	ACTTTACGTGATAGATAACGCCTTCTGGGGG	
ΔapeF	GG175	GTCAACACTTCACTTCTGTGCTACACTGG	
	GG176	CTTACCCGCATGTAATGTTGCCAGCAG	
ΔapeG	GG203	AGCCATAAACCATCAGCAAAACAACAGC	
	GG204	GCGATCAAATGGTAACCCCAAGTAC	
ΔapeH	GG231	GATTCAATCGATGCTCTGGAGCTGGG	
	GG232	TCATGTTGAGCAGGTTGCTCGCGG	
ΔapeI	GG235	CTGGCTGCTCTGTGGGGCGATAAGC	
	GG236	GATCACCACGCAGGGCGTCATC	
ΔapeJ	GG177	ACAACGGCAGCGAAGGCCAGACTG	
	GG178	CTCAGCGTCTCTGGTGTAAATTGGTGTGG	
ΔapeK	GG233	ACGCAAGCCCTTTGAGTGTCTGCTTTACC	
	GG234	CAAACCTGGTTGTCAGAAACAGCAGTATCCC	
ΔapeL	GG205	TCTGTTGAACGCATGGGGGTCAAGC	
	GG206	CCAACGTCACTACCGTTCCTGCCAGC	
ΔapeM	GG207	CCGAGCTTGCTTTACCCGACACC	
	GG208	GCCAGATGTAAGTGGTCCCATGATCAGC	
ΔapeN	GG209	CTATCCCCACTGGCCATGCCTGAGC	
	GG210	CTGATTGATGACGATCGACTCCATCTGG	
ΔapeO	GG179	CACAGAGATTCATTATGATGCCCCAAATGG	
	GG180	TGCCCCGATAATACCGGCTTTCCG	
ΔapeP	GG237	CCTTGATGAGGGTATCAGCATGTCAGC	
	GG238	ATCCCTTTGCTGGCACCGGTCAC	
ΔapeQ	GG239	ACGCATGGCATATGTCAGCCCCG	
	GG240	GGCCTTTCCTGATGGGATGTCGC	
ΔapeR	GG181	CAAATGGAAGAAGCCGCTTGAAAGAAG	
	GG182	TCGGTGATTTTTCTCAACGACATACTTTGG	

Table 5. Plasmids used in this study.

Plasmid	Genotype	Reference
pAL03	R6K γ ori, oriT, <i>sacB</i> , <i>araC</i> , <i>araBAD</i> , Km ^r	[150]
pAL03_ape_mP	markerless promoter exchange plasmid based on pAL03 with 800 bp of up- and downstream region of <i>apeB</i> start codon, Km ^r	see 7.1
pEB17_Km	R6K γ ori, oriT, <i>araC</i> , <i>araBAD</i> promoter, Km ^r	[150]
pEB17_Δ <i>apeB</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 800 bp of up- and downstream region of <i>apeB</i>	this work
pEB17_Δ <i>apeC</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 761 bp of up- and 730 bp of downstream region of <i>apeC</i>	this work
pEB17_Δ <i>apeD</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 850 bp of up- and downstream region of <i>apeD</i>	this work
pEB17_Δ <i>apeE</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 650 bp of up- and downstream region of <i>apeE</i>	this work
pEB17_Δ <i>apeF</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 650 bp of up- and downstream region of <i>apeF</i>	this work
pEB17_Δ <i>apeG</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 600 bp of up- and downstream region of <i>apeG</i>	this work
pEB17_Δ <i>apeH</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 956 bp of up- and 896 bp of downstream region of <i>apeH</i>	this work
pEB17_Δ <i>apeI</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 826 bp of up- and 782 bp of downstream region of <i>apeI</i>	this work
pEB17_Δ <i>apeJ</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 950 bp of up- and 983 bp of downstream region of <i>apeJ</i>	this work
pEB17_Δ <i>apeK</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 850 bp of up- and 818 bp of downstream region of <i>apeK</i>	this work
pEB17_Δ <i>apeL</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 857 bp of up- and 850 bp of downstream region of <i>apeL</i>	this work
pEB17_Δ <i>apeM</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 855 bp of up- and 843 bp of downstream region of <i>apeM</i>	this work
pEB17_Δ <i>apeN</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 850 bp of up- and 853 bp of downstream region of <i>apeN</i>	this work
pEB17_Δ <i>apeO</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 850 bp of up- and 825 bp of downstream region of <i>apeO</i>	this work
pEB17_Δ <i>apeP</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 823 bp of up- and 820 bp of downstream region of <i>apeP</i>	this work
pEB17_Δ <i>apeQ</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 894 bp of up- and 934 bp of downstream region of <i>apeQ</i>	this work
pEB17_Δ <i>apeR</i>	Deletion plasmid based on pEB17_Km with Km ^r , with 830 bp of up- and downstream region of <i>apeR</i>	this work

8.4.1.1. Cloning

Genomic DNA from *X. doucetiae* was isolated using the Genra Puregene Yeast/Bact kit (Qiagen). Polymerase chain reaction was (PCR) performed with Phire Hot Start II DNA polymerase (Thermo Scientific), Phusion High-Fidelity DNA Polymerase or Q5 polymerase (New England Biolabs) according to the manufacturer's instructions. Oligonucleotides were purchased from Eurofins Genomics. The Invisorb Spin DNA Extraction Kit (Strattec) was used for DNA purification from agarose gels. Plasmids were isolated with the Invisorb Spin Plasmid Mini Two Kit (Strattec). Plasmid-backbone PCRs were restriction digested with *DpnI* (New England Biolabs) following the manufacturer's protocol. All plasmids were cloned via Hot Fusion^[163] with corresponding oligonucleotides listed in Table 4. *E. coli* ST18 λ pir was used as a cloning strain for plasmids with pEB17-backbone, *E. coli* BL21 (DE3) GOLD was used instead for pACYC-DUET and pCDF-Duet backbones. The respective *E. coli* strains were electroporated with desalted Hot Fusion (MF-Millipore membrane, VSWP, 0.025 μ m). Appropriate antibiotics were used in the following final concentrations: kanamycin (50 μ g/mL), spectinomycin (50 μ g/mL), chloramphenicol (35 μ g/mL).

8.4.1.2. Construction of Deletion Mutants

Deletions *apeB-apeR* in the APE-producer (*X. doucetiae* Δ DC Δ hfqP_{BAD}*apeB*) were constructed by conjugation of *E. coli* ST18 strains, harboring the corresponding deletion plasmid with 500-900 bp of the up- and downstream flanking regions of the gene of interest (Table 5). In a first homologous recombination event, the pEB17-plasmid backbone is inserted and maintained by kan^R selection. Therefore, both strains were grown in 10 mL LB-medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl at pH 7.5) (*E. coli* ST18- λ pir was supplemented with 50 μ g/mL kanamycin and δ -aminolevulinic acid) to an OD₆₀₀ of 0.6-0.8 and harvested in a ratio of 5:1 (5 mL *X. doucetiae* strain:1 mL *E. coli* strain). To get rid of the remaining δ -aminolevulinic acid, the *E. coli* ST18 λ pir pellet was washed three times with 1 mL LB. Cells were resuspended in 60-90 μ L LB and mixed prior to pipetting them in one drop (90 μ L) on an LB agar plate. After 1 d at 30°C, cells were resuspended in 1 mL LB. The first homologous recombination event was forced by striking 200 μ L of the appropriate cell suspension on LB agar plates, supplemented with kanamycin. In a second homologous recombination, the plasmid backbone was lost by a double cross-over event, which was achieved by further cultivation of the insertion-mutants (first homologous recombination) on LB agar plates containing 12% (w/v) sucrose.

The deletions (*apeB-R*) in the APE-producer were verified through colony-PCR using oligonucleotides listed in Table 4. Therefore, half of a colony of the corresponding mutant was picked and resuspended in 50 μ L 1x Taq-buffer and heated in the microwave (600 W) for 3 min. 1 μ L of the resulting cell suspension was used for a 25 μ L PCR-reaction (10 μ M oligonucleotide fw/rv, 10 mM dNTPs, 1x taq-buffer, 0.25 μ L taq; 10x taq-buffer) using Taq DNA polymerase (New England Biolabs) according to manufacturer's protocol.

8.4.1.1. Cultivation and Extraction of APE-Producing Strains

Transesterified extracts of the APE-producer, the APE-producer mutants and the *E. coli* strains used for heterologous expression, were generated as previously described (7.1). *E. coli* strains were cultivated in XPPM-medium. For the extraction procedure of APE-lipids, the transesterification step was omitted.

8.4.1.2. HPLC-UV/MS Analysis of APE compound

Crude extract from the APE-producer and the APE-producer mutants Δ *apeA-R* was analyzed via high resolution (HR)-HPLC-ESI-UV/MS analysis using a Dionex Ultimate 3000 LC system (Thermo Fisher), coupled to an Impact II electrospray ionization mass spectrometer (Bruker) and a DAD-3000 RS UV-detector (Thermo Fisher). Internal mass calibration was achieved by injecting a 10 mM sodium formate solution (0-1.5 min, calibration segment). For Data analysis of HPLC-UV/MS-chromatograms, Compass DataAnalysis 4.3 (Bruker) was used.

Columns, solvent systems and LC- and MS-parameter, which were used upon analysis of the different compounds, are listed in Table 6 and Table 7.

Table 6. Used columns with their specifications plus conditions/parameters for analytical HPLC-MS analyses.

columns			
C18	RP C18 BEH Acquity UPLC (Waters), 50 mm x 2.1 mm x 1.7 µm		
C4	AERIS-WIDEPOR C4 (Phenomenex), 250 mm x 4.6mm x 3.6 µm		
C3	Zorbax 300SB-C3 300Å (Agilent), 150 mm x 3.0 mm x 3.5 µm		
LC-conditions			
Gradient I	flow rate 0.4 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	0.0	95	5
	2.0	95	5
	2.1	60	40
	4.0	60	40
	20.5	5	95
	29.0	5	95
	29.1	95	5
30.0	95	5	
Gradient II	flow rate 0.6 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	0.0	95	5
	2.0	95	5
	17.0	5	95
	17.1	0	100
	25.0	0	100
	25.1	95	5
	30.0	95	5
Gradient III	flow rate 0.6 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	0.0	60	40
	2.0	60	40
	17.0	5	95
	17.1	0	100
	25.0	0	100
	25.1	60	40
	30.0	60	40
Gradient IV	flow rate 1.2 mL/min, column oven temperature 55°C		
	min	solvent A (%)	solvent B (%)
	0.0	80	20
	1.0	80	20
	10.0	47	53
	10.1	0	100
	14.0	0	100
	14.1	80	20
	16.0	80	20
Gradient V	flow rate 1.2 mL/min, column oven temperature 55°C		
	min	solvent A (%)	solvent B (%)
	0.0	32	68
	1.0	32	68
	3.0	53	47
	5.0	54	46
	28.0	60	40
	28.1	97	3
	30.0	97	3
MS-parameters			
	Source		Tune
'normal'	End Plate Offset	500 V	<u>Transfer</u>
	Capillary	4500 V	Funnel1 RF 300 Vpp
	Nebulizer	3.0 Bar	Funnel2 RF 300 Vpp
	Dry Gas	8 L/min	Hexapole RF 60 Vpp
	Dry Temp	200°C	<u>Quadrupole</u>
			Ion Energy 5 eV
			<u>Collision Cell</u>
			Collision Energy 10.0 eV
			Pre Pulse Storage 5 µs
			Stepping Collision RF 500 Vpp
		Stepping Transfer Time 82.5 µs	
'APE-lipid'	see 'normal'		see 'normal' with: Stepping Collision RF 800 Vpp Stepping Transfer Time 140 µs

Table 7. Used methods for the analysis of APE_{Xd}/conjugated-FA methyl ester and APE-lipids. The best achieved conditions for separation are marked in bold. For details, see Table 6.

compound	column	solvent system	LC-condition	MS-parameter
APE-/conjugated-FA methyl ester	C18, C3 , C4	A:H ₂ O/B:ACN +0.1 % formic acid	Gradient I, II , III	'normal'
APE-lipids	C3 , C4	A:ACN/2-Prop(9:1)/B:ACN/H ₂ O(6:4) +0.2 % formic acid +10 mM ammonium formate	Gradient IV, V	'APE-lipid'

8.4.1.3. Fermentation and Crude Extraction

Fermentation was performed in collaboration with Anja Schüffler from the 'Institut für Biotechnologie und Wirkstoff-Forschung' at the University of Kaiserslautern. Two 50 L bioreactors were filled with 20 L LB-medium and inoculated with a preculture (1:100) of the APE-producer Δ *apeJ* mutant supplemented with 0.2% L-arabinose (200 mL 20% L-arabinose). The strain was cultivated for 72 h with 160 rpm, at 30°C, 4 L O₂/h and supplemented with anti-foam. Cells were harvested with a flow-through centrifuge (Eppendorf) at 60 000 rpm. The cell pellet was freeze-dried and grouted prior to extracting it with DCM/MeOH (2:1), resulting in 80 g crude extract.

8.4.1.4. Enrichment and Isolation of APE Compounds

The preparation of APE_{Xd}-methyl ester or APE-lipids was performed as described in 8.4.1.1. The purification/enrichment conditions for the conjugated-FA '399' and conjugated-FA+X are listed in Table 8. and Table 9.

Table 8. Used columns with their specifications plus conditions/parameters for the isolation process.

columns			
Silica	SNAP KP-Sil 50 g		
C18	Eclipse XDB-C18 (Agilent), 9.4 mm x 250 mm x 5 µm		
C3	Zorbax 300SB-C3 (Agilent), 9.4 mm x 250 mm x 5 µm		
Phenyl-Hexyl	Eclipse Plus Phenyl-Hexyl (Agilent), 9.4 mm x 250 mm x 5 µm		
C8 prep.	Eclipse XDB-C8 (Agilent), 21.2 mm x 250 mm x 7 µm		
sephadex	LH-20 material		
LC-conditions			
Size exclusion	self-packed 1.20 m glass-column with LH-20 material, solvent MeOH, flow rate 1 drop/s by gravity, RT		
normal phase 'Silica'	Chromatographic system: Biotage Flash-SP1		
	flow rate 30 mL/min, RT		
	solvent A: DCM, solvent B: MeOH, gradient: 20 column volumes 0-100% MeOH		
reversed phase 'Gradient I, semiprep'	Chromatographic system: Agilent 1260 Infinity		
	flow rate 4 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	7.0	20	80
	7.1	0	100
	13.0	0	100
	13.1	20	80
18.0	20	80	
Reversed phase 'Gradient II semiprep'	Chromatographic system: Agilent 1260 Infinity		
	flow rate 4 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	7.0	95	5
	7.1	0	100
	13.0	0	100
	13.1	95	5
18.0	95	5	

ff. Table 8. Used columns with their specifications plus conditions/parameters for the isolation process.

Reversed phase 'Gradient III semiprep'	Chromatographic system: Agilent 1260 Infinity		
	flow rate 5.5 mL/min, column oven temperature 40°C		
	min	min	min
	5.0	5.0	5.0
	5.1	5.1	5.1
	11.0	11.0	11.0
	11.1	11.1	11.1
	17.0	17.0	17.0
	17.1	17.1	17.1
20.0	20.0	20.0	
reversed phase 'Gradient IV semiprep'	Chromatographic system: Agilent 1260 Infinity		
	flow rate 4 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	5.0	35	65
	5.1	35	65
	15.0	20	80
	15.1	0	100
	20.0	0	100
	20.1	35	65
24.0	35	65	
reversed phase 'Gradient V semiprep'	Chromatographic system: Agilent 1260 Infinity		
	flow rate 4 mL/min, column oven temperature 40°C		
	min	solvent A (%)	solvent B (%)
	8	40	60
	15	10	90
	21	10	90
	21.1	40	60
25	40	60	
reversed phase 'Gradient I prep'	Chromatographic system: Agilent 1260 Infinity II Preparative LC/MSD System		
	flow rate 20 mL/min, RT		
	min	solvent A (%)	solvent B (%)
	8.0	50	50
	12.0	50	50
	15.0	5	95
	21.0	5	95
21.1	50	50	

Table 9. Used conditions for the isolation of conjugated-FA methyl ester and conjugated-FA+x.

compound	strain	culture volume	cultivation conditions	
conjugated-FA methyl ester	APE-producer	6 L LB	2 d, 30 °C, 180 rpm	
conjugated-FA+x	APE-producer Δ <i>apeJ</i>	40 L LB	see 8.4.1.3	
	chromatography	column	solvent system	LC-parameter
conjugated-FA methyl ester	size exclusion	sephadex	MeOH	see Table 8
	reversed phase semiprep.	C18	A:H ₂ O/B:ACN	'Gradient I'
	reversed phase semiprep.	C3	A:H ₂ O/B:ACN	'Gradient II'
	reversed phase semiprep.	Phenyl-Hexyl	A:H ₂ O/B:ACN	'Gradient III'
conjugated-FA+x	normal phase	Silica	A: DCM/B: MeOH	see Table 8
	size exclusion	sephadex	MeOH	see Table 8
	reversed phase prep.	C8 prep.	A: 2-Prop/ACN (9:1)/ B:ACN/H ₂ O(6:4)	'Gradient IV'
	reversed phase semiprep.	C3	A: 2-Prop/ACN (9:1)/ B:ACN/H ₂ O(6:4)	'Gradient V'
	reversed phase semiprep.	C3	A: 2-Prop/ACN (9:1)/ B:ACN/H ₂ O(6:4)	'Gradient I prep.'

8.4.1.5. Quenching Effect of the APE-Producer

Fluorescence measurement was performed with extracts from the APE-producer and the APE-producer Δ *apeJ*-mutant on a Fluorolog system with the assistance of Marco Kaiser (Wöhnert group, Goethe-Universität Frankfurt). Therefore, 1 mL crude extract of each strain was prepared (50 mL pellet, extracted with 30 mL MeH/DCM (2:1), as described above). Crude extract of the APE-producer was titrated (100 μ L, 200 μ L, 250 μ L) to crude extract of the APE-producer Δ *apeJ* mutant in a total volume of 2 mL and mixed by pipetting.

8.4.1.6. Thin-Layer Chromatography of APE-Extracts

Thin-layer chromatography was performed according to the protocol from Entezami *et al.* for phospholipids.^[164]

8.4.1.7. Network analysis

Sephadex fractions of transesterified APE-producer (mutant) extracts were analyzed via HPLC-UV/MS, as described in 8.4.1.2. Raw-data, obtained from DataAnalysis 4.3, were subjected to the GNPS-platform^[160] in mzXML-file format. Samples were manually grouped according to their UV/Vis-spectrum into: 1. control-group (no significant UV/Vis-signal of APE- or conjugated-FA detectable) (color in network: grey), 2. APE_{xd}-methyl esters (color network: yellow) and 3. conjugated-FA methyl esters (or hydrophobic compounds) (color network: turquoise).

If not stated otherwise, the GNPS parameters were used as default: Basic options, 'Precursor/Fragment Ion Mass Tolerance': 0.05 Da. Advanced Network Options, 'Min Pairs Cos': 0.7, 'Minimum Matched Fragment Ions': 4, 'Network TopK': 7, 'Minimum Cluster Size': 2, 'Maximum Connected Component Size': 100. Advanced Library search options, 'Maximum Analog Search Mass Difference': 500. Advanced Filtering options, 'Minimum Peak Intensity': 50.

9 Discussion and Future Work

Type II PKS systems of Gram-positive origin are well characterized and typically found in *Streptomyces*.^[23,35] Within this thesis, three PKS II systems from the Gram-negative bacteria *Xenorhabdus* and *Photorhabdus* were characterized (sections 0 and 8). Results were obtained from a combination of *in vivo* and *in vitro* experiments as well as by spectroscopic characterization. The underlying biosynthetic mechanisms are discussed below. A special emphasis was put on the APE system in *X. doucetiae*.

9.1 APEs – More than Small Molecules

Largely overlooked in HPLC-MS-based metabolomics, with the help of bioinformatics, the APE BGC was found to be one of the most widely distributed gene clusters among proteobacteria.^[138] A comprehensive global alignment divided the APE BGCs into three subfamilies, the BGC from *X. doucetiae* represents a member of subfamily 1 (Figure 23b). Associated with these BGCs are molecules known as APEs, which were introduced in this work as APE methyl esters. Small molecules with masses ranging from 300-500 Da (Figure 23a). APE compounds are thought to be located in the membrane.^[138,165] In previous studies,^[138,142,166] APEs were so far isolated from pelleted cell cultures, resulting in the structure elucidation of APE-methyl ester APE_{Ec} from *E. coli* CFT073 and Ape_{Vf} from *Vibrio fischeri*, xanthomonadins and flexirubins (Figure 23). Antismash^[167] analysis revealed that *X. doucetiae* harbors an APE BGC, which is highly homologous to the APE BGC from *E. coli* CFT073 (Figure 23).

9.1.1 APE_{Xd}-Methyl Esters from *X. doucetiae*

In order to uncover the APE-methyl esters produced by *X. doucetiae* (APE_{Xd}-methyl ester), APE derivatives were extracted as previously described^[138] and the most abundant APE_{Xd}-methyl ester from *X. doucetiae* was subsequently isolated and the structure elucidated via NMR (7.1).

Furthermore, the extracted APE_{Xd}-methyl esters were analyzed in order to determine their structural properties. The structural features of these multiple APE_{Xd}-methyl esters were investigated by adding precursor molecules to the growing cultures. It was shown that the extracted APE_{Xd}-methyl esters' chain lengths ranged from 5-7xEN and the methylation pattern from 0-2x Met (compare section 7.1 and 8). Importantly, the isolated main APE_{Xd}-methyl ester comprised 6xEN and 2xMet. These findings comply with the results obtained from *in vitro* assays with ApeO:C (7.1). Using ApeCDEFHIOPQ along with 4-HBA and malonate building blocks, resulted in the ACP-bound APE, harboring 7xEN and 0xMet (due to missing MT ApeB). These results indicate that the chain-length determination of ApeO:C is influenced by another factor since the 7xEN-bound APE-product did not account for the main product *in vivo*.

APE compound were so far only presented with a characteristic hydroxy-substitution and methylations at the aryl-ring, sometimes substituted with additional halogens (Br, Cl).^[138] Surprisingly, according to MS-analysis, besides the APE_{Xd}-methyl esters, which owe a 4-hydroxy-substitution, an amino-substitution was assigned to the aryl moiety of the APE_{Xd}-methyl esters (see section 8). Like 4-HBA, which was shown to be incorporated into APE_{Xd}-methyl ester as a precursor (7.1), this aminobenzoic acid precursor can also be obtained from the shikimate pathway,^[47] as it was already described for the PKS II-derived aurachins (Figure 11).^[123,124] However, this observation needs further investigations, i.e. feeding experiments with isotope-labeled amino-benzoate precursors, to reveal, among others, the substitution position of the amino group on the aromatic ring (*para/meta/ortho*). As it stands, this amino-substitution expands the APE moiety repertoire and may represent an alternative/additional structural feature, potentially increasing evolutionary fitness for the organism.

A plethora of unassigned APE_{Xd}-methyl esters were further detected in the HPLC-UV/MS-analysis of the sephadex fractions of the esterified extract from the APE-induced mutant (8.1). It cannot be excluded that the unassigned signals are purification artifacts, as the conjugated π -system of APE-methyl ester could be prone to isomerization and oxidation.^[147,168] In comparison, the results from Cimermancic *et al.*^[138] indicate that the respective APE BGCs from *E. coli* CFT073, *X. doucetiae* and *V. fischeri* were enough to yield the respective major products (APE_{Ec}, APE_{Xd}, APE_{Vf} in Figure 36) in a heterologous production host. Here as well, apart from the three main products, the chromatograms indicate that each strain features its own set of lesser abundant APE derivatives,^[138] which is in accordance with our observations. The above-stated indications suggest that the APE derivative formation underlies a complex mechanism which, in addition to the already known CLF, also includes auxiliary factors and will be discussed later.

9.1.2 In Small Steps from APE-Methyl Esters to APE-Lipids

Notably, besides the genes that were shown to be involved in the biosynthesis of the APE_{Xd}-moiety, the APE BGC harbors several unallocated genes (Figure 36). These additional genes indicate a further installation of the APE moiety to a lipid-like structure. According to blastp, these 'lipid-genes' encode namely for a phosphoglyceryl acyltransferase, a LolA-like protein and an MMPL-transporter, typically involved in lipid biosynthesis. In brief, LolA-like proteins mediate solubility to the hydrophobic lipids, which are assembled in the cytoplasm, while MMPL-transporter export the produced lipid compounds to their final location across the cell envelope.^[169,170] Additionally, a gene encoding for a glycosyltransferase is well-conserved.

The investigations that the APE moiety gets installed to a lipid-like structure (hereafter referred to as APE-lipid) were described within the additional results section 8. Here, the major advance pointing towards an APE-lipid was taken through the discovery of several unusual conjugated-FAs, only produced in the *X. doucetiae* APE-producer strain. These conjugated-FAs were detected in the HPLC-UV/MS-analysis next to the APE_{Xd}-methyl esters through their characteristic absorbance in the VIS-range. Here again, the major conjugated-FA was isolated, and the structure elucidated by NMR. As it was true for the APE_{Xd}-compounds, the conjugated-FA was depicted as a methyl ester, too. This led to the conclusion that a linkage between the APE and the conjugated-FA moiety to form a more complex APE-lipid exists, if the extract is not subjected to transesterification.

In order to verify this assumption, analytics to confirm APE-lipid-like structures were adapted and the extraction procedure was conducted without the transesterification step. As a result, for the first time, APE-molecules were detected as APE-lipids in MS-analyses (Figure 35).

Unfortunately, the isolation of the amphiphilic APE-lipid in total failed so far (section 8.1). Major drawbacks represented lipid-impurities of phosphatidylethanolamines due to their similar chromatographic behavior (Figure S7). Furthermore, low baseline separation of the APE-lipid derivatives and the postulated instability under light and oxygen influence, which was already reported for xanthomonadin, flexirubin or carotenoid compounds,^[168,171] have made an isolation difficult. As already mentioned, the APE-lipids are assumed to be located in the outer membranes, as it was shown for the pigment xanthomonadin.^[165] This transfer is most likely supportively co-mediated by the postulated *Mycobacterial membrane protein Large* (MmpL)-transporter encoded within the cluster (ApeM). MmpL-transporters are a subclass of *Resistance-Nodulation-Cell Division* (RND) transporters, which appear to play a major role in acting as scaffolding proteins that couple lipid synthesis and transport.^[169,172] It is suggested that the assembly of APE-lipids takes place in the cytosolic compartment or near the inner

leaflet of the plasma membrane, with the APE-lipid synthase being in the cytosolic compartment. This model would fit with an MmpL-transporter-mediated synthesis of the unusual lipids from *Mycobacteria*, derived from PKS (or NRPS) systems.^[169]

To reduce lipid-impurities during the isolation process of APE-lipids, the outer membrane can be separated from the inner membrane lipids prior to extraction.^[173] Consequently, a huge amount of the major impurities represented by the phosphatidylethanolamines were to be removed due to the asymmetric composition of the inner and outer membrane. Here, the inner membrane is composed of phospholipids and the outer membrane predominantly of a mixture of phospholipids and LPS.^[174] Furthermore, the *apeB* deletion mutant (encoding for an MT) should be used for the isolation process, as this mutant seems to produce the unmethylated product (section 8). Eventually, heat and light during the purification process ought to be reduced to the minimum, as these might be factors triggering the decay process, creating unwanted side products.^[168,171]

However, the core structural elements of the APE-lipid were successfully assigned with MS-based methods (Figure 35). The APE core was detected in MS² with an additional glucosamine substitution and due to the detection of a phosphoethanolamine head group, which displays a typical neutral loss in MS of phosphatidylethanolamines^[175], the unusual conjugated-FA moiety is proposed to be linked to a glycerol.

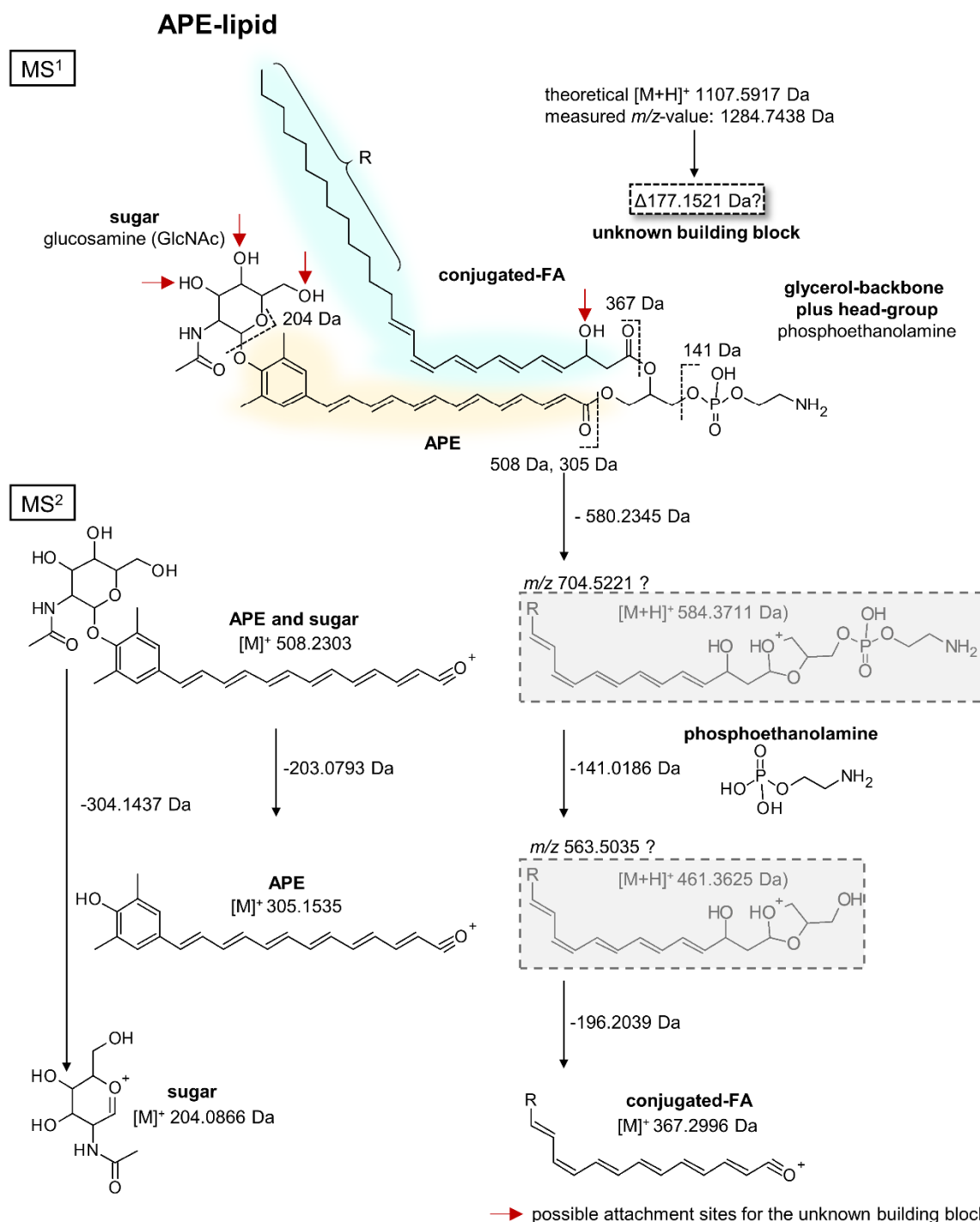


Figure 35. Proposed structure of the main APE-lipid with a detected m/z of 1285 Da. Proposed collision induced fragmentations are depicted. The linkage of the APE-sugar moiety was confirmed in MS² of m/z 1284.744 (see Figure 33) with 508.2303 $[M+H]^+$. The oxonium-ion of APE (305.1535 $[M+H]^+$), the conjugated-FA (367.2996 $[M+H]^+$), the sugar moiety (204.0866 $[M+H]^+$) and the neutral loss of the polar phosphoethanolamine head group (-141.0186 Da) were detected in the same spectra. The detected m/z 701.5221 and m/z 563.5035 were not assigned but may harbor the unknown building block.

It must be considered that the summarized masses for the APE-lipid building blocks (APE-, conjugated-FA, sugar and phosphoethanolamine head group moiety), still do not add up to the obtained overall mass of the main APE-lipid of m/z 1285 (Figure 35). The delta for the whole

compound lies at 177 Da. This may vary depending on the formed adduct (e.g. $[M+H]^+$, $[M+NH_4]^+$, $[M+Na]^+$), but stands for a yet unidentified building block. A possible attachment site for this unknown building block might be formed by the remaining hydroxy-groups of the sugar or the β -hydroxy group or the conjugated-FA moiety (red arrows, Figure 35). Thus, to this end, the proposed APE-lipid needs further structural evidence and the remaining building block has to be identified.

The combination of NMR-techniques and feeding experiments with labeled precursor molecules are highly recommended for future investigations. Such feeding experiments may comprise simple ^{13}C acetate feeding (1,2- $^{13}C_2$, 1- ^{13}C , 2- ^{13}C) or cultivation in C^{13} media to determine/proof the APE-lipid structure in total, with the feeding of selected precursor molecules (^{13}C -, ^{15}N -, or 2H -labeled or modified precursor molecules), to further distinguish the incorporation and connection of the different building blocks.^[176,177] There are many ways to conduct such feeding experiments and in the following some possibilities are exemplified.

For instance, feeding with 3-fluoro-4HBA instead of 4HBA, as it was already performed under 7.1 for the assignment of the APE_{Xd}-methyl esters, would affect all MS²-fragments harboring the APE moiety. The incorporation of a fatty acid precursor to the APE-lipid structure can be tested with feeding of ω -azido labeled C₁₃-C₁₆ fatty acids, as performed under 7.2. Here, in turn, all MS²-fragments harboring the conjugated-FA would be affected upon incorporation. With the use of both precursor feeding methods, it may be distinguishable whether the unknown building block is attached to the APE- or the conjugated-FA moiety. ^{13}C -, ^{15}N -, or 2H -labeled precursors can be used to further study the sugar and the glycerol backbone.

As phosphoethanolamine is built up from serine, the phosphoethanolamine head group can be verified with feeding of labeled ^{13}C -serine. Mass shifts of the glucosamine residue can be caused by feeding with ^{13}C -glucose. Hence, combining the results from the feeding experiments with additional MSⁿ-analyses and NMR-data would result in a fully revised APE-lipid structure.

Apart from that, the herein conducted work, represents a major step forward regarding the detection of APE-lipids and significantly changes the point of view for the hitherto believed small APE compound.^[138]

Notably, the genes between APE-subfamily 1 (*X. doucetiae* APE BGC) and 2 (*V. fischeri* APE BGC) are well conserved (Figure 36) but family 2 inherits an additional ammonia lyase to produce the aromatic starter unit and an additional halogenase encoding gene, while the halogenated product was not yet identified. In contrast, the *X. doucetiae* gene cluster of APE-subfamily 1 encodes for a protein of unknown function directly at the beginning of the gene

cluster, while the *V. fischeri* BGC encodes for an additional lipoprotein instead. It should be mentioned that all analysis with promotor exchange mutants (*X. doucetiae* APE-producer) were conducted with a promotor exchange in front of *apeB*, resulting in an APE_{Xd}-methyl ester which was double-methylated and represents the same main APE-methyl ester as for subfamily 2 (see Figure 36, APE_{Vf}). As the function of ApeA is unknown any influence on the methylation of the aryl-ring cannot be excluded.

The already described lipid genes (Figure 36, turquoise genes) are well-conserved among APE-subfamily 1 and 2 but differ among subfamily 3. It can be speculated that those genes determine the structural elements which are attached to an APE-scaffold. Interestingly, within the xanthomonadin gene cluster, two C-terminal homologs harboring the LPLAT-acyltransferase-like-domain of ApeJ were found. Thus, the domain for the glycosyltransfer is missing, typically being encoded N-terminally in the ApeJ homolog. Additionally, no homologs of the unknown proteins ApeA, ApeG and ApeN exist (Figure 36).

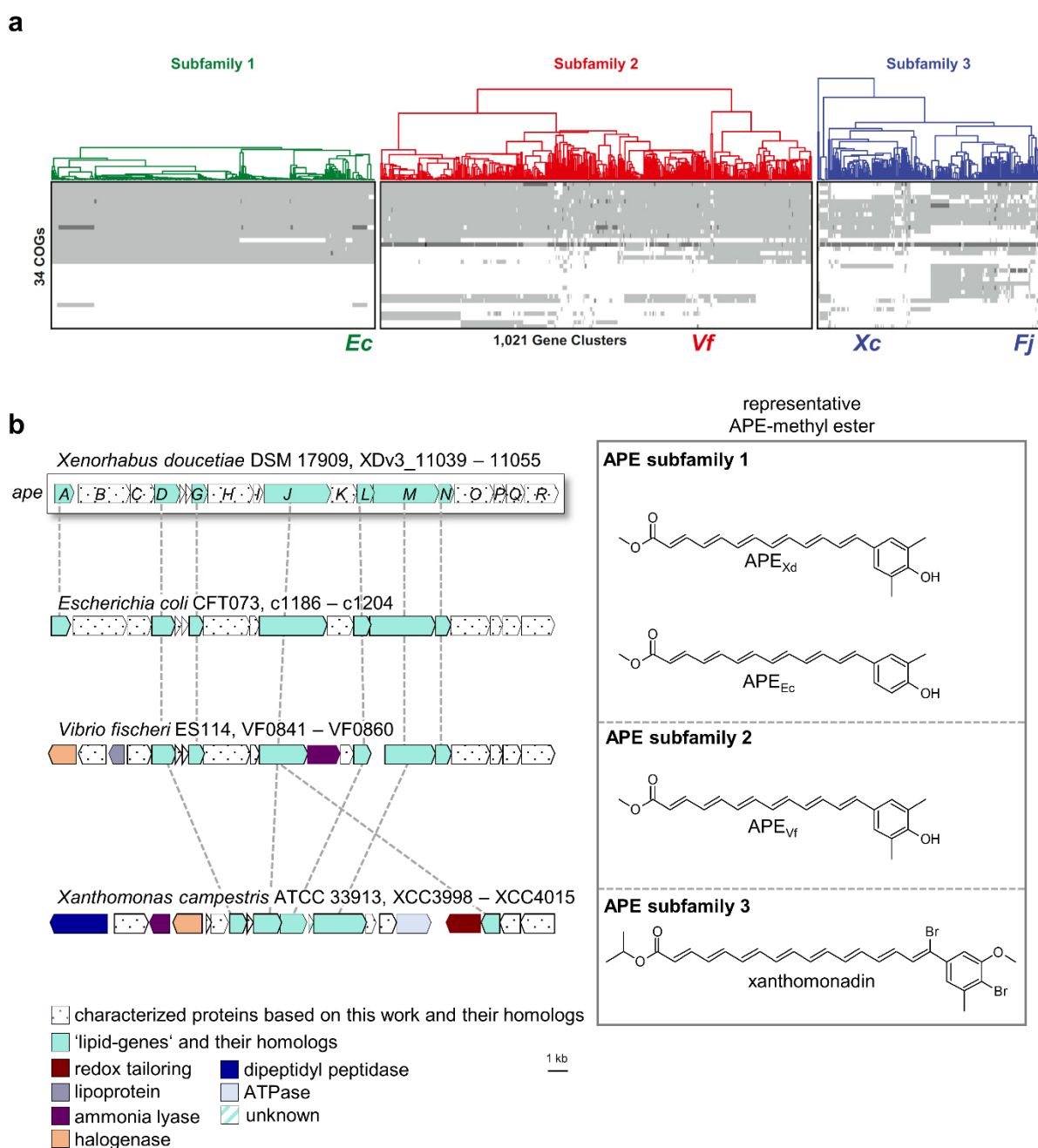


Figure 36. Relation of the APE-subfamilies. Assignment of APE BGCs to different APE-subfamilies according to their gene endowment, adapted from [145]. *Ec* = *E. coli*, *Vf* = *V. fischeri*, *Xc* = *Xanthomonas campestris*, *Fj* = *Flavobacterium johnsonii* (a) Comparison of APE-lipid genes (turquoise). Dotted lines represent an aa-sequence similarity of over 40%. In the case of the *apeJ* homologs in *Xanthomonas campestris* ATCC 33913, only the N-terminal translated regions (first 260 aa) were used for the alignment (b).

Finally, the herewith developed methods could be used to confirm the postulated APE-lipids or their esterified structural elements (APE moiety, conjugated-FA moiety) by using different representatives of the APE-subfamilies. An intriguing discrepancy can be found between the DAR connection of flexirubin-like pigments (see arcuflavin, Figure 23) and the proposed structure (Figure 35). In fact, the DAR moiety is directly linked to the carboxyl-group of the APE moiety and there is no space for adding a glycerol backbone. It cannot be excluded that the

connection between the APE-, conjugated-FA-, glycerol- and sugar moiety is different from the proposed structure of the APE-lipid in Figure 35. However, as stated above, a structural confirmation of the APE-lipid in total would clarify these uncertainties.

9.2 Multitasking APE-Synthase

The finding of an APE-lipid raises questions regarding both, the biosynthesis of the final compound itself (Figure 35) and the unusual conjugated-FA-part (Figure 28). Due to the fact that the hydrophobic conjugated-FA is not synthesized in the uninduced APE-producer, the APE-machinery is responsible for parts of its biosynthesis. At this point, it is beneficial to take the results, presented in publication under 7.1 into consideration, which addressed the *in vitro*-reconstitution of an ACP-bound APE moiety. The biosynthesis was reconstituted with the AasS ApeH and both ACPs (ApeE and ApeF) under further addition of ApeO:C (KS:CLF), ApeR (KS), ApeQ (KR) and the heterocomplex ApeI/P (DH) with addition of hydroxybenzoate and malonate units and the respective Co-factors. The role of both ACPs was distinguished using the hydroxybenzoate transfer reaction catalyzed by ApeH. Here, ApeE was assigned as starter- (4HBA-ACP) and ApeF as elongation-ACP (malonyl-ACP). In contrast, the role of both KS could not be clearly determined at that point, but the use of ApeO:C in the *in vitro*-setup produces a distinct APE-chain-length with up to seven double bonds (7xEN). The use of ApeR produces short chain-length APE product and the combination of both KS (+CLF) (ApeO:ApeC, ApeR) resulted in chain-length with up to 11 double bonds (11xEN). Supported by the structural knowledge of the APE-lipid obtained by the deletion studies with the APE BGC and NMR-data together with the HPLC-UV/MS and MS/MS analyses with APE-producer mutants (section 8), these results are interpreted as follows (Figure 37): It is proposed that the APE moiety is synthesized with the use of the ACP ApeE and AasS for loading 4HBA (or aminobenzoic acid) onto ApeE which is further elongated by ApeO:C with the interplay of the KR ApeQ and the DH-complex ApeI:P to afford in the 5-7xEN-APE products.

At the same time, it might be possible that a myristoyl moiety from primary metabolism is dehydrated by a hitherto unknown dehydratase and functions as a starter unit for the biosynthesis of the conjugated-FA. Thus, we speculate that during APE-lipid formation, the myristoyl-precursor molecule, derived from primary metabolism, is further activated by the AasS ApeH and transferred onto the second ACP ApeF, resulting in a *cis*-double bond (Figure 37). It is unknown whether the isomerization reaction occurs before the molecule is transferred onto the ACP or in the ACP-bound state. Another option might be that a β -hydroxy-FA precursor molecule gets dehydrated and isomerized by the ApeI:ApeP complex instead, comparable to the isomerization activity of FabA.^[79] The latter scenario is even more likely, as β -hydroxy-myristoyl FA represents an abundant FA within *Xenorhabdus* sp.^[178]

To date, cross-talks with the primary metabolism from FAS I/II systems are not well understood but are described for the generation of non-acetate starter units, as it was presented for the biosynthesis of frenolicin or R1128 (Figure 11).^[23,46] Additionally, with an interaction of a FAS I and a FAS II system, the very-long-chain lipids from *Mycobacterium tuberculosis*, the mycolic

acids, are produced for building up the characteristic cell wall.^[73] Here, a FAS I-synthesized-product is further processed with a FAS II system.

The ACP-bound FA moiety of the APE-lipid synthase is supposed to get elongated, reduced and dehydrated by the KS ApeR, the KR ApeQ and the DHs ApeL:P, resulting in 4xEN *trans*-double bonds (conjugated-FA). This ACP-bound conjugated-FA gets further transferred to a G3P-molecule by the G3P AT ApeD. Subsequently, this lipid-precursor is combined with the ACP-bound APE moiety by the glycosyltransferase ApeJ, which transfers a glucosamine moiety to the *para*-hydroxy/amino-group of the benzoyl-ring prior to getting transferred to the lipid-precursor. This bifunctional activity of ApeJ is postulated according to blastp and the sequence homology to Ppm1. Ppm1 is a glycosyltransferase working in dolichylphosphate biosynthesis to transfer at the one hand acyl- and on the other hand sugar moieties (Figure 37).^[179]

The exact time point of the methylation of the aryl-ring is not clarified yet. We assume that methylation does not take place at the very beginning of the biosynthesis, as no ApeB (MT) activity was observed with either 4-HBA nor 4-HB-ACP (ApeE) (7.1). But instead of a post-PKS, we rather suggest an in-line methylation due to the detection of only methylated products and the absence of non-methylated APE_{Xd}-methyl esters in early cultivation stages (Figure S8) The use of two ACP isoforms with different substrate specificities has already been found in plants, while one ACP showed specificity for short-chain substrates and the other ACP for longer ones.^[65]

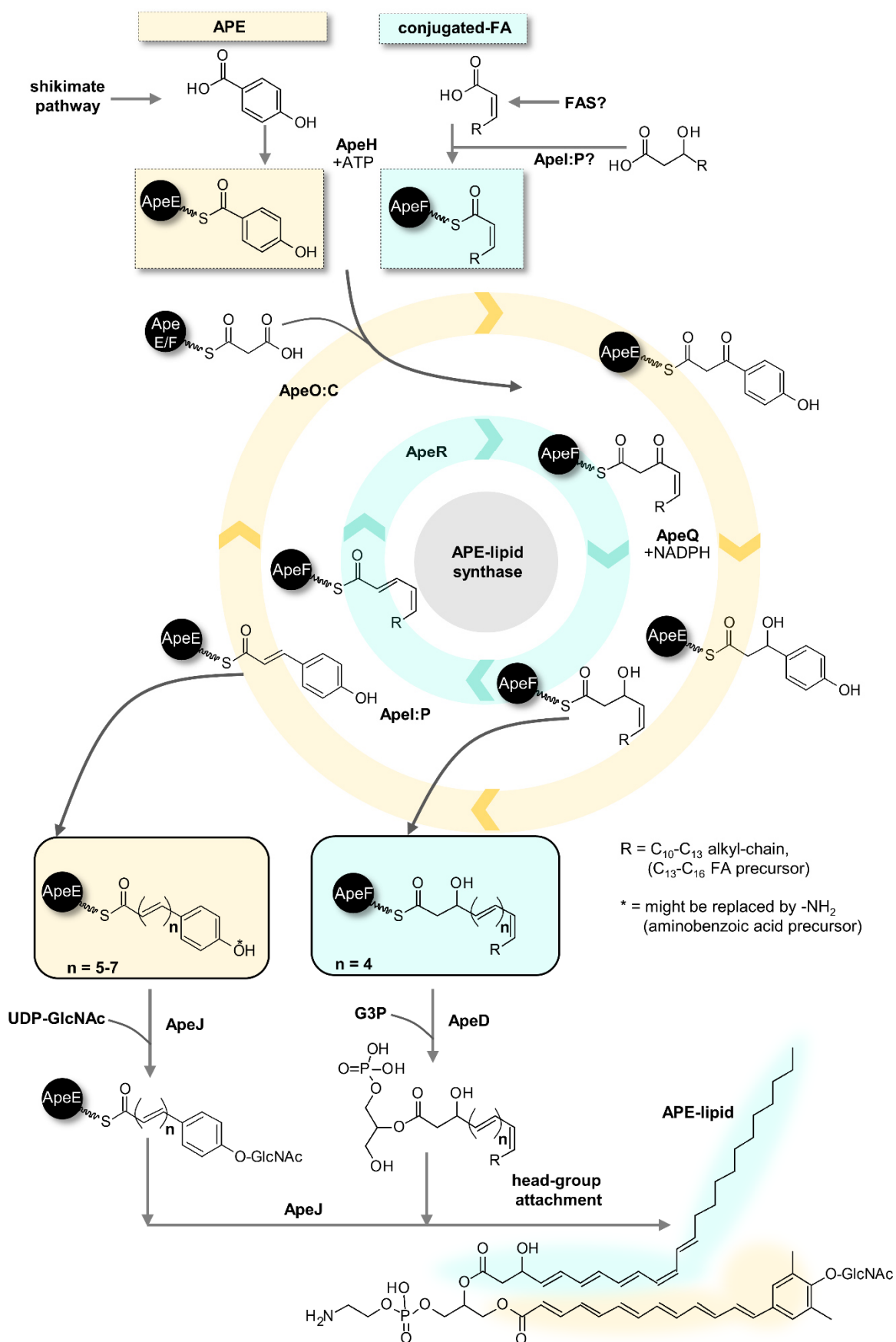


Figure 37. Proposed biosynthesis of APE-lipids. The biosynthesis-route for the APE part is highlighted in yellow, the route for the biosynthesis of the conjugated-FA moiety is marked in turquoise. The precursor derived from primary metabolism is adenylated (+ATP) and both acyl moieties are transferred

to the corresponding ACP through the action of AasS ApeH. Elongation takes place with either ApeO:C (APE part) or ApeR (conjugated-FA part) in a decarboxylative Claisen-condensation reaction with a malonate unit and results in the respective β -ketoacyl-ACP which is getting further reduced (ApeQ, NADPH) and dehydrated by the DH-complex ApeI:P. This cycle is repeated with a maximum of 7 built double-bonds (7EN) in the case of the APE part. The conjugated-FA gets not dehydrated in the last step but directly transferred to a glycerol-3-phosphate (G3P) by the G3P AT ApeD. The resulting G3P-Acyl catches the ACP-bound APE moiety with the help of the bifunctional glycosyltransferase after having attached a glucosamine-residue in *para*-substitution at the benzoyl-ring. Methylation must occur in situ, prior to the attachment of the phosphoethanolamine head-group.

The proposed biosynthesis scheme of the APE-lipid is supported by the results from deletion studies with APE-producer mutants (from additional results section 8). The deletion of *apeC* (CLF) still resulted in the synthesis of the conjugated-FA part. The same result is obtained by deleting *apeE* (ACP) and *apeJ* (glycosyltransferase) and *apeG* (membrane protein). This provides evidence that these genes are not directly involved in the biosynthesis of the conjugated-FA, but in the formation of the APE moiety or in the linkage reaction. As the exclusive production of the APE moiety was not observed, there is an indication that the APE moiety is either not synthesized or remains ACP-bound, if not being trapped by the conjugated-FA-G3P moiety.

This reaction order is in line with the hypothesis for the biosynthesis of the conjugated-FA (Figure 37). In the absence of the glycosyltransferase ApeJ or if the APE moiety is not produced (ApeK, ApeE and ApeC mutant), the conjugated-FA intermediate is accumulated. The deletion of the remaining genes resulted in the absence of an APE-lipid and led to the assumption that those genes are involved in both, the biosynthesis of the APE- and the conjugated-FA part (Figure 37). However, these hypotheses need further proof.

Clearly, the production profile of the different APE_{x_d}-methyl esters resembles the profile of the conjugated-FAs (section 8). It is therefore assumed that the chain-length of the APE moiety influences the chain-length of the conjugated-FA. Thus, it is assumed that the chain-length for building *trans*-double bonds is determined by the APE-machinery. One chain-length control point in the system is represented by the approximately 20 Å hydrophobic channel of the CLF, which is assumed to provide space for not more than 7xEN (with aryl head group) (7.1).

Since the most abundant APE_{x_d}-part only harbors 6xEN, it is suggested that the chain-length control might additionally involve downstream modifying enzymes. The exclusive support of *in vitro*-data let us previously suggest that the chain-length of the APE moiety might be controlled by the G3P-AT ApeD (7.1). Against this stands the gained knowledge from the *in vivo* studies that indicates that the glycosyltransferase ApeJ is the most likely candidate to control the APE-chain-length (compare Figure 37).

To further investigate the biosynthesis of the conjugated-FA, an *in vitro*-reconstitution with the proposed enzymes is highly recommended. On the one hand, to check whether a long precursor FA substrate gets incorporated as a starter unit that is further processed with the APE-machinery, on the other hand, to test the proposed transfer reaction by the G3P AT ApeD of the built acyl-ACP to glycerol. Therefore, a C₁₃-C₁₆-FA precursor can be loaded to ApeF with ApeH. The acylated-ApeF may further be processed in the presence of the KS ApeR, the KR ApeQ and the DHs ApeI:P to result in the ACP-bound conjugated-FA. Due to the fluorophoric system, product formation should be trackable. The transfer reaction of ApeD may be investigated by adding glycerol to the processed reaction, the cleaved acyl-glycerol product should be detectable in the supernatant. In turn, the glycosyltransferase reaction of ApeD can be investigated with *in vitro* synthesized APE-ACP (see 7.1) in the presence of UDP-GlcNAc and the synthesized acyl-glycerol product.

9.3 Physiological function of APEs – Really an Antioxidant Agent?

A linear conjugated π -systems provides high reducing potential to the APE-lipid molecules, making them potent antioxidants in lipid formations.^[144,180] Nevertheless, it is controversial if the antioxidant activity of the APE-lipids rather represents an ability than a function. The structurally closely related carotenoids or retinals provide some other independent physiological functions, i.e. screening from excessive light and spectral filtering or acting as a visual pigment in animals and as a chromophore in bacteriorhodopsin photosynthesis.^[181] Another physiological function for type I PKS-derived polyenes is assigned to the *trans*-polyene molecules amphotericin B and nystatin, produced by *Streptomyces nodosum* or *Streptomyces noursei*, respectively.^[146] In contrast to APE-lipids, these compounds are secreted by the host and enter the membrane of fungi by forming pores.^[182] These pores disturb the cell membrane by causing leakages for monovalent ions and therefore act as antifungal agents. In previous work, APE compounds were allocated as antioxidative agents, protecting the cell membrane from oxidative stress.^[144,145] While antioxidative properties were proven, additional functions remained unknown. It is conspicuous that organisms harboring an APE-gene cluster are only found within Gram-negative bacteria.^[138,183] The main difference between Gram-positive and Gram-negative bacteria lies in their cell membrane composition.^[184] Gram-positive bacteria have cell walls composed of a lipid membrane and a thick peptidoglycan-layer with teichoic acids, whereas Gram-negative bacteria possess a two-layered lipid-membrane with a thin peptidoglycan-layer with no teichoic acids but an additional lipopolysaccharide (LPS)-component not found in Gram-positive bacteria. Due to its negative charge, the LPS are prone to cationic antimicrobial peptide attacks.^[185]

The Gram-negative bacteria from *Xenorhabdus* species live in symbiosis with nematodes and together, they kill insect larvae during their symbiotic life-cycle.^[126,127] As part of the insect's

immune response, cationic antimicrobial peptides are present upon the nematodes entering the insect's hemolymph.^[186] The *trans*-conjugated acyl-chains of either the APE moiety or/and the conjugated-FA might function as a barrier for those peptides by mediating a more rigid membrane composition, thus protecting the bacteria. Furthermore, the fatty acid side chain being kinked at the *cis*-configuration (Figure 28) might act as an anchor through the opposite lipid-layer, as the *trans*-acyl chains alone (approximately 18 Å) would account for approximately half of the membrane's length. Membrane diffusion tests with fluorescence molecules as laurdan are used to measure the membrane rigidity in the presence of cationic antimicrobial peptides^[187] and could be performed with cells harboring the membrane incorporated APE-lipids in comparison with an APE-lipid deficient strain.

However, it may be hard to envision that the polar sugar moiety, of which we propose that it is located at the hydroxy-group of the APE moiety (compare Figure 38), might lay in-between the hydrophobic acyl-chains of the phospholipid membrane. Thus, it might be possible that the APE-lipids behave similar to phenolic glycolipids (PGL). PGLs are well-known from *Mycobacterium tuberculosis*, where they decorate the outermost layers of the cell wall.^[188] The aliphatic lipid-core of PGLs is hereby ω -terminated by a *para*-hydroxybenzoic acid building block and further *O*-substituted with typically one to four deoxysugars. By presenting the sugar moieties at the surface, PGLs act as virulence factors.^[169,188] A similar physiological function might be true for the APE-lipids, with the APE-sugar-part being analogously located at the outer phospholipid membrane, which represents hereby the outermost layer in Gram-negative bacteria.^[188]

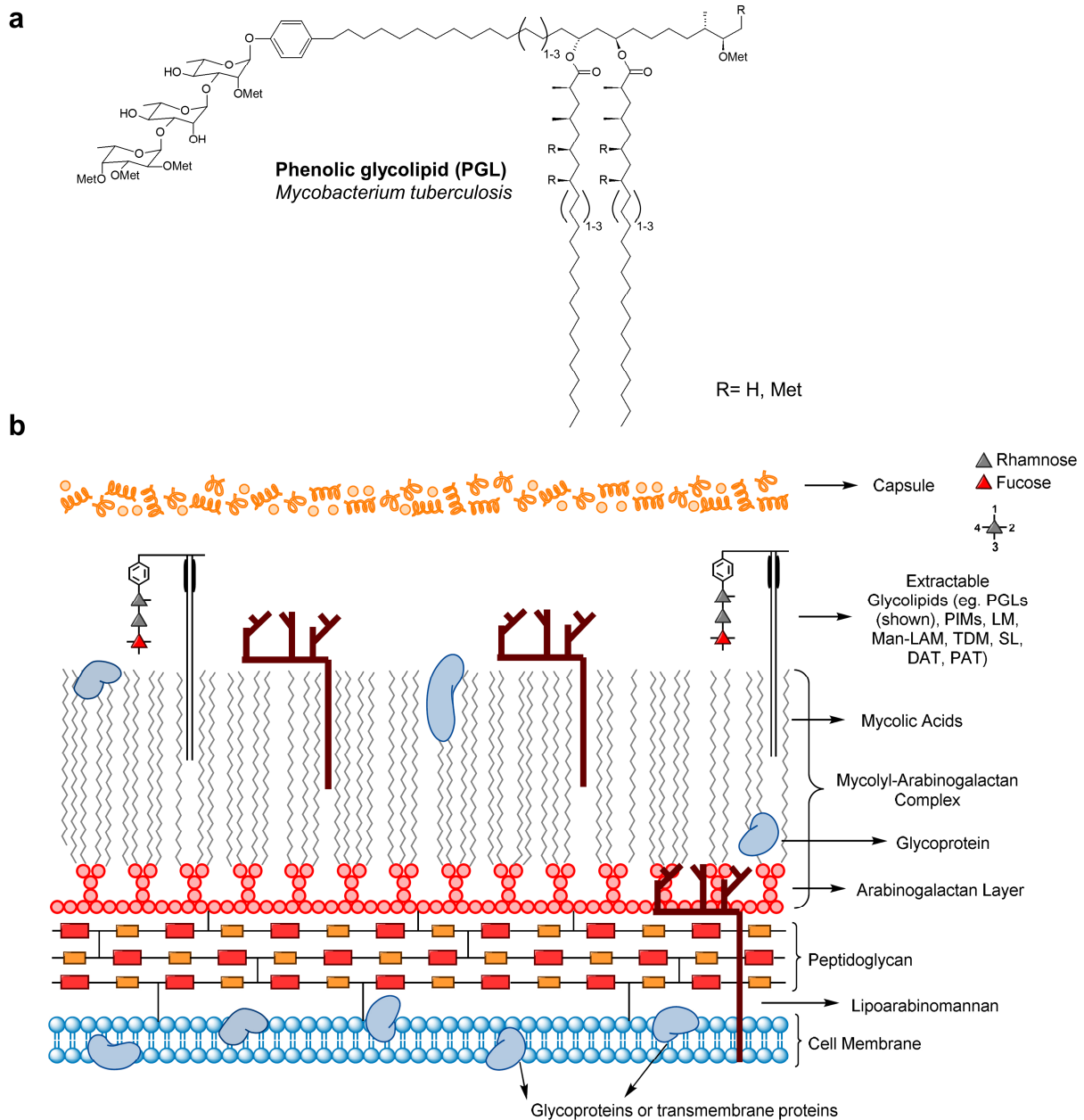


Figure 38. Structure of the phenolic glycolipids (PGLs) from *Mycobacterium tuberculosis* (a) and their localization in the cell wall. Figure b was used from reference ^[188] (b). PIM = phosphatidylinositol mannosides; LM = lipomannan; Man-LAM = mannose capped lipoarabinomannan; TDM = dimycolyl trehalose ; SL = sulfolipids; DAT = diacyltrehaloses; PAT = polyacyltrehaloses.

9.4 The APE-Lipid Synthase – Forming a Multienzyme Complex?

Within the APE BGC, blastp revealed several gene duplications for the key enzymes, ACP, KS and DH. Thus, we were interested in answering the questions of ‘why’ these gene duplications are present and ‘how’ a possible multienzyme APE-lipid synthase would look like. The role of both ACPs was already discussed under 9.2. The other two duplications (KS, DH) resulted in stable protein complexes (see 7.1). One KS (ApeO) of the cluster forms a complex with an N-terminal truncated CLF (ApeC). This truncation does not influence the overall protein folding, as observed in the X-ray data by adopting the typical thiolase fold of the KS.^[189] Typically, a CLF is encoded adjacent to the corresponding KS.^[62] Within the APE BGC, the CLF is encoded eleven genes upstream to its corresponding KS (Figure 23). The second KS (ApeR) forms a homodimeric complex (see 7.1), which is well-known for KS of FAS II systems.^[65] In native-ESI-MS analysis, ApeR seems to ‘prefer’ the monomeric state (see 7.1). This side effect is explained by high salt concentrations leading to monomerization in MS.^[189]

Albeit being six genes apart from each other, the DH (ApeI, ApeP), encoded within the APE BGC, form a heterocomplex (see 7.1). X-ray data of the complex suggest that the respective enzymes are not active in their monomeric state as the active center is built up with the interfaces of each subunit. An additional complex formation was observed between both ACPs and the TE (see 7.1). Interestingly, no hydrolytic TE activity against ACP-bound APE-molecules was detected, but shorter chain-length molecules or fatty acids were hydrolyzed. It is therefore suggested that the TE presumably functions as a proofreading enzyme as it was described for EntH involved in enterobactin biosynthesis.^[190] TE can exhibit diverse functions which can be roughly classified into families,^[191] although their function might not always be distinguishable by the primary sequence.^[192] Within publication under 7.1, we hypothesize that besides the suggested proof-reading activity, the tetrameric TE serves as a platform, being responsible for scaffolding the proposed multienzyme complex (Figure 39b). In a pull-down with tagged ApeE (ACP) analysis with enriched APE-proteins (*X. doucetiae* APE-producer), no further interactions with the ACPs were detected. Despite the tight interactions in the described complexes (KS:CLF, ApeO:ApeC; DH, ApeI:ApeP, ACP:TE ApeE:ApeK, ApeF:ApeK), other interactions with the ACPs are suggested to be more transient (unpublished data). To this end, a complex formation leading to a multienzyme complex, with the ‘remaining’ enzymes is possible but was not addressed so far. Namely the KR (ApeQ) the DH (ApeI:P) the KS:(CLF) (ApeO:C, ApeR), the AasS (ApeH) and the glycosyltransferase ApeJ may play a further role in multienzyme complex formation (Figure 37).

It should be noted that no interaction was observed in pull-down analysis with the KS ApeR and the KS:CLF complex ApeO:C (see 7.1). The highly flexible ACP:TE-complexes are

intended to scaffold a megaenzyme complex (Figure 39). To date, the only structurally described tight ACP-interactions in type II systems are complex-formations between ACP:FabZ and ACP:PPTase from FAS II.^[119,193] As it was stated in the introduction, all other complexes which were structurally confirmed, were cross-linked.^[65] Very recently, an ACP:kinase-complex from *Pseudomonas aeruginosa*, involved in LPS biosynthesis, was structurally confirmed.^[194] The kinase WaaP alone was insoluble unless the acyl-ACP was present. Thus, the authors proposed that the kinase WaaP gets stabilized with the acyl-ACP upon protein-interaction. Such an effect was also observed in the APE system, where the APE-ACPs tend to be more soluble upon co-expression with their TE (ApeK) interaction partner (unpublished data). In a tandem pulldown-analysis approach,^[195] with tagged *E. coli* AcpP, in combination with tagged marker proteins, a targeted interaction network with the *E. coli* proteome was discovered in another work. Due to the high sensitivity of nano-LC analyses, this approach allowed the detection of less tight interactions (in comparison to the tight complexes described in this work) and might be further applied to extend the knowledge about the APE-lipid synthase protein network.

However, during this work, the detection of these more transient interactions with the ACPs using native-MS methods was tried. Unpublished data (native ESI-MS measurements: Kudratullah Karimi, Morgner group) suggest an interaction of both ACPs (ApeE/F) with the AasS (ApeH), with slightly better interaction for ApeE:AasS, mainly formed between the monomeric proteins. This is in accordance with the hypothesis of the AasS being responsible for both precursor loading reactions: ApeE for loading with 4HBA/*para*-aminobenzoyl moiety and ApeF with β -hydroxy- or α,β -dehydrated alkyl moiety. It may be the case that the acylation state of the ACP may have an impact on the AasS/ACP interaction. This effect was recently faced with different NMR solution structures of acyl AcpPs^[196], but the effect of acylation still remained poorly understood. A postulated schematic overview of an interaction network of the enzymes from the APE BGC is depicted below (Figure 39). Although the multimeric forms of the several hetero-complexes and stand-alone enzymes have been addressed in this work, the following scheme represents an oversimplified interaction network, which is intended to give an overview of the discussed data and may vary under physiological conditions.

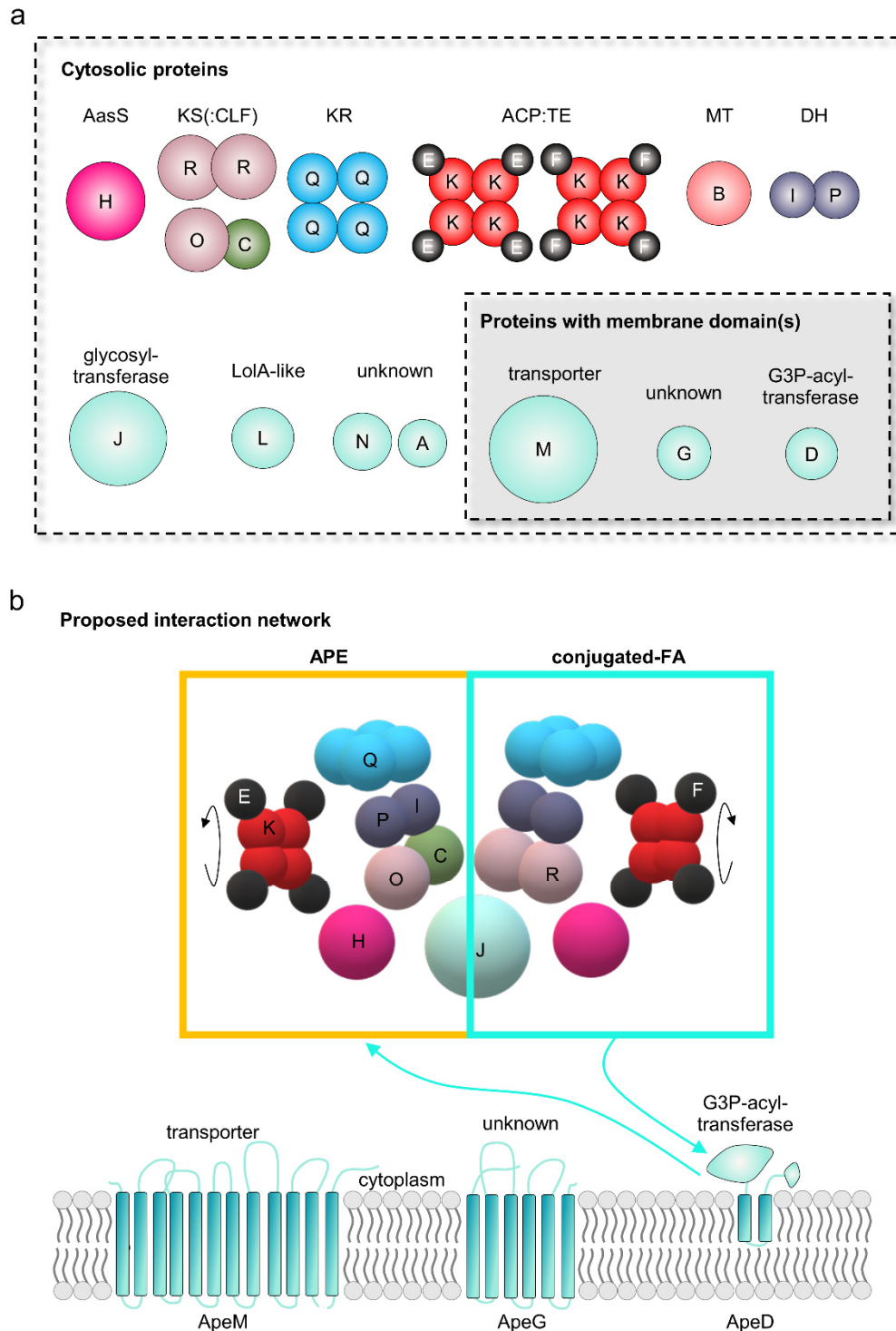


Figure 39. Proposed interaction network of APE-lipid synthase. Genes are color-coded according to the corresponding protein function, depicted throughout this work. Possible membrane domains were predicted using THHM-online-tool. The oligomerization grade of ApeB/J/L/N/A/M/D/G is unknown. The remaining proteins were analyzed throughout this work. In addition to the structurally solved complexes and the preliminary interaction data of ACP:TE, this scheme should provide a simplified overview of only one possible interaction network. Assuming a round-shaped structure for all proteins, protein sizes were relatively adapted according to their molecular weight.

For type II systems in general and throughout this work, ACPs play the central role regarding protein-protein and protein-substrate interactions.^[65] When compared with the enzyme domain

distribution in FAS I (or HR-PKS I) (compare Figure 12), it can be speculated whether the KS, the DH and KR enzymes are candidates for building the center of a multienzyme complex (Figure 39). To this end, it is proposed to perform protein pull-down analysis with KR (ApeR), DH (ApeI:ApeP) and additionally, with the bifunctional glycosyltransferase (ApeJ), as this enzyme is postulated to connect products of the APE- and conjugated-FA biosynthesis.

Notably, in FAS I systems, the KS domain plays a central role for dimerization of the homodimeric FAS I (HR-PKS) synthase (compare Figure 12).^[93] In the APE-lipid system, both KS form dimers (ApeO:C: heterodimer, ApeR: homodimer) but do not show any interaction with each other under the used conditions (7.1). According to the postulated biosynthesis of the APE-lipid, it is hypothesized that both KS, ApeO:ApeC and ApeR are responsible for the biosynthesis of either the APE- or the conjugated-FA-part (Figure 37). It can be speculated whether the multienzyme complex of the APE-lipid synthase is formed to mimic the multifunctional enzyme complex in FAS I systems (Figure 39, compare Figure 12) but with respects to the protein arrangement (no covalent linkage between the enzymes), the APE-lipid synthase might be more flexible.

It might be helpful to perform crosslinking proteomics, to catch a possible more non-covalent protein-protein interaction network. A method, also known as 'chemical cross-linking coupled with mass spectrometry (CX-MS) allows the covalent linkage of proteins close to each other, while the protein-interaction regions can be defined after tryptic digestion and MS-based protein sequencing.^[197-200]

It is possible that the flexibility of the APE-lipid synthase is necessary for producing either the APE- or the conjugated-FA-part (Figure 37). As known from FAS I systems, the product spectrum is stringent to saturated C₁₄-C₁₈ fatty acids^[74], in contrast to the broader product spectrum of FAS II systems (compare 6.2.2).^[70]

The concept of the formation of a large multienzyme complex would support the notion of a substrate channeling between the active sites, increasing the catalytic efficiency as a whole in the formed 'reaction chambers'.^[201] In type II FAS systems, the lack of this reaction chambers is thought to be balanced with higher copy numbers of the individual enzymes.^[75]

Nevertheless, a full imitation of type I systems with a protein architecture forming reaction chambers, compared to type I systems, may be hard to envision for the APE-lipid synthase. Generally, type II enzymes have active sites located at the bottom of narrow tunnels, with no opportunity for the Ppant arm to swing between them.^[201] It is therefore speculated that the APE-lipid synthase forms an 'intermediate synthase', a mixture of type I and type II systems' behavior.

Preliminary results of the *holo*-ApeF:ApeK-complex (unpublished data Maximilian Schmalhofer, Groll group) indicate a partially buried Ppant arm, being more exposed to the protein surface cave of the TE, rather than being protected in a hydrophobic pocket. These findings indicate a multienzyme complex formation and sequestering mechanism that may connect features from FAS I/PKS I, PKS II and FAS II systems (compare 6.2) and will be discussed in more detail under 9.6.

It is likely that an 'intermediate synthase' is formed due to the accommodation of the long acyl-cargo (Figure 35). In FAS II biosynthesis, an acyl-cargo with a chain-length of up to 10 carbons may be stabilized by the hydrophobic pocket of the ACP, and longer chain-lengths are known for being more susceptible to hydrolysis.^[196] While the APE moiety reaches carbon length of up to 18 carbons, the conjugated-FA reaches very-long acyl-chains of up to 26 carbons in length, being likely too long to get carried within a hydrophobic pocket of the ACP. A different type of interaction network, compared to the 'typical' interactions from type II, was discovered in the FAS II system from *Mycobacterium tuberculosis*.^[202,203] As already presented, this system is involved in producing the very-long-chain mycolic acids.^[73] The authors used a yeast-two/three-hybrid system^[202] to detect their protein interaction network. As a result, they described a change of the protein interaction-partners upon different stages of the biosynthesis, which may be a conceivable scenario for proteins of the APE-lipid synthase, too.

9.5 Special Features in IPS Biosynthesis

Being in clinical trials to work as a topical cream against psoriasis and mainly known by the name 'benvitimod',^[136] IPS and its stilbene derivatives are of special interest for natural product research.^[133] Stilbenes are a typical product of plant PKS III systems as a result of iterative elongation of cinnamoyl-derivatives with subsequent cyclization and aromatization.^[67,68] However, IPS is the only described bacterial stilbene product so far and is synthesized by a special PKS II system from *Photorhabdus* to synthesize the characteristic iso-branch.^[132,137,151] This iso-branch is derived from an alternative biosynthetic pathway through the condensation reaction of an iso-branched β -ketoacyl thioester, which is, in turn, synthesized from the BKD-pathway.^[151] This iso-branched precursor gets further cyclized with an elongated, reduced and dehydrated cinnamoyl-precursor, which serves as Michael acceptor, being necessary for subsequent reactions. Apart from the initial reaction step, which is carried out by the ammonia lyase StIA and results in the formation of cinnamic acid,^[132] the subsequent steps to synthesize the Michael acceptor remained obscure. Even though the deletion of the biosynthetically downstream enzyme StIB, which is annotated as a CoA-ligase, leads to the accumulation of cinnamic acid.^[132] Thus, it was believed so far that StIB synthesizes cinnamoyl-CoA, which is used as a precursor for further reactions.^[132]

9.5.1 Unraveling the Missing Steps in Biosynthesis of IPS

Having in mind a targeted engineering strategy to produce desired stilbene derivatives, StIB was characterized for its *in vitro*-activity with different cinnamoyl-derivatives and fatty acid-substrates (7.2). Surprisingly, StIB was more active for fatty acids than for cinnamic acid(-derivative) substrates and in fact, only minor amounts of cinnamic acid were detectable as cinnamoyl-CoA through StIB conversion. Consequently, StIB was tested for AasS-activity, already observed for ApeH with its corresponding ACP in APE biosynthesis (7.1). Congruent to our postulations, StIB also exhibited AasS activity with its cognate ACP StIE (7.2). The downstream reactions with cinnamoyl-ACP starter were confirmed to take place with enzymes from the FAS II-pathway to result in one cycle of elongation, reduction and dehydration of cinnamoyl-ACP through the activity of the KAS III-type KS *p*/FabH, the KR *p*/FabG and the DHs *p*/FabA or *p*/FabZ. Resulting in the *in vitro*-synthesis of the Michael acceptor being required for the next reaction steps (Figure 22).^[137,151] To obtain IPS, this product gets cyclized and aromatized by the unusual KS/cyclase StID and the aromatase StIC with the iso-branched β -ketoacyl thioester from the BKD pathway.^[137,151] Interestingly, within a pull-down analysis of the key enzymes StICDE, a complex formation between the ACP and the cyclase/KS was observed (unpublished data Siyar Kavakli, Bode group). According to size determination with SEC, it was not distinguishable between an (ACP)₂(KS)₂ or an (ACP)(KS)₂-complex. But due to a missing CLF-unit, a complex between a homodimeric KS^[152] and two bound ACPs

((ACP)₂(KS)₂), as already described for the cross-linked FAS II-complexes with KS-enzymes and ACP, would represent the most likely case for complex formation. However, this needs to be confirmed with precise native-MS-methods as previously performed (see 7.1). Unpublished data (Siyar Kavakli, Bode group) from a pulldown analysis of tagged-StIE and *P. laumondii*-lysate, indicates further interaction between StIE and the BKD-complex from primary metabolism. Controlled pulldown-analysis with the overproduced enzymes StIA, StIB, StIC, StID and additionally BKD- and FAB-enzymes should be performed, to create an interaction network comparable to the APE-lipid biosynthesis network (Figure 39). Due to the preferred specificity of StIB for fatty acid substrates, it cannot be excluded that StIB interacts rather with the iso-branched- than with cinnamoyl-molecules, to create the ACP-bound precursor used for the cyclization reaction. An assay using StIB and the iso-branched substrate in competition with cinnamic acid would reveal the preferred reaction of StIB, by comparing the amount of the formed acyl-ACPs.

p/FabH as a KAS III enzyme is typically known for using acyl-CoA substrates instead of ACP-bound ones^[67]. During the *in vitro*-elongation of cinnamoyl-ACP to the β -ketoacyl intermediate an excess of cinnamoyl-CoA substrate compared to cinnamoyl-ACP (1 mM cinnamoyl-CoA, 50 μ M cinnamoyl ACP) was used (7.2). The fact that the transfer reaction works with ACP-bound substrates indirectly indicates the use of the ACP-bound product rather than the CoA-intermediate as preferred *p*/FabH substrate. However, a kinetic analysis must confirm whether the acyl-ACP represents truly the favored candidate for *p*/FabH, compared to the respective acyl-CoA.

9.5.2 Gatekeeper Enzymes StIB and *p*/FabH

Due to the assumption that *E. coli* can complement the highly similar homologs *p*/FabH, *p*/FabG and *p*/FabA/Z, the heterologous production of stilbene derivatives failed so far and only the production of CHD and DAR intermediates were detectable.^[151]

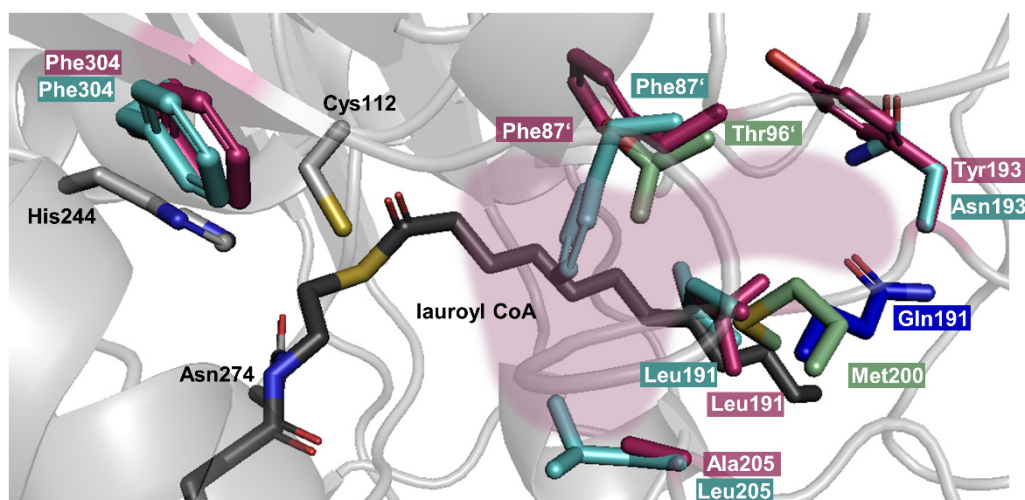
As observed in our *in vitro*-data (7.2), *ec*FabH was identified as the reason for the failure in heterologous production due to its inability to elongate cinnamoyl-ACP. In general, KAS III enzymes are well described as 'gatekeeper enzymes' in their respective pathways,^[30] predominantly in those involved in non-acetate starter loading of PKS II systems (compare 6.2.3).

Based on the *in vitro*-results, we have determined the minimal gene endowment for stilbene production and succeeded in generating stilbene products that originated from a *Photorhabdus* species in a heterologous host for the first time (7.2). Ethylstilbene was produced with heterologous expression of *stl(b)CDE* and *p*/FabH under cinnamic acid supplementation. In

the absence of the AasS StlB, ethylstilbene was still produced but in only minor amounts. Thus, the StlB addition increases the amount of ethylstilbene. At the same time, it is indicated that *E. coli* might partially complement the activation from cinnamic acid to cinnamoyl-CoA, which may be alternatively (instead of cinnamoyl-ACP) used for IPS biosynthesis (Figure 22). With the addition of the BKD-enzymes, it was further possible to produce IPS (7.2). The introduction of the downstream KR *p*/FabG resulted in a further increase of the IPS-production titer, which indicates an enhanced enzyme-substrate orchestration in the heterologous host with the homologous FAS II enzymes. These findings can be further used for engineered IPS-production to produce different derivatives by targeted mutation analysis. It may be possible to mutate *p*/FabH, to produce for example the iso-branched variant of resveratrol, which is in the discussion as a therapeutic agent,^[204] by using *para*-coumaroyl-precursors instead of cinnamic acid ones. So far, this variant cannot be produced due to discrimination of the *para*-hydroxy-substitution from both, StlB (see 7.2) and *p*/FabH (unpublished data Siyar Kavakli, Bode group).

For this purpose, crystal structures of the gatekeeper enzymes StlB and FabH would be advantageous. AMP-ligases, like StlB, have been proven difficult for crystallization attempts.^[205] A promising candidate for crystallization studies would therefore represent *p*/FabH. To get first insights into the binding pocket probabilities, even without a structure, a model using Phyre2 was created and compared with previously reported structures of FabH-homologs that are known for the accommodation of different substrates (Figure 40).^[65,85,206]

a



b

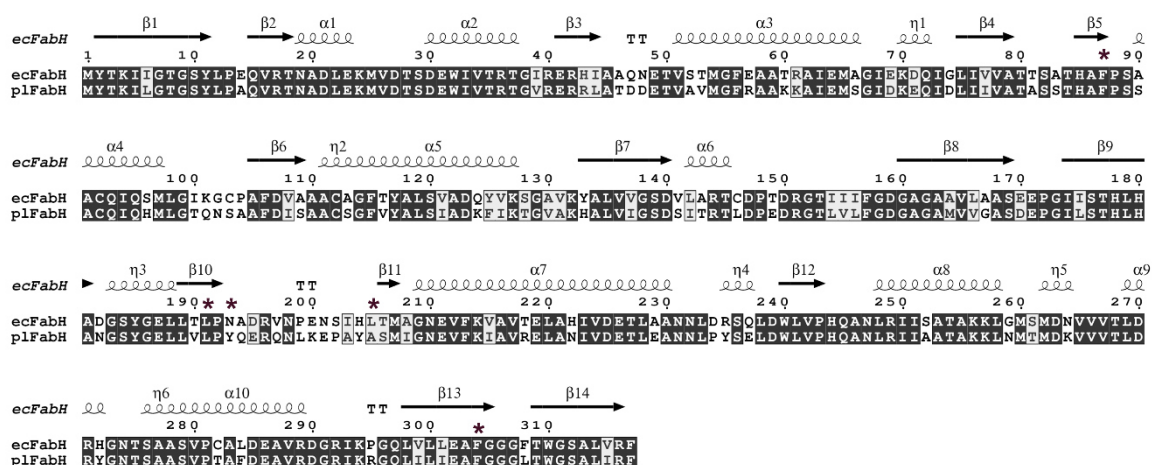


Figure 40. Proposed aa involved in binding pocket shape of modeled *p/FabH*. Residues Phe87', Leu191, Tyr193 and/or Ala205 may contribute to forming the binding pocket (colored in blurred magenta) of *p/FabH* to accommodate the aromatic cinnamoyl moiety for condensation reaction. *p/FabH* model was created using the online tool Phyre2 in intense mode. Pymol 2.3.2 was used for visualization of *p/FabH* model, main chain is depicted in cartoon and aa side chains in stick representation (main chain: grey; aa side chain: magenta) in comparison to known structures of *ecFabH* (PDB 1EBL; aa side chain: cyan) accepting straight-chain substrates,^[84] ZhuH (PDB 1MZJ; aa side chain: pale green) accepting iso-branched substrates^[206] and *mtFabH* (PDB 1U6S; aa side chain: blue) accepting very-long straight-chain substrates.^[85] *mtFabH* is shown with its lauroyl-CoA ligand (stick representation: black). Active site aa residues Cys112, His244 and Asn274, are depicted in grey (a). Alignment of *ecFabH* and *p/FabH* with a 73% sequence identity was performed with T-Coffee^[207] and visualized with ESPript3.x.^[208] aa residues, proposed to be involved in diverged substrate binding of *p/FabH*, are marked with an asterisk (b).

Even though *ecFabH* and *p/FabH* show a high sequence identity of 73% (Figure 40b), their substrate specificities are obviously different (7.2). This demonstrates that minor differences in the amino acid configuration may have an impact on the catalyzed reactions. The model of *p/FabH* was compared with structures of *ecFabH* from FAS II,^[84] accepting only C₂-C₄-straight chain acyl-CoAs, ZhuH from a type II PKS (involved in R1128 biosynthesis, see 6.2.3) accepting C₂-C₄-branched chain-substrates^[206] and *mtFabH* accepting very-long C₈-C₂₀

substrates.^[85] *mtFabH* was additionally depicted with its native lauroyl-CoA-substrate (Figure 40a).^[85] Some gatekeeper residues are known for KAS III-enzymes for starter substrate selection.^[65] A known gatekeeper residue is phenylalanine Phe87' (from the second subunit)^[65] in *ecFabH* and *p/FabH*. This residue is substituted by a threonine (Thr96') in *ZhuH*. The threonine suggests more flexibility for the iso-branched substrates in *ZhuH*.^[206] In *p/FabH*, the Phe87'-rotation is pointing away from the binding pocket, resulting in a wider pocket shape (Figure 40a), in contrast to Phe87' in *ecFabH* pointing to the inside, which is supposed to discriminate iso-branched substrates.^[209] In *p/FabH*, at residue 193, a tyrosine (Tyr193) is located and might be involved in π -stacking interactions of the aromatic cinnamoyl-substrate.^[210] In contrast, *ecFabH* instead possesses an asparagine residue (Asn193) (Figure 40a).

The Leu191-residue from *ecFabH* did not change in *p/FabH* and may serve as a restriction for long-chain acyl-substrates. The corresponding residue Gln191 in *mtFabH*, in contrast, ensures a more open shaped binding pocket to accommodate the long lauroyl-CoA substrate. In *ZhuH* (iso-branched substrates), the respective residue is occupied with a methionine (Met200), instead. The gatekeeper Phe304 in *ecFabH* was suggested to be involved in shaping the pocket at the 'entrance'.^{[65][209]} Due to the fact, that in *p/FabH* and *ecFabH* exhibit the same orientation of the Phe304, it is more likely that the phenylalanine mediates solvent and substrate sequestering, as it was described for the gating residue Phe400 in the elongation KS *FabB/FabF* (compare Figure 19).^[120]

However, also changes in unstructured regions might be considered for shaping the binding pocket. For example, the loop region close to the thiol of the CoA-substrate differs significantly between *ecFabH* and *p/FabH*. The sequence of *p/FabH* harbors a proline and proline-residues are well-known as helix-breaker.^[211] It is possible that a change in this loop of *p/FabH* provides more space, according to cinnamoyl-substrates. Additionally, the *ecFabH* Ala205 is replaced with Leu205 in *p/FabH* and may contribute to more space of the binding pocket.

Nevertheless, the *in silico* predicted diverged binding pocket of *p/FabH* needs further experimental evidence. Mutagenesis studies of the involved aa, combined with structural data of *p/FabH* will clarify the substrate binding of the bulky cinnamate in *p/FabH*.

9.6 Comparing ACPs from APE, IPS and AQ biosynthesis

In the introduction, the present knowledge regarding protein-protein interaction in type II systems was described (6.3.2). The manuscripts and publications within this work enlarge the current knowledge with biochemical and structural information of the three PKS II systems from the entomopathogenic bacteria *Xenorhabdus* and *Phothonabdus* (predominantly, 7.1, 7.2, 7.4). Only minor experimental contributions were made for the manuscript under 7.4 (AQ system), but due to the high impact regarding the structure and biochemistry of PKS II systems, the

results of the investigations regarding the minimal PKS complex were considered for this discussion section.

An ACP plays a central role during the chain-sequestering of the acyl-cargo and at the same time, as a platform for protein-protein interactions in FAS II and PKS II systems (7.1). In the sequence alignment shown below (Figure 41), ACP-candidates from five different organisms are represented for each PKS II system, including ApeF, ApeF from the APE-lipid synthase, StIE from the IPS synthase, AntF from the AQ synthase and AcpP from FAS II. Additionally, ACPs from typical type II PKS and two ACPs involved in FabH-priming are presented.

According to literature,^[29,60,110,111,212] the most likely localization for protein-protein recognition sites lays either in between helix and loop II and/or III (Figure 41). In Figure 41 the 'hot spot areas' for recognition of the AcpP with enzymes from primary metabolism^[110] are presented and can be compared with interaction sites, obtained from the AQ-minimal complex data (ACP:KS:CLF, ACP:*p*/MCAT from 7.4) (Figure 41, 1.-3.,4. primary metabolism).

Additionally, important interaction sites for the KS FabH with AcpP from FAS II are depicted (Figure 41, 5. FabH). In manuscript from 7.2, a crucial cross-activity between *p*/FabH from primary metabolism and the ACP StIE from specialized metabolism was presented. If comparing the essential interaction sites of the AcpP with the ones in StIE, it is obvious that they share high similarities and might be the reason for FabH recognition.

None of the ACPs studied within this work harbors the specific tyrosine residue, typically present in ACPs for FabH recognition during non-acetate priming reactions in PKS II systems (Figure 41, 6. orthogonal ACP FabH, see also 6.2.3).^[88] AQ-ACP (AntF) is primed with acetate through decarboxylation of malonate. IPS-ACP (StIE) and the APE-ACPs (ApeE, ApeF) are primed with an AasS enzyme with benzoyl and cinnamoyl-starter, respectively (see 7.1 and 7.2).

The AasS ApeH is highly specific to the APE-ACPs, as no acyl-product was built upon substitution with *E. coli* AcpP (unpublished data). Thus, both APE-ACP need to feature amino-acid recognition sites, which are different to the residues found in AcpP. This is in contrast to the more flexible AasS-activity of StIB, accepting both, its cognate ACP StIE and AcpP, to produce the acyl-ACP, which implies a similarity in the recognition sites for StIB in StIE and AcpP. Again, the most likely section for these differences presents the recognition helix II.

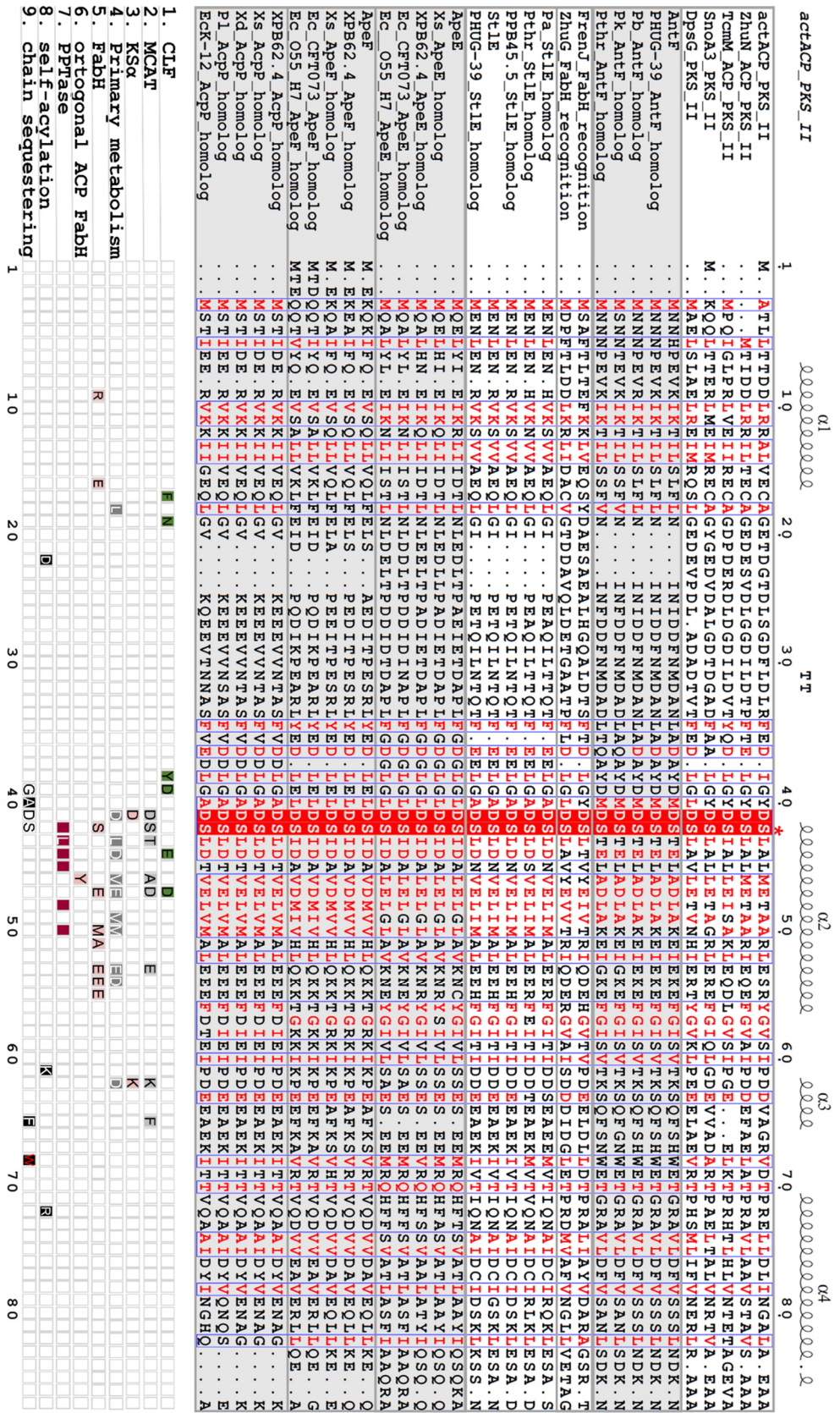


Figure 41. ACP alignment with ACP from FAS II and PKS II systems. Each type is represented by five representative ACPs, the ACP involved in orthogonal FabH recognition is exemplified with two representatives. Systems studied within this thesis are highlighted in grey. The alignment was performed with Geneious 6.1.8 and visualized with ESPript 3.0 and depicted with secondary structural elements of the actinorhodin ACP (PDB: 2K0X). The numbering holds for actinorhodin ACP. Interaction sites are summarized in 1.-9. from 7.4 and [88, 110, 212-216]. The Trip (W)-residue, which blocks the entrance of the hydrophobic pocket in Antf (7.4), is marked in red under 9.

Unpublished data from AntF (AQ biosynthesis) showed a self-malonylation in the presence of malonyl-CoA, in contrast to the APE-ACPs, where a self-malonylation was not observed (7.1). At first glance, the self-acylation reaction of an ACP makes an MCAT redundant. However, the reported induced fit-mechanism of manuscript 7.4 of the *p*/MCAT to synthesize malonyl-ACP, ensures both, the specificity and efficacy of this reaction step and is in accordance with previously reported higher turnover rates for an MCAT-catalyzed malonylation compared to the self-malonylation.^[213]

In initial reports, the self-malonylation process was reported as a phenomenon occurring only in PKS II-type ACP.^[214] However, analyses of Misra *et al.* with FAS II ACPs from *Plasmodium falciparum* and *Brassica napus* provided evidence that the catalytic self-acylation is intrinsic to the individual ACP and may be observed in type II FAS ACP, too.^[213,215,216] The work of Misra *et al.* distinguished further, in the ability for self-acylation and the transfer activity (transfer of malonate from malonyl-ACP to a *holo*-ACP) and stated the aa involved in both processes. Hence, self-acylation was described to be dependent on acidic acid residues Glu/Asp in loop 1 (between helix I and II, see Figure 41, 8. self-acylation). *E. coli* AcpP was reported to fail in self-acylation but shows transferase activity. Transferase activity is mediated by an Arg/Lys-residue in loop III and the Arg/Gln residue at the start of helix IV. According to these assumptions, AntF and ApeE should be able to perform self-acylation, which is only valid for AntF. Transferase activity might be true for ApeE but was not yet tested. StIE from IPS biosynthesis does not harbor the typical aa, neither for self-acylation nor for transacylase activity. However, none of the reactions were investigated so far for StIE.

Prior to performing *in vitro*-analyses with an ACP, initial studies should deal with the phosphopantetheinylation state after protein production in a heterologous host.

Each organism has at least one PPTase, which is needed to activate ACPs involved in fatty acid biosynthesis; PPTases are also assigned as AcpS-type PPTases. There also exist Sfp-type PPTases, which are involved in the activation of ACPs from specialized metabolism.^[217]

The heterologous production of the APE-ACPs in *E. coli* led to acylated ApeF and the *holo*-form of ApeE. It is speculated that ApeE is not further modified by the heterologous production host due to its specific interaction with AasS, while ApeF may obtain its unknown modification from interaction with FAS enzymes (see postulated biosynthesis, scheme Figure 37). *E. coli* harbors three endogeneous PPTases, AcpS, EntD and AcpT.^[218] AcpS is involved in activating ACPs from fatty acid biosynthesis, EntD in contrast, activates the ACP involved in siderophore (enterobactin) production. AcpT was reported to be highly specific for ACPs not found in the *E. coli* BL21 strain but in its ancestor.^[218] The publication under 7.1 displayed that the gene cluster, which encodes for those ACPs, is retrospectively assigned as an APE BGC. Thus,

AcpT is orphaned in the BL21 strain, but phosphopantetheinylates the APE-ACPs, while an *acpT*-deletion mutant generated *apo*-ACPs exclusively (7.1). This specificity of AcpT may be explained by the substitution of the Leu-residue, downstream the conserved serine of the ACP (DSL) with an Ile in APE-ACPs (Figure 41, 7. PPTase). This residue and some adjacent aa were shown to be important for PPTase recognition.^[29,111,212] In *X. doucetiae*, AcpT is not present but might be complemented by the promiscuous Sfp-type PPTase NgrA (*xdv3_80277*). In contrast, the production of AntF (AQ biosynthesis) requires its own PPTase during heterologous and homologous production. Here again, the Leu-residue is substituted with a threonine, which is proposed to present the 'specificity point' for the cognate PPTase AntB. The heterologous production of StIE (IPS biosynthesis) leads to a mixture of *apo/holo*-ACP. The deletion of neither *acpT* nor *entD* resulted in the complete production of *apo*-ACP (unpublished results), which implies an activation, caused by AcpS. However, the deletion of *entD* resulted in a reduced *holo*-ACP level. Due to the DSL-sequence, these observations are in tune with typical AcpS-activation ability.^[217]

In the introduction it was stated that different 'transport mechanisms' of the ACP-bound acyl-cargo depending on the different PKS/FAS-types exist (Figure 15). While possessing highly effective reaction chambers, the acyl chain is proposed to be located at the outside of the ACP but sequestered in these chambers in type I systems. In contrast to type II FAS systems, where the acyl-chain is buried within a hydrophobic pocket and is only flipped out upon protein-protein interactions (chain-flipping mechanism) (compare 6.3.2). This again stands in contrast to the transport mechanism of the acyl-cargo in type II PKS. Here, the acyl-chain is only partly protected within the ACP. It is known that the chain-elongation takes place in the sequestered channel of the KS:CLF-dimer while the ACP changes position between the MCAT and the KS:CLF-dimer to reload and transfer malonate elongation units (compare Figure 17). After reaching the final length, the acyl-chain is transferred back onto the ACP, while the partial protection in the 'surface cave' (rather than a pocket) is proposed to be advantageous for effective reactions of the downstream enzymes (compare Figure 17).^[65] In manuscript under 7.4, for the first time, this translocation process of the ACP from a minimal PKS was observed in action in AQ biosynthesis. The procedures applied revealed important protein-protein-interaction residues of the ACP:KS:CLF (AntF:AntD:AntE) and the ACP:MCAT (AntF:*p*/MCAT) complex.

One of the discoveries was the conserved Trp-residue of the ACP AntF, blocking the way for the ACP-bound acyl chain to move into a hydrophobic pocket (Figure 41, 9. chain sequestering), in the sequence of AcpP this position is substituted with an Ile, the same for StIE (IPS) while ApeE and ApeF possess Met- and Val-residues, respectively.

Preliminary structural results from the ApeF:ApeK-complex (ACP:TE) indicates an acyl-chain which is located at the outside of the ACP, turning towards a 'surface cavity' of the TE (unpublished data: Max Schmalhofer, Groll group). This result would be in accordance with the 'surface cavity' described for PKS II systems (Figure 15) and at the same time did not agree with FAS II-like behavior (compare 6.2.2.2).

In the preliminary structure, the Ile is not responsible for the surface representation but rather a Phe residue, which is pointing towards the direction of the hydrophobic pocket and may block the entrance, comparable to the Trp-residue of AntF (Figure 41, 9. chain sequestering). Recently, a GXDS-motif of the ACP was described to have an influence in chain-sequestering (Figure 41, 9. chain sequestering).^[219] Hereby, X is represented by more bulky residues in type I ACPs (acyl-chain not sequestered) but alanine residues in type II FAS-like ACPs (acyl-chain sequestered in hydrophobic pocket). At the same time, the equivalent of the Phe residue of ApeE is described with the same criteria, bulky residues are present in ACPs from type I and Ala-residues in type II FAS ACPs. Thus, this theory is in accordance with our preliminary results, with the bulky Phe blocking the hydrophobic pocket. Furthermore, both APE-ACPs possess Leu-residues in the X-position of the GXDS-motif, AntF a Met and StIE an Ala. Hence, chain-sequestering in a hydrophobic pocket for StIE may be an option, as the second aa, substituting the Phe of ApeF, displays an Ala, too (Figure 41, 9. chain sequestering).

Due to the fact, that no minimal complex is built between the ACPs from the APE system and because of the fatty-acid-like elongation mechanism (7.1), it is postulated that the APE-lipid synthase behaves not like a PKS II according to chain-sequestering.

As we proposed in 7.1, the ACP:TE (ApeK) might act as a platform for protein-protein interactions. This protein-architecture may lead to a different kind of 'chain-flipping' than known so far, which may start from the TE-platform upon protein-protein interactions.

This transport mechanism may be formed due to accommodation of the long acyl-side chains and the use of benzoic acid starter units (compare Figure 37). AcpM as an example, the ACP involved in the biosynthesis of the previously described very-long-chain mycolic acids from *Mycobacterium tuberculosis*, with chain-length ranging from C₇₀-C₉₀,^[73] shares high similarities to AcpP but harbors a characteristic C-terminal extension which is suggested to interact with the unusual long substrates.^[220] Neither ApeE nor ApeF possess such a C-terminal extension (Figure 41) for the C₂₆ substrates from APE-lipid biosynthesis (compare Figure 37).

10 Conclusion

This work deals with different PKS II systems from Gram-negative bacteria and provides not only interesting insights into the protein but also on the created product level. Remarkably, in all PKS II systems, APE (7.1 and 8), IPS (7.2) and AQ (7.3, 7.4 and 7.5)), the interaction with the primary metabolism is required. Thus, the boundaries between primary and secondary/specialized metabolism are flexible but are well-controlled by the respective biosynthesis machinery.

In section 7.3, it has been shown that small molecules, like cinnamic acid, may have a regulatory effect on the biosynthesis of AQ. This effort gives only a small idea of the complexity of the regulation of the various biosynthetic pathways, which can be based on antagonistic or synergetic effects of natural products,^[221] regardless of whether they originate from primary or specialized metabolism.

In IPS biosynthesis, it has already been shown that primary metabolism is required to synthesize the iso-branched β -ketoacyl intermediate from the BKD pathway, which in turn is cyclized to the IPS with a second intermediate serving as Michael acceptor.^[132,151,152] This work revealed that enzymes from FAS II biosynthesis enable the synthesis of this Michael acceptor (see 7.2), which has been proven through *in vitro*-reconstitution.

An interesting finding arises during investigations of the postulated CoA-ligase StIB. So far, it was assumed that cinnamic acid is activated by StIB to the acyl-CoA thioester.^[132] However, our analyses demonstrated that StIB acts as AasS and is able to load the acyl residue of adenylated cinnamic acid directly onto the ACP. This result represents an essential finding for the completion of the IPS biosynthesis scheme. StIB, as an AasS enzyme, displays one of in total two AasS enzymes discovered and biochemically characterized within this study. AasS enzymes typically occur in FAS systems in order to transfer palmitate to an ACP for lipid biosynthesis^[222,223] but have recently been of great interest for loading ACPs, by using them as biochemical tools. It is useful to note that the substrate specificities can vary depending on the origin host organism so that even bulky substrates may be accepted.^[224,225] AasS, in general, have a flexible substrate spectrum, but so far, no activation of α,β -unsaturated substrates were presented^[224,225], as it was demonstrated for StIB.

The discovery of StIB as an AasS enzyme postulated that the cinnamic acid precursor molecule is not used as CoA but rather as ACP-bound intermediate by the downstream ketosynthase FabH of *P. laumondii*. It was further demonstrated that FabH from *P. laumondii* is mandatory for the elongation reaction and cannot be replaced by the *E. coli* homolog. A model protein structure of *pIFabH* was created to identify possible aa involved in the formation of the altered binding pocket.

Conclusion

So far, no stilbene products using ACP, ARO and KS/CYC (StlCDE) with cinnamic acid supplementation have been produced in the heterologous *E. coli* host.^[132] With the gained knowledge, heterologous ethyl stilbene production was realized for the first time, by adding *p*/FabH and the AasS enzyme StlB during heterologous production in *E. coli* (7.2). Furthermore, the production of IPS with the addition of the BKD enzymes was achieved. These studies may form a promising basis for controlled engineering strategies in the production of novel stilbene derivatives.

Novel findings within the underexplored APE biosynthesis protein machineries and the corresponding compounds were obtained from the analyses conducted within this work (7.1 and 8). The biosynthesis of the APE part was investigated and successfully reconstituted *in vitro* by the interaction of nine proteins and takes place in a FAS II-like reaction manner (7.1). Among these enzymes, the second AasS, ApeH, was discovered in addition to StlB (AasS, IPS biosynthesis) described above. In contrast to the hitherto described promiscuous activity of AasS enzymes against ACPs, this enzyme showed a highly specific activity against APE-ACPs but is able to use a variety of benzoyl-derivatives as substrates.

The exact role of both KS, one which forms a heterodimer complex with a CLF (ApeO:ApeC), and the second KS (ApeR), which forms a homodimer was not unambiguously assigned under (7.1). Nevertheless, it was demonstrated that different lengths of APE products were created by the addition of different KS in an *in vitro* setup with the remaining enzymes. These observations were of particular interest with respect to the investigations at the level of the APE compounds produced in *X. doucetiae*. APE compounds were originally presented as small molecules. However, the analyses within this thesis showed a further processing of the APE molecules to an APE-lipid-like structure (section 8). Due to the conserved 'lipid-genes' within the APE BGC, it is assumed that APE molecules are generally part of APE-lipids rather than the so far presented APE-alkyl esters. It is assumed that the structure of the APE-alkyl esters APE_{Vf}, APE_{Ec} or xanthomonadine, as representative examples of this compound class, resulted from an artefact of the extraction process. It was revealed that the APE-lipid in *X. doucetiae*, in addition to the APE moiety, consists of an unusual conjugated-FA moiety, a sugar moiety and a postulated glycerol backbone with a phosphoethanolamine head group.

The combination of *in vitro* and *in vivo* studies of the APE compounds, let us speculate that the different KSs are associated with different synthesis reactions. It is assumed that the KS:CLF complex ApeO:ApeC is used to produce the APE portion and the homodimeric KS ApeR, in contrast, to produce the unusual conjugated-FA. This hypothesis can be further transferred to the two ACPs. One ACP might be used for the synthesis of the APE part, while

the other ACP will be involved in the biosynthesis of the conjugated-FA. Although the APE-lipid could not be purified as a whole, the structural elements, in combination with deletion studies, allowed us to propose the order of the biosynthesis and led us to a biosynthesis scheme for the APE-lipid. It is assumed that, in addition to the APE machinery, an interaction with a FAS machinery takes place in order to link a saturated FA side chain. It was clearly established that the *trans*-double bonds of the unusually conjugated-FA are synthesized by the APE machinery.

Based on this work, it is assumed that the glycosyltransferase ApeJ plays a crucial role in connecting the APE part with the remaining building blocks, in order to obtain the APE-lipid. The postulated biosynthesis scheme led us to the hypothesis of a highly flexible APE-lipid biosynthesis machinery (Figure 37).

Further investigations should address a general rule of product formation in the different APE families. It should be touched upon whether the conjugated-FA is a typical structural element of the APE subfamily 1 (and possibly APE subfamily 2), whereas, in the case of the APE subfamily 3, DARs are rather bound to the APE units, as is the case with flexirubin, as an example. The analytics established in this thesis for the detection of APE (-related) compounds offer access to a simple screening for such a classification.

Taken together, these results represent an exciting approach for natural product research, as the APE BGC is widespread within proteobacteria and future research may focus on further investigations of the biosynthesis of the conjugated-FA and the linkage of a glycerol backbone with the APE part.

Due to the presented structure of the APE-lipid, the exclusive activity as an antioxidant is questioned. It is speculated that APE-lipids may be responsible for enhancing the rigidity of the bacterial outer cell membrane, to protect the organism against cationic antimicrobial peptides or, similar to phenolic glycolipids, they might act as virulence factors by covering the cell membrane with the sugar moieties.^[169,188] However, the exact physiological role remains elusive.

In this work, several protein pull-down analyses have been successfully performed to investigate possible protein-protein interactions of the different PKS II systems (7.1, 7.2 and 7.4). Thus, not only the well-known KS:CLF (ApeO:ApeC) complex was discovered in the APE system, but also the unexpected DH heterodimer complex (ApeI:ApeP), as well as the unusual heterooctameric ACP:TE complex (ApeE:ApeK and ApeF:ApeK). This resulted in two X-ray structures and promising structural data for the ACP-TE complexes through the work of the Groll group. The ACP:TE complex is proposed to play a role in scaffolding the suggested APE-lipid multienzyme complex. The KS:CLF complex from APE biosynthesis is one of three

KS:CLF complexes that have been structurally characterized so far. Another complex is represented by the KS:CLF complex AntD:AntE from AQ biosynthesis. Initial investigations in the IPS system also led to the detection of protein-protein interactions. It was discovered that the special KS/CYC StID forms a complex with the ACP. Further complex formations in this rather untypical type II system are offering an interesting subject for future studies.

These results go hand in hand with the X-ray structures from the Groll group, which show the first minimal PKS by elucidating the protein structure of an ACP:KS:CLF (AntF:AntD:AntE) and an ACP:MCAT (AntF:p/MCAT) complex. These investigations represent rare examples of tight interactions within PKS II systems, or generally of those from type II systems. Only three examples have been structurally addressed so far: a KS:CLF complex from actinorhodin synthase^[61] from a PKS II system and a heterododecameric DH:ACP complex (AcpP:FabZ)^[121] as well as an ACP:PPTase complex (AcpP:AcpS)^[193], both from a FAS II systems. Within this work, not only the strong potential of pull-down analyses within type II systems or more generally in the search for potential protein- interaction partners is described, but also the key role of ACPs in the center of type II systems is strengthened.^[65]

With the herein conducted studies it can be concluded that protein partners do not necessarily have to be in genetic proximity.

With the gained biochemical and structural insights presented within this work an overall comparison of the different ACPs (APE, IPS, AQ) from type II PKS and FAS systems on the aa-sequence level was possible. We postulate that the transport mechanism for the acyl-cargo in the APE-system might vary from the typical sequestering mechanism of type II PKS or FAS systems (compare Figure 15) and may be further analyzed with targeted mutagenesis of aa proposed to be involved in this process.

Future investigations should try to solve the riddle of a possible multienzyme complex by including the native substrates. Furthermore, CX-MS-based proteomics might be useful to complement the data of X-ray crystallography and native MS measurements by catching all the interacting proteins.^[197-200] Analyses with proteins from other PKS II systems may reveal how outstanding the examples within this work are, or how well they fit into a general model.

Finally, this work represents an in-depth analysis of three different PKS II systems and elegantly leads from the protein to the product level of the different synthases and opens up new possibilities for natural product research. Through a better understanding of biosynthetic machines, we are approaching rational design for the development of novel PKS(II)-based compounds that can be used as therapeutics.

11 Literature

- [1] D. J. Newman, G. M. Cragg, K. M. Snader, *J. Nat. Prod.* **2003**, *66*, 1022.
- [2] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2007**, *70*, 461.
- [3] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2012**, *75*, 311.
- [4] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2016**, *79*, 629.
- [5] D. J. Newman in *Antibiotic drug resistance* (Eds.: J.-L. Capelo-Martínez, G. Igrejas), Wiley, Hoboken, NJ, **2019**, pp. 311–341.
- [6] P. M. Dewick, *Medicinal natural products. A biosynthetic approach*, Wiley A John Wiley and Sons Ltd. Publication, Chichester West Sussex, UK, **2009**.
- [7] C. T. Walsh, Y. Tang, *Natural product biosynthesis. Chemical logic and enzymatic machinery*, Royal Society of Chemistry, London, UK, **2017**.
- [8] C. T. Walsh, M. A. Fischbach, *J. Am. Chem. Soc.* **2010**, *132*, 2469.
- [9] J. W. Bennett, K.-T. Chung, *Adv. Appl. Microbiol.* **2001**, 163.
- [10] A. Fleming, *Br J Exp Pathol* **1929**, *10*, 226.
- [11] G. M. Cragg, D. J. Newman, K. M. Snader, *J. Nat. Prod.* **1997**, *60*, 52.
- [12] N. Ziemert, M. Alanjary, T. Weber, *Nat. Prod. Rep* **2016**, *33*, 988.
- [13] M. H. Medema, M. A. Fischbach, *Nat. Chem. Biol.* **2015**, *11*, 639.
- [14] T. Weber, H. U. Kim, *Synth Syst Biotechnol* **2016**, *1*, 69.
- [15] A. T. Aron, E. Gentry, K. L. McPhail, L. F. Nothias, M. Nothias-Esposito, A. Bouslimani, D. Petras, J. M. Gauglitz, N. Sikora, F. Vargas et al., *ChemRxiv* **2019**, DOI 10.26434/chemrxiv.9333212.
- [16] M. Wang, A. K. Jarmusch, F. Vargas, A. A. Aksenov, J. M. Gauglitz, K. Weldon, D. Petras, R. d. Silva, R. Quinn, A. v. Melnik et al., *BioRxiv* **2019**, DOI 10.1101/591016.
- [17] J. Watrous, P. Roach, T. Alexandrov, B. S. Heath, J. Y. Yang, R. D. Kersten, M. van der Voort, K. Pogliano, H. Gross, J. M. Raaijmakers et al., *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E1743-52.
- [18] L.-F. Nothias, M. Nothias-Esposito, R. da Silva, M. Wang, I. Protsyuk, Z. Zhang, A. Sarvepalli, P. Leyssen, D. Touboul, J. Costa et al., *J. Nat. Prod.* **2018**.
- [19] K. Dührkop, M. Fleischauer, M. Ludwig, A. A. Aksenov, A. v. Melnik, M. Meusel, P. C. Dorrestein, J. Rousu, S. Böcker, *Nat. Methods* **2019**, *16*, 299.
- [20] J. C. Navarro-Muñoz, N. Selem-Mojica, M. W. Mullowney, S. Kautsar, J. H. Tryon, E. I. Parkinson, E. L.C. de Los Santos, M. Yeong, P. Cruz-Morales, S. Abubucker et al., *BioRxiv* **2018**, DOI 10.1101/445270.
- [21] T. A. Scott, J. Piel, *Nat. rev. chem.* **2019**, *3*, 404.
- [22] E. S. Sattely, M. A. Fischbach, C. T. Walsh, *Nat. Prod. Rep.* **2008**, *25*, 757.
- [23] C. Hertweck, *Angew. Chem.* **2009**, *48*, 4688.
- [24] P. Tailliez, S. Pagès, N. Ginibre, N. Boemare, *Int. J. Syst. Evol. Microbiol* **2006**, *56*, 2805.
- [25] R. A. R. Machado, D. Wüthrich, P. Kuhnert, C. C. M. Arce, L. Thönen, C. Ruiz, X. Zhang, C. A. M. Robert, J. Karimi, S. Kamali et al., *Int. J. Syst. Evol. Microbiol* **2018**, *68*, 2664.
- [26] D. I. Chan, H. J. Vogel, *Biochem. J.* **2010**, *430*, 1.
- [27] A. C. Mercer, M. D. Burkart, *Nat. Prod. Rep* **2007**, *24*, 750.
- [28] D. A. Hopwood, *Chem. Rev.* **1997**, *97*, 2465.
- [29] J. Beld, D. J. Lee, M. D. Burkart, *Mol Biosyst.* **2015**, *11*, 38.
- [30] R. Nofiani, B. Philmus, Y. Nindita, T. Mahmud, *Med. Chem. Commun.* **2019**, *49*, 691.
- [31] U. Rix, C. Fischer, L. L. Remsing, J. Rohr, *Nat. Prod. Rep.* **2002**, *19*, 542.
- [32] J. E. Cronan, J. Thomas, *Methods Enzymol.* **2009**, *459*, 395.
- [33] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380.
- [34] B. J. Rawlings, *Nat. Prod. Rep.* **2001**, *18*, 190.
- [35] C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, *Nat. Prod. Rep* **2007**, *24*, 162.
- [36] Y. Katsuyama, S. Horinouchi in *Comprehensive natural products II. Chemistry and biology* (Ed.: L. Mander), Elsevier, Amsterdam, NL, **2010**, pp. 147–170.
- [37] D. E. Cane, C. T. Walsh, *Chem. Biol.* **1999**, *6*, R319-25.

- [38] C. T. Walsh, *Science* **2004**, *303*, 1805.
- [39] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468.
- [40] F. Kopp, M. A. Marahiel, *Nat. Prod. Rep* **2007**, *24*, 735.
- [41] R. M. Kohli, C. T. Walsh, *Chem. Commun.* **2003**, *7*, 297.
- [42] L. Katz, *Methods Enzymol.* **2009**, *459*, 113.
- [43] C. Bisang, P. F. Long, J. Corte´s, J. Westcott, J. Crosby, A.-L. Matharu, R. J. Cox, T. J. Simpson, J. Staunton, P. F. Leadlay, *Nature* **1999**, *401*, 502.
- [44] C. L. Bayly, V. G. Yadav, *Molecules* **2017**, *22*, 335.
- [45] T. Schwecke, J. F. Aparicio, I. Molnár, A. König, L. E. Khaw, S. F. Haydock, M. Oliynyk, P. Caffrey, J. Cortés, J. B. Lester, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7839.
- [46] B. S. Moore, C. Hertweck, *Nat. Prod. Rep.* **2002**, *19*, 70.
- [47] H. G. Floss, *Nat. Prod. Rep.* **1997**, *14*, 433.
- [48] D. A. Hopwood, *Methods Enzymol.* **2009**, *459*, xvii.
- [49] V. Simunovic, J. Zapp, S. Rachid, D. Krug, P. Meiser, R. Müller, *ChemBioChem* **2006**, *7*, 1206.
- [50] M. A. Skiba, A. P. Sikkema, W. D. Fiers, W. H. Gerwick, D. H. Sherman, C. C. Aldrich, J. L. Smith, *ACS Chem. Biol.* **2016**, *11*, 3319.
- [51] Y. A. Chan, A. M. Podevels, B. M. Kevany, M. G. Thomas, *Nat. Prod. Rep.* **2009**, *26*, 90.
- [52] S. Kosol, M. Jenner, J. R. Lewandowski, G. L. Challis, *Nat. Prod. Rep* **2018**, *35*, 1097.
- [53] E. J. N. Helfrich, J. Piel, *Nat. Prod. Rep* **2016**, *33*, 231.
- [54] R. J. Cox, *Org. Biomol. Chem.* **2007**, *5*, 2010.
- [55] W. Liu, S. D. Christenson, S. Standage, B. Shen, *Science* **2002**, *297*, 1170.
- [56] J. Ahlert, E. Shepard, N. Lomovskaya, E. Zazopoulos, A. Staffa, B. O. Bachmann, K. Huang, L. Fonstein, A. Czisny, R. E. Whitwam et al., *Science* **2002**, *297*, 1173.
- [57] W. Zhang, Y. Tang **2009**, *459*, 367.
- [58] P. Kumar, A. T. Koppisch, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **2003**, *125*, 14307.
- [59] Y. Tang, S.-C. Tsai, C. Khosla, *J. Am. Chem. Soc.* **2003**, *125*, 12708.
- [60] Xu Dong, C. D. Bailey, Christopher Williams, John Crosby, T. J. Simpson, C. L. Willis, M. P. Crump, *Chemical Science* **2016**, *7*, 1779.
- [61] A. T. Keatinge-Clay, D. A. Maltby, K. F. Medzihradzky, C. Khosla, R. M. Stroud, *Nat. Struct. Mol. Biol.* **2004**, *11*, 888.
- [62] R. McDaniel, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, *Nature* **1995**, *375*, 549.
- [63] T. P. Korman, J. A. Hill, T. N. Vu, S.-C. Tsai, *Biochemistry* **2004**, *43*, 14529.
- [64] T. P. Korman, Y.-H. Tan, J. Wong, R. Luo, S.-C. Tsai, *Biochemistry* **2008**, *47*, 1837.
- [65] A. Chen, R. N. Re, M. D. Burkart, *Nat. Prod. Rep* **2018**, *35*, 1029.
- [66] S. Okamoto, T. Taguchi, K. Ochi, K. Ichinose, *Chem. Biol.* **2009**, *16*, 226.
- [67] M. B. Austin, J. P. Noel, *Nat. Prod. Rep.* **2003**, *20*, 79.
- [68] J. Schroder, G. Schroder, *J. Biosci.* **1990**, *45*, 1.
- [69] G. Gago, L. Diacovich, A. Arabolaza, S.-C. Tsai, H. Gramajo, *FEMS Microbiol. Rev.* **2011**, *35*, 475.
- [70] S. W. White, J. Zheng, Y.-M. Zhang, Rock, *Annu. Rev. Biochem.* **2005**, *74*, 791.
- [71] C. O. Rock, J. E. Cronan, *Biochim. Biophys. Acta* **1996**, *1302*, 1.
- [72] R. Schneider, B. Brors, M. Massow, H. Weiss, *FEBS Lett.* **1997**, *407*, 249.
- [73] H. Marrakchi, M.-A. Lanéelle, M. Daffé, *Chem. Biol.* **2014**, *21*, 67.
- [74] O. Tehlivets, K. Scheuringer, S. D. Kohlwein, *Biochim. Biophys. Acta* **2007**, *1771*, 255.
- [75] M. Grininger, *Curr. Opin. Struct. Biol.* **2014**, *25*, 49.
- [76] D. A. Herbst, C. A. Townsend, T. Maier, *Nat. Prod. Rep* **2018**, *35*, 1046.
- [77] J. D. Nanson, Z. Himiari, C. M. D. Swarbrick, J. K. Forwood, *Sci. Rep.* **2015**, *5*, 14797.
- [78] J. D. Nanson, J. K. Forwood, *PLoS One* **2015**, *10*, e0141543.
- [79] M. S. Kimber, F. Martin, Y. Lu, S. Houston, M. Vedadi, A. Dharamsi, K. M. Fiebig, M. Schmid, C. O. Rock, *Int. J. STD AIDS* **2004**, *279*, 52593.

- [80] J. B. Parsons, C. O. Rock, *Prog Lipid Res.* **2013**, *52*, 249.
- [81] Y.-M. Zhang, C. O. Rock, *Nat. Rev. Microbiol.* **2008**, *6*, 222.
- [82] A. H. Delcour, *Biochim. Biophys. Acta* **2009**, *1794*, 808.
- [83] Z. Zhang, H.-X. Pan, G.-L. Tang, *F1000Res.* **2017**, *6*.
- [84] C. Davies, R. J. Heath, S. W. White, C. O. Rock, *Structure* **2000**, *8*, 185.
- [85] K.-H. Choi, R. J. Heath, C. O. Rock, *J. Bacteriol.* **2000**, *182*, 365.
- [86] G. F. Wang, T. Kuriki, K. L. Roy, T. Kaneda, *Eur. J. Biochem.* **1993**, *213*, 1091.
- [87] K. Willecke, A. B. Pardee, *Int. J. STD AIDS* **1971**, *246*, 5264.
- [88] Y. Tang, T. S. Lee, S. Kobayashi, C. Khosla, *Biochemistry* **2003**, *42*, 6588.
- [89] A. J. Waldman, E. P. Balskus, *Organic letters* **2014**, *16*, 640.
- [90] W. Zhang, B. D. Ames, S.-C. Tsai, Y. Tang, *Appl. Environ. Microbiol.* **2006**, *72*, 2573.
- [91] A. D. McCarthy, D.G. Hardie, *Trends Biochem. Sci.* **1984**, *9*, 60.
- [92] A. T. Keatinge-Clay, *Angew. Chem. Int. Ed.* **2017**, *56*, 4658-4660.
- [93] J. Zheng, C. D. Fage, B. Demeler, D. W. Hoffman, A. T. Keatinge-Clay, *ACS Chem. Biol.* **2013**, *8*, 1263.
- [94] A. L. Edwards, T. Matsui, T. M. Weiss, C. Khosla, *J. Mol. Biol.* **2014**, *426*, 2229.
- [95] J. R. Whicher, S. Dutta, D. A. Hansen, W. A. Hale, J. A. Chemler, A. M. Dosey, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith et al., *Nature* **2014**, *510*, 560.
- [96] S. Dutta, J. R. Whicher, D. A. Hansen, W. A. Hale, J. A. Chemler, G. R. Congdon, A. R. H. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith et al., *Nature* **2014**, *510*, 512.
- [97] J. Davison, J. Dorival, H. Rabeharindranto, H. Mazon, B. Chagot, A. Gruez, K. J. Weissman, *Chem. Sci.* **2014**, *5*, 3081.
- [98] D. A. Herbst, R. P. Jakob, F. Zähringer, T. Maier, *Nature* **2016**, *531*, 533.
- [99] L. Ciccarelli, S. R. Connell, M. Enderle, D. J. Mills, J. Vonck, M. Grninger, *Structure* **2013**, *21*, 1251.
- [100] I. B. Lomakin, Y. Xiong, T. A. Steitz, *Cell* **2007**, *129*, 319.
- [101] P. Johansson, B. Mulinacci, C. Koestler, R. Vollrath, D. Oesterhelt, M. Grninger, *Structure* **2009**, *17*, 1063.
- [102] T. Maier, M. Leibundgut, N. Ban, *Science* **2008**, *321*, 1315.
- [103] T. Maier, S. Jenni, N. Ban, *Science* **2006**, *311*, 1258.
- [104] A. Rittner, M. Grninger, *ChemBiochem* **2014**, *15*, 2489.
- [105] A. Witkowski, A. Joshi, S. Smith, *Biochemistry* **1996**, *35*, 10569.
- [106] M. Fischer, M. Grninger, *Beilstein J. Org. Chem.* **2017**, *13*, 1204.
- [107] J. Gajewski, R. Pavlovic, M. Fischer, E. Boles, M. Grninger, *Nat. Commun* **2017**, *8*, 14650.
- [108] M. Klaus, M. Grninger, *Nat. Prod. Rep* **2018**, *35*, 1070.
- [109] A. Rittner, K. S. Paithankar, D. Drexler, A. Himmler, M. Grninger, *Protein Sci.* **2018**, *28*, 414.
- [110] U. Yadav, R. Arya, S. Kundu, M. Sundd, *Biochemistry* **2018**, *57*, 3690.
- [111] J. Beld, H. Cang, M. D. Burkart, *Angew. Chem.* **2014**, *53*, 14456.
- [112] J. E. Cronan, *Biochem. J.* **2014**, *460*, 157.
- [113] C. S. Heil, S. S. Wehrheim, K. S. Paithankar, M. Grninger, *ChemBiochem* **2019**, *20*, 2298.
- [114] A.-L. Matharu, R. J. Cox, J. Crosby, K. J. Byrom, T. J. Simpson, *Chem. Biol.* **1998**, *5*, 699.
- [115] G. Shakya, H. Rivera, D. J. Lee, M. J. Jaremko, J. J. La Clair, D. T. Fox, R. W. Haushalter, A. J. Schaub, J. Bruegger, J. F. Barajas et al., *J. Am. Chem. Soc.* **2014**, *136*, 16792.
- [116] C. Hertweck, *Angew. Chem* **2009**, *121*, 4782.
- [117] G. Castaldo, J. Zucko, S. Heidelberger, D. Vujaklija, D. Hranueli, J. Cullum, P. Wattana-Amorn, M. P. Crump, J. Crosby, P. F. Long, *Chem. Biol.* **2008**, *15*, 1156.
- [118] G. Caldara-Festin, D. R. Jackson, J. F. Barajas, T. R. Valentic, A. B. Patel, S. Aguilar, M. Nguyen, M. Vo, A. Khanna, E. Sasaki et al., *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E6844-51.

- [119] L. Zhang, J. Xiao, J. Xu, T. Fu, Z. Cao, L. Zhu, H.-Z. Chen, X. Shen, H. Jiang, L. Zhang, *Cell Research* **2016**, *26*, 1330.
- [120] J. T. Mindrebo, A. Patel, W. E. Kim, T. D. Davis, A. Chen, T. G. Bartholow, J. J. La Clair, J. A. McCammon, J. P. Noel, M. D. Burkart, *BioRxiv* **2019**, DOI 10.1101/644518.
- [121] L. Zhang, J. Xiao, J. Xu, T. Fu, Z. Cao, L. Zhu, H.-Z. Chen, X. Shen, H. Jiang, L. Zhang, *Cell Research* **2016**, *26*, 1330.
- [122] C. Nguyen, R. W. Haushalter, D. J. Lee, P. R. L. Markwick, J. Bruegger, G. Caldara-Festin, K. Finzel, D. R. Jackson, F. Ishikawa, B. O'Dowd et al., *Nature* **2014**, *505*, 427.
- [123] B. Kunze, G. Höfle, H. Reichenbach, *J. Antibiot.* **1987**, *40*, 258.
- [124] D. Pistorius, Y. Li, A. Sandmann, R. Müller, *Mol Biosyst.* **2011**, *7*, 3308.
- [125] F. Panter, D. Krug, S. Baumann, R. Müller, *Chem. Sci.* **2018**, *9*, 4898.
- [126] E. E. Herbert, H. Goodrich-Blair, *Nat. Rev. Microbiol* **2007**, *5*, 634.
- [127] Y.-M. Shi, H. B. Bode, *Nat. Prod. Rep.* **2018**, *35*, 309.
- [128] N. J. Tobias, H. Wolff, B. Djahanschiri, F. Grundmann, M. Kronenwerth, Y.-M. Shi, S. Simonyi, P. Grün, D. Shapiro-Ilan, S. J. Pidot et al., *Nat. Microbiol.* **2017**, *2*, 1676.
- [129] N. J. Tobias, Y.-M. Shi, H. B. Bode, *Trends Microbiol.* **2018**, *26*, 833.
- [130] A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBiochem* **2007**, *8*, 1721.
- [131] A. K. Heinrich, *Regulation of natural product production of Photorhabdus luminescens via both pathway-specific and global regulators*. Doctoral Dissertation, Johann Wolfgang Goethe-Universität Frankfurt am Main, **2017**.
- [132] S. A. Joyce, A. O. Brachmann, I. Glazer, L. Lango, G. Schwär, D. J. Clarke, H. B. Bode, *Angew. Chem* **2008**, *120*, 1968.
- [133] M. Kronenwerth, A. O. Brachmann, M. Kaiser, H. B. Bode, *ChemBiochem* **2014**, *15*, 2689.
- [134] H. B. Park, P. Sampathkumar, C. E. Perez, J. H. Lee, J. Tran, J. B. Bonanno, E. A. Hallem, S. C. Almo, J. M. Crawford, *J. Biol. Chem.* **2017**, *292*, 6680.
- [135] J. Peppers, A. S. Paller, T. Maeda-Chubachi, S. Wu, K. Robbins, K. Gallagher, J. E. Kraus, *J. Am. Acad. Dermatol.* **2019**, *80*, 89-98.e3.
- [136] L. Zhao, X. Chen, L. Cai, C. Zhang, Q. Wang, S. Jing, G. Chen, J. Li, J. Zhang, Y. Fang, *J. Clin. Pharm. Ther.* **2014**, *39*, 418.
- [137] T. Mori, T. Awakawa, K. Shimomura, Y. Saito, D. Yang, H. Morita, I. Abe, *Cell Chem. Biol.* **2016**, *23*, 1468.
- [138] P. Cimermancic, M. H. Medema, J. Claesen, K. Kurita, L. C. Wieland Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy et al., *Cell* **2014**, *158*, 412.
- [139] X.-Q. Cao, J.-Y. Wang, L. Zhou, B. Chen, Y. Jin, Y.-W. He, *Mol. Microbiol.* **2018**.
- [140] T. A. Schöner, S. W. Fuchs, B. Reinhold-Hurek, H. B. Bode, *PLoS One* **2014**, *9*, e90922.
- [141] T. A. Schöner, S. W. Fuchs, C. Schönau, H. B. Bode, *Microb. Biotechnol.* **2014**, *7*, 232.
- [142] H. Reichenbach, W. Kohl, A. Böttger-Vetter, H. Achenbach, *Arch. Microbiol* **1980**, *126*, 291.
- [143] J. Paniagua-Michel, J. Olmos-Soto, M. A. Ruiz, *Methods Mol. Biol.* **2012**, *892*, 1.
- [144] T. A. Schöner, S. Gassel, A. Osawa, N. J. Tobias, Y. Okuno, Y. Sakakibara, K. Shindo, G. Sandmann, H. B. Bode, *ChemBiochem* **2016**, *17*, 247.
- [145] P. Cimermancic, M. H. Medema, J. Claesen, K. Kurita, L. C. Wieland Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy et al., *Cell* **2014**, *158*, 412.
- [146] J. O. Lampen, *Am. J. Clin. Pathol.* **1969**, *52*, 138.
- [147] E. M. Woerly, J. Roy, M. D. Burke, *Nat. Chem.* **2014**, *6*, 484.
- [148] T. A. Schöner, S. W. Fuchs, B. Reinhold-Hurek, H. B. Bode, *PLoS One* **2014**, *9*, e90922.

- [149] T. A. R. Schöner, *Biosynthesis, structure elucidation and distribution of bacterial aryl polyene pigments*. Doctoral Dissertation, Johann Wolfgang Goethe-Universität Frankfurt am Main, **2015**.
- [150] E. Bode, A. O. Brachmann, C. Kegler, R. Simsek, C. Dauth, Q. Zhou, M. Kaiser, P. Klemmt, H. B. Bode, *ChemBioChem* **2015**, *16*, 1115.
- [151] S. W. Fuchs, K. A. J. Bozhuyuk, D. Kresovic, F. Grundmann, V. Dill, A. O. Brachmann, N. R. Waterfield, H. B. Bode, *Angew. Chem.* **2013**, *52*, 4108.
- [152] T. Mori, T. Awakawa, K. Shimomura, Y. Saito, D. Yang, H. Morita, I. Abe, *Cell Chem. Biol.* **2016**, *23*, 1468.
- [153] Q. Zhou, *Identification of selected secondary metabolites from Xenorhabdus and investigation on the biosynthesis of anthraquinones from Photorhabdus*. Doctoral Dissertation, Johann Wolfgang Goethe-Universität Frankfurt am Main, **2016**.
- [154] C. J. Arthur, A. E. Szafranska, J. Long, J. Mills, R. J. Cox, S. C. Findlow, T. J. Simpson, M. P. Crump, J. Crosby, *Chem. Biol.* **2006**, *13*, 587.
- [155] N. J. Tobias, A. K. Heinrich, H. Eresmann, P. R. Wright, N. Neubacher, R. Backofen, H. B. Bode, *Environ. Microbiol.* **2017**, *19*, 119.
- [156] H. B. Bode, *Angew. Chem.* **2019**, DOI 10.1002/ange.201910563.
- [157] C. A. Smith, G. O. Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan, G. Siuzdak, *Ther. Drug Monit.* **2005**, *27*, 747.
- [158] C. Guijas, J. R. Montenegro-Burke, X. Domingo-Almenara, A. Palermo, B. Warth, G. Hermann, G. Koellensperger, T. Huan, W. Uritboonthai, A. E. Aisporna et al., *Anal. Chem.* **2018**, *90*, 3156.
- [159] T. Cajka, O. Fiehn, *Trends Analyt. Chem.* **2014**, *61*, 192.
- [160] X. Wang, Z.-F. Yuan, J. Fan, K. R. Karch, L. E. Ball, J. M. Denu, B. A. Garcia, *Mol. Cell. Proteom.* **2016**, *15*, 2462.
- [161] S. Thoma, M. Schobert, *FEMS Microbiol. Lett.* **2009**, *294*, 127.
- [162] E. Bode, *Structure, function and biosynthesis of natural products from Xenorhabdus doucetiae*. Doctoral Dissertation, Johann Wolfgang Goethe-Universität Frankfurt am Main, **2017**.
- [163] C. Fu, W. P. Donovan, O. Shikapwashya-Hasser, X. Ye, R. H. Cole, *PLoS One* **2014**, *9*, e115318.
- [164] A. A. Entezami, B. J. Venables, K. E. Daugherty, *J. Chrom. A* **1987**, *387*, 323.
- [165] W. L. Stephens, M. P. Starr, *J. Bacteriol.* **1963**, *86*, 1070.
- [166] W. C. Chun, *PHI* **2000**.
- [167] T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Brucoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben et al., *Nucleic Acids Res.* **2015**, *43*, W237-43.
- [168] K. S. Madden, F. A. Mosa, A. Whiting, *Org. Biomol. Chem.* **2014**, *12*, 7877.
- [169] C. Chalut, *Tuberculosis* **2016**, *100*, 32.
- [170] S. Okuda, H. Tokuda, *Annu. Rev. Microbiol.* **2011**, *65*, 239.
- [171] S. Liaaen-Jensen, A. G. Andrewes in *Methods in Microbiology*, Vol. 18 (Ed.: G. Gottschalk), Academic Press, London, UK, **1985**, pp. 235–255.
- [172] P. Domenech, M. B. Reed, C. E. Barry, *Infect. Immun.* **2005**, *73*, 3492.
- [173] R. E. Hancock, H. Nikaido, *J. Bacteriol.* **1978**, *136*, 381.
- [174] L. A. Clifton, M. W. A. Skoda, E. L. Daulton, A. V. Hughes, A. P. Le Brun, J. H. Lakey, S. A. Holt, *J. R. Soc. Interface* **2013**, *10*, 20130810.
- [175] K. A. Z. Berry, R. C. Murphy, *J. Am. Soc. Mass Spectrom. Chem.* **2004**, *15*, 1499.
- [176] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chemistry* **2012**, *18*, 2342.
- [177] J. Rinkel, J. S. Dickschat, *Beilstein J. Org. Chem.* **2015**, *11*, 2493.
- [178] J. D. Janse, P. H. Smits, *Lett. Appl. Microbiol.* **1990**, *10*, 131.
- [179] A. Tschumi, C. Nai, Y. Auchli, P. Hunziker, P. Gehrig, P. Keller, T. Grau, P. Sander, *J. Biol. Chem.* **2009**, *284*, 27146.
- [180] H. P. McNulty, J. Byun, S. F. Lockwood, R. F. Jacob, R. P. Mason, *Biochim. Biophys. Acta* **2007**, *1768*, 167.
- [181] A. Vershinin, *BioFactors* **1999**, *10*, 99.

- [182] R. W. Holz, *Ann. N. Y. Acad. Sci.* **1974**, 235, 469.
- [183] G. K. Auer, D. B. Weibel, *Biochemistry* **2017**, 56, 3710.
- [184] R. M. Epand, R. F. Epand, *Biochim. Biophys. Acta* **2009**, 1788, 289.
- [185] S. Ouardien, S. Brul, S. A. J. Zaat, *Front. Cell Dev. Biol.* **2016**, 4, 111.
- [186] J. C. Castillo, S. E. Reynolds, I. Eleftherianos, *Trends Parasitol.* **2011**, 27, 537.
- [187] A. D. Lúcio, C. C. Vequi-Suplicy, R. M. Fernandez, M. T. Lamy, *J. Fluoresc.* **2010**, 20, 473.
- [188] D. D. Barnes, M. L. E. Lundahl, E. C. Lavelle, E. M. Scanlan, *ACS Chem. Biol.* **2017**, 12, 1969.
- [189] M. Jenner, *Using Mass Spectrometry for Biochemical Studies on Enzymatic Domains from Polyketide Synthases*, Springer International Publishing, Cham, **2016**.
- [190] D. Leduc, A. Battesti, E. Bouveret, *J. Bacteriol.* **2007**, 189, 7112.
- [191] Y. Chen, E. E. Kelly, R. P. Masluk, C. L. Nelson, D. C. Cantu, P. J. Reilly, *Protein Sci.* **2011**, 20, 1659.
- [192] J. A. Latham, D. Chen, K. N. Allen, D. Dunaway-Mariano, *Biochemistry* **2014**, 53, 4775.
- [193] P. Dall'Aglio, C. Arthur, M. P. Crump, J. Crosby, A. T. Hadfield, *Biochemistry* **2011**, 50, 5704.
- [194] N. N. K. Kreamer, R. Chopra, R. E. Caughlan, D. Fabbro, E. Fang, P. Gee, I. Hunt, M. Li, B. C. Leon, L. Muller et al., *Sci. Rep.* **2018**, 8, 14124.
- [195] G. Butland, J. M. Peregrín-Alvarez, J. Li, W. Yang, X. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan et al., *Nature* **2005**, 433, 531.
- [196] T. Sztain, T. G. Bartholow, J. A. McCammon, M. D. Burkart, *Biochemistry* **2019**, 58, 3557.
- [197] C. Schmidt, H. Urlaub, *Curr. Opin. Struct. Biol.* **2017**, 46, 157.
- [198] A. Leitner, M. Faini, F. Stengel, R. Aebersold, *Trends Biochem. Sci.* **2016**, 41, 20.
- [199] J. Rappsilber, *J. Struct. Biol.* **2011**, 173, 530.
- [200] A. Sinz, *Mass Spectrom. Rev* **2006**, 25, 663.
- [201] J. E. Vance, *Biochemistry of Lipids, Lipoproteins and Membranes*, Elsevier Science, s.l., **2008**.
- [202] S. Cantaloube, R. Veyron-Churlet, N. Haddache, M. Daffé, D. Zerbib, *PLoS One* **2011**, 6, e29564.
- [203] C. Lefebvre, R. Boulon, M. Ducoux, S. Gavalda, F. Laval, S. Jamet, N. Eynard, A. Lemassu, K. Cam, M.-P. Bousquet et al., *Sci. Rep.* **2018**, 8, 6034.
- [204] J. A. Baur, D. A. Sinclair, *Nat. Rev. Drug Discov* **2006**, 5, 493.
- [205] Z. Zhang, R. Zhou, J. M. Sauder, P. J. Tonge, S. K. Burley, S. Swaminathan, *J. Mol. Biol.* **2010**, 406, 313.
- [206] H. Pan, S.-C. Tsai, E. S. Meadows, L. J.W. Miercke, A. T. Keatinge-Clay, J. O'Connell, C. Khosla, R. M. Stroud, *Structure* **2002**, 10, 1559.
- [207] C. Notredame, D. G. Higgins, J. Heringa, *J. Mol. Biol.* **2000**, 302, 205.
- [208] X. Robert, P. Gouet, *Nucleic Acids Res.* **2014**, 42, W320-4.
- [209] K. S. Gajiwala, S. Margosiak, J. Lu, J. Cortez, Y. Su, Z. Nie, K. Appelt, *FEBS Lett.* **2009**, 583, 2939.
- [210] K. Madhusudan Makwana, R. Mahalakshmi, *Protein Sci.* **2015**, 24, 1920.
- [211] S.-C. Li, N. K. Goto, K. A. Williams, C. M. Deber, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 6676.
- [212] J. Crosby, M. P. Crump, *Nat. Prod. Rep* **2012**, 29, 1111.
- [213] A. Misra, S. K. Sharma, N. Surolia, A. Surolia, *Chem. Biol.* **2007**, 14, 775.
- [214] C. J. Arthur, A. Szafranska, S. E. Evans, S. C. Findlow, S. G. Burston, P. Owen, I. Clark-Lewis, T. J. Simpson, J. Crosby, M. P. Crump, *Biochemistry* **2005**, 44, 15414.
- [215] A. Misra, N. Surolia, A. Surolia, *Mol Biosyst.* **2009**, 5, 651.
- [216] A. Misra, N. Surolia, A. Surolia, *IUBMB life* **2009**, 61, 853.
- [217] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep* **2014**, 31, 61.
- [218] N. R. D. Lay, J. E. Cronan, *Mol. Microbiol.* **2006**, 61, 232.

- [219] R. Farmer, C. M. Thomas, P. J. Winn, *PLoS ONE* **2019**, *14*, e0219435.
- [220] H. C. Wong, G. Liu, Y.-M. Zhang, C. O. Rock, J. Zheng, *Int. J. STD AIDS* **2002**, *277*, 15874.
- [221] L. K. Caesar, N. B. Cech, *Nat. Prod. Rep* **2019**, *36*, 869.
- [222] C. O. Rock, J. E. Cronan, *Int. J. STD AIDS* **1979**, *254*, 7116.
- [223] T. K. Ray, J. E. Cronan, JR, *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 4374.
- [224] D. J. Campopiano, *Chem. Biol.* **2014**, *21*, 1257.
- [225] J. Beld, K. Finzel, M. D. Burkart, *Chem. Biol.* **2014**, *21*, 1293.

12 Attachments

12.1 Publication: An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments

Erklärung zu den Autorenanteilen

Status: **printed**

J. Am. Chem. Soc. 2019 141 (42), 16615-16623, DOI: 10.1021/jacs.8b10776

Beteiligte Autoren: Gina L. C. Grammbitter (GLG), Maximilian Schmalhofer (MS), Kudratullah Karimi (KK), Yi-Ming Shi (YS), Tim Aaron Schöner (TAS), Nicholas J. Tobias (NJT), Nina Morgner (NM), Michael Groll (MG), Helge B. Bode (HBB)

Was hat der/die Promovierende bzw. was haben die Co-Autoren/Autorinnen beigetragen?

(1) zu Entwicklung und Planung

GLG: 65%

HBB: 25%

MG: 10%

(2) Durchführung der einzelnen Untersuchungen und Experimente

In vitro-Experimente zu ArcT und ArcV: TAS (5%)

Proteinisolation Kristallisation und Kristallisierungsexperimente ApeO:C, Apel:P: MS (20%)

nESI-MS Messung von APE-Proteinen KK (10%)

Klonierung und genetische Modifikation *Xenorhabdus*, Reinigung APE-Methylester, *In vitro*-Experimente zu *Xenorhabdus* APE-Enzymen, Fütterungsexperimente, ESI- und MALDI-MS-Analysen, Synthese und Reinigung 4HBA-CoA (GLG 65%)

(3) Erstellung der Datensammlung und Abbildungen

Phylogenie APE-Gencluster *Xenorhabdus/Photorhabdus*: NJT (1%)

Experimente zu ArcT und ArcV: GLG (1%) + TAS (3%)

nESI-Messungen APE-Enzyme: KK (10%)

Proteinstrukturen ApeO:C, Apel:P: MS, MG (20%)

Darstellung Biosynthesegencluster, Biosyntheseschema, Reinigung APE-Methylester, *In vitro*-Experimente zu *Xenorhabdus* APE-Enzymen, Fütterungsexperimente, ESI- und MALDI-MS-Analysen: GLG (60%)

NMR-Strukturaufklärung APE-Methylester: YS (5%)

(4) Analyse und Interpretation der Daten

Experimente zu ArcT und ArcV: TAS (5%)

nESI-Messungen APE-Enzyme: KK (5%) +NM (5%)

Proteinstrukturen ApeO:C, Apel:P: MS, MG (20%)

In vitro-Experimente zu *Xenorhabdus* APE-Enzymen, Fütterungsexperimente, ESI- und MALDI-MS-Analysen: GLG (60%)

NMR-Strukturaufklärung APE-Methylester: YS (5%)

(5) Verfassung des Manuskripts

GLG: 65%

HBB: 25%

MG, MS: 10%

date/place: _____

An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments

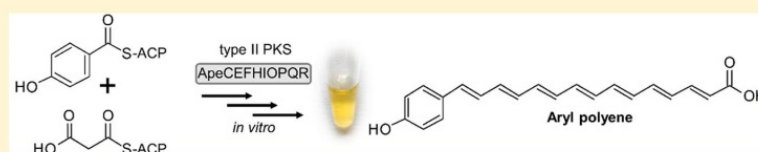
Gina L. C. Grammbitter,[†] Maximilian Schmalhofer,[‡] Kudratullah Karimi,[§] Yi-Ming Shi,[†]
 Tim A. Schöner,[†] Nicholas J. Tobias,[†] Nina Morgner,[§] Michael Groll,[‡] and Helge B. Bode^{*,†,§}

[†]Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe-Universität Frankfurt am Main and Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universität Frankfurt, Max-von-Laue-Straße 9 and 15, 60438 Frankfurt am Main, Germany

[‡]Center for Integrated Protein Science at the Department Chemie, Lehrstuhl für Biochemie, Technische Universität München, Lichtenbergstraße 4, 85748 Garching, Germany

[§]Institut für Physikalische und Theoretische Chemie, Goethe-Universität Frankfurt, Max-von-Laue-Straße 7, 60438 Frankfurt am Main, Germany

Supporting Information



ABSTRACT: Aryl polyene (APE) pigments are a widely distributed class of bacterial polyketides. So far, little is known about the biosynthesis of these compounds, which are produced by a novel type II polyketide synthase (PKS). We have identified all enzymes involved in APE biosynthesis and determined their peculiar functions. The biosynthesis was reconstituted *in vitro*, and ACP-bound intermediates were assigned for each reaction step by HPLC-MS. Native mass spectrometry experiments identified four stable complexes: the acyl-carrier proteins ApeE and ApeF bound to the thioesterase ApeK, the dehydratases ApeI and ApeP, and the ketosynthase ApeO in complex with its chain-length factor ApeC. X-ray structures of the heterodimeric ApeO:ApeC and ApeI:ApeP complexes depict striking protein–protein interactions. Altogether, our study elucidated mechanistic aspects of APE biosynthesis that unifies elements of type II fatty acid and PKS systems, but in addition includes novel enzyme complexes.

INTRODUCTION

Type II polyketide synthases (PKS II) are monofunctional enzymes forming multienzyme complexes involved in the formation of a wide range of aromatic and phenolic compounds, some of which are of great clinical importance.^{1,2} Usually such natural products are produced by Gram-positive bacteria such as *Streptomyces*, exemplified by the well-known antibiotic tetracycline.¹ Recently, a PKS II system was found in the anaerobic endospore-forming Gram-positive firmicute *Clostridium*, producing pentacyclic polyphenolic compounds named clostrubins.³ PKS II systems in Gram-negative bacteria are rare, with aurachins^{4,5} and the pyxidicyclines⁶ from myxobacteria as well as the anthraquinones produced by the γ -proteobacterium *Photorhabdus luminescens*⁷ being so far the only examples.

For aromatic PKS II-derived compounds, a so-called “minimal PKS” composed of an acyl-carrier protein (ACP), a ketosynthase:chain-length factor (KS:CLF) complex, and a malonyl-CoA:ACP transacylase (MCAT) act together with a cyclase/aromatase.^{1,2} The corresponding genes encoding the KS:CLF and ACP are often colocalized in one operon.

Polyene moieties are usually derived from type I PKS systems composed of modules that might also act iteratively as demonstrated for myxochromide.⁸ The most prominent examples are polyene macrolides such as amphotericin B and nystatin from *Streptomyces* sp., which are used as antifungal drugs.⁹ Interestingly, a PKS I system from a basidiomycete was found to produce antilarval polyene compounds using a single PKS module.¹⁰

Recently, the biosynthesis of the aliphatic polyene compound ishigamide from *Streptomyces* was shown to be derived from a PKS II system.¹¹ Interestingly, by means of informatic analyses, the biosynthetic gene cluster (BGC) of aryl polyenes (APEs) produced by a type II PKS system is one of the most common BGC among Gram-negative bacteria.^{12,13} APE-BGCs are also widespread among γ -proteobacteria and in the genus *Xenorhabdus* (Figure S1) that lives in symbiosis with *Steinernema* nematodes.^{14–16} *Xenorhabdus* APE-BGC belongs to subfamily 1 of the three APE BGC subfamilies described previously.¹² In that work closely related *ape* gene clusters from

Received: October 6, 2018

Published: March 25, 2019

Vibrio fischeri and *Escherichia coli* CFT073 were heterologously expressed and the corresponding APE structures were reported.¹² Other APEs are xanthomonadins and flexirubins produced by *Xanthomonas* sp.^{17–19} and *Flavobacteria*,²⁰ respectively. While there is some evidence that they act as antioxidative compounds similar to carotenoids,^{12,13,21–26} their biosynthesis remained unexplored.

Previously, the APE derivative arcuflavin (an APE esterified to a dialkylresorcinol) was described from *Azoarcus* sp. BH72, and its biosynthesis was proposed based on the identified BGC (Figure S2, Table S4).²⁴ Our present work addresses the *ape* gene cluster from *Xenorhabdus doucetiae*. We have identified one main APE product by NMR and other derivatives by mass spectrometry analysis and could confirm all steps in its biosynthesis by *in vitro* reconstitution experiments using the purified proteins. Moreover, native mass spectrometry revealed insights into this special PKS II system, while X-ray structure analysis of two heterodimeric enzyme complexes disclosed striking protein–protein interactions.

RESULTS AND DISCUSSION

The *ape* gene cluster from *X. doucetiae* encodes several enzymes proposed to be involved in APE biosynthesis: two acyl carrier proteins (ACPs; ApeE, ApeF), two ketosynthases (KSs (or KS α); ApeO, ApeR), a chain length factor (CLF (or KS β); ApeC), a ketoreductase (KR; ApeQ), and two dehydratases (DHs; ApeI, ApeP), as well as a putative 4-hydroxybenzoyl-CoA thioesterase (4HBA-TE; ApeK), an acyl–acyl carrier protein synthetase (AasS; ApeH), and a methyltransferase (MT, ApeB) (Figure 1a).

So far, no APE derivative has been described from *X. doucetiae*. Therefore, we inserted an arabinose-inducible P_{BAD} promoter²⁷ in front of *apeA* and *apeB*, leading to a yellow pigmentation of the induced mutant strain (Figure S3). Several approaches such as methanolysis of the yellow cell pellet, a positive KOH test²⁸ (Figure S4), and detailed MS analysis with feeding of 3-fluoro-4-hydroxy benzoic acid (3F4HBA, Figure S5) suggest that 4HBA is the natural starter unit and most likely is derived from the shikimate pathway.²⁹ It should be noted that the only known bacterial PKS II system that uses a benzoic acid precursor has hitherto been the enterocin-producing PKS II.^{30,31}

We purified the major APE product **1** for structure elucidation by NMR techniques (Table S5, Figures S6–S10). The natural product contains six double bonds, which has been identified in the above-mentioned APE from *V. fischeri*.¹² However, we could also detect APE variants with seven double bonds (Figure 1b). Moreover, the APE derivatives vary in their methylation pattern, suggesting that this tailoring reaction occurs late in the biosynthesis (Figures 1 and S5).

The biosynthesis of APE compounds is barely known, and the corresponding gene cluster harbors some interesting features. Genes encoding proteins involved in lipid biosynthesis that are characteristic for *ape* gene clusters (Figure 1a, turquoise genes) suggest that APEs are part of a larger membrane-bound structure. The biosynthetic genes for the actual APE biosynthesis are distributed across the whole cluster. Even the gene encoding for a CLF (ApeC) suggesting a PKS II system is not colocated with one of the KSs (ApeO, ApeR) of the cluster nor with one of the ACPs (ApeE, ApeF). Overall, the *ape* gene arrangement is more similar to a fatty acid synthase (FAS) II system than a typical PKS II. This correlation of cluster arrangement is reflected by the higher

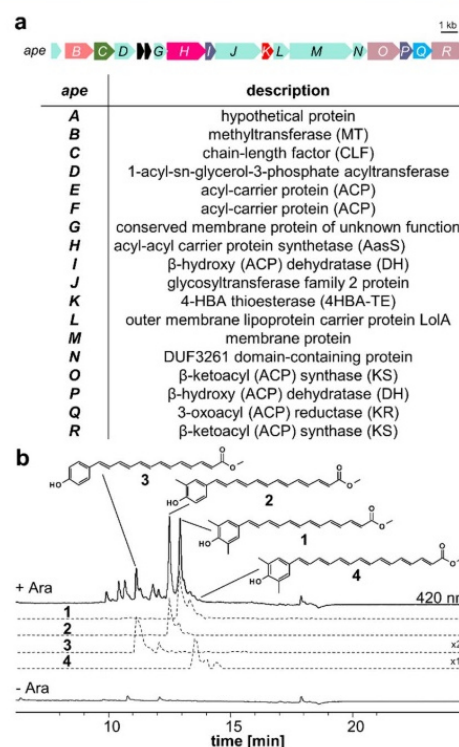


Figure 1. (a) APE biosynthetic gene cluster from *X. doucetiae*. All genes encoding proteins supposed to be involved in APE biosynthesis are color coded. Conserved genes not supposed to be involved in the biosynthesis of the APE moiety are highlighted in turquoise, and their suggested function according to BLASTp is shown in the table. (b) Arabinose induction of *X. doucetiae* DSM17909 Δ *xrdE* P_{BAD}*apeB* followed by methanolysis resulted in the identification of **1–4**. Feeding with 3F4HBA resulted in the production of **5–7** (Figure S5). For NMR data of purified **1**, see Table S5 and Figures S6–S10; for additional MS² data of **1** and **5**, see Figure S5. – Ara = without induction, + Ara = 0.2% (w/v) arabinose induction.

similarity between the linear APE compounds to fatty acids compared to typical phenolic type II polyketides. Some interesting features of the *ape* gene cluster are the duplication of genes encoding for key proteins such as ACP, KS, and DH. We were therefore interested not only in the biochemical mechanism of the biosynthesis of the APE moiety but also in the characterization of all proteins involved, trying to clarify among others the roles of a possible complementarity.

***E. coli* (DE3) BL21 PPTase AcpT Converts APE ACPs to Their *holo*-Form.** Our *in vitro* reconstitution studies started with characterizing the apo-ACPs ApeE and ApeF that should lack the phosphopantetheinyl (Ppant) prosthetic group. However, after heterologous expression and purification of the ACPs from the host *E. coli* BL21, we obtained ACPs in either *holo*- or further modified form. *E. coli* has three PPTases, which may be responsible for these post-translational modifications: AcpS (involved in fatty acid biosynthesis), EntD (involved in enterobactin biosynthesis), and AcpT. AcpT was described to activate two special ACPs,³² which apparently have been lost in *E. coli* BL21 but are found in its ancestor *E. coli* O157:H7. Both ACPs are encoded within a gene cluster, displaying 62% and 92% identity to the APE-BGCs from

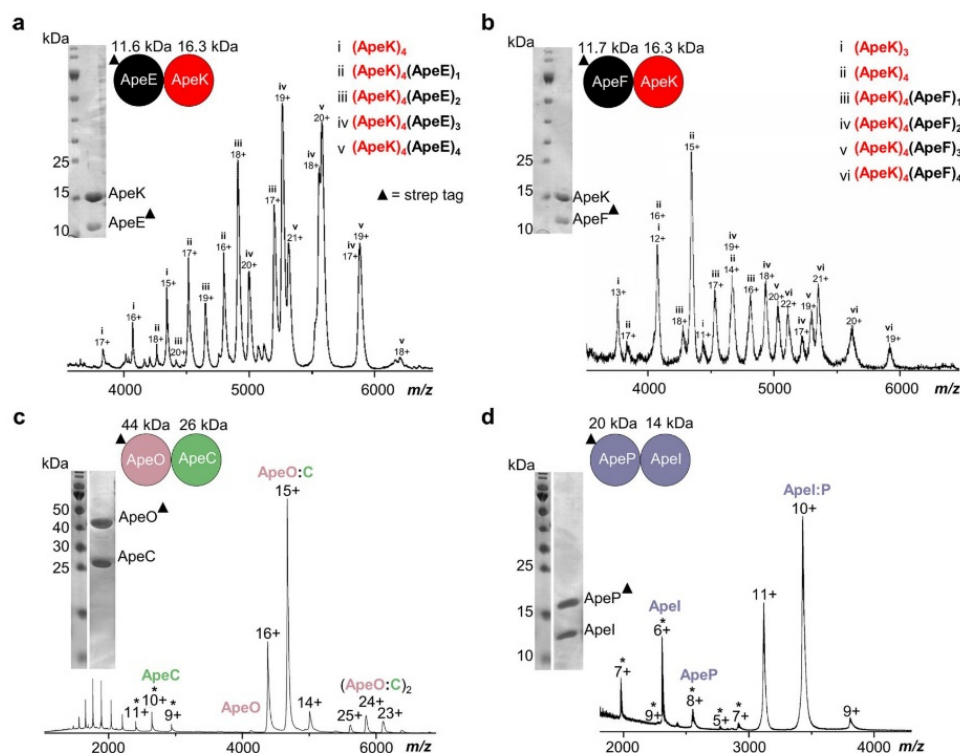


Figure 2. nESI-MS analysis of native protein complex with corresponding SDS-PAGE of ACP:TE complexes ApeE:ApeK (a) and ApeF:ApeK (b), as well as of KS:CLF complex ApeO:ApeC (c) and the DH complex ApeP:ApeI (d). All proteins were purified via Strep-Tactin affinity chromatography in a pull-down experiment and directly used for MS analysis. Strep-tagged enzymes are marked with a triangle. Corresponding monomeric signals are marked with an asterisk. The mass spectra show that ApeE:ApeK and ApeF:ApeK form heterooctameric complexes and ApeO:ApeC and ApeP:ApeI heterodimers in solution. Gas phase dissociation causes some of the proteins to monomerize. The different observed molecular species are labeled i–vi with charge states above the detected signals.

X. doucetiae and *E. coli* CFT073, respectively (ClustalW alignment, also see Table S4). In order to confirm that the orphan AcpT from *E. coli* BL21 is responsible for ApeE and ApeF modifications during heterologous production, an *acpT* deletion mutant (kindly provided by John Cronan)³² was analyzed, showing solely the production of *apo*-ApeE and *apo*-ApeF (Figure S11). Thus, the “special” ACPs from *E. coli* O157:H7 can be retrospectively assigned as ACPs, originally being encoded in an *ape* gene cluster. Interestingly, in *X. doucetiae* the APE-BGC lacks a PPTase, and therefore its activity might be complemented by the sole Sfp-type PPTase NgrA (*xdv3_80277*).

ApeH Acts as a Promiscuous AasS and Initiates the Unusual Start of APE Biosynthesis. All proteins proposed to be involved in the APE biosynthesis were purified and used for *in vitro* analyses (Figures S12, S13, and S14). APE production starts with the loading of 4HBA to the ACP. While this reaction is usually catalyzed by a CoA ligase (via adenylation and CoA thioester formation) followed by an acyltransferase (loading the activated acid to the ACP), no acyltransferase is present in the *ape* cluster. Fortunately, we could show that the ApeH directly transfers the acyl group of adenylylated 4HBA to the ACP without forming a CoA thioester intermediate. Moreover, ApeH acts on substrates with an unsaturated α,β -position, including benzoic acid derivatives with substitutions in *para*- and *meta*-positions (Figure S15;

additional ATP-PP_i exchange assay of positive substrates, see Figure S16), whereas it does not accept *ortho* benzoic acids (Table S6). From *in vivo* studies to identify additional surrogates for the production of engineered APEs it turned out that only the benzoic acid derivatives 3HBA and 4HBA, as well their fluorinated derivatives 3F4HBA and 4F3HBA were incorporated (Figure S17). The AasS enzyme activity could be confirmed for the “APE ligase”, ArcT, from *Azoarcus* sp. BH72, which was analyzed for its substrate specificity in combination with its corresponding ACP ArcV (Figure S18, Table S7). Thus, the AasS activity in APE biosynthesis might follow a general mechanism, with ApeH being a new member of the family of ligases,^{33,34} described as useful biochemical tools for controlled ACP loading.^{35,36}

Starter and Elongation ACPs Form Stable Heterooctameric Complexes with Thioesterase (TE). The two ACPs encoded within the *ape* gene cluster shared 23% sequence identity to each other. We therefore tested if one ACP serves as a starter (loaded with the 4HB moiety) and the other one as an elongation ACP (loaded with a malonyl moiety). To evaluate this hypothesis, we determined the conversion of *holo*-ApeE and ApeF to 4HB-ACP using ApeH. Even in a competitive assay containing both ACPs, the conversion of *holo*-ACP to 4HB-ACP was significantly faster for *holo*-ApeE (Figure S19), suggesting ApeE to be the starter and ApeF the elongation ACP. The malonyl moiety is assumed

to be transferred by MCAT from the fatty acid biosynthesis, as no self-malonylation³⁷ has been observed (see experimental procedures). Hence, for our *in vitro* investigations we converted ApeF to malonyl-ApeF using Sfp, a promiscuous PPTase from *Bacillus subtilis*.³⁸

The TE ApeK strongly interacts with both ACPs (ApeE and ApeF) forming ACP:TE heterooctameric complexes as deduced from size exclusion chromatography and native ESI (nESI) mass spectrometry analysis of the native protein complexes (Figure 2, Figure S13). ApeK is annotated as a 4HBA-TE. However, no measurable hydrolysis activity was detected with either 4HBA-CoA or ACP-bound APE products (Figure S20). Thus, ApeK could serve as a platform for supporting protein–protein interactions or could provide a proofreading activity for off-loading false-primed starter molecules, as indicated by its *in vitro* hydrolysis activity against CoA substrates including octanoyl-CoA (Figure S20).

ApeR and the Heterocomplex ApeO:ApeC Are Essential for Chain Initiation and Elongation of APES.

The characteristic sequences of ApeR and ApeO suggest that both proteins are KSs. The closest homologues of ApeO and ApeR are FabF from FA biosynthesis and KS putatively involved in ladderane biosynthesis,¹² while ApeC did cluster within a clade of priming KAS III enzymes (Figure S21, Table S12). The heterologous production of ApeR in *E. coli* resulted in a soluble homodimer, while ApeO can only be purified in the presence of ApeC, resulting in a stable ApeO:ApeC heterodimer (Figure 2c, Figures S12, S13, and S22). Complex formation between ApeO:ApeC and ApeR does not occur, as shown by our pull-down assays (Figure S22a). Therefore, we conclude that ApeR is similar to KS involved in type II fatty acid biosynthesis (FAS II), while ApeO:ApeC reflects a KS:CLF complex that is a general component of most type II PKS systems.^{1,2,39} In our MS² experiments, we detect the release of Ppant (Figure 3) for ApeR as well as ApeO:ApeC and chain elongation to the first β -ketoacyl product in the presence of either ApeE or ApeF. Thus, two different KS systems are involved in the APE biosynthesis that are promiscuous for ApeE and ApeF.

ApeC does not match any protein sequence deposited in the RCSB data bank. With only one solved apo KS:CLF heterodimer (PDB ID 1TQY),³⁹ we determined the crystal structure of the heterocomplex ApeO:ApeC at 1.45 Å resolution (PDB ID 6QSP, $R_{\text{free}} = 0.162$) (Figure 4). ApeO consists of two domains, each of which resembles the α - β -thiolase fold.⁴⁰ The domains adopt a five-stranded mixed β -sheet (flanked on both sides by helices) with an antiparallel and four parallel β -strands (S1 \downarrow , S5 \uparrow , S4 \downarrow , S2 \downarrow , S3 \downarrow) that are related by an improper 2-fold axis. The architecture of ApeC mimics that of ApeO. Despite significant differences in sequence length, composition (Figure S23), and structure (Figure 4), the N-terminal domains of ApeC (Ile7-Ala193) and ApeO (Ile4-Gln234) have identical topologies with an rmsd of 1.8 Å (105 C α atoms, sequence identity 12%). The C-terminal 40 residues of ApeC are composed of an antiparallel β -sheet of three strands and one helix only (mimicking half of the α - β -thiolase fold with S1' \uparrow , S5' \downarrow , S4' \uparrow).

The ApeO:ApeC complex displays an extensive dimeric interface of about 2250 Å², which is predominantly formed by the S3 strands of the ApeO and ApeC N-terminal domains, respectively. Both subunits complement each other in an antiparallel fashion, resulting in a 10-strand mixed β -sheet.

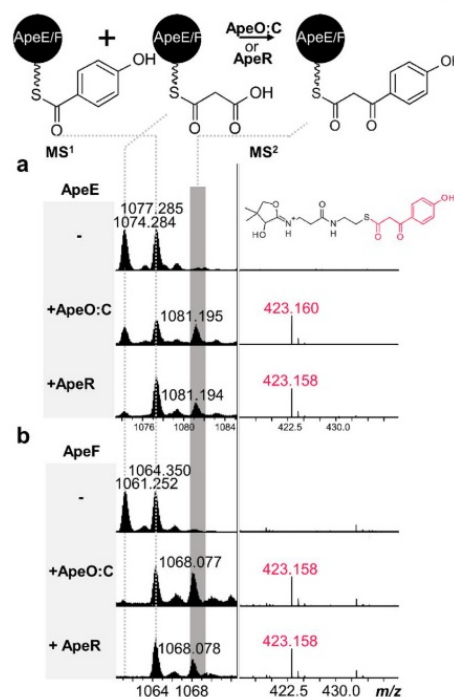


Figure 3. Chain elongation catalyzed by the KS:CLF complex ApeO:ApeC or the KS ApeR. ApeO:ApeC or ApeR was assayed with ApeE (a) or ApeF (b). 4HB-ACP (AasS reaction by ApeH) served as a starter and malonyl-ACP as an elongation ACP. The reaction was analyzed via HR-HPLC-MS in the positive mode with a Ppant ejection in MS². m/z values of the average protein masses of the 11+ charge states (MS¹) are shown. m/z 1074.284 and 1061.252 correspond to malonyl-ApeE and malonyl-ApeF in MS¹. m/z 1077.285 and 1064.350 are assigned to 4HB-ApeE and -ApeF in MS¹. The β -ketoacyl elongation products correspond to m/z 1081.19 for ApeE and 1068.07 for ApeF in MS¹ with characteristic Ppant ejection ion of m/z 423.16 in MS².

Residues along the pseudo-2-fold axis make tight interactions. The surface of the ApeO:ApeC complex illustrates a prominent entrance pore of 9 Å in length and 10 Å in width (Figure 4). The depth of the channel is about 15 Å and points toward a polar pocket formed exclusively by residues of ApeO. Cys159 forms the active site nucleophile, which catalyzes the polyketide extension via the formation of thioester-bound intermediates. ApeC provides a defined hydrophobic binding tunnel that accommodates the growing polyene chain and acts as a molecular ruler. Decarboxylation of the incoming malonyl-ACP molecule takes place at ApeO via His284 and His318. Finally, the generated carbanion attacks the thioester bond of the growing polyene ligand at Cys159.

The KR ApeQ Forms a Tetramer Composed of Dimeric Building Blocks. In the next step of APE biosynthesis, the KR ApeQ reduces the ACP-bound- β -ketoacyl group to the corresponding β -hydroxyacyl with the cosubstrate NADPH. By performing nESI mass spectra, we could show that ApeQ forms a stable tetramer (Figures S13 and S24), which is in agreement with KRAs described for PKS II as well as for FAS II systems.⁴¹ We determined the activity of ApeQ and identified ACP-bound β -hydroxyacyl groups on ApeE and ApeF via Ppant ejection in MS² (Figure S26, Table S9).

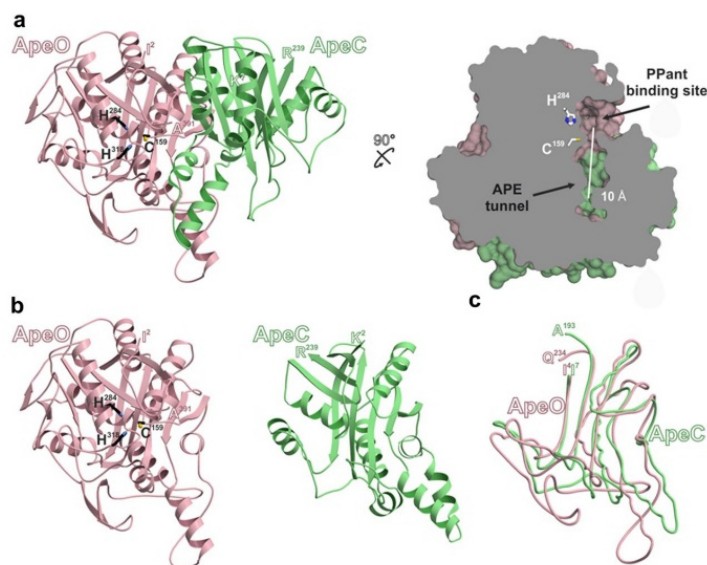


Figure 4. Crystal structure of the ApeO:C complex from *X. doucetiae* (PDB ID 6QSP, 1.45 Å resolution). (a) Ribbon plot of the heterodimeric ApeO:C complex (left panel). The sliced surface representation of the heterodimer (right panel) reveals a prominent entrance pore and a linear APE elongation tunnel (white bar corresponds to 10 Å). (b) Ribbon drawings of ApeO (left panel, 391 residues) and ApeC (right panel, 239 residues). (c) Coil representation of the N-terminal domain of ApeO (Ile4–Gln234) structurally imposed onto ApeC (Ile7–Ala193).

Dehydration in APE Biosynthesis Is Carried out by the DH Heterodimer ApeP:ApeI. The final step of each chain elongation cycle is dehydration of the ACP-linked β -hydroxy acyl group to *trans*-enoyl-ACP. In APE biosynthesis, ApeI and ApeP show low amino acid sequence similarity to previously characterized DHs from FAS and PKS systems. Strikingly, ApeI and ApeP were barely soluble in our expression assays (Figure S22b), but upon coexpression, a stable ApeI:ApeP complex could be purified in high yields (Figure 2d, Figure S13). This observation was unexpected because (i) the primary structure of ApeI and ApeP differs significantly from each other and (ii) heterodimeric DH complexes are not known for PKS and have only been observed once in a mycobacterial type II FAS system.⁴¹

Since ApeI and ApeP hardly match with deposited protein structures in the RCSB, we determined the crystal structure of the ApeI:ApeP complex at 1.85 Å resolution (PDB ID 6QSR, $R_{\text{free}} = 0.236$) (Figure 5). Interestingly, the subunits share a common topology with an rmsd of 1.6 Å (73 C α -atoms, sequence identity 12%). Both subunits adopt a typical hotdog fold⁴³ consisting of six antiparallel β strands (aligned in the order of S1 \downarrow , S2 \uparrow , S4 \downarrow , S5 \uparrow , S6 \downarrow , and S3 \uparrow) that comprise the central α helix H1 in a dome-shaped fashion. ApeP has 60 residues more than ApeI and features two additional short strands, S2a and S3a, as well as the C-terminal helix H2 (Figure 5). Complex formation in ApeI:ApeP takes place along the S3 strand of each subunit, resulting in an elongated antiparallel 12 β -sheet motif with an interaction area of about 1360 Å². This architecture is quite similar to FabA and FabZ, which form homodimers. Intriguingly, in contrast to known DHs,⁴³ the ApeI:ApeP complex exhibits only one active site that is located at the subunit interface.

Access to the catalytic center is restricted by a narrow entrance pore (width 7 Å, length 9 Å) formed by the β 3 strands that select for the flexible Ppant arm. Met92 (ApeP)

might function as gatekeeper. A tunnel of 40 Å with a diameter of 5 Å traverses the ApeI:P complex. The channel is kinked by 90° after 20 Å at the catalytic center where the β -hydroxyacyl group of the Ppant-bound intermediate is dehydrated. The trigger for this change of direction is caused by Pro40 at the beginning of helix H1 (ApeI). The front part of the channel is composed of residues from both subunits and binds the Ppant group, while the back part of the channel consists exclusively of aliphatic side chain residues from ApeP. Here the polyene moiety is bound along helix H1 (ApeP), which is ideally suited to receive the spearhead-like ligand with an aromatic moiety.

Dehydration of the β -hydroxyacyl group occurs far distant from the bulk solvent at His32 and Glu66 that belong to ApeI and ApeP, respectively. The high-resolution electron density map depicts a defined water molecule that is H-bonded to Glu66O ^{δ 1}. In addition, a second water molecule and His19 from ApeP (conserved in FabZ) might function as a proton relay system to shuttle H⁺ between His32N ^{ϵ 1} and Glu66O ^{δ 1}. Notably, there exists an unusual *cis*-peptide bond between Phe33 and His32, which might increase the mobility of the imidazole side chain while retaining its configuration.⁴⁴ On the basis of these structural insights, we suggest that the dehydration reaction is initiated by protonation of the C3-OH group of the substrate via Glu66O ^{δ 1}-H (Brønsted acid) to generate an oxonium ion as a proper leaving group. At the same time, His32N ^{ϵ 1} acts as the Brønsted base and deprotonates the C2-position of the substrate by forming the 2,3-*trans*-unsaturated polyene. With the release of the water molecule, the generated lipophilic polyene group is no longer tolerated at the polar active site and gets expelled from the ApeI:ApeP complex, thus serving as a new template for the next cycle in the APE biosynthesis.

Formation and Identification of APE Pigments by *in Vitro* Studies. In the *in vitro* combination of ApeH (AaS), ApeE and ApeF (ACPs), ApeO:ApeC (KS:CLF) as well as

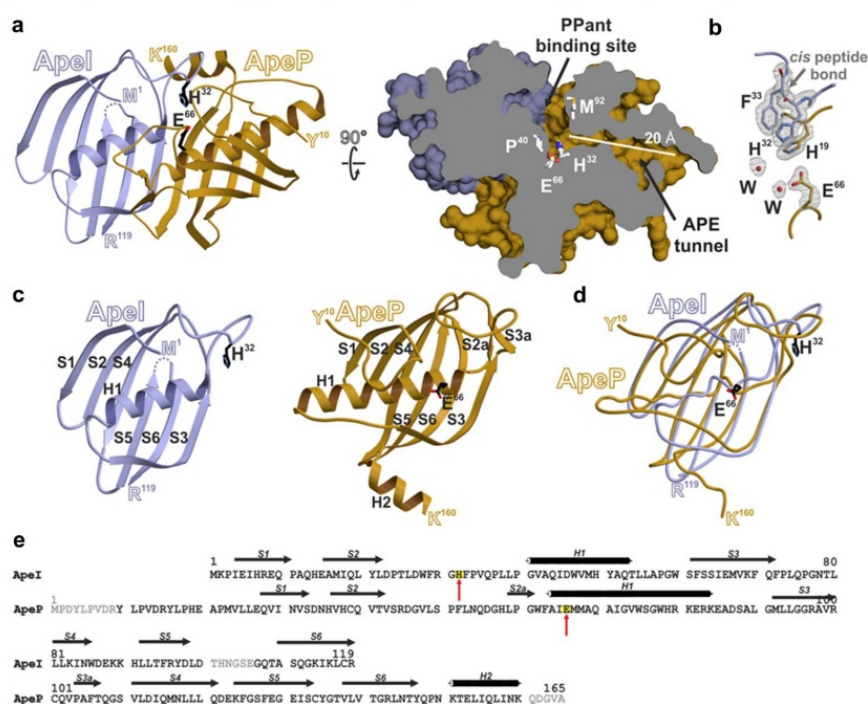


Figure 5. Crystal structure of the ApeI:ApeP complex from *X. doucetiae* (PDB ID 6QSR, 1.85 Å resolution). (a) Ribbon plot of the heterodimer (left panel). The catalytic center of the dehydratase is located at the interface of both subunits. The sliced surface representation (right panel) illustrates an elongated and narrow APE tunnel (white bar represents 20 Å). (b) Stick representation of the proton relay shuttle. The central residues including two water molecules are depicted in the $2F_o - F_c$ -electron density map (gray lattice, contoured to 1σ). The *cis* peptide bond between His32 and Phe33 in ApeI is highlighted with a gray arrow. (c) Ribbon drawings of ApeI (left panel, 119 residues) and ApeP (right panel, 165 residues). The typical hotdog fold of both subunits is shown in similar orientations. (d) Coil representation of the structural superposition of ApeI and ApeP. (e) Both subunits have unique sequences (12% identity), although their domains share a common topology. Secondary structure elements are indicated by black arrows (S, strand) and bars (H, helix). His32 and Glu66 are highlighted; residues not defined in the $2F_o - F_c$ -electron density map are colored in gray.

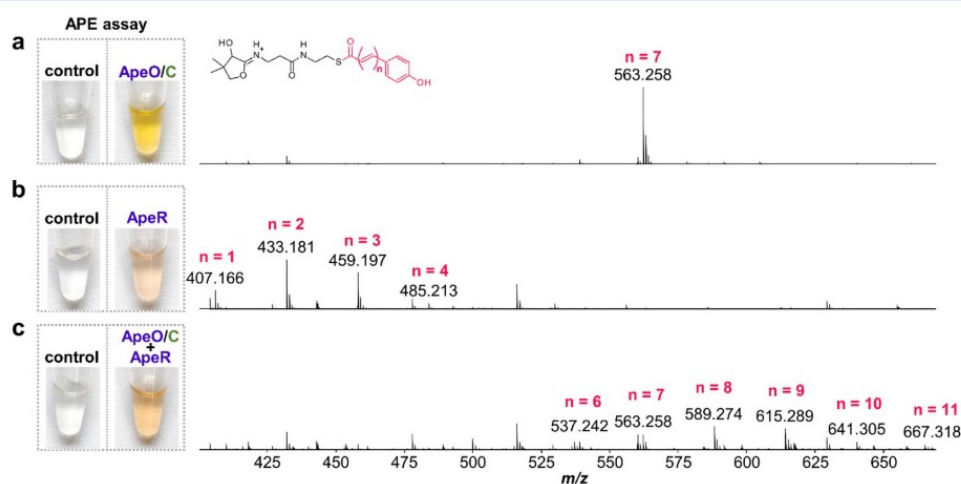


Figure 6. Ppant ejection of ACP-bound APE formation with ApeO:ApeC, ApeR, and ApeO:ApeC plus ApeR. The reactions were performed with 4HB-ApeE as starter (AasS reaction by ApeH) and malonyl-ApeE as elongation ACP by adding the essential enzymes ApeO:ApeC, ApeR, ApeQ (+NADPH), and ApeP:ApeI. Different APE chain lengths were produced by using either (a) ApeO:ApeC, (b) ApeR, or (c) ApeO:ApeC together with ApeR. In control assays no ApeP:ApeI heterodimer was added. See also Figure 27 and Tables S10 and S11.

ApeR (KS), ApeQ (KR) and ApeP:ApeI (DH), we were able to detect ACP-bound APE products, which we unambiguously

assigned in our MS² Ppant ejection assay.⁴² Fortunately, the formed polyenes are yellow so that the progress of the reaction

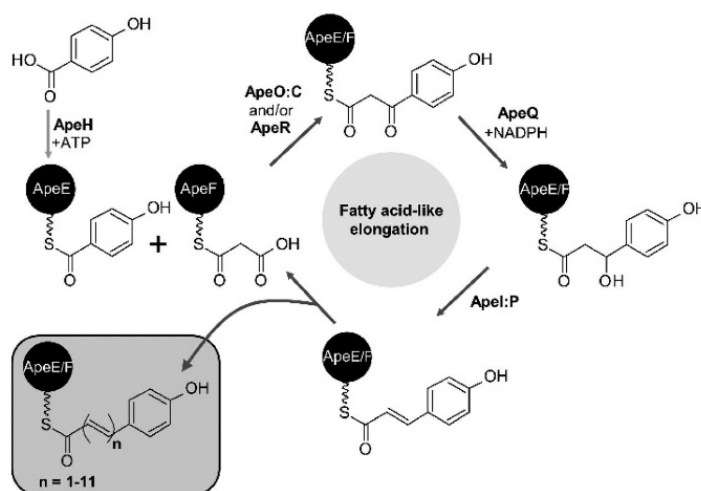


Figure 7. Suggested biosynthesis of the APE chromophore. The 4-hydroxybenzoic acid (4HBA) starter unit is transferred to ApeE by ApeH, resulting in 4HB-ApeE. The (initial) chain elongation step can be performed with malonyl-ApeF by either ApeO:ApeC or ApeR, resulting in the β -ketoacyl product. This intermediate is further reduced to the β -hydroxyacyl product by ApeQ and dehydrated to the α,β -unsaturated acyl product by ApeI:ApeP. The fatty acid-like elongation mechanism is repeated several times, resulting in APES with up to 11 double bonds dependent on the added KS (ApeO:ApeC and/or ApeR).

is well trackable. In combination with the two different KS or KS:CLF dimers, different ACP-bound APES were formed. An assay with ApeO:ApeC alone resulted in only one main APE with seven double bonds (Figure 6a) bound to ApeE, while ApeR produced short-chain APES with up to four double bonds (Figure 6b), suggesting a role in the first steps of chain elongation. A mixture of all three KS:CLF enzymes ApeCOR yields a mixture of ACP-bound APES with up to 11 double bonds (Figure 6c) (Figure S27, Tables S10 and S11).

Interestingly, seven double bonds turned out to be the maximum in our *in vivo* studies, whereas *in vitro* up to 11 double bonds could be detected. This observation is explained by a missing part *in vitro*. One might speculate that *in vivo* the glycerol acyltransferase ApeD preferentially transfers the APE products with six or seven double bonds to a glycerol moiety, thus preventing further elongation.

CONCLUSION

The present work provided detailed insights into the biosynthesis of aryl polyene pigments encoded by APE BGCs widespread in bacteria. We were able to assign a function to each component of the cluster by reconstituting the formation of the actual APE chromophore *in vitro* (Figure 7).

Interestingly, most enzymes involved in APE biosynthesis form multimers, of which some are different from the previously characterized FAS and PKS systems.

Besides the reconstitution experiments we found that ApeH fulfills the task of a promiscuous AaS enzyme, being involved in the initiation of APE biosynthesis. Furthermore, the ACPs ApeE and ApeF, which both form monomers alone, also associate with the TE-like protein ApeK to form stable heterooctamers. Intriguingly, the homodimeric ApeR and the ApeO:ApeC heterodimer individually control the chain length of the natural products, while a DH complex composed of two distinct subunits, ApeI and ApeP, was identified with the catalytic center formed by both subunits. Following our

elaborate biochemical characterization of key players in APE biosynthesis, we were able to investigate the X-ray structures of the ApeO:ApeC and ApeI:ApeP heterodimeric complexes. The crystallographic studies provided molecular insights into this fascinating type II PKS multienzyme machinery, revealing similarities to type II systems of both PKS and FAS biosynthesis.

Taken together, the elucidation of APE pigment biosynthesis combined with the crystallographic analysis is the beginning of a profound understanding of the sophisticated reaction cascades for which the AaS ApeH, the individual KS ApeR and ApeO:ApeC, the KR ApeQ, the DH ApeI:ApeP, and the ApeE:ApeK and ApeF:ApeK heterooligomers certainly play a central role. The interactions of APE enzymes identified by native mass spectrometry and crystallization let us propose that the peculiar APE biosynthesis machinery might form a transient megacomplex similar to the covalent type I PKS systems. However, further details are required, in particular the characterization of the ApeE:ApeK and ApeF:ApeK as shuttle systems and their peculiar function in the transfer of the APE moiety.

The presence of enzymes encoded in the *ape* BGC with similarity to lipid biosynthesis enzymes suggests that the ACP-bound APE intermediates might be further connected to lipids. However, this late biosynthesis and the nature of the formed lipid is beyond the scope of this study and will be addressed in the future.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b10776.

Experimental procedures and supporting tables and figures (PDF)

AUTHOR INFORMATION

Corresponding Author

*h.bode@bio.uni-frankfurt.de

ORCID

Nina Morgner: 0000-0002-1872-490X

Michael Groll: 0000-0002-1660-340X

Helge B. Bode: 0000-0001-6048-5909

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Work in the Bode and Morgner laboratories was supported by the LOEWE Schwerpunkt MegaSyn funded by the State of Hesse. The authors acknowledge the Deutsche Forschungsgemeinschaft for funding of the Impact II qToF mass spectrometer (Grant INST 161/810-1). N.M. received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC Grant agreement no. 337567 and the Cluster of Excellence Frankfurt (Macromolecular Complexes). M.G. and M.S. were supported by the SFB 749/A10. The authors are grateful to John Cronan for providing the *E. coli* BL21Δ*acpT* mutant for apo-ACP production, Michael Burkart for discussions about AasS enzymes, Carsten Kegler for providing the plasmid for Sfp production, Jonas Watzel for providing a 9.8 kDa standard protein, and Michael Karas for MALDI access. We thank the staff of the beamline X06SA (SLS, Villigen, Switzerland) for support during data collection.

REFERENCES

- Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II Polyketide Synthases: Gaining a Deeper Insight into Enzymatic Teamwork. *Nat. Prod. Rep.* **2007**, *24*, 162–190.
- Hertweck, C. Die biosynthetische Grundlage der Polyketid-Vielfalt. *Angew. Chem.* **2009**, *121*, 4782–4811.
- Herman, N. A.; Kim, S. J.; Li, J. S.; Cai, W.; Koshino, H.; Zhang, W. The Industrial Anaerobe *Clostridium Acetobutylicum* Uses Polyketides to Regulate Cellular Differentiation. *Nat. Commun.* **2017**, *8*, 1514.
- Kunze, B.; Höfle, G.; Reichenbach, H. he Aurachinons, New Quinoline Antibiotics from Myxobacteria: Production, Physico-Chemical and Biological Properties. *J. Antibiot.* **1987**, *40*, 258–265.
- Sandmann, A.; Dickschat, J.; Jenke-Kodama, H.; Kunze, B.; Dittmann, E.; Müller, R. A Type II Polyketide Synthase from the Gram-Negative Bacterium *Stigmatella Aurantiaca* is Involved in Aurachin Alkaloid Biosynthesis. *Angew. Chem., Int. Ed.* **2007**, *46*, 2712–2716.
- Panter, F.; Krug, D.; Baumann, S.; Müller, R. Self-Resistance Guided Genome Mining Uncovers New Topoisomerase Inhibitors from Myxobacteria. *Chem. Sci.* **2018**, *9*, 4898–4908.
- Brachmann, A. O.; Joyce, S. A.; Jenke-Kodama, H.; Schwär, G.; Clarke, D. J.; Bode, H. B. A Type II Polyketide Synthase Is Responsible for Anthraquinone Biosynthesis in *Photorhabdus luminescens*. *ChemBioChem* **2007**, *8*, 1721–1728.
- Wenzel, S. C.; Müller, R. Myxobacterial Natural Product Assembly Lines: Fascinating Examples of Curious Biochemistry. *Nat. Prod. Rep.* **2007**, *24*, 1211–1224.
- Brautaset, T.; Sletta, H.; Degnes, K. F.; Sekurova, O. N.; Bakke, I.; Volokhan, O.; Andreassen, T.; Ellingsen, T. E.; Zotchev, S. B. New Nystatin-Related Antifungal Polyene Macrolides with Altered Polycl Region Generated via Biosynthetic Engineering of *Streptomyces noursei*. *Appl. Environ. Microbiol.* **2011**, *77*, 6636–6643.
- Brandt, P.; García-Altares, M.; Nett, M.; Hertweck, C.; Hoffmeister, D. Induced Chemical Defense of a Mushroom by a Double-Bond-Shifting Polyene Synthase. *Angew. Chem., Int. Ed.* **2017**, *56*, 5937–5941.
- Du, D.; Katsuyama, Y.; Shin-ya, K.; Ohnishi, Y. Reconstitution of a Type II Polyketide Synthase that Catalyzes Polyene Formation. *Angew. Chem., Int. Ed.* **2018**, *57*, 1954–1957.
- Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158*, 412–421.
- Schöner, T. A.; Gassel, S.; Osawa, A.; Tobias, N. J.; Okuno, Y.; Sakakibara, Y.; Shindo, K.; Sandmann, G.; Bode, H. B. Aryl Polyenes, a Highly Abundant Class of Bacterial Natural Products, Are Functionally Related to Antioxidative Carotenoids. *ChemBioChem* **2016**, *17*, 247–253.
- Tobias, N. J.; Shi, Y.-M.; Bode, H. B. Refining the Natural Product Repertoire in Entomopathogenic Bacteria. *Trends Microbiol.* **2018**, *121*, 4782–4811.
- Shi, Y.-M.; Bode, H. B. Chemical Language and Warfare of Bacterial Natural Products in Bacteria-Nematode-Insect Interactions. *Nat. Prod. Rep.* **2018**, *35*, 309–335.
- Herbert, E. E.; Goodrich-Blair, H. Friend and foe: The Two Faces of *Xenorhabdus Nematophila*. *Nat. Rev. Microbiol.* **2007**, *5*, 634.
- Stephens, W. L.; Starr, M. P. Localization of Carotenoid Pigment in Cytoplasmic Membrane of *Xanthomonas Juglandis*. *J. Bacteriol.* **1963**, *86*, 1070–1074.
- Chun, W. C. Xanthomonadins, Unique Yellow Pigments of the Genus *Xanthomonas*. *PHI* **2000**, DOI: 10.1094/PHI-A-2000-0824-01.
- Starr, M. P.; Jenkins, C. L.; Bussey, L. B.; Andrewes, A. G. Chemotaxonomic Significance of the Xanthomonadins, Novel Brominated Aryl-Polyene Pigments Produced by Bacteria of the Genus *Xanthomonas*. *Arch. Microbiol.* **1977**, *113*, 1–9.
- Reichenbach, H.; Kohl, W.; Böttger-Vetter, A.; Achenbach, H. Flexirubin-Type Pigments in *Flavobacterium*. *Arch. Microbiol.* **1980**, *126*, 291–293.
- Britton, G. Structure and Properties of Carotenoids in Relation to Function. *FASEB J.* **1995**, *9*, 1551–1558.
- Stahl, W.; Sies, H. Bioactivity and Protective Effects of Natural Carotenoids. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2005**, *1740*, 101–107.
- Schöner, T. A.; Fuchs, S. W.; Schönau, C.; Bode, H. B. Initiation of the Flexirubin Biosynthesis in *Chitinophaga pinensis*. *Microb. Biotechnol.* **2014**, *7*, 232–241.
- Schöner, T. A.; Fuchs, S. W.; Reinhold-Hurek, B.; Bode, H. B. Identification and Biosynthesis of a Novel Xanthomonadin-Dialkylresorcinol-Hybrid from *Azoarcus* sp. BH72. *PLoS One* **2014**, *9*, e90922.
- Kumagai, Y.; Yoshizawa, S.; Nakajima, Y.; Watanabe, M.; Fukunaga, T.; Ogura, Y.; Hayashi, T.; Oshima, K.; Hattori, M.; Ikeuchi, M.; et al. Solar-panel and Parasol Strategies Shape the Proteorhodopsin Distribution Pattern in Marine *Flavobacteria*. *ISME J.* **2018**, *12*, 1329.
- Landrum, J. T.; Bone, R. A. Lutein, Zeaxanthin, and the Macular Pigment. *Arch. Biochem. Biophys.* **2001**, *385*, 28–40.
- Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. Simple “On-Demand” Production of Bioactive Natural Products. *ChemBioChem* **2015**, *16*, 1115–1119.
- Fautz, E.; Reichenbach, H. A Simple Test for Flexirubin-Type Pigments. *FEMS Microbiol. Lett.* **1980**, *8*, 87–91.
- Fuchs, G.; Boll, M.; Heider, J. Microbial Degradation of Aromatic Compounds - from one Strategy to Four. *Nat. Rev. Microbiol.* **2011**, *9*, 803–816.
- Hertweck, C.; Moore, B. S. A Plant-Like Biosynthesis of Benzoyl-CoA in the Marine Bacterium ‘*Streptomyces maritimus*’. *Tetrahedron* **2000**, *56*, 9115–9120.
- Piel, J.; Hertweck, C.; Shipley, P. R.; Hunt, D. M.; Newman, M. S.; Moore, B. S. Cloning, Sequencing and Analysis of the Enterocin Biosynthesis Gene Cluster from the Marine Isolate ‘*Streptomyces*

maritimus: Evidence for the Deraiment of an Aromatic Polyketide Synthase. *Chem. Biol.* **2000**, *7*, 943–955.

(32) Lay, N. R. D.; Cronan, J. E. A Genome Rearrangement has Orphaned the *Escherichia Coli* K-12 AcpT Phosphopantetheinyl Transferase from its Cognate *Escherichia Coli* O157:H7 substrates. *Mol. Microbiol.* **2006**, *61*, 232–242.

(33) Rock, C. O.; Cronan, J. E. Solubilization, Purification, and Salt Activation of Acyl-Acyl Carrier Protein Synthetase from *Escherichia Coli*. *Int. J. STD AIDS* **1979**, *254*, 7116–7122.

(34) Ray, T. K.; Cronan, J. E., JR. Activation of Long Chain Fatty Acids with Acyl Carrier Protein: Demonstration of a New Enzyme, Acyl-Acyl Carrier Protein Synthetase, in *Escherichia Coli*. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 4374–4378.

(35) Campopiano, D. J. ACP-AasS You Like It. *Chem. Biol.* **2014**, *21*, 1257–1259.

(36) Beld, J.; Finzel, K.; Burkart, M. D. Versatility of Acyl-Acyl Carrier Protein Synthetases. *Chem. Biol.* **2014**, *21*, 1293–1299.

(37) Arthur, C. J.; Szafranska, A.; Evans, S. E.; Findlow, S. C.; Burston, S. G.; Owen, P.; Clark-Lewis, L.; Simpson, T. J.; Crosby, J.; Crump, M. P. Self-Malonylation is an Intrinsic Property of a Chemically Synthesized Type II Polyketide Synthase Acyl Carrier Protein. *Biochemistry* **2005**, *44*, 15414–15421.

(38) Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a *Bacillus Subtilis* Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases. *Biochemistry* **1998**, *37*, 1585–1595.

(39) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradsky, K. F.; Khosla, C.; Stroud, R. M. An Antibiotic Factory Caught in Action. *Nat. Struct. Mol. Biol.* **2004**, *11*, 888–893.

(40) Mathieu, M.; Zeelen, J.; Pauptit, R. A.; Erdmann, R.; Kunau, W.-H.; Wierenga, R. K. The 2.8Å Crystal Structure of Peroxisomal 3-ketoacyl-CoA Thiolase of *Saccharomyces cerevisiae*: A Five-Layered $\alpha\beta\alpha\beta\alpha$ Structure Constructed from Two Core Domains of Identical Topology. *Structure* **1994**, *2*, 797–808.

(41) Nanson, J. D.; Forwood, J. K. Structural Characterisation of FabG from *Yersinia pestis*, a Key Component of Bacterial Fatty Acid Synthesis. *PLoS One* **2015**, *10*, e0141543.

(42) Meluzzi, D.; Zheng, W. H.; Hensler, M.; Nizet, V.; Dorrestein, P. C. Top-Down Mass Spectrometry on Low-Resolution Instruments: Characterization of Phosphopantetheinylated Carrier Domains in Polyketide and Non-Ribosomal Biosynthetic Pathways. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3107–3111.

(43) Leesong, M.; Henderson, B. S.; Gillig, J. R.; Schwab, J. M.; Smith, J. L. Structure of a Dehydratase–Isomerase from the Bacterial Pathway for Biosynthesis of Unsaturated Fatty Acids: Two Catalytic Activities in One Active Site. *Structure* **1996**, *4*, 253–264.

(44) Mayr, L. M.; Landt, O.; Hahn, U.; Schmidt, F. X. Stability and Folding Kinetics of Ribonuclease T1 are Strongly Altered by the Replacement of Cis-proline 39 with Alanine. *J. Mol. Biol.* **1993**, *231*, 897–912.

SUPPORTING INFORMATION

Contents

Experimental Section	S3
Cloning.....	S3
P _{BAD} promoter exchange in <i>Xenorhabdus doucetiae</i>	S3
APE - analytical compound production, extraction and saponification	S4
Purification of 1 for NMR analysis.....	S5
Heterologous protein production and purification for <i>in vitro</i> analysis.....	S5
Heterologous protein production and purification for crystallization	S7
Crystallization and X-ray experiments.....	S8
HR-HPLC-ESI-MS.....	S9
MALDI-MS.....	S10
ESI-MS of native protein complexes	S10
<i>In vitro</i> analysis	S11
ATP-PP _i exchange assay.....	S11
Hydrolysis reaction by ApeK	S11
Preparation of <i>crypto</i> -ACPs	S11
(Initial) chain elongation step by KS(-CLF) (ApeO:ApeC and/or ApeR, or ArcK/J).....	S12
Ketoreduction step by KR (ApeQ)	S12
APE-assay.....	S12
Phylogenetic analysis	S12
Supplementary Tables	S14
Table S1. Strains used in this study	S14
Table S2. Oligonucleotides used for plasmid construction	S15
Table S3. Plasmids used in this study	S18
Table S4. Homology of different strains with <i>ape</i> genes.....	S19
Table S5. ¹ H (500 MHz) and ¹³ C (125 MHz) NMR data assignments for APE 1 DMSO- <i>d</i> ₆	S20
Table S6. Substrate specificity of ApeH by its AasS reaction with ApeE.	S21
Table S7. Substrate specificity of ArcT from <i>Azoarcus</i> sp. BH72.....	S22
Table S8. Crystallographic data collection and refinement statistics.	S23
Table S9. KR reaction by ApeQ.....	S24
Table S10. Detected Ppant ejection ions in assay Ape-assays	S24
Table S11. Detected Ppant ejection ions in fluoro-control APE-assay	S25
Table S12. Amino acid sequences used for phylogenetic analysis of ApeO, ApeC and ApeR.	S25
Supplementary Figures	S28
Figure S1. Phylogenetic distribution of <i>ape</i> gene cluster in <i>Xenorhabdus</i> and <i>Photorhabdus</i> sp. ...	S28

Figure S2. Arcuflavin gene cluster from <i>Azoarcus</i> sp. BH72 involved in biosynthesis of arcuflavin	S29
Figure S3. Comparison of arabinose induced (+) and non-induced (-) <i>X. doucetiae ape</i> strains.	S30
Figure S4. KOH test.....	S30
Figure S5. HR-HPLC-UV-MS analysis of <i>X. doucetiae</i> Δ <i>DCΔhfq</i> _P _{BAD} <i>apeB</i> and wildtype extracts	S31
Figure S6. ¹ H NMR spectrum of 1 in DMSO- <i>d</i> ₆ .	S33
Figure S7. HSQC spectrum of 1 in DMSO- <i>d</i> ₆	S33
Figure S8. HMBC spectrum of 1 in DMSO- <i>d</i> ₆	S34
Figure S9. ¹ H- ¹ H COSY spectrum of 1 in DMSO- <i>d</i> ₆ .	S34
Figure S10. ROESY spectrum of 1 in DMSO- <i>d</i> ₆ .	S35
Figure S11. Post-translational protein modifications of the ACPs ApeE and ApeF	S35
Figure S12. SDS-PAGE of all APE enzymes tested in this work .	S36
Figure S13. Size exclusion chromatography of APE protein complexes.....	S37
Figure S14. Size exclusion chromatography of ApeE, ApeF and ApeK.	S38
Figure S15. Substrate specificity of ApeH by AasS reaction with ApeE.....	S39
Figure S16. ATP-PPi exchange assay of ApeH with benzoic acid derivatives as substrates.	S40
Figure S17. HPLC-UV analysis at 420 nm of <i>X. doucetiae</i> Δ <i>DCΔhfq</i> P _{BAD} <i>apeB</i> overproduction strain fed with different benzoic acid precursor	S40
Figure S18. Substrate specificity of ArcT from <i>Azoarcus</i> sp. BH72.	S41
Figure S19. Differences in AasS reaction of ApeH with either ApeE or ApeF	S42
Figure S20. TE hydrolysis activity of ApeK analyzed by HPLC-UV-MS.	S43
Figure S21. Phylogenetic analysis ApeO, ApeC and ApeR	S44
Figure S22. SDS-PAGE of interaction analysis between ApeO:ApeC/R	S45
Figure S23. Comparison of CLFs from aromatic (Ar), Polyene (En) and APE PKS systems.	S46
Figure S24. nESI-MS analysis of the native protein complex	S47
Figure S25. Initial chain elongation step catalyzed by ArcK:ArcJ	S48
Figure S26. KR reaction by ApeQ.	S48
Figure S27. Ppant ejection ions of ACP-bound APE-formation plus fluoro-substrate control	S49
Supplementary References	S50

Experimental Section

Cloning

For isolation of genomic DNA from *X. doucetiae* Gentra Puregene Yeast/Bact Kit (Qiagen) was used. Polymerase chain reaction (PCR) was performed with Phire Hot Start II DNA polymerase (Thermo Scientific), Phusion High-Fidelity DNA Polymerase or Q5 polymerase (New England Biolabs) according to the manufacturer's instructions. Oligonucleotides were purchased from Eurofins Genomics. The Invisorb Spin DNA Extraction Kit (Stratec) was used for DNA purification from agarose gels. Plasmid isolation was performed with the Invisorb Spin Plasmid Mini Two Kit (Stratec). Plasmid-backbone PCRs were restriction digested with *DpnI* (New England Biolabs) following the manufacturer's protocol. If not stated otherwise, all plasmids were cloned via Hot Fusion¹ with corresponding primers listed in Table S2. Promoter exchange plasmids pCEP_Km_142/143 and pCEP_Km_144/145 were constructed via Hot Fusion reaction with *NdeI* linearized backbone plasmid pCEP_Km. The promoter region was therefore amplified with primers TS142 and TS143, or TS144 and TS145, respectively. pCATI_arcV (*NcoI* and *SacI*) and pCOLA_arcJ_His_arcK (MCS1: *EcoRI* and *HindIII*, MCS2: *NdeI* and *Acc65I*) were constructed via ligation reaction with T4 DNA ligase (Sigma) of appropriate restriction digested plasmid and insert-PCR. *E. coli* DH10B and *E. coli* ST18 λ pir (pCEP_Km_142/143, pCEP_Km_144/145, pAL03_ape_mP) used as cloning strains and electroporated with desalted Hot Fusion/ligation reaction (MF-Millipore membrane, VSWP, 0.025 μ m).

P_{BAD} promoter exchange in *Xenorhabdus doucetiae*

Promoter exchange mutants *X. doucetiae* Δ xrdE_P_{BAD}apeA and *X. doucetiae* Δ xrdE_P_{BAD}apeB were constructed by conjugation with *E. coli* ST18 strains harboring the plasmid with first 300-600 bp of the gene of interest (pCEP_Km_142/143, pCEP_Km_144/145) as described previously.² As a result of a homologous recombination event the mutants possess the arabinose inducible P_{BAD} promoter system (*araC* and *araBAD*) in front of the desired gene by plasmid insertion, which has to be maintained by further antibiotic resistance selection.

Within markerless promoter exchange like in *X. doucetiae* Δ DC Δ hfqP_{BAD}apeB, the plasmid backbone is lost by a second homologous recombination event, a well-known strategy for constructing deletion mutants.³ Therefore 700-900 bp up- and downstream of the native start codon is PCR amplified and cloned flanking to the P_{BAD} promoter

cassette (*araC* and *araBAD*), resulting here in pAL03_ape_mP. The markerless promoter exchange mutants possess the arabinose inducible P_{BAD} system, but no further antibiotic selection is necessary.

Xenorhabdus was conjugated with *E. coli* ST18λpir harboring the respective promoter exchange plasmids. Therefore, both strains were grown in 10 mL LB-medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl at pH 7.5) (*E. coli* ST18-λpir was supplemented with 50 µg/mL kanamycin and δ-aminolevulinic acid) to an OD₆₀₀ of 0.6-0.8 and harvested in a ratio of 5:1 (5 mL:1 mL), respectively. To get rid of remaining δ-aminolevulinic acid, *E. coli* ST18λpir pellet was washed three times with 1 mL LB. Cells were resuspended in 60-90 µL LB and mixed, prior pipetting them in one drop (90 µL) on a LB agar plate. After 1 d at 30°C cells were resuspended in 1 mL LB, and promoter exchange mutants were selected by striking 200 µL of cell suspension on LB agar plates supplemented with kanamycin (50 µg/mL). Further selection of markerless promoter exchange mutants via double cross-over was achieved by cultivation on LB agar plates containing 12% (w/v) sucrose.

APE - analytical compound production, extraction and saponification

X. doucetiae *ape* promoter exchange mutants were grown in 200 mL LB-medium at 30°C, 200 rpm for 72 h. Media were supplemented with kanamycin (50 µg/mL), if necessary or respective feeding substances (Table S6) in a final concentration of 1 mM. The *ape* overexpression was induced with 0.2% (w/v) L-arabinose. The repression of the non-induced strain was achieved by addition of 0.2% (w/v) glucose. Pre-cultures were grown overnight in LB-medium and the following day appropriate cell material was used for inoculation of the main cultures to an OD₆₀₀ 0.2 with direct addition of arabinose (glucose) and feeding substances, respectively. Extraction and saponification was performed as described previously.⁴ Cells were harvested at 10,000 rpm for 10 min and pellet was extracted with 30 mL 1:2 methanol (MeOH)/dichloromethane (DCM) with shaking until cell debris seems to be non-pigmented and was removed at 10,000 rpm for 10 min. The supernatant of the crude extract was filtered and a saponification reaction was performed by adding 0.5 M KOH solution (½ of the crude extract volume) with stirring at room temperature for 1 h. The reaction was neutralized with 2 M H₂SO₄ solution to pH 7.0 and the organic layer was washed with brine, deionized water and dried with Na₂SO₄ in a separation funnel. The saponified crude extract was filtered and concentrated to dryness under vacuum and

stored at -20°C, if necessary. For HR-HPLC-UV-ESI-MS analysis, the extract was dissolved in 2 mL 1:2 methanol/dichloromethane and centrifuged for 30 min at 13,000 rpm.

Purification of 1 for NMR analysis

*X. doucetiae*Δ*xrdE*_P_{BAD}*apeA* was cultivated as described above, whereas extraction volumes were adjusted. In total, *ape* overexpressing mutant was cultivated in 6 L LB-medium (6x 1 L in 5 L flasks). The dried crude extract was dissolved in 15 mL 2:1 MeOH:DCM, non-pigmented precipitate was centrifuged (10,000 rpm, 10 min) and the supernatant fractionated with Sephadex LH-20 chromatography using MeOH as the eluent. Pigmented fractions were analyzed by HPLC-MS and the enriched fractions containing 1 were used for a further two-step purification on an Agilent preparative HPLC system. Dried extract was dissolved in 3 mL DMSO and passed over a C3 column (Zorbax, 250 mm x 4.6 mm), the separation was carried out using an ACN/H₂O gradient (35%-100% ACN for 55 min) at a flow rate of 3 mL/min.

Heterologous protein production and purification for *in vitro* analysis

E. coli BL21 (DE3) or *E. coli* BL21Δ*acpT* (for production of *apo*-ApeE and -ApeF) was transformed with protein production plasmids (Table S3). Protein production was achieved using an auto-induction protocol.⁵ Therefore 1 L LB-medium was supplemented with 20 mL 50 x 5052 (25% glycerol, 2.5% glucose, and 10% α-lactose monohydrate), 50 mL 20 x NPS (1 M Na₂HPO₄, 1 M KH₂PO₄, and 0.5 M (NH₄)₂SO₄), 1 mL 1 M MgSO₄ * 7 H₂O and the respective antibiotics (35 μg/mL chloramphenicol, 50 μg/mL spectinomycin, 50 μg/mL kanamycin). This auto-induction medium was inoculated with the appropriate pre-culture and incubated at 37°C and 180 rpm to an OD₆₀₀ = 0.4-0.8, cooled down for 15-45 min at 4°C and incubated overnight at 25°C and 180 rpm.

Purification of Strep-tagged proteins (ApeE:ApeK, ApeF:ApeK, ApeE, ApeF, ApeK, ApeH, ApeO:ApeC, ApeC, ApeR, ApeQ, ApeI:ApeP, ApeI, ApeP, ApeB). Cells were harvested and resuspended in Strep-tag binding buffer (400 mM NaCl, 150 mM Tris, pH 8: ApeE:ApeK, ApeF:ApeK, ApeE, ApeF, ApeO:ApeC, ApeC, ApeR, ApeI:ApeP, ApeI, ApeP, ApeB; pH 7: ApeQ, pH 9: ApeH, ApeK) and with the addition of one tablet cComplete Protease Inhibitor Cocktail (Roche) and lysozyme incubated for 20-50 min at 4°C. After cell lysis by sonication, cell debris was removed by centrifugation (20 min, 20,000 rpm, 4°C) and the supernatant was passed over a 5 mL StrepTrap HP column

(GE Healthcare) by an ÄKTApurifier system (GE Healthcare). Proteins were eluted with Strep-tag elution buffer (400 mM NaCl, 150 mM Tris, 2.5 mM d-Desthiobiotin, pH 8: ApeE:ApeK, ApeF:ApeK, ApeE, ApeF, ApeO:ApeC, ApeR, Apel:ApeP, ApeB; pH 7: ApeQ, pH 9: ApeH) and buffer exchanged in storage buffer (20 mM HEPES, 1 mM DTT, 25% (v/v) glycerol, pH 8: ApeE:ApeK, ApeF:ApeK, ApeE, ApeF, ApeO:ApeC, Apel:ApeP, ApeB; pH 7 (with 200 mM NaCl): ApeQ, pH 9: ApeH, ApeK) using PD-10 desalting columns (Amersham Biosciences). Proteins were further concentrated to 2-10 mg/ml using Amicon concentration devices (Merck) and stored in aliquots at -80 °C until needed. Due to activity lost during storage of Apel:ApeP, ApeQ and ApeH, proteins were freshly prepared if the storage exceeded more than two weeks.

Purification of His₆-tagged proteins (Sfp-SUMO, ArcV, ArcT ArcK/J). Cells were harvested and resuspended in His-tag binding buffer (300 mM NaCl, 20 mM Tris, 20 mM imidazol, pH 8.5: Sfp-SUMO, ArcK/J, pH 8 ArcV, ArcT) and with addition of one tablet cOMplete Protease Inhibitor Cocktail (Roche) and lysozyme incubated for 20-50 min at 4°C. After cell lysis by sonication, cell debris was removed by centrifugation (20 min, 20,000 rpm, 4°C) and the supernatant was passed over a 5 mL HisTrap FF column (GE Healthcare) by an ÄKTApurifier system (GE Healthcare). Proteins were eluted with His-tag elution buffer (300 mM NaCl, 20 mM Tris, 500 mM imidazol, pH 8.5) and buffer exchanged in storage buffer (20 mM HEPES, 1 mM DTT, 25% (v/v) glycerol, pH 8.5) using a PD-10 desalting column (Amersham Biosciences). In case of Cherry-fusions (ArcT and ArcV), proteins were dialyzed at 4°C against TEV cleavage buffer (300 mM NaCl, 20 mM Tris, pH 8) with addition of TEV Protease (Sigma) according to manufacturer's protocol and passed again over a 5 mL HisTrap FF column. Flowthrough fractions with ArcV and ArcT were collected, respectively, and buffer exchanged in storage buffer (20 mM HEPES, 1 mM DTT, 25% (v/v) glycerol, pH 8). All proteins were further concentrated to 2-10 mg/ml using Amicon concentration devices (Merck) and stored in aliquots at -80 °C until needed.

Size exclusion chromatography (SEC) with protein complexes ApeE, ApeF, ApeK, ApeE:ApeK, ApeF:ApeK, ApeO:ApeC, ApeR, Apel:ApeP was performed with a HiLoad 26/600 Superdex 200 pg column using a NGC system (BioRad). Proteins of Strep-tag affinity purification were loaded with a 10 mL loop onto the column and eluted with

320 mL SEC-buffer at a flow rate of 1.5 mL/min. For mass determination of protein complexes the protein standard mixture „Gel Filtration Molecular Weight Markers Kit for Molecular Weights 12,000–200,000 Da“ (Sigma Aldrich) was used following the manufacturer’s protocol.

Proteins were separated on 10-15% SDS-polyacrylamide gels and visualised with Coomassie Brilliant Blue G250.

Heterologous protein production and purification for crystallization

For the structure determination either the heterodimers Apel and ApeP or ApeO and ApeC were heterologously co-produced in BL21-Gold (DE3). Therefore the *E. coli* strains were grown at 37°C in 3 L LB (with chloramphenicol or spectinomycin) to an OD₆₀₀ of 0.8. Protein expression was induced by adding IPTG (1 mM final concentration). After a 20 h incubation at 20°C, the cells were harvested by centrifugation and stored at -20°C. The bacterial pellet was resuspended in 100 mM Tris base, 500 mM NaCl, 5 mM MgCl₂, pH 8.0 (buffer AM for Apel:ApeP) or 100 mM Tris base, 100 mM NaCl, pH 8.0 (buffer AR for ApeO:ApeC) – both supplemented with 5 mM DTT and a small amount of Pefabloc® SC (Sigma-Aldrich). After cell disruption by sonication (Branson Digital Sonifier 250) and centrifugation at 40,000 g for 30 min at 4°C the cleared lysate was applied on an equilibrated StrepTrap™ HP (GE Healthcare, 2 x 5 ml; flowrate 4 mL/min). The bound protein was washed with buffer AM or AR, followed by elution with 2.5 mM d-Desthiobiotin (Sigma-Aldrich). Further purification was achieved with size exclusion chromatography (SEC). Therefore a HiLoad 16/600 Superdex 200 pg (GE Healthcare) was equilibrated with 120 mL 20 mM Tris base, 100 mM NaCl, 1 mM MgCl₂, pH 8.0, 5 mM DTT (buffer BM for Apel:ApeP) or 20 mM Tris base, 20 mM NaCl, pH 8.0, 5 mM DTT (buffer BR for ApeO:ApeC). After protein loading the column was washed with 120 mL buffer, protein fractions collected, pooled and concentrated using a Vivaspın® 20 ultrafiltration unit (Sartorius) to a final concentration of 20 mg/mL (Apel:ApeP, 10,000 Da cut-off filter) and 18.7 mg/mL (ApeO:ApeC, 30,000 Da cut-off filter).

Crystallization and X-ray experiments

Crystals were obtained at 20°C by using sitting drop vapor diffusion method. Drops of Apel:ApeP contained a 1:1 ratio of protein (20 mg/mL) and reservoir solution (0.1 M Hepes, 30% PEG 6000, pH 7.0), whereas drops of ApeO:ApeC contained a 4:1 ratio of protein (18.7 mg/mL) and reservoir solution (1 M Na/K phosphate, pH 8.0). The crystals were vitrified in liquid nitrogen after substitution with cryoprotectant (final 30% glycerol). Diffraction images were taken at the Swiss Light Source (SLS, Paul Scherrer Institute, Villigen Switzerland: Beamline X06SA) with wavelengths of 1.0 Å (Table S8). Data processing and reduction was performed with the XDS software package.⁶ Initial phases were achieved by respective Patterson search calculations using the program PHASER.⁷ While no hits were found in the RCSB for the primary sequences of ApeP and ApeC, promising sets of coordinates could be identified in case of Apel (PDB-ID: 3ESI, uncharacterized protein of unknown function) and ApeO (PDB-ID: 2IWZ).⁸ The corresponding models 3ESI and 2IWZ included about 50% of the respective search models and were further optimized by molecular editing. The positioning of the adjusted structures resulted in clear rotation and translation solutions. In the following a first building of the models was carried out with the three-dimensional graphics program MAIN⁹, followed by rigid body and individual positioning refinements using REFMAC5.¹⁰ Fortunately, the high resolution of the respective datasets included sufficient details to overcome the initially generated model bias. The implementation of featured enhanced maps with PHENIX¹¹ in the early and middle stages greatly facilitated the model building. In addition, the automatic placement of a large number of bulk water molecules further improved the electron density map by applying ARP/wARP.¹² Eventually, the combination of all these different density modification tools ensured the assignment of secondary structural elements for ApeP and ApeC in form of a polyalanine trace. Following this approach it was possible to construct the asymmetric unit cells in iterative steps. Subsequent refinements using anisotropic corrections of temperature factors of the individual atoms (REFMAC5) resulted in an excellent electron density map. Clear sequencing of the Apel:ApeP and ApeO:ApeC complexes were now possible. Finally, the respective models were completed in successive rounds with COOT.¹³ The geometries of the coordinates were analyzed by MOLPROBITY¹⁴ and optimized for ideal stereochemistry and RMSD deviations with low bond lengths and angles (Table S8).

HR-HPLC-ESI-MS

APE-compounds from crude extraction and *in vitro*-assays were analyzed via high resolution (HR)-HPLC-ESI-UV-MS analysis using a Dionex Ultimate 3000 LC system (Thermo Fisher) coupled to an Impact II electrospray ionization mass spectrometer (Bruker) and a DAD-3000 RS UV-detector (Thermo Fisher).

Separation of APE compounds was achieved with a C3 column (Zorbax 300SB-C3 300Å, 150 mm x 3.0 mm x 3.5 µm, Agilent) by using two different gradients. ACN and H₂O containing 0.1% (v/v) formic acid were used as mobile phases:

Gradient 1: 5% ACN equilibration (0-2 min), followed by a gradient from 5-95% ACN (2-17 min, 17-25 min 100% ACN) ending with a re-equilibration step of 5% ACN (25-30 min) at a flow rate of 0.6 mL/min.

Gradient 2: 5% ACN equilibration (0-2 min), followed by a gradient from 5-40% ACN (2-4 min) and from 40-95% ACN (4-20.5 min), 20.5-29 min 95% ACN) ending with a re-equilibration step of 5% ACN (29-30 min), at a flow rate of 0.4 mL/min. For internal mass calibration a 10 mM sodium formate solution was injected.

Separation of protein samples was performed on a C3 column (Zorbax 300SB-C3 300Å, 150 mm x 3.0 mm x 3.5 µm, Agilent). ACN and H₂O containing 0.1% (v/v) formic acid were used as mobile phases at a flow rate of 0.6 mL/min. HPLC was performed with 30% ACN equilibration (0-1.5 min), followed by a gradient from 30-65% ACN (1.5-27 min) and a further elution step with 95% ACN (27-30 min). For internal mass calibration an ESI-L Mix (Agilent) was injected.

The HPLC-MS analysis was set to positive mode with a mass range of m/z 50-1500 (APE compounds) or m/z 50-2000 (proteins) and an UV at 190-800 nm. APE compounds were detected at 420 nm.

Separation of CoA ester was performed on a C18 column (ACQUITY UPLC BEH, 50 mm x 2.1 mm x 1.7 µm, Waters) using MeOH and 50 mM ammonium acetate as mobile phases. HPLC was performed at a flow rate of 0.3 mL/min with 5% MeOH equilibration (0-4 min), followed by a gradient from 5-50% MeOH (4-10 min), 10-16 min 95% MeOH) ending with a re-equilibration step of 5% MeOH (16-18 min). For internal mass calibration 10 mM sodium formate was injected. The HPLC-MS analysis was set to negative mode with a mass range of m/z 50-1200 with and an UV at 190-800 nm.

For Data analysis of UV-MS-chromatograms Compass DataAnalysis 4.3 (Bruker) was used. The theoretical average masses of proteins were calculated using Compass IsotopePattern 3.0 (Bruker).

MALDI-MS

In vitro reactions of loaded ACPs were analyzed via MALDI-ToF-MS on a Voyager-DE™ STR (Applied Biosystems, Inc.). Therefore 0.3 μ L sample was mixed with 1.7 μ L 2,5-dihydroxybenzoic acid-matrix (40 mg/mL in 35% MeCN with 0.1% trifluoroacetic acid), spotted onto a polished stainless steel target and air-dried. MALDI-MS analysis was performed in positive ion mode with a Voyager-DE STR mass spectrometer (Applied Biosystems) equipped with a nitrogen laser at 337 nm. DataExplorer 4.8 (Applied Biosystems) was used for visualization of spectra.

ATP-PP_i exchange assay was measured on a LTQ Orbitrap XL instrument (Thermo Fisher Scientific, Inc.) equipped with a nitrogen laser at 337 nm. The samples were mixed with the matrix (10 mg/mL 9-aminoacridin in acetone) while spotting them on to a polished stainless steel MALDI target in a 1:1 ratio (1 μ L each). Mass spectra were acquired in negative ion mode over a range of 500 to 540 m/z . Mass spectra for ATP-PP_i exchange analysis were obtained by averaging 100 consecutive laser shots. Data analysis was performed using Xcalibur 2.0.7 (Thermo Fisher Scientific, Inc.). ApeH activity was observed by a mass shift from m/z 514 to m/z 506, indicating the loss of the ATP-label due to PP_i-exchange.¹⁵

ESI-MS of native protein complexes

nESI-MS measurement of intact protein complexes were performed with a Synapt G2-S TOF-MS (Waters). For this, borosilicate glass needles (Harvard Apparatus Limited) were produced in house using a needle puller (Novata) and subsequently coated with platinum/palladium (80:20) in a sputter coater (Leica Microsystems). Sample aliquots for MS were buffer exchanged using Zeba Micro Spin Desalting Columns, MWCO 7 kDa (Thermo fisher) into 100 mM ammonium acetate pH 6.8 solution and stored on ice. Subsequently 3-10 μ L of the protein sample (10-100 μ M protein) were transferred into the glass capillary. For the measurements the needle voltage was adjusted to 2 kV. The protein samples were measured in positive mode with a mass range from 50-14000 m/z . The spectra were evaluated with MassLynx V4.1 (Waters), Massign (University of Oxford) and OriginPro 2017G (OriginLab Corporation).

***In vitro* analysis**

ATP-PP_i exchange assay

The ATP-PP_i exchange assay was performed as described earlier.¹⁵ Briefly, 10 μM ApeH was incubated with 1 mM γ-¹⁸O₄-ATP (Cambridge Isotope Labs), 1 mM acid, 5 mM MgCl₂, and 5 mM PP_i for 10 min at 30°C. Therefore, stock solutions of 3 mM γ-¹⁸O₄-ATP containing 15 mM MgCl₂ in 20 mM Tris (pH 8), 3 mM acid (Table S6) containing 15 mM PP_i in 20 mM Tris (pH 8) and 30 μM ApeH (in Strep elution buffer) were prepared and mixed in a 1:1:1 ratio with a final volume of 6 μL. For subsequent MALDI-MS analysis, see section “MALDI-MS”.

Hydrolysis reaction by ApeK

Hydrolysis activity of Thioesterase ApeK was performed with several CoA esters as substrates. *p*-Cumaroyl-CoA, *trans*-Cinnamoyl-CoA and Octanoyl-CoA were purchased from commercially available sources. 4-HBA CoA was synthesized and purified as described by Guo *et al.*¹⁶ ApeK reaction was performed in TE-buffer (50 mM HEPES, 1 mM DTT, 150 mM KCl, pH 8) containing 1 mM CoA ester. The reaction was started by addition of 20 μM ApeK and incubated overnight at 30°C. APE assay “ApeO:ApeC” was buffer exchanged with Amicon Ultra-0.5 30K devices (Merck Millipore) in TE buffer. All reactions were extracted with 1:1 (v/v) MeOH prior to HPLC-MS analysis.

Preparation of *crypto*-ACPs

Starter-ACP – AasS-reaction catalyzed by ApeH or ArcT. ACPs were buffer exchanged in reaction buffer (10 mM Tris, 10 mM DTT, pH 8) with PD SpinTrap G-25 columns (Amersham Biosciences) and reactions were performed using 50 μM ACP, 1 mM CoA, 0.5 μM Sfp-SUMO (in storage buffer), 2 mM ATP, 5 mM MgCl₂, 1 mM acid-substrate (Table S6, S7) solubilized as 10 mM stock in ethanol), 5-15 μM AasS (in storage buffer) by an overnight incubation at 30°C. ApeH reactions were analyzed by HR-HPLC-MS, ArcT reactions by MALDI-ToF-MS. To test ApeH activity with *holo*-ApeE and/or ApeF, the Sfp reaction with CoA was performed overnight, before adding ApeH.

Malonyl-ACP – Sfp-reaction. ACPs were buffer exchanged in reaction buffer (10 mM Tris, 10 mM DTT, pH 8) with PD SpinTrap G-25 columns (Amersham Biosciences) and reactions were performed using 50 μM ACP, 1 mM Malonyl-CoA, 0.5 μM Sfp-SUMO (in storage buffer), 5 mM MgCl₂, 1 mM acid-substrate (from 10 mM stock solution with

ethanol as solvent) by an overnight incubation at 30°C. Reactions were analyzed by HR-HPLC-MS.

Self-malonylation reaction. To test whether *holo*-ApeE or ApeF were self-malonylated, a Sfp reaction with CoA was performed overnight and subsequently buffer exchanged in reaction buffer using Amicon Ultra-0.5 30K devices (Merck Millipore). Self-malonylation was tested by adding 5 mM Malonyl-CoA and overnight incubation at 30°C. Reactions were measured by MALDI-ToF-MS.

(Initial) chain elongation step by KS(-CLF) (ApeO:ApeC and/or ApeR, or ArcK/J)

Starter-ACP assay and Elongation-ACP assay were mixed 1:3 and chain elongation reaction was performed by adding 5-15 µM KS(-CLF) (in storage buffer) and further incubation for 3-5 h at 30°C. Reactions were analyzed by HR-HPLC-MS.

Ketoreduction step by KR (ApeQ)

Ketoreduction step was performed by adding 5-15 µM KR (in storage buffer) and 0.5 mM NADPH to the chain elongation reaction. The reaction was further incubated overnight at 30°C. Reactions were analyzed by HR-HPLC-MS.

APE-assay

APE-assay was performed by adding 5-15 µM DHs (ApeI:ApeP) to ketoreductase reaction and further overnight incubation at 30 °C with 4HB-ACP as starter-ACP. In the case of the APE assays “ApeO:ApeC”, “ApeR”, “ApeO:ApeC and ApeR” the previous chain elongation step was only performed with the named KS(CLF), respectively. To overcome complexity during HR-HPLC-MS-analysis not only a control reaction without addition of DHs was performed, but also a fluoro-substrate control. The fluoro-substrate control reaction was conducted with 3F4HB-ACP instead of 4HB-ACP, resulting in a +18 mass shift during MS² analysis of corresponding Ppant ejection ions.¹⁷ All samples were analyzed by HR-HPLC-UV-MS. APE assays turned yellow and could be detected by their UV absorbance at 420 nm.

Phylogenetic analysis

Essential proteins involved in APE (ApeC/E/F/I/O/P/Q/R) biosynthesis were identified by finding matching orthologs from all sequenced *Photorhabdus* and *Xenorhabdus* genomes using the orthologous matrix stand-alone (v1.0.6).¹⁸ The phylogeny was constructed by using core genes from all *Photorhabdus* and *Xenorhabdus* genomes

as described by Tobias *et al.* (2017).¹⁹ The phylogenetic tree was visualized, pruned and annotated using iTOL v2.^{20,21}

Supplementary Tables

Table S1. Strains used in this study.

Strain	Genotype	Reference
<i>E. coli</i> DH10B	F- <i>araDJ39</i> Δ (<i>ara</i> , <i>leu</i>)7697 Δ <i>lacX74 galU galK rpsL deoR</i> ϕ 80 <i>dlacZ</i> Δ M15 <i>endA1 nupG recA1 mcrA</i> Δ (<i>mrr hsdRMS mcrBC</i>)	Invitrogen
<i>E. coli</i> BL21 (DE3) Star	F ⁻ , <i>ompT, gal, dcm, hsdSB</i> (<i>r_B- m_B-</i>), <i>lon</i> , λ (<i>DE3 [lacI, lacUV5-T7, gene1, ind1, sam7, nin5]</i>)	Invitrogen
<i>E. coli</i> BL21-Gold (DE3) Star	<i>E. coli</i> B F- <i>dcm+</i> Hte <i>ompT hsdS</i> (<i>r_B- m_B-</i>) <i>gal</i> λ (DE3) <i>endA Tet</i> ^r	Merck/Novagen
<i>E. coli</i> ST18- λ pir	Tp ^r Sm ^r <i>recA thi hsdR</i> RP4-2-Tc::Mu-Km::Tn7, λ pir phage lysogen, Δ <i>hemA</i>	²²
NDR48 (<i>E. coli</i> BL21 Δ <i>acpT</i>)	F- <i>hsdS_B</i> (<i>r_B- m_B-</i>) (λ DE3) Δ <i>acpT</i>	²³
<i>Xenorhabdus doucetiae</i> DSM17909	Wild type, amp ^r	DSMZ
<i>Xenorhabdus doucetiae</i> DSM17909 Δ <i>xrdE</i>	DSM17909 wild type with a deletion in <i>xrdE</i>	²⁴
<i>Xenorhabdus doucetiae</i> DSM17909 Δ <i>xrdE</i> _P _{BAD} <i>apeA</i>	<i>Xenorhabdus doucetiae</i> DSM17909 Δ <i>xrdE</i> with a promoter exchange in front of <i>apeA</i> ; amp ^r , kan ^r	this work
<i>Xenorhabdus doucetiae</i> DSM17909 Δ <i>xrdE</i> _P _{BAD} <i>apeB</i>	<i>Xenorhabdus doucetiae</i> DSM17909 Δ <i>xrdE</i> with a promoter exchange in front of <i>apeB</i> ; amp ^r , kan ^r	this work
<i>Xenorhabdus doucetiae</i> DSM17909 Δ DC Δ <i>hfq</i>	DSM17909 wild type with a deletion in XDD1_RS09835 (decarboxylase) and <i>hfq</i> ; amp ^r ,	²⁴
<i>Xenorhabdus doucetiae</i> DSM17909 Δ DC Δ <i>hfq</i> P _{BAD} <i>apeB</i>	<i>Xenorhabdus doucetiae</i> DSM17909 Δ DC Δ <i>hfq</i> with a markerless promoter exchange in front of <i>apeB</i> , amp ^r ,	this work
<i>Azoarcus</i> sp. BH72	Wild type	DSMZ

Table S2. Oligonucleotides used for plasmid construction. Overhangs are underlined.

Plasmid	Oligonucleotide 5' to 3'	Template	
pAL03_ape_mP	GG135	<u>TTTGGGCTAACAGGAGGAATTC</u> ATGACCTCACCC TATAATTATGAC	<i>X. doucetiae</i> DSM17909
	GG136	CAATTTGTGGAAATTC <u>CCCGGAGAGCTCTGATCTTC</u> GGAAAAACAATCC	
	GG50	GGAAATTCCTCCTGTTAGCC	pAL03
	GG23	GAGCTCTCCCGGAATTCC	
	GG139	<u>TGATGGGTAAAAAGGATCGATCCTTAATCGCCACT</u> TGAATCCTC	<i>X. doucetiae</i> DSM17909
	GG140	CCGTCAAGTTGCATAATTGTCTAGATATCCATATA <u>TCCGATAGGTAGC</u>	
	GG137	CTAGACAATTATGACAACTTGACGG	pAL03_LF
	GG138	AGGATCGATCCTTTTAACCCATC	
pCEP_Km_142/143	TS142	<u>TTTGGGCTAACAGGAGGCTAGCATATGAAAACTTC</u> ACATAAGC	<i>X. doucetiae</i> DSM17909
	TS143	<u>TCTGCAGAGCTCGAGCATGCACATATAGGTAGCAT</u> ATAGCCAG	
pCEP_Km_144/145	TS144	<u>TTTGGGCTAACAGGAGGCTAGCATATGACCTCACCC</u> CTATAATTATGAC	<i>X. doucetiae</i> DSM17909
	TS145	<u>TCTGCAGAGCTCGAGCATGCACATGCACTAAAGGC</u> GACATCAG	
pCDF_XdACP1_Strep	GG64	<u>AAGAGAACCTATACTTCCAGGGACAAGAGTTATATA</u> TTGAAATAAAACG	<i>X. doucetiae</i> DSM17909
	GG65	<u>GGTGGCAGCAGCCTAGGTTAATTATGCTTTCTGTG</u> ATTGG	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pCDF Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCDF_XdACP2_Strep	GG66	<u>AAGAGAACCTATACTTCCAGGGAGAAAAACAAAAG</u> ATTTTTCAGG	<i>X. doucetiae</i> DSM17909
	GG67	<u>GGTGGCAGCAGCCTAGGTTAATTATTGTTCTTTCAA</u> CAACTG	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pCDF Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCOLA_XdAasS_Strep	GG94	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGAGCTGG</u> TACGCCAACGGGT	<i>X. doucetiae</i> DSM17909
	GG95	<u>AGCGGTGGCAGCAGCCTAGGTTAATCAATGGGTTT</u> CATTAAATAGCTCC	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pCOLA Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCDF_XdDH2_Strep_DH1	GG114	<u>GAGAACCCTATACTTCCAGGGACCTGACTATTTACCC</u> GTGG	<i>X. doucetiae</i> DSM17909
	GG115	<u>GATTACTTTCTGTTTCGACTTAAGCATTATGCTACTC</u> CATCCTGTTTATTATG	
	GG107	TAATGCTTAAGTCGAACAGAAAG	pCDF Duet-1
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	
	GG116	<u>GTAAAGTATAAGAAGGAGATATACATATGAAACCCA</u> TTGAAATTCACC	<i>X. doucetiae</i> DSM17909
	GG117	<u>GTGGCAGCAGCCTAGGTTAATTAACGACATAACTTT</u> ATCTTTCCTTG	
	GG109	TAATTAACCTAGGCTGCTGCCAC	pCDF XdDH2_Strep
	GG108	CATATGTATATCTCCTTCTTATACTTAAC	
pACYC_CLF_Strep	GG88	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGAAAATTG</u> ACATTGGATATTACAGAC	<i>X. doucetiae</i> DSM17909
	GG89	<u>AGCGGTGGCAGCAGCCTAGGTTAATTATCGGCTCC</u> AATTCCACT	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pACYC Duet-1
	SW1	TTAACCTAGGCTGCTG	
pACYC_CLF_His	GG15	<u>ATCACACAGCCAGGATCCGAAATTGACATTGGAT</u> ATTACAGAC	<i>X. doucetiae</i> DSM17909
	GG16	<u>CAGCGGTTTCTTACCAGACTTATCGGCTCCAATTC</u> CACT	

pCDF_KS1_His	GG11	<u>ATCACCACAGCCAGGATCCGATTTATATTTCTGCCG TAGG</u>	<i>X. doucetiae</i> DSM17909
	GG12	<u>CAGCGGTTTCTTTACCAGACTTAAGCCACTCCCAAT ATCAG</u>	
pCOLA_XDDH1_ Strep	GG71	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGAAAACCC ATTGAAATTCACCG</u>	<i>X. doucetiae</i> DSM17909
	GG72	<u>AGCGGTGGCAGCAGCCTAGGTTAATCAACGACATA ACTTTATCTTTCC</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA AC</u>	pCOLA Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCOLA_XDDH2_ Strep	GG73	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGACCTGA CTATTTACCCGTG</u>	<i>X. doucetiae</i> DSM17909
	GG74	<u>AGCGGTGGCAGCAGCCTAGGTTAATTATGCTACTC CATCCTG</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA AC</u>	pCOLA Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pACYC_XdKS1_ Strep_CLF	GG110	<u>GAGAACCTATACTTCCAGGGAATTTATATTTCTGCC GTAGGTATGC</u>	<i>X. doucetiae</i> DSM17909
	GG111	<u>GATTACTTTCTGTTCCGACTTAAGCATTAGCCACTC CCAATATCAGG</u>	
	GG107	<u>TAATGCTTAAGTCGAACAGAAAG</u>	pACYC Duet-1
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA AC</u>	
	GG112	<u>GTTAAGTATAAGAAGGAGATATACATATGAAATTGA CATTGGATATTACAG</u>	<i>X. doucetiae</i> DSM17909
	GG113	<u>GTGGCAGCAGCCTAGGTTAATTATCGGCTCCAATT CCACTG</u>	
	GG109	<u>TAATTAACCTAGGCTGCTGCCAC</u>	pACYC XdKS1_ Strep
	GG108	<u>CATATGTATATCTCCTTCTTATACTTAAC</u>	
pCOLA_XdKS2_ Strep	GG86	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGAAAGCCAT AATGCGACATCCCA</u>	<i>X. doucetiae</i> DSM17909
	GG87	<u>AGCGGTGGCAGCAGCCTAGGTTAACTAAGCCCAAC GCCGATCA</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG GTGGCTCCACAT</u>	pCOLA Duet-1
	SW1	<u>GGTATATCTCCTTATTAAGTTAAAC</u>	
pACYC_XdKR_ Strep	GG74	<u>AGCGGTGGCAGCAGCCTAGGTTAATTATGCTACTC CATCCTG</u>	<i>X. doucetiae</i> DSM17909
	GG75	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGACGTTCA GTTCTCGTGACCG</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA AC</u>	pACYC Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCOLA_XdACP1_ Strep_ACP2	GG128	<u>GAGAACCTATACTTCCAGGGACAAGAGTTATATATT GAAATAAAACG</u>	<i>X. doucetiae</i> DSM17909
	GG129	<u>GATTACTTTCTGTTCCGACTTAAGCATTATGCTTTCTG TGATTGG</u>	
	GG107	<u>TAATGCTTAAGTCGAACAGAAAG</u>	pCOLA Duet-1
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGG GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA AC</u>	
	GG126	<u>GTTAAGTATAAGAAGGAGATATACATATGAAAAAC AAAAGATTTTTCAGG</u>	<i>X. doucetiae</i> DSM17909
	GG127	<u>GTGGCAGCAGCCTAGGTTAATTATTGTTCTTTCAAC AACTGTTT</u>	
	GG109	<u>TAATTAACCTAGGCTGCTGCCAC</u>	pCOLA_XdACP1_ Strep
	GG108	<u>CATATGTATATCTCCTTCTTATACTTAAC</u>	
pCOLA_XdACP1_ Strep_TE	GG130	<u>GTTAAGTATAAGAAGGAGATATACATATGCTGACTG ATCCTCGTTTTACTGCTG</u>	<i>X. doucetiae</i> DSM17909
	GG131	<u>GTGGCAGCAGCCTAGGTTAATTATCACGGCTTGAC CCCATG</u>	
	GG109	<u>TAATTAACCTAGGCTGCTGCCAC</u>	pCOLA_XdACP1_ Strep
	GG108	<u>CATATGTATATCTCCTTCTTATACTTAAC</u>	
pCOLA_XdACP2_ Strep_TE	GG66	<u>AAGAGAACCTATACTTCCAGGGAGAAAAACAAAAG ATTTTTCAGG</u>	<i>X. doucetiae</i> DSM17909
	GG134	<u>GATTACTTTCTGTTCCGACTTAAGCATTATTGTTCTTT CAACAACCTG</u>	

	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pCOLA_XdACP1_ Strep_TE
	GG107	TAATGCTTAAGTCGAACAGAAAG	
pCOLA_TE_Strep	GG77	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGACTGACT</u> <u>GATCCTCGTTTTACTGCTG</u>	<i>X. doucetiae</i> DSM17909
	GG78	<u>AGCGGTGGCAGCAGCCTAGGTTAATCACGGCTTGA</u> CCCCATG	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pCOLA Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCDF_XdMT_Strep	GG79	<u>TTCGAAAAAGAGAACCTATACTTCCAGGGAACCTCA</u> <u>CCCTATAATTATGACAAAG</u>	<i>X. doucetiae</i> DSM17909
	GG80	<u>AGCGGTGGCAGCAGCCTAGGTTAATTACTGCTTCT</u> TGCAGATCAG	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTCGAACTGCGG</u> <u>GTGGCTCCACATGGTATATCTCCTTATTAAGTTAA</u> AC	pCDF Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCATI_arcV	TS_azo0262	<u>GATCCCATGGCAACCCACACCGCCCCCGCATC</u>	<i>Azoarcus</i> sp. BH72
	TS_azo0262	<u>CTAGGAGCTCCTGCACCCGGTGGGTGGCGATG</u>	
pCOLA_azo3912 _His_azo3911	TS166	<u>ATCACCCACAGCCAGGATCCGACCCAGCCTTCACAC</u> CCGG	<i>Azoarcus</i> sp. BH72
	TS164	<u>TTAAGCATTATGCGGCCGCATCATGCCTGCAGGTC</u> GAGC	
	TS173	<u>GTATAAGAAGGAGATATACATATGAAGCCGCTTCAG</u> CAGC	<i>Azoarcus</i> sp. BH72
	TS174	<u>TTTCTTTACCAGACTCGAGGTCATCCGGCGCGCGA</u> GAAC	

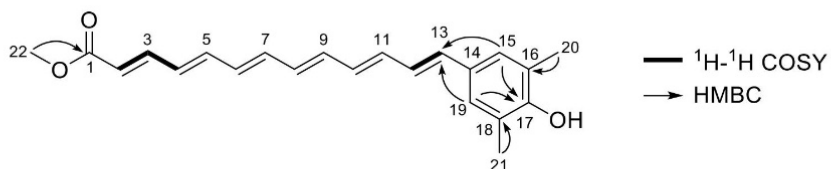
3. Plasmids used in this study.

Plasmid	Genotype	Reference
pACYC Duet-1	P15A ori, T7lac promoter, Cm ^r	Novagen
pCDF Duet-1	CDF ori, T7lac promoter, Sm ^r	Novagen
pCOLA Duet-1	ColA ori, T7lac promoter, Km ^r	Novagen
pCAT11	ColA ori, T7lac promoter, <i>cherry</i> , Km ^r	²⁵
pSUMO	ColE1 ori, T7lac promoter, <i>sumo</i> , Km ^r	²⁶
pCEP_Km	R6Kγ ori, oriT, <i>araC</i> , <i>araBAD</i> promoter, Km ^r	²
pAL03	R6Kγ ori, oriT, <i>sacB</i> , <i>araC</i> , <i>araBAD</i> , Km ^r	²
pAL03_ape_mP	markerless promoter exchange plasmid based on pAL03 with 800 bp of up- and downstream region of <i>apeB</i> startcodon, Km ^r	this work
pCEP_Km_142/143	promoter exchange plasmid based on pCEP_Km with 487 bp of downstream region of <i>apeA</i> startcodon, Km ^r	this work
pCEP_Km_144/145	promoter exchange plasmid based on pCEP_Km with 518 bp of downstream region of <i>apeB</i> startcodon, Km ^r	this work
pCDF_XdACP1_Strep	CDF ori, T7lac promoter, <i>apeE</i> , Sm ^r	this work
pCDF_XdACP2_Strep	CDF ori, T7lac promoter, <i>apeF</i> , Sm ^r	this work
pCOLA_XdAasS_Strep	ColA ori, T7lac promoter, <i>apeH</i> , Km ^r	this work
pCDF_XdDH2_Strep_DH1	CDF ori, 2xT7lac promoter, <i>apeP</i> , <i>apeI</i> , Sm ^r	this work
pACYC_CLF_Strep	P15A ori, T7lac promoter, <i>apeC</i> , Cm ^r	this work
pACYC_CLF_His	P15A ori, T7lac promoter, <i>apeC</i> , Cm ^r	this work
pCDF_KS1_His	CDF ori, T7lac promoter, <i>apeO</i> , Sm ^r	this work
pCOLA_DH1_Strep	ColA ori, T7lac promoter, <i>apeI</i> , Km ^r	this work
pCOLA_DH2_Strep	ColA ori, T7lac promoter, <i>apeP</i> , Km ^r	this work
pACYC_XdKS1_Strep_CLF	P15A ori, 2xT7lac promoter, <i>apeO</i> , <i>apeC</i> , Cm ^r	this work
pCOLA_XdKS2_Strep	ColA ori, T7lac promoter, <i>apeR</i> , Km ^r	this work
pACYC_XdKR_Strep	P15A ori, T7lac promoter, <i>apeQ</i> , Cm ^r	this work
pCOLA_XdACP1_Strep_ACP2	ColA ori, 2xT7lac promoter, <i>apeE</i> , <i>apeF</i> , Km ^r	this work
pCOLA_XdACP1_Strep_TE	ColA ori, 2xT7lac promoter, <i>apeE</i> , <i>apeK</i> , Km ^r	this work
pCOLA_XdACP2_Strep_TE	ColA ori, 2xT7lac promoter, <i>apeF</i> , <i>apeK</i> , Km ^r	this work
pCOLA_TE_Strep	ColA ori, T7lac promoter, <i>apeK</i> , Km ^r	this work
pCDF_XdMT_Strep	CDF ori, T7lac promoter, <i>apeB</i> , Sm ^r	this work
pCATI_arcV	ColA ori, T7lac promoter, <i>arcV</i> , <i>cherry</i> , Km ^r	this work
pCATI_arcT	ColA ori, T7lac promoter, <i>arcT</i> , <i>cherry</i> , Km ^r	²⁵
pCOLA_arcJ_His_arcK	ColA ori, T7lac promoter, <i>arcJ</i> , <i>arcK</i> , Km ^r	this work
pSUMO_sfp	ColE1 ori, T7lac promoter, <i>sumo</i> , <i>sfp</i> , Km ^r	²⁶

Table S4. Homology of different strains with *ape* genes. The tblastn search was performed with Geneious 6.1.8. “Grade is a percentage calculated by Geneious by combining the query coverage, e-value and identity values”, for more details see the Geneious User Manual.

<i>X. doucetiae</i> DSM17909	<i>Azoarcus</i> sp. BH72		<i>E. coli</i> CFT073		<i>E. coli</i> O157:H7 str. Sakai	
protein name (locus tag)	protein name (locus tag)	Grade %	locus tag	Grade %	locus tag	Grade %
ApeA (<i>xdd_3458</i>)	-	-	c1204	73	z4849	75
ApeB (<i>xdd_3459</i>)	-	-	c1203	83	z4850	81
ApeC (<i>xdd_3460</i>)	ArcJ (<i>azo3912</i>)	31	c1202	80	z4851	79
ApeD (<i>xdd_3461</i>)	ArcD (<i>3918</i>)	69	c1201	74	z4852	74
ApeE (<i>xdd_3462</i>)	ArcV (<i>0262</i>)	43	c1200	89	z4853	90
ApeF (<i>xdd_3463</i>)	-	-	c1199	87	z4854	87
ApeG (<i>xdd_3464</i>)	ArcU (<i>0261</i>)	24	c1198	61	z4855	71
ApeH (<i>xdd_3465</i>)	ArcT (<i>0260</i>)	66	c1197	74	z4856	75
ApeI (<i>xdd_3466</i>)	ArcS (<i>0259</i>)	29	c1196	79	z4857	78
ApeJ (<i>xdd_3467</i>)	ArcE,ArcR (<i>azo3917,0258</i>)	46, 52	c1194	82	z4858	82
ApeK (<i>xdd_3468</i>)	-	-	c1193	88	z4859	88
ApeL (<i>xdd_3469</i>)	ArcQ (<i>azo0257</i>)	35	c1192	77	z4860	77
ApeM (<i>xdd_3470</i>)	ArcP (<i>azo0256</i>)	74	c1191	80	z4861	81
ApeN (<i>xdd_3471</i>)	-	-	c1190	76	z4862	77
ApeO (<i>xdd_3472</i>)	ArcK (<i>azo3911</i>)	67	c1189	84	z4863	84
ApeP (<i>xdd_3473</i>)	ArcI (<i>azo3913</i>)	39	c1188	72	z4864	72
ApeQ (<i>xdd_3474</i>)	ArcH (<i>azo3914</i>)	75	c1187	91	z4865	91
ApeR (<i>xdd_3475</i>)	ArcK (<i>azo3911</i>)	55	c1186	90	z4866	90

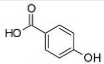
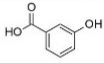
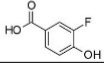
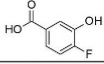
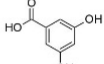
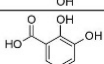
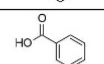
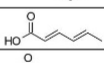
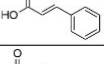
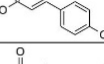
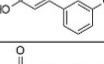
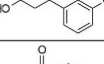
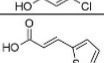
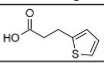
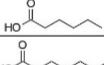
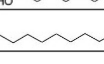


Table S5. ^1H (500 MHz) and ^{13}C (125 MHz) NMR data assignments for APE 1
DMSO- d_6 .



no.	δ_{H} (mult., J)	δ_{C}
1	-	167.1
2	5.98 (d, 15.1)	120.0
3	7.31 (dd, 15.0, 11.6)	145.3
4	6.39 (ov)	130.2
5	6.49 (ov)	132.4-136.7
6	6.49 (ov)	132.4-136.7
7	6.49 (ov)	132.4-136.7
8	6.49 (ov)	132.4-136.7
9	6.49 (ov)	132.4-136.7
10	6.49 (ov)	132.4-136.7
11	6.49 (ov)	132.4-136.7
12	6.81 (m)	132.4-136.7
13	6.52 (ov)	134.5
14	-	n.d.
15	7.07 (s)	127.3
16	-	125.0
17	-	154.0
18	-	125.0
19	7.07 (s)	127.3
20	2.15 (s)	17.2
21	2.15 (s)	17.2
22	3.66 (s)	51.8

n.d. = undetected, ov = overlapping.

Table S6. Substrate specificity of ApeH by its AasS reaction with ApeE. The AasS reaction was performed with substrates **8-25** and analyzed via HR-HPLC-MS in positive mode with a Ppant ejection in MS².

substrate	name	structure	accepted
8	4-hydroxy benzoic acid		+
9	3-hydroxy benzoic acid		+
10	3-fluoro-4-hydroxy benzoic acid		+
11	4-fluoro-3-hydroxy benzoic acid		+
12	3,5-di-hydroxy benzoic acid		+
13	2,3-di-hydroxy benzoic acid		-
14	benzoic acid		+
15	sorbic acid		+
16	<i>trans</i> -cinnamic acid		+
17	<i>para</i> -coumaric acid		+
18	3-Chloro cinnamic acid		-
19	3-(3-Chlorophenyl)propionic acid		-
20	<i>trans</i> -3-Chloro acrylic acid		-
21	3-(2-Thienyl)acrylic acid		+
22	3-(2-Thienyl)propanoic acid		-
23	caproic acid		-
24	heptanoic acid		-
25	palmitic acid		-

e S7. Substrate specificity of ArcT from *Azoarcus* sp. BH72 by its AaS ree ArcV from *Azoarcus* sp. BH72. The AaS reaction was performed with subst , 16,17,21,22, 26-32 and analyzed via MALDI-MS.

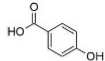
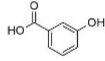
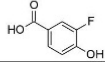
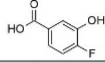
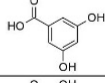
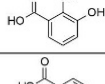
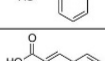
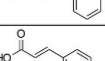
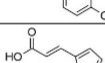
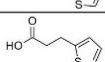
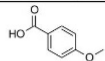
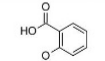
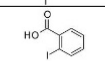
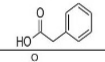
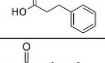
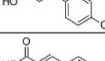
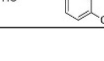

substrate	name	structure	accepted
8	4-hydroxy benzoic acid		+
9	3-hydroxy benzoic acid		+
10	3-fluoro-4-hydroxy benzoic acid		+
11	4-fluoro-3-hydroxy benzoic acid		+
12	3,5-di-hydroxy benzoic acid		+
13	2,3-di-hydroxy benzoic acid		+
14	benzoic acid		+
16	<i>trans</i> -cinnamic acid		-
17	<i>para</i> -coumaric acid		-
21	3-(2-Thienyl)acrylic acid		+
22	3-(2-Thienyl)propanoic acid		-
26	4-methoxy benzoic acid		-
27	2-methoxy benzoic acid		-
28	2-iodobenzoic acid		-
29	phenyl acetic acid		-
30	3-phenyl propionic acid		-
31	3-(4-chlorophenyl) propionic acid		-
32	4-chloro cinnamic acid		-

Table S8. Crystallographic data collection and refinement statistics.

	Apel/ApeP	ApeO/ApeC
<u>Crystal parameters</u>		
Space group	C2	P2 ₁
Cell constants	a= 144.7 Å b= 47.8 Å c= 42.7 Å	a= 57.0 Å b= 88.9 Å c= 63.0 Å
Subunits / AU ^a	1	1
Disordered regions	Apel: 101 - 107 ApeP: 1 - 9; 161 - 165	ApeO: - ApeC: 1
<u>Data collection</u>		
Beam line	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0
Resolution range (Å) ^b	30-1.85 (1.95 - 1.85)	30-1.45 (1.55 - 1.45)
No. observed reflections	68,219	358,419
No. unique reflections ^c	23,313	103673
Completeness (%) ^b	95.9 (97.6)	98.3 (96.9)
R _{merge} (%) ^{b, d}	6.9 (47.0)	4.3 (65.7)
I/σ (I) ^b	11.0 (2.7)	12.9 (2.0)
<u>Refinement (REFMAC5)</u>		
Resolution range (Å)	30-1.85	30-1.45
No. refl. working set	22,103	98,482
No. refl. test set	1,163	5,183
No. non hydrogen	2,244	5,186
Solvent (H ₂ O, PEG, glycerol)	106	328
R _{work} / R _{free} (%) ^e	19.8 / 23.6	12.8 / 16.2
r.m.s.d. bond (Å) / (°) ^f	0.007 / 1.29	0.006 / 1.34
Average B-factor (Å ²)	30.7	20.6
Ramachandran Plot (%) ^g	98.4 / 1.6 / 0	97.6 / 2.4 / 0
PDB accession code	6QSR	6QSP

^[a] Asymmetric unit

^[b] The values in parentheses for resolution range, completeness, R_{merge} and I/σ (I) correspond to the highest resolution shell

^[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections

^[d] $R_{\text{merge}}(I) = \frac{\sum_{\text{hkl}} \sum_j |I(\text{hkl})_j - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_j I(\text{hkl})_j}$, where $I(\text{hkl})_j$ is the j^{th} measurement of the intensity of reflection hkl and $\langle I(\text{hkl}) \rangle$ is the average intensity

^[e] $R = \frac{\sum_{\text{hkl}} (|F_{\text{obs}}| - |F_{\text{calc}}|)}{\sum_{\text{hkl}} |F_{\text{obs}}|}$, where R_{free} is calculated without a sigma cut off for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

^[f] Deviations from ideal bond lengths/angles

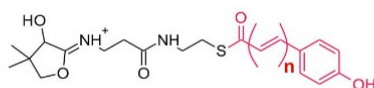
^[g] Number of residues in favored region / allowed region / outlier region

Table S9. KR reaction by ApeQ. The ACP-bound β -ketoacyl product was converted to the β -hydroxyacyl product by ApeQ with NADPH and a fluoro-substrate as control reaction. The MS-results for ApeE are displayed as well in Figure S26. The reaction was analyzed via HR-HPLC-MS in positive mode with a Ppant ejection in MS².

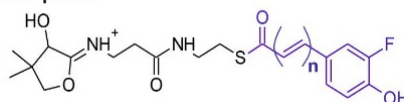
	MS ¹				MS ²			
	sum formula	theor. mass [*] [m/z]	det. mass [m/z]	Δ ppm	sum formula [M+H] ⁺	theor. mass [m/z]	det. mass [m/z]	Δ ppm
ApeE	C ₅₂₄ H ₈₁₃ N ₁₂₆ O ₁₆₉ PS ₄	1068.3500	1068.3513	1.2	C ₂₀ H ₂₉ N ₂ O ₆ S	425.1741	425.1721	4.7
	C ₅₂₄ H ₈₁₂ FN ₁₂₆ O ₁₆₉ PS ₄	1069.9855	1069.9866	1.0	C ₂₀ H ₂₈ FN ₂ O ₆ S	443.1647	443.1632	3.4
ApeF	C ₅₃₁ H ₈₃₈ N ₁₃₃ O ₁₆₇ PS ₃	1081.3733	1081.3720	1.2	C ₂₀ H ₂₉ N ₂ O ₆ S	425.1741	425.1705	8.5
	C ₅₃₁ H ₈₃₇ FN ₁₃₃ O ₁₆₇ PS ₃	1083.0088	1083.0079	0.8	C ₂₀ H ₂₈ FN ₂ O ₆ S	443.1647	443.1616	7.0

*theoretical mass of 10⁺charge state of average mass.

Table S10. Detected Ppant ejection ions in assay ApeO:ApeC, ApeO:ApeC and ApeR and ApeR.



n	sum formula [M+H] ⁺	theoretical mass [m/z]	ApeO:ApeC		ApeR		ApeO:ApeC + ApeR	
			detected mass [m/z]	Δ ppm	detected mass [m/z]	Δ ppm	detected mass [m/z]	Δ ppm
1	C ₂₀ H ₂₇ N ₂ O ₅ S	407.1635	n.d.	-	407.1657	4.6	407.1631	1.1
2	C ₂₂ H ₂₉ N ₂ O ₅ S	433.1792	n.d.	-	433.1811	4.5	433.1787	1.1
3	C ₂₄ H ₃₁ N ₂ O ₅ S	459.1948	n.d.	-	459.1972	5.2	459.1950	0.4
4	C ₂₆ H ₃₃ N ₂ O ₅ S	485.2105	n.d.	-	485.2127	4.6	485.2107	0.4
5	C ₂₈ H ₃₅ N ₂ O ₅ S	511.2261	n.d.	-	n.d.	-	511.2261	0.0
6	C ₃₀ H ₃₇ N ₂ O ₅ S	537.2418	n.d.	-	n.d.	-	537.2422	0.8
7	C ₃₂ H ₃₉ N ₂ O ₅ S	563.2574	563.2578	0.7	n.d.	-	563.2583	1.6
8	C ₃₄ H ₄₁ N ₂ O ₅ S	589.2731	n.d.	-	n.d.	-	589.2736	0.8
9	C ₃₆ H ₄₃ N ₂ O ₅ S	615.2887	n.d.	-	n.d.	-	615.2893	1.0
10	C ₃₈ H ₄₅ N ₂ O ₅ S	641.3044	n.d.	-	n.d.	-	641.3049	0.8
11	C ₄₀ H ₄₇ N ₂ O ₅ S	667.3200	n.d.	-	n.d.	-	667.3182	2.7

Table S11. Detected Ppant ejection ions in fluoro-control assay ApeO:ApeC, ApeO:ApeC and ApeR and ApeR.

n	sum formula [M+H] ⁺	theoretical mass [m/z]	ApeO:ApeC		ApeR		ApeO:ApeC + ApeR	
			detected mass [m/z]	Δppm	detected mass [m/z]	Δppm	detected mass [m/z]	Δppm
1	C ₂₀ H ₂₆ FN ₂ O ₅ S	425.1541	n.d.	-	n.d.	-	n.d.	-
2	C ₂₂ H ₂₈ FN ₂ O ₅ S	451.1697	n.d.	-	n.d.	-	451.1693	0.9
3	C ₂₄ H ₃₀ FN ₂ O ₅ S	477.1854	n.d.	-	n.d.	-	477.1848	1.2
4	C ₂₆ H ₃₂ FN ₂ O ₅ S	503.2010	n.d.	-	n.d.	-	503.2017	1.3
5	C ₂₈ H ₃₄ FN ₂ O ₅ S	529.2167	n.d.	-	n.d.	-	529.2166	0.1
6	C ₃₀ H ₃₆ FN ₂ O ₅ S	555.2323	n.d.	-	n.d.	-	555.2312	2.1
7	C ₃₂ H ₃₈ FN ₂ O ₅ S	581.2480	581.2472	1.4	n.d.	-	581.2485	0.9
8	C ₃₄ H ₄₀ FN ₂ O ₅ S	607.2636	n.d.	-	n.d.	-	607.2667	5.1
9	C ₃₆ H ₄₂ FN ₂ O ₅ S	633.2793	n.d.	-	n.d.	-	633.2803	1.6
10	C ₃₈ H ₄₄ FN ₂ O ₅ S	659.2949	n.d.	-	n.d.	-	659.2945	0.6
11	C ₄₀ H ₄₆ FN ₂ O ₅ S	685.3106	n.d.	-	n.d.	-	685.3128	3.2

Table S12. Amino acid sequences used for phylogenetic analysis of ApeO, ApeC and ApeR.²⁷

name	acc. number	strain
ActI-ORF1	CAC44200	<i>Streptomyces coelicolor</i> A3(2)
ActI-ORF2	CAC44201	<i>Streptomyces coelicolor</i> A3(2)
AknB	AAF70106	<i>Streptomyces galilaeus</i>
AknC	AAF70107	<i>Streptomyces galilaeus</i>
AknE2	AAF70109	<i>Streptomyces galilaeus</i>
AmphC	AJE44524	<i>Streptomyces nodosus</i>
AsuC13	ADI58650	<i>Streptomyces nodosus</i> subsp. <i>asukaensis</i>
AsuC14	ADI58649	<i>Streptomyces nodosus</i> subsp. <i>asukaensis</i>
AuaC	CAL48955	<i>Stigmatella aurantiaca</i> Sg a15
AuaD	CAL48956	<i>Stigmatella aurantiaca</i> Sg a15
AveA1_KS2	AAG09812	<i>Streptomyces avermitilis</i>
AviM_KS	AAK83194	<i>Streptomyces viridochromogenes</i> Tue57
BenA	CAM58798	<i>Streptomyces</i> sp. 2991200
BenB	CAM58799	<i>Streptomyces</i> sp. 2991200
BenQ	CAM58805	<i>Streptomyces</i> sp. A2991200
Cal30	ALG65306	<i>Streptomyces calvus</i>
Cal31	ALG65305	<i>Streptomyces calvus</i>
Cal32	ALG65304	<i>Streptomyces calvus</i>
Cal33	ALG65303	<i>Streptomyces calvus</i>
CalE8	AAM94794	<i>Micromonospora echinospora</i>
CalO5	AAM70355.1	<i>Micromonospora echinospora</i>
ChIA3	AAZ77696	<i>Streptomyces antibioticus</i>
ChIB1	AAZ77673	<i>Streptomyces antibioticus</i>
ChmGII_KS1	AAS79460	<i>Streptomyces bikiniensis</i>
CmmP	Q70J88	<i>Streptomyces griseus</i> subsp. <i>griseus</i>
ColC13	AIL50179	<i>Streptomyces aureus</i>
ColC14	AIL50180	<i>Streptomyces aureus</i>
ColC3	AIL50165	<i>Streptomyces aureus</i>
ColC4	AIL50166	<i>Streptomyces aureus</i>

DauA	Q55225	<i>Streptomyces</i> sp.
DHA ORF8 KS	BAA89382	<i>Moritella marina</i> ATCC 15381
DpgA	AAM80548	<i>Streptomyces toyocaensis</i>
DpsA	AAA65206	<i>Streptomyces peucetius</i>
DpsB	AAA65207	<i>Streptomyces peucetius</i>
DpsC	AAA65208	<i>Streptomyces peucetius</i>
DynE8_KS	ACB47048	<i>Micromonospora chersina</i>
Ec_FabB	EIQ69853	<i>Escherichia coli</i> EPEC C342-62
Ec_FabF	EGT67882	<i>Escherichia coli</i> O104:H4 str. C227-11
Ec_FabH	CDL30502	<i>Escherichia coli</i> ISC7
EncA	AIN46688	<i>Streptomyces qinglanensis</i>
EncB	AIN46689	<i>Streptomyces qinglanensis</i>
FscC_KS1	AAQ82564	<i>Streptomyces</i> sp. FR-008
FscD_KS1	AAQ82568	<i>Streptomyces</i> sp. FR-008
GdmAIII	AAO06918	<i>Streptomyces hygrosopicus</i>
Iga11	BAX64252	<i>Streptomyces</i> sp. MSC090213JE08
Iga12	BAX64253	<i>Streptomyces</i> sp. MSC090213JE08
JadA	AAB36562	<i>Streptomyces venezuelae</i> ATCC 10712
JadB	AAB36563	<i>Streptomyces venezuelae</i> ATCC 10712
KirAII_KS1	CAN89632.1	<i>Streptomyces collinus</i> Tu 365
LanA	AAD13536	<i>Streptomyces cyanogenus</i>
LanB	AAD13537	<i>Streptomyces cyanogenus</i>
LipPks2	ABB05103	<i>Streptomyces aureofaciens</i>
LipPks3	ABB05104	<i>Streptomyces aureofaciens</i>
LipPks4	ABB05105	<i>Streptomyces aureofaciens</i>
MdpB	ABY66019	<i>Actinomadura madurae</i>
MdpE_KS	AAQ17110	<i>Actinomadura madurae</i>
MerA_KSQ	ABJ97437	<i>Streptomyces violaceusniger</i>
MerC_KSQ	ABJ97439	<i>Streptomyces violaceusniger</i>
MonAI_KSQ	AAO65796	<i>Streptomyces cinnamonensis</i>
MtmK	CAA61990	<i>Streptomyces argillaceus</i>
MtmP	CAA61989	<i>Streptomyces argillaceus</i>
NcsB_KS	AAM77986	<i>Streptomyces carzinostaticus</i> subsp. neocarzinostaticus
NcsE_KS1	Q83TF5	<i>Streptomyces carzinostaticus</i> subsp. neocarzinostaticus
NidA2_KS1	AAC46024	<i>Streptomyces caelestis</i>
NysC_KS1	AAF71776	<i>Streptomyces noursei</i> ATCC 11455
OrfA_KS	AAK72879	<i>Schizochytrium</i> sp.
OxyA	AAZ78325	<i>Streptomyces rimosus</i>
OxyB	AAZ78326	<i>Streptomyces rimosus</i>
OxyD	AAZ78328	<i>Streptomyces rimosus</i>
OzmN	ABS90475	<i>Streptomyces albus</i>
Pa_FabF	4B7V	<i>Pseudomonas aeruginosa</i> PAO1
PctS_KS1	BAF92601	<i>Streptomyces pactum</i>
PfaA_KS	AAL01060	<i>Photobacterium profundum</i> SS9
PikAI_KS1	AAC69329	<i>Streptomyces venezuelae</i>
PikAI_KSQ	AAC69329	<i>Streptomyces venezuelae</i>
PimS3_KS1	CAC20920	<i>Streptomyces natalensis</i>
PksE	AAO25894.1	<i>Streptomyces macromomyceticus</i>
PokM1_KS	ACN64831	<i>Streptomyces diastatochromogenes</i>
RapB_KS1	CAA60459	<i>Streptomyces rapamycinicus</i> NRRL 5491
Res7	BAU09323	<i>Streptomyces roseovercillatus</i>
Sa_FabF	WP_006605824	<i>Streptomyces auratus</i>
Sc_FabF	WP_011028323	<i>Streptomyces</i> sp.
Sc_FabH	CAB62720	<i>Streptomyces coelicolor</i> A3(2)
SgcE_KS1	Q8GME1	<i>Streptomyces globisporus</i>
Sim-ORF2	AEU17884	<i>Streptomyces antibioticus</i>
Sim-ORF3	AEU17885	<i>Streptomyces antibioticus</i>
SimX5	AEU17889	<i>Streptomyces antibioticus</i>
Sky17	AEA30260	<i>Streptomyces</i> sp. Acta 2897
Sky18	AEA30261	<i>Streptomyces</i> sp. Acta 2897
Sky19	AEA30262	<i>Streptomyces</i> sp. Acta 2897
Sky22	AEA30265	<i>Streptomyces</i> sp. Acta 2897
SmcKSI	ALT05933.1	<i>Kitasatospora</i> sp. 152608
SmcKSII	ALT05934.1	<i>Kitasatospora</i> sp. 152608
SmcX5	ALT05939.1	<i>Kitasatospora</i> sp. 152608
Snoa1	Q54495	<i>Streptomyces nogalater</i>
Snoa2	Q54496	<i>Streptomyces nogalater</i>

Sp_FabF	AHA90851	<i>Streptomyces platensis</i>
Sv_FabF	WP_015033414	<i>Streptomyces venezuelae</i>
TcmK	AAA67515	<i>Streptomyces glaucescens</i>
TcmL	AAA67516	<i>Streptomyces glaucescens</i>
UrdA	Q54173	<i>Streptomyces fradiae</i>
VirA_KS1	BAF50727	<i>Streptomyces virginiae</i>
WhiE_CLF	CAB45607	<i>Streptomyces coelicolor</i> A3(2)
WhiE_KS	CAB45606	<i>Streptomyces coelicolor</i> A3(2)
ZhuA	AAG30188	<i>Streptomyces</i> sp. R1128
ZhuB	AAG30189	<i>Streptomyces</i> sp. R1128
ZhuH	AAG30195	<i>Streptomyces</i> sp. R1128

Supplementary Figures

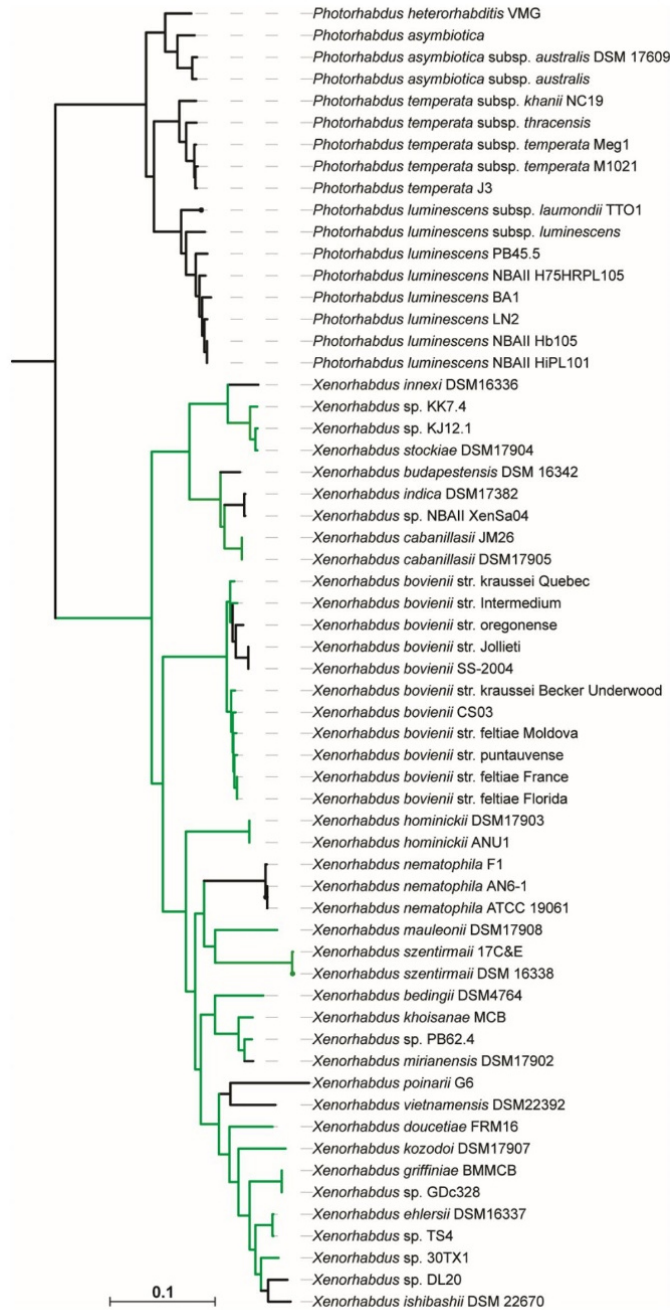


Figure S1. Phylogenetic distribution of *ape* gene cluster in *Xenorhabdus* and *Photorhabdus* sp. All strains carrying the *ape* gene cluster are displayed in green.

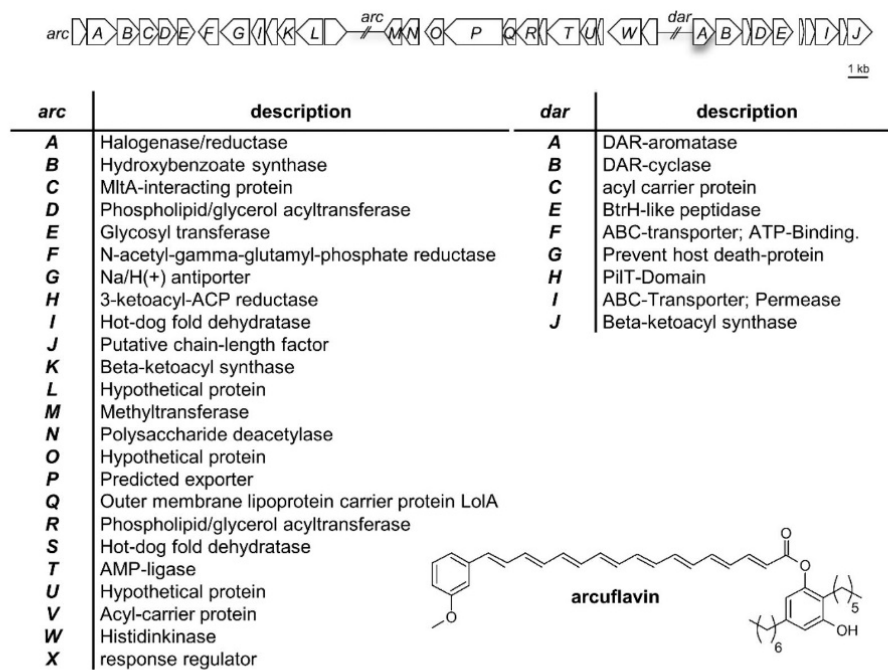


Figure S2. Arcuflavin gene cluster from *Azoarcus* sp. BH72 involved in biosynthesis of arcuflavin.²⁵

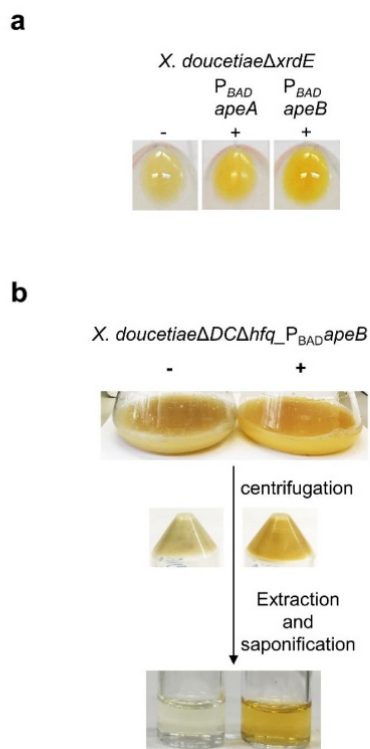


Figure S3. Comparison of arabinose induced (+) and non-induced (-) *X. doucetiae* *ape* overproduction strains. Promoter exchange in front of *apeB* results in stronger pigmentation compared to promoter exchange in front of *apeA* (a). General extraction procedure to generate APE methyl esters from crude extract (b).

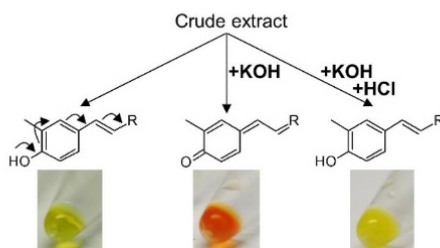


Figure S4. KOH test²⁸ of crude extract from *X. doucetiae*Δ*DC*Δ*hfq*_ P_{BAD} *apeB* overproduction strain.

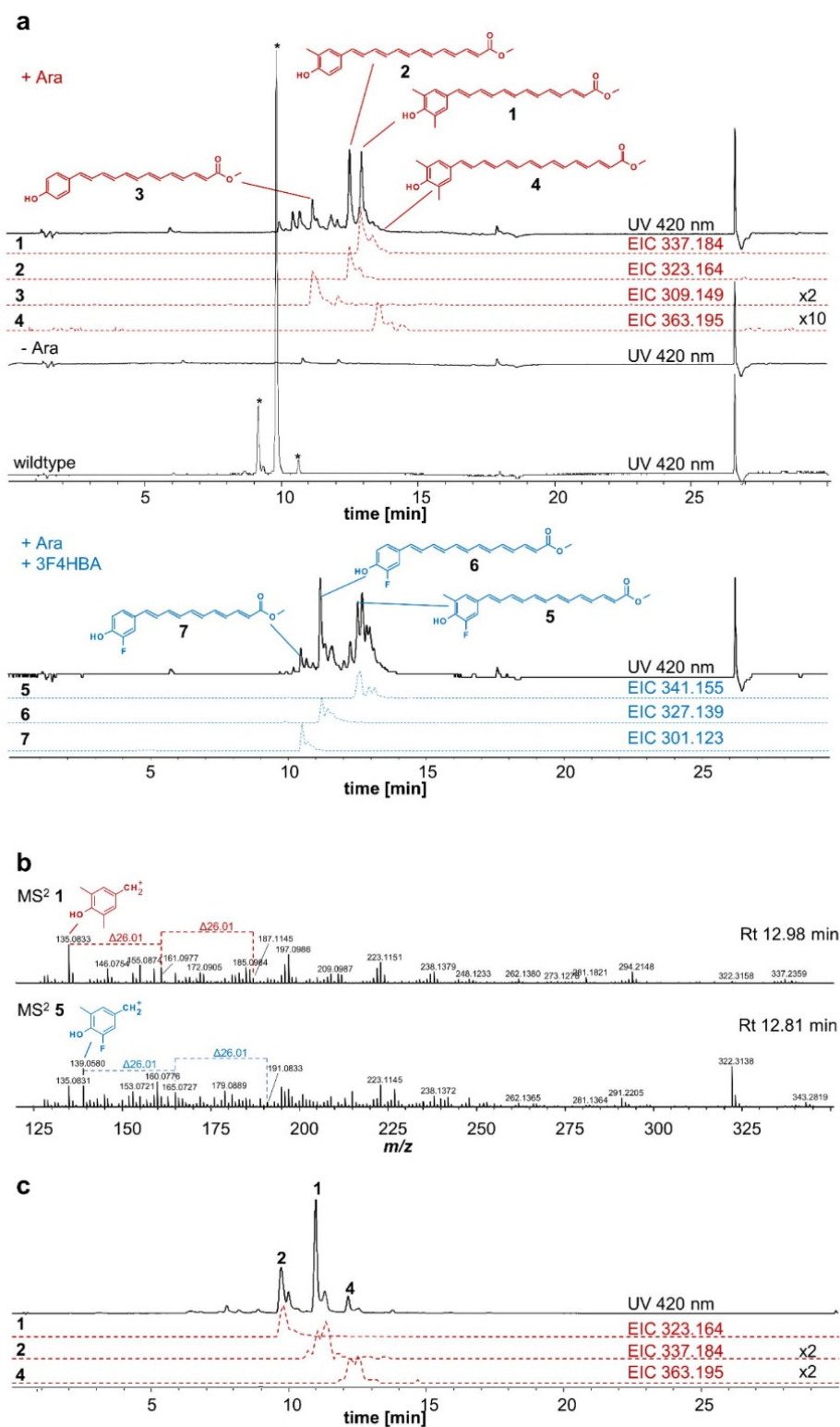


Figure S5. HR-HPLC-UV-MS analysis (positive mode) of *X. doucetiae*ΔDCΔhfq₋ P_{BADapeB} and wildtype extracts at 420 nm. Comparison of crude extracts of wildtype

and noninduced and induced promoter exchange mutant (+0.2% arabinose) vs. induced plus fed mutant (+0.2% arabinose, +1 mM 3F4HBA). EICs ($[M+H]^+ \pm 0.005$ Da) of assigned APE methyl ester obtained after methanolysis and extraction are displayed in dashed lines. Signals, marked with an asterisk represent previously described xenorhabdin derivatives only produced in the wildtype²⁴ (**a**). MS² spectra of compound **1** vs. compound **5** showing the incorporation of 3F4HBA as a precursor (**b**). Distribution between APE compounds **1**, **2** and **4** differs upon cultivation. Here, the 40-100% ACN gradient was used (see “HPLC-MS” section), explaining the retention time shifts compared to chromatogram in a) (**c**).

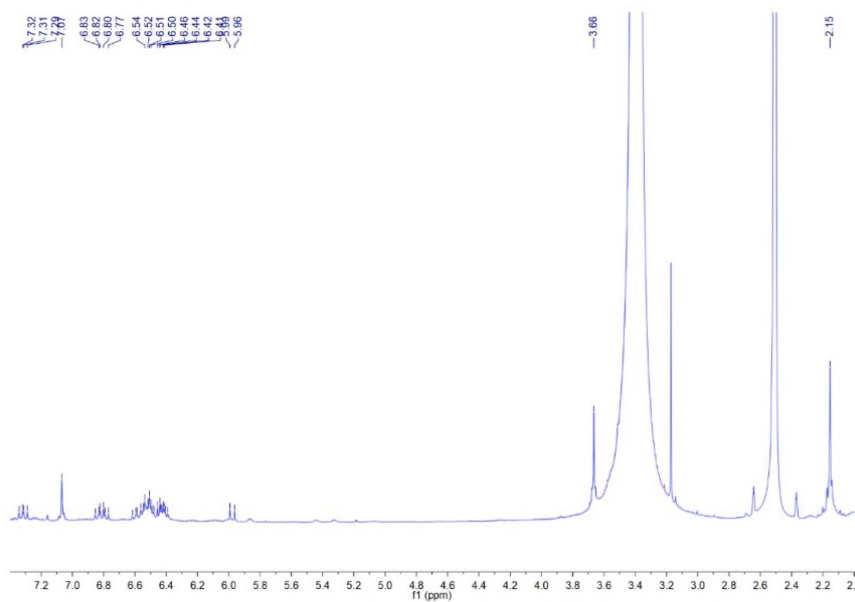


Figure S6. ^1H NMR spectrum of **1** in $\text{DMSO-}d_6$.

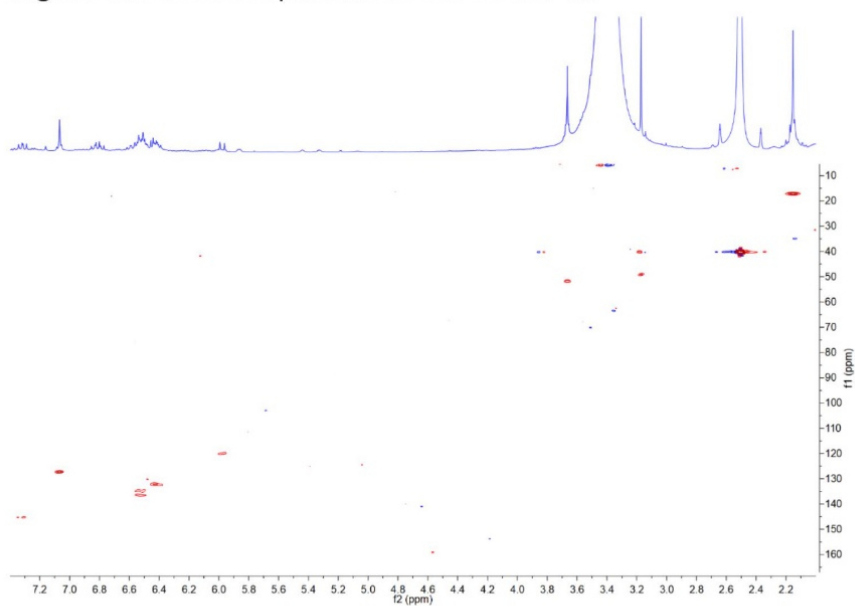


Figure S7. HSQC spectrum of **1** in $\text{DMSO-}d_6$.

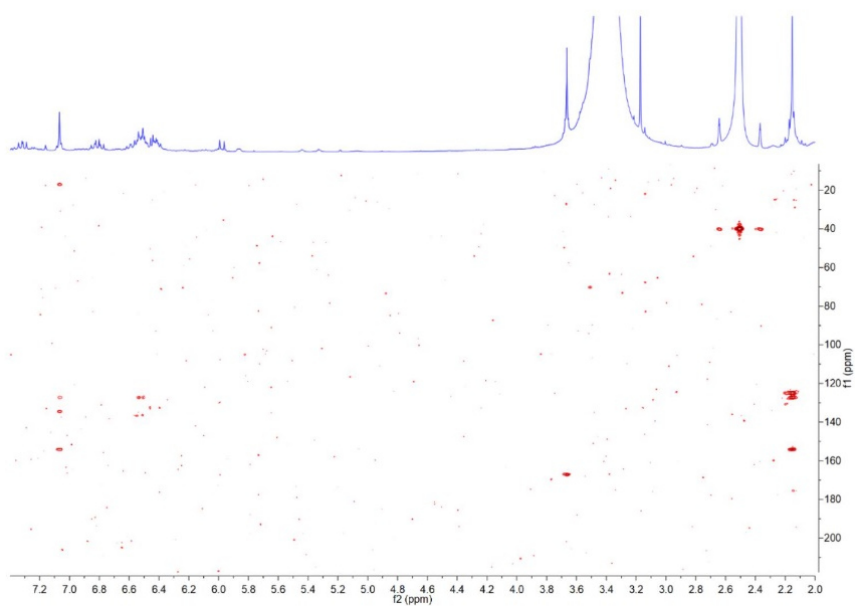


Figure S8. HMBC spectrum of **1** in DMSO-*d*₆.

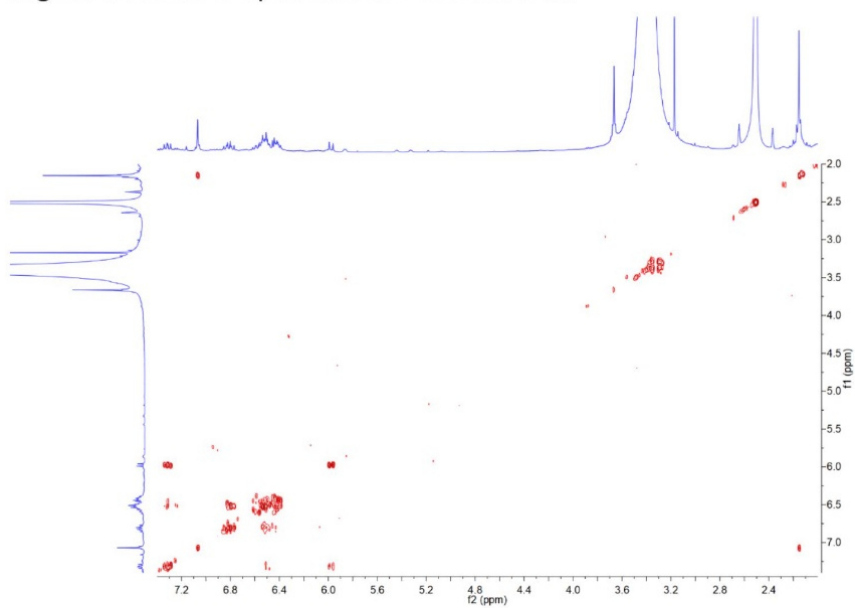


Figure S9. ¹H-¹H COSY spectrum of **1** in DMSO-*d*₆.

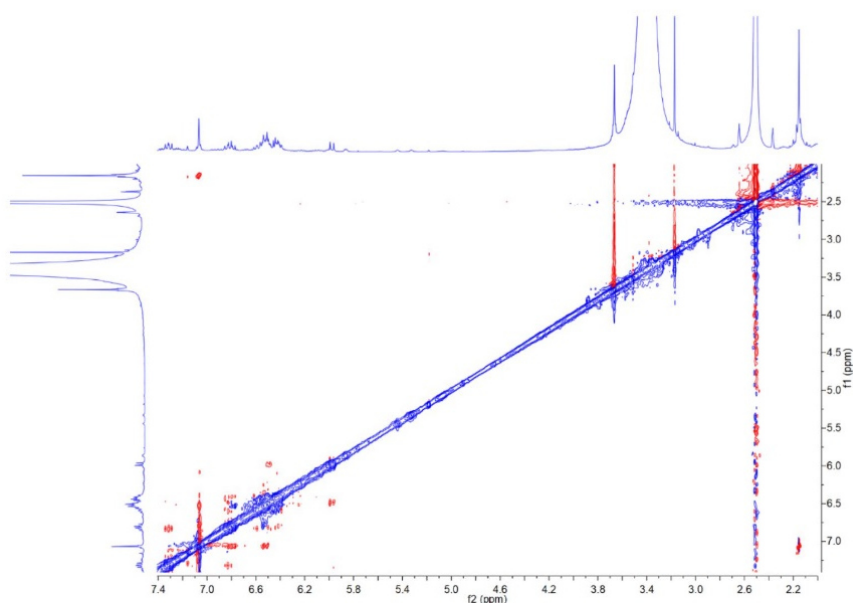


Figure S10. ROESY spectrum of **1** in DMSO- d_6 .

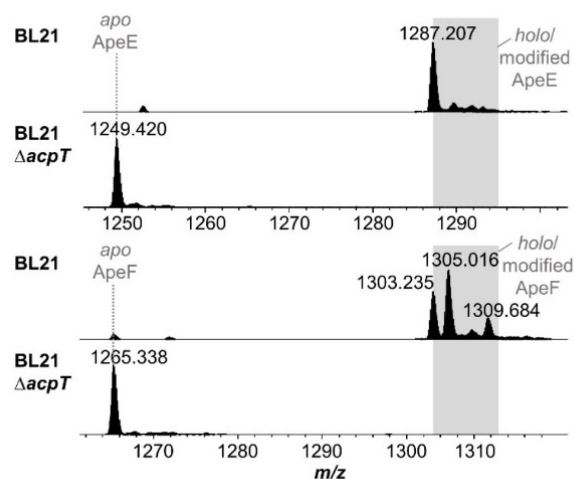


Figure S11. Post-translational protein modifications of the ACPs ApeE and ApeF by endogenous PPTase AcpT from *E. coli* BL21. Protein HR-HPLC-MS analysis of purified ApeE and ApeF after protein production in either BL21 or BL21 $\Delta acpT$. Displayed are the m/z -values of the average protein masses of the 9+ charge states (MS^1). m/z 1249.420 and 1265.338 corresponds to apo, m/z 1287.207 and 1303.235 to holo-ApeE and ApeF in MS^1 , respectively. m/z 1305.016 and 1309.684 are unknown modified variants of ApeF.

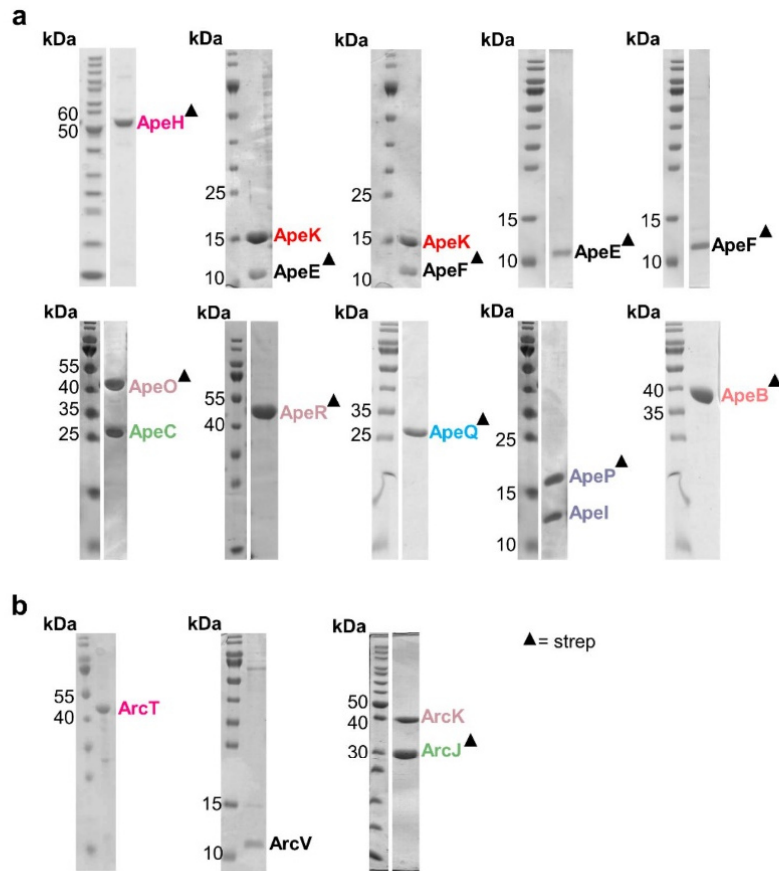


Figure S12. SDS-PAGE of all APE enzymes tested in this work isolated from *X. doucetiae* (a) and *Azoarcus* sp. BH72 (b). All proteins were purified via affinity chromatography (ApeE:ApeK, ApeF:ApeK, ApeO:ApeC, ApeI:ApeP and ArcK/J via pull-down). Strep- tagged proteins are marked with a triangle. For clarity, gel slices with multiple lines were cropped and the slice of interest was placed next to the protein ladder.

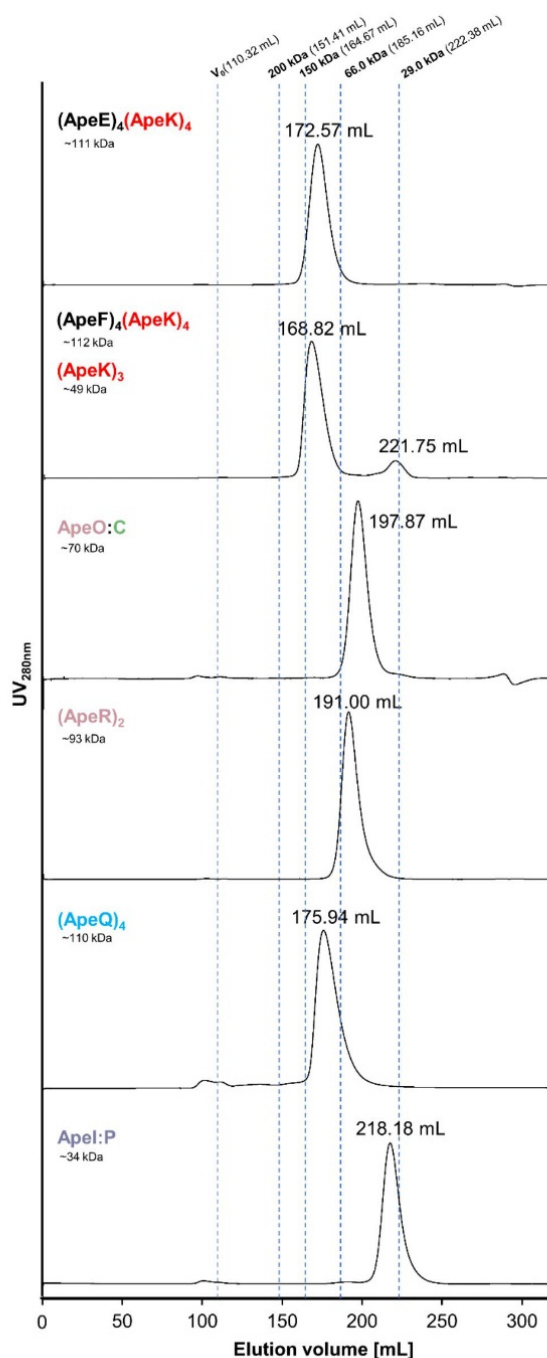


Figure S13. Size exclusion chromatography (SEC) of APE protein complexes. In dashed lines the elution volume of a standard calibration protein mixture with the corresponding molecular weight is shown. ApeO:ApeC and ApeR are retained more, than first expected from size, but based on the combination of the pull-down experiment, the native protein-MS data (Fig. 2c), and due to the fact that signals, corresponding to monomeric (or aggregated) ApeO or ApeC were missing during SEC, ApeO:ApeC and ApeR can be clearly assigned as dimer complex.

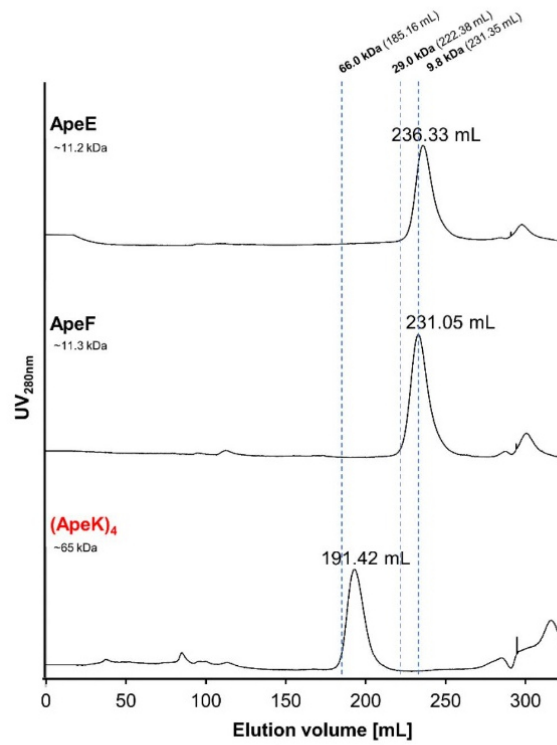


Figure S14. Size exclusion chromatography (SEC) of ApeE, ApeF and ApeK. In dashed lines the elution volume of a standard calibration protein mixture with the corresponding molecular weight is shown.

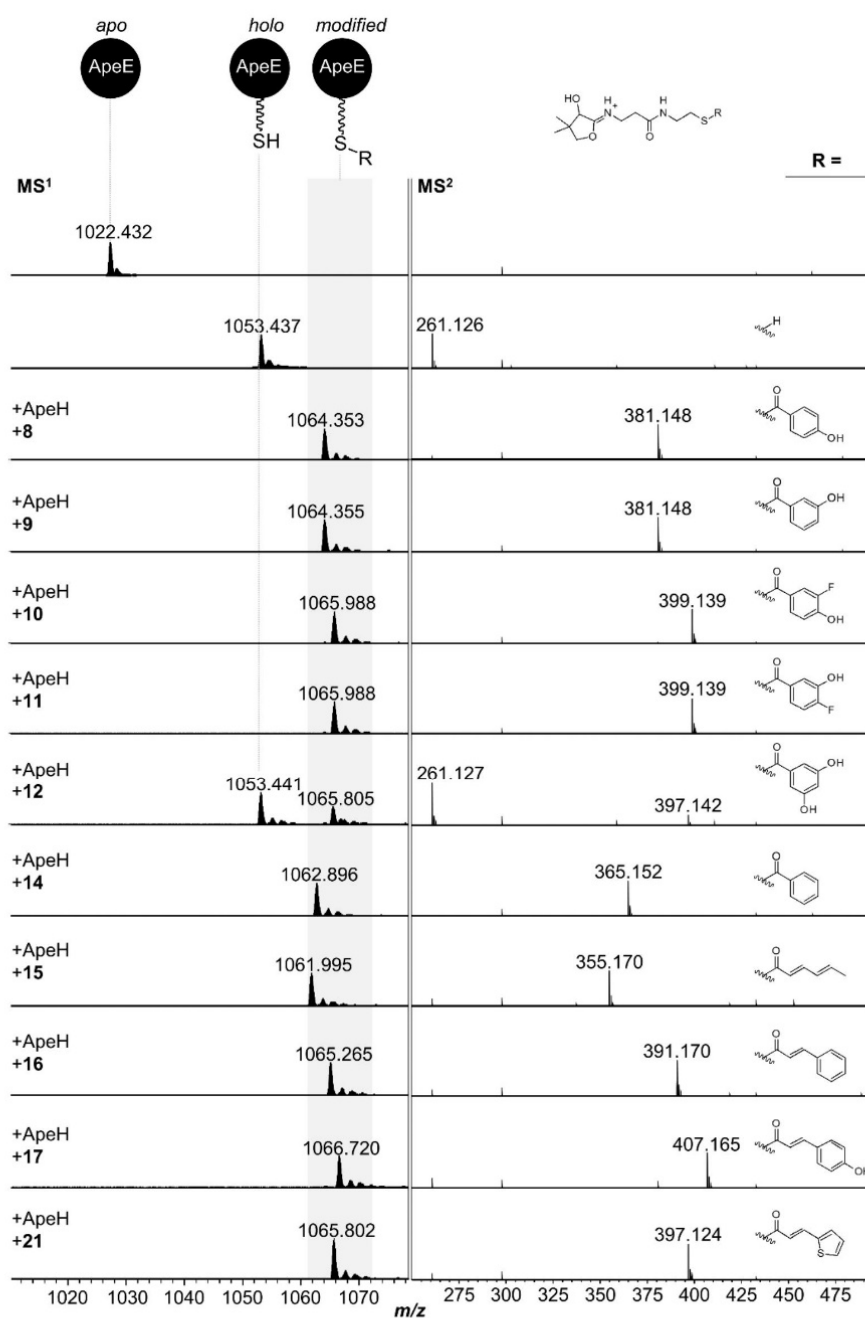


Figure S15. Substrate specificity of ApeH by AasS reaction with ApeE. The AasS reaction was performed with substrates **8-25** (Table S6) and analyzed via HR-HPLC-MS in positive mode with a Ppant ejection in MS². Displayed are the *m/z*-values of the average protein masses of the 11+ charge states (MS¹). *m/z* 1022.432 corresponds to *apo*, *m/z* 1053.437 *holo*-ApeE. *m/z* 1064.353, 1064.355, 1065.988, 1065.988, 1065.805, 1062.896, 1061.995, 1065.265, 1066.720 and 1065.802 represent *crypto*-variants of ApeE due to the AasS-loading of *holo*-ApeE by ApeH with acyl moieties of adenylated substrates **8-25** in MS¹, with corresponding Ppant ejection ions in MS², respectively. Substrates with no detectable acyl-ACP transfer are displayed in Table S6.

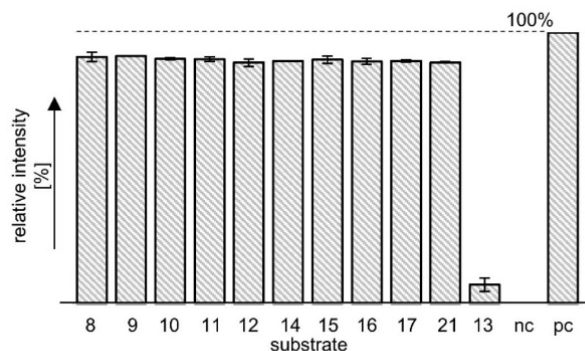


Figure S16. ATP-PPi exchange assay of ApeH with benzoic acid derivatives as substrates. The relative exchange from labeled γ - $^{18}\text{O}_4$ -ATP to unlabeled γ - $^{16}\text{O}_4$ -ATP is displayed; nc = γ - $^{18}\text{O}_4$ -ATP, pc = γ - $^{16}\text{O}_4$ -ATP.

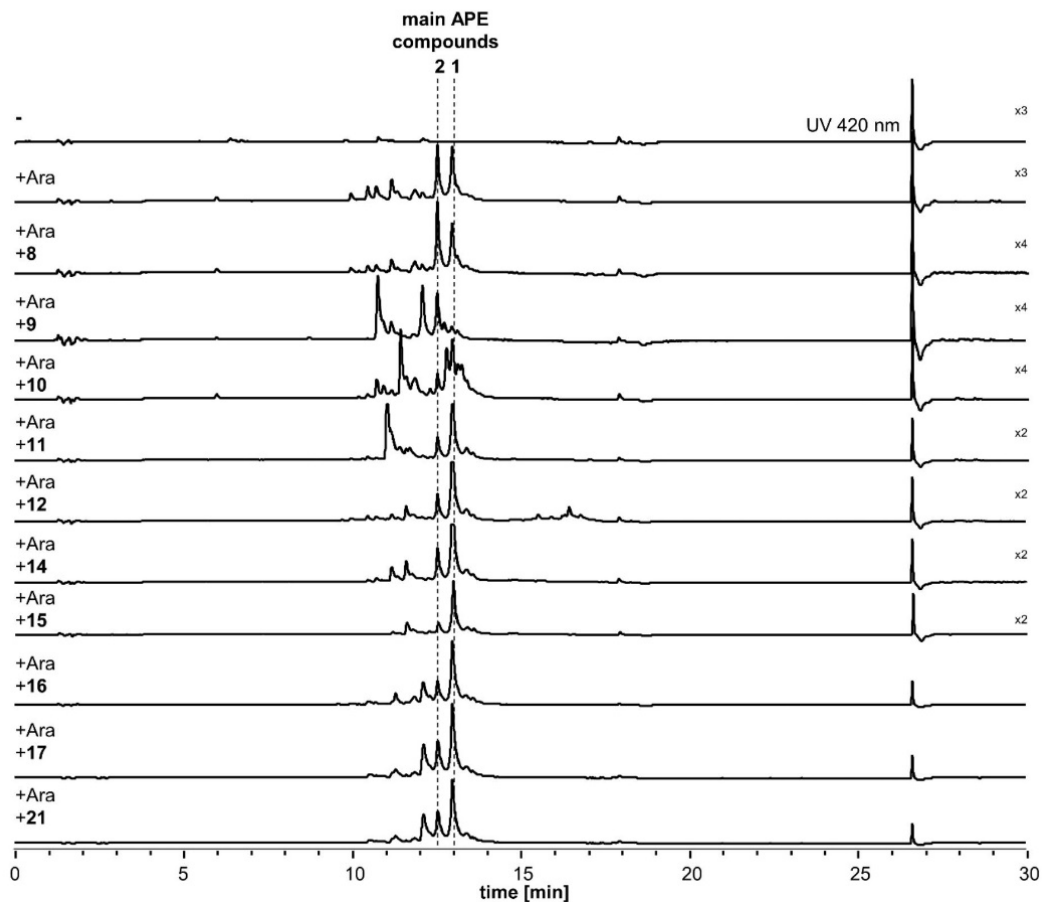


Figure S17. HPLC-UV analysis at 420 nm of *X. doucetiae* $\Delta\text{DC}\Delta\text{hfq}$ P_{BADapeB} overproduction strain fed with different benzoic acid precursor (see Table S6). Significant incorporation was obtained for **9** (3HBA), **10** (3F4HBA) and **11** (4F3HBA), whereas no incorporation was detected for precursor **12**, **14**, **15**, **16**, **17**, **21**. Main APE methyl ester **1** and **2** are displayed with dashed lines (compare Fig. S5). No statement can be made for **8** (4HBA) as its incorporation would not result in any retention time shift; - = minus induction, + Ara = plus 0.2% arabinose induction.

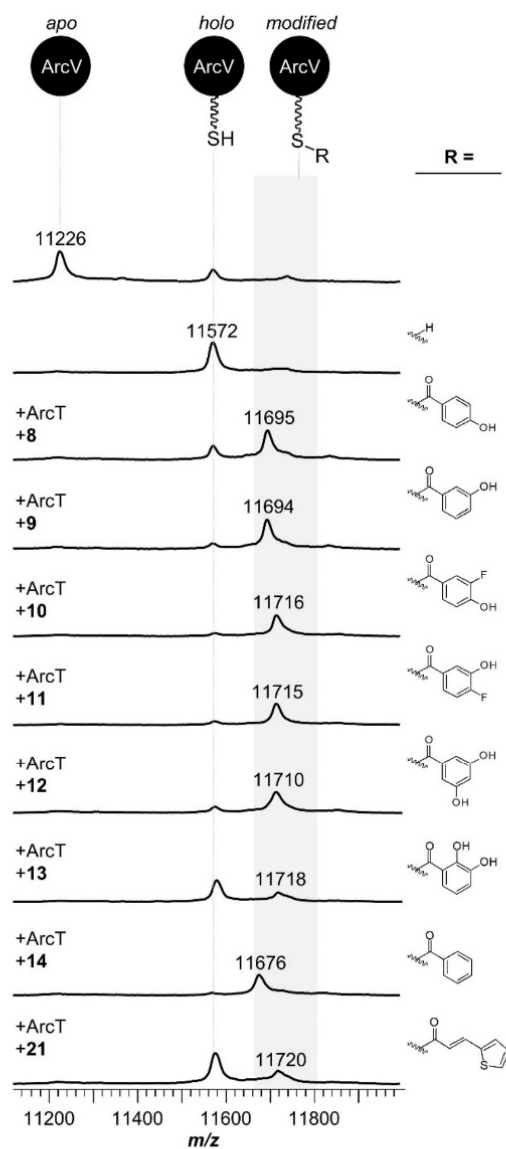


Figure S18. Substrate specificity of ArcT from *Azoarcus* sp. BH72 by AasS reaction with ArcV from *Azoarcus* sp. BH72. The AasS reaction was performed with substrates **8-14, 16,17,21,22, 26-32** (see Table S7) and analyzed via MALDI-MS. m/z 11226 corresponds to apo, m/z 11572 to holo-ArcV. m/z 11695, 11694, 11716, 11715, 11710, 11718, 11676, 11676 and 11720 represent *crypto*-variants of ArcV due to the AasS-loading of holo-ArcV by ArcT with acyl moieties of adenylated substrates **8-14, 16,17,21,22, 26-32**, substrates with no detectable acyl-ACP transfer are displayed in Table S7.

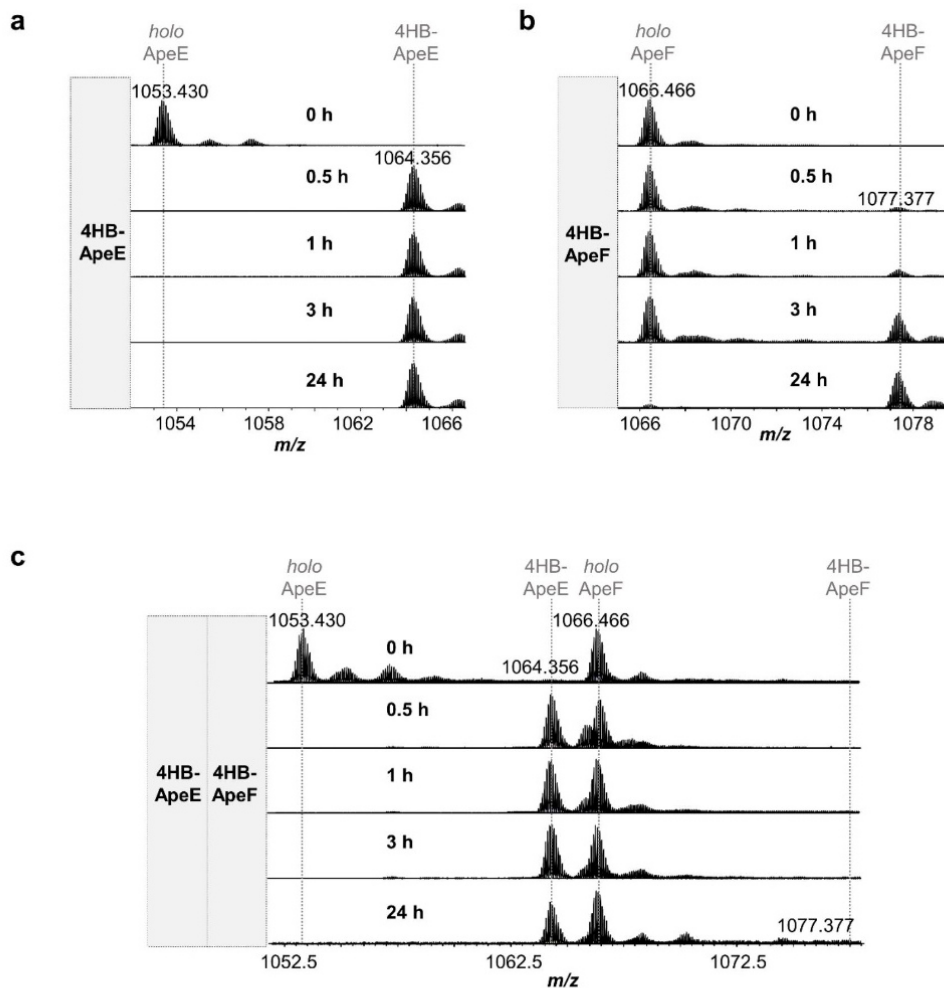


Figure S19. Differences in AasS reaction of ApeH with either ApeE (a) or ApeF (b) or competitive, with ApeE and ApeF (c). The AasS reaction of ApeH was performed with 4HBA as substrate with the corresponding *holo*-ACPs. The turnover to 4HBA-ACP was analyzed via HR-HPLC-MS in positive mode. Displayed are the m/z -values of the average protein masses of the 11+ charge states (MS^1). m/z 1053.430 and m/z 1066.466 corresponds to *holo*-ApeE or ApeF, respectively. m/z 1064.35 corresponds to 4HB-ApeE, m/z 1077.37 to 4HB-ApeF in MS^1 .

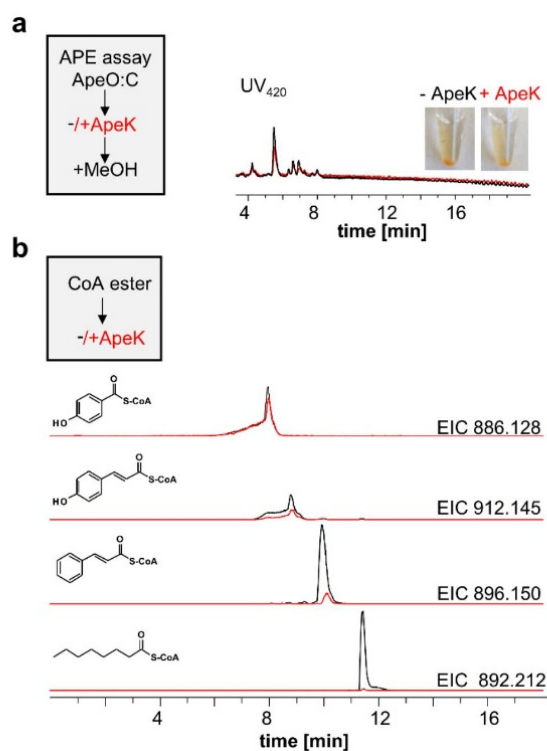


Figure S20. TE hydrolysis activity of ApeK analyzed by HPLC-UV-MS. ApeK was tested for its hydrolysis activity with ACP-bound APE from APE assay “ApeO:ApeC” monitored at 420 nm (**a**) and different CoA ester, which were displayed by their respective EICs ($[M+H]^+$, ± 0.005 Da) (**b**). Reactions with the addition of ApeK (+ApeK) are displayed in red, compared to reactions without ApeK (-ApeK). Prior analysis, the samples were extracted with MeOH which resulted in precipitated Ape-ACP and spontaneous hydrolysis products in a).

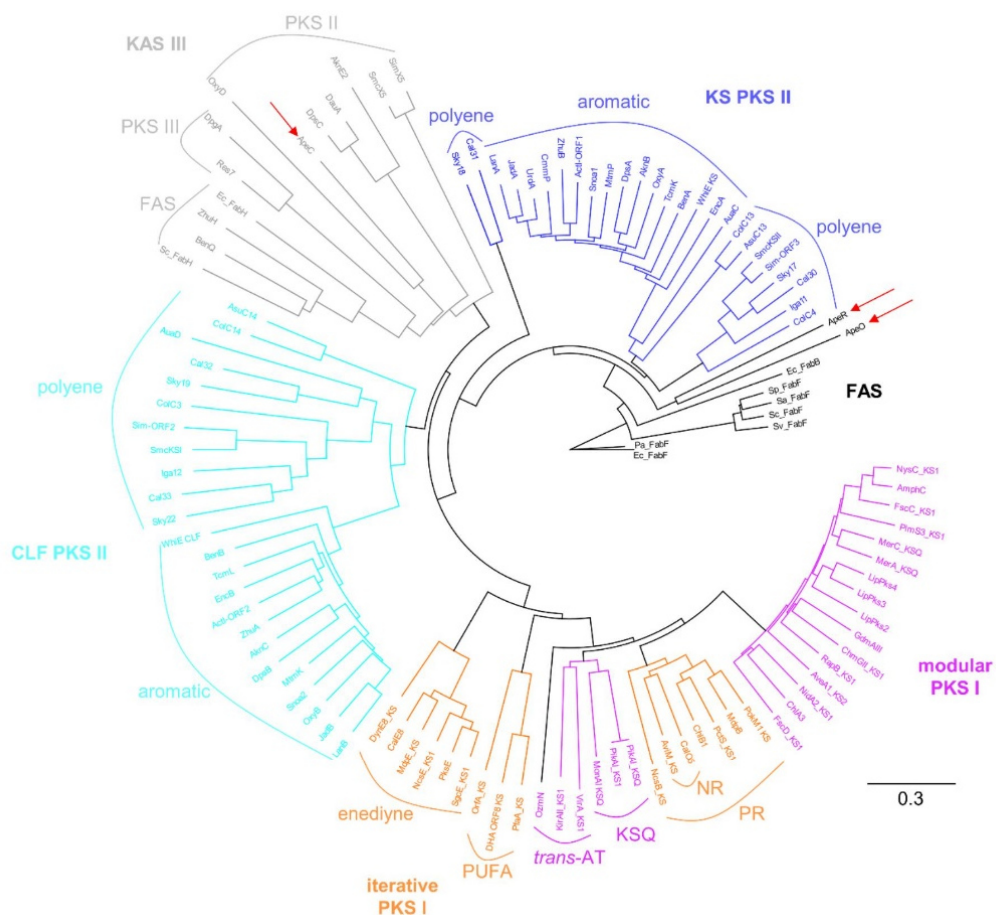


Figure S21. Phylogenetic analysis ApeO, ApeC and ApeR. Amino acid sequences are adopted from²⁷ (see Table 12), corresponding accession numbers are listed in Table S12. PUFA = polyunsaturated fatty acid, HR = highly reducing, NR = non-reducing, PR = partially reducing.

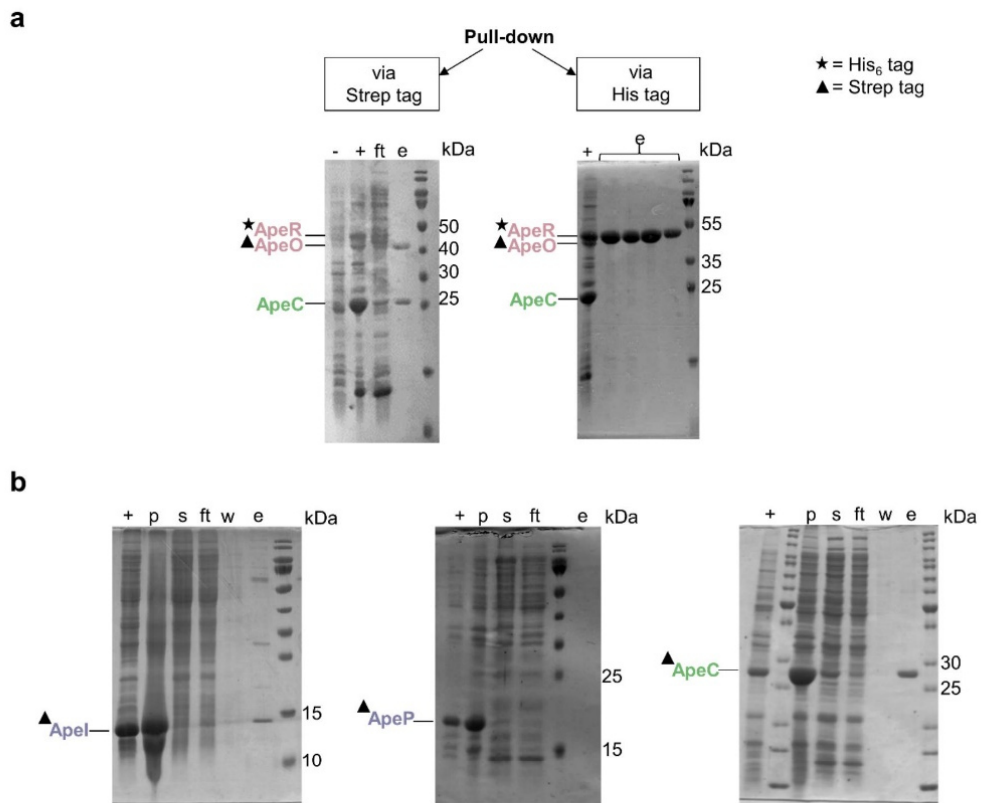


Figure S22. SDS-PAGE of interaction analysis between ApeO:ApeC/R showing an interaction of ApeO and ApeC but no interaction of ApeR with either ApeO or ApeC (**a**) and Strep-tag purification of ApeI, ApeP, ApeC, respectively (**b**). While ApeI and ApeC can be produced alone in a soluble form, ApeP is insoluble. - = without induction, + = with induction, p = pellet, s = supernatant, ft = flow through fraction, w = wash fraction, e = elution fraction.

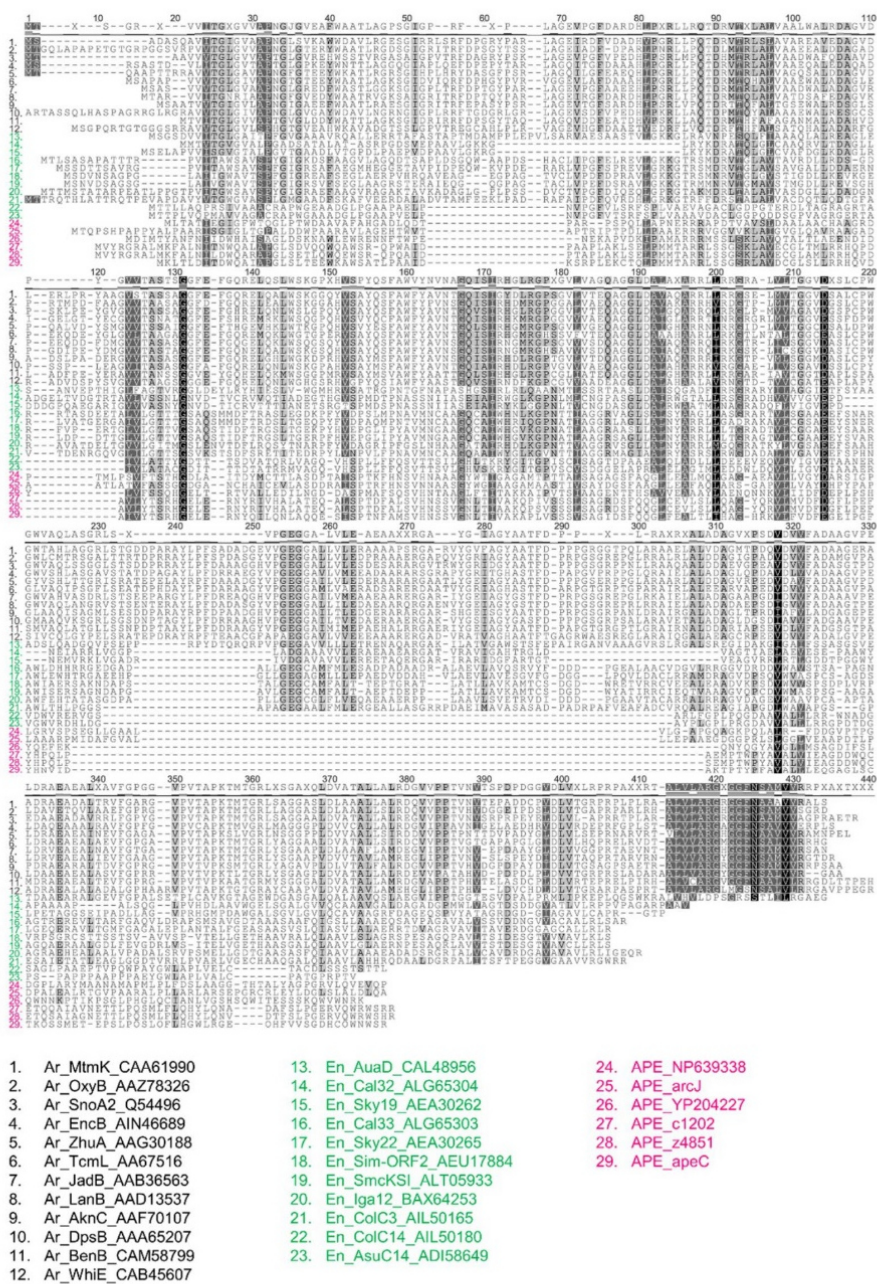


Figure S23. Comparison of CLFs from aromatic (Ar), Polyene (En) and APE PKS systems. The alignment was performed with Geneious 6.1.8 using ClustalW. Sequences 1-23 were used from²⁷, sequences from type I PKS systems were shortened to the length of typical CLF from type II PKS systems.

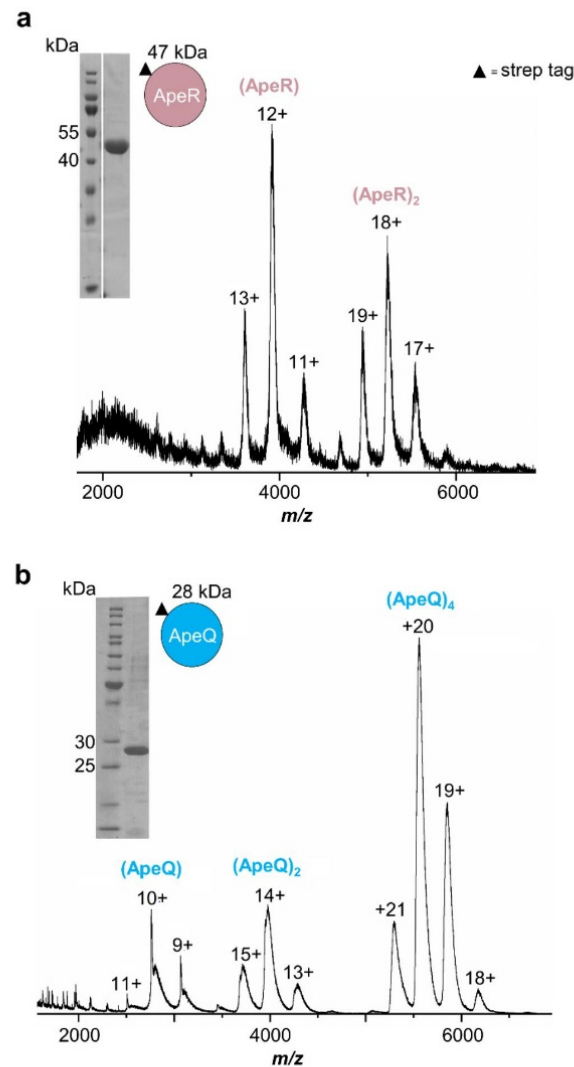


Figure S24. nESI-MS analysis of the native protein complex with corresponding SDS-PAGE of ApeR and of ApeQ. **(a)** shows monomers and the expected dimers of ApeR. **(b)** shows ApeQ in monomeric, dimeric and tetrameric state, showing that the dimer is the building block for the formation of the tetramer. All proteins were purified via Strep-Tactin affinity chromatography and directly used for MS-analysis. Strep-tagged proteins are marked with a triangle.

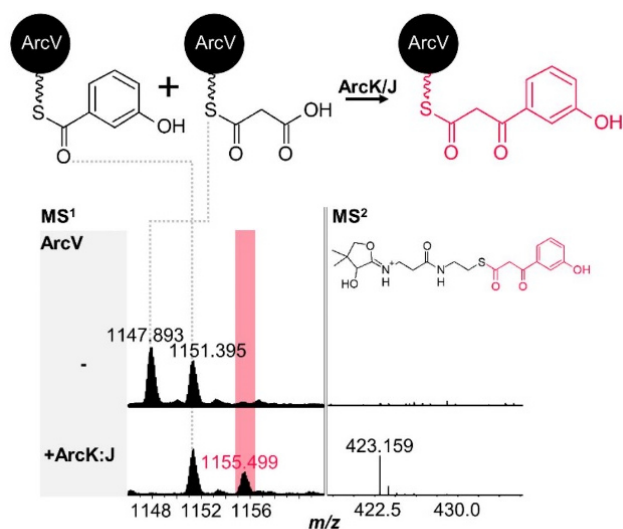


Figure S25. Initial chain elongation step catalyzed by ArcK:ArcJ. 3HB-ACP (Aas reaction by ArcT) served as starter and malonyl-ACP as elongation ACP. The reaction was analyzed via HR-HPLC-MS in positive mode with a Ppant ejection in MS². Displayed are the m/z -values of the average protein masses of the 11+ charge states (MS¹). m/z 1147.893 corresponds to malonyl-ArcV, m/z 1151.395 to 3HB-ArcV in MS¹. The β -ketoacyl elongation product corresponds to m/z 1155.499 with characteristic Ppant ejection ion of m/z 423.16 in MS².

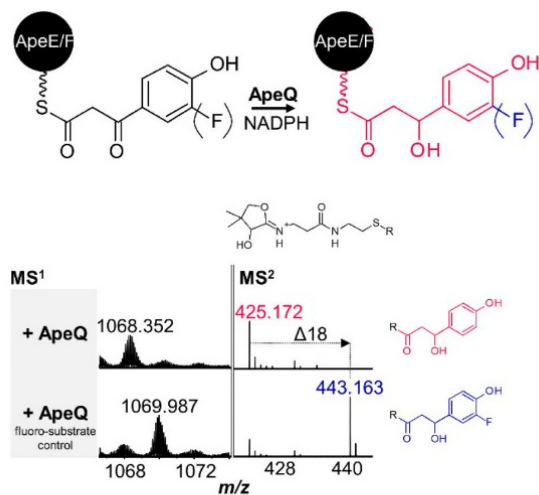


Figure S26. KR reaction by ApeQ. The ACP-bound β -ketoacyl product was converted to the β -hydroxyacyl product by ApeQ with NADPH. 3F4HB-ApeE (starter-ACP) was used in a control reaction (fluoro-substrate control) resulting in a +18 Da mass shift in MS². The reactions were analyzed via HR-HPLC-MS in positive mode with a Ppant ejection in MS². Displayed are the m/z -values of the average protein masses of the 11+ charge states (MS¹). The β -hydroxy product corresponds to m/z 1068.352 in MS¹ (fluoro-substrate control: m/z 1069.987) with the characteristic Ppant ejection ion of m/z 425.172 in MS² (fluoro-substrate control: m/z 443.163). The MS-results for ApeF are listed in Table S9.

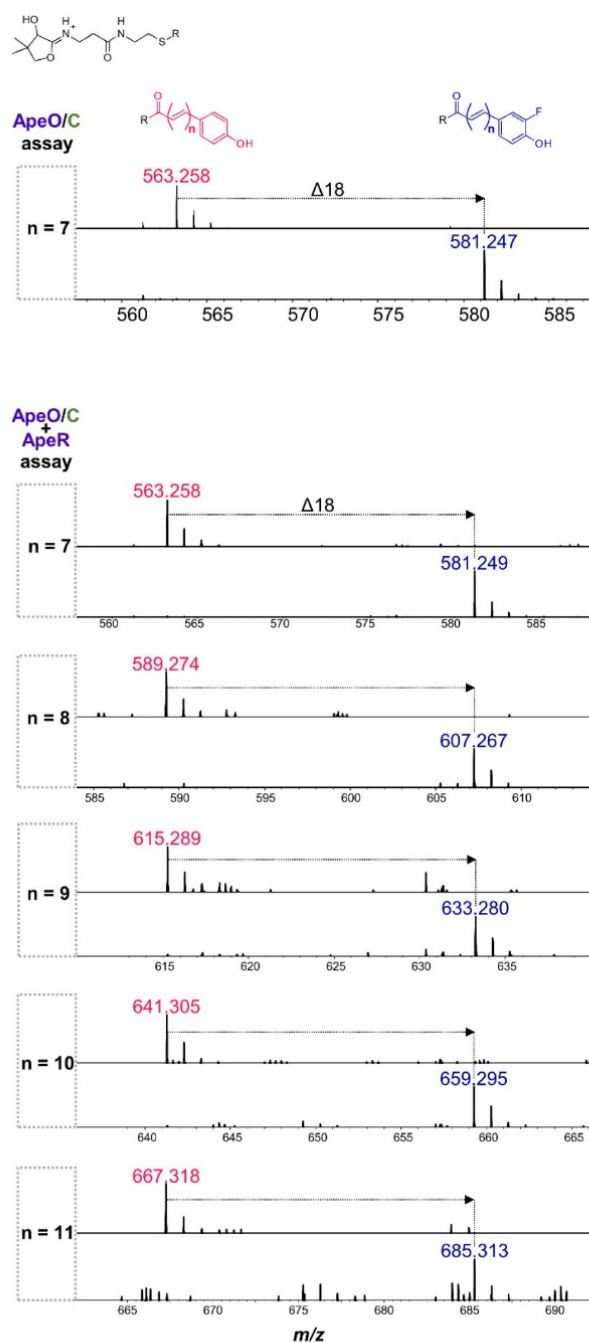


Figure S27. Main Ppant ejection ions of ACP-bound APE-formation plus fluoro-substrate control reaction with ApeO:ApeC and ApeO:ApeC plus ApeR. The reaction setup was performed with 4HB-ApeE as starter (AasS reaction by ApeH) and malonyl-ApeE as elongation ACP by adding the essential enzymes ApeO:ApeC, ApeR, ApeQ (+NADPH) and ApeI:ApeP. 3F4HB-ApeE (starter-ACP) was used in a control reaction (fluoro-substrate control) resulting in a +18 Da mass shift in MS². Different APE-chain lengths were produced by using either ApeO:ApeC and ApeO:ApeC and ApeR. See also Table S10 for more details, for fluoro-substrate control reactions see Table S11.

Supplementary References

- (1) Fu, C.; Donovan, W. P.; Shikapwashya-Hasser, O.; Ye, X.; Cole, R. H. Hot Fusion: An efficient method to clone multiple DNA fragments as well as inverted repeats without ligase. *PLoS One* **2014**, *9*, e115318.
- (2) Bode, E.; Brachmann, A. O.; Kegler, C.; Simsek, R.; Dauth, C.; Zhou, Q.; Kaiser, M.; Klemmt, P.; Bode, H. B. Simple "On-Demand" Production of Bioactive Natural Products. *Chembiochem* **2015**, *16*, 1115–1119.
- (3) Brachmann, A. O.; Joyce, S. A.; Jenke-Kodama, H.; Schwär, G.; Clarke, D. J.; Bode, H. B. A type II polyketide synthase is responsible for anthraquinone biosynthesis in *Photorhabdus luminescens*. *Chembiochem* **2007**, *8*, 1721–1728.
- (4) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J. *et al.* Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **2014**, *158*, 412–421.
- (5) Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **2005**, *41*, 207–234.
- (6) Kabsch, W. XDS. *Acta Crystallogr. D* **2010**, *66*, 125–132.
- (7) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- (8) Christensen, C. E.; Kragelund, B. B.; Wettstein-Knowles, P. von; Henriksen, A. Structure of the human beta-ketoacyl ACP synthase from the mitochondrial type II fatty acid synthase. *Protein Sci.* **2007**, *16*, 261–272.
- (9) Turk, D. MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallogr. D* **2013**, *69*, 1342–1357.
- (10) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **2011**, *67*, 355–367.
- (11) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **2010**, *66*, 213–221.
- (12) Langer, G. G.; Hazledine, S.; Wiegels, T.; Carolan, C.; Lamzin, V. S. Visual automated macromolecular model building. *Acta Crystallogr. D* **2013**, *69*, 635–641.
- (13) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **2010**, *66*, 486–501.
- (14) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **2010**, *66*, 12–21.
- (15) Phelan, V. V.; Du, Y.; McLean, J. A.; Bachmann, B. O. Adenylation enzyme characterization using gamma-¹⁸O₄-ATP pyrophosphate exchange. *Chem. Biol.* **2009**, *16*, 473–478.
- (16) Guo, Z.-F.; Sun, Y.; Zheng, S.; Guo, Z. Preferential hydrolysis of aberrant intermediates by the type II thioesterase in *Escherichia coli* nonribosomal enterobactin synthesis: Substrate specificities and mutagenic studies on the active-site residues. *Biochemistry* **2009**, *48*, 1712–1722.
- (17) Meluzzi, D.; Zheng, W. H.; Hensler, M.; Nizet, V.; Dorrestein, P. C. Top-down mass spectrometry on low-resolution instruments: Characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorg Med. Chem. Lett.* **2008**, *18*, 3107–3111.
- (18) Altenhoff, A. M.; Dessimoz, C. Inferring orthology and paralogy. *Methods Mol. Biol.* **2012**, *855*, 259–279.
- (19) Tobias, N. J.; Wolff, H.; Djahanschiri, B.; Grundmann, F.; Kronenwerth, M.; Shi, Y.-M.; Simonyi, S.; Grün, P.; Shapiro-Ilan, D.; Pidot, S. J. *et al.* Natural product diversity associated with the nematode symbionts *Photorhabdus* and *Xenorhabdus*. *Nat. Microbiol.* **2017**, *2*, 1676.
- (20) Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* **2007**, *23*, 127–128.
- (21) Letunic, I.; Bork, P. Interactive Tree Of Life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* **2011**, *39*, W475–8.
- (22) Thoma, S.; Schobert, M. An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol. Lett.* **2009**, *294*, 127–132.

-
- (23) Lay, N. R. D.; Cronan, J. E. A genome rearrangement has orphaned the *Escherichia coli* K-12 AcpT phosphopantetheinyl transferase from its cognate *Escherichia coli* O157:H7 substrates. *Mol. Microbiol.* **2006**, *61*, 232–242.
- (24) Bode, E.; He, Y.; Vo, T. D.; Schultz, R.; Kaiser, M.; Bode, H. B. Biosynthesis and function of simple amides in *Xenorhabdus doucetiae*. *Environ. Microbiol.* **2017**, *19*, 4564–4575.
- (25) Schöner, T. A.; Fuchs, S. W.; Reinhold-Hurek, B.; Bode, H. B. Identification and biosynthesis of a novel xanthomonadin-dialkylresorcinol-hybrid from *Azoarcus* sp. BH72. *PLoS One* **2014**, *9*, e90922.
- (26) Lorenzen, W.; Ahrendt, T.; Bozhüyük, K. A. J.; Bode, H. B. A multifunctional enzyme is involved in bacterial ether lipid biosynthesis. *Nat. Chem. Biol.* **2014**, *10*, 425.
- (27) Du, D.; Katsuyama, Y.; Shin-ya, K.; Ohnishi, Y. Reconstitution of a Type II Polyketide Synthase that Catalyzes Polyene Formation. *Angew. Chem.* **2018**, *57*, 1954–1957.
- (28) Fautz, E.; Reichenbach, H. A simple test for flexirubin-type pigments. *FEMS Microbiol Lett* **1980**, *8*, 87–91.

12.2 Manuscript: Biosynthesis of the Multifunctional Isopropylstilbene in *Photorhabdus laumondii* Involves Cross-Talk between Specialized and Primary Metabolism

Declaration on the contribution of the authors/Erklärung zu den Autorenanteilen

Status: submitted

Beteiligte Autoren: Gina L. C. Grammbitter (GLG), Siyar Kavakli (SK), Helge B. Bode (HBB)

Was hat der/die Promovierende bzw. was haben die Co-Autoren/Autorinnen beigetragen?

(1) zu Entwicklung und Planung

GG: 60%

SK: 10%

HBB: 30%

(2) Durchführung der einzelnen Untersuchungen und Experimente

Klonierung und Erstellung TT01-Mutante GG (10%) + SK (10%)

In vitro-Experimente mit StIB, StIE und AcpP und Click-Experiment mit gefütterter Azido-Fettsäure und entsprechende MALDI- und HPLC-ESI-MS-Analysen GG (30%)

In vitro-Experimente mit StIE und Fab-Enzymen und entsprechende HPLC-ESI-MS-Analysen: SK (30%), GG (5%)

Heterologe Stilben-Produktion SK (15%)

(3) Erstellung der Datensammlung und Abbildungen

In vitro-Experimente mit StIB, StIE und AcpP und Click-Experiment mit gefütterter Azido-Fettsäure und entsprechende MALDI- /ESI-MS-Analysen GG (30%)

Darstellung Biosynthesegencluster, Biosyntheschema, verwandte Stilbene GG (10%)

In vitro-Experimente StIE/Fab-Enzyme + ESI-MS-Analysen: SK (35%) + GG (5%)

Heterologe Stilben-Produktion SK (15%) + GG (5%)

(4) Analyse und Interpretation der Daten

In vitro-Experimente mit StIB, StIE und AcpP und Click-Experiment mit gefütterter Azido-Fettsäure und entsprechende MALDI- /ESI-MS-Analysen GG (30%)

In vitro-Experimente StIE/Fab-Enzyme + ESI-MS-Analysen: SK (30%) + GG (20%)

Heterologe Stilben-Produktion SK (15%) + GG (5%)

(5) Verfassung des Manuskripts

GLC: 60%

HBB: 30%

SK: 10%

date/place: _____

Biosynthesis of the Multifunctional Isopropylstilbene in *Photorhabdus laumondii* Involves Cross-Talk between Specialized and Primary Metabolism

Gina L. C. Grammbitter^{1‡}, Siyar Kavakli^{1‡}, Helge B. Bode^{1,2,3*}

¹ Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe-Universität Frankfurt, Frankfurt am Main, Germany

² Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universität Frankfurt, Frankfurt am Main, Germany

³ LOEWE Center for Translational Biodiversity in Genomics (TBG), Germany

‡ authors contributed equally

* corresponding author

E-mail: h.bode@bio.uni-frankfurt.de (HB)

Abstract

Isopropylstilbene (IPS) derived from the entomopathogenic bacterium *Photorhabdus* represents the only known stilbene which is not produced by a plant stilbene synthase but a bacterial PKS II synthase. The exclusive cyclization reaction, responsible for the formation of the characteristic iso-branched side-chain of the molecule, was extensively studied in the past, but some parts of the biosynthetic route remained elusive. In this study, we revealed the role of StIB that is able to acylate the ACP with cinnamate and demonstrated the elongation of this cinnamoyl-ACP intermediate with enzymes from the bacterial fatty acid biosynthesis pathway. Thus, we deciphered one of the rare cross-talks between the enzymes from primary and specialized metabolism. These insights led, for the first time, to the production of IPS in a heterologous host.

Introduction

Stilbenes are well-known natural products usually produced by type III PKS systems (or stilbene synthases, STS) in plants. The biosynthetic mechanism is well-described and involves elongation of cinnamoyl- or coumaroyl-CoA thioesters with three malonyl- extender units.^[1–3] Subsequent cyclization and aromatization leads to stilbenes, giving rise to resveratrol, pterostilbene and pinosylvin as some examples for this compound class, widespread in plants (Fig. 1).^[4]

The only non-plant stilbene is isopropylstilbene (IPS, Fig. 1, **1**) and derivatives thereof produced in *Photorhabdus* strains (Fig. 1).^[5–11] IPS is found in all *Photorhabdus* strains analyzed so far and can be regarded as a species-specific biomarker.^[12] *Photorhabdus* lives in mutualistic symbiosis with *Heterorhabditis* nematodes, acting together as an insect pathogen.^[13–15] It has been proven that IPS is required for nematode development and is important for the mutualistic symbiosis between bacterium and nematode.^[5,16] Due to its antimicrobial activity, IPS might furthermore serve as a protection against microbial competitors during insect infection. Additionally, IPS inhibits protozoa and even the mammalian soluble epoxide hydrolase, which is involved in the arachidonic acid cascade and thus in several inflammatory processes.^[5] Since it selectively modulates the cytokine cascade deep under the skin when applied as atopic cream, IPS is currently developed as treatment against psoriasis under the name tapinarof or benvitimod.^[5,17,18]

Interestingly, *Photorhabdus laumondii* utilizes a non-plant-like mechanism to produce IPS that relies on the condensation of two intermediates derived from different biosynthetic routes (Fig. 2b), which is crucial for obtaining the iso-branch.^[5,6,19] The corresponding genes are distributed within the genome and are encoded as stand-alone genes (*stIA* and *stIB*, *stICDE* and *bkdABC*) (Fig. 2a). IPS-biosynthesis starts with

the generation of cinnamic acid from phenylalanine catalyzed by a phenylalanine ammonia lyase (PAL, StIA). StIB is proposed to act as a CoA-ligase, activating cinnamic acid to cinnamoyl-CoA. The acyl moiety of cinnamoyl-CoA is suggested to be transferred on an acyl-carrier protein (ACP) followed by elongation, reduction and dehydration steps to obtain 5-phenyl-2,4-pentadienoyl-ACP. Since there is no separate operon for these steps encoded in the *Photorhabdus* genomes, these reactions might be performed by enzymes from the fatty acid biosynthesis. To date, there is no proof for this hypothesis and furthermore, no acyltransferase transferring the acyl-moiety of cinnamoyl-CoA to the ACP (StIE) was identified so far. That these reactions are most probably ACP-dependent (and performed without CoA-intermediates), was shown from analysis of a *ngrA* deletion.^[5] NgrA is an Sfp-type PPTase that is required for the biosynthesis of most PKS and non-ribosomal peptide synthetase derived peptides in *P. laumondii* TT01.^[20] The Michael acceptor 5-phenyl-2,4-pentadienoyl-ACP is further cyclized and aromatized with a β -ketoacyl isovalerate moiety derived from the branched-chain- α -keto acid dehydrogenase (BKD) complex (BkdABC), leading to IPS with the characteristic iso-branched side chain of stilbenes produced from *P. laumondii* (Fig. 2b).^[5,6] While the exclusive cyclization reaction was studied in the past with an *in vitro*-approach of the condensation (StID) and subsequent aromatization (StIC) and X-ray crystallography of the cyclase/ketosynthase (CYC/KS) StID^[19], the formation of the native substrate 5-phenyl-2,4-pentadienoyl-ACP remained elusive.

Here, we report the *in vitro*-characterization of StIB and furthermore the elongation of cinnamoyl-ACP to 5-phenyl-2,4-pentadienoyl-ACP by the orchestration of enzymes from the fatty acid biosynthesis pathway. Furthermore, we focused on stilbene/IPS production in a heterologous *E. coli* host.

Results and Discussion

We started our investigations by characterizing StIB for its postulated function as a CoA ligase. As cinnamic acid is proposed to be the natural substrate, it was conspicuous that during *in vitro* analysis with purified StIB, cinnamoyl-CoA was hardly detectable via MALDI-MS when compared to CoA derivatives generated from fatty acid substrates including chain length from C₆ to C₁₆ (Fig. S1, Table S5).

To rule out any ionization problems of cinnamoyl-CoA compared to the fatty acid acyl-CoA thioesters, we analyzed reactions with cinnamic, caproic, heptanoic and palmitic acid with HPLC-UV-MS and were able to detect the formed CoA-esters by UV (260 nm) and MS (Fig. S2). In this reaction set-up, cinnamoyl-CoA was not detectable.

We further observed that StIB accepts some cinnamic acid derivatives but discriminates substituents in the *para*-position (Table S5). These findings are in accordance with previously conducted mutasynthesis experiments, where *para*-substituted cinnamic acid derivatives did not lead to novel IPS derivatives in *P. laumondii*, while *meta*-substituted substrates were incorporated by the producer.^[21]

stIB is annotated as a FadD, a fatty acyl-CoA synthetase, converting fatty acids to acyl-CoAs for further fatty acid degradation.^[22] We surveyed the influence of *stIB* presence on the fatty acid degradation profile but were not able to detect significant differences under the used conditions (Fig. S3, S4).

Since StIB produces less cinnamoyl-CoA than fatty acyl-CoA, we tested StIB for its activity as an adenylating enzyme, thus being capable of direct-loading of adenylated cinnamoyl-precursor to *holo*-StIE. Such enzymes are known as acyl-acyl carrier protein synthetases (AasS) or AMP-ligases.^[23–26] Hence, we purified the ACP StIE and transferred it into its *holo*-form with the PPTase Sfp.^[27] When we tested StIB for AasS activity in the presence of cinnamic acid, we were indeed able to show the transferase

reaction (Fig. 3a). Hence, we suggested StIB to substitute the “missing” acyltransferase responsible for cinnamic acid activation and StIE acylation to result in cinnamoyl-StIE. Furthermore, analysis of StIB with fatty acid substrates demonstrated acyl transfer of the fatty acid substrates and also acylation of AcpP from *E. coli* (Fig. S6).

IPS biosynthesis requires the formation of 5-phenyl-2,4-pentadienoyl-ACP as substrate for the cyclization reaction catalyzed by the ketosynthase/cyclase StID (see Fig. 2b).^[5,6,19] StID only accepts unsaturated intermediates.^[6,19] Hence, the elongation of cinnamoyl-ACP and its subsequent reduction and dehydration is crucial for cyclization. Since none of the required enzymes, namely a ketosynthase, ketoreductase and a dehydratase, are encoded next to the known IPS biosynthesis enzymes, we tested enzymes from the fatty acid synthase II (FAS II) of *P. laumondii* as possible candidates for these reactions. For elongation reaction, we selected FabH, FabB and FabF as plausible ketosynthase candidates, as the special ketosynthase StID is not able to fulfill this task. FabH is usually responsible for the initial elongation of acetyl-CoA starter with malonate as elongation units in FA biosynthesis, while FabB and FabF are responsible for elongation of longer acyl-ACP supported by distinct protein-protein interactions.^[28,29] In our *in vitro* analysis with purified enzymes (Fig. S7), only FabH was able to elongate cinnamoyl-ACP with malonyl-ACP to its β -ketoacyl product, while FabB or FabF were not (Fig. S8). Further ketoreduction was performed by FabG and dehydration by FabA as well as FabZ (Fig. 3b).

Although the dehydrated product was formed even without the addition of either FabA nor FabZ, the amount of dehydrated product was increased upon the addition of a dehydratase (Fig. 3b). Furthermore, in the absence of the dehydratases, the amount

of the dehydrated side product remained stable upon 2 h incubation time, which is equivalent to the incubation time used for the assay (Fig. S9).

With the gained knowledge from our *in vitro*-assays, we focused on the production of stilbene compounds in a heterologous host. So far, only dialkylresorcinol (DAR) compounds, lacking the cinnamoyl side chain, were detected in heterologous expression experiments with StlCDE-homologs.^[30] Having now identified all key players for stilbene biosynthesis, for the first time, we successfully produced stilbene products **1** and **2** in a heterologous *E. coli* host (Fig. 4, Fig. S10). The production of **2** was achieved by co-expression of *stlCDE*, *stlB* and *p/FabH* under cinnamic acid supplementation. Although minor production of **2** was detectable even in the absence of the AasS-like enzyme StlB, the addition of StlB increased production of **2** in the heterologous *E. coli* host. Thus, both enzymes, StlB and *p/FabH* are important to produce the stilbene product, but only *p/FabH* is essential.

E. coli lacks the homologs of the BKD-complex^[31] to produce a β -keto isovalerate-moiety, which results in the characteristic iso-branch of IPS. Thus, with the introduction of the BKD-enzymes to the heterologous host, we could achieve production of **1** (Fig. 4). The addition of *p/FabG* increased the production titer of **1** and **2** even further (Fig. S11).

Conclusion

We revealed the missing parts of the IPS biosynthesis from *P. laumondii* to generate 5-phenyl-2,4-pentadienoyl-ACP, the required precursor for cyclization reaction by the KS/CYC StlD (Fig. 5). We demonstrated that StlB is able to acylate the cognate ACP StlE to form cinnamoyl-ACP. Cinnamoyl-ACP is further used by enzymes from FAS II biosynthesis for elongation, reduction and dehydration reaction, mediated by FabH, FabG and FabA or FabZ, respectively. Thus, the IPS biosynthesis is an example of

cross-talk activity between FAS II enzymes from primary metabolism and proteins from specialized metabolism. Similar cross-talk between specialized and primary metabolism and metabolites was very recently described in the formation of the rhabdoplanins formed via an 'Ugi-like' reaction, including building blocks from primary metabolism and the specialized metabolite rhabduscin.^[32] The gained knowledge from our *in vitro*-setup helped us to further select the minimal gene endowment needed for the production of stilbenes in a heterologous host. These efforts can be of great interest for further engineering strategies to produce stilbene derivatives.

Acknowledgements

Work in the Bode laboratory was supported by the LOEWE Schwerpunkt MegaSyn, funded by the State of Hesse. We acknowledge the Deutsche Forschungsgemeinschaft for funding of the Impact II qTof mass spectrometer (Grant INST 161/810-1). We thank Alexander Perez for providing the azido-C₁₆ fatty acid, Carsten Kegler and Charles O. Rock for providing the plasmids for Sfp- and AcpP-production, respectively. We further thank Michael Karas for MALDI access.

Figures

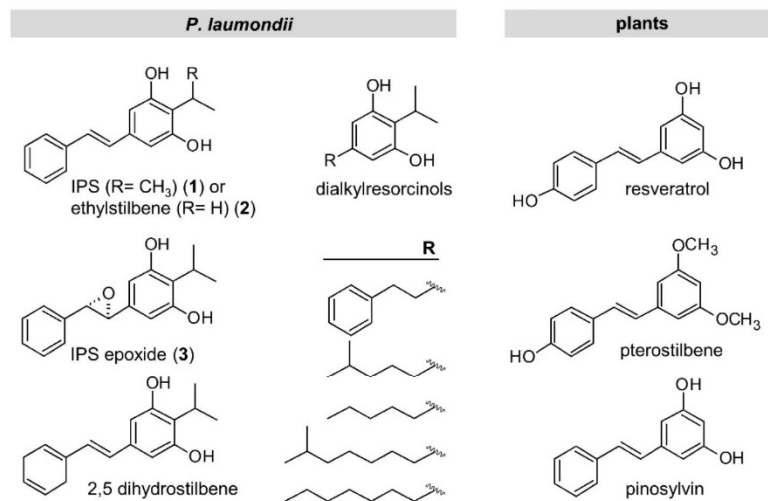


Figure 1. Selected stilbene and dialkylresorcinol (DAR) compounds produced by *P. laumondii* and stilbene representatives from plants.

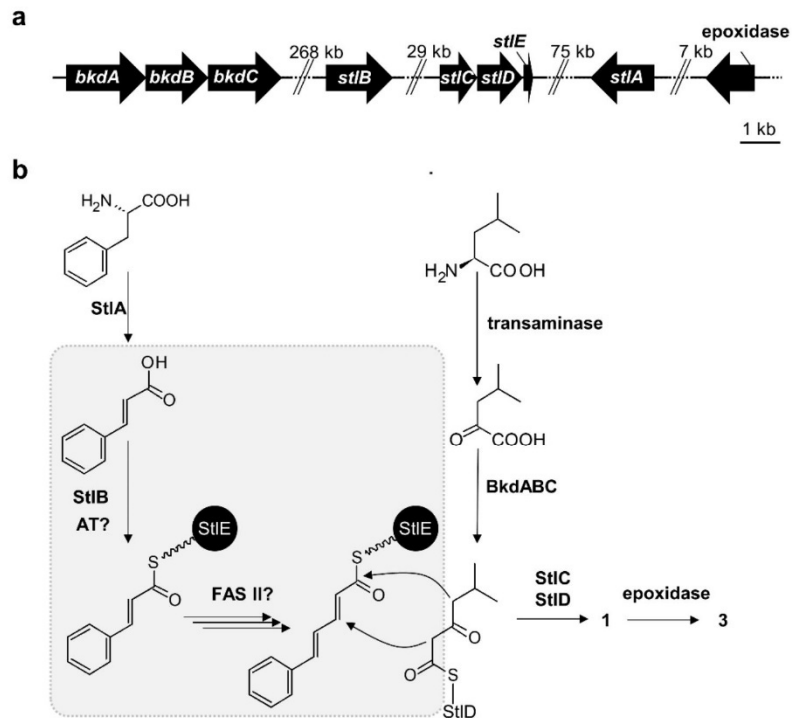


Figure 2. Genes involved in IPS production (a) and suggested IPS biosynthesis (b) in *P. laumondii*. BkdABC = branched-chain- α -keto acid dehydrogenase with ketosynthase (*plu1883-1885*), *StlB* = CoA ligase (*plu2134*), *StlC* = aromatase (*plu2163*), *StlD* = ketosynthase/cyclase (*plu2164*), *StlE* = acyl-carrier protein (*plu2165*), *StlA* = phenylalanine ammonium lyase (*plu2234*), *epoxidase* = *plu2236*^[7], AT = acyltransferase, FAS II = fatty acid synthase type II. Highlighted in grey are steps, which were investigated by *in vitro*-studies.

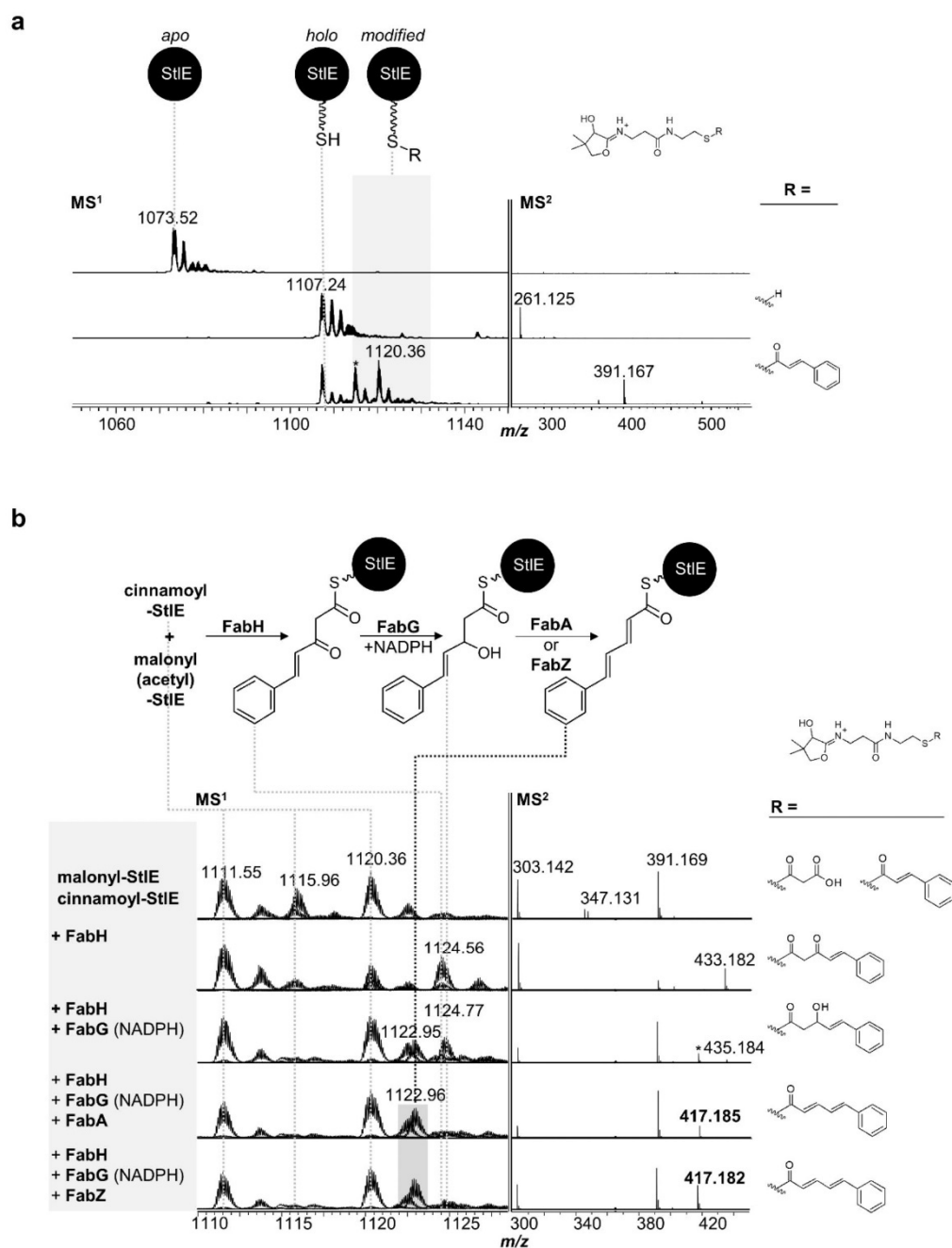


Figure 3. AaS loading of StIE catalyzed by StIB. *apo*-StIE was transferred to *holo* StIE with Sfp. *holo*-StIE was transferred to cinnamoyl-StIE by the AaS activity of StIB. Reaction marked with an asterisk indicates an unknown side product arising during overnight incubation (a). *In vitro* elongation, ketoreduction and dehydration reaction with FAS II enzymes FabH, FabG (+NADPH), FabA, FabZ from *P. laumondii* and cinnamoyl-ACP as starter and malonyl-ACP as elongation unit. Spectra were overlaid as described in the experimental section. Signal marked with an asterisk represents the dehydrated product without the addition of FabA/FabZ (b).

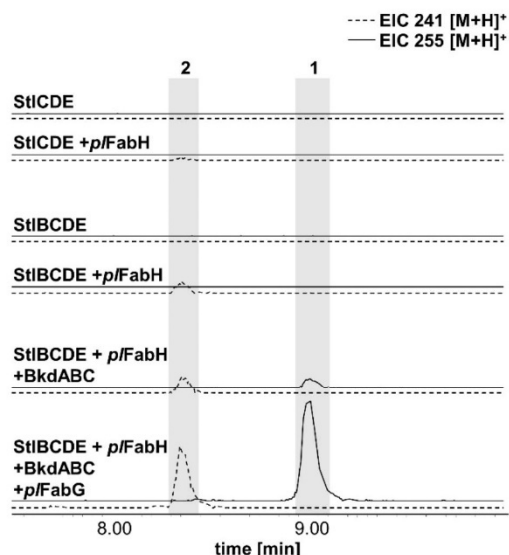


Figure 4. Heterologous stilbene production in *E. coli*. Analysis of **1** and **2** production after cultivation of *E. coli* BL21 with introduced StICDE \pm *p*/FabH, StIBCDE \pm *p*/FabH, or StIBCDE with *p*/FabH and BkdABC \pm *p*/FabG. The crude extracts of the corresponding strains were analyzed by HPLC-UV-MS in positive mode. The EIC of **2** ([M+H]⁺ 241.122 \pm 0.005 Da) and **1** ([M+H]⁺ 255.138 \pm 0.005 Da) are depicted (further MS² and UV data, see Fig. S10).

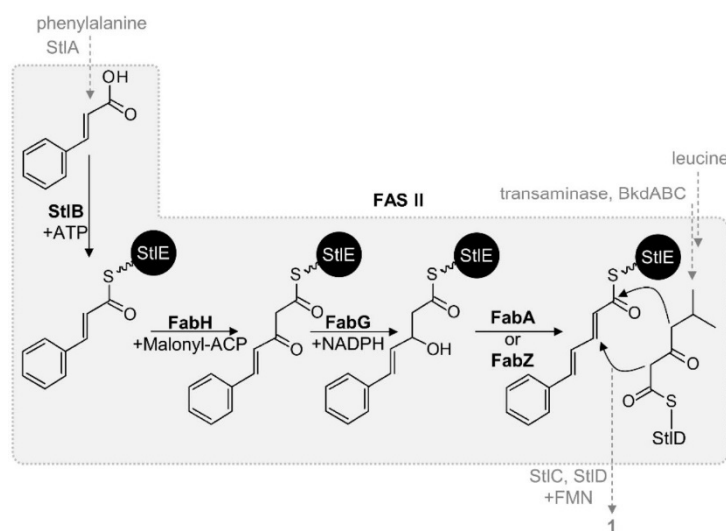


Figure 5. Extended IPS-biosynthesis. The cinnamoyl moiety of the adenylated precursor is loaded to StIE by the acyl-acyl-carrier protein synthetase (AasS) activity of StIB before being elongated with one malonate unit by FabH. The resulting β -ketoacyl product is further reduced and dehydrated by FabG and FabZ/FabA. StIB = AasS, FabH = ketosynthase; FabG = ketoreductase; FabA/FabZ = dehydratase; FAS II = fatty acid synthase type II, StIC = aromatase, StID = ketosynthase/cyclase; BkdABC = branched-chain- α -keto acid dehydrogenase with ketosynthase (for further details see Fig. 2).

References

- [1] T. P. Martins, C. Rouger, N. R. Glasser, S. Freitas, N. B. de Fraissinette, E. P. Balskus, D. Tasdemir, P. N. Leão, *Nat. Prod. Rep* **2019**.
- [2] M. B. Austin, J. P. Noel, *Nat. Prod. Rep.* **2003**, *20*, 79.
- [3] J. Schroder, G. Schroder, *J. Biosci.* **1990**, *45*, 1.
- [4] J. Chong, A. Poutaraud, P. Hugueney, *Plant Science* **2009**, *177*, 143.
- [5] S. A. Joyce, A. O. Brachmann, I. Glazer, L. Lango, G. Schwär, D. J. Clarke, H. B. Bode, *Angew. Chem* **2008**, *120*, 1968.
- [6] S. W. Fuchs, K. A. J. Bozhuyuk, D. Kresovic, F. Grundmann, V. Dill, A. O. Brachmann, N. R. Waterfield, H. B. Bode, *Angew. Chem.* **2013**, *52*, 4108.
- [7] H. B. Park, P. Sampathkumar, C. E. Perez, J. H. Lee, J. Tran, J. B. Bonanno, E. A. Hallem, S. C. Almo, J. M. Crawford, *J. Biol. Chem.* **2017**, *292*, 6680.
- [8] J. M. Crawford, S. A. Mahlstedt, S. J. Malcolmson, J. Clardy, C. T. Walsh, *Chem. Biol.* **2011**, *18*, 1102.
- [9] J. M. Crawford, R. Kontnik, J. Clardy, *Current biology : CB* **2010**, *20*, 69.
- [10] R. Kontnik, J. M. Crawford, J. Clardy, *ACS Chem. Biol.* **2010**, *5*, 659.
- [11] H. B. Park, J. M. Crawford, *J. Nat. Prod.* **2015**, *78*, 1437.
- [12] N. J. Tobias, H. Wolff, B. Djahanschiri, F. Grundmann, M. Kronenwerth, Y.-M. Shi, S. Simonyi, P. Grün, D. Shapiro-Ilan, S. J. Pidot et al., *Nat. Microbiol.* **2017**, *2*, 1676.
- [13] E. E. Herbert, H. Goodrich-Blair, *Nat. Rev. Microbiol* **2007**, *5*, 634.
- [14] Y.-M. Shi, H. B. Bode, *Nat. Prod. Rep.* **2018**, *35*, 309.
- [15] N. J. Tobias, Y.-M. Shi, H. B. Bode, *Trends Microbiol.* **2018**, *26*, 833.
- [16] A. Hapeshi, J. M. Benarroch, D. J. Clarke, N. R. Waterfield, *Microbiology* **2019**, *165*, 516.
- [17] J. Peppers, A. S. Paller, T. Maeda-Chubachi, S. Wu, K. Robbins, K. Gallagher, J. E. Kraus, *J. Am. Acad. Dermatol.* **2019**, *80*, 89-98.e3.
- [18] L. Zhao, X. Chen, L. Cai, C. Zhang, Q. Wang, S. Jing, G. Chen, J. Li, J. Zhang, Y. Fang, *J. Clin. Pharm. Ther.* **2014**, *39*, 418.
- [19] T. Mori, T. Awakawa, K. Shimomura, Y. Saito, D. Yang, H. Morita, I. Abe, *Cell Chem. Biol.* **2016**, *23*, 1468.
- [20] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep* **2014**, *31*, 61.
- [21] M. Kronenwerth, A. O. Brachmann, M. Kaiser, H. B. Bode, *ChemBiochem* **2014**, *15*, 2689.
- [22] L. Jimenez-Diaz, A. Caballero, A. Segura in *Rojo (Hg.) 2018 – Aerobic utilization of hydrocarbons*, pp. 1–23.
- [23] D. J. Campopiano, *Chem. Biol.* **2014**, *21*, 1257.
- [24] C. O. Rock, J. E. Cronan, *Int. J. STD AIDS* **1979**, *254*, 7116.
- [25] T. K. Ray, J. E. Cronan, JR, *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 4374.
- [26] J. Beld, K. Finzel, M. D. Burkart, *Chem. Biol.* **2014**, *21*, 1293.
- [27] L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585.
- [28] J. Beld, D. J. Lee, M. D. Burkart, *Mol Biosyst.* **2015**, *11*, 38.
- [29] A. Chen, R. N. Re, M. D. Burkart, *Nat. Prod. Rep* **2018**, *35*, 1029.
- [30] T. A. Schöner, S. W. Fuchs, B. Reinhold-Hurek, H. B. Bode, *PLoS One* **2014**, *9*, e90922.
- [31] Y. Fujita, H. Matsuoka, K. Hirooka, *Mol. Microbiol.* **2007**, *66*, 829.
- [32] J. Oh, N. Y. Kim, H. Chen, N. W. Palm, J. M. Crawford, *J. Am. Chem. Soc.* **2019**, *141*, 16271.

SUPPORTING INFORMATION

Experimental Section

Cloning

For isolation of genomic DNA from *P. laumondii*, Genra Puregene Yeast/Bact Kit (Qiagen) was used. Polymerase chain reaction (PCR) was performed with Phire Hot Start II DNA polymerase (Thermo Scientific), Phusion High-Fidelity DNA Polymerase, or Q5 polymerase (New England Biolabs) according to manufacturers' instructions. Oligonucleotides were purchased from Eurofins Genomics. The Invisorb Spin DNA Extraction Kit (Stratec) was used for DNA purification from agarose gels. Plasmid isolation was performed with the Invisorb Spin Plasmid Mini Two Kit (Stratec). Plasmid-backbone PCRs were restriction digested with *DpnI* (New England Biolabs) following the manufacturers' protocol. All used primers and templates and final plasmids are listed in Table S2 and S3. If not stated otherwise, all plasmids were cloned via Hot Fusion^[1] reaction with corresponding primers listed in Table S2. *plu2134* gene region was amplified using TS120 and TS121, generating a PCR product of 649 bp, which was digested (*SacI* and *SphI*) and ligated (T4 DNA ligase, Sigma) into the linearized vector (*SacI* and *SphI*) pDS132. *E. coli* S17 λ pir (pDS132_ *plu2134*) and *E. coli* DH10B (remaining plasmids) were used as cloning strains and electroporated with desalted ligation/Hot Fusion reaction (MF-Millipore membrane, VSWP, 0.025 μ m).

Generation of TT01 mutant

The *plu2134* (*stlB*)-insertion mutant of *P. laumondii* TT01 was created by conjugation with the *E. coli* S17 λ pir strain harboring pDS132_*plu2134*, which carries the first 578 bp of *plu2134*, as described earlier for creating insertion-mutants.^[2] As a result, in a homologous recombination event, the plasmid backbone is inserted within the gene of interest in the case of *plu2134*, resulting in a non-functional gene (*plu2134*-insertion, *P. laumondii* TT01::pDS132_*plu2134*).

Therefore, both strains, *P. laumondii* TT01 and *E. coli* S17 λ pir with pDS132_*plu2134*, were grown in 10 mL LB-medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl at pH 7.5) (*E. coli* S17 λ pir was supplemented with 35 μ g/mL chloramphenicol) to an OD₆₀₀ of 0.6-0.8. Cells of 1 mL culture were harvested, washed and resuspended in 400 μ L LB and mixed in a ratio of 3:1 (75 μ L TT01:25 μ L *E. coli*) prior to pipetting them in one drop on an LB agar plate. After 1 d at 30°C, cells were resuspended in 1 mL LB. Insertion mutants were selected by streaking several dilutions of cell suspension on LB-Cm^R-Rif^R (35 μ g/mL and 100 μ g/mL, respectively) agar plates. Positive insertion mutants were PCR-verified with primer TS122 and pDS132_rev leading to an 1147 bp product.

Click reaction and XAD-extraction of TT01 wildtype and mutants

The fatty acid degradation profile of *P. laumondii* was analyzed via feeding of azido palmitic acid followed by a click reaction with bicyclononyne (BCN).^[3] Therefore, pre-cultures of TT01 and mutants were grown overnight in LB-medium. The following day appropriate cell material was used for inoculation of the main cultures to an OD₆₀₀ of 0.1, supplemented with 0.1 mM azido palmitic acid and incubated for 24 h at 30°C and 200 rpm. After 2 h, 4 h, 6 h, 10 h, 20 h and 24 h, 1 mL cell culture of each strain was harvested and stored at -20°C. The click reaction of (the degraded) azido palmitic

acid was performed as previously described.^[3] Briefly, the cell pellet was resuspended in 500 μ L 1 M NaOH and incubated for 1-2 h at 85°C and 750 rpm on a thermocycler. The solution was acidified with 100 μ L 6 M HCl prior to extraction with 800 μ L hexane. 600 μ L of the organic layer was evaporated under reduced pressure. The click reaction was performed with 150 μ L of 0.2 mM BCN (in ACN) and incubated overnight at RT. A 1:10 dilution of each click reaction was analyzed via HR-HPLC-MS. The expected masses of the click products are listed in Figure S3.

To obtain a comparable growth rate of wildtype and mutant strains, the cultivation was carried out without the addition of antibiotics. After 72 h, the plasmid insertion of *P. laumondii* TT01::P_{DS132}*plu2134* was proven (no production of **1**) by HPLC-MS (Fig. S5). Therefore, the strain was cultivated in 10 mL LB as described above, with further supplementation of 2% Amberlite® XAD-16 (Sigma-Aldrich) beads. After 48 h, the XAD-beads were extracted with 10 mL MeOH, filtrated and dried under reduced pressure. Samples were further analyzed by HPLC-UV/MS.

Heterologous stilbene production

BL21 (DE3) Star (or BL21 (DE3) Star::FTRentD for StIE) was transformed with the production plasmids listed in Table S4, respectively.

For detection of **1** and **2**, extracts from 10 mL LB-culture of each strain were prepared upon XAD-bead addition. Therefore, the corresponding overnight pre-culture was used for inoculation of 10 mL LB-medium to an OD₆₀₀ of 0.2 and supplemented with 1 mM IPTG, with and without 1 mM cinnamic acid (Sigma-Aldrich) (from 100 mM cinnamic acid stock solution in ethanol) and 2% Amberlite® XAD-16 (Sigma-Aldrich) resin. After 48 h, XAD-beads of each culture were extracted with 10 mL MeOH, filtered and dried under reduced pressure. For HR-HPLC-MS measurement, the samples were dissolved in 1 mL MeOH.

Protein production and purification

E. coli BL21 (DE3) was transformed with protein production plasmids (Table S3). Protein production was achieved using an auto-induction protocol.⁵ Therefore 1 L LB-medium was supplemented with 20 mL 50 x 5052 (25% glycerol, 2.5% glucose, and 10% α -lactose monohydrate), 50 mL 20 x NPS (1 M Na₂HPO₄, 1 M KH₂PO₄, and 0.5 M (NH₄)₂SO₄), 1 mL 1 M MgSO₄ *7 H₂O and the respective antibiotics (35 μ g/mL chloramphenicol, 50 μ g/mL kanamycin, 50 μ g/mL Spectinomycin). This auto-induction medium was inoculated with the appropriate pre-culture and incubated at 37°C and 180 rpm to an OD₆₀₀ = 0.4-0.8, cooled down for 15-45 min at 4°C and incubated overnight at 25°C and 180 rpm.

Purification of Strep-tagged StIB, StIE, FabB, pIFabF, pl FabH, pIFabG, pIFabA, pIFabZ. Cells were harvested and resuspended in Strep-tag binding buffer (400 mM NaCl, 50 mM Tris, pH 8, pH 8.5 StIB, pH 7.4 StIC) and with the addition of one tablet cComplete Protease Inhibitor Cocktail (Roche) and lysozyme incubated for 20-50 min at 4°C. After cell lysis by sonication, cell debris was removed by centrifugation (35 min, 20,000 rpm, 4°C) and the supernatant was passed over a 5 mL StrepTrap HP column (GE Healthcare) by an ÄKTApurifier system (GE Healthcare). Proteins were eluted with Strep-tag elution buffer (400 mM NaCl, 50 mM Tris, 2.5 mM d-desthiobiotin, pH 8, pH 8.5 StIB, pH 7.4 StIC). All proteins were buffer exchanged in storage buffer (20 mM HEPES, 1 mM DTT, 25% (v/v) glycerol, pH 8), concentrated to 2-10 mg/ml using Amicon concentration devices (Merck, 3,000 or 10,000 Da cut-off filter) and stored in aliquots at -80 °C until needed. StIB was directly used after purification.

Purification of His₆-tagged Sfp-SUMO. Cells were harvested and resuspended in His-tag binding buffer (300 mM NaCl, 20 mM Tris, 20 mM imidazol, pH 8.5) and with addition of one tablet cComplete Protease Inhibitor Cocktail (Roche) and lysozyme

incubated for 20-50 min at 4°C. After cell lysis by sonication, cell debris was removed by centrifugation (35 min, 20,000 rpm, 4°C) and the supernatant was passed over a 5 mL HisTrap FF column (GE Healthcare) by an ÄKTApurifier system (GE Healthcare). Proteins were eluted with His-tag elution buffer (300 mM NaCl, 20 mM Tris, 500 mM imidazol, pH 8.5) and buffer exchanged in storage buffer (20 mM HEPES, 1 mM DTT, 25% (v/v) glycerol, pH 8.5) using a PD-10 desalting column (Amersham Biosciences). Proteins were further concentrated to 2-10 mg/ml using Amicon concentration devices (Merck, 3,000 Da, 10,000 Da, or 30,000 Da cut-off filter, respectively) and stored in aliquots at -80 °C until needed.

Proteins were separated on 10-15% SDS-polyacrylamide gels and visualized with Coomassie Brilliant Blue G250.

***In vitro* reactions**

CoA ligase activity assay

Reaction was performed in reaction buffer (10 mM Tris, 1 mM DTT, pH 8.5) with 1 mM CoA, 2 mM ATP, 5 mM MgCl₂, 1 mM acid-substrate (Table S5) (solubilized as 10 mM stock in ethanol), 25 μM StIB (in Strep-elution buffer) and incubated for 1 h at 37°C. Reactions were analyzed by HR-HPLC-MS.

AasS activity of StIB

StIE (ACP) or ecACP (AcpP) was buffer exchanged in reaction buffer (10 mM Tris, 1 mM DTT, pH 8.5) with PD SpinTrap G-25 columns (Amersham Biosciences) and reactions were performed using 50 μM ACP, 1 mM CoA, 0.5 μM Sfp-SUMO (in storage buffer), 2 mM ATP, 5 mM MgCl₂, 1 mM acid-substrate (Table S5) (solubilized as 10 mM stock in ethanol), 5-15 μM StIB (in storage buffer) and incubated over night at 30°C. Reactions were analyzed by HR-HPLC-MS.

Preparation of cinnamoyl (starter-ACP) and malonyl-ACP (elongation-ACP)

ACP StIE was buffer exchanged in reaction buffer (10 mM Tris, 1 mM DTT, pH 8.5) with PD SpinTrap G-25 columns (Amersham Biosciences). Reactions were performed using 50 μ M ACP, 1 mM cinnamoyl- or malonyl-CoA, 0.5 μ M Sfp-SUMO (in storage buffer) and incubated for 1 h at 30°C. Reactions were analyzed by HR-HPLC-MS.

Chain elongation, ketoreduction and dehydration with FAS II enzymes

Starter-ACP and elongation-ACP assay were mixed 1:2 and chain elongation reaction was performed by adding 10 μ M FabH or FabB or FabF and further incubated for 1 h at 30°C, prior to adding 10 μ M FabG and 1 mM NADPH for ketoreduction reaction and 10 μ M FabA or FabZ for dehydration reaction.

All ACP-bound *in vitro* reactions were analyzed using HR-HPLC-MS. MS¹ and MS² BPC and TIC spectra were summarized according to the retention time of modified ACP. If the retention time of two ACP species differed more than 0.4 min, the separately summarized spectra were overlaid. Theoretical average masses of modified ACPs were predicted using IsotopePattern (Bruker).

HR-HPLC-UV/MS

Click-reactions and XAD extracts of TT01 and TT01 mutants as well as *in vitro*-assays of CoA ligase and ACP-bound reactions were analyzed via high resolution (HR)-HPLC-ESI-UV-MS using a Dionex Ultimate 3000 LC system (Thermo Fisher) equipped with a DAD (Impact II) or MWD (micrOTOF II)-3000 RS UV-detector (Thermo Fisher) and coupled to an Impact II or micrOTOF II electrospray ionization mass spectrometer (Bruker).

Click-reactions and XAD extracts. Separation of Click-reaction and XAD extracts was achieved on a C18 column (ACQUITY UPLC BEH, 50 mm x 2.1 mm x 1.7 μ m, Waters)

using H₂O and ACN containing 0.1% (v/v) formic acid as mobile phases. HPLC was performed at a flow rate of 0.4 mL/min with 5% ACN equilibration (0-2 min) followed by a gradient from 5-95% ACN (2-14 min, 14-15 min 95% ACN) ending with a re-equilibration step of 5% ACN (15-16 min). For internal mass calibration, 10 mM sodium formate was injected. The HPLC-MS analysis was set to negative mode with a mass range of m/z 100-1200 with and an UV at 190-800 nm.

CoA ligase-reaction (Impact II). Analysis of CoA-esters was performed on a C18 column (ACQUITY UPLC BEH, 50 mm x 2.1 mm x 1.7 μ m, Waters) using MeOH and 50 mM ammonium acetate as mobile phases. HPLC was performed at a flow rate of 0.3 mL/min with 5% ACN equilibration (0-2 min), followed by a gradient from 5-95% ACN (2-14 min, 14-15 min 95% ACN) ending with a re-equilibration step of 5% ACN (15-16 min). For internal mass calibration, 10 mM sodium formate was injected.

ACP-bound reactions (microTOF II). Analysis of ACP derivatives was performed with a C3 column (Zorbax 300SB-C3 300 \AA , 150 mm x 3.0 mm x 3.5 μ m, Agilent). ACN and H₂O containing 0.1% (v/v) formic acid were used as mobile phases at a flow rate of 0.6 mL/min. HPLC was performed with 30% ACN equilibration (0-1.5 min), followed by a gradient from 30-65% ACN (1.5-27 min) and a further elution step with 95% ACN (27-30 min). For internal mass calibration, an ESI-L Mix (Agilent) was injected.

For data analysis of UV-MS-chromatograms, Compass DataAnalysis 4.3 (Bruker) was used. The theoretical average masses of StIE and ecACP were calculated using Compass IsotopePattern 3.0 (Bruker).

MALDI-MS

In vitro CoA ligase reactions were analyzed via MALDI-MS on an LTQ Orbitrap XL instrument (Thermo Fisher Scientific, Inc.) equipped with a nitrogen laser at 337 nm. The samples were mixed with the matrix (3 mg/mL in 75% ACN with 0.1% TFA) while spotting them onto a polished stainless steel MALDI target in a 1:3 ratio (0.5 μ L sample/1.5 μ L matrix) with addition of 0.25 μ L ProteoMass Normal Mass Calibration Mix (1:10 dilution, ProteoMass™ MALDI Calibration Kit, Sigma-Aldrich) for acquisition of high resolution spectra. Mass spectra were acquired in positive ion mode (FTMS) over a range of 700-1200 m/z . MS² experiments were performed in ITMS mode with ± 0.1 Da mass accuracy window, wideband activation and 23 eV collision energy. Data analysis was performed using Xcalibur 2.0.7 (Thermo Fisher Scientific, Inc.) by averaging 30-100 consecutive laser shots.

Supporting Tables

Table S1. Strains used in this study.

Strain	Genotype	Reference
<i>E. coli</i> DH10B	F ⁻ <i>araDJ39</i> Δ(<i>ara</i> , <i>leu</i>)7697 Δ <i>lacX74 galU galK rpsL deoR</i> φ8O <i>dlacZ</i> ΔM15 <i>endAI nupG recAI mcrA</i> Δ(<i>mrr hsdRMS mcrBC</i>)	Invitrogen
<i>E. coli</i> BL21 (DE3) Star	F ⁻ , <i>ompT</i> , <i>gal</i> , <i>dcm</i> , <i>hsdSB</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻), <i>lon</i> , λ(DE3 [<i>lacI</i> , <i>lacUV5-T7</i> , <i>gene1</i> , <i>ind1</i> , <i>sam7</i> , <i>nin5</i>])	Invitrogen
S17-1 λ <i>pir</i>	Tp ^f Sm ^r <i>recA thi hsdRM</i> ⁺ , RP4-2-Tc::Mu-Km::Tn7, λ <i>pir</i> phage lysogen	[4]
<i>P. laumondii</i> TT01	wild type, Rif ^R (spontaneous)	[5]
<i>P. laumondii</i> TT01::pDS132_ <i>plu2134</i>	wild type with a P _{DS132} insertion in <i>plu2134</i> , Cm ^R	this work

Table S2. Oligonucleotides used for plasmid construction and verification. Overhangs are underlined.

Plasmid		Oligonucleotide 5' to 3'	Template
pDS132_plu2134	TS120	<u>AAAGTGAGCTCCTGGCTAAAACATTATCCGG</u>	<i>P. laumondii</i> TT01
	TS121	<u>TGCCACCAGTATATTGCAGA</u>	
pACYC_plu2134_strep (stlB)	GG185	<u>TTCGAAAAAGAGAACTATACTCCAGGGAGAGAAAGTC</u> <u>TGGCTAAAACATTATC</u>	<i>P. laumondii</i> TT01
	GG186	<u>AGCGGTGGCAGCAGCCTAGGTTAACTATGCTACATTCCT</u> <u>GACCTTTTC</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pACYC Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCOLA_plu2134_strep (stlB)	GG185	<u>TTCGAAAAAGAGAACTATACTCCAGGGAGAGAAAGTC</u> <u>TGGCTAAAACATTATC</u>	<i>P. laumondii</i> TT01
	GG186	<u>AGCGGTGGCAGCAGCCTAGGTTAACTATGCTACATTCCT</u> <u>GACCTTTTC</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pCOLA Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pACYC_fabH_Strep	SK18	<u>GAGAACCTATACTCCAGGGATATACAAAAATTTAGGTA</u> <u>CAGGTAGTTATC</u>	<i>P. laumondii</i> TT01
	SK22	<u>GATTACTTTCTGTTTCGACTTAAGCATTATTA</u> <u>AAAAACGTATCAGTGCCG</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pACYC Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCOLA_bkdABC_Strep	SK23	<u>GAGAACCTATACTCCAGGGAATAGACGTACAGATCAAT</u> <u>GAAGTACC</u>	<i>P. laumondii</i> TT01
	SK24	<u>GATTACTTTCTGTTTCGACTTAAGCATTAT</u> <u>TA</u> <u>CTGTTTTGC</u> <u>AACCAG</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pCOLA Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCDF_stlB_Strep	SK25	<u>GAGAACCTATACTCCAGGGAGAGAAAGTCTGGCTAAAA</u> <u>CATTATC</u>	<i>P. laumondii</i> TT01
	SK26	<u>GATTACTTTCTGTTTCGACTTAAGCATTACTATGCTACATT</u> <u>CCTGACCTTTTC</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pCDF Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCDF_stlB_Strep_StlCDE	SK27	<u>GTTAAGTATAAGAAGGAGATATACATATGAAACGTGTTCT</u> <u>TGTAGTGTCG</u>	<i>P. laumondii</i> TT01
	SK28	<u>GTGGCAGCAGCCTAGGTTAATTAGTTAGCGGATTCCAAC</u> <u>TTTG</u>	
	GG108	<u>CATATGTATATCTCCTTCTTATACTTAAC</u>	pCDF_stlB_Strep
	GG109	<u>TAATTAACCTAGGCTGCTGCCAC</u>	
pACYC_fabH_Strep_fabG	SK29	<u>GTTAAGTATAAGAAGGAGATATACATATGAGATTAGATG</u> <u>GAAAGATTGCATTAG</u>	<i>P. laumondii</i> TT01
	SK30	<u>GTGGCAGCAGCCTAGGTTAATTAGTTCAATATACATGCCG</u> <u>CC</u>	
	GG108	<u>CATATGTATATCTCCTTCTTATACTTAAC</u>	pACYC_fabH_ Strep
	GG109	<u>TAATTAACCTAGGCTGCTGCCAC</u>	
pACYC_darABC (stlCDE)	GG1	<u>TATAAGAAGGAGATATACATATGAAACGTGTTCTTGTAGT</u> <u>G</u>	<i>P. laumondii</i> TT01
	GG8	<u>GTTTCTTTACCAGACTCGAGTTAGTTAGCGGATTCCAAC</u>	
	GG3	<u>CTCGAGTCTGGTAAAGAAAC</u>	pACYC Duet-1
	GG4	<u>CATATGTATATCTCCTTCTTATACTTAAC</u>	
pCDF_fabH_Strep	SK18	<u>GAGAACCTATACTCCAGGGATATACAAAAATTTAGGTA</u> <u>CAGGTAGTTATC</u>	<i>P. laumondii</i> TT01
	SK19	<u>AGCGGTGGCAGCAGCCTAGGTTAATTA</u> <u>AAAAACGTATCAGTGCCG</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pCDF Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	
pCDF_fabB_Strep	SK1	<u>GAGAACCTATACTCCAGGGAAGGATTTAATGAAGCGC</u> <u>GTAG</u>	<i>P. laumondii</i> TT01
	SK2	<u>AGCGGTGGCAGCAGCCTAGGTTAATTAAGCTGAATACTT</u> <u>GCTCATCAC</u>	
	GG68	<u>TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG</u> <u>GCTCCACATGGTATATCTCCTTATTAAGTTAAAC</u>	pCDF Duet-1
	SW1	<u>TTAACCTAGGCTGCTG</u>	

Attachments

pCDF_fabF_Strep	SK3	GAGAACCTATACTTCCAGGGATCTAAGCGTCGAGTAGTT G	<i>P. laumondii</i> TT01
	SK4	AGCGGTGGCAGCAGCCTAGGTTAATTAATCAGGTTTAAT CTTACGG	
	GG68	TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG GCTCCACATGGTATATCTCCTTATTAAGTTAAAC	pCDF Duet-1
	SW1	TTAACCTAGGCTGCTG	
pACYC_fabG_Strep	SK5	GAGAACCTATACTTCCAGGGAAGATTAGATGGAAGATT GCATTAG	<i>P. laumondii</i> TT01
	SK6	AGCGGTGGCAGCAGCCTAGGTTAATTAGTTCATATACAT GCCGC	
	GG68	TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG GCTCCACATGGTATATCTCCTTATTAAGTTAAAC	pACYC Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCOLA_fabA_Strep	SK7	GAGAACCTATACTTCCAGGGAGTTGATAAACTTAAATCCT ACACAAAAG	<i>P. laumondii</i> TT01
	SK8	AGCGGTGGCAGCAGCCTAGGTTAATTAGAAAAGCGCTGG TATCTTTAAAC	
	GG68	TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG GCTCCACATGGTATATCTCCTTATTAAGTTAAAC	pCOLA Duet-1
	SW1	TTAACCTAGGCTGCTG	
pCOLA_fabZ_Strep	SK9	GAGAACCTATACTTCCAGGGAAGTGATAATCATACTCTG CACATTG	<i>P. laumondii</i> TT01
	SK10	AGCGGTGGCAGCAGCCTAGGTTAACTAAACCTCACGAC GGC	
	GG68	TCCCTGGAAGTATAGGTTCTCTTTTTCGAACTGCGGGTG GCTCCACATGGTATATCTCCTTATTAAGTTAAAC	pCOLA Duet-1
	SW1	TTAACCTAGGCTGCTG	
Verification of <i>P. laumondii</i> TT01 <i>plu2134</i> - insertion mutant			
Oligonucleotide 5' to 3'			
TS121	TCTCAGCTAGCAGCATCAAT		
pDS132_rev	GATCGATCCTCTAGAGTCGACCT		

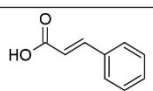
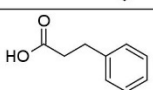
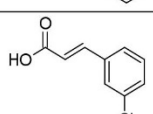
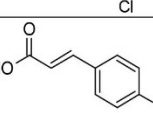
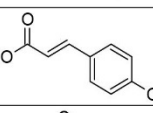
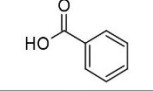
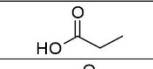
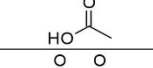
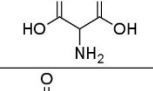
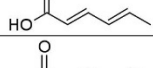
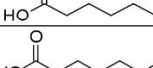
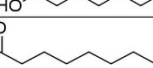
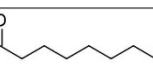
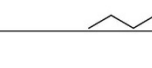
Table S3. Plasmids used in this study.

Plasmid	Genotype	Reference
pACYC Duet-1	P15A ori, T7lac promoter, Cm ^r	Novagen
pCDF Duet-1	CDF ori, T7lac promoter, Sm ^r	Novagen
pCOLA Duet-1	ColA ori, T7lac promoter, Km ^r	Novagen
pCATI4	ColA ori, T7lac promoter, <i>cherry</i> , Km ^r	[6]
pDS132	R6Kγ ori, <i>oriT</i> , <i>sacB</i> , <i>traJ</i> , <i>mob RP4</i> , Cm ^r	[7]
pSUMO_sfp	ColE1 ori, T7lac promoter, <i>sumo</i> , <i>sfp</i> , Km ^r	[8]
pET15b-AcpP	pBR322 ori, T7lac promoter, <i>acpP</i> , Amp ^r	Charles. O. Rock
pDS132_plu2134	Insertion plasmid based on pDS132 with 578 bp of <i>plu2134</i> , Cm ^r	this work
pACYC_plu2134_strep	P15A ori, T7lac promoter, <i>plu2134 (stlB)</i> , Cm ^r	this work
pCOLA_plu2134_strep	ColA ori, T7lac promoter, <i>plu2134 (stlB)</i> , Km ^r	this work
pCOLA_darC_strep	ColA ori, T7lac promoter, <i>plu2165 (stlE)</i> , Km ^r	this work
pCDF_darA_Strep_BC (<i>stlCDE</i>)	CDF ori, T7lac promoter, <i>plu2163-2165 (stlCDE)</i> Sm ^r	this work
pACYC_darA_Strep_B (<i>stlCD</i>)	P15A ori, T7lac promoter, <i>plu2163, plu2164 (stlCD)</i> , Cm ^r	this work
pACYC_darA_Strep (<i>stlC</i>)	P15A ori, T7lac promoter, <i>plu2163 (stlC)</i> , Cm ^r	this work
pCDF_darC_Strep (<i>stlE</i>)	CDF ori, T7lac promoter, <i>plu2165 (stlE)</i> , Sm ^r	this work
pCDF_darC_Strep_darB (<i>stlDE</i>)	CDF ori, T7lac promoter, <i>plu2164 (stlD) plu2165 (stlE)</i> , Sm ^r	this work
pACYC_darABC (<i>stlCDE</i>)	P15A ori, T7lac promoter, <i>plu2163-2165 (stlC)</i> , Cm ^r	this work
pCDF_fabH_Strep	CDF ori, T7lac promoter, <i>plu2835</i> , Sm ^r	this work
pCDF_fabB_Strep	CDF ori, T7lac promoter, <i>plu3184</i> , Sm ^r	this work
pCDF_fabF_Strep	CDF ori, T7lac promoter, <i>plu2831</i> , Sm ^r	this work
pACYC_fabG_Strep	P15A ori, T7lac promoter, <i>plu2833</i> , Cm ^r	this work
pCOLA_fabA_Strep	ColA ori, T7lac promoter, <i>plu1772</i> , Km ^r	this work
pCOLA_fabZ_Strep	ColA ori, T7lac promoter, <i>plu0683</i> , Km ^r	this work
pCDF_stlB_Strep_stlCDE	CDF ori, T7lac promoter, <i>plu2134 (stlB), plu2163-2165 (stlCDE)</i> , Sm ^r	this work
pACYC_fabH_Strep	P15A ori, T7lac promoter, <i>plu2835</i> , Cm ^r	this work
pCOLA_bkdABC_Strep	ColA ori, T7lac promoter, <i>plu1883-1885</i> , Km ^r	this work
pACYC_fabH_Strep_fabG	P15A ori, T7lac promoter, <i>plu2835 (fabH), plu2833 (fabG)</i> Cm ^r	this work

Table S4. *E. coli* BL21 (DE3) Star strains used for heterologous stilbene production.

Name	Introduced plasmids
StlCDE	pACYC_darABC (<i>stlCDE</i>)
StlCDE+ <i>p/FabH</i>	pACYC_darABC (<i>stlCDE</i>), pCDF_fabH_Strep
StlBCDE	pCDF_stlB_Strep_stlCDE
StlBCDE+ <i>p/FabH</i>	pCDF_stlB_Strep_stlCDE, pACYC_fabH_Strep
StlBCDE+ <i>p/FabH</i> +BkdABC	pCDF_stlB_Strep_stlCDE, pACYC_fabH_Strep, pCOLA_bkdABC_Strep
StlBCDE+ <i>p/FabH</i> +BkdABC+ <i>p/FabG</i>	pCDF_stlB_Strep_stlCDE, pACYC_fabH_Strep_fabG, pCOLA_bkdABC_Strep

Table S5. Substrate specificity of StIB from *P. laumondii* TT01 by its CoA ligase reaction with substrates **1-16**. CoA ester formation was analyzed via MALDI-MS (Fig. S1).

substrate	structure	accepted
1		+
2		+
3		+
4		-
5		-
6		-
7		-
8		-
9		-
10		+
11		+
12		+
13		+
14		+

Supporting Figures

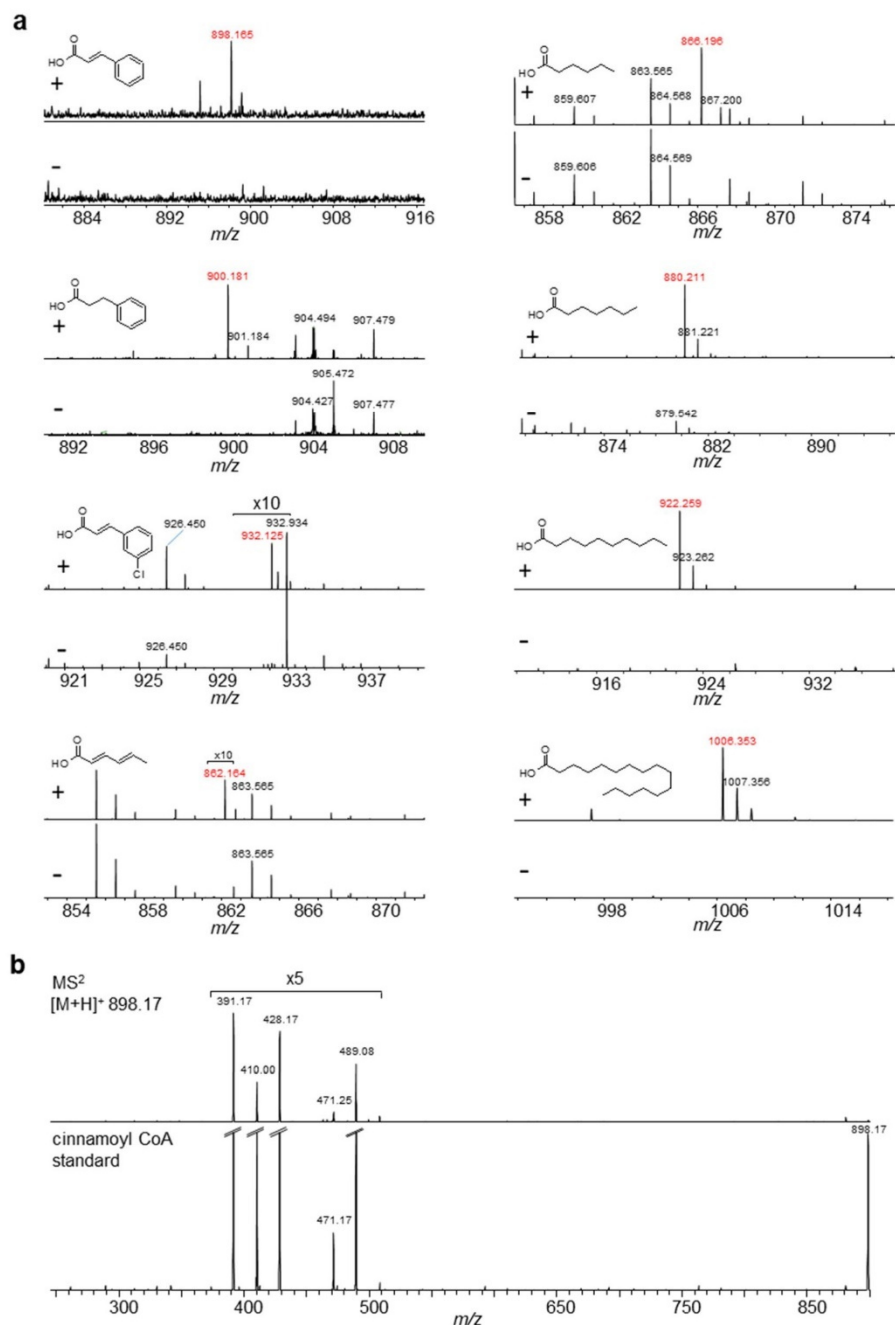


Figure S1. Detection of CoA-derivatives built upon the CoA-ligase activity of StIB. CoA ligase reaction was tested for substrates **1-16** (Table 1), CoA-esters were detected by *high resolution* MALDI-MS (**a**). Due to the weak signal of cinnamoyl-CoA in MS¹, a further MS²-experiment was performed in comparison with a cinnamoyl-CoA standard (**b**).

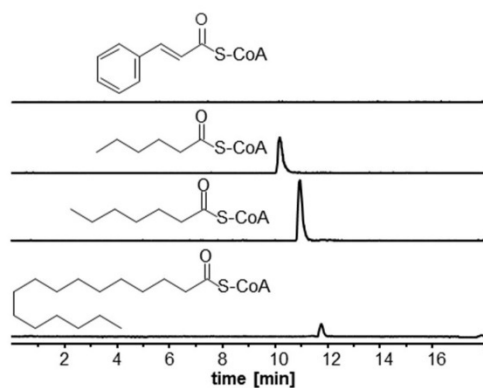


Figure S2. CoA-ligase reaction of StIB. Caproic (C₆)-, heptanoic (C₇)- and palmitic (C₁₆) acid were converted to their corresponding CoA ester by the CoA ligase activity of StIB. The conversion of cinnamic acid to cinnamoyl CoA was not detectable in this setup.

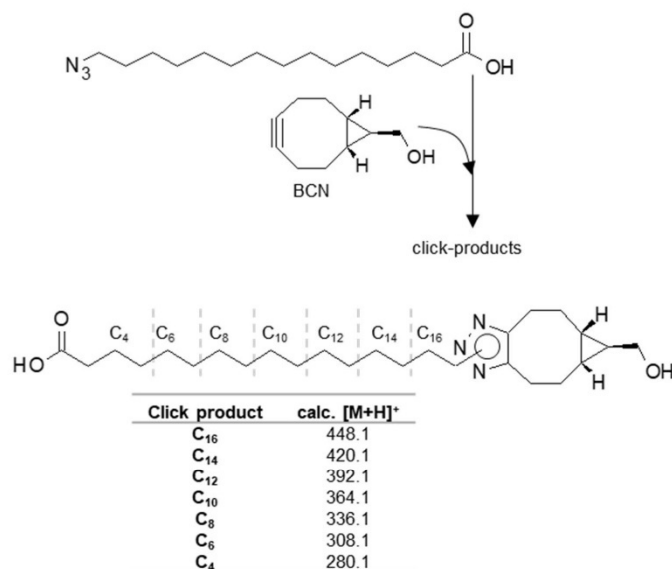


Figure S3. Expected click-products (table) after click-reaction with azido palmitic acid and BCN.^[3] HPLC-data, see Fig. S4

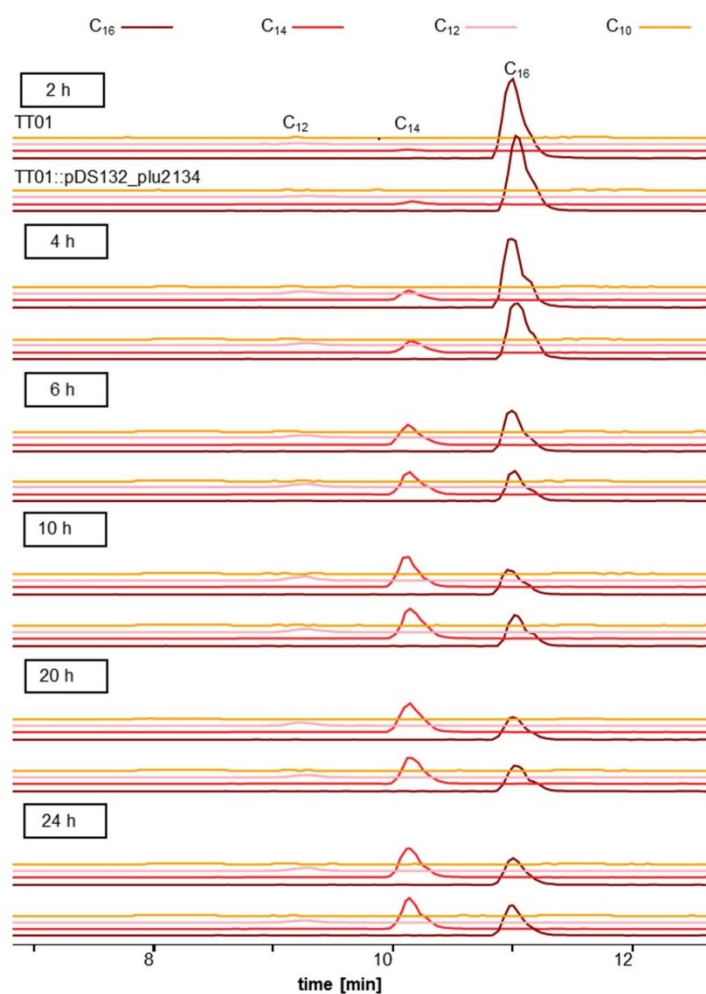


Figure S4. Fatty acid degradation profile of *P. laumondii* TT01 mutants compared to the wildtype. TT01 was fed with azido-labelled C₁₆-fatty acid and a click-reaction with BCN was performed to detect the degradation products (C₁₄, C₁₂, C₁₀, C₈) after 2 h, 4 h, 6 h, 10 h, 20 h and 24 h.

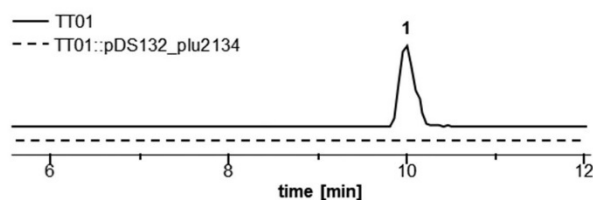


Figure S5. Growth of *P. laumondii* TT01 and insertion mutant TT01::pDS132_plu2134 without antibiotic. Plasmid insertion of TT01::pDS132_plu2134 was verified by the loss of **1** production. **1** in TT01 was detected via HPLC-UV-MS at 9.9 min with a mass of 255.1 [M+H]⁺ and its characteristic UV.

S16

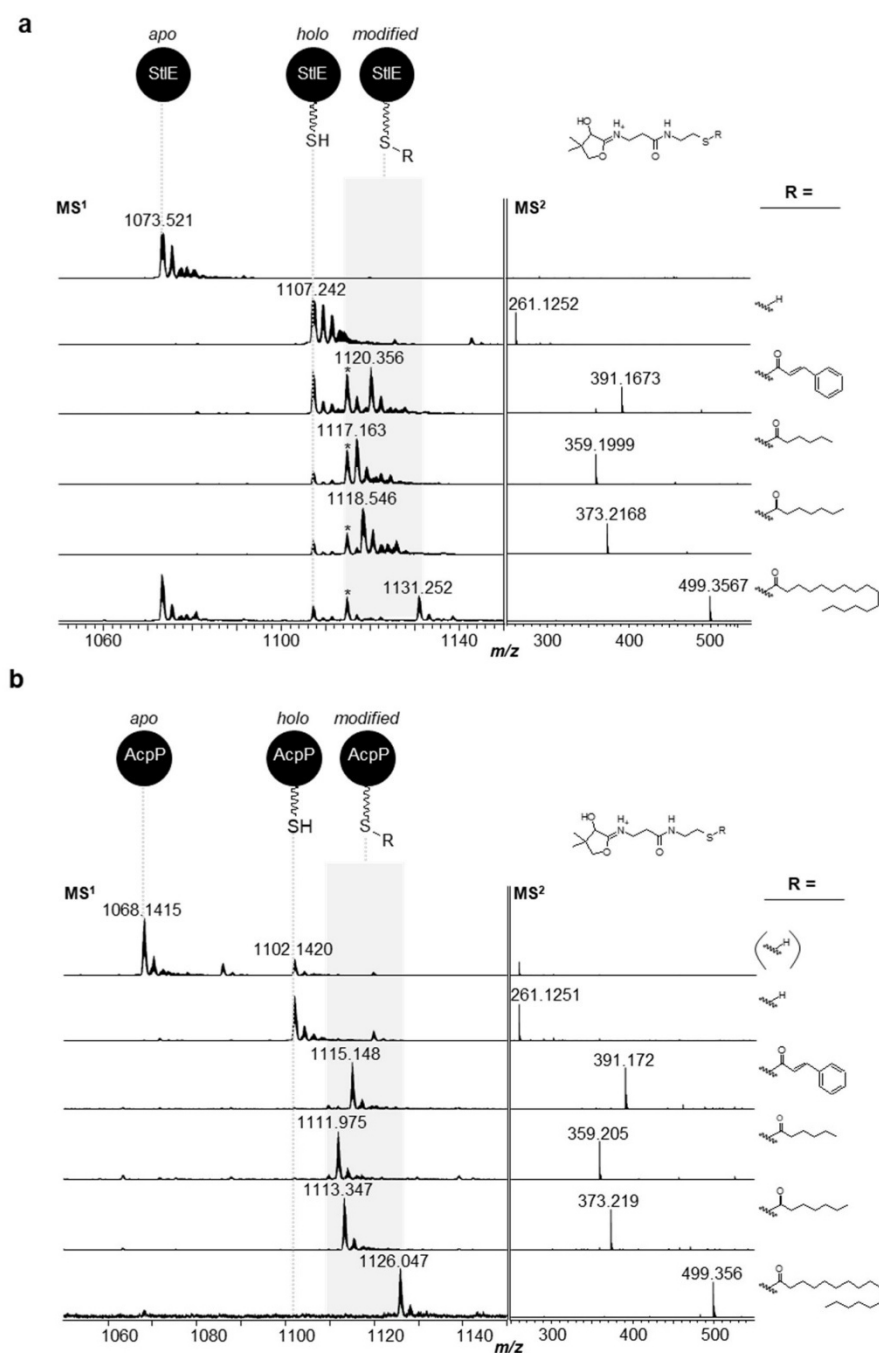


Figure S6. Loading of StIE (**a**) and AcpP (**b**) by AasS reaction of StIB with cinnamic, caproic, heptanoic and palmitic acid. The ACPs StIE and AcpP were transferred to their *holo*-forms by means of Sfp. In MS¹, the 10+ charge state of the ACPs is shown and in MS², the corresponding Ppant ejection ions are displayed. In the case of StIE loading reactions, a side product was detected during overnight reaction (corresponding signals are marked with an asterisk).

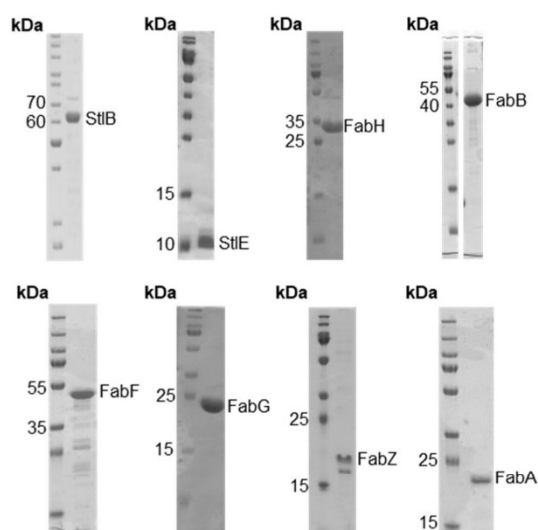


Figure S7. SDS-PAGE of strep-tagged SttB, SttE and FabH, FabB, FabF, FabG, FabZ, FabA.

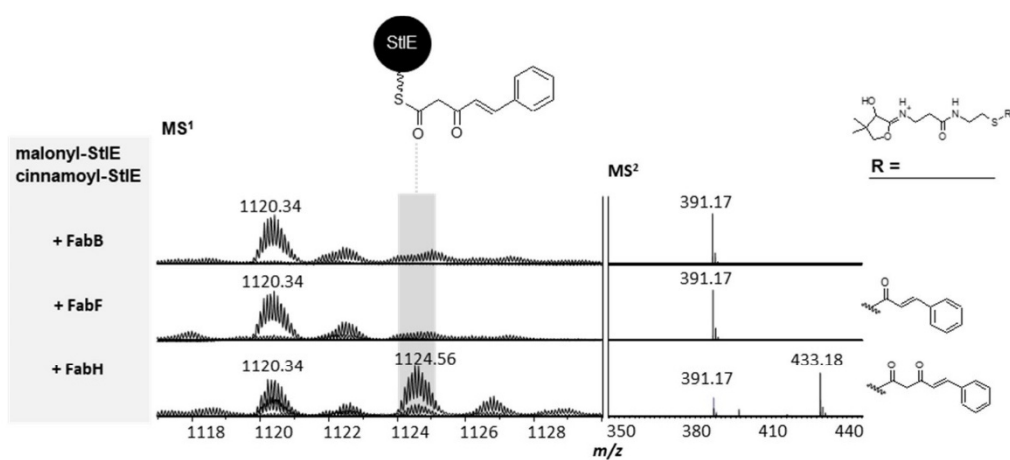


Figure S8. *In vitro* elongation with FAS II enzymes FabH, FabB and FabF and cinnamoyl-ACP as starter and malonyl-ACP as elongation unit. Spectra were overlaid as described in the experimental section.

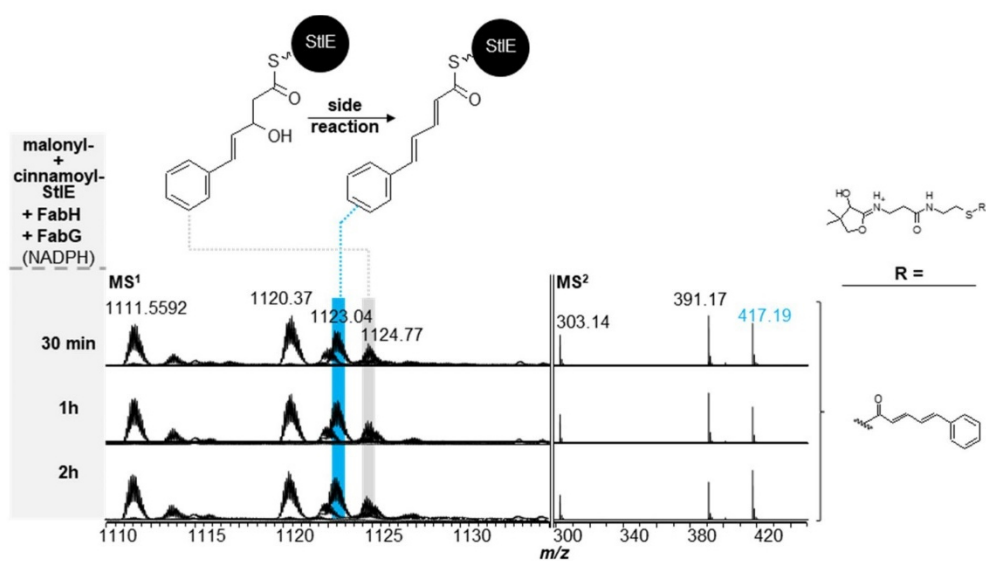


Figure S9. Time dependent *in vitro* analysis of side reaction with cinnamoyl-ACP as starter and malonyl-ACP as elongation unit being elongated and reduced by FabH and FabG (+NADPH), respectively. Spectra were overlaid as described in the experimental section. The dehydrated side product did not increase with increased reaction time.

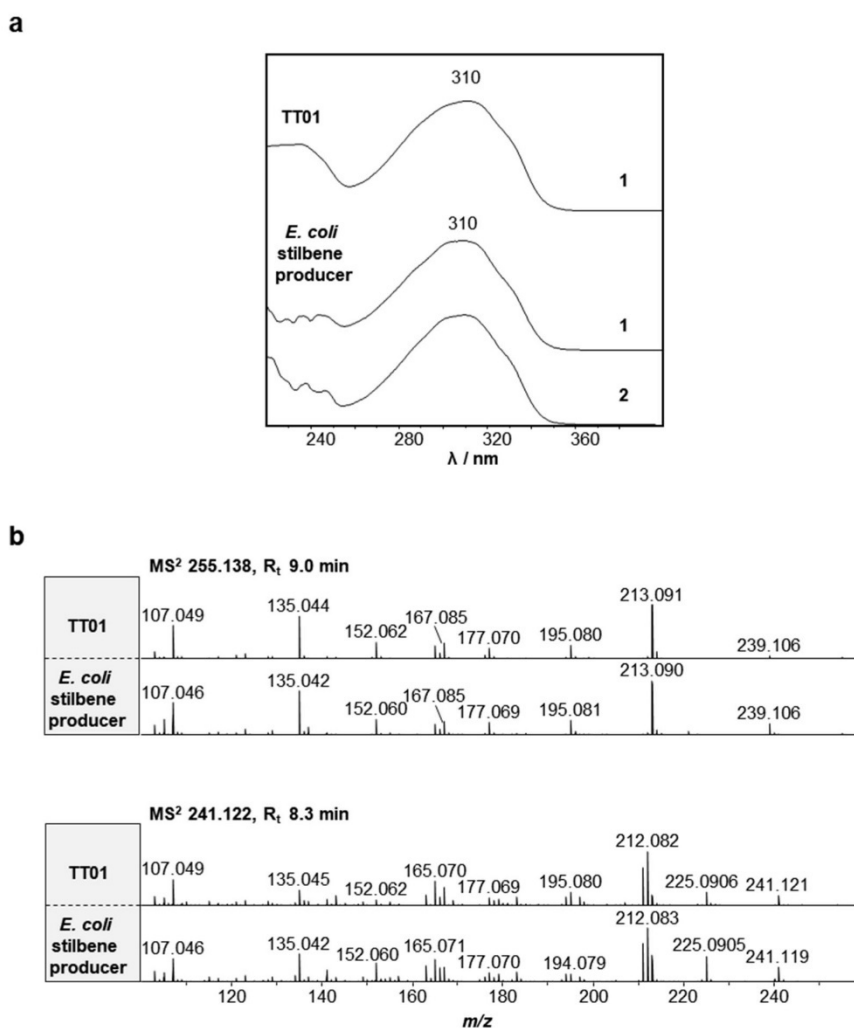


Figure S10. Heterologous stilbene production in *E. coli*. Analysis of production of **1** and **2** from crude extracts after cultivation of *E. coli* BL21 with introduced StIBCDE, *pIFabH*, *bkdABC* and *pIFabG* (*E. coli* stilbene producer) and cinnamic acid supplementation in comparison to TT01 wildtype extract. **1** and **2** were detected via HPLC-UV-MS at 9.0 min and 8.3 min with a mass of 255.138 [M+H]⁺ and 241.122 [M+H]⁺ by their characteristic UV (**a**) and fragmentation pattern (**b**), respectively.

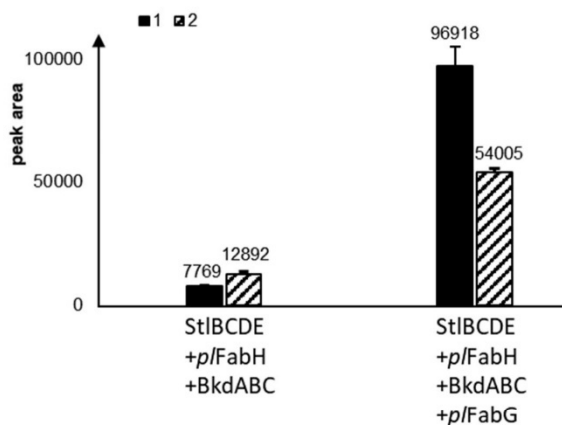


Figure S11. Comparison of heterologous *E. coli* stilbene production strains. HPLC-UV/MS analysis of crude extract from *E. coli* StIBCDE with *p/FabH* and *BkdABC* and *E. coli* StIBCDE with *p/FabH* and *BkdABC* with additional *p/FabG*. Peak areas of **1** and **2** of respective EICs were compared.

References

- [1] C. Fu, W. P. Donovan, O. Shikapwashya-Hasser, X. Ye, R. H. Cole, *PLoS One* **2014**, *9*, e115318.
- [2] A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBiochem* **2007**, *8*, 1721.
- [3] A. J. Pérez, H. B. Bode, *ChemBiochem* **2015**, *16*, 1588.
- [4] R. Simon, U. Priefer, A. Pühler, *Nat Biotechnol* **1983**, *1*, 784.
- [5] M. Fischer-Le Saux, V. Viillard, B. Brunel, P. Normand, N. E. Boemare, *Int J Syst Bacteriol* **1999**, *49*, 1645.
- [6] T. A. Schöner, S. W. Fuchs, B. Reinhold-Hurek, H. B. Bode, *PLoS One* **2014**, *9*, e90922.
- [7] N. Philippe, J.-P. Alcaraz, E. Coursange, J. Geiselmann, D. Schneider, *Plasmid* **2004**, *51*, 246.
- [8] W. Lorenzen, T. Ahrendt, K. A. J. Bozhüyük, H. B. Bode, *Nat. Chem. Biol.* **2014**, *10*, 425.

12.3 Manuscript: Anthraquinone Production is Influenced by Cinnamic Acid

Declaration on the contribution of the authors/Erklärung zu den Autorenanteilen

Status: **submitted**

Beteiligte Autoren: Gina L. C. Grammbitter (GLG), Helena Vural (HV), Nicholas J. Tobias (NJT), Helge B. Bode (HBB)

Was hat der/die Promovierende bzw. was haben die Co-Autoren/Autorinnen beigetragen?

(1) zu Entwicklung und Planung

GG: 30%

NJT, HV: 20%

HBB: 50%

(2) Durchführung der einzelnen Untersuchungen und Experimente

Fütterungsexperimente mit Zimtsäure und entsprechende HPLC-ESI-MS-Analyse GG (55%)
Erzeugung TT01-Mutante, RNA-Extraktion HV (45%)

(3) Erstellung der Datensammlung und Abbildungen

Darstellung Biosyntheschema, Fütterungsexperiment mit Zimtsäure und *in silico*-Inhibitor-Vorhersage GG (70%)

RNA-Sequenzierung NJT (30%)

(4) Analyse und Interpretation der Daten

Fütterungsexperimente mit Zimtsäure und entsprechende HPLC-ESI-MS-Analyse GG (70%)
RNA-Sequenzierung NJT (30%)

(5) Verfassung des Manuskripts

GLC: 45%

NJT: 10%

HBB: 45%

date/place: _____

Anthraquinone Production is Influenced by Cinnamic Acid

Gina L. C. Grammbitter¹, Helena Vural¹, Nicholas J. Tobias^{1,2}, Helge B. Bode^{1,2,3*}

¹ Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe-Universität Frankfurt, Frankfurt am Main, Germany

² LOEWE Center for Translational Biodiversity in Genomics (TBG), Germany

³ Buchmann Institute for Molecular Life Sciences (BMLS), Goethe-Universität Frankfurt, Frankfurt am Main, Germany

*corresponding author

h.bode@bio.uni-frankfurt.de (HB)

Abstract

Photorhabdus laumondii produces the orange-to-red pigmented anthraquinones (AQs) during the exponential growth phase through the action of type II polyketide synthase. We observed an inverse correlation between the amount of cinnamic acid (CA) used as a precursor for isopropylstilbene (IPS) and AQ production. A detailed analysis of wild type and $\Delta sttA$ strains, not producing CA, indeed suggested an inhibition of AntI, a hydrolase involved in a polyketide shortening mechanism that involves the conversion of an octa- to a heptaketide through retro-Claisen and Dieckmann reactions. Since no general regulatory effects could be observed for CA in RNAseq experiments, we propose a covalent inhibition of AntI via the Michael acceptor present in CA. The endogenous CA therefore might be a direct regulator for AQ biosynthesis.

Introduction

Photorhabdus are Gram-negative bacteria that live in mutualistic symbiosis with *Heterorhabditis* nematodes, whilst being pathogenic to insect larvae [1,2]. *Photorhabdus* harbors an elaborate specialized metabolism, producing a vast range of compounds such as antibiotics, pigments and molecules for bioluminescence [1,3]. Some of these molecules have been shown to be important for symbiosis, whilst others are involved in quorum sensing or protecting the insect cadaver from predators such as other bacteria [1,4]. Due to the production of anthraquinone (AQ) pigments, *Photorhabdus* develops its characteristic orange-to-red coloration during the exponential growth-phase [5,6]. AQs are aromatic polyketides produced by a type II polyketide synthase (PKS). The corresponding gene cluster in *P. laumondii* subsp. *laumondii* TTO1 (TTO1, renamed from previous *P. luminescens* subsp. *laumondii* TTO1 [7]) (*antA-I*, PLU_RS20740-20700, previously *plu4194-4186*) was well-studied in the past and encodes for a ketoreductase (KR), a PPTase, a cyclase (CYC), the 'minimal PKS' with a ketosynthase (KS), a chain-length factor (CLF) and an acyl-carrier protein (ACP), a CoA-ligase, a CYC/aromatase (ARO) and a hydrolase, respectively [5,8]. In previous studies, it was demonstrated that the minimal PKS requires the co-occurrence of the most-likely inactive CoA-ligase AntG in combination with the PPTase AntB to synthesize the ACP-bound octaketide (Fig. 1) [8,9]. After cyclization, the ACP-bound octaketide is hydrolyzed and at the same time, shortened by the hydrolase AntI to result in the main AQ-product **1**. The subsequent oxidation might occur spontaneously [8] and methylation by unknown methyltransferases results in the methylated derivatives **2** and **3** [5].

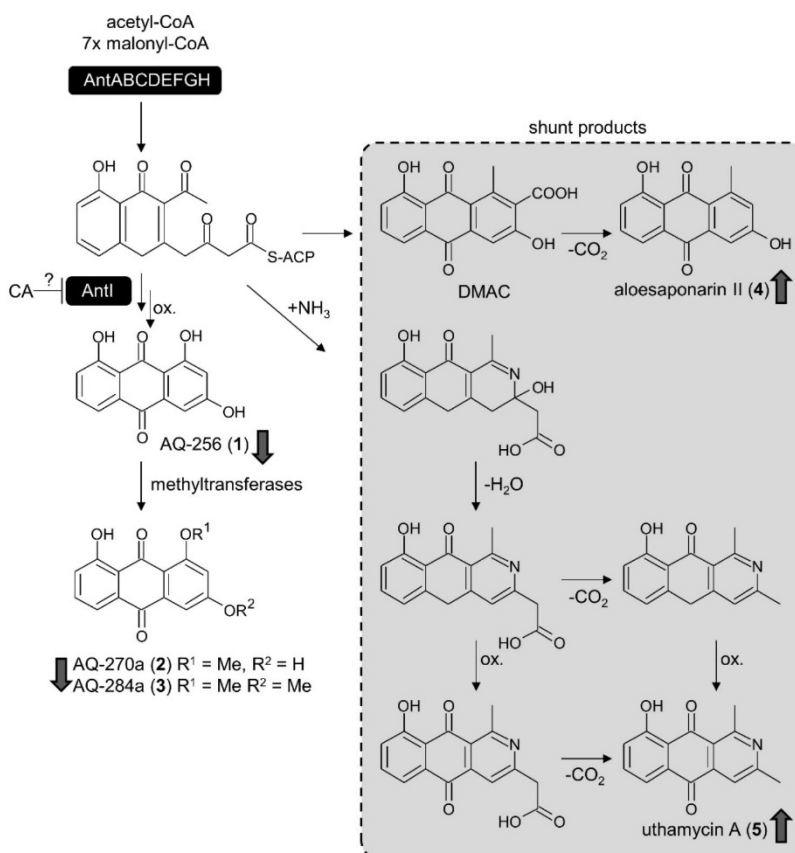


Figure 1. Anthraquinone biosynthesis [8]. Bold arrows suggest a decrease and increase of the respective compounds upon postulated inhibition of *AntI*. 3,8-dihydroxy-1-methylantraquinone-2-carboxylic acid (DMAC) [10].

Another key metabolite in TTO1 is isopropylstilbene (IPS), which has multiple biological functions^[11–13] and was recently approved as benvitimod^[14] or tapinarof^[15] against psoriasis. IPS biosynthesis requires cinnamic acid that is derived from phenylalanine via the phenylalanine ammonia lyase (PAL) *StIA*.^[11] In general, CA was described to be secreted by the bacterium at the end of the exponential growth phase^[16]. Furthermore, the abundance of CA increases if *st/A* expression is increased; *st/A* expression, in turn, is strictly regulated by a TyrR transcription factor and additionally activated by σ^S and Lrp regulators upon nutrient limitation^[17].

Results and Discussion

During our analysis of the IPS biosynthesis^[11,18], we observed dark red colonies on agar plates for TTO1 strains having a *stA* deletion. We also observed this effect in liquid culture and furthermore could restore the wild type orange color after adding cinnamic acid (CA) to the culture (Fig. 2). Since a ΔstA mutant does neither produce CA nor IPS, we also tested the effect of IPS addition but could not restore the wild type coloration (data not shown). The result from the CA addition suggested CA as a direct or indirect regulator of AQ biosynthesis and when we compared the amounts of AQs **1-3** in wild type or ΔstA , we could indeed observe an overall decrease of AQ production upon addition of CA. In addition to the general decrease in AQ-production (**1-3**) in both strains, some differences with respect to the product ratios of the main AQs were observed (Fig. 2). Upon CA supplementation, AQ-product formation in TTO1 and TTO1 ΔstA decreased strongest for AQ **1**, with a minor decrease seen for the methylated derivatives **2** and **3**.

Additionally, the known shunt products aloesaponarin II (**4**) and utahmycin A (**5**), which are derived from the last ACP-bound intermediate in AQ biosynthesis (Fig. 1) and that are also observed in strains not containing AntI^[8], showed an increased amount when CA was present (Fig 2).

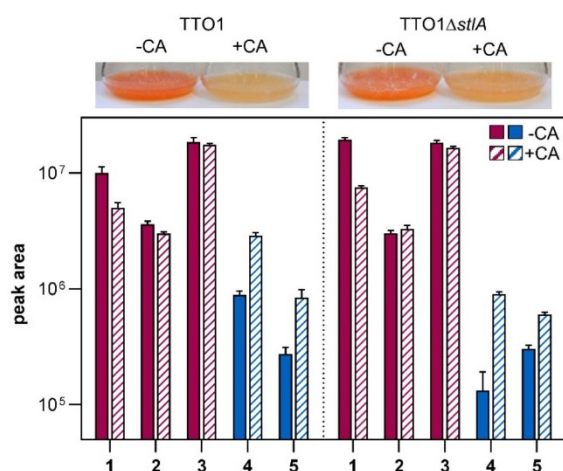


Figure 2. The feeding of CA resulted in less pigmented strains (top). HR-HPLC-MS based quantification of AQs 1-3 (magenta) and shunt products 4 and 5 (blue) after feeding with CA (below).

In *Photorhabdus*, CA is reported not just as a building block for IPS biosynthesis but can also be degraded by the phenylpropionate/cinnamate catabolic pathway, which can catabolyze CA to use it as a carbon source. HcaE is a key enzyme involved in the degradation of CA in the Hca pathway, with expression significantly increased upon addition of exogenous CA to TTO1 cultures [16]. The *hca* operon is under the control of the regulator HcaR that does not bind CA directly. Nevertheless, this might suggest other transcription factors influenced by CA.

To rule out any effects at the regulatory level, we performed RNAseq on TTO1 + EtOH (containing both CA and IPS) and TTO1ΔstIA supplemented with EtOH or IPS (dissolved in EtOH). The culture of TTO1ΔstIA + IPS therefore only lacks CA. If regulatory alterations were present, we would expect to see differences in the AQ-biosynthesis pathway in strains lacking CA (TTO1ΔstIA supplemented with IPS) compared to the wild type. IPS was recently reported to effect TTO1 pigmentation [13], but under the used conditions we were not able to see any colour change and furthermore only see significant differences in eight genes, none of which are

associated with AQ-biosynthesis (Supplementary Fig. 1). Thus, we speculate that CA might have a direct effect on enzymes involved in AQ biosynthesis and from the observed effect of CA on the production of **1-5**, AntI might be CA target (Fig. 3).

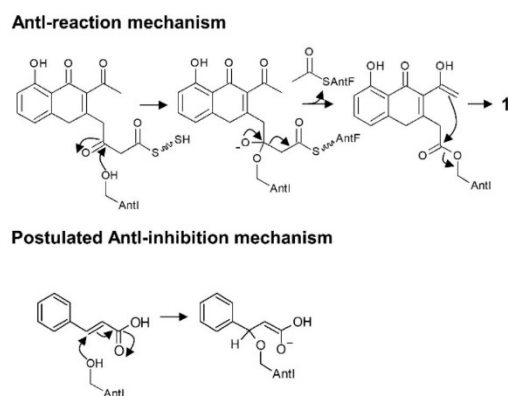


Figure 3. Reaction mechanism of AntI [8] (top). CA is proposed to inhibit AntI acting as a Michael-acceptor (bottom).

Conclusion

According to our recent work, which describes the molecular mechanism catalyzed by AntI [8], we proposed CA to act as an inhibitor acting as a Michael acceptor. Michael-acceptors are well-known known for inhibiting cysteine-proteases by covalently binding the active site cysteine [19]. Likewise, cinnamic aldehyde was described to impair melanoma cell proliferation and tumor growth by acting as a Michael-acceptor [20]. Although we were not able to obtain co-crystals of AntI with CA, we were able to show that CA influences AQ-production. The well-known shunt products aloesaponarin II (**4**) and uthamycin A (**5**) arise from the cyclized, ACP-bound, octaketide scaffold, which in turn forms the substrate for off-loading the AQ-scaffold by AntI. Due to the accumulation of those shunt products, we suggest that AntI is inhibited by CA by conjugate-addition to the conserved serine of AntI and may therefore be involved in the regulation of AQ-biosynthesis (Fig. 3). The fact that the production of methylated AQ derivatives was not as severely affected as that of the initial AQ-256 (**1**) might suggest a fast conversion of **1** into **2** and **3**, leading to similar levels of both methylated derivatives with and without CA. Currently, we try to identify the responsible methyltransferases and furthermore try to get co-crystal or other structural data showing the interaction between AntI and CA to support the observed results and our hypothesis.

Experimental section

Cultivation and metabolite extraction of *Photorhabdus*

30 mL of TTO1 [21] or TTO1 Δ stIA [11] XAD-cultures were incubated at 30°C at 180 rpm on a rotary shaker using 300 mL Erlenmeyer-flasks containing Luria-Bertani (LB) broth-medium supplemented with Amberlite XAD-16-beads (2% v/v, Sigma-Aldrich). Cultures were inoculated with a 24 h preculture to an OD₆₀₀ 0.1 and harvested following

7

72 h of incubation, reaching an OD₆₀₀ of 10-12. Cinnamic acid (CA, Sigma-Aldrich) was supplemented to a final concentration of 1 mM (100 mM CA stock solution in ethanol). Cultures not supplemented with CA had an equal volume of ethanol added. XAD-beads from each strain were separated from the supernatant by decanting. The beads were further incubated with 30 mL of MeOH under gentle agitation for 1 h at RT. The resulting extracts were dried under reduced pressure, resuspended in 3 mL of MeOH and subjected to HPLC-UV-MS analysis.

HR-HPLC-UV-MS

XAD extracts of TTO1 and TTO1-mutants were analyzed via high resolution (HR)-HPLC(-ESI)-UV-MS using a Dionex Ultimate 3000 LC system (Thermo Fisher) equipped with a DAD-3000 RS UV-detector (Thermo Fisher) coupled to an Impact II electrospray ionization mass spectrometer (ESI-MS) (Bruker).

Compound separation of XAD extracts was achieved on a C18-column (ACQUITY UPLC BEH, 50 mm x 2.1 mm x 1.7 μm, Waters) using H₂O and ACN containing 0.1% (v/v) as mobile phases. HPLC was performed at a flow rate of 0.4 mL/min with 5% ACN equilibration (0-2 min), followed by a gradient from 5-95% ACN (2-14 min), 14-15 min 95% ACN) and a re-equilibration step of 5% ACN (15-16 min). For internal mass calibration, 10 mM sodium formate was injected. The HPLC-UV-MS analysis was set to negative or positive mode with a mass range of *m/z* 100-700 with and an UV detection at 190-800 nm. AQ **1**, **2**, **3** and aloesaponarin II (**4**) were detected in negative mode, uthamycin A (**5**) was detected in positive mode.

RNA extraction and sequencing

For RNA extractions, 10 mL cultures were inoculated with overnight cultures to a starting OD₆₀₀ of 0.3. Once cultures reached OD₆₀₀ 5.0, RNA was extracted using the Qiagen RNeasy Plus Mini Kit following the manufacturer's recommendations. IPS was

8

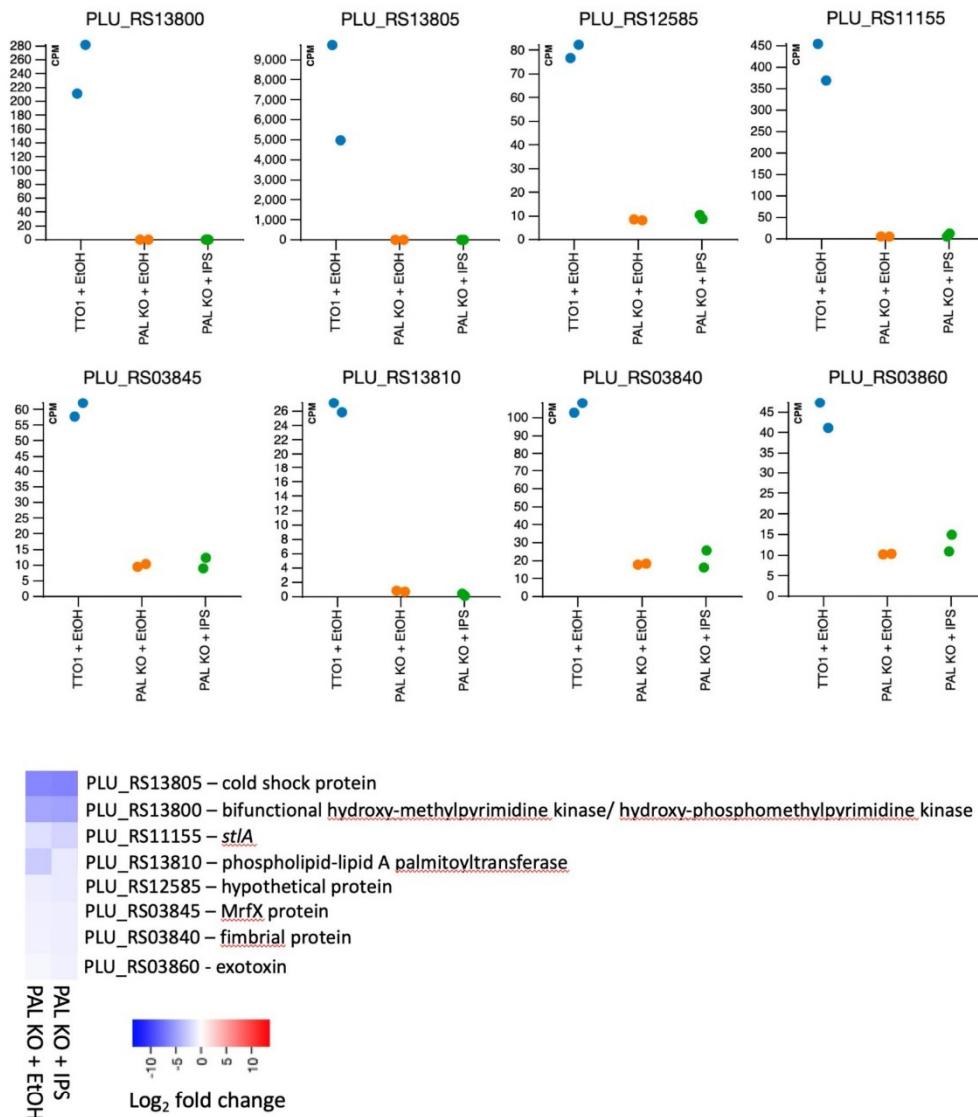
supplemented to a final concentration of 20 µg/ml. An equal volume of EtOH was added to control strains. RNA was sequenced by Novogene after bacterial rRNA depletion with RiboZero. Raw data was trimmed using Trimmomatic [22] and mapped to the *P. laumondii* genome (NC_005126.1) using bowtie2 (v2.3.4.3) [23]. The output .sam files were converted to .bam files using samtools [24]. Reads were counted using featureCounts [25] and a count matrix was then uploaded to the degust server [26] and analyzed with voom/limma. Genes with a false discovery rate less than 0.05 and a log fold change greater than 2 were considered significant.

Acknowledgements

Work in the Bode laboratory was supported by the LOEWE Schwerpunkt MegaSyn, funded by the State of Hesse. We acknowledge the Deutsche Forschungsgemeinschaft for funding of the Impact II qTof mass spectrometer (Grant INST 161/810-1) and we thank Lei Zhao for providing IPS.

References

- [1] Y.-M. Shi, H. B. Bode, *Nat. Prod. Rep.* **2018**, *35*, 309.
- [2] E. E. Herbert, H. Goodrich-Blair, *Nat. Rev. Microbiol* **2007**, *5*, 634.
- [3] a) N. J. Tobias, Y.-M. Shi, H. B. Bode, *Trends Microbiol.* **2018**, *26*, 833; b) N. J. Tobias, H. Wolff, B. Djahanschiri, F. Grundmann, M. Kronenwerth, Y.-M. Shi, S. Simonyi, P. Grün, D. Shapiro-Ilan, S. J. Pidot et al., *Nat. Microbiol.* **2017**, *2*, 1676; c) S. A. Joyce, D. J. Clarke, *Mol. Microbiol.* **2003**, *47*, 1445; d) S. A. Joyce, R. J. Watson, D. J. Clarke, *Curr. Opin. Microbiol.* **2006**, *9*, 127; e) J. Oh, N. Y. Kim, H. Chen, N. W. Palm, J. M. Crawford, *J. Am. Chem. Soc.* **2019**, *141*, 16271.
- [4] M. I. Vizcaino, X. Guo, J. M. Crawford, *J Ind Microbiol Biotechnol* **2014**, *41*, 285.
- [5] A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBiochem* **2007**, *8*, 1721.
- [6] N. J. Tobias, A. K. Heinrich, H. Eresmann, P. R. Wright, N. Neubacher, R. Backofen, H. B. Bode, *Environ. Microbiol.* **2017**, *19*, 119.
- [7] R. A. R. Machado, D. Wüthrich, P. Kuhnert, C. C. M. Arce, L. Thönen, C. Ruiz, X. Zhang, C. A. M. Robert, J. Karimi, S. Kamali et al., *Int. J. Syst. Evol. Microbiol* **2018**, *68*, 2664.
- [8] Q. Zhou, A. Bräuer, H. Adihou, M. Schmalhofer, P. Saura, G. L. C. Grammbitter, V. R. I. Kaila, M. Groll, H. B. Bode, *Chem. Sci.* **2019**, *48*, 4688.
- [9] M. Cummings, A. D. Peters, G. F. S. Whitehead, B. R. K. Menon, J. Micklefield, S. J. Webb, E. Takano, *PLoS Biol.* **2019**, *17*, e3000347.
- [10] S. Okamoto, T. Taguchi, K. Ochi, K. Ichinose, *Chem. Biol.* **2009**, *16*, 226.
- [11] S. A. Joyce, A. O. Brachmann, I. Glazer, L. Lango, G. Schwär, D. J. Clarke, H. B. Bode, *Angew. Chem* **2008**, *120*, 1968.
- [12] H. B. Park, P. Sampathkumar, C. E. Perez, J. H. Lee, J. Tran, J. B. Bonanno, E. A. Hallem, S. C. Almo, J. M. Crawford, *J. Biol. Chem.* **2017**, *292*, 6680.
- [13] A. Hapeshi, J. M. Benarroch, D. J. Clarke, N. R. Waterfield, *Microbiology* **2019**, *165*, 516.
- [14] L. Zhao, X. Chen, L. Cai, C. Zhang, Q. Wang, S. Jing, G. Chen, J. Li, J. Zhang, Y. Fang, *J. Clin. Pharm. Ther.* **2014**, *39*, 418.
- [15] J. Peppers, A. S. Paller, T. Maeda-Chubachi, S. Wu, K. Robbins, K. Gallagher, J. E. Kraus, *J. Am. Acad. Dermatol.* **2019**, *80*, 89-98.e3.
- [16] S. Chalabaev, E. Turlin, S. Bay, C. Ganneau, E. Brito-Fravallo, J.-F. Charles, A. Danchin, F. Biville, *Appl. Environ. Microbiol.* **2008**, *74*, 1717.
- [17] L. Lango-Scholey, A. O. Brachmann, H. B. Bode, D. J. Clarke, *PLoS ONE* **2013**, *8*, e82152.
- [18] G. L. C. Grammbitter, S. Kavakli, H. B. Bode, *Angew. Chem.* **submitted**.
- [19] M. M. M. Santos, R. Moreira, *Mini-Rev. Med. Chem.* **2007**, *7*, 1040.
- [20] C. M. Cabello, W. B. Bair, S. D. Lamore, S. Ley, A. S. Bause, S. Azimian, G. T. Wondrak, *Free Radic. Biol. Med.* **2009**, *46*, 220.
- [21] E. Duchaud, C. Rusniok, L. Frangeul, C. Buchrieser, A. Givaudan, S. Taourit, S. Bocs, C. Boursaux-Eude, M. Chandler, J.-F. Charles et al., *Nat Biotechnol* **2003**, *21*, 1307.
- [22] A. M. Bolger, M. Lohse, B. Usadel, *Bioinformatics* **2014**, *30*, 2114.
- [23] B. Langmead, S. L. Salzberg, *Nat. Methods* **2012**, *9*, 357.
- [24] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, *Bioinformatics* **2009**, *25*, 2078.
- [25] Y. Liao, G. K. Smyth, W. Shi, *Bioinformatics* **2014**, *30*, 923.
- [26] David Powell, Michael Milton, Andrew Perry, Kim Santos, *drpowell/degust 4.1.1*, Zenodo, **2019**.



Supplementary Figure 1. Summary of RNA seq data with plots for each of the eight significantly regulated genes (FDR < 0.05, log fold change > 2). Scale on plots represents counts per million (CPM). Heat map showing gene fold change when compared to TTO1 + EtOH showing locus tag and annotated product.

12.4 Manuscript: Structural Snapshots of the Minimal PKS System Responsible for Octaketide Biosynthesis

Declaration on the contribution of the authors/Erklärung zu den Autorenanteilen

Status: **under revision**

Beteiligte Autoren: Alois Bräuer (AB), Qiuqin Zhou (QZ), Gina L. C. Grammbitter (GLG), Maximilian Schmalhofer (MS), Michael Rühl (MR), Ville R. I. Kaila (VRK), Helge B. Bode (HBB), Michael Groll (MG)

Was hat der/die Promovierende bzw. was haben die Co-Autoren/Autorinnen beigetragen?

(1) zu Entwicklung und Planung

MG, AB: 80%
HBB, QZ: 20%

(2) Durchführung der einzelnen Untersuchungen und Experimente

Klonierung Mutagenese, Proteinreinigung Kristallisation/Kristallisationsexperimente, AB (75%)
Klonierung, heterologe Produktionen + *in vitro*-Assay minimal PKS, (HPLC-ESI/MS)-Analyse QZ (10%)
Computational modeling VRK (10%)
MCAT-Transferassay, MS-Nachweis Hexaketid, AntDEF-Mutantenanalyse durch heterologe Produktion und HPLC-ESI/MS GG (5%)

(3) Erstellung der Datensammlung und Abbildungen

Kristallisationsexperimente AB, MG, MS (75%)
Heterologe Produktionen + *in vitro*-Assay minimal PKS, (HPLC-ESI/MS)-Analyse QZ (10%)
Computational modeling VRK (10%)
MCAT-Transferassay, MS-Nachweis Hexaketid, AntDEF-Mutantenanalyse durch heterologe Produktion und HPLC-ESI/MS GG (5%)

(4) Analyse und Interpretation der Daten

Kristallisationsexperimente AB, MG, MS (75%)
Heterologe Produktionen + *in vitro*-Assay minimal PKS, (HPLC-ESI/MS)-Analyse QZ (10%)
Computational modeling VRK (10%)
MCAT-Transferassay, MS-Nachweis Hexaketid, AntDEF-Mutantenanalyse heterologe Produktion GG, MR (5%)

(5) Verfassung des Manuskripts

MG, AB, MS: 80%
HBB: 20%

date/place: _____

Title: Structural Snapshots of the *Minimal PKS System* Responsible for Octaketide Biosynthesis

Authors:

Alois Bräuer¹, Qiuqin Zhou², Gina L. C. Grammbitter², Maximilian Schmalhofer¹, Michael Rühl², Ville R. I. Kaila^{1,3}, Helge B. Bode^{2*}, Michael Groll^{1*}

Affiliations:

¹Center for Integrated Protein Science Munich (CIPSM) at the Department Chemie, Technische Universität München, Lichtenbergstraße 4, 85748 Garching, Germany.

²Molekulare Biotechnologie, Fachbereich Biowissenschaften, Goethe Universität Frankfurt, Max-von-Laue-Straße 9, Frankfurt am Main, 60438, Germany.

³Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden.

*Correspondence and requests for materials should be addressed to H.B. (email: h.bode@bio.uni-frankfurt.de) or to M.G. (email: michael.groll@tum.de)

ABSTRACT:

Type II polyketide synthases (PKS) are modular multi-enzyme complexes that produce secondary metabolites of medical relevance. Chemical backbones of such polyketides are produced by *Minimal PKS Systems* consisting of a malonyl-transacylase, an acyl-carrier protein, and an α/β -heterodimeric ketosynthase. Apart from molecular simulations, mutagenic experiments, and functional assays that characterize this invariable core, we present here x-ray structures of all ternary complexes constituting the *Minimal PKS System* for anthraquinone biosynthesis in *Photorhabdus luminescens*. We show that malonylation of the acyl-carrier protein is accompanied by major structural rearrangements in the transacylase. Principles of an ongoing chain elongation are derived from the ternary complex with a hexaketide covalently linking the heterodimeric ketosynthase with the acyl-carrier protein. Our results on the *Minimal PKS System* provide mechanistic understanding of polyketide synthases and a fundamental basis to engineer PKS pathways for future applications.

One sentence summary:

PKS core crystal structures with bound polyketides elucidate physiological chain elongation principles and complex interactions.

Main Text:

Polyketides are a structurally diverse group of bioactive secondary metabolites (1), produced by polyketide synthases (PKSs) (2). The type II PKSs contain an invariable *Minimal PKS System* core (3), with an acyl-carrier protein (ACP), a α/β heterodimeric ketosynthase (KS_{α} - KS_{β}), and a malonyl-transacylase (MCAT) (4, 5). The ACP acts as a shuttle that orchestrates the interplay between individual catalytic modules by characteristic protein:protein interactions (6). A 4'-phosphopantetheine (PPant) arm from coenzyme A (CoA) is post-translationally attached to a conserved serine of ACP^{apo} (7). This PPant group in ACP^{holo} is flexible and allows the tethered molecules to access spatially distinct enzyme-active sites. MCAT catalyzes the transfer of a malonyl unit from malonyl-coenzyme A (Mal-CoA) to ACP^{holo}, forming Mal-ACP (8), which in turn is used for polyketide synthesis by repetitive decarboxylative condensations (9). Subsequently, additional enzymes of the biosynthetic gene cluster, including ketoreductases, aromatases, cyclases and lyases convert the nascent poly- β -keto intermediate into an aromatic scaffold (10). Finally tailoring enzymes like oxygenases, methyltransferases, and glycosyltransferases modify this skeleton that distinguish the bioactivities of the respective compounds (Fig. 1 and fig. S1) (11).

The PKS machinery is key to generate a vast diversity of natural products. Due to the transient nature of the ACP-partner enzyme interactions and the chemical instability of the poly- β -keto intermediates (12), no MCAT:ACP or KS_{α} - KS_{β} -ketosynthase:ACP complex structures have been reported for type II PKS. Our major goal was to investigate principles of the *Minimal PKS System*. Producing the largest class of quinones, the central component of the anthraquinone (AQ) biosynthetic gene cluster from the Gram-negative bacterium *Phototribadus luminescens* (*Pl*) is ideally suited to study such a sophisticated system (Fig. 1A) (13). Here, we resolved the invariable core responsible for AQ biosynthesis, and elucidated protein:protein interactions of AntF (ACP) with AntDE (ketosynthase α/β -heterodimer), as well as AntF with PIMCAT (malonyl-transacylase) using x-ray crystallography. Together with structure-based mutagenesis, molecular simulations, and activity assays, the obtained ternary complexes provide fundamental insights into the *Minimal PKS System*.

RESULTS**Exploring the *Minimal PKS System***

Expression of the gene cluster *antABCDEFGH* in *Escherichia coli* (*Ec*) produces the aromatic anthraquinone **AQ-256** (Fig. 1 and tables 1 to 3) (13). This characteristic polyketide with its tricyclic ring system is derived from an octaketide, which is the primary intermediate of this type II PKS (14). We studied the minimal set of enzymes that form this octaketide via the condensation of eight malonate building blocks (fig. S2 and tables 4 to 6). In agreement to current findings (14, 15), AntF is expressed in the inactive *apo* form. Screening of AQ-proteins revealed that AntB (PPTase) and AntG (CoA-ligase) are both essential to convert AntF^{apo} into the *holo* form by the covalent attachment of PPant (Fig. 1B and fig. S3). This observation differs from all known PKS systems investigated so far (15, 16).

The heterologous expression of the minimal set of enzymes AntBDEFG in *E. coli* yields the marker molecules SEK4 and SEK4b (Fig. 1C and fig. S1), which are absent in the wild type host and form via spontaneous cyclization of ACP-released octaketides (17, 18). Surprisingly, PIMCAT (Plu2834), which transfers the malonyl group from Mal-CoA to AntF^{holo} in *P. luminescens*, is not required. Since this transfer is indispensable, we suggest that the endogenous *E. coli* orthologue *EcFabD* is capable of loading malonate onto AntF^{holo} (fig. S4) (19). With all

2

components of the *Minimal PKS System* in AQ biosynthesis, we were able to structurally and functionally analyze the interactions of the individual proteins.

AntF Adopts a Classical ACP Fold

To obtain molecular insights into the *Minimal PKS System*, x-ray structures of AntF^{apo} and AntF^{holo} were determined up to 1.85 Å resolution (AntF^{apo}, R_{free}=19.6%; PDB ID: 6SM4; AntF^{holo}, R_{free}=22.7%; PDB ID: 6SM6; 0.3 Å rmsd for 77 C^α atoms, Fig. 2, fig. S5E and table S7). Although the sequence identity of AntF to known ACPs is less than 18%, its structure is similar to ACPs from various species, e.g., the *E. coli* EcAcpP protein (AntF / EcAcpP: 1.4 Å rmsd for 74 C^α atoms, figs. S6 and S7) (20). The architecture of AntF displays a conserved α -helix bundle, comprising three amphipathic helices H1 (residues 6-17), H2 (38-52, recognition helix), and H4 (67-80), whereas H3 (58 – 63) is more dynamic and positioned perpendicular to the other helices (21). Loop region I is fully defined in the electron density map and contains a 1½-helical turn, including the PPant binding site at Ser38. The prosthetic group of AntF^{holo} is surface exposed and adopts different orientations in the two molecules of the asymmetric unit (Fig. 2A). The mobility of PPant is a likely prerequisite to transfer the covalently linked intermediates between the distinct enzymes in the PKS system (see below).

Conformational changes in P/MCAT induced by AntF^{holo}

The transfer of malonate from Mal-CoA to AntF^{holo} proceeds in two consecutive steps that involve a stable oxoester intermediate at the active site serine 92 (Ser92O') of P/MCAT. We solved the crystal structure of AntF^{holo} bound to P/MCAT to 3.3 Å resolution (R_{free} = 26.8%, PDB-ID: 6SMD, Fig. 2, fig. S5E and table S7), with P/MCAT as well as a P/MCAT:AntF^{holo} complex in the asymmetric unit. The transacylase is composed of a large α/β -hydrolase domain (residues 1 - 124 and residues 200 - 310) (22) and a subdomain (residues 124 - 200) that adopts a ferredoxin-like topology as described for acyl-phosphatases (23).

The P/MCAT:AntF^{holo} structure shows an instable assembly with a small contact area of 890 Å² (Fig. 2B and C). Complex formation takes predominately place via the recognition helix H2 of AntF (interactions are summarized in table S8). Asp37 and Thr39, located adjacent to the PPant-modified Ser38, interact with Arg287 of the transacylase, and are responsible for adjusting the prosthetic group that is fully defined in the electron density map (Fig. 2E). Protein residues of AntF in the complex maintain their original conformation as observed in the unbound ACP (0.7 Å rmsd for 77 C^α atoms). Interestingly, PPant undergoes a structural rearrangement during P/MCAT binding and accounts for 40% (360 Å²) of the buried surface between AntF and P/MCAT. Upon complex formation, the ligand adopts an extended configuration and pierces towards the active site of P/MCAT. The solvent exposed phosphate group of PPant functions as a hub that facilitates tight contacts of its dimethyl, β -alanine and cysteamine groups with the transacylase. The functionally active site Ser92O' (2.8 Å) interacts with His201 in the catalytic dyad of P/MCAT that enhances its nucleophilicity (Fig. 2E).

Structural comparison of P/MCAT:AntF^{holo} with P/MCAT alone shows a rigid body movement in the transacylase unit (Fig. 2D). While the hydrolase domain remains structurally unchanged, the ferredoxin subdomain rotates by 10° upon AntF^{holo} binding (Fig. 2D and E). This results in salt bridges between Lys189 and Arg190 of P/MCAT and Asp43 and Glu49 of AntF. In addition, the aliphatic residues Val134, Leu192, and Leu194 of P/MCAT are shifted by more than 2 Å, enabling strong interactions with the prosthetic group of AntF^{holo} that leads to a dehydration of the binding channel to accommodate PPant (Fig. 2D and E). Thus, complex formation between

AntF^{holo} and P/MCAT is accompanied by large-scale structural rearrangements that could indicate an induced fit mechanism.

P/MCAT Catalyzes Unidirectional Transacylation from Mal-CoA to AntF^{holo}

Remarkably, even in the absence of P/MCAT, octaketide formation takes place in the *antBDEFG* engineered *E. coli* strain (Fig. 1C and fig. S1). The *EcFabD* orthologue (PDB ID: 2G1H) of the host that performs the same reaction shares a sequence identity of 75% and might be the substitute for our *Minimal PKS System*. Its structural superposition with P/MCAT reveals 0.8 Å rmsd for 306 C^α atoms (fig. S4). Moreover, the x-ray study of Oefner *et al.* on *EcFabD* contains a bound malonate at Ser92O^γ originating from Mal-CoA (24). In both MCAT structures, the active site is located between the hydrolase and the ferredoxin domains. Ser92O^γ is positioned at the conserved GHSXG motif by forming a dyad with His201N^{ε2} (2.7 Å) (22). The hydrogen-bond between the carboxylate of Mal and His201 enhances the nucleophilicity of the active site Ser92O^γ, whereas Arg117 is located at the bottom of the specificity pocket and stabilizes the incoming Mal-CoA substrate by a bidentate salt-bridge. Altogether, these interactions facilitate transacylation via an oxoester intermediate (Fig. 3A) and prevent decarboxylation of the building block.

The transfer of the malonyl moiety to MCATs does not require structural rearrangements to keep the channel in its open state (Fig. 2D) (24). Accordingly, the generated CoA is loosely bound and surrounded by a vast number of water molecules. In contrast, binding of AntF^{holo} to P/MCAT induces the rotation of the ferredoxin domain, resulting in strong interactions with the AntF-bound PPant (Fig. 2D and E). Due to its close proximity to the active center, the deprotonated sulfhydryl group of PPant attacks the carbonyl carbon atom (C₁) of the Ser92O^γ:Mal-oxoester resulting in an (O⁻,S,O)-hemi-thio-orthoester-olate (Fig. 3A). In this arrangement, His201 could act as a Brønsted acid to promote the formation and release of Mal-AntF. This conformational change in P/MCAT:AntF^{holo} might suppress the back reaction, which is in agreement with our MS-coupled *in vitro* experiments. While Ser92O^γ of P/MCAT and AntF^{holo} are immediately esterified by Mal-CoA, malonyl-transfer from Mal-AntF to P/MCAT was not recorded within 2 hours in the absence of Mal-CoA (fig. S8).

A Hexaketide is Covalently Bound to AntDE

To probe the chain elongation mechanism in the *Minimal PKS System*, we determined the crystal structure of AntDE (KS_α-KS_β-ketosynthase) at 2.7 Å resolution (R_{free}=26.8%; PDB ID: 6SMO, fig. S5E and table S9). Despite their low sequence identity of 26% (fig. S9), AntD and AntE share the same fold (1.4 Å rmsd for 243 C^α atoms). Both proteins adopt a five-layered structure of alternating α-helices and β-sheets (αβαβ motif) (25, 26) that can be divided into two parts (27). Each of these sub-domains is composed of a 10-stranded mixed β-sheet that complement each other in an antiparallel fashion with a contact surface of 2920 Å² per subunit (fig. S5A). The central cavity of AntDE is accessible through a narrow 15 Å long channel (fig. S5B). Cys176 at the N-terminus of helix α6 of AntD forms the putative nucleophile and is located in proximity to the heterodimeric interface, and His317 and His354 that reside 4-5 Å from Cys176, are important for decarboxylation of the malonyl moiety (28). Strikingly, the 2F_o-F_c map depicts additional electron density at Cys176, suggesting that the active site of AntDE has been covalently modified during its recombinant expression (Fig. 4D and E). Indeed, ESI-MS analysis of tryptic digested AntDE identified the chemical transformation at Cys176 to be a hexaketide, which is absent in the C176A mutant (Fig. 5A, fig. S10, tables S10 and S11). Together, these results demonstrate that Cys176 of AntD

acts as the catalytic nucleophile and the thioester-bound hexaketide is formed by consecutive condensation reactions using endogenous enzymes of the *E. coli* host system.

The highly reactive hexaketide in AntDE is stabilized by surrounding protein moieties. Specifically, the 15 Å-long bound metabolite is trapped in a narrow channel, preventing its spontaneous intramolecular cyclization. In order to probe the energetics and conformational constraints of the bound hexaketide, we performed quantum chemical (DFT and QM/MM MD) simulations (see Supporting Information). Our calculations suggest that the ligand remains structurally stable in the binding channel, supporting the spatial arrangement of its heteroatoms (Fig. 5C and D). While the thioester at C₁ adopts an *anti*-configuration and interacts with His354 (AntD), the residual part of the molecule is arranged planar in *syn*. Moreover, the calculations suggest that the sp³-hybridized C₂ atom has a dihedral angle of around -140° between the two carbonyl groups at C₁ and C₃, consistent with the 2F_o-F_c-electron density map. The C₃ atom was modeled in a keto form that is stabilized by hydrogen bonds with the AntD main chain atoms Cys176NH and Phe413NH. C₅ can adopt a keto or an enol configuration, because of π-interactions via Phe413 and the lack of protein coordination for its oxygen atom. Three extender units of the hexaketide (C₄ – C₇ and C₁₁ – C₁₂) are modeled as enol units with vinyl alcohols at positions C₅, C₇, and C₁₁ form hydrogen-bonds with the AntE carbonyl oxygens Val129O and Ala130O. These double bonds between C₄ – C₅, C₆ – C₇ and C₁₀ – C₁₁ are stabilized by π-stacking with Trp133 (AntE), Phe215 (AntD), and Phe134 (AntE), respectively (Fig. 5D). Interestingly, in QM/MM dynamics simulations we observe intramolecular, Grotthuss-type proton transfers between the keto and enol-units suggesting that energetic barrier for proton shifts is small (Fig. 5C).

One of the general scaffold generated by PKS II systems is an octaketide, whereas our heterologous expression of AntDE in *E. coli* yields AntD-bound hexaketides (see above). We extended the ligand by one acyl-moiety bound to Cys176 (AntD) in our quantum chemical models. These calculations reveal that the modeled heptaketide fills the entire polyketide binding channel of AntDE leading to small bending of the terminal keto unit (Fig. 5E). The added acetyl group is in close proximity to the bottom of the specificity pocket formed by Trp107 and Trp133 of AntE, while its C₁₃ carbonyl oxygen and C₁₄ methyl group lack defined protein interactions. Moreover, the introduced building block is tilted by 109° to avoid steric clashes with Leu210 (AntD), therefore restricting the resonance stabilization of the growing polyketide chain to C₅ - C₁₁. This results in accumulation of a molecular strain in the ligand (fig. S11A and D), suggesting that the subsequent condensation faces high energetic barriers and is sterically disfavored. An octaketide can be modeled into the binding cavity, but induces subtle structural changes in the surrounding protein framework (fig. S11C). These findings support the *in vitro* findings showing that Mal-AntF^{holo} is essential for octaketide formation in AQ biosynthesis. Indeed, the expression of AntDE with ACPs of closely related type II PKS such as ActI-ORF3 (ACP from actinorhodin biosynthesis) (29) or RemC (ACP from resistomycin biosynthesis) (30) cannot substitute AntF and fail to produce octaketides (table S6).

Structural Insights into the Ternary AntDE:F Complex

Early NMR-studies revealed dynamic ACP proteins, which might be indispensable for their mutual interactions (31). Our x-ray structures of AntDE in complex with AntF demonstrate that neither the acyl-carrier protein nor the ketosynthases undergo significant rearrangements (AntF^{holo}: 0.9 Å rmsd for 70 C^α atoms, AntDE: 0.3 Å rmsd for 749 C^α atoms). Interactions between AntDE and AntF in the AntDE:F^{apo} (R_{free}=26.8%; PDB ID: 6SMO; 2 AntDE, 1 AntDE:F^{apo} in the AU, table S9) as well as the AntDE:F^{holo} (R_{free}=25.5%; PDB ID: 6SMP; 4 AntDE:F^{holo} in the AU, table S9) crystal structures are identical, arguing for the physiological assembly (Fig. 4). As expected, the contact areas between AntD:F^{apo} (240 Å²) and AntE:F^{apo} (560 Å²) are small and only few prominent polar

5

interactions are observed between AntF and its ketosynthase AntDE (summarized in tables S12 and S13). For instance, exchange of Asp43 to alanine in AntF still promotes formation of SEK4 and SEK4b, whereas the D37A mutation abolishes octaketide formation. Altogether, these findings prove the importance of single contacts in the AntDE:F complex and their interdigitating network. (Figs. 4B and 5B).

While the catalytically active KS_{α} subunit of ketosynthases is conserved and exchangeable (fig. S12), the chain length factor KS_{β} is unique (fig. S13) (32), suggesting that the specific interactions between AntF^{holo} and AntDE are a prerequisite for its function (table S13). In agreement, AntF forms more interactions with AntE than with AntD. In particular, the polar residues Glu40 and Asp43 (AntF) interact with Arg61, Arg64, Lys65 (AntE), and Asp37 (AntF) is coordinated to His282 (AntD). To this end, Arg61 plays a key role for the stability of the assembly by its interplay with Glu40, as well as with Asn18 and Tyr34 of AntF. Accordingly, the R61A mutant in AntE fails to produce SEK4 and SEK4b (Fig. 5B). Thus, the structural and functional studies reveal that AntDE forms a stable assembly due to the large hydrophobic core, whereas its interactions with AntF are predominantly polar and depend only on few protein:protein interactions between the ACP and the heterodimeric KS_{α} - KS_{β} -ketosynthase. Taken together, these features of the *Minimal PKS System* allow for a fast and adjustable delivery of malonyl precursors to the ketosynthase machinery.

The catalytic site of AntDE is buried deep inside the heterodimer at the bottom of a 15 Å-long tunnel that houses the prosthetic group (Fig. 4C). PPant increases the contact area in AntDE:F^{holo} by 30% (AntD:F^{holo}: 200 Å², AntE:F^{holo}: 520 Å² and AntD:PPant: 320 Å²), stabilizing the complex and providing a possible explanation on to how AntDE distinguishes between AntF^{apo} and AntF^{holo}. All protein interactions with PPant are exclusively transmitted via AntD. The phosphate group of the ligand is solvent exposed and highly mobile with its residual part occupying a narrow channel with less than 5 Å diameter. An extensive network of residues including Leu218, Ala220, Thr284, Asp285, Leu286, Ser319 and Thr321 of AntD forms interaction with the pantothenic residue (Fig. 4E, tables S12 and S13). The tight coordination of the carbonyl oxygen of the β-alanine moiety of PPant via Ser319O^γ (2.9 Å) and Thr321O^γ (2.8 Å) is most striking. The function of these residues is supported by the S319A / T321A double mutation, which abolishes formation of SEK4 and SEK4b (Fig. 5B). In contrast, the cavity around Cys176 is spacious (250 Å³) and facilitates the flexible cysteamine group to transfer the malonyl building block from PPant to the ketosynthases.

AntDE:F^{holo} Reveals a Late Reaction Intermediate of Chain Elongation

The x-ray structures of AntDE:F^{apo} and AntDE:F^{holo} were obtained by co-crystallizing the individual proteins with the active site nucleophile Cys176 of AntD carrying the hexaketide. In AntDE:F^{apo}, the conformation of the hexaketide is identical to that in the AntDE complex. In contrast, the 2F_o-F_c-electron density map of AntDE:F^{holo} shows that the thiol group of the cysteamine moiety is covalently linked to the carbonyl carbon atom of the thioester at Cys176. Hereby, a (O,S,S)-hemi-dithio-orthoester-olate in sp³ configuration is formed with the alkoxide occupying the oxyanion hole by strong hydrogen-bonds with Cys176NH and Phe413NH (Fig. 4E and F), illustrating a late reaction intermediate in the chain elongation process. In this state, the oxygen atom at C₁ is flipped by 180° so that all carbonyl groups of the polyketide are oriented in *syn*. This conformational switch pushes the entire polyketide into the binding channel (Fig. 4C) and disrupts protein:ligand interactions. As a result, the hexaketide appears more diffuse in the 2F_o-F_c-electron density map, while PPant and the orthoester-olate remain well defined. These findings are supported by our quantum chemical models, suggesting that the ligand is

stabilized by protonation of the carbonyl group, which compensates the electrostatic repulsion between the carbonyl oxygen at C₃ and the alkoxide (fig. S11B).

Principles of Octaketide formation in the *Minimal PKS System*

In the *Minimal PKS System* of AQ biosynthesis, the *PIMCAT* module transfers malonate onto AntF^{holo}, which in turn delivers the attached building block either as starter or extender unit to AntDE. Decarboxylation of the malonyl moiety involves a classical Brønsted acid / base catalysis, producing an enolate that is transferred from PPant to the active site cysteine of the ketosynthase. Covalently bound to AntDE the growing chain matures to a polyketide in iterative rounds.

The initiation reaction starts with protonation of Mal-AntF by His317 and His354 of AntDE (His-His relay) leading to CO₂ release (Fig. 3B). The generated enol converts to an acetyl group by capturing the proton from the active site cysteine of the ketosynthase. The resulting Cys176S⁻ attacks the thioester of PPant by forming a (O,S,S)-hemi-dithio-orthoester-olate intermediate that covalently links AntF^{holo} to AntDE (Figs. 3B and 4F). The change in carbon-configuration at C₁ (sp³ versus sp²) is accompanied by PPant rearrangements allowing the olate to interact with the oxyanion hole. Subsequently, PPant's sulphur atom is protonated, which causes AntF^{holo} release and restores the carbon configuration at C₁ back to sp². Concomitantly, a 180° flip of the thioester leads to a displacement from the oxyanion hole, thereby completing the acetylation of AntDE.

The following reactions are Claisen condensations (Fig. 3B). The previously formed acetyl group at Cys176 of AntDE acts as a substrate for the first elongation step to connect a new malonyl unit. Similar to the initiation reaction, the incoming building block is decarboxylated. The resulting enolate attacks the thioester of the acetyl group forming acetoacetate-AntF, which is immediately transferred back to Cys176 of AntDE. During this reaction step the acetyl moiety is pushed into the binding channel, thereby destabilizing the growing polyketide due to the lack of appropriate protein interactions. Because of the covalent linkage with the ketosynthase and PPant, the ligand shifts back and displaces the olate out of the oxyanion hole. As a consequence, the orthoester-olate decomposes and AntF^{holo} is released, resulting in the covalently bound acetoacetate group (diketide) at AntDE.

The remaining chain elongation cycles follow the same mechanistic principle. However, once the heptaketide forms, the lack of further space in the AntDE binding channel prevents accommodation of an additional malonate unit (Fig. 4C and D). Therefore, in the last chain elongation round, the orthoester-olate between AntF^{holo} and AntDE lacks stabilization, and the covalent bond to the ketosynthase is thus likely to break, leading to release of the final octaketide-AntF product from the *Minimal PKS System* (Fig. 6).

Similarities and Differences of ACPs in Natural Product Biosynthesis

Our x-ray structures of the acyl-carrier protein AntF^{holo} unbound and in complex with *PIMCAT* or AntDE provide fundamental insights into synergies within the *Minimal PKS System*. As revealed in the present study, ACPs function is transmitted via rigid body motion (fig. S7), where their designated target proteins are predominantly recognized via the N-terminal part of helix H2. High temperature factors in this region relative to the ternary complexes indicate increased flexibility, which enables rapid recognition of the binding unit. In all AntF-structures, the conformations of Asp37 and Glu40, which are conserved among ACPs from various species, are identical and coordinate basic amino acids of the corresponding proteins (figs. S6 and S7). Specifically, the phosphopantetheine arm helps to select proper binding partners with a 15 Å deep cavity next to a characteristic

positively charged surface. Interactions with PPant are solely provided by the client and enlarge the overall contact area of the protein:protein interfaces as revealed in the *PIMCAT:AntF^{holo}* and *AntDE:F^{holo}* complex structures. In addition, we could show that ACPs are partly interchangeable. For example, the *E. coli* system malonates *AntF^{holo}* and chain elongation at AntDE can be substituted by *EcAcpP* up to the hexaketide. On the other hand, ACPs own pathway specific characteristics, because neither *EcAcpP* nor ACPs of actinorhodin or resistomycin biosynthesis can replace AntF to produce octaketides. Thus, ACPs are key of the *Minimal PKS System* by combining a rigid framework with partial mobility to produce a diverse set of individual polyketides.

DISCUSSION:

Our study provides detailed mechanistic insights into biosynthesis of polyketides, which are medically important natural products. In the *Minimal PKS System* for AQ biosynthesis, malonate is transferred to the acyl-carrier protein *AntF^{holo}* by the transacylase *PIMCAT*. Mal-AntF serves as substrate for initiation and chain elongation reactions leading to the synthesis of octaketides (Fig. 6). Crystallographic snapshots of a hexaketide bound to the KS_{α} subunit AntD and a physiologically reaction intermediate linking *AntF^{holo}* to AntDE depict this chain elongation process in action. The emerging polyketide is predominantly protected by the KS_{β} subunit AntE (25). Our molecular simulations confirm that C₁, C₃ and C₉ atoms of the growing chain are stabilized in carbonyl groups and the less reactive enols between C₄ – C₇ and C₁₀ – C₁₁ prevent intra- and intermolecular side reactions. After completion of octaketide biosynthesis by AntDE, the highly reactive nascent chain is transferred to *AntF^{holo}* and liberated from the *Minimal PKS System* for further processing (13, 14).

Although MCAT is part of the *Minimal PKS System*, it is not encoded in the gene cluster (3). Interestingly, its function can be replaced by orthologous transacylases, e.g., from *E. coli* during heterologous expression of *AntF^{holo}*. The promiscuity of MCAT is surprising as the transfer of the malonyl building block onto the ACP occurs via several reaction steps and, as revealed in the *PIMCAT:AntF^{holo}* structure, requires a rotation of its ferredoxin subdomain. Although ACPs adopt a common fold, the variations in their primary structure is significant, whereas MCATs share high sequence identity among different species. However, the characteristic interaction pattern of *AntF^{holo}* with *PIMCAT* via PPant could explain why a chimeric *Minimal PKS System* is catalytically active. Furthermore, a single MCAT protein hub might catalyze a large number of biosynthetic pathways, since Mal-ACP constitutes a central building block for all polyketides and fatty acids. This suggests that ACPs and MCATs from distinct organisms have common structural features that complement each other and establish a functional core of diverse PKS biosynthetic machineries.

The structure of *AntDE:F^{holo}* identified a small but contact-rich motif between the ACP and the KS_{α} - KS_{β} -ketosynthase. Intriguingly, this interaction site appears to be conserved, as heterologous AntDE expression in *E. coli* results in the formation of a covalently bound premature hexaketide. These findings demonstrate that an endogenous ACP is able to adapt the function of *AntF^{holo}*, but chain elongation to the octaketide has to overcome high energetic barriers and requires a genuine shuttle machinery. Thus, components of polyketide chain elongation, at least in AQ biosynthesis, are only partially interchangeable. Yet, such pathway-specific protein-protein interactions might allow bacteria to coordinate their multiple *Minimal PKS Systems* at the same time and to precisely control the production of diverse polyketides composed of the universal malonyl building block (2).

So far detailed information is available for individual parts of PKS, non-ribosomal peptide and fatty acid synthase machineries, but understanding the overall interplay between these unique pathways remains challenging. As

shown, minor structural changes control the selective recruitment of appropriate binding partners and intermediates at the right time. Altogether, we discovered principles of polyketide biosynthesis facilitating future engineering and modulation of PKS pathways for a broad range of applications.

REFERENCES AND NOTES:

1. D. A. Hopwood, Genetic Contributions to Understanding Polyketide Synthases. *Chem. Rev.* **97**, 2465-2498 (1997).
2. C. Hertweck, The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed. Engl.* **48**, 4688-4716 (2009).
3. R. McDaniel, S. Ebert-Khosla, H. Fu, D. A. Hopwood, C. Khosla, Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc. Natl. Acad. Sci. U S A* **91**, 11542-11546 (1994).
4. C. W. Carreras, A. M. Gehring, C. T. Walsh, C. Khosla, Utilization of enzymatically phosphopantetheinylated acyl carrier proteins and acetyl-acyl carrier proteins by the actinorhodin polyketide synthase. *Biochemistry* **36**, 11757-11761 (1997).
5. W. Bao, E. Wendt-Pienkowski, C. R. Hutchinson, Reconstitution of the iterative type II polyketide synthase for tetracenomycin F2 biosynthesis. *Biochemistry* **37**, 8132-8138 (1998).
6. A. Chen, R. N. Re, M. D. Burkart, Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions. *Nat. Prod. Rep.* **35**, 1029-1045 (2018).
7. R. H. Lambalot, C. T. Walsh, Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *J. Biol. Chem.* **270**, 24658-24661 (1995).
8. L. Serre, E. C. Verbree, Z. Dauter, A. R. Stuitje, Z. S. Derewenda, The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5-Å resolution. Crystal structure of a fatty acid synthase component. *J. Biol. Chem.* **270**, 12961-12964 (1995).
9. C. Khosla, R. S. Gokhale, J. R. Jacobsen, D. E. Cane, Tolerance and specificity of polyketide synthases. *Annu. Rev. Biochem.* **68**, 219-253 (1999).
10. S. C. Tsai, The Structural Enzymology of Iterative Aromatic Polyketide Synthases: A Critical Comparison with Fatty Acid Synthases. *Annu. Rev. Biochem.* **87**, 503-531 (2018).
11. S. Sundaram, C. Hertweck, On-line enzymatic tailoring of polyketides and peptides in thiotemplate systems. *Curr. Opin. Chem. Biol.* **31**, 82-94 (2016).
12. C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **24**, 162-190 (2007).
13. A. O. Brachmann *et al.*, A type II polyketide synthase is responsible for anthraquinone biosynthesis in *Photobacterium luminescens*. *Chembiochem.* **8**, 1721-1728 (2007).
14. Q. Zhou *et al.*, Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photobacterium luminescens*. *Chem. Sci.* **10**, 6341-6349 (2019).
15. M. Cummings *et al.*, Assembling a plug-and-play production line for combinatorial biosynthesis of aromatic polyketides in *Escherichia coli*. *PLoS Biol.* **17**, e3000347 (2019).
16. C. T. Walsh, A. M. Gehring, P. H. Weinreb, L. E. Quadri, R. S. Flugel, Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* **1**, 309-315 (1997).
17. J. Dreier, A. N. Shah, C. Khosla, Kinetic analysis of the actinorhodin aromatic polyketide synthase. *J. Biol. Chem.* **274**, 25108-25112 (1999).
18. H. Fu, D. A. Hopwood, C. Khosla, Engineered biosynthesis of novel polyketides: evidence for temporal, but not regiospecific, control of cyclization of an aromatic polyketide precursor. *Chem. Biol.* **1**, 205-210 (1994).
19. R. G. Summers, A. Ali, B. Shen, W. A. Wessel, C. R. Hutchinson, Malonyl-coenzyme A:acyl carrier protein acyltransferase of *Streptomyces glaucescens*: a possible link between fatty acid and polyketide biosynthesis. *Biochemistry* **34**, 9389-9402 (1995).
20. T. A. Holak, S. K. Kearsley, Y. Kim, J. H. Prestegard, Three-dimensional structure of acyl carrier protein determined by NMR pseudoenergy and distance geometry calculations. *Biochemistry* **27**, 6135-6142 (1988).
21. D. I. Chan, H. J. Vogel, Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **430**, 1-19 (2010).
22. D. L. Ollis *et al.*, The alpha/beta hydrolase fold. *Protein Eng.* **5**, 197-211 (1992).

23. A. Pastore, V. Saudek, G. Ramponi, R. J. Williams, Three-dimensional structure of acylphosphatase. Refinement and structure analysis. *J. Mol. Biol.* **224**, 427-440 (1992).
24. C. Oefner, H. Schulz, A. D'Arcy, G. E. Dale, Mapping the active site of Escherichia coli malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 613-618 (2006).
25. A. T. Keatinge-Clay, D. A. Maltby, K. F. Medzihradzky, C. Khosla, R. M. Stroud, An antibiotic factory caught in action. *Nat. Struct. Mol. Biol.* **11**, 888-893 (2004).
26. G. L. C. Grammbitter *et al.*, An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments. *J. Am. Chem. Soc.*, (2019). (in press, doi.org/10.1021/jacs.8b10776)
27. M. Mathieu *et al.*, The 2.8 Å crystal structure of peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: a five-layered alpha beta alpha beta alpha structure constructed from two core domains of identical topology. *Structure* **2**, 797-808 (1994).
28. T. Robbins, J. Kapilivsky, D. E. Cane, C. Khosla, Roles of Conserved Active Site Residues in the Ketosynthase Domain of an Assembly Line Polyketide Synthase. *Biochemistry* **55**, 4476-4484 (2016).
29. H. Fu, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, Engineered Biosynthesis of Novel Polyketides: Dissection of the Catalytic Specificity of the act Ketoreductase. *J. Am. Chem. Soc.* **116**, 4166-4170 (1994).
30. K. Jakobi, C. Hertweck, A gene cluster encoding resistomycin biosynthesis in *Streptomyces resistomycificus*; exploring polyketide cyclization beyond linear and angucyclic patterns. *J. Am. Chem. Soc.* **126**, 2298-2299 (2004).
31. Y. Kim, J. H. Prestegard, A dynamic model for the structure of acyl carrier protein in solution. *Biochemistry* **28**, 8792-8797 (1989).
32. D. H. Sherman, E. S. Kim, M. J. Bibb, D. A. Hopwood, Functional replacement of genes for individual polyketide synthase components in *Streptomyces coelicolor* A3(2) by heterologous genes from a different polyketide pathway. *J. Bacteriol.* **174**, 6184-6190 (1992).

Acknowledgments: We thank the staff of beamline X06SA at the Paul Scherrer Institute (SLS, Villigen) for assistance during data collection, Dr. Sebastian Fuchs for MALDI-MS analysis and Prof. Dr. Michael Karas for MALDI-MS access.

Funding: This work was supported by the DFG within the SPP 1617 (H.B) and SFB1035 (project B12 to V.R.I.K), and the LOEWE program of the state of Hesse as part of the MegaSyn research cluster (H.B.).

Author contributions: H.B. and M.G. initiated and supervised the project. Q.Z. generated the expression constructs and established the activity assays to reconstitute the *Minimal PKS System*, G.G. and M.R. performed MS analyses, A.B. cloned mutants, purified and crystallized proteins, M.G. determined x-ray structures, A.B., M.S., V.K. and M.G. analyzed structures, V.K. carried out molecular simulations and M.G. wrote the paper with input from all authors.

Competing interests: The authors declare no competing financial interests.

Data and materials availability: Crystallographic data were deposited in the RCSB Protein Data Bank with the PDB IDs: 6SM4, 6SM6, 6SMD, 6SMO and 6SMP. All other data are available in the text or in the supplementary materials.

Supplementary Materials:

Materials and Methods
 Figures S1-S13
 Tables S1-S13
 References 1-24

MAIN FIGURES

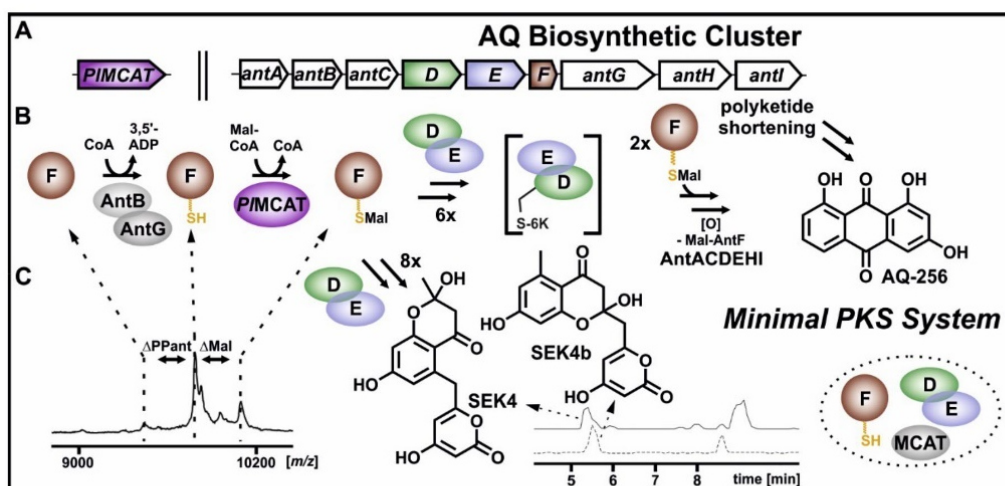


Fig. 1. Overview of the *Minimal PKS System*. (A) Type II PKS gene cluster of anthraquinone (AQ) biosynthesis in *P. luminescens*. The genes *antD*, *antE* and *antF* are shown in capitals and are colored in green, blue and brown, respectively. *P/MCAT* is not part of the gene cluster and indicated in magenta. (B) Schematic progression of the reaction steps resulting in the product **AQ-256**. The hexaketide (6K) is trapped in the AntDE complex (marked in brackets). The reaction is completed by polyketide shortening via AntI (14). (C) Analysis of the octaketide shunt products using HPLC/MS upon heterologous expression of AntBDEFG in *E. coli*. Extracted ion chromatograms for SEK4 (continuous line, m/z 301 [M+H]⁺) and SEK4b (dashed line, m/z 319 [M+H]⁺) are shown.

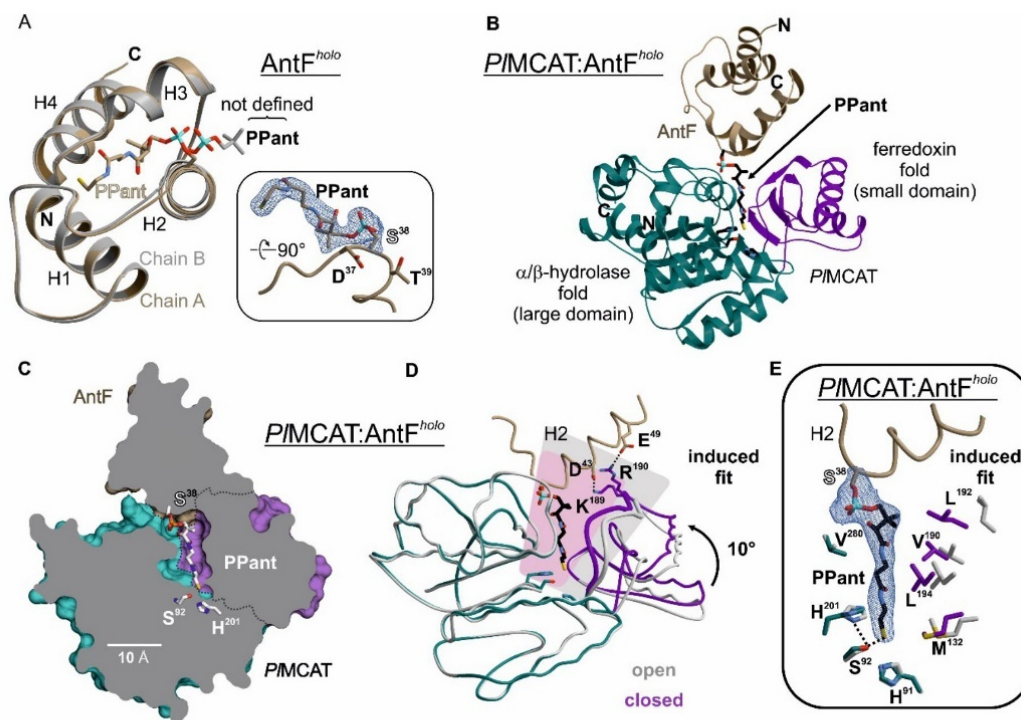


Fig. 2. *AntF^{holo}* alone and in complex with *PIMCAT*. (A) The asymmetric unit consists of two *AntF^{holo}* molecules (chain A colored in brown and B in grey). The prosthetic group (sticks) is surface exposed and flexible. The box displays the $2F_o - F_c$ electron density map (blue mesh countered to 1σ) for PPant of chain A. (B) Ribbon drawing of the *PIMCAT:AntF^{holo}* complex. *PIMCAT* is composed of a large α/β -hydrolase domain (teal) and a small ferredoxin subdomain (magenta). (C) The sliced surface representation of *PIMCAT:AntF^{holo}* reveals a prominent entrance that positions the covalently bound PPant in close proximity to the active site of *PIMCAT*. (D) The structural superposition of *PIMCAT* (open state, gray background) with *PIMCAT* bound to *AntF^{holo}* (closed state, magenta background) illustrates rotation of the ferredoxin domain by 10° . The structural rearrangement facilitates transacylation from Mal-CoA to *AntF^{holo}*. (E) Zoom in of PPant in the *PIMCAT:AntF^{holo}* complex. The thiol group of PPant is hydrogen-bonded to the active site of *PIMCAT* (black dots). Residues shown in grey (sticks) depict the transacylase in the unbound open conformation and highlight the conformational rearrangements upon binding of *AntF^{holo}*.

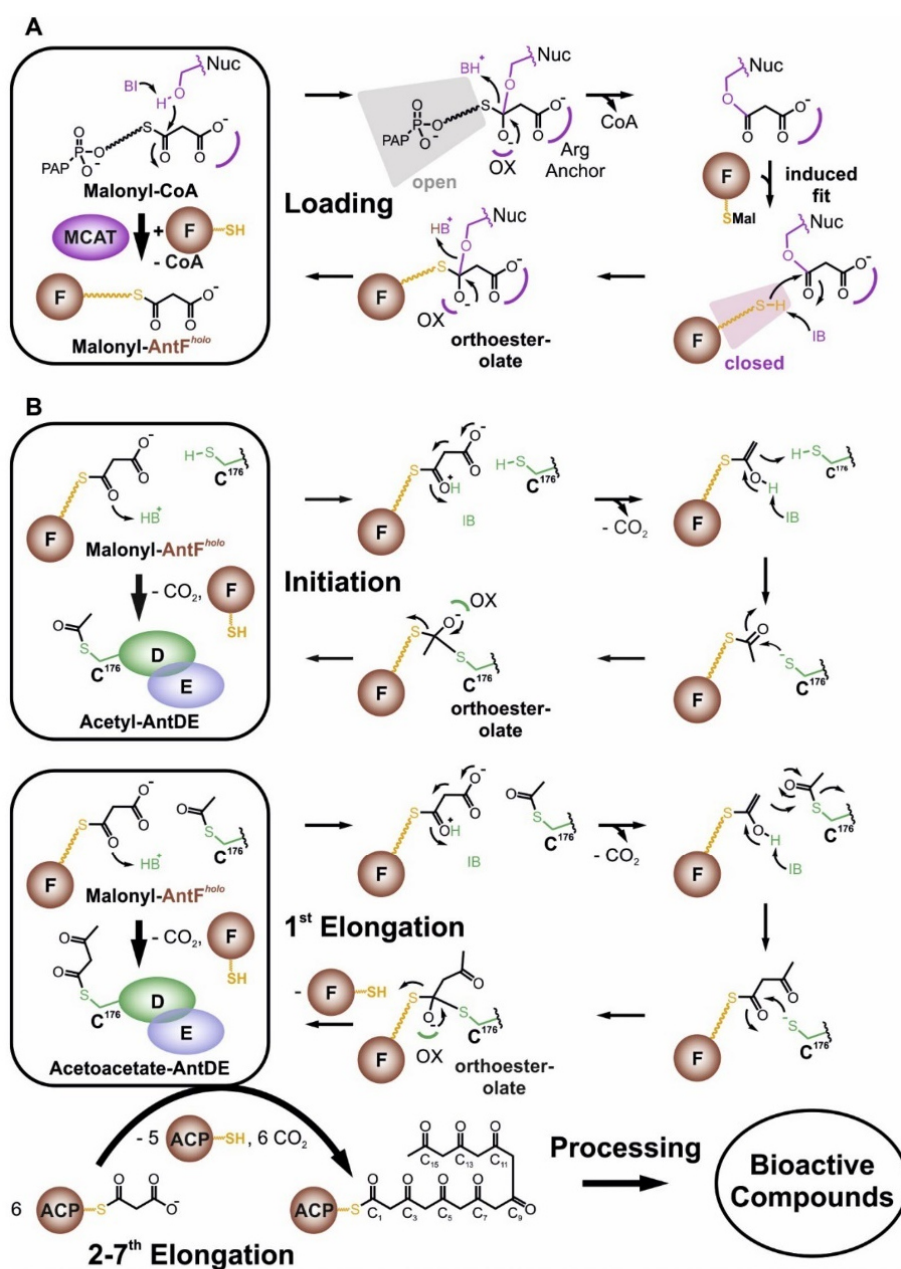


Fig. 3. Proposed reaction mechanism of malonate loading and polyketide chain elongation. (A) Loading of AntF^{holo} via an (O,S,O)-hemi-thio-orthoester-olate (labelled as orthoester-olate) at P₁MCAT (OX = oxyanion hole; Nuc = active site nucleophile of P₁MCAT, Arg-Anchor = specificity pocket stabilizing the carboxylate of the malonyl transfer unit). PAP = adenosine-3'-5'-bisphosphat (B) Octactide formation takes place by chain elongation of eight malonyl precursor units which are delivered one by one to the AntDE ketosynthase complex via AntF^{holo}. Initiation and chain elongation reactions proceed in subsequent steps via (O,S,S)-hemi-dithio-orthoester-olate.

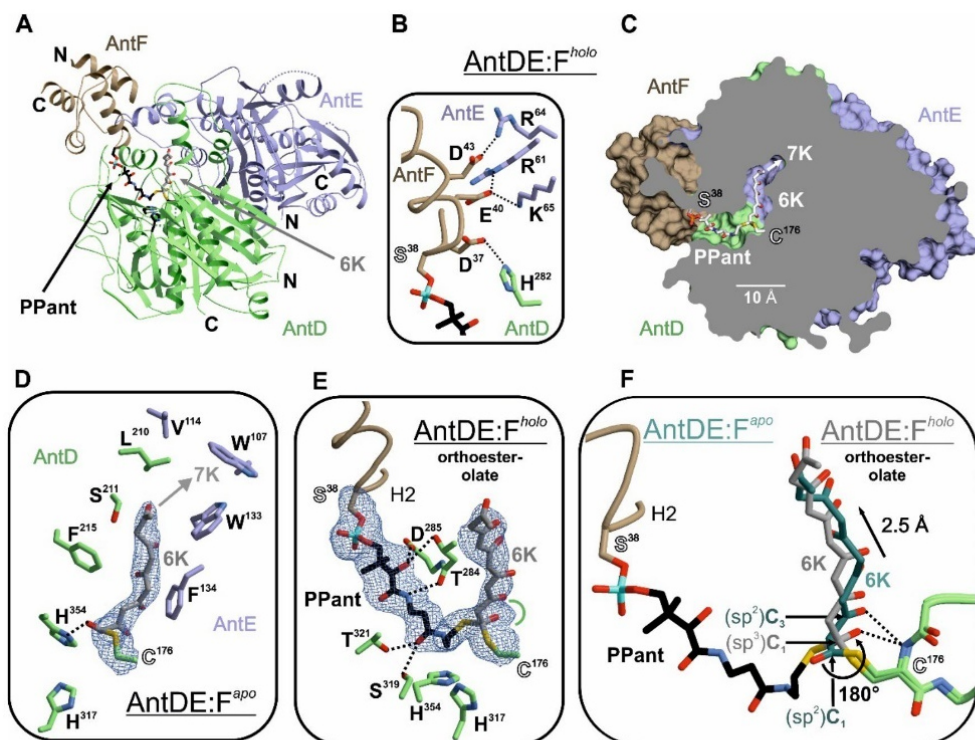


Fig. 4. Crystal structures of AntDE:F^{apo} and AntDE:F^{holo} . (A) Ribbon drawing of the AntDE:F^{holo} complex. The PPant (black) and the bound hexaketide (6K, grey) are shown as sticks. (B) Prominent side chain interactions of the interface (hydrogen-bonds are depicted as black dots). (C) The sliced surface of AntDE:F^{holo} with the PPant bound 6K intermediate (white sticks) that covalently links AntF^{holo} to AntDE. In the chain elongation channel is space left to accommodate a heptaketide (7K, emphasized by a white arrow). (D) Close-up view of residues facing the polyketide binding channel. 6K is bound as a sp^2 -hybridized thioester at Cys176 of AntD and illustrated by its $2F_o-F_c$ electron density map. (E) Zoom-in of the AntDE:F^{holo} complex. Here, 6K forms an sp^3 -hybridized (O,S,S)-hemi-dithio-orthoester-olate and lacks defined stabilization upon position C_3 . The conformational rearrangement at C_1 of 6K is facilitated by characteristic interactions between Ser119 and Thr321 of AntD with the carbonyl oxygen atom of the β -alanine moiety of PPant. (F) The structural superposition of AntDE (gray) with AntDE:F^{apo} (teal) highlights the different hybridization states of 6K causing a ligand mobility of 2.5 Å between both structures.

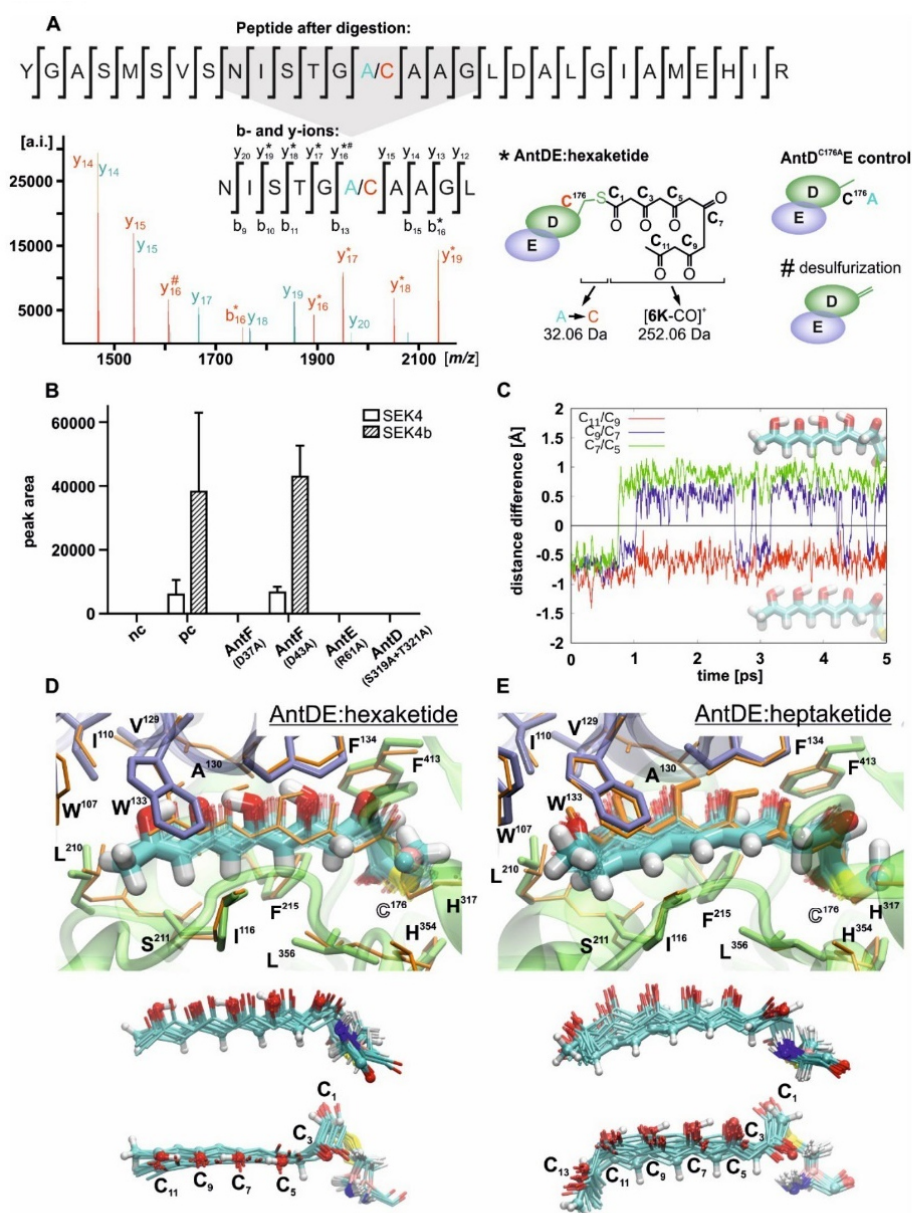


Fig. 5. The AntDE bound hexaketide (6K) is derived from consecutive chain elongation steps. (A) Trypsin digestion experiments combined with HPLC/MS collision-induced dissociation fragmentation analysis between wild type AntDE and mutant AntD(C176A) identify that Cys176 of AntD is modified by 6K. Ions containing the ligand are highlighted with asterisks. Note, the y₁₆-ion of AntDE occurs additionally in a desulfurized form (hash key). **(B)** Identification of important protein-protein interactions of the *Minimal PKS System* by analysis of SEK4 and SEK4b production of *E. coli* strains expressing AntBDEFG variants (nc = negative control, AntBDEFG; pc = positive control, AntBDEFG). **(C)** Proton transfer dynamics between enol units C₁₁-OH and C₉-OH (red); C₉-OH and C₇-OH (green); and C₇-OH and C₅=O (blue) during QM/MM dynamics of the AntDE:hexaketide system. **(D)** Structural snapshot from 5 ps QM/MM dynamics of the AntDE:hexaketide and **(E)** AntDE:heptaketide system. The optimized DFT model is superimposed and presented in orange sticks. Side and top views of the hexaketide during the QM/MM dynamics is shown below. See also fig. S11.

1b

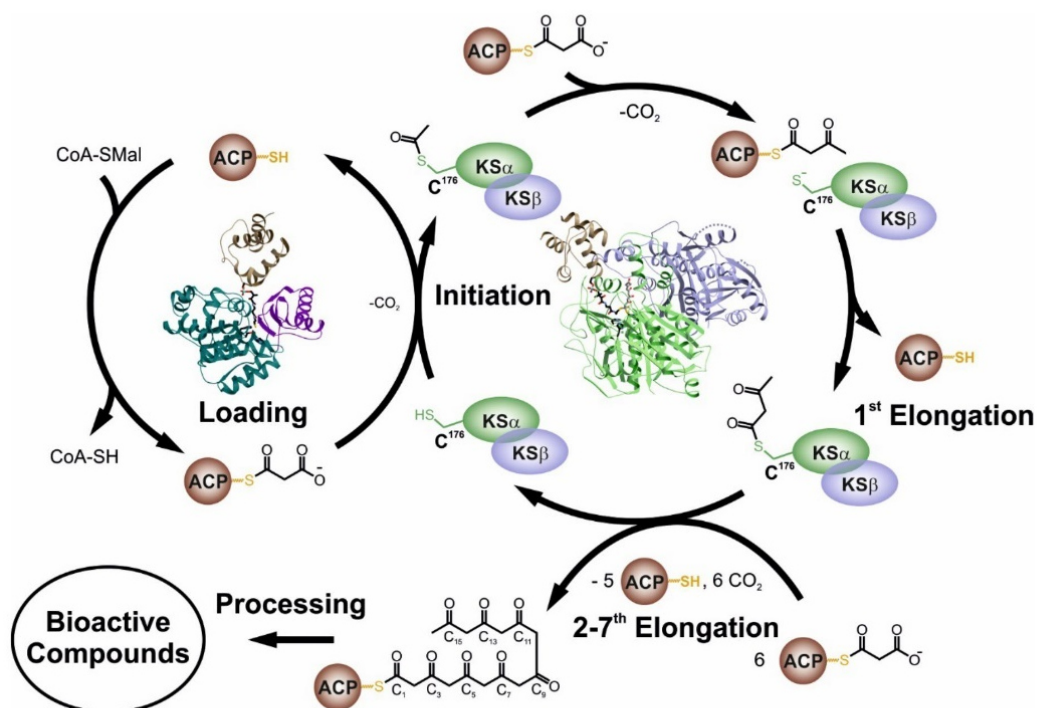


Fig. 6. The Minimal PKS System is responsible for octaketide biosynthesis. Universal to all type II polyketide synthases, transient complex formation of the ACP unit takes place with MCAT (left circle) or KS_α-KS_β-ketosynthase (right circle). Once the 8th chain elongation step is completed the derived octaketide remains bound to the ACP moiety. Depending on the source organism, the key intermediate is further processed to the demanded bioactive compounds.

Supplementary Materials for
**Structural Snapshots of the *Minimal PKS System* Responsible for Octaketide
Biosynthesis**

Alois Bräuer¹, Qiuqin Zhou², Gina L. C. Grammbitter², Maximilian Schmalhofer¹, Michael Rühl², Ville R. I. Kaila^{1,3}, Helge B. Bode^{2*}, Michael Groll^{1*}

*Correspondence to: h.bode@bio.uni-frankfurt.de (H.B.); michael.groll@tum.de (M.G.)

This PDF file includes:

Materials and Methods
Figs. S1 to S13
Tables S1 to S13

Materials and Methods

Cloning and mutagenesis

Genomic DNA from *P. luminescens* TT01 (*I*) was isolated with Gentra® Puregene® kit (Qiagen) according to the protocol for Gram-negative bacteria. The gene *plu2834* (*PIMCAT*) was amplified using primers listed in table S1. The PCR product was digested with restriction enzymes *PstI/EcoRI* and ligated into pCOLADuet vector, generating ZQ37 (pCOLA-His₆-Plu2834). ZQ46 (pCOLA-His₆-AntF^{apo}), ZQ47 (pCOLA-His₆-AntG) and ZQ48 (pCOLA-His₆-AntB) were constructed applying the same procedure. ZQ56 (pACYC-AntDEFG) was obtained using PCR product of *antDEFG* digested with *EcoRI/PstI* and ligated into pACYCDuet. For the generation of ZQ32 (pCOLA-His₆-AntD_AntE), amplified *antE* was digested with *NdeI/Acc65I* and ligated into pCOLADuet vector, resulting in the plasmid pCOLA_AntE. The PCR product of *antD* was digested with *EcoRI/PstI* and ligated into pCOLADuet-AntE, resulting in ZQ32 (pCOLA-His₆-AntDE). ZQ65 (pACYC-His₆-AntF_AntDEFG) was prepared in two steps. The PCR product of *antF* was digested with *PstI/EcoRI* and ligated into pACYCDuet vector, generating the plasmid pACYCDuet_AntF. The PCR product of the genes *antDEFG* was digested with *BglII/Acc65I* and ligated into pACYCDuet-AntF, yielding ZQ65 (pACYC-His₆-AntF_AntDEFG). ZQ81 (pACYCDuet-His₆-RemC_AntDE; *remC*: ACP from the resistomycin biosynthesis) and ZQ82 (pACYCDuet-His₆-ActI-ORF3_AntDE; *actI-ORF3*: ACP from the actinorhodin biosynthesis) were cloned using artificial genes *actI-ORF3** and *remC**, synthesized by the company Life Technologie™ after codon optimization for *E. coli* expression. Therefore, *antDE* genes were amplified and the PCR fragments digested with *BglII/Acc65I* for ligation into pACYCDuet to obtain pACYCDuetAntDE. The artificial genes were digested with *EcoRI/PstI* and ligated into pACYCDuetAntDE, resulting in ZQ81 (pACYCDuet-His₆-RemC_AntDE) and ZQ82 (pACYCDuet-His₆-ActI-ORF3_AntDE).

DNA sequence of the synthesized gene encoding ACP from the actinorhodin biosynthesis (*actI-ORF3**):

```
EcoRI GAATTC CATGGCAACCCTGCTGACCACCGATGATCTGCGTCGTGCACTGGTT
GAATGTGCCGGT GAAACCGATGGCACC GATCTGAGCGGTGATTTTCTGGATCTGCGT
TTTGAAGATATTGGCTATGATAGCCTGGCACTGATGGAAACCGCAGCACGTCTGGAA
AGCCGTTATGGTGTTAGCATTCCGGATGATGTTGCAGGTCGTGTTGATACACCGCGT
GAACTGCTGGATCTGATTAATGGTGCCTGCAAGCAGCATAACTGCAGPstI
```

DNA sequence of the synthesized gene encoding ACP from the resistomycin biosynthesis (*remC**):

```
EcoRI GAATTC CATGGATGTTTCGTGAACAGGTTTTTGCAGTTATGCGTGGTGTGGT
TTTGATGGTGATGAACCTGGCAGATGGTAAACGTCTGGCAGAAGATCTGGATATTGAT
AGCACCGA ACTGGTTGAAATTGTTGTTGCACTGGAACAGCATTGATATTACCGTT
GATGCAGATGCAGAAGGTGGTTTTACCACCGTTGGTGATCTGGTTGGTTGTGTTGGT
CGTCTGCTGAGCGCAGGTAGCGCAGCACGTGGTTAACTGCAGPstI
```

The gene encoding for Sfp was cloned into pCDFDuet-1 using *NdeI/Acc65I* resulting in ZQ78. ZQ62 encoding for AntDEFGH and ZQ76 encoding for AntABC were constructed with pACYCDuet and pCDFDuet, respectively. *antB*, *antABC*, *antAB*, *antBC* PCR products were amplified and cloned in pSU18 with *NdeI/SacI* to result in ZQ6 (pSUAntB), ZQ10 (pSUAntABC), ZQ11 (pSUAntAB) and ZQ16 (pSUAntBC), respectively. Additionally, *antDEF*, *antDEFG*, *antDEFGH* and *antDEFGHI* genes were amplified and ligated into pJET1.2/blunt, generating the

ZQ1 (pJET1.2/blunt-AntDEF), ZQ9 (pJET1.2/blunt-AntDEFG), ZQ12 (pJET1.2/blunt-AntDEFGH) and ZQ13 (pJET1.2/blunt-AntDEFGHI).

Mutations in the wild type genes of AntF (ZQ56_AntF_D37A, ZQ56_AntF_D43A), AntD (ZQ32_AntD_C176A, ZQ56_AntD_S319A_T321A), and AntE (ZQ56_AntE_R61A) were introduced via QuikChange[®] mutagenesis (Agilent Technologies, Waldbronn, Germany). *E. coli* DH10B or BL21 (DE3) was transformed with corresponding plasmids resulting in *E. coli* strains (ES) listed in tables S2 and S3.

Heterologous production of SEK4 and SEK4b

BL21 (DE3) Star was transformed with ZQ56_AntF_D37A, ZQ56_AntF_D43A, ZQ56_AntE_R61A and ZQ56_AntD_S319A_T321A and ZQ48 (AntB, PPTase), respectively, resulting in ES101, ES102, ES103 and ES104. ZQ56 (AntDEFG, minimal PKS plus CoA-ligase) in combination with ZQ48 (ES24) leads to the production of SEK4 and SEK4b and was used as a positive control; ZQ56 (ES26) does not yield in any SEK production and served as a negative control.

SEK4 and SEK4b were obtained from 5 mL LB-culture. Therefore, the corresponding overnight pre-culture was used for inoculation of 10 mL LB-medium to an OD₆₀₀ of 0.2 and supplemented with 1 mM isopropyl-thiogalactopyranoside (IPTG). After 48 h incubation at 30°C and 200 rpm, 5 mL cell culture was acidified with concentrated HCl to pH 4 and extracted 1:1 (v/v) with ethyl acetate. The organic layer was dried under reduced pressure and dissolved in 1 mL methanol for *high resolution* (HR)-HPLC/MS measurement using a Dionex Ultimate 3000 LC system (Thermo Fisher) that is equipped with a DAD-3000 RS UV-detector (Thermo Fisher) and coupled to an Impact II electrospray ionization mass spectrometer (Bruker). Compound separation was performed with a C18 column (ACQUITY UPLC BEH, 50 mm × 2.1 mm × 1.7 μm, Waters) using ACN and H₂O containing 0.1 % formic acid as mobile phases at a flow rate of 0.4 mL/min with a 14 min gradient of 5-95 % ACN. For internal mass calibration a 10 mM sodium formate solution was injected. The HPLC/UV/MS analysis was operated in positive mode with a mass range of *m/z* 100-700 and an UV at 190-800 nm.

ActI-ORF3 and RemC are two ACPs in the type II PKS systems for actinorhodin and resistomycin biosynthesis, respectively. Both ACPs originate from *Streptomyces*, which is phylogenetically close to *P. luminescens*. Therefore, we selected ActI-ORF3 and RemC for coexpression with AntDE (table S3). ZQ81 and ZQ82 were transformed into *E. coli* BL21 (DE3) containing ZQ78. However, polyketide shunt products were not identified in these engineered strains (table S6). Therefore, AntF is essential for octaketide biosynthesis in our studied *Minimal PKS System* and cannot be replaced by the orthologues ActI-ORF3 and RemC.

Nano-LC/MS to confirm the presence of a hexaketide bound to Cys176 of AntD

Sample preparation was adapted from Kulak *et al.* (2). Therefore, 5 μL of 100 μg/mL protein solution (AntDE or AntD^{Cys176A}E in 25 mM Tris pH 8.5) was mixed with 20 μL 6 M guanidinium chloride (GdmCl) in 100 mM Tris pH 8.5 with 10 mM DTT and incubated for 5 min at 95°C. Reduced thiols were alkylated by adding 2.5 μL of 400 mM chloroacetamide. Prior 20 min incubation at RT, the samples were diluted with 25 mM Tris pH 8.5 with 10 % acetonitrile (ACN) to obtain a final GdmCl concentration of 0.6 M. Proteins were digested with 1 μg Trypsin/LysC (sequencing grade, Promega) overnight at 37°C under gentle agitation. Digestion was stopped by adding trifluoroacetic acid to a final concentration of 0.1 %. Peptides were desalted on multi-stop-

and-go tips (StageTips) containing three C18-disks. The eluted peptides were dried and resuspended in 15 μ L 50 % ACN containing 0.1 % formic acid for nano-LC measurement (~25 μ g protein per analysis).

Nano-LC/MS was performed on a Thermo Scientific Q Exactive Plus system equipped with an ultra-high performance liquid chromatography unit (Thermo Scientific Dionex Ultimate 3000) and a Nanospray Flex Ion Source (Thermo Scientific). Peptides were separated with an in-house packed column (100 μ m i.d. x 30 cm x 2.4 μ m separated C18-resins) with ACN and H₂O containing 0.1 % formic acid as mobile phases using a gradient from 4 %-30 % ACN for 30 min, followed by 80 % ACN for 60 min and 60 % ACN for 30 min, with a flow rate of 300 nL/min. The mass scan range was set to m/z 300–2000. The resulting MS data were edited with mMass version 5.5.0.

Analysis of polyketides by HPLC/MS

Experiments and analysis were performed as described in Zhou *et al.* (3).

Culture growth and protein purification

The recombinant *E. coli* strains ES20 (AntF^{apo}), ES23 (AntF^{holo}, posttranslational modification was proved by MALDI-TOF), ES19 (AntG), ES18 (plu2834/MCAT), ES21 (AntDE), and ES100 (AntD^{C176A}E) were grown in shaking flasks containing 3 L of lysogeny broth supplemented with the respective antibiotic. The cultures were incubated at 37°C with shaking to an OD₆₀₀ of 0.6. Cells were cooled at 4°C for 30 min and protein expression was induced with 0.1 mM IPTG at 16°C for 16 h. Cells were harvested by centrifugation, washed with 0.9 % (w/v) NaCl and stored at -20°C.

Frozen bacterial cell mass was thawed in 50 mM Tris, pII 7.5, containing 500 mM NaCl and 20 mM imidazole (buffert A, 5 ml per 1 g pellet). The cells were disrupted by sonification (Branson Digital Sonifier 250, G.Heinemann, Schwäbisch Gmünd, DE). The resulting suspension was centrifuged at 21,000 g for 30 min at 4°C using a SIGMA 4K15 centrifuge (Sigma Aldrich, DE). The supernatant was applied to a column of HisTrap™ HP column (GE Healthcare Life Science, Uppsala, Sweden; column volume, 5 mL), which had been equilibrated with buffer A (flow rate, 5 mL/min) using ÄKTA prime plus (GE Healthcare, Uppsala, Sweden). The column was washed with buffert A (50 mL) and buffer A containing 50 mM imidazole hydrochloride (25 mL). The protein was eluted with buffer A containing 200 mM imidazole hydrochloride. Fractions were combined, concentrated and applied to a column of HiLoad™ 16/60 Superdex™ 75 pg (GE Healthcare, Uppsala, Sweden), which had been equilibrated with 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM dithiothreitol (DTT), and developed with a flow rate of 1.6 mL/min. The solution was concentrated to a final concentration of 15 mg/mL for AntF^{apo}/AntF^{holo} ($\epsilon_{\text{apo/}} \text{AntF} = 6990 \text{ M}^{-1} \text{ cm}^{-1}$), of 20 mg/ml for P/MCAT ($\epsilon_{\text{P/MCAT}} = 34950 \text{ M}^{-1} \text{ cm}^{-1}$), and 23 mg/ml for AntDE ($\epsilon_{\text{AntDE}} = 59250 \text{ M}^{-1} \text{ cm}^{-1}$) using a 10k Amicon Ultra Centrifugal Filter Device (Millipore, Billerica, MA), flash frozen with liquid nitrogen and stored at -80°C.

Crystallization and structure determination

Crystals of AntF^{apo}, AntF^{holo}, P/MCAT:AntF^{holo}, AntDE:F^{apo} and AntDE:F^{holo} were grown at 20°C by using sitting drop vapor diffusion method. **AntF^{apo}**: 0.2 μ L AntF^{apo} (15 mg/ml in 20 mM Tris pH 7.5, 100 mM NaCl) + 0.2 μ L 1 M LiCl, 0.1 M citric acid pH 4.0-4.5, 30-33 % PEG6000; mountable crystals grew within two days. **AntF^{holo}**: 0.2 μ L AntF^{holo} (15 mg/ml in 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT) + 0.2 μ L 0.1 M citric acid pH 5.0, 3.2 M (NH₄)₂SO₄; mountable

crystals grew within 5-7 days. **P/MCAT:AntF^{holo}**: 0.2 μ L protein (70 % (v/v) of 20 mg/ml P/MCAT in 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT + 30 % (v/v) of 15 mg/ml AntF^{holo} in 20 mM Tris pH 7.5, 100 mM NaCl) + 0.2 μ L 1.8-2.0 M Na-K-phosphate pH 7.3-7.5, 3-4 % (v/v) polypropylene glycol P400; mountable crystals grew within four weeks; several rounds of seeding were necessary. **AntDE:F^{apo}**: 0.2 μ L protein (75 % (v/v) of 22 mg/ml AntDE in 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT + 25 % (v/v) 15 mg/ml AntF^{apo} in 20 mM Tris pH 7.5, 100 mM NaCl) + 0.35 μ L 29 % PEG3350, 0.4 M (NH₄)₂SO₄, 0.1 M BisTris pH 6.2; mountable crystals grew within three weeks; several rounds of seeding were necessary. **AntDE:F^{holo}**: 0.2 μ L protein (75 % (v/v) 22 mg/ml AntDE in 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT + 25 % (v/v) 15 mg/ml AntF^{holo} in 20 mM Tris pH 7.5, 100 mM NaCl) + 0.35 μ L 23 % PEG3350, 0.3 0.5 M (NH₄)₂SO₄, 0.1 M BisTris pH 6.0; mountable crystals grew within three weeks; several rounds of seeding were necessary.

Each crystal was soaked for 30 seconds in its reservoir solution containing 25 % (v/v) PEG400 (AntF^{apo}, AntF^{holo}, AntDE:F^{apo}, AntDE:F^{holo}) or 25 % (v/v) glycerol (P/MCAT:AntF^{holo}) and vitrified in liquid nitrogen. Diffraction datasets were recorded using synchrotron radiation of $\lambda = 1.0 \text{ \AA}$ at the beamline X06SA, Swiss Light Source (SLS), Paul Scherrer Institut, Villigen, Switzerland. Collected reflections were processed with the program package XDS (4). Starting phases were received by Patterson search techniques applying PHASER (5) and using coordinates of EcACP (PDB ID: 2FAD) (6), EcMCAT (PDB ID: 2G1H) (7), and the actinorhodin K α -K β of *Streptomyces coelicolor* (PDB ID: 1TQY) (8) as appropriate models, respectively. Model building was performed with MAIN (9) and Coot (10). Repetitive building and refinement cycles by REFMAC5 (11) resulted in well-defined 2F_o-F_c-electron density maps that depict protein-protein interactions, the PPant cofactor attached to Ser38 of AntF, as well as the poly- β -keto hexaketide intermediate covalently bound to Cys176 of AntD. The AntDE:F^{holo} ternary complex revealed a late reaction intermediate of chain elongation with the hexaketide covalently linking AntDE with AntF^{holo}. Geometric constraints for each ligand were calculated with SYBYL-X. Water molecules were placed automatically using ARP/wARP (12). The quality of the stereochemistry was confirmed by the Ramachandran plot. All statistics are summarized in the Supplementary Tables S7 and S9.

DFT structure optimization and QM/MM dynamics

Quantum chemical DFT models were constructed based on our solved x-ray structure of AntDE (PDB ID: 6SMO). The ketide was modeled as a hexaketide or heptaketide with C₇, C₉, and C₁₁ carbons in their respective enol form, and including Ile116, Ala119, Gly175 (backbone), Cys176, Leu210, Ser211, Ile212, Ser214, Phe215, His317, His354, Leu356, Ala357, Gly412 (backbone), Phe413 of AntD and Trp107, Ile110, Val129, Ala130, Thr131, Ala132, Trp133, Phe134, of AntE. His317 was considered in its protonated (HisH⁺) form, whereas all other residues were used in their standard protonation states. Protein residues were cut at the C β position. In addition, backbone atoms were included for protein residues that formed close contacts with the ketide. The complete ketide unit and hydrogen atoms within the protein part of the molecule were optimized at the B3LYP-D3/def2-SVP/ $\epsilon=4$ level (13-17). The models comprise $N=358$ for the hexaketide (6K) and $N=363$ atoms for the heptaketide (7K). The isolated systems including Cys176 of AntD in complex with either 6K or 7K were optimized at the B3LYP-D3/def2-SVP/ $\epsilon=4$ level.

QM/MM models of AntDE were constructed based on our solved x-ray structure (PDB ID: 6SMO). The protein was embedded in a water-ion environment with 100 μ M NaCl concentration.

The model was further trimmed to include a 30 Å radius around the polyketide binding cavity, leading to a system with around 14,000 atoms. The 6K/7K subsystem was modeled at the B3LYP-D3/def2-SVP level of theory (13-17) by introducing an atom between C^α-C^β atoms of Cys176 (AntD), and electrostatically embedded in the surrounding system that was treated at the CHARMM36 theory level (18). Furthermore, a QM/MM model for the octaketide was constructed following the same principles as for the 6K- and 7K. After QM/MM minimization, 5 ps QM/MM MD simulations were performed with a 1 fs timestep at $T=310$ K, with $r>10$ Å beyond the ketide cavity fixed.

DFT models of AntDE:F^{holo} (PDB ID: 6SMP) with $N=443$ atoms were optimized at the B3LYP-D3/def2-SVP/ $\epsilon=4$ level (13-17). In addition to residues of the AntDE model, the system included Ser319, Thr321, Thr284, Asp285 of AntD. A stable (O⁻,S,O)-hemi-thio-orthoester-olate bond was obtained by protonating the carbonyl oxygen by deprotonation of His317.

All QM calculations were performed with TURBOMOLE v. 6.6 (19), CHARMM v. 38b1 (20) was utilized to propagate the classical, and the QM/MM dynamics was studied using an in-house version of the CHARMM/TURBOMOLE interface (21). Visualization of the calculated data was carried out with VMD (22).

AntF-*P*IMCAT transfer assay

The unidirectional reaction of the malonate-moiety by *P*IMCAT to AntF^{holo} was analyzed by transfer assays. Therefore, 50 μM AntF^{holo} and *P*IMCAT were incubated with 1 mM Mal-CoA in 50 mM Tris buffer (pH 7.5) that yielded Mal-AntF. The reaction was incubated for 8 h at 30°C, while further incubation did not yield higher amount of Mal-AntF. Mal-CoA was removed by washing the reaction mixture eight times with 50 mM Tris (pH 7.5), using an Amicon Ultra-0.5 30K device (Merck Millipore, MWCO 3 kDa; centrifugation was carried out at 14 000 x *g* for 20 min). After each washing cycle, fresh *P*IMCAT (50 μM) was added and incubated for 2 h. Reactions were stopped by adding 10 % ACN with 0.1 % formic acid in a 1:1 ratio (v/v). HR-HPLC-MS analysis was performed using a Dionex Ultimate 3000 LC system (Thermo Fisher) that is equipped with a MWD-3000 RS UV-detector (Thermo Fisher) and coupled to a micrOTOF II electrospray ionization mass spectrometer (Bruker). Back transfer of the malonyl moiety from Mal-AntF to *P*IMCAT was not recorded.

HPLC/MS analyses of protein samples were performed using a C3 column (Zorbax 300SB-C3 300 Å, 150 mm x 3.0 mm x 3.5 μm, Agilent) with ACN and H₂O containing 0.1 % formic acid as mobile phase at a flow rate of 0.6 mL/min and applying a 25.5 min gradient of 30-65 % ACN. For internal mass calibration an ESI-L Mix (Agilent) was injected. The HPLC-MS analysis was operated in positive mode in a mass range of m/z 50-1500.

Data analyses of UV-MS-chromatograms were performed with Compass DataAnalysis 4.3 (Bruker). The theoretical average masses of proteins were calculated using Compass IsotopePattern 3.0 (Bruker).

UTL-MALDI-TOF MS

E. coli strains ES5, ES7 and ES8 (Table S3) were used for heterologous expression of *ant* genes. Collected cell pellets were treated with 10 μg/ml lysozyme and 2.5 U/ml benzonase in PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4). For the preparation of the UTL substrate, 10 μL of a solution, consisting of 1 part of saturated α -cyano-4-hydroxycinnamic acid (CHCA) in 66 % ACN and 0.1 % trifluoroacetic acid (TFA) combined with 3 parts of isopropanol, were spread on the UTL-plate. In the following, cell lysate of the distinct

expression strains was mixed with a matrix solution (1:10, 0.5 μ L sample with 4.5 μ L matrix solution in CHCA, 3 parts of formic acid and 2 parts of isopropanol) and spotted according to Fenyo *et al.* (23).

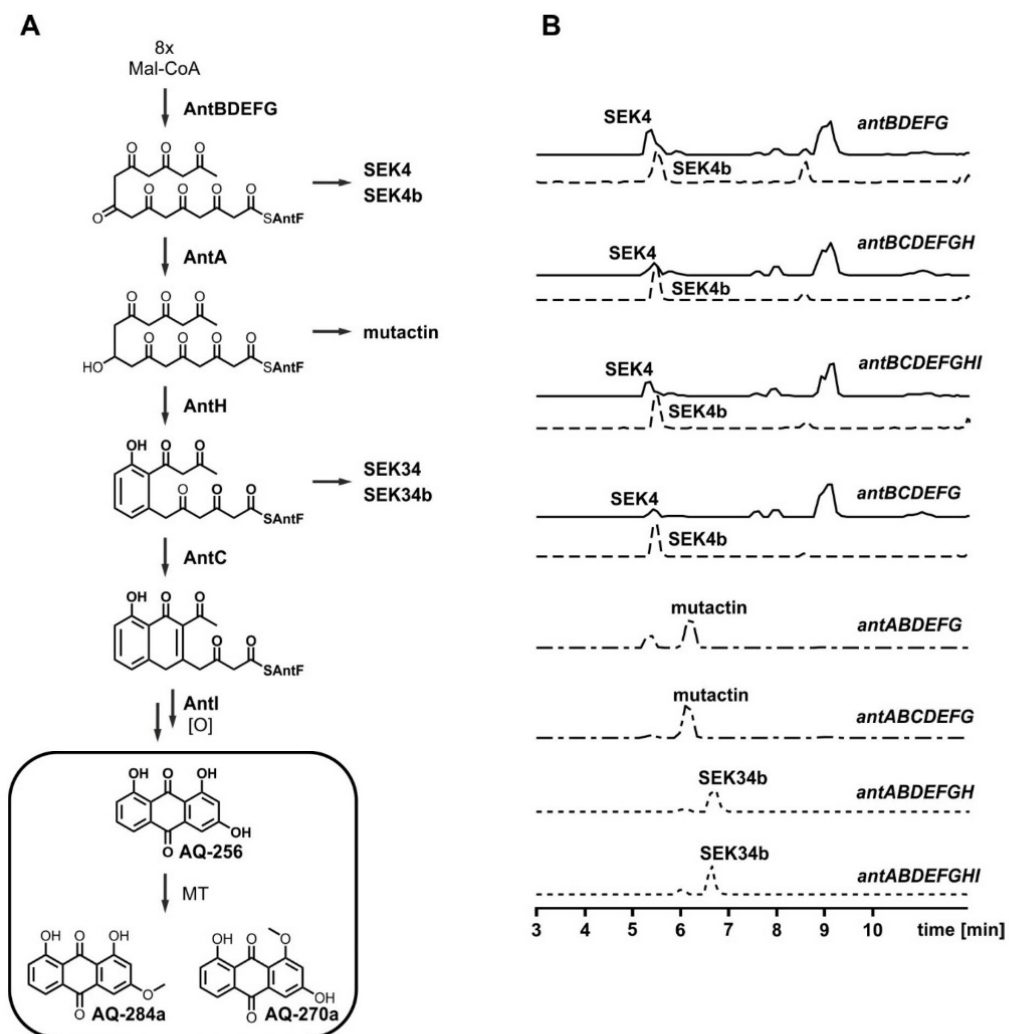


Fig. S1. Identification of genes involved in AQ biosynthesis.

(A) Biosynthesis of anthraquinones (AQs) in *P. luminescens*. Heterologous expression of the individual genes in *E. coli* is accompanied by the formation of shunt products. (B) HPLC/MS of *E. coli* strains expressing different *ant* genes. Shown are extracted-ion chromatograms (EICs) for SEK4 (continuous line, m/z 301 $[M+H]^+$), SEK4b (dashed line, m/z 319 $[M+H]^+$), mutactin (broken line, m/z 303 $[M+H]^+$) and SEK34b (dotted line, m/z 285 $[M+H]^+$). Used strains, standards and methods of identification adapted from Zhou *et al.* (3).

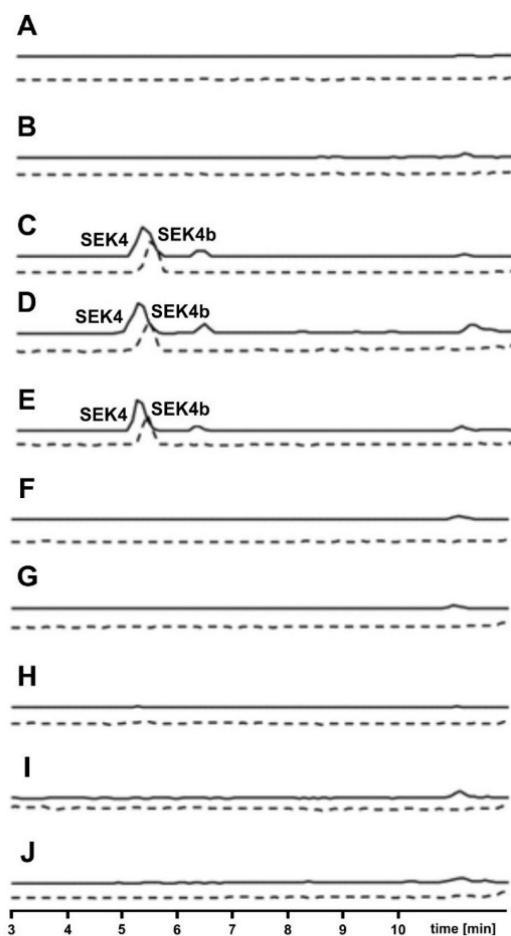


Fig. S2. HPLC/MS data from *in vitro* biosynthesis of SEK4 and SEK4b.

The individual graphs represent EICs for SEK4 (m/z 301 $[M+H]^+$) and SEK4b (dashed line, m/z 319 $[M+H]^+$). (A) AntF^{holo} + buffer; (B) AntDE + P/IMCAT + AntG + AntF^{apo} Mal-CoA + acetyl-CoA + buffer; (C) AntDE + P/IMCAT + AntG + AntF^{holo} + Mal-CoA + acetyl-CoA + buffer; (D) AntDE + P/IMCAT + AntG + AntF^{holo} + Mal-CoA + buffer; (E) AntDE + P/IMCAT + AntF^{holo} + Mal-CoA + buffer; (F) AntDE + P/IMCAT + Mal-CoA + buffer; (G) AntDE + AntF^{holo} + Mal-CoA + buffer; (H) P/IMCAT + AntF^{holo} + Mal-CoA + buffer; (I) AntDE + P/IMCAT + AntF^{holo} + buffer; (J) AntDE + P/IMCAT + AntF^{holo} + Mal-CoA. The combination of AntDE, P/IMCAT, AntF^{holo}, and Mal-CoA (as seen in E) already produce SEK4 and SEK4b.

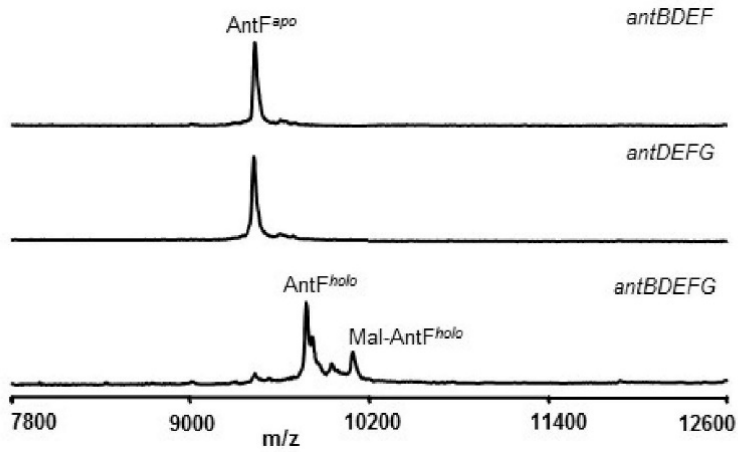


Fig. S3. The heterodimer AntBG transfers PPant from CoA forming AntF^{holo}.

UTL-MALDI-TOF MS results from *E. coli* cells expressing *ant* genes. AntB (PPTase) and AntG (CoA ligase) are essential to produce SEK4 and SEK4b. These two proteins convert AntF^{apo} into its *holo*-form using CoA as substrate.

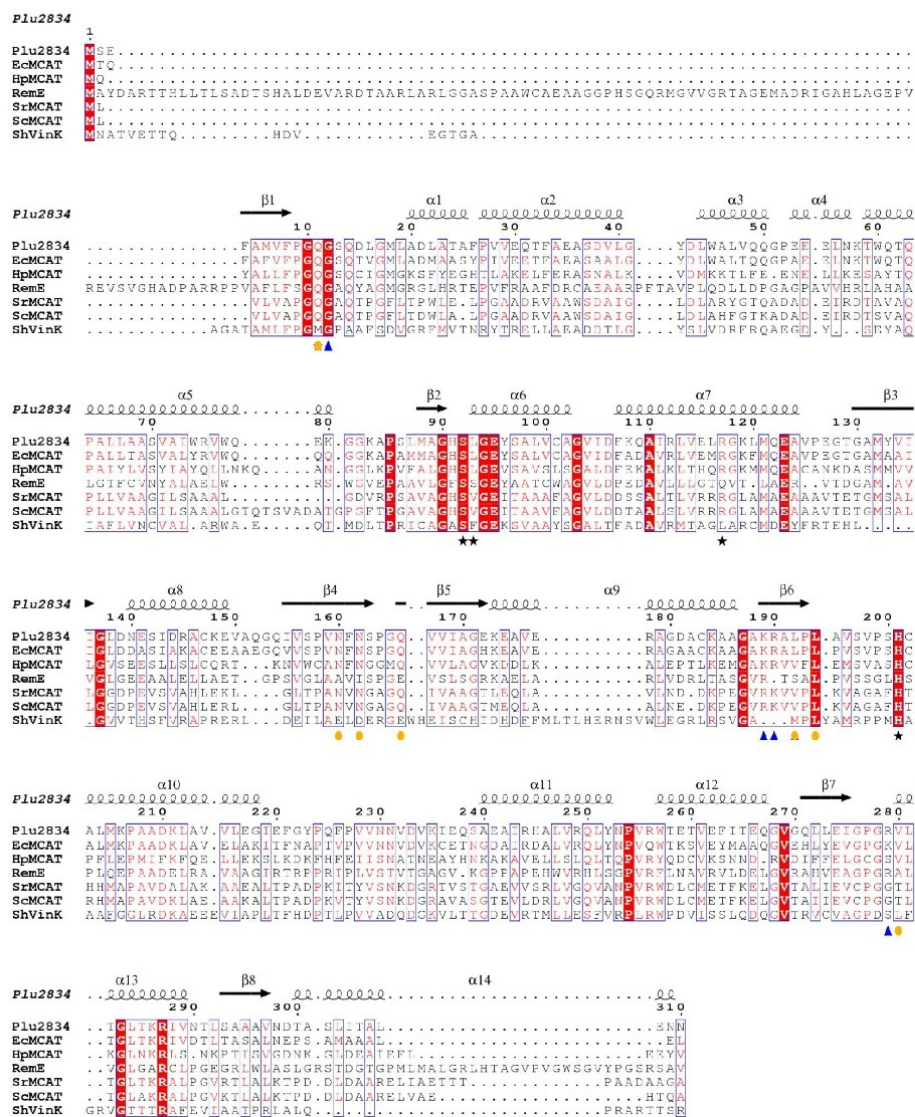


Fig. S4. Sequence alignment of PIMCAT (Plu2834).

WP_011147054.1 and selected MCATs are compared. *EcMCAT* (*Escherichia coli*, FabD, WP_000191372, PDB ID: 2G2Z), *HpMCAT* (*Helicobacter pylori*, WP_077231645, PDB ID: 2H1Y), RemE (resistomycin biosynthesis, *Streptomyces resistomycificus*, WP_053190554), *SrMCAT* (from FAS, *Streptomyces resistomycificus*, WP_030040117), *ScMCAT* (*Streptomyces coelicolor*, WP_011028322, PDB ID: 1NM2), *ShVinK* (*Streptomyces halstedii*, vicenistatin biosynthesis acyltransferase, PDB ID: 5CZD). Active site residues are highlighted with asterisks. Protein:protein interactions of AntF^{holo} and its partner proteins within the *Minimal PKS System* are highlighted by blue triangles and the PPant contact region by orange circles.

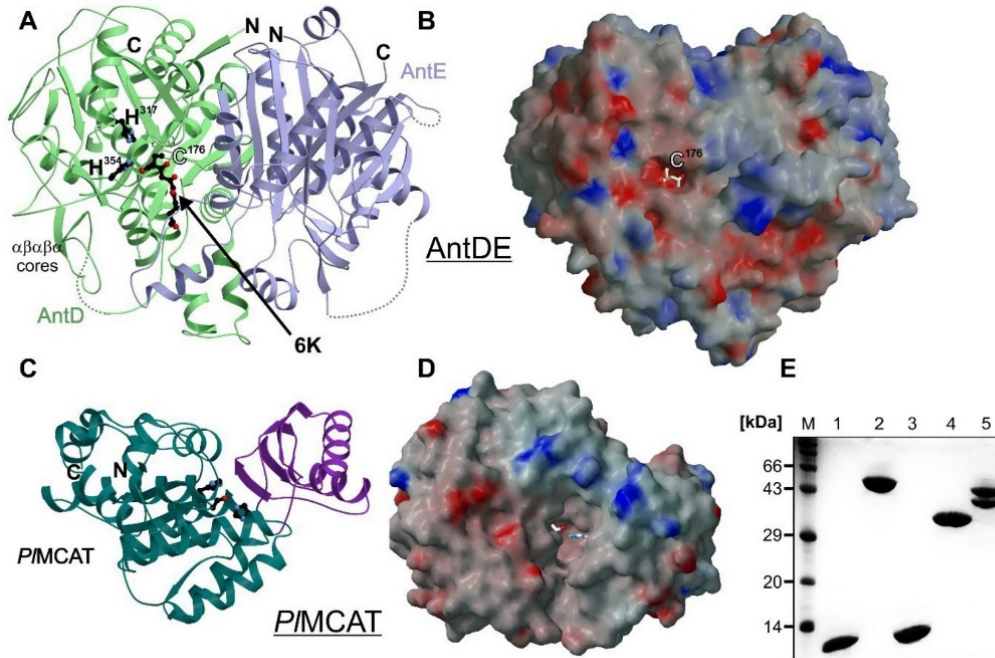


Fig. S5. Crystal structure of AntDE and PIMCAT. (A/C) Ribbon drawing of AntDE and PIMCAT, respectively. Color coding and labelling are according to Fig. 2 and Fig. 3, respectively. The bound hexaketide (6K, black) is shown as sticks. (B/D) Surface representation of AntDE. Negative and positive electrostatic potentials are contoured from $-20k_B T/e$ (red) to $+20k_B T/e$ (blue). (E) SDS-PAGE gel (12 %) of purified proteins stained with Coomassie blue. Protein marker (M), His₆-AntF^{apo} (9 kDa, lane 1), His₆-AntG (60 kDa, lane 2), His₆-AntF^{holo} (9 kDa, lane 3), His₆-PIMCAT (35 kDa, lane 4), His₆-AntDE (47 kDa, 41 kDa, lane 5).

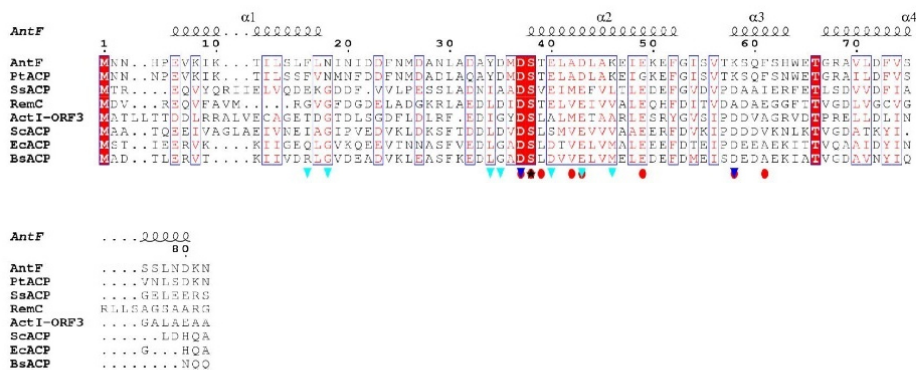


Fig. S6. Sequence alignment of AntF.

WP_011148293.1 and selected ACPs. *PtACP* (*Photorhabdus temperata*, WP_021325342), *SsACP* (*Streptococcus suis*, WP_029175793), *RemC* (*Streptomyces resistomycificus*, resistomycin biosynthesis, WP_030043019), *ActI-ORF3* (*Streptomyces coelicolor*, actinorhodin biosynthesis WP_003973889, PDB ID: 2AF8), *ScACP* (*Streptomyces coelicolor*, from FAS, WP_003976417.1, PDB ID: 2CNR), *EcACP* (*Escherichia coli*, WP_059329731, PDB ID: 2FAD), *BsACP* (*Bacillus subtilis*, WP_003154310, PDB ID: 1F80). $\alpha 2$ is the recognition helix, the PPant binding site at Ser38 is indicated with an asterisk. Protein:protein interactions between AntD and AntE are highlighted by turquoise and the *PIMCAT* contact region by red circles.

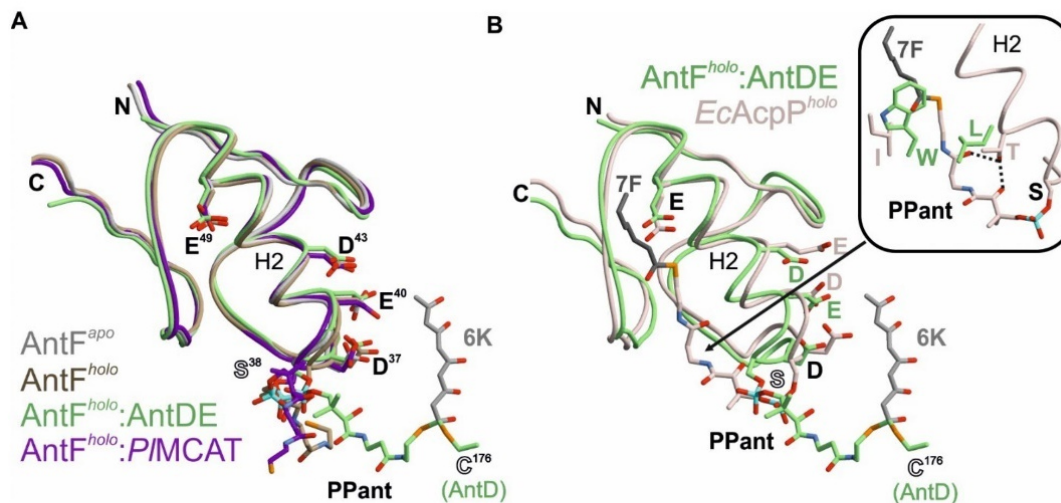


Fig. S7. Acyl-carrier-protein AntF acts as rigid body object.

(A) The structural superposition (coil representations) of AntF^{apo} (gray), AntF^{holo} (brown), AntF^{holo} in complex with PIMCAT (magenta) and AntDE (green) including side chain orientations of Asp37, Glu40, Asp43 and Glu49 reveals a perfect match. In contrast, the PPant group adopts individual positions in each structure. (B) AntF and AcpP from the *E. coli* type II fatty acid synthase (FAS, gray) share a common fold including the conformation of identical (black) and conserved (color-specific) residues that interact with their respective partner proteins. However, there are differences between AntF and AcpP. Fatty acid intermediates are covalently bound to the same ACP that delivers the growing acyl chain to respective clients during each cycle of elongation. During transport, the evolving fatty acid is protected and immobilized in a hydrophobic cavity as seen in the AcpP structure (PDB ID: 2FAD) bound to heptanoic acid (7F) (6). The zoom-in box shows that AntF lacks such binding pocket due to a tryptophan residue. Moreover, PPant cannot be stabilized in this conformation since threonine is replaced by leucine (H-bonds are drawn as black dots). This peculiarity might explain rapid formation of the shunt products SEK4 and SEK4b once the highly reactive octaketide is bound to AntF^{holo}.

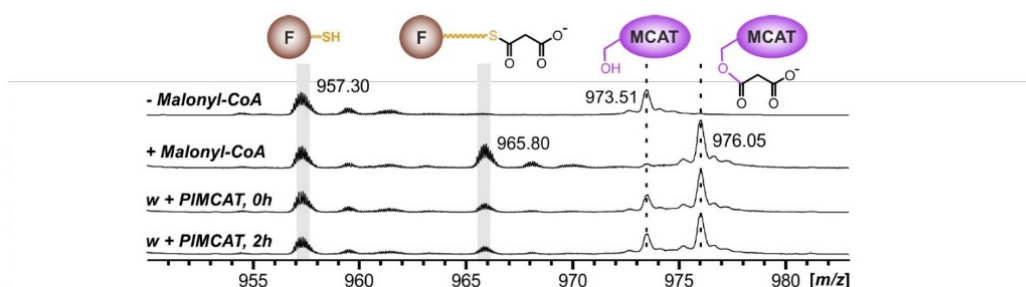


Fig. S8. *P/MCAT* catalyzes malonyl-transfer from Mal-CoA to AntF^{holo}.

P/MCAT instantly transfers the malonyl building block to AntF^{holo} using Mal-CoA as designated substrate. The transacylation takes place by a covalently bound Mal-*P/MCAT* oxoester-intermediate as proven by HPLC/MS analysis. In contrast, after extensive washing (w) steps and addition of fresh *P/MCAT*, formation of Mal-*P/MCAT* is not observed in presence of Mal-AntF (back reaction). These findings agree with the crystallographic data and suggest a unidirectional malonyl transfer mechanism. Analyses were performed between 0 and 2 hours (h). The MS¹-mass spectra illustrate 10⁺-charge states for AntF^{holo} and Mal-AntF as well as 34⁺-charge states for *P/MCAT* and Mal-*P/MCAT*.

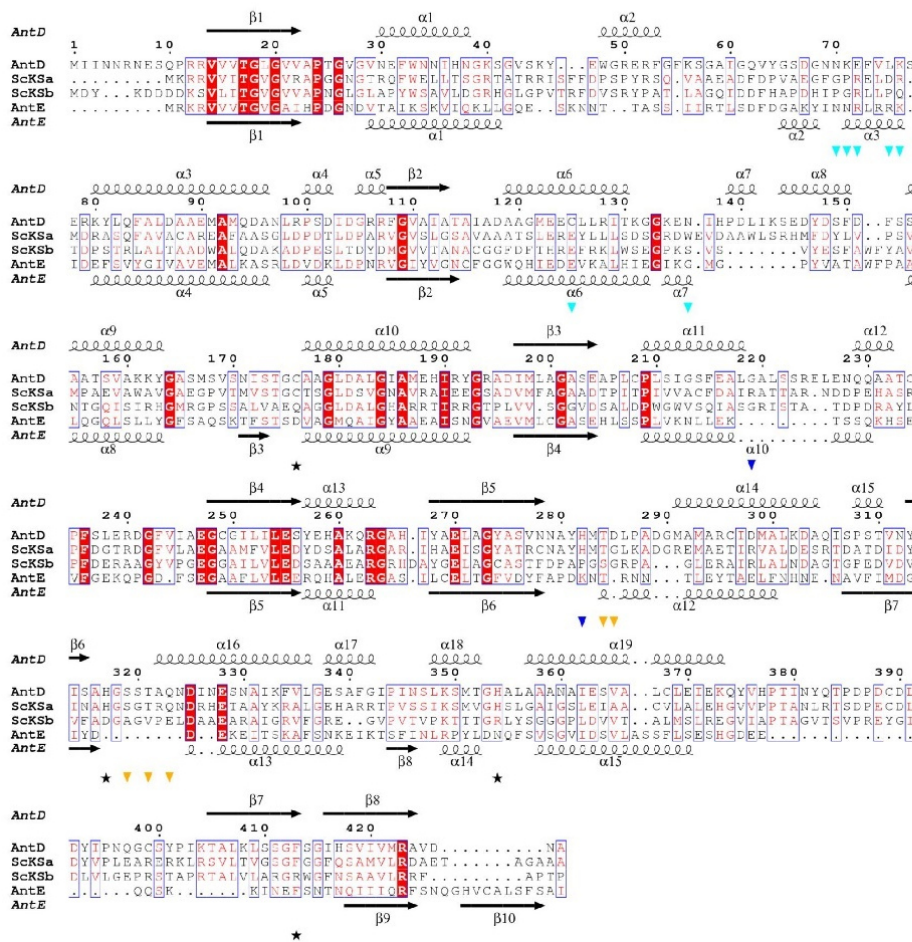


Fig. S9. Sequence alignment of KS α (KS) and KS β (CLF) proteins.

AntD (KS α) and AntE (KS β) from *P. luminescens*, ScKSa (KS α) and ScCLF (KS β) from *Streptomyces coelicolor* (PDB ID: 1TQY). Active site residues are indicated with asterisks. AntD and AntE residues interacting with AntF are highlighted by blue and cyan triangles; PPant interactions are shown by orange triangles.

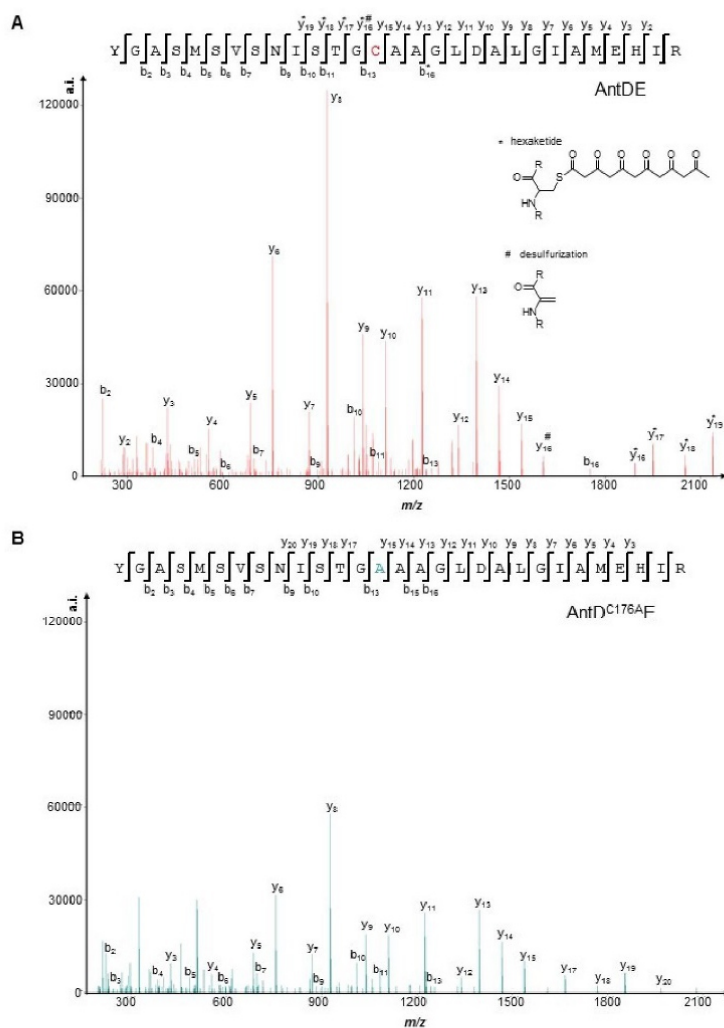


Fig. S10. Hexaketide bound to AntD.

MS-analysis identifies Cys176 of AntD to be modified by a bound hexaketide. (A) Spectra of tryptic digestion experiments of wild type AntDE in comparison with the AntD^{C176A}E mutant. (B). Matched b- and y-ESI-fragments are marked within the peptide sequence (list of fragments with corresponding masses are summarized in Tables S10 and S11). Characteristic desulfurization event of the modified Cys176 causes a mass shift towards dehydroalanine in MS². Experiments have been performed according to Rühl *et al.* (24). The modified cysteine is marked with an asterisk and dehydroalanine with a hash.

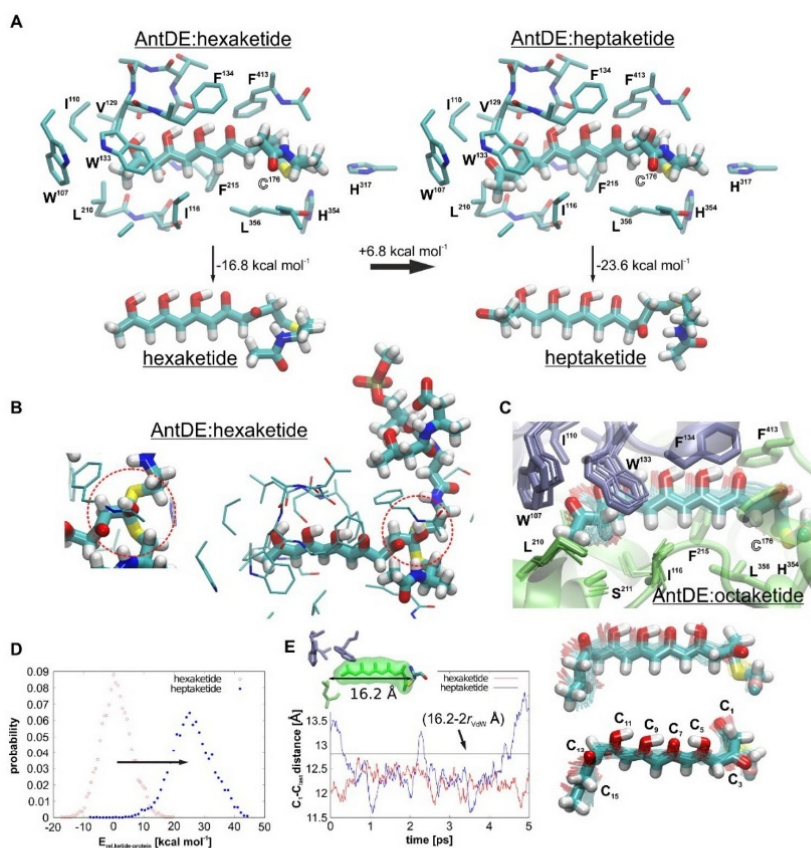


Fig. S11. DFT structure optimization and QM/MM dynamics

(A) DFT optimized model of the AntDE:hexaketide system (*top, left*) and the isolated Cys176-hexaketide model (*top, right*). $16.8 \text{ kcal mol}^{-1}$ (B3LYP-D3/def2-TZVP/ $\epsilon=4$) of electronic energy is released upon optimizing the Cys176-hexaketide from its strained protein conformation. DFT model of the AntDE:heptaketide system (*bottom, left*) and the isolated Cys176-heptaketide (*bottom, right*). $23.6 \text{ kcal mol}^{-1}$ (B3LYP-D3/def2-TZVP/ $\epsilon=4$) electronic energy is released upon optimizing the Cys176-heptaketide from its strained protein conformation. The difference in strain contributions in the DFT models is $6.8 \text{ kcal mol}^{-1}$. (B) Optimized DFT model of AntDE:F^{holo}. A stable thio-orthoester-olate intermediate could be obtained upon protonation of the carbonyl group, which compensates the electrostatic repulsion between the carbonyl oxygen at C₃ and the alkoxide (*inset*). (C) QM/MM dynamics of AntDE:octaketide. (D) Relative interaction energy distribution of the hexaketide (red open circles) and heptaketide (blue closed circle) with the surrounding AntDE-protein. The QM-MM interaction energy is more strongly stabilized in the AntDE-hexaketide system relative to the heptaketide, indicating accumulation of a molecular strain during this chain elongation process. (E) *Edge-to-edge* distances between C₁-C₆ (hexaketide, in red) and C₁-C₇ (heptaketide, in blue) during QM/MM dynamics. The black line shows the substrate cavity length in AntDE, measured from the Cys176 (C^α) to Leu210 (C^γ) with $r=16.2 \text{ \AA}$. After subtracting, the typical carbon atom radius ($r_{vdW}=1.7 \text{ \AA}$) is twice as much.

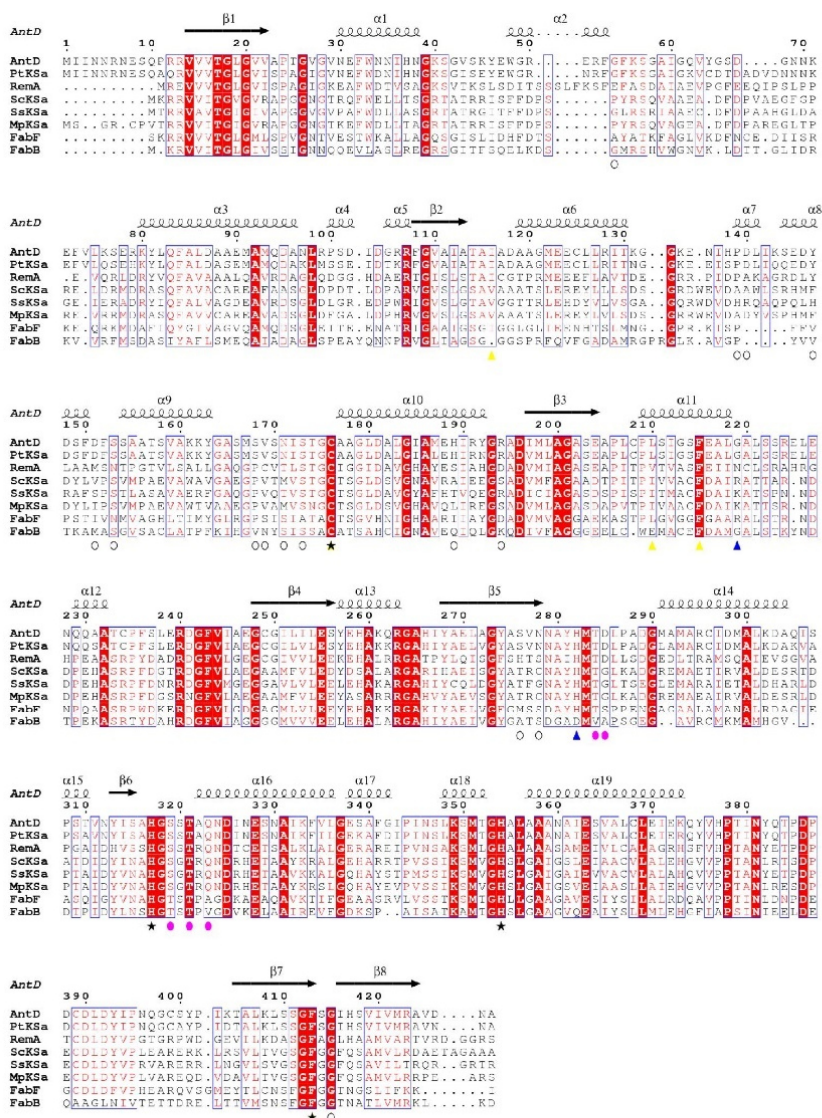


Fig. S12. Sequence alignment of AntD and selected KS α s.

PtKS α (*Photorhabdus temperata*, WP_021325340), *RemA* (or *SrKS α* , *Streptomyces resistomycificus*, resistomycin biosynthesis, WP_030043016.1), *ScKS α* (*Streptomyces coelicolor*, PDB ID: 1TQY), *SsKS α* (*Streptomyces* sp., SEF09529.1), *MpKS α* (*Micromonospora purpureochromogenes*, SCF16573.1), *FabF* (or *EcKS I*, PDB ID: 1MLA), *FabB* (or *EcKS II*, 1KAS). Active site residues are indicated with asterisks, KS β (AntE) contacts with black circles, ACP (AntF) contacts with blue triangles, phosphopantetheine arm contacts with magenta circles, and polyketide contacts with yellow triangles.

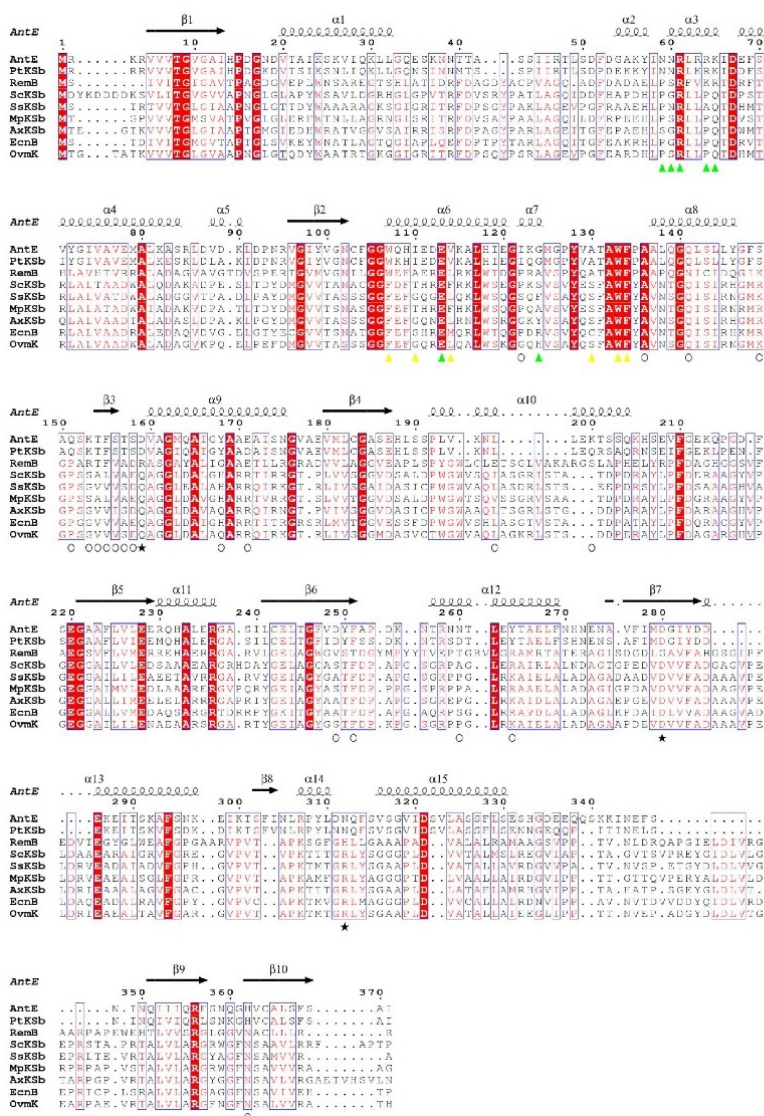


Fig. S13. Sequence alignment of AntE and selected KS β s.

PtKS β (*Photorhabdus temperata*, WP_021325341.1), *RemB* (or *SrKS β* , *Streptomyces resistomycificus*, resistomycin biosynthesis, ADK37_07640), *ScKS β* (*Streptomyces coelicolor*, 1TQY), *SsKS β* (*Streptomyces* sp., WP_093502894.1), *MpKS β* (*Micromonospora purpureochromogenes*, WP_088961753.1), *AxKS β* (*Amycolatopsis xylanica*, WP_091297626.1), *EcnB* (AAF81729.1), *OvnK* (WP_030791003.1). KS β (AntE) contacts are indicated with black circles, ACP (AntF) contacts with green triangles, polyketide contacts or gate residues with yellow triangles, and gates with blue squares. Notably, only the first two thirds of AntE show homology to other KS β enzymes.

Table S1.
Plasmids and primers used in this work.

Plasmid	Primer sequence 5' → 3'
ZQ1/pJET1.2/blunt-AntDEF	ataatagaatcgaatgataaataaacagaaatg atagccgacgtaattttatcgtttaaac
ZQ6/pSUAntB	tactaacatagacgatalttctttatcatctgatt gtcaatgacgacttactctctttgtccttataatctcg
ZQ9/pJET1.2/blunt-AntDEFG	actgacagataacgatttc acattcctggccattat
ZQ10/pSUAntABC	tagttccatagaaatagcctttattaccgg cttgaagtcgaccaltgggtalatgaaatctcttt
ZQ11	tagttccatagaaatagcctttattaccgg ttctcagtcgacttcccaaaatcacaatctatagg
ZQ12	actgacagataacgatttc ctattgggttlatlittattatcatct
ZQ13/pJET1.2/blunt-AntDEFGHI	actgacagataacgatttc ttaccatcgcatgatatt
ZQ16	tactaacatagacgatalttctttatcatctgatt cttgaagtcgaccaltgggtatagaaatctcttt
ZQ32/pCOLA-His ₆ -AntD_AntF	agtagacatagcgtaaaagagttgctgttacc tttagaggtaacctaaatagctgaaaaactcaacgc laatlagaattccgtgataaataaacagaaatgaatctcaacc
ZQ32_AntD_C176A	caccaactgcagccggtaacgacaactcttttacg gccctgccgacgtcgggttgatggttactgac gtcagtaacatatacaaccggagctgaggcggc
ZQ37/pCOLA-His ₆ -Plu2834	acatgtaattccatgctgaattgcaaatggtatt aatctcctgcagattgccclcaattatttcca
ZQ46/pCOLA-His ₆ -AntF ^{apo}	aaaaatgaattccatgaataatcatccagaagtaaa ttaattctgcagcagtttaattttatcgtttaaact
ZQ47/pCOLA-His ₆ -AntG	aaacaggaattccatgaactaatctctatgtttaca tctacactgcagattcattatgattcctcaatg
ZQ48/pCOLA-His ₆ -AntB	tactaagaattccatggacgatattttatcat tactaagaattccatggacgatattttatcat
ZQ56/pACYC-AntDEFG	laatlagaattccgtgataaataaacagaaatgaatctcaacc tctacactgcagattcattatgattcctcaatg
ZQ56_AntF_D37A	gtcagc taattccgtagaagccatataatgcatcagc gctgatccctatgataaggctctacggaattagctgac
ZQ56_AntF_D43A	ctcaatctctttgccaaggcagctaatccgtagaatc gattctacggaattagctgcttggcaaaagagattgag
ZQ56_AntF_R61A	aaattcctcaattttacgtttaaagcgttaffgataatctcgcgccatcg cgalggcgcgaaatatacaataacgcttaagacgtaaaatgatgaattt
ZQ56_AntD_S319A_T321A	gattcgtaatatcattttgcgctgcagaagccatgagcactaataataattac gtaaattatattagctctatggcgcctctcagcgcgaaaatgataataacgaatc
ZQ65/pACYC-His ₆ -AntF_AntDEFG	laattaaagatctagtgataaataaacagaaatgaatct tctacaggtacattcattatgattcctcaatg aaaaatgaattccatgaataatcatccagaagtaaa
ZQ76	tttaattctgcagcagtttaattttatcgtttaaactt tagttcgaattccatgaaatagcctttattaccgg Atgacactgcagttattataatattgcgaccactc
ZQ78/pCDFDuetSfp	agccgacatagaaatgacggaattt ccgaatggtacctatataaagctctctgtacg
ZQ81/pACYCDuet-His ₆ -RemC_AntDE	aaaaatgaattccatgaataatcatccagaagtaaa ttaattctgcagcagtttaattttatcgtttaaactt
ZQ82/pACYCDuet-His ₆ -ActI-ORF3_AntDE	aaaaatgaattccatgaataatcatccagaagtaaa ttaattctgcagcagtttaattttatcgtttaaactt

Table S2.

Generated plasmids. * genes with selected mutation. ** codon optimized for *E. coli* expression.

Plasmid	Genotype
pJET1.2/blunt	pMB1ori, Ap ^r
pSU18	P15A ori, Cm ^r , <i>lacZ</i> promoter
pACYC Duet-1	P15A ori, Cm ^r , T7 <i>lac</i> promoter
pCOLA Duet-1	ColA ori, Km ^r , T7 <i>lac</i> promoter
pCDF Duet-1	CDF ori, Sm ^r , T7 <i>lac</i> promoter
pDS132	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i>
ZQ1	pMB1 ori, (Apr), <i>antDEF</i> with native promoter
ZQ6	P15A ori, Cm ^r , <i>antB</i> , <i>lacZ</i> promoter
ZQ9	pMB1 ori, (Ap ^r), <i>antDEFG</i> with native promoter
ZQ10	P15A ori, Cm ^r , <i>antABC</i> , <i>lacZ</i> promoter
ZQ11	P15A ori, Cm ^r , <i>antAB</i> , <i>lacZ</i> promoter
ZQ12	pMB1ori, (Ap ^r), <i>antDEFGH</i> with native promoter
ZQ13	pMB1ori, (Ap ^r), <i>antDEFGHI</i> with native promoter
ZQ16	P15A ori, Cm ^r , <i>antBC</i> , <i>lacZ</i> promoter
ZQ32	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antD</i> , <i>antE</i>
ZQ32_AntD_C176A	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antD*</i> , <i>antE</i>
ZQ37	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>plu2834</i>
ZQ46	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antF</i>
ZQ47	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antG</i>
ZQ48	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antB</i>
ZQ48	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antB</i>
ZQ56	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antDEFG</i>
ZQ56_AntF_D37A	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antDEF*G</i>
ZQ56_AntF_D43A	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antDEF*G</i>
ZQ56_AntE_R61A	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antDE*FG</i>
ZQ56_AntD_S319A_T321A	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antD*EFG</i>
ZQ62	P15A ori, Cm ^r , T7 <i>lac</i> promoter, <i>antDEFGH</i>
ZQ65	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antF</i> , <i>antDEFG</i>
ZQ76	CDF ori, SM ^r , T7 <i>lac</i> promoter, <i>antABC</i>
ZQ78	CDF ori, Smr, T7 <i>lac</i> promoter, <i>sfp</i>
ZQ81	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antDE</i> , <i>remC**</i>
ZQ82	P15A ori, Cmr, T7 <i>lac</i> promoter, <i>antDE</i> , <i>actI-ORF3**</i>
pDS_plu4186	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i> , partial <i>plu4186</i>
pDS_4188	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i> , partial <i>plu4188</i>
pDS_plu4192	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i> , partial <i>plu4192</i>

Table S3.Generated *E. coli* strains (ES).

Strain	Genotype
ES5	DH10B::ZQ1 (pJET1.2/blunt-AntDEF), ZQ6 (pSUAntB)
ES7	DH10B::ZQ9 (pJET1.2/blunt-AntDEFG)
ES8	DH10B::ZQ9 (pJET1.2/blunt-AntDEFG), ZQ6 (pSUAntB)
ES10	ES10 DH10B::ZQ13 (pJET1.2/blunt-AntDEFGHI), ZQ16 (pSUAntABC)
ES12	ES12 DH10B::ZQ9 (pJET1.2/blunt-AntDEFG), ZQ11 (pSUAntAB)
ES13	ES13 DH10B::ZQ9 (pJET1.2/blunt-AntDEFG), ZQ10 (pSUAntABC)
ES14	ES14 DH10B::ZQ12 (pJET1.2/blunt-AntDEFGH), ZQ11 (pSUAntAB)
ES15	ES15 DH10B::ZQ13 (pJET1.2/blunt-AntDEFGHI), ZQ11 (pSUAntAB)
ES17	DH10B::ZQ13 (pJET1.2/blunt-AntDEFGHI), ZQ10 (pSUAntABC)
ES18	BL21 (DE3)::ZQ37 (pCOLA-His ₆ -Plu2834)
ES19	BL21 (DE3)::ZQ47 (pCOLA-His ₆ -AntG)
ES20	BL21 (DE3)::ZQ46 (pCOLA-His ₆ -AntF ^{apo})
ES21	BL21 (DE3)::ZQ32 (pCOLA-His ₆ -AntD_AntE)
ES23	BL21 (DE3)::ZQ65 (pACYC-His ₆ -AntF_AntDEFG), ZQ48 (pCOLA-His ₆ -AntB)
ES24	BL21 (DE3)::ZQ56 (pACYC-AntDEFG), ZQ48 (pCOLA-His ₆ -AntB)
ES26	BL21 (DE3)::ZQ56 (pACYC-AntDEFG)
ES67	BL21 (DE3)::ZQ81 (pACYCDuet-His ₆ -RemC_AntDE), ZQ78 (pCDFDuetSfp)
ES68	BL21 (DE3)::ZQ82 (pACYCDuet-His ₆ -ActI-ORF3_AntDE), ZQ78 (pCDFDuetSfp)
ES100	BL21 (DE3)::ZQ32_AntD_C176A
ES101	BL21 (DE3)::ZQ56_AntF_D37A, ZQ48 (pCOLA-His ₆ -AntB)
ES102	BL21 (DE3)::ZQ56_AntF_D43A, ZQ48 (pCOLA-His ₆ -AntB)
ES103	BL21 (DE3)::ZQ56_AntE_R61A, ZQ48 (pCOLA-His ₆ -AntB)
ES104	BL21 (DE3)::ZQ56_AntD_S319A_T321A, ZQ48 (pCOLA-His ₆ -AntB)

Table S4.

E. coli strains (ES) with coexpressed proteins and major polyketide products in agreement with Zhou *et al.* (3).

Strain	Proteins	Major products
ES8	AntBDEFG	SEK4, SEK4b
ES10	AntBCDEFGHI	SEK4, SEK4b
ES12	AntABDEFG	mutactin
ES13	AntABCDEF	mutactin
ES14	AntABDEFGH	SEK34b
ES15	AntABDEFGHI	SEK34b
ES17	AntABCDEF	AQ-256
ES11	AntBCDEFG	SEK4, SEK4b
ES9	AntBCDEFGH	SEK4, SEK4b

Table S5.

Summary of components used for *in vitro* assays including HPLC/MS results. The *Minimal PKS System* consists of *P/MCAT*, AntF^{holo}, and AntDE. Mal-CoA acts as the sole substrate.

	Assay	a	b	c	d	e	f	g	h	i
monomer	acetyl-CoA (4 mM)	-	✓	✓	-	-	-	-	-	-
	Mal-CoA (4 mM)	-	✓	✓	✓	✓	✓	✓	✓	-
	His ₆ -AntG	-	-	✓	✓	-	-	-	-	-
	His ₆ - AntF ^{apo}	-	✓	-	-	-	-	-	-	-
Minimal PKS System	MCAT (His ₆ - <i>P/MCAT</i> ; 0.1 μM)	-	✓	✓	✓	✓	✓	-	✓	✓
	ACP (His ₆ -AntF ^{holo} ; 1 μM)	✓	-	✓	✓	✓	-	✓	✓	✓
	KS_α-KS_β (His ₆ -AntDE; 0.1 μM)	-	✓	✓	✓	✓	✓	✓	-	✓
	Tris-HCl buffer (50 mM, pH = 7.5)	✓	✓	✓	✓	✓	✓	✓	✓	✓
	total volume / μl	50	50	50	50	50	50	50	50	50
polyketide	SEK4/SEK4b	-	-	✓	✓	✓	-	-	-	-

Table S6.

E. coli strains (ES) with coexpressed proteins and polyketide products.

Strain	Proteins	Octaketide	Conclusion
ES5	AntBDEF	no	AntG essential to form AntF ^{holo}
ES7	AntDEFG	no	AntB essential to form AntF ^{holo}
ES8	AntBDEFG	SEK4/SEK4b	AntBG essential to form AntF ^{holo}
ES67	AntDE + RemC/Sfp (= ACP ^{holo})	no	AntF essential to form octaketide
ES68	AntDE + ActI-ORF3/Sfp (= ACP ^{holo})	no	AntF essential to form octaketide
ES17	AntABCDEFGHI	AQ-256	

Table S7.

X-ray data collection and refinement statistics.

	AntF^{apo}	AntF^{holo}	PIMCAT:AntF^{holo}
Crystal parameters			
Space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2 ₁
Cell constants	a=58.7 Å; c=150.0 Å	a=59.1 Å; c=151.2 Å	a=79.6 Å; b=84.6 Å c=158.1 Å
Subunits / AU ^a	2 2 AntF ^{apo}	2 2 AntF ^{holo}	3 1 PIMCAT:AntF ^{holo} 1 PIMCAT
Disordered regions (AntX ^Y : X=protein, Y=chain)	AntF ^A : 1-3, 82 AntF ^B : 1-3, 81-82	AntF ^A : 1, 81-82 AntF ^B : 1-2, 80-82	AntF ^A : 1-2, 81-82 PIMCAT ^B : - PIMCAT ^C : 1
Data collection			
Beam line	X06SA, SLS	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0	1.0
Resolution range (Å) ^b	30-1.85 (1.95-1.85)	30-2.4 (2.5-2.4)	30-3.3 (3.4-3.3)
No. observations	166229	58471	77604
No. unique reflections ^c	23127	11070	16593
Completeness (%) ^b	99.2 (99.8)	99.0 (99.9)	99.3 (99.5)
R _{merge} (%) ^{b, d}	4.3 (55.6)	5.5 (49.4)	13.3 (47.4)
I/σ (I) ^b	24.5 (3.6)	18.2 (3.6)	9.6 (3.1)
Refinement (REFMAC5)			
Resolution range (Å)	15-1.85	15-2.4	15-3.3
No. refl. working set	21891	10411	15708
No. refl. test set	1153	548	827
No. non hydrogen	1359	1294	5286
No. of ligand atoms	-	42	21
Solvent	132	30	-
R _{work} /R _{free} (%) ^e	15.8 / 19.6	18.8 / 22.7	22.9 / 26.8
r.m.s.d. bond (Å) / (°) ^f	0.006 / 1.042	0.006 / 1.012	0.007 / 0.98
Average B-factor (Å ²)	37.2	60.2	74.7
Ramachandran Plot (%) ^g	100 / 0 / 0	98.6 / 1.4 / 0	98.1 / 1.8 / 0.1
PDB accession code	6SM4	6SM6	6SMD

^[a] Asymmetric unit^[b] The values in parentheses for resolution range, completeness, R_{merge} and I/σ (I) correspond to the highest resolution shell^[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections^[d] $R_{\text{merge}}(I) = \frac{\sum_{\text{hkl}} \sum_j |I(\text{hkl})_j - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_j I(\text{hkl})_j}$, where $I(\text{hkl})_j$ is the j^{th} measurement of the intensity of reflection hkl and $\langle I(\text{hkl}) \rangle$ is the average intensity^[e] $R = \frac{\sum_{\text{hkl}} (|F_{\text{obs}}| - |F_{\text{calc}}|)}{\sum_{\text{hkl}} |F_{\text{obs}}|}$, where R_{free} is calculated for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections^[f] Deviations from ideal bond lengths/angles^[g] Percentage of residues in favored/allowed/outlier region

Table S8.

Interaction between AntF^{holo} and PIMCAT.

AntF ^{holo}		PIMCAT		
Residue	Atom	Residue	Atom/ interaction	distance (Å)
D37	O ^{δ2}	R287	N ⁿ	2.6
D37	O ^{δ1}	R287	N ⁿ	3.1
T39	O ^γ	R287	N ⁿ	2.3
A42	C ^β	L192	VdW	4.2
A42	C=O	R190	N ⁿ	2.6
D43	O ^δ	K189	N ^ζ	2.6
E49	O ^ε	R190	N ⁿ	2.8
K58	N ^ε	G12	C=O	3.3
F61	phenyl	R190	Cat-π	3.6
PPant	Thiol SH	S92	O ^γ	2.9
PPant	Thiol SH	M132	S ^δ	4.4
PPant	cysteamine C ^β	Q11	O ^ε	4.4
PPant	cysteamine C ^α	H91	C ^ε	3.7
PPant	cysteamine NH	Q11	C=O	3.4
PPant	β-alanine C=O	N162	N ^δ	2.6
PPant	β-alanine C ^β	L194	VdW	4.2
PPant	β-alanine C ^β	V280	VdW	4.3
PPant	β-alanine C ^α	V134	VdW	3.5
PPant	β-alanine C ^α	L194	VdW	4.5
PPant	β-alanine NH	Q166	O ^ε	2.7
PPant	pantoic acid C=O	K58	N ^ε	3.0
PPant	pantoic acid OH	Q11	O ^ε	3.3
PPant	pantoic acid Me ¹	V134	VdW	4.5
PPant	pantoic acid Me ¹	L192	VdW	4.3
PPant	pantoic acid Me ¹	I136	VdW	4.7
PPant	pantoic acid Me ²	K58	VdW	3.4

Table S9.

X-ray data collection and refinement statistics

	AntDE:F ^{apo}	AntDE:F ^{holo}
Crystal parameters		
Space group	P4 ₁	P2 ₁
Cell constants	a=107.7 Å; c=201.3 Å	a=108.3 Å; b=135.9 Å; c=148.6 Å; β=97.1°
Subunits / AU ^a	7 1 AntDE:F ^{apo} 2 AntDE	12 4 AntDE:F ^{holo}
Disordered regions (AntX ^Y : X=protein, Y=chain)	AntF ^A : 1-5, 77-82 AntD ^B : 1-9, 65-73, 428 AntE ^C : 33-43, 333-346 AntD ^D : 1-9, 65-73, 428 AntE ^E : 35-39, 333-348 AntD ^F : 1-9, 65-74, 428 AntE ^G : 36-40, 333-345	AntF ^A : 1-4, 77-82 AntD ^B : 65-73 AntE ^C : 34-43, 332-344 AntF ^D : 1-4, 77-82 AntD ^E : 65-73 AntE ^F : 34-43, 332-344 AntF ^G : 1-4, 77-82 AntD ^H : 65-78 AntE ^I : 34-44, 332-344 AntF ^J : 1-4, 77-82 AntD ^K : 8-10, 65-79 AntE ^L : 34-43, 332-344
Data collection		
Beam line	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	1.0
Resolution range (Å) ^b	30-2.7 (2.8-2.7)	30-2.9 (3.0-2.9)
No. observations	180954	285455
No. unique reflections ^c	61596	92717
Completeness (%) ^b	98.2 (98.8)	97.9 (98.4)
R _{merge} (%) ^{b, d}	12.3 (59.0)	9.7 (54.3)
I/σ (I) ^b	8.8 (2.3)	11.4 (2.4)
Refinement (REFMAC5)		
Resolution range (Å)	15-2.7	15-2.9
No. refl. working set	58164	87418
No. refl. test set	3061	4600
No. non hydrogen	18065	25689
No. of ligand atoms	63	192
Solvent	151	81
R _{work} /R _{free} (%) ^e	23.7 / 26.8	21.9 / 25.5
r.m.s.d. bond (Å) / (°) ^f	0.007 / 1.008	0.007 / 1.018
Average B-factor (Å ²)	45.5	61.8
Ramachandran Plot (%) ^g	97.0 / 2.9 / 0.1	97.1 / 2.8 / 0.1
PDB accession code	6SMO	6SMP

^[a] Asymmetric unit^[b] The values in parentheses for resolution range, completeness, R_{merge} and I/σ (I) correspond to the highest resolution shell^[c] Data reduction was carried out with XDS and from a single crystal. Friedel pairs were treated as identical reflections^[d] $R_{\text{merge}}(I) = \frac{\sum_{\text{hkl}} \sum_j |I(\text{hkl})_j - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_j I(\text{hkl})_j}$, where $I(\text{hkl})_j$ is the j^{th} measurement of the intensity of reflection hkl and $\langle I(\text{hkl}) \rangle$ is the average intensity^[e] $R = \frac{\sum_{\text{hkl}} (|F_{\text{obs}}| - |F_{\text{calc}}|)}{\sum_{\text{hkl}} |F_{\text{obs}}|}$, where R_{free} is calculated for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections^[f] Deviations from ideal bond lengths/angles^[g] Percentage of residues in favored/allowed/outlier region

Table S10.

List of matched b- and y-ESI-fragments identifying a bound hexaketide at Cys176 of AntD (marked with an asterisk). MS-analyses obtained from tryptic digested AntDE. MS² analysis resolved the desulfurization of the modified Cys176 to dehydroalanine.

Sequence	#	b _{calculated}	b _{detected}	Error [Da]	b _{calculated}	b _{detected}	Error [Da]
		Hexaketide			Desulfurization		
Y	1	163.063	n.d.				
G	2	221.092	221.092	0.000			
A	3	292.129	292.130	0.001			
S	4	379.161	379.1614	0.000			
M	5	510.202	510.203	0.001			
S	6	597.234	597.236	0.002			
V	7	696.302	696.305	0.002			
S	8	783.334	n.d.				
N	9	897.377	897.383	0.006			
I	10	1010.461	1010.468	0.007			
S	11	1097.493	1097.495	0.002			
T	12	1198.541	n.d.				
G	13	1255.562	1255.578	0.016			
C*	14	1610.635	n.d.		1324.584	1324.586	0.003
A	15	1681.672	n.d.				
A	16	1752.709	1752.870	0.161			
G	17	1809.731	n.d.				
L	18	1922.815	n.d.				
D	19	2037.842	n.d.				
A	20	2108.879	n.d.				
L	21	2221.963	n.d.				
G	22	2278.984	n.d.				
I	23	2392.068	n.d.				
A	24	2463.105	n.d.				
M	25	2594.146	n.d.				
E	26	2723.188	n.d.				
H	27	2860.247	n.d.				
I	28	2973.331	n.d.				
R	29	3129.433	n.d.				

Detection of the hexaketide as a bound thioester with *m/z* 286.1993 (error: 0.0093 Da)

Sequence	# (+1)	y _{calculated}	y _{detected}	Error [Da]	y _{calculated}	y _{detected}	Error [Da]
		Hexaketide			Desulfurization		
Y	29	3147.443	n.d.				
G	28	2984.380	n.d.				
A	27	2927.358	n.d.				
S	26	2856.321	n.d.				
M	25	2769.289	n.d.				
S	24	2638.249	n.d.				
V	23	2551.217	n.d.				
S	22	2452.148	n.d.				
N	21	2365.116	n.d.				
I	20	2251.073	n.d.				
S	19	2137.989	2138.149	0.160			
T	18	2050.957	2051.096	0.139			
G	17	1949.910	1950.050	0.140			
C*	16	1892.888	1893.046	0.158	1606.837	1606.858	0.021
A	15	1537.816	1537.811	0.005			
A	14	1466.778	1466.772	0.007			
G	13	1395.741	1395.740	0.001			
L	12	1338.720	1338.719	0.001			
D	11	1225.636	1225.637	0.001			
A	10	1110.609	1110.610	0.001			
L	9	1039.572	1039.573	0.001			
G	8	926.488	926.488	0.000			
I	7	869.466	869.465	0.001			
A	6	757.382	757.382	0.001			
M	5	685.345	685.345	0.000			
E	4	554.305	554.305	0.000			
H	3	425.262	425.263	0.001			
I	2	288.203	288.203	0.000			
R	1	175.119	n.d.				

n.d. = not detectable

Table S11.

List of matched b- and y-ESI-fragments of the AntD^{C176A}E mutant (marked with an asterisk).

Sequence	#	b _{calculated}	b _{detected}	Error [Da]
Y	1	163.063	n.d.	
G	2	221.092	221.092	0.000
A	3	292.129	292.130	0.001
S	4	379.161	379.163	0.002
M	5	510.202	510.259	0.057
S	6	597.234	597.288	0.055
V	7	696.302	696.307	0.005
S	8	783.334	n.d.	
N	9	897.377	897.408	0.030
I	10	1010.461	1010.462	0.001
S	11	1097.493	n.d.	
T	12	1198.541	n.d.	
G	13	1255.562	1255.548	0.015
A*	14	1326.599	1326.601	0.001
A	15	1397.637	1397.725	0.089
A	16	1468.674	1468.763	0.089
G	17	1525.695	n.d.	
L	18	1638.779	n.d.	
D	19	1753.806	n.d.	
A	20	1824.843	n.d.	
L	21	1937.927	n.d.	
G	22	1994.949	n.d.	
I	23	2108.033	n.d.	
A	24	2179.070	n.d.	
M	25	2310.110	n.d.	
E	26	2439.153	n.d.	
H	27	2576.212	n.d.	
I	28	2689.296	n.d.	
R	29	2845.397	n.d.	

Sequence	# (+1)	y _{calculated}	y _{detected}	Error [Da]
Y	29	2863.408	n.d.	
G	28	2700.344	n.d.	
A	27	2643.323	n.d.	
S	26	2572.286	n.d.	
M	25	2485.254	n.d.	
S	24	2354.213	n.d.	
V	23	2267.181	n.d.	
S	22	2168.113	n.d.	
N	21	2081.081	n.d.	
I	20	1967.038	n.d.	
S	19	1853.954	1853.944	0.010
T	18	1766.922	1766.918	0.004
G	17	1665.874	1665.883	0.009
A*	16	1608.853	n.d.	
A	15	1537.816	1537.809	0.007
A	14	1466.778	1466.787	0.009
G	13	1395.741	1395.742	0.000
L	12	1338.720	1338.712	0.008
D	11	1225.636	1225.638	0.002
A	10	1110.609	1110.607	0.002
L	9	1039.572	1039.572	0.000
G	8	926.488	926.488	0.001
I	7	869.466	869.467	0.001
A	6	756.382	756.381	0.001
M	5	685.345	685.344	0.001
E	4	554.305	554.307	0.002
H	3	425.262	425.263	0.001
I	2	288.203	288.203	0.000
R	1	175.119	n.d.	

n.d. = not detectable

Table S12.

Interaction between AntF^{holo} and AntDE. Asymmetric unit contains 4 ternary complexes. Interactions are only shown for chains A, B and C.

AntF		AntDE ^a		distance (Å)
Residue	Atom	Residue	Atom	
F16	phenyl	R64 ^a	Cat- π	3.3
N18	C $^{\beta}$	R61 ^a	VdW	3.9
N18	N $^{\delta}$	N59 ^a	O $^{\delta}$	2.8
N18	O $^{\delta}$	N60 ^a	N $^{\delta}$	3.4
Y34	C=O	R61 ^a	N $^{\eta 1}$	2.8
Y34	C=O	R61 ^a	N $^{\eta 2}$	2.9
D37	O $^{\delta}$	H282	N $^{\epsilon}$	3.0
T39	C $^{\gamma}$	L218	VdW	3.8
T39	C $^{\gamma}$	P127 ^a	VdW	3.9
E40	O $^{\epsilon 1}$	R61 ^a	N $^{\epsilon}$	3.0
E40	O $^{\epsilon 2}$	R61 ^a	N $^{\eta 1}$	3.1
E40	O $^{\epsilon 1}$	K65 ^a	N $^{\zeta}$	3.0
A42	C $^{\beta}$	P127 ^a	VdW	3.6
D43	C $^{\beta}$	P127 ^a	VdW	3.8
D43	O $^{\delta}$	R64 ^a	N $^{\eta}$	2.7
K46 ^b	N $^{\zeta}$	G124 ^a	C=O	3.3
K46 ^b	N $^{\zeta}$	E113 ^a	O $^{\epsilon}$	3.2
PPant	cysteamine SH	F413	π -interaction	3.5
PPant	β -alanine C=O	S319	O $^{\gamma}$	2.9
PPant	β -alanine C=O	T321	O $^{\gamma}$	2.8
PPant	β -alanine C $^{\beta}$	L286	VdW	4.3
PPant	β -alanine C $^{\alpha}$	T321	VdW	3.8
PPant	β -alanine NH	T284	C=O	2.9
PPant	pantoic acid OH	T284	C=O	3.0
PPant	pantoic acid OH	D285	O $^{\delta}$	2.6
PPant	pantoic acid Me ¹	L218	VdW	4.0
PPant	pantoic acid Me ²	A220	VdW	4.5

^[a] Interaction with AntE

^[b] K46 is poorly defined in the 2F_o-F_c electron density map.

Table S13.

Interaction between AntF^{F_{apo}} and AntDE. Asymmetric unit contains 1 ternary AntDE:F^{F_{apo}} (chains A, B, C) complex and two isolated AntDE (chains D, E; chains F, G) KS_α-KS_β-ketosynthase.

AntF ^{F_{apo}}		AntDE ^a		distance (Å)
Residue	Atom	Residue	Atom	
F16	phenyl	R64 ^a	Cat- π	4.1
N18	C ^{β}	R61 ^a	VdW	4.3
N18	N ^{δ}	N59 ^a	O ^{δ}	4.1
N18	O ^{δ}	N60 ^a	N ^{δ}	4.1
Y34	C=O	R61 ^a	N ^{η1}	3.0
Y34	C=O	R61 ^a	N ^{η2}	3.4
D37	O ^{δ}	H282	N ^{ϵ}	3.4
T39	C ^{γ}	L218	VdW	3.9
T39	C ^{γ}	P127 ^a	VdW	3.9
E40	O ^{ϵ1}	R61 ^a	N ^{ϵ}	2.8
E40	O ^{ϵ2}	R61 ^a	N ^{η1}	2.8
E40	O ^{ϵ1}	K65 ^a	N ^{ζ}	2.8
A42	C ^{β}	P127 ^a	VdW	4.2
D43	C ^{β}	P127 ^a	VdW	3.7
D43	O ^{δ}	R64 ^a	N ^{η}	2.6
K46 ^b	N ^{ζ}	G124 ^a	C=O	-
K46 ^b	N ^{ζ}	E113 ^a	O ^{ϵ}	-
K46 ^b	N ^{ζ}	M125 ^a	C=O	2.4
K46 ^b	N ^{ζ}	I122 ^a	C=O	2.3

^[a] Interaction with AntE

^[b] K46 is well defined in the 2F_o-F_c electron density map. However, K46N ^{ζ} coordinates two repulsive carbonyl oxygen atoms, which is likely caused by crystal packing.

References

1. E. Duchaud *et al.*, The genome sequence of the entomopathogenic bacterium *Photobacterium luminescens*. *Nat. Biotechnol.* **21**, 1307-1313 (2003).
2. N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj, M. Mann, Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319-324 (2014).
3. Q. Zhou *et al.*, Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photobacterium luminescens*. *Chem. Sci.* **10**, 6341-6349 (2019).
4. W. Kabsch, Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795-800 (1993).
5. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
6. A. Roujeinikova *et al.*, Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *J. Mol. Biol.* **365**, 135-145 (2007).
7. C. Oefner, H. Schulz, A. D'Arcy, G. E. Dale, Mapping the active site of *Escherichia coli* malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 613-618 (2006).
8. A. T. Keatinge-Clay, D. A. Maltby, K. F. Medzihradszky, C. Khosla, R. M. Stroud, An antibiotic factory caught in action. *Nat. Struct. Mol. Biol.* **11**, 888-893 (2004).
9. D. Turk, MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1342-1357 (2013).
10. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501 (2010).
11. G. N. Murshudov *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355-367 (2011).
12. G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171-1179 (2008).
13. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648-5652 (1993).
14. C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B Condens. Matter* **37**, 785-789 (1988).
15. A. Schäfer, C. Huber, R. Ahlrichs, Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* **100**, 5829-5835 (1994).
16. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 1-29 (2010).
17. A. Klamt, G. Schuurmann, COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans.*, **2**, 799-805.
18. J. Huang, A. D. MacKerell, Jr., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135-2145 (2013).

19. R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **162**, 165-169 (1989).
20. B. R. Brooks *et al.*, CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545-1614 (2009).
21. S. Riahi, C. N. Rowley, The CHARMM-TURBOMOLE interface for efficient and accurate QM/MM molecular dynamics, free energies, and excited state properties. *J. Comput. Chem.* **35**, 2076-2086 (2014).
22. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33-38, 27-38 (1996).
23. D. Fenyo *et al.*, MALDI sample preparation: the ultra thin layer method. *J. Vis. Exp.*, 192 (2007).
24. M. Ruhl *et al.*, Elucidation of chemical modifier reactivity towards peptides and proteins and the analysis of specific fragmentation by matrix-assisted laser desorption/ionization collision-induced dissociation tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **33**, 40-49 (2019).

12.5 Publication: Molecular Mechanism of Polyketide Shortening in Anthraquinone Biosynthesis of *Photorhabdus laumondii*

Declaration on the contribution of the authors/Erklärung zu den Autorenanteilen

Status: **printed**

Chem. Sci., 2019, 10 (25), 6341-6349, DOI: 10.1039/C9SC00749K

Qiuqin Zhou (QZ), Alois Bräuer (AB) H el ene Adihou (HA), Maximilian Schmalhofer (MS), Patricia Saura (PS), Gina L. C. Grammbitter (GLG), Ville R. I. Kaila (VRK), Michael Groll (MG) and Helge B. Bode (HBB)

Was hat der/die Promovierende bzw. was haben die Co-Autoren/Autorinnen beigetragen?

(1) zu Entwicklung und Planung

HBB, QZ: 50%

MG, AB: 50%

(2) Durchf uhrung der einzelnen Untersuchungen und Experimente

Proteinreinigung Antl, Kristallisation/Kristallisationsexperimente Antl (AB) (35%)

Klonierung, molekularbiologische/biochemische/bioinformatische und HPLC-MS-basierte Untersuchungen QZ (45%), GG (5%)

Chemische Synthese 5% (HA)

Computational modeling VRK, PA (10%)

(3) Erstellung der Datensammlung und Abbildungen

Kristallisationsexperimente (AB, MG, MS) (40%)

molekularbiologische/biochemische/bioinformatische und HPLC-MS-basierte Untersuchungen QZ (45%), GG (5%)

Computational modeling VRK, PA (10%)

(4) Analyse und Interpretation der Daten

Kristallisationsexperimente (AB, MG, MS) (40%)

molekularbiologische/biochemische/bioinformatische und HPLC-MS-basierte Untersuchungen QZ (45%), GG (5%)

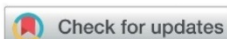
Computational modeling VRK, PA (10%)

(5) Verfassung des Manuskripts

HBB, QZ: 55%

MG, AB: 45%

date/place: _____

Cite this: *Chem. Sci.*, 2019, 10, 6341

All publication charges for this article have been paid for by the Royal Society of Chemistry

Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photorhabdus luminescens*†

Qiuqin Zhou,^a Alois Bräuer,^b H el ene Adihou,^a Maximilian Schmalhofer,^{id b} Patricia Saura,^{id b} Gina L. C. Grammbitter,^a Ville R. I. Kaila,^{id b} Michael Groll^{id *b} and Helge B. Bode^{id *a}

Anthraquinones, a widely distributed class of aromatic natural products, are produced by a type II polyketide synthase system in the Gram-negative bacterium *Photorhabdus luminescens*. Heterologous expression of the *antABCDEFGHI* anthraquinone biosynthetic gene cluster in *Escherichia coli* identified AntI as an unusual lyase, catalysing terminal polyketide shortening prior to formation of the third aromatic ring. Functional *in vitro* and *in vivo* analysis of AntI using X-ray crystallography, structure-based mutagenesis, and molecular simulations revealed that AntI converts a defined octaketide to the tricyclic anthraquinone ring *via* retro-Claisen and Dieckmann reactions. Thus, AntI catalyses a so far unobserved multistep reaction in this PKS system.

Received 13th February 2019
Accepted 8th May 2019

DOI: 10.1039/c9sc00749k

rsc.li/chemical-science

Introduction

Polyketide natural products attract significant attention due to their remarkable biological and pharmacological activities. Their structural diversity derives from only a few building blocks assembled by different polyketide synthase systems (PKS):^{1,2} type I PKS represent elaborate multifunctional enzyme complexes responsible for the synthesis of macrolides including erythromycin and candicin,³ whereas type III PKS catalyse the formation of stilbenes, chalcones, and pyrones.⁴ On the other hand, type II PKS constitute discrete proteins that act in an iterative manner to synthesise polycyclic aromatic compounds such as anthraquinones (AQs), tetracyclines and doxorubicins.^{5,6} First, a specific acyl carrier protein (ACP) is loaded with an α -carboxylated precursor, and the activated unit is then transferred onto the corresponding ketosynthase. Next, iterative elongation cycles with malonyl-coenzyme A building blocks form the octaketide framework.¹ Subsequently, ketoreductases, cyclases and aromatases convert the unstable and highly reactive octaketide into the corresponding target compounds. Notably, the variability of the natural products can be further increased by tailoring enzymes, which are not part of

the gene cluster and therefore can be difficult to identify. Yet, these catalysts are specific and perform only minor modifications such as alkylations, hydroxylations, or oxidations, while preserving the octaketide skeleton.

Type I and type III PKS subclasses are well-known from bacteria, fungi, and plants.¹ In contrast, type II PKS have only been characterised in Gram-positive *Streptomyces* and other actinomycetes.⁵ We identified a biosynthetic gene cluster encoding a type II PKS in the Gram-negative entomopathogenic bacterium *Photorhabdus luminescens*, an enzyme system that enables the biosynthesis of AQs.⁷ Notably, along with the antimicrobial aurachins⁸ and pyxidicyclines⁹ from myxobacteria, AQs represent another fascinating example of the type II PKS machinery. Here, we have introduced the anthraquinone gene cluster *antABCDEFGHI* from *P. luminescens* in *E. coli*. Surprisingly, the formed AQ-256 product exhibits a heptaketide framework, which is unusual since polyphenolic polyketides derived from bacterial type II PKS systems are generally octaketides.¹ Our present study reveals that AntI shortens the octaketide intermediate 1 to the final AQ heptaketide. Although this type of reaction is uncommon, a related mechanism has been described for Ayg1p¹⁰ from *Aspergillus fumigatus* and WdYg1p¹¹ from *Wangiella dermatitidis*. Both fungal proteins are part of PKS systems involved in the biosynthesis of melanin pigments.

We provide insights into the molecular mechanism of AntI *via* biochemical, crystallographic, computational, and functional characterisations, and show how this enzyme catalyses an elimination reaction, followed by cyclisation to the anthraquinone ring. Taken together, our combined findings elucidate an interesting type of polyketide shortening mechanism.

^aMolekulare Biotechnologie, Fachbereich Biowissenschaften, Buchmann Institute for Molecular Life Sciences (BMLS), Goethe Universit at Frankfurt, Max-von-Laue-Str. 15, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany. E-mail: h.bode@bio.uni-frankfurt.de

^bCenter for Integrated Protein Science Munich (CIPSM), Department of Chemistry, Technische Universit at M unchen, Lichtenbergstra e 4, 85748 Garching, Germany. E-mail: michael.groll@tum.de

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9sc00749k

Results and discussion

The gene cluster encoding the anthraquinone biosynthetic machinery (*antABCDEFGHI*) in *P. luminescens* was heterologously expressed in *E. coli* (strain EC^{AQ}). This approach allowed detection of the central metabolite 1,3,7-trihydroxyanthracene-9,10-dione (AQ-256), a tricyclic aromatic compound not found in wild-type *E. coli* (Fig. 1). Yet, methylated derivatives such as AQ-270a (one methyl group) and AQ-284a (two methyl groups) were identified in the natural producer *P. luminescens* (Fig. 2). As these molecules were not detected in our engineered EC^{AQ} strain, this suggests that the responsible methyltransferases are absent in the *ant* cluster, and therefore must be located elsewhere in the *Phototrhhabdus* genome. Notably, the EC^{AQ}-system is accompanied by several shunt products (SP) such as SP1, SP^{EC}1 and SP^{EC}2 (Fig. 2). One reason for this observation might be that the expression level of the individual genes in the heterologous host is not balanced. Nevertheless, all shunt products could be traced back to the common precursor molecule **1** harboring an octaketide skeleton as its basic structural element, which is covalently bound to the acyl carrier protein AntF *via* a phosphopantetheine (PPT) prosthetic group (see below). To elucidate the formation of these distinct compounds, we selectively removed individual genes from our EC^{AQ} operon. We focused on classical type II PKS enzymes such as the ketoreductase AntA (C9 carbonyl reduction), aromatase AntH (first cyclisation, C7/C12), and cyclase AntC (second cyclisation, C5/C14). The resulting intermediates were identified by UV-VIS spectroscopy and mass spectrometry (ESI Fig. 1–3, Tables 1–2†). Mutants constructed in *P. luminescens* (ESI Fig. 4†) and labelling experiments in *E. coli* (ESI Fig. 3†) allowed assignment of the reaction cascade based on known biosynthetic pathways^{1,2} (ESI Fig. 5†).

Most polyphenolic polyketides derived from bacterial type II PKS systems are octaketides.¹ In this respect, AQ-256 with its heptaketide framework is exceptional. Moreover, the local alignment search in the NCBI databank revealed that AntI (Uniprot: Q7MZT8) is predominantly found in *Phototrhhabdus*

species. Thus, AntI is a promising candidate for polyketide shortening, followed by a unique cyclisation of the third aromatic ring in AQ biosynthesis (Fig. 1). In order to confirm our hypothesis, we generated a deletion mutant of *antI* in *P. luminescens* (PL^{ΔantI}). As expected, the modified strain lost the ability to produce AQ-256 and its methylated derivatives (Fig. 2), and only the nitrogen-containing octaketide derivative SP1 was formed. Next, we removed *antI* from our engineered *E. coli* strain (EC^{AQΔantI}) and analysed the resulting product pattern. As expected, AQ-256 is no longer present. However, in contrast to the *Phototrhhabdus* Δ*antI* mutant, which produced only the SP1 metabolite, we now observed several compounds. *Via* detailed characterisation of the EC^{AQΔantI} lysate, we identified SP1–5, SP^{EC}1, and SP^{EC}2, all of which could unambiguously be shown to derive from molecule **1** (Fig. 2, ESI Fig. 5†). In summary, (i) the dihydropyridine moiety of SP1 results from **1** by assimilation of intracellular NH₃; (ii) SP4 is formed from SP1 after dehydration and decarboxylation; (iii) utahmycin A¹² (SP5) is an oxidation product of SP4; (iv) 3,8-dihydroxy-1-methyl-anthraquinone-2-carboxylic acid (DMAC, SP^{EC}1) is the oxidized product of an aldol-condensation of **1**; and (v) aloesaparin II (SP^{EC}2) results from the spontaneous decarboxylation of SP^{EC}1. The assignment of these shunt products strongly indicates that AntI orchestrates the cyclisation of the third ring system in the AQ biosynthesis. In this reaction, 4-(3-acetyl-5-hydroxy-4-oxo-1,4-dihydronaphthalen-2-yl)-3-oxobutanoic acid (**1**) bound to AntF *via* the PPT-prosthetic group possibly acts as the AntI substrate that is converted to 8-hydroxy-1,10-dihydroanthracene-2,3,9(4H)-trione (**2**, Fig. 1). We thus propose that AntI catalyses the last step in this reaction cascade, whereas formation of the aromatic quinone moiety of AQ-256 occurs in an enzyme-independent manner by keto-enol-tautomerism and spontaneous oxidation.

The rapid turnover of the precursor molecule **1** in combination with the resulting large variety of different shunt products achieved in our modified *E. coli* system, encouraged us to additionally validate octaketide intermediate **1** as the true AntI substrate. To this end, the biosynthesis of the well-known blue

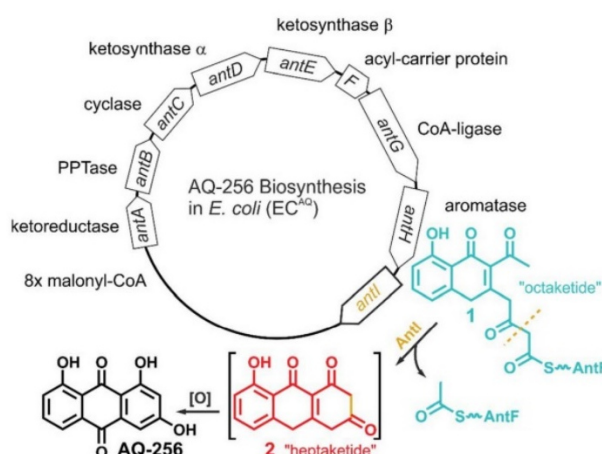


Fig. 1 Overview of the *ant* biosynthetic gene cluster from *Phototrhhabdus luminescens*. The AQ biosynthesis is highlighted with a focus on AntI catalyzed polyketide shortening.

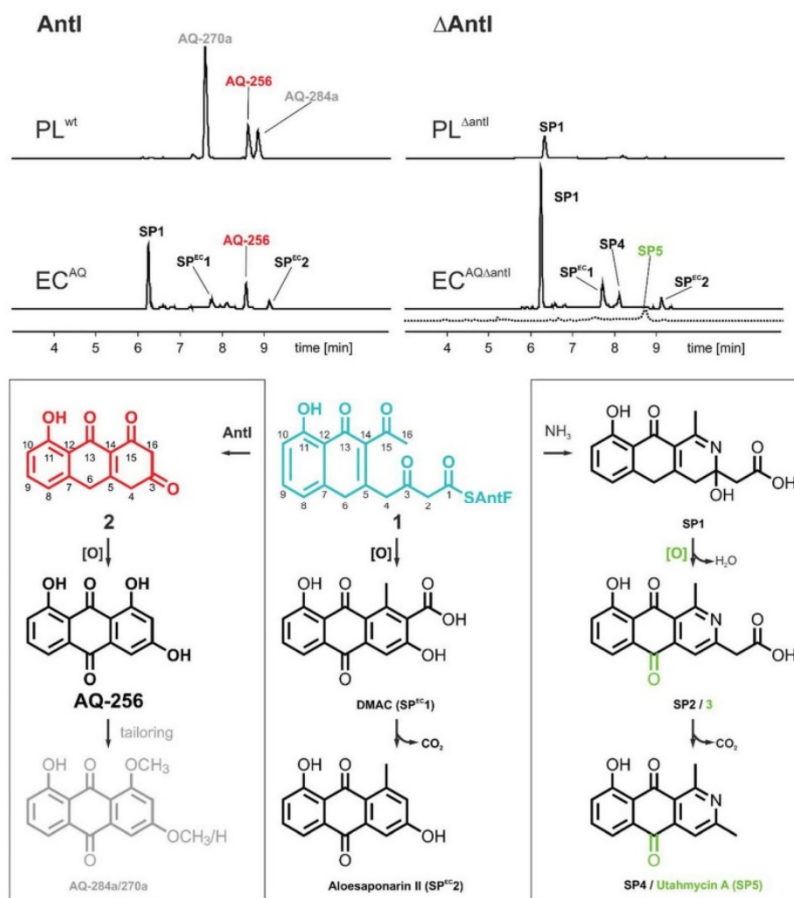


Fig. 2 HPLC/UV (420 nm, solid line) and HPLC/MS (EIC m/z 254 for SP5, dashed line) analysis for polyketide production in the presence and absence of AntI in *E. coli* (EC) and *Photorhabdus luminescens* (PL). Identified key metabolites (top) and suggested biosynthesis pathways to these compounds (bottom) are shown.

pigment actinorhodin from *Streptomyces coelicolor* appeared to be ideal, as it transforms **1** into the octaketide 4-dihydro-9-hydroxy-1-methyl-10-oxo-3-*H*-naphtho-[2,3-*c*]-pyran-3-(*S*)-acetate (**S-DNPA**). In this case, the stereospecific ketoreductase (Red1) acts as a tailoring enzyme by recognising the metabolite and reducing its 3 β -keto group to a secondary alcohol.¹³ This peculiarity led us to extend the EC^{AQΔantI} construct *via* a codon-optimized *actVI ORF1* gene, encoding Red1, which resulted in the EC^{AQΔantI+red1} strain. In this variant, AQ-256 was no longer produced, but we indeed detected the compound **S-DNPA** (Fig. 3a), confirming the activity of Red1 in our model system. These results further prove that octaketide **1** is formed, and then subsequently shortened and cyclised by AntI.

With the identification of AntI's substrate (**1**), we wanted to obtain a deeper mechanistic insight into the catalytic properties of this fascinating enzyme. AntI was therefore heterologously expressed in *E. coli*, purified, and crystallised with one subunit in the asymmetric unit. Native diffraction data were collected to 1.85 Å resolution, and the structure was phased by single-wavelength anomalous dispersion using selenomethionine

labelling (PDB ID: 6HXA, $R_{\text{free}} = 0.209$, ESI Table 3†). The structure reveals that AntI functions as a homodimer, with each subunit comprising two domains (Fig. 4a) that are connected by a short linker sequence (residues 130–134). The N-terminal domain (residues 7–130) comprises an α -helix bundle ($\alpha 1$ – $\alpha 5$) that orchestrates the arrangement of the two subunits in a back-to-back orientation with a contact area of 2055 Å². The C-terminal domain (residues 134–381) contains a twisted eight-stranded β -sheet ($\beta 1$ – $\beta 8$), which is flanked by 7 α -helices ($\alpha 6$ – $\alpha 12$) exhibiting an α/β -hydrolase fold (Fig. 4c).¹⁴

AntI forms a rod-shaped particle (Fig. 4b) with its N- and C-terminal domains enclosing a conspicuous intramolecular substrate binding cavity with approximate dimensions of 8 × 8 × 20 Å³. A DALI search¹⁵ demonstrated similarity of AntI with 2,6-dihydroxy-pseudo-oxynicotine hydrolase (DHPON; PDB ID: 2JBW, with a sequence identity of 23%, N-domain: Z-score = 10.6, RMSD = 2.5 Å; C-domain: Z-score = 30.5, RMSD = 1.9 Å),¹⁶ a C-C bond cleaving α/β -hydrolase from the Gram-positive bacterium *Arthrobacter nicotinovorans* that is involved in nicotine degradation (Fig. 4c). The structural comparison of AntI with DHPON

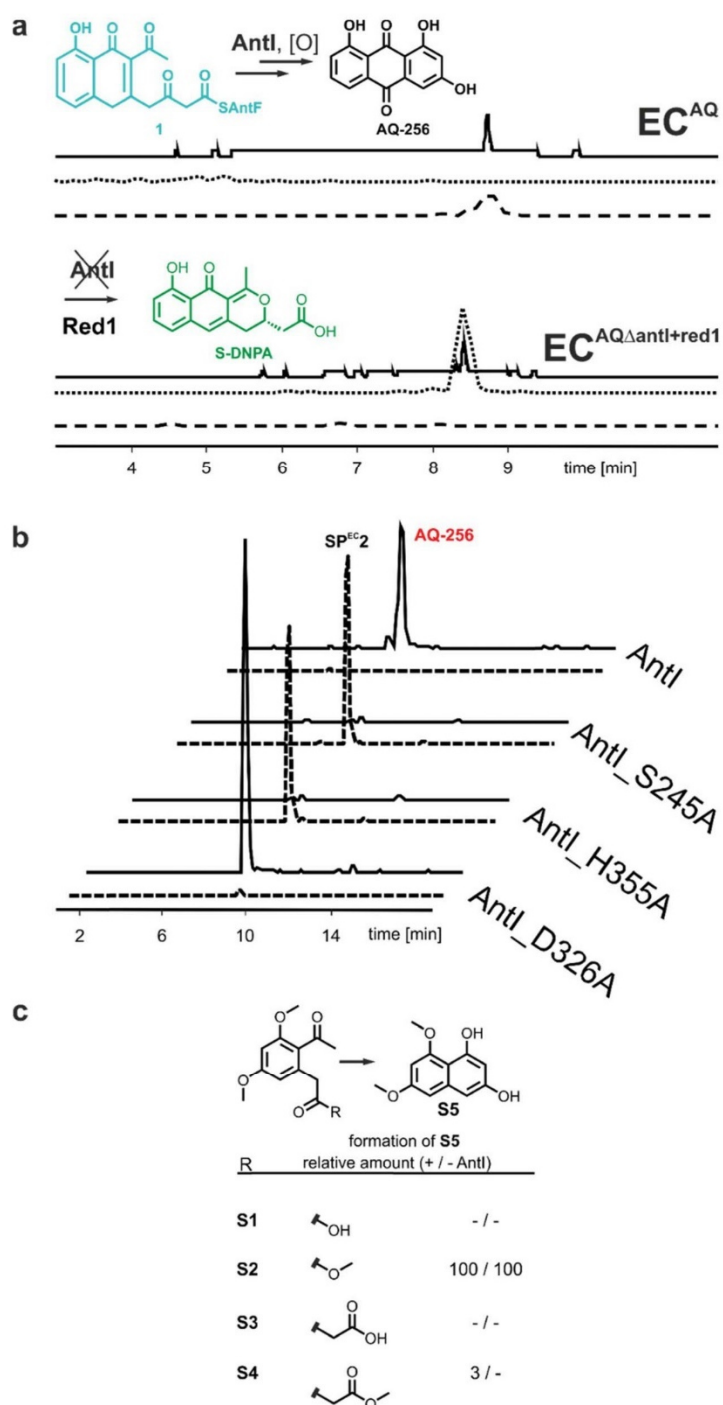


Fig. 3 (a) HPLC analyses of *E. coli* expressing *antABCDEF**GH*I and *antABCDEF**GH* + *actVI-ORF1* (encoding Red1). Besides UV traces (420 nm, solid line), EIC for AQ-256 (m/z 255 [M - H]⁻, dashed line) and S-DNPA (m/z 287 [M + H]⁺, dotted line) are shown. (b) In HPLC/MS analyses of *E. coli* expressing *antABCDEF**GH* with the AntI-mutants S245A or H355A, metabolite 1 is no longer detected, but compound SP^{EC}2 (m/z 253 [M - H]⁻, dashed line) is formed. Mutation of D326 to alanine did not alter catalysis. (c) Conversion of model compounds S1–S4 into S5.

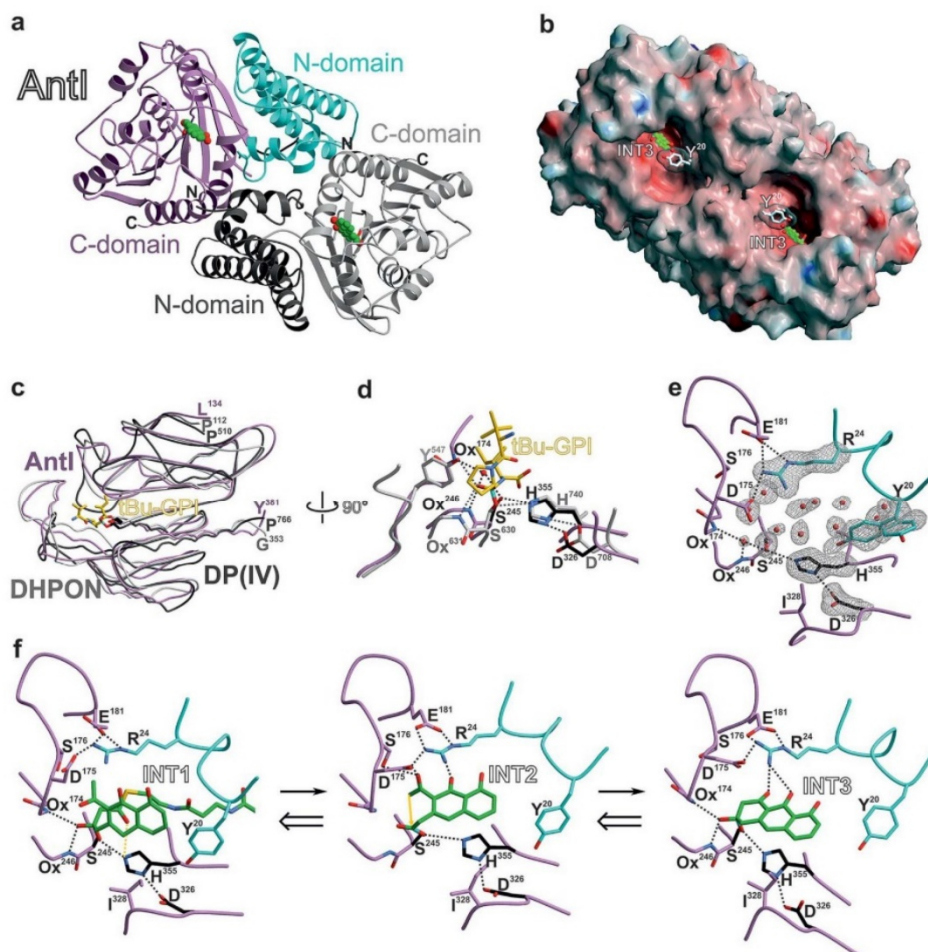


Fig. 4 Crystal structure of Antl. (a) The homodimer (ribbon plot) is depicted with a modelled intermediate (INT3, balls-and sticks, carbon atoms in green). N- and C-domains of subunit A are coloured in pink and cyan, respectively; subunit B is shown in grey. (b) Surface representation of Antl. The active site of the lyase is solvent exposed. Colours highlight negative and positive electrostatic potentials contoured from $-40k_B T/e$ (red) to $+40k_B T/e$ (blue). (c) Superposition of the α/β -hydrolase fold of Antl (residues 174–381, pink, PDB ID: 6HXA) with domains of 2,6-dihydroxy-pseudo-oxynicotine hydrolase (DHPON, light grey, PDB ID: 2JBW)¹⁶ and the human dipeptidyl peptidase DP(IV)-tBu-GPI complex (grey, PDB ID: 2AJB).¹⁷ (d) Structural overlay of Antl with the active site of DP(IV) (C-atoms in grey) covalently bound to the tBu-GPI surrogate. Notably, the tetrahedral coordination fits into the active site of Antl, and identifies its putative oxyanion hole (O_x), which is occupied by a water molecule. In Antl, the main chain atoms of Leu174NH and Phe246NH might form the oxyanion hole, while in DP(IV) Tyr631N and Tyr547OH take over this function. Hydrogen bonds are drawn as dashed lines. (e) Close up view of the active site of Antl showing the $2F_o - F_c$ electron density map (grey meshes contoured to 1.0σ) for the catalytic triad, Tyr20 (alternative conformation which might function as the gate keeper) and Arg24. The abundance of well-defined water molecules (red spheres) is a clear indication for the possible substrate binding site. (f) Proposed multistep reaction sequence of Antl including a retro-Claisen reaction and a Dieckmann condensation. The INT1, INT2, and INT3 intermediates were modelled based on the DP(IV)-tBu-GPI complex. The ligand-bound Antl structures remains dynamically stable during the MD simulations (see ESI†). The calculations were performed in the reverse direction from INT3 to INT1.

reveals the catalytic centre of a serine-type protease composed of Ser245, Asp326, and His355. In the active site, Ser245 is positioned at the tip of a sharp turn between sheet β_5 and helix α_8 , called the nucleophilic elbow,¹⁴ and points towards the helix dipole generated at the N-terminus of helix α_8 (Fig. 4d). The mutation of Ser245 or His355 to alanine confirmed the catalytic importance of these residues (Fig. 3b). In contrast, the Asp326Ala mutant remains active. Thus, Asp326 polarises His355, but

proton transfer from Ser245 to His355 is possible also without the aspartate residue. Inspection of the $2F_o - F_c$ -electron density map illustrates that the catalytic centre is enclosed by a cluster of bulk solvent molecules, the release of which might facilitate substrate binding (Fig. 4e). Consistent with the asymmetric hydrophilic rim of **1**, which is composed of three oxygen atoms, the specificity pocket exhibits an amphiphilic character. The hydrophobic half of this pocket comprises Phe27, Ile244, Ile326,

Val356, Leu358, and Ile361, whereas the polar part includes Arg24, Asp175, Ser176, Glu181, and Asp287. The central opening into the chamber is formed by Tyr20, Pro281, and Asp327. Interestingly, the sidechain of Tyr20 is observed in two alternative conformations (Fig. 4e), suggesting a regulatory role for this residue in allowing the substrate to access the active centre.

We aimed to obtain detailed insights into the reaction cascade of AntI. While the elimination reaction could clearly be assigned to AntI, the enzyme-dependency of the cyclisation resulting in the anthracene ring system still required experimental validation. We therefore synthesised surrogates **S1** and **S2** representing non-cyclised derivatives of the heptaketide intermediate (Fig. 3c, ESI Fig. 6†): **S1** harbours a free carboxylate moiety, whereas the methoxy group of **S2** mimics the sidechain of Ser245. Interestingly, **S1** is stable in solution and in the presence of AntI, whereas **S2** undergoes rapid ring closure to **S5** even without AntI (Fig. 3c). These findings indicate that an oxo-ester with AntI has to be formed, to enable intermolecular cyclization. Hence, AntI is involved in the multistep biosynthetic route by catalyzing polyketide shortening and subsequent ring formation.

The next goal was to confirm the retro-Claisen reaction for shortening the octaketide skeleton to a heptaketide. In contrast to the Dieckmann condensation, this requires a strong nucleophile, which might represent the rate-limiting step in AntI catalysis. To this end, we determine whether AntI works like a tailoring enzyme that uses intermediate 1 as its substrate after hydrolysis from the PPT arm (containing a free carboxylate), or whether 1 is still covalently bound to its ACP AntF. We synthesised 2-carboxymethyl-benzoic acid (surrogate **S3**), as well as its corresponding methyl ester **S4** which mimics metabolite 1. Both molecules remain stable in solution and in the presence of AntI. This observation is in contrast to the spontaneous enzyme-independent aromatisation of ester surrogate **S2** (see above). Thus, compared to the Dieckmann condensation, we assume that AntF is mandatory for the retro-Claisen reaction.

To obtain mechanistic insights into the structure and dynamics of these reaction intermediates, we performed molecular simulations based on our experimentally solved structure of AntI. As a starting point for these computations, we used a surrogate pose of the hydrolytic domain of the homologous porcine dipeptidyl peptidase IV (DP(IV), residues 511–766, PDB ID: 2AJB, RMSD of 145 C^α-atoms = 2.7 Å, sequence identity 17%). DP(IV) was solved with the tripeptide *tert*-butyl-Gly-Pro-Ile (*t*Bu-GPI), a surrogate that covalently binds in a sp³-configuration to the catalytic Ser630O^γ by forming an orthoester.¹⁷ Hereby, the alkoxide moiety of the inhibitor is negatively charged due to its strong interaction with the oxyanion hole. The tetrahedral coordination of *t*Bu-GPI resembles the expected transition state structure with the alcoholate stabilised by hydrogen bonds to the backbone of Tyr631NH (3.1 Å) and the sidechain of Tyr547OH (2.9 Å). Intriguingly, superimposing the α/β-hydrolase fold of DP(IV) with AntI reveals an excellent structural overlap of the catalytic residues (Fig. 4c). We observe that in AntI the putative oxyanion hole is occupied by a well-defined water molecule (Fig. 4d) that forms hydrogen-bonds to the backbone of Leu174NH (2.7 Å) and Phe246NH (3.2 Å).

In accordance with the structural similarities of the active site in AntI and DP(IV), we modelled the putative reaction intermediates **INT1**–**INT3** and refined the protein–ligand complexes by classical molecular dynamics (MD) simulations. The procedure was initiated in reverse order, starting from a fully formed anthracene ring system at the same site as the tetrahedral hemiacetal carbon atom of *t*Bu-GPI is bound. Thus, we first modelled the late product state **INT3** which adopts a configuration having only one chiral centre out of plane with respect to the surrounding sp²-hybridised atoms. We find that **INT3** remains stably positioned at the oxyanion hole, the catalytic centre, and Arg24 during our 100 ns MD simulations. Interestingly, Tyr20, which is located at the entry gate of the substrate binding-pocket, switches between an open and a closed conformation during the MD simulations (ESI Fig. 7†). This finding is consistent with the two alternative positions of this residue in the X-ray structure (Fig. 4e). Furthermore, our simulations suggest that **INT3** forms van der Waals contacts with Leu284 and Ile328, as well as hydrogen bonds between the biaryl hydroxyl and keto groups *via* Arg24. In addition, the aromatic ring system of **INT3** is stabilised from both sides by cation–π interactions with Arg24 and by π-stacking with His355 (Fig. 4f, left panel). These insights allowed us to analyse the Dieckmann condensation in more detail.

INT3 was used as the starting pose for modelling the **INT2** state by following the enzymatic reaction cycle in reverse order. Thus, we opened the third ring of **AQ-256** between C3 and C16, leading to an acetyl group and an ester bond at the nucleophile Ser245O^γ (**INT2**, Fig. 4f, central panel). The **INT2** pose remains stable during the MD simulations, with less than 2 Å fluctuation in the atomic positions of **INT2** relative to **INT3** (ESI Fig. 7†). However, the **INT2** ligand has an increased flexibility that leads to rotation of the acetyl group by 45°. In addition, the carbonyl oxygen coordinates to the sidechain of Ser176, forming close contacts with Asp175 that favour the enol tautomer of the ligand. In accordance with the AntI X-ray structure, the sidechain of Asp175 is in proximity to Asp287 (2.5 Å) and, together with Glu181, forms an ion-pair with Arg24. This conformation facilitates deprotonation of the enol tautomer, *via* proton transfer to His355. Subsequently, the generated alkoxide might act as the nucleophile, initiating the Dieckmann condensation by intramolecular ring closure (Fig. 5a). The reaction takes place *via* a Bürgi–Dunitz angle in which the carbon atom in α-position (C16) is perfectly oriented to form a covalent bond with the carbonyl group (C3) of the acyl–AntI complex leading to **INT3** (Fig. 4f, right panel). Starting from **INT2**, this attack is regioselective from the *re*-face, while the backside is shielded from the catalytically active centre to exclude saponification that would result in shunt products.

We prolonged **INT2** by an additional acetyl group, and converted the ester at the active site Ser245 back to sp³-hybridisation (C3) (**INT1**, Fig. 4f, left panel). In doing this, we expected that the flexible, *ca.* 20 Å long PPT chain, introduces some structural uncertainty into the model. Intriguingly, the MD simulations of this state suggest that the oxyanion hole is rigid and dynamically stable over time, allowing us to properly refine the orientation of the introduced acyl–thioester moiety in **INT1**.

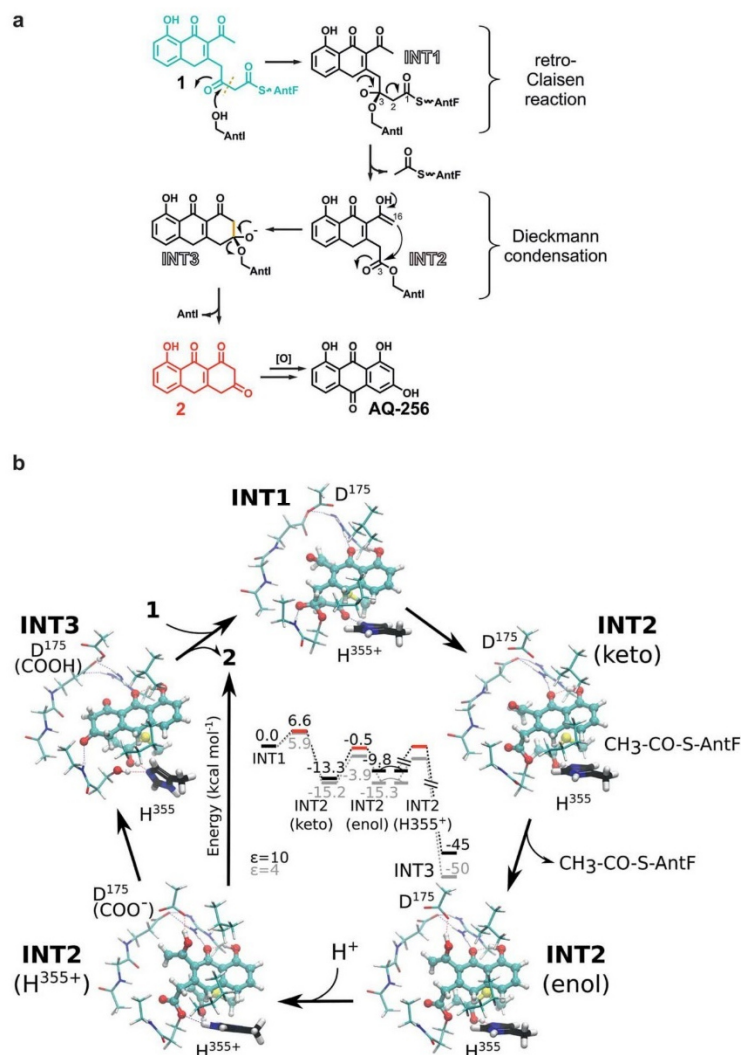


Fig. 5 (a) Proposed reaction scheme for the conversion of **1** into AQ-256 by AntI. (b) Quantum chemical DFT calculations (B3LYP-D3/def2-TZVP/ $\epsilon = 4/10$) on a putative mechanism starting from **INT1**. See main text for detailed description.

The steric hindrance from the PPT arm and the adjacent acetyl group at C14 significantly increased in our calculations, relative to the **INT2** and **INT3** states. The predictions suggest that **INT1** may re-orient towards the catalytically active serine residue in a new position induced by the tetrahedral configuration near the oxyanion hole (Fig. 4f, left panel). This conformation directs the carbonyl oxygen of the thioester at C1 towards His355, which might initiate the breakage of the adjacent C2–C3 bond. Due to the anionic charge distribution within the active site, both nitrogens of His355 might be protonated in the late state of **INT2**. In order to probe the energetics of these putative **INT1**–**INT3** intermediates, we performed quantum chemical density functional theory (DFT) calculations, starting from an active site model system constructed from an MD-relaxed snapshot of **INT1** (Fig. 5b). The DFT calculations illustrate that the proton

transfer from His355 to the carbonyl oxygen of the thioester group has a barrier of less than 10 kcal mol⁻¹, and is exergonic by 13–15 kcal mol⁻¹, indicating that the reaction is both kinetically and thermodynamically feasible. Hereby hydrolysis of the ester at the nucleophilic Ser245 is suppressed, and the subsequent Dieckmann condensation favoured. For the keto–enol tautomerisation forming **INT2**, we obtain a reaction barrier of ca. 12 kcal mol⁻¹ leading to an energetically degenerate state at ca. –10 kcal mol⁻¹. Proton transfer from the enol-tautomer to the nearby Asp175 increases the nucleophilic character of the enol, hence favouring the ring formation that takes place semi-concertedly. The cyclisation reaction succeeds *via* a Bürgi–Dunitz angle in which the carbon atom in α -position (C16) is perfectly oriented to form a covalent bond with the carbonyl group (C3) of the acyl–AntI complex leading to **INT3**. Due to

multiple bond formations and cleavage processes, the exact barrier for the INT2 \rightarrow INT3 reaction could not be determined, but the entire trajectory is energetically strongly exergonic by more than 40 kcal mol⁻¹. Additionally, we find that the process is favoured by re-protonation of the Ser245 nucleophile via His355, which regenerates the active site residue. Taken together, our sampled configurations and putative DFT models manifest that release of an acetyl-AntF induce shortening of the octaketide backbone and as a consequence, forms the third aromatic ring in AQ-256.

Conclusion

Anthraquinone derivatives may contribute to the toxicity of *Photorhabdus* extracts against other organisms such as insects, fungi and bacteria. In most strains of *Photorhabdus*^{18,19} corresponding gene clusters have been identified (ESI Fig. 8†), implying an important function of these natural products from this genus. In such pathways, the lyase AntI is exceptional because it shortens polyketides. The reaction mechanism of the lyase provides a unique route to the formation of the tricyclic aromatic ring (Fig. 5a) – a structural feature typically introduced by dedicated cyclases in type II PKS pathways.^{1,20} Interestingly, AntI and its homologues are also found in type II PKS-encoding biosynthetic gene clusters of cyanobacteria and other Gram-negative bacteria (ESI Fig. 8†), indicating that polycyclic aromatic polyketides are more abundant than previously thought. Furthermore, the amino acids around the proposed active centre of Ayg1p from *Aspergillus fumigatus*¹⁰ and WdYg1p from *Wangiella (Exophiala) dermatitidis*¹¹ are identical to AntI (ESI Fig. 9 and 10†), and therefore they can be clearly assigned as lyases. Intriguingly, these two enzymes are produced in pathogenic fungi and play a key role in the biosynthesis of the virulence factor 1,8-dihydroxynaphthalene-melanin. Thus, the structural and functional characterisation of AntI described here opens up a hitherto unexplored mechanism for an exciting class of catalysts.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors are grateful to Dr Sebastian Fuchs and Dr Darko Kresovic for MALDI-MS measurements and discussion. The Tsai and Brady labs are acknowledged for polyketide standards and the Ichinose lab for providing original data about S-DNPA. We thank the staff of the beamline X06SA at the Paul Scherrer Institute, SLS, Villigen (Switzerland) for assistance during data collection. This work was supported by the DFG within the SPP 1617 (H.·B.·B.), SFB 749 (M. G.), SFB 1035 (V. R. L. K.) and the LOEWE program of the state of Hesse as part of the MegaSyn research cluster (H.·B.·B.).

References

- 1 C. Hertweck, The Biosynthetic Logic of Polyketide Diversity, *Angew. Chem., Int. Ed.*, 2009, **48**, 4688–4716.
- 2 J. Staunton and K. J. Weissman, Polyketide biosynthesis: a millennium review, *Nat. Prod. Rep.*, 2001, **18**, 380–416.
- 3 C. Khosla, D. Herschlag, D. E. Cane and C. T. Walsh, Assembly line polyketide synthases: mechanistic insights and unsolved problems, *Biochemistry*, 2014, **53**, 2875–2883.
- 4 M. B. Austin and J. P. Noel, The chalcone synthase superfamily of type III polyketide synthases, *Nat. Prod. Rep.*, 2003, **20**, 79–110.
- 5 C. Hertweck, A. Luzhetskyy, Y. Rebets and A. Bechthold, Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork, *Nat. Prod. Rep.*, 2007, **24**, 162.
- 6 Z. Zhang, H.-X. Pan and G.-L. Tang, New insights into bacterial type II polyketide biosynthesis, *F1000Research*, 2017, **6**, 172–212.
- 7 A. O. Brachmann, *et al.*, A Type II Polyketide Synthase is Responsible for Anthraquinone Biosynthesis in *Photorhabdus luminescens*, *ChemBioChem*, 2007, **8**, 1721–1728.
- 8 A. Sandmann, *et al.*, A Type II polyketide synthase from the Gram-negative Bacterium *Stigmatella aurantiaca* is involved in Aurachin alkaloid biosynthesis, *Angew. Chem., Int. Ed.*, 2007, **46**, 2712–2716.
- 9 F. Panter, D. Krug, S. Baumann and R. Müller, Self-resistance guided genome mining uncovers new topoisomerase inhibitors from *myxobacteria*, *Chem. Sci.*, 2018, **9**, 4898–4908.
- 10 I. Fujii, *et al.*, Hydrolytic polyketide shortening by ayg1p, a novel enzyme involved in fungal melanin biosynthesis, *J. Biol. Chem.*, 2004, **279**, 44613–44620.
- 11 M. H. Wheeler, *et al.*, New Biosynthetic Step in the Melanin Pathway of *Wangiella (Exophiala) dermatitidis*: Evidence for 2-Acetyl-1,3,6,8-Tetrahydroxynaphthalene as a Novel Precursor, *Eukaryotic Cell*, 2008, **7**, 1699–1711.
- 12 J. D. Bauer, R. W. King and S. F. Brady, Utahmycins A and B, azaquinones produced by an environmental DNA clone, *J. Nat. Prod.*, 2010, **73**, 976–979.
- 13 T. Itoh, *et al.*, Actinorhodin biosynthesis: structural requirements for post-PKS tailoring intermediates revealed by functional analysis of ActVI-ORF1 reductase, *Biochemistry*, 2007, **46**, 8181–8188.
- 14 D. L. Ollis, *et al.*, The α/β hydrolase fold, *Protein Eng., Des. Sel.*, 1992, **5**, 197–211.
- 15 L. Holm and L. M. Laakso, Dali server update, *Nucleic Acids Res.*, 2016, **44**, W351–W355.
- 16 C. Schleberger, P. Sachelaru, R. Brandsch and G. E. Schulz, Structure and Action of a CC Bond Cleaving α/β -Hydrolase Involved in Nicotine Degradation, *J. Mol. Biol.*, 2007, **367**, 409–418.
- 17 M. Engel, *et al.*, Rigidity and Flexibility of Dipeptidyl Peptidase IV: Crystal Structures of and Docking Experiments with DPPIV, *J. Mol. Biol.*, 2006, **355**, 768–783.

Edge Article

- 18 N. J. Tobias, *et al.*, Natural product diversity associated with the nematode symbionts *Photorhabdus* and *Xenorhabdus*, *Nat. Microbiol.*, 2017, **2**, 1676–1685.
- 19 N. J. Tobias, *et al.*, Genome comparisons provide insights into the role of secondary metabolites in the pathogenic phase of the *Photorhabdus* life cycle, *BMC Genomics*, 2016, **17**, 537.
- 20 T. Taguchi, *et al.*, Bifunctionality of ActIV as a Cyclase-Thioesterase Revealed by *in Vitro* Reconstitution of Actinorhodin Biosynthesis in *Streptomyces coelicolor* A3(2), *ChemBioChem*, 2017, **18**, 316–323.

Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2019

Supplementary Information

Material and methods

Cultivation of strains. All *E. coli* DH10B strains (Supplementary Table 6), *E. coli* BL21 (DE3) without any overexpression plasmids and *Photobacterium* strains (Supplementary Table 7) used in this study were cultivated in liquid or solid Luria-Bertani (LB, pH 7.0) medium at 30°C. All *E. coli* BL21 (DE3) strains with different overexpression plasmids (Supplementary Table 5) were cultivated in a standard cultivation condition for protein expression (method described below). The cultivation condition for construction of *Photobacterium luminescens* mutants was described below. Chloramphenicol (34 µg mL⁻¹), ampicillin (100 µg mL⁻¹), kanamycin (50 µg mL⁻¹), streptomycin (50 µg mL⁻¹) and rifampicin (50 µg mL⁻¹) were used for the selection of strains with corresponding resistant markers. For cultivation of *E. coli* strain ES16 in ¹⁵N or ¹³C labeled medium, ISOGRO-¹⁵N or ISOGRO-¹³C medium was prepared according to manufacturer's instructions (Sigma-Aldrich). Isopropyl-β-D-1-thiogalactopyranoside (IPTG, 0.1 mM) was used to induce the heterologous expression of *ant* genes in *E. coli* BL21 (DE3).

General methods in molecular biology. All methods used in molecular biology were conducted according to standard procedures and manufacturer's instructions. All oligonucleotides (primers) obtained from Sigma-Aldrich were listed in Supplementary Table 4. All Plasmids constructed in this work were confirmed by sequencing at the SeqIT GmbH (Germany, Kaiserslautern) also listed in Supplementary Table 5. Polymerase chain reactions (PCRs) were performed using the phusion high-fidelity polymerase (Thermo-scientific). DNA isolation was performed with GeneJET™ Gel Extraction Kit (Fermentas). Plasmid isolation was performed with GeneJET™ Plasmid Miniprep Kit (Fermentas). Transformation of plasmids into *E. coli* was carried out using electroporation protocol for *E. coli* in an electroporation cuvette with a width of 1 mm (1250V, 25 µF, 200Ω). Genomic DNA was isolated using Gentra® Puregene® kit (Qiagen) according to the protocol for Gram-negative bacteria. Plasmid ZQ80 was constructed using an artificial gene (synthesized by Life Technologies™) flanked with restriction sites *EcoRI/PstI*.

Construction of *E. coli* strains with different combinations of *ant* genes. Two strategies were used to combine different *ant* genes in this work. The first one is the combination of pJET1.2/blunt and pSU18 based plasmids. The second one is the

1

combination of pCOLA Duet-1, pACYC Duet-1 and pCDF Duet-1 based plasmids. All *E. coli* strains used in this work (with the respective plasmids) were listed in Supplementary Table 4. Plasmids derived from pJET1.2/blunt and pSU18 vector were transformed into *E. coli* DH10B for heterologous expression of *ant* genes. Plasmids derived from Duet-vectors were transformed into *E. coli* BL21 (DE3) for IPTG induced heterologous expression of *ant* genes.

Heterologous expression of genes in *E. coli* DH10B: for analysis of produced polyketides. First of all, different DNA fragments were amplified using primers in Supplementary Table 4, resulting in *antB*, *antABC*, *antAB*, *antBC* PCR products. The PCR products were then digested with restriction enzymes *NdeI/SacI* and subsequently ligated into pSU18 vector, generating plasmids (Supplementary Table 5) ZQ6, ZQ10, ZQ11 and ZQ16, respectively. The plasmid pSUsfp was cloned using the same method. Later, *antDEFG*, *antDEFGH* and *antDEFGHI* genes were amplified with the primers in Supplementary Table 4. The resulted PCR products were directly ligated into pJET1.2/blunt, generating the plasmids ZQ9, ZQ12 and ZQ13 in Supplementary Table 5. The plasmid pSUMtaA was described previously.¹ All generated plasmids were selectively transformed into *E. coli* DH10B for heterologous expression, yielding strains listed in Supplementary Table 6. The polyketides found in the bacteria culture were listed in Supplementary Table 1.

Cloning, expression, purification, crystallization and structure determination of His₆-AntI His₆-AntI (*Photorhabdus luminescens*, Uniprot: Q7MZT8, expression vector ZQ40 [Supplementary Table 5]) was heterologously expressed in *E. coli* BL21(DE3) gold. Cells were grown in 6 liter LB-cultures in the presence of kanamycin (50 mg/L) at 37°C. Protein production was induced at an OD₆₀₀ of 0.6 by the addition of IPTG to a final concentration of 1 mM. Subsequently, cultures were incubated overnight at 20°C. Cells were harvested by centrifugation and stored at -20°C until further use. SeMet-labeled protein was expressed in the same strain using a protocol previously described².

For protein purification, cell pellets were resuspended in buffer A (50 mM tris hydrochloride, 500 mM NaCl, pH 7.5) and 20 mM imidazole. Cell disruption was carried out by sonication (Branson Digital Sonifier 250, G. Heinemann, Schwäbisch Gmünd, Germany). The resulting suspension was centrifuged at 30,000 g for 30 min at 4°C

using a SIGMA 4K15 centrifuge (Sigma Aldrich, Germany). The supernatant was applied to a 5 mL HisTrap™ HP column (GE Healthcare Life Science, Uppsala, Sweden), which had been equilibrated with buffer A (flow rate, 5 mL/min) using an ÄKTA prime plus system (GE Healthcare, Uppsala, Sweden). The column was washed with 50 mL of buffer A containing 50 mM imidazole. Next, the protein was eluted with buffer A containing 200 mM imidazole. Fractions were combined, concentrated and directly applied to a HiLoad™ 16/60 Superdex™ 200 column (GE Healthcare, Uppsala, Sweden), which had been equilibrated with 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM dithiothreitol (DTT). The obtained peak fractions were analysed by SDS PAGE, pooled and concentrated to a final concentration of 10 mg/mL using Amicon Ultra Centrifugal Filter Device (Millipore, Billerica, USA). Protein not directly applied to crystallographic experiments was flash frozen with liquid nitrogen and stored at -80°C.

Native and SeMet labeled His₆-AntI crystals were grown by the sitting drop vapor diffusion method at 20°C combining 0.2 µL protein (10 mg/mL in 20 mM Tris pH 7.5, 100 mM NaCl, 5 mM DTT) and 0.2 µL reservoir (0.1 M NaAc, 0.1 M HEPES pH 7.5, 22% PEG3350). Mountable crystals grew within one week. Native diffraction datasets were collected up to 1.85 Å resolution using synchrotron radiation at the X06SA-beamline, SLS, Villigen, Switzerland (see Supplementary Table 3). Recorded reflections were processed with the program package XDS³. His₆-AntI crystallized in the orthorhombic space group C222₁ with cell parameters of a = 54 Å, b = 154 Å, c = 91 Å, indicating one AntI-subunit in the asymmetric unit with a solvent content of 41%. Experimental phases were obtained by single anomalous dispersion (SAD) methods using the peak absorption wavelength of selenium derivatized AntI crystal ($\lambda = 0.9798$ Å, $f' = -8.05$; $f'' = 5.52$). The program package SHELXD⁴ located 10 heavy atom sites using a dataset recorded to 1.9 Å resolution. Subsequent SHARP-SAD-phasing⁵ and solvent flattening with the program DM⁶ resulted in an electron density map with phases at about 2.5 Å resolution. The quality was sufficient to model secondary structure elements by polyaniline residues. The initial model was transferred to the native dataset with a resolution of 1.85 Å by applying rigid body- and preliminary positional-refinement using REFMAC5⁷. The resulting electron density map allowed unambiguous identification of the entire AntI-sequence. The model was completed in iterative rounds with the three-dimensional graphic programs MAIN⁸ and COOT⁹. Temperature factors were anisotropically refined with restraints between bonded atoms using translation / libration / screw motion-parameters, yielding crystallographic

values of $R_{\text{cryst}} = 0.168$ and $R_{\text{free}} = 0.204$ (see Supplementary Table 3). Coordinates were confirmed to have superb stereochemistry in the Ramachandran plot with 98.1% of residues in the most favored region and 1.9% of residues in the additionally allowed regions. The asymmetric unit cell contains one AntI subunit with the six N-terminal and the three C-terminal amino acid residues being structurally disordered.

Heterologous expression of genes in *E. coli* BL21 (DE3): for mutagenesis of AntI.

ZQ40_S245A, ZQ40_D326A, ZQ40_D327A and ZQ40_H355A harboring the mutated gene *antI**, were generated (Supplementary Table 5). The site-directed mutation was achieved by ligation of two amplified PCR fragments, using the designed primer pair binding around the mutated region (Supplementary Table 4) and a primer pair binding at *ColA ori* (pCOLA Duet-1). The plasmid ZQ40 harboring the original *antI* gene was used as PCR template. The primer pair binding at *ColA ori* is 5'-[Phos]GTGGATTTAGATATCGAGAGTC-3' and 5'-TAATTCTCAGTTACCAATGGC-3'. For coexpression, plasmid ZQ62 encoding AntDEFGH and plasmid ZQ76 encoding AntABC were constructed using primers in Supplementary Table 4. Transformation of plasmids ZQ40, ZQ62 and ZQ76 into *E. coli* BL21 (DE3) led to strain ES53. Transformation of plasmids encoding mutated AntI* into *E. coli* BL21 (DE3) containing plasmids ZQ62 and ZQ76 led to strains ES59, ES61, ES62 and ES64 in Supplementary Table 6. HPLC-MS analysis is shown in Figure 3b.

Heterologous expression of genes in *E. coli* BL21 (DE3): for coexpression with RED1.

RED1 encoded originally by *actVI-ORF1* in *Streptomyces coelicolor* A3(2) was investigated in order to find out whether the polyketide intermediate **1** produced by AntA-H can serve as a substrate for the stereospecific ketoreduction.¹⁰ For this experiment, a gene encoding RED1 optimized for expression in *E. coli* was synthesized (Life Technologies™). The artificial gene flanked with restriction sites *EcoRI/PstI* was digested and ligated into pCOLA Duet-1 vector, resulted in plasmid ZQ80 (Supplementary Table 5). Plasmid ZQ62 encoding AntDEFGH and Plasmid ZQ76 encoding AntABC (described above) were used for coexpression with plasmid ZQ80 in *E. coli*, resulting in strain ES65 in Supplementary Table 6.

DNA sequence of the synthesized artificial gene encoding ketoreductase RED1 from the actinorhodin biosynthesis (*actVI-ORF1*):

^{EcoRI}GAATTCCATGAGCACCGTTACCGTTATTGGTGCAGGCACCATTGGTCTGGG
 TTGGATTAACCTGTTTAGCGCACGTGGTCTGACCGTTCGTGTTAATAGCCGTCG
 TCCGGATGTTTCGTGTTGTTTCATGAAGCACTGGAAGTGTAGTCCGGGTCG
 TGTTGATGAACTGGCAGCACGTATTGAATATGAACCGGATGTGGGTCGTGCAGT
 TGCCGGTGCAGATGTTGTTAGCGAAAATGCACCGGATGATCTGCCGCTGAAAC
 AGCGTCTGTTTGCAGAAATTGGTGCCGCAGCACCGGATCATGCACTGGTTCTGA
 GCAGCACAGCAAAGTCTGCCGGATGAACTGAGCCGTGATATGCCTGGTCCT
 GGTCGTCTGGTTGTTGCACATCCGTTTAAATCCGCCTCATATTGTTCCGCTGGT
 GAAGTTGTTTCGTGGTGAACGTACCGATCCGGAAGCAGTGAACGTACCCTGGC
 ATTTCTGGCAAGCGTTGGTCGTACACCGGTTGTTGTGCGTCGTGCACTGCCTG
 GTTTTGCAGCAAATCGTCTGCAAAGCGCACTGCTGCGTGAAAGCATTTCATCTGG
 TTCTGGAAGGTGTTGTTACCGTTGAAGAACTGGATCGTATTGTTACCGATAGTAT
 TGGCCTGCGTTGGAGCACAATTGGTCCGTTTCATGCATTTTCATCTGGGTGGTGG
 TCCGGGTGGTCTGCGTAAATGGCTGGAACATCTGGGTAGCGGTCTGGAACAGG
 GTTGGCGTGGTCTGGGCCAGCCTGCACTGACACCGCAGGCAGTTGAAGCCCT
 GGTGACAGACCGAAGCAGCATATGGTCATCGTCCGTATGCAGAACTGGTTC
 GTGATCGTGATGATCGTCATCTGGCCGTTCTGGCAGCCCTGGAACGCACCGAA
 CAGCCGCAAGAAGAAACCAATAACTGCAG^{PstI}

Construction of *P. luminescens* mutants. Insertion mutants (*antI::cat*) and the deletion mutant ($\Delta antC$) were constructed as described previously.¹¹

Analysis of polyketides in bacteria culture by HPLC-MS. For production analysis, extracts from 10 ml of culture were prepared. Usually, 1 % overnight pre-culture was used for inoculation and 2 % AMBERLITE™ XAD-16 (Sigma Aldrich) was added at the same time for the absorption of hydrophobic secondary metabolites. AMBERLITE™ XAD-16 resin beads were collected with a sieve (or filter paper) after 72 h, washed with small amount of water and extracted with 20 mL of methanol. The methanol extract was dried and dissolved in 1 mL methanol, which could be then diluted (1:10-dilution) for the analysis in HPLC-MS. Additionally, the remaining bacteria culture without AMBERLITE™ XAD-16 and bacteria cells was extracted using 10 mL ethyl acetate/acetic acid (99:1). After remove of all the solvent, the dried extract was dissolved in 1 mL methanol, which could be then diluted (1:10-dilution) for the analysis in HPLC-MS.

Analysis of extracts was performed on Dionex ULTIMATE 3000 HPLC system with a photodiode array detector in the range of 200-600 nm and an Acquity UPLC BEH C18 column (1.7 μm , Waters), which is coupled to Bruker AmaZon X iontrap mass spectrometer using electrospray ionization at positive and negative mode. Solvent A: water with 0.1 % formic acid and solvent B: acetonitrile with 0.1 % formic acid. Gradient A: 5 % solvent B for 2 min, increasing to 95 % solvent B in 12 min, 95 % solvent B for 3 min. Gradient B: 10 % solvent B for 4 min, increasing to 25 % in 9 min, increasing to 95 % solvent B in 1 min, 95 % solvent B for 3 min. Flow rate: 0.6 mL min⁻¹. Gradient A was used as standard gradient for most analysis in this work. Gradient B was used for separation of **SEK4**, **SEK4b**, **mutactin** and **SEK34** (Supplementary Table 2).

Biochemical assays. Hydrolase AntI activity assays were performed with synthesized model compounds **S1-S4** (Supplementary Figure 6). The methyl esters (**S2** and **S4**) were hydrolyzed *in situ* with 250 μM NaOH solution with a final concentration of 2 mM for synthetic analogs. The assays with the four model compounds (2.5 μM , **S1-S4**) were realized with 250 μM His₆-AntI in Tris-HCl (pH 8). The same assays were performed without His₆-AntI as negative controls. After incubation for 3 h at 30°C, the reaction was stopped by the addition of 200 μL ACN to the mixture. After centrifugation, the clear supernatants were analysed by HPLC/MS analysis quantifying the formation of **S5** (m/z 221 [M+H]⁺). Results showed that model compounds **S1** and **S3** (Figure 3c) could not be converted by His₆-AntI. For the conversion of **S2** to **S5**, no significant difference with or without AntI could be observed even for shorter incubation times of 1 h.

Synthesis of S2. 145 mg of methyl 2-(3,5-dimethoxyphenyl)acetate (**i**) were dissolved in 5 ml of acetic acid, then 540 μl of acetic anhydride and 3 drops of HClO₄ were added to the mixture. A yellow precipitate appeared. The mixture was heated to 35°C for 7 min. The completion of the reaction was monitored by TLC. The acid was quenched with ice and 20 ml of saturated NaHCO₃. The aqueous layer was extracted 3 times with 10 ml of Et₂O. The organic layers were combined and extracted with 20 ml of brine and then dried over Na₂SO₄. The solvents were removed under vacuum. The crude product was purified by flash chromatography with a gradient from 40-80 % EtOAc in hexane over 10 column volumes. After evaporation of the solvents, 120 mg of **S2** were

obtained as yellow oil. ^1H NMR (500 MHz, CDCl_3) δ 6.31 (dd, $J = 30.1, 2.1$ Hz, 2H), 3.76 (s, 3H), 3.74 (s, 3H), 3.62 (s, 2H), 3.61 (s, 3H), 2.46 (s, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 203.77, 171.77, 161.61, 159.45, 134.97, 123.64, 108.27, 97.50, 55.62, 55.42, 51.97, 39.10, 32.24. HRMS (ESI) Calcd for $\text{C}_{13}\text{H}_{17}\text{O}_5$: 253.1070 $[\text{M}+\text{H}]^+$, Found: 253.1089 $[\text{M}+\text{H}]^+$.

Synthesis of ii. 565 mg of methyl ester **i** (2.7 mmol, 1 equiv.) were dissolved in 8.5 ml of dry toluene under nitrogen. The mixture was then cooled down to -78°C and 2.8 ml of Me_3Al in hexane (2M, 2.8 mmol, 1.03 eq.) was added. 1.55 g of tert-butyl((1-methoxyvinyl)oxy)dimethylsilane (7.57 mmol, 2.83 equiv.) were dissolved in 2 ml of toluene. This solution was slowly added to the reaction mixture. The mixture was warmed to 0°C and stirred for 7 hours. The mixture was quenched with 100 ml of saturated aqueous solution of NaHCO_3 , and then extracted 3 times with 30 ml of EtOAc. The organic layers were combined and washed once with brine, then dried over Na_2SO_4 . The solvent removed under vacuum and the residue oil was purified by flash chromatography using a gradient from 5-40 % of EtOAc in hexane over 10 column volumes. The two diastereoisomers of **ii** were isolated, 250 mg of diastereoisomer A (22% of yield) and 120 mg of diastereoisomer B (11%) were obtained both as yellow oil. Isomer A: ^1H NMR (250 MHz, CDCl_3) δ 6.44 (d, $J = 2.3$ Hz, 2H), 6.21 (t, $J = 2.3$ Hz, 1H), 3.64 (s, 6H), 3.54 (s, 3H), 3.20 (s, 3H), 3.02 (dd, $J = 33.6, 13.1$ Hz, 2H), 2.44 (dd, $J = 45.5, 14.2$ Hz, 2H), 0.78 (s, 9H), 0.02 (d, $J = 9.1$ Hz, 6H). Isomer B: ^1H NMR (250 MHz, CDCl_3) δ 6.44 (d, $J = 2.3$ Hz, 2H), 6.21 (t, $J = 2.2$ Hz, 1H), 3.64 (s, 6H), 3.54 (s, 3H), 3.20 (s, 3H), 3.02 (dd, $J = 33.6, 13.1$ Hz, 2H), 2.44 (dd, $J = 45.4, 14.2$ Hz, 2H), 0.78 (s, 10H), 0.02 (d, $J = 9.1$ Hz, 6H). ^{13}C NMR (63 MHz, CDCl_3) δ 170.33, 160.30, 139.51, 108.96, 100.80, 98.73, 55.28, 51.44, 49.05, 44.50, 41.52, 25.86, 18.34, -2.67, -2.90.

Synthesis of iii. 160 mg of **ii** was dissolved in 16 ml of MeOH, and 8 ml of HCl 6M were added at 0°C . The mixture was stirred 30 min at 0°C then the ice bath was removed and the mixture was stirred for 1 hour at room temperature. The mixture was quenched with 100 ml of NaHCO_3 and extracted three times with 25 ml of AcOEt. The organic layers were combined, washed with brine and dried over Na_2SO_4 . The solvents were removed under vacuum and the crude was purified by flash chromatography. The expected compound **iii** (63 mg) was obtained as yellow oil with

a yield of 60%. ¹H NMR (500 MHz, CDCl₃) δ 6.37 (d, *J* = 2.2 Hz, 1H), 6.34 (d, *J* = 2.2 Hz, 2H), 3.77 (s, 6H), 3.73 (s, 2H), 3.71 (s, 3H), 3.46 (s, 2H). ¹³C NMR (126 MHz, CDCl₃) δ 197.78, 165.06, 158.59, 132.71, 105.02, 104.81, 96.86, 52.80, 52.77, 49.84, 47.84, 45.18.

Synthesis of S4. 30 mg of **iii** were dissolved in 0.190 ml of acetic acid, then 180 μl of acetic anhydride and two drops of HClO₄ were added to the mixture. The mixture color turned to orange. It was heated to 35°C for 5 min. The completion of the reaction was monitored by TLC. The acid was quenched with ice and 20 ml of saturated NaHCO₃. The aqueous layer was extracted 3 times with 10 ml of Et₂O. The organic layers were combined and washed with 10 ml of brine and then dried over Na₂SO₄. The solvents were removed under vacuum. The crude product was purified by flash chromatography with a gradient from 40-80 % EtOAc in hexane over 10 column volumes. After evaporation of the solvents, 12 mg of **S4** (yield: 36%) were obtained as a slight purple solid. ¹H NMR (500 MHz, CDCl₃) δ 6.35 (d, *J* = 2.2 Hz, 1H), 6.23 (d, *J* = 2.2 Hz, 1H), 3.78 (s, 3H), 3.77 (s, 2H), 3.75 (s, 3H), 3.67 (s, 3H), 3.51 (s, 2H), 2.41 (s, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 203.00, 199.83, 167.92, 162.03, 160.14, 135.60, 122.76, 108.83, 97.69, 55.62, 55.46, 52.31, 48.59, 48.27, 32.29. HRMS (ESI) Calcd for C₁₅H₁₉O₆: 295.1176 [M+H]⁺, Found: 295.1184 [M+H]⁺.

Multiple sequence alignments of hydrolases. All performed multiple sequence alignments were calculated using the ClustalW-algorithm¹² applying the standard parameters.

Molecular dynamics (MD) simulations. Classical molecular dynamics (MD) simulations were performed with compound **1** modelled in different poses into the crystal structure of the *apo*-form of AntI. The ligand-protein system was solvated and neutralized in a water-ion box with a 150 mM KCl concentration, comprising in total ca. 59,000 atoms. The CHARMM36 force field¹³ was employed to model the protein, the water molecules, and the ions. Different reaction intermediates (substrate, Bürgi-Dunitz intermediate, product, see main text) of the ligand were parametrized using a combination of the CGenFF force field¹⁴ and *in-house* DFT-based parameters computed at B3LYP/def2-SVP level. The MD simulations were performed at *T* = 310 K and *p* = 1 bar in an *NPT* ensemble, using an integration timestep of 2 fs. Long-range

electrostatic interactions were treated by the Particle Mesh Ewald (PME) method. All residues were modeled in their standard protonation states, except for His355 and Asp287, which were modeled in their protonated forms. The different ligand poses were relaxed for 100 ns. The MD simulations were performed with NAMD2¹⁵, and VMD¹⁶ was used for analysis and visualization.

DFT calculations. DFT models were built based on classically relaxed structures of the substrate:AntI complex. The DFT models comprised the aromatic rings of the substrate and the S-acyl moiety of its PPT-arm, the sidechains of protein residues Ser245, Arg24, Asp175, Asp287, Leu284, Ile328, His355, and the backbone atoms of residues Phe246 and Leu174 that form the putative oxyanion-hole. Sidechains were cut at the C α -C β bond, except for the arginine, which was cut at the C γ -C δ position, and saturated with hydrogen atoms. The DFT models comprised ca. 140 atoms. The terminal parts of the protein residues were kept fixed during optimizations to simulate the constraints arising from the rigid protein framework. Geometry optimizations were performed at the B3LYP-D3/def2-SVP/ $\epsilon=10$ ¹⁷⁻²¹, and single point energy evaluations were performed at B3LYP-D3/def2-TYZVP with $\epsilon=10$ and $\epsilon=4$. Reaction pathways were optimized using a *chain-of-states* method^{22,23}. All QM calculations were performed with TURBOMOLE version 7.1-7.3²⁴.

Supplementary Table 1. *E. coli* strains with coexpressed proteins and major polyketide products.

Strain	Proteins	Major products
ES8	AntBDEFG	SEK4, SEK4b
ES10	AntB-I	SEK4, SEK4b
ES12	AntABDEFG	mutactin
ES13	AntABCDEFG	mutactin
ES14	AntABDEFGH	SEK34b
ES15	AntABDEFGHI	SEK34b
ES16	AntABCDEFGH	SP^{EC1}, SP1, SP4, SP5
ES17	AntABCDEFGHI	AQ-256, SP1, SP^{EC1}
ES65	AntABCDEFGH + Red1	S-DNPA

Supplementary Table 2. Identification of polyketides using HPLC-UV-MS or NMR.

	[M+H] ⁺	[M-H] ⁻	RT ^A / min	RT ^B /min	identified with
SEK4	301 ^C		5.3	6.5	authentic standard (Tsai lab) , UV
SEK4b	319		5.4	7.1	authentic standard (Tsai lab), UV
mutactin	303		6.3	10.2	authentic standard ($\Delta antC$ mutant), UV
SEK34	285 ^D		6.0	9.5	UV
SEK34b	285		6.8		UV
DMAC (SP^{EC1})		297	7.8		isolation and NMR, UV
SP1	302		6.2		UV, labeled media
SP4	240		8.2		labeled media
utahmycin A (SP5)	254		8.4		authentic standard (Brady lab) ²⁵
AQ-256		255	8.7		standard (TT01 wild type), UV
S-DNPA	287		8.5		UV and MS; Ichinose lab ¹⁰

A: HPLC gradient A; B: HPLC gradient B (described above); C: SEK4-H₂O; D: SEK34-H₂O

Supplementary Table 3. Crystallographic data collection and refinement statistics.

	Antl	Antl [SeMet (peak)]
<u>Crystal parameters</u>		
Space group	C222 ₁	C222 ₁
Cell constants	a= 54.1 Å b= 154.3 Å c= 91.2 Å	a= 54.1 Å b= 154.5 Å c= 91.2 Å
Subunits / AU ^a	1	1
<u>Data collection</u>		
Beam line	X06SA, SLS	X06SA, SLS
Wavelength (Å)	1.0	0.979
Resolution range (Å) ^b	30-1.85 (1.95 - 1.85)	30-1.9 (2.0 - 1.9)
No. observed reflections	221,122	220,792
No. unique reflections ^c	33,019 [*]	57,339 [#]
Completeness (%) ^b	99.7 (100)	98.4 (99.5)
R _{merge} (%) ^{b, d}	5.7 (56.0)	6.4 (51.6)
I/σ (I) ^b	18.9 (3.0)	11.7 (2.7)
<u>Refinement (REFMAC5)</u>		
Resolution range (Å)	30-1.85	
No. refl. working set	31,340	
No. refl. test set	1,649	
No. non hydrogen	3,253	
Solvent (H ₂ O, PEG, glycerol)	205	
R _{work} / R _{free} (%) ^e	16.8 / 20.4	
r.m.s.d. bond (Å) / (°) ^f	0.003 / 0.79	
Average B-factor (Å ²)	37.1	
Ramachandran Plot (%) ^g	98.1 / 1.9 / 0	
PDB accession code	6HXA	

^[a] Asymmetric unit

^[b] The values in parentheses for resolution range, completeness, R_{merge} and I/σ (I) correspond to the highest resolution shell

^[c] Data reduction was carried out with XDS and from a single crystal. *Friedel pairs were treated as identical reflections; #Friedel pairs were treated as individual reflections

^[d] $R_{\text{merge}}(I) = \frac{\sum_{hkl} \sum_j |I(hkl)_j - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_j I(hkl)_j}$, where $I(hkl)_j$ is the j^{th} measurement of the intensity of reflection hkl and $\langle I(hkl) \rangle$ is the average intensity

^[e] $R = \frac{\sum_{hkl} | |F_{\text{obs}}| - |F_{\text{calc}}| |}{\sum_{hkl} |F_{\text{obs}}|}$, where R_{free} is calculated without a sigma cut off for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections

^[f] Deviations from ideal bond lengths/angles

^[g] Number of residues in favored region / allowed region / outlier region

Supplementary Table 4. Primers used for the plasmid construction and genotype verification. Restriction sites are underlined.

plasmid	primer 5' to 3'	vector
ZQ6	TACTAACATAT <u>GGACGATATTTCTTTATCATCTGATT</u> GTCAAT <u>GTCGACT</u> ATTACTCATCTTTGTTCCTTATAATCTCG	pSU18
ZQ9	ACTTGCAGATAACGATTTTC ACATTCCTGGCCATTTAT	pJET1.2/blunt
ZQ10	TAGTTCCATATGAAATATGCCTTTATTACCGG CTTGAAGTCGACCCATTGGGTATATGAAATCTCTTT	pSU18
ZQ11	TAGTTCCATATGAAATATGCCTTTATTACCGG TTCTCAGTCGACTTCCCAAAAATCACAATCTATAGG	pSU18
ZQ12	ACTTGCAGATAACGATTTTC CTATTGGGTTTATTTTATTATTCATCT	pJET1.2/blunt
ZQ13	ACTTGCAGATAACGATTTTC TTACCATCGCGATGTATATT	pJET1.2/blunt
ZQ16	TACTAACATAT <u>GGACGATATTTCTTTATCATCTGATT</u> CTTGAAGTCGACCCATTGGGTATATGAAATCTCTTT	pSU18
ZQ40	AAAATGAATTCATGAATAATAAAAATAAACCCAATAGA GTATACTGCAGTCAATTAACCTTTTTTATAGCCA	pCOLA Duet-1
ZQ40_S245A	[Phos]CTTTTGGTGGTTATTTTGC CAATTCCTAAGAAACAAAGAAG	pCOLA Duet-1
ZQ40_D326A	[Phos]GAAAGTGAAAAATTAGATCAAC GCGATCTGAAATATATCATCTAA	pCOLA Duet-1
ZQ40_D327A	[Phos]CTATATTTTCAGATCGATAAAGTG CATCTAATTCACCATGAAC	pCOLA Duet-1
ZQ40_H355A	[Phos]CTGTTTGCTTAAATAAAAATAAACG CAGCCTCTGATTCATAACATAA	pCOLA Duet-1
ZQ62	TAATTGAATTCGGTGATAATAAATAACAGAAATGAATCTCAACC TCCTTTCTGCAGGTTACTAAATACGAGTGTCTAACCCT	pACYC Duet-1
ZQ76	TAGTTCGAATTCATGAAATATGCCTTTATTACCGG ATGACACTGCAGTTATTATAATATTGCGACCACTC	pCDF Duet-1
ZQ80	No primers, more information in method section.	pCOLA Duet-1
pDS_plu4186	GGTCAAGCATGCCTGGGTGATAGCTATATTAATATCG TTCACCTGAGCTCCCAATCTGAAACTTGATCAT	pDS132
pDS_plu4188	AACTTAGCATGCCTCCGCAATCTATTGCTAAC CGAGATGAGCTCCCAAGTGGCAAACCACTC	pDS132
pDS_plu4192	11	pDS132
primers for verification PCR		
vP_plu4186_Fw	GTGATTCAGTAAAAGTCATTTATAATG	
vP_plu4186_Rv	GCCAGTTAATACCTCAGCAG	
vP_plu4188_Fw	GCGCTTTAGTAATCAAGGTC	
vP_plu4188_Rv	GCTGAGAATTGATTTTAATTACG	
vP_plu4192_Fw	TACCTTATGGATTTCAAGATGC	
vP_plu4192_Rv	AACTCTTTGTTATTGCCATCAC	
pDS132fw	GATCGATCCTCTAGAGTCGACCT	
pDS132rv	ACATGTGGAATTGTGAGCGG	

Supplementary Table 5. Plasmids used in this work. *: gene *antI* with selected mutation. **: codon optimized for *E. coli* expression.

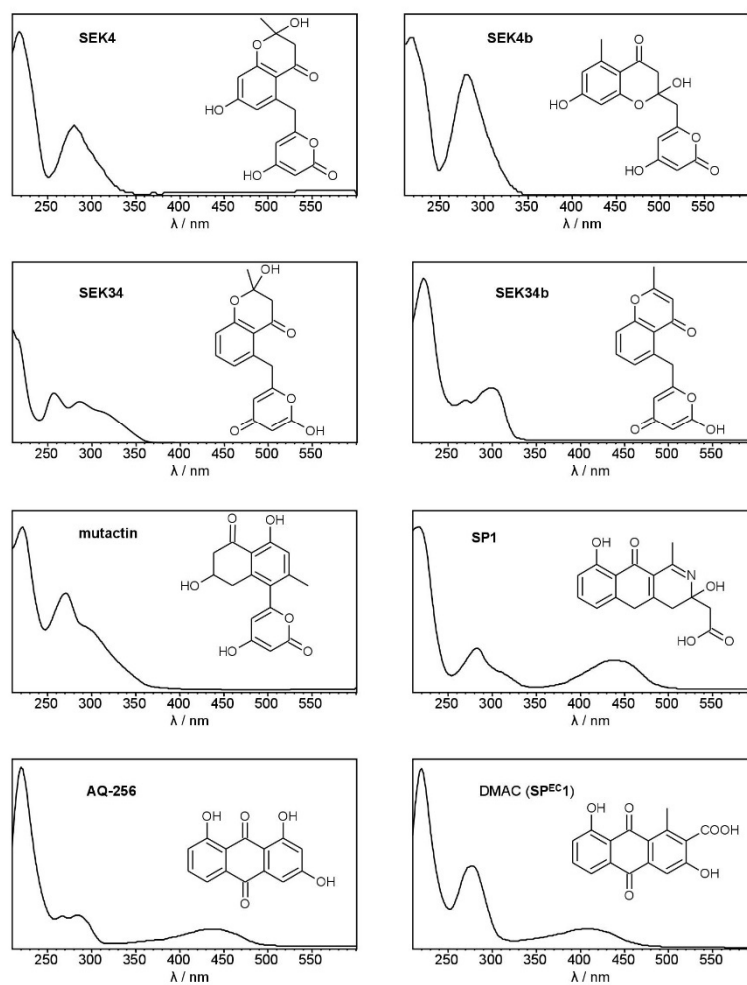
plasmid	genotype	reference
pJET1.2/blunt	pMB1ori, Ap ^r	Fermentas
pSU18	P15A ori, Cm ^r , <i>lacZ</i> promoter	²⁶
pACYC Duet-1	P15A ori, Cm ^r , T7 <i>lac</i> promoter	Novagen
pCOLA Duet-1	ColA ori, Km ^r , T7 <i>lac</i> promoter	Novagen
pCDF Duet-1	CDF ori, Sm ^r , T7 <i>lac</i> promoter	Novagen
pDS132	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i>	²⁷
ZQ6	P15A ori, Cm ^r , <i>antB</i> , <i>lacZ</i> promoter	this work
ZQ9	pMB1 ori, (Ap ^r), <i>antDEFG</i> with native promoter	this work
ZQ10	P15A ori, Cm ^r , <i>antABC</i> , <i>lacZ</i> promoter	this work
ZQ11	P15A ori, Cm ^r , <i>antAB</i> , <i>lacZ</i> promoter	this work
ZQ12	pMB1ori, (Ap ^r), <i>antDEFGH</i> with native promoter	this work
ZQ13	pMB1ori, (Ap ^r), <i>antDEFGHI</i> with native promoter	this work
ZQ16	P15A ori, Cm ^r , <i>antBC</i> , <i>lacZ</i> promoter	this work
ZQ40	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antI</i>	this work
ZQ40_S245A	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antI</i> *	this work
ZQ40_D326A	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antI</i> *	this work
ZQ40_D327A	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antI</i> *	this work
ZQ40_H355A	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>antI</i> *	this work
ZQ62	P15A ori, Cm ^r , T7 <i>lac</i> promoter, <i>antDEFGH</i>	this work
ZQ76	CDF ori, SM ^r , T7 <i>lac</i> promoter, <i>antABC</i>	this work
ZQ80	ColA ori, Km ^r , T7 <i>lac</i> promoter, <i>actVI-ORF1</i> **	this work
pDS_plu4186	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i> , partial <i>plu4186</i>	this work
pDS_4188	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i> , partial <i>plu4188</i>	this work
pDS_plu4192	<i>pir</i> dependent, Cm ^r , <i>oriT</i> , <i>oriV</i> , <i>sacB</i> , partial <i>plu4192</i>	this work

Supplementary Table 6. *E. coli* strains (ES) used in this work.

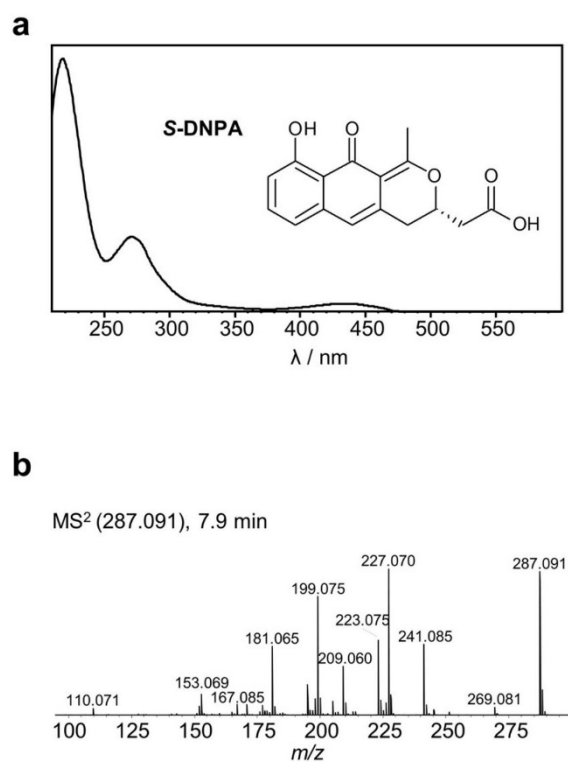
strain	genotype	reference
<i>E. coli</i> DH10B		28
<i>E. coli</i> BL21 (DE3)		Novagen
<i>E. coli</i> s17-1 λ pir		29
ES8	DH10B::ZQ9, ZQ6	this work
ES10	DH10B::ZQ13, ZQ16	this work
ES12	DH10B::ZQ9, ZQ11	this work
ES13	DH10B::ZQ9, ZQ10	this work
ES14	DH10B::ZQ12, ZQ11	this work
ES15	DH10B::ZQ13, ZQ11	this work
ES16	DH10B::ZQ12, ZQ10	this work
ES17	DH10B::ZQ13, ZQ10	this work
ES36	BL21 (DE3)::ZQ40	this work
ES53	BL21 (DE3)::ZQ62, ZQ76, ZQ40	this work
ES59	BL21 (DE3)::ZQ62, ZQ76, ZQ40_S245A	this work
ES61	BL21 (DE3)::ZQ62, ZQ76, ZQ40_D326A	this work
ES62	BL21 (DE3)::ZQ62, ZQ76, ZQ40_D327A	this work
ES64	BL21 (DE3)::ZQ62, ZQ76, ZQ40_H355A	this work
ES65	BL21 (DE3)::ZQ62, ZQ76, ZQ80	this work
ES27	s17-1 λ pir::pDS_plu4186	this work
ES28	s17-1 λ pir::pDS_4188	this work
ES29	s17-1 λ pir::pDS_4192	this work

Supplementary Table 7. *P. luminescens* TT01 wild type and mutants with genotypes and major products

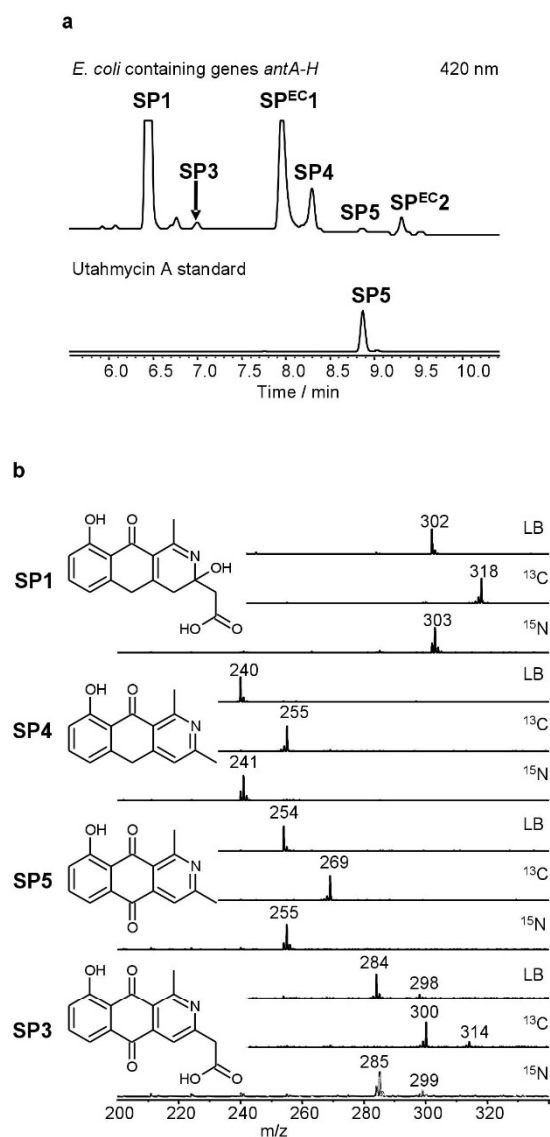
TT01 strain	Genotype	Major products	Reference
WT	<i>P. luminescens</i> wild type	AQ-256, AQ-270a, AQ-284a	30
<i>antI</i> ::cat	<i>antI</i> inactivated by plasmid insertion	SP1	this work
$\Delta antC$	<i>antC</i> deletion	SEK34, SEK34b	this work
$\Delta antH$	<i>antH</i> deletion	mutactin	11
<i>antD</i> ::cat	<i>antD</i> inactivated by plasmid insertion	no polyketide	11



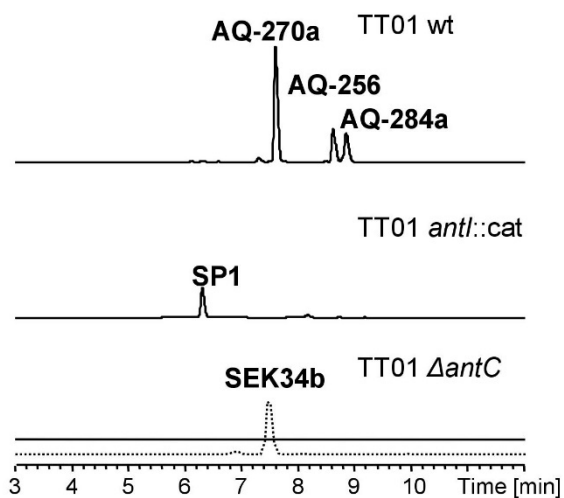
Supplementary Figure 1. UV spectra (λ : 210-600 nm) and structures of major shunt products produced by *E. coli* DH10B with *ant* genes.



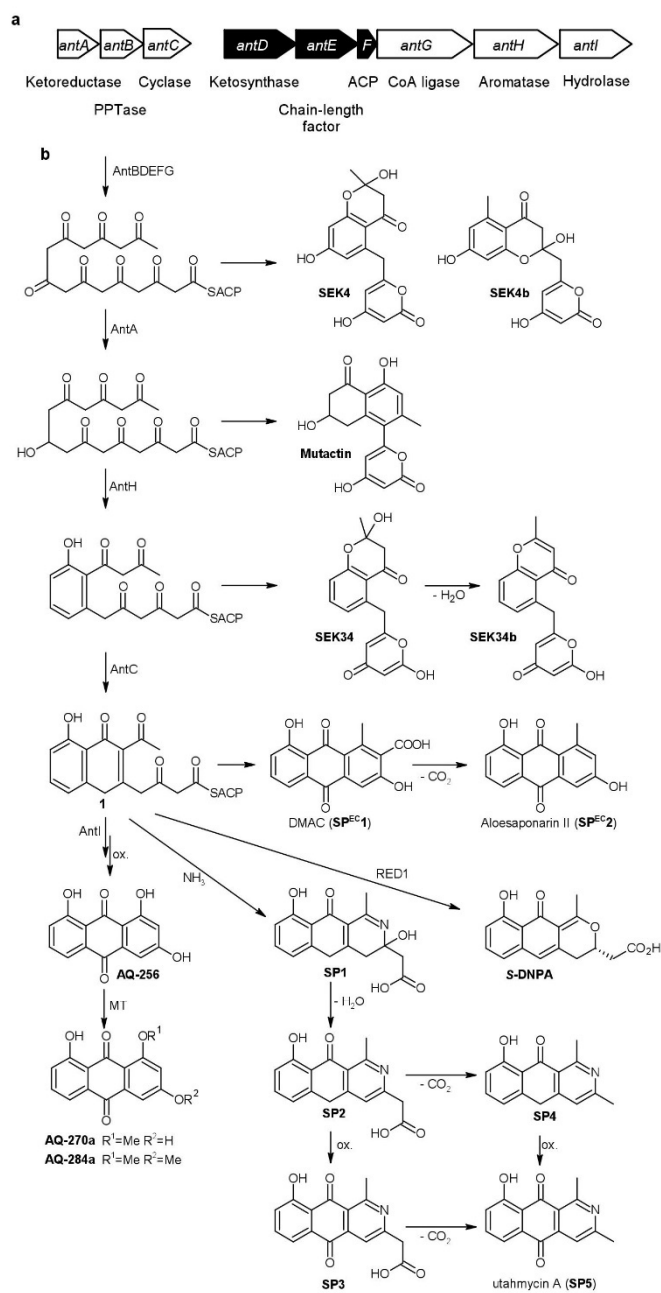
Supplementary Figure 2. **a)** UV (λ : 210-600 nm) and **b)** MS² spectra of S-DNPA produced by *E. coli* BL21 (DE3) with *antA-H* and the gene encoding for RED1. The shown data corresponds well with published data from the Ichinose lab.¹⁰



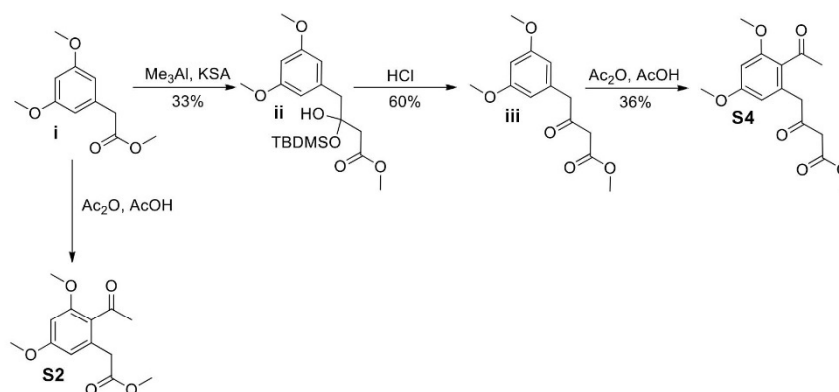
Supplementary Figure 3. (a) HPLC-UV analyses (420 nm) of extracts of heterologous expression of genes *antABCDEFGH* in *E. coli* DH10B (top) and authentic utahmycin A standard (bottom). **(b)** Determination of the number of carbon and nitrogen atoms for **SP1**, **SP4**, **SP5** and **SP3** using cultivation of the *E. coli* DH10B strain ES16 with genes *antABCDEFGH* in standard growth medium (LB medium, ^{14}N , ^{12}C , ^{16}O , ^1H), ^{15}N labeled medium (^{15}N , ^{12}C , ^{16}O , ^1H), or ^{13}C labeled medium (^{14}N , ^{13}C , ^{16}O , ^1H) as described previously.³¹



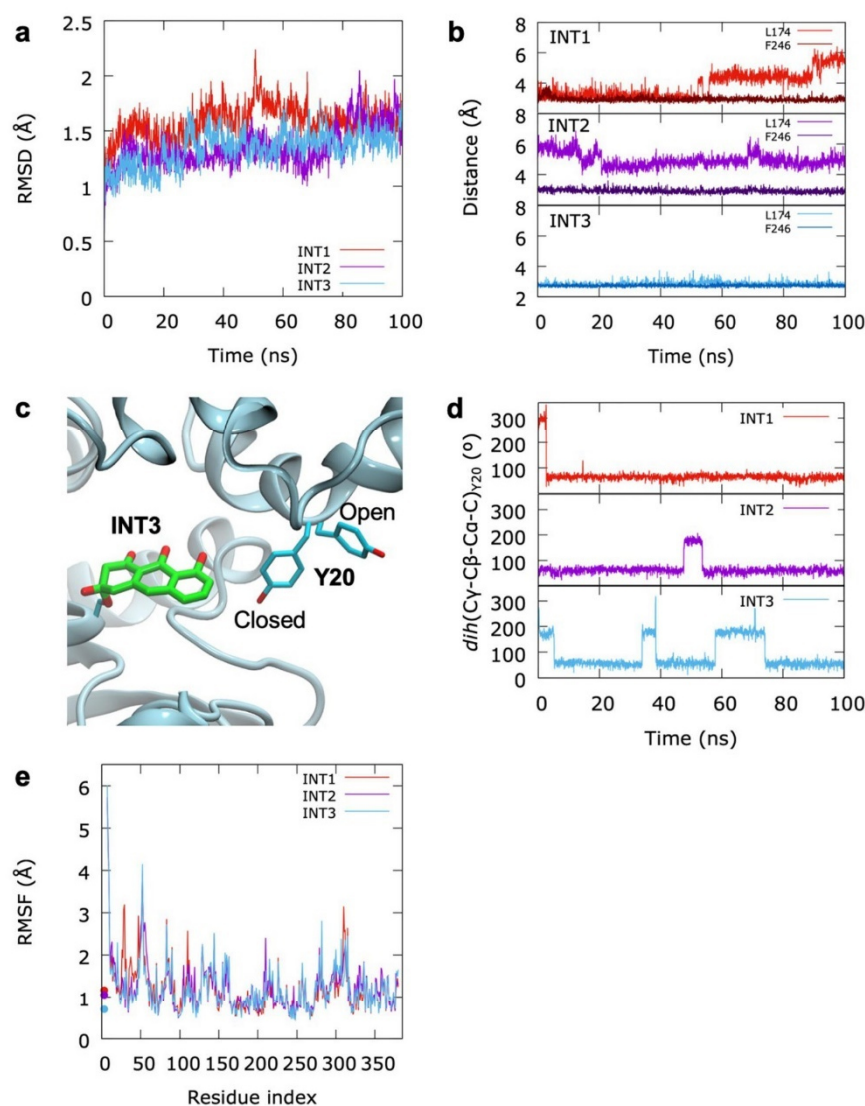
Supplementary Figure 4. HPLC-UV analyses (420 nm) of extracts of *P. luminescens* TT01 wild type, mutant *antI::cat* and mutant $\Delta antC$. In both mutants, the wt AQs were not produced. A new shunt product **SP1** could be identified in the mutant *antI::cat*. In the mutant $\Delta antC$, shunt product **SEK34b** could be identified with EIC (dotted line, m/z 285 $[M+H]^+$). See also Supplementary Table 7.



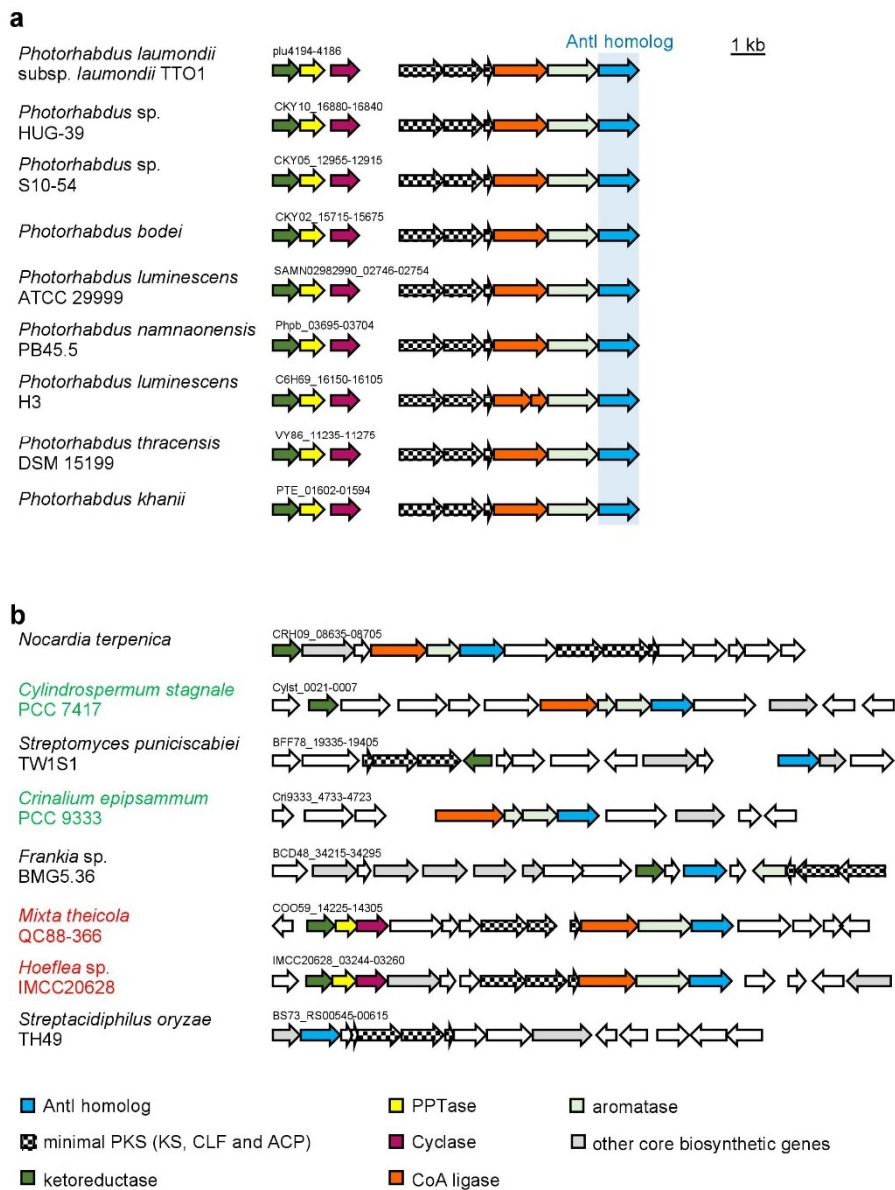
Supplementary Figure 5. Biosynthesis gene cluster *antABCDEFGHI* for the production of anthraquinones from *P. luminescens* (a) and proposed biosynthesis of all identified compounds (b) following classical type II PKS pathways.^{32,33}



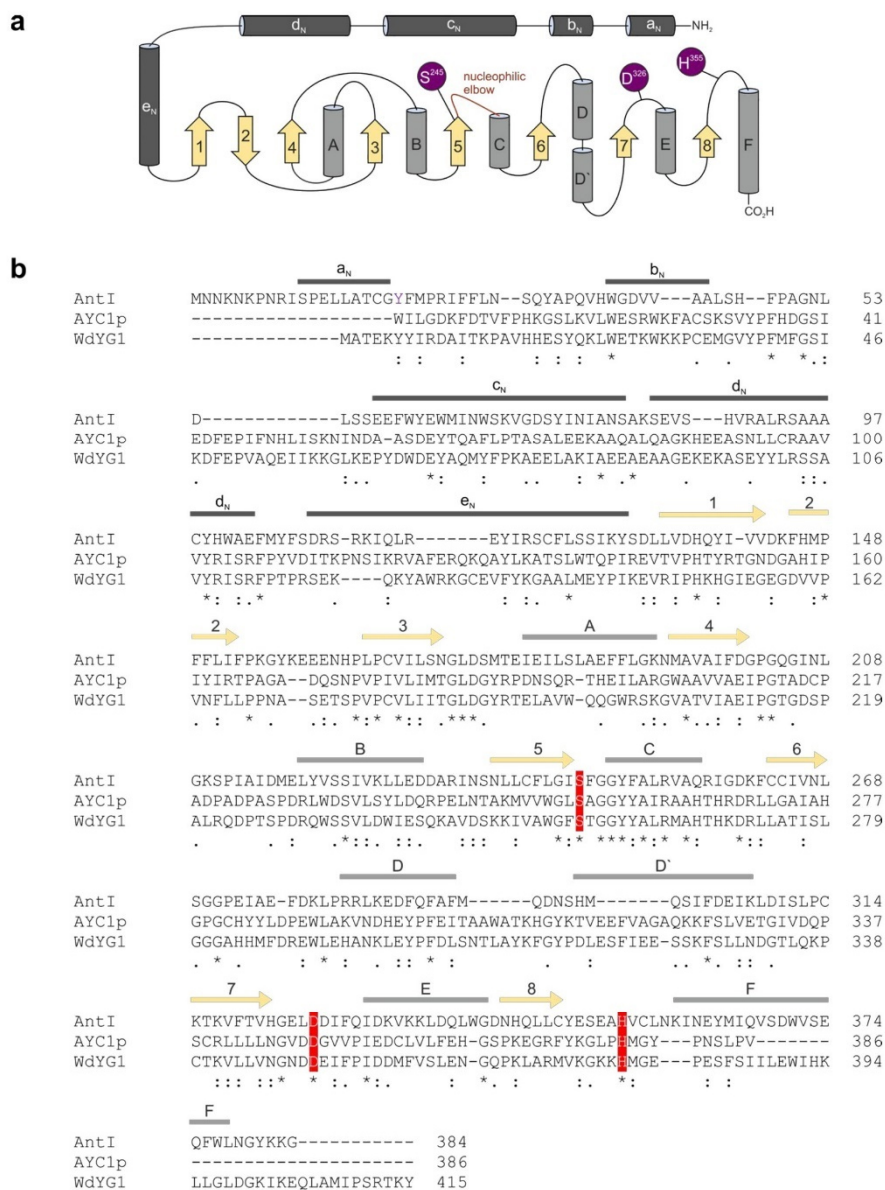
Supplementary Figure 6. Synthesis of model compounds **S2** and **S4**. KSA: tert-butyl((1-methoxyvinyl)oxy)dimethylsilane; Me_3Al : trimethylaluminium; AcOH: acetic acid; Ac_2O : acetic anhydride.



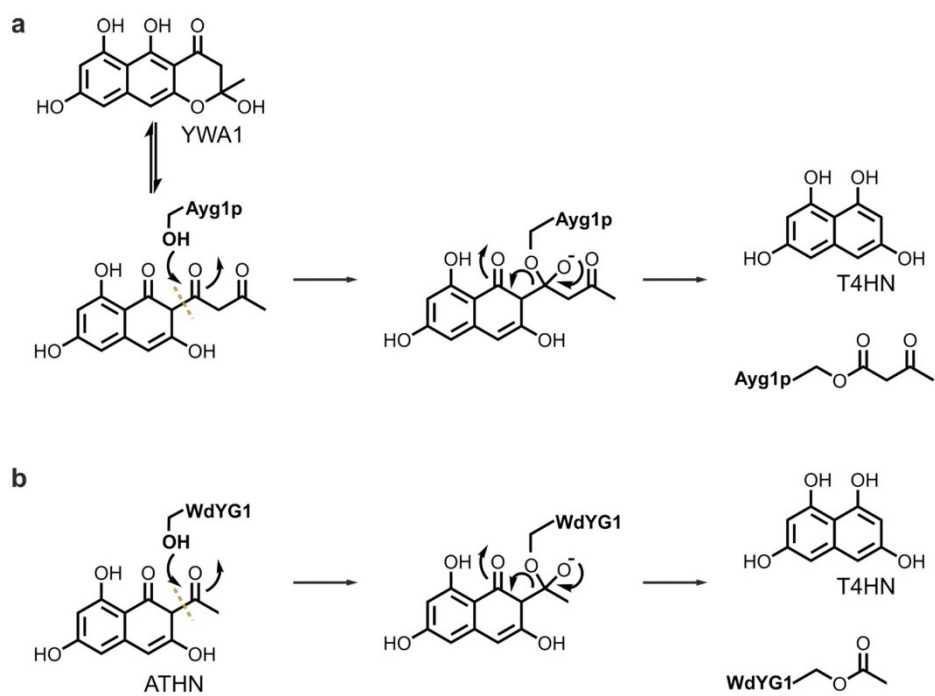
Supplementary Figure 7. Classical molecular dynamics simulations of intermediates INT1, INT2, and INT3, modeled in the active site of AntI. **(a)** RMSD for protein backbone atoms. **(b)** Distance between the oxanion hole and backbone Leu174/Phe246. **(c)** Open and closed conformations of Tyr20 observed during the MD simulations. **(d)** Dynamics of the Tyr20 C γ -C β -C α -C diedral angle during the MD simulations. A diedral angle $> 100^\circ$ / $< 100^\circ$ corresponds to the open/closed conformations, respectively. **(e)** RMSF of protein residues (red/purple/blue lines), and the ligand, shown as a dot at residue index zero.



Supplementary Figure 8. Biosynthesis gene clusters related to *ant* from *P. luminescens* from other *Photorhabdus* strains (a) and other bacteria (b). Cyanobacteria are shown in green and other Gram-negative bacteria in red.

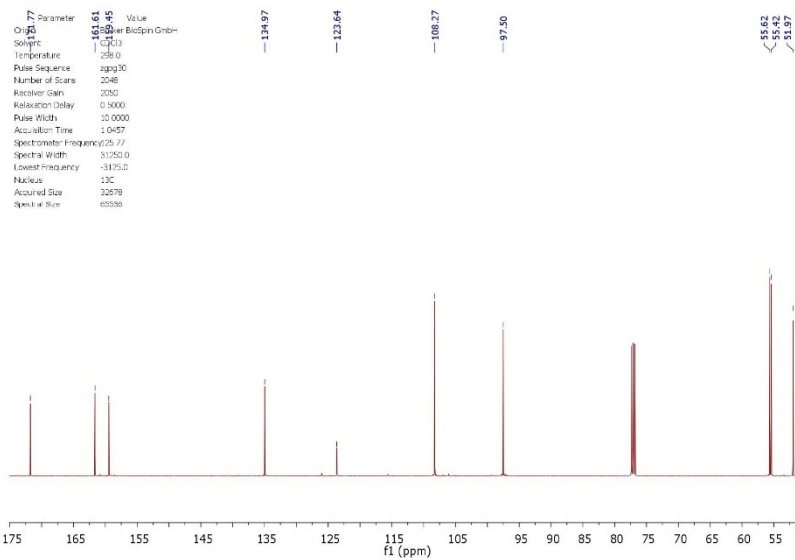
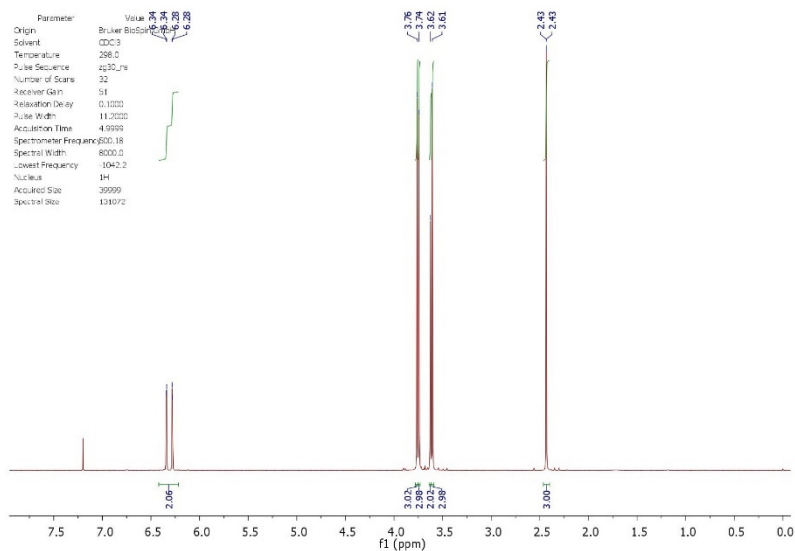
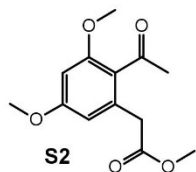


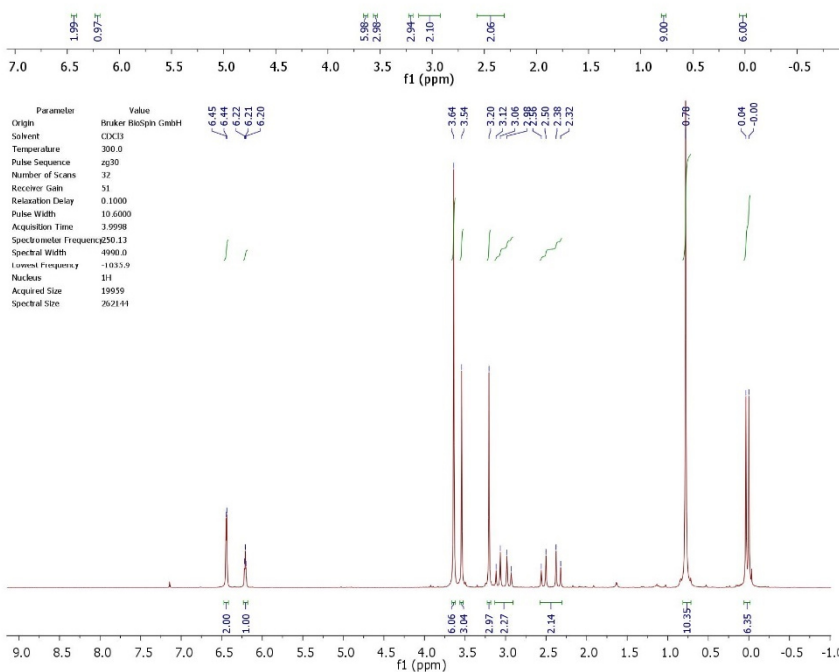
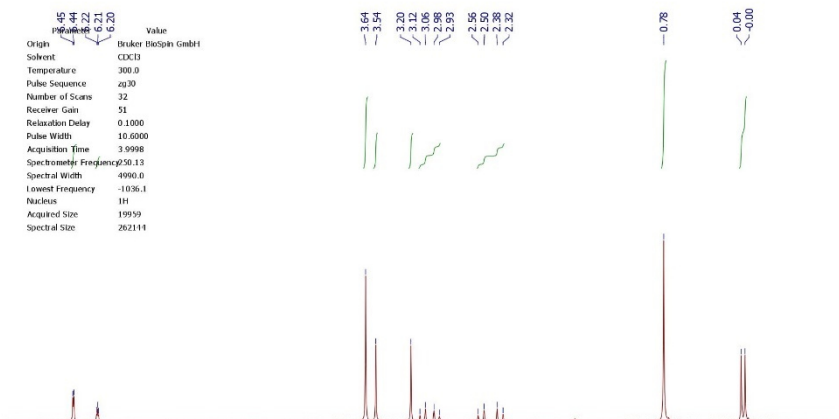
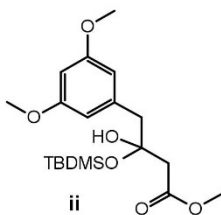
Supplementary Figure 9. AntI adopts α/β hydrolase fold (N-terminal domain in dark grey; C-terminal domain in grey). Helices and β -strands are represented as cylinders and arrows respectively. The nucleophilic elbow and the catalytic residues are highlighted. (a) Folding topology of AntI. (b) Multiple protein sequence alignment of AntI with AYG1p from *Aspergillus fumigatus*³⁴ and WdYG1 from *Exophiala dermatitidis*³⁵ using Clustal Omega.

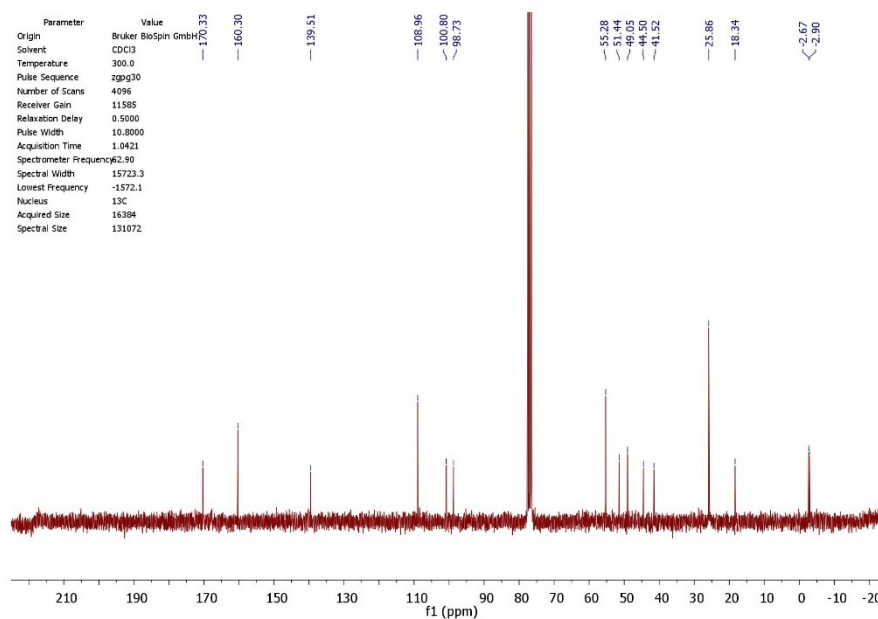


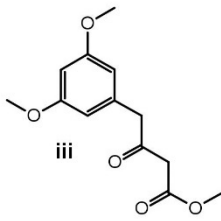
Supplementary Figure 10. Proposed polyketide shortening mechanism catalysed by Ayg1p (**a**) and WdYG1 (**b**) from *Aspergillus fumigatus* and *Exophiala dermatitidis*, respectively.

NMR spectra

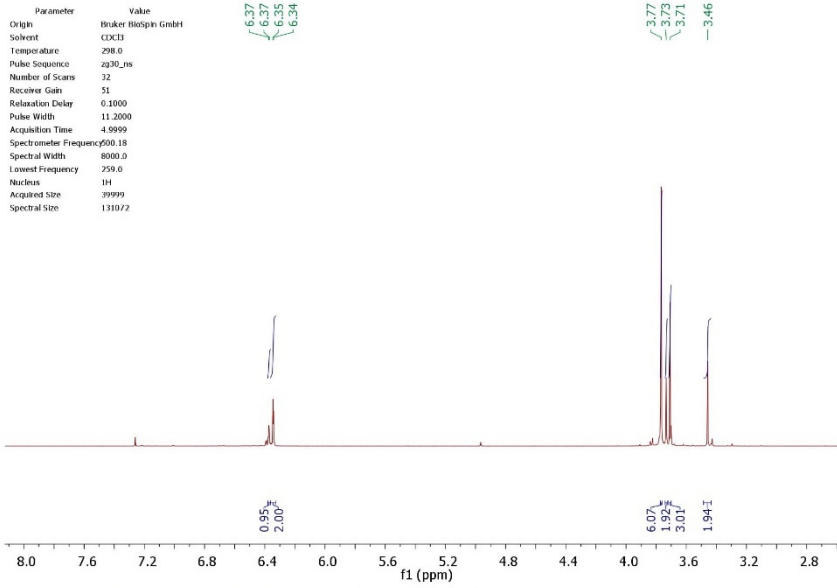




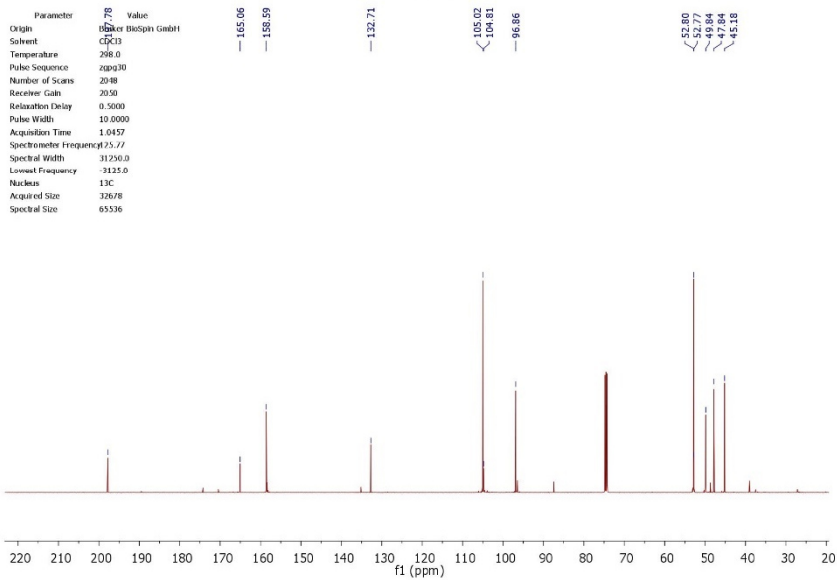


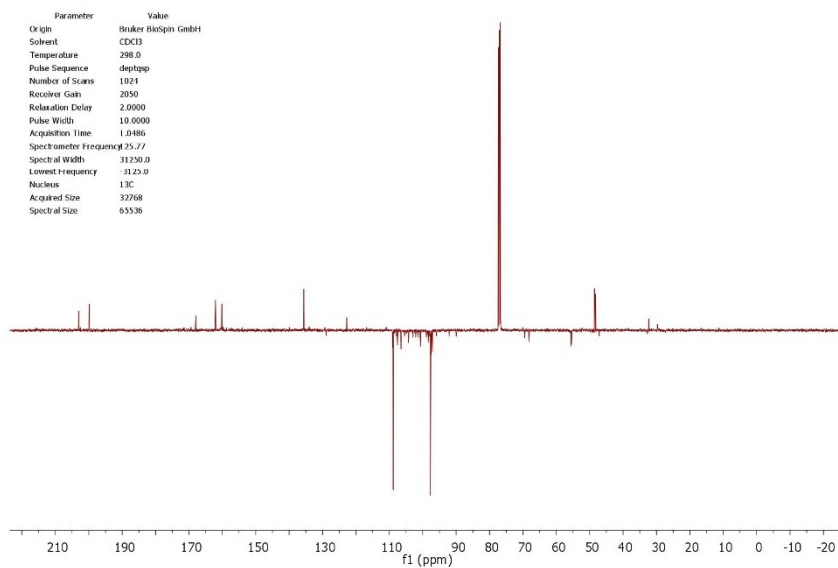
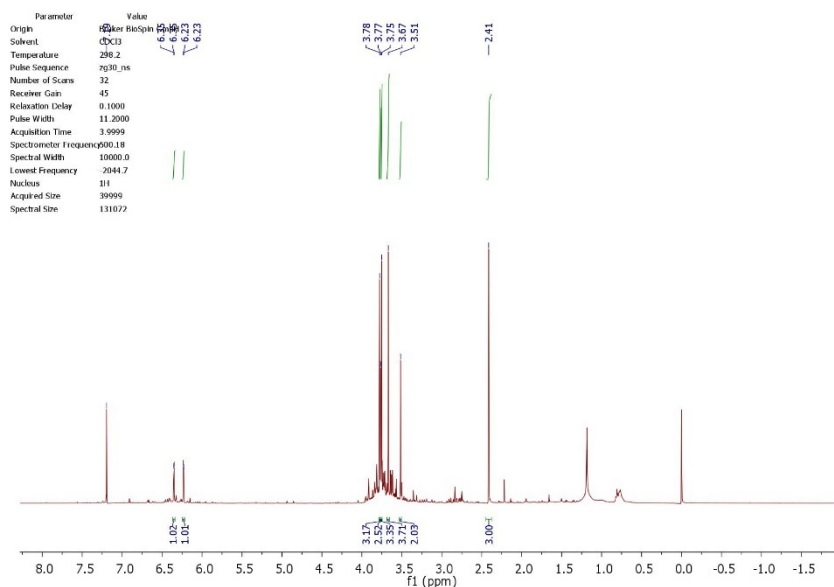
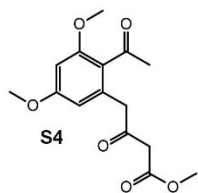


Parameter	Value
Origin	Bruker Biospin GmbH
Solvent	CDCl3
Temperature	298.0
Pulse Sequence	zg30_mh
Number of Scans	32
Receiver Gain	51
Relaxation Delay	0.1000
Pulse Width	11.2000
Acquisition Time	4.9999
Spectrometer Frequency	500.18
Spectral Width	8060.0
Lowest Frequency	259.0
Nucleus	¹ H
Acquired Size	39999
Spectral Size	131072



Parameter	Value
Origin	Bruker Biospin GmbH
Solvent	CDCl3
Temperature	298.0
Pulse Sequence	zgpg30
Number of Scans	2048
Receiver Gain	3050
Relaxation Delay	0.5000
Pulse Width	10.0000
Acquisition Time	1.0457
Spectrometer Frequency	125.77
Spectral Width	31230.0
Lowest Frequency	3125.0
Nucleus	¹³ C
Acquired Size	32678
Spectral Size	65536





References

1. Gaitatzis, N., Hans, A., Müller, R. & Beyer, S. The *mtaA* Gene of the Myxothiazol Biosynthetic Gene Cluster from *Stigmatella aurantiaca* DW4/3-1 Encodes a Phosphopantetheinyl Transferase that Activates Polyketide Synthases and Polypeptide Synthetases. **129**, 119–124 (2001).
2. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**, 105–124 (1993).
3. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132 (2010).
4. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr., A, Found. Crystallogr.* **64**, 112–122 (2008).
5. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
6. Cowtan, K. *Joint CCP4 ESF-EACBM Newsletter on Protein Crystallography (Daresbury Laboratory, Warrington, UK)*. **31**, 34–38 (1994).
7. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367 (2011).
8. Turk, D. MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallogr D Biol Crystallogr* **69**, 1342–1357 (2013).
9. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
10. Itoh, T. *et al.* Actinorhodin biosynthesis: structural requirements for post-PKS tailoring intermediates revealed by functional analysis of ActVI-ORF1 reductase. *Biochemistry* **46**, 8181–8188 (2007).
11. Brachmann, A. O. *et al.* A Type II Polyketide Synthase is Responsible for Anthraquinone Biosynthesis in *Photorhabdus luminescens*. *ChemBioChem* **8**, 1721–1728 (2007).
12. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
13. Huang, J. & MacKerell, A. D., Jr. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
14. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **91**, NA–NA (2009).
15. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
16. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–8–27–8 (1996).
17. Becke, A. D. Density - functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **98**, 5648–5652 (1993).
18. Lee, C., Yang, W. & Parr, R. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev., B Condens. Matter* **37**, 785–789 (1988).
19. Schäfer, A., Horn, H. & Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *The Journal of Chemical Physics* **97**, 2571–2577 (1992).
20. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **132**, 154104–20 (2010).

21. Klamt, A. & Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* 799–805 (1993). doi:10.1039/P29930000799
22. Plessow, P. Reaction Path Optimization without NEB Springs or Interpolation Algorithms. *J. Chem. Theory Comput.* **9**, 1305–1310 (2013).
23. E, W., Ren, W. & Vanden-Eijnden, E. String method for the study of rare events. *Phys. Rev., B Condens. Matter* **66**, 251–4 (2002).
24. Ahlrichs, R., Bär, M., Häser, M., Horn, H. & Kölmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chemical Physics Letters* **162**, 165–169 (1989).
25. Bauer, J. D., King, R. W. & Brady, S. F. Utahmycins a and B, azaquinones produced by an environmental DNA clone. *J. Nat. Prod.* **73**, 976–979 (2010).
26. Bartolomé, B., Jubete, Y., Martínez, E. & la Cruz, de, F. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**, 75–78 (1991).
27. Philippe, N., Alcaraz, J.-P., Coursange, E., Geiselmann, J. & Schneider, D. Improvement of pCVD442, a suicide plasmid for gene allele exchange in bacteria. *Plasmid* **51**, 246–255 (2004).
28. Grant, S. G., Jessee, J., Bloom, F. R. & Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4645–4649 (1990).
29. Simon, L. D., Randolph, B., Irwin, N. & Binkowski, G. Stabilization of proteins by a bacteriophage T4 gene cloned in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2059–2062 (1983).
30. Duchaud, E. *et al.* The genome sequence of the entomopathogenic bacterium *Photobacterium luminescens*. *Nat Biotechnol* **21**, 1307–1313 (2003).
31. Bode, H. B. *et al.* Determination of the absolute configuration of peptide natural products by using stable isotope labeling and mass spectrometry. *Chem. Eur. J.* **18**, 2342–2348 (2012).
32. Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chem. Int. Ed.* **48**, 4688–4716 (2009).
33. Hertweck, C., Luzhetskyy, A., Rebets, Y. & Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* **24**, 162 (2007).
34. Fujii, I. *et al.* Hydrolytic polyketide shortening by *ayg1p*, a novel enzyme involved in fungal melanin biosynthesis. *Journal of Biological Chemistry* **279**, 44613–44620 (2004).
35. Wheeler, M. H. *et al.* New Biosynthetic Step in the Melanin Pathway of *Wangiella (Exophiala) dermatitidis*: Evidence for 2-Acetyl-1,3,6,8-Tetrahydroxynaphthalene as a Novel Precursor. *Eukaryotic Cell* **7**, 1699–1711 (2008).

12.6 Supplemental information: Tracing the Full-length APE

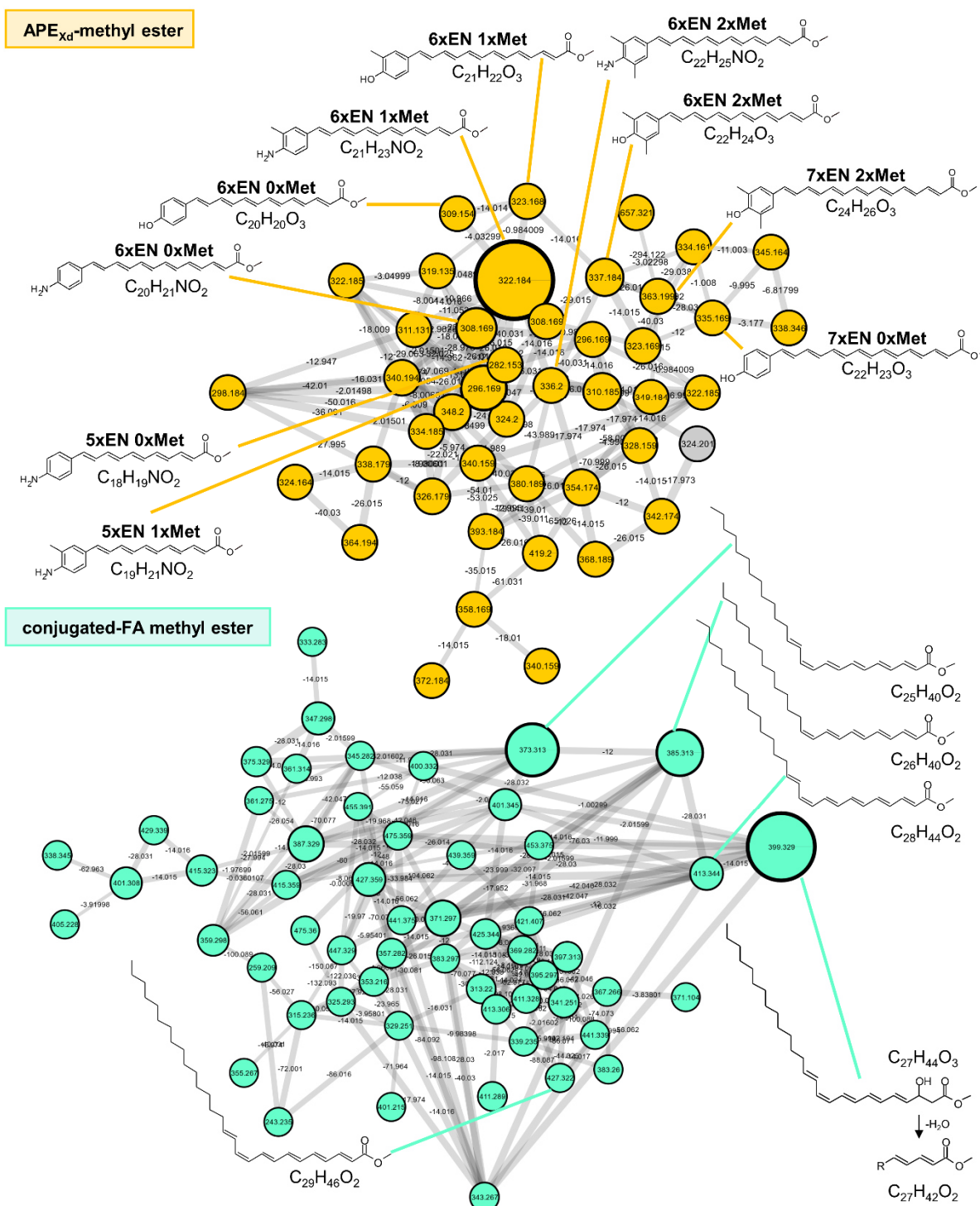


Figure S1. Network analysis of APE_{Xd}- and conjugated-FA methyl esters. Samples from size exclusion chromatography fractions were analyzed through HPLC-UV/MS. The resulting data were subjected to the GNPS-platform^[160] to create an MS²-based network. Samples were grouped manually according to their APE- and conjugated-FA- characteristic absorbance (8.1) in HPLC-UV/MS-analysis: F70-118 = APE_{Xd}-methyl ester, yellow; F17-33 = conjugated-FA methyl ester, turquoise; F1-16 = control fractions (without obvious absorbance of APE_{Xd}- or conjugated-FA-methyl ester), grey. Nodes are marked with the parent mass; the size of the node corresponds to the respective intensity of the parent mass in MS¹. Edges are marked with the delta Da between two MS¹-values; the edge thickness is cosine dependent. See also 8.2 for further MS- and HPLC-UV analysis.

Table S1. Assigned APE_{Xd}- and conjugated-FA-methyl ester. *If two signals appeared during HPLC-UV/MS analysis, the stronger signal is marked in bold here. For corresponding network analysis and survey view data of the selected analysis, see Figure S1 and Figure 30.

compound	sum formula [M]	theoretical mass [m/z]	detected mass [m/z]	Δppm	Rt
hydroxy-APE _{Xd} -methyl ester					
6xEN 0xMet	C ₂₀ H ₂₀ O ₃	309.1485	309.1491	2	12.3
6xEN 1xMet	C ₂₁ H ₂₂ O ₃	323.1642	323.1651	3	11.73, 12.65*
6xEN 2xMet	C ₂₂ H ₂₄ O ₃	337.1798	337.1798	0	13.10
7xEN 0xMet	C ₂₂ H ₂₂ O ₃	335.1642	335.1638	1	13.15, 14.86*
7xEN 2xMet	C ₂₄ H ₂₆ O ₃	363.1915	363.1942	4	13.8
amino-APE _{Xd} -methyl ester					
5xEN 0xMet	C ₁₈ H ₁₉ NO ₂	282.1489	282.1481	3	9.67
5xEN 1xMet	C ₁₉ H ₂₁ NO ₂	296.1645	296.1645	0	11.38
5xEN 2xMet	C ₂₀ H ₂₃ NO ₂	310.1802	310.1803	0	11.94
6xEN 0xMet	C ₂₀ H ₂₁ NO ₂	308.1652	308.1645	2	10.59, 11.29*
6xEN 1xMet	C ₂₁ H ₂₃ NO ₂	322.1802	322.1802	0	12.65
6xEN 2xMet	C ₂₂ H ₂₅ NO ₂	336.1960	336.1960	1	12.73
conjugated-FA					
373	C ₂₅ H ₄₀ O ₂	373.3101	373.3082	5	15.25
385	C ₂₆ H ₄₀ O ₂	385.3101	385.3080	5	15.35
399	C ₂₇ H ₄₂ O ₂	399.3258	399.3237	5	15.46
413	C ₂₈ H ₄₄ O ₂	413.3414	413.3393	5	15.87
427	C ₂₉ H ₄₆ O ₂	427.3571	427.3548	5	16.19

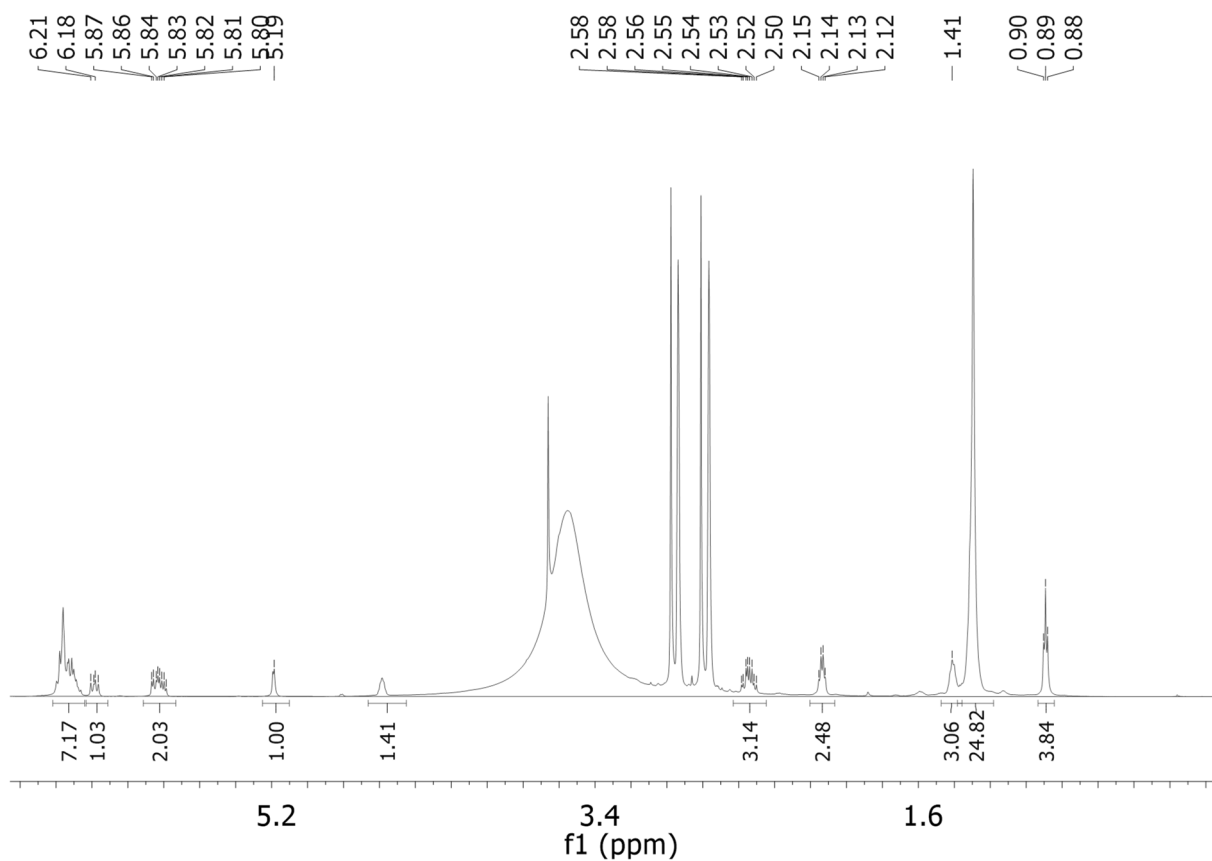


Figure S2. ^1H NMR (500 MHz) spectrum for conjugated-FA 399 in methanol- d_4 .

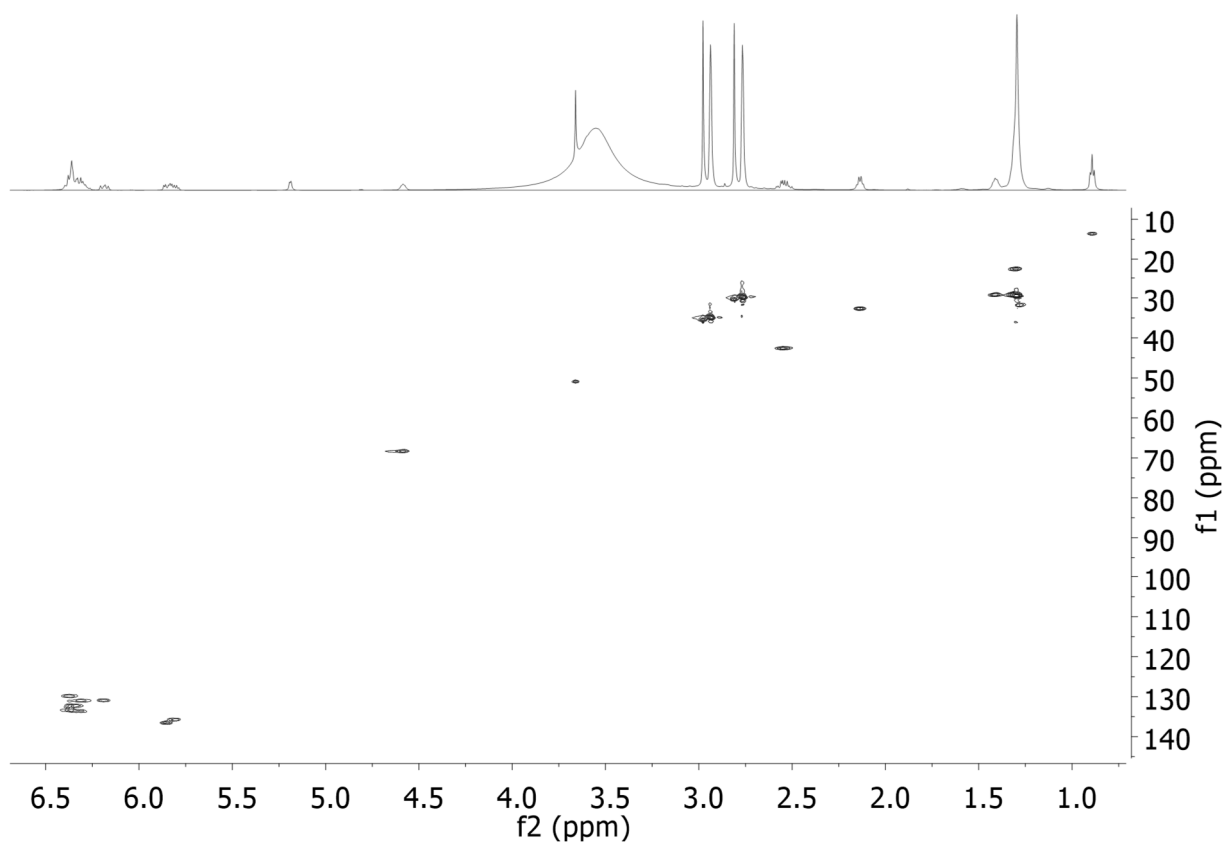


Figure S3. HSQC (500 MHz) spectrum for conjugated-FA 399 in methanol- d_4 .

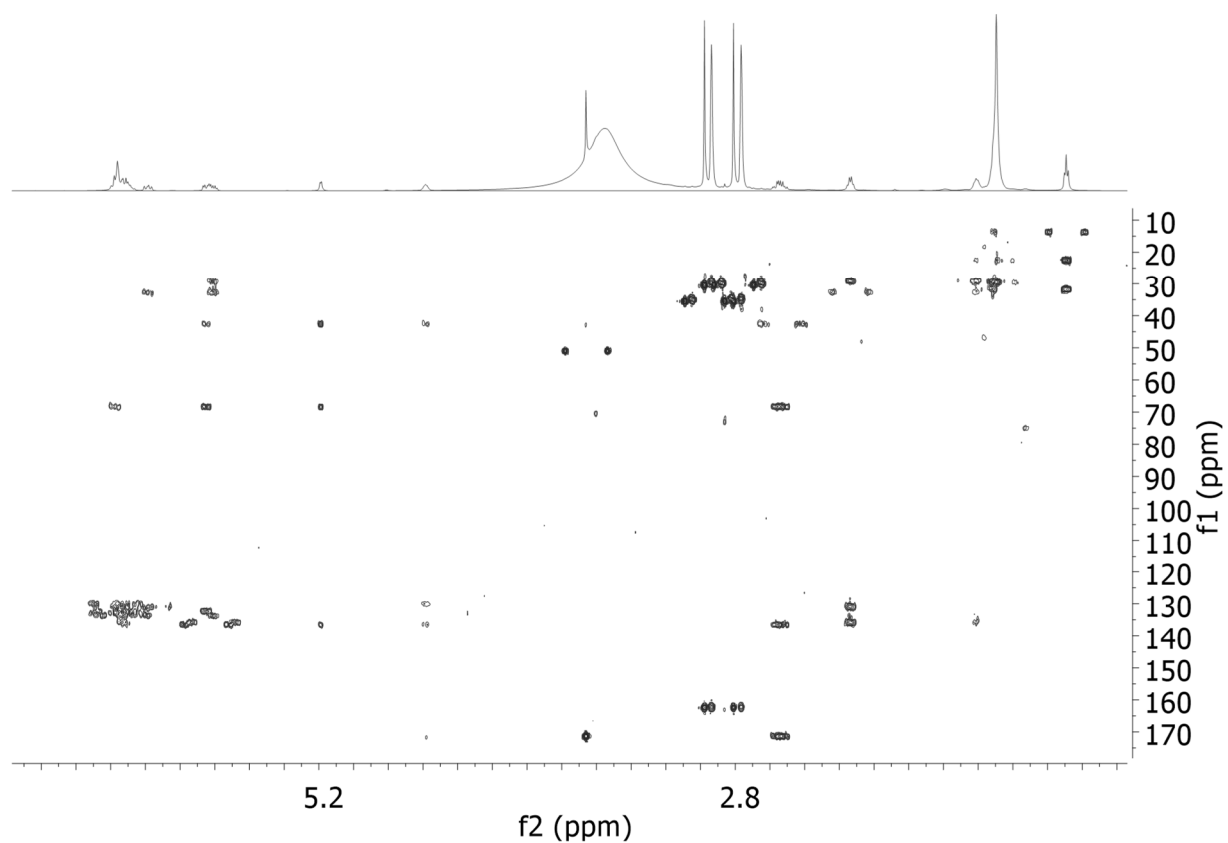


Figure S4. HMBC (500 MHz) spectrum for conjugated-FA 399 in methanol- d_4 .

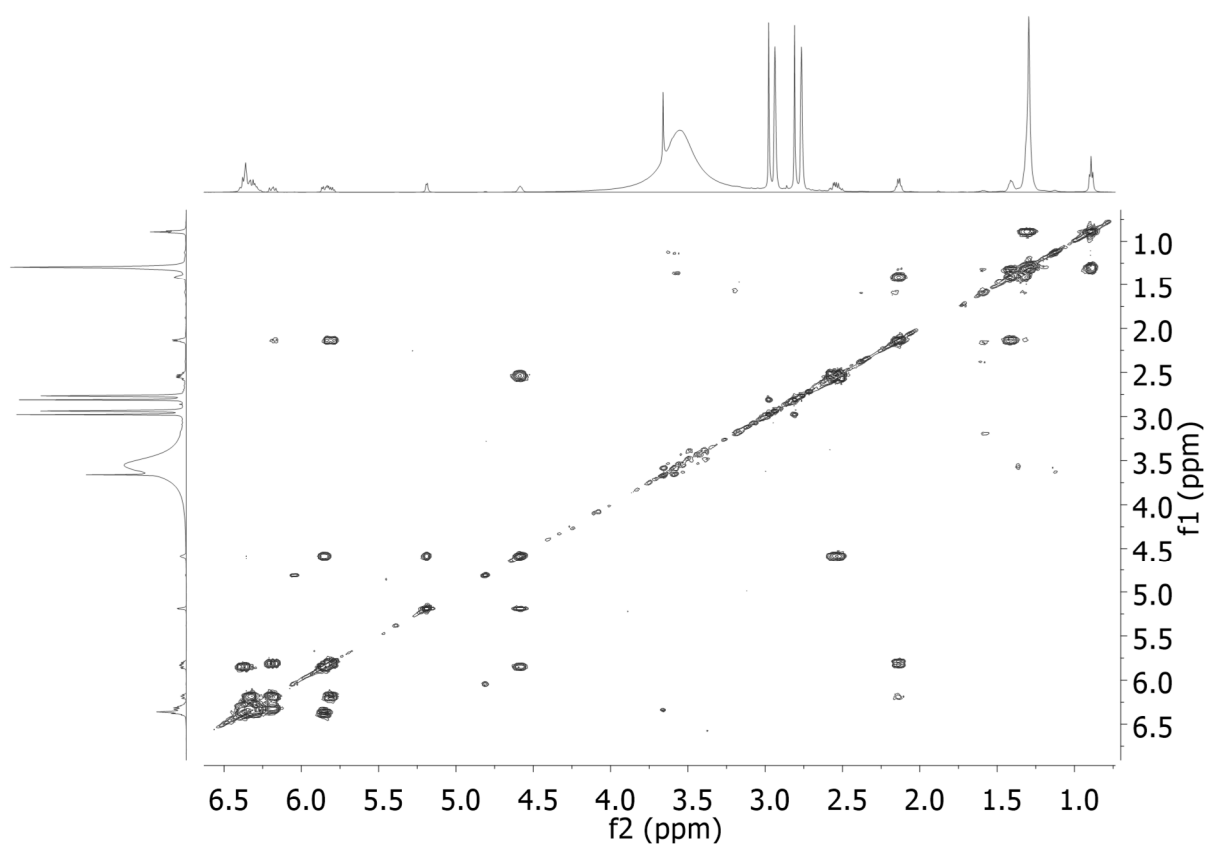


Figure S5. ^1H - ^1H COSY (500 MHz) spectrum for conjugated-FA 399 in methanol- d_4 .

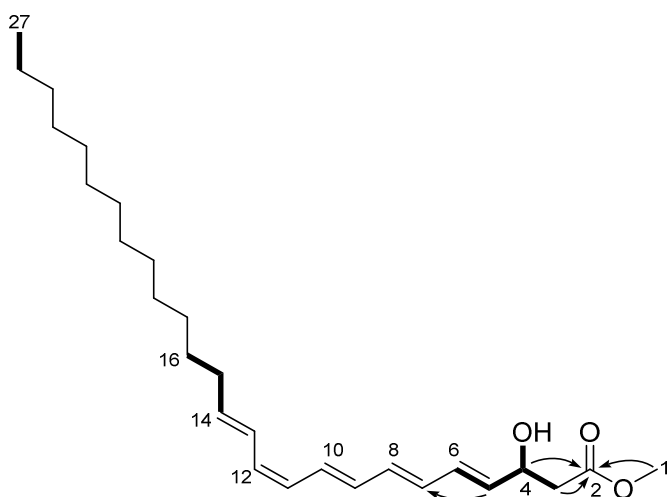


Figure S6. 2D NMR correlations of conjugated-FA 399.

Table S2. ^1H (500 MHz) and ^{13}C NMR data assignments for conjugated-FA 399 in methanol- d_4 .

No.	δ_{H} (mult., J)	δ_{C}
1	3.66 (s)	51.0, CH ₃
2	-	171.6, C
3	2.54 (dq, 14.8, 6.7)	42.6, CH ₂
4	4.59 (m)	68.4, CH
5	6.33 (overlapped)	132-133, CH
6	6.33 (overlapped)	132-133, CH
7	6.33 (overlapped)	132-133, CH
8	6.33 (overlapped)	132-133, CH
9	6.33 (overlapped)	132-133, CH
10	6.33 (overlapped)	132-133, CH
11	5.85 (brdd, 15.0, 5.6)	136.5, CH
12	5.81 (brdd, 14.8, 7.3)	135.8, CH
13	6.33 (overlapped)	132-133, CH
14	6.19 (dd, 14.8, 10.2)	132-133, CH
15	2.14 (dd, 13.9, 6.8)	32.7, CH ₂
16	1.41 (m)	29-30, CH ₂
17	1.30 (overlapped)	29-30, CH ₂
18	1.30 (overlapped)	29-30, CH ₂
19	1.30 (overlapped)	29-30, CH ₂
20	1.30 (overlapped)	29-30, CH ₂
21	1.30 (overlapped)	29-30, CH ₂
22	1.30 (overlapped)	29-30, CH ₂
23	1.30 (overlapped)	29-30, CH ₂
24	1.30 (overlapped)	29-30, CH ₂
25	1.30 (overlapped)	29-30, CH ₂
26	1.30 (overlapped)	29-30, CH ₂
27	0.89 (t, 6.7)	14.1, CH ₃
4-OH	5.19 (d, 4.4)	-

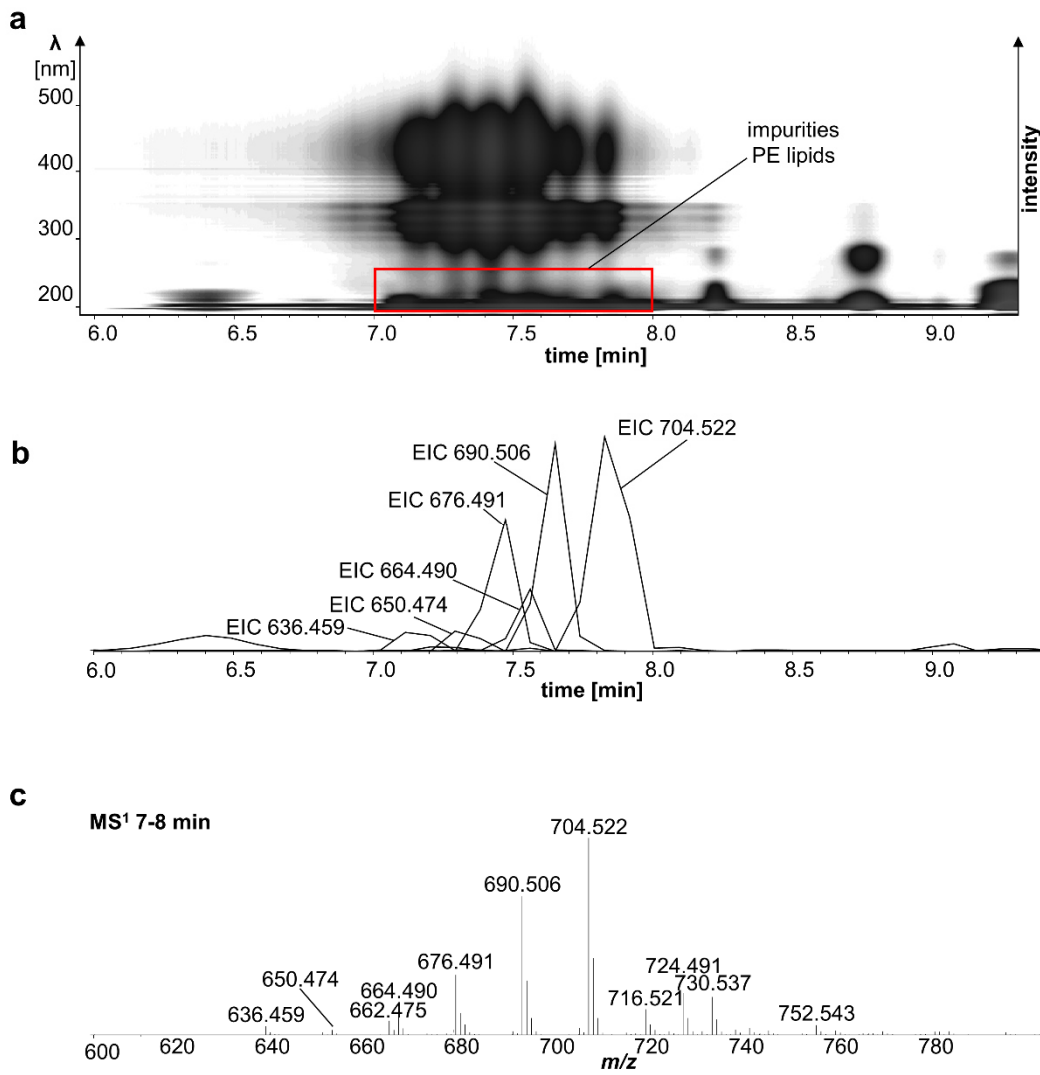


Figure S7. HPLC-UV/MS data of observed lipid-impurities.

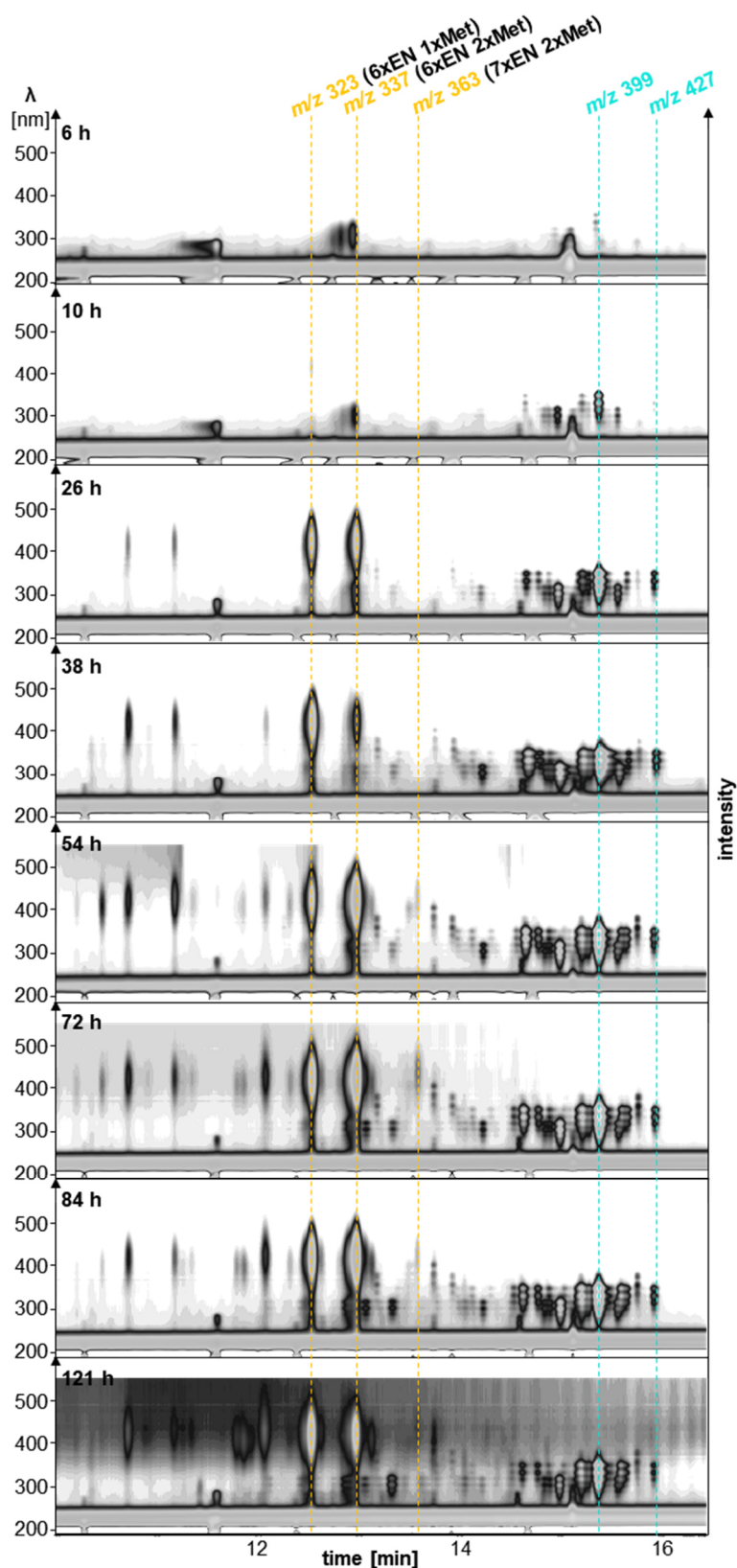


Figure S8. HPLC-UV/MS time course analysis (survey view) of the production of APE_{Xd}-methyl esters of the APE-producer.

13 List of Publications

Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photorhabdus luminescens*

Qiuqin Zhou, Alois Bräuer, Hélène Adihou, Max Schmalhofer, Patricia Saura, Gina L. C. Grammbitter, Ville R. I. Kaila, Michael Groll, and Helge B. Bode

Chemical Science **2019** 10 (25), 6341-6349

DOI: 10.1039/c9sc00749k

An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments

Gina L. C. Grammbitter, Maximilian Schmalhofer, Kudratullah Karimi, Yi-Ming Shi, Tim A. Schöner, Nicholas J. Tobias, Nina Morgner, Michael Groll, and Helge B. Bode

Journal of the American Chemical Society **2019** 141 (42), 16615-16623

DOI: 10.1021/jacs.8b10776

Fabclavine biosynthesis in *X. szentirmaii*: shortened derivatives and characterization of the thioester reductase FclG and the condensation domain-like protein FclL

Sebastian L. Wenski, Diana Kolbert, Gina L. C. Grammbitter, and Helge B. Bode

Journal of Industrial Microbiology and Biotechnology **2019** 46 (3-4), 565-572.

DOI: 10.1007/s10295-018-02124-8

14 Record of Conferences

09/2016 – VAAM workshop Freiburg – poster presentation

09/2017 – VAAM workshop Tübingen – poster presentation (poster prize)

04/2018 – VAAM Wolfsburg – poster presentation

07/2018 – ASP Kentucky, Lexington – oral presentation

09/2018 – ECNP Frankfurt am Main – poster presentation

09/2019 – VAAM workshop Jena – poster presentation

VAAM = Verein für Allgemeine und Angewandte und Mikrobiologie; ASP = American Society of Pharmacognosy; ECNP = European Conference on Natural Products

15 Erklärung und Versicherung

Erklärung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den

(Unterschrift)

Versicherung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation über

„Characterization of PKS II systems from entomopathogenic bacteria“

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.

Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den

(Unterschrift)