

Dissertation  
zur Erlangung des Doktorgrades  
der Naturwissenschaften

## **Structure and Function of Megasyntases**

Vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie  
der Johann Wolfgang Goethe-Universität  
in Frankfurt am Main

von

**Mirko Joppe**

aus Bad-Homburg v.d. Höhe

Frankfurt am Main (2021)

(D30)

Vom Fachbereich Biochemie, Chemie und Pharmazie der Johann Wolfgang  
Goethe-Universität als Dissertation angenommen

Dekan : Prof. Dr. Clemens Glaubitz

Gutachter: Prof. Dr. Martin Grininger  
Prof. Dr. Helge B. Bode

Datum der Disputation:

## Abbreviations

2D	two dimensional
3D	three dimensional
3-EP	3-ethyl phenol
3-MP	3-methyl phenol
3-PrP	3-propyl phenol
6-ESA	6-ethyl salicylic acid
6-MSA	6-methyl salicylic acid
6-PrSA	6-propyl salicylic acid
6-PSA	6-pentyl salicylic acid
<i>A. terreus</i>	<i>Aspergillus terreus</i>
<i>A. aculeatus</i>	<i>Aspergillus aculeatus</i>
Å	Ångström
ACP	acyl carrier protein
amp	ampicillin
APS	ammonium persulfate
AT	acyltransferase
ATP	adenosine triphosphate
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
bla	β-lactamase
bp	base pair
BIS-TRIS	2,2-Bis-(hydroxymethyl)-2,2',2''-nitrilotriethanol
BSA	bovine serum albumin
ChIB2	discrete acyl carrier protein
ChIB3	KASIII-like protein
CoA	coenzyme A
COSY	2D NMR correlation spectroscopy
CV	column volume
DCM	dichloromethane
DEB	6-deoxyerythronolide B
DEBS	6-deoxyerythronolide B synthase
DH	dehydratase
DM1-4	dimerization module 1-4
DMAP	4-dimethylaminopyridine
DMPK	drug metabolism and pharmacokinetics

DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	exempli gratia, Latin for “for example”
EDC×HCl	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide-hydrochloride
EDTA	ethylenediaminetetraacetic acid
EIC	extracted ion chromatogram
EM	electron microscopy
ER	enoylreductase
ESI	electron spray ionization
ET	electron tomography
EtOAc	ethyl acetate
FA	fatty acid
FAS	fatty acid synthase
FDA	food and drug administration
F-Mal-CoA	2-fluoromalonyl coenzyme A
F-MM-CoA	2-fluoro-2-methylmalonyl coenzyme A
G418	geneticin
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum correlation
i.e.	<i>id est</i> , that is
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
kan	kanamycin
kanMX4	Gene for kanamycin resistance
KASIII	3-oxoacyl-(acyl-carrier protein) synthase
KR	ketoreductase
KS	ketosynthase
lacI	Lac repressor protein
LB	lysogeny broth
LD	linker domain
LDD	loading didomain
Mal-CoA	malonyl coenzyme A

MALS	multiangle light scattering
MAT	malonyl/acetyl transferase
MBP	maltose binding protein
mFAS	murine fatty acid synthase
MM-CoA	2-methylmalonyl coenzyme A
MPT	malonyl palmitoyl transferase
MS	mass spectrometry
MSAS	methyl salicylic acid synthase
NADH	nicotinamide adenine dinucleodide
NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhancement spectroscopy
<i>Npt</i>	4'-phosphopantetheinyl transferase from <i>Streptomyces platensis</i>
NRPS	non-ribosomal peptide synthase
NTA	nitrilotriacetic acid
OD	optical density
<i>P. patulum</i>	<i>Penicillium patulum</i>
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PenPaMSAS	MSAS from the organism <i>penicillium patulum</i>
PIKS	pikromycin synthase
PKS	polyketide synthase
PPT	phosphopantetheinyl transferase of yeast FAS
PPTase	phosphopantetheinyl transferase
Pr-TAL	4-hydroxy-6-propyl-2-pyrone
P-TAL	4-hydroxy-6-pentyl-2-pyrone
PVDF	polyvinylidene difluoride
rcf	relative centrifugal force
RFU	relative fluorescence unit
<i>S. antibioticus</i>	<i>Streptomyces antibioticus</i>
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>

<i>S. erythraea</i>	<i>Saccharopolyspora erythraea</i>
SanChIB3	KASIII-like protein from <i>Streptomyces antibioticus</i>
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
select-fluor	1-chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octan-bis-(tetrafluoroborate)
SerChIB3	KASIII-like protein from <i>Saccharopolyspora erythraea</i>
SerMSAS	MSAS from the organism <i>Saccharopolyspora erythraea</i>
<i>Sfp</i>	4'-phosphopantetheinyl transferase from <i>Bacillus subtilis</i>
spc	spectinomycin
TAL	triacetic acid lactone, 4-hydroxy-6-methyl-2-pyrone
TB	terrific broth
TE	thioesterase
TEMED	NNN'N'-tetra-methylethylenediamine
THF	tetrahydrofuran
TKL	triketide lactone
TM	trimerization module
TPP	thiamine pyrophosphate
TRIS	Tris-(hydroxymethyl)-aminomethan
TSA	thermal shift assay
U	Unit, defined as the incorporation of 1 $\mu\text{mol}$ malonyl-CoA per minute
UFF	universal force field
UV	ultraviolet
WT	wild type
x-ray	electromagnetic radiation
$\alpha$ -KGDH	$\alpha$ -ketoglutarate dehydrogenase complex
$\psi$ KR	pseudo-ketoreductase
$\psi$ ME	pseudo-methyltransferase

# Zusammenfassung

## Naturstoffe und ihre Fabriken: die Megasyntasen

Naturstoffe sind aufgrund ihrer biologischen Funktionen zentrale Moleküle des Lebens. Sie sind biotechnologisch und pharmakologisch äußerst relevant und werden als Medikamente sowie für die Medikamentenentwicklung genutzt.<sup>1</sup> Fettsäuren und Polyketide sind zwei Hauptklassen der Naturstoffe, welche in der Natur aus einem gemeinsamen Pool von einfachen Substraten synthetisiert werden. Die Biosynthese wird oft von kovalent verknüpften Enzymen übernommen, den sogenannten Megasyntasen. Diese katalysieren stereo- und regioselektive chemische Reaktionen, wobei das Substrat und die Zwischenprodukte immer an das Protein gebunden sind. Zusätzlich weisen die einzelnen Enzyme häufig eine große Substratspezifität auf. Diese Mechanismen verhindern unerwünschte Nebenreaktionen und erklären die hohe Effizienz der Megasyntasen. Trotz der unterschiedlichen chemischen Strukturen und damit unterschiedlichen biologischen Funktionen der Fettsäuren und Polyketiden, folgt die Biosynthese der Fettsäure-Syntasen (FASs) und Polyketid-Syntasen (PKSs) ähnlichen Mechanismen.

## Die Naturstoffsynthese von Megasyntasen

Während dieser Biosynthese übernimmt die Acyltransferase (AT) eine zentrale Rolle, indem sie Starter- sowie Verlängerungssubstrate selektiert, welche in die Naturstoffe eingebaut werden. Dabei werden die selektierten Substrate auf das Acyl-Carrier-Protein (ACP) übertragen. Dieses kann jedoch erst nach der Aktivierung durch die Phosphopantethein-Transferase (PPT), Substrate sowie die Zwischenprodukte der Synthese zu den jeweiligen Enzymen transferieren. Die Verlängerung des Startersubstrates mit dem Verlängerungssubstrat wird durch die Ketosynthase (KS) katalysiert, wobei ein ACP-gebundener  $\beta$ -Ketoester entsteht. Dieser wird in Abhängigkeit der zu betrachtenden Megasyntase durch prozessierende Domänen, wie die Ketoreduktase (KR), die Dehydratase (DH) und die Enoylreduktase (ER), sukzessiv reduziert. Diese Reaktionskaskade wird wiederholt ausgeführt, um das Intermediat weiter zu verlängern und zu funktionalisieren. Anschließend wird das Endprodukt von der Megasyntase, typischerweise durch eine Thioesterase (TE), freigesetzt. Die Megasyntasen werden in iterative und modulare Systeme unterteilt. Während die Enzyme der iterativen Systeme mehrere Reaktionen katalysieren können, werden die Enzyme der modularen Systeme typischerweise nur einmalig pro synthetisiertem Molekül eingesetzt. Wie der Name impliziert, sind letztere in Modulen aufgebaut, welche alle notwendigen Enzyme für die Kettenverlängerung sowie zusätzlich variierenden Prozessierungsdomänen besitzen. Nach einer Kettenverlängerung sowie Funktionalisierung wird das Zwischenprodukt anschließend auf das nächste Modul übertragen. Dementsprechend kann die Biosynthese der modularen Systeme als

vektoriell beschrieben werden. Trotz ihrer Unterschiede sind die beschriebenen Megasyntasen evolutionär verwandt und neue Erkenntnisse einer Klasse können aufgrund ihrer strukturellen und funktionellen Gemeinsamkeiten oft auf die andere Klasse übertragen werden.

Das enorme Potenzial dieser Enzyme spiegelt sich auch in der großen Anzahl an Studien auf diesem Gebiet wider. Neben der strukturellen und mechanistischen Aufklärung wurden die Megasyntasen für die Biosynthese von relevanten Naturstoffen und das Protein-Engineering eingesetzt, mit dem Ziel, die enzymatische Effizienz zu verbessern sowie das breite Produktspektrum durch gezielte Derivatisierung zu erweitern. Dabei ist die Biosynthese der Fettsäuren und Polyketiden aufgrund der Generierung von Kohlenstoffverbindungen sowie funktionellen Gruppen besonders relevant. Insbesondere die außergewöhnliche Anzahl an chiralen Zentren innerhalb der Polyketide zeigt, warum die Biosynthese dieser Verbindungen eine ökonomische Alternative zur anspruchsvollen chemischen Totalsynthese ist.

### Derivatisierung von Naturstoffen durch Proteinengineering

Damit die synthetisierten Verbindungen industriell und vor allem pharmazeutisch relevant sind, ist häufig die Derivatisierung dieser Naturstoffe erforderlich.<sup>1</sup> Die Fettsäure- und Polyketid-Derivate können dabei direkt durch Biosynthese mit Hilfe von Megasyntasen erzeugt werden. Dies erfordert jedoch meist gezieltes Protein-Engineering aufgrund der intrinsischen Substratspezifitäten der Enzyme. Allerdings führt das Protein-Engineering oft zu dem Verlust der enzymatischen Aktivität, da es an einem grundlegenden Verständnis der strukturellen und funktionellen Eigenschaften der Enzymklassen mangelt. Um Verbindungen durch die Biosynthese gezielt herstellen zu können, ist es daher notwendig, die Megasyntasen und ihre Enzyme in allen Details zu charakterisieren, um Richtlinien für effiziente Designs von veränderten Systemen aufzustellen. Aus dieser Motivation heraus wurden verschiedene Megasyntasen, darunter die FAS aus Hefe und Säugetieren, die Methylsalicylsäure-Synthase (MSAS) und die 6-Desoxyerythronolid-B-Synthase (DEBS), strukturell und funktionell charakterisiert. Das Ziel war es diese Enzyme für das Protein-Engineering zu nutzen, um die Biosynthese von neuen Naturstoffen zu ermöglichen.

### Eine schnelle Reinigungsstrategie für die Hefe-FAS

Die Charakterisierung von Megasyntasen *in vitro* ist im Vergleich weniger komplex und ermöglicht die Untersuchung von Strukturen sowie funktionellen Eigenschaften des gesamten Enzyms, aber auch von einzelnen Domänen. Auf der anderen Seite ist es oft eine große Herausforderung, optimale Bedingungen für die Herstellung reproduzierbarer Protein-Proben in hoher Qualität zu finden. In der vorliegenden Doktorarbeit wurde eine vektorbasierte Expressions- und Reinigungsstrategie für die



Hefe-FAS etabliert, welche gezielte Veränderungen des Enzyms vereinfacht und daher besonders gut für Mutationsstudien geeignet ist. Nach der Expression konnte das Enzym mittels Affinitäts- und Größenausschlusschromatographie innerhalb von fünf Stunden gereinigt werden. Als Qualitätskontrollen wurden SEC, SDS-PAGE, TSA und vor allem NADPH-Verbrauch verwendet, welche die Eignung der angewandten Strategie demonstrierten, um reine FAS-Proben mit reproduzierbaren enzymatischen Aktivitäten zu erhalten. Obwohl die Verwendung eines Tags für die Affinitätschromatographie zugängliche Termini und damit ein strukturabhängiges Design erfordert, kann die Reinigungsstrategie in Kombination mit den Qualitätskontrollen als Leitfaden für andere makromolekulare Komplexe verwendet werden.

### Denaturierung von FAS an der Luft-Wasser-Grenzfläche während der Kryo-EM-Grid-Präparation

Die FAS-Probe konnte schließlich für die Strukturaufklärung mittels Kryo-Elektronenmikroskopie (EM) verwendet werden. Während der strukturellen Untersuchung wurde beobachtet, dass der größte Teil der FAS-Partikel teilweise beschädigt vorlag, was die hochauflösende Strukturaufklärung erschwerte. Die Qualität der FAS-Proben konnte durch die Veränderung der Reinigungsstrategie mit verschiedenen Puffer- oder Zellysemethoden nicht weiter verbessert werden. Das Problem wurde dementsprechend nicht bei der Reinigung, sondern bei der Probenvorbereitung für die Kryo-EM vermutet. Diese Idee wurde durch die Beobachtung von unbeschädigten Partikeln während der EM mittels negativer Färbemethode unterstützt. Aufgrund dessen wurden die Grids, welche für die Kryo-EM vorbereitet wurden, mittels Kryo-ET analysiert. Dabei wurde beobachtet, dass die meisten Partikel an die Wasser-Luft-Grenzfläche mit einer bevorzugten Orientierung adsorbieren. Die beschädigte Seite der FAS war immer in Richtung der Luft-Grenzfläche ausgerichtet und als Folge denaturiert oder teilweise entfaltet.

### Graphen-Grids verhindern die Denaturierung von FAS-Proteinen

Um die Beschädigung der FAS-Proben zu verhindern, wurde die Verwendung von Graphen-Grids etabliert, welche mit 1-Pyren-Carbonsäure hydrophilisiert wurden. Da die FAS-Partikel die Graphen-Wasser-Grenzfläche gegenüber der Wasser-Luft-Grenzfläche bevorzugten, konnte die Unversehrtheit der FAS-Probe, während der Grid-Präparation sichergestellt werden. Die Verwendung von Graphen-Grids könnte auch für andere makromolekulare Komplexe große Vorteile gegenüber herkömmlichen Grids aufweisen und damit die Strukturaufklärung mittels Kryo-EM voranbringen.

Diese Arbeit wurde in eLife publiziert: D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *eLife*, 8, e42747

### Neue strukturelle Erkenntnisse über die Hefe-FAS

Der große Erfolg der etablierten Methode konnte anhand der vollständigen Struktur der Hefe-FAS in hoher Auflösung (PDB: 6TA1, 3,1 Å Auflösung) demonstriert werden. Diese beinhaltet die PPT an der Außenseite der fassförmigen Struktur, deren Elektronendichte in früheren Studien aufgrund hoher Flexibilität nur schwach ausgeprägt war. Deshalb benötigte die vollständige Strukturklärung der Hefe-FAS in früheren Studien eine individuelle Strukturklärung der PPT, welche in diese Elektronendichte modelliert wurde. Die Positionierung der PPT früherer Studien konnte mithilfe der vollständigen Strukturaufklärung mittels Kryo-EM bestätigt werden. Zusätzlich wurde eine Phosphorylierungsstelle an Ser1440 sowie die Bindung des Co-Faktors NADPH an die KR gezeigt. Eine weitere Phosphorylierungsstelle wurde an Ser1827 vermutet, konnte anhand der Struktur aufgrund von zu niedriger lokaler Auflösung, jedoch nicht bestätigt werden.

Diese Arbeit wurde in *IUCrJ* publiziert: Joppe, M., D'Imprima, E., Salustros, N., Paithankar, K. S., Vonck, J., Grininger, M., & Kühlbrandt, W. (2020). The resolution revolution in cryoEM requires high-quality sample preparation: a rapid pipeline to a high-resolution map of yeast fatty acid synthase. *IUCrJ*, 7(2), 220-227.

### Charakterisierung von zwei Phosphorylierungsstellen der Hefe-FAS

Eine Mutationsstudie wurde durchgeführt um die Phosphorylierung und deren Relevanz an den Positionen Ser1440 sowie Ser1827 zu untersuchen. Beide Stellen wurden entweder zu Alanin- oder zu Glutaminsäure mutiert, wobei letztere die negative Ladung der Phosphatgruppe imitiert. Beide Mutationen an der Position Ser1827 haben die Geschwindigkeit der Fettsäuresynthese signifikant verringert, wobei die Glutaminsäure-Mutation einen größeren Einfluss hatte. Sofern eine Phosphorylierung an dieser Position vorliegt, könnte sie zur Drosselung der FAS verwendet werden, um den enormen Substratverbrauch zu bremsen. Die Mutation könnte dabei die Phosphopantetheinylierung des ACPs beeinträchtigen, da Ser1827 sich an der Außenseite der PPT-Domäne befindet. Im Gegensatz dazu hatten beide Mutationen an Ser1440 einen vernachlässigbaren Effekt auf die Fettsäuresynthese, weswegen die Phosphorylierung an dieser Position anhand der Analytik nicht bewertet werden konnte. Zusammengefasst lässt sich sagen, dass keine der beiden Mutationsstellen die angestrebte enzymatische Effizienz der Hefe-FAS deutlich verbessern konnte. Jedoch konnte gezeigt werden, dass die etablierte Reinigungsstrategie für das Protein-Engineering sowie die Charakterisierung von Mutanten sehr gut geeignet ist.

### Die Hefe-FAS ist auch ohne die strukturellen Domänen TM und DM2 aktiv

Zusätzlich wurden strukturelle Domänen der Hefe-FAS untersucht, um Richtlinien für das Protein-Engineering zu erstellen und Limitierungen des Enzyms zu überwinden.

Es wurde gezeigt, dass die Trimerisierungs- (TM) und Dimerisierungsdomäne (DM2) vom Enzym entfernt werden können, ohne die Assemblierung zu stören und dabei die Fettsäuresynthese zu beeinträchtigen. Damit sind diese Domänen potenziell für den Austausch mit zusätzlichen funktionalisierenden Domänen geeignet, um das Produktspektrum der FAS für biotechnologische Anwendungen zu erweitern.

Diese Arbeit wurde in *Scientific reports* publiziert: Fischer, M., Joppe, M., Mulinacci, B., Vollrath, R., Konstantinidis, K., Kötter, P., Ciccarelli, L., Vonck, J., Oesterhelt, D. & Grninger, M. (2020). Analysis of the co-translational assembly of the fungal fatty acid synthase (FAS). *Scientific reports*, 10(1), 1-13.

### 6-MSA ist eine Vorstufe vieler bioaktiver Verbindungen

Die MSAS hingegen produziert 6-MSA, eine Vorstufe für relevante bioaktive Verbindungen mit antibiotischen, antitumoralen und antibakteriellen Wirkungen. Zusätzlich kann 6-MSA durch eine Decarboxylase in 3-Methylphenol (3-MP) umgesetzt werden. Strukturell verwandte Verbindungen, wie 3-Ethylphenol (3-EP) und 3-Propylphenol (3-PrP), werden als Lockstoffe in Fallen gegen die Tsetsefliege eingesetzt, um deren Population zu kontrollieren. Die Fliege überträgt Parasiten der Gattung *Trypanosomen*, welche als Folge die Schlafkrankheit beim Menschen sowie die Krankheit Trypanosomiasis bei Tieren verursachen. Die Lockstoffe 3-EP und 3-PrP werden traditionell chemisch synthetisiert. Kürzlich wurde jedoch von der Biosynthese dieser Stoffe durch die MSAS in Kombination mit einer Decarboxylase berichtet. Dabei wurden nicht natürliche Startersubstrate wie Propionyl-CoA und Butyryl-CoA von einem Hefestamm bereitgestellt, aus denen die MSAS 6-Ethylsalicylsäure (6-ESA) sowie 6-Propylsalicylsäure (6-PrSA) synthetisieren konnte. Anschließend konnten diese durch Decarboxylierung in 3-EP und 3-PrP umgesetzt werden. Die Ausbeuten von 3-EP und 3-PrP waren im Vergleich zu 3-MP deutlich geringer, was die Spezifität des Enzyms für das natürliche Startersubstrat Acetyl-CoA demonstriert. Um den Grund für die Spezifität herauszufinden und die Ausbeuten von 3-EP und 3-PrP durch Protein-Engineering zu erhöhen, wurde MSAS in der vorliegenden Arbeit biochemisch charakterisiert.

### MSAS aus *S. erythraea* kann in *E. coli* exprimiert werden

Zunächst wurde MSAS aus *Saccharopolyspora erythraea* (SerMSAS) verwendet, welche aus *Actinomyces* stammt und daher für die heterologe Expression in *E. coli* potenziell geeignet ist. Für die chromatographische Reinigung wurden verschiedene Designs mit Tags an beiden Termini untersucht. Dabei konnte eine deutlich schlechtere Proteinqualität bei der Verwendung eines N-terminalen Tags festgestellt werden. Daher wurde ein C-terminaler Tag verwendet und eine Strategie für die Reinigung des Enzyms etabliert. Die Qualität des Enzyms konnte durch die Analysemethoden SDS-PAGE, SEC, TSA bestätigt werden. Zusätzlich wurde die

posttranslationale Aktivierung des ACP durch einen Fluoreszenz Assay nachgewiesen.

### SerMSAS ist in ein Gencluster eingebunden

Im Gegensatz zu Experimenten *in vivo* konnte jedoch keine Synthese von 6-MSA *in vitro* beobachtet werden. Folglich wurde die Primärstruktur durch ein Alignment mit verwandten MSAS untersucht. Dabei wurde festgestellt, dass ein konserviertes Motiv fehlt, welches für die Freisetzung von 6-MSA als notwendig beschrieben wurde. Zusätzlich konnte gezeigt werden, dass SerMSAS in einem Gencluster integriert ist, in welchem zusätzlich ein KASIII-ähnliches Protein (ChIB3), ein diskretes Carrier-Protein (ChIB2) und ein Cytochrom P450 enthalten ist. Eine phylogenetische Analyse des KASIII-ähnlichen Proteins ließ vermuten, dass es sich um eine ACP-abhängige Transferase handelt. Des Weiteren wurde aufgrund der großen Ähnlichkeit zu einem Enzym aus der Chlorothricin Synthese angenommen, dass ChIB3 das Produkt 6-MSA nach der Biosynthese von MSAS auf ein diskretes Carrier-Protein überträgt. Um diese Theorie zu untersuchen, konnten Reinigungsstrategien für die Proteine ChIB3, ChIB2 sowie das alleinstehende ACP von MSAS etabliert werden. Dafür wurden verschiedene Konstrukte entworfen und Codon-Optimierungen durchgeführt, um genügend Proteinmengen für die Analysen zu erhalten. Diese zeigten, dass SerMSAS in Anwesenheit von ChIB3 und ChIB2 *in vitro* aktiv ist und 6-MSA synthetisiert. Zusätzlich wurde die Transferfunktion von ChIB3 untersucht und gezeigt, dass es 6-MSA von einem künstlichen Ester, welche das ACP nachahmt, auf ChIB2 übertragen kann. Zusammengefasst bestätigten die gesammelten Daten die aufgestellten Theorien, dass SerMSAS sein Produkt 6-MSA nicht freisetzen kann und dieses stattdessen von ChIB3 auf das diskrete Carrier-Protein übertragen wird, auf welchem weitere Modifikationen stattfinden können. Weiterhin wurde SerMSAS mittels EM untersucht und es konnten bereits vielversprechende Ergebnisse erzielt werden. Im Gegensatz dazu erschwerten die Enzymeigenschaften von SerMSAS, das System für die angestrebten Engineering-Ansätze zu nutzen, da es nicht in der Lage ist sein Produkt freizusetzen.

### MSAS aus *P. patulum* setzt nicht native Startersubstrate deutlich langsamer um

Daher wurde der Fokus auf eine verwandte MSAS aus *Penicillium patulum* (PenPaMSAS) gesetzt. Die Reinigungsstrategie von SerMSAS konnte erfolgreich auf PenPaMSAS angewendet werden, um genügend Enzym mit hoher Reproduzierbarkeit zu erhalten. Anschließend konnte eine Methode etabliert werden, um die enzymatische Aktivität von MSAS *in vitro* zu analysieren. Dabei wurde der Verbrauch von NADPH photometrisch untersucht sowie die Produkte mittels HPLC-MS verifiziert. Die Umsatzraten für die 6-MSA-Synthese des gereinigten Enzyms waren im Vergleich zur Literatur etwas höher, was für eine sehr gute Proteinqualität spricht. Zusätzlich wurde die Verwendung nicht natürlicher Startersubstrate untersucht und gezeigt, dass Butyryl-CoA und Hexanoyl-CoA

deutlich langsamer umgesetzt werden. Anhand dieser Ergebnisse konnte die Spezialisierung von PenPaMSAS für das natürliche Substrat Acetyl-CoA belegt werden. Außerdem wurden die Umsatzraten für die Synthesen von Nebenprodukten in Abwesenheit von NADPH untersucht. Diese sind im Vergleich zu den jeweiligen Salicylsäuren signifikant niedriger für Acetyl-CoA, jedoch etwas höher für Butyryl-CoA und Hexanoyl-CoA. Die Ergebnisse deuten darauf hin, dass längere Acylketten schlechtere Substrate für die KR sind. Jedoch wurden Charakterisierungen einzelner Enzyme benötigt, um den Grund für langsamere Umsatzraten von nicht natürlichen Startersubstraten zu untersuchen.

#### Die AT von PenPaMSAS ist spezifisch für ihr natives Startersubstrat

Es wurde mit der Charakterisierung der AT begonnen, da sie für die Beladung der Substrate verantwortlich ist. Nach dem Design und der Reinigung der individuellen KS-AT-Domäne sowie des ACP konnte gezeigt werden, dass die AT nicht natürliche Substrate wie Propionyl-CoA und Hexanoyl-CoA 400-mal langsamer im Vergleich zu Acetyl-CoA transferiert. Aufgrund der Substratspezifität kann die AT gegen diese nicht natürlichen Substrate diskriminieren und ist zumindest ein Hauptfaktor, wenn nicht allein verantwortlich, für deren langsamere Umsatzgeschwindigkeiten durch die MSAS.

#### Die Spezifität der PenPaMSAS AT kann durch Änderungen des aktiven Zentrums verändert werden

Um die Spezifität dieser Domäne durch Protein-Engineering zu verändern, wurden mehrere Mutanten untersucht und schließlich ein vielversprechender Kandidat mit höheren Umsatzraten für die Beladung von Butyryl-CoA (~200%) und Hexanoyl-CoA (~300%) gefunden.

#### Die promiskuitive MAT der Säugetier-FAS kann fluoridierte Substrate auf das DEBS-ACP übertragen

Im Gegensatz zu der AT von MSAS konnte für die Malonyl/Acetyl-Transferase (MAT) der Säugetier-FAS in vergangenen Arbeiten gezeigt werden, dass sie viele Substrate mit hohen Umsatzgeschwindigkeiten übertragen kann. Außerdem können Punktmutationen die Spezifität der Domäne verändern, was die Plastizität, aber auch die Robustheit der Domäne zeigt. Aufgrund dieser faszinierenden Eigenschaften wurde die MAT als ein vielversprechender Kandidat für das Protein-Engineering verwendet. Im Gegensatz dazu sind ATs von PKSs Gatekeeper für die Auswahl von Substraten und weisen daher meist hohe Substratspezifitäten auf. In der vorliegenden Arbeit wurde ein Domänenaustausch von ATs aus PKS Systemen mit der MAT untersucht, um eine regiospezifische Derivatisierung von Polyketiden zu ermöglichen und erstmals PKS/FAS-Hybride zu beschreiben. Der Fokus lag dabei auf der Umsetzung von fluoridierten Substraten, weil dadurch pharmakologisch

relevante fluoridierte Polyketide durch die Biosynthese synthetisiert werden können. Nach der chemischen Synthese von Fluormalonyl-CoA (F-Mal-CoA) konnte gezeigt werden, dass die MAT exzellente Transferkinetiken für dieses Substrat aufweist. Als PKS Modellsystem wurde das sechste Modul der DEBS verwendet, da es in der Literatur sehr gut beschrieben und in *E. coli* zugänglich ist. Anschließend wurde die MAT-katalysierte Übertragung von Substraten auf das ACP von DEBS untersucht und mit der nativen AT von DEBS Modul 6 verglichen. Es konnte gezeigt werden, dass die Spezifitätskonstante der MAT-katalysierten Reaktion im Vergleich zur AT von DEBS um 2 bis 3 Größenordnungen höher war. Der Grund hierfür liegt an den hohen Umsatzgeschwindigkeiten der MAT und beweist, dass die MAT Substrate auf das ACP von DEBS sehr schnell übertragen kann.

### Das Design von DEBS/FAS-Hybriden

Anschließend wurden DEBS/FAS-Hybride über einen Domänen austausch hergestellt. Da die AT in beiden Systemen mit der KS über einen Linker verbunden ist, wurden zwei Hybride entworfen, in welchem die AT mit oder ohne diesen Linker ausgetauscht wurde. Nach der Reinigung der Hybride sowie der unveränderten PKS wurden die enzymatische Aktivitäten mithilfe chemisch hergestellter Startersubstrate untersucht. Die Umsatzgeschwindigkeiten sowie die hergestellten Produkte zeigten, dass beide Hybride in der Lage waren,  $\delta$ -Lactone (TKLs) mit dem Verlängerungssubstrat Methylmalonyl-CoA (MM-CoA) zu synthetisieren. Jedoch wies der AT-Austausch ohne den angrenzenden Linker bessere Enzymeigenschaften auf, was anhand von Oligomerisierungszuständen sowie enzymatischen Aktivitäten gezeigt wurde. Aufgrund dessen wurde anschließend nur dieses Enzym für weitere Arbeiten verwendet. Das Design kann vermutlich als Richtlinie für die Herstellung anderer Hybride verwendet werden. Neben MM-CoA war der DEBS/FAS-Hybrid auch in der Lage, die nicht natürlichen Substrate Malonyl-CoA (Mal-CoA) und insbesondere F-Mal-CoA zu nutzen, um C2-Derivate der TKLs zu produzieren. Im Gegensatz dazu konnte die unveränderte PKS keine Derivate synthetisieren, was den großartigen Erfolg der Methode demonstriert.

### Der DEBS/FAS-Hybrid synthetisiert Makrolacton-Derivate

Mit dem katalytisch aktiven DEBS/FAS-Hybrid sollten schließlich auch Makrolacton-Derivate hergestellt werden, welche Vorstufen bioaktiver Verbindungen sind. Unter Verwendung von Pentaketid- und Hexaketid-Startersubstraten aus der verwandten Pikromycin-Synthese, welche von der Arbeitsgruppe David Sherman bereitgestellt wurden, konnte die Herstellung von 10-Deoxymethynolid und Narbonolid unter Verwendung der unveränderten PKS mit dem natürlichen Substrat MM-CoA gezeigt werden. Diese Verbindungen sind Vorstufen für die Antibiotika Methymycin/Neomethymycin und Pikromycin/Narbomycin. Im Gegensatz dazu konnte der DEBS/FAS-Hybrid auch Mal-CoA und F-Mal-CoA verwenden, um demethylierte und fluoridierte Derivate zu synthetisieren, was anhand von

NADPH-Verbrauch sowie HPLC-MS Analysen bestätigt wurde. Überraschenderweise waren die Umsatzraten für die Umsetzung von Mal-CoA etwas höher im Vergleich zu MM-CoA und nur geringfügig langsamer für F-Mal-CoA. In Gegenwart des Hexaketid-Startersubstrats betrug die Umsatzgeschwindigkeit des Hybriden von F-Mal-CoA ungefähr die Hälfte des unveränderten Enzyms mit dem natürlichen Verlängerungssubstrat. Aufgrund dessen konnte gezeigt werden, dass die Methode für die Herstellung fluorierter Verbindungen besonders gut geeignet ist.

#### Der DEBS/FAS-Hybrid synthetisiert 2S-Fluor-10-Deoxymethynolid aus FMM-CoA

Für weitere Analysen wurde die Biosynthese und Reinigung der demethylierten und insbesondere fluorierten Verbindungen in Milligramm-Ausbeuten etabliert. Im Gegensatz zu einer demethylierten Verbindung konnten die fluorierten Verbindungen allerdings nicht isoliert werden, was vor allem an geringen Ausbeuten sowie dem Auftreten von Nebenreaktionen lag. Durch die Identifizierung eines fluorierten Nebenproduktes konnte gezeigt werden, dass die TE das Elongationsprodukt nicht ausreichend zyklisieren kann und stattdessen durch Hydrolyse vom Enzym entfernt. Um die Ausbeuten zu erhöhen, wurde versucht diese Nebenreaktion durch die Herstellung einer Fluormethyl-Disubstitution zu verhindern. Diese könnte ebenfalls einen positiven Einfluss auf die chemische Stabilität des fluorierten Polyketids haben. Der Einbau einer solchen Gruppe wäre durch die Beladung des Enzyms mit Fluormethylmalonyl-CoA (F-MM-CoA) möglich. Nach chemischer Synthese konnte demonstriert werden, dass der DEBS/FAS-Hybrid dieses Substrat im Vergleich zu F-Mal-CoA mit deutlich höheren Geschwindigkeiten umsetzen kann und das entsprechende Produkt wurde mittels HPLC-MS verifiziert. Zusätzlich konnte das fluorierte 10-Desoxymethynolid-Derivat isoliert und durch NMR untersucht werden. Diese zeigte den stereoselektiven Einbau des chemisch hergestellten F-MM-CoA Racemats. Neben der außerordentlichen Relevanz des Produktes ist die Biosynthese einer di-substituierten Verbindung auch für den KS-Mechanismus aufschlussreich. Dieser muss entweder in einem konzertierten oder in einem zweistufigen Decarboxylierungs-Additionsmechanismus stattfinden. Im Gegensatz dazu kann die C-C-Bindung nicht vor der Decarboxylierung aufgebaut werden. In Abwesenheit von NADPH wurde ein Cyclohexanon-Seitenprodukt gefunden, ähnlich zu der Reaktion mit F-Mal-CoA. Auch hier kann das Elongationsprodukt durch die TE nicht zyklisiert werden und wird stattdessen hydrolysiert. Das Hydrolyseprodukt wird anschließend decarboxyliert und das Cyclohexanon durch Aldolkondensation gebildet. Zusammengefasst konnte gezeigt werden, dass der AT-Austausch von PKSs mit der MAT die Biosynthese von regiospezifischen Polyketidderivaten ermöglicht, darunter hochrelevante fluorierte Verbindungen wie die Einführung einer Fluormethyl-Disubstitution.

Diese Arbeit wurde in *bioRxiv* publiziert: Rittner, A., Joppe, M., Schmidt, J. J., Mayer, L. M., Heid, E., Sherman, D. H., & Gringer, M. (2021). Directed biosynthesis of fluorinated polyketides. *bioRxiv*.

Neue Erkenntnisse über Megasyntasen

Die vorliegende Arbeit trägt zum Verständnis der FAS- und PKS-Systeme bei. Es wurden Expressionsstrategien etabliert, welche die strukturelle und funktionelle Charakterisierung von Megasyntasen ermöglichen. Weiterhin konnten Richtlinien und Methoden für das Protein-Engineering entworfen werden, welche die Biosynthese relevanter Naturstoffderivate ermöglichen. Darüber hinaus wurde die Rolle der AT als Gatekeeper für die Selektion von Substraten in PKSs gezeigt. Durch das Protein-Engineering, wie ortsspezifische Mutagenese des aktiven Zentrums sowie Domänenaustausch, konnte die Substratspezifität von Megasyntasen verändert werden. Schließlich konnte durch das Design eines effizienten PKS/FAS-Hybriden unter anderem ein fluoriertes Macrolacton biosynthetisch hergestellt werden. Diese pharmakologisch relevante Verbindung demonstriert das enorme Potenzial des Engineerings von Proteinen der Naturstoffsynthese, da effiziente Katalysatoren für die Biosynthese von Naturstoffderivaten zu Verfügung gestellt werden können.



## Abstract

In this thesis, we characterized megasyntases such as fatty acid synthases (FASs) and polyketide synthases. The obtained insights into structure and function were used to engineer such systems to produce new-to-nature compounds.

The *in vitro* characterization of megasyntases requires reproducible access to these enzymes in high quality. Therefore, we established purification strategies for the yeast FAS and the methylsalicylic acid synthase (MSAS) from *Saccharopolyspora erythraea* (SerMSAS) and applied the latter one on MSAS from *Penicillium patulum* (PenPaMSAS) and on 6-deoxyerythronolide B synthase (DEBS) module 6. With the purified samples, we were able to obtain initial structural data for SerMSAS and solve the complete structure of the yeast FAS (PDB: 6TA1). On the example of the yeast FAS, we could show that the sample can suffer from adsorption to the water-air interface during the grid preparation for electron microscopy and presented how the use of graphene-based grids can overcome this problem. The combined structural and functional analysis of the yeast FAS showed that the structural domains trimerization module and dimerization module 2 are not essential for the assembly of the whole system. Therefore, they can potentially be used for domain exchange approaches. The in-depth functional analysis of SerMSAS revealed that not SerMSAS itself releases the product, but a 3-oxoacyl-(acyl-carrier protein) synthase like enzyme within the gene cluster transfers 6-methyl salicylic acid from SerMSAS to another carrier protein for subsequent modifications. In contrast, we showed that PenPaMSAS can release its product by hydrolysis and that non-native substrates can be incorporated although at significantly slower turnover rates compared to the native starter substrate. Our further investigation demonstrated that the substrate specificity of the acyltransferase (AT) is a critical factor for the incorporation of non-native substrates.

With the insight from the functional and structural characterization, we engineered megasyntases for the biosynthesis of natural product derivatives. We targeted the AT of PenPaMSAS for active site mutagenesis and discovered a mutant which can transfer non-native substrates significantly faster (~200-300%). Additionally, the malonyl/acetyl transferase (MAT) of the mammalian FAS was used as a promising target for protein engineering because of its previously reported properties including polyspecificity, fast transfer kinetics, robustness, and plasticity. We showed that the MAT can transfer fluorinated substrates and accept the acyl carrier protein of DEBS module 6. By exchanging the substrate specific AT of DEBS with the polyspecific MAT of the mammalian FAS, we demonstrated an efficient DEBS/FAS hybrid and an optimal truncation site for the applied ATs. In contrast to the wild type system, the DEBS/FAS enzyme was able to synthesize demethylated and fluorinated derivatives. The production and purification of a fluoro-methyl-disubstituted polyketide was of particular interest, as it has a high potential for the generation of new drugs and shows the potential of protein engineering. Furthermore, the incorporation of the

disubstituted substrate had important implication in the mechanistic details of the ketosynthase-mediated C-C bond formation.

## Table of Contents

<b>Abbreviations</b> .....	<b>III</b>
<b>Zusammenfassung</b> .....	<b>VII</b>
<b>Abstract</b> .....	<b>XVII</b>
<b>Chapter 1: Introduction</b> .....	<b>- 1 -</b>
1.1 Background.....	- 1 -
1.2 Fatty acid and polyketide biosynthesis.....	- 2 -
1.2.1 Acyltransferases.....	- 6 -
1.2.2 Ketosynthases.....	- 8 -
1.2.3 Thioesterases.....	- 9 -
1.3 Type I FASs.....	- 10 -
1.3.1 Fungal FAS.....	- 13 -
1.3.2 Mammalian FAS.....	- 15 -
1.4 Type I PKSs.....	- 16 -
1.4.1 Methylsalicylic acid synthase.....	- 17 -
1.4.2 6-Deoxyerythronolide B synthase.....	- 19 -
1.4.3 Pikromycin synthase.....	- 21 -
1.5 Engineering of megasyntases.....	- 22 -
1.6 Aim of the thesis.....	- 24 -
<b>Chapter 2: Results and analysis of the data</b> .....	<b>- 25 -</b>
2.1 Characterization of the yeast fatty acid synthesis.....	- 25 -
2.1.1 Establishment of a purification strategy for yeast FAS.....	- 25 -
2.1.2 Sample preparation for cryo-EM.....	- 27 -
2.1.3 Structural determination of the yeast FAS.....	- 36 -
2.1.4 Mutational study of yeast FAS phosphorylation sites.....	- 38 -
2.1.5 Investigating scaffolding domains of the yeast FAS.....	- 40 -
2.2 Characterization of a non-releasing MSAS from <i>S. erythraea</i> .....	- 42 -
2.2.1 Expression and purification of SerMSAS in <i>E. coli</i> .....	- 43 -
2.2.2 <i>In vivo</i> 6-MSA synthesis.....	- 46 -
2.2.3 <i>In vitro</i> 6-MSA synthesis.....	- 47 -
2.2.4 Sequence analysis and expression of SerChIB3.....	- 48 -
2.2.5 Expression of SerMSAS-ACP(0) and SerChIB2.....	- 53 -

2.2.6 <i>In vitro</i> 6-MSA synthesis with SerChIB3 .....	- 55 -
2.2.7 SerChIB3 can transfer 6-MSA moieties .....	- 56 -
2.2.8 Initial electron microscopy data .....	- 57 -
2.3 Engineering of PenPaMSAS towards 6-PrSA synthesis .....	- 58 -
2.3.1 Expression and purification of PenPaMSAS .....	- 60 -
2.3.2 Characterization of PenPaMSAS .....	- 60 -
2.3.3 Purification of dissected PenPaMSAS domains .....	- 65 -
2.3.4 Specificity of the AT .....	- 67 -
2.3.5 Engineering of the AT .....	- 68 -
2.3.6 Characterization of the mutants .....	- 71 -
2.4 Directed biosynthesis of fluorinated polyketides .....	- 71 -
2.4.1 Characterization of the MAT .....	- 72 -
2.4.2 Design and purification of DEBS/FAS hybrids .....	- 74 -
2.4.3 Characterization of DEBS/FAS hybrids .....	- 76 -
2.4.4 Analysing the stereochemistry of compound 27 .....	- 83 -
<b>Chapter 3: Discussion .....</b>	<b>- 91 -</b>
3.1 Natural products of megasynthases are pre-cursors for bioactive compounds .....	- 91 -
3.2 Purification strategies are required for structural- and biochemical characterization .....	- 92 -
3.3 Specificities of megasynthases .....	- 93 -
3.4 Protein engineering usually decreases enzymatic activity .....	- 94 -
3.5 Engineering of ATs .....	- 95 -
3.6 Guidelines for engineering .....	- 96 -
3.7 Semi-synthesis and biosynthesis .....	- 97 -
3.8 Impact of our findings on enzyme mechanisms .....	- 98 -
<b>Chapter 4: Summary .....</b>	<b>- 99 -</b>
<b>Chapter 5: Outlook .....</b>	<b>- 106 -</b>
<b>Chapter 6: Experimental procedure .....</b>	<b>- 110 -</b>
6.1 Material and methods .....	- 110 -
6.1.1 Cloning of plasmids .....	- 110 -
6.1.2 Protein expression and purification .....	- 113 -
6.1.3 SDS-PAGE .....	- 117 -

6.1.4 Thermal-shift-assay (TSA).....	- 118 -
6.1.5 Enzymatic activity <i>in vitro</i> (NADPH consumption) .....	- 118 -
6.1.6 Analysing protein denaturation at the air-water interface.....	- 120 -
6.1.7 Native-PAGE Western blot to investigate the assembly of the yeast FAS.....	- 120 -
6.1.8 Analysing the phosphopantetheinylation by CoA-488 .....	- 121 -
6.1.9 Extraction of 6-MSA, 6-PrSA, 6-PSA and 4-hydroxy-6-alkyl-2-pyrone-	122 -
6.1.10 Mass spectroscopic analysis of ACPs .....	- 122 -
6.1.11 Characterization of SerChIB3 in its ability to transfer 6-MSA moieties from 6-MSA-SNAC to ChIB2 .....	- 122 -
6.1.12 Enzymatic activity <i>in vitro</i> (4-hydroxy-6-alkyl-2-pyrone assay) .....	- 123 -
6.1.13 Characterization of ATs .....	- 123 -
6.1.14 Enzymatic activity for the synthesis of non-reduced TKLs.....	- 125 -
6.1.15 Verification of TKLs and macrolactones using HPLC-MS.....	- 125 -
6.1.16 Analysis and purification of macrolactones using HPLC-UV .....	- 125 -
6.1.17 Blue native PAGE.....	- 126 -
6.2 Chemical synthesis and biosynthesis.....	- 126 -
6.2.1 Synthesis of 6-MSA-SNAC.....	- 127 -
6.2.2 Synthesis of natural diketide SNAC (2) .....	- 127 -
6.2.3 Synthesis of F-Mal-CoA.....	- 128 -
6.2.4 Synthesis of F-MM-CoA .....	- 130 -
6.2.5 Biosynthesis and analysis of compound 16.....	- 131 -
6.2.6 Biosynthesis and analysis of compound 17/25.....	- 133 -
6.2.7 Biosynthesis and analysis of compound 21 .....	- 134 -
6.2.8 Biosynthesis and analysis of compound 27 .....	- 136 -
<b>Chapter 7: Appendix .....</b>	<b>- 139 -</b>
7.1 Supplementary figures .....	- 139 -
7.2 Supplementary equations .....	- 155 -
7.2.1 Linear fit.....	- 155 -
7.2.2 Exponential fit.....	- 155 -
7.2.3 Michealis-Menten approximation.....	- 155 -
7.2.4 Michealis-Menten equation.....	- 155 -
7.2.5 Global fit .....	- 155 -

7.3 NMR spectra .....	- 156 -
7.4 Primers and plasmids.....	- 170 -
7.5 Literature.....	- 174 -
7.6 Statement of personal Contributions.....	- 197 -
7.6.1 Scientific work.....	- 197 -
7.6.2 Writing process.....	- 199 -
7.7 Acknowledgement.....	- 204 -
7.8 Danksagung.....	- 204 -
7.9 Eidesstattliche Erklärung.....	- 205 -
7.10 Curriculum Vitae .....	- 206 -

# Chapter 1: Introduction

## 1.1 Background

Natural products have been one of the most relevant sources of drugs over the last several decades and are used to treat a wide range of diseases.<sup>1,2</sup> Reflecting their relevance, around 32% of all FDA-approved drugs in the years from 1981 to 2019 are natural products or their derivatives and the number would be even higher if natural product-inspired drugs (used as lead compounds) would be taken into account.<sup>2</sup> This high relevance can be explained by the fact that their molecular scaffolds have evolved to interact with biomacromolecules such as proteins and nucleic acids.<sup>1</sup> However, only 6% of the 32% are unaltered natural products, mainly because of limited compound availability, poor solubility and lack of metabolic stability.<sup>1</sup> This huge discrepancy demonstrates the relevance of natural product derivatization as a method to develop bioactive compound into clinical drugs. The development of new or improved drugs requires the understanding of biological activities and target interactions. The fluorination of compounds is a promising strategy for lead structure optimization,<sup>3,4</sup> able to improve drug metabolism and pharmacokinetics (DMPK), efficacy, bioavailability, chemical stability and reduce side effects.<sup>1,3,5-7</sup>

The fatty acids and polyketides are two major classes of natural products with a wide range of biological activities and both are derived from the acetate pathway. Fatty acids are de novo biosynthesized by fatty acid synthases (FASs). Fatty acids are central molecules of the primary metabolism used for cell walls, membranes, protein modification, as energy storage and building blocks.<sup>8,9</sup> Besides their biological activity, fatty acids are also biotechnologically interesting as platform chemicals for detergents, lubricants, soaps, plastics, coatings, biofuels precursors and others.<sup>10,11</sup> Traditionally fatty acids are obtained by animal fats and plant oils, but because of environmental and socioeconomic concerns, biosynthesis within an engineered microbial strain becomes more and more relevant.<sup>12</sup> Even if the biosynthetic production of fatty acids is more costly, it also offers major advantages; i.e., the independence from feedstocks and the possibility to specifically synthesize a functionalized fatty acid with higher value, like fatty alcohols, alkanes or alkenes, which may not be available in nature.<sup>11,12</sup> The production of those compounds *in vivo* requires metabolic- and protein engineering and therefore has been a focus for many years. The chemically related polyketides are a diverse class of secondary metabolites produced by polyketide synthases (PKSs). More than 10.000 compounds are known to date with biological roles as pigments, virulence factor, infochemical and for defending the habitant.<sup>13,14</sup> Currently many polyketides are used as antibiotic (e.g., erythromycin, Biaxin, Zithromax), anticancer (e.g., epothilone), antifungal antiparasitic drugs (e.g., avermectin) and cholesterol-lowering drugs (e.g., lovastatin)

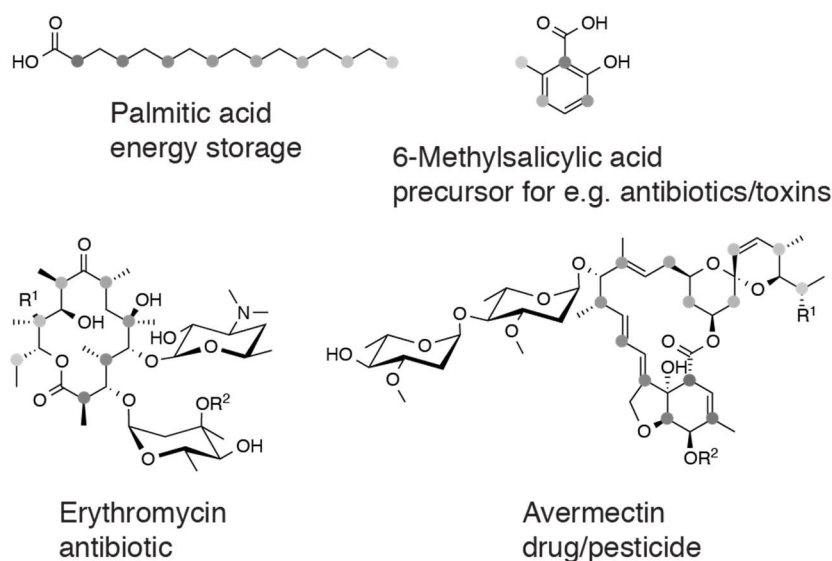
as well as immunosuppressants (e.g., rapamycin, FK506,).<sup>13–16</sup> Reflecting their relevance, since 2008, 20% of the natural-product derived entities launched into the clinic were polyketides.<sup>16</sup> On the other hand the majority of polyketides are not used as drugs, probably as they evolved to interact with biomacromolecules from the microbiome, but not from the human body.<sup>16</sup> Therefore, to find new medicines, there is a high demand in discovering new polyketide scaffolds and to derivatize them to improve properties as discussed above.<sup>16</sup> Polyketides often possess extraordinary structural complexity. Following their discovery, they were traditionally produced by highly challenging chemical total synthesis over multiple steps at overall negligible yields, high costs and large amounts of solvent and by-product wastes. In nature, these complex structures are assembled from very simple substrates by specific enzymatic catalysis. Following biosynthetic approaches, polyketides could be received within few steps and at environmentally friendly conditions, which could decrease drug prices immensely. However, derivatization by biosynthetic strategies, that is expanding the structural diversity of polyketides by regioselective modification, is very challenging and requires elaborate protein engineering.<sup>13,17,18</sup>

Engineering of multidomain-enzymes like FASs or PKSs is very challenging and requires deep knowledge about enzyme kinetics and protein properties like structures, conformational variability, protein-protein and domain-domain interactions and molecular mechanisms. Acquiring protocols and guidelines for FAS and PKS engineering is tempting, because they could improve the access to relevant and highly complex chemicals from cheap sources in an economically efficient and environmentally friendly way. Therefore, there is a need for in-depth characterization of these enzymes and the translation of this information into protein engineering applications.

## 1.2 Fatty acid and polyketide biosynthesis

Fatty acids and polyketides are major classes of natural products and both are derived from the acetate pathway. In nature, they are produced by FASs and PKSs. The extraordinary relevance of those enzymes can be explained by their synthetic capability to use acetate units, as one of the simplest building blocks in nature, to assemble them to fatty acids and polyketides with partly high structural complexity (figure 1). During the biosynthesis, carbon-carbon bonds are generated and  $\beta$ -ketoester processed, giving compounds with accumulated functional groups and chiral centres.<sup>13</sup>

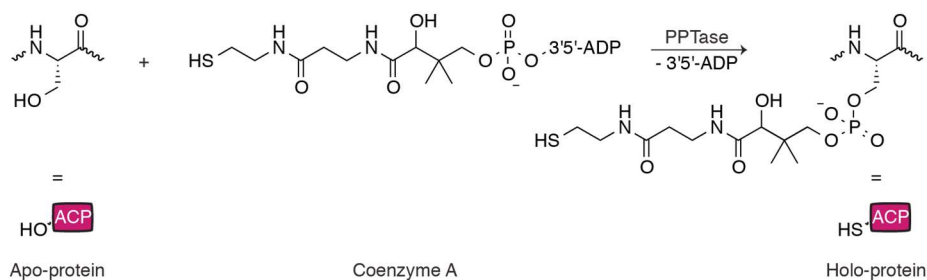




**Figure 1** | Chemical structures and diversity of fatty acids and polyketides demonstrated on palmitic acid, 6-methylsalicylic acid, erythromycin and avermectin. The order of assembled building blocks during the biosynthesis is illustrated by grey dots with increasing grey scale starting with the first building block as lightest colour.

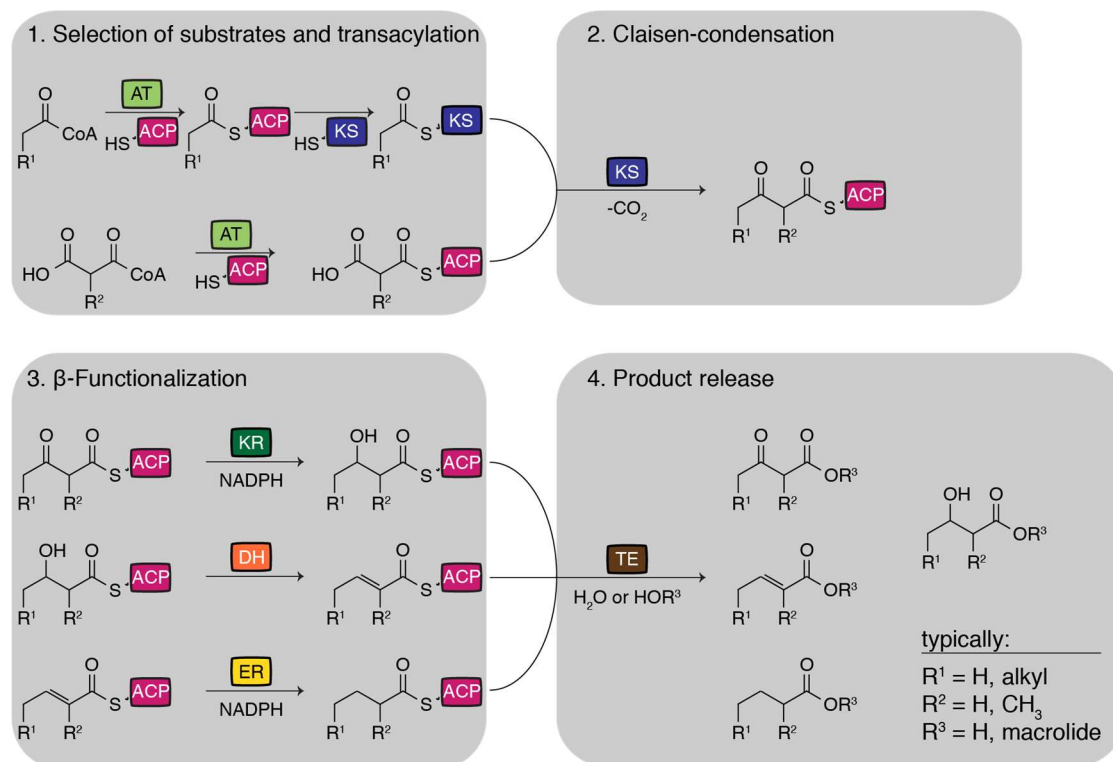
Although fatty acids and polyketides are different in their chemical complexity, they are chemically and structurally related. Strictly speaking, fatty acids can be categorized as a polyketide subclass. The common synthetic origin of fatty acids and polyketides is due to the synthesis by FASs and PKSs that share an evolutionary ancestor. Therefore, they have similar catalytical mechanisms and use a common pool of starter- and elongation substrates.<sup>13,15,19</sup>

The fatty acid and polyketide biosynthesis is performed by one set of catalytic domains supported by an acyl carrier protein (ACP) that shuttles substrates and intermediates to the active sites. Therefore, the ACP must form productive domain-domain and protein-protein interactions during biosynthesis. To be able to accept acyl moieties, the active site serine of the apo-ACP has to be posttranslationally modified with a 4-phosphopantetheine to form the active holo-ACP (figure 2).<sup>14</sup>



**Figure 2** | The phosphopantetheinylation of apo-ACP to form the active holo-ACP is catalysed by a phosphopantetheinyl transferase (PPTase). The PPTase acylates the conserved active serine with 4-phosphopantetheine.

A functional cycle in FASs/PKSs proceeds as follows: The substrates are usually loaded by acyltransferases (ATs) on the holo-ACP. Consequently AT domains have a major role as gatekeepers in selecting and transferring starter and elongation substrates (figure 3).<sup>20</sup> The ACP bound acyl-moiety is shuttled and loaded to the ketosynthase (KS), where it is then Claisen-condensed with an ACP bound extender substrate. In doing so, the starter moiety gets elongated by at least two carbon atoms to create an ACP bound  $\beta$ -ketoester. The release of  $\text{CO}_2$  is the driving force of this reaction.<sup>21</sup> The name polyketide refers to the assembly of  $\beta$ -ketones and consequently the required domains for its synthesis, ACP, AT and KS, build the minimal set within PKSs.<sup>22</sup> Subsequently a set of sequentially acting reductive domains, including a ketoreductase (KR), a dehydratase (DH) and an enoylreductase (ER), can further process the  $\beta$ -ketoester. The KR catalyses the nicotinamide adenine dinucleotide phosphate (NADPH) mediated reduction to form a  $\beta$ -hydroxyl group, which subsequently is tailored to a double bond after elimination of water by the DH. The double bond can finally be further reduced to the fully saturated acyl-chain by another NADPH mediated reduction by the ER. In contrast to the fatty acid synthesis where all reductive reactions are performed before the intermediate enters the next round of extension, the processing steps are optional for PKSs, leading to a more complex pattern of functionalization. The acyl-chain gets extended and processed multiple times until the final product is released from the enzyme, usually as free acids or lactones catalysed by thioesterases (TEs). In contrast to fatty acids, polyketides are a class of large chemical and structural diversity due to optional and variable processing, and furthermore by the selection of different substrates, variable polyketide release mechanisms and post-PKS modifications.<sup>13</sup>



**Figure 3** | Typical chemical reactions during the fatty acid and polyketide synthesis, catalysed by FASs and PKSs. Both enzyme classes use a similar set of domains, although PKSs can include further specialized domains (for more information see reference).<sup>16,21</sup> In this scheme, enzymatic reactions are divided in the categories: substrate loading, carbon-carbon bond formation,  $\beta$ -ketoester reduction and product release (grey boxes). ACP: acyl-carrier-protein; AT: acyltransferase; KS: ketosynthase; KR: ketoreductase; DH: dehydratase; ER: enoylreductase; TE: thioesterase. The figure is adapted from *Rittner et al.*<sup>23</sup>

FASs and PKSs are usually classified into different types depending on their architecture. In type I systems, catalytic domains are covalently fused to large multifunctional enzymes, also referred to as megasynthases. In contrast, FASs and PKSs made up from discrete domains are termed type II systems. Additionally, type III was assigned to chalcone synthases and is solely a category for PKSs where the polyketide is never attached to the enzyme. In the following, the focus will be on type I systems as the main targets in this thesis. This group can be further subdivided into iterative and modular working systems and describes if the synthase uses their catalytic domains repetitively or not. All known type I FASs are iterative systems and their domains perform a defined number of reactions until the final product is released from the enzyme. Similarly, there are iteratively working PKSs categorized into highly-reducing, partially-reducing and non-reducing PKSs, according to the presence and absence of processing domains. The class of type I PKSs also includes modular systems in which each domain usually performs a single reaction during the biosynthesis. Here modules are covalently and non-covalently connected

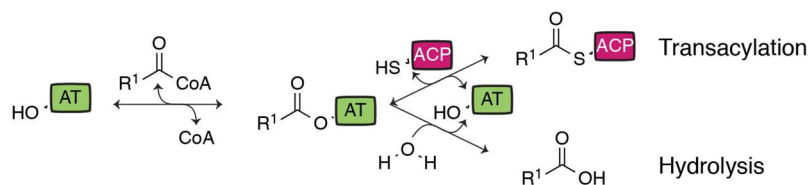
to form linear assembly lines that elongate and process the growing chain module-by-module in vectorial manner, involving a translocation reaction of the polyketide from one to the next module.<sup>13,14</sup> Consequently, the number of modules correlates with the number of chain elongations and therefore the final product can be predicted by the enzyme architecture (and vice versa), known as the principle of collinearity. Although there are some exceptions to the rule, like module skipping and iterative use of modules, it implicates the great potential of this enzyme class for rational reprogramming by genetic manipulation.<sup>13,17,24</sup> The origin of the vectorial synthesis is still not known, but recently a turnstile mechanism was proposed, which inhibits an iterative use from the KS.<sup>17,25</sup> In recent years, several polyketide derivatives have been produced by protein engineering. However, yields have usually been low, probably because important features for the biosynthesis like substrate specificities, interactions between domains or modules, and the ACP mediated intermediate shuttling, are disturbed in engineering PKSs (for more information see chapter 1.5).<sup>26</sup> Finally, it is noted that besides the above described cis-AT PKSs, also trans-AT PKSs are known that lack a covalently attached AT, but operate with discrete ATs.<sup>13</sup> Trans-AT PKSs have not been investigated in this thesis, and in the following the term PKS is used for type I cis-AT systems.

The classification in general helps us to reflect on this huge family of enzymes, but is not always useful as there are many exceptions and hybrids of classes, indicating that FASs and PKSs originate from a common ancestor.<sup>13,19</sup> In the following, ATs, KSs and TEs are described in more detail as the enzymatic domains relevant for this thesis.

### 1.2.1 Acyltransferases

ATs have an important role as gatekeepers within FASs and PKSs, as they select the starter- and extender substrates that are transferred to the sulfhydryl group of the ACP and finally incorporated into the product. As a consequence, ATs have been a major target for engineering, including domain exchanges, active site mutagenesis and AT cross-complementation.<sup>17,27-32</sup> The enzyme is built up by an  $\alpha/\beta$ -hydrolase-like core domain and a smaller ferredoxin-like structure. The active site serine usually contains a conserved GHSXG motif and lies at a sharp turn, termed nucleophilic elbow, in a gorge between those two subdomains. The nucleophilicity of the serine is enhanced by a helix dipole-moment, created by the architecture of the elbow, and furthermore by hydrogen bonding to a histidine forming a serine-histidine catalytic dyad.<sup>20,33</sup> The transacylation is performed in a double replacement mechanism also referred to as “ping-pong” mechanism. In the first step the enzyme is occupied with an acyl-CoA and then the active site serine nucleophilically attacks the carbonyl group to bind the acyl-moiety and to liberate coenzyme (CoA). In the second step, the ACP docks at the AT and finally the acyl moiety is transferred to the ACP by a nucleophilic attack from the sulfhydryl group. Beside the ACP, also water can

attack the acylated enzyme forming the free acid, but usually this side reaction is slow compared to transacylation (figure 4).<sup>34,35</sup> The transacylation reaction is an equilibrium reaction so that wrongly loaded substrates, which cannot be condensed by KS, can be transferred back to CoA.<sup>20</sup>

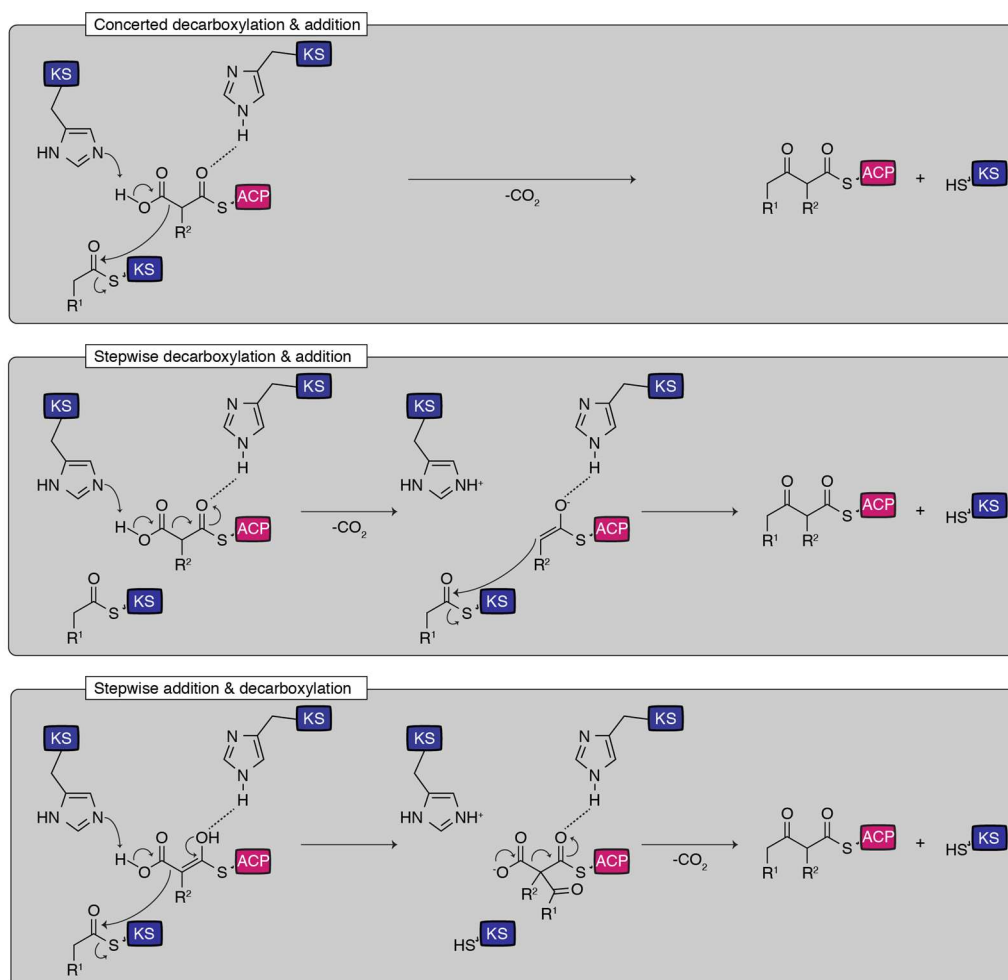


**Figure 4** | Illustration of the AT catalysed transfer from an acyl moiety of acyl-CoA to the ACP. Enzymatic hydrolysis is shown as a side reaction. ACP: acyl-carrier-protein; AT: acyltransferase.

ATs either transfer starter- or extender substrates, and in some cases both in competitive manner. There are four sequence regions described that form the substrate binding pocket and affect the specificity; the RVDVVQ motif, the GHSXG motif, the YASH motif and the C-terminal region.<sup>20</sup> In the mammalian FAS, both starter and extender substrates are transferred by the AT, commonly referred to as malonyl/acetyl transferase (MAT). The enzyme was also shown to be promiscuous to other acyl-moieties.<sup>34</sup> In contrast, the fungal FAS uses two ATs, one domain for the acetyl starter substrate (AT termed acetyl transferase (AT)) and another one for the import of the malonyl extender substrate as well as for the export of the final palmitoyl product (AT termed malonyl/palmitoyl transferase (MPT)).<sup>36</sup> Furthermore, yeast FAS was demonstrated to be able to use various priming substrates.<sup>37</sup> Similar to the mammalian FAS, there are ATs from iterative PKSs predicted to use starter and extender substrates, for example the AT of the methylsalicylic acid synthase (MSAS), which can also accept various starter substrates.<sup>38,39</sup> In contrast, ATs from modular PKSs usually contain an AT specific for starter substrates, referred to as loading-ATs, and an AT specific for the extender substrate, referred to as extender-ATs. These ATs are embedded differently into the megasynthase, as the loading-ATs are positioned at the beginning of the assembly lines included into loading-modules. Although these ATs have a preference for their native substrate, it was demonstrated that some of them are substrate tolerant.<sup>40,41</sup> In contrast, extender-ATs are embedded into elongation modules and usually have narrow substrate specificity.<sup>20,42</sup> Besides substrate specificity, a successful transacylation requires productive interaction between AT and ACP.<sup>17,27</sup>

## 1.2.2 Ketosynthases

The KS catalysed carbon-carbon bond forming reaction represents the key step in polyketide- and fatty acid synthesis. The domain forms a thiolase-like fold, containing an  $\beta\beta\alpha\beta\beta$  topology, and exists as a dimer. The Claisen condensation reaction is catalysed involving a cysteine and usually two histidine in the active centre. As in ATs, the active site is located at the N-terminus of a  $\alpha$ -helix in the nucleophilic elbow, enhancing the nucleophilicity of the cysteine by the helix-dipole. Although the KS is one of the most studied domains within PKSs and FASs, the reaction mechanism is still not fully understood. However, the reaction contains three steps, including the acyl-transfer from ACP to the KS, the decarboxylation of an extender substrate and the condensation with the acyl moiety attached to the KS. Three different mechanisms are possible, which vary in the order of steps in the C-C bond forming reaction: (i) concerted decarboxylation and nucleophilic addition, (ii) stepwise decarboxylation with subsequent nucleophilic addition and (iii) stepwise nucleophilic addition with subsequent decarboxylation reaction (figure 5).<sup>20,43</sup>



**Figure 5** | Possible mechanisms for the Claisen condensation, catalysed by the KS. Acylation of the KS by acyl-ACP takes place before condensation and is not illustrated in this figure. Then the KS catalyse the chain elongation with an extender substrate for which three distinct mechanisms are possible. In the first mechanism the decarboxylation and the nucleophilic addition of the extender

substrate is performed in concerted manner. In the second mechanism the extender substrate gets decarboxylated to an enolate intermediate, which subsequently attacks nucleophilic the acyl-KS moiety. Finally, in the third mechanism, the two steps are swapped in sequence and the extender substrate is added to the acyl-KS followed by decarboxylation. Figure is adapted from Blaquiere *et al.*<sup>43</sup> and was published by Rittner and Joppe *et al.*<sup>44</sup>

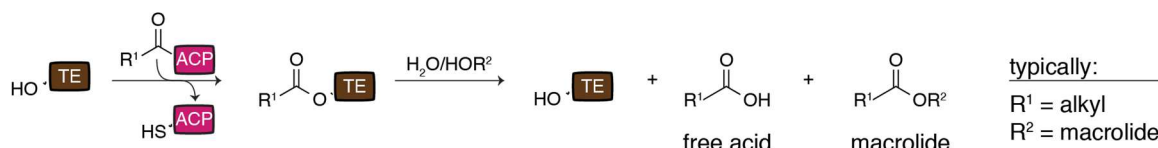
The substrate specificity of the KS plays a crucial role for the biosynthesis of fatty acids and polyketides. In mammalian FASs, the KS is specific for fully saturated acyl moieties and the turnover rates increase with chain length until C12 and decrease then again for longer chains. Therefore, the KS is responsible for the specific production of palmitic acid as elongation only proceed when the substrate is fully reduced and fits to the substrate binding pocket.<sup>45,46</sup> In contrast, KSs from the 6-deoxyerythronolide B synthase (DEBS) showed more relaxed substrate specificity and can accept  $\beta$ -ketoester with various  $\beta$ -functionalization.<sup>42</sup>

### 1.2.3 Thioesterases

The TEs perform the release of covalently attached fatty acids and polyketides from the enzyme as free acids, esters, macrolactone or macrolactams. Active offloading is crucial, because it prevents stalling of intermediates and consequently inactivation of enzymes.<sup>47</sup> The architecture of TEs is build up as a  $\alpha/\beta$ -hydrolase fold with several  $\beta/\alpha/\beta$  motifs and contains a central seven-stranded  $\beta$ -sheet. TE domains usually contain a serine, histidine, and an aspartate in the active site. The nucleophilicity of the active site serine is increased by the arrangement and is positioned in a nucleophilic elbow in a turn between the C-terminus of a  $\beta$ -strand and the N-terminus of a  $\alpha$ -helix.<sup>20,47</sup> The release of fatty acids and polyketides is performed in a two-step mechanism. After docking of the ACP to the TE, the active site serine gets acylated by acyl-ACP and the attached acyl-TE intermediate is subsequently unloaded by attack of an O-, N-, or C-nucleophile (figure 6). Consequently a productive ACP-TE interaction is required to deliver the substrate to the active site of the TE.<sup>47</sup>

The substrates specificity of TEs is important to ensure specific product synthesis and is determined by the shape of the substrate binding pocket and the interactions to the intermediate. These key properties determine when a substrate is loaded on the TE and how the substrate is unloaded from the enzyme.<sup>47</sup> In the human FAS, the substrate channel forms a hydrophobic groove which explains the specificity for a palmitoyl substrate. In contrast, TEs from the DEBS and pikromycin synthase (PIKS) contain an unusually long substrate channel, also consistent with the structure of their products. Key interactions to several functional groups can anchor the polyketide chain to bring the C13-OH in vicinity to the acyl-O-serine, which is required for the macrolactonization.<sup>20</sup> The DEBS TE can tolerate some substrate variabilities,

especially in the vicinity of the hydroxy group, to synthesize the macrolactone but other non-native substrates are hydrolysed from the enzyme as a free acid.<sup>42,47</sup> Particularly important for the cyclization is the D-configuration of the C13-OH, whereas the respective L-configured diastereomer is released as a free acid.<sup>47</sup> In similar manner, a recent study about the PIKS TE demonstrated that shunted intermediates can be cyclized, but diastereomers are unloaded as free acids or cyclohexanones.<sup>48</sup>



**Figure 6** | TE catalysed release of fatty acid and polyketide intermediates from enzymes. The acyl-moiety gets transferred to the active site serine of TEs and subsequently released by nucleophilic attack.<sup>47</sup> Figure is adapted from Heil *et al.*<sup>10</sup>

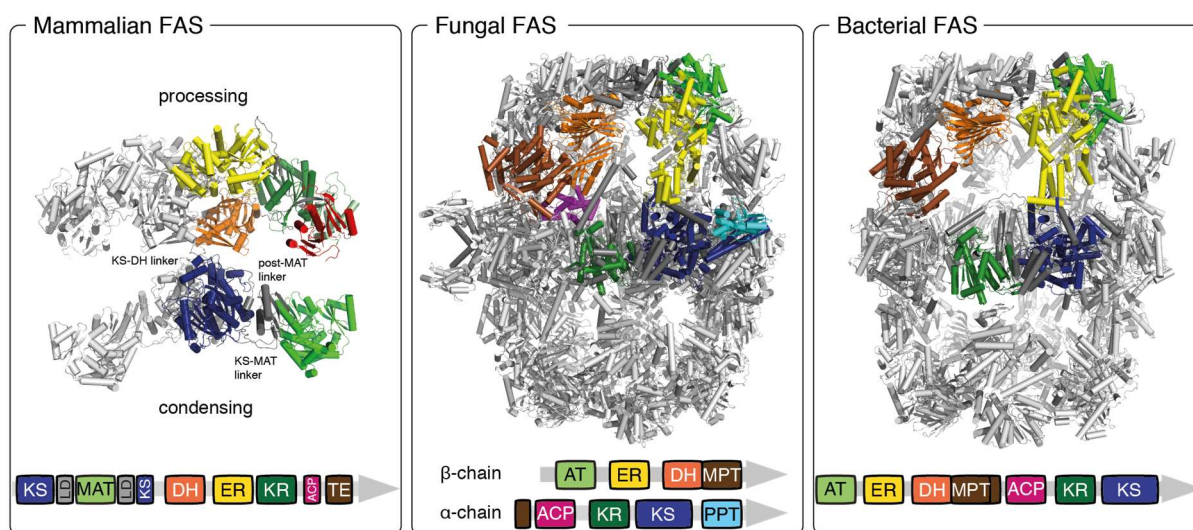
### 1.3 Type I FASs

Type I FASs assemble to large complexes with all enzymes provided as domains on multifunctional polypeptide chains. Within the protein complex, the catalytic domains for fatty acid synthesis are in spatial proximity and consequently the substrate shuttling of ACPs requires only short distances compared to type II systems.<sup>49,50</sup> It is considered that the compartmentalisation improves kinetic efficiency, prevents labile or reactive intermediates from side reactions and is beneficial during protein expression.<sup>23,49,51</sup>

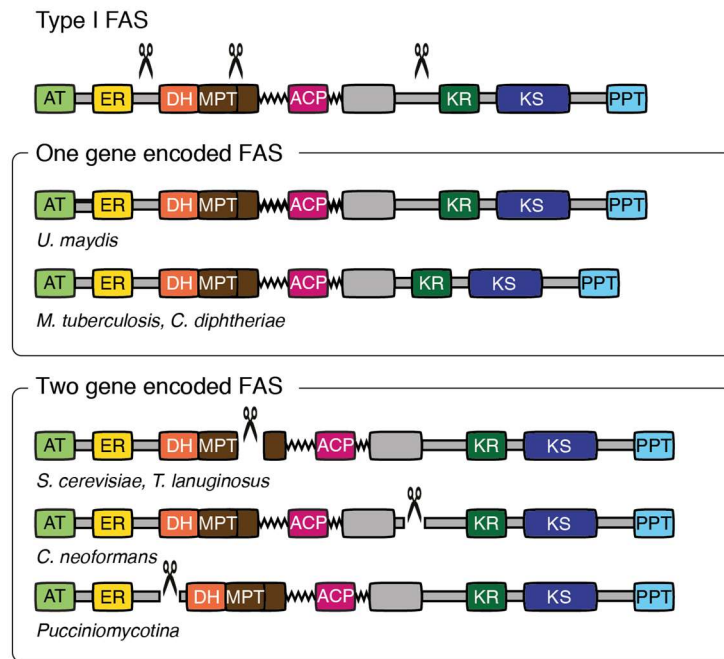
During evolution, type I FASs suggested to evolve from discrete enzymes of the type II FAS system and are a result of gene fusion events including the insertion of scaffolding domains to stabilize their structures.<sup>23,49,50,52</sup> Despite a common ancestor, there are three different architectures known for the type I FASs (figure 7). (i) The mammalian FAS forms a 540 KDa “X-shaped” homodimer with only 10% of amino acids used as scaffolding elements.<sup>53</sup> (ii) In contrast, the fungal FAS is built up as a 2.6 MDa homohexameric ( $\alpha_6$ ) or heterododecameric ( $\alpha_6\beta_6$ ) barrel-shaped complex with D3-symmetry and about 50% of amino acid are involved in scaffolding elements. (iii) The bacterial FAS forms a minimized version of the fungal FAS and consists of a 1.9 MDa homohexameric ( $\alpha_6$ ) complex, missing the PPT domain and structural domains and comprise about 35% of scaffolding elements. The bacterial FAS is always assembled from one polypeptide chain contrary to the fungal FAS, where the domains are distributed either on one or two polypeptide chains. Consequently, it is considered that the fungal FAS was developed by addition of PPT



and structural domains from the bacterial FAS to form a one-gene encoded fungal FAS. The two-gene encoded fungal FAS, in turn, suggested to be a result of gene-splitting at different positions (figure 8).

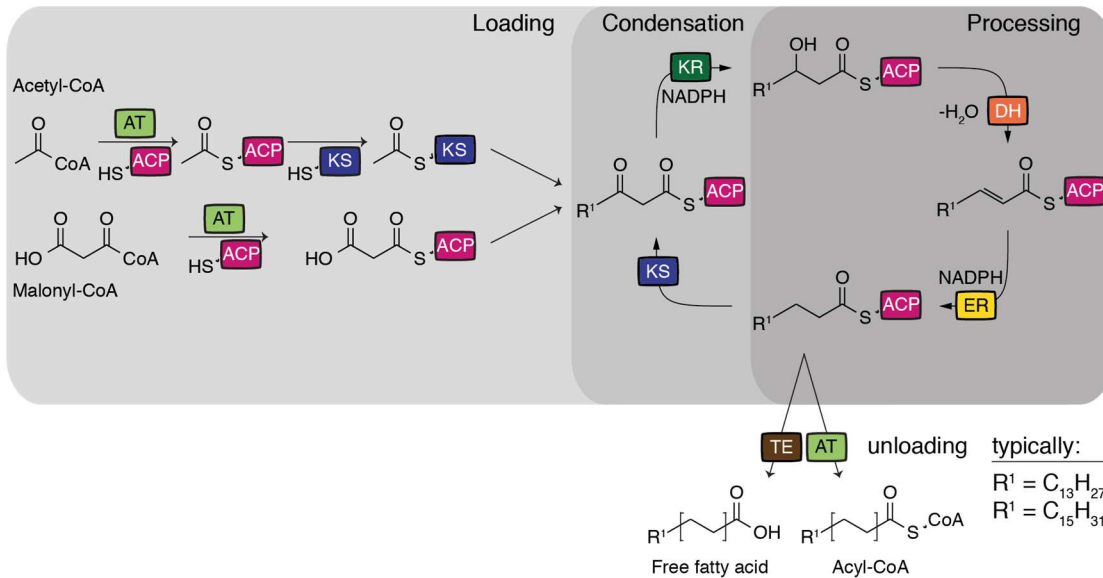


**Figure 7** | Comparison of the different type I FAS architectures. The structures of the mammalian FAS (PDB: 2VZ8), fungal FAS (PDB: 3HMJ) and bacterial FAS (PDB: 4V8L) are illustrated in cartoon presentation.<sup>36,53–55</sup> One set of catalytic active domains is highlighted in each structure according to the colours of the domain representation below and scaffolding domains are highlighted in dark grey. Figure is adapted from Gringer.<sup>49</sup>



**Figure 8 |** Comparison of the domain organisation of bacterial and fungal FASs. Catalytical active domains are highlighted in colours and structural domains are depicted in grey. Scissors indicating gene splitting events. Figure is adapted from Grninger.<sup>49</sup>

Although FASs appear in different kingdoms of life with various architectures, the fatty acid synthesis is chemically highly conserved (figure 9).<sup>10</sup> They use acetyl-CoA, malonyl-CoA (Mal-CoA), and NADPH to produce fatty acids or CoA-Ester in a repetitive cycle. The growing acyl-chain intermediate is always fully reduced before it enters another round of extension giving the alkyl unit. One major difference between the different type I FASs is the offloading of the fatty acid acyl chain from the ACP that usually contains 16 or 18 carbon atoms after the last cycle. In the mammalian FAS the acyl-chain gets hydrolysed with water by a TE domain to release the acyl chain as a free fatty acid. In contrast, the bacterial and fungal FAS use an AT to transfer the intermediate to CoA.<sup>10</sup>

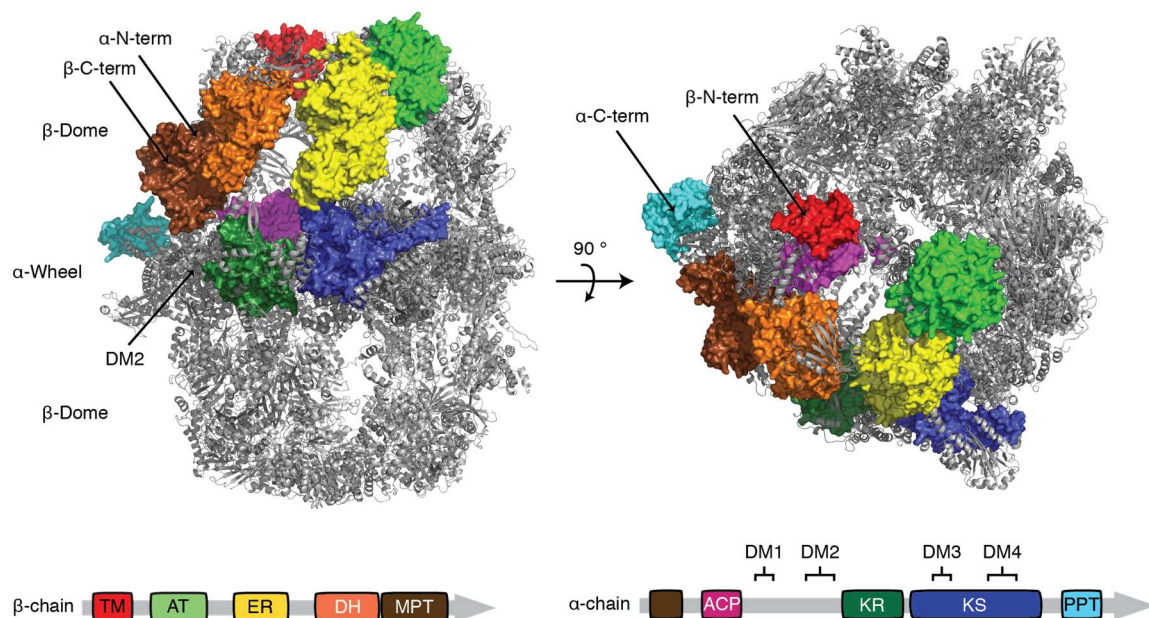


**Figure 9** | Illustration of the fatty acid biosynthesis. The moieties acetyl and malonyl are transferred to the ACP by an AT. After the acetyl-moiety is loaded on the KS, the chain gets elongated by malonyl-ACP in a Claisen condensation reaction. The produced  $\beta$ -keto intermediate is successively reduced by KR, DH, and ER to the fully saturated acyl-chain. Then the acyl-chain is transferred back to the KS for another round of extension. This cycle usually repeats 7-8 times until the chain gets terminated by either hydrolysis by a TE or transacylation to a CoA by an AT. Figure is adapted from Gringer.<sup>49</sup>

### 1.3.1 Fungal FAS

The fungal FAS was first solved in structure in 2006 by X-ray crystallography and later also with cryo electron microscopy (cryoEM) for the organisms *Saccharomyces cerevisiae* (*S. cerevisiae*), *Thermomyces lanuginosus* and *Rhodospiridium toruloides*.<sup>36,54,56–59</sup> The fungal FAS builds a 2.6 MDa D<sub>3</sub>-symmetric barrel-shaped structure, comprising a central  $\alpha$ -wheel, made up by six  $\alpha$ -polypeptide chains, that is flanked by two  $\beta$ -domes each formed by three  $\beta$ -polypeptide chains (figure 10). The  $\alpha_6\beta_6$  heterododecameric FAS barrel is a rigid cage with minor conformational flexibilities<sup>49</sup> and the structure is stabilized with dimerization modules (DM1-4) at the central wheel and a trimerization module (TM) closing the  $\beta$ -domes at the apical sides.<sup>9,36</sup> The biosynthesis of fatty acids is performed within the two reaction chambers on each side of the barrel, each containing three full sets of catalytic domains. The ACP is anchored to the MPT and the central hub of the equatorial wheel with two flexible linkers in the inside of the reaction chamber. Substrate shuttling is achieved by the flexibility of the ACP domain.<sup>49,57,58</sup> All enzymatic domains involved in the fatty acid cycle have their binding tunnels directed to the inside of the barrel-shaped structure<sup>36,56,57,60,61</sup>, allowing the ACP to approach the

active sites. The 4-phosphopantetheinyl transferase (PPTase of yeast FAS termed PPT) is the only domain that was found attached at the outside of the barrel-shaped structure.<sup>54</sup> In all available structures the density for the PPT was weak because of its high flexibility. In 2009, the crystal structure of the PPT domain, crystallized as a separate protein, was solved and docked into this weak density to obtain a fully updated structural model.<sup>54</sup> The PPT is responsible for the post-translational activation of the ACP domain and because of the spatial separation, the ACP/PPT interaction must occur in a phosphopantetheinylation active intermediate during the assembly.<sup>62</sup> The PPT domain is active as a dimer and trimer, which implies that the intermediate must have a trimeric, dimeric or heterotetrameric ( $\alpha_2\beta_2$ ) architecture.<sup>49,54</sup> As discussed, it is considered that the two gene encoded fungal FAS has evolved from the bacterial FAS and consequently it was assumed that an early interaction of the  $\beta$ -chain with the  $\alpha$ -chain is required for the assembly.<sup>49</sup> This was confirmed in a recent study, where co-translational assembly was demonstrated by ribosome profiling and showed that the  $\beta$ -polypeptide chain interacts with the nascent  $\alpha$ -chain during translation to form  $\alpha\beta$ -hetero-dimers and initiate assembly.<sup>63</sup> This explains an earlier observation which showed that the  $\beta$ -chain has an influence on the transcription of the  $\alpha$ -chain, which also suffers degradation in the absence of the  $\beta$ -chain.<sup>64</sup> The assembly of the fungal FAS is biotechnologically and medically relevant, as it provides guidelines for engineering fungal FASs as well as paves the way for an assembly-based inhibition strategy.



**Figure 10** | Structure of the yeast FAS (PDB: 6TA1). One set of catalytic active domain is highlighted, according to the colours of the domain representation below. Note that the DM2 is not

resolved in the structure and therefore its position is indicated with an arrow. Additionally, also the termini of both chains are highlighted with arrows.

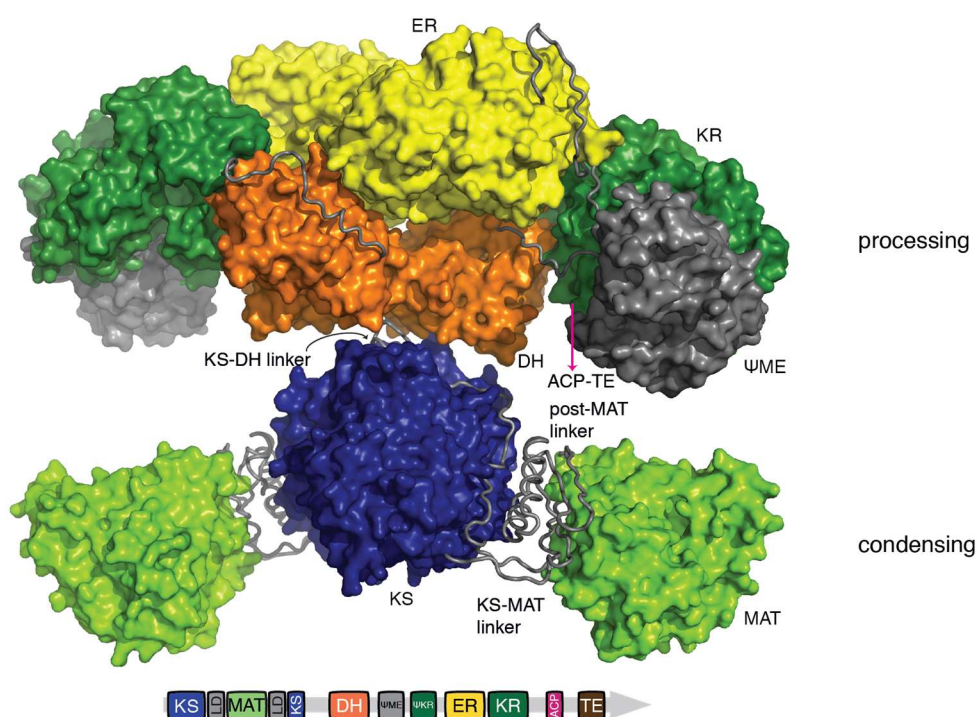
### 1.3.2 Mammalian FAS

The mammalian FAS (from porcine) was structurally solved by x-ray crystallography as a 540 KDa intertwined “X-shaped” homodimer of two polypeptide chains (figure 11).<sup>46,53</sup> It is segregated into a condensing part, containing the KS and MAT, and a processing part, containing DE, ER, KR, ACP and TE. The condensing and the processing part are linked with each other in a narrow waist, referred to as the KS-DH linker.<sup>9,10</sup> The domains ACP and TE could not be traced in electron density of the mammalian FAS x-ray structure, but structural information is available on the separate proteins.<sup>65,66</sup> The ACP shuttling of substrates and intermediates is achieved by high conformational flexibilities. As demonstrated by mutational studies, the ACP can interact with the KS and MAT from the other polypeptide chain.<sup>67</sup> This observation was further confirmed by negative staining EM, revealing the swivelling and swinging of condensing and processing parts around the narrow waist.<sup>49,68,69</sup> The KS-MAT structure of murine FAS (mFAS), also solved by x-ray crystallography<sup>34</sup>, has a similar architecture as the porcine FAS. The two MAT domains are positioned at the outside of the condensing part and consequently exist as monomeric domains. They are connected to the central KS dimer via a KS-MAT  $\alpha/\beta$  fold linker domain (LD). The LD contains a post-KS and post-MAT linker with the latter running along the perimeter of the KS dimer and ending at the narrow waist region.<sup>9</sup>

The MAT of the mammalian FAS transfers the starter acetyl-CoA, but also the elongation substrate Mal-CoA to the ACP domain. The active site is positioned in a cleft between the  $\alpha/\beta$ -hydrolase-like fold and the ferredoxin-like subdomain. An arginine residue at the bottom of the active site tunnel interacts with the carboxylic group of extender substrates and consequently was assigned to be crucial for the bifunctional role.<sup>46,70</sup> The specificity was shown to be relaxed as the MAT is able to transfer various starter- and extender substrates.<sup>34</sup> The promiscuity was attributed to the conformational variability of the  $\alpha/\beta$ -hydrolase and the ferredoxin-like subdomain. Rigid body movement and rotational movement was observed for the ferredoxin-like subdomain as well as significant main-chain torsion angles of several amino acids within the linker to this subdomain. Additionally the conserved arginine residue has high positional variability as it contains a less extended H-bond network compared to related ATs.<sup>34,46</sup>

The relaxed substrate specificity makes this domain an excellent target for protein engineering, as it allows loading of a variety of substrates on FASs or PKSs. Additionally, the MAT shows fast transacylation rates combined with low hydrolytic

activity, demonstrating the efficiency of this domain. The domain can be expressed as a standalone domain and with point mutations, indicating the robustness of the fold and the possibility to use it as a trans-acting AT. The robustness of a protein fold is a major requirement to change substrate specificities through protein engineering because it allows to design stable mutants.<sup>71</sup> Furthermore, the position of the monomeric MAT at the outside of the FAS structure makes it an easy target for domain exchange approaches, which has already been successfully demonstrated for short chain ketone production by Yuzawa et al.<sup>28,34,46</sup>



**Figure 11** | Domain structure and architecture of the mammalian FAS (PDB: 2VZ8).<sup>53</sup> The structure is divided into a condensing and a processing part, connected by KS-DH linker in the narrow waist. The condensing part contains the KS and the MAT, connected by a KS-MAT linker. The processing part includes the DH, ER, KR, ACP and TE with the latter two not solved in the x-ray structure. Figure is adapted from Rittner et al.<sup>23</sup>

## 1.4 Type I PKSs

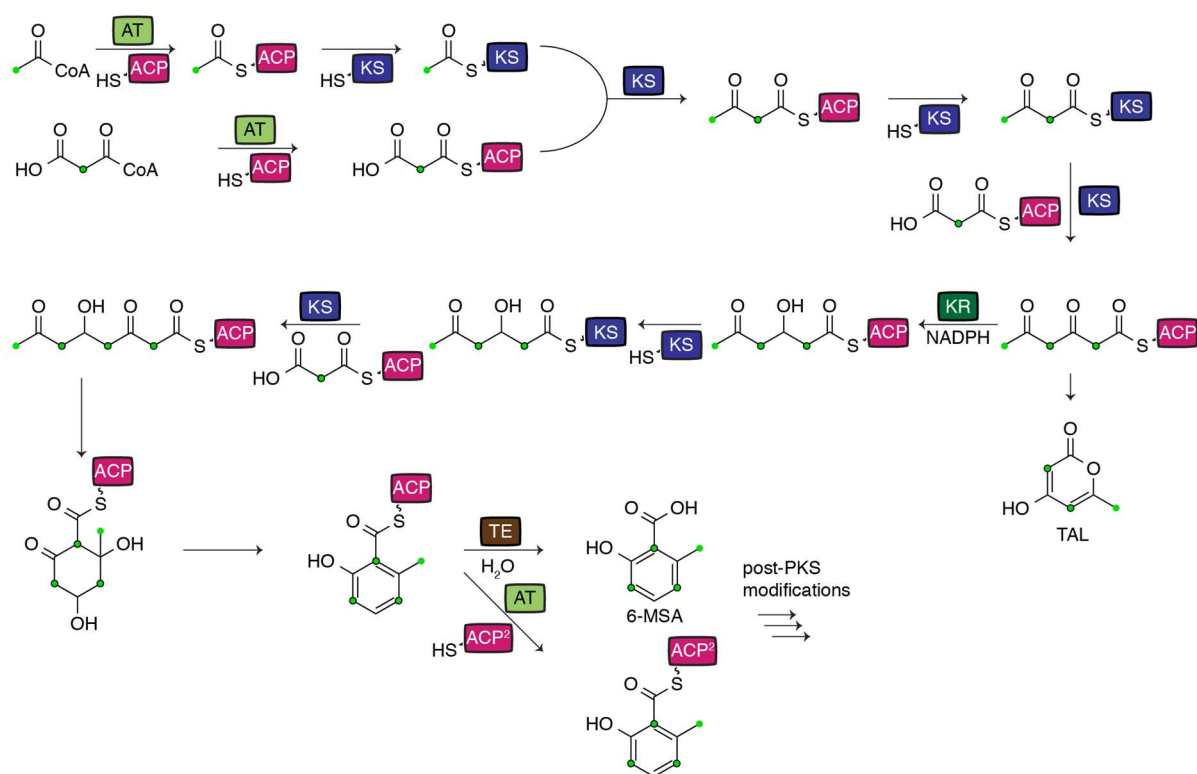
Type I PKSs are involved in the biosynthesis of many important natural products, including patulin, chlorothricin, erythromycin and pikromycin.<sup>13,72</sup> In nature, these polyketides are produced by either iterative or modular PKSs and have been targeted in many studies, because of the high relevance of their products.<sup>17</sup> While structural information was obtained for truncated domains, full-length proteins and modules

have not yet been described at high resolution. To date, two models for the architecture of PKSs have been described with considerable differences: (i) The mammalian FAS is evolutionary related to PKSs and it has been used as a structural model for PKSs, assuming that it represents a fully reducing protein version.<sup>17</sup> Many substructures of individual domains and multidomain constructs, particularly from the mycocerosic acid synthase obtained by x-ray crystallography, supported the similarity of FASs and PKSs.<sup>73,74</sup> Recently, also the lovastatin synthase structure obtained by cryo-EM, suggested a mammalian FAS like structure of type I PKSs.<sup>75</sup> (ii) The second structural model for PKSs derives from structural studies on the fifth module of the PIKS, obtained by cryo-EM. Here a different arrangement of domains in an overall arched conformation was presented.<sup>17,73,75,76</sup>

### 1.4.1 Methylsalicylic acid synthase

6-Methylsalicylic acid (6-MSA) is a secondary metabolite used as a precursor molecule for the biosynthesis of biological active compound like patulin, chlorothricin, maduropeptin and neocarzinostin with antibiotic, antitumor and antibacterial activities<sup>72,77-80</sup>. It is produced by the MSAS, a partly reducing iterative PKS using one molecule of acetyl-CoA and one of NADPH, and three molecules of Mal-CoA for 6-MSA biosynthesis (figure 12).<sup>81</sup> MSAS consists of domains KS, AT, TE, KR and ACP residing on a single polypeptide chain (from N to C).<sup>82</sup> Similar to the mammalian FAS, the AT of MSAS is bifunctional and responsible for the selection and transfer of the starter- and elongation substrates.<sup>39</sup> The starter acetyl unit is elongated twice to the triketide, which is subsequently reduced by the KR followed by another elongation step. 6-MSA formation finally results from aldol-condensation under release of water, although the exact mechanism is still not known. ACP-bound 6-MSA is eventually either cleaved off by hydrolysis or directly transferred to another ACP where the polyketide can be tailored to the final compound.<sup>83-87</sup> The TE is involved in the product release, which was demonstrated for MSAS from *Aspergillus terreus* (*A. terreus*) with an artificial 6-MSA ester.<sup>82</sup> *Moriguchi et al.* suggested that the hydrolytic cleave-off of 6-MSA is catalysed by an aspartate and a histidine within the TE active site. However, it is still speculated that the domain also acts as a product template domain involved in cis-trans isomerization or cyclization.<sup>72</sup> In the absence of NADPH, the unstable triketide is omitted from further elongation and spontaneously released as triacetic acid lactone (TAL).<sup>81</sup> The timing of the reduction within the 6-MSA biosynthesis, but also the evidence that TAL is produced as a shunt product, indicates strict specificities of the KS and KR.<sup>87,88</sup> KR specificity was demonstrated for the related R-mallein synthase, which only accepted a diketide and tetraketide substrate.<sup>89,90</sup>

In the 1970s, MSAS from *Penicillium patulum* (*P. patulum*) has been purified in its native host and later also from *Escherichia coli* (*E. coli*) and yeast. As purified protein, MSAS was characterized in its basic function.<sup>81,91</sup> Remarkably, 6-MSA was recently produced in an overall yield of about 2 g/L in an engineered *S. cerevisiae* strain.<sup>92</sup> Additionally, production of m-cresol was achieved by co-expression of MSAS with a decarboxylase in an engineered yeast strain.<sup>92</sup> MSAS from *P. patulum* was also able to convert non-native starter substrates *in vitro*, but at significantly slower turnover rates, which was confirmed by the synthesis of m-cresol derivatives *in vivo*.<sup>38,92</sup> For the latter approach, non-native starter substrates were provided by the yeast strain. Specifically, propionyl-CoA and butyryl-CoA were provided by metabolic engineering leading to 6-ethyl salicylic acid (6-ESA) and 6-propyl salicylic acid (6-PrSA), respectively, which were subsequently decarboxylated to 3-ethylphenol (3-EP) and 3-propylphenol (3-PrP). The yields of 3-EP and 3-PrP were significantly lower compared to 3-methylphenol (3-MP), demonstrating that the enzyme has high preference for its natural starter substrate acetyl-CoA.<sup>93</sup>



**Figure 12 |** Biosynthesis of 6-MSA. In the first step the AT loads a starter acetyl moiety to the ACP, which is subsequently transferred to the KS. Then AT-mediated loading of a malonyl extender unit to the ACP is followed by the first chain elongation of the starter acetyl unit with malonyl-ACP. This Claisen elongation is repeated until the triketide is formed, which is subsequently reduced by the KR with the co-factor NADPH. TAL is formed as a shunt product if NADPH is absent. After another elongation step with malonyl-ACP, 6-MSA-ACP is formed by aldol-condensation and dehydration (mechanistic details are still unknown). Finally, 6-MSA is either released from the enzyme by TE-mediated hydrolysis or by AT mediated transfer to another ACP. In both cases the polyketide is tailored to final bioactive compounds (post-PKS modifications). Green dots visualize positions

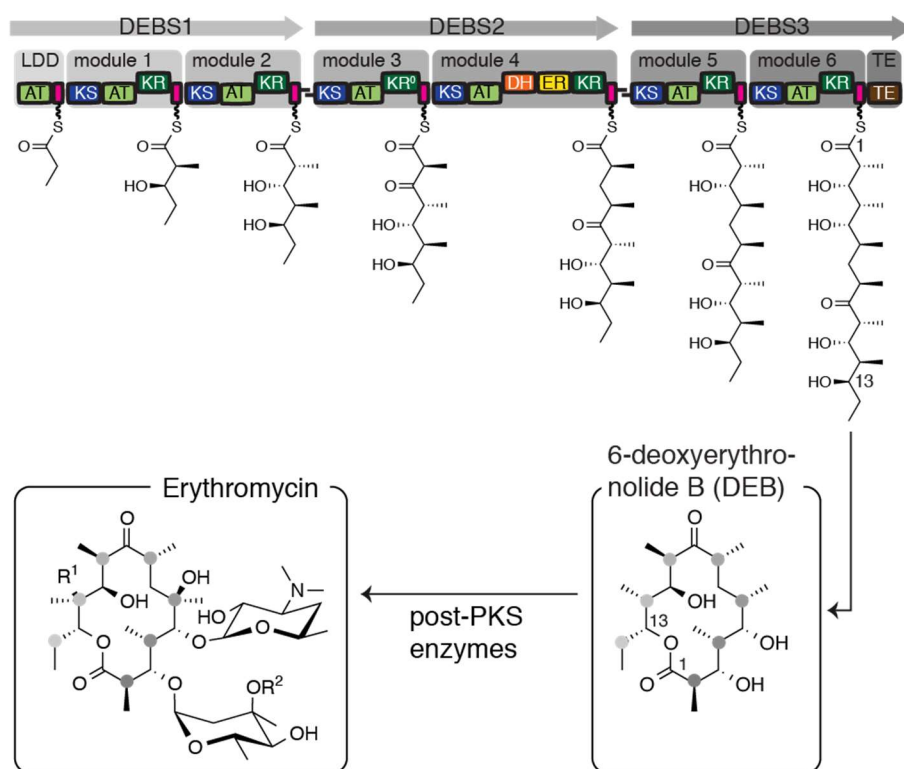


determined by the choice of starter substrates (green dot) and elongation substrates (green dots with black circles).

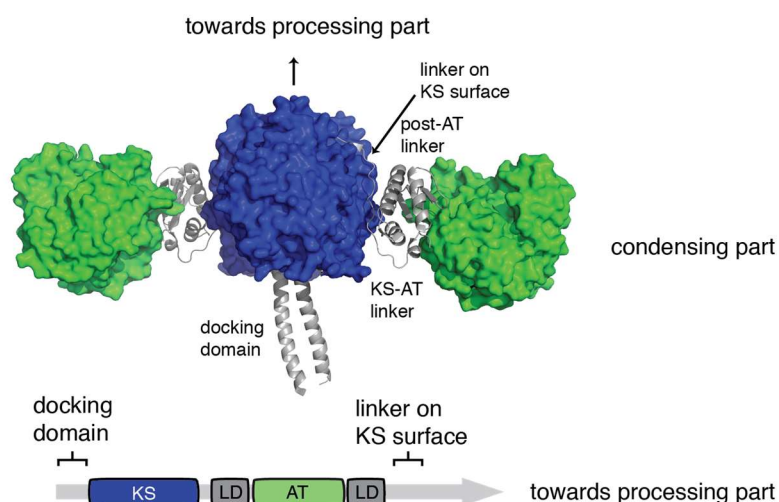
### 1.4.2 6-Deoxyerythronolide B synthase

DEBS is the currently best understood modular PKS. It produces 6-deoxyerythronolide B (DEB), the parent molecule of the antibiotic erythromycin.<sup>94</sup> DEBS consists of six modules distributed on three polypeptide chains, referred to as DEBS1, DEBS2 and DEBS3. Each polypeptide chain harbours two modules. DEBS1 contains an additional N-terminal loading didomain (LDD) and DEBS3 an additional C-terminal TE for macrolactonization (figure 13). The three polypeptide chains interact through docking domains that assemble the polypeptide chain (DEBS1-3) in the right order. The loading domain is responsible for accepting the starter propionyl-CoA unit and consist of a specialized AT and an ACP. In each module, the propionyl unit gets successively elongated with a methylmalonyl moiety and the condensation product is optionally processed. To do so, each module contains a minimal set of domains for the chain extension, i.e., AT, KS, and ACP, as well as optionally the domains KR, DH, and ER for the processing of the  $\beta$ -keto group. The elongated and processed intermediate is subsequently translocated to the downstream module for another elongation step. The substrate is passed through all six extension modules and finally cyclised by the C-terminal TE to the macrolactone DEB.<sup>95</sup>

During the past years, structural data has been made available by x-ray crystallography for several individual domains of DEBS; most interestingly for this thesis, the KS-AT structure of module 5 (figure 14).<sup>74</sup> The KS-AT didomain is structurally similar to the mammalian FAS and consist of a monomeric AT, connected to the homodimeric KS by a KS-AT linker region. The narrow substrate specificity of the AT domain for methylmalonyl-CoA (MM-CoA) originates for a specialized binding pocket, including Arg667 for coordinating the  $\alpha$ -carboxylate, Tyr742, Gln643 and His745 for controlling the transfer of 2S-substrates, and a Tyr742 for interacting with the  $\alpha$ -methyl group.



**Figure 13** | Biosynthesis of DEB by the modular PKS DEBS. DEB synthesis is initiated by the transfer of a propionyl moiety to the ACP of the loading domain, and its successive elongation and processing by the action of six modules. Finally, the C-terminal TE domain performs the lactonization to DEB, which is subsequently converted to erythromycin by post-PKS enzymes. Grey arrows indicate the three polypeptide chains. Each module is highlighted by grey boxes with increasing grey scale. Docking domains are responsible for the interaction of module 2 with module 3 and module 4 with module 5 and are highlighted with black lines. The order of assembled building blocks during the biosynthesis is illustrated by grey dots with increasing grey scale starting with the first building block in light grey. The figure was created together with Alexander Rittner (Goethe-University, Frankfurt).

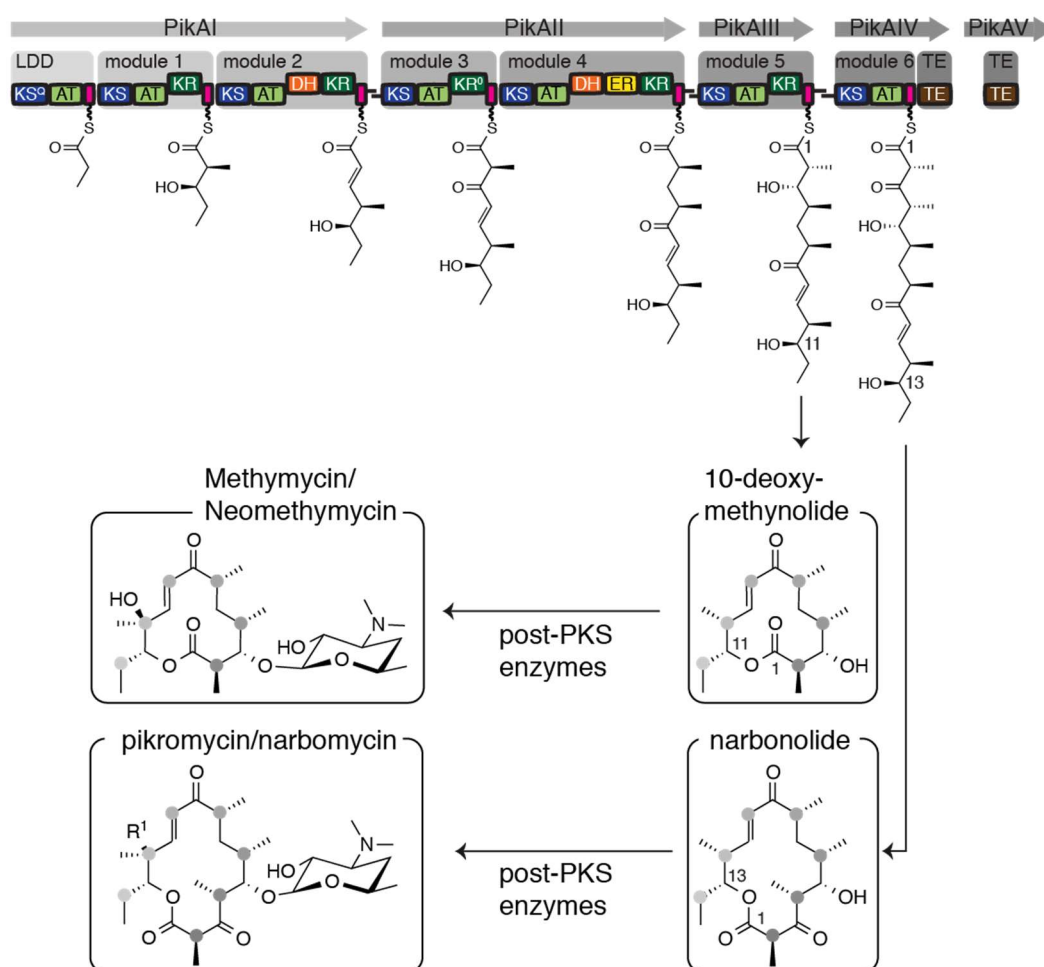


**Figure 14** | Domain structure and architecture of the KS-AT didomain of DEBS module 5 (PDB: 2HG4).<sup>74</sup> The structure contains two monomeric ATs, connected to the dimeric KS by a LD, consisting of a KS-AT and a post-AT linker. Subsequently, a linker on the KS surface leads to the processing part which is indicated with an arrow. Additionally, the structure comprises a docking domain, responsible to interact with the C-terminal docking domain of module 4.

### 1.4.3 Pikromycin synthase

The PIKS produces 10-deoxymethynolide and narbonolide, the parent molecules of the antibiotics YC-17, methymycin, neomethymycin, narbomycin and pikromycin.<sup>96</sup> PIKS is a modular PKS and in the following only differences to DEBS are discussed. PIKS consists of six extension modules like DEBS, but it contains an additional C-terminal TE domain (figure 15). In contrast to DEBS, the six modules and the additional TE are distributed on five polypeptide chains, PikAI, PikAII, PikAIII, PikAIV and PikAV. PikAI and PikAII contain the loading module and the first four modules. Module 5, module 6 and the additional TE are each provided on separated polypeptide chains. Unlike DEBS, PIKS contains a KS in the loading module in which the active site serine is replaced by a glutamine. For the animal FAS it was shown that this mutation eliminates the condensation reaction and increases the decarboxylation rate of malonyl moieties by more than two orders of magnitude.<sup>97</sup> Therefore, the serine to glutamine mutation improves the availability of starter substrate by decarboxylation of malonyl and methylmalonyl to acetyl and propionyl, respectively. The KS action is supported by a N-terminal AT that is specialized to transacylate extender substrates.<sup>20,70</sup> PIKS is differently equipped with processing domains, leading to the different functionalization of the product. Module 2 contains an additional DH and in module 6 the KR is absent. A unique feature of PIKS is the

ability to synthesize a 12-membered and 14-membered macrolactone. This is possible by an alone-standing (separate) TE protein (PikAV) that hydrolysis the hexaketide in trans before further elongated by module 6.<sup>96</sup>



**Figure 15** | Biosynthesis of 10-deoxymethynolide and narbonolide by PKS. Product synthesis is initiated by decarboxylation of the methylmalonyl moiety to receive propionyl, which is then successively elongated by the action of six extension modules. Finally, the C-terminal TE domain or a separate TE perform lactonization. The macrolactones are further converted to methymycin, neomethymycin, narbomycin and pikromycin by post-PKS enzymes. Grey arrows indicate the five polypeptide chains, and each module is highlighted with a grey box in increasing grey scale. Docking domains are responsible for the interaction of module 2 with module 3, module 4 with module 5 and module 5 with module 6 and are highlighted with black lines. The order of assembled building blocks during the biosynthesis is illustrated by grey dots with increasing grey scale starting with the first building block in light grey. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>

## 1.5 Engineering of megasynthases

Protein engineering of megasynthases has enormous potential as it enables biosynthesis of novel natural product derivatives with putative technological and

medicinal properties. To date, the lack of detailed knowledge about enzyme kinetics, as well as structural and conformational properties prevents the generation of widely applicable protocols and therefore, protein engineering approaches often fail.<sup>17</sup> In the following, general engineering strategies will be presented with the main focus on the fungal FAS and ATs from PKS systems.

The majority of engineering strategies targeted the modulation of protein-protein interactions, the mutagenesis of active sites, and the replacement, insertion or removal of domains or complete modules.<sup>17,27</sup> The fungal FAS is a biotechnologically relevant scaffold for the biosynthesis of platform chemicals, oleochemicals and biofuels and various studies used the fungal FAS to produce fatty acids and fatty alcohols.<sup>10,11</sup> Engineering approaches demonstrated that the product spectrum of the yeast FAS can be enlarged by active site mutagenesis or by insertion of additional domains, which enabled biosynthesis of short fatty acids and even polyketides.<sup>10,98–100</sup>

PKSs have been a major target for engineering with the objective to produce novel polyketides. PKS engineering offers the possibility to access complex polyketide scaffolds without challenging synthetic chemical methods and protocols.<sup>27</sup> Many studies approached the ATs as the gatekeepers selecting the substrates incorporated into the final product. Usually ATs from PKSs have narrow substrate specificities and do not allow the loading of non-native substrates.<sup>20,42</sup> Therefore, the ATs of loading modules were exchanged in various studies, either by just replacing the AT domain of the module or by replacing the whole loading module.<sup>17</sup> Analogously, domain exchanges of ATs from extender modules have been reported. Recently, Yuzawa *et al.* presented several designs for the replacement of an AT that varied in their junction site. The AT was exchanged with the full, with parts of, or completely without the adjacent KS-AT LD and the resulting chimeras had different enzymatic efficiencies. Yuzawa *et al.* concluded that the best strategy for AT exchanges includes the whole KS-AT linker.<sup>28</sup> Other studies used a combination of an AT knockout with the addition of a trans-ATs to create polyketide derivatives. As probably most interesting example, this strategy led to the fluorination of simple polyketides by loading enzymatically synthesized fluoromalonyl-moieties.<sup>30</sup> Finally, other strategies approached AT design by active site engineering, demonstrating that mutagenesis of the binding pocket can alter substrate specificities. Several studies showed successful incorporation of non-native substrates with different PKSs, including PIKS and DEBS.<sup>29,31</sup>

It is worth mentioning that the loading of non-native substrates can affect condensation, processing and termination reactions so that also the specificities of the downstream domains need to be considered in AT engineering.<sup>31,48,101</sup>

## 1.6 Aim of the thesis

Engineering of megasynthases, such as PKSs and FASs, gives access to the biosynthesis of natural product derivatives, with designed and improved properties for biotechnological and medicinal applications. However, engineering approaches often lead to loss of enzymatic activity, mainly caused by a lack of knowledge about the key structural and functional properties of these enzymes. The goal of this thesis was to establish broadly applicable guidelines for the engineering of megasynthases by both characterizing as well as the rationally designing representatives of the FAS and PKS family. The megasynthases of interest in this thesis were fungal and mammalian FAS, MSAS and DEBS:

(1) The yeast FAS represents a biotechnologically relevant scaffold for the biosynthesis of fatty acids. Our goal was to establish a vector-based expression system that allows straight-forward access to purified FAS variants in sufficiently high quality for structural and biochemical characterization. The rapid availability of structural data could support engineering approaches, which leads to increased success rates in the design of enzymatically active enzymes with desired product output. For engineering fungal FAS, additional enzymes should be attached to the scaffold and this thesis had the goal to identify suitable positions for enzyme fusions.

(2) MSAS is an iterative PKS synthesizing 6-MSA as a precursor for biological active compounds with antibiotic, antitumor and antibacterial activities. Additionally, 6-MSA can be enzymatically converted to 3-MP which is structurally related to compounds like 3-EP and 3-PrP.<sup>93</sup> These biologically active compounds are used in traps for the Tsetse-fly to control its population. The Tsetse fly transmits the parasites of the *trypanosomes* genus causing human sleeping sickness and animal trypanosomiasis, a deadly disease that threatens 70 million people and leads to the depletion of feedstocks in sub-Saharan Africa.<sup>102,103</sup> To harness MSAS as a catalyst for the biosynthesis of 3-EP and 3-PrP, the megasynthase has to be engineered to not just accept acetyl-CoA as substrate, but also longer starter substrates like propionyl-CoA and butyryl-CoA. Therefore, we targeted MSAS for biochemical characterization with a focus on substrate specificities of individual domains.

(3) In the final project of this thesis, we focused on the well described MAT from the mammalian FAS as a target for engineering. A previous study in our group showed that the MAT can transacylate a broad range of substrates with fast turnover rates. In contrast, ATs from PKS systems generally have strict substrate specificities and are comparably slow. We aimed to create PKS/FAS hybrids that are based on a PKS module in which the cognate AT is replaced by the MAT of mFAS. In doing so, we intended to integrate the superior enzymatic features of the MAT into modular PKSs for the incorporation of non-native and especially fluorinated substrates into the polyketide synthesis.

## Chapter 2: Results and analysis of the data

During the PhD various projects were processed with different megasynthases. This chapter describes the results, analysis and interpretation of the data and is categorized into subchapter focusing on the individual projects on FASs and PKSs. The chapter contains published and so far unpublished data.

The first subchapter (chapter 2.1.1) describes a protocol, established to prepare high quality yeast FAS samples for structural analysis. The sample was used as model system to analyse and improve sample preparation for cryo-EM (chapter 2.1.2). The improved protocol led to a complete structural model of the yeast FAS (chapter 2.1.3), which was utilized for a mutational study on suggested phosphorylation sites (chapter 2.1.4). Additionally, structural domains of the yeast FAS were engineered with the aim to remodel and expand the yeast FAS scaffold (chapter 2.1.5). The MSAS was targeted for engineering to produce 6-MSA derivatives. MSAS from *Saccharopolyspora erythraea* (*S. erythraea*, MSAS from that organism referred as SerMSAS) was characterized (chapter 2.2) but turned out to be unsuited for our purposes, because it does not release its product from the protein scaffold. In contrast, MSAS from *P. patulum* (PenPaMSAS) was able to release 6-MSA and used for engineering with the objective to improve the biosynthesis of 6-MSA derivatives by targeting the specificity of the AT (chapter 2.3). Finally, a modular PKS was engineered to synthesize new-to-nature compounds of higher complexity. Specifically, we were able to engineer a module of the PKS DEBS by exchanging its AT with the polyspecific murine MAT. The DEBS/FAS hybrid enabled site-specific fluorination of a macrolactone (chapter 2.4).

### 2.1 Characterization of the yeast fatty acid synthesis

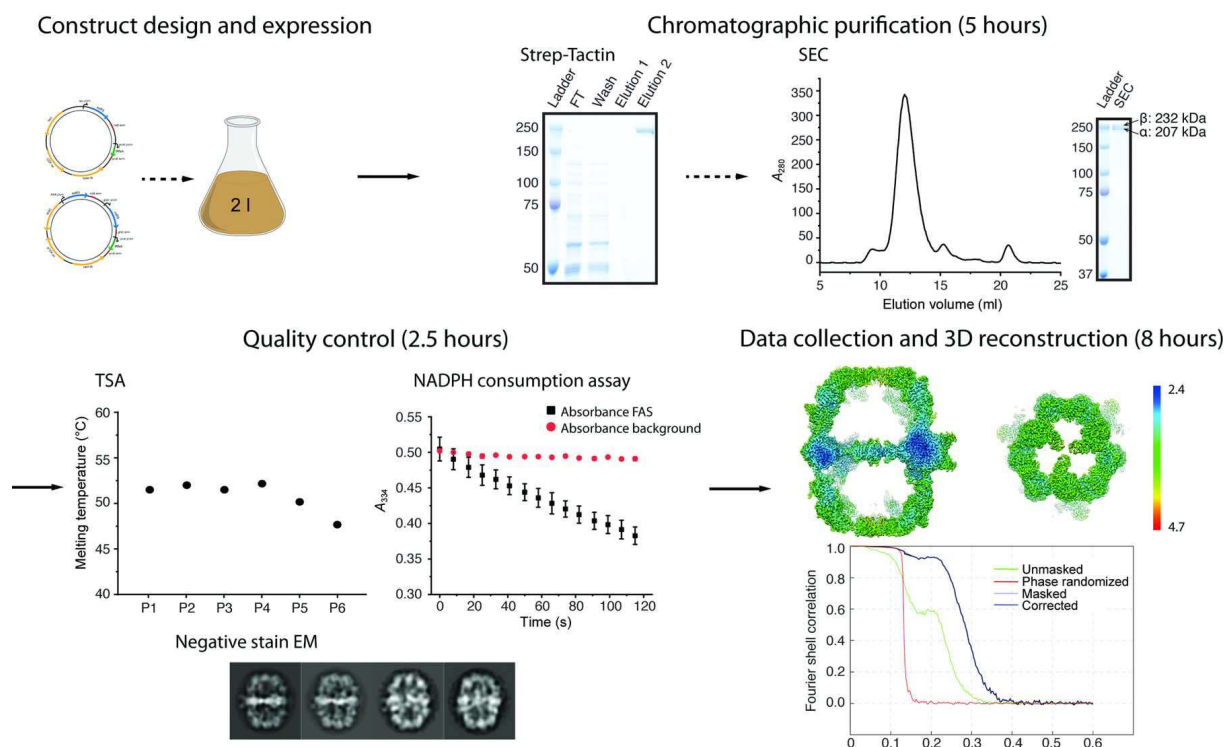
#### 2.1.1 Establishment of a purification strategy for yeast FAS

At the outset of the study, we aimed for establishing a streamlined protocol for the rapid production of purified yeast FAS for structural determination by high-resolution cryoEM. We chose a plasmid-based system for the simple modification of yeast FAS at the DNA level. To keep as closely as possible to physiological conditions, we provided the FAS encoding genes on two single copy number pRS shuttling-vectors (pRS313, pRS315) under the native promoter. A Strep-tag was fused at the accessible C-terminus of the  $\beta$ -chain for rapid chromatographic purification.<sup>104</sup> The position is at the outside of the barrel within the MPT domain (figure 10, chapter 1.3.1) and was used before.<sup>58,61,63,105</sup> We expressed the yeast FAS in an engineered *S. cerevisiae* strain (BY.PK1238\_1A\_KO) with the FAS encoding genes

replaced by a kanMX4 cassette.<sup>99</sup> With this strategy, we were able to use the fatty acid synthesis of the expressed yeast FAS as an auxotrophic marker. Lethal phenotypes that may be induced by FAS mutations are tolerated when external fatty acids are supplied. After cultivation, we lysed the cells by bead disruption and purified the enzyme with affinity and size exclusion chromatography (SEC). The whole purification procedure took less than 5 hours and the quality of the enzyme was subsequently analysed within 2.5 hours by sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE), TSA and enzymatic activity (figure 16).<sup>106</sup> We received an average yield of  $1.4 \pm 0.4$  mg enzyme, which was pure, as demonstrated by SDS-PAGE, and eluted as a monodisperse fraction in the correct appeared molecular weight from SEC. The thermal shift assay (TSA) showed that the enzyme is stable in the purification buffer and higher melting temperatures indicated that phosphate buffers are more suitable for the purification compared to Tris-HCl and distilled water. The enzymatic activity over nine independent samples was  $2100 \pm 300$  mU/mg and is within the range of the best reported purifications, demonstrating that the vector-based strategy, the attachment of a tag and the chromatographic purification is suitable for this enzyme.<sup>59,62,107,108</sup> Additionally, the comparatively low standard deviation demonstrates high reproducibility. To improve the protein quality further, we tried enzymatic cell lysis and different buffer for the purification. Based on similar or lower enzymatic activity, we concluded that the sample quality could not be further improved (data not shown).

In summary, we established a rapid purification strategy for the yeast FAS, which enables fast mutagenesis studies.





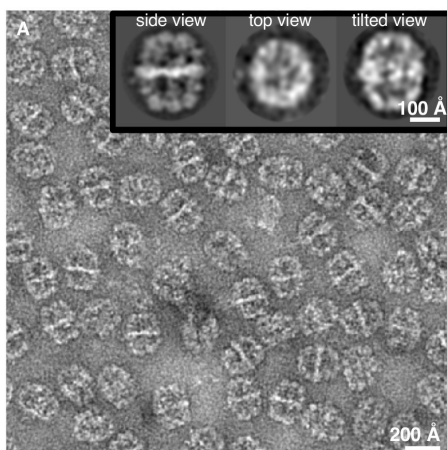
**Figure 16** | Establishment of plasmid-based purification strategy for the yeast FAS. After expression in an engineered *S. cerevisiae* strain (BY\_PK1238\_1A\_KO) using pRS-vectors, FAS was purified in two chromatographic steps with a Strep-tactin column followed by SEC on a S6 Increase 10/300 GL column (in 100 mM sodium phosphate, pH 6.5). The pure sample was received within 5 hours and the quality was analysed within 2.5 hours by SDS-PAGE, TSA and enzymatic activity. The melting temperatures of the yeast FAS in different buffer was analysed by TSAs and varied less than 0.5°C within technical replications (P1: 100 mM NaPi, pH 6.5; P2: 100 mM NaPi, pH 7.4, P3: 100 mM NaPi, pH 8.0, P4: 100 mM NaPi, pH 7.4, 100 mM sodium chloride; P5: 100 mM Tris-HCl, pH 7.4; P6: distilled water). The enzymatic activity was investigated in three technical replicates by NADPH consumption. Shown is a typical experiment with error bars. The specific activity of the sample was  $2310 \pm 48$  mU/mg. Negative staining and cryo-EM experiments were performed and analysed by Edoardo D’Imprima (Max Planck Institute for Biophysics, Frankfurt), which confirmed the excellent quality of the sample and led to a complete map of the yeast FAS at high resolution (3.1 Å). The figure was published by Joppe and D’Imprima *et al.*<sup>109</sup>

### 2.1.2 Sample preparation for cryo-EM

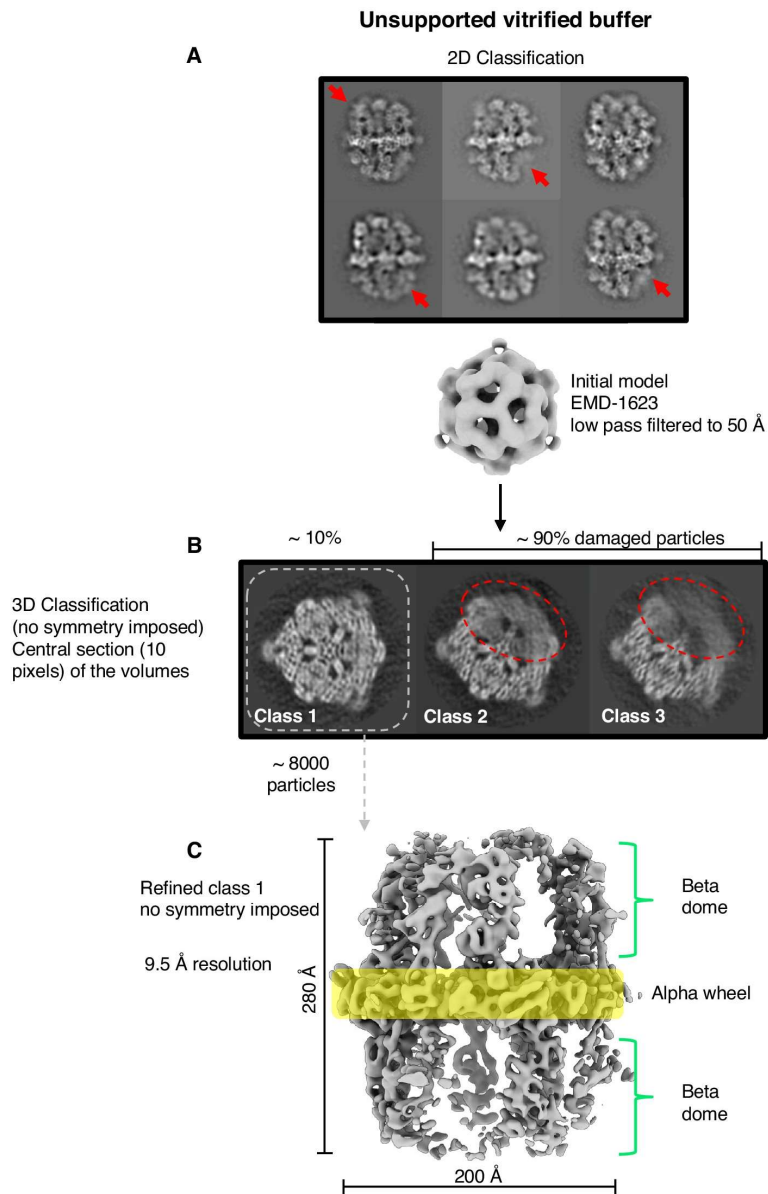
As described by Werner Kühlbrandt, cryo-EM revolutionized structural biology, because of the development of new cameras, detectors, and processing software, and is today the method of choice for large macromolecular complexes like the yeast FAS.<sup>110</sup> Nowadays, it is possible to solve large biomolecule structures at near-atomic resolution, so that cryo-ET is a main driver to expand our understanding about these large complexes. The big advantage of EM is that the sample can directly be applied on grids and has not to be crystallized. Unfortunately, the large complexes are often

unstable and exist in different conformations. Although another huge advantage in cryo-EM is that different conformations can be detected, it also includes the risk that damages, caused by sample preparation, can be misinterpreted as conformational variability. Consequently, damaging the macromolecular complexes during protein and cryo-EM sample preparation must be prevented.<sup>111</sup> In cryo-EM, the percentage of picked particles can be used as indication for good sample preparation. However, many solved structures were received by small percentages of picked particles; in some cases even below 20%.<sup>112–116</sup> With the objective to investigate the cryo-EM sample preparation, we thought that the yeast FAS would be a perfect model system, because the structure has been solved several times at moderate to high resolution since 2006<sup>36,54,56–59</sup>, so that damages of the structure during sample preparation can be easily detected.

We used the established protocol for the preparation of the FAS sample and, because of the high enzymatic activity, we were confident that the quality of the sample is excellent before grid preparation. We applied the purified enzyme on grids and performed negative-staining experiments with 2% (w/v) sodium silicotungstate (Agar Scientific, Stansted, UK) and cryo-EM experiments (glow-discharged Quantifoil R2/2 holey carbon grids, Quantifoil Micro Tools, Jena, Germany). The resulting micrographs and the 2D-classifications from negative stain data showed that the complex was undamaged and confirmed the excellent sample quality (figure 17). In contrast, 2D- and 3D- reconstruction of cryo-EM experiments intriguingly revealed that around 90% of all particles were damaged. The resulting map of the residual 10% (~8000 particles) was with 9.5 Å poorly resolved (figure 18). The particles lacked 33%-50% of electron densities, especially at the outer part of the  $\beta$ -domes. Since the biochemical characterization and the negative-staining EM experiments indicated excellent protein quality prior, we assumed that grid preparation for cryo-EM data collection caused the protein damage.



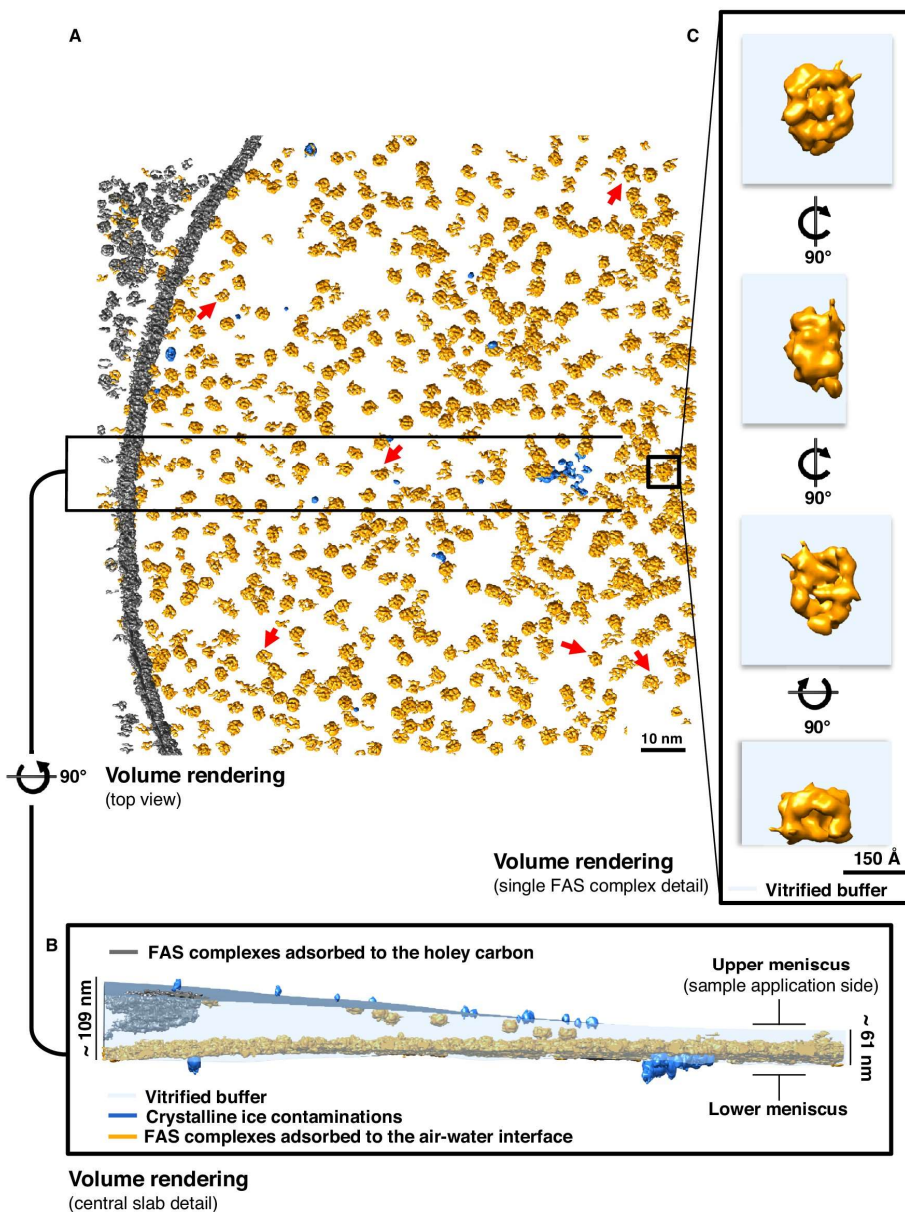
**Figure 17** | Representative micrograph of negatively stained FAS with different representation of 2D class averages. Negative staining EM experiments were performed and analysed by Edoardo D’Imprima (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D’Imprima and Floris *et al.*<sup>117</sup>



**Figure 18** | Single-particle cryo-EM experiments were performed in unsupported vitrified buffer. (A) 2D-classification from the cryo-EM experiments revealed missing or weak densities, especially in the  $\beta$ -domes (red arrows). (B) 3D-classification, focused on the  $\alpha$ -wheel, showed that about 90% of all particles were damaged (red dashes). The residual 10% were used to reconstruct a poorly resolved map. (C) The resulting map was reconstructed from ~8000 particles and was limited to a resolution of 9.5 Å with no symmetry imposed. Cryo-EM experiments were performed and analysed by Edoardo D’Imprima (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D’Imprima and Floris *et al.*<sup>117</sup>

In following our hypothesis that grid preparation caused the particle damage, we performed cryo-electron tomography (ET) experiments with the grids prepared for cryo-EM. The results confirmed the cryo-EM experiments and showed damaged FAS particles in all imaged areas (figure 19). Furthermore, the cryo-ET experiments

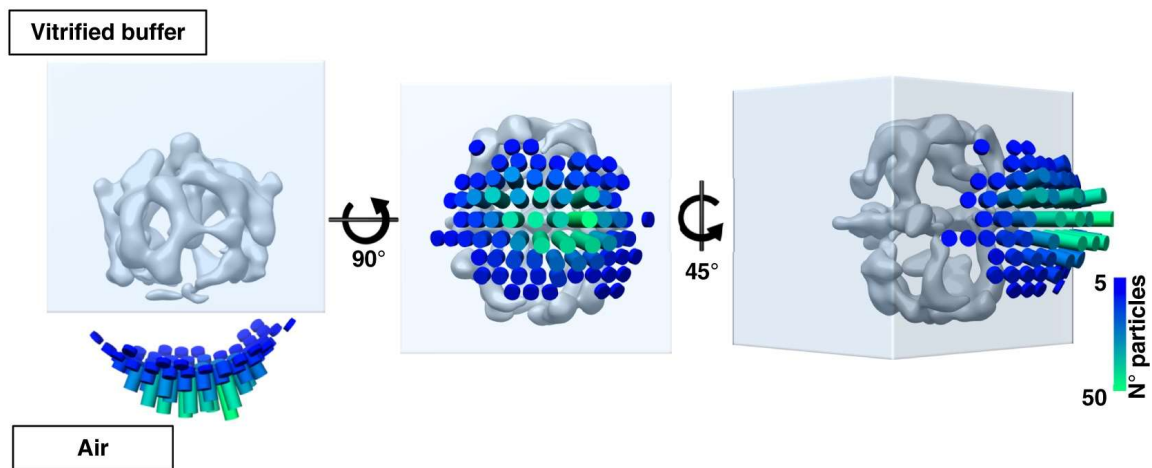
revealed that most FAS particles adsorb to the air-water interfaces on both sides. While one side was preferred and densely occupied, the other side contained just a small number of FAS particles. Unaffected from the side, nearly all adsorbed particles were damaged and the most appeared flattened. A reconstruction of the FAS particles at the interface revealed that one side of the complex was complete, while the other side lacked electron densities. We suggested that during grid preparation, FAS particles adsorb with its long axis to the air-water interfaces and were destroyed or unfolded as a result from the air-exposure.



**Figure 19** | Cryo-ET experiments of FAS particles in unsupported vitrified buffer. (A) Typical tomogram of a grid hole occupied with FAS particles. Red arrows highlighting damaged FAS particles (B) The side-view section of a typical tomogram of a grid hole shows that the FAS particles adsorb to the air-water interface. (C) Different orientations of a single FAS particle within the tomogram shows

that one side is damaged. Cryo-ET experiments were performed and analysed by Davide Floris (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D'Imprima and Floris *et al.*<sup>117</sup>

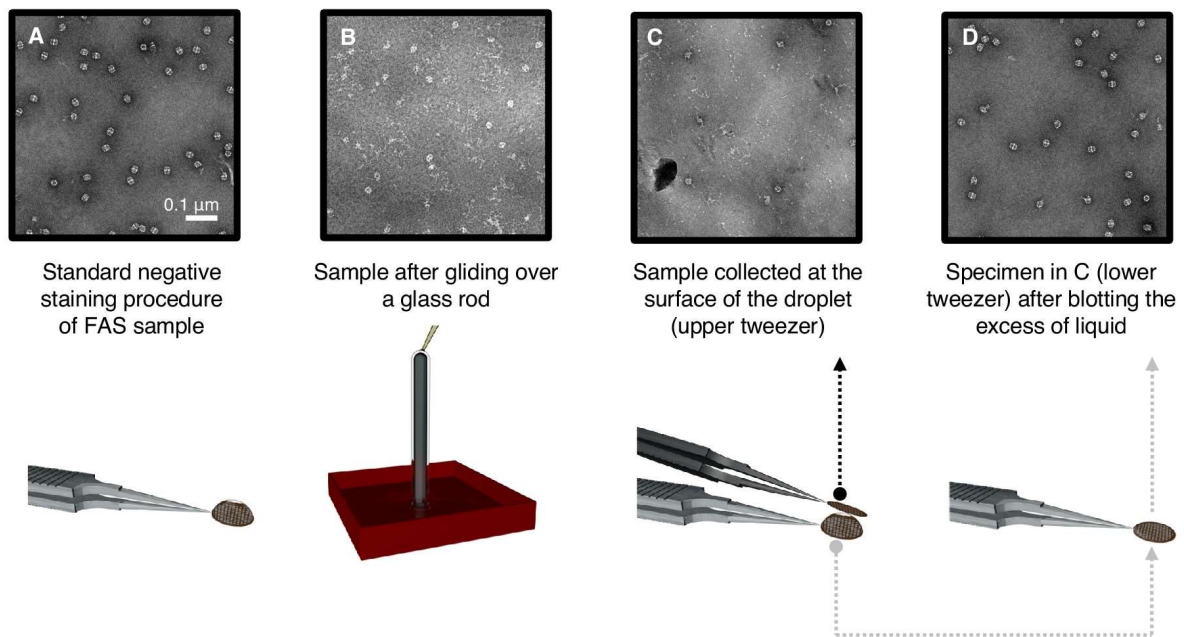
The orientation of the broken particles with respect to the interface was further analysed. The interface was determined, by fitting a surface through all particles. Then the orientation of the particles was calculated by comparing the normal vector of this interface with a vector from the centre of the particles toward its flattened side. The analysis revealed that the damaged side was oriented towards the air-water interface and demonstrated that the adsorption is the reason for protein damages (figure 20).



**Figure 20** | Orientation of the damaged FAS particles with respect to the air-water interface. The angular distribution of the particles revealed that the damaged side was oriented towards the air-water interface. The analysis was performed by Ricardo Sánchez (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D'Imprima and Floris *et al.*<sup>117</sup>

We performed further experiments to investigate the impact of the exposure of yeast FAS to the air-water interface. In a first experiment, we bubbled air through the FAS solution. This treatment causes large air-water interfaces particularly by bubbles at the surface. Negative-stain micrographs showed that the particles were destroyed completely (not shown). As positive controls, we always investigated the untreated FAS solution with our standard procedure, which confirmed the structural integrity of the solution (figure 21 A). In a second experiment, we poured the protein solution over a glass rod. The micrograph showed that only few particles stayed intact (figure 21 B). We received similar results for FAS solution incubated on a carbon-coated grid, which was subsequently collected from the top of the droplet with a second grid (figure 21 C). Using that procedure, we picked up the particles, which adsorbed to the air-water interface. In contrast, FAS particles adsorbed to the first carbon coated grid remained intact (figure 21 D). The experiments clearly confirmed our hypothesis that exposure to air is harmful to the protein. The air-water interface

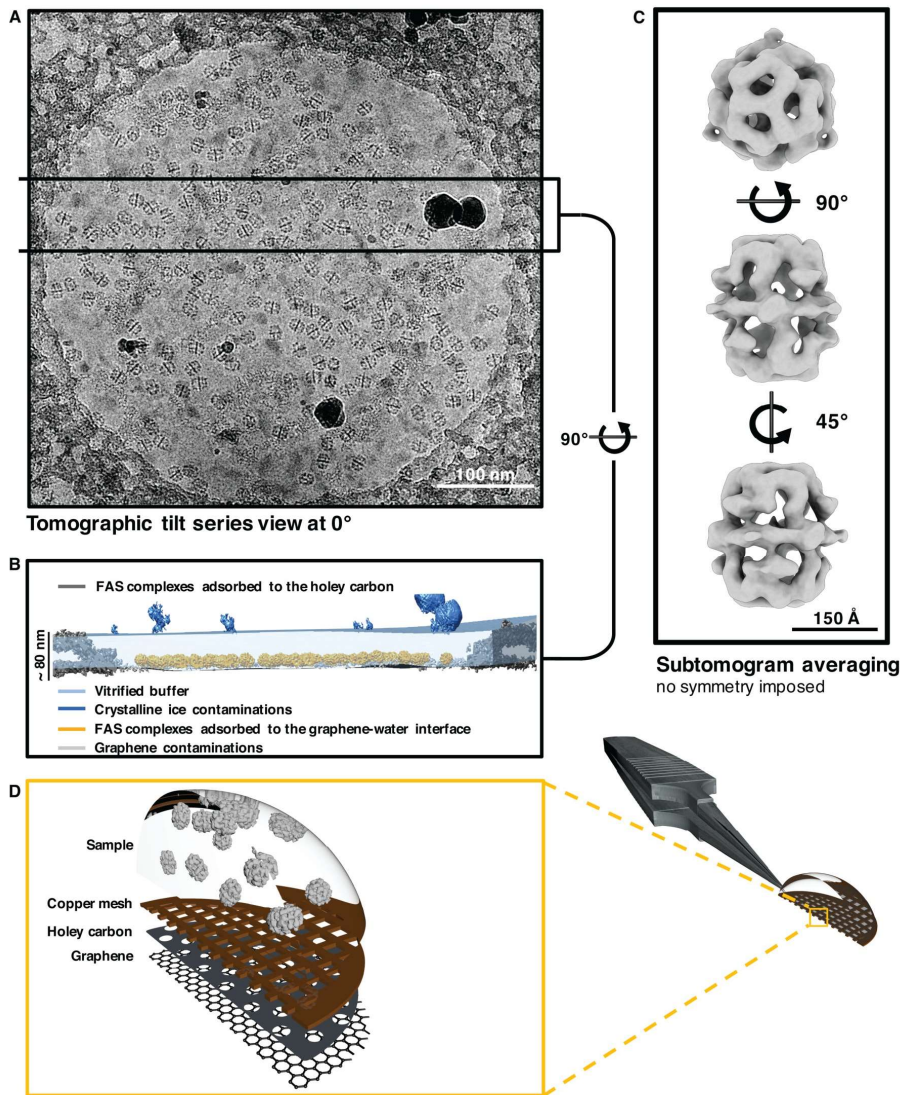
was suggested to cause preferred orientation, unfolding and denaturation before.<sup>118</sup> Furthermore, a recent study of 31 different protein demonstrated that most of them also adsorb to the air-water interface.<sup>119</sup>



**Figure 21** | Structural integrity of yeast FAS, exposed to air, was investigated by EM using the negative-staining method. (A) The untreated sample was analysed with the standard procedure as a positive control. (B) The FAS solution was glided over a glass rod, which led to heterogeneity in the micrographs. (C) The FAS solution was incubated on a grid, and then the top of the solution was picked up with a second grid. While the FAS particles were undamaged on the first grid (D), major heterogeneities were visible on the sample collected with the second grid (C). Negative staining EM experiments were performed and analysed by Edoardo D’Imprima (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D’Imprima and Floris *et al.*<sup>117</sup>

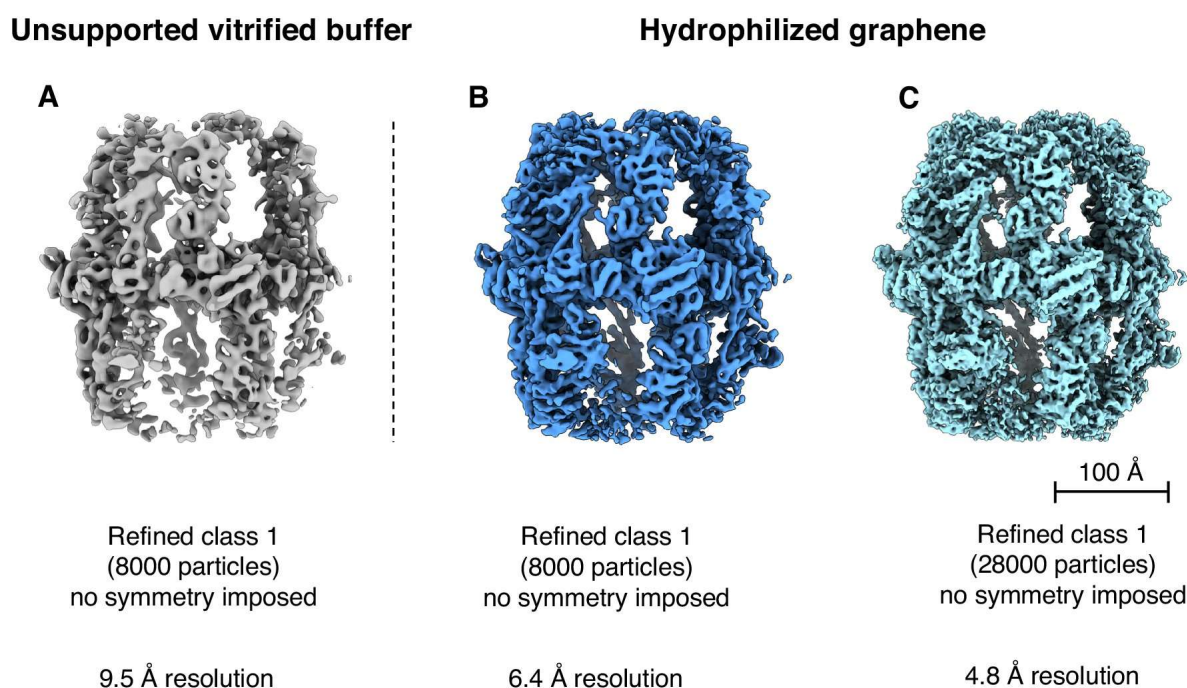
In summary, several experiments showed that yeast FAS particles get destroyed during cryo-EM sample preparation and, consequently, we searched for a better method to prepare the grids. Since the proteins can diffuse to the air-water interface in milliseconds, the problem cannot be solved with faster grid preparation.<sup>120</sup> However, a physical support can potentially prevent that the protein interacts with the air-water interface, which should then result in a better sample quality. We decided to use graphene, as it is a very thin and strong material, an excellent conductor, stable under a 300 kV electron beam and electron transparent to 2.13 Å.<sup>121–123</sup> On the other hand it is very hydrophobic and prevents that the protein can be evenly distributed on the grid. Therefore, we decided to improve the hydrophilicity by plasma etching and by non-covalent chemical doping with 1-pyrenecarboxylic acid.<sup>124,125</sup> In doing so, the number of particles per area can be adjusted by the 1-pyrenecarboxylic acid concentration. A freshly prepared FAS solution was applied on a hydrophilized graphene grid (Quantifoil R0.6/1 holey carbon grids) and analysed with cryo-ET. The results clearly show that the FAS particles were arranged on the graphene-water

interface and not on the air-water interface (figure 22). We performed sub-tomogram averaging and multi-reference classification on 2090 hand-picked subvolumes, which showed that all particles were structurally intact. In summary, the hydrophilized graphene grids prevents denaturation as FAS particles prefer the graphene-water interface over the air-water interface.



**Figure 22** | Cryo-ET experiments of FAS particles with physical support, using hydrophilized graphene grids. (A) Typical tomogram of a grid hole occupied with FAS particles. (B) The side-view section of a typical tomogram shows that the FAS particles prefer the graphene-water interface over the air-water interface. (C) Different orientations of subtomogram averaging demonstrating the structural integrity of the FAS. (D) Composition of the graphene grids. The FAS solution was applied from the uncoated side. Cryo-ET experiments were performed and analysed by Davide Floris (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D'Imprima and Floris *et al.*<sup>117</sup>

Next, we investigated whether the grids are suitable for high-resolution cryo-EM. The full data set of about 28,000 particles led to a resolution of 4.8 Å and 4.0 Å with no symmetry (C1) or D3 symmetry applied (figure 23). To compare the physical support to unsupported vitrified buffer, we randomly selected 8,000 particles from the dataset, received from the graphene grids, and obtained a map with a resolution of 6.4 Å. In contrast, the undamaged 10% of the traditional grids led to a much lower resolution (9.5 Å), demonstrating that the hydrophilized graphene grids are more suitable for the yeast FAS. Furthermore, all particles of the dataset using hydrophilized graphene could be used for 3D reconstruction, resulting in shorter data collections. Since the adsorption was also shown for other proteins, the presented protocol might be also beneficial for other macromolecular complexes.



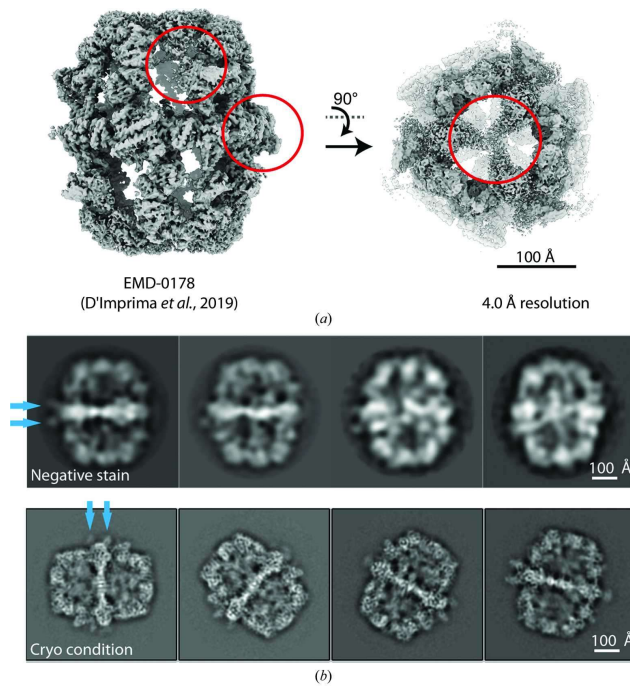
**Figure 23** | Comparison of the FAS maps, received from cryo-EM with traditional grids in unsupported vitrified buffer and hydrophilized graphene grids using a physical support. (A) All 8,000 undamaged particles (about 10% of the whole dataset) from the dataset using unsupported vitrified buffer were reconstructed to a 9.5 Å with no symmetry applied. (B) In comparison the same number of particles using hydrophilized graphene yielded a 6.4 Å map and about 28,000 particles a 4.8 Å map (C), demonstrating that the physical support is beneficial for the sample. Cryo-EM experiments were performed and analysed by Edoardo D’Imprima (Max Planck Institute for Biophysics, Frankfurt). The figure was published by D’Imprima and Floris *et al.*<sup>117</sup>

With the established protocol, we were able to improve the resolution and the integrity of the map. Unfortunately, the map still revealed parts of the  $\beta$ -domes that are poorly resolved and lacks the PPT domain at the outside of the barrel (figure 24 a, red circles). Most likely this is caused by remaining structural heterogeneity, because of local denaturation or unfolding during. Most relevant is the absence of electron density for the PPT domain, which might be particularly prone to



unfolding, as it forms a trimer as a stand-alone protein<sup>54</sup> but is held as monomer at the perimeter of the mature FAS complex.<sup>56</sup> Also other yeast FAS structures might suffer from PPT unfolding, because they are lacking electron densities for the PPT domain at the outside of the barrel.<sup>36,56,58,61</sup> The PPT might be damaged during protein purification, crystallization, grid preparation or concentration using a semipermeable membrane. Therefore, the structural integrity of the PPT can be used as a quality measure in EM experiments.

To address the challenge of keeping PPT intact during sample preparations, we adapted our protocol slightly. First, we included an incubation step of yeast FAS with 1 mM Mal-CoA and 1 mM NADPH for 5 minutes at room temperature. This treatment allows the FAS to complete (but not to start) elongation of fatty acids, which might then reduce sample heterogeneity. Second, we avoided the final protein concentration step through a semipermeable membrane, which was possible by the need for lower protein concentrations with the improved protocol for the grid preparation. The quality of the data was much better, judged by the presence of the PPT domain in class averages, received from negative-staining and cryo-EM experiments (figure 24 b and figure 25 in comparison to figure 17-18). We note that data were obtained from just a few sample preparations, and more preparations would be necessary for a deeper analysis. Further, the experimental set-up could not clarify whether improved protein quality results from the incubation of substrates with Mal-CoA or from omitted protein concentration through a semipermeable membrane. The biochemical analysis of the quality of PPT is difficult, because the PPT domain is only required for phosphopantetheinylation during the assembly, but not involved further in fatty acid synthesis, so that the quality of PPT cannot be judged by the fatty acid synthetic activity of the yeast FAS. Finally, the domain does not contribute to the thermal or oligomeric stability, because of its position at the outside of the barrel, so that it also escapes TSA measurements.



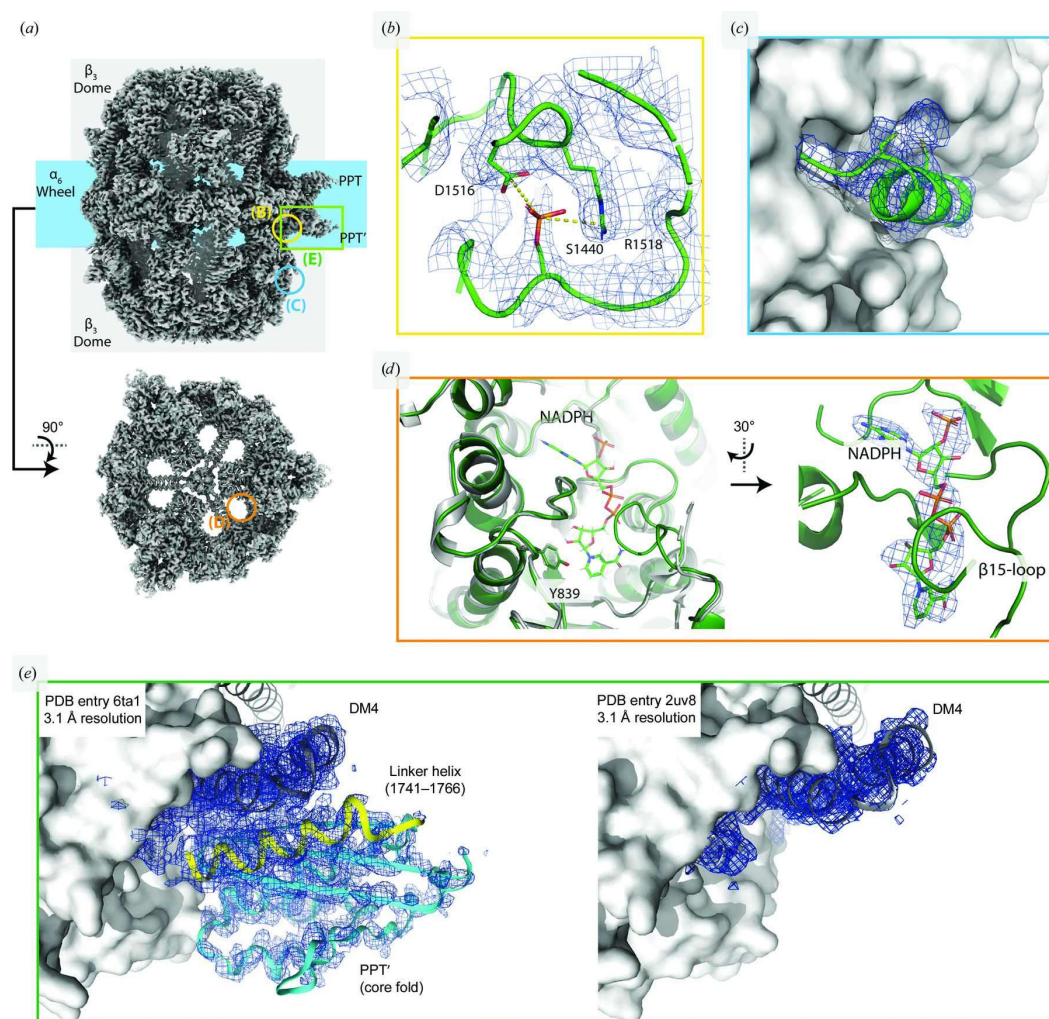
**Figure 24** | Investigation of different FAS preparations. a) The 3D-reconstructions of the untreated sample, which was concentrated through a semipermeable membrane features parts of the  $\beta$ -domes that are poorly resolved and lacks the PPT domain at the outside of the barrel b) 2D-Class averages of the improved yeast FAS preparation. The final concentrations step was avoided, and the sample was treated with 1 mM Mal-CoA and 1 mM NADPH for 5 minutes at room temperature prior to grid-preparation. Class average of negative-stain and cryo-EM showed the PPT at the outside of the barrel (indicated with blue arrows) and overall better structured  $\beta$ -domes. Cryo-EM and negative staining experiments were performed and analysed by Edoardo D'Imprima (Max Planck Institute for Biophysics, Frankfurt). The figure was published by Joppe and D'Imprima *et al.*<sup>109</sup>

### 2.1.3 Structural determination of the yeast FAS

With the established rapid and non-invasive purification strategy and the protocol to prevent denaturation at the air-water interface, we were able to elucidate the structure of the yeast FAS with cryo-EM at a high resolution of 3.1 Å. We used almost all particles (77%), demonstrating the excellent quality of the sample and furthermore reduced instrument time immensely. Similar protocols can be used for other macromolecular complexes.

As described in previous structures, received by X-ray crystallography and cryo-EM, yeast FAS assembles to a 2.6 MDa D3-symmetric barrel-like complex, comprised by an  $\alpha$ -wheel built from six  $\alpha$ -polypeptide chains that is flanked by two  $\beta$ -domes formed by three  $\beta$ -polypeptide chains each.<sup>36,53,56,57</sup> The rigid cage of the heterododecameric FAS barrel is stabilized with DM1-4 at the central wheel and a TM, closing the  $\beta$ -domes at the apical sides. Compared to older structures, the received FAS model (figure 25) revealed additional density at Ser1440 of the  $\alpha$ -chain, indicating phosphorylation, similarly as observed in a large-scale phosphorylation screen

before.<sup>126</sup> The phosphate group might interact with Asp1516 and Arg1518. Additionally, we found density at Cys820 and Cys824 at the  $\beta$ -chain, which are probably malonylated due to the high concentration of Mal-CoA used for sample preparation. Another feature of the map is the binding of NADPH to the active site of the KR domain, which is exposed to the inner surface in the near of the  $\beta$ 15 loop. In contrast, NADPH was not found in the active site of the ER domain. Furthermore, the excellent sample preparation led to a full map, including the PPT domain with its linker domain, whereas the electron density of those domains was missing in older structures.<sup>36</sup>



**Figure 25** | Details of the 3.1 Å map of the yeast FAS (PDB: 6TA1). (a) Overview of the whole map in top- and side-view. Sections of the map, which are enlarged in panel (b)-(e) are highlighted with coloured circles and squares. (b) Additional density was found at position Ser1440, which indicates phosphorylation at this position. The phosphate might interact with Asp1516 and Arg1518. (c) Additional density was also found at Cys820 and Cys824, which might be malonylated from Mal-CoA. (d) NADPH binds to the active site of the KR. (e) Comparison of DM4 (grey), PPT (cyan) and the linker

helix (yellow) in the cryo-EM map with a map received from X-ray crystallography at a similar resolution (PDB: 2UV8).<sup>36</sup> In contrast to the fully resolved cryo-EM map, there are no densities for the linker helix and the PPT in the map, received by X-ray. Cryo-EM experiments were performed and analysed by Edoardo D'Imprima (Max Planck Institute for Biophysics, Frankfurt). The figure was published by Joppe and D'Imprima *et al.*<sup>109</sup>

### 2.1.4 Mutational study of yeast FAS phosphorylation sites

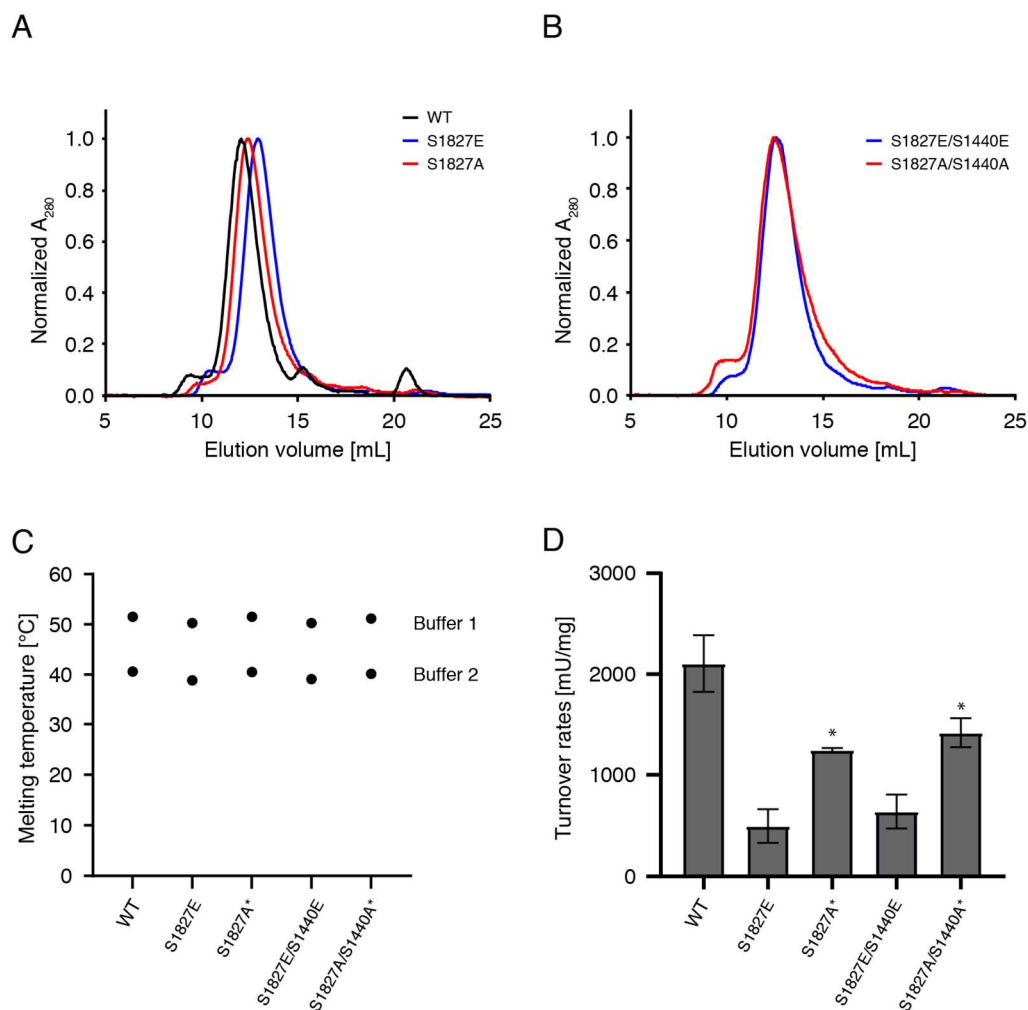
Post-translational phosphorylation is estimated to affect about 30% of all proteins, which is essential for the regulation of cellular processes.<sup>126,127</sup> This modification is reversible and takes place at specific serine, threonine or tyrosine residues.<sup>126</sup> Conformational changes of proteins can be initiated by phosphorylation. The phosphate group adds two negative charges to the protein, which can change structures locally and globally, putatively modulating the substrate binding and catalytic activity. Further, the phosphate group can undergo hydrogen bond formation, leading to new specific interactions.<sup>21</sup>

With our rapid protein preparation protocol in hand, we sought to evaluate the phosphorylation of Ser1440 in its enzyme kinetic impact. In addition, we decided to also evaluate Ser1827 as a phosphorylation site, which was identified recently by Martinez-Montanes *et al.*<sup>128</sup> Both sites were analysed by mutation to glutamic acid and alanine. The cloning, expression, purification, and analysis of the mutants was performed by the student Martin Schwalm, supervised by me.

Ser1440 is located at the DM4, positioned at the outside of the KS domain.<sup>109,126</sup> DM4 stabilizes the PPT at the perimeter of the FAS barrel. As seen in structural data, the phosphate group at Ser1440 interacts with Asp1516 and Arg1518 and this motif was found at high conservation in other fungal FASs (figure S1 A, amino acids are highlighted with light blue boxes).<sup>49,109</sup> Ser1827, also highly conserved, is located at the outside of the PPT domain at the C-terminal end of  $\alpha$ 55, pointing away from the barrel.<sup>54</sup> It is positioned relative close to the CoA binding site (6-8 Å) and has larger distance to the predicted ACP binding site, to active site residues (>12 Å) and the dimeric interface (>12 Å).<sup>54,129</sup> The nearby Lys1826 and the backbone amid group of Gly1829 possibly stabilize the phosphate (figure S1 B, amino acids highlighted with light blue boxes).

We performed site-directed mutagenesis to change both phosphorylated serines either to an alanine or to a glutamic acid with the latter one mimicking the negative charge of a phosphate. We prepared overall four mutated constructs, S1827E, S1827A, S1440E/S1827E and S1440A/S1827A, in which Ser1440 was only mutated in addition to the Ser1827 mutation. Expression and purification were performed analogously to chapter 2.1.1 with no fatty acids supplemented to the medium. The yeast cultures grew unaffectedly, implying that the engineered FASs were still active.

After purification with affinity- and size-exclusion chromatography (figure 26 A-B), we yielded 0.5-0.8 mg of all mutants, comparable with wild type (WT) protein ( $1.4 \pm 0.4$  mg). A TSA showed only minor differences in the melting temperatures, indicating that all mutants are stable (figure 26 C). Mutations at the position Ser1827 decreased the specific activities of FAS to 24 % for S1827E and 59% for the S1827A (figure 26 D). These data suggest that phosphorylation of Ser1827 down-regulates FAS activity. A molecular mechanism for this effect remains elusive. Phosphorylation may occur prior or during assembly and influence the degree of phosphopantetheinylation of the ACP<sup>130</sup>. The phosphate might induce structural changes and therefore impact the activity of the PPT, as for example the formation of the active PPT dimer during assembly<sup>130</sup> and the substrate binding of CoA and ACP. Mutations at position Ser1440 slightly improved the fatty acid synthesis to 129 % for the glutamic acid mutation, compared to S1827E, and 113% for the alanine mutation, compared to S1827A (figure 26 D). Both increases are within the error range of the measurements and should be addressed in further experiments for clarification. Therefore, the phosphorylation on this position could not be evaluated.



**Figure 26** | Evaluation of two phosphorylation sites of the yeast FAS. (A) SEC on a S6 Increase 10/300 GL column of yeast FAS WT and mutants at position Ser1827 (in 100 mM sodium phosphate, pH 6.5). (B) SEC on a S6 Increase 10/300 GL column of yeast FAS mutants at position Ser1440 (in 100 mM sodium phosphate, pH 6.5). (C) Average melting temperatures in two different buffers, analysed by TSA (buffer 1: 100 mM NaPi, pH 6.5; buffer 2: 100 mM citric acid, 250 mM sodium chloride, pH 5.0). Glutamic acid mutations were analysed with two biological samples and alanine mutations with just one, indicated with an asterisk. Each biological sample was analysed in three technical replicates. Melting points varied by less than 1 °C. (D) Average turnover rates of each FAS construct, analysed by NADPH consumption. Glutamic acid mutations were analysed with two biological samples and alanine mutations with just one, indicated with an asterisk. Each biological sample was analysed in three technical replicates and the standard deviations of technical replicates is provided as error bars. For comparison WT activity of eight biological samples is shown with error bars. Data were collected by Martin Schwalm.

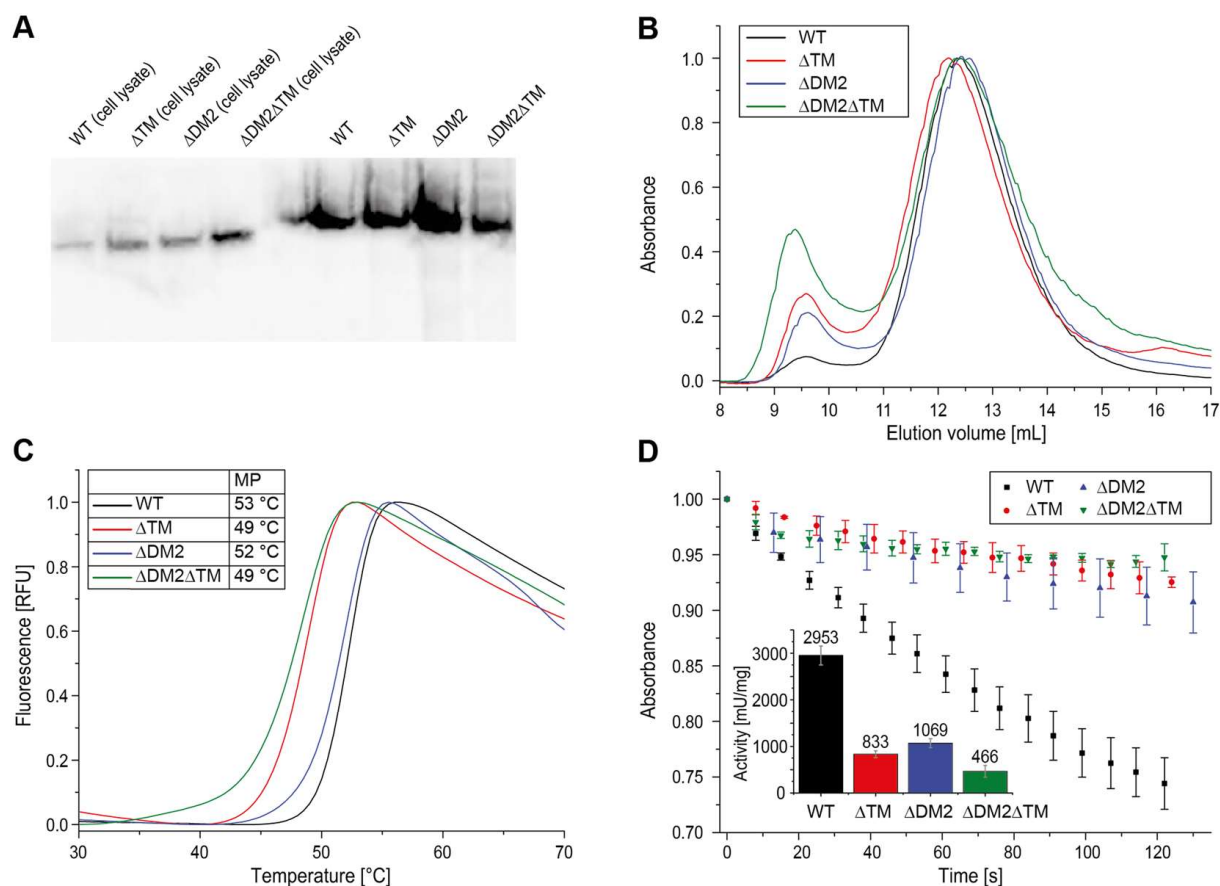
### 2.1.5 Investigating scaffolding domains of the yeast FAS

The yeast FAS is a highly efficient fatty acyl-CoA producing protein. However, due to its rigid structure and enclosed reaction chambers, the scaffold can hardly be remodelled, so that enlarging the product spectrum towards non-fatty acid

compounds is challenging. In seeking to overcome these limitations, we investigated the impact of scaffolding domains (also called insertion elements) for assembly, stability, and catalytic efficiency. In case the domains are not essential for these processes, they might be suitability for exchange with other domains, i.e., they might be used as fusion sites for catalytic domains. We focused on the insertion elements TM and DM2. The TM closes the apical sides of the  $\beta$ -domes, whereas the DM2 is positioned at the outside of the barrel between the dimeric interface of the KR. The project was performed during the bachelor thesis of Alexander Rill, supervised by Manuel Fischer and me.<sup>131</sup>

We designed three different constructs by deleting the TM, the DM2 and both. We purified all constructs from *S. cerevisiae* with the same strategy, described in chapter 2.1.1. We verified the successful assembly of the different constructs by native PAGE Western Blot (figure 27 A). This was further confirmed by SEC during the purification of proteins, showing that all constructs elute at approximately the same volume of about 12.5 mL (figure 27 B). In all constructs there is a minor peak at 9.5 mL, which likely originates from aggregates or higher oligomeric states. This peak is increased for the mutants compared to WT protein, especially for the TM deletions, indicating compromised protein stability. These results correlate very well to the TSA, which shows slightly lower melting temperatures for the mutants, compared to the WT protein (figure 27 C). The lower stability of the mutants is also reflected by the decreased fatty acid synthesis, analysed by NADPH consumption assays (figure 27 D). The difference in the stability of the TM and DM2 deletion can be explained by the position of the domains. The TM is responsible for the trimerization of the FAS and as such very relevant for the protein stability. In addition, its vicinity to the AT and the ER might influence the catalytic efficiency of those domains. In contrast, the DM2 simply expands the dimeric interface of the KR and therefore is less relevant for structural integrity. In conclusion, our experiments demonstrate that the deletion of these domains does not affect the assembly or abolishes the catalytic activity. These domains are likely involved in the stabilization of the FAS barrel after its assembly. This is in line with the evolutionary development of fungal and bacterial FASs. Not all representatives of this protein family possess a TM and DM2 domain, but all assemble to a barrel-shaped structure.<sup>49,52,130</sup>

In summary, we showed that the TM and DM2 are not essential for the yeast FAS, which indicates that these domains are suitable for domain exchange approaches. The lowered catalytic efficiency likely originates from decreased protein stability and may be restored by well-designed domain exchanges.



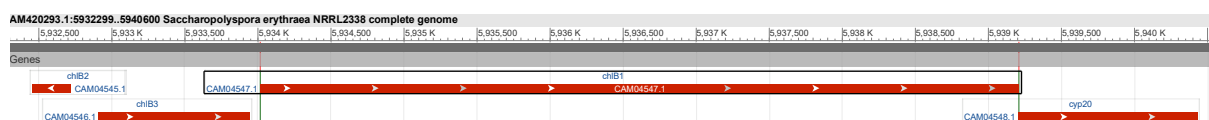
**Figure 27** | Investigating the impact of scaffolding elements on the yeast FAS assembly, stability, and catalytic efficiency. The insertion elements TM and DM2 were deleted by sequence and ligation independent cloning and then expressed in *S. cerevisiae* strain (FAS deletion). (A) Cell lysates (left) and purified proteins (right) of FAS WT and mutants were analysed with native PAGE Western Blot using an anti-FAS antiserum.<sup>132</sup> (B) FAS WT and mutants were purified with SEC on a Superose 6 Increase 10/300 GL (in the buffer 100 mM sodium phosphate, 200 mM sodium chloride, pH 6.5). The absorbance was normalized with respect to the highest peak. (C) FAS WT and mutants were analysed for thermostability using TSA in the purification buffer (100 mM sodium phosphate, 200 mM sodium chloride, pH 6.5). Illustrated are typical melting curves and average melting temperatures (MP), with the latter one varying less than 0.5°C in technical replications. (D) The enzymatic activity of FAS WT and its mutants was determined by NADPH consumption at 334 nm over 120 seconds. Illustrated is the NADPH absorbance over time after background correction (without Mal-CoA) with error bars and the calculated activities in mU/mg as bar diagram with error bars from three technical replicate per construct. The figure was published by Fischer *et al.*<sup>130</sup> Data were collected by Alexander Rill.<sup>131</sup>

## 2.2 Characterization of a non-releasing MSAS from *S. erythraea*

At the outset of this study, we aimed for accessing sufficient amounts of the iterative PKS MSAS as a target for structural and functional investigations and for producing 6-MSA derivatives by protein engineering. We identified SerMSAS as a promising candidate for characterization because we were able to express enough enzyme in



*E. coli*. The following data demonstrate that SerMSAS was unable to release its own product and therefore not suitable for the aspired engineering approach. A deeper look into the *S. erythraea* genome identified two genes directly upstream and one gene directly downstream to the SerMSAS gene (figure 28). The upstream genes encode for a KASIII-like protein (CAM04546.1), annotated as SerChIB3, and a discrete ACP or peptide carrier protein (CAM04545.1), annotated as SerChIB2. The downstream gene encodes for a putative cytochrome P450 oxidase (CAM04548.1). The gene cluster has been previously described, but the structure of the product remained elusive.<sup>133</sup> Our data suggest that the product of SerMSAS (gene product of SerChIB1) is transferred by SerChIB3 to the discrete carrier protein SerChIB2, instead of hydrolysis. The putative Cytochrome P450 oxidase is probably responsible for further modifications, like described for *Aspergillus clavatus*.<sup>86</sup> It can be assumed that 6-MSA, produced by SerMSAS, is processed to a yet uncharacterized natural product.

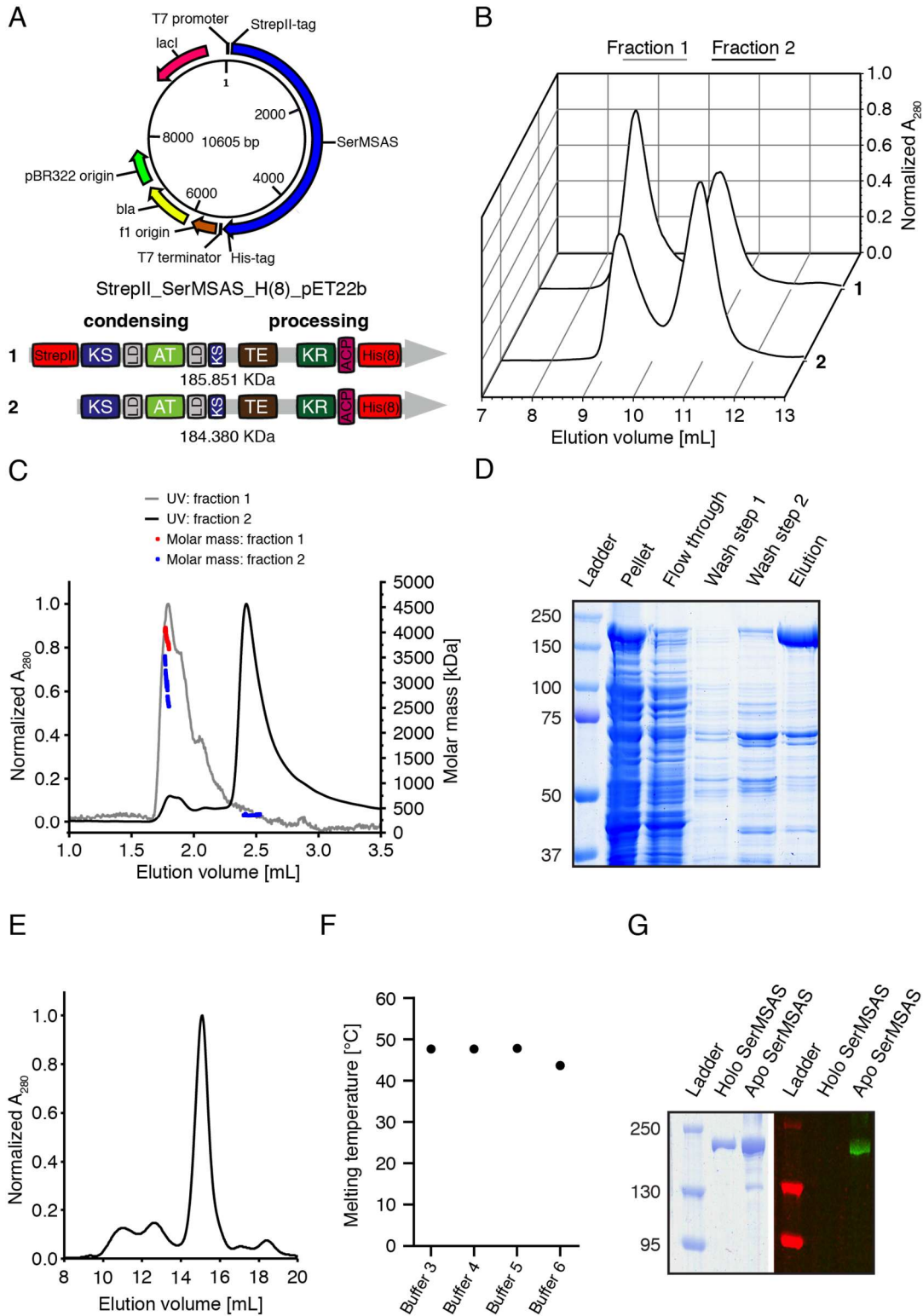


**Figure 28** | Visualisation of the *S. erythraea* genome in range of the targeted gene (GenBank: AM420293.1). The gene of interest is highlighted by a black box (SerMSAS annotated as ChIB1, CAM04547.1). Upstream open reading frames encode the KASIII-like ChIB3 (CAM04546.1) and ChIB2 (CAM04545.1), and the downstream gene encodes the P450 Cytochrome C oxidase (CAM04548.1). The section of the genome was taken from <https://www.ncbi.nlm.nih.gov/>.

### 2.2.1 Expression and purification of SerMSAS in *E. coli*

To express and purify SerMSAS for structural and functional characterization, the respective gene was cloned from genomic DNA into a bacterial pET22b expression plasmid, containing a N-terminal StrepII-tag and a C-terminal His-tag for purification (figure 29). The cloning was performed previously in our group. For a related MSAS from *A. terreus*, it was reported that the N-terminal-tagged protein was inactive.<sup>82</sup> Taking this into account, we removed the N-terminal StrepII-tag by site-directed mutagenesis to yield an additional construct (figure 29 A). Each vector was used for co-expression with a *Sfp* encoding vector in *E. coli*, kindly received from Alexander Rittner (Goethe-University, Frankfurt). After cell lysis with French press and centrifugation of the cells, supernatant was purified by affinity chromatography with either a Ni-NTA (immobilized metal affinity chromatography (IMAC), indicating nickel columns in this thesis) or a Strep-tactin column followed by SEC. Both strategies provided MSAS in high purity and the SEC profile showed two peaks in different ratio, depending on the construct (figure 29 B). For determination of the oligomeric state,

both peaks of construct one were pooled and analysed with SEC-MALS by high performance liquid chromatography (HPLC) (figure 29 C). We calculated the mass ( $M_r$ ) from both peaks using the ASTRA software (version 7.1.0.29, theoretical mass of the monomer: 185,851 Da). The first peak had a mass of 4.1-2.5 MDa, indicating heterogenic distribution of higher oligomeric states. The second peak had an average mass of 362 KDa, which is 1.95 times the mass of the monomer, suggesting that SerMSAS exists as a dimer. In contrast MSAS from *P. patulum* was described as a tetrameric protein, judged by gel filtration.<sup>134</sup> We concluded that the N-terminal tag has indeed a negative influence on proteins quality, because of a higher tendency to form aggregates. We proceeded with construct 2 and optimized its purification. Based on *Lowry et al.*, we changed the buffer to high amount of sodium chloride and used a different column for SEC, which appeared more suitable for higher molecular weight proteins.<sup>135</sup> Using this strategy, we had access to high amounts (5-8 mg/L) of pure protein with a small tendency to form aggregates, judged by SDS-PAGE and SEC (figure 29 D-E). Furthermore, a TSA confirmed the stability of SerMSAS in all buffers used for purification, validating the established purification strategy (figure 29 F). The assay was performed by Christian Gusenda during his bachelor thesis.<sup>136</sup> Since the phosphopantetheinylation of the ACP is crucial for the enzymatic activity, it was investigated *in vitro* in the same bachelor thesis.<sup>136</sup> We expressed and purified SerMSAS with or without *Sfp*, resulting in the holo- or apo-form. Both proteins were incubated with *Sfp* and fluorescence marked CoA (CoA-488), which can only be transferred to the apo- und not to the holo-form. The absence of a fluorescent band for the holo protein and the presence of a fluorescence marked band for the apo protein demonstrated complete activation of the ACP. Next, we focused on the enzymatic activity of SerMSAS.

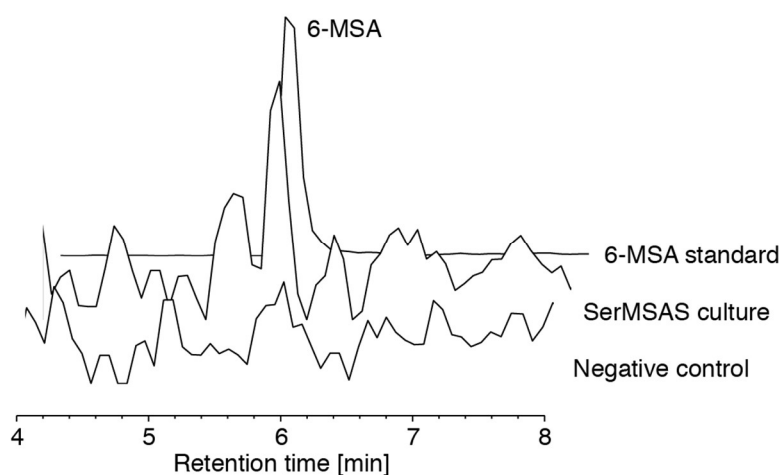


**Figure 29** | Establishment of a purification strategy for SerMSAS. (A) Representative vector map for a pET22b derived expression system with a N-terminal StreptII-tag and a C-terminal His(8)-tag. The domain structure of the two SerMSAS constructs used for purification is mapped below. (B) SEC of the two SerMSAS constructs on a S200 Increase 10/300 GL column (in 100 mM sodium phosphate,

pH 7.6, 100 mM sodium chloride, 20% glycerol and 1 mM EDTA). Different elution times for the two constructs most likely originate from different settings. (C) HPLC of the two fractions of construct 1 from figure b on a Yarra-SEC-X300 (Phenomenex) coupled to a MiniDAWN TREOS (Wyatt technologies) for MALS experiments (in 100 mM sodium phosphate, pH 7.6, 100 mM sodium chloride, 20% glycerol and 1 mM EDTA). Shown is the normalized absorbance (left axis) and calculated mass from MALS experiments (right axis) as a function of the elution volume. The model Zimm was used for LS analysis with the ASTRA software (version 7.1.0.29). (D) A representative Coomassie-stained SDS-PAGE with samples taken during the purification of SerMSAS by IMAC. SerMSAS construct 2 was co-expressed with *Sfp*. (E) SEC on a S6 Increase 10/300 GL column of SerMSAS construct 2 (in 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol and 1 mM EDTA). (F) Melting temperatures from a TSA with SerMSAS construct 2. All buffers used for purification were analysed (buffer 3: 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol and 1 mM EDTA; buffer 4: 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol and 20 mM imidazole; buffer 5: 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol and 60 mM imidazole; buffer 6: 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol and 300 mM imidazole). (G) Analysis of the phosphopantetheinylation in co-expression with *Sfp*, using in-gel fluorescence (right panel) and Coomassie-stained SDS-PAGE (left panel). SerMSAS, expressed without *Sfp* and received as apo-protein was used as a positive control. Both proteins were incubated with 320 mM MgCl<sub>2</sub>, 1.13 μM purified *Sfp*, 1 mM DTT and 10 μM CoA-488 for one hour at 22 °C prior to treatment with Laemmli buffer. SDS-PAGE, monitored with fluorescence detection and Coomassie staining is shown.

### 2.2.2 *In vivo* 6-MSA synthesis

The enzymatic activity of SerMSAS *in vivo* was investigated by Trinetri Goel during her bachelor thesis.<sup>137</sup> After co-expression of SerMSAS with *Sfp* and centrifugation of the cells, the medium was acidified, extracted with ethyl-acetate and analysed for 6-MSA by HPLC-MS. Like expected, 6-MSA was found in the supernatant (figure 30). 6-MSA was supplied to our standard medium as a positive control (6-MSA standard) and co-expression of SerChIB2 was used as a negative control. After validation of 6-MSA synthesis *in vivo*, we focused on the activity *in vitro*.

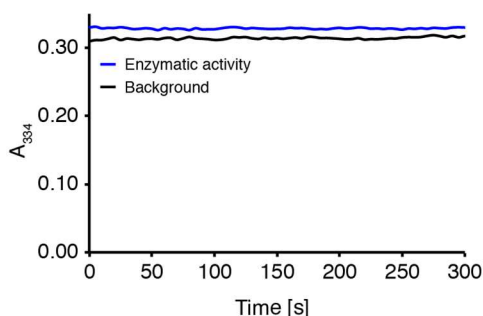


**Figure 30** | Enzymatic activity of SerMSAS *in vivo*. The medium of a SerMSAS culture was analysed for 6-MSA by HPLC-MS [extracted ion chromatogram (EIC for **6-MSA**): [M-H]<sup>-</sup> m/z = 151.04

$\pm 0.10$ ]. We added 6-MSA to standard medium as a positive control and used a SerChIB2 culture as a negative control. Data were normalized with respect to the 6-MSA peak. The negative control was normalized with the same factor as the SerMSAS sample.

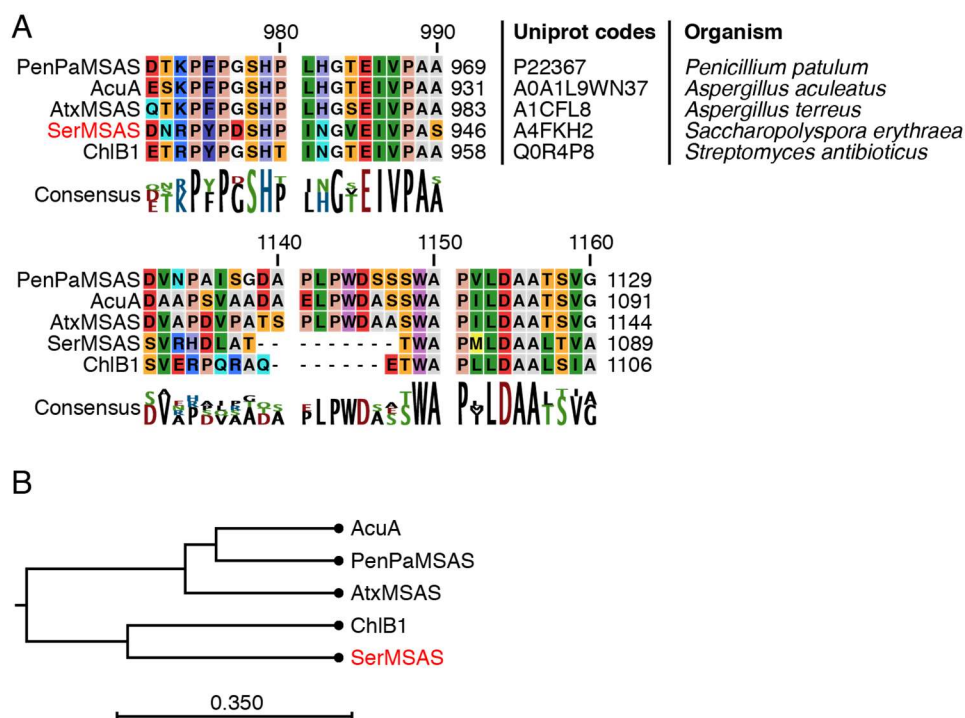
### 2.2.3 *In vitro* 6-MSA synthesis

We analysed the *in vitro* enzymatic activity photometrically, by incubating SerMSAS with the substrates acetyl-CoA, Mal-CoA and NADPH. The decrease of NADPH was monitored by its absorption. 6-MSA synthesis could not be observed, based on the absence of NADPH consumption (figure 31).



**Figure 31** | Analysis of SerMSAS's activity *in vitro*. SerMSAS ( $1 \mu\text{M}$ ) was incubated with  $200 \mu\text{M}$  acetyl-CoA,  $200 \mu\text{M}$  Mal-CoA and  $50 \mu\text{M}$  NADPH for five minutes (in  $50 \text{ mM}$  potassium phosphate, pH 7.0, 10% glycerol, 1 mM DTT and 0.4 mg/mL BSA). Background was taken in the absence of SerMSAS. Absorption of NADPH at 334 nm was measured every five seconds with the Nanodrop 2000C (Thermo Scientific). Data points were connected for visualization.

We performed a sequence alignment to other MSASs due to the observation that SerMSAS was not active *in vitro*. Moriguchi *et al.* shows for the related MSAS from *A. terreus* that the hydrolytic cleavage of 6-MSA requires a histidine (His972) at position 972 and suggests that an aspartate (Asp1129) at position 1129 is also involved.<sup>82</sup> In the alignment (figure 32 A), we noticed that the histidine at position 972 (alignment number 979) in *A. terreus* lies within a conserved motif (PXPXSHXXXGXEIVP) in all analysed synthases. In contrast the aspartate (Asp1129) was only conserved (LPWDXXSWAP) for MSAS from *P. patulum*, *A. terreus* and *Aspergillus aculeatus* (*A. aculeatus*), all reported or suggested to release 6-MSA.<sup>81,82,138</sup> For MSAS from *Streptomyces antibioticus* (*S. antibioticus*), missing the Asp-containing motif, it was demonstrated that 6-MSA is not hydrolysed, but instead transacylated by a KAS-III like protein, annotated with ChIB3 (SanChIB3; similar to SerChIB3 in the *S. erythraea* gene cluster).<sup>83,84</sup> A phylogenetic analysis on the five synthases, that have been characterized so far, shows that proteins cluster in two branches, in line with the presence of the Asp-containing motif and their different origin (bacterial vs. fungal, figure 32 B).



**Figure 32** | (A) Amino acid alignment of different MSASs with Uniprot accession codes and origin. The alignment was created by CLC Main Workbench (version 6.9.1). The targeted enzyme is highlighted in red. Shown are motifs within the reported range for product hydrolysis in *A. terreus* (His972, alignment number 979; Asp1129, alignment number 1145).<sup>82</sup> Indeed, a conserved motif was found for the histidine in all analysed synthases. A conserved motif for the aspartate was found in MSASs from *P. patulum*, *A. aculeatus* and *A. terreus*, which is absent in *S. erythraea* and *S. antibioticus*. (B) Phylogenetic analysis (phylogram layout) of the selected MSASs, created by CLC Main Workbench (version 6.9.1) using the UPGMA algorithm.

Based on *in vivo* and *in vitro* experiments and the sequence alignment, we suggested that 6-MSA is produced by the synthase, but cannot be released by hydrolysis, as described for MSAS from *S. antibioticus* that is involved in the chlorothricin biosynthesis.<sup>83,84</sup> On this basis, we focused on the characterization of SerChIB3 and SerChIB2 from the SerMSAS gene cluster (figure 28, chapter 2.2).

## 2.2.4 Sequence analysis and expression of SerChIB3

On the basis of the SerChIB3 gene position and the studies of the related SanChIB3, we concluded that the analysis of the KASIII-like protein is of key importance for understanding the function of SerMSAS.<sup>83,84,139</sup>

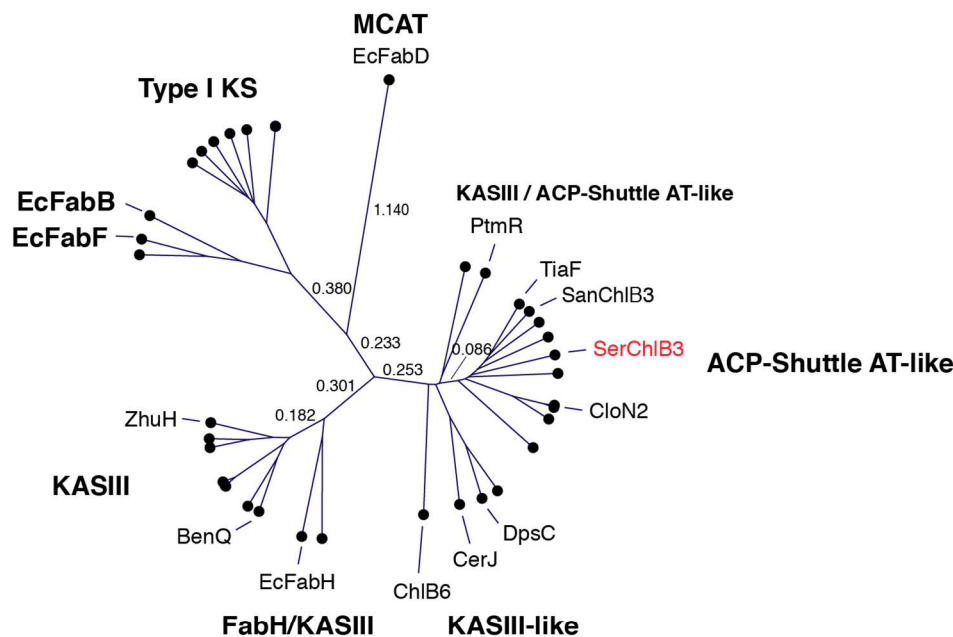
We performed a phylogenetic analysis of SerChIB3 similar to *Abugrain et al.* and *Bretschneider et al.* (figure 33 A) and analysed active site residues like *Jackson et al.* (figure 33 B).<sup>140–142</sup> On the basis of this analysis, we suggested SerChIB3 to be an ACP-dependent transferase rather than a condensing protein, because it is closely

related to TiaF, SanChIB3 and CloN2, all performing or suggested to perform esterification using ACP bound substrates. These proteins feature a Cys-His-Asp catalytic triad, like CerJ, ChIB6 and PtmR. In contrast, more distant related condensing enzymes like EcFabH or BenQ carrying a Cys-His-Asn catalytic triad (Cys-His-His for FabB and FabF) supporting the idea that SerChIB3 has rather AT than KS activity.<sup>35,139,142–153</sup>

In the phylogenetic analysis, SerChIB3 is arranged with KAS III enzymes identified to catalyse ester formation. C-O bond-formation was shown for CerJ, TiaF, CloN2, PtmR and ChIB6.<sup>84,140,141</sup> CerJ transfers a malonyl moiety from Mal-CoA onto the sugar moiety of cervimycin, but was also capable to use the derivatives MM-CoA and dimethylmalonyl-CoA.<sup>140</sup> CloN2, involved in the clorobiocin pathway, was found to transfer a pyrrole-2-carboxyl moiety to a desoxysugar unit, most likely the pyrrole-2-carboxyl is tethered to an ACP.<sup>154,155</sup> TiaF is suggested to transfer an orsellinic acid moiety to a 2-O-methyl-D-rhamnose residue.<sup>156</sup> PtmR was found to transfer 6-MSA with relaxed substrate specificity from an iterative PKS to an aminocyclopentitol unit.<sup>141</sup> In contrast ChIB6, involved in the chlorothricin biosynthesis, transfers an ACP tethered 6-MSA moiety to a demethylsalicylylchlorothricin.<sup>84</sup> The thioester formation between ACP and 6-MSA is catalysed by another KAS III protein, annotated as SanChIB3.<sup>83,84,87</sup> DpsC, a closely related enzyme to CerJ, was reported to use a propionyl starter substrate to initiate the daunorubicin biosynthesis. It's function as AT or KS and whether the enzyme has strict or narrow substrate specificity remains elusive.<sup>142,149,150</sup>

As the enzymes show specificities for different substrates, we analysed motifs to predict the substrate for SerChIB3. The analysis of middle pocket residues and active site cap unveiled similarities between transferases of aromatic units with a conserved glycine, a rather small unpolar amino acid and mainly a phenylalanine as middle pocket residue. Additionally, they share a conserved (besides of ChIB6) glycine as an open active site cap. These features are also present in SerChIB3, indicating the transfer of an aromatic unit rather than a small substrate, most likely excluding a DpsC like priming function. Additionally, a CerJ-like function can be ruled out for SerChIB3, since it misses the residues for coordinating the carboxylic group. Supported by the phylogenetic analysis and the amino acid alignment, we suggested SerChIB3 to transfer 6-MSA from SerMSAS to the discrete carrier protein SerChIB2, similar to SanChIB3.<sup>83,84</sup>

A



B

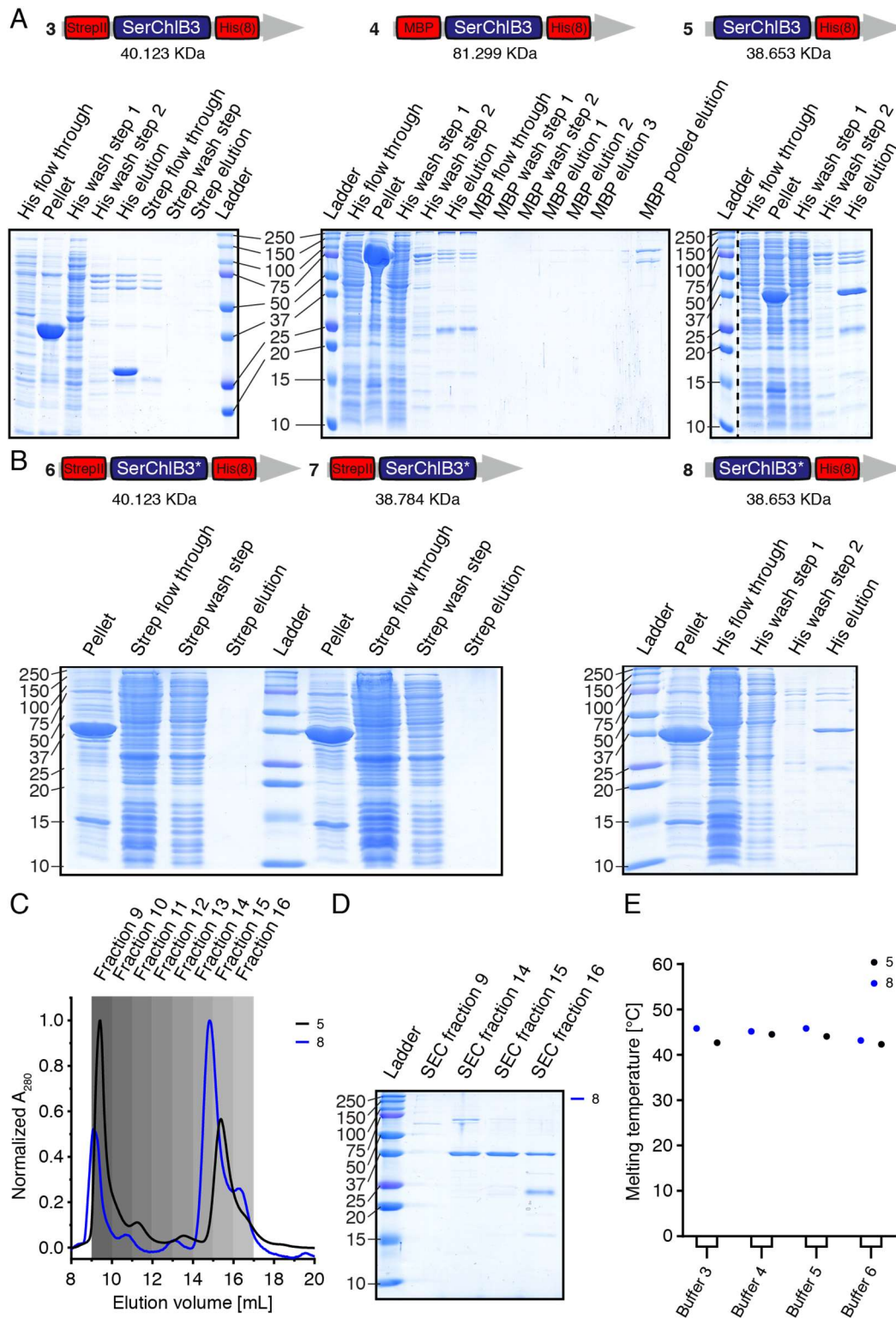
Protein	Active site residues				Middle pocket residues			Active site cap	Relative Substrate size	Malonyl binding in CerJ	
EcFabH	C112	H244	N274	S279	A111	F157	E200	-	Small ↓ Large	A111	I156
ZhuH	C121	H257	N287	S293	A120	F166	-	-		A120	L165
BenQ	C124	H261	N292	S297	A123	F169	-	-		A123	L168
DpsC	S118	P265	H297	D302	V117	-	L217	L93		V117	Y165
CerJ	C116	P264	H295	D300	R115	-	L216	L91		R115	Y164
CloN2	C113	H265	H296	D301	G112	L161	Y213	G88		G112	F160
ChIB6	C113	A264	H296	D301	G112	L160	V214	E88		G112	A161
PokM2	C112	M259	H291	D296	G111	V160	F212	G88		G111	L159
SanChIB3	C113	M264	H296	D301	G112	A160	F212	G88		G112	-
SerChIB3	C113	A264	H296	D301	G112	V160	F212	G88		G112	-
PtmR	C113	L262	H294	D299	G112	L160	M212	G88		G112	-
TiaF	C113	M261	H293	D298	G112	A160	F212	G88		G112	-
CalO4	C113	M261	H296	D298	G112	L160	F212	G88		G112	-
AviN	C113	M260	T292	D297	G112	G160	F211	G88		G112	-

**Figure 33** | Phylogenetic and active site residue analysis of the KASIII-like protein ChIB3. The target protein is highlighted in red. (A) Phylogenetic analysis (radial layout) of selected enzymes from the KAS III class was performed by CLC Main Workbench (version 6.9.1) using the UPGMA algorithm. A detailed list of all used enzymes is shown in figure S2. (B) Alignment of active site residues, middle pocket residues, active site cap and malonyl binding residues with ascending relative substrate size. The – indicate missing amino acid according to the alignment.

We focused on the purification of SerChIB3 to evaluate our idea about its function. We cloned the corresponding gene into a pET22b expression plasmid, containing a N-terminal StrepII-tag and a C-terminal His-tag for affinity purification (construct 3, figure 34 A). The protein was expressed, but was found to form insoluble inclusion bodies, implying stability or solubility problems (figure 34 A). Similar issues were reported for the related enzyme CerJ and attributed to the fusion of an N-terminal



tag.<sup>140</sup> We designed two more constructs, by changing the Strep-tag to a MBP tag (construct 4) to improve the solubility in water<sup>157,158</sup> and by removal of the N-terminal Strep-tag (construct 5). After purification with affinity chromatography, we received approximately 0.5 mg and 0.9 mg of construct 4 and 5 respectively. Judged by SDS-PAGE, both proteins were received with major impurities and high tendency to form aggregates (figure 34 A). We continued by performing codon optimization using the tool Jcat (<http://www.jcat.de/>, rho-independent transcription terminators were avoided, date 01.06.2018)., The codon optimized synthetic gene was purchased from ThermoFisher (GeneArt). The gene was cloned into a pET22b derived expression vector, designing three constructs with (i) both a N-terminal StrepII-tag and a C-terminal His-tag (construct 6), (ii) with just a N-terminal StrepII-tag (construct 7) and (iii) with just a C-terminal His-tag (construct 8). These constructs were cloned, expressed, purified and analysed during a bachelor thesis by Christian Gusenda.<sup>136</sup> After purification with affinity chromatography, we only received soluble protein for construct 8, confirming N-terminal tagging to be unfavourable for SerChIB3 (figure 34 B). Again, insoluble protein was found for all constructs 6-8. However, codon optimization improved the yield for the His-tagged protein from 0.9 mg/L to 2.7 mg/L. We continued with both C-terminal his-tagged proteins for all following applications. Both proteins were further purified with SEC on a S200 increase 10/300 GL column (figure 34 C). By using two chromatographic purification steps, we were able to receive pure SerChIB3, as judged by SDS-PAGE (figure 34 D). We compared the elution time to protein standards (ferritin, aldolase, ovalbumin) and estimated the apparent molecular weight of 65 KDa (calculated molecular weight of monomer is 39 KDa), suggesting SerChIB3 to be a dimer similar to DpsC and CerJ.<sup>140,142</sup> SEC runs were performed by Christian Gusenda during his bachelor thesis.<sup>136</sup> Next, we tested the stability of SerChIB3 in the purification buffers with a TSA. SerChIB3 seems to be less stable compared to SerMSAS, estimated on a lower melting temperature (figure 34 E). Nonetheless SerChIB3 is stable in all buffers used for purification, showing that the purification strategy of SerMSAS is also suitable for SerChIB3.

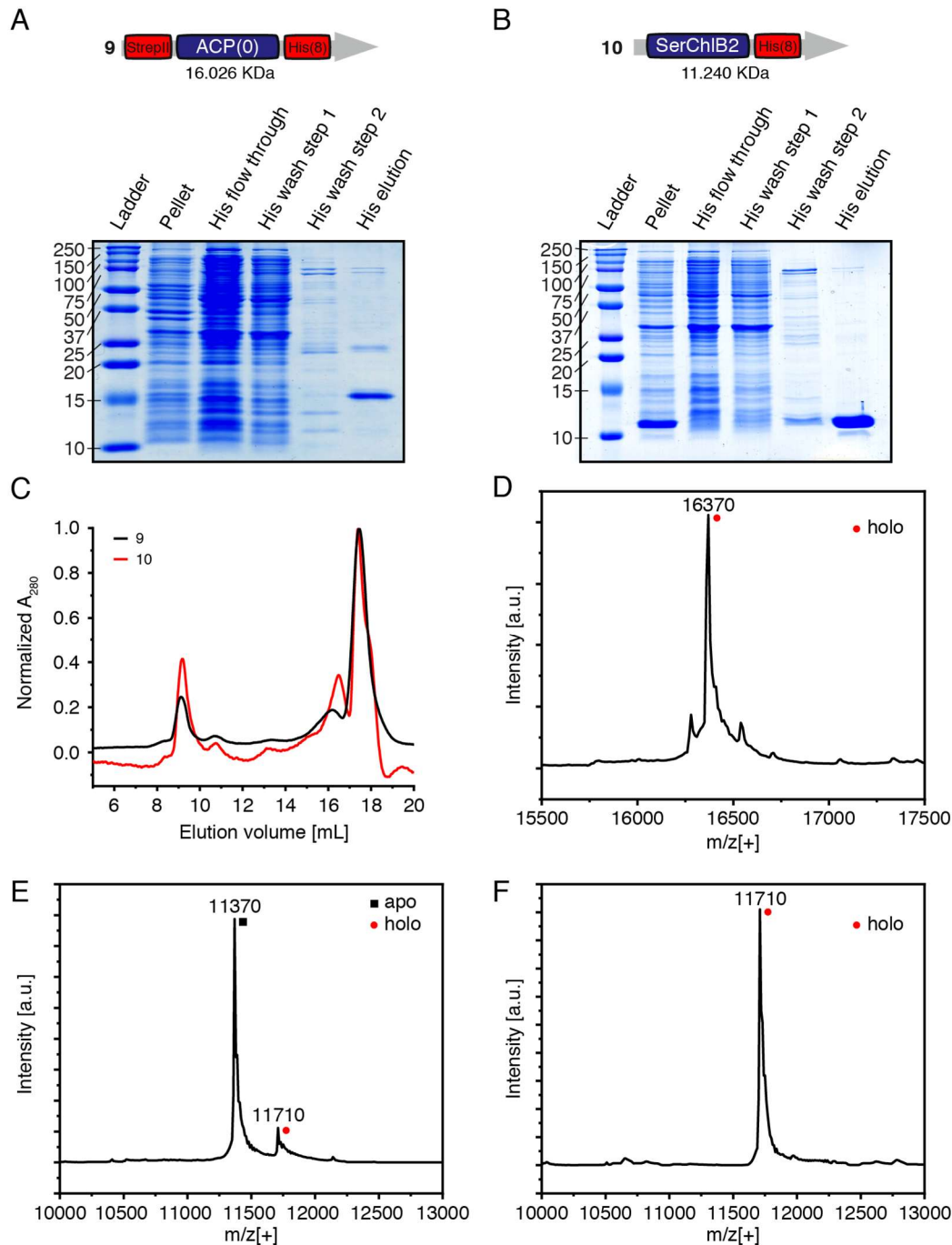


**Figure 34** | Establishment of a purification strategy for SerChIB3. Data were received within three bachelor theses, supervised by me.<sup>136,137,159</sup> (A) Coomassie-stained SDS-PAGE with samples taken during SerChIB3 purification using appropriate affinity chromatography strategies for the different constructs. Gel of construct 5 was modified for the sake of visualisation. Complete gel is illustrated in figure S3 (B) Coomassie-stained SDS-PAGE with samples taken during the codon optimized SerChIB3 purification using appropriate affinity chromatography strategies for the different constructs. Codon optimization is indicated with an asterisk. (C) SEC of SerChIB3 constructs 5 and 8 on a S200 Increase 10/300 GL column (in 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20%

glycerol and 1 mM EDTA). (D) Coomassie-stained SDS-PAGE of the SEC fractions of construct 8. (F) Melting temperatures of a TSA with SerChIB3 construct 5 and 8. Assay was performed in technical triplicates (n = 3). Errors were smaller than 1°C and are not shown. All buffers used for purification were analysed (buffer 3: 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 1 mM EDTA; buffer 4: 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 20 mM imidazole; buffer 5: 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 60 mM imidazole; buffer 6: 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 300 mM imidazole).

### 2.2.5 Expression of SerMSAS-ACP(0) and SerChIB2

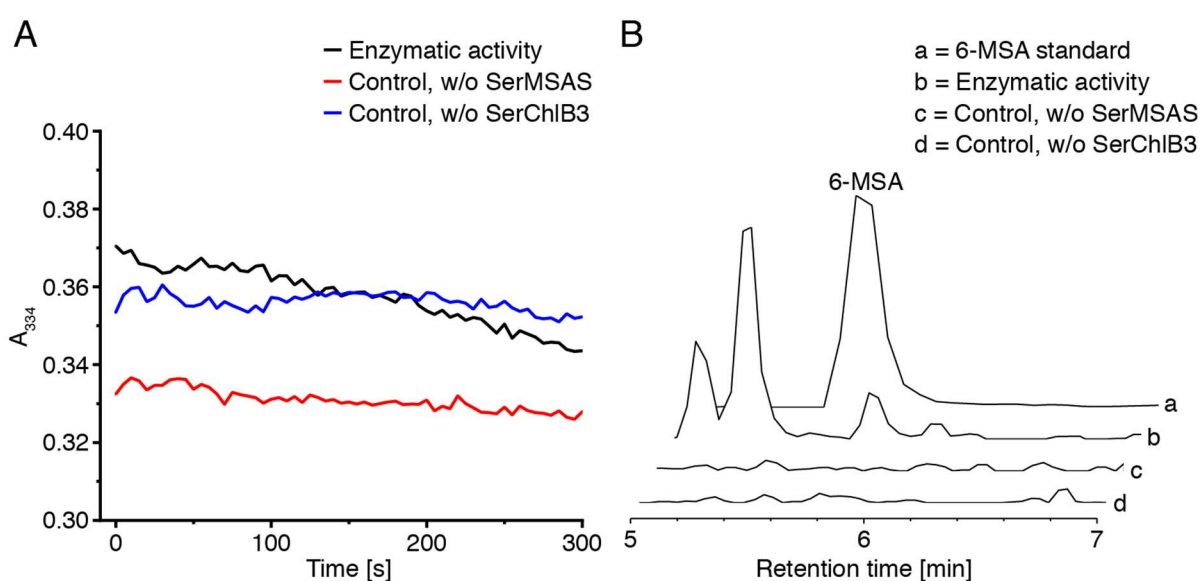
To be able to evaluate the transfer of 6-MSA moieties *similar to He and Yi et al.*<sup>83,84</sup>, we required purified SerMSAS-ACP(0), i.e. ACP of MSAS as separate protein, and SerChIB2. Both genes were cloned into pET22b derived expression plasmids. The discrete protein SerChIB2 was directly cloned from genomic DNA. In contrast, the gene encoding for ACP(0) had to be dissected from the SerMSAS, which was performed by Franziska Stegemann (Goethe-University, Frankfurt). Both proteins were co-expressed with the PPTase *Sfp* in *E. coli* and purified analogously to SerMSAS, except that the S200 Increase 10/300 GL column was used for SEC (figure 35 A-C). ACP(0) and ChIB2 could be purified in sufficient amount of 2 mg and 28 mg respectively. Codon optimization of ACP(0) further improved the yield to 24 mg (see figure S4 for purification and ESI analysis; codon optimization by Franziska Stegemann, protein purification and analysis during bachelor theses of Christian Gusenda, Christian Schreiber und Trinetri Goel<sup>136,137,159</sup>). Mass spectroscopic analysis of ACP(0), performed and analysed by Khanh Vu Huu (Goethe-University, Frankfurt), showed complete phosphopantetheinylation regardless if the gene was codon optimization or not (figure 35 D, figure S4 C). This was expected since SerMSAS was also fully activated by *Sfp* (figure 29G, chapter 2.2.1). However, SerChIB2 was not fully phosphopantetheinylated by *Sfp* (figure 35 E). In contrast co-expression with the PPTase *Npt* led to complete activation of SerChIB2 (figure 35 F; this work was again performed during the three bachelor theses supervised by me<sup>136,137,159</sup>).



**Figure 35** | Establishment of a purification strategy for SerMSAS-ACP(0) and SerChIB2. Data were received within three bachelor theses supervised by me.<sup>136,137,159</sup> (A) Coomassie-stained SDS-PAGE with samples taken during ACP(0) purification, co-expressed with *Sfp*, by IMAC. (B) Coomassie-stained SDS-PAGE with samples taken during SerChIB2 purification, co-expressed with *Sfp*, by IMAC. (C) SEC of construct 9 and 10 on a S200 Increase 10/300 GL column (in 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 1 mM EDTA). (D) Mass spectroscopic analysis of ACP(0), co-expressed with *Sfp*, was performed with ESI in positive mode (SYNAPT G-2, waters). The theoretical mass for the holo-form is 16,366 Da. (E) Mass spectroscopic analysis of SerChIB2, co-expressed with *Sfp*, was performed with ESI in positive mode (SYNAPT G-2, waters). The theoretical mass for the apo- and holo-form is 11,371 Da and 11,711 Da, respectively. (F) Mass spectroscopic analysis of SerChIB2, co-expressed with *Npt*, was performed with ESI in positive mode (SYNAPT G-2, waters). The theoretical mass for the holo-form is 11,711 Da.

### 2.2.6 *In vitro* 6-MSA synthesis with SerChIB3

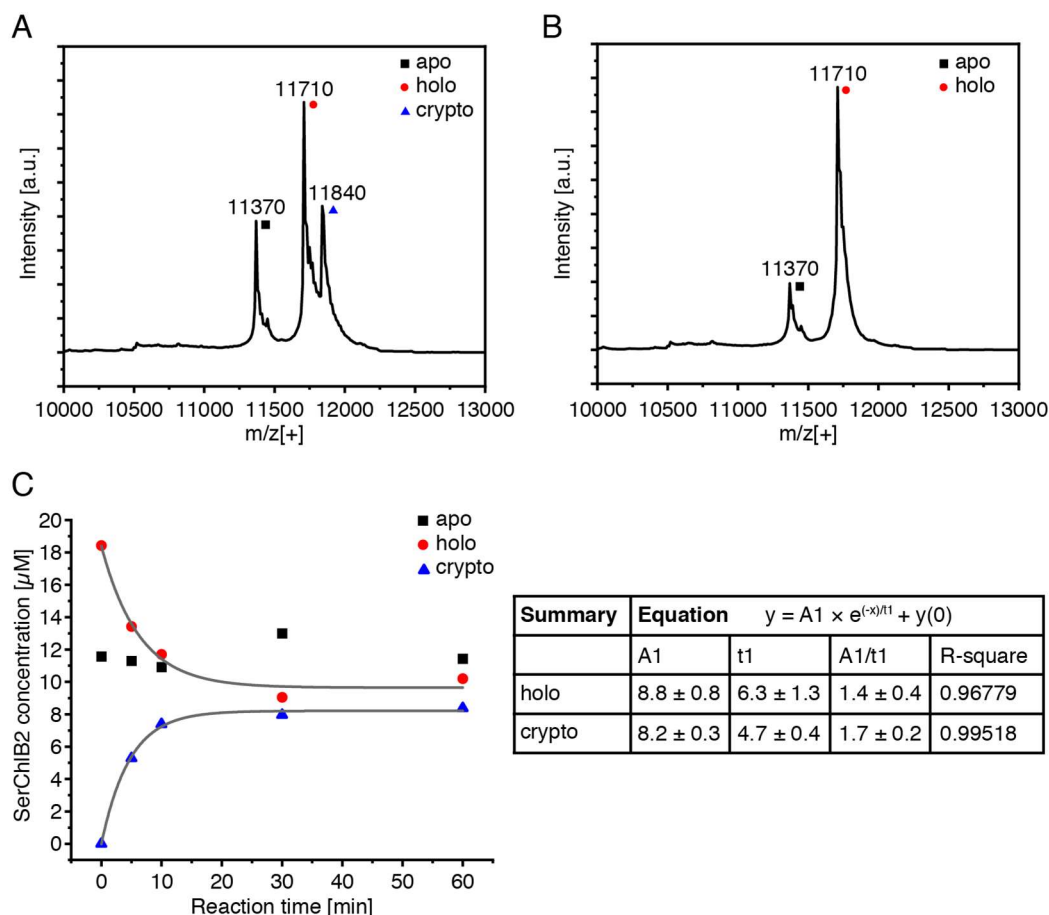
With the access to purified SerChIB3 and SerChIB2, the SerMSAS activity *in vitro* was investigated (by Trinetri Goel).<sup>137</sup> All three proteins were incubated with the substrates acetyl-CoA, Mal-CoA and NADPH. The NADPH consumption was monitored by decrease of its absorbance at 334 nm and converted to concentrations by using an extinction coefficient of  $6,220 \text{ M}^{-1}\text{cm}^{-1}$ . In this experimental design, 6-MSA was produced with a turnover rate of  $0.26 \text{ min}^{-1}$  (figure 36 A, experimental data were corrected by the background in the absence of SerMSAS). Furthermore 6-MSA was verified by HPLC-MS after acidification (figure 36 B). The experimental data indicate a terminating function of SerChIB3 like it was discussed in chapter 2.2.4, since no 6-MSA synthesis was observed in the absence of SerChIB3 (figure 31, chapter 2.2.3).



**Figure 36 |** SerMSAS activity *in vitro* in the presence and absence of SerChIB3. Data were received within the bachelor thesis of Trinetri Goel.<sup>137</sup> (A) Analysis of the SerMSAS enzymatic activity by NADPH consumption *in vitro*. SerMSAS ( $0 \mu\text{M}$  or  $2 \mu\text{M}$ ) was incubated with  $20 \mu\text{M}$  SerChIB2,  $0 \mu\text{M}$  or  $2 \mu\text{M}$  SerChIB3,  $100 \mu\text{M}$  acetyl-CoA,  $100 \mu\text{M}$  Mal-CoA and  $50 \mu\text{M}$  NADPH for five minutes (in  $200 \text{ mM}$  potassium phosphate (pH 7.0),  $100 \text{ mM}$  sodium chloride,  $100 \text{ mM}$  potassium chloride,  $10\%$  glycerol,  $1.3 \text{ mM}$  DTT). The background sample was performed without SerMSAS or SerChIB3. Absorption of NADPH at  $334 \text{ nm}$  was measured every five seconds with the Nanodrop 2000C (Thermo Scientific). Data points were connected for visualization. (B) All reaction mixtures from A were extracted after acidification and analysed with HPLC-MS [(EIC):  $[\text{M}-\text{H}]^- \text{ m/z} = 151.04 \pm 0.10$ ]. The negative control was performed without SerMSAS or SerChIB3. 6-MSA standard from figure 30 of chapter 2.2.2 was shown for comparison. Data were normalized with respect to the highest peak. The negative controls were normalized with the same factor as the enzymatic activity sample.

### 2.2.7 SerChIB3 can transfer 6-MSA moieties

We aimed for 6-MSA-ACP(0) to address the predicted transfer function of SerChIB3. Unfortunately we failed to produce the thiophenyl ester of 6-MSA (6-MSA-SPh) for direct acylation of SerMSAS-ACP(0) as well as to acylate CoA for synthesis of 6-MSA-CoA, which can be used to convert the apo-protein to 6-MSA-ACP(0) by using a PPTase.<sup>83</sup> In contrast 6-MSA-SNAC was successfully synthesised analogue to *He et al.* by esterification of 6-MSA with N-acetycysteamine, using *N*-ethyl-*N'*-(3-dimethylaminopropyl)-carbodiimide (EDC-HCl) and 4-(dimethylamino)pyridine (DMAP) as coupling agents.<sup>83</sup> The synthesis was performed by Chris Gusenda during his bachelor thesis.<sup>136</sup> We used 6-MSA-SNAC as substrate, mimicking the 6-MSA-ACP(0).<sup>39</sup> The substrate was then incubated with SerChIB2 and with or without SerChIB3 for 30 minutes at room temperature. We used SerChIB2, co-expressed with *Sfp* in all transfer experiments as the fully phosphopantetheinylated SerChIB2 by co-expression with *Npt* was not established to that date. However, we increased the holo to apo ratio by incubation of SerChIB2 with *Sfp* and CoA *in vitro* (figure 37 B in comparison to figure 35 E of chapter 2.2.5). Mass spectroscopic analysis of SerChIB2 confirmed the predicted transfer function of SerChIB3, because 6-MSA-SerChIB2 (crypto-form) was only generated in the presence of SerChIB3 (figure 37 A-B). In another experiment, we tried to address turnover rates of SerChIB3 and quenched the reaction at certain times (figure 37 C, figure S5).<sup>83,84</sup> SerChIB3 was found to transfer 6-MSA moieties from the substrate 6-MSA-SNAC to SerChIB2 with a turnover rate of  $3.9 \pm 0.8 \text{ min}^{-1}$ . This turnover rate is much slower, compared to the similar enzyme ChIB3 from *S. antibioticus*, which transfers 6-MSA moieties from the MSAS-ACP to another carrier protein with a turnover rate of  $73 \text{ min}^{-1}$ .<sup>83,84</sup> The huge discrepancy was reported before and was attributed to the missing protein-protein interactions, due to the replacement of ACP with a chemical analogue.<sup>83</sup> However the data show clearly that SerChIB3 is capable to transfer 6-MSA moieties with a much higher turnover rate compared to the 6-MSA production by MSAS, described in the previous chapter. The experimental work was performed by Chris Gusenda during his bachelor thesis and ESI measurements and analysis was performed by Khanh vu Huu (Goethe-University, Frankfurt).<sup>136</sup>

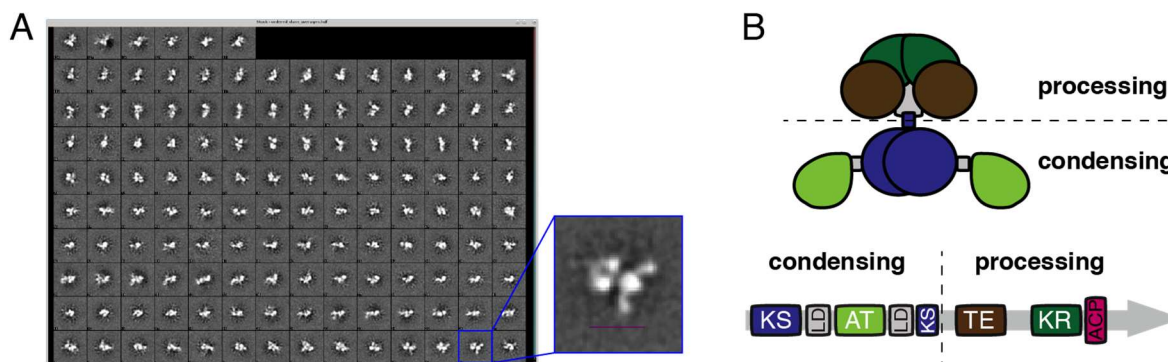


**Figure 37** | Transacylation of 6-MSA moieties by SerChIB3. Data were received from the bachelor thesis of Chris Gusenda supervised by me.<sup>136</sup> (A) Mass spectroscopic analysis of SerChIB2 (33  $\mu$ M), incubated with 0.2  $\mu$ M SerChIB3 and 3.3 mM 6-MSA-SNAC for 30 minutes at room temperature, was performed with ESI in positive mode (SYNAPT G-2, waters). The theoretical mass for the apo-, holo- and crypto-form (6-MSA-SerChIB2) is 11,371 Da, 11,711 Da and 11,845 Da respectively. (B) The negative control was performed like in A but in absence of SerChIB3. (C) Transfer of 6-MSA-SNAC to SerChIB2 was monitored with ESI in positive mode (SYNAPT G-2, waters). The reaction contained 30  $\mu$ M SerChIB2, 0.4  $\mu$ M SerChIB3 and 15.0 mM 6-MSA-SNAC and was quenched at specific times (0 min, 1 min, 5 min, 10 min, 30 min, 60 min). OriginPro 8.5 was used to integrate the apo-form (11,300-11,500 Da), holo-form (11,650-11,800 Da) and crypto-form (11,800-11,900 Da) and final concentrations were calculated by multiplying the ratio of each form with the initial SerChIB2 concentration. Shown is the concentration of each form at different times. The turnover rate was calculated by an exponential decay function with OriginPro 8.5, including the increase of the crypto-form and the decrease of the holo-form.

## 2.2.8 Initial electron microscopy data

Structural information is important for rational protein design and is currently not available for MSAS. For that reason, we performed initial EM experiments with the

purified SerMSAS construct 2. We performed negative staining experiments and analysed them with 2D classifications, which already showed promising class averages (figure 38 A). An arrangement of SerMSAS was postulated according on the data (figure 38 B), indicating that the structure of SerMSAS could be similar to the model of the mycocerosic acid synthase<sup>73</sup> and is rather not comparable to the model of the PIKS.<sup>160</sup>



**Figure 38** | EM experiments with SerMSAS. (A) 2D classification from negative stain images using EM. Grid preparation, EM experiments and the processing of the data was performed by Eduardo D’Imprima (Max Planck Institute for Biophysics, Frankfurt). (B) Postulated, arrangement of the SerMSAS domains.

### 2.3 Engineering of PenPaMSAS towards 6-PrSA synthesis

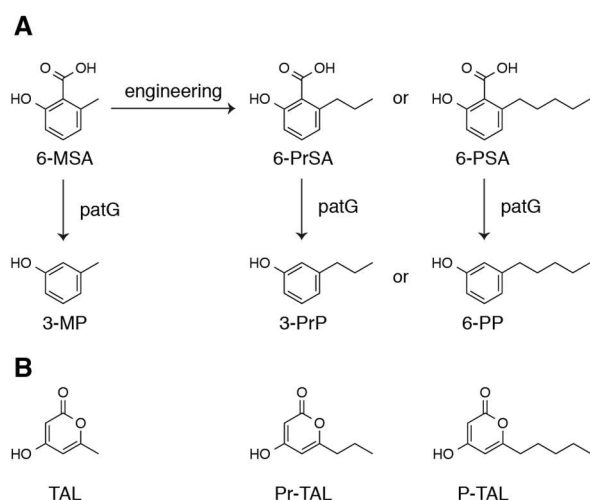
Given that SerMSAS was characterized in this work as a non-releasing enzyme, we focused on a MSAS from a different organism as a target for structural and functional characterization and engineering for synthesis of 6-MSA derivatives. PenPaMSAS has been purified already 1970<sup>81</sup> in its native host and later also in *E. coli* and yeast, and characterized in its basic function.<sup>91</sup> The initial focus was on SerMSAS, because we were not able to purify this protein in *E. coli* previously. However, applying the purification strategy for SerMSAS on PenPaMSAS eventually enabled high amount of active protein.

The aim of this project was the production of 3-MP derivatives in yeast, more specifically 3-PrP and 3-pentyl phenol (3-PP) (figure 39 A). 3-PrP is an attractant for the Tsetse-fly. Applied in traps, this compound allows to control the population and therefore reduces the risk of transmitting the sleeping sickness.<sup>102</sup> 3-PP may have similar properties, although not described in literature. Biosynthesis of 3-PrP and 3-PP in yeast is tempting, because it could deliver a simple, cheap, and decentralized access to the compounds that could easily be used by rural tribes. Besides that, alkylphenols are valuable organic chemicals, and sustainable access to those compounds is technologically relevant. We sought to produce 3-PrP and 3-PP by yeast fermenting sugar and delivering the substrates acyl- and Mal-CoA for



PenPaMSAS. For 3-PrP and 3-PP, PenPaMSAS needs to accept the non-native starter substrates butyryl-CoA and hexanoyl-CoA to produce 6-PrSA and 6-pentyl salicylic acid (6-PSA), respectively, which are then converted to 3-PrP and 3-PP by the decarboxylase PatG (figure 39 A). Using that strategy, an overall yield of about 2 g/L 6-MSA, 589 mg/L m-cresol, 12.5 mg/L 3-EP and 2.6 mg/L 3-PrP was recently achieved in an engineered strain of *S. cerevisiae*, showing the feasibility of this approach.<sup>92,93</sup> However, this metabolic engineering approach just led to limited 3-PrP production, because the natural starter substrate acetyl-CoA, present as a central metabolite, presumably outcompetes butyryl-CoA loading. This demonstrates that engineering of PenPaMSAS is required to improve the yield of 3-PrP.<sup>93</sup> Moreover PenPaMSAS was also described *in vitro* to convert butyryl-CoA and hexanoyl-CoA at significantly slower rates (87%, ~94% respectively) compared to the native substrate acetyl-CoA.<sup>38</sup> The reason for this significantly slower reaction is not described in literature, but in principle reflects a narrow substrate specificity of at least one catalytic domain.

In this study, we aimed to discover the domain(s) imposing substrate specificity to PenPaMSAS and enlarge specificity bottlenecks by rational protein engineering. Specifically, we focused on producing 6-PrSA and 6-PSA by the improved acceptance of longer starter units. We started with the characterization of the AT, since it is responsible for the selection of the substrates. Engineered PenPaMSAS should then pave the way to improved production of 3-PrP and 3-PP *in vivo* following the approach of *Hitschler et al.*<sup>93,161</sup> PatG is able to produce 6-MP and 3-PrP from 6-MSA and 6-PrSA *in vivo*, but so far nothing is reported about the specificity of this enzyme, which will not be addressed in this study.<sup>92,93,161</sup>

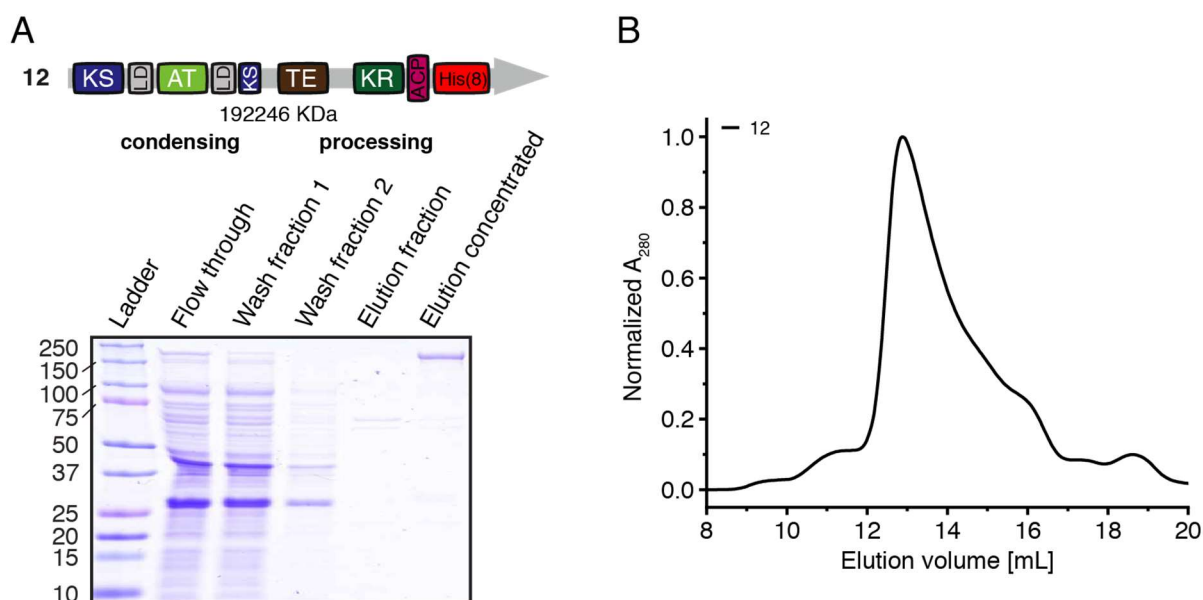


**Figure 39** | Target compounds in this study. (A) MSAS can produce 6-MSA out of one moiety acetyl-CoA, one moiety NADPH and three moieties Mal-CoA. 6-PrSA and 6-PSA can be synthesized if acetyl-CoA is replaced by butyryl-CoA and hexanoyl-CoA, respectively.<sup>38</sup> PatG is able to decarboxylate 6-MSA to 3-MP, 6-PrSA to 3-PrP and supposed to decarboxylate 6-PSA to 6-PP.<sup>92,93,161</sup>

(B) During the biosynthesis of MSA and its derivatives, the triketide intermediate can also be lactonized to the shunt products TAL, Pr-TAL and P-TAL, particularly when absence of NADPH hinders the final elongation to the tetraketide intermediate.

### 2.3.1 Expression and purification of PenPaMSAS

For purification of PenPaMSAS, we applied the design and the established purification strategy for SerMSAS. PenPaMAS was cloned into a pET22b derived expression plasmid with only a C-terminal His-tag for chromatographic purification. Alexander Rittner performed the design and cloning of the construct. The construct was expressed and purified accordingly to the established strategy from SerMSAS to yield  $2.1 \pm 0.6$  mg/L PenPaMSAS after chromatographic purification by IMAC and SEC (figure 40 A-B).



**Figure 40** | Purification of PenPaMSAS. **(A)** Domain structure of PenPaMSAS and Coomassie-stained SDS-PAGE with samples taken during its purification by IMAC. PenPaMSAS was co-expressed with *Sfp* **(B)** SEC of construct 12 on a S6 Increase 10/300 GL column (in 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 1 mM EDTA).

### 2.3.2 Characterization of PenPaMSAS

After successful purification of PenPaMSAS in *E. coli*, we first evaluated the reported activity for different starter substrates.<sup>38</sup> To do so, we established a protocol for monitoring the activity of the synthase, by analysing the NADPH consumption fluorometrically with a plate reader. With this approach, we can investigate PenPaMSAS activity independently from the applied substrate. After calibrating the fluorescence read-out to the molar concentration of NADPH (figure 41 A), we

analysed PenPaMSAS activity with the native substrates acetyl-CoA and Mal-CoA. With the validation that the protein was active, we decided to optimize the buffer composition. In a first round of buffer optimisation, we used 60 mM potassium phosphate as a fixed component and analysed the influence of glycerol and sodium chloride. No significant difference in turnover rates was observed in the absence and presence of 10% glycerol (figure 41 B, buffer 7-8). However, we decided to use 10% glycerol, because of its capability to stabilize enzymes.<sup>162</sup> In contrast, turnover rates decreased to 55% and 19% in the presence of 100 mM and 450 mM sodium chloride (figure 41 B, buffer 8-10) and sodium chloride was omitted consequently. In a second round of buffer optimisation, we optimized the bovine serum albumin (BSA) concentration. Turnover rates of PenPaMSAS decreased to 80% in the absence of BSA (figure 41 C, buffer 11 and 8), but no significant difference was visible when the BSA concentration was increased from 0.1 mg/mL to 0.4 mg/mL (figure 41 C, buffer 8 and 12). We chose the lower amount of BSA, because it is easier to avoid bubbles during pipetting. Finally, buffer 8 turned out to be best suited from our optimisation with the final concentrations of 60 mM potassium phosphate, 0.1 mg/mL BSA and 10% glycerol in all following experiments.

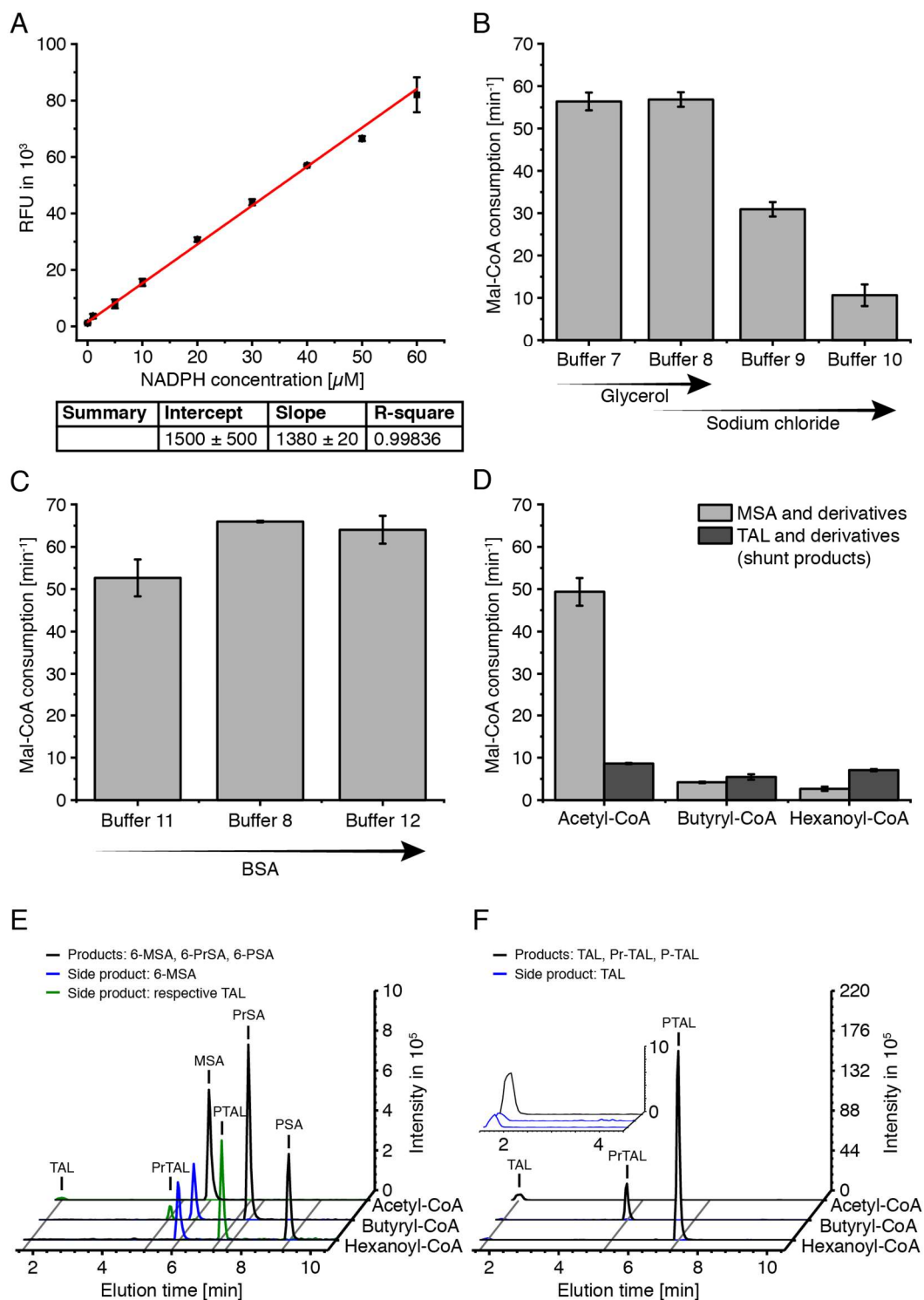
We utilized the established protocol to compare the enzymatic activity in the presence of the different starter substrates acetyl-CoA, butyryl-CoA and hexanoyl-CoA, similar to *Dimroth et al.*<sup>38</sup> When incubated with Mal-CoA, MSAS is capable to convert the different starter substrates to 6-MSA, 6-PrSA and 6-PSA in the presence of NADPH and to the shunt products TAL, Pr-TAL and 6-TAL in the absence of NADPH (figure 41 D). TAL production was analysed by its absorbance in cuvettes at 298 nm. For simplicity, we assumed that the extinction coefficient of  $2,540 \text{ M}^{-1}\text{cm}^{-1}$  reported for TAL is valid also for the derivatives.<sup>163,164</sup> In contrast to the biosynthesis of salicylic acids, the production of lactones does not require the domains KR and TE but occurs via spontaneously lactonization from the triketide intermediate. By analysing and comparing production rates of both products with different starter substrates, we could get valid information about the substrate specificities of PenPaMSAS. That's why all production rates are presented in consumed molecules of Mal-CoA per minute, corresponding to one condensation step of the enzyme.

We finally recorded 6-MSA productions with a turnover rate of  $49 \pm 4 \text{ min}^{-1}$ . This rate is slightly faster compared to literature, indicating that the introduction of a C-terminal His-tag and the purification strategy are suitable for this enzyme.<sup>81</sup> This rates are, however, slightly slower compared to those received when establishing the protocol, which we explained by different approaches in background correction; i.e., taking background in absence of protein or in absence of starter substrate. In comparison, we received a higher background in the absence of starter substrates and decided to use this as a better background correction for all following assays. The shunt product

TAL is produced significantly slower (17%) compared to 6-MSA, which is in good agreement with the literature.<sup>81</sup> The slower production rate of the shunt product was expected because it reflects a side reaction of 6-MSA production that should be hindered under physiological (reducing) conditions. 6-PrSA and 6-PSA are produced at rates of 8% and 5% of the 6-MSA synthesis, which is also in good agreement with the literature.<sup>81</sup> This tendency has been also reported for *in vivo* studies, showing significant lower yields for the decarboxylated products 3-EP and 3-PrP compared to 3-MP.<sup>93</sup> The shunt product synthesis is slightly higher for butyryl-CoA (130%) and significantly higher for hexanoyl-CoA (265%) compared to 6-PrSA and 6-PSA. This indicates that the processing of the triketide intermediate by domains KR, KS and TE becomes slower when the starter substrate chain length increases, so that in turn spontaneous lactonization becomes more prominent.

The identity of products from all assays with the different starter substrates were extracted and analysed with HPLC-MS. The correct salicylic acid products in the presence of NADPH (figure 41 E) and the correct lactone shunt products in the absence of NADPH (figure 41 F) could be verified (black lines), demonstrating the feasibility of both assays (see figure S6 for compound list). Surprisingly, 6-MSA was found as side product in assays with butyryl-CoA and hexanoyl-CoA in the presence of NADPH and TAL as side product in the absence of NADPH (figure 41 E-F, blue lines). We speculate that either PenPaMSAS can decarboxylate Mal-CoA to acetyl-CoA or substrates or protein preparations contained acetyl-CoA impurities. The ACP, purified as a standalone protein, was shown to be acetylated (figure 42, chapter 2.3.3). This indicates that the production of 6-MSA and TAL, in the absence of acetyl-CoA as a substrate, results at least in part from purified acetylated enzyme (figure 42 D, chapter 2.3.3). As mentioned above, we also received lower background for negative controls in the absence of enzyme, again indicating that impurities in purified proteins can cause the TAL and 6-MSA production in absence of acetyl-CoA as a substrate. We note that, production rates of figure 41 D from all products and shunt products may therefore be affected systematically. Since we only analysed data from the time span of 5-10 minutes, where acetyl moieties originating from impurities in protein preparations should have been consumed already, systematic errors should be small. We finally identified the respective lactone shunt products as side product in assays conducted in presence of NADPH (TAL, Pr-TAL and P-TAL), especially with butyryl-CoA and hexanoyl-CoA as starter substrates (figure 41 E, green lines). This finding is correlating with the lactone production rates observed in non-reductive assays (figure 41D). We note that the peak intensity of the products and shunt products is not correlating with the production rates. This is not surprising, because peak intensities recorded in HPLC-MS must not correlate with its concentrations due to different ionizableness of the compounds, but also due to differences in sample preparation caused by different extractability from the assay solution. For example, we speculate that compounds with higher polarity like TAL will tend to remain in the aqueous phase, different to P-TAL that has better solubility in the organic solvent.

In conclusion, characterization of PenPaMSAS showed the slower conversion of butyryl-CoA and hexanoyl-CoA, compared to acetyl-CoA. The data indicate that at least one domain is limiting the synthesis of these derivatives and protein engineering is required to improve their production rates. The appearance of the shunt products during 6-PrSA and 6-PSA synthesis suggests that the KR-domain is not capable to reduce the triketide sufficiently fast and engineering of the KR-domain may be a suited strategy for improved yields. We note that a narrow specificity of KRs has been reported for other iterative PKSs before.<sup>90</sup> However, we first targeted the AT for engineering, because the overall lower rates for longer-chain substrates (for both lactone and salicylic acid production) indicate a generally compromised acceptance of butyryl- and hexanoyl-CoA by PenPaMSAS. Analysis of the AT was done on the basis of a coupled enzyme assay that was established in our lab before.<sup>34</sup>



**Figure 41** | Establishment of a protocol and a suited buffer system to probe PenPaMSAS with different starter substrates in the presence and absence of NADPH **(A)** Fluorometric NADPH calibration in 384 white well plates with excitation and emission wavelengths of 348 nm and 476 nm, respectively. **(B)** Analysing buffers with different amounts of glycerol and sodium chloride. [buffer 7: 60  $\mu$ M potassium phosphate (pH7.6), 0.1 mg/mL BSA; buffer 8: 60  $\mu$ M potassium phosphate (pH 7.6), 0.1 mg/mL BSA, 10% glycerol; buffer 9: 60  $\mu$ M potassium phosphate (pH 7.6), 0.1 mg/mL BSA, 10% glycerol, 100 mM sodium chloride; buffer 10: 60  $\mu$ M potassium phosphate (pH 7.6), 0.1 mg/mL BSA, 10% glycerol, 450 mM sodium chloride]. NADPH consumption was measured fluorometrically for

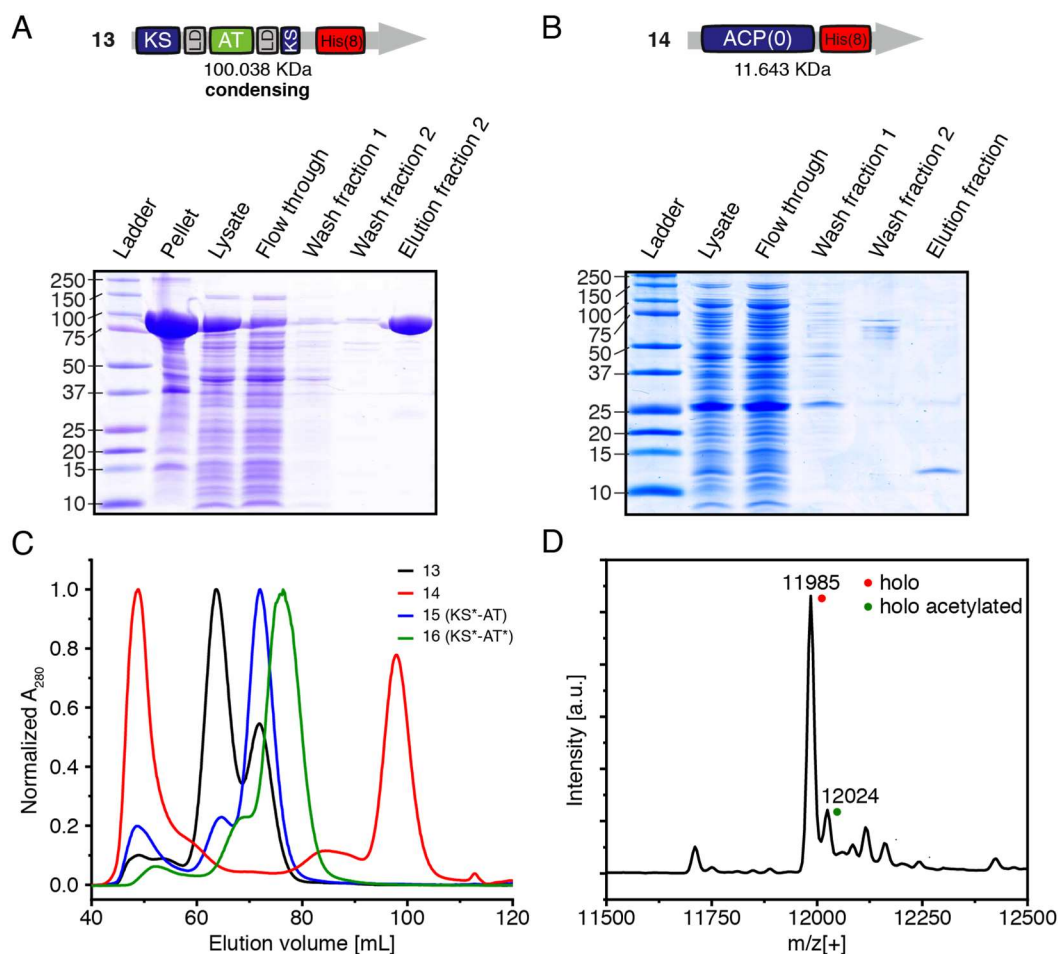
10 minutes at 25°C in 384 white well plates with excitation and emission wavelengths of 348 nm and 476 nm, respectively. Final concentrations in the assays were 0.125  $\mu$ M PenPaMSAS, 60  $\mu$ M NADPH, 200  $\mu$ M acetyl-CoA and 200  $\mu$ M Mal-CoA. The background correction was performed in the absence of enzyme. **(C)** Analysing buffers with different amounts of BSA [buffer 11: 60  $\mu$ M potassium phosphate (pH 7.6), 10% glycerol; buffer 8: 60  $\mu$ M potassium phosphate (pH 7.6), 0.1 mg/mL BSA, 10% glycerol; buffer 12: 60  $\mu$ M potassium phosphate (pH 7.6), 0.4 mg/mL BSA, 10% glycerol]. NADPH consumption was measured fluorometrically for 10 minutes at 25°C in 384 white well plates with excitation and emission wavelengths of 348 nm and 476 nm, respectively. Final concentrations in the assays were 0.125  $\mu$ M PenPaMSAS, 60  $\mu$ M NADPH, 200  $\mu$ M acetyl-CoA and 200  $\mu$ M Mal-CoA. The background correction was performed in the absence of enzyme **(D)** Comparing the turnover of Mal-CoA by PenPaMSAS when incubated with the different starter substrates acetyl-CoA, butyryl-CoA and hexanoyl-CoA in the presence and absence of NADPH. NADPH consumption was measured fluorometrically for 10 minutes at 25°C in 384 white well plates with excitation and emission wavelengths of 348 nm and 476 nm, respectively. Final concentrations in the assays were 0.125  $\mu$ M PenPaMSAS for acetyl-CoA and 1  $\mu$ M PenPaMSAS for butyryl-CoA and hexanoyl-CoA, 60  $\mu$ M NADPH, 200  $\mu$ M acetyl-CoA or 200  $\mu$ M butyryl-CoA or 200  $\mu$ M hexanoyl-CoA and 200  $\mu$ M Mal-CoA. Only data from 5-10 minutes were used for calculating the turnover rates, since fluctuations were visible during the first 5 minutes. The shunt product TAL and its derivatives were analysed by the increase of its absorbance at 298 nm for 10 minutes at room temperature. Final concentrations in the assays were 0.5  $\mu$ M PenPaMSAS, 200  $\mu$ M acetyl-CoA or 200  $\mu$ M butyryl-CoA or 200  $\mu$ M hexanoyl-CoA and 200  $\mu$ M Mal-CoA. The background correction was performed in the absence of the starter substrate. **(E)** All reaction mixtures for the 6-MSA, 6-PrSA and 6-PSA synthesis from D were extracted after acidification and analysed with HPLC-MS [(EIC for **6-MSA**): [M-H]<sup>-</sup> m/z = 151.04; (EIC for **6-PrSA**): [M-H]<sup>-</sup> m/z = 179.05; (EIC for **6-PSA**): [M-H]<sup>-</sup> m/z = 207.10; (EIC for **TAL**): [M-H]<sup>-</sup> m/z = 125.02; (EIC for **Pr-TAL**): [M-H]<sup>-</sup> m/z = 153.05; (EIC for **P-TAL**): [M-H]<sup>-</sup> m/z = 181.08. **(F)** All reaction mixtures for the TAL, Pr-TAL and P-TAL synthesis from D were extracted and analysed with HPLC-MS [(EIC for **TAL**): [M+H]<sup>+</sup> m/z = 127.04; (EIC for **Pr-TAL**): [M+H]<sup>+</sup> m/z = 155.07; (EIC for **P-TAL**): [M+H]<sup>+</sup> m/z = 183.10].

### 2.3.3 Purification of dissected PenPaMSAS domains

To be able to characterize the AT in enzyme kinetic properties, we designed a KS-AT didomain and a standalone ACP(0) by dissecting the PenPaMSAS encoding gene. We also designed a KS-knockout (C204G) of the KS-AT didomain to control whether the KS-domain has a significant influence on the transacylation of CoA-esters, and a KS- and AT-knockout (C204G, S653G) used as a negative control.

All constructs were purified with the established purification strategy for SerMSAS yielding pure proteins (62 mg/L for KS-AT; 23 mg/L for KS\*(C204G)-AT; 28 mg/L KS\*(C204G)-AT\*(S653G); 4 mg/L for ACP(0)) after IMAC (figure 42 A-B). We used the HiLoad 16/600 Superdex 200 for SEC (figure 42 C), showing two peaks for the KS-AT at an elution volume of 64 mL and 72 mL. We used a calibration with protein standards to estimate the mass of both peaks (239 KDa for 64 mL; 116 KDa for 72 mL) and assigned the first peak to the dimer and the second to the monomer. The dimer peak was dominant for the non-mutated protein, while constructs

KS\*(C204G)-AT and KS\*(C204G)-AT\*(S653G) showed higher portions of monomeric species. We speculate that the mutation on the KS is disturbing the dimeric interface. The ACP(0) elutes at 98 mL, which corresponds to an apparent monomeric protein. ACP(0) showed another peak at an elution volume of 49 mL, indicating tendency for aggregation. Aggregation could be significantly reduced when the protein is less concentrated during purification (data not shown). However, we were able to purify all proteins in sufficient amount and mass spectroscopic analysis of ACP(0), co-expressed with *Sfp*, showed complete phosphopantetheinylation (figure 42 D).



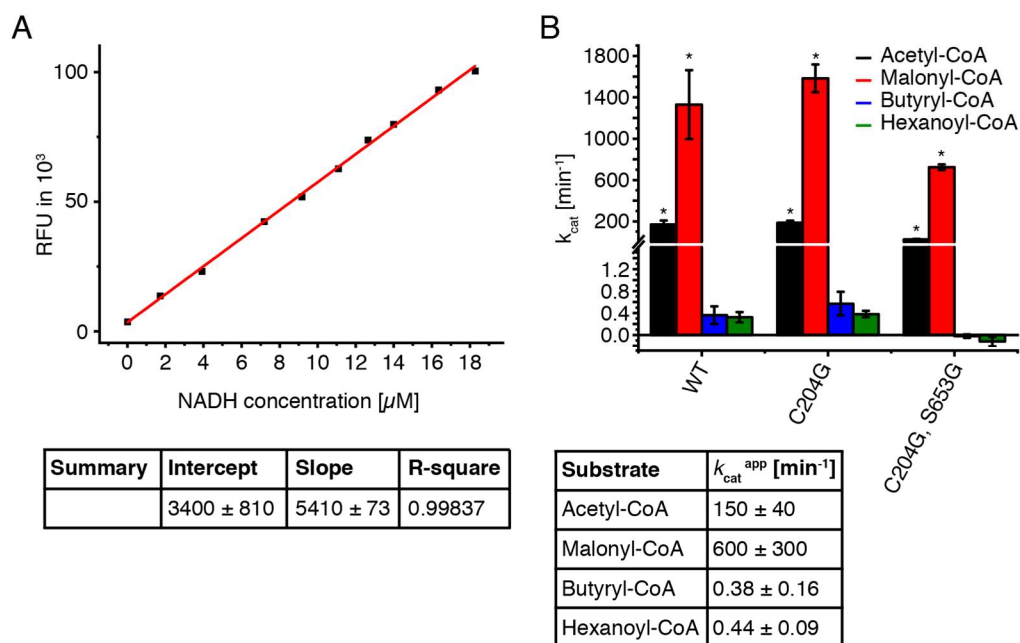
**Figure 42** | Purification of KS-AT, KS\*(C204G)-AT, KS\*(C204G)-AT\*(S653G) and ACP(0). (A) Domain structure of the KS-AT didomain and Coomassie-stained SDS-PAGE with samples taken during purification by IMAC. (B) Coomassie-stained SDS-PAGE with samples taken during ACP(0) purification by IMAC. (C) SEC of constructs 13-16 on a HiLoad 16/600 Superdex 200 column (in 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 10% glycerol and 1 mM EDTA). (D) Mass spectroscopic analysis of ACP(0), co-expressed with *Sfp*, was performed with ESI in positive mode (SYNAPT G-2, waters). The theoretical mass for the holo-form and acetylated holo-form is 11,984 Da and 12,026 Da respectively.



### 2.3.4 Specificity of the AT

With the purified KS-AT didomain and the standalone ACP(0), we were able to address the specificity of the AT, using the established enzyme coupled *in vitro* assay. In this experiential set up, liberated CoA as a consequence of AT-mediated transacylation leads to the reduction of nicotinamide adenine dinucleotide (NAD<sup>+</sup>), which can be monitored fluorometrically (see figure 43 A for calibration of NADH, see figure 59 for visualization of the assay).<sup>34,165</sup> Since also the enzymatic hydrolysis of acyl-CoAs leads to free CoA, all data were corrected to NADH increase in the absence of ACP(0). To be able to compare the transacylation of butyryl-CoA and hexanoyl-CoA to the native substrates acetyl-CoA and Mal-CoA, we analysed the turnover at fixed substrate and ACP concentrations. Since the KS domain can perform a transacylation reaction (ping step) as the AT, we compared the protein WT KS-AT with the KS-knockout (C204G). We observed no significant difference in transacylation (figure 43 B), indicating that the KS mediated transacylation is negligible for PenPaMSAS.<sup>46</sup> Surprisingly, turnover rates for the KS/AT-double-knockout (C204G, S653G) were approximately 11% and 43% for the substrates acetyl-CoA and Mal-CoA compared to the KS-AT WT, but no turnover was observed for the substrates butyryl-CoA and hexanoyl-CoA (figure 43 B). Consequently we performed more controls with the enzymes WT KS-AT, KS-knockout (C204G), KS/AT-double-knockout (C204G, S653G) and without enzyme in the presence and absence of ACP(0) (figure S7). The data show that different side reactions occur during the transacylation assay, dependent on the used substrate. Spontaneous non-enzymatic hydrolysis with water was observed for all substrates but was very slow for Mal-CoA and much faster for acetyl-CoA and especially for butyryl-CoA and hexanoyl-CoA (figure S7, -ACP). Compared to the spontaneous hydrolysis, enzymatic hydrolysis is negligible for acetyl-CoA and Mal-CoA and is slow for butyryl-CoA and hexanoyl-CoA, (figure S7, comparison between WT-ACP and -ACP). Additionally, non-enzymatic ACP acylation occurs with acetyl-CoA and especially Mal-CoA but was not observed for butyryl-CoA and hexanoyl-CoA (figure S7, comparison between +ACP and -ACP). Non-enzymatic ACP acylation explains the turnover of acyl-CoAs by the KS/AT-double-knockout, while the protein itself was not transacylation-active (figure S7, comparison of KS\*-AT\* + ACP and + ACP). In conclusion, the observed turnover rates are mainly comprised by the self-acylation of ACP with acetyl- and malonyl- moieties and the spontaneous hydrolysis of acetyl-CoA, butyryl-CoA and hexanoyl-CoA. We note that turnovers rates in figure 43 B include ACP-acylation, because NADPH signals were just corrected by the enzymatic and non-enzymatic hydrolysis. Finally, we corrected data by the ACP acylation (data recorded with KS/AT-double knockout; table in figure 43 B). The resulting hydrolysis-corrected and self-acylation-corrected turnover rates indicate 4 times faster transacylation of Mal-CoA than acetyl-CoA. We speculate that these results are reflecting the higher need for Mal-CoA in the

biosynthesis of 6-MSA (one acetyl moiety compared to three malonyl moieties). Turnover rates for butyryl-CoA and hexanoyl-CoA were found to be approximately 400 times slower compared to the native starter substrate acetyl-CoA, indicating a strict specificity of the AT (figure 43 B) and a function as gatekeeper for substrate selection. Our main goal, the production of 3-PrP and 3-PP *in vivo* requires the introduction of butyryl- and hexanoyl- moieties. Our data show that engineering of the AT is required to succeed in this endeavour. Ideally, an engineered AT has high turnover rates for butyryl- and hexanoyl- moieties, while suppressing the loading of the ubiquitously available acetyl-CoA.



**Figure 43 |** Characterization of the AT of PenPaMSAS. (A) Fluorometric NADH calibration for the enzyme coupled assay. Linear fitting was performed with OriginPro 8.5. (B) Comparison of KS-AT, KS(C204G) and KS(C204G)-AT(S653G) transacylation rates. Final concentration in the assays were 60  $\mu$ M ACP(0) and 80  $\mu$ M acyl-CoA. The enzyme concentration was adjusted to the turnover rates for the analysed substrate and was 5 nM for acetyl-CoA, 0.5 nM for Mal-CoA, and 1  $\mu$ M for butyryl-CoA and hexanoyl-CoA. Background measurements were performed in the absence of ACP. Asterisks indicate that the turnover rate was not corrected by the ACP self-acylation. The table represents the turnover rates of the WT AT, which were corrected by the enzymatic and non-enzymatic hydrolysis in the absence of ACP(0) and by the ACP self-acylation in the presence of the KS/AT-double-knockout (C204G, S653G).

### 2.3.5 Engineering of the AT

Given that the high substrate specificity of the PenPaMSAS AT limits the production of 6-MSA derivatives, we started to engineer the AT with the objective to change its specificity (or make it at least more tolerant) to longer starter substrates like butyryl-CoA and hexanoyl-CoA. The engineering should not affect Mal-CoA transfer.

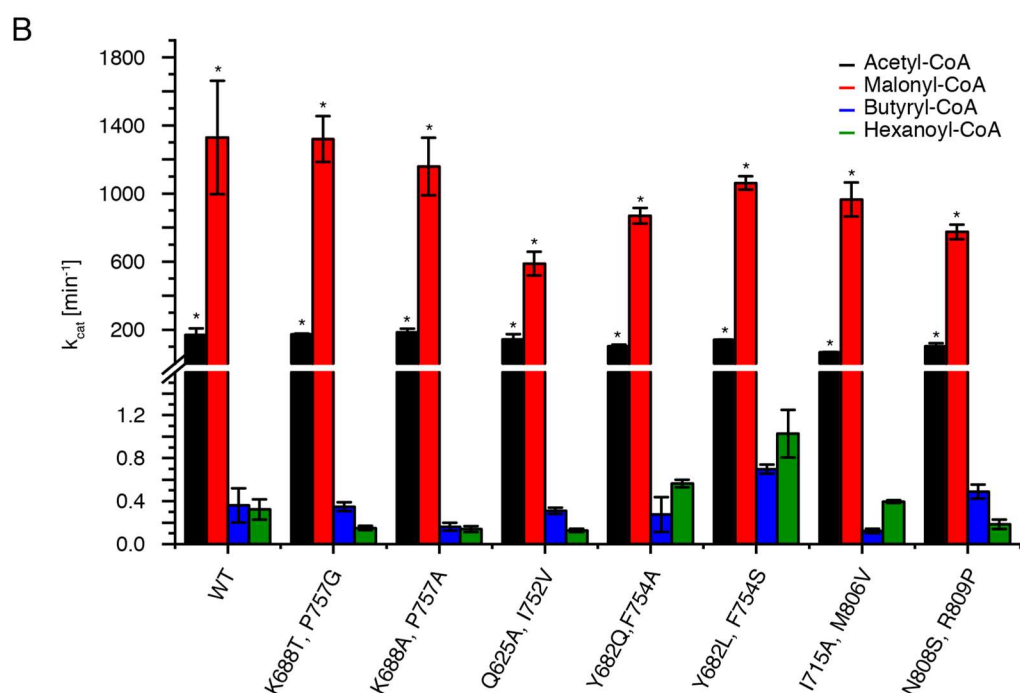
Positions for engineering were selected based on literature<sup>29,32,34,40,46,166–168</sup>, an amino acid alignment (figure S8) and an structural alignment of ATs of avermectin (PDB: 4RL1), PKS13 (PDB: 3TZW, palmitoyl bound state: PDB: 3TZY, carboxypalmitoyl bound state: PDB: 3TZZ) and mouse FAS (PDB: 5MY0, octanoyl bound state: PDB: 6ROP) and a swiss-model<sup>169–173</sup> of the KS-AT didomain of PenPaMSAS, received from UniProt (accession code: P22367, model based on the template 2QO3 on the PDB). We had to use a homology model as no authentic structural information was available for PenPaMSAS. Amino acids were selected for mutation based on our expectations to induce conformational changes, create more space for longer starter substrates and stabilize intermediates with longer acyl chains. We further submitted the PenPaMSAS homology model to the platform FunLib to receive information to the stability of mutants.<sup>174</sup> It is worth to mention that the platform is warning from the use of homology models as inputs, because computational calculations are sensitive to atomic details.<sup>174</sup> We decided to analyse double point mutations to save time and resources compared to single point mutation.

The WT KS-AT was used as cloning template (and not the KS-knockout), because of similar transacylation rates to the KS(C204G)-AT construct (figure 43 B, chapter 2.3.4). Mutants were expressed and purified as described in chapter 2.3.3. Only 7 of the 10 mutants could be purified, indicating destabilization by the amino acid exchanges. Destabilization was also visible by aggregation peaks in the SEC profiles of some double mutants (Y682Q, F754A; I715A, M806V; K688T, P757G; K688A, P757G). Nevertheless, the dimer peak was dominant for all mutants and the fractions of the dimeric peak were pooled for protein analysis (figure 44 A, figure S9). We screened all mutants in transacylation properties with a defined acyl-CoA and ACP(0) concentration of 60  $\mu$ M acyl-CoA and 80  $\mu$ M ACP(0). Characterization was performed as described in chapter 2.3.4 and thus data were not corrected for ACP self-acylation (figure 44 B). The Y682L, F754S double mutant turned out to perform best, showing higher transacylation rates for butyryl-CoA (~200%) and hexanoyl-CoA (~300%) compared to the WT protein. In contrast, turnover of the native substrates acetyl-CoA and Mal-CoA was slightly slower (~80%), indicating that the amino acids exchanges increase promiscuity, but also shifts specificities to longer acyl moieties. Additionally, the N808S, R809P double mutant showed slightly improved transfer of butyryl-CoA (~135%) in addition to decreased rates of the native substrates (~60%). Despite this success, the improvement of the AT for butyryl-CoA is still not sufficient for our needs, because in both cases the transfer remains 200 times slower compared to the native substrate acetyl-CoA. As another interesting mutant, the Q625A, I752V double mutant showed similar turnover rate for the transfer of Mal-CoA as the KS/AT-double-knockout (compare figure 44 B to figure 43 B), indicating negligible malonyl transfer. This can be explained by the position of the Gln625 (in the homology model) nearby a conserved Arg678 (figure S8, alignment number 186).

Since Arg678 is likely responsible for coordinating the carboxyl group of the malonyl, Gln625 may interfere in this interaction.<sup>34,70</sup> In conclusion, data show that engineering of the AT can change the substrate specificity, but further improvement in the transacylation of butyryl-CoA and hexanoyl-CoA vs. the transacylation of acetyl-CoA is necessary. Additionally, structural information would be very helpful for the design of new mutations.

A

Mutation	Rosetta energy difference to WT	Purification yield [mg/L]
WT	0	62
K688T, P757G	-1035	5.6
K688A, P757A	374	4.7
Q625A, I752V	849	10.7
Q625F, I752L	1076	0
Y682Q, F754A	1603	1.1
Y682L, F754S	-1872	9.4
I715A, M806V	6160	1.8
I715G, M806A	nd	0
N808S, R809P	nd	32.5
VNR807-809SST	nd	0



**Figure 44** | Characterization of the AT mutants. (A) Comparison of the Rosetta energy difference of the wildtype protein and the mutants to the protein yield after expression and purification. (B) Comparison of transacylation rates of the WT protein and mutants. Final concentration in the assays were 60  $\mu$ M ACP(0) and 80  $\mu$ M acyl-CoA. The enzyme concentration was adjusted to the turnover rates for the analysed substrate and was 5 nM for acetyl-CoA, 0.5 nM for malonyl-CoA, and 1  $\mu$ M for butyryl-CoA and hexanoyl-CoA. Background measurements were performed in the absence of ACP. Asterisks indicate that the turnover rate was not corrected by the ACP acylation.

### 2.3.6 Characterization of the mutants

Promising candidates of the AT mutant screen of chapter 2.3.5 like Y682L/F754S, N808S/R809P should be tested for their ability to synthesize 6-MSA, 6-PrSA, 6-PSA and their respective shunt products as described in chapter 2.3.2. This is necessary to exclude that other domains are also limiting the production of the target compounds. If so, another round of engineering will be required. This work is currently under progress.

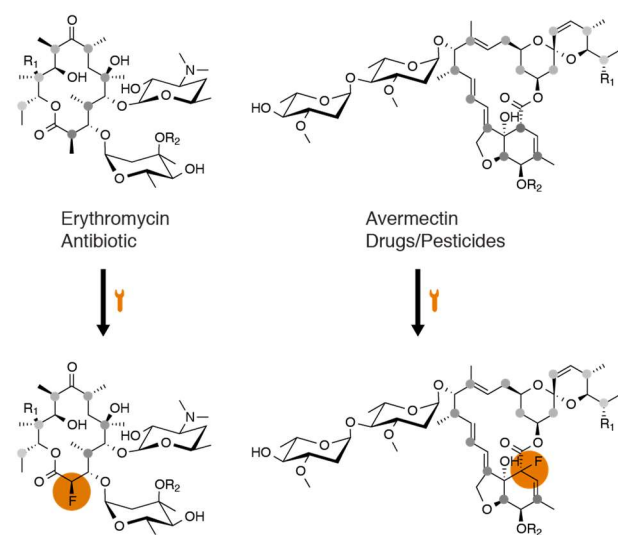
## 2.4 Directed biosynthesis of fluorinated polyketides

As described in chapter 1.1, about 32% of all drugs, that were FDA approved between 1981 to 2019, are natural products or its derivatives<sup>2</sup>. The high representation of natural compounds probably originates from their scaffolds having evolved to interact with biomacromolecules<sup>1</sup>. However, only 6% of the compounds are unaltered natural products, demonstrating the importance of derivatization the compounds to make them suitable for therapeutic treatment.<sup>1</sup>

With the motivation to unlock the dormant potential of polyketides and with the expertise in FAS and PKS engineering in the lab, we aimed at establishing a broadly applicable approach for the regioselective biosynthetic derivatization of polyketides. We focused on modular PKSs, as they are best suited for regioselective modification due to the vectorial (module-by-module) synthesis.<sup>13,17,24</sup> We chose the sixth module of DEBS with its C-terminal TE domain as our model system, because it is one of the most intensely studied PKS modules. This module also allows producing relevant DEB derivatives with appropriate starter substrates.<sup>94</sup>

Fluorine has been widely used in medicinal chemistry for lead structure optimization as its electronegativity and small size can strongly impact molecular properties such as bioavailability, metabolic stability and protein-ligand interactions.<sup>4</sup> About 25% of all small molecule based drugs contain at least one fluorine atom, demonstrating the high relevance of fluorine for drug development.<sup>175</sup> In contrast, methods for the biosynthetic derivatization of polyketides with fluorine are still rare.<sup>30</sup> The introduction of a fluorine group into a polyketide scaffold requires the selection and transfer of a fluorinated extender substrate by the AT. Unfortunately, ATs from modular PKSs are normally substrate specific and do not allow the loading of non-native (and non-canonical) substrates.<sup>20,42</sup> The loading of non-native and non-canonical substrates was so far achieved just in a few approaches. Yuzawa *et al.* exchanged the AT of DEBS module 6 (DEBS M6) as well as the AT from the lipomycin PKS with several ATs from PKS modules to produce ketones.<sup>28</sup> In another recent work, Geyer *et al.* could demonstrate that extender ATs can have relaxed specificity allowing the

incorporation of even non-canonical substrates.<sup>176</sup> Finally, Walker *et al.* showed in a pioneering work in 2013 that an alone-standing AT from a trans-AT PKS can transacylate fluoromalonyl-CoA (F-Mal-CoA) to DEBS M6 carrying a functional knock-out of the cognate AT. However, in all these examples, the incorporation of non-native and non-cognate extender substrates was possible just at low efficiency and low turnover rates. Additionally, a specialized AT for fluorinated substrates remained elusive. We hypothesized that the highly promiscuous MAT of mFAS would be capable to efficiently transfer fluorinated substrates. The enzyme was characterized in our lab and has been shown to transfer various acyl-CoAs with high transacylation rates.<sup>34</sup> Furthermore, the fact that the mammalian FAS is structurally and chemically related to PKSs convinced us to perform a domain exchange between the AT of DEBS M6 with the MAT with the aim to expand the product spectrum of DEBS towards fluorinated polyketides.<sup>53,76,177</sup> Inserting the MAT in the PKS assembly line allows the regioselective derivatization of polyketides at even numbered compound positions (general approach illustrated in figure 45).



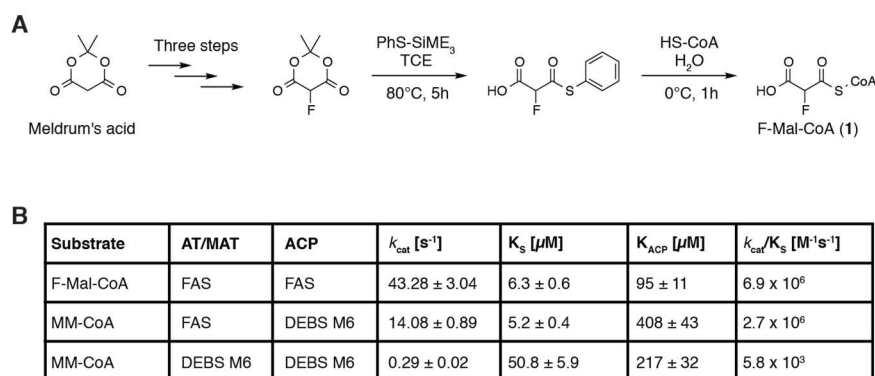
**Figure 45** | Illustration of the desired approach, exemplified on two relevant polyketides. The positions of erythromycin and avermectin, which can be derivatized by exchanging the intrinsic AT with the MAT are marked by grey dots. One position of each polyketide drug was chosen to visualise the approach for fluorination by modulating the terminal modules of the respective assembly lines.

### 2.4.1 Characterization of the MAT

Site-specific biocatalytic fluorination requires the introduction of F-Mal-CoA as extender substrate and consequently we chemically synthesized the substrate following a route by Saadi and Wennemers and Dunn *et al.* (figure 46 A, for quality control and stability analysis see figure S10-S11).<sup>178,179</sup> The chemical synthesis was performed by Alexander Rittner (Goethe-University, Frankfurt). In the next step, we characterized the MAT in its ability to transfer the F-Mal moiety to the ACP, following

an enzyme coupled assays developed by Molnos *et al.* and used for modular PKSs before by Dunn *et al* (see also chapter 2.3.4).<sup>179,180</sup> For the analysis we used the dissected didomain construct KS(C161G)-MAT and the stand-alone ACP from mFAS (see figure 48 B for purified proteins, chapter 2.4.2) and monitored the transfer of acyl-moieties.<sup>34</sup> Characterization of the MAT revealed excellent transfer kinetics of the MAT for F-Mal-CoA (figure 46 B, for hydrolysis rates see figure S11 C). In comparison to the native extender substrate Mal-CoA, the  $k_{cat}$  and  $k_{cat}/K_M$  for F-Mal-CoA were 36% and 49%, respectively (figure 46B,  $K_S$  and  $K_{ACP}$  are the  $K_M$  for the respective substrate and ACP).<sup>34</sup> Demonstrating that the MAT is able to transfer fluorinated substrates, we began to explore the interaction between the MAT of FAS and the ACP of DEBS M6, again using the enzyme coupled assay (see figure 48 B of chapter 2.4.2 for purified proteins, design of ACP6 was analogously to Kim *et al.*<sup>181</sup>). In contrast to the ACP of the mFAS, DEBS ACP6 had to be co-expressed with *Npt* instead of *Sfp* to be fully phosphopantetheinylated (see figure S12). As a reference for transacylation efficiencies, we used MM-CoA as the cognate substrate of the DEBS M6 AT. The experiments demonstrated that the MAT can interact with the ACP of DEBS M6 with a specificity constant which is 2-3 orders higher in magnitude than the native interaction of DEBS AT6, which can be explained by the high transacylation rates of the MAT.

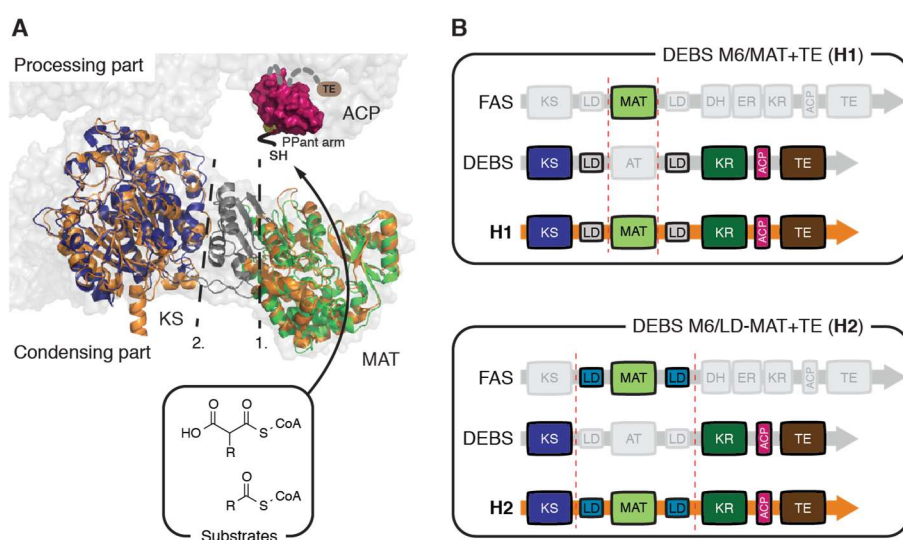
The experimental work and the analysis of the data was done by Lara Maria Mayer, supervised by Alexander Rittner (Goethe-University, Frankfurt). The data were globally fitted by using the Michealis-Menten-assumptions with equation 7. The individual fits are shown in figure S13-S15.



**Figure 46 |** Chemical synthesis of F-Mal-CoA and characterization of the MAT in its ability to transacylate F-Mal-CoA and DEBS-ACP6 as substrates. (A) F-Mal-CoA was synthesized in five steps following a route of Saadi and Wennemers and Dunn *et al.*<sup>178,179</sup> (B) Absolute kinetic parameters for MAT- and DEBS AT6-mediated transfer. The figure was adapted from Rittner and Joppe *et al.*<sup>44</sup>

## 2.4.2 Design and purification of DEBS/FAS hybrids

Showing that MAT can transacylate fluorinated substrates and collaborate with DEBS ACP6, we focused on the design of DEBS/FAS hybrids. As there were no structural information for DEBS M6 available, we used the solved structure of the related KS-AT didomain of module 5 (PDB: 2HG4). We superimposed the structure with the KS-MAT of mFAS (PDB: 5MY0) and could identify two possible junction sites, one at the interface of the KS and the LD and another one at the interface of the LD and the (M)AT (figure 47 A). Therefore, we decided to generate two different DEBS/FAS hybrids, either exchanging the AT without or with its adjacent LD, yielding hybrid 1 (**H1**, DEBS M6 MAT hybrid) and hybrid 2 (**H2**, DEBS M6 LD-MAT hybrid), respectively (figure 47 B, further information on figure S16). Different to Yuzawa *et al.*, we attributed the terminal long  $\alpha$ -helix to the AT and not to the LD, based on the observation that the  $\alpha$ -helix is also present in alone-standing type II AT.<sup>28</sup>

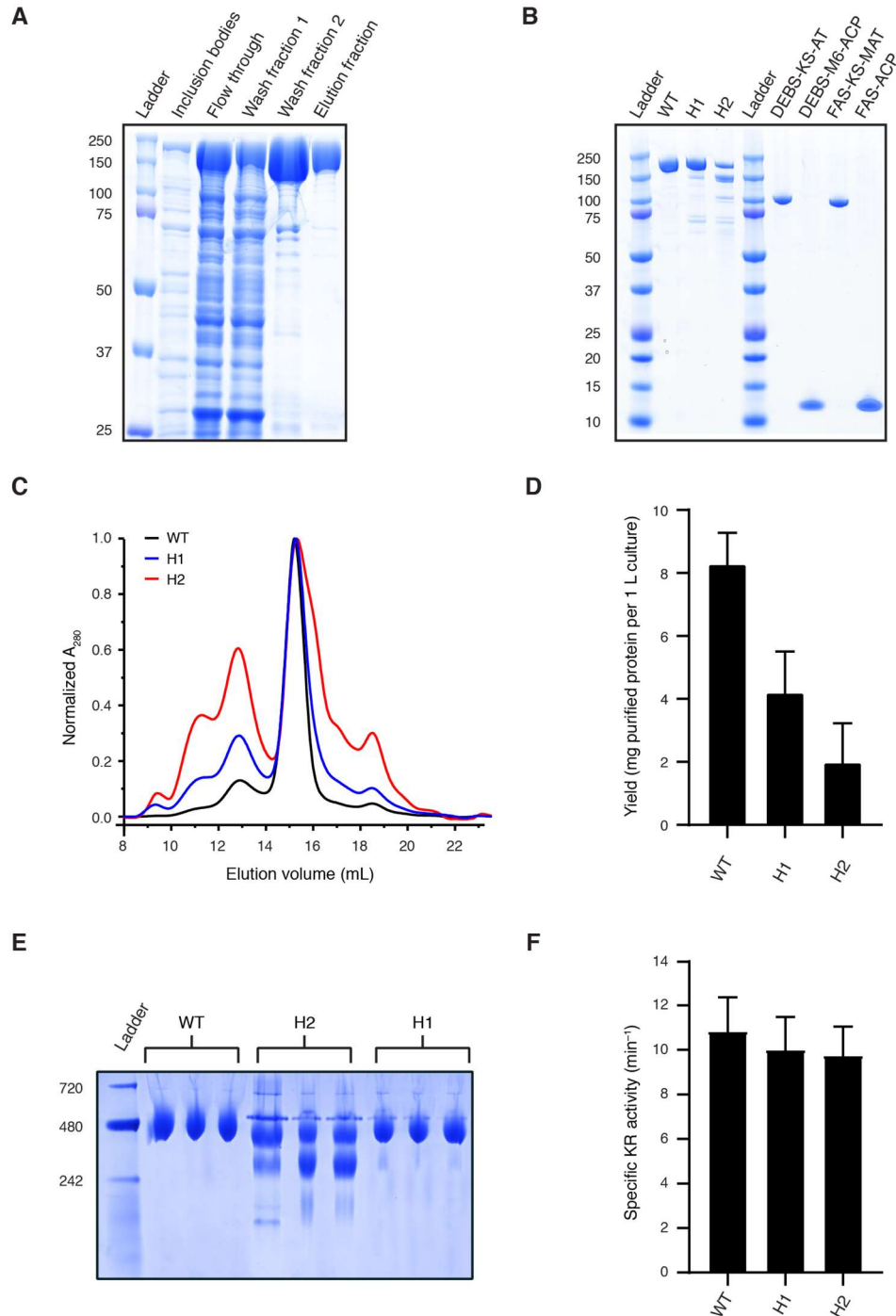


**Figure 47** | Design of the DEBS/FAS hybrids (further information on figure S16). (A) Acyl transfer of the mFAS MAT and its embedment to the KS-MAT didomain. The structure of the KS-MAT (blue = KS; dark grey = LD; green = MAT) is shown as cartoon and the ACP in surface depiction (purple, PDB: 2PNG) and both were embedded into the surface structure of the porcine FAS (light grey; PDB: 2VZ9).<sup>34,53,65</sup> KS and AT of the DEBS module 5 KS-AT didomain were superimposed individually with the murine structure, demonstrating structural conservation of these domains.<sup>74</sup> (B) The DEBS/FAS hybrids were created by exchanging AT6 (**H1**) or LD-AT6 (**H2**) with MAT or LD-MAT, respectively. The figure was adapted from Rittner and Joppe *et al.*<sup>44</sup>

We expressed the DEBS M6 **WT** and both hybrids in *E. coli* and purified them using Ni-NTA resin (figure 48 A) followed by SEC (figure 48 B-C). The hybrids were received in similar amounts as the **WT**, but the overall yields of the dimeric species decreased to 50% for **H1** and 25% for **H2**, because of different oligomerization ratios (figure 48 C-D). Furthermore, blue native PAGE of the purified dimeric species revealed that **WT** protein and **H1** were stable as dimers, whereas **H2** was partially degraded or disassembled (figure 48 E). To monitor the fitness of the hybrid constructs, we recorded KR activity in a NADPH consumption assay with the



substrate *trans*-1-decalone. The specific activity of **H1** (92%) and **H2** (90%) were similar to the **WT** protein ( $10.8 \pm 1.5 \text{ min}^{-1}$ , figure 48 F). We assumed that the KR and other domains are correctly folded in both hybrids, but **H2** suffers from destabilization of the dimeric architecture, as suggested by SEC and native PAGE.



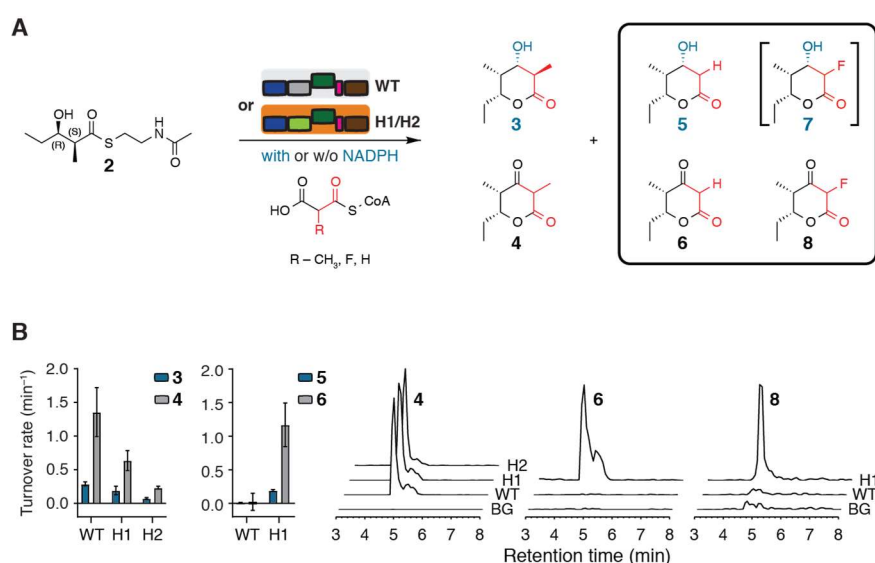
**Figure 48** | Purification and characterization of DEBS M6 (**WT**) and the DEBS/FAS hybrids **H1** and **H2**. (A) Typical Coomassie-stained SDS-PAGE with samples taken during the purification of DEBS M6 (**WT**) by IMAC. DEBS M6 was co-expressed with *Sfp* and received as pure protein. The

pellet was analysed for inclusion bodies, demonstrating that WT protein has little tendency to form aggregates. (B) The purity of all enzymes for this project was evaluated with Coomassie-stained SDS-PAGE (NuPAGE 4-12 % Bis-Tris, Thermo Fisher). (C) SEC of **WT** protein and hybrids revealed higher tendency for **H1** and especially **H2** to form higher oligomeric species. However, the dimeric species was the main peak for all proteins. A shoulder at higher elution volumes at the dimer peak of **H2** indicates the presence of the monomer. The absorbance was normalized to the highest peak. (D) Comparison of the yields of **WT** protein and hybrids after both purification steps. The average of three expression cultures ( $n = 3$ ), co-expressed with *Sfp*, is illustrated as bar diagram with standard deviations give as errors. (E) Native PAGE (3-12 % Bis-Tris, Thermo Fisher) of the dimeric species demonstrates that **WT** protein and **H1** are stable as a dimer, whereas **H2** exists as a mixture of dimer and monomer. (F) KR specific assay using *trans*-1-decalone demonstrates that the enzymatic activity of **H1** and **H2** was 92% and 90% in comparison to the **WT** protein ( $10.8 \pm 1.5 \text{ min}^{-1}$ ). The assay was performed in biological and technical triplicates ( $n = 3$ ) and the average values are illustrated as bar diagram with standard deviations give as errors. Final substrate concentrations in the assay were  $0.3 \mu\text{M}$  enzyme,  $2 \text{ mM}$  *trans*-1-decalone and  $60 \mu\text{M}$  NADPH. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>

### 2.4.3 Characterization of DEBS/FAS hybrids

With the purified **WT**, **H1** and **H2** and the understanding that the MAT can transfer acyl-moieties to the ACP of DEBS M6, we evaluated the overall enzymatic activities. After chemical synthesis of a suitable N-acetylcysteamine starter substrate for DEBS, we investigated the production of  $\delta$ -lactones (triketide lactones, TKLs) with MM-CoA as the native extender substrate of DEBS (figure 49).<sup>182</sup> We monitored the NADPH consumption fluorometrically during the production of compound **3** to determine turnover rates for all enzymes. The experiment showed that **H1** had slightly reduced turnover rates (67% of **WT**) in comparison to the **WT** ( $0.29 \pm 0.03 \text{ min}^{-1}$ ), whereas **H2** was significantly slower (25% of **WT**). Analysis of the products by HPLC-MS revealed that a mixture of reduced and non-reduced TKLs was produced under the assay conditions as previously reported (compound **3** and **4**, figure S18 A, figure S17 for compound list).<sup>28</sup> Therefore, we performed a second assay in the absence of NADPH and established a method to determine the turnover rates for the production of compound **4**, by monitoring the concentrations of MM-CoA and CoA by HPLC-UV (figure 49 B, figure S18 B). This was done by Elia Heid during his master thesis, supervised by me.<sup>183</sup> As expected, turnover rates for compound **4** were higher compared to compound **3**, as we avoided the synthesis of a product mixture, which also explains why the ratios between the hybrids to the **WT** were slightly lower. **H1** showed 47% and **H2** 17% of the **WT** enzymatic activity ( $1.35 \pm 0.36 \text{ min}^{-1}$ ). Additionally, the higher turnover rates for compound **4** revealed that the KR was rate limiting for the synthesis of compound **3**. We decided to continue the project with **H1** only, because of its significant higher turnover rates compared to **H2**, its better expression yields and dimeric stability. We tested whether **WT** and **H1** were able to produce C2-derivatives of the TKLs **3** and **4**, by using Mal-CoA and especially F-Mal-CoA as extender substrates. Investigation of the turnover rates and product analysis showed that **H1** is capable to produce the reduced (**5**) and unreduced (**6**)

demethylated TKLs, whereas the **WT** cannot synthesize any derivative. This is the final proof for the ability of MAT to preserve its properties as polyspecific enzyme inserted into a PKS module fold, which also demonstrate the unique success of our approach. Specifically, turnover rates with Mal-CoA to receive the reduced TKL **5** were identical to the rates with MM-CoA leading to **3**, while the unreduced TKL **6** was synthesized much faster than **4** (184%). Construct **H1** was able to produce the unreduced fluorinated TKL (**8**) in presence of F-Mal-CoA as extender substrate, as shown by HPLC-MS analysis. Unfortunately, fast enzymatic hydrolysis of F-Mal-CoA hampered the HPLC-UV based determination of the turnover rates (for further information see figure S19). Furthermore, **H1** was not able to produce the reduced fluorinated TKL (**7**), demonstrating a strict substrate specificity of the KR as reported before.<sup>30</sup>



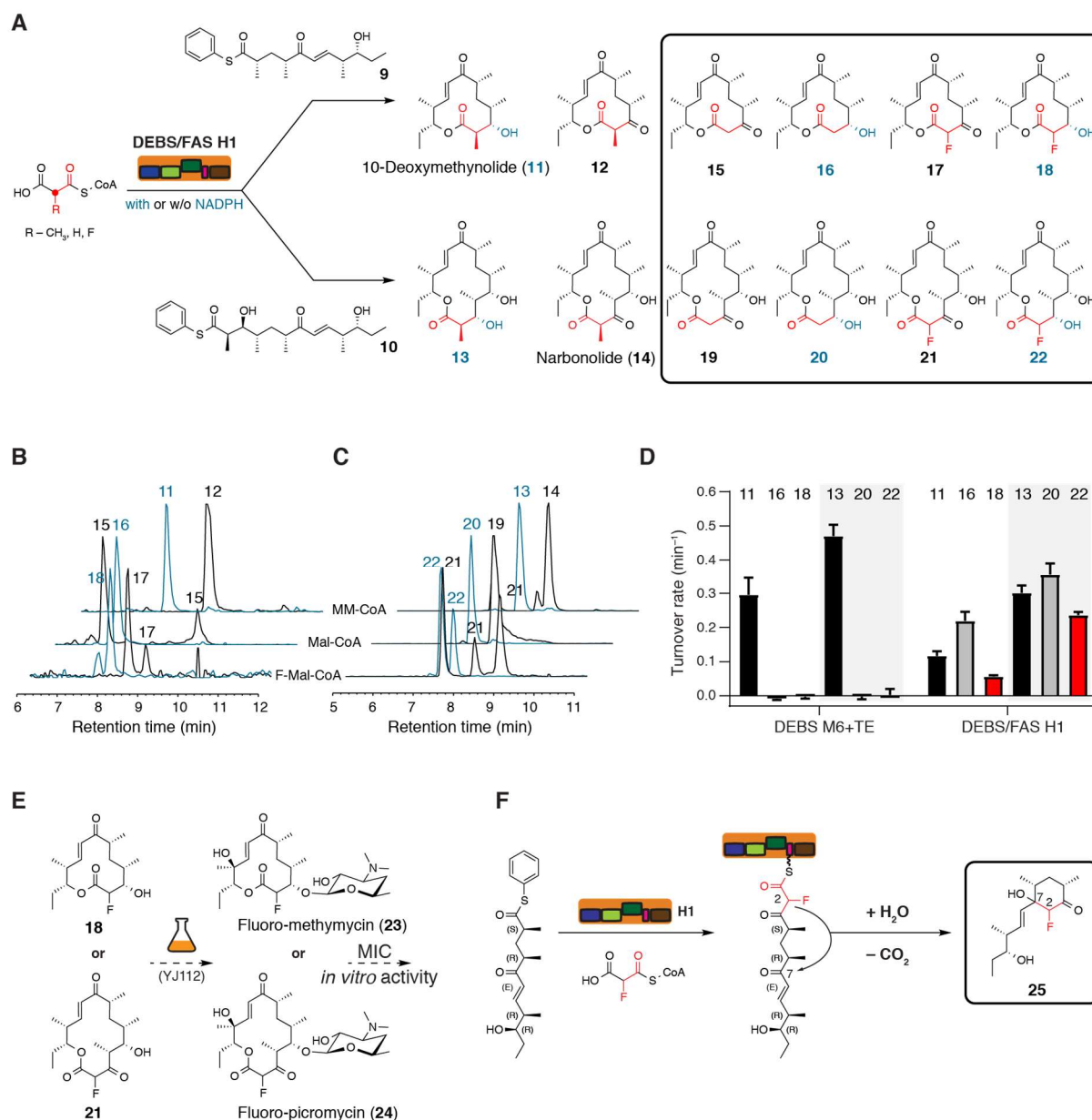
**Figure 49** | Characterization of DEBS **WT** and DEBS/FAS hybrids **H1** and **H2** in its ability to produce TKLs. (A) Reaction scheme of the enzyme mediated synthesis of TKLs from natural diketide (**2**) with the elongation substrates MM-CoA, Mal-CoA and F-Mal-CoA in the presence and absence of NADPH. Only DEBS/FAS **H1** was able to synthesise the demethylated and fluorinated TKLs, highlighted by a black box. Compound **7** was not synthesized by any enzyme, indicated by compound **7** put in brackets. (B) Turnover rates for the synthesis of compound **3** and **4**, monitored by NADPH (for compound **3** and **5**, n = 3) or CoA-ester (for compound **4** and **6**, n = 2) consumption. Final concentrations in the assays were 5 mM of compound **2**, 200  $\mu$ M elongation substrate, 0 or 60  $\mu$ M NADPH and 4  $\mu$ M enzyme. For turnover rates, see figure S17. NADPH consumption for compounds **3** and **5** were measured in three biological and technical triplicates and average values are shown with standard deviation given as errors. Turnover rates for compounds **4** and **6** were measured in two biological replicates and includes CoA-ester consumption as well as the liberation of free CoA. Average values are shown with standard deviation given as errors Product formation was confirmed by HPLC-MS for all non-reduced TKLs [EIC for **4**: [M-H]<sup>-</sup> m/z = 169.12, EIC for **6**: [M-H]<sup>-</sup> m/z = 155.16, EIC for **8**: [M-H]<sup>-</sup> m/z = 173.11] and for the reduced TKLs **3** and **5** (data not shown). Product peaks for compound **4** (MM-CoA with various constructs) were normalized to the highest peak. The figure was

adapted from Rittner and Joppe *et al.*<sup>44</sup> Figure includes data received from the master thesis of Elia Heid.<sup>183</sup>

Finally, we aimed at employing **H1** to produce DEB macrolactones, derivatized at position C2. Unfortunately, the appropriate starter substrate (hexaketide substrate mimic, see figure 13 of chapter 1.4.2) was not available, but we thankfully received pentaketide (**9**) and hexaketide (**10**) substrate mimics from the closely related pikromycin pathway from the lab of David Sherman (University of Michigan, Ann Arbor). A single elongation of these substrates can convert compounds **9** and **10** to 10-deoxymethynolide (**11**) and narbonolide (**14**), respectively, the direct precursors for the antibiotics YC-17/methymycin/neomethymycin and pikromycin/narbomycin.<sup>184</sup> A recent study by Koch *et al.* showed that DEBS M6 converts both substrates **9** (~6.6% conversion to macrolactone) and **10** (~34%). The authors also show that the DEBS M6 TE is a bottleneck in the macrolactonization of the PIKS-related compounds.<sup>185</sup> Since we aimed at validating our approach as a proof-of-concept in the synthesis of macrolactones **11** and **14**, derivatized at the position C2, we did not optimize **H1** in the TE domain. Finally, we confirmed that DEBS M6 is capable to convert both substrates (**9** and **10**), by using **WT** and **H1** with MM-CoA, as the native elongation substrate of DEBS and PIKS. In contrast to PIKS, DEBS M6 contains a KR and consequently we were able to produce both the unreduced and the reduce derivatives (figure 50). All products (**11-14**) were confirmed by HPLC-MS analysis (figure 50 B-C, figure S17). Additionally, turnover rates of all reduced products were determined by NADPH consumption, demonstrating that **H1** is slightly slower (64%) for the cognate MM-CoA than the **WT** (figure 50 D). Next, we used Mal-CoA and F-Mal-CoA as extender substrates. Turnover rates and HPLC-MS analysis showed that the DEBS/FAS hybrid **H1** was able to synthesize the demethylated and the fluorinated analogs **15-22**, while the **WT** was not able to generate any derivatives (figure 50 A-D, figure S20) similar to the TKL results. The turnover rates for the reduced macrolactones showed that the **H1**-mediated synthesis is faster with Mal-CoA (**16** and **20**) and only slightly slower for F-Mal-CoA (**18** and **22**) as compared to MM-CoA (**11** and **13**). It is worth emphasizing, that the turnover rate for the full length fluorinated macrolactone **22** by **H1** was 50% of the **WT** turnover rate, utilizing the native MM-CoA substrate for production of **13**. The high efficiency of **H1** in producing the fluorinated product again demonstrates the success of the approach.

The presented approach can give access to a library of macrolactones and those compounds are relevant for medicinal treatment when further transformed to the macrolides. This is possible by in-cell biotransformation with an engineered strain, as previously reported (figure 50 E, exemplified on fluorinated macrolactones), which has, however, not yet been performed with fluorinated species.<sup>184</sup> To prepare the macrolactones for biotransformations, we scaled-up the **H1**-mediated synthesis with Mal-CoA for compound **16** and received about 0.5 mg pure product (~59% conversion of **9**, for further information see 5.2.5). In contrast, the scaled-up reactions for the fluorinated compounds **17**, **18** and **21** led to very low amounts of purified

products and to the production of side products (for further information see 5.2.6 and 5.2.7). We note that we were able to identify the cyclohexanone product **25** as the main product in the reaction to the putative compound **17** (figure 50 F). The formation of the cyclohexanone has been reported before, as formed by either aldol condensation to the six-membered ring, followed by hydrolysis and decarboxylation, or by hydrolysis with subsequent decarboxylation and nucleophilic attack.<sup>48,186</sup> In either way, the product **25** indicates a problem of the TE to perform the macrolactone formation, like previously reported, and successful semi-synthesis of compounds (like compound **17**) require TE engineering.<sup>185</sup> Further, we assumed that the stability of the fluorinated compounds is low, because of the inductive effect of the fluorine on the  $\alpha$ -carbon and an expected high acidity of the  $\alpha$ -proton.

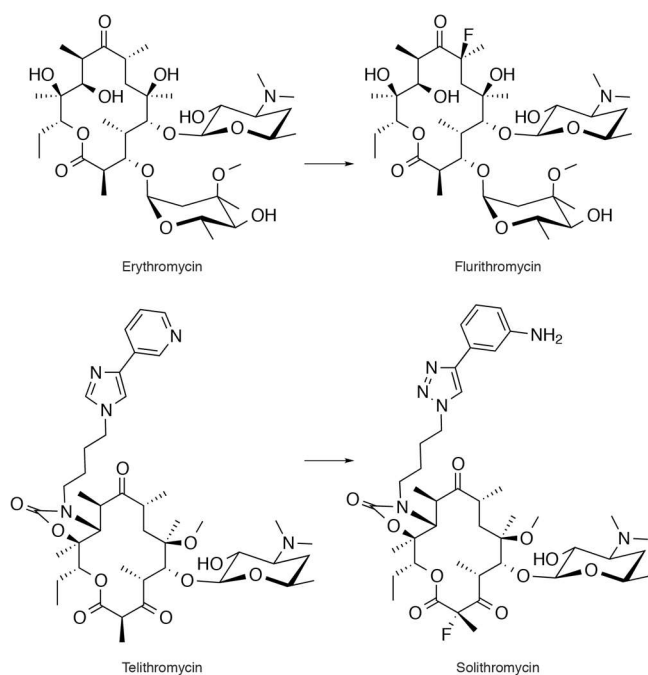


**Figure 50** | Characterization of DEBS **WT** and DEBS/FAS hybrids **H1** in its ability to produce macrolactones. (A) Reaction scheme of the enzyme-mediated synthesis of macrolactones from starter substrates (**9** and **10**) with the elongation substrates MM-CoA, Mal-CoA and F-Mal-CoA in the presence and absence of NADPH. For simplicity, only the **H1**-mediated reaction is shown, as the **WT** was not able to synthesise the demethylated and fluorinated derivatives, highlighted by a black box. (B) Conversion of pentaketide starter substrate **9** with MM-CoA, Mal-CoA and F-Mal-CoA to the macrolactones **11,12** and **15-18** by **H1** in the presence (blue) and absence (black) of NADPH, verified with HPLC-MS [EIC for **11**:  $[M+Na]^+$   $m/z = 319.11$ , EIC for **12**:  $[M+Na]^+$   $m/z = 317.09$ , EIC for **15**:  $[M+Na]^+$   $m/z = 303.08$ , EIC for **16**:  $[M+Na]^+$   $m/z = 305.09$ , EIC for **17**:  $[M+Na]^+$   $m/z = 321.07$ , EIC for **18**:  $[M+Na]^+$   $m/z = 323.08$ ]. Data were normalized to the highest peak. (C) Conversion of hexaketide starter substrate **10** with MM-CoA, Mal-CoA and F-Mal-CoA to the macrolactones **13,14** and **19-22** by **H1** in the presence (blue) and absence (black) of NADPH, verified with HPLC-MS [EIC for **13**:  $[M+Na]^+$   $m/z = 377.16$ , EIC for **14**:  $[M+Na]^+$   $m/z = 375.15$ , EIC for **19**:  $[M+Na]^+$   $m/z = 361.13$ , EIC for **20**:  $[M+Na]^+$   $m/z = 363.14$ , EIC for **21**:  $[M+Na]^+$   $m/z = 379.12$ , EIC for **22**:  $[M+Na]^+$   $m/z = 381.13$ ]. Data were normalized to the highest peak. (D) Turnover rates for the synthesis of all reduced compounds, monitored by NADPH consumption. All assays were performed in biological and technical replicates

(n = 3) and error bars indicate standard deviations. Final concentrations in the assay were 1 mM **9** or **10**, 200  $\mu$ M extender substrate, 60  $\mu$ M NADPH, 4  $\mu$ M enzyme (**WT** or **H1**). (E) Concept of biotransformation of macrolactones to the bioactive compounds, outlined on the fluorinated compounds **18** and **21** (as a project outlook).<sup>184</sup> (F) Identification of compound **25** as side product of the reaction to the putative compound **17**. The figure was created together with Alexander Rittner (Goethe-University, Frankfurt). Parts of the figure were adapted from Rittner and Joppe *et al.*<sup>44</sup> Figure included data received from the master thesis of Elia Heid.<sup>183</sup>

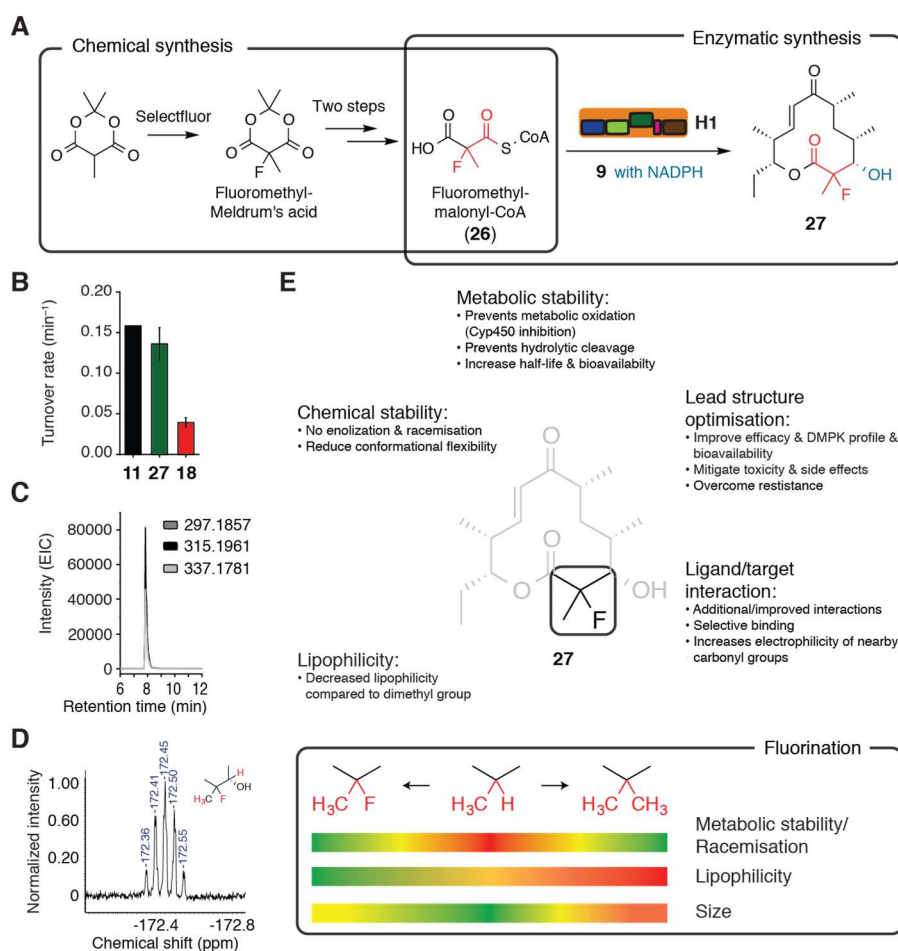
While our results demonstrate that biosynthesis of fluorinated polyketides is in principle possible with our approach, we realized that other fluorinated malonyl-CoAs, leading to more stable compounds, should be used as extender substrates. The development of the drug candidates fluorithromycin and solithromycin (figure 51) indicates that a disubstitution with a fluoro and a methyl-group (MeFC group) at an  $\alpha$ -carbon has high potential for derivatization. It was reported that the introduction of this group leads to improved binding to the target and enhanced the chemical stability as it prevents keto-enol tautomerism.<sup>187,188</sup> The enzyme-mediated introduction of a MeFC-group requires the incorporation of 2-fluoro-2-methylmalonyl-CoA (F-MM-CoA) as extender substrate. To date, there is a single study about the KS-catalyzed Claisen condensation of a disubstituted substrate (dimethylmalonyl-CoA) reported in literature, which motivated us to probe the chain elongation of a full module (DEBS M6) with F-MM-CoA.<sup>189</sup> The chemical synthesis of F-MM-CoA was performed by Alexander Rittner (Goethe-University, Frankfurt), similar to the established route for F-Mal-CoA (figure 52 A). We used F-MM-CoA for the **H1**-mediated conversion of the pentaketide starter substrate and received the desired macrolactone **27**, demonstrated by NADPH consumption and HPLC-MS (figure 52 B-C). The turnover rate for the production was  $0.14 \pm 0.02 \text{ min}^{-1}$ , which was much faster compared to F-Mal-CoA as elongation substrate and only slightly slower compared to MM-CoA (figure 52 B). We scaled-up the reaction and optimized conditions, demonstrating that lower pH values (6.8 compared to 7.1) improve the reaction slightly from 5% to 6%, while an increased temperature from about 23-25°C to 37°C led to a significant drop of the yield to 2.4% (pH 7.1). We combined all reactions with a total amount of 16.7 mg pentaketide (**9**) used and received about 1.1 mg (~6%) of the desired compound **27** after purification, shown by NMR analysis (figure 52 D, for further information see 5.2.8). Although the reaction conditions were different, the yield of **27** is similar to the reported yield by DEBS model 6, utilizing the native substrate MM-CoA, which demonstrates that the MAT swap does not impact the overall macrolactone yield.<sup>185</sup> Additionally, we found a side product **28**, most likely originating from the elongation product after KR-reduction and subsequent acetal or hemiacetal formation (chemical formula  $\text{C}_{20}\text{H}_{33}\text{FO}_6$ , estimated on HPLC-HRMS experiments, figure S21 A). Although the yield of the side product **28** was about 50%, estimated by  $^1\text{H-NMR}$ , the compound could not be isolated in sufficient amounts and quality for 2D-NMR analysis (for further information see 6.2.8). Similar to experiments with F-Mal-CoA as extender substrate, the results

from the condensation with F-MM-CoA indicate that the TE cannot cyclize the elongation product of the pentaketide substrate (**9**) properly. We note that the **H1**-mediated elongation of **9** with F-MM-CoA in the absence of NADPH led to the cyclohexanone product **29**, demonstrated by HPLC-HRMS, which points at limitation of the TE (figure S21 B). However, the successful semi-synthesis and purification of compound **27** demonstrates the capability of the hybrid in producing stable fluorinated polyketides (figure 52 E). The successful condensation with F-MM-CoA has also mechanistic implications. The reaction is indicative of a KS-catalyzed chain elongation that proceed via an initial decarboxylation to the enolate with subsequent nucleophilic C-C bond formation. Alternatively, decarboxylation and C-C bond formation may also proceed in concerted manner. In the light of the F-MM-CoA being condensed, the KS-mediated condensation reaction cannot follow a two-step mechanism, where the C-C bond formation precedes the decarboxylation (see figure 5, chapter 1.2.2)



**Figure 51** | The synthetic compounds flurithromycin, telethromycin and solithromycin are erythromycin analogues with improved pharmacological properties. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



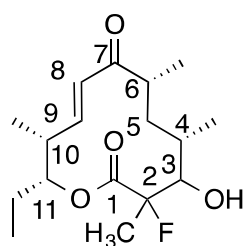


**Figure 52** | Derivatization of polyketides with a MeFC-group, exemplified on 10-deoxymethynolide. (A) Chemical synthesis of F-MM-CoA and its conversion to a fluorinated 10-deoxymethynolide derivative (**27**) by the DEBS/FAS hybrid **H1** with the pentaketide starter substrate **9** in the presence of NADPH. (B) Turnover rates for the **H1**-mediated synthesis of the reduced macrolactones **11**, **27** and **18** using pentaketide starter substrate **9** with the elongation substrates MM-CoA (black bar), F-MM-CoA (green bar) or F-Mal-CoA (red bar). (C) HPLC-HRMS analysis confirmed that compound **27** was synthesized by **H1**. [EIC for **27**:  $[M+H]^+$   $m/z = 315.1961$ ,  $[M+Na]^+$   $m/z = 337.1781$ ,  $[M+H-H_2O]^+$   $m/z = 297.1857$ ]. (D) The purified product **27** was further verified with NMR, illustrated by <sup>19</sup>F-NMR. (E) Potential lead structure optimization of polyketides with MeFC-groups, exemplified on 10-deoxymethynolide. See the cited references for further information.<sup>3,5-7</sup> The figure was created together with Alexander Rittner (Goethe-University, Frankfurt). Parts of the figure were adapted from Rittner and Joppe *et al.*<sup>44</sup>

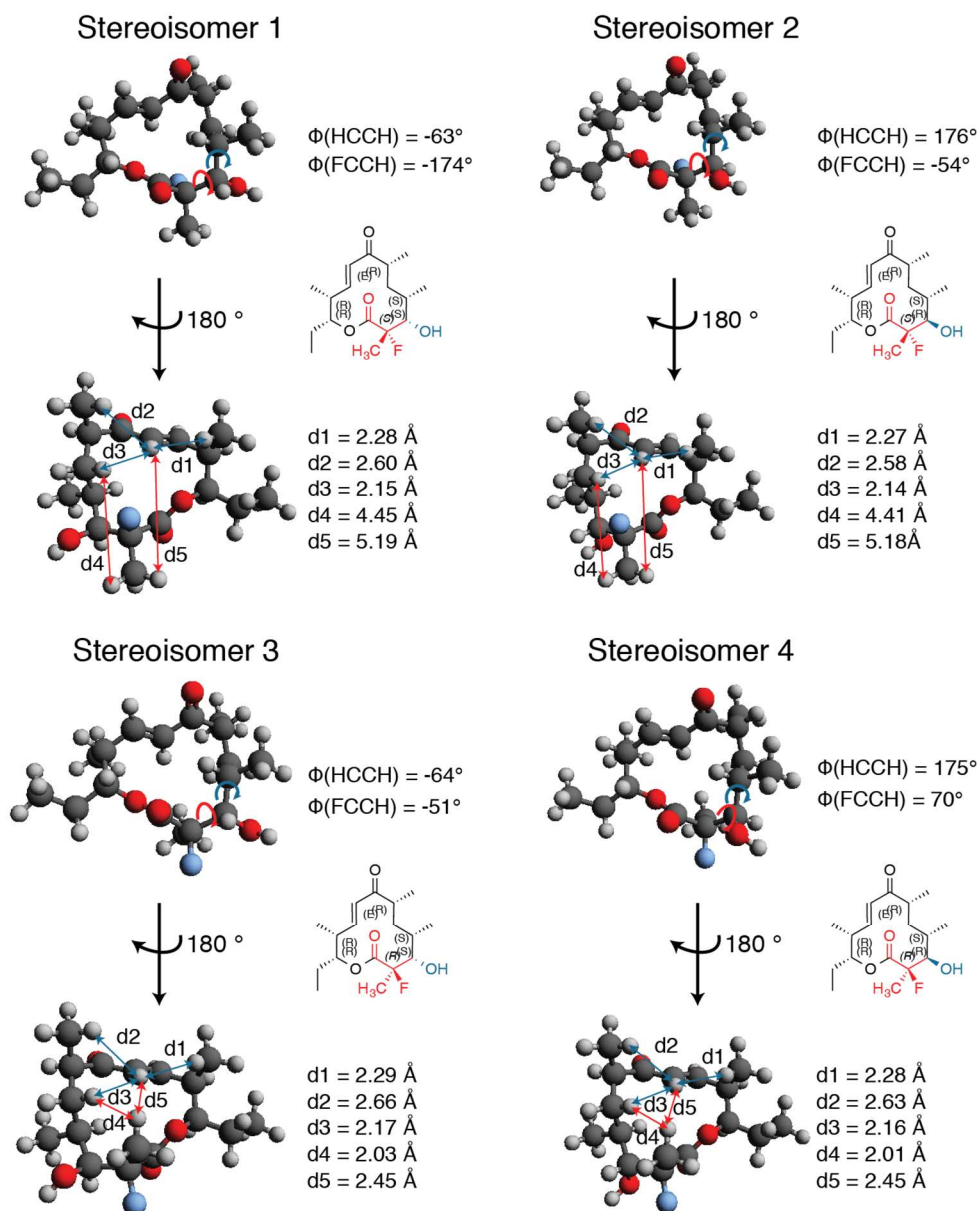
## 2.4.4 Analysing the stereochemistry of compound 27

Compound **27** was analysed in absolute configuration at position 2 and 3 (figure 53) by <sup>1</sup>H-NMR, NOESY and <sup>19</sup>F-NMR spectroscopy. To be able to compare the results to 3D-structures, we used the software Avogadro with UFF force field (Steepest decent) algorithm and no constraints to generate structural models for all

stereoisomers (see figure 54). The starting point of all models was the structure of 10-deoxymethynolide, solved by crystallization of the PIKS TE (PDB: 2hfk, E4H). Then, the models were analysed for characteristic distances (d1-d5) in the different stereoisomers. We calculated distances between the hydrogen at C8 (carbon atom at position 8) to the closest hydrogens (methyl groups at C6 and C9, methylene group at C5) and to hydrogens from the methyl group at C2 (see figure 54). We also measured distances between the hydrogens from the methyl group at C2 to the hydrogen of the methylene group at C5. Finally, we calculated dihedral angles between the hydrogen of the CHOH group at C3 to the hydrogen of the CHCH<sub>3</sub> group at C4 (HCCH angle, blue arrow) and to the fluorine at C2 (HCCF angle, red arrow).



**Figure 53** | Chemical structure of compound **27** to visualize the numbering of the carbon atoms. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure 54** | Structural models of all possible stereoisomers were created by the software Avogadro (version 1.2.0) using UFF force field (Steepest decent) algorithm and no constraints with the structure of 10-deoxymethynolide used as template (PDB: 2hfk, E4H). The chemical structures were created with ChemDraw (version 14.0) and illustrated for clarity. The models were used to calculate characteristic distances with Avogadro and shown with blue arrows (d1-d3) for observed NOE signals and red arrows for non-existent NOE signals (d4-d5). Furthermore, dihedral torsion angles for HCCH (C3-C4, blue arrow) and HCCF (C2-C3, red arrow) were also calculated with the software Avogadro. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>

To determine which stereoisomer was synthesized, we used the coupling constants of the  $^1\text{H-NMR}$  signal at 3.68 ppm, corresponding to the hydrogen at C3 (dd,  $^3J = 26.41, 0.9 \text{ Hz}$ , 1H). We used an improved Karplus equation, taking electronegativities of the substituents into account, to calculate the coupling constant

of each stereoisomer model (table 1, equation 1).<sup>190,191</sup> Values for the substituents electronegativity were received from Altona and we used the value of CH<sub>2</sub>F and CH<sub>2</sub>X (0.65) for the CCH<sub>3</sub>FC group and CH<sub>2</sub>C group, as the exact groups were not listed.<sup>190</sup> Only the calculated theoretical coupling constants for stereoisomer 1 and 3 fitted to the experimental one (<sup>3</sup>J<sub>HH</sub> = 0.9 Hz), whereas the values for stereoisomer 2 and 4 vary widely. Additionally, we also calculated the <sup>3</sup>J<sub>HF</sub> coupling constant for each stereoisomer, by using a similar Karplus equation, which was developed for CHHF torsion angles (equation 2)<sup>192</sup>. Values for the substituents electronegativity were received again from Altona and we used the value of CHCH<sub>3</sub>X for the CHCH<sub>3</sub>C group (0.6) as the exact group was not listed (table 2).<sup>190</sup> Once more, the calculated coupling constants for stereoisomers 2 and 4 did not correlate with the experimental (<sup>3</sup>J<sub>FH</sub> = 26.41 Hz). Consequently, we concluded that compound **27** has an 3S configuration. Unfortunately, we could not exclude stereoisomer 1 or 3 with that analysis. Although the calculated <sup>3</sup>J<sub>FH</sub> coupling constant for stereoisomer 1 fitted better to the experimental (0.04 Hz), the value for stereoisomer 3 differed only slightly (4.00 Hz). Therefore, the isomer cannot be excluded, especially because we used structural models for the analysis.

**Table 1** | Calculated HCCH dihedral angles ( $\phi$ ) between C3 and C4 and experimental and theoretical <sup>3</sup>J<sub>HH</sub> coupling constants using equation 1.  $\lambda_i$  is the electronegativity and  $s_i$  the sign factor of the substituent.<sup>190,191</sup> The table was published by Rittner and Joppe *et al.*<sup>44</sup>

HCCH	$\phi$	<sup>3</sup> J <sub>HH</sub> [Hz]	OH: $\lambda_1$ ; $s_1$	CCH <sub>3</sub> FC: $\lambda_2$ ; $s_2$	CH <sub>2</sub> C: $\lambda_3$ ; $s_3$	CH <sub>3</sub> : $\lambda_4$ ; $s_4$
Exp.		0.9				
Isomer 1	-63	1.28	1.33; 1	0.65; -1	0.65; 1	0.8; -1
Isomer 2	176	12.47	1.33; -1	0.65; 1	0.65; 1	0.8; -1
Isomer 3	-64	1.09	1.33; 1	0.65; -1	0.65; 1	0.8; -1
Isomer 4	175	12.40	1.33; -1	0.65; 1	0.65; 1	0.8; -1

**Table 2** | Calculated HCCF dihedral angles ( $\phi$ ) between C2 and C3 and experimental and theoretical  $^3J_{HF}$  coupling constants using equation 2.  $\lambda_i$  is the electronegativity and  $\xi_i$  the sign factor of the substituent and  $a_{FCC}$  and  $a_{HCC}$  the respective bond angles.<sup>190,192</sup> The table was published by Rittner and Joppe *et al.*<sup>44</sup>

HCCF	$\phi$	$^3J_{HF}$ [Hz]	OH: $\lambda_1$ ; $s_1$	CHCH <sub>3</sub> C: $\lambda_2$ ; $s_2$	COOR: $\lambda_3$ ; $s_3$	CH <sub>3</sub> : $\lambda_4$ ; $s_4$	$a_{FCC}$ ; $a_{HCC}$ (°)
Exp.		26.41					
Isomer 1	-174	26.37	1.33; -1	0.6; 1	0.42; 1	0.8; -1	111; 104.9
Isomer 2	-54	11.34	1.33; 1	0.6; -1	0.42; 1	0.8; -1	111; 107.5
Isomer 3	-51	22.41	1.33; -1	0.6; 1	0.42; -1	0.8; 1	105.9; 102.7
Isomer 4	70	7.99	1.33; 1	0.6; -1	0.42; -1	0.8; 1	105.9; 107.8

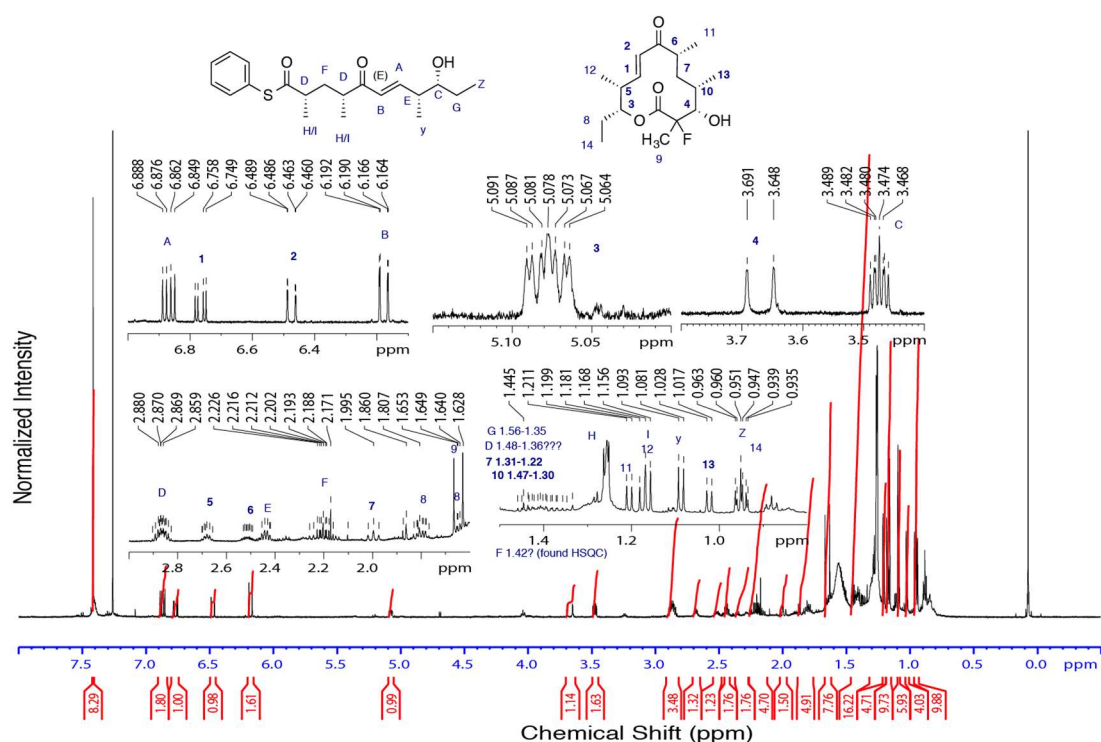
$$3J(H, H) = 14.64 \cos^2(\phi) - 0.78 \cos(\phi) + 0.58 + \sum_i \lambda_i [0.34 - 2.31 \cos^2(s_i(\phi) + 18.40 |\lambda_i|)] \quad \text{Equation 1}$$

$$3J(H, F) = 40.61 \cos^2(\phi) - 4.22 \cos(\phi) + 5.88 + \sum_i \lambda_i [-1.27 - 6.20 \cos^2(\xi_i(\phi) + 0.20 \lambda_i)] - 3.72 \left[ \frac{(a_{FCC} + a_{HCC})}{2} - 110 \right] \cos^2(\phi) \quad \text{Equation 2}$$

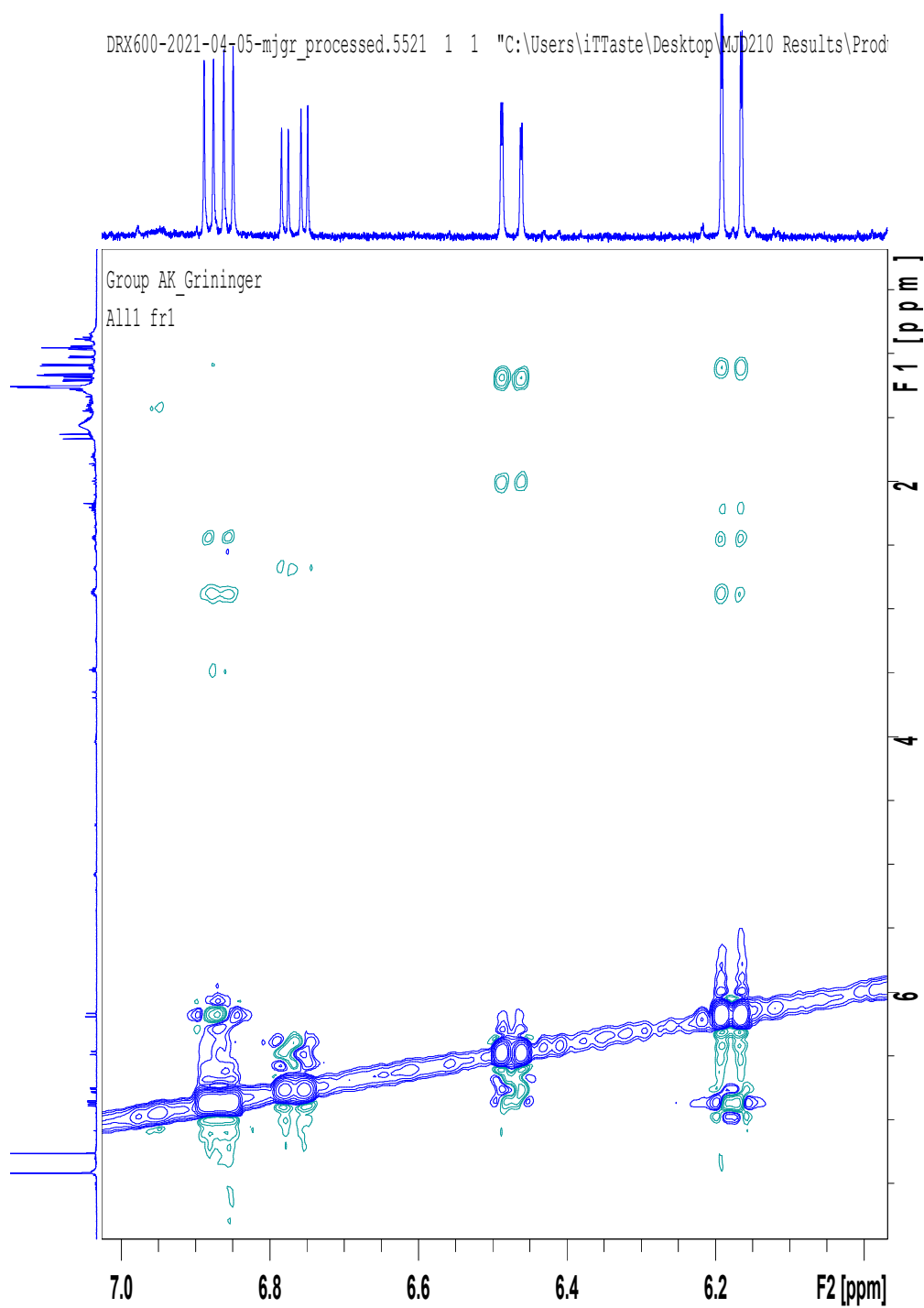
with  $\lambda_i$  = the electronegativity of the substituent,  $s_i$  and  $\xi_i$  = sign factor of the substituent<sup>190,191</sup>,  $a_{FCC}$  and  $a_{HCC}$  = bond angles.<sup>192</sup>

We performed another biosynthesis and purification of **27** to receive enough compound for a NOESY experiment. Unfortunately, we received a mixture of product and educt (see figure 55) as both compounds could not be separated by flash chromatography. Additionally, the educt was not fully consumed as the incubation time was reduced to 3 hours. However, we performed a NOESY experiment with the product/educt mixture and focused on the hydrogen of the alkene group at C8

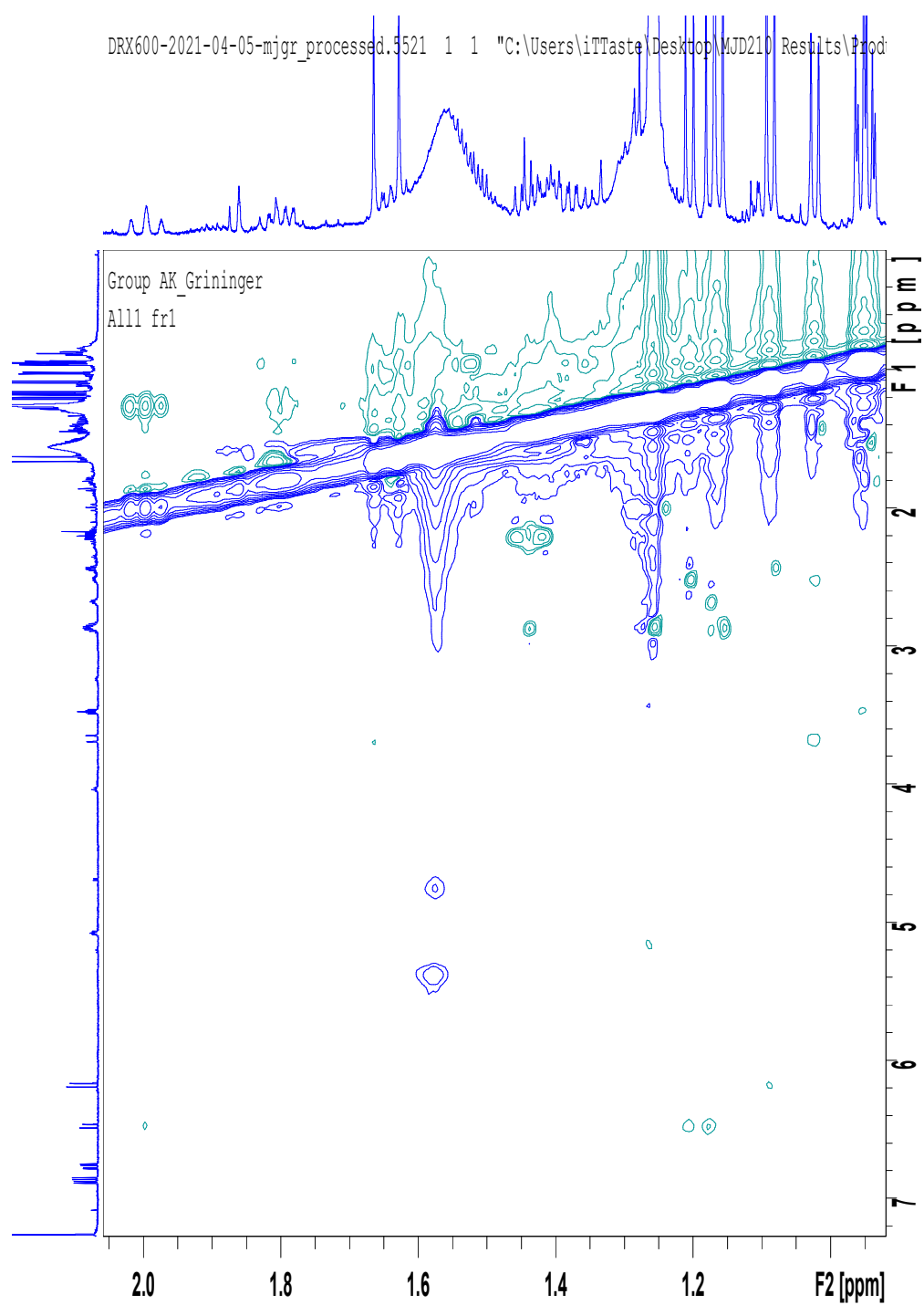
( $^1\text{H}$  NMR peak at 6.47 ppm). We found NOE signals of the alkene group at C8 to the methylgroup at C10 (1.18 ppm, distance about 2.3 Å), to the methylgroup of C6 (1.21 ppm, distance about 2.6 Å) and to the methylene group at C5 (2.00 ppm, distance about 2.2 Å, figure 56-57). In contrast, there are no NOE signals of the methylgroup at C2 ( $^1\text{H}$  NMR peak at 1.63 ppm) to the alkene group at C8 nor to the methylene group at C5. The distance to these groups is 2.45 Å and 2.03 Å in stereoisomer 3 and consequently these signals should be observed if this isomer is formed. Consequently, we could demonstrate that **H1** synthesized stereoisomer 1 with a 2S, 3S configuration. Moreover, this shows that the hybrid is capable to convert the F-MM-CoA racemate into a stereo pure compound. We thankfully received support from Julia Wirmer-Bartoschek for the analysis (Goethe-University, Frankfurt).



**Figure 55** |  $^1\text{H}$ -NMR analysis shows that we receive a product/educt mixture after purification. All peaks were assigned with the help of the COSY and HSQC experiment (data not shown). The NMR spectrum was processed with TopSpin (version 4.1.1) and the respective structures were created by ChemDraw (version 14.0). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure 56** | Selected zoom of the NOESY experiment with the purified product/educt mixture. The NMR spectrum was processed with TopSpin (version 4.1.1). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure 57** | Selected zoom of the NOESY experiment with the purified product/educt mixture. The NMR spectrum was processed with TopSpin (version 4.1.1). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



## Chapter 3: Discussion

Megasynthases are responsible for the biosynthesis of bioactive natural products or their precursors including fatty acids and polyketides, which are biotechnologically and pharmacologically relevant. Although natural products are traditionally synthesized chemically, biosynthesis presents an economical alternative to produce these complex scaffolds. Derivatization of these compounds can further improve properties to use them as platform chemicals and especially as drugs. In general, regio- and stereospecific derivatives can be obtained directly by harnessing megasynthases as catalyst, which, however, often requires challenging protein engineering. To utilize the enormous potential of megasynthases and to produce targeted compounds by e.g., microbial fermentation, there is a need to characterize the structural and functional properties of those enzymes in full depth. Detailed knowledge on megasynthases can then inform protein engineering approaches to modulate specificities and avoid invasive designs that harm protein quality and compromise enzymatic activity. The following section is focused on protein engineering and related aspects will be discussed.

### 3.1 Natural products of megasynthases are pre-cursors for bioactive compounds

The fungal FAS, as a part of the primary metabolism, produces fatty acid CoA esters. In contrast to fatty acids, which need to be activated by an acyl-CoA synthase and ATP, the CoA-ester can directly enter other biosynthetic pathways. For example, fatty acyl-CoAs can be converted into triacylglycerol for energy storage in muscles and fatty tissue, or into phospholipids and sphingolipids that are both essential building blocks for membranes.<sup>21</sup>

Like fatty acids, 6-MSA is as a precursor for bioactive compounds. We showed that PenPaMSAS can generate 6-PrSA as the propyl-derivative of 6-MSA. The biosynthesis of 3-PrP, as the main goal of the engineering study on PenPaMSAS, relies on decarboxylation of 6-PrSA, which may require the modulation of the substrate specificity of the decarboxylase PatG in further studies. We also demonstrated that SerMSAS is integrated into a gene cluster and that its product is likely transferred to another carrier protein (SerChlB2) for further modifications. This carrier protein can either be an ACP, derived from PKSs, or a PCP, involved in non-ribosomal peptide synthases (NRPS). The latter one was proposed for the MSAS from a different organism (*Streptomyces gandocaensis*) in a recent study.<sup>193</sup> Additionally, *S. erythraea* has two loci for the synthesis of siderophores, a hydroxamate and a proposed catechol-type siderophore.<sup>133</sup> Because of the structural

relationship of 6-MSA to those catechol-type siderophores, we hypothesized that SerMSAS could also be involved in its biosynthesis.

Like 6-MSA, products from modular PKSs are often further modified to be biologically active. The products of DEBS and PIKS are processed by various enzymes attaching hydroxy, methyl groups and most importantly amino sugars, relevant to improve water solubility.<sup>184,194,195</sup> In this thesis, we were able to synthesize and purify a fluorinated derivative of 10-deoxymethynolide through protein engineering. Further studies will enlarge this approach by additional fluorinated polyketides and optimize tailoring steps to receive bioactive macrolides.

In conclusion, protein engineering of megasynthases enables the generation of new-to-nature compounds. To be biologically active, these compounds often require further modifications by additional enzymes. Therefore, substrate specificities of tailoring enzymes must be considered to generate new bioactive compounds through biosynthesis.

### **3.2 Purification strategies are required for structural- and biochemical characterization**

Characterization *in vitro* benefits from a lower complexity and allows to directly investigate enzyme kinetic constants, substrate consumption and (side) product formation. However, it can be very challenging to find conditions for the expression and purification of the enzymes in sufficient yields and high and consistent quality. The purification procedures should be as rapid and little invasive to preserve enzymatic activity.

In this work, we established a purification strategy for the megasynthases yeast FAS and SerMSAS and applied the latter one on PenPaMSAS and DEBS M6. In both strategies we used affinity tags to enable fast chromatographic purification. The introduction of affinity tags requires careful designs adjusted to the protein architectures. In yeast FAS, a Strep-tag carrying gene was inserted at the accessible C-terminus of the  $\beta$ -chain. For SerMSAS, we designed two different constructs. We demonstrated that the insertion of a tag at the N-terminus significantly decreases protein quality. We concluded that this terminus is not accessible and attaching additional sequences may disturb proper folding. Finally, the insertion of a C-terminal His-tag enabled the purification of PenPaMSAS and DEBS M6 in good yields and quality.

Key to successful purification protocols is the constant monitoring of protein quality and enzymatic activity. To address this issue, we demonstrated different strategies to investigate the quality of protein samples. The purity of a sample can be easily

investigated by SDS-PAGE and its oligomeric state by SEC, SEC-MALS or blue native PAGE. Since unfolded, aggregated or denatured enzyme leads to compromised activity, enzymatic assays combined with product analysis are powerful tools to monitor and optimize protein preparations. In this thesis, we demonstrated that the establishment of purification strategies is essential for structural and functional characterization and showed two purification strategies which can potentially be applied on other megasynthases.

### 3.3 Specificities of megasynthases

Substrate specificities of megasynthases are essential features to ensure the biosynthesis of specific natural products. In this thesis, we investigated two bi-functional ATs, responsible to transfer both starter- and elongation substrates to the enzyme. Consequently, both enzymes possess a conserved arginine required to stabilize the carboxylic group of extender substrates (see figure S8, alignment number 186).<sup>46,70</sup> The MAT of the mammalian FAS has been demonstrated to be polyspecific and we were further able to demonstrate that this enzyme can also transfer fluorinated substrates. In contrast, we demonstrated that the AT of PenPaMSAS is specialized for its native starter substrate acetyl-CoA. The  $k_{cat}$  of the enzyme is about 1.6 times lower in transferring Mal-CoA, 5.5 times lower in transferring acetyl-CoA and 1700 times lower in transferring butyryl-CoA compared to the MAT at comparable conditions.<sup>34</sup> This indicates that the AT of PenPaMSAS has evolved to discriminate against non-native starter substrates. We claim that the difference in substrate specificities can be explained by key interaction between protein and substrates and the conformational variability at the binding site. Unfortunately, structural information is only available for the MAT, preventing structural comparisons of the mammalian FAS MAT and the AT of PenPaMSAS.

Although the lack of structural information makes engineering approaches less precise, we were able to alter the substrate specificity of the AT through active site mutagenesis. Additionally, we successfully synthesized 6-MSA and 10-deoxymethynolide derivatives by PenPaMSAS and the DEBS/FAS hybrid, indicating that all involved enzymes can at least tolerate non-native substrates or intermediates. Although PenPaMSAS can incorporate non-native starter substrates, turnover rates were much slower (5-8%) compared to the native one. This implies that the protein is sufficiently substrate tolerant to accept non-native substrates or intermediates, but at least one domain hampers the efficient production of derivatives. The appearance of lactones as side products, especially in the presence of non-native substrates, indicates that the KR is such a bottleneck and unable to reduce the triketide as efficiently as the native one. Similarly, also the DEBS/FAS

hybrid can synthesize derivatives by accepting non-native/non-canonical substrates. Here, we observed enzymatic turnover for various extender substrates, which indicate substrate tolerant KS, KR and TE domains. However, we also observed cyclohexanone side products, especially in the absence of NADPH, indicating that the TE domain cannot cyclize non-reduced intermediates and instead hydrolyze the bound compounds. We hypothesized that the binding pocket of the TE does not accept a keto group at the C3-position, or that the spatial arrangement of the intermediate prevents bringing the C11-OH group near the C1-atom.

In conclusion, characterization of substrate specificities is crucial to improve our understanding of megasynthases as it can explain enzymatic properties like kinetics, efficiencies, product spectra and the appearance of side products. Besides that, detailed analysis allows us to create a library with enzymes, characterized in kinetic constants for various substrates. This library can then be further expanded by protein engineering e.g., active site mutagenesis. Such a library already exists (<https://www.brenda-enzymes.org/>) and is beneficial to find appropriate enzymes to synthesize target compounds with engineered megasynthases. The MAT of the mammalian FAS is an excellent candidate for this library because it can transfer many substrates with fast kinetics.

### **3.4 Protein engineering usually decreases enzymatic activity**

In this thesis, we present engineered megasynthases that can produce novel compounds. However, our data showed that most of those systems have significantly decreased enzymatic activities. This tendency was clearly visible for all investigated systems and engineering approaches. We showed that affinity tags, active site mutagenesis, domain dissections, and domain exchanges (creating non-cognate interfaces) can decrease enzyme stabilities as well as affect protein folding and assembly. Therefore, there is a need to combine engineering approaches with structural data and biochemical characterization based on classical enzymology, to learn how engineering can affect stability and enzymatic efficiencies. Such a combined study can valuable model that guide the design of new systems with increasing complexity. Structural data alone would be very helpful regarding designed evolution approaches.

### 3.5 Engineering of ATs

In this thesis, we demonstrated that engineering of ATs is a promising strategy and enables biosynthesis of polyketide derivatives in mg scale.

We targeted PenPaMSAS for a designed evolution approach with the objective to broaden or alter the substrate specificities of the AT for longer starter substrates. Due to the lack of authentic structural data, engineering was based on a homology model and alignments to design the mutants. While we were able to improve the ability of the AT to incorporate non-native substrates by 200-300%, the enzyme remained slow, catalysing the transacylation 200 times slower compared to the native starter substrate.

Additionally, we targeted the MAT of the mammalian FAS for protein engineering, because of its previously reported properties such as polyspecificity, fast transfer kinetics, robustness, and plasticity.<sup>34</sup> Specifically, we were able to design DEBS/FAS hybrids of excellent protein quality and catalytic properties. This was possible by exchanging the MAT without the adjacent linker region, creating a chimeric interface between the MAT and LD. We further left the subsequent linker on the KS surface unaltered. In contrast, a recent study on a DEBS/PKS hybrid suggesting the exchange of the AT with its adjacent LD.<sup>28</sup> Unlike this study, we defined enzyme boundaries differently, especially for the AT. We attributed the terminal long  $\alpha$ -helix to the AT and not to the LD, because this motif is also present in type II systems. Dissection of this helix could lead to inactive or instable enzyme and explain the huge discrepancy in catalytic efficiency of the hybrids in the recent study in comparison to our data. To summarize, we recommend that the ATs should be exchanged without the LD between the KS and AT and without the linker on the KS surface.

In conclusion, we demonstrated that the MAT is a very promising candidate for engineering. We identified a truncation site, which allows domain exchange approaches. A previous study was able to show that a single point mutation is able to change the substrate specificity, which demonstrate the plasticity of the fold.<sup>23</sup> For applying MAT in *in vivo* biosynthesis, the narrowing of the substrate specificity to discriminate against substrates present in the host cells will be important. As the MAT can be used for designed evolution, it is possible to create a library of specialized MAT mutants with enhanced specificity for targeted substrates. In combination with the identified truncation site, such a library has enormous potential to enable biosynthesis of targeted compounds in engineered PKSs.

### 3.6 Guidelines for engineering

In this thesis, we presented various engineering approaches including active site mutagenesis and domain exchanges. We showed that the introduction of point mutation can alter the substrate specificities and therefore is a good strategy to enable access to new-to-nature compounds. There are two ways to create specialized enzymes for targeted substrates:<sup>196</sup>

- 1.) Changing the substrate specificities of specialized enzymes
- 2.) Narrowing the specificities of polyspecific enzymes

Specialized enzymes normally have very low turnover rates as they must discriminate against non-native substrates.<sup>196</sup> As described in chapter 3.4, mutations usually decrease the enzymatic activities even more. Therefore, it may be beneficial to use polyspecific enzymes with good properties like robustness, plasticity, and fast kinetics for mutational approaches. Those enzymes can tolerate more point mutations before being destabilized critically, similarly as initial fast kinetics may rather preserve reasonable turnover rates in specialized variants. Therefore, there is a good reason to investigate polyspecific domains and to specialize them for targeted substrates with mutational studies.

Since megasynthases are large polypeptides that carry catalytic activities in enzymatic domains, engineering of FASs and PKSs generally also affects domains in permanent or transient interaction. Particularly, the exchange, insertion, or removal of domains requires the consideration of the overall architectures of the synthase. Careful design allowed us to exchange the DEBS M6 AT with the MAT of the mammalian FAS, as discussed in 3.5:

- 3.) ATs of mammalian FASs and PKSs should be exchanged without the adjacent LD and without altering the linker on the KS surface
- 4.) The terminal long  $\alpha$ -helix was assigned to the AT and should not be altered in domain exchanges

While the MAT domain of mammalian FAS has excellent properties and is suitable for mixing with PKSs, the yeast FAS and its enzymatic domains are less suited for engineering approaches. The limitations of yeast FAS mainly arise from the rather closed barrel-like structure as well as the tight embedment of enzymatic domains, scaffolding domains, and folds. In this thesis, two scaffolding domains, TM and DM2, were evaluated as putative attachment sites for further enzymatic domains capable to process and tailor fatty acid CoA esters. We note that the positions are at the outside of the barrel, so that attached domains could not be addressed with ACP directly. The compartmentalization of such domains at the FAS barrel could however be beneficial to improve kinetic efficiencies, local concentrations, limit side reactions and regulate protein expression.<sup>49,51</sup> We showed that the dissection of TM and DM2 is

possible, although it led to decreased enzymatic activity and lower thermal stability. The position of the TM at the apical sides of the  $\beta$ -domes can theoretically be used to fuse an additional enzyme, but it must fit into a constrained area where it may disturb proper assembly or folding of the domains AT and ER. The DM2 is positioned at the outside of the barrel in between the dimeric interface of the KR. This scaffold is rather accessible and the exchange with dimer forming enzymatic domains seem possible.

- 5.) The structural domains TM and DM2 of the yeast FAS can be used to fuse additional functionalizing enzymes

In conclusion, there is a huge demand for engineering guidelines to simplify the creation of new synthases for the biosynthesis of new-to-nature compounds and natural compound derivatives. We present guidelines for domain exchanges of ATs and showed how the yeast FAS scaffold can be expanded. Our guidelines should be considered as starting points in engineering projects but need to be adapted to the megasynthase architectures and structural properties of the scaffold.

### **3.7 Semi-synthesis and biosynthesis**

In this thesis, we showed that domain exchanges with a polyspecific domain are powerful for *in vitro* semi-synthesis. Hybrids can be treated with different substrate at a time, including non-canonical ones. By doing so, a huge library of derivatives can be synthesized, which can then be used for potency tests. However, semi-synthesis requires the availability of stable starter substrates, which are often challenging to synthesize. In contrast, *in vivo* biosynthesis can exploit the full potential of megasynthases, as these enzymes usually convert simple substrates, which can be provided by the expression host and overcome the dependency of complex starter substrates. However, *in vivo* approaches require strategies like metabolic engineering to provide the non-native/non-canonical substrates by the expression host. In the case of fluorinated compounds, Chang and co-workers have recently presented promising strategies by feeding cheap fluoromalonate building blocks or by using a fluoroacetate pathway in an engineered strain.<sup>30</sup> Further, as discussed above, *in vivo* approaches often depend on specialized enzymes, which can discriminate against native substrates of the expression host.

### 3.8 Impact of our findings on enzyme mechanisms

The successful elongation of acyl chains with di-substituted substrates also helps clearing skies of the molecular mechanism of the KS-mediated condensation reaction. Three mechanisms have been discussed for the condensation reaction (see figure 5, chapter 1.2.2). Our data on the production of di-substituted polyketides implies that the KS-mediated condensation cannot be performed via a C-C bond formation preceding decarboxylation, as this would lead to a carbon with five bonds. Therefore, the KS catalyzed Claisen-condensation must either occur in a concerted manner or in a sequential two-step addition-decarboxylation mechanism. Furthermore, we showed that elongation of di-substituted substrates by DEBS/FAS hybrid can also lead to cyclohexanone side product formation similar as reported.<sup>48</sup> We could clarify the molecular mechanism for this reaction. Here, the elongated di-substituted intermediate is first hydrolyzed from the enzyme, either by the TE or spontaneously, and the side product is then formed by subsequently decarboxylation and aldol-condensation as reported for the rosamicin PKS.<sup>186</sup> Overall, our findings expand the knowledge about mechanistic aspect of the domains KS and TE, which will help engineering approaches in future.



## Chapter 4: Summary

In this thesis, we focused on selected FASs and PKSs for structural and functional characterization with the objective to use them for the biosynthesis of new bioactive natural compounds. The biosynthesis of fatty acids and polyketides is especially relevant as these enzymes form C-C bonds, introduce functional groups and generate stereoinformation in enantioselective manner. In particular, the extraordinary number of chiral centers within polyketide scaffolds makes biosynthesis an economical alternative compared to the often-challenging chemical total synthesis. It has emerged that derivatization of natural products is often required to generate natural products of industrial and especially pharmaceutical relevance.<sup>1</sup> Fatty acid and polyketide derivatives can be directly obtained by biosynthesis using megasynthases, which usually requires protein engineering to remodel substrates specificities or protein-protein and domain-domain interactions. Protein engineering often leads to loss of enzymatic activity that is largely caused by a lack of knowledge about key structural, conformational and functional properties. Consequently, to enable the engineering of catalytically efficient megasynthases, there is a need to characterize megasynthases and their enzymes in full details and to create guidelines for efficient designs. With this motivation, we worked with different iterative and modular megasynthases, more specifically the yeast and mammalian FAS, the MSAS and the DEBS for structural and functional characterization and the objective to use these proteins in engineering approaches.

### A rapid purification strategy for yeast FAS

The *in vitro* characterization of megasynthases is challenging in establishing conditions to produce the proteins and its enzymatic domains in required protein amounts and at high quality. In this thesis, we established a vector-based protocol for the yeast FAS as a convenient purification strategy for mutational studies. After expression, we were able to purify the enzyme by affinity- and SEC within 5 hours. Quality control by SEC, SDS-PAGE, TSA and NADPH consumption assays demonstrated the suitability of the applied strategy to deliver pure yeast FAS with high enzymatic activities in reproducible manner. The presented purification strategy combined with quality controls can be applied to other macromolecular complexes.

### Protein denaturation at the air-water interface during cryo-EM grid preparation

The quality of the purified yeast FAS was furthermore investigated by cryo-EM. During the structural investigation, we observed that electron densities were partially missing for most FAS particles, which prevented structural determination at high resolution. After confirming that the quality of the purified FAS could not be increased during the purification, we concluded that the problem may occur during cryo-EM grid

preparation. This idea was supported by the observation of undamaged particles on grids prepared with the negative-staining method. We analyzed the cryo-EM grids with cryo-ET and demonstrated that most particles adsorbed to the water-air interface with a preferred orientation. Thereby, the damaged side was pointing towards the air interface and consequently was denatured or partially unfolded.

### Graphene grids prevent denaturation of FAS particles

We highlighted the use of graphene grids, rendered hydrophilic with 1-pyrene carboxylic acid, which prevented destruction during grid preparation as FAS particles preferred the graphene-water interface over the air-water interface. The use of graphene grids had major advantages over common grids for the applied enzyme and may be beneficial for other macromolecular complexes.

The work was published in *Elife*: D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *Elife*, 8, e42747

### New structural insights of yeast FAS

The success of the applied strategy was highlighted by a solving the complete yeast FAS structure at high resolution (PDB: 6TA1, 3.1 Å resolution), including the PPT domain at the outside of the barrel. The electron density of the PPT was weak in previous structures because of high flexibilities and, previously, a complete structure of yeast FAS was just available by modeling the PPT, solved as separate protein, inside this density. We validated the positioning of the PPT and revealed a phosphorylation site at Ser1440 and binding of the cofactor NADPH to the KR. An additional phosphorylation site was suggested on Ser1827 but could not be confirmed by the structure, as the position was not resolved at atomic resolution.

The work was published in *IUCrJ*: Joppe, M., D'Imprima, E., Salustros, N., Paithankar, K. S., Vonck, J., Grininger, M., & Kühlbrandt, W. (2020). The resolution revolution in cryoEM requires high-quality sample preparation: a rapid pipeline to a high-resolution map of yeast fatty acid synthase. *IUCrJ*, 7(2), 220-227.

### Characterization of two phosphorylation sites of yeast FAS

We performed a mutational study to investigate Ser1440 and Ser1827 with the objective to evaluate the phosphorylation and to examine their relevance on enzymatic activity. Both serines were changed to either alanine or to glutamic acid, with the latter mimicking the negative charge of the phosphate group. We showed that both mutations at position Ser1827 significantly decreased the fatty acid synthesis. As the glutamic acid mutant led to a larger drop in enzymatic activity, we suggested that the phosphorylation may down-regulate FAS activity. The mechanistic

details for the effect of the glutamate mutation remained elusive, and further work is necessary in this regard. Ser1827 is located at the outside of the PPT and the phosphate may impair phosphopantetheinylation of the ACP during the assembly process. However, such a process would not be reversible, meaning that such phosphorylation would render the FAS particle less active during its lifetime. In contrast, both mutations at Ser1440 slightly improved FAS activity.

#### Yeast FAS is active without the scaffolding domains TM and DM2

Finally, we characterized structural domains of the yeast FAS and demonstrated that the TM and DM2 can be truncated from the enzyme without harming the assembly of yeast FAS. Our data demonstrates that these domains are good candidates for engineering approaches like domain exchanges. Finally, with our work on yeast FAS, we could demonstrate that the established purification strategy is suitable for mutant characterization and protein engineering approaches.

The work was published in *Scientific reports*: Fischer, M., Joppe, M., Mulinacci, B., Vollrath, R., Konstantinidis, K., Kötter, P., Ciccarelli, L., Vonck, J., Oesterhelt, D. & Grninger, M. (2020). Analysis of the co-translational assembly of the fungal fatty acid synthase (FAS). *Scientific reports*, 10(1), 1-13.

#### MSAS from *S. erythraea* can be expressed in *E. coli*

This thesis was further focussed on MSAS, as second protein class, with the aim of its biochemical characterization. We suggested that MSAS from *S. erythraea* is suited for heterologous expression in *E. coli* since it originates from the bacteria Actinomycetes. We designed two differently tagged enzymes for chromatographic purification and showed that the use of an N-terminal tag disturbs the protein stability and significantly reduces the protein quality. We proceeded with the C-terminal tag and established a purification strategy to yield high amounts of protein. The quality of the enzyme was confirmed by SDS-PAGE, SEC and TSA, and the post-translational activation of the ACP was demonstrated by a fluorometric assay. The production of 6-MSA by MSAS was observed when expressing SerMSAS in *E. coli*, but we did not observe any enzymatic activity of SerMSAS *in vitro*.

#### SerMSAS is a non-releasing enzyme and involved in a gene cluster

We analyzed the primary sequence by alignment to related MSASs and showed that a conserved motif is missing which was previously described to catalyze the release of 6-MSA from its enzyme. Additionally, SerMSAS is integrated in a gene cluster, including a KASIII-like protein, a discrete carrier protein and a putative cytochrome P450. A phylogenetic analysis of the KASIII-like protein suggested ChIB3 to be an

ACP-dependent transferase with high similarities to an enzyme from the chlorothricin pathway, responsible for the transfer of 6-MSA from the synthase to a discrete carrier protein. We suggested that our ChIB3 has a similar function and established purification strategies for ChIB3, ChIB2 and the dissected ACP from MSAS. We designed various constructs with different tags and performed codon-optimization to receive sufficient protein to investigate our idea. Of note, in the presence of ChIB3 and ChIB2, SerMSAS was active to produce 6-MSA *in vitro*. Furthermore, ChIB3 was able to transfer 6-MSA moieties from an artificial ester, mimicking 6-MSA-ACP, to the discrete carrier protein ChIB2. Our data demonstrates that the product 6-MSA is not released from the enzyme and most-likely transferred by ChIB3 to ChIB2 for further modifications to a yet unidentified natural product. SerMSAS was also targeted for structural determination and initial data already showed promising 2D-class averages. Despite the interesting data, the non-product releasing properties of SerMSAS revealed that this system was not suitable for our aspired engineering approach.

#### MSAS from *P. patulum* converts non-native starter substrates significantly slower

Consequently, we turned our interest to a related MSAS from *P. patulum*. The purification strategy from SerMSAS was successfully applied to yield sufficient enzyme with high reproducibility. We established a method to analyze the enzymatic activity of PenPaMSAS *in vitro*, by tracking the consumption of NADPH photometrically and combined this approach with product verification by HPLC-MS. The purified PenPaMSAS had slightly higher turnover rates for 6-MSA synthesis compared to the literature, demonstrating the suitability of the applied purification strategy. We reported the conversion of the non-native starter substrates butyryl-CoA and hexanoyl-CoA at significantly slower turnover rates, demonstrating that the enzyme is specialized for its native substrate. Additionally, we analyzed the production of shunt products in the absence of NADPH to draw conclusions about substrate specificities of certain domains. Turnover rates were significantly lower for acetyl-CoA and slightly higher for butyryl-CoA and hexanoyl-CoA compared to the respective salicylic acids, indicating that the triketide intermediate is converted more slowly by KR, KS and TE when the acyl-chain length increases. However, the characterization of individual enzymes is required to investigate the reason for the slower conversion of non-native starter substrates.

#### The AT of PenPaMSAS is specific for its native starter substrate

We analyzed the AT as the responsible domain for the loading of substrates. After designing and purifying dissected KS-AT and ACP proteins, we demonstrated that the AT loads non-native substrates, such as propionyl-CoA and hexanoyl-CoA 400 times slower compared to acetyl-CoA. The data revealed that the AT discriminates against non-native substrates and is responsible or at least a major factor for significantly slower conversion.

### The specificity of the PenPaMSAS AT can be altered by active site engineering

With the objective to change the specificity of this domain through protein engineering, we designed several mutants based on related mutational studies and alignments to other ATs. Screening mutants in activity revealed a promising candidate with higher turnover rates for the loading of butyryl-CoA (~200%) and hexanoyl-CoA (~300%), demonstrating that the substrate specificity can be affected by active site engineering.

### The promiscuous MAT can transfer fluorinated substrates to the DEBS ACP

In contrast to MSAS, the AT of the mammalian FAS was demonstrated to be very promiscuous. Previous studies in our group showed that the MAT can transfer every provided substrate with fast rates and that point mutations can alter the substrate specificities. Because of these fascinating properties, we targeted the MAT as a promising candidate for protein engineering. In contrast to the MAT, ATs from PKS systems usually have very narrow substrate specificities and act as gatekeeper for the selection of substrates. We suggested that the exchange of ATs from PKSs with the MAT would allow site-specific derivatization of polyketides. Our main goal was the incorporation of fluorinated substrates, as this could lead to fluorinated compounds with high pharmaceutical relevance. After chemical synthesis of F-Mal-CoA, we demonstrated that the MAT had excellent transfer kinetics for this substrate. We chose DEBS M6 as PKS model system because it is well described in literature and accessible in *E. coli*. We characterized both ATs from the mammalian FAS and DEBS M6, to analyze whether the MAT can transfer acyl-moieties to the ACP of DEBS M6. The data showed that the specificity constant for the MAT mediated transfer was two to three orders higher of magnitude compared to the AT of DEBS, which can be explained by the high transacylation rates of the MAT. In conclusion, the characterization of the MAT revealed that the domain is suitable to transfer fluorinated substrates to the ACP of DEBS.

### The design of DEBS/FAS hybrids

To harness the polyspecific MAT for the biosynthesis of polyketide derivatives, we designed DEBS/FAS hybrids by exchanging the AT of DEBS M6 with the MAT. The AT of both systems is connected to the KS by a linker domain (LD) and consequently we designed two DEBS/FAS hybrids by exchanging the ATs with or without their linker region. To be able to analyze enzymatic activities of the hybrids, we synthesized a substrate which can be converted by DEBS M6 and validated that both hybrids were active. By comparing oligomerization and enzymatic activities, we showed that the AT exchange without the adjacent LD led to better enzyme properties and consequently the work was continued with that enzyme only. In

addition to MM-CoA, the DEBS/FAS hybrid was able to use the non-native substrates Mal-CoA and of particular interest F-Mal-CoA to produce C2 derivatives of TKLs as demonstrated by NADPH consumption and HPLC-MS. In contrast, the WT protein could not synthesize any derivatives, demonstrating the unique success of this approach.

### The DEBS/FAS hybrid synthesizes macrolactone derivatives

With the catalytically active hybrid, we aimed to produce macrolactone derivatives as pre-cursors for bioactive compounds. We used pentaketide and hexaketide substrates from the related pikromycin pathway and confirmed that DEBS M6 can convert these substrates by using the WT enzyme with the native substrate of DEBS (MM-CoA) to produce 10-deoxymethynolide and narbonolide. These are the precursors for the antibiotics YC-17/methymycin/neomethymycin and pikromycin/narbomycin. In contrast to the WT, the DEBS/FAS hybrid was able to synthesize demethylated and fluorinated derivatives confirmed by NADPH consumption and HPLC-MS. Surprisingly, turnover rates for the conversion of Mal-CoA were higher compared to MM-CoA and only slightly slower for F-Mal-CoA. Thereby, the full length fluorinated macrolactone was synthesized with half of the turnover rate compared to the WT, utilizing the native MM-CoA substrate. This result demonstrates the relevance of the DEBS/FAS hybrid for the biosynthesis of fluorinated compounds.

### DEBS/FAS mediated conversion of FMM-CoA yields 2S-fluoro-10-deoxymethynolide

For further investigation, we aimed to produce demethylated and especially fluorinated compounds in milligram yields. Unfortunately, the targeted fluorinated compounds could not be purified, mainly because of low yields and the occurrence of side reactions. We identified a fluorinated side product, indicating that the elongation product of pentaketide with F-Mal-CoA could not be cyclized by the TE and instead forms a cyclohexanone because of hydrolysis, decarboxylation, and aldol reaction. To increase yields, we tried to prevent this side reaction by introducing fluoro-methyl-disubstitutions (MeCF-group), which should also be beneficial to enhance the chemical stability of the fluorinated compound. The incorporation of such a group would be possible by loading F-MM-CoA to the enzyme. After chemical synthesis, we showed that the DEBS/FAS hybrid can convert this substrate at significantly higher turnover rates compared to F-Mal-CoA and we verified the product by HPLC-MS. Additionally, we isolated the 10-deoxymethynolide derivative for structural analysis by NMR, which revealed the stereospecific introduction of the chemically synthesized F-MM-CoA racemate by the DEBS/FAS-hybrid. Beside its extraordinary relevance from a product point of view, the disubstituted compound is indicative for the KS-mechanism which occurs either in a concerted or in a two-step decarboxylation-addition mechanism and not in a C-C bond formation preceding decarboxylation. Similar to the F-Mal reaction, we determined a cyclohexanone side

product in the absence of NADPH, demonstrating that the elongation product cannot be cyclized by the TE and is hydrolyzed instead. The hydrolyzation product is then subsequently decarboxylated and cyclohexanone is formed by aldol condensation. Our data demonstrates that the AT exchange of PKS systems with the MAT of the mammalian FAS enables the biosynthesis of site-specifically modified polyketide derivatives, highlighted by relevant fluorinated compounds especially by the introduction of MeFC-groups.

The work was published in *bioRxiv*: Rittner, A., Joppe, M., Schmidt, J. J., Mayer, L. M., Heid, E., Sherman, D. H., & Grninger, M. (2021). Directed biosynthesis of fluorinated polyketides. *bioRxiv*.

### New insights into megasynthases

This thesis contributes to the understanding of FAS and PKS systems. We established expression strategies and enabled access to megasynthases for structural and functional characterization. This allowed us to demonstrate engineering approaches and propose guidelines for megasynthase engineering to use them for the biosynthesis of novel natural products. Moreover, we characterized the ATs as gatekeeper for the selection of substrates in PKSs and showed that the substrate specificity can be altered by protein engineering such as active site mutagenesis and domain exchanges. Finally, the creation of a first PKS/FAS hybrid allowed the biosynthesis of a pharmaceutically relevant fluorinated macrolactone among others, demonstrating the huge potential of engineered megasynthases to yield novel natural products.

## Chapter 5: Outlook

The presented work contributes significantly to the emerging field of megasynthases. With a focus on protein engineering, several follow-up projects should be addressed which will be discussed separately for the individual enzymes in the following.

As the characterization of the structural domains TM and DM2 of the yeast FAS revealed that these domains are not essential for the assembly or fatty acid synthesis, they should be targeted for engineering approaches. Using TM and DM2 as attachment sites for domains that can functionalize fatty acid CoA esters could enable to synthesis fatty acid derivatives such as aldehydes, alcohols, or alkenes. Especially domain exchanges with the DM2 are very promising, because of their position at the outside of the barrel.

SerMSAS was characterized as a non-releasing iterative PKS and initial data indicates that a KASIII-like enzyme (SerChIB3), as part of the gene cluster, transfers 6-MSA from SerMSAS to another carrier protein for further modifications. Additionally, based on an amino acid alignment to related MSASs, we highlighted that a conserved motif in the TE region is missing which was previously predicted to be crucial for the product release.<sup>82</sup> Further experiments are required to validate this idea. To investigate if the enzyme has indeed a non-releasing TE region, a domain exchange with related enzymes like *P. patulum* or *A. terreus* could be helpful. Alternatively, the TE region from the related enzymes can also be added separately to the SerMSAS reaction solution. In future experiments, acylation of ChIB2 with 6-MSA moieties should be analyzed in the presence of SerMSAS, ChIB3, acetyl-CoA, Mal-CoA and NADPH which would validate our idea about the synthesis and transfer of 6-MSA moieties. It would be of great interest to identify the yet unknown natural product of this gene cluster. This could be achieved by analyzing the natural products of the organism in combination with metabolic engineering approaches, deleting relevant genes. With the discovered product, a mechanism for the biosynthesis could be established and validated by loading additional enzymes with 6-MSA moieties by 6-MSA-ChIB2.

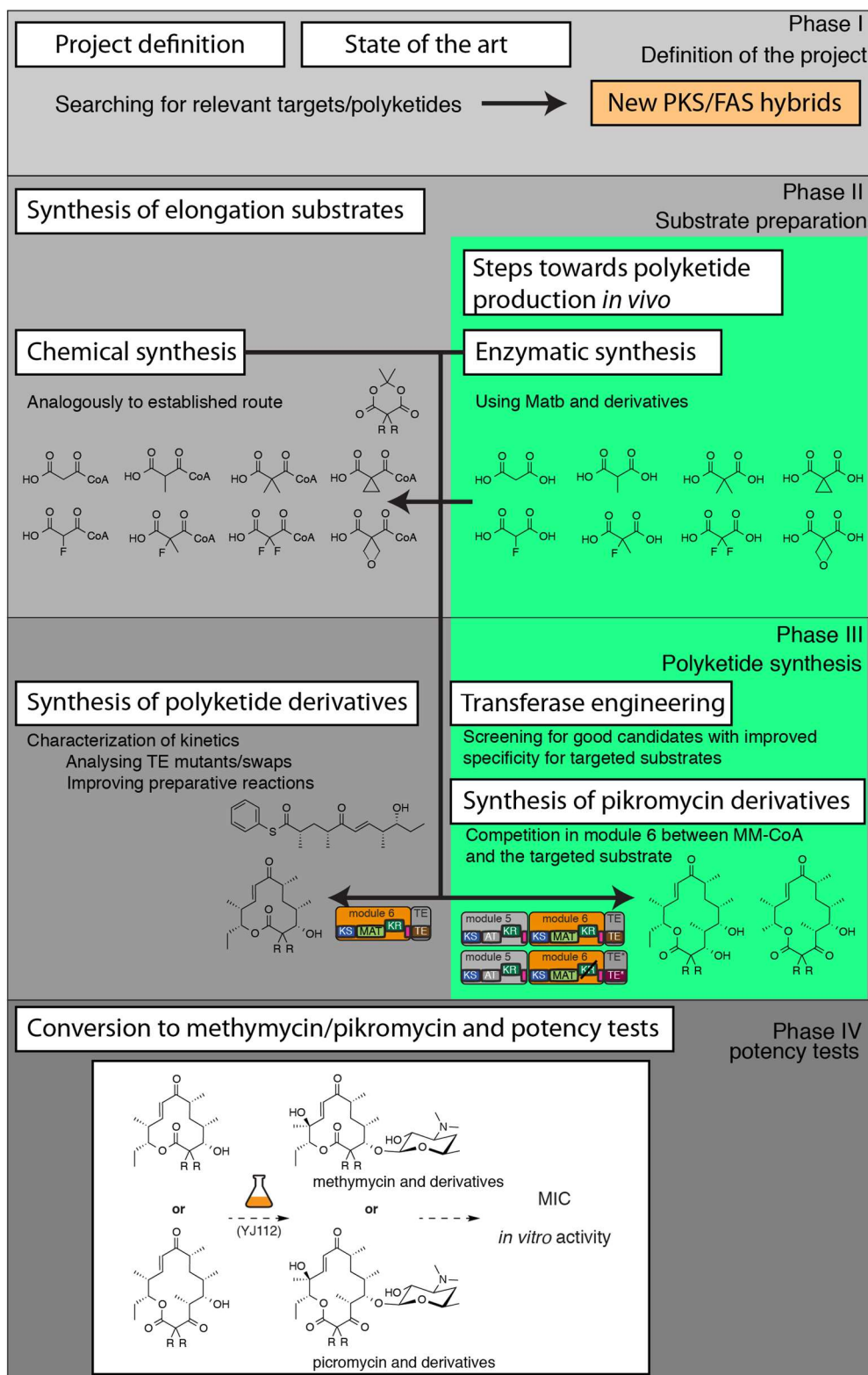
Because of the non-releasing properties, we turned our interest to PenPaMSAS with the objective to synthesize 6-MSA derivatives. We identified the AT as a bottleneck, limiting the incorporation of non-native starter substrates because of its substrate specificity. With an attempt to overcome this limitation by active site mutagenesis, we discovered a mutant which can transfer non-native substrates significantly faster (200-300%). The data demonstrates that the strategy can enhance turnover rates. However, further mutations are required to improve the system for the main goal to produce high amounts of 6-MSA derivatives *in vivo*. As the native substrate acetyl-CoA is vital and cannot be excluded from the production cells, the non-native substrate must be transferred with similar or faster rates to achieve higher yields of these derivatives. Otherwise PenPaMSAS would discriminate against non-native



substrates and mainly produce the native product. To make a mutational study easier, a structure of the AT would be very helpful, especially in a butyryl- or hexanoyl-bound state. This would further enhance our knowledge about ATs, because it would enable us to analyze the substrate binding pocket and conformational variability. By comparison to other ATs, we could investigate the reason for the substrate specificities and probably get an idea of how these enzymes have evolved to gain such properties. Furthermore, a structure could help us to find a truncation site to exchange this domain with related enzymes, able to transfer longer starter substrates.

The MAT of the mammalian FAS was successfully exchanged with the AT of DEBS M6. We demonstrated that the resulting DEBS/FAS hybrid can efficiently synthesize macrolactone derivatives, in particular fluorinated polyketides. These results show the great potential of this project for further development (figure 58). The successful purification of two DEBS/FAS hybrids, with different chimeric interfaces and catalytic efficiencies can help to enhance our understanding on how to design those interfaces. Here, structural information would be very helpful to investigate how altered interaction and conformational variability can affect stabilities and efficiencies. Furthermore, the demonstrated approach of the DEBS/FAS hybrids should be applied on other PKS systems to synthesize derivatives of bioactive compounds. Most interesting are polyketides which failed in clinical phases as they are well studied and resemble excellent references to judge the impact in fluorination. However, our DEBS/FAS hybrid should be the main target for various enzyme- and product orientated studies. With the established strategy for the chemical synthesis of F-Mal-CoA and F-MM-CoA, we can synthesize a library of substrates which likely can be incorporated in the polyketide scaffold. Alternatively, an enzymatic pathway could be established to synthesize these substrates which would be beneficial for the long-time objective to produce polyketide derivatives by biosynthesis *in vivo*. This approach requires to provide non-native/non-canonical substrates inside an engineered strain which can be achieved by incorporating the established enzymes through metabolic engineering. However, the library of substrates can be used to synthesize 10-deoxymethynolide derivatives analogously to the established strategy which should be converted to YC-17, methymycin and neomethymycin derivatives for potency testing. Furthermore, the substrates could also be used with a DEBS3 enzyme, with only one of the two modules bearing the MAT. In the presence of the native MM-CoA and an additional non-native elongation substrate we would be able to synthesize narbonolide derivatives. Three challenges need to be addressed for such an approach. First, the KR of module 6 needs to be knocked-out to ensure a keto-group at C3. Secondly, an exchange of the intrinsic TE with the PIKS TE may be required to enable cyclization in the presence of this C3 keto-group. Lastly, the MAT needs to be engineered towards higher specificities for non-native substrates to be able to incorporate the target substrate by discrimination against the native MM-CoA.

By addressing the latter, we would be able to create a library of specialized MAT variants for various substrates. Most interestingly are those mutants which discriminate against MM-CoA as it is required for the macrolactone synthesis. Additionally, mutants which discriminate against Mal-CoA are needed for approaches *in vivo*, because Mal-CoA is required for the fatty acid synthesis and cannot be excluded from the production host. The library of MAT mutants is highly relevant for the PKS field as the specialized ATs can be used to incorporate specific substrates at certain positions, which will ultimately lead to novel natural products.



**Figure 58** | Outlook in exchanging the MAT with ATs from PKSs.

## Chapter 6: Experimental procedure

### 6.1 Material and methods

Materials were purchased from **Takahara** (CloneAmp HiFi PCR Premix), **Serva** (BSA), **Merck** (plasmid pET22b, potassium chloride, H<sub>2</sub>O<sub>2</sub>, acetic acid), **Clontech** (In-Fusion HD Cloning Kit (Takara Bio), Stellar Competent Cells, His60 Ni-NTA Superflow Resin), **Merck Millipore** (Amicon ultra centrifugal filters), **Carl Roth** (glas beads (0.25-0.5 mm), Tween20, Na<sub>2</sub>HPO<sub>4</sub>, TRIS, NaH<sub>2</sub>PO<sub>4</sub>, K<sub>2</sub>HPO<sub>4</sub>, KH<sub>2</sub>PO<sub>4</sub>, DTT, LB-Media, LB-Agar, Glucose, ammonium persulfate, potassium permanganate, tricine, magnesium chloride hexahydrate, ampicillin sodium salt, isopropyl-β-D-thiogalactopyranoside, glycerol), **Agilent** (BL21gold(DE3)), **Qiagen** (Ni-NTA affinity resin), **IBA** (5 mL Strep-Tactin columns, desthiobiotin), **GE Healthcare** (Superdex 200 increase 10/300 GL, Superdex 200 HiLoad 16/60, Superose 6 increase 10/300 GL, ÄKTA Basic, ÄKTA Explorer), **New England Biolabs** (Dpn1, EcoRI, NdeI), **Thermo Fisher Scientific** (UltiMate 3000 UHPLC system, Synchronis™ aQ-C18-LC column, Sypro Orange (Invitrogen)), **Roche Diagnostics GmbH** (protease inhibitor cocktail, NADPH), **BMG Labtech** (CLARIOstar), **Amico** (French Pressure Cell Press), **New Brunswick** (Incubation Shaker Innova 44), **Eppendorf Research** (Micropipettes, Thermomixer comfort for 1.5 mL reaction tubes), **Savant** (Speed Vac Concentrator), **Becton, Dickinson company** (Yeast extract, peptone, Agar-Agar), **Alfa Aesar** (lithium acetate, imidazole), **Fluka** (PEG-1500, ethyl chloroformate), **AppliChem** (DNAse, PBS, sodium hydroxide), **VWR** (sodium chloride, TEMED, 96% ethanol, methanol), **Apollo Scientific** (6-MSA), **Caelo** (ascorbic acid) and **Sigma** (betaine, β-NAD<sup>+</sup>, NADH, NADPH, D-sorbitol, myristic acid, palmitic acid, stearic acid, single stranded DNA from *salmon testis* (10 g/L), α-ketoglutarate dehydrogenase, α-ketoglutaric acid, thiamine pyrophosphate, ammonium phosphate, EDTA, G418, spectinomycin dihydrochloride pentahydrate, kanamycin sulphate, all CoA-esters except F-Mal-CoA and F-MM-CoA, CoA, EDC×HCl, propionaldehyde (propanal), lithium hydroxide, NEt<sub>3</sub>, select-fluor, methyl-Meldrum's acid, Phenyl-trimethylsilylsulfid, sodium metabisulfite, BIS-TRIS, Coomassie G-250, ε-aminocaproic acid).

#### 6.1.1 Cloning of plasmids

Cloning of plasmids was performed accordingly to the described procedure in the following. The used primers and applied annealing temperature are listed in table S2-S3 (chapter 7.4). PCR was usually performed in a similar fashion with a pre-incubation step at 98°C for 180 seconds, followed by 30 repeated cycles of incubation at 98°C for 10 seconds, incubation at the annealing temperature for 20 seconds and incubation at 72°C for 20 seconds + 5 seconds/1000bp. Then a final incubation step was performed at 72°C for 420 seconds. After PCR, all templates were digested with Dpn1. Point mutations were introduced by PCR with primers. In

case of unsuccessful PCRs or large templates, two separated PCRs were performed and inserted by the In-Fusion HD Cloning Kit (sequence and ligation dependent cloning) after purification. Sequence and ligation independent cloning was also used when DNA sequences were exchanged within plasmids. All reaction mixtures were purified by gel electrophoresis and DNA was extracted and purified with the Wizard® SV Gel and PCR Clean-up System (Promega) or NucleoSpin® Gel and PCR Clean-up kit (Machery and Nagel). Then Stellar™ Competent Cells were transformed with 5-10% of the PCR- or assembly master-mix and distributed on agar plates with the appropriate antibiotic. After overnight incubation at 37°C, cells were used to inoculate 5-8 mL cultures. Plasmids were isolated with the GeneJET Plasmid Miniprep Kit (Thermo Scientific). Plasmid sequences were validated by sequencing, performed by Microsynth Seqlab (Göttingen, Germany)

### Yeast FAS:

S1827E point mutation for construct pMJD095 was inserted with PCR using the template MF319d. pMJD095 was then used as template to introduce S1440E point mutation using PCR with 1 M betaine as additive to yield plasmid pMJD096. The vectors pMJD097 and pMJD098 were produced by sequence and ligation independent cloning using the template MF319d and pMJD097, respectively.

TM and DM2 deletions were created by sequence and ligation independent cloning with the template MF639k1 (or pMF014) and MF319d (or pMF011) yielding pMF025 (TM deletion, removal of 2-137 on the  $\beta$ -chain) and pBAR1 (DM2 deletion, removal of 541-598 on the  $\alpha$ -chain).

### SerMSAS

For the SerMSAS construct without Strep-tag, MW001 was used as a vector and the tag was removed by sequence and ligation independent cloning to yield pMJD065.

### SerChIB3

SerChIB3 was cloned from genomic DNA of *S. erythraea* and the insert was introduced into a vector by sequence and ligation independent cloning with vector DNA received from a PCR using the template MW001 to yield pMJD063. To remove the Strep-tag, pMJD063 was used as a vector and the tag was removed by sequence and ligation independent cloning to yield pMJD068. The MBP-tag was introduced by inserting the SerChIB3 sequence into an MPT-tag carrying vector (pMK60, kindly received from Maja Klaus) by sequence and ligation independent cloning to yield pMJD069. The codon optimized gene received from ThermoFisher (GeneArt) was amplified and the insert was introduced into different vectors by sequence and

ligation independent cloning. The vectors contained Strep-tag and His-tag (pMJD078), only Strep-tag (pMJD080) or only His-tag (pMJD079) and was received from PCRs using the template MW001.

The plasmid pMJD090 was only used as a cloning template. For its creation, a (GGSG)<sub>2</sub>-linker was introduced in between the C-terminal end of the SerChIB3 sequence and the His-tag by sequence and ligation independent cloning with DNA received from a PCR using the template pMJD079.

### SerChIB2

SerChIB2 was cloned from genomic DNA of *S. erythraea* and sequence and ligation independent cloning was performed with a vector DNA received from a PCR using the template pMJD065 to yield pMJD066.

### PenPaMSAS KS-AT

The KS-AT gene was cloned from pAR014 and sequence and ligation independent cloning was performed with a vector DNA received from a PCR using the template pMJD090 to yield pISI011. Point mutations were introduced by sequence and ligation independent cloning using DNA received from a PCR with the template pISI011.

### PenPaMSAS ACP(0)

The PenPaMSAS\_ACP(0) gene was cloned from pAR014 and sequence and ligation independent cloning was performed with a vector DNA received from a PCR using the template pMJD090 to yield pISI014.

### *Npt* (UniProt code: A0A1Y2MXW0):

The *Npt* gene was kindly received from Kyra Geyer by the (Erb laboratory, Max Planck Institute for terrestrial Microbiology, Marburg) and after amplification the insert was introduced into a vector by sequence and ligation independent cloning with a vector DNA from a PCR using the template pAR357 to yield pMJD091

### DEBS/FAS hybrids **H1** and **H2**

DEBS/FAS hybrid H1 (pMJD076) and H2 (pMJD077) were created by sequence and ligation cloning with vector DNA received from a PCR with the template pBL18 and inserts, received from amplification of pAR264 (mFAS construct).

### DEBS M6-KS(C1661G)-AT

DEBS M6-KS(C1661G)-AT (pAR432) was created by sequence and ligation dependent cloning with two fragments of amplified DNA received from a PCR using the template pBL18. Cloning was performed by Alexander Rittner (Goethe-University, Frankfurt) and is shown for the sake of completion.

### DEBS M6-ACP-W

DEBS ACP6 plasmid was designed like reported by *Kim et al.* with an additional tryptophane for photometrical concentration determination.<sup>181</sup> The insert was created by amplification of pBL18 and incorporated into an linearized pET-28a plasmid (Merck Millipore) with sequence and ligation dependent cloning. The enzyme NdeI and EcoRI were used for digestion.

## **6.1.2 Protein expression and purification**

The concentrations of the protein solutions were calculated from the absorbance at 280 nm, which was determined with the NanoDrop 2000c (Thermo Scientific). The extinction coefficients were calculated from the primary sequence without the N-terminal formylmethionine with the software ProtParam (Expasy, <https://web.expasy.org/protparam/>) or by CLC Main Workbench (version 6.9.1).

### Yeast FAS

We established an expression and purification strategy for the yeast FAS, including a Strep-tag integrated to the C-terminus of the  $\beta$ -chain. We used a FAS-deficient strain of *S. cerevisiae* (BY\_PK1238\_1A\_KO), kindly received from Manuel Fischer<sup>130</sup>, and transformed it with 1-2  $\mu$ g of plasmids, harbouring the  $\alpha$ - and  $\beta$ -chain of FAS. The FAS-deficient strain was grown in YPD medium (1% yeast extract (BD), 2% pepton (BD), 2 % glucose) containing 200 mg/L G418 and 50 mg/L free fatty acids (myristic acid, palmitic acid, stearic acid) until a OD<sub>600</sub> of 0.8 and transformation was performed on ice with sterile solutions, using the method by Schiestl and Gietz<sup>197</sup>. For transformation, 50 mL of cells were centrifugated (3200 g, 4°C, 5 min) and washed with 4 mL water. After centrifugation (3200 g, 4°C, 5 min), cells were incubated with 3 mL 0.1 M LiOAc for 5 minutes at 24°C. After centrifugation (5000 g, 4°C, 15 s), cells were re-suspended with 912  $\mu$ L water, 432  $\mu$ L 1M LiOAc, 2880  $\mu$ L PEG-1500 (50%) and 30  $\mu$ L of SS-salmon-sperm-DNA (10 g/L), with the latter one being incubated at 98°C for 5 minutes prior to the usage. Then cells were divided in six parts and the plasmids were added followed by incubation at 30 minutes at 30°C and 20 minutes at 42°C. Transformed cells were washed with water and grown on YPD-agar (1% yeast extract (BD), 2% pepton (BD), 2 % glucose, 2% agar-agar (BD))

with 200 mg/L G418 and after 1-2 day of incubation at 30°C, a single colony was taken to inoculate a pre-pre culture of 5 mL YPD medium containing 200 mg/L G418. The culture was grown for 1-2 days at 30°C and used to inoculate a 50 mL pre-culture at a theoretical OD<sub>600</sub> of 0.2. After 2.5 days of incubation at 30°C, a 2L culture with the same medium was inoculated at a theoretical OD<sub>600</sub> of 0.5 and grown 11.5 h at 30°C. Cells were harvested by centrifugation, washed with water, and resuspended in the purification buffer (100 mM sodium phosphate, pH 6.5) containing two tablets of cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche) and one spatula tip of DNase. Glass beads (0.25-0.5 mm) were added and cells were lysed by bead disruption using Fastprep24 (Biomedicals) 10 times for 15 seconds. After centrifugation of glass beads (3200 g, 4°C, 10 min) and insoluble protein (20000 g, 4°C, 40 min), supernatant was filtered through a 0.22 µm PVDF membrane (Ahlstrom) and poured on 5 mL Strep-tactin columns for purification. After a wash step with 40 mL of purification buffer, FAS was eluted with the same buffer containing desthiobiotin (2.5 mM final concentration), concentrated with an Amicon ultra centrifugal filter, filtered through a 0.22 µm PVDF membrane (Ahlstrom) and further purified on a S6 Increase 10/300 GL column.

### SerMSAS

SerMSAS was usually co-expressed with *Sfp* from *Bacillus subtilis* (*B. subtilis*) (pAR357) in *E. coli* BL21gold(DE3) cells. Transformation of cells was performed as described from the manufacturer and after overnight incubation on LB-Agar containing 100 µg/mL ampicillin and 50 µg/mL spectinomycin, a single colony was used to inoculate a 20 mL pre-culture with LB medium (Lennox), containing the same antibiotics. After overnight incubation at 37°C, the pre-culture was used to inoculate 1L TB-medium (12 g tryptone (BD), 24 g yeast extract (BD), 5 mL glycerol (87%), 2.31 g KH<sub>2</sub>PO<sub>4</sub>, 16.43 g K<sub>2</sub>HPO<sub>4</sub>), containing 100 µg/mL ampicillin and 50 µg/mL spectinomycin. The main culture was grown until a OD<sub>600</sub> of 0.5-0.8 and then cooled at 4°C for 20-30 minutes. Afterwards, the culture was induced with 0.25 µM IPTG and subsequently incubated for 16 hours at 20°C. Cells were harvested by centrifugation and resuspended in the purification buffer (100 mM sodium phosphate, pH 7.6, 100 mM sodium chloride, 20% glycerol, 1 mM EDTA, 20 mM imidazole was used additionally when a His-tag was used; during the thesis the buffer was optimized to 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA 20 mM imidazole). After addition of one spatula tip of DNase1 and cell lysis using French press, insoluble parts were centrifugated (50000 g, 40 min) and supernatant was transferred to a Strep-Tactin-column or Ni-NTA-column (MgCl<sub>2</sub> was added to a final concentration of 2 mM when a His-tag was used). The Strep-Tactin-column was washed with 8 CV purification buffer and proteins were eluted with the same buffer containing desthiobiotin (2.5 mM final concentration). The Ni-NTA-column was washed with 5 CV purification buffer, followed by 2 CV with increased imidazole concentration (60 mM) and proteins were eluted with the same buffer containing 300 mM imidazole. The elution fractions were concentrated with



appropriate Amicon ultra centrifugal filter, filtered through a 0.22  $\mu$ M PVDF membrane (Ahlstrom) and further purified on a S200 Increase 10/300 GL column or S6 Increase 10/300 GL column.

### SerChIB3

SerChIB3 was expressed and purified with the same strategy as used for SerMSAS but without co-expression with *Sfp* and consequently spectinomycin was avoided. Additionally, the optimized buffer was used for purification on the Ni-NTA column (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA, 20 mM imidazole). When a Strep-tag was used for purification, we used the same buffer without imidazole or transferred the elution fraction of the Ni-NTA purification directly on the Strep-tactin column when both tags were used. The column was washed with 8 CV of the same buffer and protein was eluted by addition of 2.5 mM desthiobiotin. For purification using the MBP-tag, we transferred the elution fraction of the Ni-NTA purification on an amylose column. The column was washed with 8 CV of the same buffer as used for the purification on a Strep-tactin column and the protein was eluted in three fractions by addition of 10 mM maltose.

### SerMSAS-ACP(0) and SerChIB2

ACPs were co-expressed with *Sfp* from *B. subtilis* (pAR357) or *Npt* from *Streptomyces platensis* (pMJD091) and purified with the same strategy as used for SerMSAS on a Ni-NTA column in the purification buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA 20 mM imidazole).

### PenPaMSAS

PenPaMSAS was co-expressed with *Sfp* from *B. subtilis* (pAR357) and purified with the same strategy as used for SerMSAS on a Ni-NTA column in the purification buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA 20 mM imidazole).

### PenPaMSAS KS-AT

KS-AT didomain was expressed and purified with the same strategy as used for SerMSAS on a Ni-NTA column in the purification buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA 20 mM imidazole) but without co-expression with *Sfp* and consequently spectinomycin was avoided.

### PenPaMSAS ACP(0)

ACP(0) was co-expressed with *Sfp* from *B. subtilis* (pAR357) and purified with the same strategy as used for SerMSAS on a Ni-NTA column in the purification buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA 20 mM imidazole).

### KS(C161G)-AT of mFAS

Expression and purification of mFAS KS(C161G)-AT didomain was performed by Lara Maria Mayer, supervised by Alexander Rittner and is shown for the sake of completion. The didomain was purified with a similar strategy as used for SerMSAS with the following changes. Expression was performed without *Sfp* and consequently spectinomycin was avoided. Additionally, 1% (w/v) glucose was used in all pre-cultures. Furthermore, a different buffer was used for purification (50 mM potassium phosphate, 200 mM potassium chloride, 10 % glycerol, 1 mM EDTA, 30 mM imidazole, pH 7.0). After lysis, the protein was purified on a Ni-NTA-column with just one wash step, using 5 CV of the purification buffer, followed by elution by increasing the imidazole concentration to 300 mM. The eluate was then poured on a strep-tactin column for another purification step. After a wash step with 5 CV of a strep-wash buffer (250 mM potassium phosphate, 10 % glycerol, 1 mM EDTA, pH 7.0), the didomain was eluted with 2.5 CV of the elution buffer containing additionally 2.5 mM D-desthiobiotin. Finally, the protein was purified in a final step on a Superdex 200 increase 10/300 GL (250 mM potassium phosphate, 10 % glycerol, 1 mM EDTA, pH 7.0).

### ACP of mFAS and DEBS M6-ACP

ACP of mFAS was co-expressed bicistronically with *Sfp* from *B. subtilis* (pAR357), whereas DEBS M6-ACP was co-expressed with *Sfp* or with *Npt* from *Streptomyces platensis* (pMJD091) on separate plasmids. Both expressions were performed similar beside that only ampicillin (100 µg/mL) was used for the mFAS ACP expression, while DEBS ACP6 expression required kanamycin (50 µg/mL) and spectinomycin (50 µg/mL). The expression and purification of ACPs were performed with the same strategy as used for SerMSAS on a Ni-NTA column with the following changes. For the pre-cultures 1% (w/v) glucose was used, and the purification was performed with a different buffer (50 mM sodium phosphate, 200 mM sodium chloride, 20 % glycerol, 1 mM EDTA, 30 mM imidazole, pH 7.4). After lysis, the protein was purified on a Ni-NTA-column with just one wash step, using 5 CV of the purification buffer, followed by elution by increasing the imidazole concentration to 300 mM. Finally, the protein was purified in a final step on a Superdex 200 16/60 or 26/60 with the buffer (50 mM potassium phosphate, 200 mM NaCl, 10 % glycerol, 1 mM EDTA, 1 mM DTT, pH 7.4). The expression and purification of the ACPs was performed by Lara Maria Mayer and me and is shown for the sake of completion.

Phosphopantetheinylation of DEBS ACP6 was analysed by ESI mass spectrometry, performed by Khanh Vu Huu and Kudratullah Karimi from the Morgner group (Goethe-University, Frankfurt), which demonstrated that only co-expression with *Npt* led to complete activation. Therefore, all experiments were performed with this protein.

### DEBS M6 (WT) and DEBS/FAS hybrids H1 and H2

All proteins were co-expressed with *Sfp* from *B. subtilis* (pAR357) and purified with the same strategy as used for SerMSAS on a Ni-NTA column in the optimized purification buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA 20 mM imidazole). Purification on the Ni-NTA column was followed by SEC on the Superdex 200 10/300 GL, Superdex 200 16/60 or Superose 6 Increase 10/300 GL, dependent on the amount of protein (in buffer: 50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA).

### **6.1.3 SDS-PAGE**

SDS-PAGE was performed to investigate the purity of proteins. The composition of the gel is listed in the following table. Each gel contained a stacking gel, and the amount of acrylamide was adjusted to the size of the analysed protein. The samples usually contained 1 µg of enzyme, 6 µL of laemmli loading buffer (375 mM Tris-HCl, pH 6.8, 50% glycerol, 10% SDS, 10% mercaptoethanol, 0.01% bromophenol blue, 50 mM EDTA) and filled up with water to 20 µL. The flow through contained 1 µL sample, wash fractions 14 µL sample and the amount of elution fraction was adjusted to its concentration to obtain 1 µg protein. To analyse inclusion bodies, the pellet was washed twice (1xPBS, 2% triton-X, 5 mM EDTA), resolved in a urea buffer (1xPBS, 8 M urea) and 2 µL of supernatant were taken. After incubation at 98°C for 5 min, 10 µL of each sample was applied to the gel and electrophoresis was performed at 200 V for 45 min.

**Table 3** | SDS-PAGE Gels were prepared according to the following composition

Material	4% stacking gel	7% separation gel (50-500 kDa)	10% separation gel (20-300 kDa)	12% separation gel (10-200 kDa)	15% separation gel (3-100 kDa)
Acrylamide stock	0.53 mL	3 mL	4 mL	4.8 mL	6 mL
Separation buffer (1.5 M Tris-HCl, pH 8.8)		3 mL	3 mL	3 mL	3 mL
Stacking buffer (0.5 M Tris-HCl, pH 6.8)	1 mL				
Water	2.47 mL	6 mL	5 mL	4.2 mL	3 mL
10% APS	0.02 mL	0.1 mL	0.1 mL	0.1 mL	0.1 mL
TEMED	0.01 mL	0.01 mL	0.01 mL	0.01 mL	0.01 mL

### 6.1.4 Thermal-shift-assay (TSA)

To analyse the thermal stability, TSA was performed in white Multiplate PCR plates 96 well (BioRAD) and sealed with tape (iCycler IQ optical tape (BioRAD)). The reaction solution contained 2  $\mu$ L of protein solution (0.8-1.2 mg/L), 21  $\mu$ L buffer and 2  $\mu$ L 62.5X SYPRO Orange protein gel stain (Thermo Fisher Scientific, Invitrogen). The fluorescence was measured from 5°C to 95°C in 0.5°C/min steps with the excitation wavelength set to 450-490 nm and emission wavelength set to 560-580 nm. The melting temperatures was calculated with the software CFX Maestro 1.0.

### 6.1.5 Enzymatic activity *in vitro* (NADPH consumption)

#### Yeast FAS:

The enzymatic activity of the yeast FAS was analysed in UV transparent flat bottom wells (Corning). A total volume of 120  $\mu$ L was used containing 0.7  $\mu$ g FAS solution, 417  $\mu$ M acetyl-CoA, 500  $\mu$ M Mal-CoA and 250  $\mu$ M NADPH in buffer (200 mM potassium phosphate, pH 7.3, 1.75 mM DTT, 0.03 mg/mL BSA) and background measurements were performed in the absence of Mal-CoA. The absorption was measured at 334 nm and converted to concentration with a NADPH calibration curve.

#### SerMSAS:

The enzymatic activity of SerMSAS was analysed in UV-cuvette (UV-Cuvette micro, BRAND) with the Nanodrop 2000C (Thermo Scientific). Thereby the absorption of NADPH was analysed at 340 nm for five minutes. The final concentrations of substrates, enzymes and buffer compositions are described in the respective figure legends. The adsorption was converted to concentration by using the NADPH extinction coefficient (6220 M<sup>-1</sup>cm<sup>-1</sup>).

### PenPaMSAS:

We established a fluorometric assay for plate reader read-out to investigate the enzymatic activity and to be able to compare different conditions simultaneously. We used 384 white well plates (Greiner Bio-one) and the microplate reader (ClarioStar, BMG Labtech) with excitation and emission wavelengths set to 348 nm and 476 nm and analysed the NADPH consumption for 10 minutes at 25°C. The final concentrations of substrates, enzymes and buffer compositions are described in the respective figure legends. The fluorescence was converted to concentration with a NADPH calibration curve.

### Specific KR activity of DEBS M6 KR

We established a fluorometric assay with plate reader read-out to investigate the enzymatic activity of DEBS KR6, based on a previous report.<sup>198</sup> We used 384 white well plates (Greiner Bio-one) and the microplate reader (ClarioStar, BMG Labtech) with excitation and emission wavelengths set to 348 nm and 476 nm (gain: 1301; focal height: 12.4 mm; flashes: 17; orbital averaging: off) and analysed the NADPH consumption for 3 minutes at 25°C. Two solutions were prepared as 4-fold concentrated stocks, whereas the *trans*-1-decalone solution was prepared as 2 fold concentrated stock in the assay buffer (400 mM phosphate buffer, 20% glycerol, 2 mM DTT, 1 mM EDTA, 0.8% DMSO, pH 7.2).<sup>28</sup> The solutions were pre-incubated at 25°C and then 5 µL of the 1.2 µM enzyme stock solution (containing the DEBS M6 construct) and 5 µL of the 240 µM NADPH stock solution were mixed with 10 µL of the 4 mM *trans*-1-decalone solution. After background correction (absence of enzyme), the decrease of NADPH was converted to concentrations with a NADPH calibration curve. All enzymes were analysed in technical and biological triplicates. Final concentrations were 0.3 µM enzyme, 60 µM NADPH, 2 mM *trans*-1-decalone.

### DEBS M6 enzymatic activity

We established a fluorometric assay with plate reader read-out to investigate the enzymatic activity of DEBS KR6 during the synthesis of TKLs and macrolactones. We used 384 white well plates (Greiner Bio-one) and the microplate reader (ClarioStar, BMG Labtech) with excitation and emission wavelengths set to 348 nm and 476 nm (gain: 1301; focal height: 12.4 mm; flashes: 17; orbital averaging: off) and analysed the NADPH consumption for 13-20 minutes at 25°C. All ingredients were prepared as 4-fold concentrated stocks, while the NADPH and CoA-ester solutions were combined to a 2-fold concentrated stock. The same assay buffer was used as for the KR specific assay but without DTT, as it destabilizes F-Mal-CoA (400 mM phosphate buffer, 20% glycerol, 1 mM EDTA, 0.8% DMSO, pH 7.2).<sup>28</sup> After pre-incubation at 25°C, 5 µL of the two solutions, containing 16 µM enzyme

(DEBS M6 construct), 20 mM natural diketide SNAC **2** or 4 mM pentaketide **9** or 4 mM hexaketide **10** were mixed with 10  $\mu$ L of the combined solution, containing 120  $\mu$ M NADPH and 400  $\mu$ M extender CoA-ester. Prior to the mixing, the hexaketide **10** was deprotected similar as previously reported.<sup>199</sup> For that, 4 mM of the substrate was mixed with 25 mM ascorbic acid and 1 mM sodium metabisulfite and incubated for 20 min with light irradiation at 365 nm. After background correction (absence of enzyme for **2** and absence of extender substrate for **9** and **10**), the decrease of NADPH was converted to concentrations with a NADPH calibration curve. All enzymes were analysed in technical and biological triplicates. Final concentrations were 4  $\mu$ M enzyme, 5 mM **2** or 1 mM **9** or 1 mM **10**, 200  $\mu$ M extender CoA-ester and 60  $\mu$ M NADPH. Turnover rates for the comparison of F-MM-CoA, MM-CoA, Mal-CoA (figure 52 B, chapter 2.4.3) were performed with slightly different conditions and just in technical triplicates (higher enzyme concentration 6.4  $\mu$ M compared to 4  $\mu$ M, different buffer 250  $\mu$ M potassium phosphate, 10 % glycerol, pH 7.0).

### 6.1.6 Analysing protein denaturation at the air-water interface

To simulate the air exposure, we performed three different experiments in triplicates with a FAS solution at 0.01 mg/mL in the purification buffer (100 mM sodium phosphate, pH 6.5). After the procedure, each sample was analysed with negative-stain EM, which was performed by Edoardo D'Imprima (Max Planck Institute for Biophysics, Frankfurt). The untreated FAS solution was used as a positive control. In the first experiment, we bubbled air for about 10 seconds through the FAS solution. For the second experiment, we sealed a 5 cm 100  $\mu$ L intraMARK disposable glass micropipette (Brand, Wertheim, Germany) on both ends and passed 200  $\mu$ L of the FAS solution over it. In the third experiment we incubated 20  $\mu$ L of the FAS solution for 15 second on a carbon-coated grid and collected the FAS particles, which adsorbed on the air-water interface with a second grid, by touching the top of the drop.

### 6.1.7 Native-PAGE Western blot to investigate the assembly of the yeast FAS

To analyse the FAS assembly from cell lysate, we transformed the engineered *S. cerevisiae* FAS deletion strain with the respective plasmid as described in chapter 6.1.2 (expression of yeast FAS). Then we cultivated the cells in YPD medium supplied with 50 mg/mL fatty acids and 200 mg/L G418 (1% yeast extract (BD), 2% pepton (BD), 2 % glucose, myristic acid, palmitic acid, stearic acid) over night at 30°C. This pre-culture was used to inoculate a fresh culture with the same medium to a theoretical OD(600) of 0.2, which was cultivated at 30°C until an OD(600) of 1-2. The cells were harvested by centrifugation (3200 g, 4°C, 8 min) and washed with 40 mL ice cold water. After another centrifugation step (3200 g, 4°C, 5 min), the cells

were resuspended with 5 mL/OD(600) lysis buffer (20 mM Tris-HCl, pH 7.6, 1.1 M sorbitol, 1 mM DTT, 57 units/mL zymolyase (AMS Biotechnology). After incubation for about 2 h at 30°C, cells were harvested again by centrifugation (3200 g, 4°C, 5 min). The cells were resuspended in homogenization buffer (50 mM Tris-HCl, pH 7.6, 1.1 M glycerol, 5 mM EDTA, 1% of EDTA-free Protease Inhibitor Cocktail (Roche)) and incubated for 2 h on ice. After centrifugation (3200 g, 4°C, 15 min), supernatant was concentrated to a total protein concentration of 1-5 mg/mL using a 100 kDa cut-off Amicon® centrifuge filter (Merck Millipore). For comparison, we used the same concentrations for the purified proteins. Each sample was then mixed with Blue Native sample buffer (2x, Serva Electrophoresis GmbH) and loaded on a 3-12 % Native-Page Bis-Tris gel (Life Technologies, Novex). Electrophoresis (150 V, 2 h, ice cooling) was performed with the Blue Native buffer system from Serva (Serva Electrophoresis GmbH). After blotting onto a PVDF membrane using electro-transfer (25 V, 1h, ice cooling) in the buffer (0.5 M bicine, 0.5 M Bis-Tris, 20 mM EDTA, 10% methanol), the membrane was blocked with I-Block (Thermo Fisher Scientific). We used rabbit-anti FAS antiserum (1:5000 dilution)<sup>132</sup> and horseradish peroxidase goat anti-rabbit IgG (Thermo Fisher Scientific, Pierce, 200 ng/mL) to detect the FAS after 5 min incubation with peroxidase substrate at 25°C (Carl Roth GmbH) and luminescence for 12-20 seconds on the Fusion-SL (Vilber Lourmat). The analysis was performed by Alexander Rill, supervised by Manuel Fisher and me.<sup>131</sup>

### **6.1.8 Analysing the phosphopantetheinylation by CoA-488**

Phosphopantetheinylation of the ACP from SerMSAS was investigated by its ability to accept a fluorescence marked acyl-moiety from CoA-488 (New England Biolabs), transferred by *Sfp*. The reaction mixture contained 4 µg protein, 1 mM DTT, 320 mM MgCl<sub>2</sub>, 10 µM CoA-488, 1 µM *Sfp* in the reaction buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA) and incubated for 60 minutes at 25°C. Subsequently 20 µL of laemmli loading buffer was added and heated up to 98°C for 5 minutes and gel electrophoresis was performed accordingly to chapter 6.1.3. Before treatment with Instant Blue, the fluorescence of the gel was analysed with Fusion SL (Vilber Lourmat) at the wavelength set to 488 nm. The holo-form cannot accept the fluorescence acyl-moiety and consequently full activation leads to absence of a fluorescence band. The apo-form was used as a positive control of the assay, because it can accept the fluorescence moiety and should have a fluorescence band in the gel.

### **6.1.9 Extraction of 6-MSA, 6-PrSA, 6-PSA and 4-hydroxy-6-alkyl-2-pyrone**

To validate the 6-alkyl salicylic acid production *in vivo* and *in vitro*, we extracted the cultivated medium (50 mL) or the reaction mixture (60-100  $\mu$ L) 2-3 times with ethyl acetate (10 mL for cultures and 400  $\mu$ L for *in vitro* reaction mixtures) after acidification with hydrochloric acid to pH 1. In contrast, reaction solutions for 4-hydroxy-6-alkyl-2-pyrone synthesis were not acidified, because of their prospected instabilities in acidic solution. The organic phases were combined and the solvent was removed. The residual was resolved in methanol (1000  $\mu$ L for cultures, 50  $\mu$ L for *in vitro* reaction mixtures), centrifugated (20817 g, 4°C) for five minutes and analysed with HPLC-MS using the Ultimate 3000 LC (Dionex) system with a Acquity UPLC BEH C18 column (2.1 x 50 mm, particle size 1.7  $\mu$ m, Waters) and AmaZonX (Bruker) for ESI. The column was equilibrated with 5% acetonitrile in water and samples were separated using a linear gradient from 5-95% within 16 min.

### **6.1.10 Mass spectroscopic analysis of ACPs**

To investigate the mass of ACPs, we changed the buffer to 100 mM ammonium phosphate (pH 7.5) by using 10 K 0.5 mL Amicon® centrifuge filter (Merck Millipore) and analysed ACPs with ESI in positive mode (SYNAPT G-2, waters). The sensitivity of the method is approximately 10 Da.

### **6.1.11 Characterization of SerChIB3 in its ability to transfer 6-MSA moieties from 6-MSA-SNAC to ChIB2**

We analysed SerChIB3 in its ability to transfer 6-MSA moieties from 6-MSA-SNAC to ChIB2. We synthesized 6-MSA-SNAC as a mimic of the SerMSAS-ACP(0) and incubated this substrate with SerChIB3 in the presence of SerChIB2 at room temperature. All reactions were performed in the assay buffer (50 mM sodium phosphate, pH 7.6, 450 mM sodium chloride, 20% glycerol, 1 mM EDTA) and concentrations of substrates and proteins are described in the appropriate figure. The negative control was performed in the absence of SerChIB3. After incubation, we changed the buffer to 100 mM ammonium phosphate (pH 7.5) by using 10 K 0.5 mL Amicon® centrifuge filter (Merck Millipore) and analysed SerChIB2 with ESI in positive mode (SYNAPT G-2, waters). To investigate kinetics, we precipitated the proteins by addition of 4-fold volume acetone (-20°C) to the reaction mixture at specific time points (0 min, 5 min, 10 min, 30 min, 60 min). Then SerChIB2 was resuspended in the ammonium acetate buffer and analysed with ESI.



### **6.1.12 Enzymatic activity *in vitro* (4-hydroxy-6-alkyl-2-pyrone assay)**

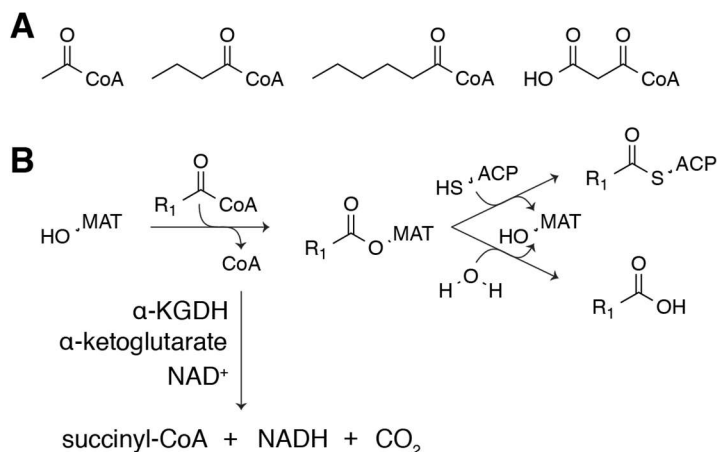
#### PenPaMSAS:

The enzymatic activity of PenPaMSAS in its ability to produce 4-hydroxy-6-alkyl-2-pyrone was analysed in UV-cuvette (UV-Cuvette micro, BRAND) with the Nanodrop 2000C (Thermo Scientific). Thereby the absorption of 4-hydroxy-6-alkyl-2-pyrone was analysed at 298 nm for ten minutes in the assay buffer (60 mM potassium phosphate, 0.1 mg/mL BSA and 10% glycerol). The final concentrations of substrates and enzymes are described in the respective figure legends. The adsorption was converted to concentration by using the TAL extinction coefficient for all products ( $2540 \text{ M}^{-1}\text{cm}^{-1}$ ).

### **6.1.13 Characterization of ATs**

#### Characterization of PenPaMSAS-AT

An enzyme coupled assay was used to investigate the AT-mediated transfer of acyl-moieties from CoA-ester to ACPs. During this reaction, CoA gets liberated, which is subsequently converted to succinyl-CoA by  $\alpha$ -KGDH and in the process NADH is generated, which was measured fluorometrically (figure 59). 384-well Small Volume HiBase Microplates (Greiner Bio-one) were used with 5  $\mu\text{L}$  of four 4-fold concentrated solutions in the assay buffer (60 mM sodium phosphate, pH 7.6, 10% glycerol, 0.1 mg/mL BSA). Solution one contained the respective CoA-ester. Solution two contained 8 mM  $\alpha$ -ketoglutaric acid, 1.6 mM TPP, 1.6 mM  $\text{NAD}^+$  and 60 mU/100  $\mu\text{L}$   $\alpha$ -KGDH. Solution three contained the respective ACP and solution four the AT with 0.4 mg/mL BSA. The enzyme concentration was adjusted to the turnover rates for the respective CoA-ester and final concentrations for the enzyme, ACP and CoA-ester are described in the figure legends. The reaction mixture was mixed with a pipette and equidistant kinetic measurement were taken for five minutes every 5-8 seconds at  $25^\circ\text{C}$  with the microplate reader (ClarioStar, BMG Labtech) and the settings: 348-20 nm, emission: 476-20 nm, 1500 gain, 11.9 mm focal, 17 flashed, orbital averaging off. The background measurement was either performed in the absence of ACP or in the presence of the knockout, indicated in the figure legend. The fluorescence was converted to concentration with a NADH calibration curve.



**Figure 59 |** Enzyme coupled assay for the characterisation of the AT in its ability to transfer CoA-ester. A) Acetyl-CoA, butyryl-CoA, hexanoyl-CoA and Mal-CoA was used as substrates for the characterisation of the PenPaMSAS AT. B) Transacylation of the AT follows a ping-pong-bi-bi mechanism. In the first step (ping-step) the AT gets acylated from CoA-Ester and CoA is liberated. In the second step (pong-step) the acyl moiety is transferred to the ACP or hydrolysed by water. Transacylation of AT can be tracked by an enzyme-coupled assay using  $\alpha$ -KGDH to convert liberated CoA to succinyl-CoA and  $\text{CO}_2$ . During this process NADH is generated from  $\text{NAD}^+$  and  $\alpha$ -Ketoglutarate. The increase of NADH can be measured fluorometrically.

### Characterization of mFAS MAT and DEBS M6-AT

Characterization of mFAS MAT and DEBS M6-AT was performed similar as described with the following changes. A different assay buffer was used (50 mM sodium phosphate, pH 7.6, 10% glycerol, 0.1 mg/mL BSA, 1 mM EDTA, filtered and degassed). The solution was pipetted in the order solution 1, solution 2 and solution 3, followed by manual mixing. Solution 4 was then added by injection, using the dispenser. The final concentration of the ingredients was slightly different (50 mM sodium phosphate, pH 7.6, 10% glycerol, 0.03 mg/mL BSA, 1 mM EDTA, 2 mM  $\alpha$ -ketoglutaric acid, 0.4 mM TPP, 0.4 mM  $\text{NAD}^+$  and 15 mU/100  $\mu\text{L}$   $\alpha$ -KGDH, 1-5 nM FAS MAT or 250 nM DEBS M6-AT, 10-400  $\mu\text{M}$  ACP, 0.1-130  $\mu\text{M}$  of the respective CoA-ester. The analysis was performed for 5 min with equidistant measurement every 5 seconds in technical replicates ( $n = 3$ ) and at least in biological replicates ( $n \geq 2$ ). The background was recorded in the absence of AT and subtracted from the actual measurement. The enzyme-mediated hydrolysis rates were relatively low compared to the transacylation and therefore not subtracted. The fluorescence was converted to concentration with a NADH calibration curve and initial velocities were determined for eight CoA-ester concentrations at four fixed ACP concentrations. All according data for the AT with the respective CoA-ester were fitted globally using equation 7 with the program OriginPro (version 8.5). The characterization was performed by Lara Maria Mayer, supervised by Alexander Rittner (Goethe-University, Frankfurt) and is shown for the sake of completion.

#### 6.1.14 Enzymatic activity for the synthesis of non-reduced TKLs

We established an assay to investigate turnover rates for the synthesis of non-reduced TKLs in the absence of NADPH, by monitoring the consumption of extender CoA-ester and the liberation of free CoA using HPLC-UV. All ingredients were prepared as 4-fold concentrated stocks in the assay buffer (400 mM phosphate buffer, 20% glycerol, 1 mM EDTA, 0.8% DMSO, pH 7.2). 20  $\mu$ L of each solution, containing 16  $\mu$ M enzyme (DEBS M6 construct), 20 mM natural diketide SNAC **2** and 800  $\mu$ M extender were mixed with 20  $\mu$ L assay buffer and incubated for 60 minutes at 25 °C. 40  $\mu$ L samples were taken after 0, 5, 10, 15, 20, 30 and 60 minutes and quenched with 5  $\mu$ L perchloric acid (70%) on ice. After mixing, the solution was directly neutralized with 5  $\mu$ L NaOH (10 mM). Then hydroxybutyryl-CoA was added to a final concentration of 50  $\mu$ M as an internal standard. The samples were spun (4°C, 20000 rcf, 10 min) and 30  $\mu$ L of the supernatant was transferred to HPLC-vials. 20  $\mu$ L of each sample were analysed with HPLC-UV on a Synchronis aQ-c18-column (4.6 x 250 mm, particle size: 5  $\mu$ M, ThermoFisher Scientific) using a Dionex UltiMate 3000 RSLC with a UV-detector. After equilibration of the column with 95% buffer A (200 mM ammonium acetate, pH 6.0) and 5% buffer B (methanol), the CoA-ester were separated using a linear gradient from 5-20% B over 15 minutes and an increased gradient from 20-60% B over 8 minutes. After a final wash step at 60% B for 3 minutes, the column was re-equilibrated for the next sample. The absorbance of the CoA-ester were detected at 260 nm and peak areas were converted into concentration by using the internal standard. The concentration of the CoA-ester and free CoA were plotted against the reaction time and fitted exponentially with OriginPro 8.5 (OriginLab, USA) using equation 4. The turnover rates were obtained from the initial velocities of the function by its slope at 0 min.

#### 6.1.15 Verification of TKLs and macrolactones using HPLC-MS

To verify TKLs and macrolactones, we extracted all reaction mixtures with ethyl acetate (2 x 400  $\mu$ L for TKLs, 3 x 300  $\mu$ L for macrolactones). The organic phases were combined, and the solvent removed *in vacuo* using a SpeedVac. The samples were dissolved with 50  $\mu$ L methanol and transferred into HPLC vials after centrifugation (20 min, 20000 rcf, 4°C). HPLC-MS measurements were performed with the Ultimate 3000 LC system (Dionex) on a Acquity UPLC BEH C18 column (2.1 x 50 mm, particle size: 1.7  $\mu$ m, Waters). The separation was performed with a linear gradient from 5% acetonitrile in water to 95% within 16 minutes. ESI measurements were performed on a AmaZonX or Impact II qTof (HRMS).

#### 6.1.16 Analysis and purification of macrolactones using HPLC-UV

HPLC-UV was performed to analyse macrolactone synthesis and for the first purification step for compound **17**. We used the Dionex UltiMate 3000 RS UHPLC

system with a RS Diode Array UV detector. The separation was performed on a Synchronis aQ-c18 column (4.6 x 250 mm, particle size: 5  $\mu$ m, ThermoFisher Scientific) with a linear gradient from 5% to 95% acetonitrile in water within 15 minutes and another 5 minutes at 95% acetonitrile. A linear gradient was used to change the mobile phase back to 5% acetonitrile within 1 minute and the column was re-equilibrated for five minutes. We performed fractionation with the UltiMate 3000 Autosampler within the elution time of 3-16 minutes with 20 seconds per fraction.

### 6.1.17 Blue native PAGE

Blue native PAGE was performed by Martin Schwalm, supervised by Alexander Rittner and is shown for the sake of completion. Native PAGE was performed with 3-12 % Bis-Tris gels (Thermo Fisher, Invitrogen) in an Invitrogen chamber with the anode buffer (50 mM Bis-Tris, pH 7.0) and cathode buffer (50 mM tricine, 15 mM Bis-Tris, pH 7.0, 0.02% (w/v) Coomassie G-250). Typical 2-10  $\mu$ g of protein was prepared in the loading buffer (5 mM Bis-Tris, 50 mM  $\epsilon$ -aminocaproic acid, 4% glycerol, pH 7.0) and applied on the gel. After 1 h electrophoresis at 100 V, the cathode buffer was changed to cathode buffer II (50 mM tricine, 15 mM Bis-Tris, pH 7.0) and electrophoresis was performed 2-2.5 hours at 200 V. The gel was stained for 5 minutes in staining solution (0.4% Coomassie G-250, 50% MeOH, 10% acetic acid). Then the gel was destained several times with destaining solution (10% EtOH, 10% acetic acid). The nativeMark (Thermo Fisher Scientific) ladder was used as reference.

## 6.2 Chemical synthesis and biosynthesis

All reactions were stirred magnetically and analysed by thin layer chromatography, using 0.25 mm silica gel plates (60F-254, Merck) with UV light and  $\text{KMnO}_4$  (with heat) for visualization. For chromatographic separation, we used silica gel (60, particle size: 0.040-0.060 mm, Merck). Mass spectra were obtained by electrospray ionization on Surveyor (ThermoFisher, HRMS on MSQ). For NMR, samples were solved in deuterated solvents (Deutero GmbH) and analysed with Bruker DPX250, Bruker AV400, Bruker AV500 or Bruker DRX600 and chemical shifts are expressed in parts per million (ppm) relative to methanol ( $\delta$  3.31), chloroform ( $\delta$  7.27), DMSO ( $\delta$  2.51) or water ( $\delta$  4.79). Multiplicities are described with the abbreviations: s = singlet, d = doublet t = triplet, q = quartet, m = multiplet, dd = doublet of doublet, dq = doublet of quartet, ddd = doublet of doublet of doublet. Chemicals were received from Merck, Carl-Roth and abcr GmbH and used without further purification. All reactions were performed under Argon, except when water was used as solvent.

### 6.2.1 Synthesis of 6-MSA-SNAC

6-MSA (500 mg, 3.29 mmol), DMAP (100 mg, 0.822 mmol), EDC×HCl (750 mg, 4.80 mmol) and *N*-acetylcysteamine (379  $\mu$ L, 3.84 mmol) was dissolved in 42 mL DCM and stirred for 24 hours at 25°C. The reaction mixture was washed three times with brine and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated *in vacuo* and the crude product was purified by column chromatography (silico gel 60, 0.040-0.060 mm particle size, Merck) with a CHCl<sub>3</sub>/MeOH gradient (50:1, 40:1, 30:1, 20:1). The synthesis was performed during a bachelor thesis by Christian Gusenda.<sup>136</sup>

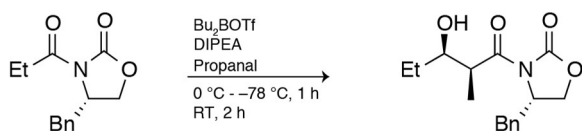
**Yield:** 22 % (182 mg), colourless solid

**R<sub>F</sub>-value** (TLC; chloroform/methanol: 20/1): 0.71

**<sup>1</sup>H-NMR** (250 MHz, methanol-4*d*):  $\delta$  7.13 (t, <sup>3</sup>J = 7.5 Hz, 1H), 6.72-6.67 (m, 2H), 3.46 (t, <sup>3</sup>J = 7.5 Hz, 2H), 3.20 (t, <sup>3</sup>J = 7.5 Hz, 2H), 2.26 (s, 3H), 1.96 (s, 3H)

**MS** (ESI-): found 252.23, for [M-H]<sup>-</sup> calculated 252.07

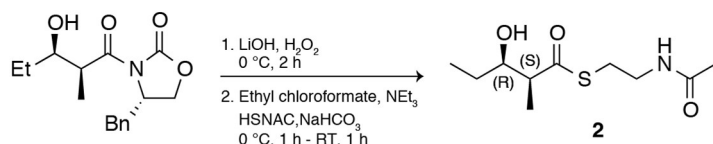
### 6.2.2 Synthesis of natural diketide SNAC (2)



(4*S*, 2'*S*, 3'*R*)-3-(2'-Methyl-3'-hydroxypentanoyl)-4-benzyl-2-oxazolidinone was synthesized as described by Sharma *et al.*<sup>200</sup> Briefly an Evans auxiliary was used for an aldol-condensation with propanal. As the product was not clean after column chromatography, it was further purified by trituration with EtOAc and n-hexane.

**Yield:** 69 % (4.3 g), white crystals.

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.37-7.21 (m, 5H), 4.75-4.69 (m, 1H), 4.27-4.19 (m, 2H), 3.90-3.86 (m, 1H), 3.81 (dq, *J* = 2.7, 7.0 Hz, 1H), 3.27 (dd, *J* = 3.4, 13.5 Hz, 1H), 2.81 (dd, *J* = 9.4, 13.4 Hz, 1H), 1.65-1.43 (m, 2H), 1.27 (d, *J* = 7.0 Hz, 3H), 0.99 (t, *J* = 7.5 Hz, 3H) ppm.



In the next step (2S, 3R) 3-hydroxy-2-methylpentanoyl-S-N-acetylcysteamine thioester (**2**) was synthesized by two steps following the routes by Sharma *et al.* and Peter *et al.*<sup>200,201</sup> After aldol condensation, the Evans auxiliary was almost quantitatively hydrolysed with LiOH and H<sub>2</sub>O<sub>2</sub> to obtain the acid. The acid (220 mg, 1.66 mmol) was dissolved in 5 mL THF and cooled to 0 °C. To the stirring solution, 148  $\mu$ L ethyl chloroformate (168 mg, 1.55 mmol) and 215  $\mu$ L NEt<sub>3</sub> (157 mg, 1.55 mmol) were added and further stirred for 45 minutes at 0 °C. Then, a solution of 224  $\mu$ L N-acetylcysteamine (251 mg, 2.00 mmol) and 54 mg NaHCO<sub>3</sub> in 2 mL H<sub>2</sub>O were added and the reaction mixture was stirred for 1 hour at room temperature. After extraction with EtOAc (3 x 20 mL), combined organic phases were dried with anhydrous MgSO<sub>4</sub> and filtered. The solvent was removed in vacuo and the crude product was purified by column chromatography (gradient: DCM:Acetone 10:1, 8:1, 6:1, 4:1).

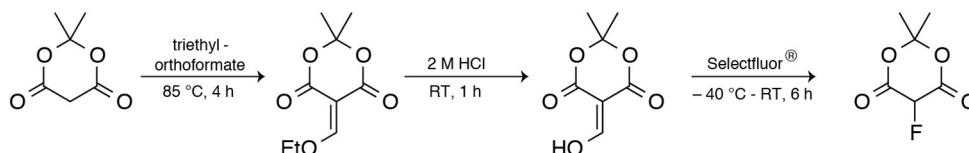
**Yield:** 33 % (128 mg), clear liquid.

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 5.76 (s, br, 1H), 3.88-3.83 (m, 1H), 3.53-3.41 (m, 2H), 3.10-2.99 (m, 2H), 2.78-2.72 (m, 1H), 1.98 (s, 3H), 1.57-1.43 (m, 2H), 1.23 (d,  $J$  = 7.0 Hz, 3H) 0.99 (t,  $J$  = 7.4 Hz, 3H) ppm.

**MS** (ESI<sup>+</sup>): found 234.12, for [M+H]<sup>+</sup> calculated 234.12; found 256.24, calculated for [M+Na]<sup>+</sup> 256.10.

### 6.2.3 Synthesis of F-Mal-CoA

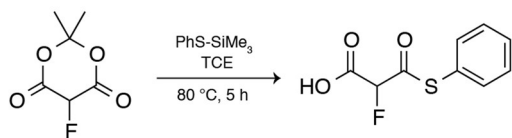
Synthesis of F-Mal-CoA was performed by Alexander Rittner (Goethe-University, Frankfurt) and is shown for the sake of completion.



Firstly, we used a described method by Saadi and Wennemers, to synthesize fluoro-Meldrum's acid in three steps.<sup>178</sup>

**Yield** (over three steps): 44 % (5.1 g), colourless solid.

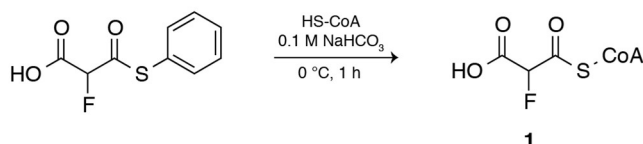
**<sup>1</sup>H-NMR** (400 MHz, C<sub>2</sub>D<sub>6</sub>OS):  $\delta$  = 6.69 (d,  $J$  = 41.6 Hz, 1H), 1.82 (s, 3H), 1.71 (s, 3H) ppm.



Then, we used the fluoro-Meldrum's acid with trimethylsilylthiophenol to obtain the fluoromalonoyl-thiophenyl half-ester by using the butyllithium method described in *Saadi and Wennemers*.<sup>178</sup>

**Yield:** 53 % (0.7 g), yellowish powder.

**<sup>1</sup>H-NMR** (400 MHz, C<sub>2</sub>D<sub>6</sub>OS):  $\delta$  = 7.64-7.50 (m, 3H), 7.49-7.46 (m, 2H), 5.95 (d,  $J$  = 47.2 Hz, 1H) ppm.



In the last step we used the fluoromalonoyl-thiophenyl half-ester for transacylation of free CoA with the fluoromalonate moieties to yield F-Mal-CoA. We adapted a method reported by Dunn et al.<sup>179</sup> Briefly, free coenzyme A (CoA) (30 mg, 38  $\mu$ mol) and fluoromalonoyl thiophenyl half-ester (44 mg, 205  $\mu$ mol) were dissolved in 1 mL NaHCO<sub>3</sub> at 0°C and the pH was adjusted to 8-9 with 1 mL ice cold 0.2 M NaOH. The reaction mixture was stirred for 1 hour at 0°C and then transferred into a 50 mL falcon tube, containing 15 mL acetone, cooled to -20°C prior to usage. After centrifugation (10 min, 3000 rcf, 4°C), the supernatant was discarded, and the pellet was washed with 1-2 mL acetone (-20°C). After another centrifugation step (10 min, 3000 rcf, 4°C), the pellet was dried in vacuo and then dissolved in water to determine the yield by its absorbance at 260 nm with the extinction coefficient 16400 M<sup>-1</sup>cm<sup>-1</sup>.<sup>202</sup>

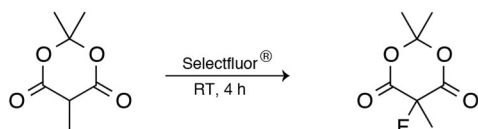
**Yield:** 81 %, white crystals.

**<sup>1</sup>H-NMR** (400 MHz, D<sub>2</sub>O):  $\delta$  = 8.54 (s, 1H), 8.25 (s, 1H), 6.18-6.13 (m, 1H), 5.33 (d,  $J$  = 50.3 Hz, 1H), 4.83-4.79 (m, 2H), 4.58-4.53 (m, 1H), 4.25-4.17 (m, 1H), 3.99 (s, 1H), 3.83-3.76 (m, 1H), 3.55-3.47 (m, 1H), 3.47-3.39 (m, 2H), 3.37-3.30 (m, 2H), 3.10-3.03 (m, 2H), 2.47-2.37 (m, 2H), 0.85 (s, 3H), 0.70 (s, 3H) ppm.

**HRMS** (ESI+): found 872.1138, calculated for  $[M+H]^+$  872.1141.

### 6.2.4 Synthesis of F-MM-CoA

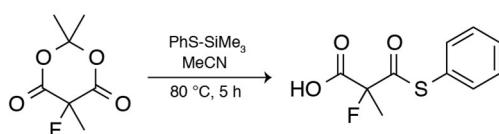
Synthesis of F-MM-CoA was performed by Alexander Rittner (Goethe-University, Frankfurt) and is shown for the sake of completion. The synthesis was performed with a similar strategy as for the F-Mal-CoA synthesis.



In the first step, methyl-Meldrum's acid was fluorinated with Selectfluor<sup>®</sup>. Methyl-Meldrum's acid (10 g, 61.3 mmol) and Selectfluor<sup>®</sup> (25.2 g, 67.5 mmol) were dissolved in 500 mL anhydrous acetonitrile and the reaction mixture was stirred for 4.5 hours at room temperature. After evaporation of the solvent *in vacuo*, 200 mL DCM and 100 mL water were added. Then phases were separated, and the aqueous phase was extracted four times with 75 mL DCM. The organic phases were combined, washed with brine, dried with anhydrous MgSO<sub>4</sub> and filtered. After evaporation of the solvent *in vacuo*, the crude product was purified by column chromatography (gradient: hexane:ethylacetate 10:1, 5:1, 3:1, 2:1, 1:1, 1:2, 1:3)

**Yield:** 9 g (83 %, 96% purity), white crystals.

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 1.93 (d,  $J$  = 24 Hz, 3H), 1.85 (s, 3H), 1.80 (s, 3H) ppm. **<sup>19</sup>F-NMR** (300 MHz, CDCl<sub>3</sub>):  $\delta$  = -153.50 (q,  $J$  = 24 Hz) ppm.



In the next step, we used the 5-fluoro-5-methyl-Meldrum's acid with trimethylsilylthiophenol to obtain the 2-fluoro-2-methylmalonyl-thiophenyl halfester by using the butyllithium method described in *Saadi and Wennemers*.<sup>178</sup> 5-Fluoro-5-methyl-Meldrum's acid (1.95 g, 11.1 mmol) and trimethylsilylthiophenol (2.4 g, 2.5 mL, 13.2 mmol) were dissolved in 10 mL anhydrous acetonitrile and the reaction mixture was stirred for 5 hour under reflux and argon atmosphere. After evaporation of the solvent *in vacuo*, 10 mL ice cold saturated NaHCO<sub>3</sub> was added, followed by 10 mL water and 80 mL DCM. Phases were separated and the aqueous phase was extracted three times with 30 mL DCM. The organic phases were

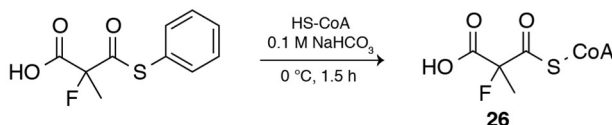


combined, dried over anhydrous  $\text{MgSO}_4$  and filtered. After evaporation of the solvent *in vacuo*, the crude product was purified by column chromatography (gradient: hexane:ethylacetate 4:1+0.5% AcOH, 2:1+0.5% AcOH)

**Yield:** 220 mg (9 %), colourless oil.

**$^1\text{H-NMR}$**  (250 MHz, MeOD):  $\delta = 7.48\text{-}7.41$  (m, 5H), 1.77 (d,  $J = 22$  Hz, 3H) ppm.

**$^{19}\text{F-NMR}$**  (300 MHz,  $\text{CDCl}_3$ ):  $\delta = -156.4$  (q,  $J = 22$  Hz) ppm



In the last step we used the 2-fluoro-2-methylmalonyl-thiophenyl half-ester for transacylation of free CoA with the 2-fluoro-2-methylmalonate moieties to yield F-MM-CoA. We adapted a method reported by Dunn et al.<sup>179</sup> Briefly, free CoA (200 mg, 255  $\mu\text{mol}$ ) and 2-fluoro-2-methylmalonyl-thiophenyl half-ester (200 mg, 876  $\mu\text{mol}$ ) were dissolved in 1 mL  $\text{NaHCO}_3$  at  $0^\circ\text{C}$  and the pH was adjusted to 8-9 with 1 mL ice cold 0.2 M NaOH. The reaction mixture was stirred for 1.5 hours at  $0^\circ\text{C}$  and then transferred into a 50 mL falcon tube, containing 20 mL acetone, cooled to  $-20^\circ\text{C}$  prior to usage. After centrifugation (10 min, 3000 rcf,  $4^\circ\text{C}$ ), the supernatant was discarded, and the pellet was washed with 1-2 mL acetone ( $-20^\circ\text{C}$ ). After another centrifugation step (10 min, 3000 rcf,  $4^\circ\text{C}$ ), the pellet was dried *in vacuo* and then dissolved in water to determine the yield by its absorbance at 260 nm with the extinction coefficient  $16400 \text{ M}^{-1}\text{cm}^{-1}$ .<sup>202</sup>

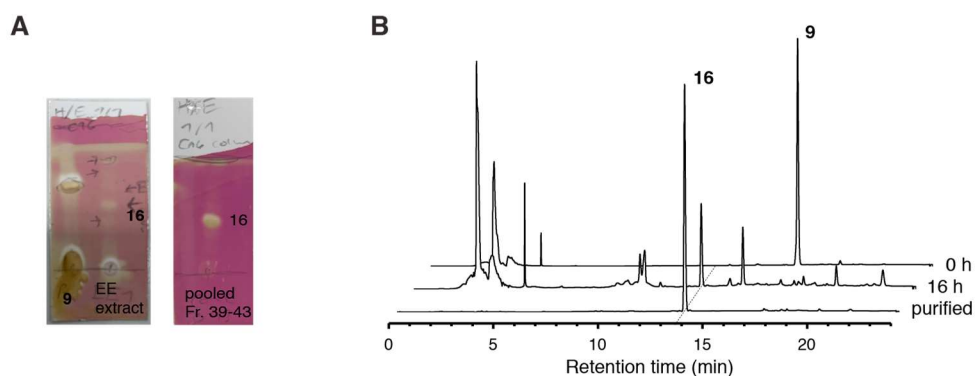
**Yield:** 93 %, white crystals

**$^1\text{H-NMR}$**  (300 MHz,  $\text{D}_2\text{O}$ ):  $\delta = 8.57$  (s, 1H), 8.28 (s, 1H), 6.19 (d,  $J = 6.9$  Hz 1H), 4.62-4.56 (m, 1H), 4.28-4.21 (m, 2H), 4.02 (s, 1H), 3.86-3.80 (m, 1H), 3.58-3.52 (m, 1H), 3.48-3.44 (m, 2H), 3.40-3.33 (m, 2H), 3.09-3.03 (m, 2H), 2.48-2.42 (m, 2H), 1.71 (d,  $J = 22$  Hz, 3H), 0.88 (s, 3H), 0.73 (s, 3H) ppm.  **$^{19}\text{F-NMR}$**  (300 MHz,  $\text{D}_2\text{O}$ ):  $\delta = -146.7$  (q,  $J = 22$  Hz, 1H) ppm.

### 6.2.5 Biosynthesis and analysis of compound 16

We scaled-up the reaction for structural analysis of compound **16** with NMR. We used a total volume of 10 mL with the final concentrations of 300  $\mu\text{M}$  pentaketide substrate **13** (1.05 mg, dissolved in DMSO), 400  $\mu\text{M}$  Mal-CoA and 500  $\mu\text{M}$  NADPH

were dissolved in the buffer (250 mM potassium phosphate, 10 % glycerol, pH 7). Finally, 5  $\mu$ M **H1** (9.16 mg, dissolved in 250 mM potassium phosphate, 10 % glycerol, pH 7) was added to the reaction mixture, followed by careful mixing and incubation for 16 hours at 25°C. To monitor the progress of the reaction, 20  $\mu$ L of the reaction mixture was taken and enzyme was precipitated with 60  $\mu$ L methanol. After centrifugation (20000 rcf, 4°C, 10 min) supernatant was injected into the HPLC-UV system. After overnight incubation, the reaction mixture was transferred to a 50 mL falcon tube and extracted 5 times with 10 mL ethyl acetate. Phases were separated by centrifugation (1-5 min, 3000°rcf) and the organic phases were combined, and the solvent was removed *in vacuo*. The crude product was furthermore dried by azeotropic evaporation with toluene and then adsorbed to silica gel for chromatographic purification (gradient: hexane:EtOAc 20:1, 10:1, 7:1, 5:1, 3:1, 2:1, 1:1).



**Figure 60** | Biosynthesis and analysis of compound **16**. (A) Analysis of extraction and purification procedure using TLC (hexane:EtOAc 1:1, staining with  $\text{KMnO}_4$  and heat). (B) Monitoring the progress of the reaction and analysis of the purification procedure using HPLC-UV.

**Yield:** ~500  $\mu$ g (~59%), white solid

**R<sub>F</sub>-value** (hexane:EtOAc: 1:1): 0.44

**Retention time** (HPLC; water/ACN): 13.74 min

**$^1\text{H}$  NMR** (600 MHz, Chloroform-*d*)  $\delta$  6.76 (dd,  $J = 15.7, 5.4$  Hz, 1H), 6.44 (dd,  $J = 15.8, 1.2$  Hz, 1H), 5.02 (ddd,  $J = 8.3, 6.0, 2.3$  Hz, 1H), 4.05-4.02 (ddd,  $J = 11.2, 5.2, 1.2$  Hz, 1H), 2.71-2.54 (m, 4H), 1.74-1.56 (m, 4H), 1.30-1.28 (m, 1H), 1.25 (d,  $J = 7.0$  Hz, 3H), 1.13 (d,  $J = 6.9$  Hz, 3H), 1.03 (d,  $J = 6.2$  Hz, 3H), 0.93 (t,  $J = 7.4$  Hz, 3H) ppm.  **$^{13}\text{C}$  NMR** (500 MHz, Chloroform-*d*):  $\delta$  170.3, 147.5, 125.5, 74.5, 72.9, 45.3, 39.0, 38.1, 33.0, 32.5, 25.0, 17.7, 17.2, 10.3, 9.4 ppm.

**MS (HR-ESI<sup>+</sup>)** found 283.1900, for  $[\text{M}+\text{H}]^+$  calculated 283.1910; found 305.1720, calculated for  $[\text{M}+\text{Na}]^+$  305.1729; found 265.1796, calculated for  $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$  265.1804.

### 6.2.6 Biosynthesis and analysis of compound 17/25

We scaled-up the reaction for structural analysis of compound **17** with NMR. We used a total volume of 50 mL with the final concentrations of 600  $\mu\text{M}$  pentaketide substrate **13** (10.46 mg, dissolved in DMSO), 4000  $\mu\text{M}$  F-Mal-CoA and 1000  $\mu\text{M}$  NADPH were dissolved in the buffer (250 mM potassium phosphate, 10 % glycerol, pH 7). Finally, 10  $\mu\text{M}$  **H1** (90.57 mg, dissolved in 250 mM potassium phosphate, 10 % glycerol, pH 7) was added to the reaction mixture, followed by careful mixing and incubation for 22.5 hours at 25°C. To monitor the progress of the reaction, 20  $\mu\text{L}$  of the reaction mixture was taken and enzyme was precipitated with 60  $\mu\text{L}$  methanol. After centrifugation (20000 rcf, 4°C, 10 min) supernatant was injected into the HPLC-UV system. After overnight incubation, the reaction mixture was transferred to two 50 mL falcon tube and extracted 5 times with 20 mL ethyl acetate. Phases were separated by centrifugation (1-5 min, 3000°rcf) and the organic phases were combined and the solvent was removed *in vacuo*. The crude product was dissolved in ~500  $\mu\text{L}$  MeOH:CHCl<sub>3</sub> (9:1) and purified with the uHPLC system (Thermo Fisher) on a Synchronis aQ-C18-LC column as described in the method section. Based on <sup>19</sup>F-NMR analysis, all fractions containing the major fluorinated compound were combined. The solvent was removed *in vacuo* and the product was further dried by azeotropic evaporation with toluene. Then the product was adsorbed to silica for another chromatographic purification (gradient: hexane:EtOAc 20:1, 10:1, 7:1, 5:1, 3:1, 2:1, 1:1). We received compound **25** instead of compound **17** as described in chapter 2.4.3.

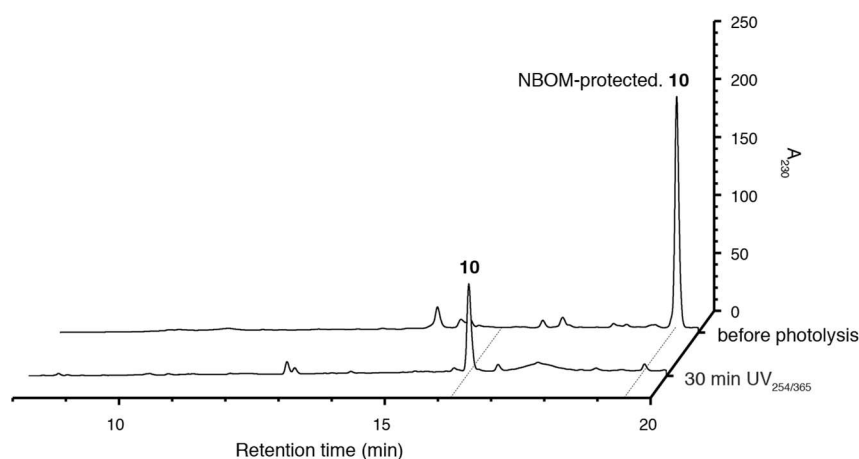
**Retention time** (HPLC; water/ACN): 13.23 min

**<sup>1</sup>H NMR** (500 MHz, Chloroform-*d*):  $\delta$  5.70 (dd,  $J = 15.6, 8.0$  Hz, 1H), 5.51 (dd,  $J = 15.7, 0.8$  Hz, 1H), 4.84 (d,  $J = 48.4$  Hz, 1H), 3.41 (ddd,  $J = 9.0, 5.5, 3.7$  Hz, 1H), 2.54-2.47 (m, 1H), 2.41-2.34 (m, 1H), 2.07-2.00 (m, 1H), 1.92-1.87 (m, 1H), 1.66 (q,  $J = 12.8$  Hz, 1H), 1.61-1.55 (m, 1H), 1.44-1.35 (m, 1H), 1.11 (d,  $J = 6.4$  Hz, 3H), 1.08 (d,  $J = 6.8$  Hz, 3H), 0.98-0.95 (ovlp m, 6H). **<sup>19</sup>F NMR** (500 MHz, Chloroform-*d*):  $\delta$  -206.21(d,  $J = 46,8$  Hz). **<sup>13</sup>C NMR** from HSQC (500 MHz, Chloroform-*d*):  $\delta$  134.1, 131.5, 94.8/96.4, 76.5, 42.4, 42.2, 38.7, 36.9, 27.0, 13.4, 15.2, 10.4, 13.6

**MS (HR-ESI+)**: found 273.1859, for  $[\text{M}+\text{H}]^+$  calculated 273.1866; found 295.1677, calculated for  $[\text{M}+\text{Na}]^+$  295.1686; found 255.1754, calculated for  $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$  255.1761; found 567.3462, calculated for  $[\text{M}_2+\text{Na}]^+$  567.3474

### 6.2.7 Biosynthesis and analysis of compound **21**

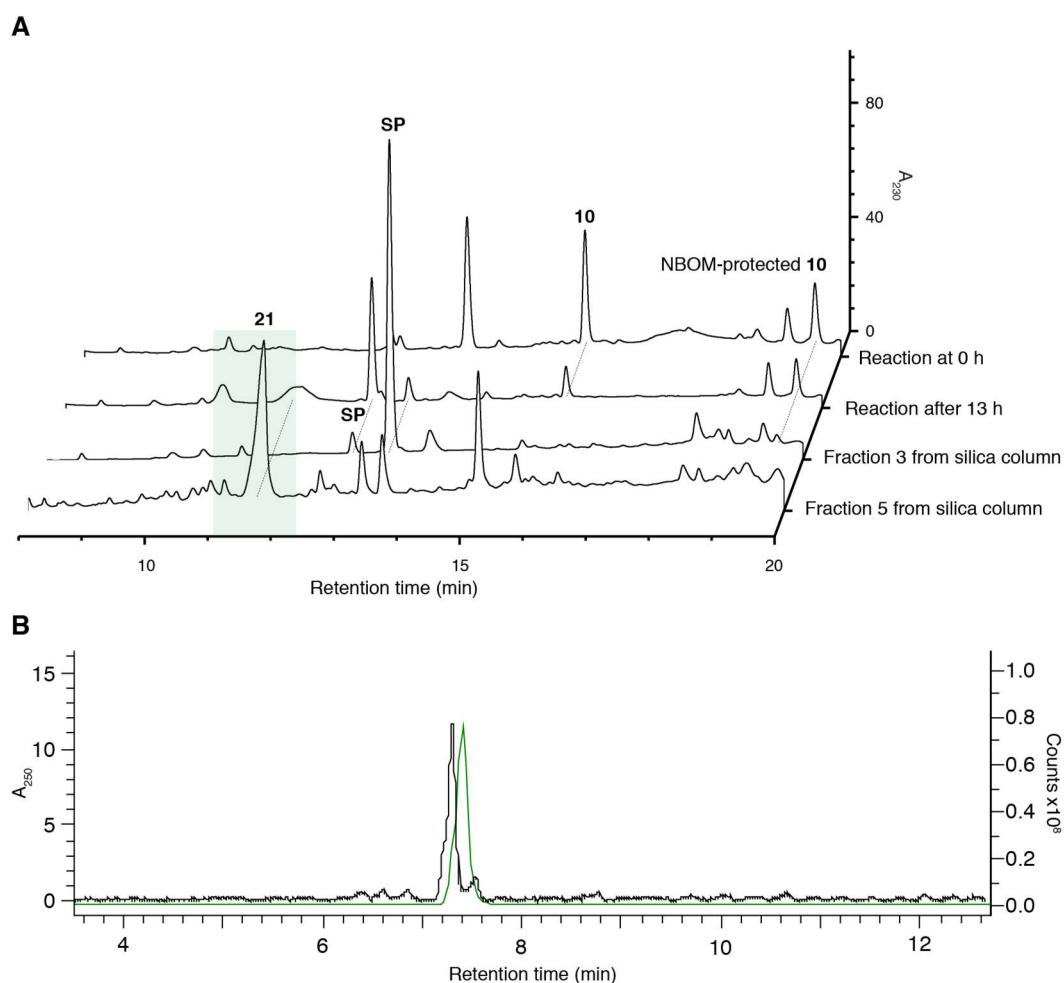
We scaled-up the reaction for structural analysis of compound **21** with NMR. In the first step, the hexaketide educt **10** was generated from the NBOM-protected substrate by photolysis as described by Hansen *et al.*<sup>199</sup> Final concentrations of the reaction mixture were 1.2 mM NBOM-protected substrate (3.6 mg), 25 mM ascorbic acid and 1 mM sodium metabisulfite. All ingredients were dissolved in 5 mL water and irradiated for 30 min under UV light (260 nm/395 nm). We monitored the deprotection by HPLC, demonstrating that the NBOM-protection group was almost quantitatively removed after 30 min.



**Figure 61** | Generation of educt **10**, by photolysis of the NBOM-protection group. HPLC-UV analysis showed that the protection group was almost quantitatively removed after 30 min.

The reaction mixture of the hexaketide substrate **10** was transferred into a 50 mL falcon tube. A total volume of 20 mL was used for the biosynthesis of compound **21** with the final concentrations of 300  $\mu$ M hexaketide substrate **10**, 2 mM F-Mal-CoA and 10  $\mu$ M **H1** in the assay buffer (168 mM potassium phosphate, 6.7 % glycerol, 6.25 mM ascorbic acid, 0.25 mM sodium bisulfite, pH 7). The reaction mixture was carefully mixed and incubated for 16 hours at 25°C. To monitor the progress of the reaction, 20  $\mu$ L of the reaction mixture was taken and enzyme was precipitated with 60  $\mu$ L methanol. After centrifugation (20000 rcf, 4°C, 10 min), supernatant was injected into the HPLC-UV system. After overnight incubation, the reaction mixture was transferred to a 50 mL falcon tube and extracted 6 times with 25 mL ethyl acetate. Phases were separated by centrifugation (1-5 min, 3000°rcf) and the organic phases were combined, and the solvent was removed *in vacuo*. The crude product was further dried by azeotropic evaporation with toluene. As the extraction with EtOAc was not sufficient, we repeated the extraction procedure after acidifying the reaction solution to pH 1-2. Then the product was adsorbed to silica for chromatographic purification (gradient: hexane:EtOAc 20:1, 10:1, 7:1, 5:1, 3:1, 2:1, 1:1, 1:2, EtOAc + 0.5% TFA). Compound **21** was eluted in the last fraction, containing EtOAc and 0.5% TFA. The product fraction was pooled, filtered and the

solvent was removed *in vacuo*. Product **21** further purified with the uHPLC system (Thermo Fisher) on a Synchronis aQ-C18-LC column, as described in the method section. Fractions, containing compound **21** (11-12.33 min) were collected and the purified product was shown by HPLC-MS. Unfortunately, the amount was not enough for structural analysis by NMR. Furthermore, we found an unknown side product during chromatographic purification, which likely originates from an elongation of educt **10** with F-Mal-CoA (figure 62).



**Figure 62** | Biosynthesis and purification of compound **21**. (A) The reaction and purification were monitored by HPLC-UV, as described in the method section. The green bar indicates the collected fractions of **21**. During chromatographic purification we found an unknown side product (SP). (B) HPLC-MS analysis showed the successful synthesis and purification of compound **21**. UV absorbance at 250 nm is shown by the black line and MS analysis by the green line [EIC: **21** [M+Na]<sup>+</sup> m/z = 379.13].

**Retention time** (HPLC; water/ACN): 11.5-12.2 min

**MS (HR-ESI+):** found 357.2072, for  $[M+H]^+$  calculated 357.2078; found 379.1887, calculated for  $[M+Na]^+$  379.1897; found 339.1965, calculated for  $[M-H_2O+H]^+$  339.1972.

### 6.2.8 Biosynthesis and analysis of compound 27

We scaled-up the reaction for structural analysis of compound **27** with NMR. We used a total volume of 50 mL with the final concentrations of 600  $\mu$ M pentaketide substrate **13** (10.46 mg, dissolved in DMSO), 800  $\mu$ M F-MM-CoA and 1000  $\mu$ M NADPH were dissolved in the buffer (250 mM potassium phosphate, 10 % glycerol, pH 7). Finally, 7  $\mu$ M **H1** (63.40 mg, dissolved in 250 mM potassium phosphate, 10 % glycerol, pH 7) was added to the reaction mixture, followed by careful mixing and incubation for 16 hours at 25°C. After overnight incubation, the reaction mixture was transferred to two 50 mL falcon tube and extracted 5 times with 20 mL ethyl acetate. Phases were separated by centrifugation (1-5 min, 3000<sup>o</sup>rcf) and the organic phases were combined, and the solvent was removed *in vacuo*. The crude product was furthermore dried by azeotropic evaporation with toluene and then adsorbed to silica gel for chromatographic purification (gradient: hexane:EtOAc 20:1, 10:1, 7:1, 5:1, 3:1, 2:1, 1:1).

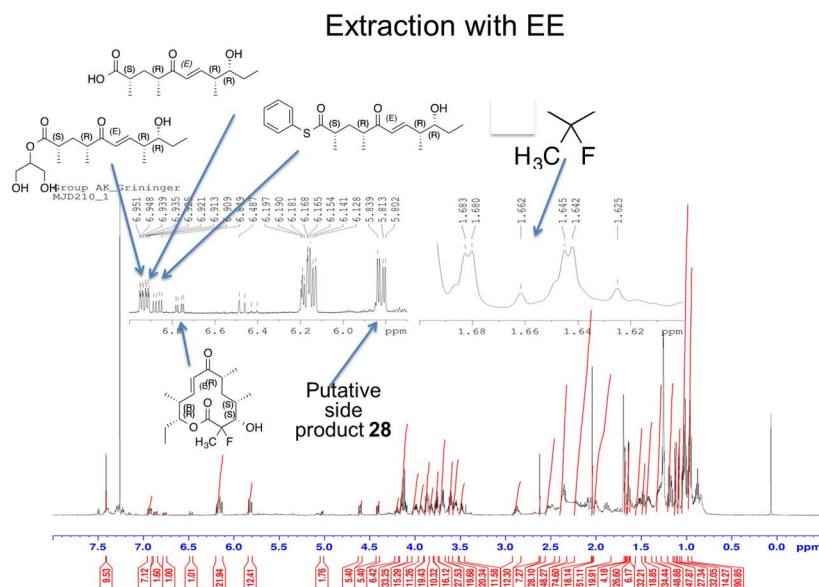
To improve reaction conditions, we performed 20 mL reactions with the final concentrations 300  $\mu$ M pentaketide substrate **13** (2.09 mg, dissolved in DMSO), 600  $\mu$ M F-MM-CoA, 500  $\mu$ M NADPH and 8.3  $\mu$ M **H1** (30.07 mg, dissolved in 250 mM potassium phosphate, 10 % glycerol, pH 7) dissolved in the assay buffer (250 mM potassium phosphate, 10 % glycerol, pH 6.8-7.3). We changed the pH value to 6.8 and 7.3 and increased the reaction temperature to 37°C for the pH value of 7.0. With <sup>1</sup>H-NMR as analytics and the focus on the alkene group, we could easily detect 5 different compounds which are associated with the educt **13** (figure 63). We identified the educt, the hydrolysed educt, the hydrolysed and esterified educt (glycerol ester), the product and an unknown side product **28**. Furthermore, a focus on the putative MeFC-group in the <sup>1</sup>H-NMR, but also the <sup>19</sup>F-NMR confirmed the product and the formation of at least one fluorinated side product, which most likely originates from an elongation of educt **13** with F-MM-CoA (figure 63-64). We performed additional HPLC-HRMS analytics which confirmed all described compounds (data not shown). With the assumption that we extracted all compounds, associated with the educt **13**, we can calculate the yield of the three different reactions. The analytics demonstrated that lower pH values (6.8 compared to 7.1) improved the reaction slightly from 5% to 6%, while an increased temperature from about 23-25°C to 37°C led to a significant drop of the yield to 2.4% (pH 7.1), most likely due to the lack of enzyme stability at this temperature (data not shown). Lastly, it is worth to mention that the putative side product was the main product within the reaction with a yield of more than 50%.

**Yield:** 6% (estimated on NMR analytics)

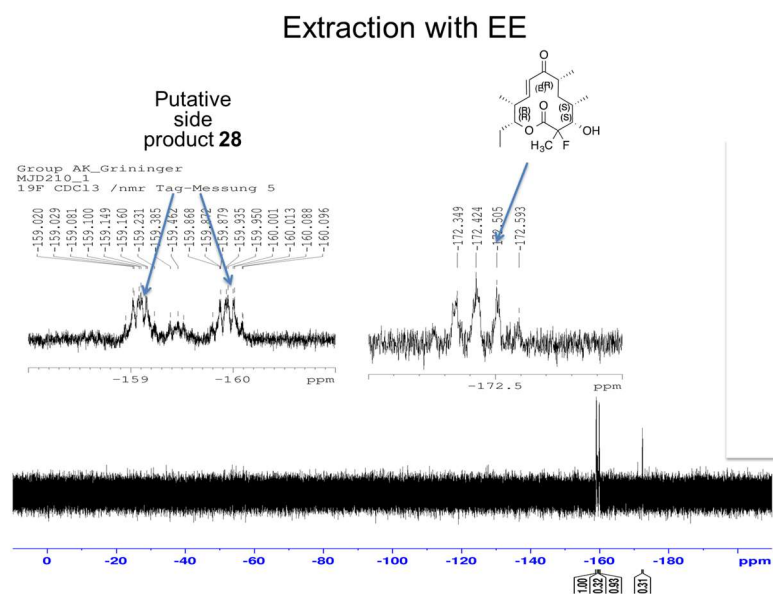
**Retention time (HPLC; water/ACN):** 14.65 min

**$^1\text{H}$  NMR** (500 MHz, Chloroform-*d*):  $\delta$  6.78 (dd,  $J = 15.8, 5.4$  Hz, 1H), 6.49 (dd,  $J = 15.8, 1.2$  Hz, 1H), 5.09 (ddd, 8.0, 5.4, 1.8 Hz, 1H), 3.68 (dd,  $J = 26.3, 0.9$  Hz, 1H), 2.73-2.66 (m, 1H), 2.57-2.48 (m, 1H), 2.03-1.98 (m, 1H), 1.85-1.77 (m, 1H), 1.69-1.62 (m, 1H), 1.66 (d,  $J = 22$  Hz, 3H), 1.34-1.25 (m, 2H), 1.22 (d,  $J = 7$  Hz, 3H), 1.19 (d,  $J = 6.9$  Hz, 3H), 1.04 (d,  $J = 6.7$  Hz, 3H), 0.96 (t,  $J = 7.3$  Hz, 3H)  **$^{19}\text{F}$  NMR** (500 MHz, Chloroform-*d*):  $\delta$  -172.45 (q,  $J = 22,5$  Hz)  **$^{13}\text{C}$  NMR** from HSQC:  $\delta$  146.0, 125.9, 78.2, 75.2, 45.0, 37.7, 32.7, 31.8, 25.1, 22.7, 17.7, 17.5, 10.2, 9.3

**MS (HR-ESI+)** found 315.1961, for  $[\text{M}+\text{H}]^+$  calculated 315.1972; found 337.1781, calculated for  $[\text{M}+\text{Na}]^+$  337.1791; found 297.1857, calculated for  $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$  297.1866



**Figure 63** |  $^1\text{H}$ -NMR of the ethyl acetate extract identified 5 different compounds, which are associated with the educt **13**.  $^1\text{H}$ -NMR spectrum of compound **27** was performed on a 600 MHz instrument (Bruker) with the solvent  $\text{CDCl}_3$ .

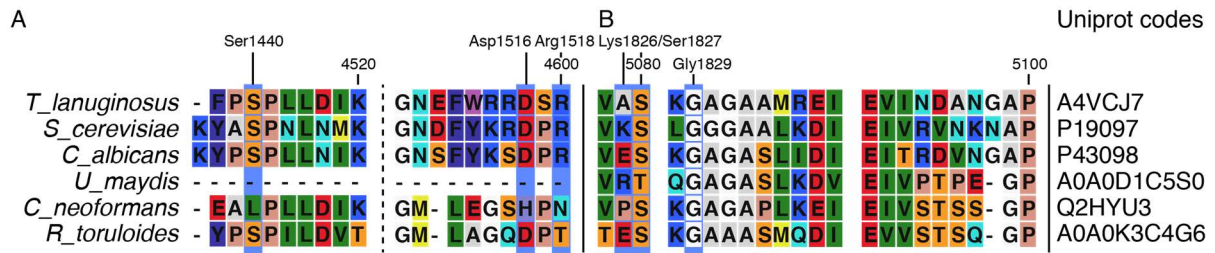


**Figure 64** |  $^{19}\text{F}$ -NMR of the ethyl acetate extract identified at least two different fluorinated compounds.  $^{19}\text{F}$ -NMR spectrum of compound **27** was performed on a 300 MHz instrument (Bruker) with the solvent  $\text{CDCl}_3$ .



## Chapter 7: Appendix

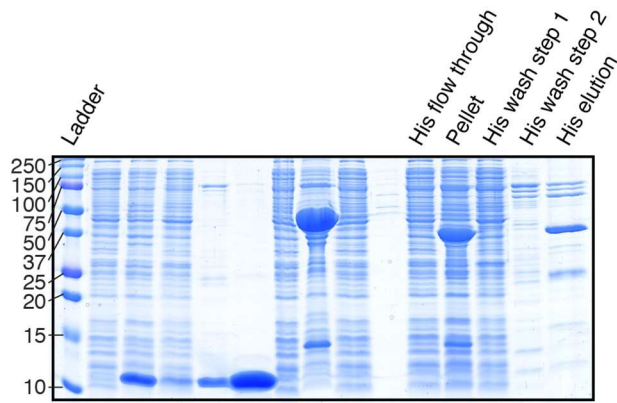
### 7.1 Supplementary figures



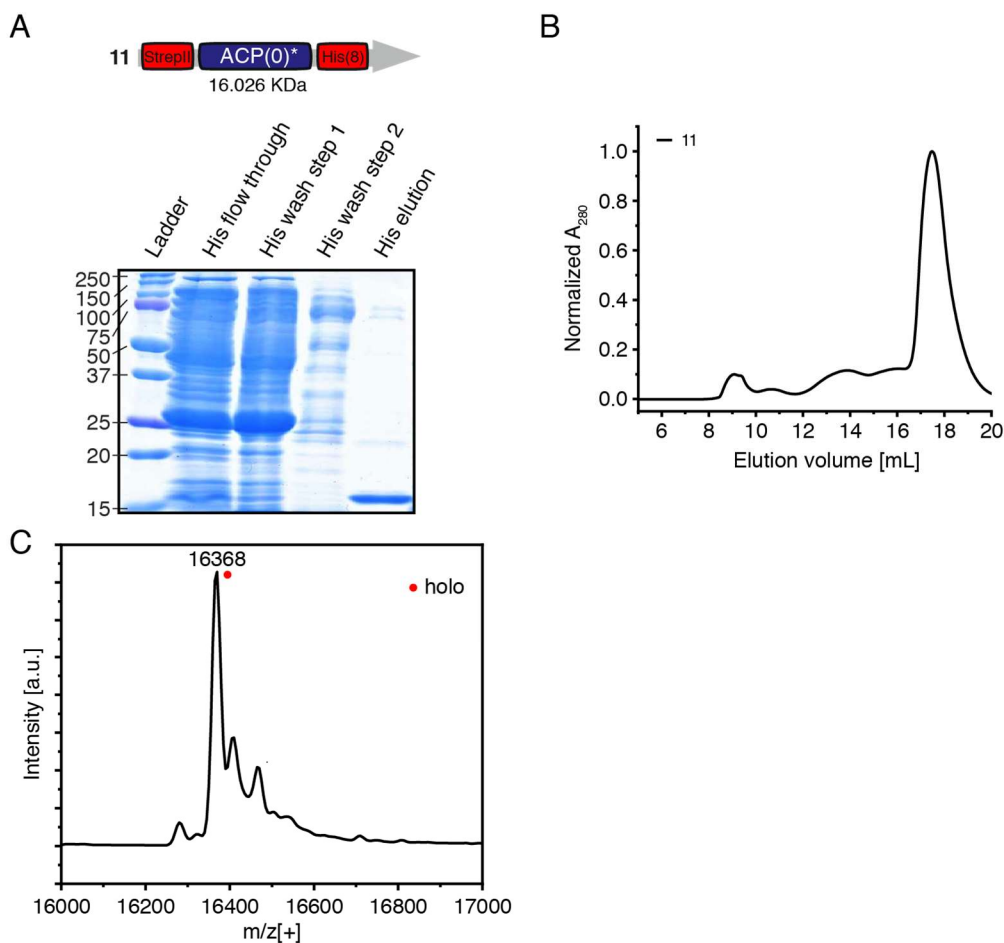
**Figure S1** | Amino acid alignment of different FASs with Uniprot accession codes. The alignment was created by CLC Main Workbench (version 6.9.1). Amino acids able to interact with the phosphate are highlighted with light blue boxes. (A) Amino acid alignment of the Ser1440 motif. (B) Amino acid alignment of the Ser1827 motif.

Abbreviation	Protein type	Uniprot codes	Organism
PtmR	KASIII/ACP-Shuttle AT-like	B7TWM5	<i>Streptomyces pactum</i>
	KASIII/ACP-Shuttle AT-like	H6ACY8	<i>Streptomyces antibioticus</i>
TiaF	ACP-Shuttle AT-like	E9LIM9	<i>Dactylosporangium aurantiacum</i> <i>subsp. hamdenensis</i>
CalO4	ACP-Shuttle AT-like	Q8KND4	<i>Micromonospora echinospora</i>
SanChIB3	ACP-Shuttle AT-like	Q0R4P5	<i>Streptomyces antibioticus</i>
	ACP-Shuttle AT-like	E3FQ73	<i>Stigmatella aurantiaca</i>
PokM2	ACP-Shuttle AT-like	C0JWB2	<i>Streptomyces diastatochromogenes</i>
	ACP-Shuttle AT-like	Q9F8U5	<i>Streptomyces rishiriensis</i>
CloN2	ACP-Shuttle AT-like	Q8GHC0	<i>Streptomyces roseochromogenus</i>
	ACP-Shuttle AT-like	Q2I767	<i>Streptomyces sp. Tu6071</i>
AviN	ACP-Shuttle AT-like	Q93KW7	<i>Streptomyces viridochromogenes Tue57</i>
SerChIB3	ACP-Shuttle AT-like	A4FKH1	<i>Saccharopolyspora erythraea</i>
EcFabB	FabB	P0A953	<i>Escherichia coli</i>
EcFabF	FabF	P0AAI5	<i>Escherichia coli</i>
	FabF	Q5SL80	<i>Thermus thermophilus</i>
	KAS III	P72392	<i>Streptomyces coelicolor</i>
	KAS III	Q5PT44	<i>Streptomyces echinatus</i>
	KAS III	Q54206	<i>Streptomyces glaucescens</i>
	KAS III	Q2MGB8	<i>Streptomyces griseus</i>
	KAS III	O68915	<i>Streptomyces roseofulvus</i>
BenQ	KAS III	B1GSM6	<i>Streptomyces sp. A2991200</i>
ZhuH	KAS III	Q9F6D4	<i>Streptomyces lividans</i>
	KAS III	B6SEH3	<i>Streptomyces sp. CM020</i>
ChIB6	KAS III-like	Q0R4P2	<i>Streptomyces antibioticus</i>
	KAS III-like	Q9L548	<i>Streptomyces galilaeus</i>
DpsC	KAS III-like	Q54816	<i>Streptomyces peucetius</i>
CerJ	KAS III-like	J7EPG1	<i>Streptomyces tendae</i>
EcFabD	MCAT (Malonyl-CoA-AT)	P0AAI9	<i>Escherichia coli</i>
	Type I KS	Q9KIV4	<i>Streptomyces antibioticus</i>
	Type I KS	O339542	<i>Streptomyces fradiae</i>
	Type I KS	Q67G56	<i>Streptomyces griseoruber</i>
	Type I KS	Q7WTD6	<i>Streptomyces nanchangensis</i>
	Type I KS	D3Y1I3	<i>Streptomyces sp. 307-9</i>
	Type I KS	A0FCL1	<i>Streptomyces violaceusniger</i>
EcFabH	FabH/KAS III	P0A6R0	<i>Escherichia coli</i>
	FabH/KAS III	P20582	<i>Pseudomonas aeruginosa</i>

**Figure S2 |** Enzyme list used for the alignment and phylogenetic of the KAS III like proteins (figure 33) with classification, Uniprot accession codes and origin.

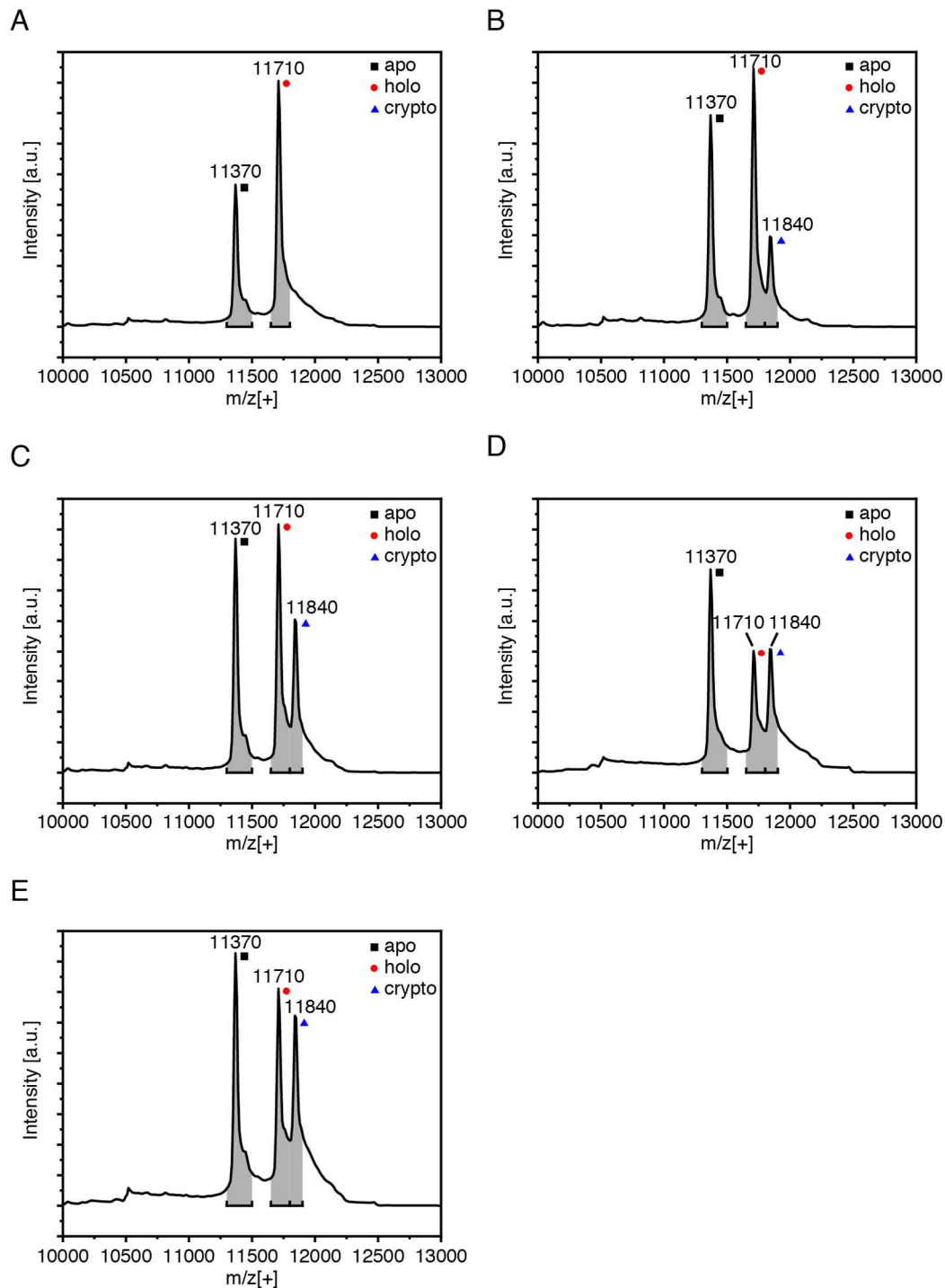


**Figure S3** | Complete Coomassie-stained SDS-PAGE of figure 34



**Figure S4** | Purification and ESI analysis of codon optimized SerMSAS-ACP(0)\*. The asterisk was used to clarify that the gene was codon optimized. Data were received from three bachelor theses, supervised by me.<sup>136,137,159</sup> (A) Coomassie-stained SDS-PAGE with samples taken during the purification of ACP(0)\* by IMAC. ACP(0)\* was co-expressed with *Sfp*. (B) SEC of ACP(0)\* on a S200 Increase 10/300 GL column (in 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 1 mM EDTA). (C) Mass spectroscopic analysis of ACP(0)\*, co-expressed with *Sfp*, was

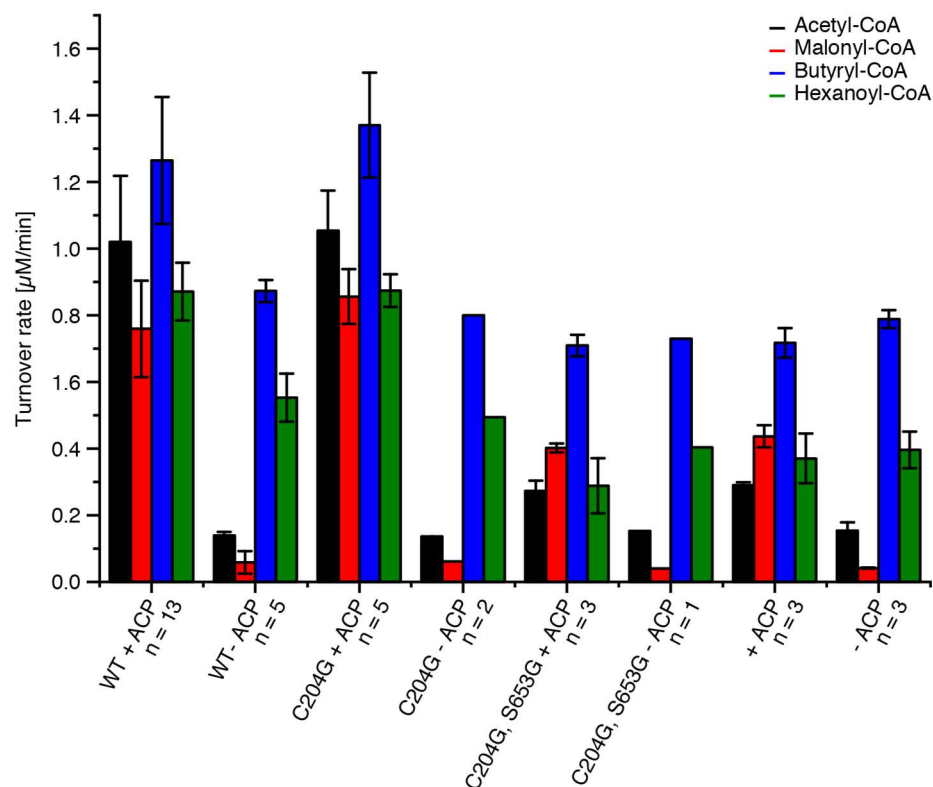
performed with ESI in positive mode (SYNAPT G-2, waters). The theoretical mass for the holo-form is 16,366 Da.



**Figure S5 |** Mass spectrometry analysis of SerChIB2 during the transfer assay of figure 37 C, monitored with ESI in positive mode (SYNAPT G-2, waters). The reaction contained 30  $\mu$ M SerChIB2, 0.4  $\mu$ M SerChIB3 and 15.0 mM 6-MSA-SNAC and was quenched at specific times (0 min, 1 min, 5 min, 10 min, 30 min, 60 min). OriginPro 8.5 was used to integrate the apo-form (11,300-11,500 Da), holo-form (11,650-11,800 Da) and crypto-form (11,800-11,900 Da). Data were received from the bachelor thesis of Christian Gusenda.<sup>136</sup>

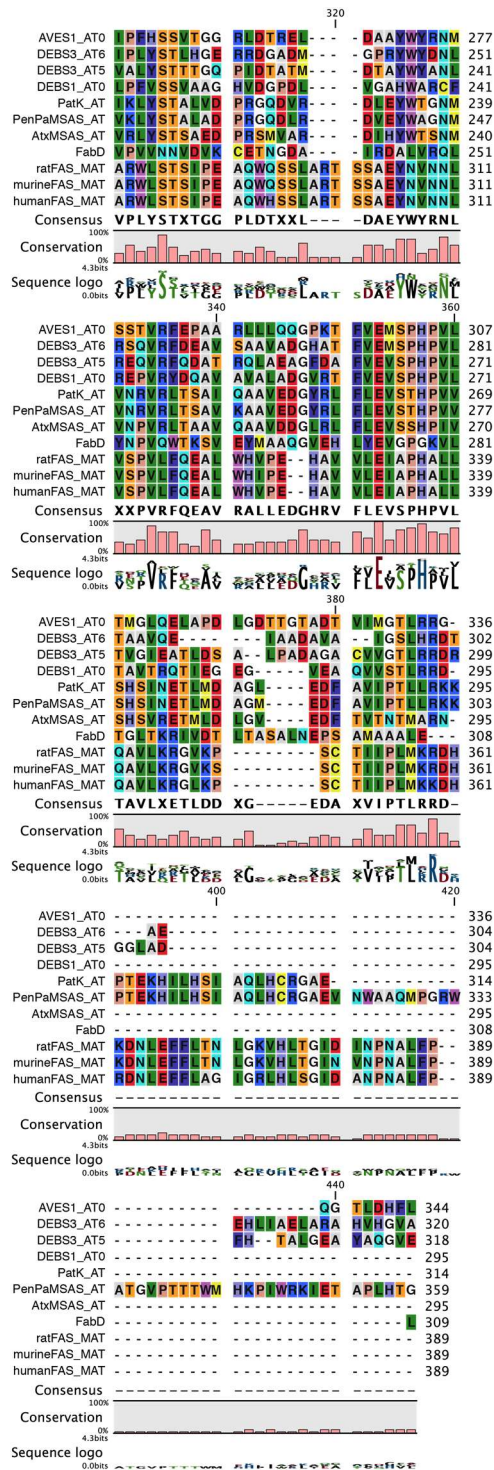
compound number	chemical formula	calculated mass [M+H] <sup>+</sup>	calculated mass [M-H] <sup>-</sup>	observed mass [M+H] <sup>+</sup>	observed mass [M-H] <sup>-</sup>	retention time [min]
6-MSA	C <sub>8</sub> H <sub>8</sub> O <sub>3</sub>	153,05783	151,04217	nd	151.04	5.9
6-PrSA	C <sub>10</sub> H <sub>12</sub> O <sub>3</sub>	181,08783	179,07217	nd	179.05	7.5
6-PSA	C <sub>12</sub> H <sub>16</sub> O <sub>3</sub>	209,11783	207,10217	nd	207.10	9.1
TAL	C <sub>6</sub> H <sub>6</sub> O <sub>3</sub>	127,03783	125,02217	127.04	125.02	1.8
Pr-TAL	C <sub>8</sub> H <sub>10</sub> O <sub>3</sub>	155,06783	153,05217	155.07	153.05	5.3
P-TAL	C <sub>10</sub> H <sub>14</sub> O <sub>3</sub>	183,09783	181,08217	183.10	181.08	7.2

**Figure S6** | Compound list of all biosynthesized products by PenPaMSAS, verified by HPLC-MS analysis (figure 41, chapter 2.3.2).

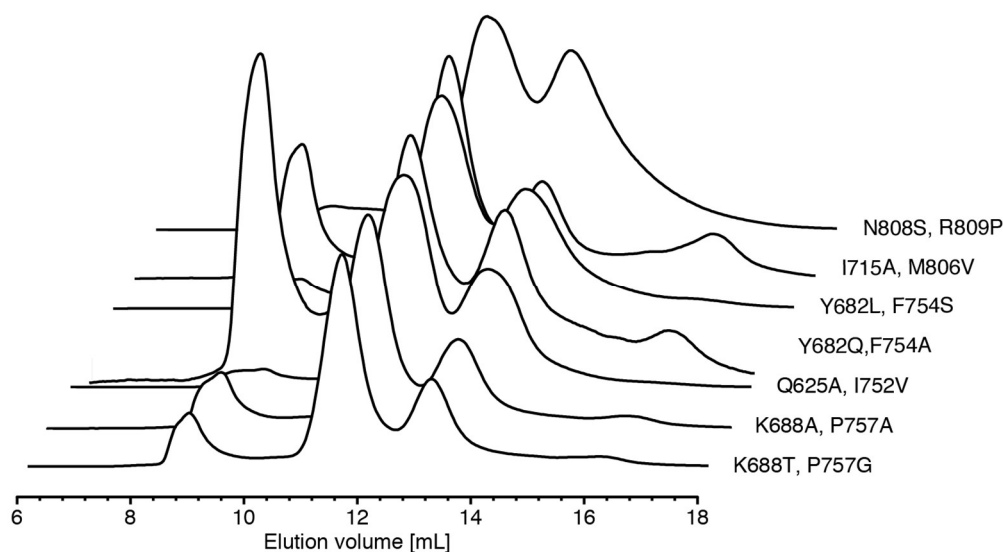


**Figure S7** | Investigation of side reactions within the transacylation assay. Comparison of acyl-CoA turnover rates by transacylation or hydrolysis of the wildtype protein, KS\*(C204G)-AT, KS\*(C204G)-AT\*(S653G) and without enzyme in the presence and absence of ACP with the substrates acetyl-CoA, Mal-CoA, butyryl-CoA and hexanoyl-CoA. Final concentration in the assays were 60 µM or 0 µM ACP(0) and 80 µM acyl-CoA. The enzyme concentration was adjusted to the turnover rates for the analysed substrate and was 5 nM for acetyl-CoA, 0.5 nM for Mal-CoA, and 1 µM for butyryl-CoA and hexanoyl-CoA.

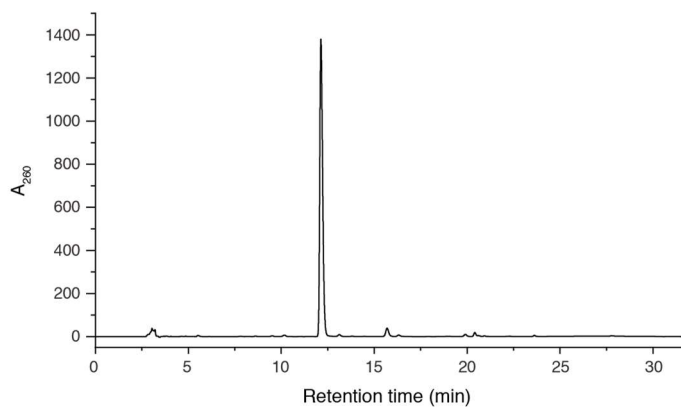




**Figure S8** | Sequence alignment of ATs of avermectin (AT0, isobutyrate, 2-methylbutyrate), DEBS AT6 (MM-CoA), DEBS AT5 (MM-CoA), DEBS loading module (AT0, propionyl-CoA), MSAS from *Penicillium expansum* (PatK; acetyl-CoA, Mal-CoA), PenPaMSAS (acetyl-CoA, Mal-CoA), MSAS from *A. terreus* (acetyl-CoA, Mal-CoA), FabD (Mal-CoA), rat FAS (polyspecific), mFAS (polyspecific), human FAS (polyspecific) with its specificity in brackets and UniProt accession codes. The alignment was created by CLC Main Workbench (version 6.9.1).

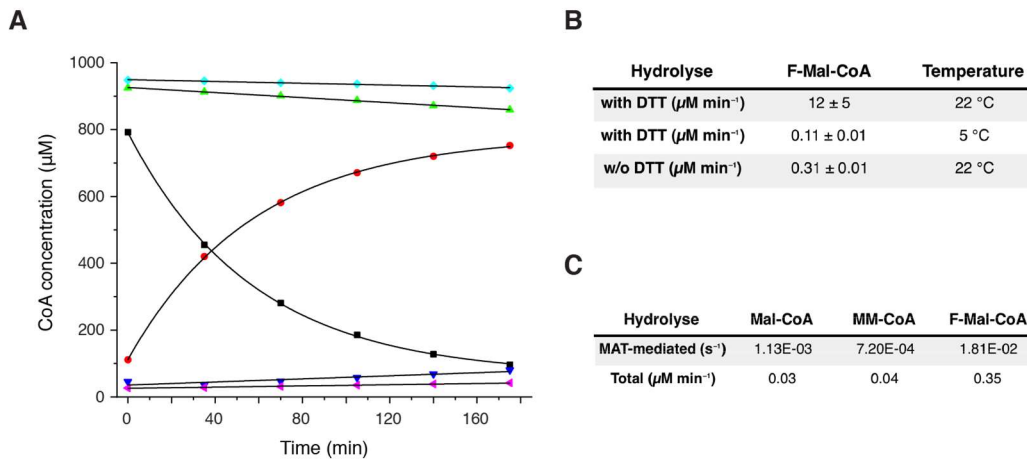


**Figure S9** | SEC of all stable constructs used for the mutational study. SEC was performed on a Superdex 200 increase 10/300 GL column (in 50 mM sodium phosphate (pH 7.6), 450 mM sodium chloride, 20% glycerol and 1 mM EDTA)

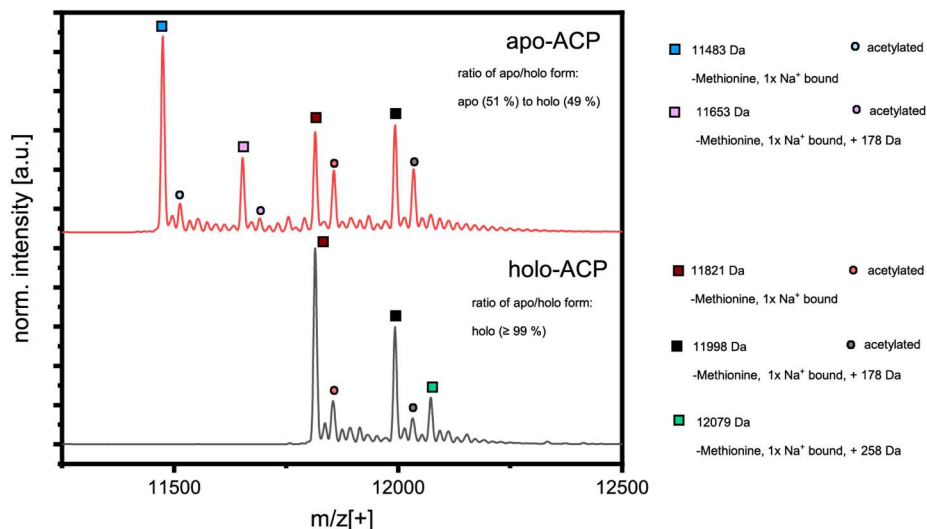


**Figure S10** | Quality control of chemically synthesized F-Mal-CoA. F-Mal-CoA analysis using HPLC-UV (described in material and methods) demonstrates almost quantitative conversion of free CoA (retention time: 15.5 min). The retention time of F-Mal-CoA is with 12.1 min similar to Mal-CoA (12.0 min, data not shown). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



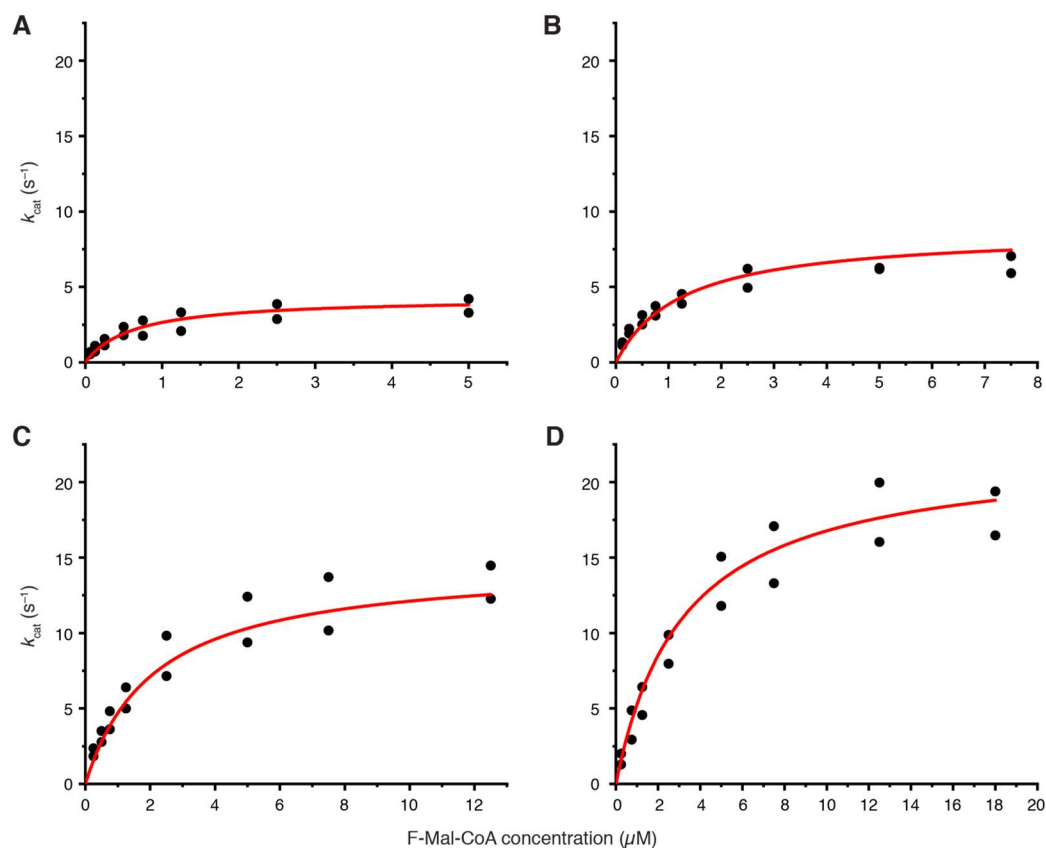


**Figure S11** | Investigation of F-Mal-CoA stability with and without DTT at different temperatures. (A) the stability of F-Mal-CoA was monitored with HPLC-UV and peak areas were converted to concentration with a CoA calibration curve. F-Mal-CoA degradation (black) and its corresponding CoA liberation (red) in buffer with DTT at 22 °C was fitted with an exponential decay function in OriginPro 8.5. Degradation without DTT at 5 °C (cyan: F-Mal-degradation, purple: CoA-liberation) and at 22 °C (green: F-Mal-degradation, blue: CoA-liberation) was fitted by linear regression (buffer: (400 mM phosphate buffer, 20 % (v/v) glycerol, 2 mM or 0 mM DTT, 1 mM EDTA, 0.8 % DMSO (pH 7.2)). (B) Initial slope average of F-Mal-degradation and CoA-liberation of the corresponding condition. The analysis revealed that the substrate is not stable in buffer containing DTT and much less prone for hydrolysis at lower temperatures. (C) Hydrolysis rates of the extender substrates determined by the  $\alpha$ KGDH assay. Final substrate concentrations in the assays were 20  $\mu\text{M}$  X-CoA and 0.2  $\mu\text{M}$  KS<sup>-</sup>-MAT. Fast F-Mal-CoA degradation by DTT was the reason why DTT was omitted in assays when using this substrate. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>

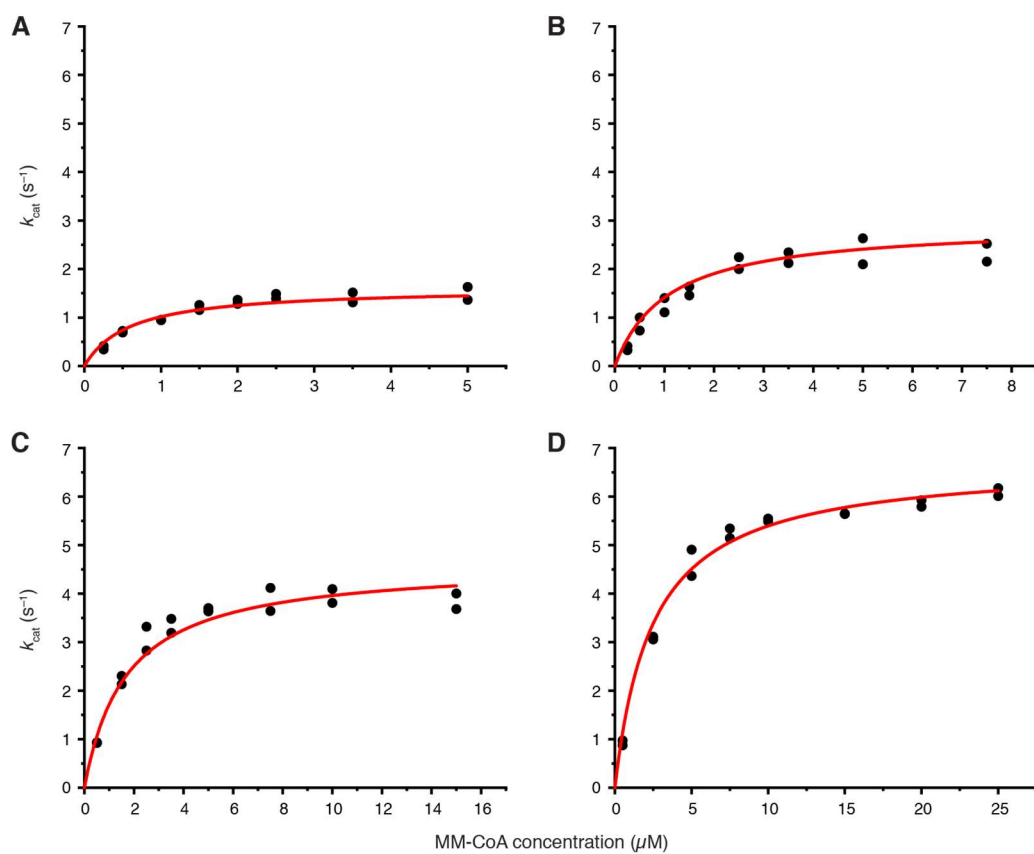


**Figure S12** | Mass spectroscopic analysis of DEBS ACP6 for quality control. ACP was co-expressed with *Sfp* (red) or *Npt* (black) and analysed with ESI in positive mode (SYNAPT G-2, waters). Co-expression with *Sfp* led to a mixture of 49 % of phosphopantetheinlated (holo-form, brown square, theoretical mass: 11,816.10 Da) and 51 % apo-form (blue square, theoretical mass = 11,475.76 Da),

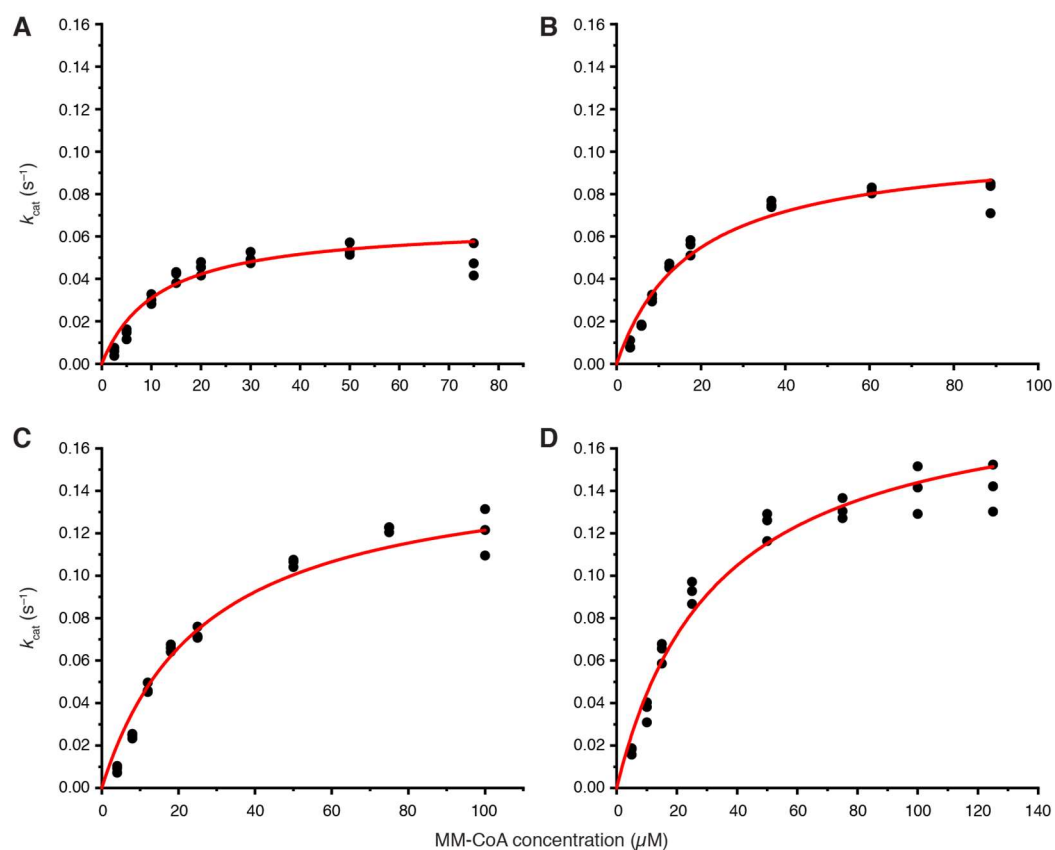
while co-expression with *Npt* resulted in quantitative conversion to the holo-form. A mass shift of 178 Da (purple square for the apo-form and black square for the holo-form) and 258 Da (cyan square for the holo-form) could be explained by His-tag phosphogluconoylation in *E. coli*, whereas a mass shift of 42 Da corresponds to the acetylated form (circles with the corresponding colour).<sup>203</sup> The ESI experiments were performed and analysed by Khanh Vu Huu and Kudratullah Karimi from the Morgner group (Goethe-University, Frankfurt). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S13** | MAT mediated transacylation of fluoromalonyl moieties to FAS ACP at four ACP concentrations: (A) 11  $\mu\text{M}$ , (B) 24  $\mu\text{M}$ , (C) 49  $\mu\text{M}$ , (D) 100  $\mu\text{M}$ . Initial velocities are plotted against F-Mal-CoA concentration and fitted globally using the Michealis-Menten equation ( $n = 2$ ). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



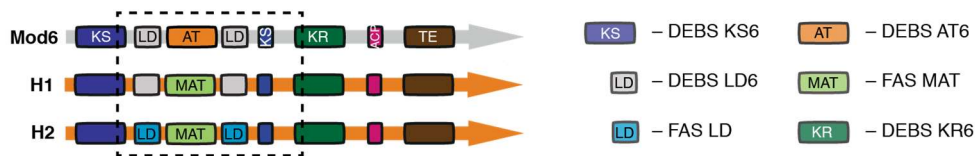
**Figure S14** | MAT mediated transacylation of methylmalonyl moieties to DEBS ACP6 at four ACP concentrations: (A) 53  $\mu\text{M}$ , (B) 107  $\mu\text{M}$ , (C) 200  $\mu\text{M}$ , (D) 373  $\mu\text{M}$ . Initial velocities are plotted against MM-CoA concentration and fitted globally using the Michealis-Menten equation ( $n = 2$ ). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S15** | DEBS AT6 mediated transacylation of fluoromalonyl moieties to DEBS ACP6 at four ACP concentrations: (A) 64  $\mu M$ , (B) 119  $\mu M$ , (C) 238  $\mu M$ , (D) 406  $\mu M$ . Initial velocities are plotted against MM-CoA concentration and fitted globally using the Michealis-Menten equation ( $n = 3$ ). The figure was published by Rittner and Joppe *et al.*<sup>44</sup>

## DEBS3 1881

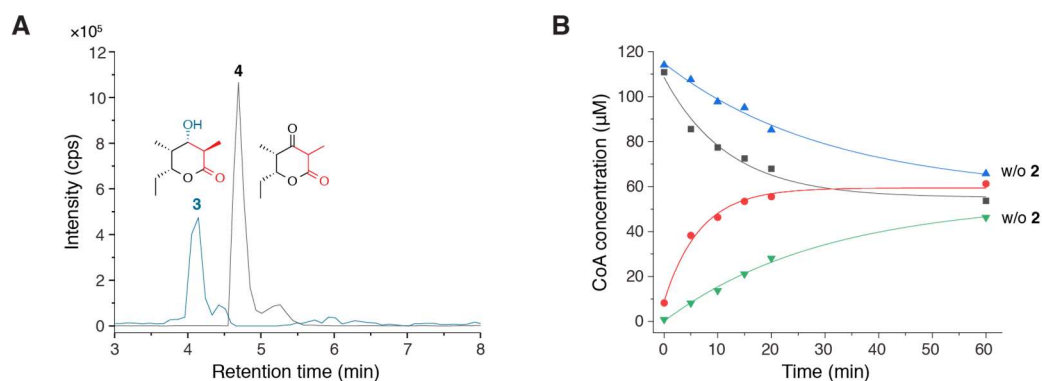
Mod6	VELAEAVSPW	PPAADGVRRA	GVSAGFVSGT	NAHVIIAEPP	EPEPLPEPGP	VGVLAAANSV
H1	VELAEAVSPW	PPAADGVRRA	GVSAGFVSGT	NAHVIIAEPP	EPEPLPEPGP	VGVLAAANSV
H2	VELAEAVSPW	PPAADGVRRA	GVSAGFVSGT	NAHVIIAEPP	TRQ- - - - AP	APTAHAALPH
Mod6	PVLLSARTET	ALAAQARLLE	SAVDDSVPLT	ALASALATGR	AHLPRRAALL	AGDHEQLRGG
H1	PVLLSARTET	ALAAQARLLE	SAVDDSVPLT	ALASALATGR	AHLPRRAALL	AGDHEQLRGG
H2	LLHASGRTLE	AVQDLLEQGR	QHSQDLAFVS	MLNDIAATPT	AAMPFRGYTV	LGVEGRVQ- -
Mod6	LRAVAEGVAA	PGATTGTASA	GGVV- FVFPG	QGAQWEGMAR	GLLSVPVFAE	SIAECDAVLS
H1	LRAVAEGVAA	PGATTGTASN	KRPLWFICSG	MGTQWRGMGL	SLMRLDSFRE	SILRSDEAVK
H2	- - - - -	- - EVQQVSTN	KRPLWFICSG	MGTQWRGMGL	SLMRLDSFRE	SILRSDEAVK
Mod6	EVAGFSASEV	LEQRPDAPSL	ERVDVVQPVV	FSVMVSLARL	WGACGVSPSA	VIGHSQGEIA
H1	PLGVKVSDDL	LST- - DERTF	DDIVHAFVSL	TAIQIALIDL	LTSVGLKPDG	IIGHSLGEVA
H2	PLGVKVSDDL	LST- - DERTF	DDIVHAFVSL	TAIQIALIDL	LTSVGLKPDG	IIGHSLGEVA
Mod6	AAVVAGVLSL	EDGVRVVALR	AKALRALAGK	GGMVSLAAPG	ERARALIAPW	EDRISVAAVN
H1	CGYADGCLSQ	REAVLAAYWR	GQCICKAHLR	PG- - SMAAVG	LSWEECKQRC	PAGVVPACHN
H2	CGYADGCLSQ	REAVLAAYWR	GQCICKAHLR	PG- - SMAAVG	LSWEECKQRC	PAGVVPACHN
Mod6	SPSSVVVSGD	PEALAEVLR	CEDEGVRRAK	LPVDYASHSR	HVEE- - - - I	RETI LADLDG
H1	SEDVTI I SGP	QAAVNEFVEQ	LKQEGVFAKE	VRTGGLAFHS	YFMELIAPT	LQALKKVIRE
H2	SEDVTI I SGP	QAAVNEFVEQ	LKQEGVFAKE	VRTGGLAFHS	YFMELIAPT	LQALKKVIRE
Mod6	ISARRAAI PL	YSTLHGERRD	GADMGPRYWY	DNLRSQVRFD	EAVSAAVADG	HATFVEMSPH
H1	PRPRSARWLS	TSIPEAQWQS	SLARTSSAEY	NVNNLVSPVL	FQEALWHIPE	HAVVLEIAPH
H2	PRPRSARWLS	TSIPEAQWQS	SLARTSSAEY	NVNNLVSPVL	FQEALWHIPE	HAVVLEIAPH
Mod6	PVLTAAVQEI	AADAVAIGSL	HRDTAEEHL-	- - IAELARAH	VHGVAVDWRN	VFP- - - - AA
H1	ALLQAVLKRQ	VKSSCTI I PL	MKRDKDNL	FFLTNLGKVH	LTGVAVDWRN	VFP- - - - AA
H2	ALLQAVLKRQ	VKSSCTI I PL	MKRDKDNL	FFLTNLGKVH	LTGINVNPNA	LFPPVEFPAP
Mod6	PPVALPNYPF	EPQRYWLAP	VSDQLADSR	RVDWRPLATT	PVDLEGGFLV	HGSAPESLTS
H1	PPVALPNYPF	EPQRYWLAP	VSDQLADSR	RVDWRPLATT	PVDLEGGFLV	HGSAPESLTS
H2	RGTPLPNYPF	EPQRYWLAP	VSDQLADSR	RVDWRPLATT	PVDLEGGFLV	HGSAPESLTS



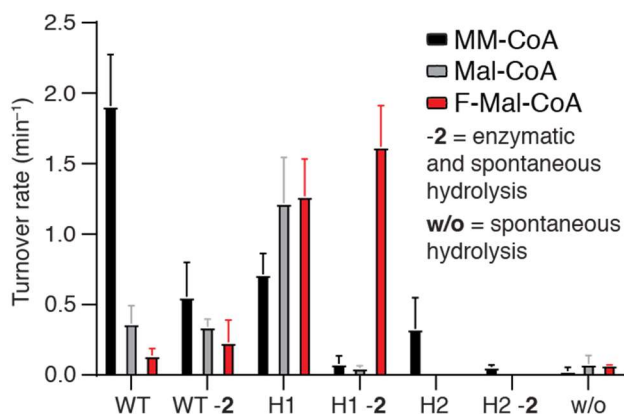
**Figure S16** | Further information on the DEBS/FAS hybrid design. Sequence alignment of DEBS M6 (WT) with both DEBS/FAS hybrids (H1 and H2). H1 was created by exchange of the AT6 with MAT, whereas in H2 the AT6 was exchanged with the adjacent LD by LD-MAT. The sequence was coloured accordingly to the domain architecture below. The figure was published by Rittner and Joppe *et al.*<sup>44</sup>

compound number	chemical formula	turnover rate (min <sup>-1</sup> )	calculated mass	observed mass	observed mass (HR)	retention time [min]
3	C9H16O3	WT = 0.29 ± 0.03 H1 = 0.19 ± 0.06 H2 = 0.07 ± 0.02	[M+H] <sup>+</sup> = 173.1178 [M+Na] <sup>+</sup> = 195.0997 [M-H <sub>2</sub> O+H] <sup>+</sup> = 155.1072 [M-H] <sup>+</sup> = 171.1021	[M+H] <sup>+</sup> = 173.03 [M+Na] <sup>+</sup> = 194.98 [M-H <sub>2</sub> O+H] <sup>+</sup> = 155.01	nd	4,1
4	C9H14O3	WT = 1.4 ± 0.4 H1 = 0.6 ± 0.1 H2 = 0.23 ± 0.02	[M+H] <sup>+</sup> = 171.1021 [M+Na] <sup>+</sup> = 193.0841 [M-H <sub>2</sub> O+H] <sup>+</sup> = 153.0916 [M-H] <sup>+</sup> = 169.0865	[M+H] <sup>+</sup> = 171.00 [M+Na] <sup>+</sup> = 192.96 [M-H <sub>2</sub> O+H] <sup>+</sup> = 152.96 [M-H] <sup>+</sup> = 169.12	nd	4,8
5	C8H14O3	WT = 0.012 ± 0.003 H1 = 0.19 ± 0.02	[M+H] <sup>+</sup> = 159.1021 [M+Na] <sup>+</sup> = 181.0841 [M-H <sub>2</sub> O+H] <sup>+</sup> = 141.0916 [M-H] <sup>+</sup> = 157.0865	[M+Na] <sup>+</sup> = 181.03	nd	3,3
6	C8H12O3	WT = 0.0 ± 0.1 H1 = 1.2 ± 0.3	[M+H] <sup>+</sup> = 157.0865 [M+Na] <sup>+</sup> = 179.0684 [M-H <sub>2</sub> O+H] <sup>+</sup> = 139.0759 [M-H] <sup>+</sup> = 155.0708	[M-H] <sup>+</sup> = 155.16	nd	4,6
7	C8H13FO3	nd	[M+H] <sup>+</sup> = 177.0927 [M+Na] <sup>+</sup> = 199.0747 [M-H <sub>2</sub> O+H] <sup>+</sup> = 159.0822 [M-H] <sup>+</sup> = 175.0771	-	-	-
8	C8H11FO3	nd	[M+H] <sup>+</sup> = 175.0771 [M+Na] <sup>+</sup> = 197.0590 [M-H <sub>2</sub> O+H] <sup>+</sup> = 157.0665 [M-H] <sup>+</sup> = 173.0614	[M-H] <sup>+</sup> = 173.11	nd	4,8
11	C17H28O4	WT = 0.30 ± 0.05 H1 = 0.12 ± 0.01 H1 = 0.158 ± 0.001*	[M+H] <sup>+</sup> = 297.2066 [M+Na] <sup>+</sup> = 319.1886 [M-H <sub>2</sub> O+H] <sup>+</sup> = 279.1960 [M-H] <sup>+</sup> = 295.1910	[M+H] <sup>+</sup> = 297.11 [M+Na] <sup>+</sup> = 319.11 [M-H <sub>2</sub> O+H] <sup>+</sup> = 279.11	[M+H] <sup>+</sup> = 297.2056 [M+Na] <sup>+</sup> = 319.1876 [M-H <sub>2</sub> O+H] <sup>+</sup> = 279.1950	8,1
12	C17H26O4	nd	[M+H] <sup>+</sup> = 295.1910 [M+Na] <sup>+</sup> = 317.1729 [M-H <sub>2</sub> O+H] <sup>+</sup> = 277.1804 [M-H] <sup>+</sup> = 293.1753	[M+H] <sup>+</sup> = 295.09 [M+Na] <sup>+</sup> = 317.09 [M-H <sub>2</sub> O+H] <sup>+</sup> = 277.08	[M+H] <sup>+</sup> = 295.1904 [M+Na] <sup>+</sup> = 317.1720 [M-H <sub>2</sub> O+H] <sup>+</sup> = 277.1790	9,1
13	C20H34O5	WT = 0.47 ± 0.03 H1 = 0.30 ± 0.02	[M+H] <sup>+</sup> = 355.2485 [M+Na] <sup>+</sup> = 377.2304 [M-H <sub>2</sub> O+H] <sup>+</sup> = 337.2379 [M-H] <sup>+</sup> = 353.2328	[M+H] <sup>+</sup> = 355.14 [M+Na] <sup>+</sup> = 377.16 [M-H <sub>2</sub> O+H] <sup>+</sup> = 337.14	[M+H] <sup>+</sup> = 355.2478 [M+Na] <sup>+</sup> = 377.2293 [M-H <sub>2</sub> O+H] <sup>+</sup> = 337.2369	8,0
14	C20H32O5	nd	[M+H] <sup>+</sup> = 353.2328 [M+Na] <sup>+</sup> = 375.2148 [M-H <sub>2</sub> O+H] <sup>+</sup> = 335.2223 [M-H] <sup>+</sup> = 351.2172	[M+Na] <sup>+</sup> = 375.15 [M-H <sub>2</sub> O+H] <sup>+</sup> = 335.13	[M+H] <sup>+</sup> = 353.2318 [M+Na] <sup>+</sup> = 375.2136 [M-H <sub>2</sub> O+H] <sup>+</sup> = 335.2211	7.4 / 8.8
15	C16H24O4	nd	[M+H] <sup>+</sup> = 281.1753 [M+Na] <sup>+</sup> = 303.1573 [M-H <sub>2</sub> O+H] <sup>+</sup> = 263.1647 [M-H] <sup>+</sup> = 279.1597	[M+H] <sup>+</sup> = 281.07 [M+Na] <sup>+</sup> = 303.08 [M-H <sub>2</sub> O+H] <sup>+</sup> = 263.08	[M+H] <sup>+</sup> = 281.1741 [M+Na] <sup>+</sup> = 303.1559 [M-H <sub>2</sub> O+H] <sup>+</sup> = 263.1637	7.2 / 9.5
16	C16H26O4	WT = -0.009 ± 0.004 H1 = 0.22 ± 0.03	[M+H] <sup>+</sup> = 283.1910 [M+Na] <sup>+</sup> = 305.1729 [M-H <sub>2</sub> O+H] <sup>+</sup> = 265.1804 [M-H] <sup>+</sup> = 281.1753	[M+Na] <sup>+</sup> = 305.09 [M-H <sub>2</sub> O+H] <sup>+</sup> = 265.09	[M+H] <sup>+</sup> = 283.1899 [M+Na] <sup>+</sup> = 305.1719 [M-H <sub>2</sub> O+H] <sup>+</sup> = 265.1796	7,5
17	C16H23FO4	nd	[M+H] <sup>+</sup> = 299.1659 [M+Na] <sup>+</sup> = 321.1478 [M-H <sub>2</sub> O+H] <sup>+</sup> = 281.1553 [M-H] <sup>+</sup> = 297.1502	[M+H] <sup>+</sup> = 299.08 [M+Na] <sup>+</sup> = 321.07 [M-H <sub>2</sub> O+H] <sup>+</sup> = 281.06	[M+H] <sup>+</sup> = 299.1642 [M+Na] <sup>+</sup> = 321.1457 [M-H <sub>2</sub> O+H] <sup>+</sup> = 281.1537	8.4 / 8.9
18	C16H25FO4	WT = 0.003 ± 0.005 H1 = 0.056 ± 0.004 H1 = 0.04 ± 0.01*	[M+H] <sup>+</sup> = 301.1815 [M+Na] <sup>+</sup> = 323.1635 [M-H <sub>2</sub> O+H] <sup>+</sup> = 283.1710 [M-H] <sup>+</sup> = 299.1659	[M+H] <sup>+</sup> = 301.10 [M+Na] <sup>+</sup> = 323.08 [M-H <sub>2</sub> O+H] <sup>+</sup> = 283.08	[M+H] <sup>+</sup> = 301.1807 [M+Na] <sup>+</sup> = 323.1625 [M-H <sub>2</sub> O+H] <sup>+</sup> = 283.1700	7.7 / 8.0
19	C19H30O5	nd	[M+H] <sup>+</sup> = 339.2172 [M+Na] <sup>+</sup> = 361.1991 [M-H <sub>2</sub> O+H] <sup>+</sup> = 321.2066 [M-H] <sup>+</sup> = 337.2015	[M+Na] <sup>+</sup> = 361.12 [M-H <sub>2</sub> O+H] <sup>+</sup> = 321.12	[M+H] <sup>+</sup> = 339.2160 [M+Na] <sup>+</sup> = 361.1980 [M-H <sub>2</sub> O+H] <sup>+</sup> = 321.2056	8,1
20	C19H32O5	WT = 0.002 ± 0.008 H1 = 0.35 ± 0.04	[M+H] <sup>+</sup> = 341.2328 [M+Na] <sup>+</sup> = 363.2148 [M-H <sub>2</sub> O+H] <sup>+</sup> = 323.2223 [M-H] <sup>+</sup> = 339.2172	[M+Na] <sup>+</sup> = 363.14 [M-H <sub>2</sub> O+H] <sup>+</sup> = 323.13	[M+Na] <sup>+</sup> = 363.2122 [M-H <sub>2</sub> O+H] <sup>+</sup> = 323.2200	7,5
21	C19H29FO5	nd	[M+H] <sup>+</sup> = 357.2078 [M+Na] <sup>+</sup> = 379.1897 [M-H <sub>2</sub> O+H] <sup>+</sup> = 339.1972 [M-H] <sup>+</sup> = 355.1921	[M+Na] <sup>+</sup> = 379.12 [M-H <sub>2</sub> O+H] <sup>+</sup> = 339.11	[M+H] <sup>+</sup> = 357.2062 [M+Na] <sup>+</sup> = 379.1886 [M-H <sub>2</sub> O+H] <sup>+</sup> = 339.1962	7.4 / 8.2 / 8.9
22	C19H31FO5	WT = 0.00 ± 0.02 H1 = 0.24 ± 0.01	[M+H] <sup>+</sup> = 359.2234 [M+Na] <sup>+</sup> = 381.2054 [M-H <sub>2</sub> O+H] <sup>+</sup> = 341.2128 [M-H] <sup>+</sup> = 357.2078	[M+Na] <sup>+</sup> = 381.13 [M-H <sub>2</sub> O+H] <sup>+</sup> = 341.13	[M+H] <sup>+</sup> = 359.2228 [M+Na] <sup>+</sup> = 381.2045 [M-H <sub>2</sub> O+H] <sup>+</sup> = 341.2120	7.4 / 7.7
25	C15H25FO3	nd	[M+H] <sup>+</sup> = 273.1866 [M+Na] <sup>+</sup> = 295.1686 [M-H <sub>2</sub> O+H] <sup>+</sup> = 255.1761 [M <sub>2</sub> +Na] <sup>+</sup> = 567.3474	nd	[M+H] <sup>+</sup> = 273.1859 [M+Na] <sup>+</sup> = 295.1677 [M-H <sub>2</sub> O+H] <sup>+</sup> = 255.1754 [M <sub>2</sub> +Na] <sup>+</sup> = 567.3462	7,0
27	C17H27FO4	0.14 ± 0.02*	[M+H] <sup>+</sup> = 315.1972 [M+Na] <sup>+</sup> = 337.1791 [M-H <sub>2</sub> O+H] <sup>+</sup> = 297.1866 [M-H] <sup>+</sup> = 313.1815	nd	[M+H] <sup>+</sup> = 315.1961 [M+Na] <sup>+</sup> = 337.1781 [M-H <sub>2</sub> O+H] <sup>+</sup> = 297.1857	7,8
28	C20H33FO6	nd	[M+H] <sup>+</sup> = 389.2340 [M+Na] <sup>+</sup> = 411.2159 [M-H <sub>2</sub> O+H] <sup>+</sup> = 371.2234	nd	[M+H] <sup>+</sup> = 389.2331 [M+Na] <sup>+</sup> = 411.2148 [M-H <sub>2</sub> O+H] <sup>+</sup> = 371.2225	8.2 / 9.5
29	C16H27FO3	nd	[M+H] <sup>+</sup> = 287.2023 [M+Na] <sup>+</sup> = 309.1842 [M-H <sub>2</sub> O+H] <sup>+</sup> = 269.1917 [M <sub>2</sub> +Na] <sup>+</sup> = 595.3787	nd	[M+H] <sup>+</sup> = 287.2017 [M+Na] <sup>+</sup> = 309.1836 [M-H <sub>2</sub> O+H] <sup>+</sup> = 269.1913 [M <sub>2</sub> +Na] <sup>+</sup> = 595.3779	7,0

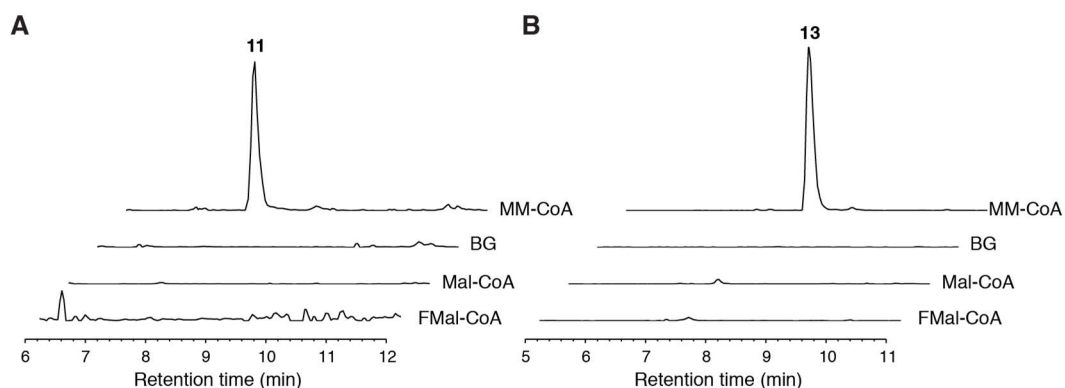
**Figure S17** | Compound list of all biosynthesized products by H1, verified by HPLC-MS analysis. Turnover rates highlighted with an asterisk were performed with slightly different conditions (higher enzyme concentration 6.4 μM compared to 4 μM, different buffer: 250 μM potassium phosphate, 10 % glycerol, pH 7.0, see chapter 6.1.5). The figure was adapted from Rittner and Joppe *et al.*<sup>44</sup>



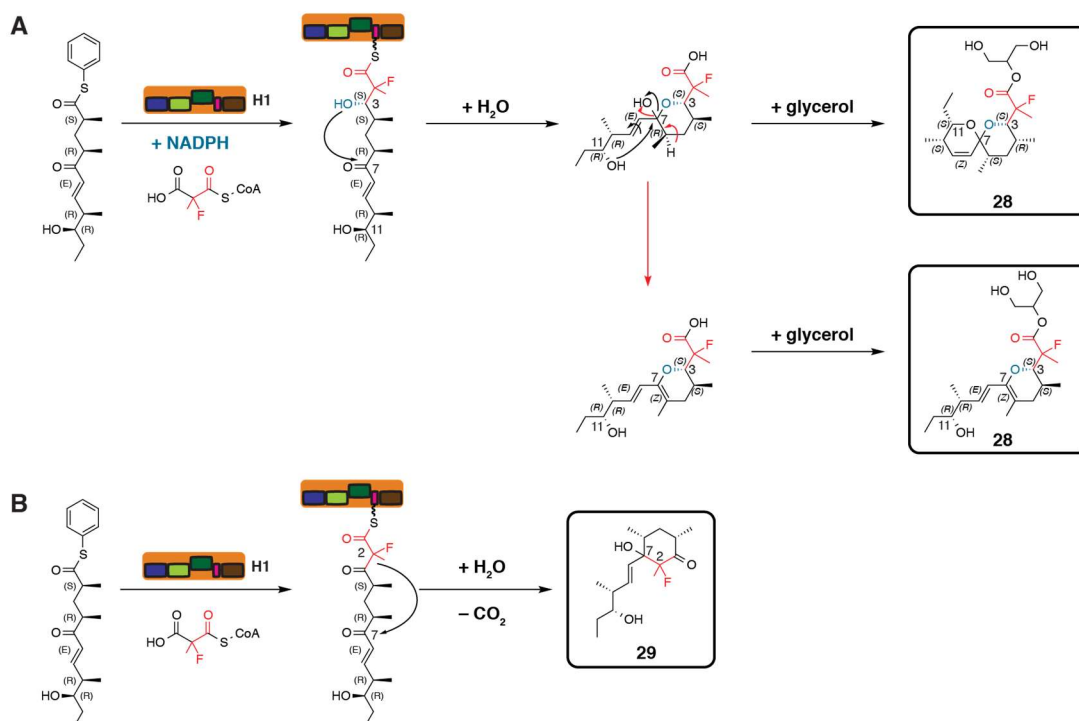
**Figure S18** | Further information on the functionality of DEBS M6. (A) Conversion of natural diketide SNAC (2) with MM-CoA in the presence of NADPH led to a mixture of reduced and unreduced TKLs shown by HPLC-MS [EIC: **3** [M+Na]<sup>+</sup> *m/z* = 194.98 and **4** [M-H]<sup>-</sup> *m/z* = 169.12]. (B) Typical CoA-consumption assay in the absence of NADPH. Conversion of natural diketide SNAC (2) with extender substrate was monitored by HPLC-UV, exemplified by DEBS M6 (WT) and MM-CoA. The product formation was determined by the consumption of extender substrate and the liberation of CoA at defined time points. The peak areas of the absorbance ( $A_{280}$ ) were converted to concentrations with an internal hydroxybutyryl-CoA standard and fitted with an exponential decay function using OriginPro 8.5. We used the sum of extender substrate and CoA concentration of each sample as quality control and outliers were exclude. Data were received from the master thesis of Elia Heid.<sup>183</sup> The figure was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S19** | Further information on the CoA consumption Assay. H2 was only measure with MM-CoA. Data were received from Elia Heid and the figure is adapted from the master thesis.<sup>183</sup>



**Figure S20** | HPLC-MS analysis of the reductive DEBS M6 (**WT**) mediated chain extension of **9** and **10** using the extender substrates MM-CoA, Mal-CoA and F-Mal-CoA revealed that WT was only capable to introduce MM-CoA (A) HPLC-MS analysis of the reductive DEBS M6 (**WT**) mediated chain extension of **9** using the extender substrates MM-CoA, Mal-CoA and F-Mal-CoA. The negative control (**BG**) was performed without **9** and is shown as a negative control for the chain elongation with MM-CoA. All data were normalized with respect to the highest peak of the H1-mediated reaction (see figure 50 B, chapter 2.4.3). [EIC: **11**  $[M+Na]^+$   $m/z = 319.11$ , **16**  $[M+Na]^+$   $m/z = 305.09$ , **18**  $[M+Na]^+$   $m/z = 323.08$ ]. (B) HPLC-MS analysis of the reductive DEBS M6 (**WT**) mediated chain extension of **10** using the extender substrates MM-CoA, Mal-CoA and F-Mal-CoA. The negative control (**BG**) was performed without **10** and is shown as a negative control for the chain elongation with MM-CoA. All data were normalized with respect to the highest peak of the H1-mediated reaction (see figure 50 C, chapter 2.4.3). [EIC: **13**  $[M+Na]^+$   $m/z = 377.16$ , **20**  $[M+Na]^+$   $m/z = 363.14$ , **22**  $[M+Na]^+$   $m/z = 381.13$ ]. Data were received from Elia Heid and the figure is adapted from the master thesis.<sup>183</sup>



**Figure S21** | Postulated mechanism and possible side product (**28**, **29**) of the **H1**-mediated chain elongation of **9** with F-MM-CoA in the presence (A) and absence (B) of NADPH. The prediction of side products is based on HPLC-HRMS experiments (see compound list, figure S17 for exact masses).



<sup>1</sup>H-NMR of compound 28 can be seen in chapter 7.3. The figure was adapted from Rittner and Joppe *et al.*<sup>44</sup>

## 7.2 Supplementary equations

### 7.2.1 Linear fit

$$y = a + bx \quad \text{[equation 3]}$$

### 7.2.2 Exponential fit

$$y = A1 \times e^{\frac{-x}{\tau_1}} + y(0) \quad \text{[equation 4]}$$

### 7.2.3 Michealis-Menten approximation



If reverse reaction does not occur, such as product concentration is too low.

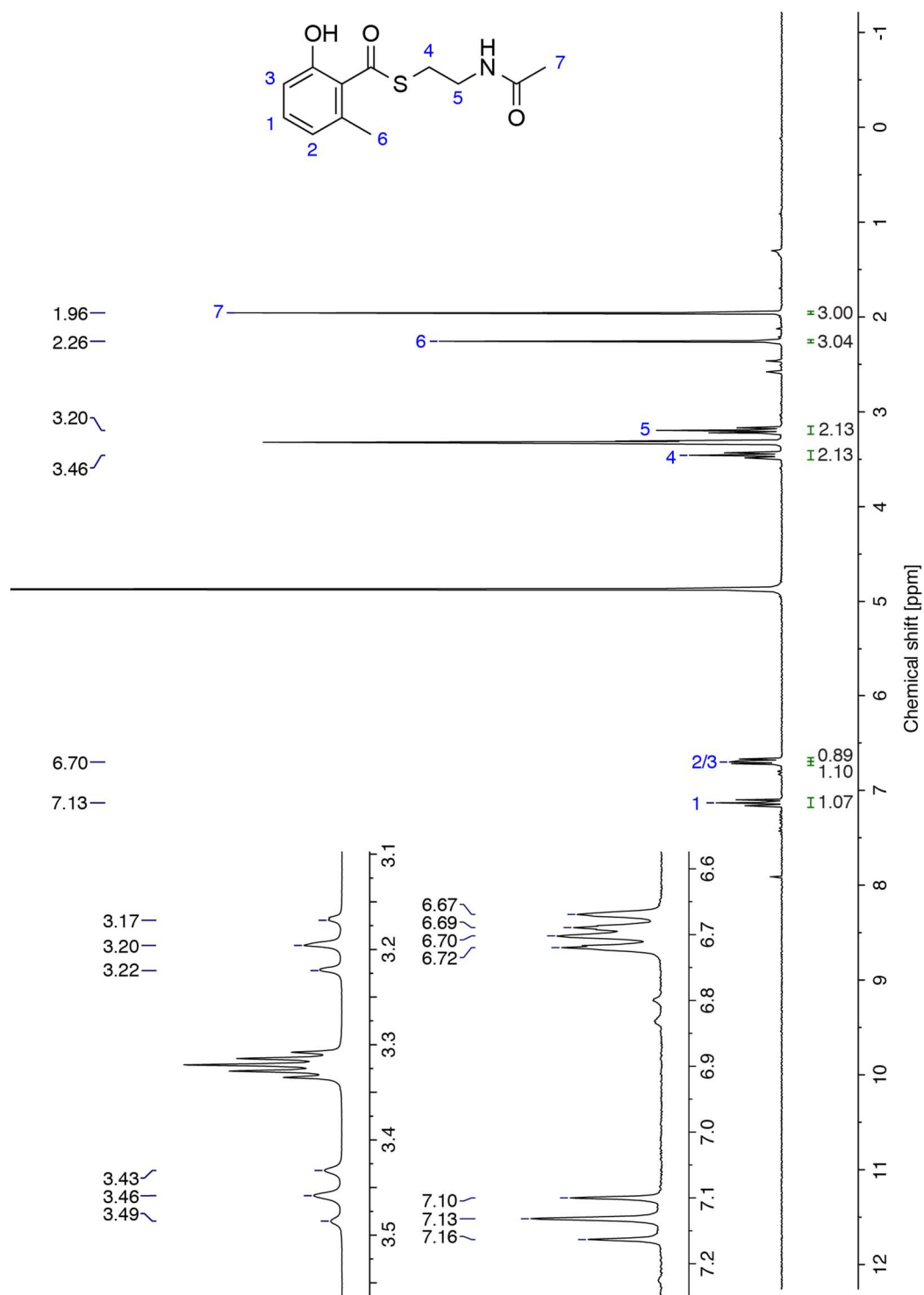
### 7.2.4 Michealis-Menten equation

$$v_0 = v_{max} \frac{[S]}{[S] + K_M}; K_M = \frac{k_{-1} + k_2}{k_1}; v_{max} = k_{cat} [E]_0 \quad \text{[equation 6]}$$

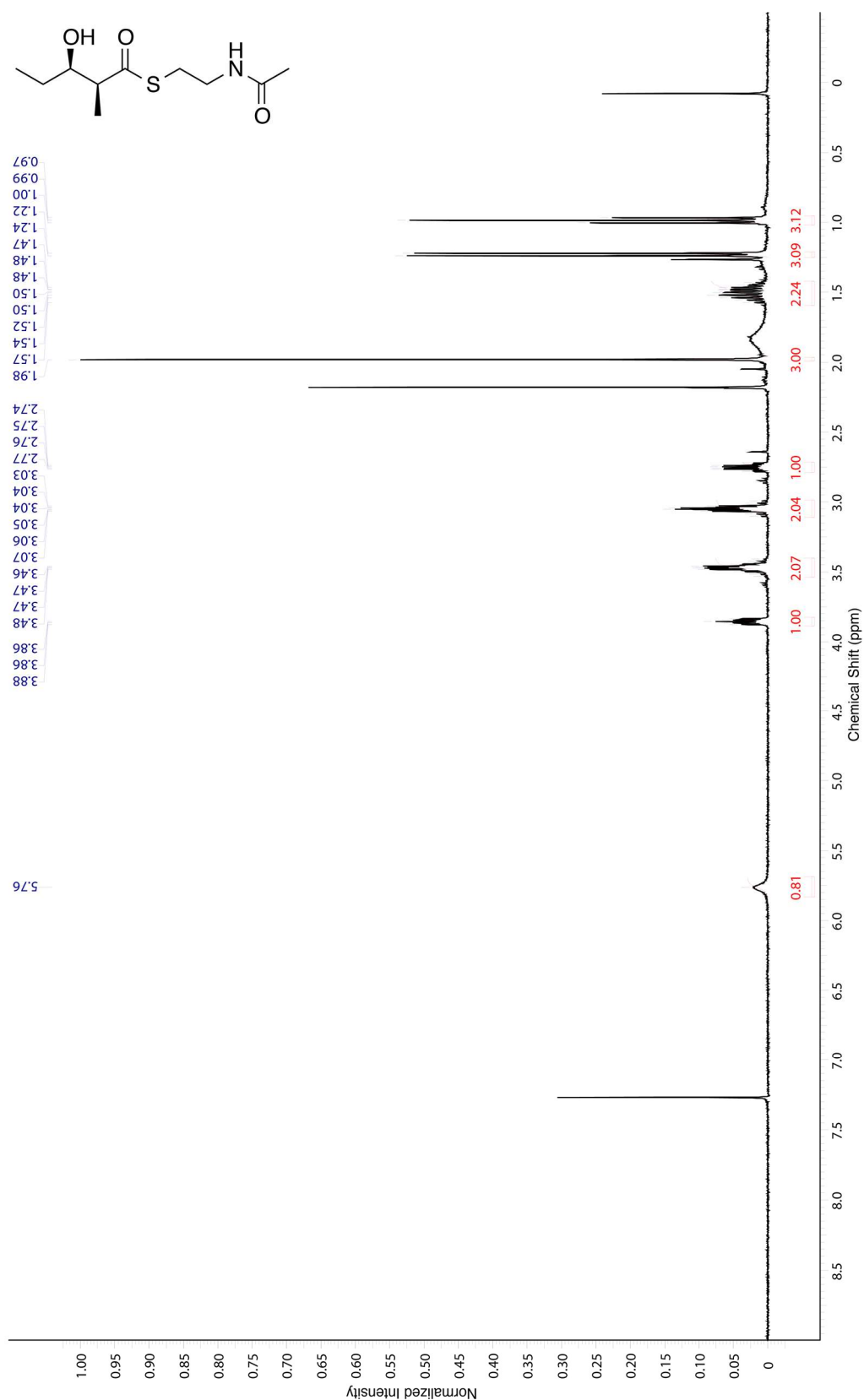
### 7.2.5 Global fit

$$v = \frac{k_{cat} [AT_0] [XCoA] [ACP]}{[XCoA]K_{ACP} + [ACP]K_S + [XCoA] [ACP]} \quad \text{[equation 7]}$$

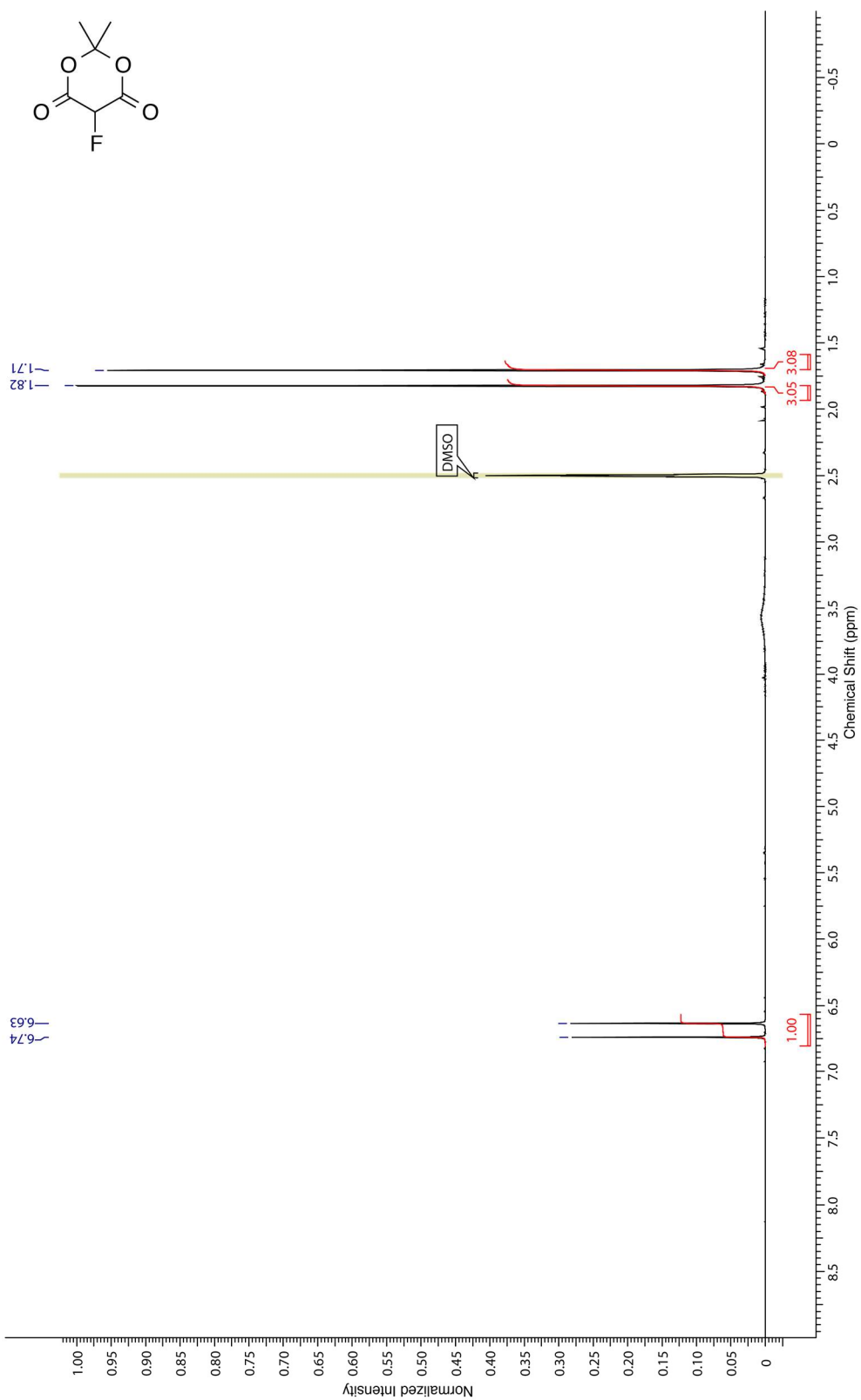
## 7.3 NMR spectra



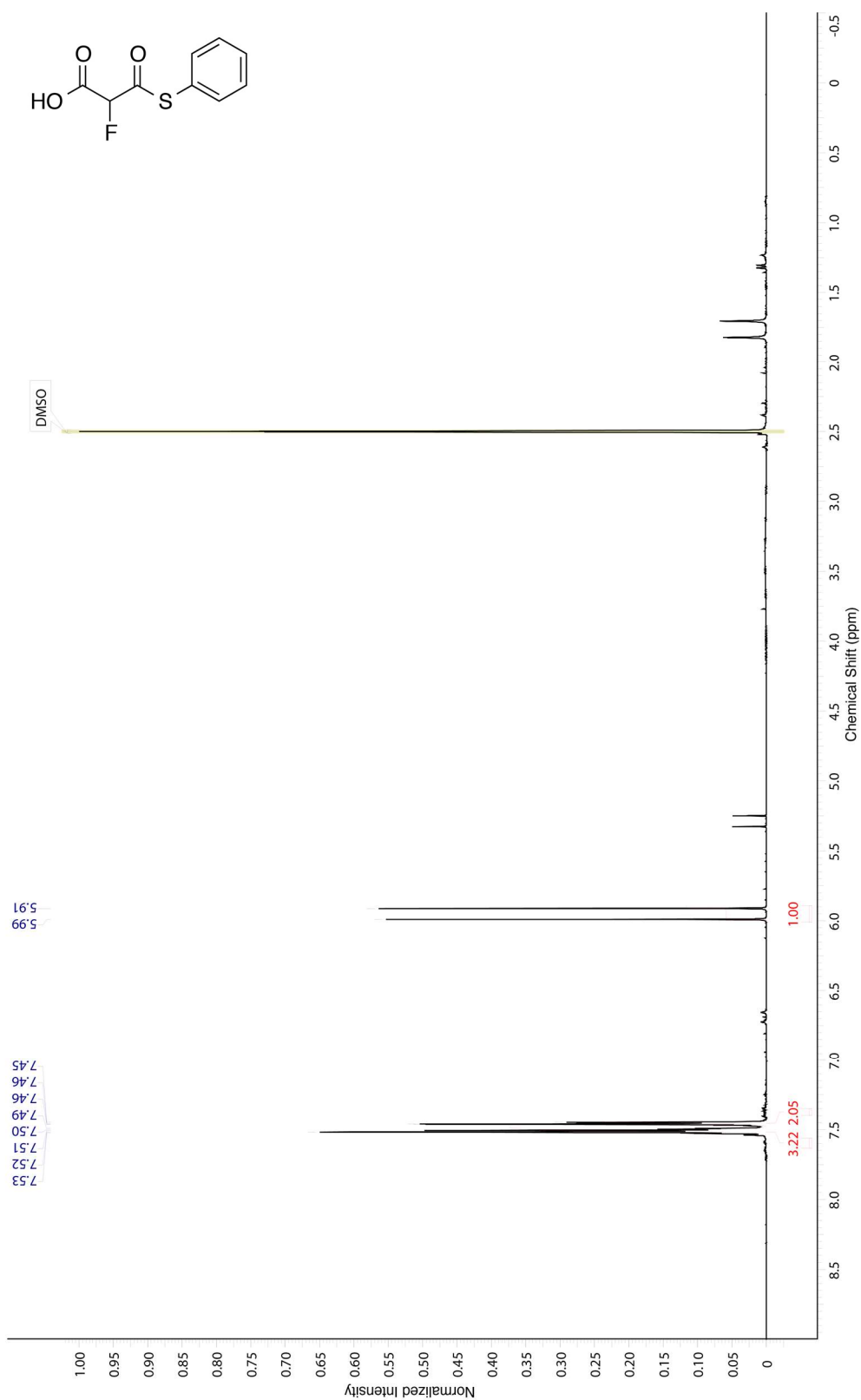
**Figure S22** |  $^1\text{H-NMR}$  spectrum of 6-MSA-SNAC was performed on a 250 MHz instrument (Bruker) with the solvent methanol- $d_4$ .



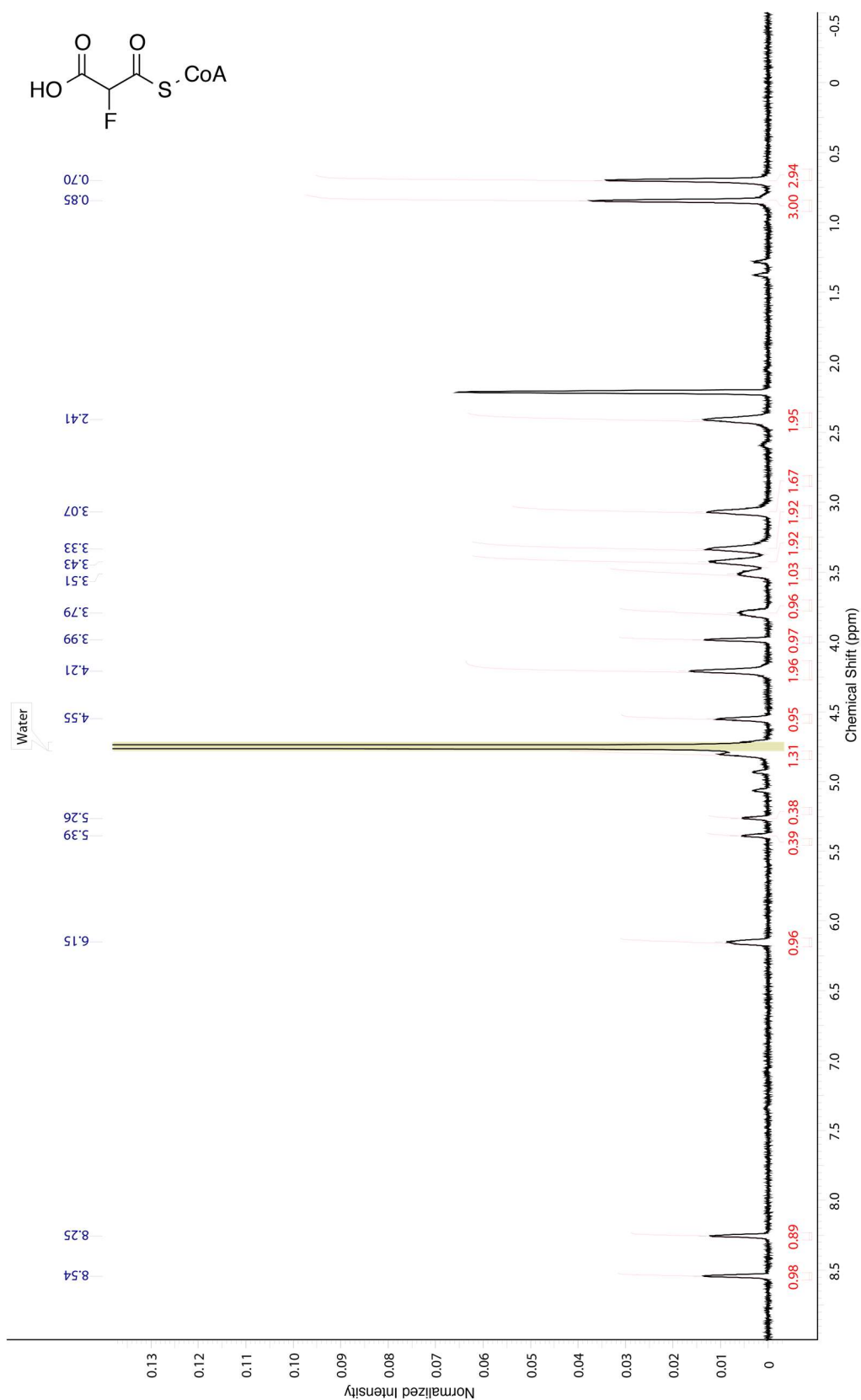
**Figure S23** | <sup>1</sup>H-NMR spectrum of natural diketide SNAC (**2**) was performed on a 400 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



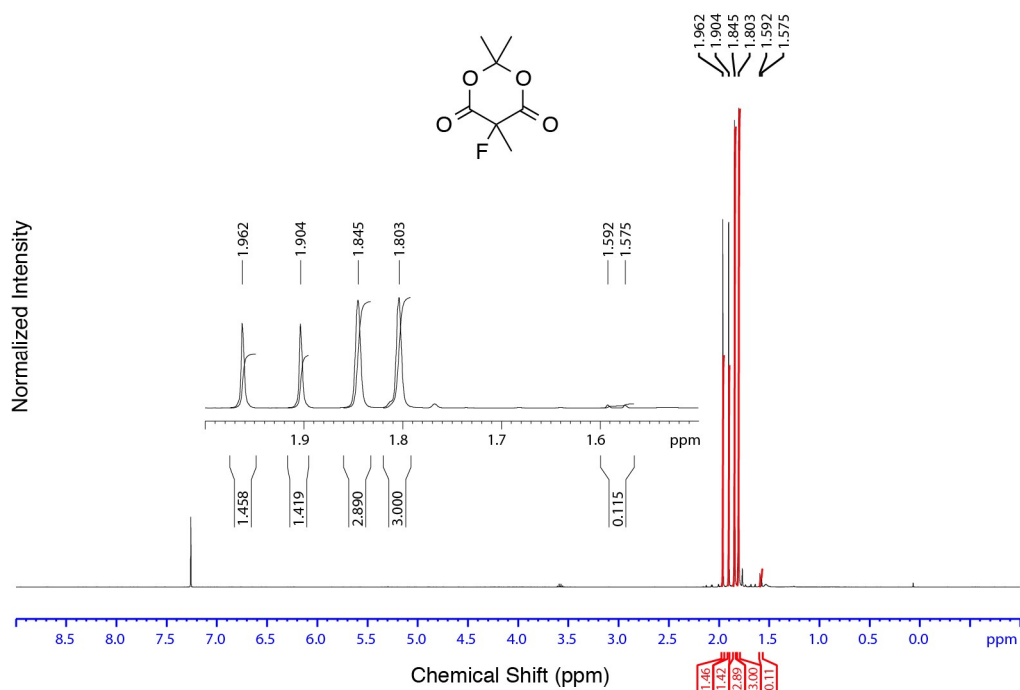
**Figure S24** |  $^1\text{H-NMR}$  spectrum of fluoro-Meldrum's acid was performed on a 400 MHz instrument (Bruker) with the solvent  $\text{C}_2\text{D}_6\text{OS}$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



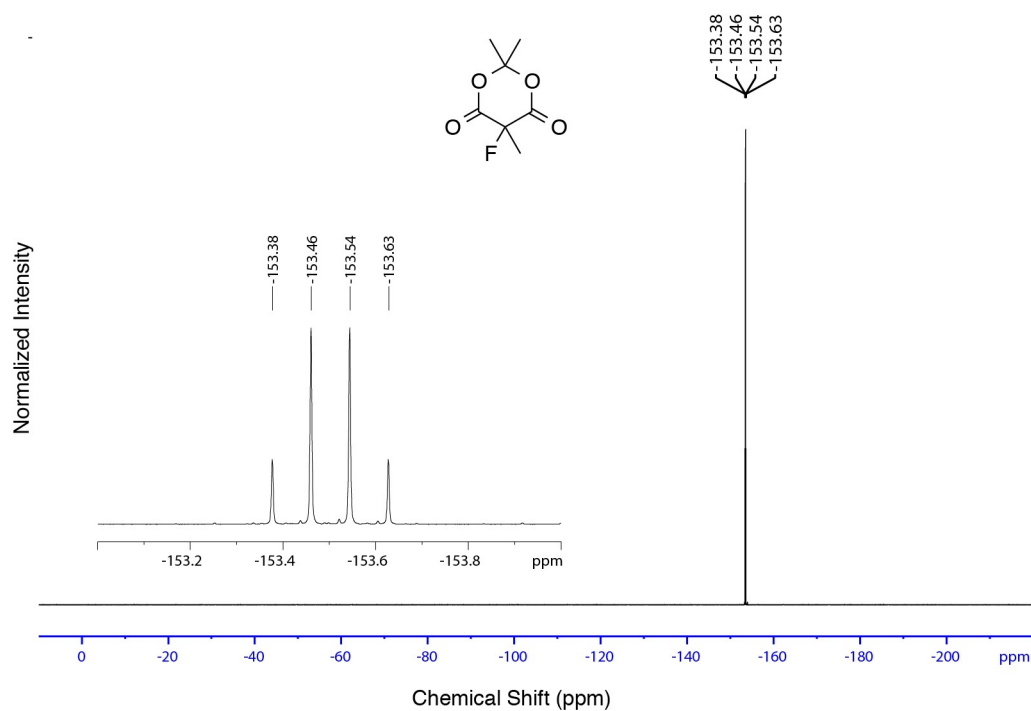
**Figure S25** | <sup>1</sup>H-NMR spectrum of fluoromalonyl-thiophenyl halfester was performed on a 400 MHz instrument (Bruker) with the solvent C<sub>2</sub>D<sub>6</sub>OS. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



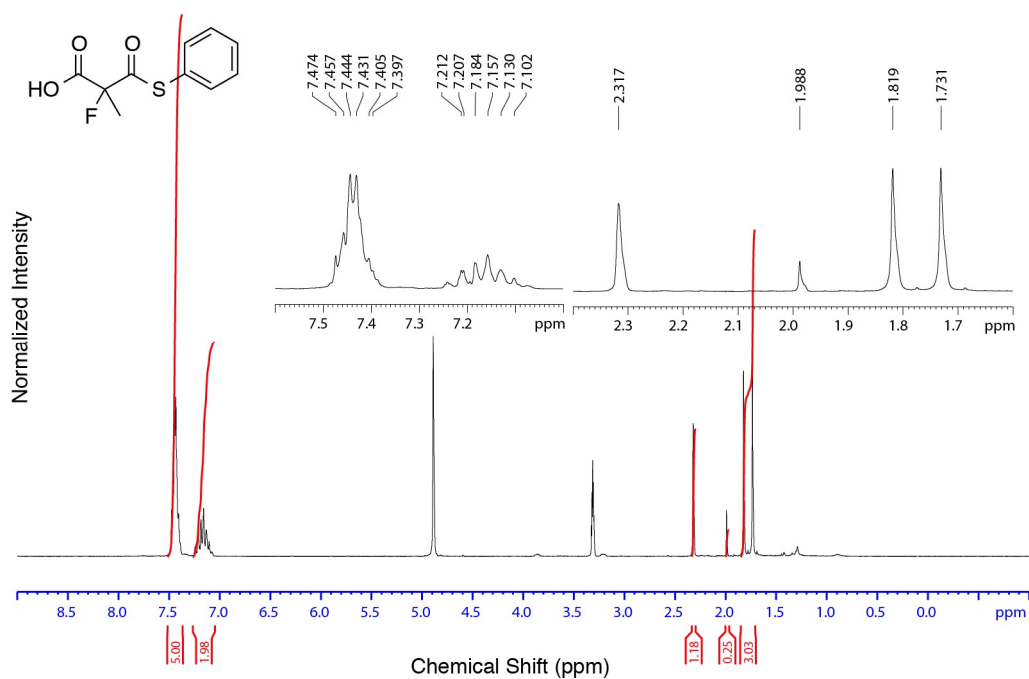
**Figure S26** | <sup>1</sup>H-NMR spectrum of F-Mal-CoA was performed on a 400 MHz instrument (Bruker) with the solvent D<sub>2</sub>O. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



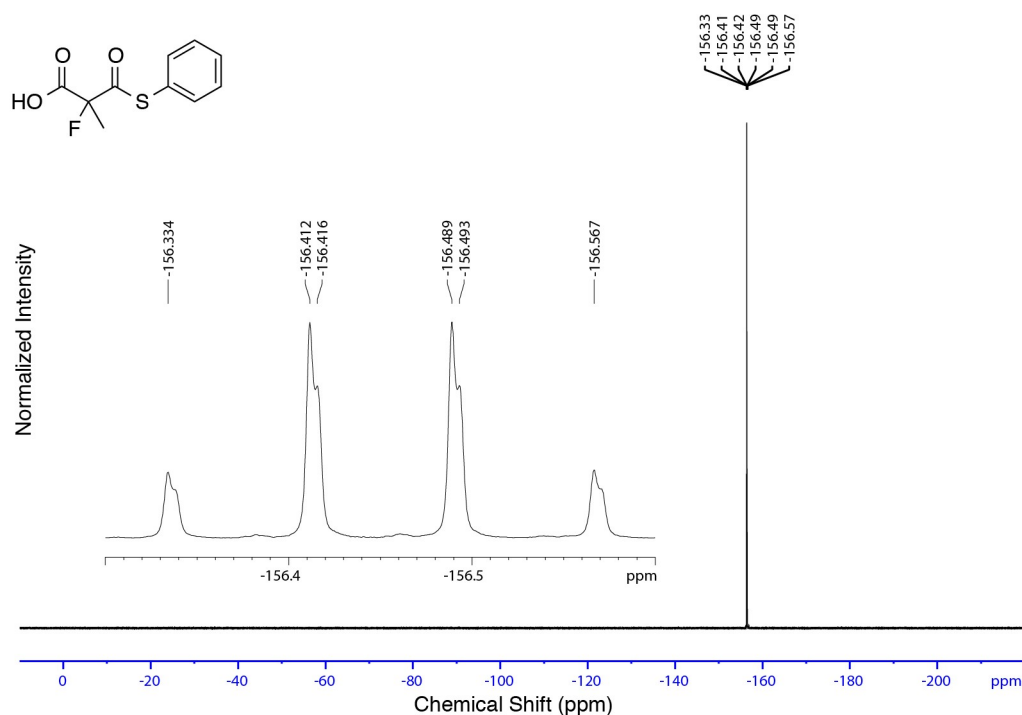
**Figure S27** | <sup>1</sup>H-NMR spectrum of 5-fluoro-5-methyl-Meldrum's acid was performed on a 400 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S28** | <sup>19</sup>F-NMR spectrum of 5-fluoro-5-methyl-Meldrum's acid was performed on a 300 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>

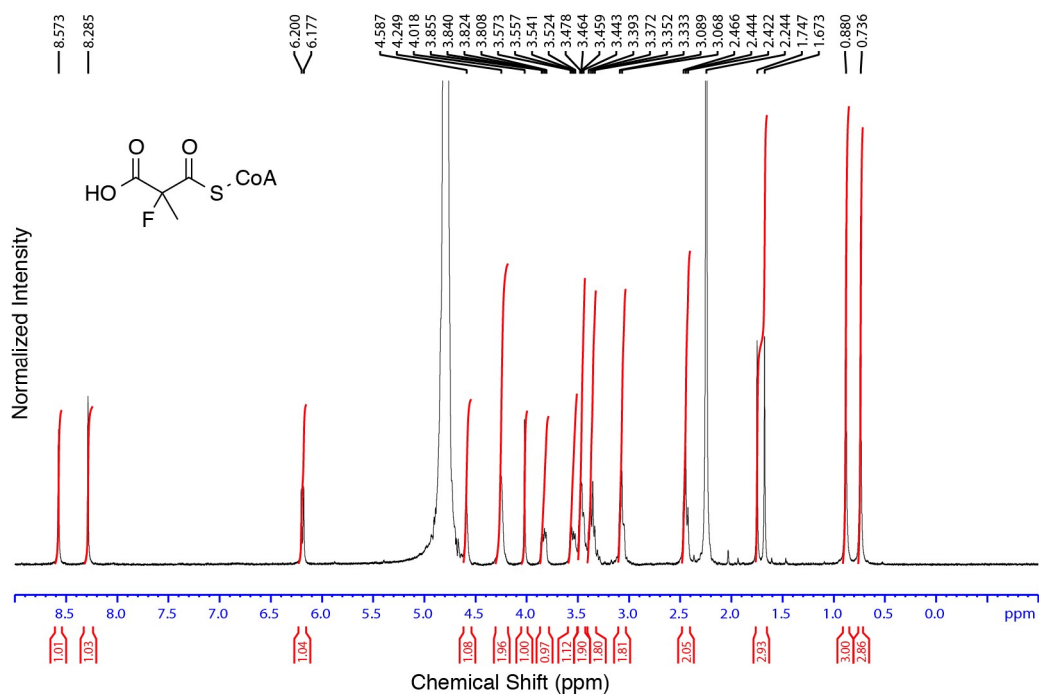


**Figure S29** | <sup>1</sup>H-NMR spectrum of 2-fluoro-2-methylmalonyl-thiophenyl halfester was performed on a 250 MHz instrument (Bruker) with the solvent methanol-4d. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>

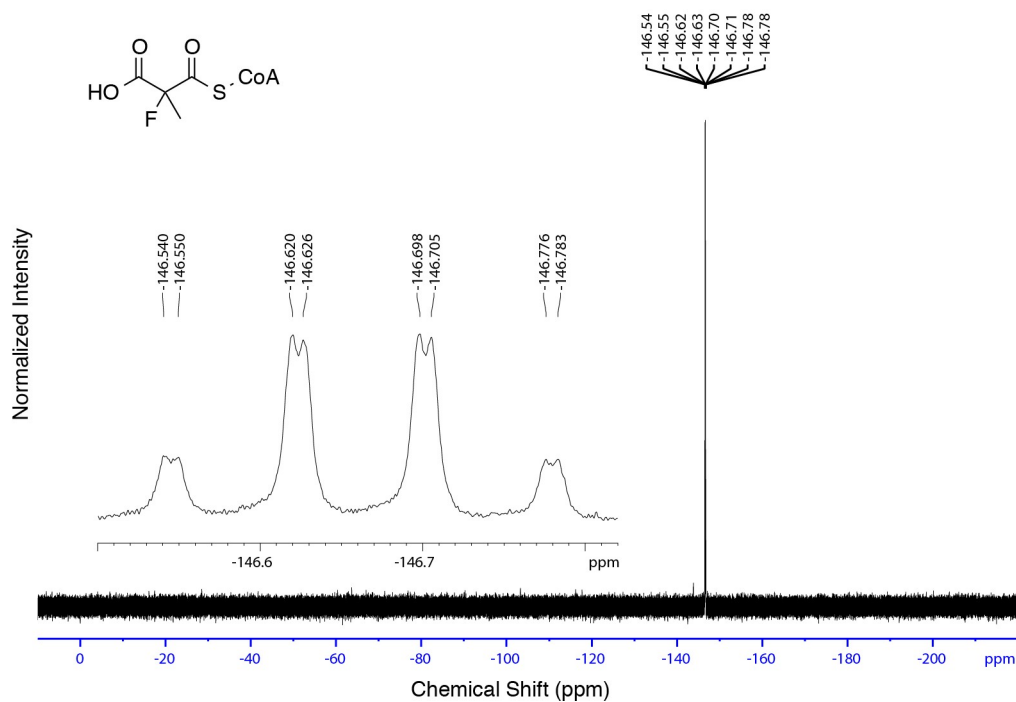


**Figure S30** | <sup>19</sup>F-NMR spectrum of 2-fluoro-2-methylmalonyl-thiophenyl halfester was performed on a 300 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>

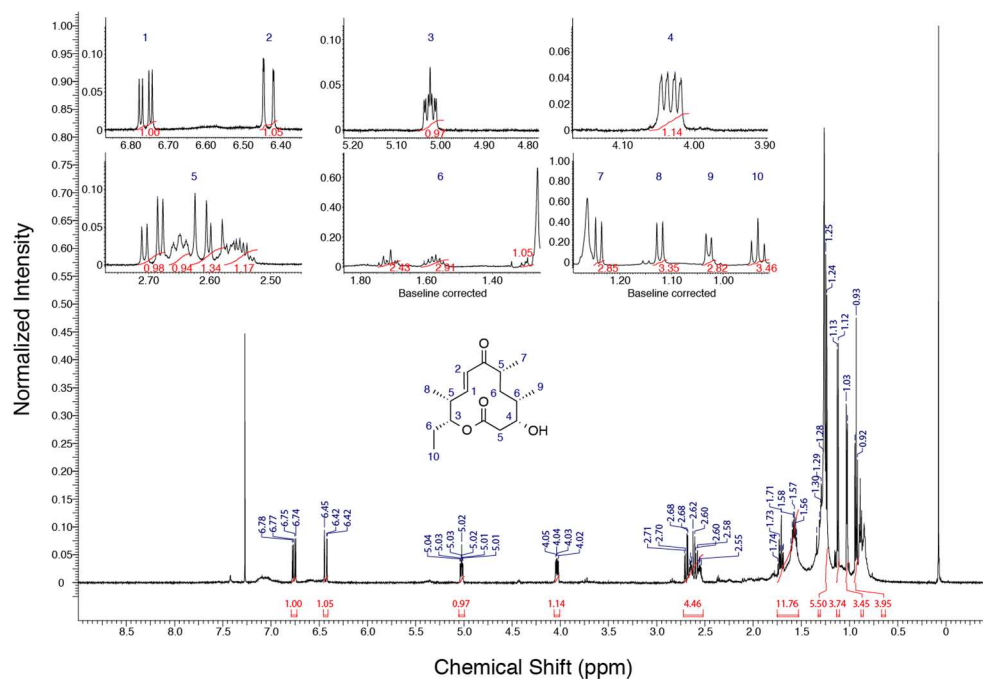




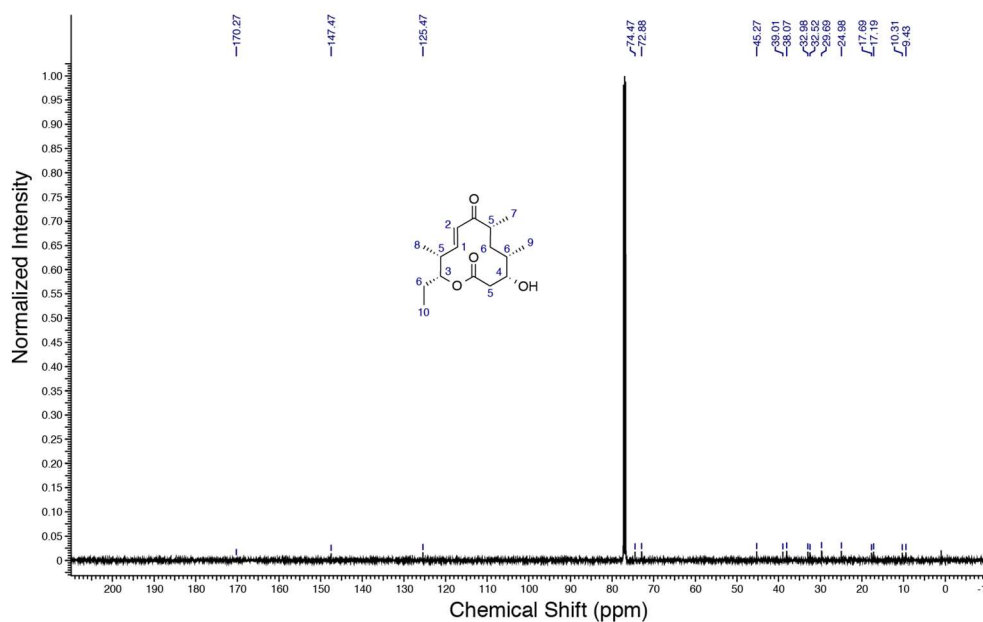
**Figure S31** |  $^1\text{H-NMR}$  spectrum of F-MM-CoA was performed on a 300 MHz instrument (Bruker) with the solvent  $\text{D}_2\text{O}$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



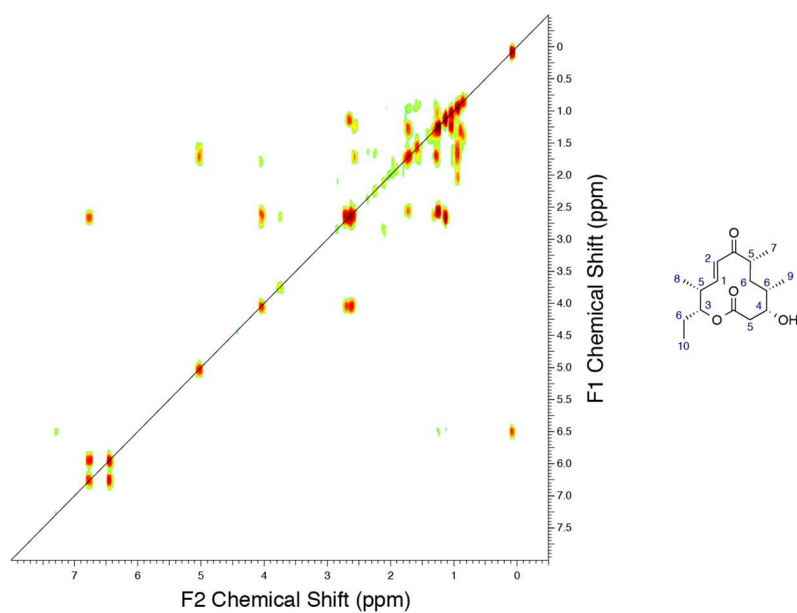
**Figure S32** |  $^{19}\text{F-NMR}$  spectrum of F-MM-CoA was performed on a 300 MHz instrument (Bruker) with the solvent  $\text{D}_2\text{O}$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



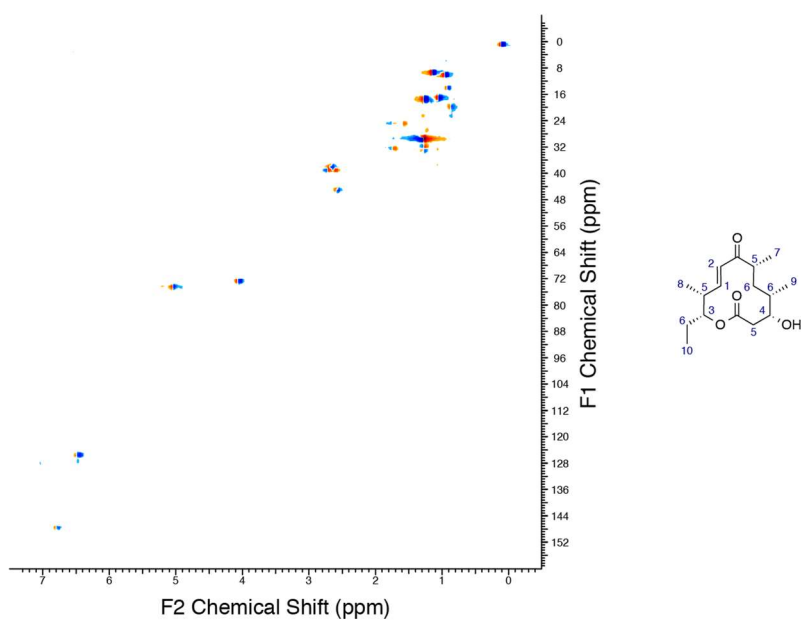
**Figure S33** | <sup>1</sup>H-NMR spectrum of compound **16** was performed on a 600 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



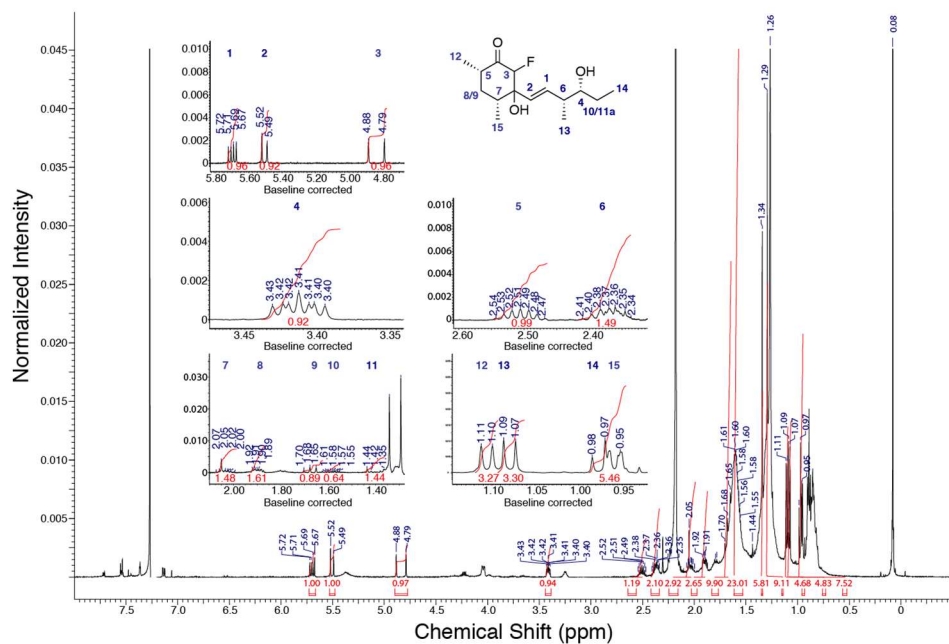
**Figure S34** | <sup>13</sup>C-NMR spectrum of compound **16** was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



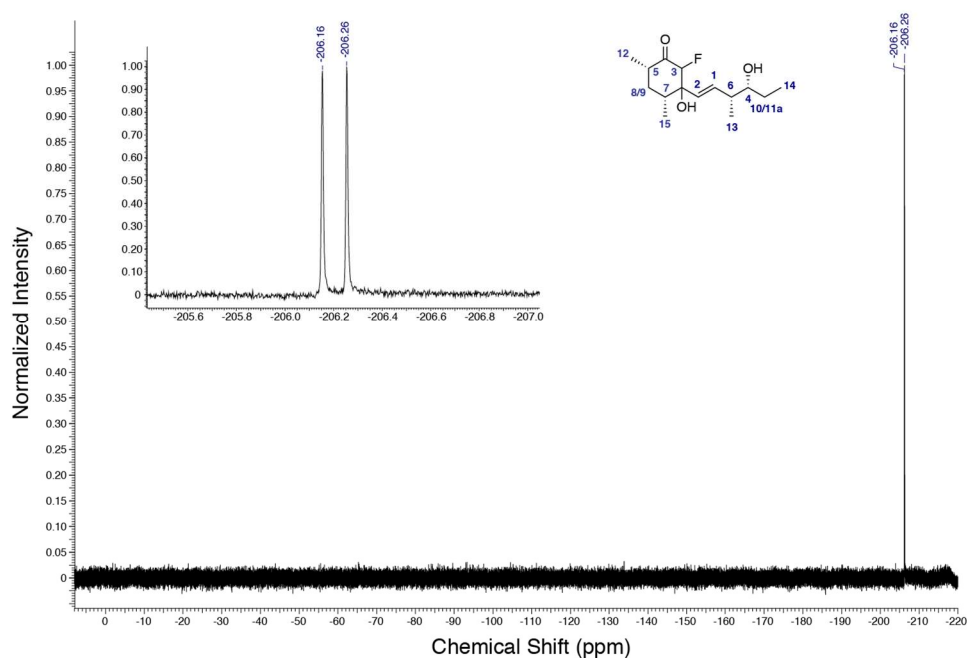
**Figure S35** | 2D-COSY spectrum of compound **16** was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



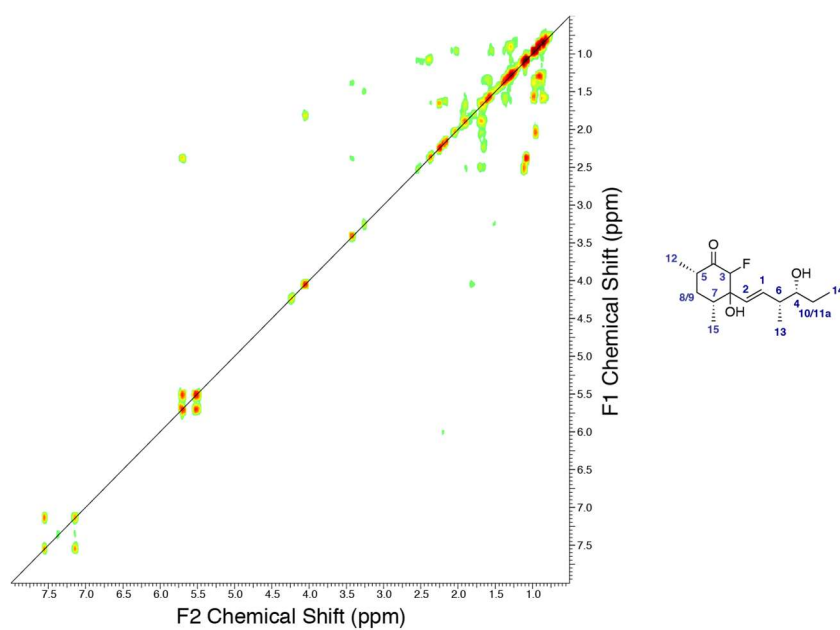
**Figure S36** | 2D-HSQC spectrum of compound **16** was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



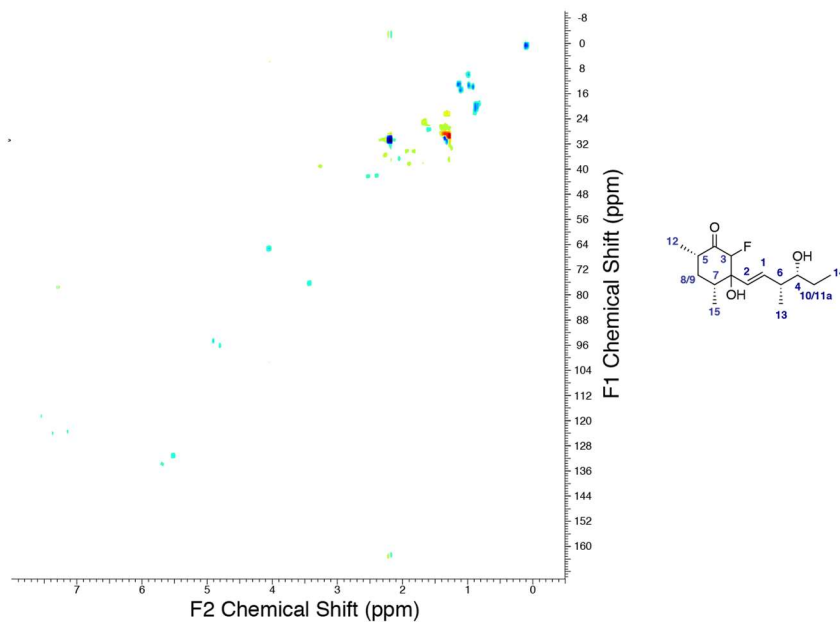
**Figure S37** |  $^1\text{H-NMR}$  spectrum of compound **17** was performed on a 500 MHz instrument (Bruker) with the solvent  $\text{CDCl}_3$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



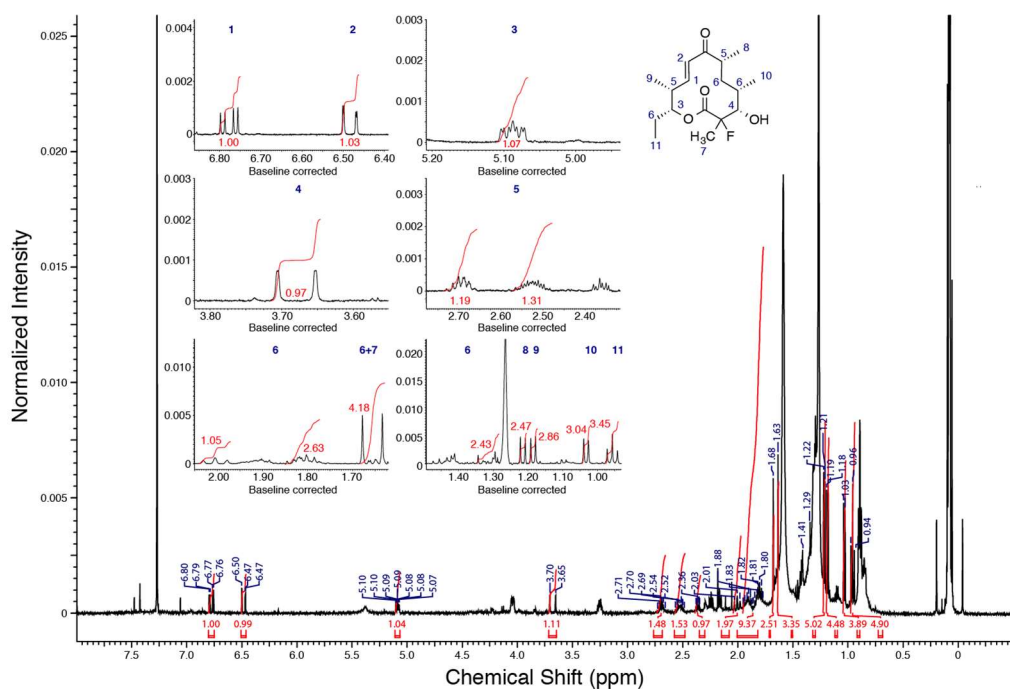
**Figure S38** |  $^{19}\text{F-NMR}$  spectrum of compound **17** was performed on a 500 MHz instrument (Bruker) with the solvent  $\text{CDCl}_3$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



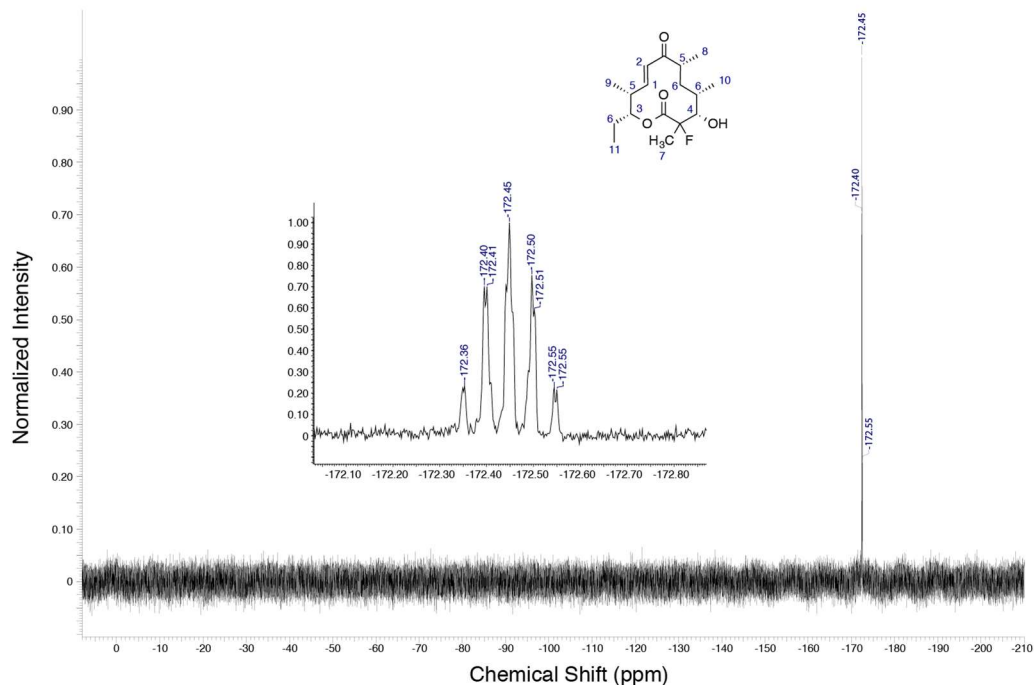
**Figure S39** | 2D-COSY spectrum of compound **17** was performed on a 500 MHz instrument (Bruker) with the solvent  $\text{CDCl}_3$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



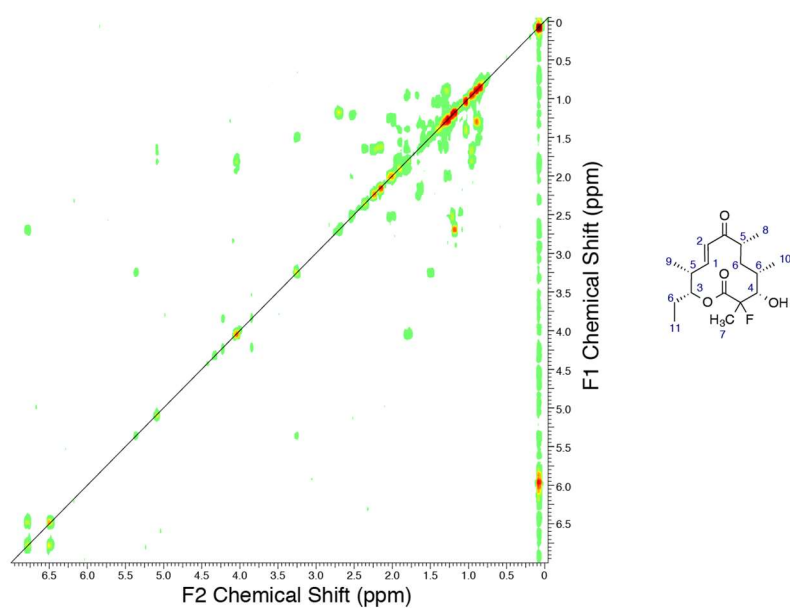
**Figure S40** | 2D-HSQC spectrum of compound **17** was performed on a 500 MHz instrument (Bruker) with the solvent  $\text{CDCl}_3$ . NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



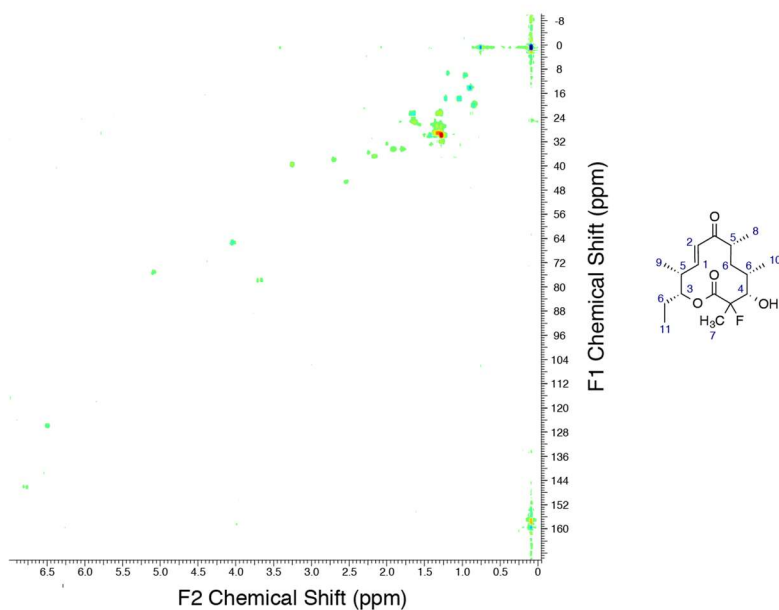
**Figure S41** | <sup>1</sup>H-NMR spectrum of compound 27 was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S42** | <sup>19</sup>F-NMR spectrum of compound 27 was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S43** | 2D-COSY spectrum of compound **27** was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>



**Figure S44** | 2D-HSQC spectrum of compound **27** was performed on a 500 MHz instrument (Bruker) with the solvent CDCl<sub>3</sub>. NMR spectrum was published by Rittner and Joppe *et al.*<sup>44</sup>

## 7.4 Primers and plasmids

**Table S1** | Plasmid list with used primers and applied annealing temperature. Asterisks indicate codon optimized genes for the organism *E. coli*. Plasmid sequences were confirmed by sequencing (Seqlab)

Plasmid number	Construct name	Used primers with applied annealing temperature in brackets
pMJD063	pET22b_StrepII_SerChIB3_His(8)	All primers had the number PRMJD063 ChIB3_fwd+ ChIB3_rev for insert (62°C) pET22b_fwd+ pET22b_rev for vector (62°C)
pMJD065	pET22b_SerMSAS_His(8)	All primers had the number PRMJD061 noStrep_serMSAS fwd+noStrep_serMSAS rev (55°C)
pMJD066	pET22b_ChIB2_His(8)	PRMJD066+PRMJD067 for vector (65°C) PRMJD068+PRMJD069 for insert (64°C)
pMJD068	pET22b_SerChIB3_His(8)	PRMJD074+PRMJD075 (63°C)
pMJD069	pET22b_MBP_SerChIB3_His(8)	PRMJD076+PRMJD077 for insert (63°C) PRMJD078+PRMJD079 for vector (61°C)
pMJD076 H2	pET22b_DEBS_M6_TE_mFASm_LDAT	PRMJD101+PRMJD102 for vector (62-68°C) PRMJD087+PRMJD088 for insert (62°C)
pMJD077 H1	pET22b_DEBS_M6_TE_mFASm_AT	PRMJD105+PRMJD106 for vector (62-68°C) PRMJD091+PRMJD092 for insert (62°C)
pMJD078	pET22b_StrepII_SerChIB3*_His(8)	PRMJD107+PRMJD108 for insert (62°C) PRMJD109+PRMJD110 (62°C)
pMJD079	pET22b_SerChIB3*_His(8)	PRMJD111+PRMJD108 for insert (62°C) PRMJD109+PRMJD112 for vector (62°C)
pMJD080	pET22b_StrepII_SerChIB3*	PRMJD107+PRMJD108 for insert (62°C) PRMJD113+PRMJD110 (62°C)
pMJD090	pET22b_SerChIB3*_Linker_His(8)	Was only used as a cloning template PRMJD135+MJD136 (56°C)
pMJD091	pCDF_Npt	PRMJD136+PRMJD137 for insert (62°C) PRMJD138+PRMJD139 for vector (64°C)
pMJD094	pET28a_DEBS M6-ACP-W-H8	MJD145+MJD146 (65°C)
pMJD095	pRS313_FAS2_(S1827E)	PRMJD147+PRMJD148 (59°C)
pMJD096	pRS313_FAS2_(S1827E/S1440E)	PRMJD149+PRMJD150 (61°C)
pMJD097	pRS313_FAS2_(S1827A)	PRMJD151+AR26 (63°C), PRMJD152+AR27 (59°C)
pMJD098	pRS313_FAS2_(S1827A/S1440A)	PRMJD153+AR26 (61°C), PRMJD154+AR27 (63°C)
MF319d or pMF011	pRS313_FAS2 ( $\alpha$ -chain)	Kindly received from Manuel Fischer
MF639k1 or pMF014	pRS315_FAS1_Strep ( $\beta$ -chain)	Kindly received from Manuel Fischer
pMF025	pRS315_FAS1_Strep $\Delta$ 2-137 ( $\beta$ -chain)	MF717_for+AR26 (59°C) MF717_rev+AR27 (60°C)
pBAR1	pRS313_FAS2 $\Delta$ 541-598 ( $\alpha$ -chain)	pBAR001_for+AR26 (60°C) pBAR002_rev+AR27 (60°C)
MW001	pET22b_StrepII_SerMSAS_His(8)	Kindly received from Jan Gajewski
pFS26	pET22b_StrepI_SerMSAS-ACP_His(8)	Kindly received from Franziska Stegemann
pFS123	pET22b_StrepI_SerMSAS-ACP*_His(8)	Kindly received from Franziska Stegemann
pAR014	pET22b_PenPaMSAS_His(10)	Kindly received from Alexander Rittner
pAR070	pET22b_StrepI-mFAS-KS(C161G)-MAT-H8	Kindly received from Alexander Rittner <sup>34</sup>
pAR264	StrepI_Not1_mFASm_H8_pET22b	Kindly received from Alexander Rittner <sup>34</sup>
pAR352	pET22b_StrepII-mACP-H8-RBS-Sfp	
pAR357	pCDF-1b_Sfp	Kindly received from Alexander Rittner
pAR432	DD2_DEBS M6-KS(C1661G)-AT6-H6	Created by Alexander Rittner AR719+AR722 AR721+AR720
pMK60	pET22b_MBP_DEBSACP1_His(8)	Kindly received from Maja Klaus
pISI011	pET22b_PenPaMSAS_KS-AT_His(8)	PRMJD157+PriSI012-rev for vector (59°C) PriSI036-f+PriSI037-r for insert (59°C)
pISI014	pET22b_PenPa_ACP_native-2_L_His(8)	PRMJD157+PriSI012-rev for vector (59°C) PriSI042-f+PriSI041-r for insert (61°C)



pSI017	pET22b_PenPaMSAS_KS-AT_H(8)_K688T/P757G	PrSI050_fwd+PrSI049_rev (60°C) PrSI051_fwd+PrSI052_rev (60°C)
pSI018	pET22b_PenPaMSAS_KS-AT_H(8)_K688A/P757A	PrSI050_fwd+PrSI049_rev (60°C) PrSI053_fwd+PrSI054_rev (60°C)
pSI019	pET22b_PenPaMSAS_KS-AT_H(8)_Q625A/I752V	PrSI056_fwd+PrSI055_rev (60°C) PrSI057_fwd+PrSI058_rev (60°C)
pSI020	pET22b_PenPaMSAS_KS-AT_H(8)_Q625F/I752L	PrSI056_fwd+PrSI055_rev (60°C) PrSI059_fwd+PrSI060_rev (60°C)
pSI021	pET22b_PenPaMSAS_KS-AT_H(8)_Y682Q/F754A	PrSI062_fwd+PrSI061_rev (60°C) PrSI063_fwd+PrSI064_rev (60°C)
pSI022	pET22b_PenPaMSAS_KS-AT_H(8)_Y682L/F754S	PrSI062_fwd+PrSI061_rev (60°C) PrSI065_fwd+PrSI066_rev (60°C)
pSI023	pET22b_PenPaMSAS_KS-AT_H(8)_I715A/M806V	PrSI068_fwd+PrSI067_rev (60°C) PrSI069_fwd+PrSI070_rev (60°C)
pSI024	pET22b_PenPaMSAS_KS-AT_H(8)_I715G/M806A	PrSI068_fwd+PrSI067_rev (60°C) PrSI071_fwd+PrSI072_rev (60°C)
pSI025	pET22b_PenPaMSAS_KS-AT_H(8)_N808S/R809P	PrSI073_fwd+PrSI074_rev (60°C)
pSI026	pET22b_PenPaMSAS_KS-AT_H(8)_VNR807-809SST	PrSI073_fwd+PrSI075_rev (60°C)
pBL18	DEBS M6	Kindly received from the Khosla lab
pMEH004	PenPaMSAS_KS-AT_H(8)_S653G	PRMEH005+AR27 (67°C) PRMEH002+AR26 (67°C)
pMEH005	PenPaMSAS_KS-AT_H(8)_C204G	PRMEH006+AR27 (65°C) PRMEH004+AR26 (65°C)
pMEH006	PenPaMSAS_KS-AT_H(8)_C204G/S653G	PRMEH006+PRMEH002 (65°C) PRMEH005+PRMEH004 (65°C)

Table S2 | Primer list

Primer number	Primer sequence
PRMJD061 noStrep_serMSAS fwd	CATATGAGCACGCAGCCGG
PRMJD061 noStrep_serMSAS rev	CTGCGTGCTCATATGTATATCTCCTTCTTAAAGTTAAAC
PRMJD063_ ChIB3 fwd	GCGCCGGATCCCCTGCTTCCGAAACGTTTCATC
PRMJD063_ ChIB3 rev	GAGCTTGACCGAGTCGGCGAC
PRMJD063_ pET22b fwd	GACTCGGTCAAGCTCGAGCATCATCACCAC
PRMJD063_ pET22b rev	CACGGGATCCGGCGCCTTTTTT
PRMJD066	GAAGCGAACTCGAGCATCATCACCACC
PRMJD067	GTCTACGCGCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTTC
PRMJD068	CATATGCGCGTAGACGAGATCGAG
PRMJD069	CTCGAGTTTCGCTTCGCTCGACAG
PRMJD074	GAGATATACATATGCGTGCTTCCGAAACGTTTC
PRMJD075	GATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTTC
PRMJD076	CGGATCCCCTGCTTC
PRMJD077	CATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTTC
PRMJD078	GGAGATATACATATGAAATCGAAGAAGGTAAACTG
PRMJD079	GAAGCACGGGATCCGTAGTGTGCCG
PRMJD087	acacggcaggccc
PRMJD088	aggagtcctcgggg
PRMJD091	aacaagcggcactctg
PRMJD092	tgtgaggtgcaccttgc

## Chapter 7: Appendix

PRMJD101	ccccgagggactcctCTGCCCAACTACCCGTTTCGAG
PRMJD102	caggggcctgccgtgtCGGGGGCTCGGCGAT
PRMJD105	caaggtgcacctcacaGGCGTGGCCGTGGAC
PRMJD106	gagtgggcgcttggGAGGCGGTTCCGGTG
PRMJD107	CGTGCTTCTGAAACCTTCATCTC
PRMJD108	TTAACAGAGTCAGCAACGATACG
PRMJD109	GCTGACTCTGTAAACTCGAGCATCATCACCACC
PRMJD110	GGTTTCAGAAGCACGGGATCCGGCGCCTTTTTTC
PRMJD111	ATGCGTGCTTCTGAAACCTTC
PRMJD112	TTCAGAAGCACGCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTTC
PRMJD113	GCTGACTCTGTAAATGAGATCCGGCTGCTAACAAAG
PRMJD134	GAACCTCCGCCAGAACCTCCCTCGAGTTTAAACAGAGTCAGC
PRMJD135	TTCTGGCGGAGGTTCTGGCCATCATCACCACCACC
PRMJD136	TAATTAACCTAGGCTGCTGCCAC
PRMJD137	GGTATATCTCCTTATTAAAGTTAAACAAAATTATTTTC
PRMJD138	ATAAGGAGATATACCATGATTGAGAAGTTACTCC
PRMJD139	AGCCTAGGTTAAttactCGAGTGCGGCCCG
PRMJD145	CCGCGCGGCAGCCATATGTGGGCGGCCCCCG
PRMJD146	ACGGAGCTCGAATTCAGAGCTGCTGTCCTATGTGGTCCG
PRMJD147	GTCAAGGAATTAGGCGGTGGTG
PRMJD148	GCCTAATTCCTTGACGCCTAAG
PRMJD149	GTATGCTGAACCAAACCTTGAACATG
PRMJD150	GTTTGGTTCAGCATACTTAACACTG
PRMJD151	GTCAAGGCTTTAGGCGGTGGTG
PRMJD152	GCCTAAAGCCTTGACGCCTAAG
PRMJD153	GTATGCTGCTCCAAACTTGAACATG
PRMJD154	GTTTGGAGCAGCATACTTAACACTG
PRMJD157	CTCGAGGGAGGTTCTGGC
AR26	GAGGACCGAAGGAGCTAACC
AR27	GGTTAGCTCCTTCGGTCCTC
AR719	CTCGAGCACCACCACCAC
AR720	GTGGTGGTGCTCGAGGCTGTCGGCGAGCTG
AR721	CACGGTGGACACGGCGGGCTCGTCGTCGTTGGTGG
AR722	CGCCGTGTCCACCG
MF717_for	ATGGACAAAAAATCCAACCTCTGCTC
MF717_rev	GGATTTTTTGTCCATAATGAGCGAATAATAGAGC
pBAR001	AGAATCTATTCCGTTCTTACATTTGAGAAAGAAG
pBAR00	AACGGAATAGATTCTTTGGTGATTGGACCAC
PRMEH002	GTGTCCAATCACGGCCTGTG
PRMEH004	CGCCGCATCCACGGC
PRMEH005	GCCGTGATTGGACACGGTGTGGGCGAAATCGCGGC
PRMEH006	GCCGTGGATGCGGCGGGAGCTTCGTCTCTCGTCGCGC
PrISI012	ATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTTC
PrISI036-f	GAAGGAGATATACATATGCATTCCGCTGCAAC
PrISI037-r	AGAACCTCCCTCGAGACCAGTGTGAAGCGGC

PrISI041-r	AGAACCTCCCTCGAGTTTGGCAAGCTTCTCAGC
PrISI042-f	GAAGGAGATATACATATGGGTTCCAGCGACGCG
PrISI049_rev	GCCATAACCTGGCGGTAC
PrISI050_fwd	ACCTTGAATGGGCTGGTCG
PrISI051_fwd	CGCCAGGTTATGGGCACCGGCGGCATGATCCTTGTC
PrISI052_rev	CAGCCATTCAAGGTGCCGCTGTGGAAGGCAATATCACTC
PrISI053_fwd	CGCCAGGTTATGGGCGCGGGCGGCATGATCCTTGTC
PrISI054_rev	CAGCCATTCAAGGTGCCGCTGTGGAAGGCAATATCACTC
PrISI055_rev	TACCCGATCAGACGACTCAAAG
PrISI056_fwd	GCCTTCCACAGCCCAAC
PrISI057_fwd	TCGTCTGATCGGGTAGCGATCCTTACCTATGTCATGCAAATTGG
PrISI058_rev	TGGGCTGTGGAAGGCCACATCACTCTTCACAGTGAATGTCTTC
PrISI059_fwd	TCGTCTGATCGGGTATTTATCCTTACCTATGTCATGCAAATTGG
PrISI060_rev	TGGGCTGTGGAAGGCCAGATCACTCTTCACAGTGAATGTCTTC
PrISI061_rev	CAGCAACGCTCTGCGG
PrISI062_fwd	CACAGCCCAACCTTGAATGG
PrISI063_fwd	CGCAGAGCGTTGCTGCAGCGCCAGGTTATGGGCAAG
PrISI064_rev	CAAGGTTGGGCTGTGCGCGCAATATCACTCTTCACAGTG
PrISI065_fwd	CGCAGAGCGTTGCTGCTGCGCCAGGTTATGGGCAAG
PrISI066_rev	CAAGGTTGGGCTGTGGCTGGCAATATCACTCTTCACAGTG
PrISI067_rev	GGCCACAACCAGGTCTGATC
PrISI068_fwd	GTCAACCGCGTGCGTCTC
PrISI069_fwd	GACCTGGTTGTGGCCGCGGACTCATCGCCATCTTCTCG
PrISI070_rev	ACGCACGCGGTTGACCACGTTGCCGGCCAGTATTC
PrISI071_fwd	GACCTGGTTGTGGCCGCGGACTCATCGCCATCTTCTCG
PrISI072_rev	ACGCACGCGGTTGACCACGTTGCCGGCCAGTATTC
PrISI073_fwd	GTGCGTCTCACTTCGGCTG
PrISI074_rev	CGAAGTGAGACGCACCGGGCTGACCATGTTGCCGGCCC
PrISI075_rev	CGAAGTGAGACGCACCGGTGCTGCTCATGTTGCCGGCCAGTATTC

## 7.5 Literature

1. Wang, S., Dong, G. & Sheng, C. Structural Simplification of Natural Products. *Chem. Rev.* **119**, 4180–4220 (2019).
2. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod* **34** (2020).
3. Talele, T. T. Natural-Products-Inspired Use of the *gem* -Dimethyl Group in Medicinal Chemistry. *J. Med. Chem.* **61**, 2166–2210 (2018).
4. Hagmann, W. K. The Many Roles for Fluorine in Medicinal Chemistry. *J. Med. Chem.* **51**, 4359–4369 (2008).
5. Fernandes, P., Martens, E., Bertrand, D. & Pereira, D. The solithromycin journey—It is all in the chemistry. *Bioorg. Med. Chem.* **24**, 6420–6428 (2016).
6. Rozatian, N., Beeby, A., Ashworth, I. W., Sandford, G. & Hodgson, D. R. W. Enolization rates control mono- versus di-fluorination of 1,3-dicarbonyl derivatives. *Chem. Sci.* **10**, 10318–10330 (2019).
7. Burkhard, J. A., Wuitschik, G., Rogers-Evans, M., Müller, K. & Carreira, E. M. Oxetanes as Versatile Elements in Drug Discovery and Synthesis. *Angew. Chem. Int. Ed.* **49**, 9052–9067 (2010).
8. Beld, J., Lee, D. J. & Burkart, M. D. Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Mol. Biosyst.* **11**, 38–59 (2015).
9. Herbst, D. A., Townsend, C. A. & Maier, T. The architectures of iterative type I PKS and FAS. *Nat. Prod. Rep.* **35**, 1046–1069 (2018).
10. Heil, C. S., Wehrheim, S. S., Paithankar, K. S. & Grininger, M. Fatty Acid Biosynthesis: Chain-Length Regulation and Control. *Chembiochem* **20**, 2298–2321 (2019).

11. Marella, E. R., Holkenbrink, C., Siewers, V. & Borodina, I. Engineering microbial fatty acid metabolism for biofuels and biochemicals. *Curr. Opin. Biotechnol.* **50**, 39–46 (2018).
12. Kang, M.-K. & Nielsen, J. Biobased production of alkanes and alkenes through metabolic engineering of microorganisms. *J. Ind. Microbiol. Biotechnol.* **44**, 613–622 (2017).
13. Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chem. Int. Ed.* **48**, 4688–4716 (2009).
14. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18**, 380–416 (2001).
15. Nivina, A., Yuet, K. P., Hsu, J. & Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases: Focus Review. *Chem. Rev.* **119**, 12524–12547 (2019).
16. Weissman, K. J. Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.* **33**, 203–230 (2016).
17. Klaus, M. & Grninger, M. Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **35**, 1070–1081 (2018).
18. Ang, E. L., Sun, H., Liu, Z. & Zhao, H. Recent advances in combinatorial biosynthesis drug discovery. *Drug Des. Devel. Ther.* **9**, 823–833 (2015).
19. Müller, R. Don't Classify Polyketide Synthases. *Chem. Biol.* **11**, 4–6 (2004).
20. Smith, S. & Tsai, S.-C. The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat. Prod. Rep.* **24**, 1041–1072 (2007).

21. Berg, J. M., Tymoczko, J. L., Gatto, G. J. & Stryer, L. *Stryer Biochemie*. (Springer Berlin Heidelberg, 2018). doi:10.1007/978-3-662-54620-8.
22. *The IUPAC Compendium of Chemical Terminology: The Gold Book*. (International Union of Pure and Applied Chemistry (IUPAC), 2019). doi:10.1351/goldbook.
23. Rittner, A., Paithankar, K. S., Drexler, D. J., Himmler, A. & Grninger, M. Probing the modularity of megasynthases by rational engineering of a fatty acid synthase Type I: Probing the Modularity of Megasynthases by Rational Engineering of a Fatty Acid Synthase Type I. *Protein Sci.* **28**, 414–428 (2019).
24. Cane, D. E., Walsh, C. T. & Khosla, C. Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63–8 (1998).
25. Lowry, B., Li, X., Robbins, T., Cane, D. E. & Khosla, C. A Turnstile Mechanism for the Controlled Growth of Biosynthetic Intermediates on Assembly Line Polyketide Synthases. *ACS Cent. Sci.* **2**, 14–20 (2016).
26. Menzella, H. G. *et al.* Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **23**, 1171–1176 (2005).
27. Musiol-Kroll, E. & Wohlleben, W. Acyltransferases as Tools for Polyketide Synthase Engineering. *Antibiotics* **7**, 1–35 (2018).
28. Yuzawa, S. *et al.* Comprehensive *in Vitro* Analysis of Acyltransferase Domain Exchanges in Modular Polyketide Synthases and Its Application for Short-Chain Ketone Production. *ACS Synth. Biol.* **6**, 139–147 (2017).

29. Koryakina, I. *et al.* Inversion of Extender Unit Selectivity in the Erythromycin Polyketide Synthase by Acyltransferase Domain Engineering. *ACS Chem. Biol.* **12**, 114–123 (2017).
30. Walker, M. C. *et al.* Expanding the Fluorine Chemistry of Living Systems Using Engineered Polyketide Synthase Pathways. *Science* **341**, 1089–1094 (2013).
31. Kalkreuter, E., CroweTipton, J. M., Lowell, A. N., Sherman, D. H. & Williams, G. J. Engineering the Substrate Specificity of a Modular Polyketide Synthase for Installation of Consecutive Non-Natural Extender Units. *J. Am. Chem. Soc.* **141**, 1961–1969 (2019).
32. Zhang, F. *et al.* Structural Insights into the Substrate Specificity of Acyltransferases from Salinomycin Polyketide Synthase. *Biochemistry* **58**, 2978–2986 (2019).
33. Hol, W. Effects of the  $\alpha$ -helix dipole upon the functioning and structure of proteins and peptides. *Adv. Biophys.* **19**, 133–165 (1985).
34. Rittner, A., Paithankar, K. S., Huu, K. V. & Grninger, M. Characterization of the Polyspecific Transferase of Murine Type I Fatty Acid Synthase (FAS) and Implications for Polyketide Synthase (PKS) Engineering. *ACS Chem. Biol.* **13**, 723–732 (2018).
35. Finzel, K., Lee, D. J. & Burkart, M. D. Using Modern Tools To Probe the Structure-Function Relationship of Fatty Acid Synthases. *ChemBioChem* **16**, 528–547 (2015).
36. Jenni, S. *et al.* Structure of Fungal Fatty Acid Synthase and Implications for Iterative Substrate Shuttling. *Science* **316**, 254–261 (2007).

37. Pirson, W., Schuhmann, L. & Lynen, F. The Specificity of Yeast Fatty-Acid Synthetase with Respect to the 'Priming' Substrate. Decanoyl-CoA and Derivatives as 'Primers' of Fatty-Acid Synthesis in vitro. *Eur. J. Biochem.* **36**, 16–24 (1973).
38. Dimroth, P., Ringelmann, E. & Lynen, F. 6-Methylsalicylic Acid Synthetase from *Penicillium patulum*. Some Catalytic Properties of the Enzyme and Its Relation to Fatty Acid Synthetase. *Eur. J. Biochem.* **68**, 591–596 (1976).
39. Richardson, M. T., Pohl, N. L., Kealey, J. T. & Khosla, C. Tolerance and Specificity of Recombinant 6-Methylsalicylic Acid Synthase. *Metab. Eng.* **1**, 180–187 (1999).
40. Wang, F. *et al.* Structural and Functional Analysis of the Loading Acyltransferase from Avermectin Modular Polyketide Synthase. *ACS Chem. Biol.* **10**, 1017–1025 (2015).
41. Liou, G. F., Lau, J., Cane, D. E. & Khosla, C. Quantitative Analysis of Loading and Extender Acyltransferases of Modular Polyketide Synthases †. *Biochemistry* **42**, 200–207 (2003).
42. Khosla, C., Gokhale, R. S., Jacobsen, J. R. & Cane, D. E. Tolerance and Specificity of Polyketide Synthases. *Annu. Rev. Biochem.* **68**, 219–253 (1999).
43. Blaquiere, N., Shore, D. G., Rousseaux, S. & Fagnou, K. Decarboxylative ketone aldol reactions: development and mechanistic evaluation under metal-free conditions. *J Org Chem* **74**, 6190–8 (2009).
44. Rittner, A. *et al.* *Directed biosynthesis of fluorinated polyketides*.  
<http://biorxiv.org/lookup/doi/10.1101/2021.07.30.454469> (2021)  
doi:10.1101/2021.07.30.454469.



45. Witkowski, A., Joshi, A. K. & Smith, S. Characterization of the Interthiol Acyltransferase Reaction Catalyzed by the  $\beta$ -Ketoacyl Synthase Domain of the Animal Fatty Acid Synthase †. *Biochemistry* **36**, 16338–16344 (1997).
46. Rittner, A., Paithankar, K. S., Aaron Himmler & Grininger, M. Type I fatty acid synthase trapped in the octanoyl-bound state. *Protein Sci.* **29**, 589–605 (2020).
47. Horsman, M. E., Hari, T. P. A. & Boddy, C. N. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Nat. Prod. Rep.* **33**, 183–202 (2016).
48. Hansen, D. A., Koch, A. A. & Sherman, D. H. Identification of a Thioesterase Bottleneck in the Pikromycin Pathway through Full-Module Processing of Unnatural Pentaketides. *J. Am. Chem. Soc.* **139**, 13450–13455 (2017).
49. Grininger, M. Perspectives on the evolution, assembly and conformational dynamics of fatty acid synthase type I (FAS I) systems. *Curr. Opin. Struct. Biol.* **25**, 49–56 (2014).
50. McCarthy, A. D. & Hardie, D. G. Fatty acid synthase — an example of protein evolution by gene fusion. *Trends Biochem. Sci.* **9**, 60–63 (1984).
51. Sweetlove, L. J. & Fernie, A. R. The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nat. Commun.* **9**, 1–12 (2018).
52. Bukhari, H. S. T., Jakob, R. P. & Maier, T. Evolutionary Origins of the Multienzyme Architecture of Giant Fungal Fatty Acid Synthase. *Structure* **22**, 1775–1785 (2014).
53. Maier, T., Leibundgut, M. & Ban, N. The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* **321**, 1315–1322 (2008).

54. Johansson, P. *et al.* Multimeric Options for the Auto-Activation of the *Saccharomyces cerevisiae* FAS Type I Megasyntase. *Structure* **17**, 1063–1074 (2009).
55. Boehringer, D., Ban, N. & Leibundgut, M. 7.5-Å Cryo-EM Structure of the Mycobacterial Fatty Acid Synthase. *J. Mol. Biol.* **425**, 841–849 (2013).
56. Lomakin, I. B., Xiong, Y. & Steitz, T. A. The Crystal Structure of Yeast Fatty Acid Synthase, a Cellular Machine with Eight Active Sites Working Together. *Cell* **129**, 319–332 (2007).
57. Gipson, P. *et al.* Direct structural insight into the substrate-shuttling mechanism of yeast fatty acid synthase by electron cryomicroscopy. *Proc. Natl. Acad. Sci.* **107**, 9164–9169 (2010).
58. Lou, J. W., Iyer, K. R., Hasan, S. M. N., Cowen, L. E. & Mazhab-Jafari, M. T. Electron cryomicroscopy observation of acyl carrier protein translocation in type I fungal fatty acid synthase. *Sci. Rep.* **9**, 1–8 (2019).
59. Fischer, M. *et al.* Cryo-EM structure of fatty acid synthase (FAS) from *Rhodospiridium toruloides* provides insights into the evolutionary development of fungal FAS: Structure of *Rhodospiridium Toruloides* FAS I. *Protein Sci.* **24**, 987–995 (2015).
60. Leibundgut, M., Jenni, S., Frick, C. & Ban, N. Structural Basis for Substrate Delivery by Acyl Carrier Protein in the Yeast Fatty Acid Synthase. *Science* **316**, 288–290 (2007).
61. Johansson, P. *et al.* Inhibition of the fungal fatty acid synthase type I multienzyme complex. *Proc. Natl. Acad. Sci.* **105**, 12803–12808 (2008).

62. Fichtlscherer, F., Wellein, C., Mittag, M. & Schweizer, E. A novel function of yeast fatty acid synthase: Subunit  $\alpha$  is capable of self-pantetheinylation. *Eur. J. Biochem.* **267**, 2666–2671 (2000).
63. Shiber, A. *et al.* Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* **561**, 268–272 (2018).
64. Schuller, H.-J., Fortsch, B., Rautenstrauss, B., Wolf, D. H. & Schweizer, E. Differential proteolytic sensitivity of yeast fatty acid synthetase subunits alpha and beta contributing to a balanced ratio of both fatty acid synthetase components. *Eur. J. Biochem.* **203**, 607–614 (1992).
65. Ploskon, E. *et al.* A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *J Biol Chem* **283**, 518–528 (2008).
66. Chakravarty, B., Gu, Z., Chirala, S. S., Wakil, S. J. & Quiocho, F. A. Human fatty acid synthase: Structure and substrate selectivity of the thioesterase domain. *Proc Natl Acad Sci U A* **101**, 15567–15572 (2004).
67. Rangan, V. S., Joshi, A. K. & Smith, S. Mapping the functional topology of the animal fatty acid synthase by mutant complementation in vitro. *Biochemistry* **40**, 10792–10799 (2001).
68. Brignole, E. J., Smith, S. & Asturias, F. J. Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat. Struct. Mol. Biol.* **16**, 190–197 (2009).
69. Asturias, F. J. *et al.* Structure and molecular organization of mammalian fatty acid synthase. *Nat Struct Mol Biol* **12**, 225–232 (2005).

70. Rangan, V. S. & Smith, S. Alteration of the Substrate Specificity of the Malonyl-CoA/Acetyl-CoA:Acyl Carrier Protein S -Acyltransferase Domain of the Multifunctional Fatty Acid Synthase by Mutation of a Single Arginine Residue. *J. Biol. Chem.* **272**, 11975–11978 (1997).
71. Tóth-Petróczy, Á. & Tawfik, D. S. The robustness and innovability of protein folds. *Curr. Opin. Struct. Biol.* **26**, 131–138 (2014).
72. Parascandolo, J. S. *et al.* Insights into 6-Methylsalicylic Acid Bio-assembly by Using Chemical Probes. *Angew Chem Int Ed Engl* **55**, 3463–3467 (2016).
73. Herbst, D. A., Jakob, R. P., Zähringer, F. & Maier, T. Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **531**, 533–537 (2016).
74. Tang, Y., Kim, C.-Y., Mathews, I. I., Cane, D. E. & Khosla, C. The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci.* **103**, 11124–11129 (2006).
75. Wang, J. *et al.* Structural basis for the biosynthesis of lovastatin. *Nat. Commun.* **12**, 1–10 (2021).
76. Dutta, S. *et al.* Structure of a modular polyketide synthase. *Nature* **510**, 512–517 (2014).
77. Malinowski, J. T., Sharpe, R. J. & Johnson, J. S. Enantioselective Synthesis of Pactamycin, a Complex Antitumor Antibiotic. *Science* **340**, 180–182 (2013).
78. Pache, W. & Chapman, D. Interaction of antibiotics with membranes: Chlorothricin. *Biochim. Biophys. Acta BBA - Biomembr.* **255**, 348–357 (1972).
79. Hanada, M. *et al.* Maduropeptin, a complex of new macromolecular antitumor antibiotics. *J. Antibiot. (Tokyo)* **44**, 403–414 (1991).

80. Li, B. *et al.* Dissection of patulin biosynthesis, spatial control and regulation mechanism in *Penicillium expansum*. *Environ. Microbiol.* **21**, 1124–1139 (2019).
81. Dimroth, P., Walter, H. & Lynen, F. Biosynthese von 6-Methylsalicylsäure. *Eur. J. Biochem.* **13**, 98–110 (1970).
82. Moriguchi, T., Kezuka, Y., Nonaka, T., Ebizuka, Y. & Fujii, I. Hidden Function of Catalytic Domain in 6-Methylsalicylic Acid Synthase for Product Release. *J. Biol. Chem.* **285**, 15637–15643 (2010).
83. He, Q.-L. *et al.* Dissection of Two Acyl-Transfer Reactions Centered on Acyl-S-Carrier Protein Intermediates for Incorporating 5-Chloro-6-methyl- O -methylsalicyclic Acid into Chlorothricin. *ChemBioChem* **10**, 813–819 (2009).
84. Yi, X. *et al.* Insights into the Functionalization of the Methylsalicyclic Moiety during the Biosynthesis of Chlorothricin by Comparative Kinetic Assays of the Activities of Two KAS III-like Acyltransferases. *Chin. J. Chem.* **37**, 821–826 (2019).
85. Jia, X.-Y. *et al.* Genetic Characterization of the Chlorothricin Gene Cluster as a Model for Spirotetronate Antibiotic Biosynthesis. *Chem. Biol.* **13**, 575–585 (2006).
86. Artigot, M. P. *et al.* Molecular cloning and functional characterization of two CYP619 cytochrome P450s involved in biosynthesis of patulin in *Aspergillus clavatus*. *Microbiology* **155**, 1738–1747 (2009).
87. Zhang, Q., Pang, B., Ding, W. & Liu, W. Aromatic Polyketides Produced by Bacterial Iterative Type I Polyketide Synthases. *ACS Catal.* **3**, 1439–1447 (2013).
88. Kage, H. *et al.* Chemical chain termination resolves the timing of ketoreduction in a partially reducing iterative type I polyketide synthase. *Org. Biomol. Chem.* **13**, 11414–11417 (2015).

89. Sun, H. *et al.* Synthesis of ( *R* )-Mellein by a Partially Reducing Iterative Polyketide Synthase. *J. Am. Chem. Soc.* **134**, 11924–11927 (2012).
90. Soehano, I. *et al.* Insights into the programmed ketoreduction of partially reducing polyketide synthases: stereo- and substrate-specificity of the ketoreductase domain. *Org Biomol Chem* **12**, 8542–8549 (2014).
91. Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C. & Barr, P. J. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci.* **95**, 505–509 (1998).
92. Hitschler, J. & Boles, E. De novo production of aromatic m-cresol in *Saccharomyces cerevisiae* mediated by heterologous polyketide synthases combined with a 6-methylsalicylic acid decarboxylase. *Metab. Eng. Commun.* **9**, e00093 (2019).
93. Hitschler, J., Grininger, M. & Boles, E. Substrate promiscuity of polyketide synthase enables production of tsetse fly attractants 3-ethylphenol and 3-propylphenol by engineering precursor supply in yeast. *Sci. Rep.* **10**, 1–11 (2020).
94. Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J. & Leadlay, P. F. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **348**, 176–178 (1990).
95. Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A. & Cane, D. E. Structure and Mechanism of the 6-Deoxyerythronolide B Synthase. *Annu. Rev. Biochem.* **76**, 195–221 (2007).
96. Xue, Y., Zhao, L., Liu, H. -w. & Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity. *Proc. Natl. Acad. Sci.* **95**, 12111–12116 (1998).

97. Witkowski, A., Joshi, A. K., Lindqvist, Y. & Smith, S. Conversion of a  $\beta$ -Ketoacyl Synthase to a Malonyl Decarboxylase by Replacement of the Active-Site Cysteine with Glutamine †. *Biochemistry* **38**, 11643–11650 (1999).
98. Gajewski, J. *et al.* Engineering fatty acid synthases for directed polyketide production. *Nat. Chem. Biol.* **13**, 363–365 (2017).
99. Gajewski, J., Pavlovic, R., Fischer, M., Boles, E. & Grininger, M. Engineering fungal de novo fatty acid synthesis for short chain fatty acid production. *Nat. Commun.* **8**, 14650 (2017).
100. Zhu, Z. *et al.* Expanding the product portfolio of fungal type I fatty acid synthases. *Nat. Chem. Biol.* **13**, 360–362 (2017).
101. Koch, A. A. *et al.* A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J. Am. Chem. Soc.* **139**, 13456–13465 (2017).
102. Baader, S., Podsiadly, P. E., Cole-Hamilton, D. J. & Goossen, L. J. Synthesis of tsetse fly attractants from a cashew nut shell extract by isomerising metathesis. *Green Chem* **16**, 4885–4890 (2014).
103. Kennedy, P. G. E. Update on human African trypanosomiasis (sleeping sickness). *J. Neurol.* **266**, 2334–2337 (2019).
104. Schmidt, T. G. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.* **2**, 1528–1535 (2007).
105. Shpilka, T. *et al.* Fatty acid synthase is preferentially degraded by autophagy upon nitrogen starvation in yeast. *Proc Natl Acad Sci U A* **112**, 1434–9 (2015).

106. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289–298 (2006).
107. Kolodziej, S. J., Penczek, P. A., Schroeter, J. P. & Stoops, J. K. Structure-Function Relationships of the *Saccharomyces cerevisiae* Fatty Acid Synthase: THREE-DIMENSIONAL STRUCTURE. *J. Biol. Chem.* **271**, 28422–28429 (1996).
108. Wieland, F., Renner, L., Verfurth, C. & Lynen, F. Studies on the Multi-Enzyme Complex of Yeast Fatty-Acid Synthetase. Reversible Dissociation and Isolation of Two Polypeptide Chains. *Eur. J. Biochem.* **94**, 189–197 (1979).
109. Joppe, M. *et al.* The resolution revolution in cryoEM requires high-quality sample preparation: a rapid pipeline to a high-resolution map of yeast fatty acid synthase. *IUCrJ* **7**, 220–227 (2020).
110. Kuhlbrandt, W. The Resolution Revolution. *Science* **343**, 1443–1444 (2014).
111. Chari, A. *et al.* ProteoPlex: stability optimization of macromolecular complexes by sparse-matrix screening of chemical space. *Nat. Methods* **12**, 859–865 (2015).
112. Zhu, S. *et al.* Structure of a human synaptic GABAA receptor. *Nature* **559**, 67–72 (2018).
113. Kim, Y. & Chen, J. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science* **359**, 915–919 (2018).
114. Takizawa, Y. *et al.* Cryo-EM structure of the nucleosome containing the *ALB1* enhancer DNA sequence. *Open Biol.* **8**, 170255 (2018).
115. Bai, X. *et al.* An atomic structure of human  $\gamma$ -secretase. *Nature* **525**, 212–217 (2015).



116. Yan, Z. *et al.* Structure of the Na v 1.4- $\beta$ 1 Complex from Electric Eel. *Cell* **170**, 470-482.e11 (2017).
117. D’Imprima, E. *et al.* Protein denaturation at the air-water interface and how to prevent it. *eLife* **8**, e42747 (2019).
118. Glaeser, R. M. & Han, B.-G. Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophys. Rep.* **3**, 1–7 (2017).
119. Noble, A. J. *et al.* Routine single particle CryoEM sample and grid characterization by tomography. *eLife* **7**, e34257 (2018).
120. Taylor, K. A. & Glaeser, R. M. Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. *J. Struct. Biol.* **163**, 214–223 (2008).
121. Geim, A. K. & Novoselov, K. S. The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007).
122. Sader, K., Stopps, M., Calder, L. J. & Rosenthal, P. B. Cryomicroscopy of radiation sensitive specimens on unmodified graphene sheets: Reduction of electron-optical effects of charging. *J. Struct. Biol.* **183**, 531–536 (2013).
123. Pantelic, R. S., Meyer, J. C., Kaiser, U. & Stahlberg, H. The application of graphene as a sample support in transmission electron microscopy. *Solid State Commun.* **152**, 1375–1382 (2012).
124. Russo, C. J. & Passmore, L. A. Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas. *Nat. Methods* **11**, 649–652 (2014).

125. Pantelic, R. S., Fu, W., Schoenenberger, C. & Stahlberg, H. Rendering graphene supports hydrophilic with non-covalent aromatic functionalization for transmission electron microscopy. *Appl. Phys. Lett.* **104**, 134103 (2014).
126. Li, X. *et al.* Large-scale phosphorylation analysis of alpha-factor-arrested *Saccharomyces cerevisiae*. *J Proteome Res* **6**, 1190–7 (2007).
127. Ptacek, J. *et al.* Global analysis of protein phosphorylation in yeast. *Nature* **438**, 679–684 (2005).
128. Martínez-Montañés, F. *et al.* Phosphoproteomic Analysis across the Yeast Life Cycle Reveals Control of Fatty Acyl Chain Length by Phosphorylation of the Fatty Acid Synthase Complex. *Cell Rep.* **32**, 108024 (2020).
129. Bunkoczi, G. *et al.* Mechanism and Substrate Recognition of Human Holo ACP Synthase. *Chem. Biol.* **14**, 1243–1253 (2007).
130. Fischer, M. *et al.* Analysis of the co-translational assembly of the fungal fatty acid synthase (FAS). *Sci. Rep.* **10**, 1–13 (2020).
131. Alexander Rill. Assemblierung der Fettsäuresynthase in Pilzen unter biotechnologischen Gesichtspunkten. (Johann Wolfgang Goethe-Universität Frankfurt am Main, 2017).
132. Egner, R. *et al.* Tracing intracellular proteolytic pathways. Proteolysis of fatty acid synthase and other cytoplasmic proteins in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* **268**, 27269–76 (1993).
133. Oliynyk, M. *et al.* Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat. Biotechnol.* **25**, 447–453 (2007).

134. Spencer, J. B. & Jordan, P. M. Purification and properties of 6-methylsalicylic acid synthase from *Penicillium patulum*. *Biochem. J.* **288**, 839–846 (1992).
135. Lowry, B. *et al.* *In Vitro* Reconstitution and Analysis of the 6-Deoxyerythronolide B Synthase. *J. Am. Chem. Soc.* **135**, 16809–16812 (2013).
136. Christian Gusenda. Analyse einer iterativen Polyketidsynthase aus Bakterien. (Johann Wolfgang Goethe-Universität Frankfurt am Main, 2018).
137. Trinetri Saraswati Maria Goel. Synthese von aromatischen Verbindungen über iterative Polyketidsynthese. (Johann Wolfgang Goethe-Universität Frankfurt am Main, 2019).
138. Petersen, L. M., Holm, D. K., Gotfredsen, C. H., Mortensen, U. H. & Larsen, T. O. Investigation of a 6-MSA Synthase Gene Cluster in *Aspergillus aculeatus* Reveals 6-MSA-derived Aculinic Acid, Aculins A-B and Epi-Aculin A. *ChemBioChem* **16**, 2200–2204 (2015).
139. Nofiani, R., Philmus, B., Nindita, Y. & Mahmud, T. 3-Ketoacyl-ACP synthase (KAS) III homologues and their roles in natural product biosynthesis. *MedChemComm* **10**, 1517–1530 (2019).
140. Bretschneider, T. *et al.* A ketosynthase homolog uses malonyl units to form esters in cervimycin biosynthesis. *Nat. Chem. Biol.* **8**, 154–161 (2012).
141. Abugrain, M. E., Brumsted, C. J., Osborn, A. R., Philmus, B. & Mahmud, T. A Highly Promiscuous  $\beta$ -Ketoacyl-ACP Synthase (KAS) III-like Protein Is Involved in Pactamycin Biosynthesis. *ACS Chem. Biol.* **12**, 362–366 (2017).
142. Jackson, D. R. *et al.* Structural and Functional Studies of the Daunorubicin Priming Ketosynthase DpsC. *ACS Chem. Biol.* **13**, 141–151 (2018).

143. Davies, C., Heath, R. J., White, S. W. & Rock, C. O. The 1.8 Å crystal Structure and active-site architecture of  $\beta$ -ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* **8**, 185–195 (2000).
144. Qiu, X. *et al.* Crystal Structure of  $\beta$ -Ketoacyl-Acyl Carrier Protein Synthase III. *J. Biol. Chem.* **274**, 36465–36471 (1999).
145. Feng, Y. & Cronan, J. E. *Escherichia coli* Unsaturated Fatty Acid Synthesis. *J. Biol. Chem.* **284**, 29526–29535 (2009).
146. Edwards, P., Sabo Nelsen, J., Metz, J. G. & Dehesh, K. Cloning of the *fabF* gene in an expression vector and in vitro characterization of recombinant *fabF* and *fabB* encoded enzymes from *Escherichia coli*. *FEBS Lett.* **402**, 62–66 (1997).
147. von Wettstein-Knowles, P., Olsen, J. G., McGuire, K. A. & Henriksen, A. Fatty acid synthesis. Role of active site histidines and lysine in Cys-His-His-type beta-ketoacyl-acyl carrier protein synthases. *FEBS J.* **273**, 695–710 (2006).
148. Huang, W. *et al.* Crystal structure of  $\beta$ -ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes. *EMBO J* **5**, 1183–1191 (1998).
149. Bao, W., Sheldon, P. J. & Hutchinson, C. R. Purification and Properties of the *Streptomyces peucetius* DpsC  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase III that Specifies the Propionate-Starter Unit for Type II Polyketide Biosynthesis †. *Biochemistry* **38**, 9752–9757 (1999).
150. Sheldon, P. J., Wendt-Pienkowski, E. & Hutchinson, C. R. The *Streptomyces peucetius* dpsC Gene Determines the Choice of Starter Unit in Biosynthesis of the Daunorubicin Polyketide. *J. Bacteriol.* **181**, 4690–4695 (1999).

151. Xu, Z., Schenk, A. & Hertweck, C. Molecular Analysis of the Benastatin Biosynthetic Pathway and Genetic Engineering of Altered Fatty Acid–Polyketide Hybrids. *J. Am. Chem. Soc.* **129**, 6022–6030 (2007).
152. Pan, H. *et al.* Crystal Structure of the Priming  $\beta$ -Ketosynthase from the R1128 Polyketide Biosynthetic Pathway. *Structure* **10**, 1559–1568 (2002).
153. Meadows, E. S. & Khosla, C. In Vitro Reconstitution and Analysis of the Chain Initiating Enzymes of the R1128 Polyketide Synthase †. *Biochemistry* **40**, 14855–14861 (2001).
154. Freitag, A., Wemakor, E., Li, S.-M. & Heide, L. Acyl Transfer in Clorobiocin Biosynthesis: Involvement of Several Proteins in the Transfer of the Pyrrole-2-carboxyl Moiety to the Deoxysugar. *ChemBioChem* **6**, 2316–2325 (2005).
155. Xu, H., Kahlich, R., Kammerer, B., Heide, L. & Li, S.-M. CloN2, a novel acyltransferase involved in the attachment of the pyrrole-2-carboxyl moiety to the deoxysugar of clorobiocin. *Microbiology* **149**, 2183–2191 (2003).
156. Xiao, Y. *et al.* Characterization of Tiacumicin B Biosynthetic Gene Cluster Affording Diversified Tiacumicin Analogues and Revealing a Tailoring Dihalogenase. *J. Am. Chem. Soc.* **133**, 1092–1105 (2011).
157. Maina, C. V. *et al.* An Escherichia coli vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene* **74**, 365–373 (1988).
158. Raran-Kurussi, S., Keefe, K. & Waugh, D. S. Positional effects of fusion partners on the yield and solubility of MBP fusion proteins. *Protein Expr. Purif.* **110**, 159–164 (2015).

159. Christian Schreiber. Studie von iterativen Polyketidsynthesen. (Johann Wolfgang Goethe-Universität Frankfurt am Main, 2018).
160. Whicher, J. R. *et al.* Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **510**, 560–564 (2014).
161. Snini, S. P. *et al.* The gene PatG involved in the biosynthesis pathway of patulin, a food-borne mycotoxin, encodes a 6-methylsalicylic acid decarboxylase. *Int. J. Food Microbiol.* **171**, 77–83 (2014).
162. Bradbury, S. L. & Jakoby, W. B. Glycerol as an Enzyme-Stabilizing Agent: Effects on Aldehyde Dehydrogenase. *Proc. Natl. Acad. Sci.* **69**, 2373–2376 (1972).
163. Tang, S.-Y. *et al.* Screening for Enhanced Triacetic Acid Lactone Production by Recombinant *Escherichia coli* Expressing a Designed Triacetic Acid Lactone Reporter. *J. Am. Chem. Soc.* **135**, 10099–10103 (2013).
164. Xie, D. *et al.* Microbial synthesis of triacetic acid lactone. *Biotechnol. Bioeng.* **93**, 727–736 (2006).
165. Dunn, B. J., Cane, D. E. & Khosla, C. Mechanism and Specificity of an Acyltransferase Domain from a Modular Polyketide Synthase. *Biochemistry* **52**, 1839–1841 (2013).
166. Bergeret, F. *et al.* Biochemical and Structural Study of the Atypical Acyltransferase Domain from the Mycobacterial Polyketide Synthase Pks13. *J. Biol. Chem.* **287**, 33675–33690 (2012).
167. Sundermann, U. *et al.* Enzyme-directed mutasynthesis: a combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem. Biol.* **8**, 443–50 (2013).

168. Bravo-Rodriguez, K. *et al.* Predicted incorporation of non-native substrates by a polyketide synthase yields bioactive natural product derivatives. *Chembiochem* **15**, 1991–7 (2014).
169. Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *ELECTROPHORESIS* **30**, S162–S173 (2009).
170. Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. & Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* **7**, 1–15 (2017).
171. Studer, G. *et al.* QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics* **36**, 1765–1771 (2020).
172. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
173. Bienert, S. *et al.* The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res.* **45**, D313–D319 (2017).
174. Khersonsky, O. *et al.* Automated Design of Efficient and Functionally Diverse Enzyme Repertoires. *Mol Cell* **72**, 178-186.e5 (2018).
175. Muller, K., Faeh, C. & Diederich, F. Fluorine in Pharmaceuticals: Looking Beyond Intuition. *Science* **317**, 1881–1886 (2007).
176. Geyer, K., Sundaram, S., Sušnik, P., Koert, U. & Erb, T. J. Understanding Substrate Selectivity of Phoslactomycin Polyketide Synthase by Using Reconstituted in Vitro Systems. *ChemBioChem* **21**, 2080–2085 (2020).

177. Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **11**, 660–670 (2015).
178. Saadi, J. & Wennemers, H. Enantioselective aldol reactions with masked fluoroacetates. *Nat. Chem.* **8**, 276–280 (2016).
179. Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E. & Khosla, C. Comparative Analysis of the Substrate Specificity of *trans*- versus *cis*- Acyltransferases of Assembly Line Polyketide Synthases. *Biochemistry* **53**, 3796–3806 (2014).
180. Molnos, J., Gardiner, R., Dale, G. E. & Lange, R. A continuous coupled enzyme assay for bacterial malonyl–CoA:acyl carrier protein transacylase (FabD). *Anal. Biochem.* **319**, 171–176 (2003).
181. Kim, C.-Y. *et al.* Reconstituting Modular Activity from Separated Domains of 6-Deoxyerythronolide B Synthase †. *Biochemistry* **43**, 13892–13898 (2004).
182. Wu, N., Kudo, F., Cane, D. E. & Khosla, C. Analysis of the Molecular Recognition Features of Individual Modules Derived from the Erythromycin Polyketide Synthase. *J. Am. Chem. Soc.* **122**, 4847–4852 (2000).
183. Elia Heid. Engineering of Polyketide Synthases for the Production of Small Organic Molecules. (Johann Wolfgang Goethe-Universität Frankfurt am Main, 2020).
184. Hansen, D. A. *et al.* Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **135**, 11232–11238 (2013).
185. Koch, A. A. *et al.* Probing Selectivity and Creating Structural Diversity Through Hybrid Polyketide Synthases. *Angew Chem Int Ed* **132**, 13677–13682 (2020).



186. Awakawa, T. *et al.* Salinipyrone and Pacificanone Are Biosynthetic By-products of the Rosamicin Polyketide Synthase. *ChemBioChem* **16**, 1443–1447 (2015).
187. Donald, B. J., Surani, S., Deol, H. S., Mbadugha, U. J. & Udeani, G. Spotlight on solithromycin in the treatment of community-acquired bacterial pneumonia: design, development, and potential place in therapy. *Drug Des. Devel. Ther.* **11**, 3559–3566 (2017).
188. Zhanel, G. G. *et al.* Solithromycin: A Novel Fluoroketolide for the Treatment of Community-Acquired Bacterial Pneumonia. *Drugs* **76**, 1737–1757 (2016).
189. Poust, S. *et al.* Divergent Mechanistic Routes for the Formation of gemDimethyl Groups in the Biosynthesis of Complex Polyketides. *Angew Chem Int Ed* **54**, 2370–2373 (2015).
190. Altona, C. Vicinal Coupling Constants and Conformation of Biomolecules. in *Encyclopedia of Magnetic Resonance* (ed. Harris, R. K.) (John Wiley & Sons, Ltd, 2007). doi:10.1002/9780470034590.emrstm0587.
191. Haasnoot, C. A. G., de Leeuw, F. A. A. M., de Leeuw, H. P. M. & Altona, C. The relationship between proton-proton NMR coupling constants and substituent electronegativities. II—conformational analysis of the sugar ring in nucleosides and nucleotides in solution using a generalized Karplus equation. *Org. Magn. Reson.* **15**, 43–52 (1981).
192. Thibaudeau, C., Plavec, J. & Chattopadhyaya, J. A New Generalized Karplus-Type Equation Relating Vicinal Proton-Fluorine Coupling Constants to H–C–C–F Torsion Angles. *J. Org. Chem.* **63**, 4967–4984 (1998).

193. Park, S. R. *et al.* Discovery of cahuitamycins as biofilm inhibitors derived from a convergent biosynthetic pathway. *Nat. Commun.* **7**, 1–11 (2016).
194. Staunton, J. & Wilkinson, B. Biosynthesis of Erythromycin and Rapamycin. *Chem. Rev.* **97**, 2611–2630 (1997).
195. Jung, W. S. *et al.* Enhanced Heterologous Production of Desosaminyll Macrolides and Their Hydroxylated Derivatives by Overexpression of the pikD Regulatory Gene in *Streptomyces venezuelae*. *Appl. Environ. Microbiol.* **74**, 1972–1979 (2008).
196. Khersonsky, O. & Tawfik, D. S. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annu Rev Biochem* **79**, 471–505 (2010).
197. Schiestl, R. H. & Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **16**, 339–46 (1989).
198. Dutler, H., Kull, A. & Mislin, R. Fatty Acid Synthetase from Pig Liver. 2. Characterization of the Enzyme Complex with Oxidoreductase Activity for Alicyclic Ketones as a Fatty Acid Synthetase. *Eur. J. Biochem.* **22**, 213–217 (1971).
199. Hansen, D. A., Koch, A. A. & Sherman, D. H. Substrate Controlled Divergence in Polyketide Synthase Catalysis. *J. Am. Chem. Soc.* **137**, 3735–3738 (2015).
200. Sharma, K. K. & Boddy, C. N. The thioesterase domain from the pimaricin and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates. *Bioorg. Med. Chem. Lett.* **17**, 3034–3037 (2007).
201. Peter, D. M. *et al.* Screening and Engineering the Synthetic Potential of Carboxylating Reductases from Central Metabolism and Polyketide Biosynthesis. *Angew Chem Int Ed Engl* **54**, 13457–61 (2015).

202. Nagi, M. N., Cook, L., Suneja, S. K., Osei, P. & Cinti, D. L. Spectrophotometric assay for the condensing enzyme activity of the microsomal fatty acid chain elongation system. *Anal. Biochem.* **179**, 251–261 (1989).
203. Geoghegan, K. F. *et al.* Spontaneous  $\alpha$ -N-6-Phosphogluconoylation of a “His Tag” in *Escherichia coli*: The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **267**, 169–184 (1999).

## 7.6 Statement of personal Contributions

### 7.6.1 Scientific work

The projects of this thesis were performed in collaborative work. The individual contributions of cooperation partners, colleagues or students are described below in more detail for clarity. Also, students contributed to this thesis during internships in the lab under my supervision, and details are listed below, too.

#### 2.1 Characterization of the yeast fatty acid synthesis

In this project, I kindly received the FAS1-FAS2 deletion strain (*S. cerevisiae* strain, BY.PK1238\_1A\_KO) from Manuel Fisher (AK Grininger), which was produced in cooperation with Peter Kötter.<sup>130</sup> Furthermore, the design and cloning of the pRS shuttling vectors containing to FAS encoding gene fused to a Strep-tag were performed by Manuel Fischer. All EM and ET experiments were performed by Edoardo D’Imprima and Davide Floris (supported by Janet Vonck and Werner Kühlbrandt; all Max Planck Institute for Biophysics, Frankfurt), including preparation of the grids, data processing and structure solution. The orientation of FAS particles of the cryo-ET was investigated by Ricardo Sanchez (Max Planck Institute for Biophysics, Frankfurt). The phosphorylation sites were cloned, expressed, and analysed by Martin Schwalm during an internship in the lab, supervised by me. The scaffolding domains were characterized by Alexander Rill during his bachelor thesis supervised by Manuel Fischer and me.<sup>131</sup> The respective constructs were designed by Manuel Fischer and me.

## 2.2 Characterization of a non-releasing MSAS from *S. erythraea*

The experimental work was mainly performed during three bachelor theses by Christian Gusenda, Trinetri Goel and Christian Schreiber, all supervised by me.<sup>136,137,159</sup> The Strep- and his-tagged encoding plasmid for SerMSAS was designed and cloned previously in our lab. The ESI measurements of the ACPs and their analysis were performed by Khanh vu Huu (supported by Nina Morgner, all Goethe-University, Frankfurt). Codon optimization of ACPs were designed and cloned by Franziska Stegemann (Goethe-University, Frankfurt). EM experiments and analysis was performed by Edoardo D'Imprima (supported by Janet Vonck and Werner Kühlbrandt; all Max Planck Institute for Biophysics, Frankfurt). HPLC-MS measurements were kindly performed by the Bode-Group (Goethe-University, Frankfurt).

## 2.3 Engineering of PenPaMSAS towards 6-PrSA synthesis

The experimental work was mainly performed by Ilka Siebels under my supervision (AK Grininger). The design and cloning of the PenPaMSAS encoding plasmid was performed by Alexander Rittner (Goethe-University, Frankfurt). Trinetri Goel performed the cloning of the AT- and KS-knockouts and was contributing to protein expression, purification, analysis and in setting up the assay to investigate the AT during an internship in the lab supervised by me. The ESI measurements of the ACPs and analysis were performed by Khanh vu Huu (supported by Nina Morgner, Goethe-University, Frankfurt). Elia Heid performed activity assays of PenPaMSAS in different buffers, cloned, expressed and analysed AT- and KS-knockouts and investigated the enzymatic activity of the AT as pre-tests during his master thesis supervised by me.<sup>183</sup> HPLC-MS measurements were kindly performed by the Bode-Group (Goethe-University, Frankfurt). I planned and supervised the whole project supported by Martin Grininger.

## 2.4 Directed biosynthesis of fluorinated polyketides

The project was processed in close collaboration with Alexander Rittner (AK Grininger). The dissected KS-MAT-didomain and the stand-alone ACP of the mFAS, as well as the KS-AT didomain of DEBS M6 were designed and cloned by Alexander Rittner. The expression and purification of KS-MAT-didomain and the stand-alone ACP of the mFAS, as well as of the KS-AT didomain and the stand-alone ACP of DEBS M6 was performed by Lara Maria Mayer. The characterization of the MAT of mFAS and AT of DEBS M6 was also performed by her. Lara Maria Mayer was supervised by Alexander Rittner. The DEBS/FAS-hybrids were designed together with Alexander Rittner. The synthesis of fluorinated extender substrates was performed by Alexander Rittner, including the NMR analysis of the respective compounds. The biosynthesis of unreduced TKLs and their analysis was performed during the master thesis of Elia Heid, supervised by me.<sup>183</sup> Additionally, the analytical

macrolactone reactions were performed together with Elia Heid. The scaled-up reactions, the purification and the analysis of the macrolactones were performed together with Alexander Rittner. HPLC-MS measurements were kindly performed by the Bode-Group (Goethe-University, Frankfurt). Global fitting of the (M)AT-mediated transacylation data (equations 7) was developed by Alexander Rittner and Aaron Himmler. Blue native PAGE analysis was performed by Martin Schwalm, supervised by Alexander Rittner.

## 7.6.2 Writing process

Parts of the sections Results, Discussion and Experimental Procedures (and the according Supplementary information) have been published before. The chapters, that are based on those publications are described below, and a statement to the contribution of colleagues or cooperation partners is attached. It is indicated in figure legends, when figures of this thesis were published before and prepared by others (coauthors or cooperation partners). Whenever a figure or a table (or parts thereof) has been taken from a previous publication, it is stated in the figure legend as well as (optionally) a copyright permission or a co-author agreement is attached (see table S3). Whenever figures were adapted or used from publication, it is also stated explicitly in the figure legend. It is further stated in the figure legends when raw data for the preparation of figures were received from others.

**Table S3** | Obtained licence number (order number) for the used figures in the thesis

Figure number	License number/order number
3, 11	5073080914574
6	5073091406877
7-9	5073100319703

## 2.1 Characterization of the yeast fatty acid synthesis

The presented chapter 2.1.1-2.1.3 and the respective experimental section are based on two manuscript published in IUCrJ and eLife (adapted by me for this thesis). For clarity, the papers are shortly summarized in the following.

Chapter 2.1.5 and the respective experimental sections are based on a manuscript published in Scientific reports with a focus on the assembly of the yeast FAS (adapted by me for this thesis). For clarity, the paper is shortly summarized in the following.

#### 2.4 Directed biosynthesis of fluorinated polyketides

The presented chapter 2.4, the respective experimental section and the according supplementary information are based on a manuscript available as preprint in bioRxiv (adapted by me for this thesis). For clarity, the paper is shortly summarized in the following.

#### **The resolution revolution in cryoEM requires high-quality sample preparation: a rapid pipeline to a high-resolution map of yeast FAS**

Joppe, M.; D'Imprima, E.; Salustros, N.; Paithankar, K. S.; Vonck, J.; Grininger, M.; Kühlbrandt, W. The Resolution Revolution in CryoEM Requires High-Quality Sample Preparation: A Rapid Pipeline to a High-Resolution Map of Yeast Fatty Acid Synthase. *IUCrJ* **2020**, 7 (2), 220–227.

published

DOI: <https://doi.org/10.1107/S2052252519017366>.

URL: <http://journals.iucr.org/m/issues/2020/02/00/eh5006/index.html>

The yeast FAS was a major target for structural and functional characterizations during the thesis. This publication presents a rapid pipeline for purification, quality control, cryo-EM experiments, data collection and 3D reconstruction of the yeast FAS. A vector-based protocol for cloning, expression and purification was established, which can serve as a guideline for other macromolecular complexes. By non-invasive affinity chromatography and SEC, it was possible to purify the FAS within 5 hours. The quality of purified FAS solution can then be validated within 2.5 hours, by SDS-PAGE, SEC, TSA and enzymatic activity. The value of this protocol was demonstrated in solving the yeast FAS at 3.1 Å, treated with NADPH and Mal-CoA prior to cryo-EM experiments. The model includes additional density at Ser1440, indicating that this serine is phosphorylated. Additional density was also found at Cys820 and Cys824, which are probably malonylated due to the high concentration of Mal-CoA used for sample preparation. Another feature of the model is the binding of NADPH to the active site of the KR domain. Furthermore, this publication presents initial negative-staining EM experiments, indicating that centrifugation through a semipermeable membrane harms the protein, particularly the

PPT domain positioned at the outside of the barrel. Avoiding this step improved protein quality and densities at previously poorly resolved domains. The presented pipeline is beneficial for mutational studies, e.g. to inform protein engineering.

Author contribution: In this project, I established a protocol for rapid yeast FAS purification and validated the protocol by evaluating protein quality. I prepared every protein sample for cryo-EM, performed protein quality checks, analyzed data and reviewed and edited the final draft of the manuscript.

### **Protein denaturation at the air-water interface and how to prevent it**

D'Imprima, E.; Floris, D.; Joppe, M.; Sánchez, R.; Grininger, M.; Kühlbrandt, W. Protein Denaturation at the Air-Water Interface and How to Prevent It. *eLife* **2019**, *8*, e42747.

published

DOI: <https://doi.org/10.7554/eLife.42747>.

URL: <https://elifesciences.org/articles/42747>

This publication reveals protein adsorption to the air-water interface as a major challenge during sample preparation for cryo-EM, on the example of the yeast FAS. 2D and 3D classifications in cryo-EM data processing revealed a major damage in the protein structure of about 90% of particles. Many of the damaged particles showed missing densities of one third to one half of the barrel-shaped fold. In contrast characterization during protein purification revealed a pure, stable and fully active enzyme. Cryo-ET experiments confirmed the damage and additionally showed that most of the affected particles adsorbs to the air-water interface. The damaged side of the FAS particles was generally oriented towards the air layer. This publication highlights the use of graphene grids, rendered hydrophilic with 1-pyrene carboxylic acid, to avoid adsorption to the air-water interface. Cryo-ET experiments showed that nearly all particles prefer the graphene-water interface and stayed undamaged, which was also observed by single particle cryo-EM experiments. Finally, this publication presents a high resolution (4.0 Å) D3-symmetrized map of the yeast FAS from a small data set.

Author contribution: I used established protocols to purify FAS for cryo-EM and cryo-ET. I extended the quality control, described in chapter 2.1, by establishing blue

native PAGE. I evaluated the data, prepared figures, wrote parts of the paper (on protein purification) and reviewed and edited the final draft.

### **Analysis of the co-translational assembly of the fungal fatty acid synthase (FAS)**

Fischer, M.; Joppe, M.; Mulinacci, B.; Vollrath, R.; Konstantinidis, K.; Kötter, P.; Ciccarelli, L.; Vonck, J.; Oesterhelt, D.; Grninger, M. Analysis of the Co-Translational Assembly of the Fungal Fatty Acid Synthase (FAS). *Sci Rep* **2020**, *10* (1), 895.

published

DOI: <https://doi.org/10.1038/s41598-020-57418-8>.

URL: <https://www.nature.com/articles/s41598-020-57418-8>

This publication provides insights into the assembly of the two-gene encoded fungal FAS at the molecular level, confirming and extending a previous work on the co-translation assembly of multidomain proteins.<sup>63</sup> We characterized FAS mutants in its ability to complement a FAS deletion strain in FA-limited liquid cultures or in medium without fatty acids, analysed structural integrity by Western-blot and determined enzymatic activity *in vitro*. In doing so, we identified key residues (K2, E8) for FAS assembly and assigned the role of structural domains during the assembly process. This publication also shows that the two-gene encoded FAS can be fused to a single-gene encoded as well as a three-gene encoded FAS, indicating that the co-translational assembly may not be sequence- or site-specific. Furthermore, *in vitro* experiments of purified  $\alpha$ -chain and dissected  $\alpha$ -chain indicated that the essential post-translational modification of the ACP by the PPT happens within a dimeric sub-structure. This publication reveals the spatial and temporal coordination of three key-processes during the assembly, the co-translational interaction of  $\alpha$  and  $\beta$  to form the MPT, the post-translational modification of the ACP within in a dimeric sub-structure and the barrel enclosure to form the fully matured D3 symmetric  $\alpha_6\beta_6$  FAS. Based on our findings guidelines for remodelling the fungal FAS scaffold to enlarge its product spectrum for biotechnological applications were presented.

Author contribution: I designed constructs deleted in structural domains together with Manuel Fischer. I supervised Alexander Rill, who cloned and purified FAS constructs (wildtype and mutants), further characterized by native PAGE Western Blot, SEC, TSA and enzymatic activity. I analyzed the data, prepared figures and reviewed and edited the final draft.



**Directed biosynthesis of fluorinated polyketides**

Rittner, A., Joppe, M., Schmidt, J. J., Mayer, L. M., Heid, E., Sherman, D. H., & Grninger, M. (2021). Directed biosynthesis of fluorinated polyketides. *bioRxiv*.

Published as preprint in bioRxiv

doi: <https://doi.org/10.1101/2021.07.30.454469>

<https://www.biorxiv.org/content/10.1101/2021.07.30.454469v1.abstract>

This publication presents an approach to broaden the product spectrum of PKSs by swapping its native ATs with the highly promiscuous MAT of the mFAS. The approach is performed with the sixth module of the DEBS. Two different DEBS/FAS hybrids were presented, which were created by exchanging the MAT with or without its adjacent linker region. By comparing both hybrids, we showed that the introduction of a chimeric interface in between the MAT and the linker region led to better enzyme properties. In contrast to the WT system, the DEBS/FAS hybrid was capable to incorporate non-native extender substrates like Mal-CoA, F-Mal-CoA and F-MM-CoA, monitored by turnover rates and HPLC-MS. The highlight of this publication is the site-specific fluorination of 10-deoxymethynolide as fluorinated compounds are pharmacologically highly demanded. In conclusion, we could demonstrate that exchanging ATs of PKSs with the polyspecific MAT can enable the production of polyketide derivatives.

Author contribution: Alexander Rittner and I designed the DEBS/FAS hybrids. The hybrids were expressed, purified, and characterized by me. I designed and performed the cloning of module 6 ACP and characterized the protein. I synthesized the diketide SNAC to evaluate the enzymatic activity of the hybrids by NADPH consumption and HPLC-MS and established a CoA-consumption assay to analyze enzymatic activity in the absence of NADPH with HPLC-UV together with Elia Heid during his master thesis. Alexander Rittner and I performed the semi-synthesis, purification and analysis of the 10-deoxymethynolide and narbonolide derivatives. The analysis of the stereochemistry was performed by me with help from Julia Wirmer-Bartoschek (Goethe-University, Frankfurt). I evaluated data and prepared figures, wrote parts of the manuscript (and supporting information) and reviewed and edited the manuscript.

## 7.7 Acknowledgement

I would like to thank Alexander Rittner, Manuel Fischer, Franziska Stegemann, Maja Klaus from the Grininger Group (Goethe-University, Frankfurt) for providing plasmids and strains for cloning and expression of enzymes. I am grateful to Khanh Vu Huu and Kudratullah Karimi from the Morgner group (Goethe-University, Frankfurt) for performing MS-analysis of ACPs and Kyra Geyer from the Erb laboratory (Max Planck Institute for terrestrial Microbiology, Marburg) for providing the gene encoding for *Npt* (UniProt code: A0A1Y2MXW0). I would like to thank the Bode group (Goethe-University, Frankfurt) for performing HPLC-MS analysis, Gabriele Sentis (Goethe-University, Frankfurt) for support with NMR experiments and Julia Wirmer-Bartoschek (Goethe-University, Frankfurt) for support in NMR analysis. I am grateful to Edoardo D'Imprima, Davide Floris and Ricardo Sanchez (Max Planck Institute for Biophysics, Frankfurt) for EM and ET measurements and analysis. I thank Micheal Thumm (Georg August University Göttingen) for providing the anti-FAS serum and Jurema Schmidt and Veysel Erdel (Goethe-University, Frankfurt) for setting up the Native-PAGE for the yeast FAS. The plasmid encoding for DEBS M6, was kindly received from the Khosla laboratory (Stanford University, California)

## 7.8 Danksagung

Ein ganz besonderer Dank gilt meinem Doktorvater Martin Grininger für die Möglichkeit an vielen interessanten Projekten mitwirken und vor allem sie mitgestalten zu dürfen. Die vielen tiefgründigen Diskussionen über komplexe Themen waren immer eine Bereicherung für die die Bewältigung wissenschaftlicher Herausforderungen.

Weiterhin möchte ich mich bei Helge Bode für die Übernahme des Zweitgutachtens bedanken.

Außerdem möchte ich mich bei der gesamten Arbeitsgruppe Grininger (G-Force) bedanken. Dabei geht ein besonderer Dank an Ilka Siebels. Du hast es nicht nur immer geschafft alles bereit zu stellen, was ich für meine Arbeiten brauchte, sondern warst auch immer eine Unterstützung in wissenschaftlichen und sonstigen Fragen. Vor allem habe ich die regelmäßigen Kaffeepausen immer sehr genossen. Zusätzlich möchte ich mich herzlich bei Florian Bourdeaux und Alexander Rittner bedanken. Neben der genialen wissenschaftlichen Unterstützung konnten wir in den Abendstunden im Labor über alles quatschen. Alex mit dir hatte ich sehr großartige Nächte, entweder im Labor, während der Pandemie oder bei einem Herrengedeck im Sternchen. Auch möchte ich mich bei Karthik Paithankar für die technische Unterstützung sowie seinen Rat in strategischen Angelegenheiten bedanken.

Weiterhin möchte ich mich für die gewissenhafte Mitarbeit meiner Bachelor- und Masterstudierenden bedanken: Alexander Rill, Christian Gusenda, Christian Schreiber, Trinetri Goel, Elia Heid, Martin Schwalm und Simon Reiners.

Des Weiteren gilt mein dank Khanh Vu Huu und Kudratullah Karimi für die wissenschaftliche Mitarbeit und den Kaffeeklatsch während den Messungen. Unserer gemeinsame Vietnamreise (Khanh's Hochzeit im Dschungel) wird mir unvergessen bleiben.

Zuletzt möchte ich einen herzlichen Dank an meine Freunde und Familie aussprechen. Besonders hervorheben möchte ich meinen Gym-Partner Mirko Jamin, Spin-move Chris (Lenk), meinen Gitarrenlehrer Sören Moos und Verena Röhrig für das Korrekturlesen meiner Doktorarbeit. Weiterhin möchte ich mich bei meinen engen Freunden Dominik Göbel und David Drexler bedanken, besonders für die witzigen Dodecin-Sessions. Ein ganz persönlicher Dank gilt meiner Familie, insbesondere meiner Mutter, meinem Vater, meinem Bruder und meiner Freundin für die jahrelange Unterstützung.

## 7.9 Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den  (Mirko Joppe)

### Versicherung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation über „Structure and Function of Megasyntases“ 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  (Mirko Joppe)

## 7.10 Curriculum Vitae

**Mirko Joppe**

Master of Science  
in Chemistry

E-Mail: Mirko-Joppe@web.de  
Phone: +491628599304  
Address Mörfelder Landstraße 94  
60598 Frankfurt am Main  
Date of birth: July 8<sup>th</sup>, 1989



### Current position

---

01/2017-: **Ph.D. student**  
Goethe University, Frankfurt

**Thesis title:** Structure and function of megasynthases

**Concrete goal:** Engineering of fatty acid synthases and polyketid synthases for product derivatization

12/2021 Estimated graduation

### Education & work experience

---

06/2016-12/2016 Fraport Skyliners e.V., Basketball coaching in schools

07/2016-10/2016 Hill & Knowlton Strategies GmbH, healthcare communications

02/2014-03/2016 **M. Sc. Chemistry (1.2)**  
Goethe University Frankfurt  
Focus: Organic- and biochemistry  
Thesis title: Structure elucidation and function of an iterative polyketide synthase  
Tutor: Prof. Dr. Grininger

10/2010-02/2014 **B. Sc. Chemistry (1.5)**  
Goethe University Frankfurt  
Thesis title: Photoschaltbare Fluoreszenzfarbstoffe zur hochauflösenden Mikroskopie mittels Fluktuationsanalyse  
Tutor: Prof. Dr. Mike Heilemann

09/2009-05/2010 VzF Kindergarten, community service

2009 **Abitur (2.2)**  
Adolf-Reichwein-Gesamtschule

## Skills

---

### Scientific

Instruments:	<u>Operator:</u> HPLC, multiangle light scattering, size exclusion chromatography and French press
Cloning:	Polymerase chain reaction, Gibson assembly, digestion, and ligation
Strong background	Protein purification, biochemical and biophysical characterization
Organic chemistry:	Synthesis, flash chromatography, RP-HPLC, NMR-Spectroscopy

### Computer

Data analysis and presentation	MS Office, Adobe Illustrator CS5, Origin 2020, Chromeleon 6.8, Astra 7.1, professional scientific writing
--------------------------------	---

### Language

English	Business fluent, regular presentation and group meetings in English
German	Native speaker

## Research focus & publications

---

Characterization and engineering of enzymes was the focus during my Ph.D. Additionally, I worked on structural determination with electron microscopy in cooperation with the Max-Planck-Institute of biophysics. I am highly interested in protein engineering, research of small molecules and method development.

2021	Rittner, A. <sup>‡</sup> , <u>Joppe, M.<sup>‡</sup></u> , Schmidt, J. J., Mayer, L. M., Heid, E., Sherman, D. H., & Grninger, M. (2021). Directed biosynthesis of fluorinated polyketides. <i>bioRxiv</i> .
2020	<u>Joppe, M.<sup>#</sup></u> ; D'Imprima, E. <sup>#</sup> ; Salustros, N.; Paithankar, K. S.; Vonck, J.; Grninger, M.; Kühlbrandt, W. The Resolution Revolution in CryoEM Requires High-Quality Sample Preparation: A Rapid Pipeline to a High-Resolution Map of Yeast Fatty Acid Synthase. <i>IUCrJ</i> <b>2020</b> , <i>7</i> (2), 220–227. <a href="https://doi.org/10.1107/S2052252519017366">https://doi.org/10.1107/S2052252519017366</a> .
2020	Fischer <sup>#</sup> , M.; <u>Joppe, M.</u> ; Mulinacci, B.; Vollrath, R.; Konstantinidis, K.; Kötter, P.; Ciccarelli, L.; Vonck, J.; Oesterhelt, D.; Grninger, M. Analysis of the Co-Translational Assembly of the Fungal Fatty Acid Synthase (FAS). <i>Sci Rep</i> <b>2020</b> , <i>10</i> (1), 895. <a href="https://doi.org/10.1038/s41598-020-57418-8">https://doi.org/10.1038/s41598-020-57418-8</a> .
2019	D'Imprima <sup>#</sup> , E.; Floris <sup>#</sup> , D.; <u>Joppe, M.</u> ; Sánchez, R.; Grninger, M.; Kühlbrandt, W. Protein Denaturation at the Air-Water Interface and How to Prevent It. <i>eLife</i> <b>2019</b> , <i>8</i> , e42747. <a href="https://doi.org/10.7554/eLife.42747">https://doi.org/10.7554/eLife.42747</a>

## Conferences

---

3<sup>rd</sup> European Conference on Natural Products     Poster title: Assembly of a biofactory - Insights from fungal fatty acid synthase for engineering a megasynthase compartment

MegaSyn symposium on megasynthases     Talk title: Directed biosynthesis of fluorinated polyketides

## Personal interests & additional qualifications

---

2002-:                      Basketball player  
2017                      Course on time and project management during Ph.D.  
2005-2018              Basketball referee with licence  
2005-2016              Basketball coach with licence



GOETHE   
UNIVERSITÄT  
FRANKFURT AM MAIN

# URKUNDE

Herr  
**Mirko Joppe**

geboren am 08.07.1989 in Bad-Homburg v.d.Höhe  
hat am 27.02.2014 die Prüfung im  
Bachelor-Studiengang Chemie  
mit der Gesamtnote

**sehr gut (1,5)**

bestanden.

Aufgrund dieser Prüfung verleiht  
der Fachbereich Biochemie, Chemie und Pharmazie  
den akademischen Grad

**Bachelor of Science (B.Sc.)**

Frankfurt am Main, 27.02.2014

  
Der Dekan  
Prof. Dr. Thomas Pristner

  
  
Vorsitzender des Prüfungsausschusses  
Prof. Dr. Michael Göbel