Engineering Animal Fatty Acid Synthase Towards Polyketide Synthases

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie (FB 14)
der Johann Wolfgang Goethe -Universität
in Frankfurt am Main

von

Alexander Rittner

aus Bonn

Frankfurt 2018

(D30)

vom Fachbereich Biochemie, Chemie und Pharmazie (FB 14) der Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

1. Gutachter: Prof. Dr. Martin Grininger
2. Gutachter: Prof. Dr. Helge Bode

Datum der Disputation:

Zusammenfassung

Naturstoffe sind wertvolle Quellen für biologisch aktive Verbindungen, die als Arzneimittel verwendet werden können. Dabei basiert die Produktion rein auf Biokatalyse, die oft von sogenannten Megaenzymen durchgeführt wird. Ein wichtiger biosynthetischer Weg ist der Acetatweg, der die Polyketid- und Fettsäuresynthese einschließt, und eine der größten Klassen chemisch unterschiedlicher Naturstoffe hervorbringt. Diese haben aufgrund ihrer antibakteriellen, antimykotischen, anthelmintischen, immunsuppressiven und antitumoralen Eigenschaften medizinische Relevanz.

Der Aufbau von Fettsäuren folgt dabei dem gleichen Grundprinzip wie auch die Biosynthese von Polyketiden. Kleine Vorläufermoleküle, normalerweise Acetyl-CoA, werden selektiert und durch sukzessiven Anbau von Elongationssubstraten, wie z. B. Malonyl-CoA, auf eine ausgewählte Kettenlänge (typischerweise C16) verlängert. Die Auswahl der Substrate erfolgt durch Acyltransferasen (AT) die für die Beladung des Enzyms verantwortlich sind. Während der Synthese bleiben alle Intermediate kovalent an das Enzym gebunden, bevorzugt an das Acyl-Carrier-Protein (ACP), welches die Reste zu den entsprechenden aktiven Zentren transportiert. In einem ersten Schritt wird der Acetyl-Rest in der Ketoacylsynthase (KS) mittels einer decarboxylierenden Claisen-Kondensation mit einem Malonyl-ACP Ester verknüpft. Der dabei entstandene β-Ketoester wird anschließend durch drei weitere Reaktionsschritte modifiziert. Zuerst wird die β-Keto-Gruppe mit Nicotinamidadenindinukleotid-phosphat (NADPH) durch die Ketoreduktase (KR) zur Hydroxyl-Gruppe reduziert. Daraufhin wird durch eine Dehydratase (DH) die Eliminierung von Wasser katalysiert, bevor die entstehende Doppelbindung wiederum mit NADPH durch die Enoylreduktase (ER) zur gesättigten Kohlenstoffkette reduziert wird. Der hergestellte Acyl-ACP Ester wird nun in sich wiederholenden Zyklen, die oben eingeführten katalytischen Zentren verwendend, um jeweils eine C_2-Einheit verlängert bis die finale Kettenlänge erreicht ist. Typischerweise wird die Fettsäuresynthese durch Hydrolyse des Thioesters mittels einer Thioesterase (TE) beendet, kann jedoch auch, abhängig vom

entsprechenden Fettsäuresynthasentyp, durch Transacylierung, wie z.B. bei bakteriellen Fettsäuresynthasen, terminiert werden.

Die strukturelle Organisation der einzelnen aktiven Zentren bestimmt die Klassifizierung der Fettsäure- und Polyketidsynthasen (PKS) in zwei bzw. drei unterschiedliche Typen. Sind alle Enzyme individuell kodiert und interagieren als separate Proteine, wird vom Typ II gesprochen, wohingegen im Typ I alle katalytischen Domänen auf einer Polypeptidkette organisiert sind und somit sogenannte Multidomänenproteine oder Megasynthasen bilden. In PKS wird das gleiche Repertoire an katalytischen Funktionen verwendet wie bei Fettsäuresynthasen, wobei die große chemische Vielfalt der Polyketide im Wesentlichen durch Auswahl verschiedenen Substrate, die Variation der Anzahl der Kondensationsschritte und der An- oder Abwesenheit von modifizierenden Enzymen (KR, DH, ER) erreicht wird. Die einzelnen Domänen werden in Funktionseinheiten zusammengefasst, den sogenannten Modulen, die sich in erster Linie im modifizierenden Teil unterscheiden. Somit existieren vier verschiedene Module: das α- (KS+AT+ACP), β- (KS+AT+KR+ACP), γ- (KS+AT+KR+DH+ACP) und δ-Modul (KS+AT+KR+DH+ER+ACP), die im Gegensatz zur Fettsäuresynthese dazu führen, dass entsprechende, funktionelle Gruppen in Polyketiden erhalten bleiben. Ein Modul kann wie in Fettsäuresynthasen mehrfach (iterativ) oder nur einmalig (modular) verwendet werden, was in riesigen, Megadalton großen Assemblierungen von Modulen resultieren kann. Im Rahmen dieser Arbeit wurde ausschließlich mit Typ I Fettsäuresynthase- oder PKS-Systemen gearbeitet, wobei der Schwerpunkt mit Abstand auf der Tierfettsäuresynthase (FAS) aus Mäusen (mFAS) lag.

Aufgrund der hohen strukturellen und funktionellen Ähnlichkeit zwischen PKS und FAS kann die FAS als Paradigma für die gesamte Klasse von multifunktionellen Enzymen dienen. Um das biosynthetische Potential von FAS voll ausschöpfen zu können, ist ein guter Zugang zum Enzym von wesentlicher Bedeutung. In dieser Hinsicht bleibt *Escherichia coli* ein unangefochtener heterologer Wirt aufgrund niedriger Kulturkosten, schneller Mutagenesezyklen und relativ einfacher Handhabung. Überraschenderweise wurde noch keine effiziente

Expressionsstrategie für eine FAS in *E. coli* berichtet, da sich herausgestellt hat, dass die einzige, publizierte Strategie nicht reproduzierbar ist.

Wir begannen unsere Analyse mit der Suche nach einem geeigneten FAS-Homolog, das unsere Anforderungen an hohe Proteinqualität, ausreichende Ausbeute und garantierte Funktionalität erfüllt. Nach einem umfangreichen Screening verschiedener Varianten, Kultivierungsbedingungen und Koexpressionsstrategien identifizierten wir die FAS aus Mäusen als unser Protein der Wahl. Unsere etablierte Aufreinigungsstrategie, die auf der Verwendung von kurzen künstlichen Peptidsequenzen an beiden Termini basiert, gewährleistete einen reproduzierbaren und ausreichenden Zugang zum Protein in ausgezeichneter Qualität. Das Enzym wurde weiter biochemisch charakterisiert, einschließlich einer enzymkinetischen Analyse der Fettsäuresynthese und einer Untersuchung, ob verschiedene Acyl-CoA Substrate als Startermoleküle fungieren können. Damit wurde unser Repertoire von handhabbaren Megaenzymen um die mFAS erweitert, was den Weg zu einer Nutzung der katalytischen Effizienz in Bezug auf mikrobielle Synthese von maßgeschneiderten Verbindungen ebnet.

Mit dem Fokus, das Verständnis für die Funktionsweise solcher Megaenzyme zu vertiefen, anstatt entsprechende Biosyntheseprodukte zu analysieren, haben wir uns mit der Frage beschäftigt, ob mFAS selbst in eine PKS umgewandelt werden kann oder ob Eigenschaften von mFAS für die Entwicklung von PKS genutzt werden können. Dieser Ansatz wurde auf drei Komplexitätsebenen von der Funktion einzelner Domänen über die Organisation von Domänen zu Modulen bis hin zum Zusammenspiel zweier Module in bimodularen Konstrukten durchgeführt.

Die Fettsäuresynthese beginnt mit der Beladung der FAS mit Acyleinheiten, die von einer Domäne namens Malonyl/Acetyltransferase (MAT) durchgeführt wird. Diese Domäne wurde aufgrund ihrer ungewöhnlichen bifunktionalen Spezifität näher untersucht und detailliert charakterisiert. Unsere Analyse umfasste strukturelle und funktionelle Aspekte, wie z. B. das Lösen von Kristallstrukturen in zwei unterschiedlichen Acyl-gebundenen Zuständen und eine enzymkinetische Beschreibung der Hydrolyse- und Transacylierungsreaktionen für zwölf exemplarische CoA-Ester. Zu diesem Zweck wurde ein Assay etabliert, bei dem

unter Verwendung der α-Ketoglutaratdehydrogenase als ein gekoppeltes Enzym, das beim Transfer der Acyl-Reste auf die ACP-Domäne entstehende, freie Coenzym A in Nicotinamidadenindinukleotid (NADH) umgewandelt wird. Die Zunahme der Fluoreszenz von NADH wurde mithilfe eines Mikroplattenlesers kontinuierlich verfolgt und mit dem Verbrauch des Substrats korreliert. Aus individuellen initialen Reaktionsgeschwindigkeiten bei bestimmten Substratkonzentrationen wurden scheinbare und absolute kinetische Parameter mit den Methoden von Michaelis-Menten und Hanes-Woolf bestimmt. Diese Daten enthüllten eine ausgedehnte Substratambiguität der MAT-Domäne, die zuvor nicht in diesem Ausmaß beschrieben worden war. Weiterhin konnten wir die Domäne aus ihrem strukturellen Verbund mit der KS- und der Linker Domäne lösen und somit als eigenständiges Enzym exprimieren, was eine große Robustheit der MAT-Faltungseinheit gezeigt hat. Des Weiteren konnte die Substratspezifität innerhalb einer Mutagenesestudie mit 42 Varianten gezielt verändert werden mit dem Ziel jeweils spezifische AT für unterschiedliche Substrattypen zu generieren. Mit dieser Plastizität zusammen mit der Robustheit der Faltungseinheit sind die beide Kriterien für Proteinevolvabilität erfüllt, weswegen diese Domäne als ein vielseitiges Werkzeug für PKS-Engineering in potenziellen FAS/PKS Hybridsystemen eingesetzt werden kann.

Auf der höheren Komplexitätsebene untersuchten wir die architektonische Variabilität der mFAS-Faltungseinheit, was eine fundamentale Basis für eine breitere biosynthetische Anwendung darstellt. Wir konnten alle vier Modultypen rekonstruieren, die in typischen modularen PKS auftreten, was einen hohen Grad an Modularität innerhalb der Faltungseinheit bestätigte. Die einzelnen Module wurden hinsichtlich ihres oligomeren Zustandes mithilfe von Größenausschlusschromatographie und ihrer Stabilität durch Schmelzkurven charakterisiert. Nicht nur die strukturelle, sondern auch die funktionelle Integrität dieser Module wurde unter Verwendung der Triacetsäurelactonbildung und der Ketoreduktaseaktivität validiert. Insbesondere letztere Analyse ermöglichte es, Effekte der Veränderung des prozessierenden Teils durch entsprechende enzymkinetische Parameter zu quantifizieren. Interessanterweise wurde die

partielle Aktivität der KR-Domäne durch Abstraktion der TE-Domäne stärker beeinflusst als durch Deletion der benachbarten, interagierenden ER-Domäne.

Im letzten Schritt haben wir unseren Fokus über ein einzelnes Modul hinaus erweitert und die mFAS-Faltungseinheit zur Konstruktion von bis zu 380 kDa großen bimodularen Konstrukten verwendet. Bei diesem Ansatz wurde eine N-terminale Ladedidomäne angehängt, die eine zusätzliche MAT- und ACP-Domäne enthielt. Zwei Konstrukte konnten in exzellenter Qualität exprimiert und gereinigt werden, mit deren Hilfe der Einfluss einer veränderten Gesamtarchitektur auf die Fettsäuresynthese untersucht wurde. Durch einen Vergleich mit entsprechenden Kontrollen konnte in der Tat ein funktioneller Effekt des zusätzlichen Lademoduls nachgewiesen werden. Diese Konstrukte dienen daher als potentielles Modellsystem um in Zukunft eine umfassende Analyse des zugrunde liegenden molekularen Mechanismus zum Übergang von einer iterativen zu einer vektoriellen Synthese *in vitro* durchzuführen, welcher bisher nicht gut verstanden ist.

In dieser Arbeit wurde die Expression der Tierfettsäuresynthase aus Mäusen in Bakterien etabliert und gezeigt, dass dieses Multidomänenprotein exzellente Eigenschaften besitzt um *in vitro* untersucht werden zu können. Wir konnten im Rahmen dieser These unser Verständnis über die Funktionsweise einzelner Domänen bis hin zur Modularität der mFAS-Faltungseinheit vertiefen und wichtige Erfahrungswerte zum Verhalten dieser Familie der Megasynthasen gewinnen. Diese Erkenntnisse sind direkt auf die medizinisch relevanten PKS übertragbar und fungieren als Grundlage zur Kreation von FAS/PKS-Hybridsystemen. Auf diese Weise hoffen wir neue Katalysatoren für mikrobielle Biosynthese zu schaffen, die effizient komplexe Polyketide derivatisieren oder produzieren können, die durch konventionelle Totalsynthese nicht zugänglich sind.

In Dankbarkeit für die jahrelange Unterstützung

meinen Eltern und meiner Freundin Lan-Na

gewidmet

Table of Contents

1.1 Abbreviations

Å – Ångström

A domains – adenylation domains

aaRS – aminoacyl-tRNA-synthetase

ACP – acyl carrier protein

αKGDH – α-ketoglutarate dehydrogenase complex

AM3L, AM11L – KS-ACP linker of avermectin synthase module 3 or 11

amp – ampicillin

AT – acyltransferase

ATP – adenosine triphosphate

AVES – avermectin synthase

BB – backbone

bla – b-lactamase

bp – base pair

BSA – bovine serum albumin

CoA – coenzyme A

CPM – 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin

Cpn – chaperonin

cryoEM – cryo electron microscopy

CV – column volume

DCM – dichloromethane

DD – docking domain

DE – dimerization element

DEBS – 6-deoxyerythronolide B synthase

DH – dehydratase

DNA – deoxyribonucleic acid

DTT – dithiothreitol

E. coli – Escherichia coli

EDTA – Ethylenediaminetetraacetic acid

EPR – electron paramagnetic resonance

ER – enoylreductase

ESI – electron spray ionization

FA – fatty acid

FabD – malonyl CoA-acyl carrier protein transacylase from E. coli

FAS – fatty acid synthase

FASN – gene for animal fatty acid synthase

FRET – fluorescence resonance energy transfer

hFAS – human fatty acid synthase

HPLC – high performance liquid chromatography

HR-PKS – highly-reducing polyketide synthase

IMAC – immobilized metal affinity chromatography

IPTG – isopropyl-β-D-thiogalactopyranoside

KAS I – 3-oxoacyl-[acyl-carrier-protein] synthase 1 from E. coli

KR – ketoreductase

KS – Ketosynthase

lacI – Lac repressor protein

LB – lysogeny broth

LD – linker domain

LDD – loading didomain

M. bakeri – *Methanosarcina bakeri*

M. jannaschii – *Methanocaldococcus jannaschii*

M. mazei – *Methanosarcina mazei*

MAT – malonyl-/acetyltransferase

MBP – maltose binding protein

mFAS – murine fatty acid synthase

MPI – Max Planck Institute

MS – mass spectrometry

MSA – methylsalicyclic acid

MSAS – methylsalicylic acid synthase

NADH – nicotinamide adenine dinucleotide

NADPH – nicotinamide adenine dinucleotide phosphate

NMR – nuclear magnetic resonance

NR-PKS – non-reducing polyketide synthase

NRPS – nonribosomal peptide synthetase

NTA – nitrilotriacetic acid

OD_{600} – optical density at a wavelength of 600 nm

P. furiosus – *Pyrococcus furiosus*

pAT – post-AT linker

PCR – polymerase chain reaction

PDB – protein data bank

PEG – polyethylene glycol

ΨKR – pseudo-ketoreductase

PKS – polyketide synthase

ΨME – pseudo-methyltransferase

PR-PKS – partially-reducing polyketide synthase

RAPS – rapamycin synthase

RBS – ribosome binding site

rcf – relative centrifugal force

RFU – relative fluorescence unit

RMSD – root mean square deviation

S. avermitilis – Streptomyces avermitilis

S. erythraea – Saccharopolyspora erythraea

SDS-PAGE

SEC – size exclusion chromatography

Sf9 – Spodoptera frugiperda

Sfp – 4'-phosphopantetheinyl transferase

ST – sulfotransferase

TAG – amber stop codon

TAL – triacetic acetic lactone

TB – terrific broth

TE – thioesterase

TF – Trigger factor

TPP – thiamine pyrophosphate

TRX – thioredoxin

1.1 Abstract

Natural products are valuable sources for biologically active compounds, which can be utilized as pharmaceuticals. Thereby, the synthesis is based purely on biosynthetic grounds often conducted by so-called megaenzymes. One major biosynthetic pathway is the acetate pathway including polyketide and fatty acid synthesis, which encompass one of the largest classes of chemically diverse natural products. These have medicinal relevance due to their antibacterial, antifungal, anthelmintic, immunosuppressive and antitumor properties.

Due to the high structural and functional similarity between polyketide synthases and type I animal fatty acid synthases (FASs), FAS can serve as a paradigm for the whole class of multifunctional enzymes. To fully exploit the biosynthetic potential of FASs, a good access to the enzyme is of essential importance. In this regard, *Escherichia coli* remains an unchallenged heterologous host due to low culturing costs, particularly fast mutagenesis cycles and relatively easy handling. Surprisingly, no sufficient expression strategy for an animal FAS in *E. coli* has yet been reported, as it turned out that the only approach was not reproducible.

We commenced our analysis with searching for an appropriate FAS homolog that fulfills our requirements of high protein quality, sufficient yield and ensured functionality. After extensive screening of different variants, culturing conditions and co-expression strategies, we identified the murine FAS (mFAS) as our protein of choice. The established purification strategy using tags at both termini led to a reproducible and sufficient access to the protein in excellent quality. The enzyme was further biochemically characterized including an enzyme kinetic investigation of fatty acid synthesis and an examination whether different acyl-CoA substrates can serve as priming units. This adds mFAS to our repertoire of manageable megaenzymes paving the way to exploit the catalytic efficiency in regards of microbial custom-compound synthesis.

With a strong focus on deepening our understanding of the working mode of such megaenzymes, rather than analyzing respective biosynthetic products, we have addressed the question whether mFAS itself can be engineered towards PKSs or whether properties of mFAS can be exploited to engineer PKSs. This approach was conducted on three levels of complexity from function of individual domains via organization of domains to form modules to the interplay of two modules in bimodular constructs.

Fatty acid synthesis begins with the loading of acyl moieties onto the FAS, which is conducted by a domain called malonyl-/acetyltransferase (MAT). This domain was in-

depth characterized due to its important role of choosing the substrates that are built in the final compound. Our analysis comprised structural and functional aspects providing crystal structures of two different acyl-bound states and kinetic parameters for the hydrolysis and transacylation reaction using twelve exemplary CoA-esters. For this purpose, we have successfully established a continuous fluorometric assay using the α-ketoglutarate dehydrogenase as a coupled enzyme, which converts the liberated coenzyme A into Nicotinamide adenine dinucleotide. These data revealed an extensive substrate ambiguity of the MAT domain, which had not been reported to that extent before. Further, we could demonstrate that the fold fulfills both criteria for the evolvability of an enzyme by expressing MAT in different structural arrangements (robustness) and by altering the substrate ambiguity within a mutagenesis study (plasticity). Taken these aspects together, we are persuaded that the MAT domain can serve as a versatile tool for PKSs engineering in potential FAS/PKS hybrid systems.

On the higher level of complexity, we investigated the architectural variability of the mFAS fold, which constitutes a fundamental basis for a broader biosynthetic application. We could rebuild all four module types occurring in typical modular PKSs confirming a high degree of modularity within the fold. Not only structural, but also functional integrity of these modules was validated by using triacetic acid lactone formation and ketoreductase activity. Especially the latter analysis, made it possible to quantify effects of the engineering within the processing part by respective enzyme kinetic parameters. Expanding our focus beyond a singular module, we have utilized the mFAS fold for designing up to 380 kDa large bimodular constructs. In this approach, a loading didomain was attached N-terminally containing an additional MAT and acyl carrier protein (ACP) domain. Two constructs could be expressed and purified in excellent quality to investigate the influence of an altered overall architecture on fatty acid synthesis. By comparison with appropriate controls, a functional effect of the additional loading module could indeed be proven in the bimodular systems. Those constructs allow a comprehensive analysis of the underlying molecular mechanism in the future and serve as a potential model system to study the transition from iterative to vectorial polyketide synthesis *in vitro*.

Chapter I: Introduction

1.1 Background

Natural products are valuable sources for biologically active compounds, which can be utilized as pharmaceuticals. In fact, although only about 1 % of all known molecules are natural products, almost two thirds of all drugs currently in use are at least in part derived from natural sources.[1] This bias is commonly explained with the argument that natural products have co-evolved with living organisms and were thus selected over a long period of time. But, on the contrary, with the excessive use of certain compounds in nature, living organisms adapt to their altered environment, which can cause resistances. This aspect gains more and more in importance as the (over)use of for example antibiotics in hospitals and livestock breeding created multi-resistant human pathogens, which demands the discovery of new effective remedies.[1]

Life includes a continuous transformation and interconversion of a vast amount of organic compounds. To adjust the speed of chemical reactions to the needs of the host, enzymes have evolved that are macromolecular biological catalysts. Over a period of time, cells have fine-tuned the interplay of thousands of enzymes to generate biosynthetic pathways. Despite a tremendous versatility among different species, the primary metabolism creating compounds like carbohydrates, proteins, lipids and nucleic acids was found to be evolutionary conserved. Out of the essential metabolic pathways, certain species, especially soil bacteria, fungi and plants, have evolved secondary metabolisms yielding secondary metabolites, which serve as a pool for pharmaceuticals. The most important biosynthetic pathways for such compounds are the non-ribosomal peptide synthesis, the shikimic acid pathway, the mevalonate pathway and the acetate pathway. Especially, the latter three are synthetically very interesting as during the process new carbon-carbon bonds are formed.

There is hardly any limit for the structural complexity, which can be generated with these biosynthetic pathways. A good example for this aspect is Maitotoxin (**Figure S1**), which is a 3422 Da polyketide-derived polycyclic ether with 164 carbon atoms, 68 oxygen atoms and countless chiral centres. The compound, produced by *Gambierdiscus toxicus*, a marine dinoflagellate, is among the largest known nonpolymeric natural products and displays an extremely high toxicity (LD50 is 50 ng/kg in mice).[2] Disregarding structural complexity, organic chemists challenge nature in the field of total synthesis trying to

synthesize these compounds from simple and cheap building blocks. Although in the case of Maitotoxin even K. C. Nicolaou, one of the most renowned organic chemists of the total synthesis field, surrendered due to funding issues,[3] many polyketides have been successfully synthesized over the last decades. In the case of erythromycin, a macrolide antibiotic, Robert B. Woodward, who was awarded the Nobel Prize in Chemistry in 1965, reported the complete total synthesis already in 1981.[4-6] With all respect to the ingenious synthetic routes, most total syntheses suffer from negligible yields and high costs, which often limits the application to the academic field.

A more economical approach to obtain derivatized, biologically active compounds is provided by the methods of semi-synthesis and rational biosynthetic pathway engineering. In semi-synthesis, the final natural products, isolated from their natural or a heterologous source, are utilized as precursors for regioselective organic-chemical synthesis. This approach has created many new semi-synthetic compounds, like e.g. tigecycline, paclitaxel and artemether,[7-9] which are already approved as pharmaceuticals. However, semi-synthesis has its limitations due to the necessity of regioselective and stereoselective reactions, which are barely available for complex molecules with many functional groups and stereogenic centers. Hence, a rather new, very promising approach for the modification of existing or the creation of new lead compounds inserts the desired alterations already during the process of production by engineering the responsible enzymes. Unfortunately, those are often very sophisticated themselves, and sufficient rational engineering fails due to a lack of knowledge of protein properties and enzyme kinetics.

The rational engineering of megaenzymatic assembly lines, like modular PKSs, is virtually as challenging for an enzymologist as the total synthesis of natural products is for an organic chemist. Once established though, altered biosynthetic pathways can provide cheap sources for pharmaceuticals in the long term, which can aid in antagonizing evolving multiresistent pathogens.

1.2 Principle of Polyketide Synthesis

Polyketides constitute a large class of natural products grouped together on purely biosynthetic grounds.[10] Their chemical structure is very diverse including saturated fatty acids, polyacetylene, prostaglandine, macrolide antibiotics and many aromatic compounds, e.g. anthraquinones and tetracyclines. Many of these natural products have

a high medicinal relevance as pharmaceuticals due to their biological activity with antibacterial, antifungal, anthelmintic, immunosuppressive and antitumor properties.

A

B

class	R_1	R_2
A	OH	CH_3
B	H	CH_3
C	OH	H
D	H	H

class	R_1	R_2	X–Y	class	R_1	R_2	X–Y
A_{1a}	CH_2CH_3	CH_3	CH=CH	B_{1a}	CH_2CH_3	H	CH=CH
A_{1b}	CH_3	CH_3	CH=CH	B_{1b}	CH_3	H	CH=CH
A_{2a}	CH_2CH_3	CH_3	CH–CH(OH)	B_{2a}	CH_2CH_3	H	CH–CH(OH)
A_{2b}	CH_3	CH_3	CH–CH(OH)	B_{2b}	CH_3	H	CH–CH(OH)
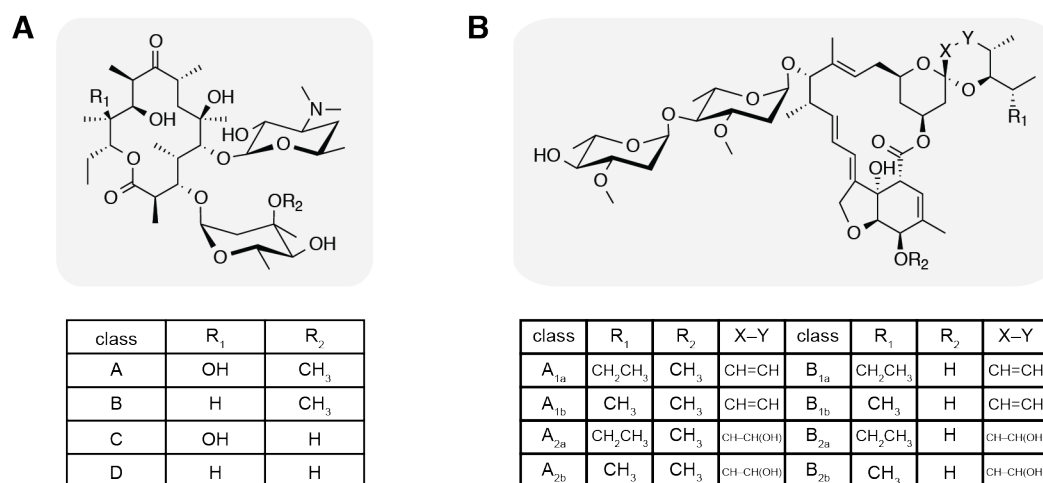
Figure 1: Chemical structures of select erythromycins (A) and avermectins (B). Rs indicate possible functional modifications leading to different members of the respective families. Structures were drawn with the software ChemDraw.

Although polyketide compounds are miscellaneous (for examples, see **Figure 1**), the enzymatic systems responsible for the *de novo* biosynthesis share many similarities including the utilization of common precursors, similar chemistry, similar structures and overall architectural design. These enzymes are called polyketide synthases (PKSs) and are classified due to their different structural frameworks. Multimodular PKSs (type I) consist of one or more large multidomain polypeptides where the growing polyketide chain is sequentially passed from one active site to the next.[11] When all catalytic domains are utilized once in a stepwise fashion they are normally referred to as "modular" type I PKSs, whereas megasynthases are called "iterative" type I PKSs, when active sites are used repetitively. In the type II system all catalytic components are freestanding monofunctional polypeptides, which nevertheless can also assemble as non-covalent complexes.[12] In both cases, the nascent polyketide intermediate remains permanently covalently attached to the protein until the final product is released. A third type of PKSs (type III) is fundamentally different in this manner, as the polyketide is never directly attached to the protein. This class contains very simple β-ketoacylsynthase proteins, e.g. chalcone and stilbene synthases, which do not require any further functionalities.[13]

1. Selection of substrates:

priming substrate
acyl
elongating substrate
malonyl

2. Claisen-condensation:
β-keto-
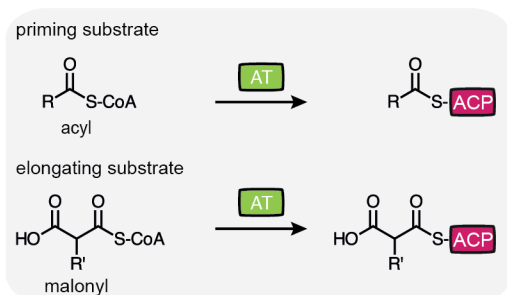
3. β-Carbon processing:
a) β-hydroxy-
b) enoyl-
c) acyl
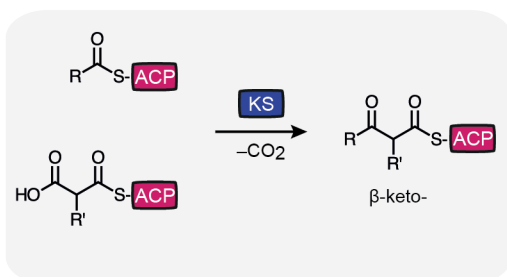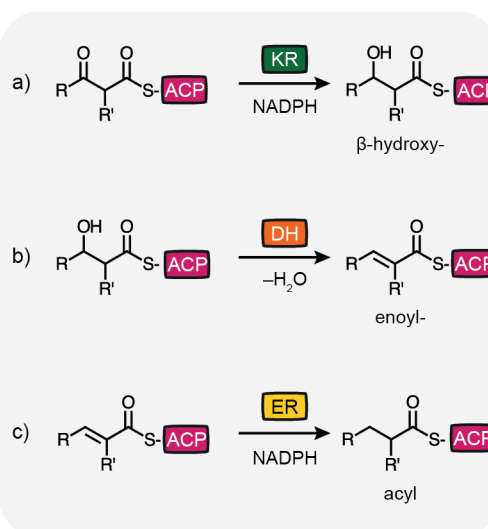
4. Termination:
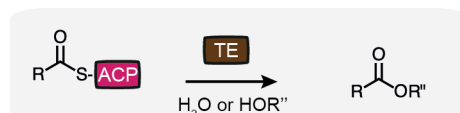
typicallly: R = CH₃, C₂H₅
R' = H, CH₃
R'' = H, macrolide

Figure 2: Catalytic activities in PKSs. Introduction of commonly used functionalities in PKSs. The abbreviations for domains refer to: AT – acyltransferase; KS – ketosynthase; KR – ketoreductase; DH - dehydratase; ER – enoylreductase; TE – thioesterase; ACP – acyl carrier protein. The color code of domains is consistently used within this thesis.

All PKSs, regardless of the type and regardless of being modular or iterative, utilize a common set of functionalities (Figure 2), which are either present as freestanding enzymes or as domains of a multidomain protein. Polyketide synthesis begins with the selection of precursors, which are delivered as active CoA-esters. These precursors can be classified into priming/loading and elongating/extending substrates, which are typically acetyl- or propionyl-CoA and dicarboxylic acid derived malonyl- or methylmalonyl-CoA, respectively. This job is performed by acyltransferases (ATs), which catalyze the transfer of the acyl-moiety to an acyl carrier protein (ACP), which then shuttles the polyketide intermediates to all active sites. After the ACP-bound acyl chain has been transferred to the active site of a ketosynthase (KS), a KS-mediated decarboxylative Claisen-condensation takes place with an ACP-bound extending moiety, e.g. malonyl-ACP. This leads to the formation of an ACP-bound β-ketoester, which has been formally elongated by two carbon atoms. Subsequently, the β-keto position can be

modified by processing enzymes, which act sequentially. The first step is a NADPH-mediated reduction of the keto-group to a secondary alcohol by a ketoreductase (KR). The hydroxyl-group can be eliminated by a dehydratase (DH) leading to the loss of water and the creation of a double bond, which can be further reduced with NADPH to a saturated acyl-chain by the enoylreductase (ER). The last step of polyketide synthesis is the release of the final product by thioesterases (TEs), which catalyze either the hydrolysis or an esterification of the thioester. It shall be again highlighted that during the whole biosynthesis until release of the final product, the growing polyketide remains covalently bound to the protein, either to the ACP, KS or TE domain.
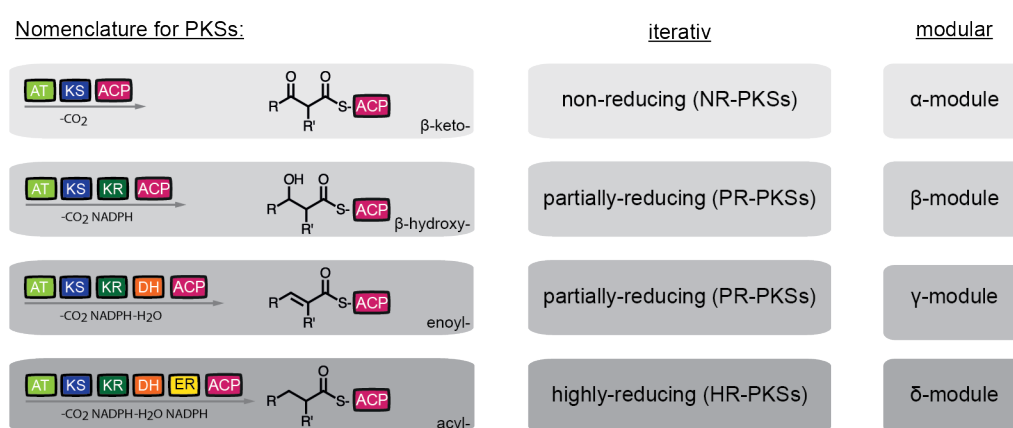
Figure 3: Nomenclature for PKSs. R and R' are defined as in **Figure 2**. For reviews, see references.[13, 14]

The great diversity of polyketide compounds is basically accomplished by selection of different substrates by ATs, a variation in the number of condensation steps and the presence or absence of modifying enzymes (KR, DH, ER). In modular PKSs, the minimal functional unit, which is called module, consists of the three domains KS, AT and ACP. This α-module selects the substrates and condenses an acyl-chain with an extender unit (**Figure 3**) leading to a keto-group at the β-position. Depending on the presence of modifying domains, modules are further classified as β-module (KS+AT+KR+ACP), γ-module (KS+AT+KR+DH+ACP) and δ-module (KS+AT+KR+DH+ER+ACP), which determine the remaining functional group as a hydroxyl-group, a double bond and a saturated acyl chain, respectively. As soon as modification of the β-position (also α and γ positions can be altered, e.g. the ER domain determines enantiomers of respective branches at the α-position) is accomplished, the

acyl-chain is passed on to the next, downstream KS domain, where further elongation cycles occur.

Similarly, iterative PKSs are divided into three classes, namely non-reducing (NR-PKS), partially-reducing (PR-PKS) and highly-reducing (HR-PKS), each possessing a set of enzymes commonly employed in that class.

1.2.1 DEBS as the Textbook Example of a Modular PKS

For the first time erythromycin A (**Figure 1A**) was isolated from *Saccharopolyspora erythraea* (*S. erythraea*) in 1952 and already in the same year a report on clinical trials was published assigning antibacterial properties to the natural product.[15, 16] For an excellent review on the pharmacology and clinical use of erythromycin, see reference.[17, 18] The compound contains a 14-membered macrocycle composed entirely of propionate units, both as priming and elongating units, the former via propionyl-CoA and the latter via methylmalonyl-CoA. The core ring structure is called 6-deoxyerythronolide B (**Figure 4**) and is entirely produced by one modular type I PKS, named 6-deoxyerythronolide B synthase or in short DEBS. In common with many antibacterial macrolides, biological activity is only gained after further modifications, which include hydroxylation, methylation and the attachment of sugar moieties.
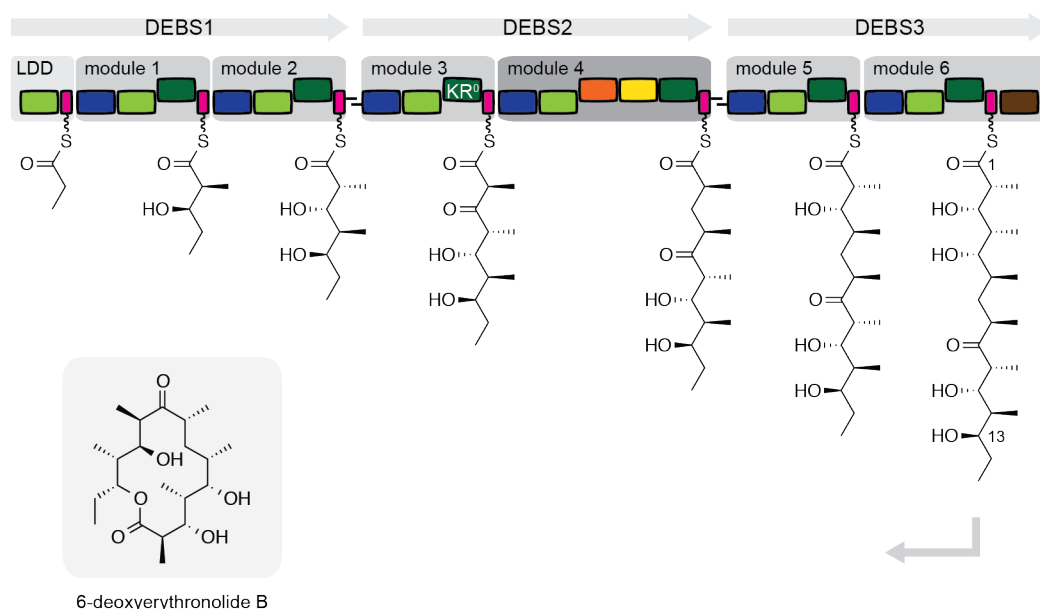
Figure 4: Architecture of the 6-deoxyerythronolide B synthase (DEBS). Each box represents an enzymatic domain in the PKS multifunctional polypeptide (abbreviations as introduced above). Modules are highlighted by grey background. Adapted from reference.[19]

The 6-deoxyerythronolide B synthase contains three polypeptide chains (DEBS 1 (365 kDa), DEBS 2 (374 kDa) and DEBS 3 (331 kDa)) encoded by three genes: *ery*A1, *ery*A2 and *ery*A3.[20, 21] It has a linear sequence of six modules (five β-modules and one δ-module, see section 1.2), which are responsible for six elongation steps and a loading didomain (LDD) associated with the first module. The product is released from the PKS by a thioesterase attached to the C-terminus of module 6, which catalyzes the lactonization via the alcohol function at C-13. Module 3 generates a β-ketoester due to an inactivation of the integral KR domain, defining the module only structurally as a β-module.

Due to a vast amount of studies conducted on the protein DEBS, mainly in the laboratories of Peter Leadlay and Chaitan Khosla, the enzyme has become the textbook example for modular type I PKSs.

1.2.2   AVES as a Second Example of Modular PKSs

The avermectins (**Figure 1B**) have no antibacterial activity but possess anthelmintic, insecticidal and acaricidal properties, and these are exploited in human and veterinary medicine.[10] The discovery and characterization of this series of drugs was mainly achieved by William C. Campbell and Satoshi Ōmura, who were recently awarded with the Nobel Price in Physiology or Medicine in 2015.

The avermectins are 16-membered macrolides containing oxygen heterocycles fused to the macrolide. Additional cyclization reactions lead to the very characteristic spiroketal system, which arise from a combination of a ketone and two alcohol functions. Further oxidation and glycosylation reactions accomplish the active product (**Figure 1B**).

Biosynthesis occurs in *Streptomyces avermitilis* (*S. avermitilis*), where all necessary enzymes are encoded in one gene cluster. It can be classified into three stages as followed: formation of the polyketide-derived initial aglycon, modification of the initial aglycon to generate avermectin aglycons and final glycolysation to generate avermectins.[22] In the following, we focus upon the formation of the initial polyketide compound, as it is conducted by a single modular PKS, named avermectin synthase (AVES; **Figure 5**).
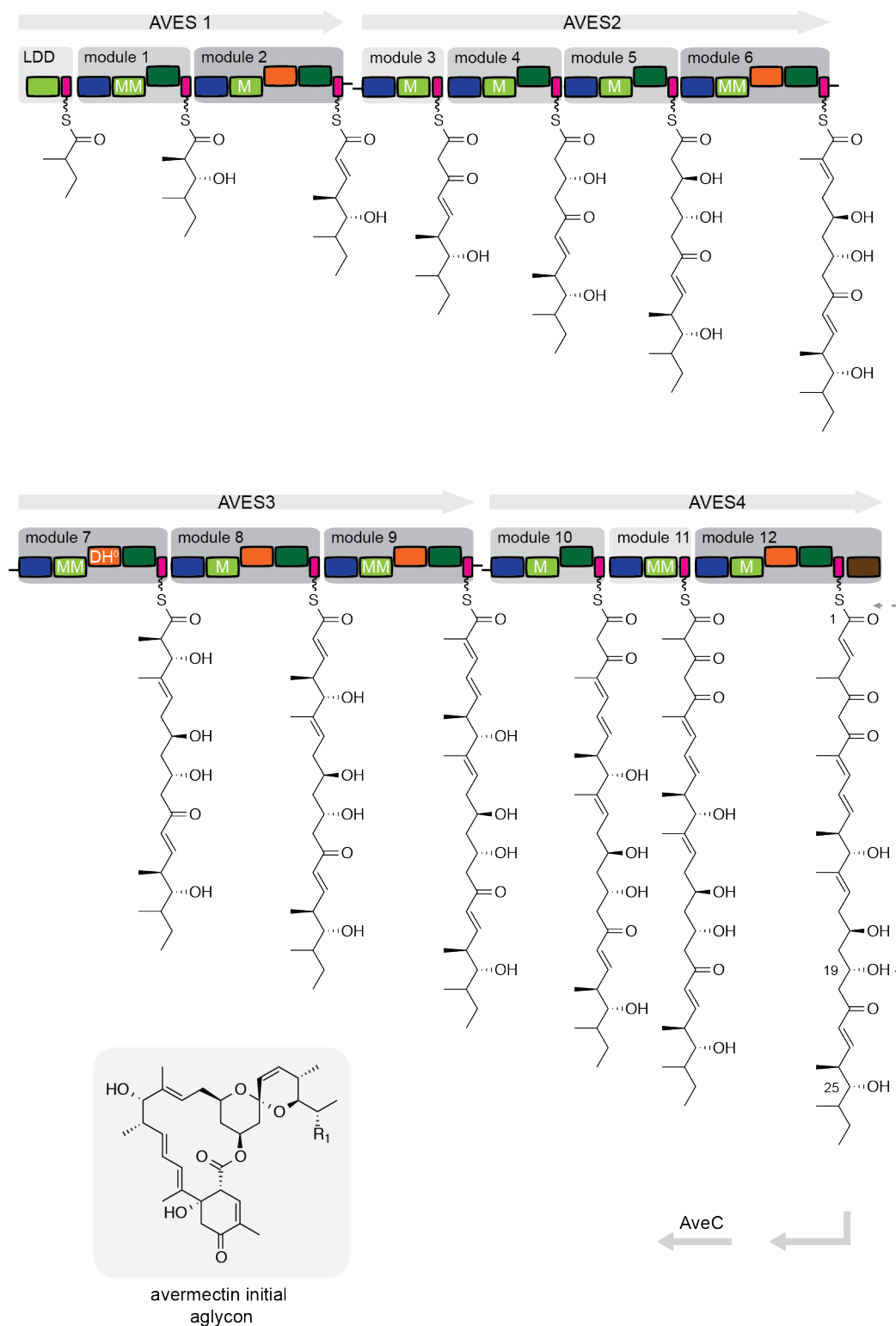
Figure 5: Domain structure of the avermectin PKS. Each box represents an enzymatic domain in the PKS multifunctional polypeptide (abbreviations as introduced above). Adapted from reference.[22]

AVES is composed of four polypeptide subunits, further organized into twelve modules. AVES 1 (414 kDa) contains the loading didomain plus two partially reducing modules 1 and 2. AVES 2 is with 666 kDa the largest polypeptide chain containing modules 3-6 and AVES 3 (575 kDa) is built up by modules 7-9, respectively. Synthesis is accomplished by AVES 4 (510 kDa) containing further three modules (10-12) for polyketide chain extension and the C-terminal thioesterase domain, necessary for cyclization and hence release of the completed polyketide (initial aglycon). During synthesis, one starting acyl unit, seven molecules malonyl-CoA and five molecules methylmalonyl-CoA are consumed, which has been revealed by labelling studies already in the 1980s.[23] Typically, two different starting acyl units are utilized, either isobutyryl-CoA (component a) or 2-methylbutyryl-CoA (component b) yielding a mixture of compounds. The accountable loading-AT domain has been characterized structurally and kinetically, revealing further substrate promiscuity.[24] The B-series of avermectins, the 5-hydroxy analogues of the A-series (**Figure 1**B), is used on agricultural crops to control mites and insects being the most active antiparasitic compound in this manner.

1.3 Fatty Acid Synthesis and Animal Fatty Acid Synthase

Fatty acids are key compounds of cell membranes, metabolism and signaling, and are obtained either directly from the diet, or synthesized *de novo* in a repeating cyclic reaction catalyzed by fatty acid synthases (FASs). The chemistry of fatty acid (FA) synthesis is largely conserved from bacteria to eukaryotes. Acetyl- and malonyl-CoA compounds are first loaded by the malonyl-/acetyltransferase (MAT) and then condensed by a KS domain in a decarboxylative Claisen condensation. Subsequently, the β-ketoester is modified by three active sites in the KR, DH and ER domains. A first cycle delivers a fully reduced ACP-bound butyryl group, which is elongated with a malonyl moiety in further cycles to a select chain length (typically C16). During this multi-step catalytic process, the acyl carrier protein (ACP) domain shuttles the fatty acid cargo between the different enzymatic domains (**Figure 6**).
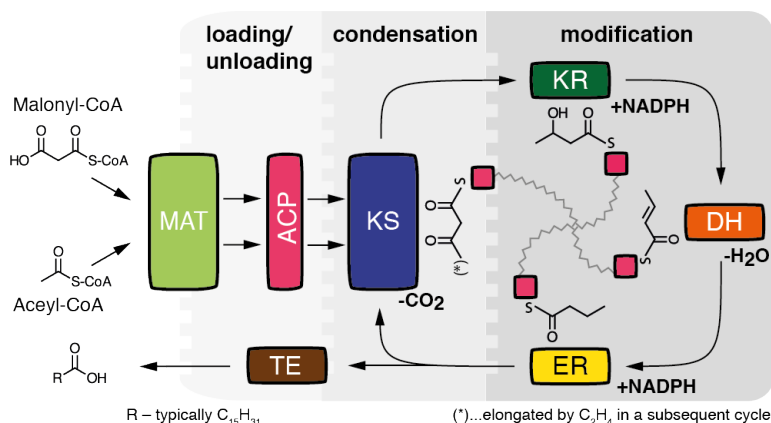
Figure 6: Schematic depiction of the animal fatty acid synthesis. Compartmentalized synthesis by FAS, with the loading function of MAT domain put into focus. The elongation cycle is divided in condensation performed by KS and modification of the β-carbon catalyzed by the domains KR, DH and ER. The flexible attachment of ACP via linkers is abstracted by the zigzag lines for elongation cycle. Abbreviations as introduced in the text. Adapted from reference.[25]

The enzymatic functions of fatty acid synthesis are contained in remarkably different structural frames. In general, type II FAS systems are comprised of separate enzymes and found mostly in plants, bacteria and in mitochondria. In case of eukaryotes, cytosolic *de novo* fatty acid synthesis is performed by large multidomain type I FAS.[26, 27] For example, fungal FASs exhibit large barrel-shaped complexes of about 2.6 MDa, in which the synthesis is enclosed in reaction chambers. In contrast, animal FASs show an open, X-shaped architecture of about 540 kDa.[28-31]

Animal FASs, including also human FAS, are present in all eukaryotic, multicellular organisms. In the terminology of PKSs, animal FAS is an example of an iterative type I PKS, and can hence be treated as a subset of type I PKSs.[14] More precisely, they belong to the third class of iterative type I PKSs, namely the HR-PKSs (section 1.2). A KS-based phylogentic analysis revealed a close evolutionary relation of animal FAS to fungal and eubacterial iterative PKSs I as well as modular type I PKSs and very likely they have shared a common ancestor.[32] Based on the presence of the full set of catalytic activities, this ancestor may have even been quite similar to the animal FAS fold.[14]

Human FAS is ubiquitously expressed in all tissues, but the expression level is highly dependent on the cell type and on the diet.[33, 34] Most prominently, human FAS occurs in breast, adipose and soft tissues, where it catalyzes the production of saturated straight-chain fatty acids, predominantly palmitic acid (C_{16}). Yet, the product spectrum can be altered in lactating mammary glands (shown for rat) to produce medium chain fatty acids, characteristic of milk fat, by utilizing separate, discrete thioesterases (termed

thioesterase II to distinguish from the integral thioesterase I).[35] Furthermore, animal FAS is capable of producing methyl-branched fatty acids with even and odd numbers of carbon atoms in specialized tissues. This has been shown, for example, by a comprehensive study on total lipids in Harderian glands of guinea pigs, in which up to 82 mol% of the esterified fatty acids were branched.[36] All of the methyl branches were found at the even-numbered carbon atoms and it has been already stated that the methyl-groups originated from the usage of methylmalonyl-CoA during fatty acid elongation. Similar findings were also made in specialized glands of many birds (sebaceous and uropygial), in which the major fraction of fatty acids was found to be even multiple methyl-branched.[37] Despite the versatile usage of FASs in various animals, the fatty acid product spectrum is exclusively limited to saturated acyl-chains.

1.3.1   Historical and Enzyme Kinetic Description

Investigating the biosynthesis of fatty acids has a long history going back to the beginning of the 20th century.[38] Inspired by pioneer work on lipid biochemistry by Konrad Bloch and Feodor Lynen, who were awarded with the Nobel Price in Physiology or Medicine in 1964, several laboratories had already purified the corresponding, eukaryotic, high molecular-mass enzymes from yeast and animal tissues by the late 60s. The first *in vitro* studies were conducted with animal FASs extracted from pigeon and rat liver and further examples followed from pig, chicken and goose.[39-42] Already by the 1980s, all domains were identified and functionally characterized and assays for partial and overall FAS activity were available. Functional investigation of the native enzyme was completed by deciphering the steady-state kinetics of FAS in the laboratories of John W. Porter and Gordon Hammes.[43, 44] Their findings shall be introduced briefly in the following:

$$Ac-CoA + 7\,Mal-CoA + 14\,NADPH + 14\,H^+ \qquad\qquad [\,1\,]$$
$$\rightarrow palmitic\ acid + 8\,CoA + 7\,CO_2 + 6\,H_2O + 14\,NADP^+$$
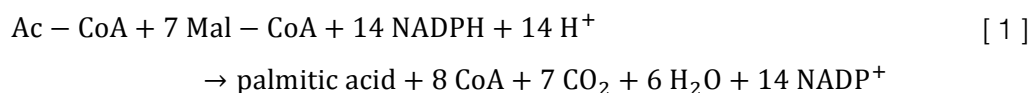
The overall reaction of FAS catalysing the production of palmitic acid is shown in equation [1]. It is an enzyme-catalyzed reaction involving three different substrates, namely NADPH, acetyl-CoA (Ac-CoA) and malonyl-CoA (Mal-CoA). Kinetics for such reactions were developed and introduced mainly by W. Wallace Cleland,[45, 46] and the reaction can be conveniently monitored photometrically by following the consumption of NADPH by the decrease in the absorbance at 334 nm. By measuring over a wide range of substrate

concentrations Hammes *et al*. could quantitatively determine the steady-state parameters and reported the rate law for the steady state initial velocity v:[44]

$$v = \frac{k_{cat}\,[E]_0}{1 + \frac{K_M}{[Mal-CoA]}\left(1 + \frac{[Ac-CoA]}{K_{A,i}}\right) + \frac{K_A}{[Ac-CoA]}\left(1 + \frac{[Mal-CoA]}{K_{M,i}}\right) + \frac{K_N}{[NADPH]}} \qquad [\,2\,]$$

In equation [2], k_{cat} is the turnover number in terms of NADPH consumption, $[E_0]$ is the total enzyme concentration, the K_j are Michaelis constants and the $K_{j,i}$ are inhibition constants. The parameter j refers to the different substrates M – malonyl-CoA, A – acetyl-CoA and N – NADPH, respectively. Interestingly, the two substrates acetyl- and malonyl-CoA act both as substrates and inhibitors, which leads to the practical consequence that the maximal observable velocity (at standard conditions) is only 60 % of $k_{cat}\,[E_0]$.

Further advances in the field of animal FAS were achieved by Stuart Smith, who analyzed the acceptance of noncanonical substrates, e.g. phenylacetyl-CoA, and succeeded to express rat FAS heterologously in *Spodoptera frugiperda* (Sf9) insect cells.[47, 48] Heterologous expression provided an easy access to variants of FAS and this was exploited in innumerous comprehensive mutagenesis studies conducted from the mid 1990s to the beginnings of 2000s. Therein, separated domains were investigated kinetically, specificities of domains were altered and it was reported that mutations in one polypeptide chain can be complemented by the other half of the FAS dimer.[49-52] Although many of the studies addressed the question of how the multidomain architecture is organized, the structure remained elusive until 2006.

1.3.2   Crystal Structure of the Porcine FAS

In 2006, Maier *et al*. succeeded to determine the crystal structure of native porcine FAS at 4.5 Å resolution.[53] This resolution was sufficient enough to properly place secondary-structure elements of most parts into the crystallographic map revealing the architecture of an animal FAS. This was a big breakthrough, as porcine FAS serves also as a paradigm for type I PKSs and granted the first access to structural information for this class of multifunctional enzymes. Only two years later, an improved crystal structure at 3.2 Å resolution was reported from the same lab.[31]
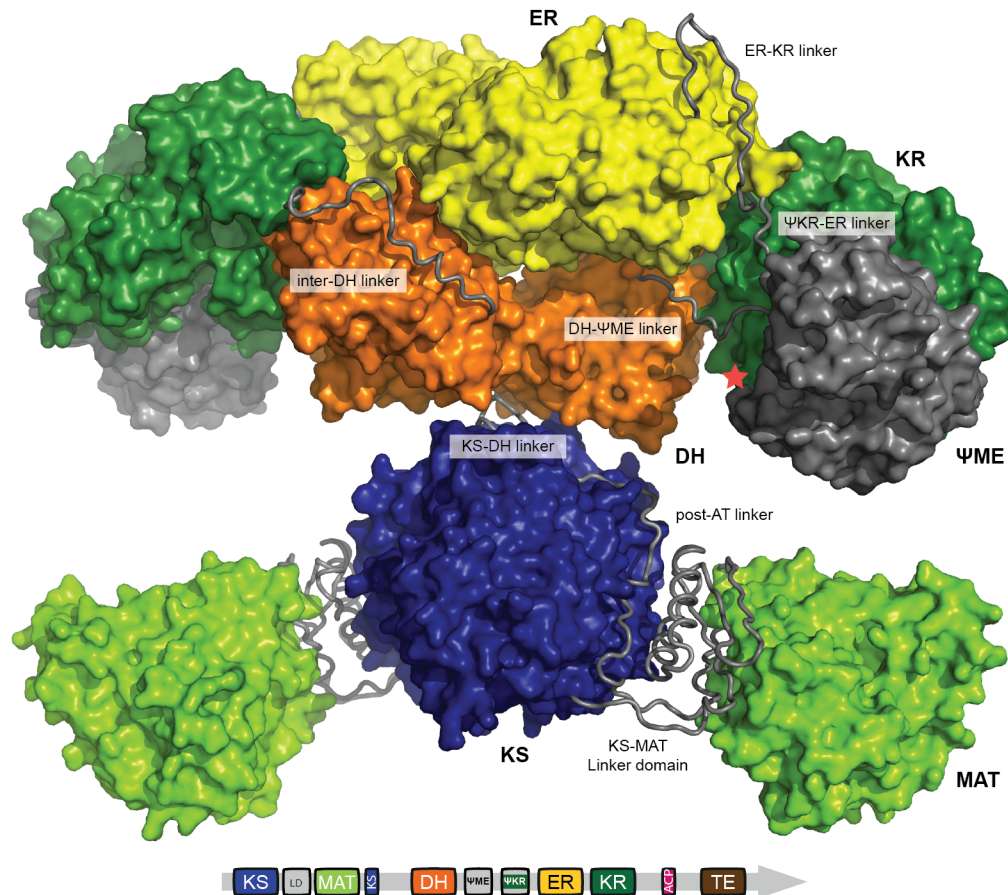
Figure 7: Crystal structure of the porcine FAS. Surface representation of individual domains with inter-domain linkers shown as tubes (porcine FAS, 2vz9).[31] Coloring as introduced in the domain organization of FAS. Red asterisk indicates the C-terminus of the KR domain, where the ACP and TE domains are flexibly tethered, which remained unresolved in the crystal structure.

Porcine FAS assembles into an intertwined dimer adopting an "X" shape with slightly broken C2 symmetry (Figure 7). It is segregated into a lower condensing part, containing the domains KS and MAT, and an upper portion including the domains DH, ER and KR, responsible for β-carbon processing. These two parts are conformationally loosely connected forming only tangential contacts and hence show a remarkable degree of relative positional variability.[54] Additional to the catalytic domains, there are three structural domains present, one is located in between KS and MAT, which is called linker domain (LD), and two at the periphery of the processing part. The first of these domains is homologous to the methyltransferase family and is thus named pseudo-methyltransferase (ΨME). The second is called pseudo-ketoreductase (ΨKR) as it represents a truncated KR fold, which dimerizes with the functional KR domain facilitating its catalytic activity within one chain of the FAS dimer. The C-terminal domains ACP and TE are flexibly tethered to the KR domain (asterisk in Figure 7) and remained unresolved in the crystal structure.

In contrast to fungal FAS,[29] in which almost 50 % of the total sequence is used for scaffolding, animal FAS invests only about 25 % of its total sequence for non-catalytic structural elements. The main fraction (16 %) is utilized for the before mentioned lateral domains ΨME and ΨKR, whereas the remaining ~9 % appear in linkers between the domains. Structurally, the core folds of the active sites closely resemble related domains in PKSs as well as their freestanding counterparts in type II FASs.

Within a multienzyme complex it is essential how functional subunits are connected, as such linkers must allow a sufficient interaction between reaction partners. Besides the structured LD, which connects KS and MAT, various characteristic linkers are present in FAS, which are highlighted in **Figure 7**. The inter DH linker connects the two subdomains of DH forming a catalytic active pseudo-dimer within a polypeptide chain. The C-terminus of DH is linked to the ΨME domain and the linker forms a β-sheet with a stretch of residues in the ΨKR-ER linker. These two linkers are not located on the surface of the protein being buried between the KR, ΨKR and ΨME domain highlighting their structural role for the integritiy of the ΨKR/KR fold. ER is connected to KR via a long, surface exposed linker, in which the linker length indicates the oligomeric state of ER in also homologous PKSs.[55] Finally, the post-AT linker joining MAT and DH shall be highlighted as it plays an essential role in FAS and related PKSs. It is composed of three different parts, the N-terminal residues participate in the LD fold, whereas the C-terminal part forms the connection between the condensing and modifying part. The median part is located on the surface of KS, participating in the domain fold and bears a very conserved motif of residues, which is utilized to identify KS domains in huge gene clusters of PKSs.

The two polypetides dimerize through a contact area of 5400 $Å^2$,[31] which mainly involves residues from the domains KS, DH and ER. The main contribution (area of 2600 Å) to the dimer interface arises from homophilic interactions of KS, which is the only domain that requires dimerization for catalytic activity.

1.4 AT Domains

AT domains play an important role in all PKSs, as they are responsible for the selection of priming and elongating substrates from a potpourri of organic compounds in a cell. The select moiety appears in the final product and thus ATs regioselectively determine many positions, i.e. at all even carbons in the polyketide, participating in shaping the carbon skeleton. Especially, the priming substrate can originate from different

pathways and has been found in huge structural diversity.[56] Here, it shall be mentioned, that loading of a PKS can also be achieved by adenylation domains (A domains) from NRPSs in hybrid systems, which activate acids with adenosine triphosphate (ATP) and offer the activated species to the megaenzyme. A respective example is MicC being part of the micacocidin synthase, which utilizes hexanoic acid as a starter unit.[57]

The fold of all AT domains is structurally very similar, independent of their occurrence in modular proteins or as freestanding enzymes in type II systems or *trans*-AT PKSs, modular type I PKSs with separated ATs, and irrespective of their specificity for loading or extending units. They are characterized by an α/β-hydrolase-like core domain and an appended smaller subdomain with a ferredoxin-like fold. In a cleft in between the two subdomains lies the active site.

1.4.1 Catalytic Mechanism and Enzyme Kinetics

Acyltransferases catalyze the transfer of an acyl-moiety from a CoA-thioester to an ACP-thioester, which subsequently shuttles the acyl-chain to other active sites. Hence,
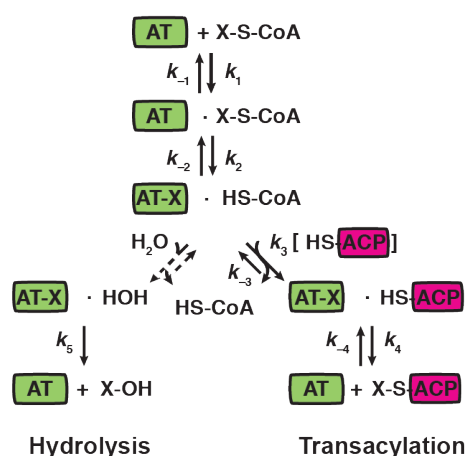
Figure 8: Simplified scheme of the ping-pong bi-bi mechanism of MAT. The acyl-moiety of the CoA-ester is transferred to the enzyme to form an acyl-enzyme intermediate (MAT-X; active site serine-acylated), with the concomitant release of the free CoA (ping). The ACP then interacts with the loaded enzyme intermediate forming the acylated ACP and regenerating the native enzyme (pong). The left branch represents the common side reaction of hydrolysis, which can occur when water is utilized as a nucleophile attacking the acyl-enzyme intermediate.

ATs utilize two substrates (CoA-ester, ACP) converting them into two products (CoA, ACP-ester), which has been classified by Cleland as bi-bi reactions.[45, 46] Such reactions facilitate different catalytic mechanisms and are further classified depending on the existence of a non-covalent or a covalent substrate enzyme state. In the case of AT, the acyl-moiety is temporarily bound to an active site serine creating an acyl-enzyme state (Figure 8), which refers to the subgroup of ping-pong reactions.

Similar to the mechanism of well-studied serine proteases, the active site's nucleophile of an AT domain attacks the carbonyl group of the acyl-CoA leading to the formation of a covalent enzyme-bound acyl moiety and the release of free CoA (ping step). In a second step, the thiol group of an ACP nucleophilicly attacks the carbonyl

group of the acyl-enzyme ester leading to the regenerated enzyme and an acylated ACP (pong step). Common to such reactions, tetrahedral intermediates occur, which are close to the transition states and has to be stabilized by the enzyme, which is accomplished by the existence of a so-called oxyanion whole. Furthermore, the nucleophilicity of the active site's serine is increased by a conserved histidine. Whereas the histidine in serine proteases is further activated by the conjugated base of an acidic residue (catalytic triad), ATs from PKSs utilize backbone carbonyl groups, which is called a catalytic dyad. Although ATs have evolved to transfer acyl-moieties, the acyl-enzyme state may be insufficiently protected leading to the side reaction of hydrolysis (**Figure 8**).

Considering the underlying catalytic mechanism and using steady state assumptions, the rate law for a ping-pong reaction can be obtained:[58]

$$v = \frac{V_{max}\,[X-CoA]\,[ACP]}{K_m^{ACP}[X-COA] + K_m^{X-CoA}[ACP] + [X-CoA][ACP]} \qquad [\,3\,]$$

Where X refers to any preferred acyl moiety. If the value for concentration of ACP [ACP] is fixed, the equation basically reduces to a simple Michaelis-Menten equation (compare equation [8]) with apparent kinetic constants:

$$v = \frac{k_{cat}^{app}\,[E]_0[X-CoA]}{K_m^{X-CoA,app} + [X-CoA]\left(1 + \frac{K_m^{ACP}}{[ACP]}\right)} \qquad [\,4\,]$$

A graphical method to analyze the Michaelis-Menten function is double reciprocal plotting in a Lineweaver-Burk plot. A special characteristic of ping-pong reactions is the appearance of a nest of parallel lines for varying the fixed concentrations of ACP. This resembles the behaviour of uncompetitive inhibitors and leads to increasing apparent Michaelis- *and* rate constants with increasing concentrations of ACP.

1.4.2   Substrate Specificities of AT Domains

ATs can occur with a variety of distinct substrate specificities and are organized in different structural frameworks. Whereas, bacterial type II fatty acid synthesis utilizes FabD (*E.* coli) a freestanding enzyme with very strict malonyl-CoA specificity,[59] there are basically two different types present in modular *cis*-PKSs. These can be classified into loading-ATs and extender-ATs, which differ in their structural embedment into the multidomain architecture and their tolerance, i.e. specificity, for certain substrates or group of substrates. Loading-ATs occur only once at the very beginning of a modular PKS and structurally resemble the independent fold of FabD.[24] Typically, these ATs are characterized by a rather broad substrate tolerance, which has been shown for the loading domains of DEBS, AVES and the rapamycin synthase (RAPS).[24, 60, 61]

Nevertheless, loading-ATs exclude the transfer of extender-substrates and show a preference for their native substrate.

In modular *cis*-PKSs, the extender-ATs are always structurally joined with the KS domain, connected by a linker domain. Most commonly, exclusively malonyl-CoA or methylmalonyl-CoA are used as extender-units (see **Figure 4** and **Figure 5**) by very specialized ATs, though it shall be mentioned that specific kinetic data is barely available. Further, also extender-ATs are known being specialized for other modifications of the α-position of the dicarboxilic acid moiety, like ethylmalonyl-CoA.[62] Current research aims at making also artificial modifications of the α-position available in noncanonical CoA-esters and at engineering the corresponding ATs.[63]
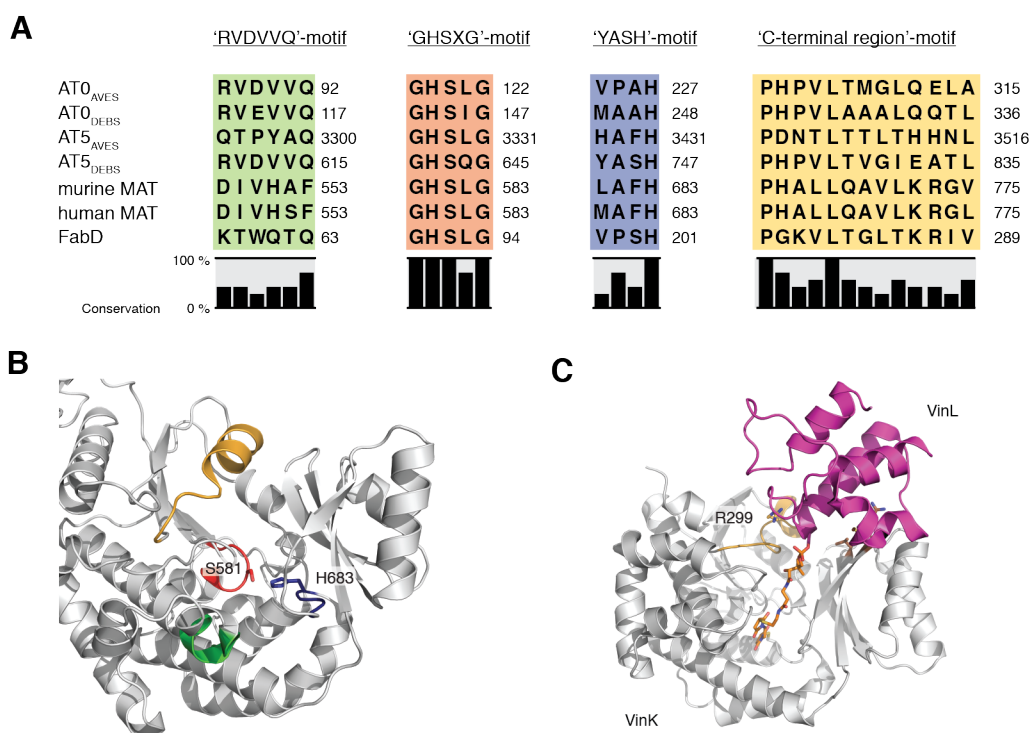
Figure 9: Substrate specificity of ATs. (A) Sequence alignments of various ATs in regards to the four sequence motifs characterizing substrate specificity. Utilized primary sequences refer to loading-ATs: AT0$_{AVES}$ (Q9S0R8) and AT0$_{DEBS}$ (Q03131), specialized extender-ATs: AT5$_{AVES}$ (Q9S0R7), AT5$_{DEBS}$ (Q03133) and FabD (P0AAI9), and bifunctional-ATs: murine MAT (P19096) and human MAT (P49327). (B) Cartoon depiction of the human MAT (3hhd) with sequence motifs colored as above. Active site serine (S581) is shown in sticks representation.[64] (C) Crystal structure of the VinK-VinL complex (5czd) in cartoon representation.[65] The ACP (VinL) was chemically cross-linked (sticks, orange) to the AT (VinK). The authors characterized the binding interface and identified crucial residues including R299 in the 'C-terminal' region (brightorange).

Past studies indicated the existence of mainly four sequence elements that are important for determining the substrate specificity of AT domains.[38] These sequence motifs are summarized in **Figure 9** and shall be introduced in detail. (1) The 'RVDVVQ' motif, 30 residues upstream of the active site serine, is located in a 35 Å-long amphiphatic helix that participates in defining the substrate-binding pocket. It was first recognized in the rapamycin synthase and has been further specified since.[66] ATs specific for methylmalonyl-CoA are characterized by the stretch of residues [RQSED]V[DE]VVQ, whereas the motif ZTX$[AT][QE] is found in malonyl-ATs, in which Z refers to a hydrophilic residue and $ to an aromatic residue, respectively.[62] This consensus sequence has reliable predictive value for the assignment of substrate specificities in ATs of PKSs. Consistently, the two representative extender-ATs from the chosen examples, are predicted to be specific for methylmalonyl-CoA ($AT5_{DEBS}$) and malonyl-CoA ($AT5_{AVES}$), which can be confirmed in the respective polyketide product (see **Figure 4** and **Figure 5**). Interestingly, the 'RVDVVQ' motif is also present in both depicted loading-ATs ($AT0_{DEBS}$ and $AT0_{AVES}$), where it cannot be used to predict substrate specificity. ATs from FAS (murine and human MAT) display a mixed sequence character being consistent with the general acceptance of both substrates, albeit malonyl-CoA is the preferred substrate.

(2) The 'GHSXG' motif contains the active site serine and forms the loop in between the structural elements of the 'nucleophile elbow'. Sequence alignments and information from crystal structures indicate that residue X adjacent to the catalytic serine plays a role in determining specificity. Malonyl-specific ATs utilize branched hydrophobic residues (leucine in $AT5_{AVES}$), whereas a less bulky residue (glutamine in $AT5_{DEBS}$) is preferred in methylmalonyl-specific ATs.

(3) The 'YASH' motif contains the active site histidine forming the catalytic dyad. This motif appears 100 residues downstream of the active serine and is located in the connecting loop in between the ferredoxin-like and the α/β-hydrolase subdomain. Gokhale and co-workers refined the motifs as [YVW]ASH for methylmalonyl-specific ATs and [HTVY]AFH for malonylspecific ATs (compare YASH in $AT5_{DEBS}$ and HAFH in $AT5_{AVES}$). In FAS the corresponding residues are [MILV]AFH (LAFH in murine MAT and MAFH in human MAT).[62]

(4) The 'C-terminal region' motif refers to a stretch of residues in a α-β-α structural motif at the C-terminal end of the α/β-hydrolase subdomain. The first α-helix of this motif is located at the protein's surface and forms part of the substrate pocket entrance. No defined consensus sequence for this motif is known, but it has been shown that swapping this segment can influence the substrate specificity.[67]

Furthermore, it is worth mentioning, that this α-helix also plays a role in defining the AT-ACP interface, which is crucial for the sufficient recognition of an ACP.[68] This aspect has gained importance for the PKS field, as it has been shown, that matching protein interfaces are necessary in modular PKSs and that these are fine-tuned within modules.[69] Hence, also ACP can be treated as a substrate and its respective specificity for an AT is determined by matching protein interfaces.[65]

1.5 KS Domain

The central aspect of natural polyketide assembly is the formation of carbon-carbon bonds. This task is accomplished by a decarboxylative Claisen-condensation taking place in the active site of the KS domain. In fact, this domain is the only essential catalytic site for polyketide synthesis as proven by the type III system. Also in modular PKSs, the minimal functional unit is generated by AT, ACP and KS in a α-module (section 1.2). Not only the role of KS during catalysis is extraordinary, but also the overall architecture of PKSs is prominently determined by the KS interface contributing the largest area for dimerization.[31, 70] Despite significant differences in the architecture of PKSs, all KSs share a common fold, chemical mechanism and also the primary sequence is relatively conserved, which makes the domain most suitable for phylogenetic trees.[32] Besides all similarities, the substrate specificity of KSs differs considerably from very tolerant, e.g. the six KSs in DEBS,[71] to relatively restricted in FASs, which only accept saturated acyl chains for elongation.

Crystal structures of the KS domain revealed an overall fold characteristic for the thiolase-superfamilliy. Each domain consists of two subdomains that exhibit the same topology of alternating layers of α-helices and β-sheets (called $\alpha/\beta/\alpha/\beta/\alpha$ sandwich motif) and together form a homodimeric contact. In human FAS, this interface is created by 401 van der Waals contacts, four salt bridges and 31 hydrogen bonds.[64] The catalytic machinery is comprised of a cysteine as well as two histidines, which is called catalytic triad accordingly. Comparable to AT domains, the nucleophilicity of the active site's residue, in this case a cysteine, is increased by the positioning on a 'nucleophilic elbow' such that the positive dipole of the α-helix decreases its pK_a value.

Mechanism 1: Concerted Decarboxylation-Addition
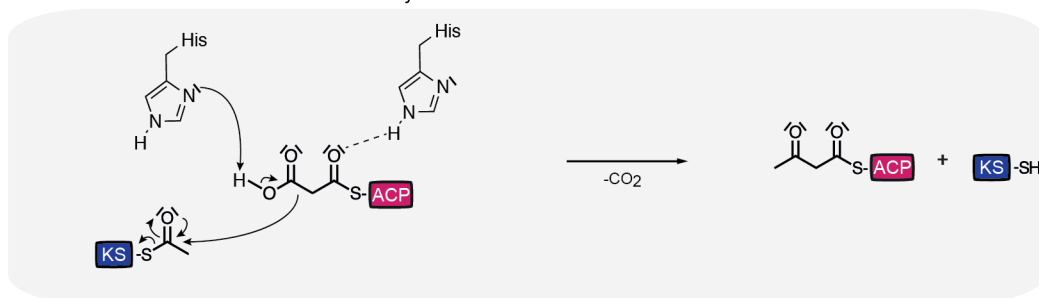
Mechanism 2: Stepwise Decarboxylation-Addition
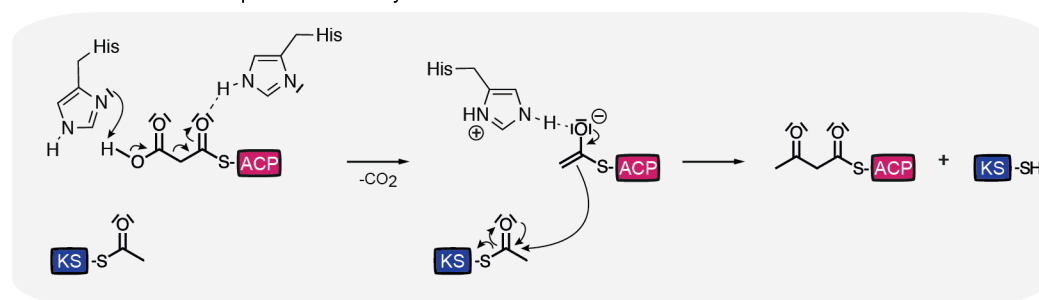
Mechanism 3: Stepwise Addition-Decarboxylation
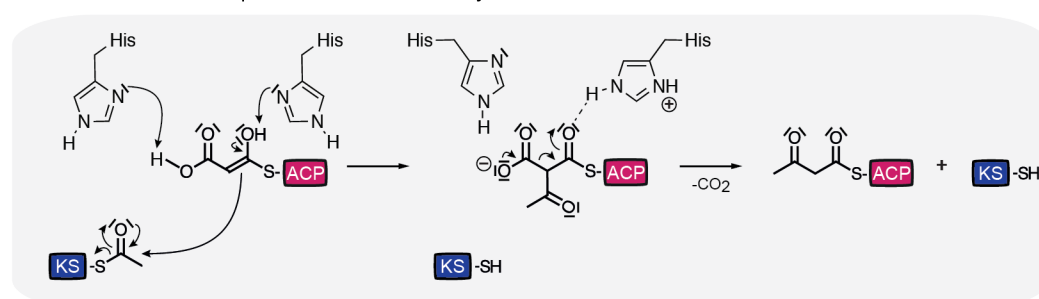
Figure 10: Possible mechanistic pathways for the enzymatic decarboxylative Claisen condensation. For simplicity only the reaction between malonyl and acetyl moieties is shown, but accordingly applicable for other extender and acyl units. Scheme is based on the given reference.[72] Arrows indicate the movement of electrons during progression of the chemical reaction.

Catalytically, the mechanism of KS can be divided into two steps, a transthioesterification reaction leading to the acyl-enzyme intermediate and a decarboxylative Claisen-condensation, which generates the carbon-carbon bond regenerating a liberated active site. In the first step, KS is acylated by either a priming unit loaded by loading-ATs or MAT or by the growing processed polyketide intermediate, which was passed from the upstream module. In iterative PKSs and FASs, the KS domain is repetitively loaded with processed intermediates (saturated acyl chains in FAS) until a certain chain length is reached, which exceeds the size of the substrate binding pocket. As soon as the KS-mediated condensation rate drops, the intermediate is rather released by an associated domain, like a TE domain in FAS, than further elongated. The

second step of catalysis is way more complex and a clear mechanism has not yet been firmly established.[14] Basically, there are three distinct possibilities how the two essential events of decarboxylation and Claisen-condensation can occur (**Figure 10**) and still remain a matter of debate. The first postulated mechanism earmarks a decarboxylation of the malonyl moiety with a concerted nucleophilic attack at the carbonyl group of the acyl enzyme intermediate. This mechanism demands the stabilization of a complex transition state to generate the ACP-bound β-ketoester. Alternatively, two stepwise mechanisms have also been proposed, one requires an initial decarboxylation step to produce the active enolate species followed by electrophilic trapping by the cysteine-bound acyl unit. The other involves an acid-base catalyzed enolization of the malonyl moiety with a subsequent nucleophilic addition of the C-nucleophile to the acyl group followed by decarboxylation. For an elaborate compilation of references for all theories it is referred to the study of Blaquiere *et al.*[72] It shall be mentioned that mechanism 2 is predominantly preferred in the PKS and FAS field with the variation that bicarbonate is supposed to be produced instead of carbon dioxide, which has been demonstrated for a KS variant of rat FAS.[50]

1.6   Motivation for/Hypothesis behind this Thesis

It is generally believed that animal FASs and PKSs share a common ancestor due to the fact that both enzyme classes utilize collective precursors, operate with a similar chemistry and share common domain structures and overall modular architecture. However, animal FAS occurs in a different kingdom, namely the animalia, in which it fulfils a discriminative role being essential in the primary metabolism. Hence, this enzyme has been submitted to a different selection pressure in evolution, which may have led to unique characteristics, such as high catalytic efficiency, the evolution of an unusual bifunctional AT domain (MAT) and a very restrict KS domain. Furthermore, the FAS field has a long tradition making animal FAS a well-described enzyme on a high level of molecular understanding.

Our hypothesis is that animal FAS, due to the before mentioned reasons, may well serve as a *practical* model system to study the working mode of PKSs. Animal FAS has been structurally elucidated by X-ray crystallography, can be handled as a protein *in vitro* and FA synthesis has been kinetically described. The FAS fold exemplifies a complete δ-module bearing all necessary functionalities (including a TE domain) of typical modules of PKSs allowing to study the modularity of such structural arrangements. Furthermore,

certain unique catalytic properties might be practically exploited and applied in hybrid FAS/PKS-systems to create more efficient and sophisticated microbial factories.

1.7 Aim of the Thesis

The overall goal of this thesis was to establish an expression and purification strategy for an animal FAS, best in *Escherichia coli* (*E. coli*) to gain access to this model system for (modular) PKSs. Once inaugurated, the wild-type enzyme shall be analyzed structurally and kinetically in depth with a focus on the "special" domains MAT and KS to lay the foundations for a rational design towards PKSs. As our laboratory was just launched, this must include the establishment of appropriate enzymatic assays to study functionalities and kinetics as well as the organization of processes enabling structural elucidation by X-ray crystallography.

Aiming at engineering animal FAS towards modular PKSs, three levels of complexity have to be faced: i) Specificities and catalytic efficiencies of single domains have to be figured out to test for advantageous properties, which might be exploited in hybrid FAS/PKS systems. This also includes the amenability of such domains for a directed alteration of specificities to broaden their applicability. ii) All possible organizations of domains within a typical module of PKSs (from α-δ modules) have to be rebuilt to evaluate structural and functional integrity of the remaining domains, which is a crucial aspect in terms of modularity. Furthermore, functional communication between these domains must be maintained. iii) It should be possible to connect several such modules to establish an overall architecture beyond a single module. On this level effective communication between modules rather than within a module is crucial and has to be guaranteed.

Chapter II: Results and Analysis of the Data

2.1 Animal Fatty Acid Synthase as a Model System to Investigate Modular Polyketide Synthases

At the outset of this study, we had to establish an easy and sufficient access to an animal FAS in high quality as a model system. At this stage, two important decisions had to be made: we had to choose an appropriate expression system and we had to decide, which homologue of an animal FAS was to be used. Although most of the recent kinetic and biochemical data was collected on rat FAS, expressed in *Spodoptera frugiperda* (*Sf-9*) insect cells in the Smith laboratory,[48] we did not want to follow this strategy for several reasons: insect cell culture was not established in our lab nor in our institute, maintenance of an insect cell culture is very cost-intensive, mainly due to the necessity to buy high quality medium and, most importantly, changing a construct is very time consuming, because the creation of every individual baculovirus takes about four weeks.[73] Taking these criteria into account, *E. coli* is still an unchallenged heterologous host due to low costs and particularly fast mutagenesis cycles making it the organism of choice. Although an animal FAS might seem to be far to complex and huge for such simple expression systems, our decision was supported by a study in the Wakil laboratory, reporting a successful expression strategy for human FAS in *E.* coli.[74]

Choosing the best-suited homolog of an animal FAS is very challenging, as it is difficult to anticipate heterologous expression sufficiency and *in vitro* protein properties. Nevertheless, based on the before mentioned study and the highest impact, we focused on human FAS. This homologue is a member of the animal FAS family and hence suited to investigate the relationship to polyketide synthases, plus it serves as a potential drug target due to its connection to widespread diseases.[75, 76]

2.1.1 Expression of Human FAS in *E. coli*

The project was initiated by cloning human *FASN* genes into bacterial expression plasmids. Two different genes were used, a synthetic, codon optimized gene was purchased from Mr. Gene (GeneArt, ThermoFisher) and a verified cDNA clone was ordered at Source BioScience. Both genes were cloned into a pET22b derived expression vector, generating a N-terminal StrepI-tag and a C-terminal His-tag (Figure 11). Cloning of the synthetic gene encoding human FAS (hFASe will be used for

the codon optimized and hFASh for the native sequence) vector was performed by Ronnald Vollrath, which served as a template for following constructs. Additionally, both genes were cloned into the vector pMAL-c5G (New England BioLabs) generating a N-terminal fusion of the maltose binding protein (MBP) and a C-terminal His-tag according to Jayakumar *et al.*.[74]
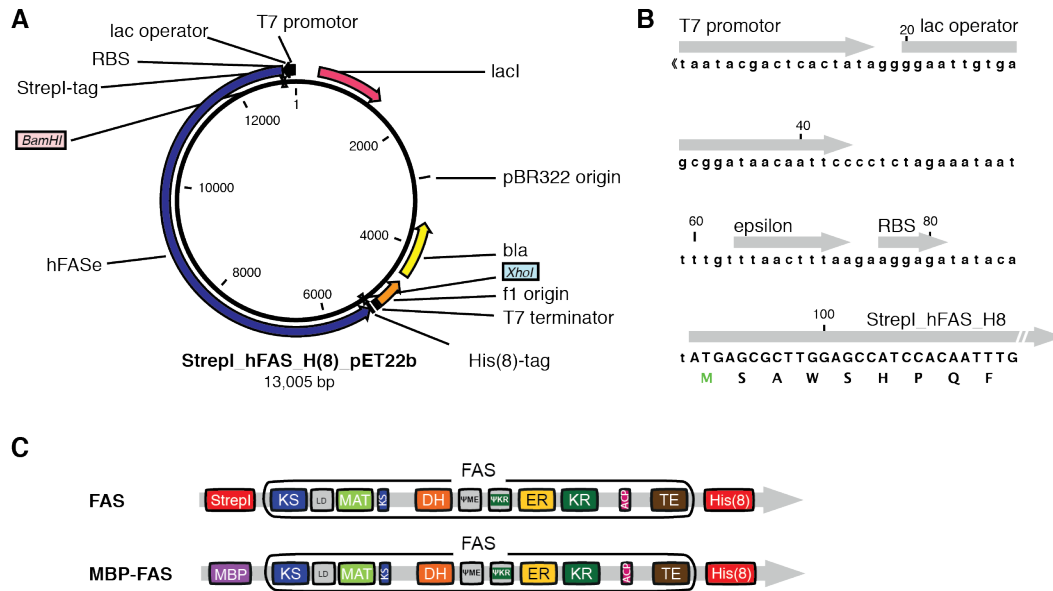
Figure 11: Expression strategy for animal FAS in *E. coli*. (A) Representative vector map of a pET22b derived expression vector generating an N-terminal StrepI-tag and a C-terminal His-tag. This organization was used for the majority of constructs. lacI – Lac repressor protein; bla – β-lactamase; RBS – ribosome binding site. (B) Sequence of the 5' regulatory region of the multiple cloning site of pET22b and the 5' part of the N-terminally Strep-tagged hFAS. (C) Domain organization of FAS constructs containing different tags. Abbreviations were introduced before, except MBP corresponding to the maltose binding protein fused to the N-terminus.

Contrary to Jayakumar's report, all attempts to express the human FAS in *E. coli* were either unsuccessful or yielded aggregated material, as also stated elsewhere.[64, 74] Various expression strategies were tested from using different fusion proteins, cell lines, expression temperatures, concentrations of inducer IPTG to different cell densities when induced. Representative results are summarized in **Figure 12**.
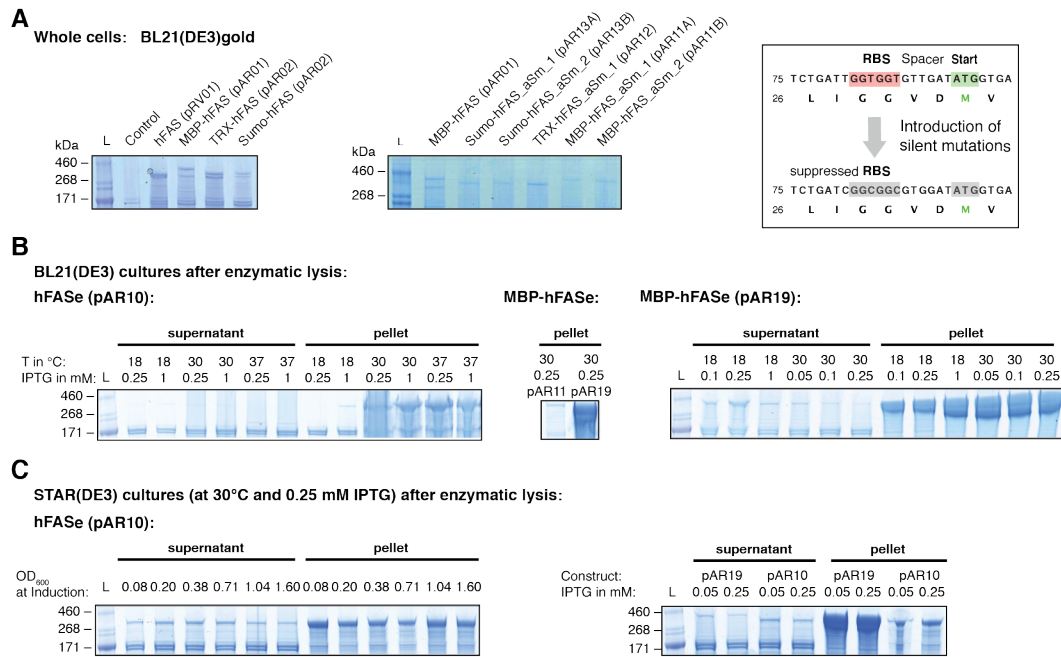
Figure 12: Investigating expression strategies for human FAS in *E. coli*. (A) First test-expression of different fusion constructs of hFASe expressed in Bl21(DE3)gold cells. Whole cells were loaded on the gel to analyze the sufficiency of expression. Due to the appearance of a second band for all constructs (left panel), an alternative translation start was identified and suppressed by introducing silent mutations in the ribosome binding site (right panel). The alternative start codon of M32 is highlighted in green. (B) Influence of temperature and inducer concentrations on the expression of hFASe (left panel) and MBP-hFASe (right panel) in Bl21(DE3). Additionally, two different expression plasmids for MBP-hFASe (pAR11 and pAR19) were tested (middle). Supernatants and pellets after centrifugation of enzymatic lysed cells were loaded. (C) Influence of the cell density at induction on the expression of hFASe (left panel) and MBP-hFASe (right panel) in STAR(DE3)BL21 cells. All shown Coomassie-stained SDS-PAGE gels are NuPage 4-12 % bis-tris gradient gels.

The first expression of four different constructs of hFASe (N-terminal StrepI-tag, MBP-, TRX- and Sumo-fusion) in Bl21gold cells was analyzed via SDS-PAGE, performed on whole cells (**Figure 12A**). Every lane showed a faint band corresponding to the expected sizes (hFAS: 278 kDa; MBP-hFAS: 320 kDa; TRX-hFAS: 290 kDa; Sumo-hFAS: 287 kDa) implying the expression of full-length polypeptide chains by the bacterium. Interestingly, a second, very prominent band appeared at a slightly smaller, but similar size in all lanes indicating truncated protein. Due to the fact that all different N-terminal fused constructs resulted in the same truncated protein, we concluded that either a strong cleavage site or an alternative translation start codon must exist at the beginning of hFAS. Indeed, there is a methionine at position M32 with an appropriate strong translation initiation site and two silent mutations in the alternative ribosome binding site abolished the production of truncated FAS (**Figure 12A**).

Based on the finding that *E. coli* is capable of translating the human *FASN* gene into a full-length polypeptide chain regardless of a N-terminal fusion, we next investigated whether hFAS is also properly folded. Test-expressions in medium scale (50 mL) were performed to simultaneously screen the impact of different expression conditions. Cells were lysed mildly by lysozyme treatment in a stabilizing buffer and the insoluble fraction was separated by centrifugation (see 4.1.1). Both fractions (inclusion bodies were dissolved in 8 M urea) were analyzed by SDS-PAGE. Unfortunately, most of the hFAS was expressed insolubly independent of the tested expression conditions and also the N-terminal fusion of MBP did not increase the fraction of soluble protein significantly (**Figure 12B**). Interestingly, the total amount of produced protein was dramatically increased by the MBP-fusion, but only when the plasmid pAR19 was used for expression (**Figure 12B**; middle panel). This plasmid is based on the vector pMAL, which contains the whole regulatory region of *malE* under a TAC promoter. Furthermore this plasmid does not have a N-terminal tag and directly starts with the *malE* sequence.

In order to increase the fraction of soluble protein, we tested several approaches. First, the improved *E. coli* BL21 strain Star™ was used, which offers a better mRNA stability due to a mutation in the *rne131* gene. Additionally, the cell density at the time point of IPTG induction was varied and the results are shown in **Figure 12C**. From the Coomassie-stained SDS-PAGE, it seems that expression in Star™ cells resulted in a slightly increased soluble fraction, especially, when cells were induced at an OD_{600} between 0.2-0.7. Again, a fusion with MBP did not increase the soluble fraction of protein. Nevertheless, the main fraction remained insoluble.

A last attempt to support a soluble expression of hFAS in *E. coli* was tried by co-expression with molecular chaperones. These are known to assist the folding of polypeptide chains and the assembly of macromolecular structures. The family of chaperones is very diverse and so is their function, but different chaperone systems were reported to facilitate heterologous expression of especially mammalian proteins in *E. coli*.[77, 78] Indeed, a potential role of the human chaperonin TRiC/CCT was assigned to the folding process of hFAS.[79, 80] Without going into a detailed analysis of how chaperones work, we decided to test a broad variety of chaperones from *E. coli*: Trigger factor (TF), DnaK-/DnaJ/GrpE, GroEL/GroES, ClpB and IbpA/IbpB plus chaperones of the thermophilic organism *Pyrococcus furiosus*: prefoldin (pfuPfdB), chaperonin (Cpn) and PF1883.

The genes for all chaperones were either obtained from plasmids, provided by Prof. Hartl (MPI Martinsried) or from genomic DNA (*E. coli*: self-prepared (see 4.1.1); *P.*

furiosus: DSMZ (DSM Nr.: 3638). We decided to use the pETcoco system (Merck) as co-expression vector, as it allows a switchable copy number in the cell from low to medium copy. Every chaperone system consisting of more than one gene was organized polycistronically. Usually, the 5' regulatory region of pET22b from the XbaI restriction site was used as translation initiation site between the genes. The effect of co-expression was again tested in 50 mL test cultures with subsequent separation of the soluble and insoluble fraction (**Figure 13**).
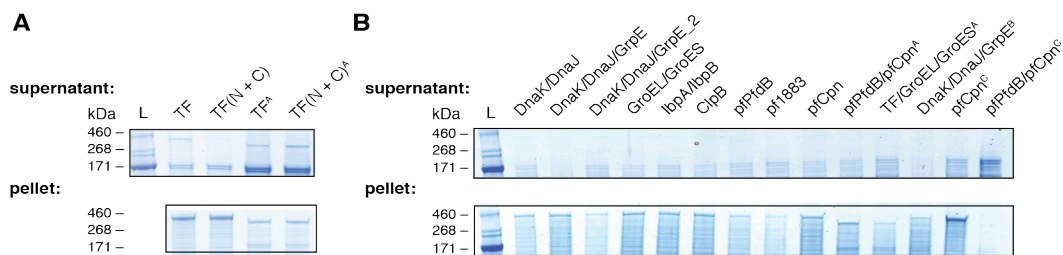
Figure 13: Co-expression of hFAS with chaperones. (A) Co-expression of MBP-hFAS and hFAS (indicated by superscript A) in Bl21gold(DE3) cells with trigger factor. Coomassie-stained SDS-PAGE of the supernatant and pellet after enzymatic lysis of 50 mL cultures (4 h/30 °C) are shown. TF(N + C) refers to a truncated version containing the fused N- and C- terminal part of the trigger factor. (B) Co-expression of MBP-hFAS and hFAS (indicated by superscript A) in Bl21gold(DE3) and Star(DE3) (indicated by superscript B) cells with various chaperones from *E. coli* and *P. furiosus*. Coomassie-stained SDS-PAGE of the supernatant and pellet after enzymatic lysis of 50 mL cultures (4 h/30 °C) are shown. Superscript C indicates co-expression at 47 °C.

No chaperone system, neither from *E. coli* nor from *P. furiosus*, was able to prevent aggregation of hFAS. Further, hardly any improved solubility, as traced by the proportion of soluble protein, could be observed. Only co-expression with TF seemed to result in the appearance of small amounts of hFAS in the soluble fraction (**Figure 13A**). Unfortunately, the amount of total protein loaded on the gel was too low, which was judged by the thickness of bands at ca. 171 kDa representing the largest inherent proteins of *E. coli*. Therefore, it is difficult to draw a final conclusion, but a major effect of these chaperone systems on hFAS folding can be excluded.

Nevertheless, due to the appearance of small bands in the soluble fractions of our test expressions, we decided to test an expression on a larger scale. It is well known that the scale of a bacterial culture influences expression yields and also Jayakumar *et al.* used optimal conditions of a fermenter (15 L scale) to express hFAS.[74, 81]
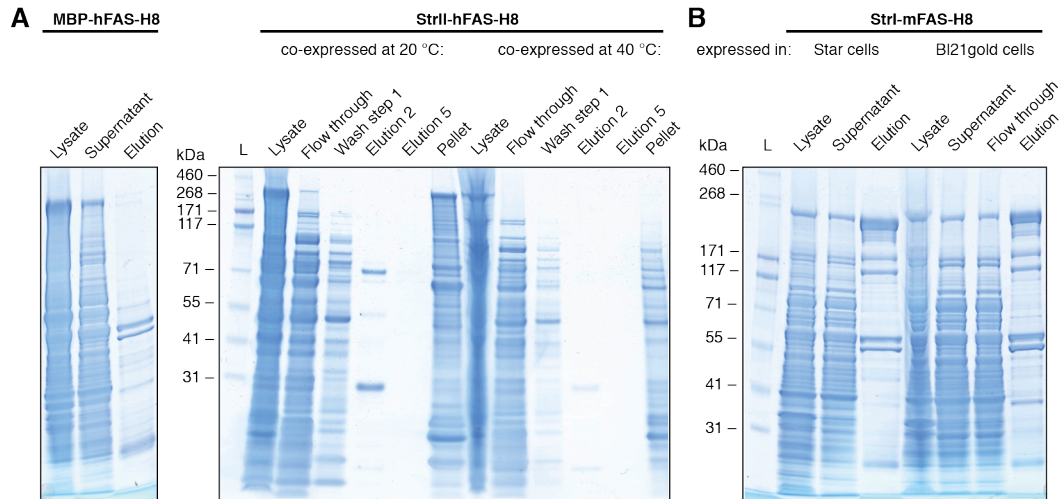
Figure 14: Expression of human and murine FAS on a 1-2 L scale. (A) Expression of MBP-hFAS in a 2 L Bl21gold(DE3) culture and co-expression of hFAS with chaperones (pfuPfdA, pfuPfdB and pfuCpn bearing a mutation to increase activity at lower temperatures (MAmut)[82] in 1 L cultures at 20 °C and 40 °C. (B) Expression and purification of mFAS, expressed in 2 L cultures of Star(DE3) or BL21gold(DE3) cells. Coomassie-stained SDS-PAGE of the cell lysis (French press) and purification are shown.

It can clearly be seen that MBP-hFAS is expressed at full-length in *E. coli* (**Figure 14A**). Comparing the lane with lysate to the supernatant shows a dramatic decrease of protein after centrifugation reflecting the expected high portion of insoluble protein. Nevertheless, a sufficient band is still present in the supernatant fraction, but this protein could not be purified by immobilized metal affinity chromatography (IMAC). This finding implies that either the C-terminal His-tag was inaccessible or that hFAS was not properly folded, although remaining in the supernatant. A third possibility could be that hFAS denatured during IMAC due to instability in the elution buffer (300 mM imidazole) or simple time reasons. hFAS without a N-terminal fusion domain behaved similarly and again no beneficial effect of chaperone co-expression was detected.

Based on this data, we concluded that hFAS could not be considered as an appropriate model system to study PKSs as expression in *E. coli* is too insufficient. Taking into account that also point mutations have to be tolerated and that those may further decrease expression yields, we focused on finding a more suitable homolog. A promising candidate seemed to be the murine FAS (mFAS) sharing 81 % identity to hFAS and a clone with cDNA was available at Source BioScience. The same cloning strategy was applied (see **Figure 11**) resulting in a construct with a N-terminal StrepI-tag and a C-terminal His-tag. The first attempt to express and purify mFAS in *E.* coli is shown in

Figure 14B, where the appearance of a respective band after elution from IMAC was very promising.

2.1.2 Quality Control and Characterization of Murine FAS

Based on the fact, that mFAS is expressed solubly in *E. coli*, we further worked out a purification strategy, which was optimized for a two-step affinity chromatography. The fraction of soluble protein was applied to Ni-NTA-columns and subsequently a second chromatography was performed with the elution fraction from IMAC using Strep-Tactin material. This procedure provided mFAS in high purity (**Figure 15**), was relatively fast (from lysis to final elution ca. 5 h) and did not involve any extreme buffer conditions. A typical yield of 1.5-2 mg of purified protein was achieved from 1 L cell culture. Interestingly, mFAS has only little tendency to aggregate compared to hFAS analyzed by inclusion bodies work up (pellet lane in **Figure 15A**).
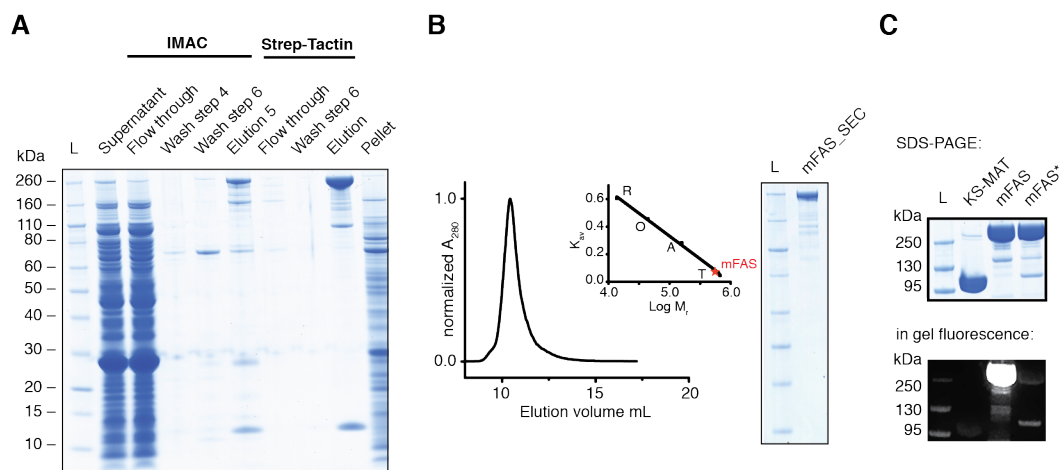
Figure 15: Purification and quality control of murine FAS. (A) Tandem affinity purification strategy using IMAC and subsequent Strep-Tactin chromatography. A representative Coomassie-stained SDS-PAGE of mFAS, co-expressed with Sfp, is shown. (B) SEC of purified mFAS with absorption normalized to the highest peak. Peak corresponds to an apparent molecular weight of 601 kDa (theoretical weight: 554 kDa). Calibration is shown in inset; thyroglobulin (T), aldolase (A), ovalbumin (O) and ribonuclease (R). (C) In-gel fluorescence for monitoring the degree of phosphopantetheinylation in co-expressions with Sfp (indicated by *; lower panel) and Coomassie-stained SDS-PAGE gel (upper panel). Proteins were incubated with 10 mM $MgCl_2$, 10 µM purified Sfp and 1 µM CoA 647 for 1 h at 37 °C. SDS-PAGE gel (NuPAGE 4-12 % Bis-Tris) of fluorescently labeled constructs is shown after fluorescence detection and Commassie-staining. For monitoring phosphopantetheinylation, KS-MAT (pAR70 (see section 2.3.1); not containing ACP) was used as a negative, and FAS, expressed as apo-protein was used as positive control.

The integrity and oligomeric state was further investigated by size exclusion chromatography (SEC). As reported by Smith *et al.*,[52] the monomeric and dimeric states are in a temperature and buffer (ionic strength) dependent equilibrium. After incubation at 37 °C in the appropriate buffer (see section 4.1.2), the main fraction of mFAS was in a dimeric, potentially active state (**Figure 15B**). No significant aggregation peak was ever seen in SEC.

Furthermore, we aimed at finding out, whether *E. coli* is also capable of performing the essential posttranslational modification of the ACP. Hence, a fluorescent *in vitro* phosphopantetheinylation assay was carried out to quantify the degree of modification.[83] In brief, purified Sfp (available as stock in our lab) is used to transfer a fluorescent CoA-derivate, usually CoA-647, to murine ACP *in vitro*. Only protein in the apo-state can be fluorescently labelled and the degree of labelling can be quantified using SDS-PAGE. As shown in **Figure 15C**, in gel fluorescence was detected and the appearance of a strong fluorescent band for mFAS indicates, that *E. coli* does not sufficiently modify ACP. This assay can normally not rule out that a minor fraction of ACP has been modified, but it is confirmed later within this thesis (section 2.2.3, **Figure 24E**) that the bacterium expresses mFAS *entirely* in the apo-state due to inactivity of the protein in an activity assay. Though, mFAS can be obtained in the holo-state by co-expression with Sfp (section 4.1.4). Comparing the corresponding band to a negative control (KS-MAT without any ACP; see section 2.2) allowed quantification of phosphopantetheinylation of more than 98 %, after co-expression with Sfp.

Having pure, dimeric, post-translationally modified mFAS at hand, paved the way to functionally characterize the multidomain protein. Two assays are usually used in literature, namely a NADPH consumption assay at the presence of malonyl- and acetyl-CoA indicating the production of fatty acids (section 4.1.18) and the production of triacetic lactone (TAL) by condensation of malonyl- and acetyl-CoA in the absence of NADPH (section 4.1.20).
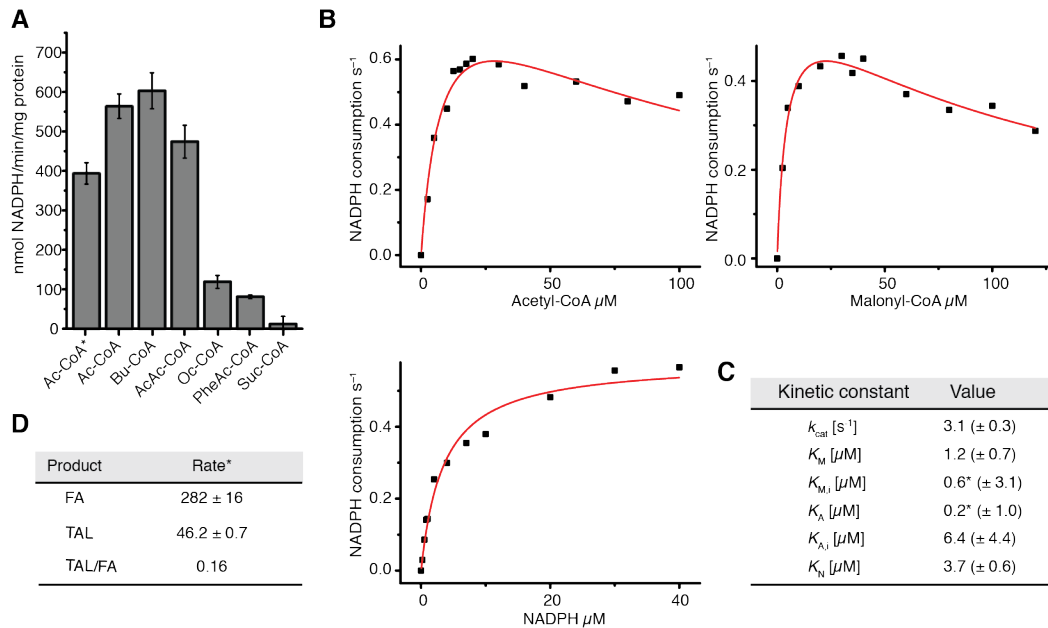
Figure 16: Activity of murine FAS. (A) Specificity of mFAS for different starter units. NADPH consumption assay was performed as described in section 4.1.18. Substrate concentrations were 100 µM X-CoA, 100 µM malonyl-CoA and 40 µM NADPH, respectively. The asterisk indicates usage of a higher concentration of potassium phosphate (350 mM). Error bars show the standard deviation between three biological replicates (B) Typical plots of the initial velocity divided by the total enzyme concentration versus the concentration of acetyl-CoA (20 µM malonyl-CoA and 20 µM NADPH), malonyl-CoA (20 µM acetyl-CoA and 20 µM NADPH) and NADPH (30 µM malonyl-CoA and 15 µM acetyl-CoA). NADPH consumption was monitored fluorometrically. The curves are the best fit to equation [7] of all data. Global fits were performed with the Origin.(C) Steady-state kinetic parameters for the best fit of three curves determined in biological triplicates by Aaron Himmler with optimized conditions within his Master thesis.[84] The asterisk indicates that the ratio of $K_A/K_{M,I}$ is well defined, but the absolute value of either is not.[44] (D) Comparison of the consumption rate of malonyl-CoA of mFAS for FA and TAL synthesis. The rate is given in nmol malonyl-CoA consumed per minute and mg protein. For FA synthesis, the rate of NADPH consumption was divided by two and the rate of TAL production was multiplied by two, respectively. The TAL production rate was 23 ± 2.0 nmol/min/mg protein (n = 3).

Although monitoring the consumption of NADPH is very simple, one has to evaluate many parameters, i.e. buffer conditions, protein stability and substrate concentrations, to assay the activity of mFAS properly. Most of the work to establish the set-up was performed within the supervised master thesis of Aaron Himmler.[84] First of all, we determined an appropriate enzyme concentration (20 nM) to measure in the linear range over five minutes. At such low protein concentrations mFAS loses catalytic activity quite rapidly in the assay buffer, why we had to screen for stabilizing reagents. We identified PEG 400 to have a stabilizing effect and hence we prepared all solutions containing enzyme with 20 % (v/v) PEG 400, which led to a final concentration of 5 % (v/v) PEG 400 during FA synthesis. Further, we found out, that high salt concentrations,

either added sodium or potassium chloride or potassium phosphate, reduced the NADPH consumption rate (compare first two columns in **Figure 16A**).

Utilizing optimized conditions, mFAS showed a specific activity of 564 ± 31 nmol/min/mg protein when using acetyl- and malony-CoA (for a comparison to reported values, see section 3.1). Further, the ability of mFAS to accept other priming substrates was tested (**Figure 16A**) revealing that the loading with butyryl-CoA led to slightly higher NADPH consumption rates (603 ± 46 nmol/min/mg protein). This finding agrees well with a study from Abdinejad *et al.*, who compared different mammalian and avian FASs, and demonstrated that selected mammalian FASs prefer butyryl-CoA as priming substrate in contrast to avian FASs.[42] Also substrates acetoacetyl-, octanoyl- and phenylacetyl-CoA were accepted as priming substrates, though at significantly reduced rates, especially for the latter two. Succinyl-CoA cannot serve as a loading substrate in the wild-type protein.

In order to obtain steady-state kinetic parameters for mFAS, we performed an analysis as reported by Cox *et al.*.[44] All three substrates (acetyl-CoA, malonyl-CoA and NADPH) were individually varied while keeping the other two substrates constant and initial velocities were plotted accordingly (**Figure 16B**). All data was fit with equation [2] and parameters were determined by global fitting with Origin as described in section 4.1.19. The plots demonstrate the inhibitory effects of both substrates acetyl- and malonyl-CoA, as they are both loaded via the MAT domain. A comprehensive analysis in biological triplicates was performed by Aaron Himmler and obtained parameters are listed in **Figure 16C**.

Finally, the production of TAL in the absence of NADPH was measured and we obtained a rate of 23 ± 2.0 nmol/min/mg protein for mFAS using the substrates acetyl- and malony-CoA. For a better comparison, turnover rates were converted into consumption rates of malonyl-CoA per minute and mg protein. The analysis yielded a value of 282 ± 16 nmol malonyl-CoA/min/mg protein for FA synthesis and a value of 46.2 ± 0.7 nmol malonyl-CoA/min/mg for TAL production of mFAS, respectively (**Figure 16D**). These values confirm a preference of the KS domain for fully reduced acyl chains (production of TAL/FA = 0.16) and hence a higher efficiency of mFAS under reducing conditions in the presence of NADPH.

2.2 Modularity of the Multidomain Architecture of mFAS

Goal of this project was to investigate whether animal FAS fulfils the criteria of modularity. This is a widely-used concept, which has been applied interdisciplinarily (e.g. in computer systems), and may well serve to describe adaptability of biological systems in general, which is the basis of evolution.[85] In the context of modular PKSs, Prof. Khosla has phraised the definition: "A modular system may be defined as a multi-component system that can be divided into smaller subsystems, which interact with each other and can be predictably interchanged for functional flexibility and variety".[86] Animal FAS is such a multi-component system providing all catalytic activities necessary for fatty acid synthesis on one polypeptide chain, where domains are structurally organized in one unit, a module. Modularity in the before mentioned definition shall be examined critically within the animal FAS fold by demonstrating the possibility of deleting domains without affecting the integrity of the remaining ones.
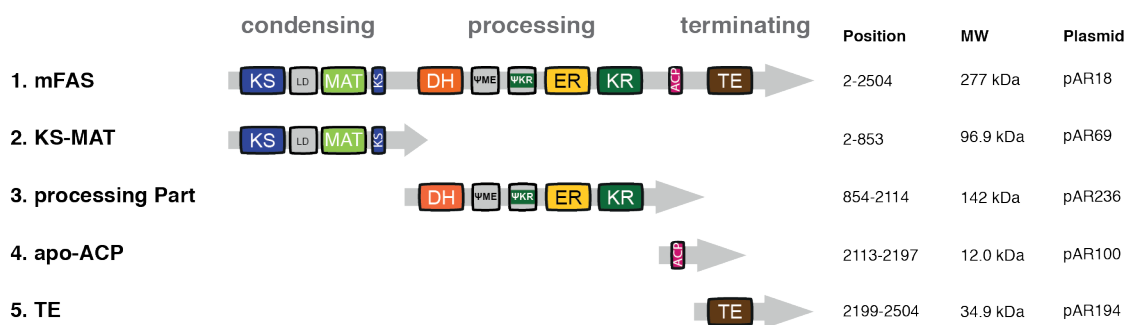
Figure 17: Modularity of the FAS fold. Domain organization of murine FAS and first engineered constructs to investigate modular characteristics. Constructs were designed using sequence alignments of animal FASs (**Figure S2**) and the crystal structure of porcine FAS (pdb: 2vz9).[31] All amino acid positions/numbers refer to the wild-type murine FAS. Molecular weights were calculated with CLC Main Workbench without the N-terminal formyl-methionine.

This project was initiated by dividing mFAS into basically three parts, namely the condensing, processing and terminating part (Figure 17). All constructs were again designed with a N-terminal StrepI-tag and a C-terminal His-tag facilitating usage of the established tandem purification strategy. Additionally, the ACP was provided as a separate domain with the exception of having a N-terminal StrepII-tag to reduce the size of the construct. It turned out during the thesis that this tag was not sufficient to bind ACP to Strep-Tactin material and hence ACP purification was accomplished by IMAC and

SEC. A detailed description of how ACP is expressed and purified as apo or holo domain is shown in section 2.3.1.4.

The condensing part, termed KS-MAT herein, is not only the heart of fatty acid synthesis, but also played the central role for this thesis. Designing the construct was relatively easy, as it starts with the N-terminus and ends at the waist region of FAS, which can easily be identified by typical linker amino acids (serine, alanine, glycine, proline) in the sequence alignment (**Figure S2**). Additionally, a crystal structure of the highly homologous human KS-MAT didomain served as template.[64] KS-MAT expressed quite well with typical yields of 12-15 mg purified protein from a 1 L *E. coli* culture, which was roughly ten times the yield of mFAS. This construct was used for crystallisation (section 2.4.1 and 2.4.3) and for kinetic studies on the domains MAT (section 2.3.2) and KS (section 2.3.4). Furthermore, the corresponding plasmid (pAR69) served as template for site directed mutagenesis to engineer the specificity of the MAT domain (section 2.5).

The processing part, containing residues from the hinge region until the end of KR (854-2114), was of minor importance, as studies on this part were performed on constructs based on the whole mFAS (**Figure 19**). Nevertheless, also the processing part was divided into sub-constructs (pAR236-241) and all expressed very well (**Figure S3**). Indeed, the processing part expressed approximately three times better than the condensing part and 38 mg protein were purified from 1 L culture. Consequently, this construct was used to investigate the effect of an acetylation of Lys1988 (introduced by amber codon suppression, section 2.6) on KR activity within the scope of a master thesis by David Drexler under my supervision.[87]

The thioesterase controls fatty acid chain length by terminating the iteration as counterpart of the KS domain. The separated domain (thioesterase I) expressed very well (ca. 100 mg/1 L culture), but only when the N-terminal StrepI-tag was removed. The tag was not accessible in the first construct (pAR169) and led to an increased tendency to aggregate. The two thioesterase I constructs together with two thioesterase II constructs (pAR115, pAR193) were investigated by Daniel Beyer in a bachelor thesis under my supervision.[88] Unfortunately, access to analytics and instruments was limited at that time, making the results difficult to interpret. Nevertheless, the strategy to quantify the released thiol after CoA-ester hydrolysis with 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM) seemed promising and a strong substrate inhibition could be identified.[89] Utilizing well-established assays for enzyme kinetics, which will be introduced later within this thesis (coupled enzyme system in section 2.3.1.3 and methylumbelliferone

esters in section 2.3.1.2),[90] would make the thioesterase an easy target to investigate, but this was not anymore scope of this thesis.
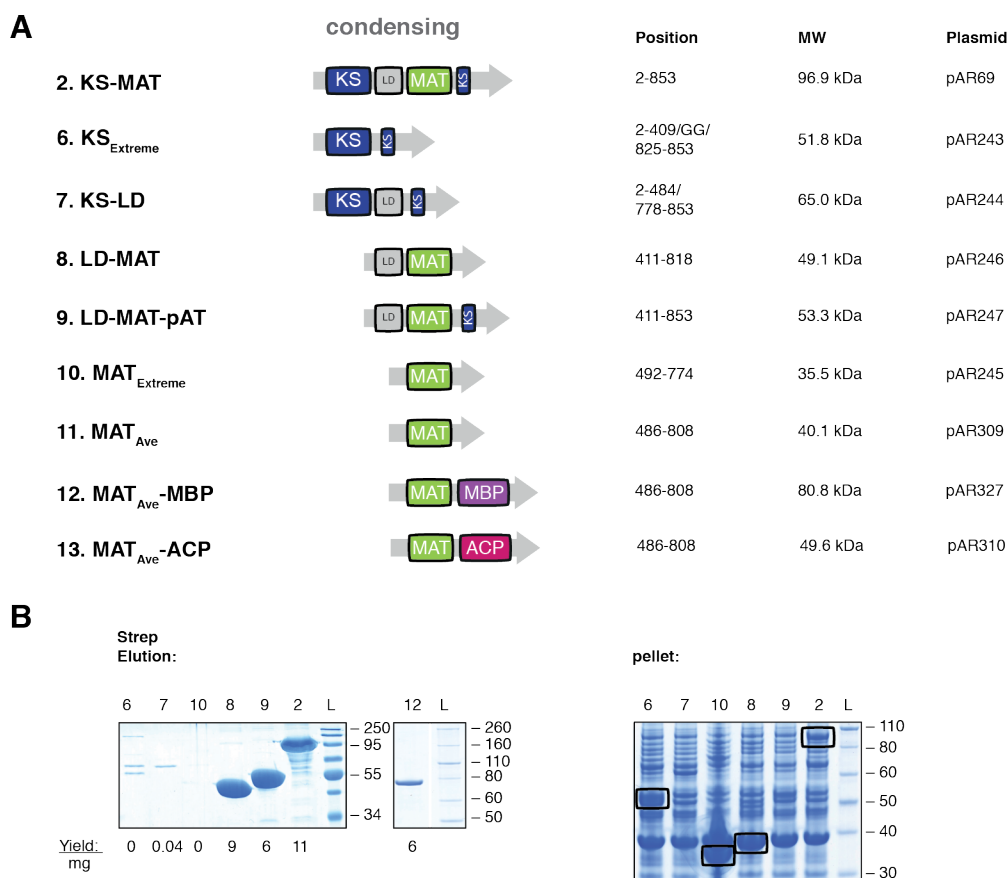
Figure 18: Modularity of the condensing part. (A) Domain organization of different constructs containing KS or MAT. The design is based on the crystal structure of murine KS-MAT (section 2.4.1; pdb: 5my0) and avermectin loading AT (pdb: 4rl1).[24] All amino acid positions/numbers refer to the wild-type murine FAS. Molecular weights were calculated with CLC Main Workbench without the N-terminal formyl-methionine. (B) Coomassie-stained SDS-PAGE of purified and aggregated truncated constructs based on the condensing part. Aggregated protein is highlighted by black boxes. Protein yields refer to purified protein from 1 L cell culture.

Furthermore, we were interested in modularity of the condensing part. Main goal in this regard was to release MAT from the rigid KS-MAT fold. Such KS-independent ATs appear in PKSs as loading ATs or as *trans*-ATs. At first, constructs were designed, which contained both domains with or without the linker domain (LD) (**Figure 18**). It turned out that the KS domain could not easily be released from MAT and both constructs without (**6**) or with LD (**7**) hardly expressed. On the contrary, MAT was also expressible without the KS domain. Two constructs containing LD with (**8**) or without the post AT linker (**9**) expressed well with yields between 5–10 mg of purified protein from a 1 L culture. Then

MAT was released from the LD utilizing structural information of the avermectin loading AT (**Figure S4**).[24] One extreme shortened construct (492-774; **10**) lacking a part of a β-sheet (780-783) and the long C-terminal α-helix (792-806) was not properly folded and occurred in inclusion bodies (**Figure 18B**). Also a slightly longer construct (486-806; **11**) matching the fold of the avermectin loading AT was difficult to express. Nevertheless, this MAT_{Ave} fold (486-806) was properly expressed when either a MBP (**12**) or an ACP (**13**) was attached as domain to the C-terminal helix (see **Figure S4** for linker sequence). Besides the expressibility as soluble proteins in *E. coli*, constructs **8** and **12** were further characterized and proven to be functional (see section 2.3.1.5).

2.2.1   Engineering of PKS-like Reducing Modules with mFAS

Within the AT-containing modular type I PKSs (*cis*-AT PKSs) there are essentially four common module types, namely α- (KS-AT-ACP), β- (KS-AT-KR-ACP), γ- (KS-AT-DH-KR-ACP) and δ-modules (KS-AT-DH-ER-KR-ACP) (section 1.2).[14] Within this section it shall be investigated, whether the animal FAS fold allows creation of such modules taking the architecture of PKSs as templates. In this section the focus is set on reducing modules.

As before mentioned the main difference between module types lies in the organization of the processing part. To understand the architecture of this part and hence the design for PKS-like modules, we have to take a closer look on the porcine crystal structure (**Figure 19A**). The KR domain is not only the first catalytic site during β-carbon processing, but also plays a central role in structuring the processing part. It serves as a hub with all domains directly attached to the surface of the domain without much interaction between each other. The domains DH and ψME are positioned N-terminally to the ψKR domain and the ER domain is inserted between the structural ψKR and functional KR part (**Figure 19B**).
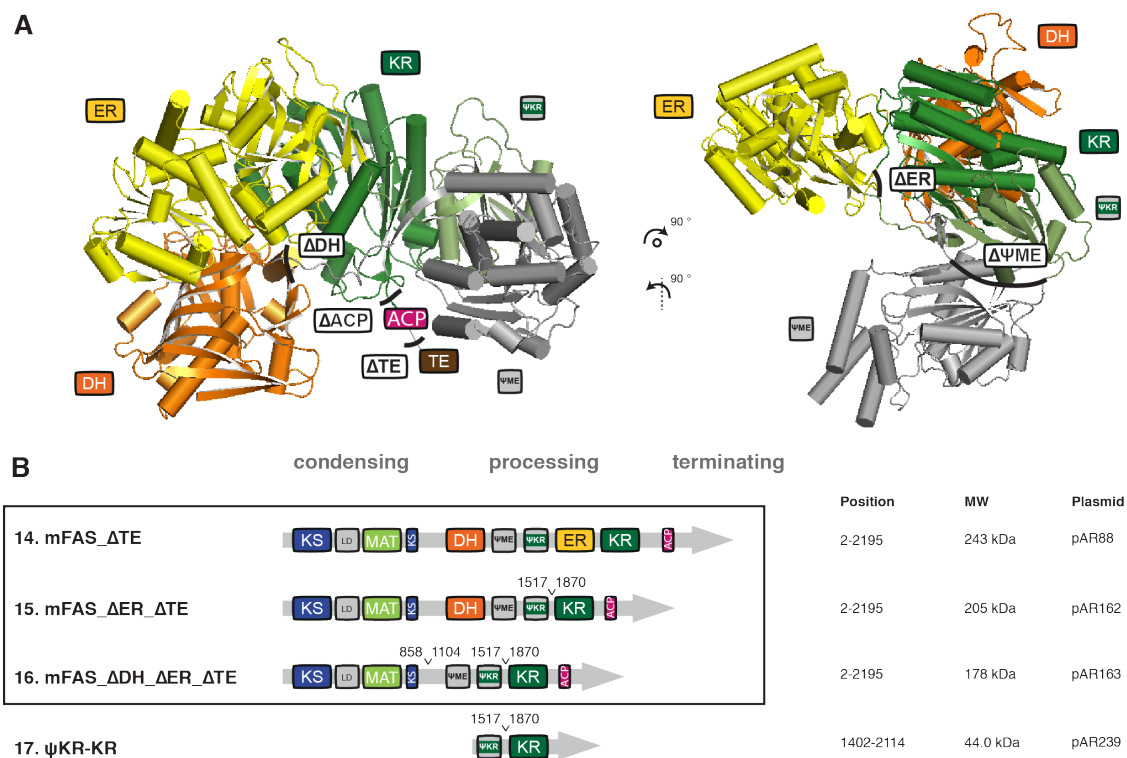
Figure 19: Engineering reducing, PKS-like modules. (A) Structure of the monomeric processing part (porcine FAS; pdb: 2vz9) in cartoon depiction.[31] The KR pays a central structural role in the architecture and serves as a dock for all other domains. Vertical black lines (left panel) indicate sites for the deletion of the corresponding domain and longitudinal black lines (right panel) represent designed linkers to delete ER and ψME. (B) Domain organization of designed constructs representing highly reducing (construct **14**; δ-module) and partially reducing (construct **15**; γ-module and construct **16**; β-module) modules of PKSs.

Construction of the δ-module (**14**; 2-2195) was straightforward, as only a deletion of the terminating TE domain was needed. Although this domain is not apparent in the porcine crystal structure, it has been crystallized as a separated protein and therefore the N-terminal boarder of this domain is well known.[91] Furthermore, TE is structurally independent only connected by a flexible linker to ACP. This linker was left in the designed construct to attach the C-terminal His-tag (protein yield: 2 mg/1 L culture).

Construct **14** was used as a template to generate the γ-module (**15**; 2-1517+1870-2195). From a protein engineering point of view, this construct was the most interesting one, as it required a deletion of the ER domain. ER is inserted in the primary sequence between ψKR and KR and contributes to the dimeric interface of mFAS implying a potentially problematic deletion. This was not the case. Residues P1517 and S1870 at the N- and C-terminus of ER appear in a long, accessible linker between the structural and functional KR and are only 5.7 Å (distance between Cα-atoms) apart from each other. They were simply linked resulting in **15** in good yields (2 mg/1 L culture).

The β-module was created by deleting the DH domain in construct **15**. This was accomplished by linking S858 with S1104 utilizing the natural linker of the waist region of mFAS plus the full DH to ψME linker. Construct **16** again expressed slightly better than the before designed modules with a typical yield of 2.5 mg per 1 L cell culture.

Having expressed and purified all three desired reducing PKS-like modules, we needed further insights into the integrity and functionality of the created folds. The size exclusion chromatogram (**Figure 20B**) revealed peak profiles of monodisperse protein and very little aggregation. Construct **14** looked similar to the whole mFAS (**Figure 15B**) with only a small shoulder for the monomeric fraction. With deletion of the ER domain the ratio between dimeric and monomeric protein changed towards a pronounced fraction of the monomeric state. Unfortunately, these constructs were not assembled by incubation at 37 °C and hence a quantification was not possible. Furthermore, stability of the created reducing modules was investigated by a thermal shift assay and melting temperatures were compared to mFAS. Interestingly, melting temperatures differed in only a few degrees (buffer conditions have a higher influence) implying maintained stability. Deletion of the TE domain even seemed to slightly increase stability of the mFAS fold.
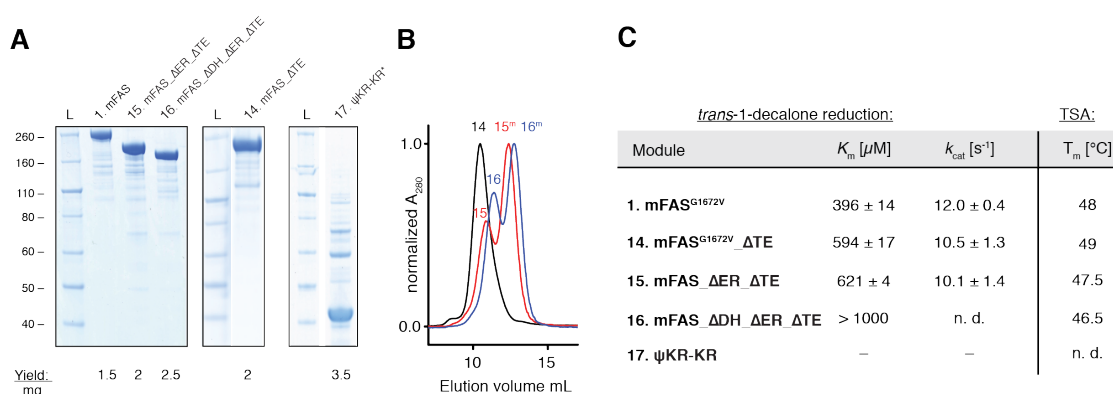
Figure 20: Integrity of PKS-like reducing modules. (A) Coomassie-stained SDS-PAGE of purified reducing modules (β-, γ- and δ-) with typical yields of purified protein from 1 L cell culture. Star at construct **17** indicates incomplete purification after a single Strep-Tactin affinity chromatography. (B) SEC of purified constructs **14**, **15** and **16** with absorption normalized to the highest peak. Samples **15** and **16** were not incubated at 37 °C before the run. Peaks correspond to apparent molecular weights of 584 kDa, 490 (237) and 378 (200) kDa, respectively (theoretical weights: **14**: 486 kDa, **15**: 410 kDa and **16**: 356 kDa). (C) Specific KR activity determined by a *trans*-1-decalone reduction assay and data from a thermal shift assay.[92-94] This data is taken from a master thesis by David Drexler under my supervision.[87] The ER domain was knocked out by G1672V. n.d. – not determined and – means no detectable activity. Errors refer to standard deviation from technical replicates (n = 4).

Besides providing data on protein's integrity, we intended to quantify the effect of our engineering. In terms of the given definition on modularity (section 2.2) the remaining domains have to retain their function. Hence, we decided to utilize the specific KR activity as a measure to evaluate our engineering, as this domain remains present in all reducing modules. An assay using *trans*-1-decalone to investigate specific KR activity was invented in the early 70s and we established this assay in our lab.[92, 93] The NADPH binding sites of the ER domain containing constructs were altered by mutation G1672V in the Rossmann fold.[52] Michaelis constants were determined for mFAS^{G1672V} to be 396 ± 14 µM with turnover rates of 12.0 ± 0.4 s^{-1}, reported values for FAS purified from pig liver were 150 µM and 0.73 s^{-1}, respectively.[92] Interestingly, deletion of the TE domain in construct **14** influenced K_m (594 ± 17 µM) and k_{cat} (10.5 ± 1.3 s^{-1}), although both domains are not directly connected nor a structural interaction of both folds is known. Furthermore, deletion of the vicinal ER domain (**15**) did not further affect (within error bars) activity of KR (K_m: 621 ± 4 µM and k_{cat} (10.1 ± 1.4 s^{-1})). Construct **16** was still active, but a deletion of the DH domain dramatically impaired KR activity. The thermal shift assay and the *trans*-1-decalone assay to investigate enzyme kinetics of reducing modules was performed by David Drexler under my supervision.[87]

Finally, we also tried to delete the ψME domain, which is a typical structural domain of FASs, that is normally not present in PKSs. Based on the crystal structure of porcine FAS, the residues S1132 and P1402 were identified and linked. The respective construct (pAR229) expressed very poorly, which indicated a close structural connection between the ψME domain and the ψKR domain. Accordingly, the ψKR/KR dimer (**17**) as a separate protein was not active anymore and following this strategy would require further improvements of the construct.

2.2.2 Connection of MAT and ACP Creating an α-Module

In order to generate a PKS-like α-module one has to connect the ACP to the condensing part via a flexible linker (**Figure 21**). This task is rather straightforward, as relatively independent domains are supposed to be fused, which reduces the problem to a matter of linker lengths. For guiding the engineering of initial constructs, we used module 3 and module 11 (both α-modules) of the avermectin synthase (**Figure 5**; AVES 2: Q9S0R7; AVES 4: Q79ZM6) as templates. The C-terminal ends of the post-AT linkers (anchor point for the desired linker) were identified by localising a conserved motif occurring in most PKSs (KS3$_{AVES}$: YPFQHHHYW; KS11$_{AVES}$: YAFERERFW).[95] The N-terminal ends of the ACP$_{AVES}$ domains were determined by homology to the ACP domain

of DEBS in module 2 (ACP2$_{DEBS}$), which has been structurally elucidated by nuclear magnetic resonance (pdb: 2ju2).[96] Accordingly, the linkers are 31 and 37 amino acids long in avermectin module 3 (M3L$_{AVES}$) and module 11 (M11L$_{AVES}$), respectively (for sequence, see **Figure 21B**). To acquire nucleotide sequences, we have ordered the genomic DNA from *S. avermitilis* and amplified linkers by PCR.
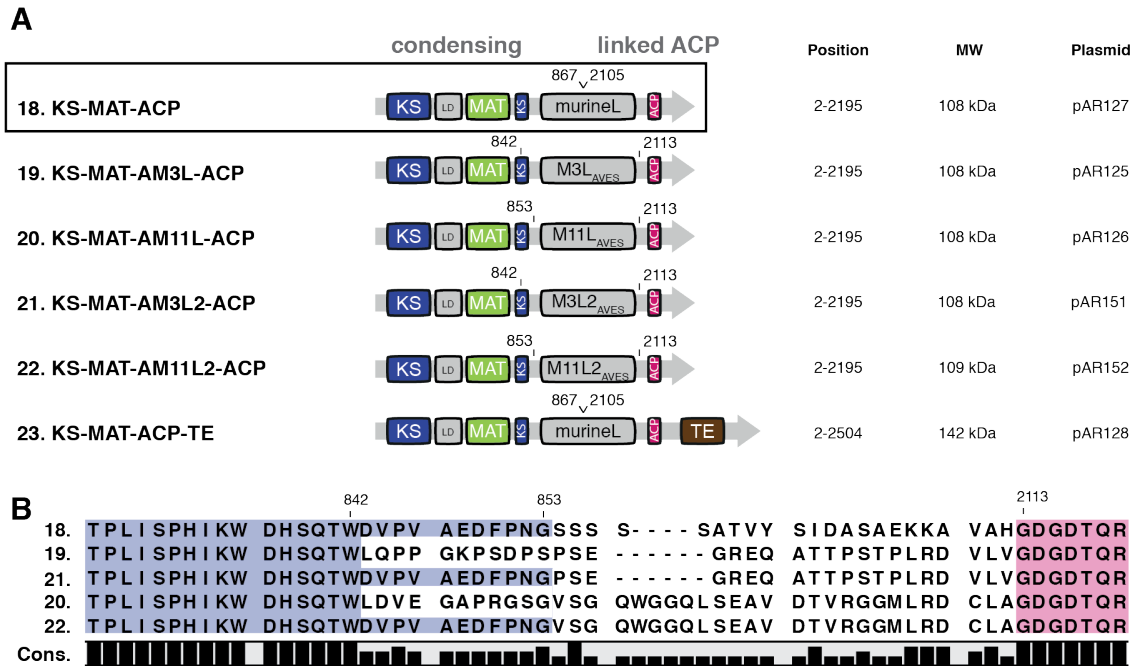
Figure 21: Creation of α-modules. (A) Domain organization of covalently linked ACP to KS-MAT constructs representing PKS-like α-modules. AVES served as a template and linkers from module 3 and 11 were tested in the mFAS derived construct. (B) Sequence alignment of linked constructs focusing on sequences of the linker region. Blue highlight indicates the murine KS domain (post-AT linker) and purple coloring shows the beginning of murine ACP.

Following the same approach, we identified domain boarders for murine KS and ACP domains. As ACPs from FASs differ in the length of helix 1 compared to ACPs from PKSs, the structural information of rat ACP (pdb: 2png) was used to identify the N-terminus of the domain.[97] For a first set of constructs, we determined the C-terminal end of the conserved motif in the post-AT linker as C-terminal end of the KS domain and hence anchor point for linkers (**Figure 21B**). Inserting linkers M3L$_{AVES}$ and M11L$_{AVES}$ created constructs **19** and **20**. Further, a construct was designed only utilizing sequence from mFAS. It was cloned like a deletion combining the first β-sheet of the DH domain with the natural KR-ACP linker to connect ACP (construct: **18**; linker length: 33 aa). These three constructs were analyzed biochemically (**Figure 22**). A co-expression of construct **18** with Sfp reduced the protein yield significantly by four-fold. Besides low

expression yields, constructs **19** and **20** suffered from decreased stability, low dimeric fraction of protein and no activity in a TAL assay (**Figure 22**), which was unexpected as implying an impact of linkers on these properties.
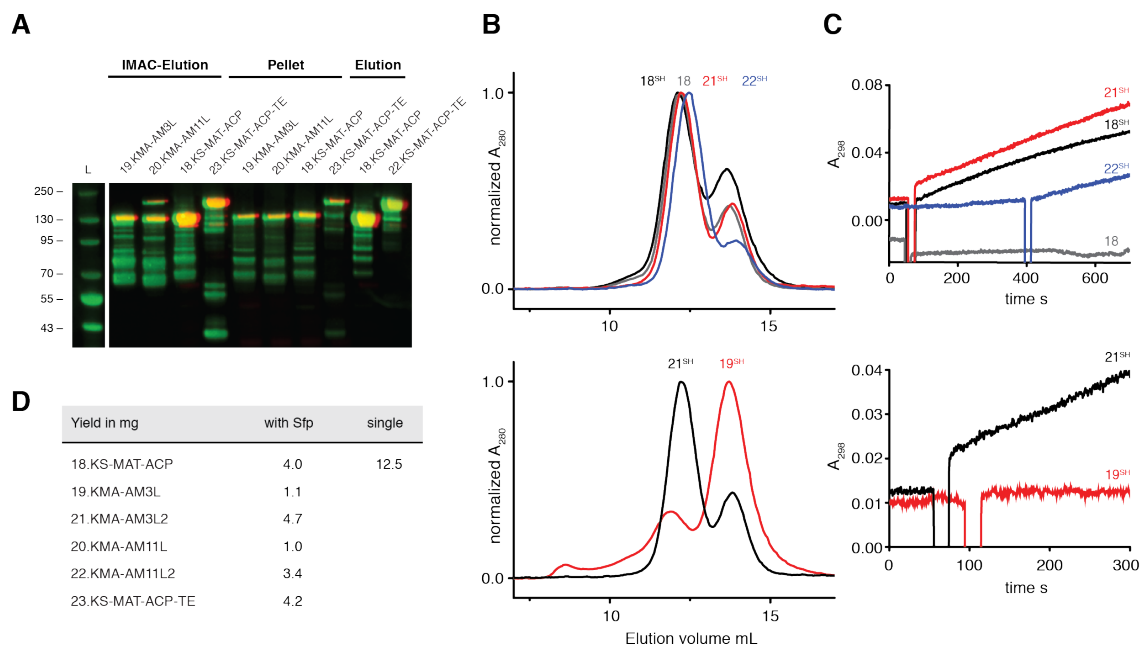
Figure 22: Biochemical analysis of PKS-like α-module. (A) Western blot of α-modules using fluorescently labeled antibodies. A merged picture is shown in false colors. Antibodies against both tags were used simultaneously, anti-Strep is colored in red and anti-His in green, respectively. (B) SEC of purified constructs **18**, **19, 21,** and **22** with absorption normalized to the highest peak. (C) Qualitative TAL production assay. HS in superscript indicates post-translational modified protein after co-expression with Sfp. (D) Yields of purified protein from 1 L cell culture in milligrams. Yield of **18** was three times higher, when not co-expressed with Sfp.

Based on the fact that both constructs with avermectin linkers performed weakly, we concluded that also the amino acids after the conserved motif of the post-AT linker (will be discussed in section 3.5.2; **Figure 48**) are from importance for the integrity of the KS domain. Exchanging the first 11 amino acids in constructs **19** and **20** to the natural ones created two new constructs **21** and **22 (Figure 21B)**. These behaved accordingly to construct **18** in terms of expressibility, running behaviour in SEC and activity in a TAL assay. Specific activity was further verified using apo**18** as negative control showing no activity in a TAL assay (**Figure 22C**).

Construct **18** turned out to be an important construct and was further used in other projects. Insights into linker design and implications of these findings on the polyketide synthase field are particularly discussed in section 3.5.2.

2.2.3 Generation of Bimodular Constructs Connecting a Loading Didomain to mFAS

In the last chapter, we have demonstrated a remarkable tolerance of the mFAS fold for structural alterations allowing the construction of PKS-like modules. Herein, the possibility to create constructs beyond one module shall be investigated testing the limits of *E. coli's* expression capacities. The simplest possible bimodular system would contain a loading module fused to the N-terminus of the FAS fold. By intending to separate the loading and elongation step of fatty acid synthesis, such constructs can give insight into the intermodular transfer of intermediates. Hence, we commenced with searching for loading modules, which can be fused N-terminally to the mFAS fold.
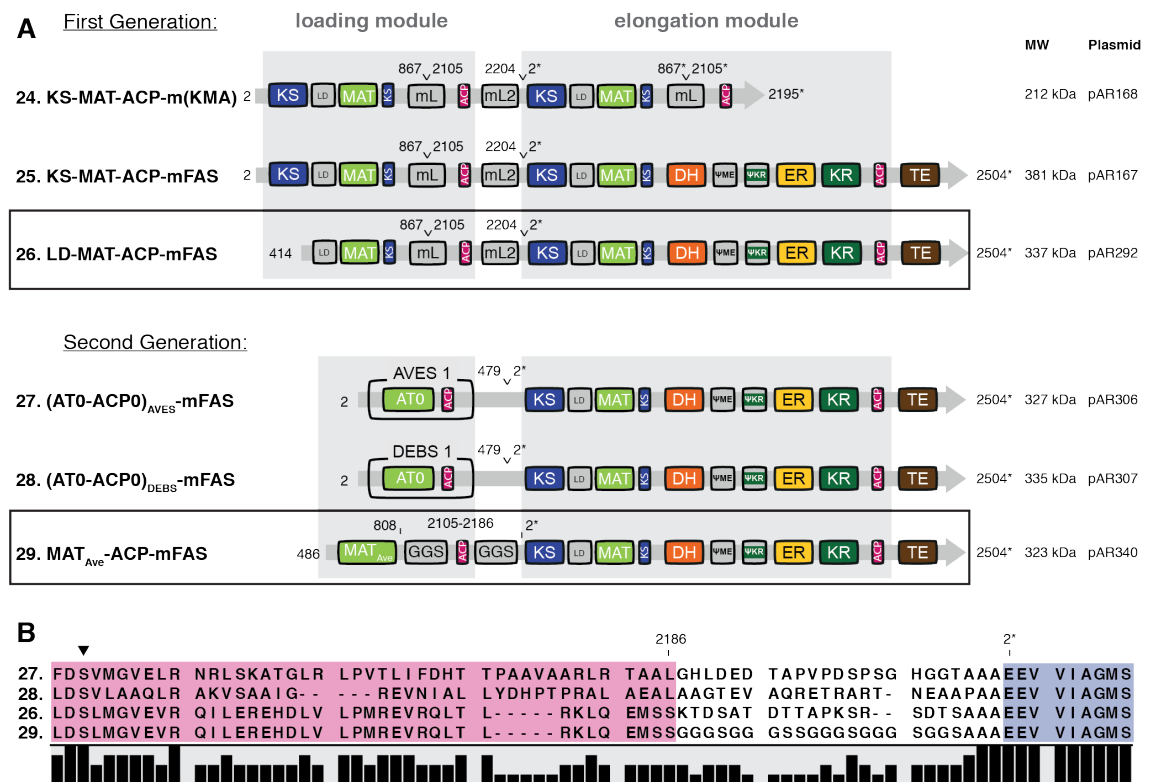
Figure 23: Bimodular constructs having a separated loading module. Domain organization of bimodular constructs. Loading and elongation of fatty acid synthesis is intended to take place in two different modules. mFAS is used as elongation module and different loading modules are fused N-terminally. Constructs **25**, **26** and **29** use parts from the previous chapters (**18**, **9+4** and **13**) exclusively using mFAS sequence. As the identical sequence occurs twice, the numbering of the second module again starts at 1 indicated by a star. Additionally, loading modules from AVES (AVES 1: Q9S0R8) and DEBS (DEBS 1: Q03131) were utilized. (B) Sequence alignment of select constructs showing sequences of the ACP to KS linker region. Blue highlight indicates the murine KS domain (conserved motif: EEVVI, compare KS1$_{DEBS}$: EPVAV) and purple colouring shows different attached ACPs. Small arrow indicates the site of phosphopantetheinylation.

The central question in such an approach is how to connect the C-terminus of a loading module, i.e. the ACP domain, to the KS domain of mFAS. Again, PKSs were exploited as templates, which show different architectures for loading modules. The simplest solution can be found in DEBS and AVES connecting a single loading didomain (AT and ACP) to the first module. These loading ATs were functionally investigated and it was found that they are specific for loading substrates, but posses a rather broad substrate spectrum.[24, 98] Another solution is provided for example by the Pikromycin synthase (PikA), which uses a modified α-module (KS, AT and ACP) with a decarboxylating KS domain and an AT domain specific for extender substrates.

Both opportunities were pursued, creating bimodular constructs shown in **Figure 23**. On the one hand, we have used constructs **18** and **9+4** from the previous chapter to create constructs **24**, **25** and **26**. These constructs exclusively contain amino acid sequences from mFAS. The linker between the C-terminus of ACP and the N-terminus of KS was built out of the natural ACP-TE linker with three alanine residues added C-terminally. The linker length resembles typical lengths from PKS modules. On the other hand, two actual loading didomains from AVES and DEBS were attached to the KS domain of mFAS utilizing their natural ACP-KS linkers (**27** and **28**). In both cases, the respective KS domain was exchanged using the conserved sequence at the N-terminus of the domain for orientation (**Figure 23**B). The nucleotide sequences were amplified from genomic DNA (AVES) or vector pBP144 (DEBS), the latter kindly provided by the Khosla lab. The last, probably most advanced construct **29** combines information from both approaches. MAT_{Ave} linked to murine ACP (**13**) was fused to the KS domain by an artificial GGS linker with the length of the $ACP0\text{-}KS1_{AVES}$ linker. This construct closely resembles typical design of respective PKSs.
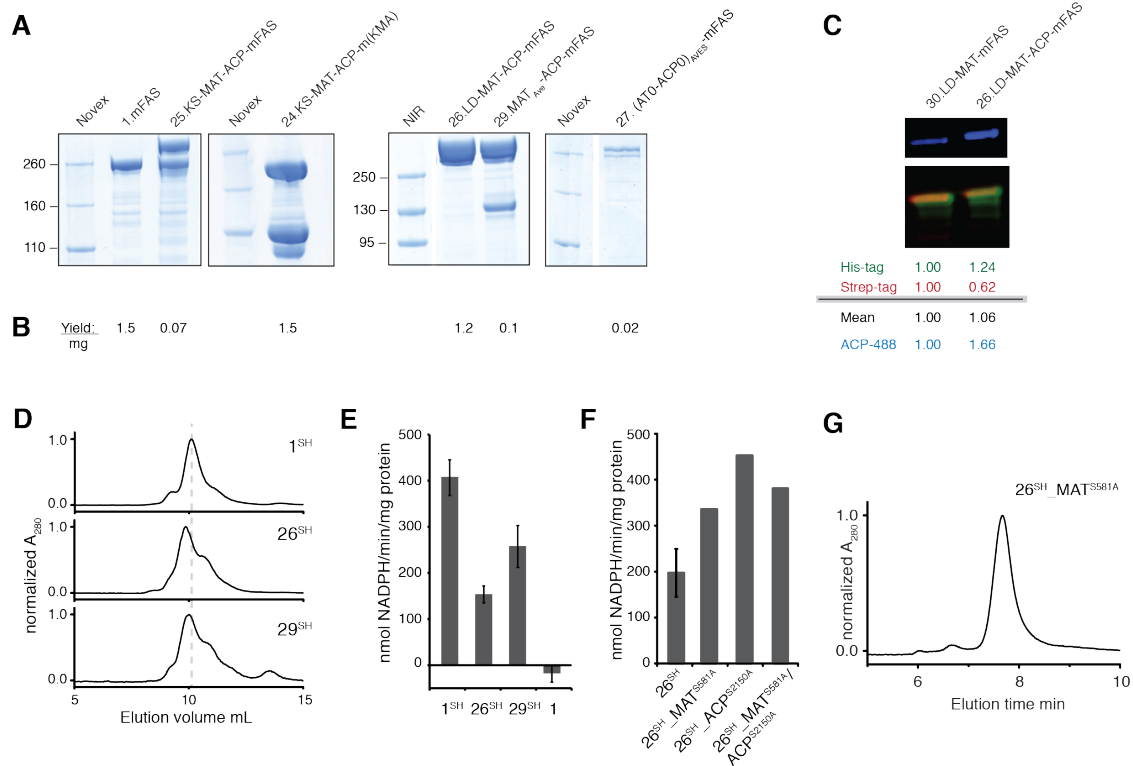
Figure 24: Integrity and functionality of bimodular constructs. (A) Coomassie-stained SDS-PAGE of purified bimodular constructs. Despite using tandem affinity purification, strong bands for truncated proteins remained in preparations of **24** and **25** implying the existence of heterodimers. (B) Typical yields of purified protein from 1 L cell culture in milligrams. Numbers refer to total protein in purified samples including truncations. (C) Accessibility of ACPs in bimodular construct **26**. Quantitative western blot after *in vitro* phosphopantetheinylation with CoA-488. A merged picture is shown in false colours. Antibodies against both tags were used simultaneously, anti-Strep is coloured in red and anti-His in green, respectively. Blue channel refers to fluorescently labelled ACPs. Normalized intensities for all channels are shown below using a control **30** (LD-MAT-mFAS) as reference. (D) SEC of purified (two affinity chromatographies) constructs after co-expression with Sfp with absorption normalized to the highest peak. Peaks (1^{SH}), (26^{SH}) and (29^{SH}) correspond to apparent molecular weights of 697 kDa, 784 kDa and 727 kDa, respectively. Peak shoulders indicate the presence of monomers in all samples, which are in temperature and buffer dependent equilibrium with the dimer. (E) FAS activity addressed by the specific oxidation of NADPH. The consumption of NADPH was monitored fluorometrically. FAS without Sfp co-expression shows no activity. (F) Control experiment for compromised activity of LD-MAT-ACP-mFAS (26^{SH}). NADPH consumption was determined for three knocked out constructs: MAT^{S581A}, ACP^{S2150A} and MAT^{S581A}/ACP^{S2150A} by measuring the decrease of absorbance at 340 nm. Error bar of the wild-type reflects standard deviation from four biological replicates. Data of controls were collected in technical triplets with a standard deviation of 33, 36 and 57 nmol/min/mg protein for MAT^{S581A}, ACP^{S2150A} and MAT^{S581A}/ACP^{S2150A}, respectively. (G) Analytical SEC of tested variants with HPLC to proof structural integrity of all samples (for other runs, see appendix **Figure S5**). Peaks correspond to a dimeric oligomeric state.

The designed bimodular constructs were tested for their expressibility in *E. coli* and subsequently purified (**Figure 24A**). Unfortunately, the constructs using natural PKS

loading didomains were not (DEBS: **28**) or hardly (AVES: **27**) expressible. Additionally, constructs **24** and **25** using the murine α-module as loading domain suffered from strong proteolysis and truncated proteins could not be separated despite utilizing the tandem affinity purification strategy. This is an indication for the occurrence of heterodimers and optimization of these constructs would need engineering of the site of proteolysis in the KS domain, which was not further pursued. The two constructs with murine didomains (MAT and ACP) **26** and **29** (for an overview, see **Figure S6**) were expressible and could be sufficiently purified, although the engineered MAT$_{Ave}$ domain reduced protein yields significantly. Nevertheless, both constructs could be analyzed by SEC and showed the characteristic mFAS chromatogram with slightly increased monomeric shoulders (**Figure 24D**). Furthermore, both constructs were functionally compared to mFAS in a NADPH consumption assay (**Figure 24E**). Construct **26** showed approximately 40 % (153 ± 18 nmol/min/mg) of the wild-type mFAS activity (406 ± 39 nmol/min/mg) and construct **29** 60 % (257 ± 45 nmol/min/mg), respectively. Both constructs were tested in an *in vitro* phosphopantetheinylation assay to confirm quantitative modification (data not shown) excluding this reason for reduced activity. As a control mFAS without phosphopantetheinylation did not have any activity.

Last but not least, we have tested the accessibility of both ACPs in construct **26** by fluorescently labelling the ACPs *in vitro* (**Figure 24C**). Compared to a control construct possessing only one ACP (**30**; pAR291), the signal was 1.66-fold of the control implying that the two apo-ACPs were accessible.

To evaluate whether the additional loading module interferes structurally or catalytically with fatty acid synthesis, we tested three controls of the LD-MAT-ACP-mFAS construct providing either loading module domain MAT or ACP or both as functional knockouts (**Figure 24F–G** and appendix **Figure S5**). All three controls regained activity of wild-type mFAS as read out by NADPH consumption rates. These data are interesting, since they imply the structural integrity of module 1 (the integral mFAS fold) in the bimodular constructs. The comprised activity of the non-mutated wild-type bimodular construct (no knock-out in MAT and ACP domains) further demonstrates a functional communication of the loading module with the integral mFAS fold. An influence of monomeric contaminations was ruled out by revisiting the used samples with analytical SEC with HPLC.

2.2.4 Quantifying the Amount of Particles of Select Constructs in the Cell

Being interested in a rough number of molecules per *E. coli* cell, we have performed a quantitative western blot on whole cells (4 mL expression cultures). This number includes all full-length translated polypeptide chains regardless of their folding state.
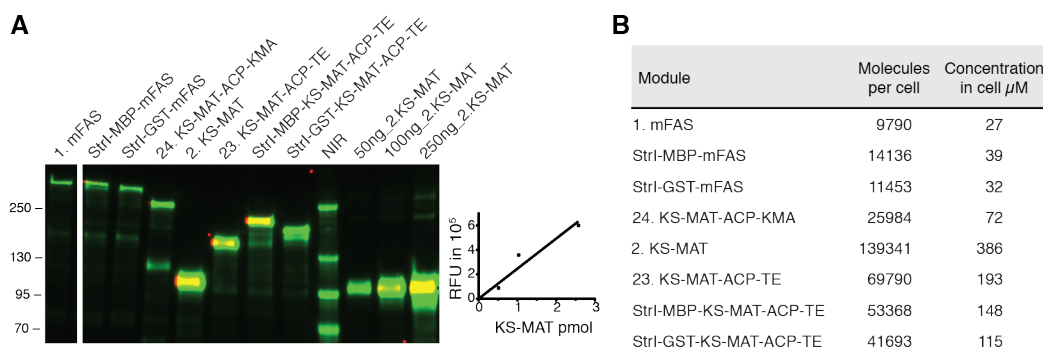
Figure 25: Rough estimate of molecules per cell for select constructs. (A) Western blot of select constructs using fluorescently labeled antibodies. A merged picture is shown in false colors. Antibodies against both tags were used simultaneously, anti-Strep is colored in red and anti-His in green, respectively. The anti-Strep signal is rather low likely due to the use of an old batch of antibody. A rough calibration was done using purified KS-MAT as standard. (B) Calculated number of molecules per cell and rough intracellular concentration in *E. coli*. Counts were converted into molar amount, multiplied by Avogadro's constant and divided by the rough amount of cells, assuming 8×10^8 cells in 1 mL medium with an optical density (OD$_{600}$) of 1. For the intracellular concentration, counts were converted into molar amount, divided by the rough amount of cells and divided by an averaged cell volume of 0.6 µm^3 (6×10^{-16} L).[99]

Approximately 8000 molecules of mFAS per *E. coli* cell existed after the chosen expression conditions (0.25 mM IPTG; 20 °C; ON). This number compares to a rough number of 6800–72000 ribosomes per cell depending on the doubling time of the *E. coli* culture.[100] As seen before (section 2.1.1) fusion proteins do not significantly increase the yield when N-terminally Strep-tagged. The condensing part expressed at 14-fold higher yields than the whole mFAS.

2.3   Functional Studies on the Condensing Part

The condensing part plays the central role in fatty acid and polyketide synthesis. It is responsible for selecting substrates (MAT/AT) and for coupling those by formation of a new C-C bond (KS). As outlined in the next chapter, structural features as well as substrate specificities of these domains are of decisive difference in FASs and PKSs. Specificity and (probably) efficiency of MAT are in contrast to typical ATs from PKSs, which transfer either loading or extender substrates.[38] Also the KS domain of FAS with its very narrow substrate spectrum and fast iterative catalysis differs from typical, more unspecific KS domains in PKSs. The development of FAS towards custom-compound biosynthesis would certainly require an adaptation of specificities of both domains. It is therefore of utmost importance to establish methods and protocols that can provide enzyme kinetic data to make specificity and efficiency of domains accessible. The following chapter describes the foundation of this task comprising with the establishment of good analytics for partial activities and in depth characterization of the wild-type.

All experiments determining KS or MAT kinetics were performed with the KS-MAT didomain with an appropriate knock out of the other domain unless otherwise stated. The respective mutations were C161G (pAR70) in the KS domain and S581A (pAR159) in the MAT domain.

2.3.1   Establishment of Analytics for the Transferase Function

Many methods to investigate the function of MAT are provided in literature: starting from monitoring overall fatty acid synthesis (section 2.1.2),[44] which provides indirect insight into the function of MAT, to methods for recording specific enzyme activity from quantifying thiols or free CoA by reagents or HPLC to using isotope labelling strategies.[24, 98, 101-103] As there is no standard technique defined in the field, we commenced this project by comparing different methods to find out, which fulfils our requirements best. Our focus was to conveniently determine specific enzyme kinetic constants of the individual domain that reflect native constants during fatty acid synthesis as well as possible. Such detailed information on kinetic properties of the domain, best for different substrates, have not been available in literature.

2.3.1.1.MAT Occupancy as a Measure for Domain's Specificity

At first, we have tested the usage of thiol quantifying agents. Taking reported K_m-values of MAT in a single digit micromolar range into account,[51] it was clear that sensitivity of Ellman's reagent was not sufficient to quantify thiols at such low concentration. This problem can be solved by using fluorometric reagents to quantify thiols, which are characterized by a much higher sensitivity. An example of such compounds is ThioGlo[TM], which has been used to investigate hydrolysis rates and occupancies of PikAIV, a module of the pikromycin synthase.[104] This reagent was not easily available and consequentially, we have replaced it by 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM), which has comparable properties and had been already used to analyze overall fatty acid synthesis.[105-107] The disadvantage of thiol quantifying agents is that no transthioesterification kinetics can be measured, as the number of free thiol groups does not change during the reaction. Further, these compounds are typically applied in end point assays, as the protein has to be precipitated to reduce unspecific background. Consequently, only occupancies of domains and hydrolysis rates can be measured in a discontinuous manner.

A protocol was established based on Smith's reports using 6 % perchloric acid to stop the reaction and precipitate the protein (see section 4.1.8).[108] Briefly, the protein of interest was incubated with an excess of CoA-ester, centrifuged and the supernatant was transferred to 96-well plates. The solution was neutralized with 1 M KOH and free CoA was quantified by reaction with CPM. To reach the detection limit, 2 µM of protein had to be used. The establishment of the assay was performed with FabD (pAR90), because of its stability and known high substrate specificity, which was confirmed by the assay (**Figure 26**). MAT (pAR70) was treated accordingly. Interestingly, even with high excess of substrates, a maximum occupancy of only 0.56 ± 0.12 (acetyl-CoA) and 0.57 ± 0.08 (malonyl-CoA) was reached (**Figure 26**). Only when incubated longer with acetyl-CoA, the signal for released free CoA exceeded the amount of protein implying hydrolysis of the acyl-enzyme state. Such substoichiometric occupancies were also reported before for MAT,[108] and ATs of PKSs.[109]
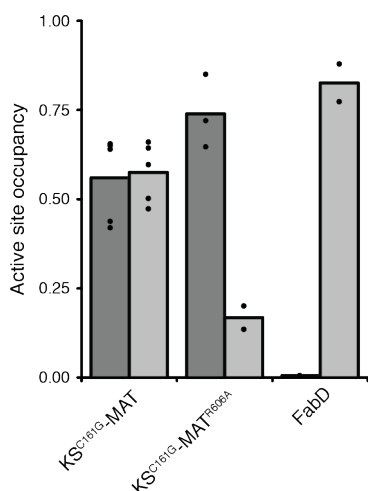
Figure 26: Occupancy of MAT and FabD with acetyl- and malonyl-CoA. The ATs were incubated with acetyl- (dark grey) and malonyl-CoA (light grey) and the released free CoA was quantified by CPM. Occupancy was calculated by calibration of the signal and division by the protein concentration. Every spot represents an independent measurement.

To verify the results and to evaluate if this method is sufficient to judge MAT variants, a known mutant was analyzed, which is supposed to have enhanced acetyl- and impaired malonyl-transfer (R606A).[51] Indeed, the mutant possessed altered occupancies with a level of acetylation increased to 0.74 and a level of malonylation decreased to 0.17, respectively. Also FabD showed the reported,[59] very strict specificity for malonyl-CoA.

This result seemed promising, but the method did not give insight into enzyme kinetics and therefore lacked important information necessary to evaluate mutants (K_m, k_{cat} and k_{cat}/K_m). Furthermore, the results for other, unknown variants were by far not that unambiguous and the procedure was practically very challenging as small changes in pH-values influence the fluorescence properties of CPM. Conditions and other precipitation methods were optimized within the scope of a master thesis by Khanh Vu Huu under my supervision.[110] Unfortunately, interpretation of the results remained challenging and we focused on establishing a more suited method.

2.3.1.2. Transfer of Acetyl Moieties from 4-Methylumbelliferyl Acetate

In an early study from Jacks *et al.* the hydrolytic activity of lipases was investigated with coumarin esters,[90] and their usage has already been expanded to study the human TE I.[111] Inspired by this approach, we wanted to test whether this strategy can also be applied to investigate the MAT domain, which shares some common features with both other enzyme classes. The method is based on the fact that esters of 4-methylumbelliferone do not fluoresce unless cleaved to release the fluorophore. Consequently, we synthetically esterified 4-methylumbelliferone with acetyl (**C1**) and malonyl moieties (**C2**) (for synthesis, see section 4.2). Chemical structures and emission spectra of three coumarin derivatives are shown in **Figure 27**.
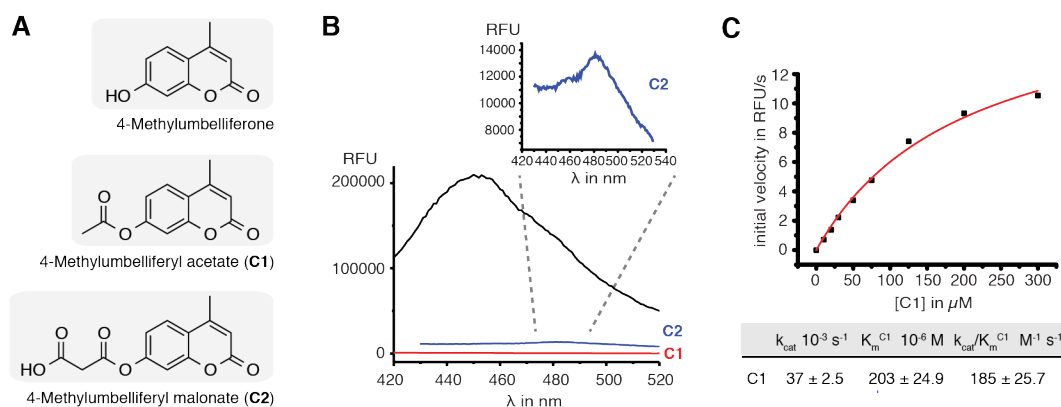
Figure 27: Coumarin derivatives and their usage to study MAT. (A) Chemical structures of 4-methylumbelliferone and respective synthesized esters. (B) Emission spectra of coumarin derivatives in ethanol. Compounds were excited at 360 ± 15 nm. Inlet shows the magnified spectrum of **C2** with a shifted maximum to a higher wavelength. (C) Plot of the reaction velocities as a function of **C1** (Michaelis-Menten plot) for MAT. Kinetic constants were determined using the Michaelis-Menten fit function with the software Origin. Due to limitation of the solubility of **C1** in water, the appropriate concentration range (conc. ≥ 5 K_m) could not be fulfilled meaning the kinetic constants have to be considered with reasonable care. Shown errors reflect the square from fit function.

As expected, compound **C1** shows no fluorescence, when excited with UV-light, but it cannot be stored in aqueous buffers as non-enzyme mediated hydrolysis takes place. Interestingly, **C2** (not reported in literature) retains residual fluorescence with an emission maximum shifted to 480 nm. The acceptance of MAT for both compounds serving as acyl-donors was tested. In contrast to **C2** (not accepted), **C1** can indeed be used to load the MAT domain. Given this fact, **C1** was used to examine enzyme kinetics and kinetic constants were determined. The turnover rate was relatively slow (k_{cat} = 37 ± 0.25 × 10^{-3} s^{-1}) with a Michaelis constant of 203 ± 0.25 µM. Both constants have to be considered with reasonable care due to a limitation in solubility of **C1** in water. Thus, the highest concentration of **C1** is limited to 300 µM and the real K_m and k_{cat} are probably higher. Nevertheless, signal to noise ratio was fantastic adding this compound to our toolbox for investigation of the MAT domain. Select variants were analyzed by Khanh Vu Huu and respective kinetic constants were determined (master thesis).[110] Interpretation of the results is very challenging as this compound class is structurally very distinct from the native substrates. However, **C1** is very cheap and experimental application is comparatively simple making this assay a useful alternative to judge protein integrity by comparing relative values.

Finally, also the corresponding thioester derivatives were synthesized and tested. Besides the fact that MAT did not accept these as acyl donors, the acetyl ester also

possessed a higher fluorescence than the hydrolysed compound 4-methyl-7-thioumbelliferone.

2.3.1.3.　　Enzyme-coupled HS-CoA Quantification

A very convenient way to visualize the progress of a reaction is to pair the reaction of interest with a second enzymatic reaction that can be followed by absorption or fluorescence.[58] Typically, the product of the first reaction is consumed as substrate for the coupled enzymatic reaction. Such coupled enzymatic systems often facilitate common, well-characterized enzymes from glycolysis or the citric acid cycle. Using an enzyme-coupled reaction has major advantages, as it offers a continuous readout and the product of the reaction of interest is withdrawn from the equilibrium making the interpretation of kinetic constants much easier.

An assay was developed in the histone acetylation field quantifying the released free CoA after reaction of acetyl-CoA with lysines of histones with the enzyme α-ketoglutarate dehydrogenase (αKGDH).[112] This enzyme decarboxylates α-ketoglutarate to an enzyme bound succinyl moiety, which is transferred to HS-CoA by transthioesterification creating succinyl-CoA. The enzyme is regenerated by electron transfer via two steps to nicotinamide adenine dinucleotide (NAD⁺) generating NADH, which can be detected fluorometrically. This assay was adapted by Molnos *et al.* to study FabD (AT from bacterial type II fatty acid synthase) and further optimized by Dunn *et al.* to investigate AT3 from DEBS.[69, 101] In both cases the enzyme mediated transfer from a CoA-ester bound moiety to ACP was measured, which requires high amounts of purified, independent ACP, which is entirely phophopantetheinylated.

2.3.1.4.　　Optimization of ACP Expression and Purification

As before mentioned, ACP is needed as an independent domain to investigate MAT kinetics with a αKGDH coupled assay. Hence, we designed a construct with a freestanding, separated ACP domain based on the structure of rat ACP (pdb: 2png). The first one (pAR75) contained a N-terminal StrepII-tag, which turned out to be insufficient for purification via Strep-Tactin columns. The construct was recloned (pAR100) to contain an additional C-terminal His-tag, which has been used ever since. After IMAC ACPs were further purified by SEC to provide the protein in good purity and yield (10 mg/1 L co-expression culture after SEC).

Although vast amounts of ACP are needed to perform detailed kinetic analysis of MAT, it turned out that more importantly a high, reproducible quality of ACP has to be guaranteed to ensure correct results on kinetic constants. ACP is hardly or not modified

by *E. coli*, why completeness of phosphopantetheinylation becomes a matter of co-expression strategies with Sfp. We tried several approaches with co-expression vectors (Sfp on pETcoco-1 (pAR01H) and on pCDF-1b (pAR357)), with a bicistronic organization (pAR352; Sfp gene behind ACP gene) and with expression in BAP1 cells (Sfp is inserted in the genome),[113] at different temperatures and analyzed the purified protein by HPLC-MS (**Figure 28**). Expression in BAP1 cells was not sufficient to provide entirely modified ACP (peak at 11991;**Figure 28D** and E), neither expressed at 20 °C or 37 °C over night. Also co-expression with Sfp from the pETcoco vector only resulted in fully activated protein when expressed at 37 °C over night (**Figure 28F**). This had been the method of choice, but a second peak in mass spectrometry at 12244 indicated a fraction of truncated ACP without the N-terminal serine (*N*-formylmethionine is always cleaved). To achieve more homogenous preparations, pAR100 was recloned and a translation cassette containing Sfp was installed 3' of the *ACP* gene. This construct performed best resulting in fully activated full-length protein, when expressed at 20 °C (**Figure 28B**). Nevertheless, two other peaks appeared in mass analysis at 12373 and 12509, which could not be avoided. These peaks most likely correspond to acetyl-ACP (modified by acetyl-CoA instead of free CoA in the cell) and a modification distal to the active thiol group (see succinylation of 12509, **Figure 29E**). A mass shift of 178 Da could be explained by phosphogluconoylation of a His-tag in *E. coli*.[114] Routinely, 12 L expression cultures of the bicistronic vector were prepared in one batch and combined purified.
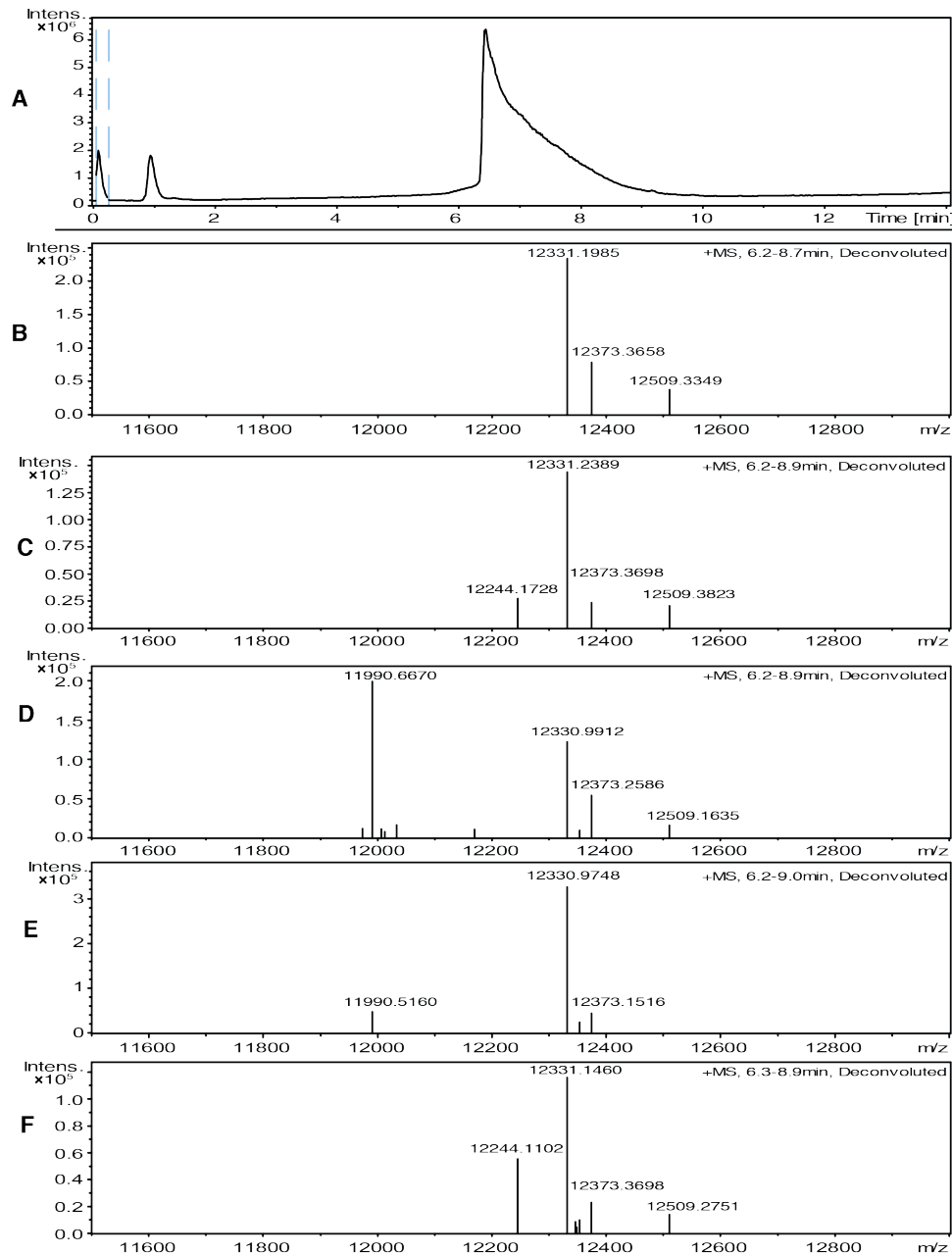
Figure 28: HPLC-MS analysis of murine ACP. (A) Representative reverse phase HPLC run of purified murine ACP (expressed from bicistronic construct at 20 °C). The protein eluted between 6 and 9 min. (B)-(F) Mass analysis (positive mode) of ACP co-expressed with Sfp under various conditions: bicistronically, at 20 °C (B) and 37 °C (C) over night; in BAP1 cells at 20 °C (D) and 37 °C (E) over night and with pETcoco at 37 °C (F) over night. Primary peaks correspond to the post-translationally modified protein without the N-terminal formyl-methionine (theoretical mass: 12330.80) or apo-ACP without the N-terminal formyl-methionine (theoretical mass: 11990.45 Da). Minor peaks correspond to the ACP after cleavage of the second N-terminal residue (serine) (theoretical mass: 12243.71 Da) or an acetylated ACP (theoretical mass: 12372.85 Da). Peak at 12509 Da corresponds to an unknown modification of the main fraction ACP (12331 Da), which was tested to exclude a modification of the active thiol. A mass shift of 178 Da could be explained by phosphogluconoylation of a His-tag in *E. coli.* [114]

2.3.1.5.   Setting up the Fluorometric αKGDH Assay in Our Lab

Having well-established proteocols for expression and purification of the KS-MAT didomain as well as the separated ACP domain, we comprised to set up the αKGDH assay in our lab. This could be conveniently executed after we had acquired our own microplate reader, which made it possible to carefully adjust all paramters. The αKGDH enzyme is extracted from pig liver and can be purchased from Sigma. Pure NADH was used to set up the method at the instrument and subsequently for calibration (section 4.1.10; **Figure 29A**). Basically, we followed the protocol of Molnos *et al*. and Dunn *et al*.,[69, 101] preparing four solutions containing: MAT in a BSA (for stabilization) buffer, the αKGDH assay buffer, CoA-ester substrates in concentrations from 0.2-5 K_m and ACP in concentrations from 50-300 µM. Solution 4 (ACP) was loaded into the dispenser and was used to start the reaction. The enzyme concentration (0.4-4 nM) was adjusted in a way that all curves from different substrate concentrations remained linear in the chosen time frame (**Figure 29B**).[58]
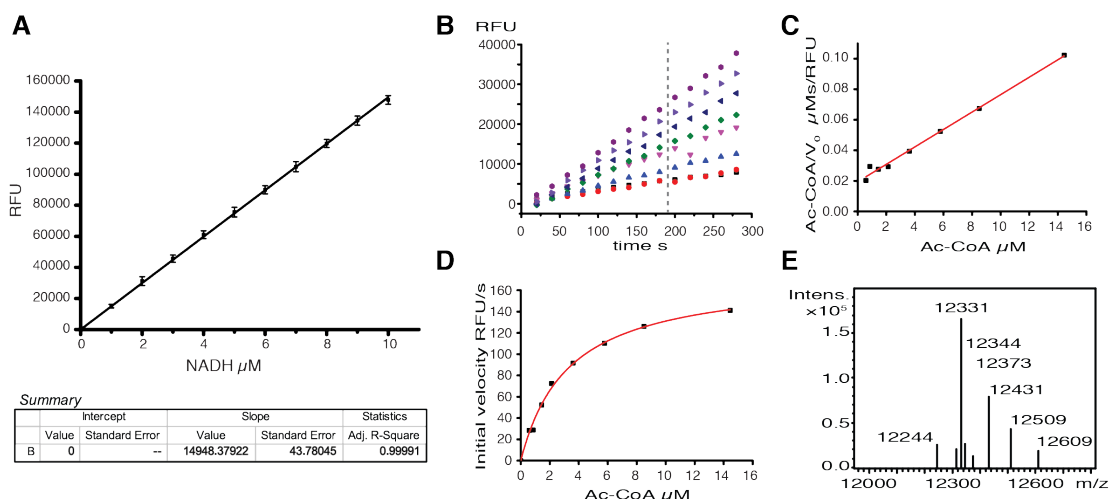
Figure 29: Establishment of the αKGDH assay. (A) Calibration of αKGDH assay with NADH. Calibration was performed to convert relative fluorescence units (RFU) into concentrations (µM). Error bars represent the standard deviation between three independent measurements. (B) Progress curves for MAT-mediated acetyl-transfer to ACP (100 µM). Starting concentrations of acetyl-CoA were 0.6, 0.9, 1.4, 2.1, 3.6, 5.8, 8.5 and 14.4 µM (C) Corresponding plot of the acetyl-CoA concentrations divided by the reaction velocities as a function of the acetyl-CoA concentrations (Hanes-Woolf fit; section 5.2.3) (D) Corresponding plot of the reaction velocities as a function of the acetyl-CoA concentrations (Michaelis-Menten fit) (E) HPLC-MS analysis of succinyl-ACP. Mass analysis (positive mode) of ACP (co-expressed at 37 °C, compare **Figure 28F**) after MAT mediated succinyl-transfer. Shifts of 100 mass units can be detected for 12244, 12331 and 12509.

By carefully evaluating the progress curve for one substrate concentration over a longer period, our attention was caught as the detected curve remained linear even when more substrate was transferred than CoA-ester had been added. This phenomenon could only be explained by the fact that also succinyl-CoA, the product of the coupled enzyme system, was transferred by MAT to ACP. This was unexpected as nothing comparable had been reported before. Hence, we analyzed ACP after MAT-catalyzed succinyl-transfer by HPLC-MS and indeed, the corresponding mass shifts of 100 were found (**Figure 29C**). Based on this finding, the assay was set up in a way that data was only recorded to a consumption of substrates up to 20 %. Relatively low K_m at µM concentrations made it necessary to extend ranges at low substrate concentrations up to a turnover of 50 %. This relatively high cut off was chosen to ensure a good signal to noise ratio for low CoA-ester concentrations. Furthermore, data points were only used until 180 s due to fast succinyl-CoA hydrolysis generating an increasing background.

Data analysis was performed with Microsoft Excel, because it allows convenient copying of data and therefore made individual evaluation of many data sets possible. An example for a typical analysis is presented in **Figure S10**. Michaelis constants and rate constants were determined with the Hanes-Woolf method (section 5.2; equation [10])(**Figure 29C**). This linear regression method is known for its even weighting of data points, and turned out to well represent data in our case.[58] Data quality was checked by also plotting via the Lineweaver-Burk- (equation [9]) and Eadie-Hofstee-method (equation [11]) and non-linear regression (**Figure 29D**).

The αKGDH assay provided very high quality data and facilitated determination of kinetic constants. It served as the method of choice to analyze the wild-type MAT (section 2.3.2), variants of MAT (section 2.5) and KS domains (section 2.3.4).

2.3.2 Wild-type MAT Kinetics

The following section examines transfer kinetics of the wild-type MAT using different CoA-esters addressing the question of polyspecificity. This part of the thesis is taken from a manuscript written with the help of Karthik Paithankar and Martin Grininger.[115] All kinetic investigations of MAT were performed with the αKGDH assay using the KS mutant C161G (pAR70).
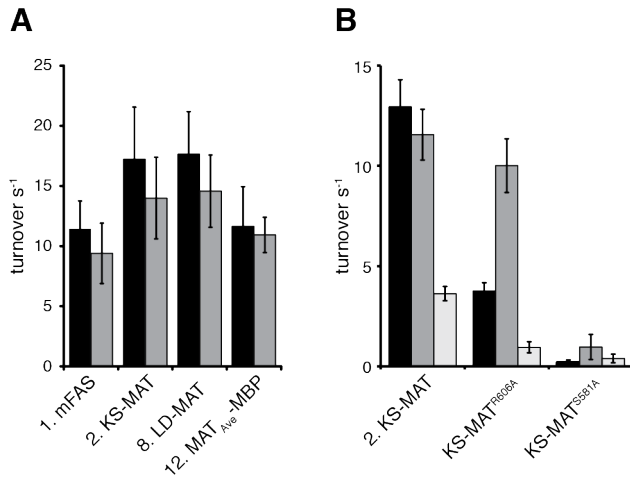
Figure 30: Partial activity of MAT in the αKGDH-assay. (A) Effect of truncation on MAT activity. (B) Verification of αKGDH assay with select variants. All constructs were investigated at fixed substrate concentrations of 20 µM CoA-ester and 60 µM ACP. Malonyl-, acetyl- and palmitoyl-transfer is shown as black, grey and light grey bars, respectively. All data collected in technical triplets. Error reflects standard deviation from four biological replicates (n = 4).

The MAT-mediated substrate hydrolysis rate was examined to enable appropriate background subtraction for transacylation (section 1.4.1, **Figure 8**). The hydrolysis rate was extremely low with turnover rates of $9.8 \pm 1.7 \times 10^{-3}$ s^{-1} and $9.3 \pm 0.8 \times 10^{-3}$ s^{-1} for malonyl- and acetyl-CoA, respectively. The rate determined for acetyl-CoA hydrolysis agreed well to the value previously reported for rat FAS (4.1 ± 0.8 nmol/min/mg protein (3.2×10^{-3} s^{-1})),[108] was about 3–4 magnitudes slower than the transacylation reaction and was hence ignored in background correction. Depending on the used CoA-ester, the non-enzymatic catalyzed hydrolysis or the self-acylation rate of mouse ACP was significant,[116] and had to be subtracted as background (see methods).

We investigated the impact of truncations of the full length FAS onto MAT activity (section 2.2; **Figure 18**). The apparent turnover rate for transacylation at fixed CoA-ester and ACP concentrations was determined. All constructs showed a slightly faster turnover for malonyl than for acetyl transfers (**Figure 30A**), whereas a faster acetyl-transfer was previously reported for the rat MAT.[51] Interestingly, both the KS-MAT didomain and the LD-MAT construct possess rate constants, which are about 50 % higher than for the full length or MAT$_{Ave}$-MBP construct. This phenomenon might be caused by competition of the interaction of ACP and MAT with the intrinsic ACP of FAS or the MBP domain of the MAT$_{Ave}$-MBP fusion. The mutant protein S581A was inactive, confirming specificity of MAT-mediated transacylation (**Figure 30B**).
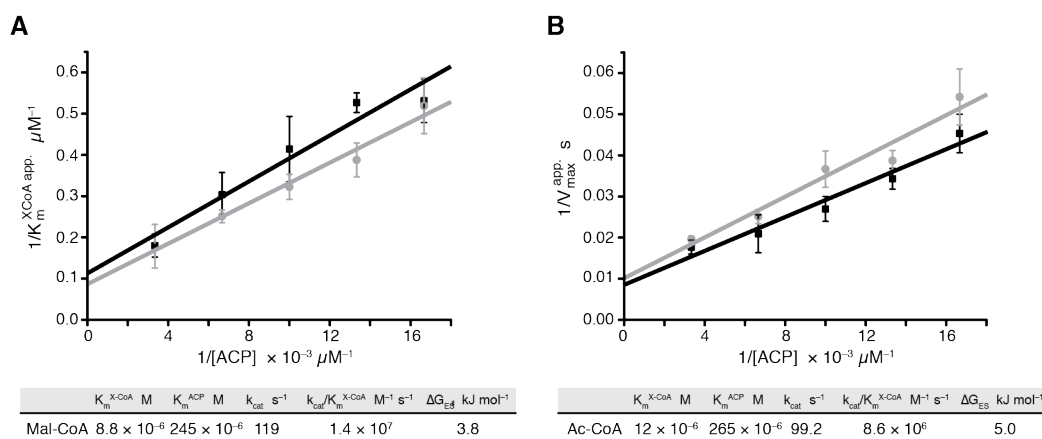
Figure 31: Acetyl- and malonyl-transacylation kinetics of the MAT domain. Double-reciprocal plots of the Michaelis-Menten constants K_m (A) and the initial rates k_{cat} (B) are presented as functions of reciprocal ACP concentrations. Data collected on malonyl-CoA are shown in black, acetyl-CoA in grey, and linear fits are accordingly. Error bars reflect the standard deviation between three biological replicates (n = 3). A biological replicate was defined by three independently expressed MAT preparations and different batches of purified, high quality ACP (section 2.3.1.4), mainly from bicistronic expression with Sfp. For establishing the fidelity of the assay, every measurement was performed in technical triplicates. Tables underneath graphs summarize absolute kinetic constants calculated from the double-reciprocal plot. Michaelis-Menten constants for ACP were calculated from both plots and averaged (245 µM and 245 µM for malonyl-transfer; 284 µM and 246 µM for acetyl-transfer).

Exploiting the ping-pong bi-bi mechanism of MAT,[46] (section 1.4.1) allowed addressing absolute steady-state kinetic parameters for malonyl- and acetyl-transfer by double reciprocal fits (**Figure 31**). Apparent K_m and k_{cat} constants were determined at five different ACP concentrations by fitting the raw data with the Hanes-Woolf equation (equation [10]; **Figure S10**). Individual reciprocal plotting of these constants against the reciprocal ACP concentration resulted in the absolute K_m of 8.8 µM and the absolute k_{cat} of 119 s^{-1} for malonyl-CoA, and 12 µM (K_m) and 99.2 s^{-1} (k_{cat}) for acetyl-CoA. Furthermore the K_m for ACP was calculated to 245 µM for malonyl-CoA and to 265 µM for acetyl-CoA. Both values agreed well indicating good quality of the experimental data. A K_m of 25 µM determined for the ecACP to FabD interaction reveals a one order of magnitude lower affinity for the type I system.[101]

2.3.3 Transfer of Noncanonical CoA-esters

We further aimed at elucidating the role of MAT in the reported broad acceptance for unusual CoA-esters of the purified animal FAS.[47, 117] With the αKGDH-assay as a reliable tool in hand, we determined apparent steady-state kinetic parameters at a fixed ACP concentration of 60 µM for ten additional CoA-esters (**Table 1**). Remarkably, while

being most specific for malonyl-CoA, the MAT domain was capable of transferring all tested CoA-esters, except palmitoyl-CoA, with specificity constants in the same order of magnitude. As a second native extender substrate of animal FAS, MAT reveals an approximately two-fold lower efficiency for methylmalonyl-CoA than for malonyl-CoA. Although malonyl-CoA is transferred four-fold faster and hence the transition state is better stabilized, this is compensated by a higher binding affinity to methylmalonyl-CoA. This finding may explain that under certain conditions, i.e. when substrate levels are sufficiently high in the cell, methylmalonyl-CoA can be used for fatty acid synthesis.[37]

Table 1: Kinetic analysis of the transacylation reaction with unusual CoA-esters at a fixed acceptor concentration of 60 µM ACP

Substrate	K_m^{app} (µM)	k_{cat}^{app} (s^{-1})	k_{cat}/K_m ($10^6 \times$ M^{-1} s^{-1})	Hydrolysis rate ($10^{-3} \times$ s^{-1})
Malonyl-CoA	1.28 ± 0.13	15.6 ± 1.4	12.2 ± 1.7	9.8 ± 1.7
Methylmalonyl-CoA	0.62 ± 0.07	4 ± 0.8	6.5 ± 1.4	6 ± 1.2
Acetyl-CoA	1.63 ± 0.1	13.7 ± 2.1	8.4 ± 1.4	9.3 ± 0.8
Butyryl-CoA	1.64 ± 0.16	10.8 ± 0.3	6.6 ± 0.7	8.5 ± 1.7
Octanoyl-CoA*	0.73 ± 0.28	4.1 ± 0.3	5.6 ± 2.2	6.1 ± 1
Palmitoyl-CoA**	4.8 ± 1.1	2.5 ± 0.5	0.5 ± 0.2	16 ± 1.4
Acetoacetyl-CoA*	0.86 ± 0.03	7.3 ± 0.4	8.5 ± 0.5	13 ± 2.5
Hydroxybutyryl-CoA	1.72 ± 0.39	6.1 ± 1.1	3.6 ± 1	20.4 ± 2.7
Crotonyl-CoA	1.36 ± 0.13	6 ± 1.1	4.4 ± 0.9	24 ± 2
Phenylacetyl-CoA*	0.96 ± 0.1	6 ± 0.2	6.2 ± 0.7	21.3 ± 2.3
Methylbutyryl-CoA	2.06 ± 0.36	7.2 ± 0.9	3.5 ± 0.7	10.4 ± 1.3
Succinyl-CoA	2.1 ± 0.17	13.4 ± 0.6	6.4 ± 0.6	28.7 ± 2.7

All parameters were determined in technical and biological (independent expressed and purified MAT preparations) triplets.

* and ** indicate that 2 nM and 4 nM MAT were used in the assay, because of low turnover rates.

Subsequently, we focused on the effect of the carbon chain length on substrate loading. Acetyl-CoA is the preferred substrate judged on the basis of k_{cat} and on the specificity constant k_{cat}/K_m. Interestingly, also longer chains as octanoyl and even palmitoyl can be transferred, although palmitoyl showed relatively poor transacylation kinetics. As differing from earlier data collected on the rat FAS MAT,[51] we confirmed the specific palmitoyl-transfer with the S581A mutant (**Figure 30B**).

The intermediates of the first condensation and modification cycle of fatty acid synthesis were also tested. Interestingly, acetoacetyl-CoA, the condensation product of acetyl and malonyl, turned out to be accepted as a substrate by MAT with similar transfer efficiency as acetyl-CoA. Hydroxybutyryl- and crotonyl-CoA were less favored and possess very similar kinetic constants.

Expanding the MAT transacylation to a spectrum of unusual loading substrates, phenylacetyl-CoA and methylbutyryl-CoA were tested,[47] both being accepted with specificity constants of $6.2 \pm 0.7 \times 10^6$ M^{-1} s^{-1} and $3.5 \pm 0.7 \times 10^6$ M^{-1} s^{-1}, respectively. Interestingly, with this value, MAT is 4-5 orders of magnitude more efficient than AT0$_{AVES}$ (reported 29.6 ± 3.8 mM^{-1} min^{-1}),[24] although in that study a N-acetylcysteamine thioester was used. Finally succinyl-CoA, a C4-dicarbonic acid derived thioester well known for its central role in the citric acid cycle, was tested and is well accepted by MAT. Apparent K_m and k_{cat} were determined to be 2.1 ± 0.17 µM and 13.4 ± 0.6 s^{-1}.

The transfer of succinyl-CoA, being the product of the coupled enzyme αKGDH, resulted in several *in vitro* cycles of transferring substrates to ACP (section 2.3.1.5). Unexpected high amounts of generated NADH were measured resulting in changed transfer kinetics after the substrate was consumed, rather than reaching an equilibrium.

To ensure that rates for non-native CoA-esters originate from transacylation and not from a possible increased hydrolysis rate, the MAT-mediated hydrolysis rate was also determined for every substrate. Although many non-native CoA-esters, for example, crotonyl- and succinyl-CoA, were hydrolyzed up to three times faster than acetyl- or malonyl-CoA, the apparent k_{cat} of the transacylation reaction remained 10^2–10^3 orders of magnitude faster. Hence, we can rule out a "proof reading" function of MAT by hydrolysis, which will be discussed in detail in section 3.2.1.[24, 118]

2.3.4 Expanding the Usage of the αKGDH-Assay to Study KS Kinetics

The first step in KS-catalyzed Claisen condensation is the transacylation of an acyl-moiety from acetyl-ACP to the active site cysteine. Considering this similarity to the mechanism of MAT and exploiting our gathered experience with such reactions, we aimed at expanding the usage of the αKGDH-assay to also investigate KS kinetics. In respect of the MAT transacylation assay, we wanted to investigate the ability of KS to transfer an acyl-chain from a CoA-ester to ACP with subsequent quantification of the released free CoA. Transacylation kinetics of the rat KS domain had been studied before by investigating the transfer from CoA-esters to pantetheine or N-acetylcysteamine and vice versa.[49] Therein, the quantification of acyl-products was achieved by HPLC-MS.

Utilization of the continuous coupled enzyme assay would offer a convenient way to investigate the KS domain allowing an in-depth kinetic characterization of the domain.

In contrast to MAT and typical KS domains from PKSs, KS from FAS is supposed to have a rather narrow substrate spectrum (section 1.5), which should already be apparent in the transthioesterification step according to Witkowski *et al.*.[49] They further reported a Gaussian distribution of turnover rates in respect of the chain length of saturated acyl intermediates with a maximum at eight carbon atoms. Similarly to functional studies on the MAT domain, we have used the KS-MAT[S581A] didomain with a MAT knockout (pAR159) for analyzing kinetics of the KS domain. Initial parameters were optimized by comparison with the double kockout KS[C161G]-MAT[S581A] (pAR160), which was conducted together with Aaron Himmler. We could demonstrate a KS-mediated transacylation of octanoyl moieties from CoA-esters to ACP and the expected specificity of the KS domain was qualitatively confirmed by comparing to the transacylation rates of other substrates, e.g. malonyl-CoA serving as a negative control (**Figure 32A**). Aaron proceeded with optimizing the set-up (section 4.1.14) and determining absolute kinetic constants for the transacylation of octanoyl moieties via the KS domain within his master thesis, which is summarized in **Figure 32B** and **C**.[84]

Kinetic Parameter	Michaelis-Menten fit	Lineweaver-Burk fit
k_{cat} s^{-1}	1.1 ± 0.1	1.1 ± 0.1
$K_m^{Oc\,CoA}$ μM	126 ± 6.1	122 ± 2.6
$k_{cat}/K_m^{Oc\text{-}CoA}$ M^{-1} s^{-1}	$9.0 \pm 0.8 \times 10^3$	$9.1 \pm 0.5 \times 10^3$
K_m^{ACP} μM (k_{cat} vs[ACP])	22 ± 5.3	18 ± 2.1
K_m^{ACP} μM ((K_m vs[ACP])	4 ± 1.0	2.4 ± 0.3
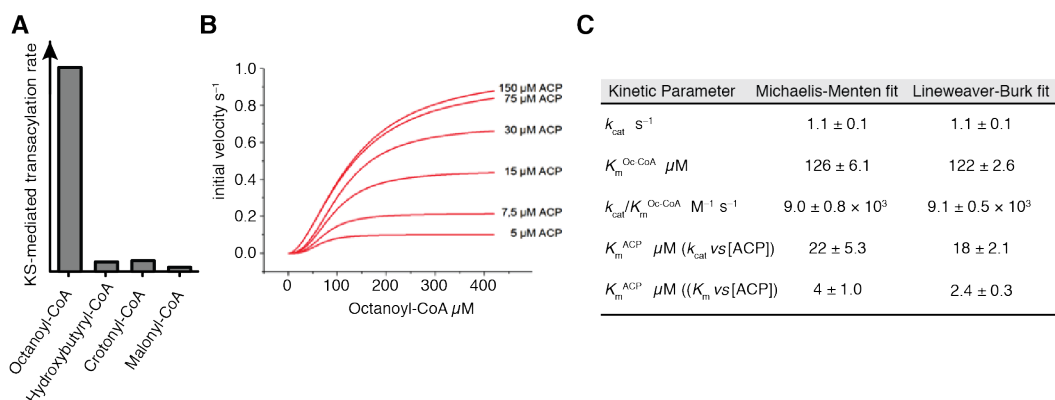
Figure 32: KS-mediated transacylation from CoA-esters to ACP. (A) Qualitative KS-mediated transacylation of acyl moieties from CoA-ester (250 µM for octanoyl-CoA and 500 µM for other substrates) to ACP (75 µM) measured with the αKGDH-assay (section 4.1.10). (B) Initial velocities were plotted against the octanoyl-CoA concentration at six fixed ACP concentrations. Data was fit by the Hill equation rather than the Michaelis-Menten equation as data points showed a sigmoidal increase. Measurements were performed by Aaron Himmler. (C) Absolute kinetic constants for the KS-mediated transfer of octanoyl moieties. Apparent K_m and k_{cat} values were plotted against the ACP concentration and fit with the Michaelis-Menten equation and the reciprocal of apparent K_m and k_{cat} values were plotted against the reciprocal ACP concentrations and fit with the Lineweaver-Burk equation (equation [9]).

Following the typical approach, we have plotted the initial rates versus the octanoyl-CoA concentrations at six fixed ACP concetrations to determine kinetic parameters for the transthioesterification of octanoyl moieties. As data points could not properly be fit with the Michaelis-Menten function, we used the Hill equation (section 5.2.5; equation [12]) for fitting instead, which well described the sigmoidal shape of the data points. Typically, the Hill coefficients were around two, which is shown and analyzed in detail in Aaron's master thesis.[84] We interpreted this deviation from standard Michaelis-Menten kinetics with cooperativity between both active sites of the KS domain. Although this finding is rather unexpected and to the best of my knowledge, cooperativity in the KS dimer has not yet been described in literature, it is not unlikely that both active sites catalytically influence each other. The acyl binding pockets of both KS domains are connected and merge at the dimeric interface, as will be described in section 2.4.3.1.

We determined absolute kinetic constants by plotting apparent K_m and k_{cat} values individually versus the ACP concentrations and fitted with the MIchaelis-Menten function (**Figure 32C**). Similar to the approach for the MAT domain (section 2.3.2), we have also applied double-reciprocal plotting with subsequent fitting with the Lineweaver-Burk equation (equation [9]), which gave comparable results. For fitting with the Michaelis-Menten function, the k_{cat} was determined to be 1.1 ± 0.1 s^{-1} and the K_m for octanoyl-CoA was 126 ± 6.1 µM, respectively ($k_{cat}/K_m = 9.0 \pm 0.8$ s^{-1} mM^{-1}). These values compare relatively well with the reported ones (bar plot: k_{cat} = ca. 1.75 s^{-1} and k_{cat}/K_m = ca. 4.2 s^{-1} mM^{-1}), although a triple mutant of rat FAS (S581A, S2151A, S2302A) was used at 37 °C with pantetheine as acceptor, instead of ACP.[49]

These results confirm, that only minor adaptations are required to also use the αKGDH-assay for investigating KS-mediated transthioesterification, which further broadens the applicability of the assay. Nevertheless, the relatively low acylation rate of the KS domain (ca. 1 s^{-1} for transacylation of octanoyl moieties to ACP) by CoA-esters demonstrates that this is not the naturally occurring way during FA synthesis, as loading is performed by the MAT domain and acyl chains are normally offered as ACP-esters to the KS domain. Hence, one should try using pantetheine-esters as donors to obtain faster turn over rates, which might better reflect rates of the native system, which will be discussed in section 3.2.3. The assay is good enough at the current state though to screen variants of the KS domain for altered specificities and efficiencies, which has already been tested in Aaron's master thesis.

2.4   Structural Studies on the Condensing Part

After kinetic characterization of both MAT and KS domains in the preceding chapter, we further aimed at also gaining structural insight into these domains. There is plenty information provided by literature for homologous structures, for example: porcine FAS (pdb: 2vz9)[31], human KS-MAT didomain (pdb: 3hhd)[64] and for domains from PKS: AT0$_{AVES}$ (pdb: 4rl1)[24] and DEBS KS-AT didomains of module 3 (pdb: 2qo3)[119] and module 5 (pdb: 2hg4)[120]. Nevertheless, structures of substrate-bound states are rare, for example from type II FAS of *E. coli*: FabD with malonyl bound (pdb: 2g2z)[121] and KAS I (FabB) with octanoyl bound (pdb: 2bui)[122] or AT of DynE with malonyl bound (pdb: 4amp)[118] and these domains differ either in overall architecture or substrate specificity. Therefore, we tried to solve the first structures of murine KS-MAT didomain by X-ray crystallography in substrate bound states. We focused on two different substrates, which will be scope of the following sections: 2.4.1 (malonyl-CoA) and 2.4.3 (octanoyl-CoA). This shall give further information on the origin of characteristic specificities and may aid in rational alteration of these (section 2.5).

2.4.1   Crystal Structure of KS-MAT Didomain with Bound Malonyl Moiety

This part of the thesis is again taken from a manuscript written with the help of Karthik Paithankar and Martin Grininger.[115] Furthermore, it shall be highlighted that Kathik Paithankar solved the crystal structure and I performed model verification and optimization plus the analysis. Khanh Vu Huu and Aaron Himmler participated in finding appropriate crystallization conditions and setting up crystallization plates.

The purified mouse KS-MAT didomain was crystallized, and crystals were soaked with malonyl-CoA. X-ray diffraction data were collected up to a resolution of 2.9 Å, and the resulting structural model refined to R/R$_{free}$ of 0.18/0.24 (**Table 2**). The asymmetric unit was found to contain four molecules (referred to as chains A, B, C and D; **Figure S7**), arranged as two biological dimers wedged in the cleft between the KS and the linker domain (**Figure 33A**). The individual KS-MAT dimers are observed in a conformation that superimposes well with the KS-MAT architectures of the natively purified porcine FAS[31] and the corresponding didomain structure of human FAS[64] (average root-mean-square deviation (RMSD) (Cα) of biological dimers to porcine FAS of 1.6 and 1.7 Å; human KS-MAT superimposes to porcine FAS with average RMSD 1.3

and 1.4 Å).[123] These didomains show remarkable resemblance to the KS-AT substructures of PKS.[70, 120]

Table 2: Crystallographic data collection and refinement statistics for murine KS-MAT crystal soaked with malonyl-CoA

X-ray diffraction source	SLS X06SA
Wavelength (Å)	0.97852
Temperature (K)	100
Space group	$C222_1$
a, b, c (Å)	147, 354.4, 217.4
α, β, γ (°)	90, 90, 90
Resolution range (Å)	50 – 2.9 (2.95 – 2.9)
Total no. of reflections	787148
Unique reflections	120224
Completeness	96 (98)
Redundancy	6 (6)
$<I/\sigma(I)>$	9 (1)
R_{meas}	0.2 (2)
Wilson B factor (Å^2)	46
R_{work}/R_{free}	0.18/0.22
No. of residues (A, B, C, D)	853, 848, 849, 853
RMSD bond lengths (Å)	0.01
RMSD bond angles (deg)	1.7

In the structure, the N-terminal 407 residues (1-407) form the KS domain, which is completed by the residue stretch 859–886 (referred to as the post-AT linker). Linker domains LD and LD2, of residue ranges 408-487 and 807-823, separate KS from the MAT domain. The MAT domain of residue range 488-806 is comprised of a α/β-hydrolase core interspersed with a ferredoxin-type fold (616-684), which is located at the outer wings of the structure (**Figure 33**). The four molecules in the asymmetric unit show overall similar structure and superimpose with a RMSD (Cα) of 1, 0.8, and 0.9 Å (for chains A, B and C *vs*. chain D). As observed in other similar X-ray crystal structures,[64] the subdomains of four molecules of mouse FAS KS-MAT undergo rigid-body movement (**Figure 33C**). A KS domain based superposition (average RMSD (Cα) 0.1 Å over residue range 1-407) indicates the variability of relative KS vs. MAT positioning. A superposition of Cα atoms of the MAT α/β-hydrolase domain (MAT without the ferredoxin-like fold, average RMSD (Cα) is less than 0.4 Å over residue ranges 488-615 and 685-806) shows pronounced conformational flexibility for the ferredoxin-like fold, represented by a

rotational movement of the ferredoxin-like fold of 10.6, 6.3 and 10.9 degrees (values given for chains A, B and C *vs*. chain D) (**Figure 33C**).
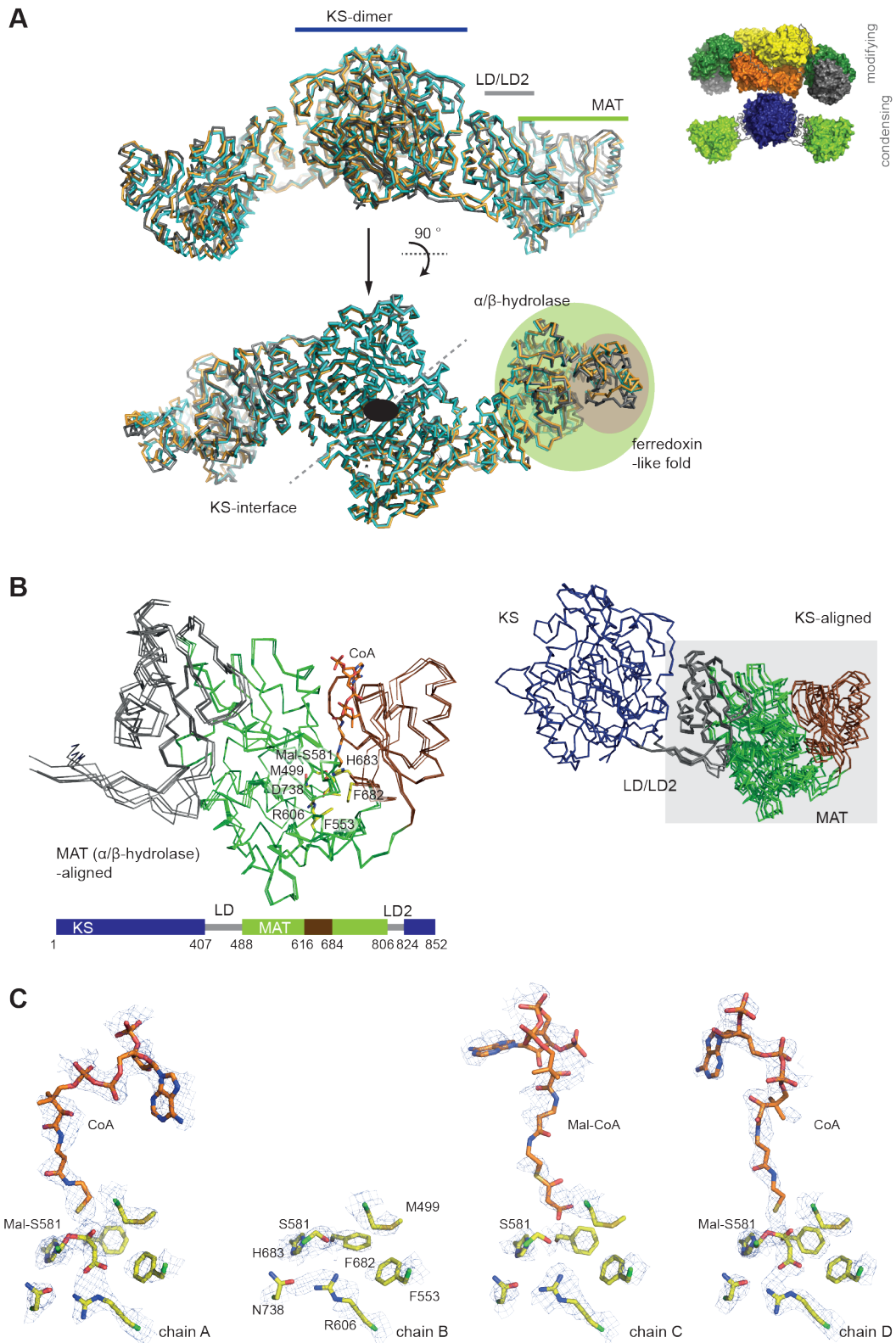
Figure 33: Overall structure of murine KS-MAT didomain. (A) Superposition of the two biological dimers of KS-MAT onto the porcine FAS KS-MAT dimer. For clarity, the full X-ray structural model of porcine FAS, excluding the domains ACP and TE that were not traceable in electron density,[31] is shown as inset. Porcine FAS KS-MAT is shown in grey, and the two biological dimers in orange (chain A and B) and cyan (chain C and D). Subdomains and folds are indicated by labels and coloring. (B) Superposition of the four KS-MAT modules in the asymmetric unit aligned via the KS domains (residue range 1–407; upper panel) and of the MAT (including LD/LD2) aligned via their α/β-hydrolase domain (residue ranges 488–615 and 685–806; lower panel). Domains and folds are color coded as depicted in the attached cartoon. Cofactor and residues are shown for chain D (malonyl-loaded state). (C) Binding site of murine MAT in its apo-state (chain B), in complex with malonyl-CoA (chain C) and in its acyl-enzyme state (chain A and D). Electron density of an unbiased feature-enhanced map of the active site regions is contoured at 1σ.[124] View 70° rotated (y-axis) as compared to **Figure 33B**.

Although the four molecules in the asymmetric unit superimpose well onto one another, we found striking differences in substrate binding. Upon soaking KS-MAT crystals with malonyl-CoA, the malonyl group was transferred onto the MAT domain yielding the malonyl-enzyme covalent complex in two of the four molecules (chain A and D; see **Figure 34A**), confirming MAT activity in the crystallized conformation. In chain C, one molecule malonyl-CoA is non-covalently bound at the entrance of the active site tunnel, whereas no substrate density could be detected in chain B. Since there is a large amount of void space available to allow substrates to enter active sites, crystal packing can be discounted as reasoning for the observed differences in malonyl-CoA binding. Rather, subtle local structural differences in relative positioning of α/β-hydrolase and ferredoxin-like subdomains may explain the observed variations in substrate binding. The subunits in acyl-enzyme states (chains A and D) show striking differences in binding the nucleotide part of the CoA moiety. Whereas the position in chain D resembles the one observed in FabD, the nucleotide part in chain A is directed away from the ferredoxin-like fold mainly due to interaction of the nucleobase and the alpha-phosphate with R787 of the hydrolase fold.
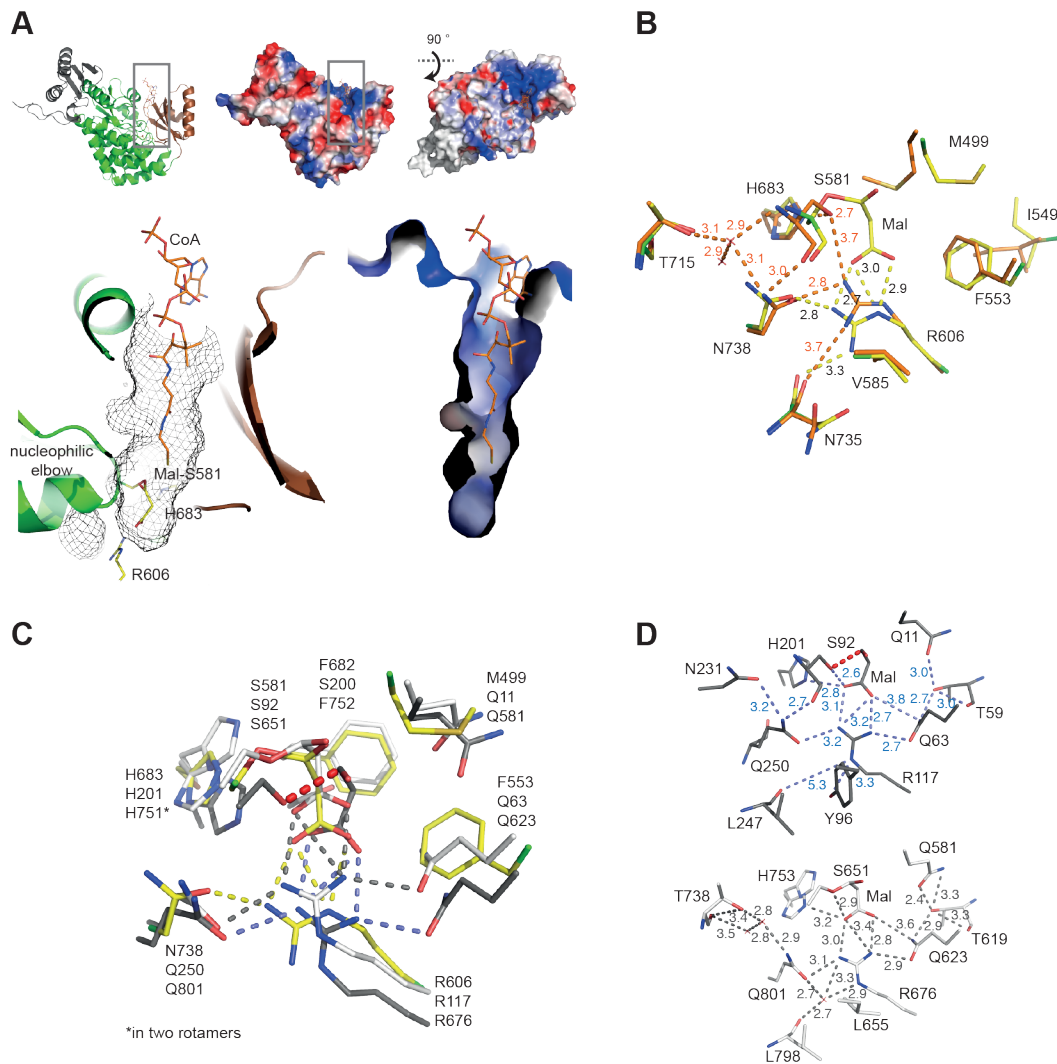
Figure 34: MAT binding site architecture. (A) Acyl-CoA binding cavity as comprised by the α/β-hydrolase and the ferredoxin-like domain. The binding cavity is shown in mesh representation (left panel) and with surfaces colored in electrostatic *potential (right).* The cavity was drawn with default settings of *PyMOL (vacuum electrostatics function; The PyMOL Molecular Graphics System, Schrödinger, LLC.).* Figures of the LD-MAT substructure in the upper panel are provided for overview. (B) Comparison of the binding site of the acyl-enzyme state (chain D) with the apo-state (chain A, orange) of human MAT. Distances indicate that N738 is the only residue in H-bond distance in the apo state. Superposition has been performed with α/β-hydrolase domains (residue ranges 488–615 and 685–806; pdb: 3hhd).[64] Residue numbers are identical for both proteins. H-bonding to the guanidinium group is indicated by dashed lines. (C) Malonyl-loaded enzyme states of MAT (chain D, residues in yellow) in superposition with malonyl transferases FabD (pdb: 2g2z, grey, H-bonds in blue dashed lines),[121] and AT of DynE8 (pdb: 4amp, white, H-bonds in grey).[118] Superposition indicates side-on *vs.* end-on bidentate coordination of arginine and the carboxyl group of serine bound malonyl. (D) H-bond networks of active site arginines are different in FabD (top) and in AT of DynE8 (bottom) to MAT (see **Figure 34**B).

2.4.2 Mechanistic Insight into MAT-mediated Substrate Transfer

The MAT entry tunnel and active site are located in the cleft between both subdomains, and are characterized by a narrow shape and a hydrophobic character (**Figure 34B**).[31, 125] The active site serine (S581) is positioned in a strand-turn-helix motif (identified by the consensus sequence GHSXG), stated the 'nucleophile elbow',[126] where its nucleophilicity is enhanced by a helix dipole-moment.[127] Furthermore, the tightness of the strand-turn-helix motif induces the nucleophilic residue to adopt energetically unfavorable main chain torsion angles (located in the generously allowed region of the Ramachandran plot), and imposes steric restrictions on residues located in its proximity. This is very important for transition state stabilization by the oxyanion hole.[126] The catalytic dyad is accomplished by the presence of H683 at short distance, as the general acid/base catalyst of the transferase reaction. Here, the N_τ of the imidazole ring accepts the S581 hydroxyl group proton, whereas the N_π is in hydrogen bond distance to backbone carbonyls of residues N738, L739 and S741. The malonyl-enzyme state allows localizing the oxyanion hole. The malonyl-ester's carbonyl oxygen is in hydrogen bond distance to the backbone amide of M499 in chain A and D.[125] Intriguingly, the distance to the backbone amide of L582 is 3.6 Å in both chains (A and D), which is slightly longer than a typical hydrogen bond distance. This may indicate that the trapped acyl-enzyme conformational state has already slightly moved away from the transition state conformation.

Another key residue for transferring extender substrates is R606,[51] which is located at the bottom of the active site tunnel. In chains A and D, the acyl-enzyme state binding pocket show the typical bidendate interaction formed between the carboxyl of the malonyl moiety and the guanidinium group of R606. Contrary to structural data on the malonyl transferases FabD[121] and DynE8[118] where an end-on coordination is observed, the active site arginine R606 in murine KS-MAT forms a side-on bidentade interaction (**Figure 34C**). Additionally, R606 is stabilized by hydrogen bonds to the side-chain carbonyl of N738 and backbone carbonyl of N735. Although the overall electron densities of chain B and C are less well defined in this region, R606 is observed pointing towards S581. Such a rotamer conformation is also observed in all (unoccupied) active sites of human KS-MAT (**Figure 34C**). In this conformation, R606 forms hydrogen bonds only to the side chain carbonyl of N738, indicating the repositioning of the guanidinium moiety upon malonyl loading. In addition, R606 of MAT is embedded in an H-bond network that is less extended than observed in FabD and DynE. A strongly positionally conserved

phenylalanine (F553, murine FAS numbering), which replaces glutamine of FabD and DynE, is largely responsible for the reduced H-bonding of R606 (compare **Figure 34D**).

2.4.3 Crystal Structure of KS-MAT Didomain with Bound Octanoyl Moiety

Having found appropriate crystallization conditions for the KS-MAT didomain, we further aimed at also trapping KS in an acyl-bound state. The most promising approach for this purpose seemed to be soaking with octanoyl-CoA being the most suitable substrate judged on turnover rates in interthiol transfer.[49] Following the established protocol, we soaked crystals and collected X-ray diffraction data up to a resolution of 2.9 Å. The structural model was obtained by using molecular replacement with/facilitating our malonyl-bound KS-MAT model (pdb: 5my0) and refined to R/R_{free} of 0.16/0.22 (**Table 3**). Similarly to our former crystal structure, the asymmetric unit was found to contain four polypeptide chains (A-D) arranged as two biological dimers interacting in the cleft between KS and the linker domain.

Table 3: Crystallographic data collection and refinement statistics for murine KS-MAT crystal soaked with octanoyl-CoA

Sample	Wild-type
Soak	Octanoyl-CoA
Diffraction source	SLS X06SA
Wavelength (Å)	0.97852
Temperature (K)	100
Space group	$C222_1$
a, b, c (Å)	147.4 354.0 218.5
α, β, γ (°)	90, 90, 90
Resolution range (Å)	50 – 2.88 (2.9 – 2.8)
Total number of reflections	1688104
Unique reflections	121669
Completeness	94 (70)
Redundancy	14 (12)
$<I/\sigma(I)>$	18 (3)
R_{meas}	0.14 (1)
Wilson B factor (Å^2)	66
R_{work}/R_{free}	0.16/0.22
Residues	1-852
(A, B, C, D)	1-611,613-673,680-852
	1-635,638-672,675-852
	1-852

Structural data revealed that the octanoyl-moiety was transferred onto the KS domain in all four chains yielding the octanoyl-enzyme covalent complex. Furthermore, an octanoyl-group was also transferred to S581 of MAT in chain A and D generating the acyl-enzyme state with an additional octanoyl-CoA molecule being bound non-covalently in chain D. The analysis of structural data and their impact on catalysis will be described separately in the following starting with the KS domain.

2.4.3.1. Description of the Octanoyl-bound KS domain

The KS domain possesses a thiolase fold, which contains an architecture of alternating layers of α-helices and β-sheets ($\alpha/\beta/\alpha/\beta/\alpha$) (see **Figure 33A**, chain B). It forms a dimeric interface, which is the largest contribution of a domain to the dimerization of FAS. The entry to the active site is located at the side of the KS domain, where a small vestibule is formed containing all active site residues allowing access to the nucleophilic cysteine (C161) (**Figure 35B**). The substrate binding tunnels of the protomers further extend along the dimer interface to merge at the two-fold axis. Interestingly, the acyl binding pocket forms an extra-cavity in the middle of the binding pocket, right at the omega-carbon atom of the bound octanoyl-moiety allowing in principle two binding modes for longer acyl chains.

All four molecules in the asymmetric unit align very well in a KS domain based superposition with an average RMSD (backbone atoms) of 0.2 Å over the residue range 1–407. The position and conformation of all active site residues, taken from a mechanistic study on homologous KAS I (FabB),[122] are essentially the same (**Figure 35C**, left panel). Only the bound octanoyl-chain shows a certain degree of positional variability due to an unconstraint rotational freedom of the single bonds of the alkyl chain. Similarly to the equivilant position in MAT (S581), the active site nucleophile C161 (in this case a cysteine) is positioned in a nucleophile elbow, which enhances nucleophilicity by a helix dipole-moment. Indeed, a study on a thiolase has shown,[128] that such arrangements can decrease the thiol's pk_a to 5.4 leading to a deprotonation of the thiol in water. The bound octanoyl-chain allows localization of the oxyanion hole, which is created by backbone amides of residues C161 and F395. In all chains, the carbonyl's oxygen is in hydrogen bond distance (2.8–3.2 Å) to the corresponding amides implying a ready-to-react conformation. Furthermore, there are two essential, conserved histidines (H293 and H331) in the active site responsible for creation of the C-nuclephile. The N_τ of H331 is in hydrogen bond distance (3.0–3.5 Å) to C163, whereas N_π accepts hydrogen bonds from backbone amides of P332 and E333. Interestingly, this structural arrangement implies a

different role of this histidine, compared to the catalytic dyad of MAT, being a general acid in catalysis. This role has been confirmed by mutant studies assigning both histidines a relatively modest effect on *trans*-thioesterification and hence thiol activation, but being essential for decarboxylation and condensation, respectively.[122]

Next, we were interested in conformational changes induced by the loaded octanoyl-chain. For this purpose, all four chains of 5my0 (unbound KS) were aligned to chain A (serving as the representative chain) in a KS domain based superposition (residues 1–407) (**Figure 35C**, middle panel). Again, average RMSD were small (0.2 Å, backbone atoms) allowing good interpretation of differences in side chain conformations or shifts. Most prominently, upon binding, the stretch of residues FGF (F393, G394, F395) is slightly shifted and reorganized. This results in a small displacement of F395 by 0.8 Å plus a rotation of the side chain by more than 90°. The rotamer of F395 in the unbound state would clash into the octanoyl-chain. Furthermore, the active site H331 is slightly tilted, whereas other active side residues align perfectly well.

As before mentioned, there is a high resolution structure (2.4 Å) of KAS I in an octanoyl-bound state available,[122] which is responsible for elongating C4 to C16 in the bacterial type II fatty acid synthesis. Active site residues and their positions are highly conserved (**Figure 35C**, right panel). Interestingly, F392 in KAS I adopts a different conformation compared to equivalent F395 in mFAS. In addition, a glutamate (E342), which was thought to participate in catalysis by stabilizing a water molecule, is not conserved and exchanged by alanine (A340) in KS excluding an equivalent role in FAS.
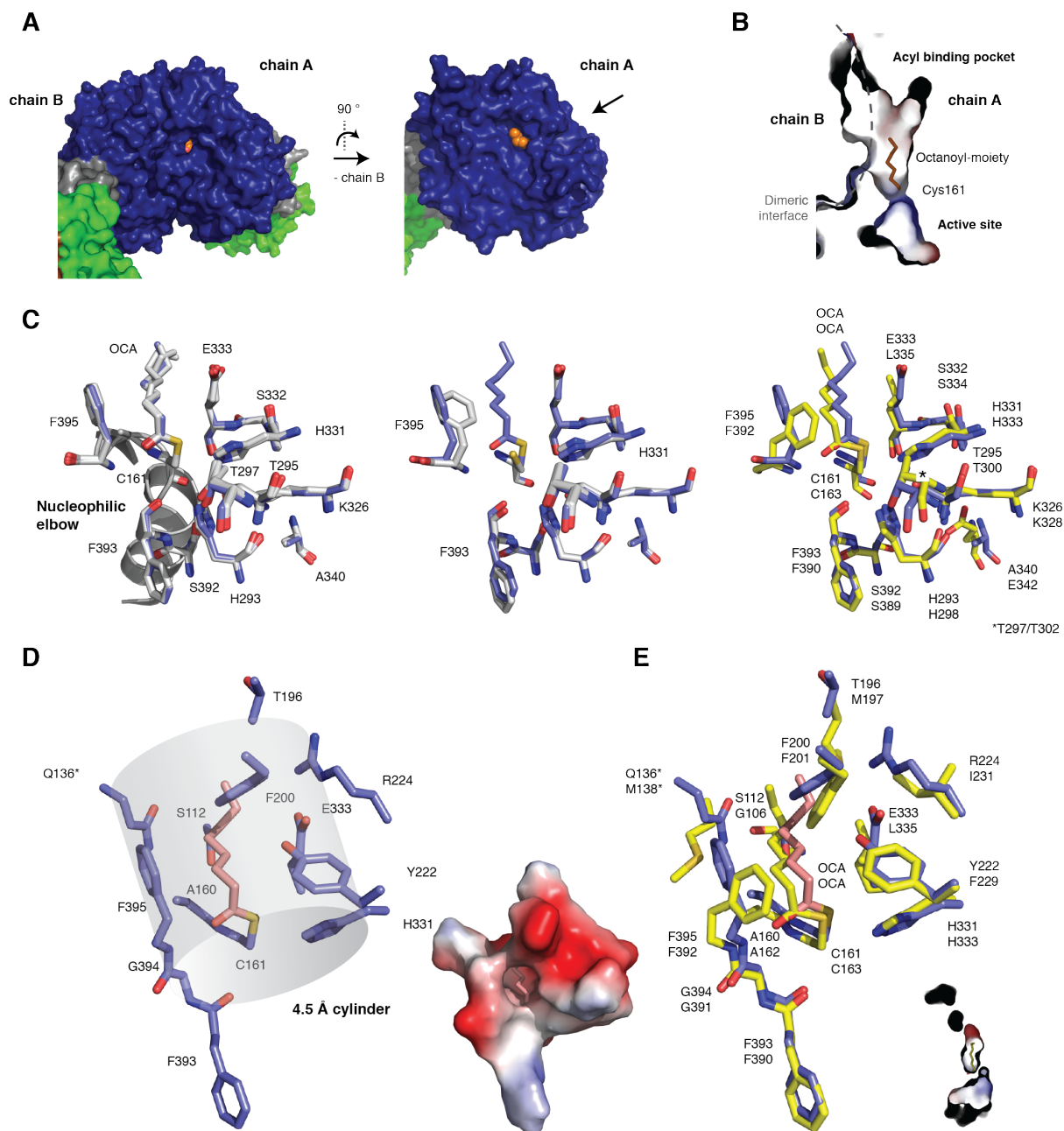
Figure 35: KS binding site architecture. (A) Side view on KS domain focusing on the active site entry. A surface depiction is shown in the usual color code. The view is rotated and a monomer (chain A) is shown to highlight the position of a bound octanoyl-moiety (orange) in the inner acyl binding pocket. The arrow indicates the active site entry. (B) Active site and acyl binding cavity of KS. In addition to the substrate binding cavity between the dimer interface a small side chamber is visible in the monomer. The binding cavity is shown with surfaces colored in electrostatic potential. (C) Active site of KS showing residues, which were reported to participate in catalysis. All four chains with an octanoyl-moiety bound were aligned by a KS based superposition (residues 1–407) (left panel). All residues adopt essentially the same conformation with a certain degree of freedom for the distal carbon atoms of the octanoyl-moiety. The middle panel shows a KS based superposition of octanoyl-bound active site of KS (chain A) with the four active sites of unoccupied KS domains (5my0). F395 is slightly shifted and the side chain is rotated by more than 90°. The right panel shows a superposition of the

active site with equivalent residues of KAS I (yellow). (D) Acyl binding pocket of KS. All residues within a distance of 4.5 Å to the octanoyl chain were identified and are shown. Analysis was performed with MacPymol. The inset depicts the surfaces of those residues colored in electrostatic potential forming the entrance to the acyl binding pocket. (E) Comparison of acyl binding pocket to KAS I. The corresponding residues from KAS I (yellow) are superimposed indicating deviations in the cavity's geometry, mainly induced by the appearance of two methionines (M138* and M197).

To gain some information about the origin of the narrow substrate specificity of KS from FAS, we analyzed, which residues are in direct contact of the bound acyl-chain. For this purpose, we searched for residues, which are within a distance of 4.5 Å to any atom of the bound octanoyl-moiety. This cylinder is schematically depicted in **Figure 35D** identifying thirteen residues (S112, Q136* (other dimer), A160, C161, T196, F200, Y222, R224, H331, E333, F393, G394 and F395). These residues alone are responsible for a very narrow entrance to the acyl binding pocket, which is rather hydrophobic at the inside. A comparison with the corresponding residues in KAS I indicates a good spatial conservation (**Figure 35E**). Furthermore, all three prominent aromatic residues (F200, Y222, F395) are also present in KAS I with the exception that a phenylalanine appears at position 222 (F201, F229, F393). This leads to a similar narrow, rather hydrophobic appearance of the acyl binding pocket in KAS I. Though, it shall be highlighted, that the overall binding pocket differs from KS of mFAS. It adopts a conformation that perfectly fits the octanoyl chain and neither the acyl binding pocket extends via the dimeric interface nor a side chamber can be found. Two methionines (M138* and M197) mainly account for this different appearance.

2.4.3.2. Description of the Octanoyl-bound MAT Domain

Similar to the previously described malonyl-bound structure, soaking with Oc-CoA also resulted in covalently bound enzyme states in chain A and D. This was very exciting, as it structurally confirmed the capability of MAT to transfer octanoyl-moieties (see section 2.3.3).

We commenced our analysis by aligning all four chains of the unit cell by a α/β-hydrolase based superposition (backbone atoms (BB) of 488–615; see **Figure 36A**). Again, all molecules superimpose very well with RMSDs (BB) of 0.6, 0.5, and 0.3 Å (for chains B, C and D *vs*. chain A). This structural alignment revealed remarkable differences in the relative positioning of the ferredoxin-like subdomain with local shifts of up to 7 Å between corresponding residues. The major differences appeared between the two octanoyl-bound active sites in chain A and D. In the chains, the covalently bound octanoyl-moieties are positioned in the cleft between both folds and the alkyl-chains point either towards the front in chain A or to the right side in chain D. This leads to dramatic changes in the active site's geometry compared to unbound or malonyl-bound states. Chain A adopts a very open conformation forming a cleft between the α/β-hydrolase and ferredoxin-like subdomains (**Figure 36**B). Such an open conformation has not been observed before in a MAT domain of type I FASs. Nevertheless, it is comparable to the open conformation observed in crystal structures of FabD (e.g. 1mla or 2g2z). In contrast, chain D adopts the classical, closed conformation leading to a long and narrow entry tunnel to the active site. What is very unusual though, is the occurrence of a second entry tunnel to the active site between the subdomains at the back. This tunnel accommodates the bound octanoyl-chain and procedes to the protein's surface. Furthermore, a hydrophobic cavity between the two subdomains is formed, which accommodates the additional octanoyl-chain from a non-covalently bound Oc-CoA. A search for all residues within 4 Å distance to the six carbon atoms (C3-C8) of the octanoyl chain revealed residues: I610, L615, L680, A681, F682, H683 and F686. The stretch of residues LAFH (680–683) is the most dominant motif for defining the specificity of ATs in PKSs (see section 1.4.2).[38]
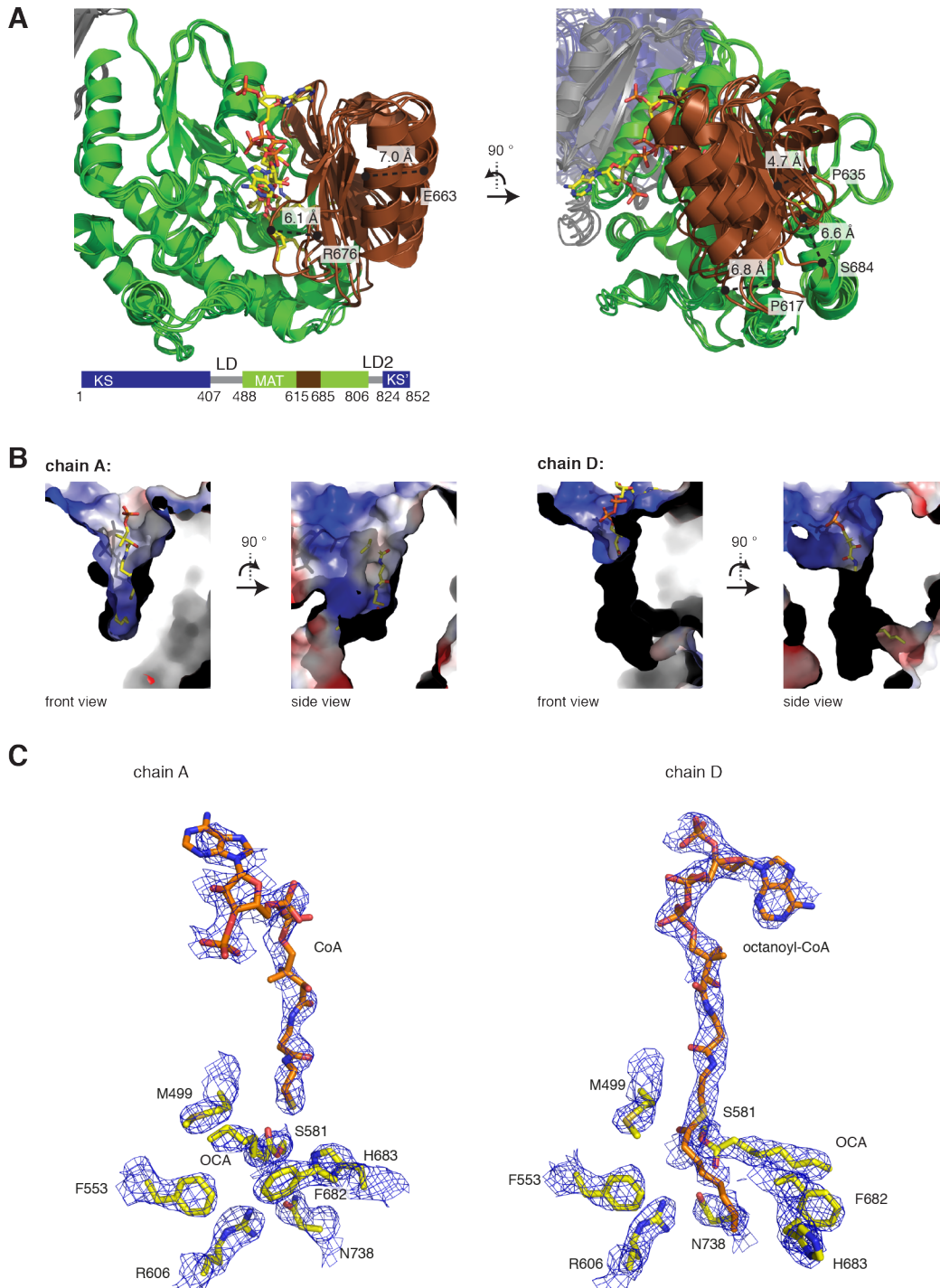
Figure 36: Architecture of octanoyl-bound active site of MAT. (A) Hydrolase fold based alignment of all four chains from the octanoyl-CoA soaked unit cell. MAT is shown in cartoon depiction with the subdomains coloured according to the commonly used colour code. Measurements between corresponding residues indicate a relative movement of the ferredoxin-like fold with local changes up to 7 Å. (B) Geometry of the active site and entry tunnel. A surface depiction of active sites of chain A and D is shown in two different views. Used colours indicate the surface electrostatics. Chain A adopts a very open conformation with a gorge shape comparable to FabD, whereas chain D forms the classical closed, narrow entry tunnel normally found in other FAS structures. (C) Active site of MAT in chain A and D.

Important active site residues are shown in sticks with the observed electron density shown in blue mesh. Additionally, clear electron density for covalently bound octanoyl-groups (OCA) to S581 was found in both chains. Furthermore, non-covalently bound CoA was found in chain A and octanoyl-CoA in chain D, respectively.

In the following, we take a closer look at important residues of the active site. The electron density was well defined facilitating the placement of all side chains (**Figure 36**C). As described before, the electron density clearly precedes from the active site's serines towards the front of the protein (chain A) and to the right (chain D), where the covalently bound octanoyl-chains were placed. R606 is again in hydrogen bond distance to the side chain amide of N738 as previously seen (compare **Figure 34**C). Interestingly, we observe R606 in the same conformation in both chains, which is different to the rotamer in unbound or malonyl-bound active sites. Hence, this is the third rotameric state we have seen in our crystal structures underlining the conformational freedom of this important residue in MAT. Although electron density at this position was rather weak, Maier *et al*. reported R606 in this particular conformation in the whole porcine FAS (2vz9) (**Figure S8**).[31] Furthermore, dramatic differences were found for residues F682 and H683 in the C-terminal connecting loop between the ferredoxin-like and the α/β-hydrolase subdomains. These residues will be discussed separately later on.

The non-covalently bound CoA in chain A and octanoyl-CoA in chain D also adopt two different conformations similar to previous observations. Electron density for CoA in chain A is clearly visible, but not defined enough at the surface to identify one binding state of the sugar and nucleobase part. Both parts point towards the α/β-hydrolase fold implying a rather loose interaction with the predominant positive surface of the α/β-hydrolase domain. On the contrary, the nucleobase of octanoyl-CoA in chain D is bound at a precise position between the two subdomains (**Figure 37**A). This conformation reveals a very interesting binding pocket for the nucleobase, which utilizes two different charged interactions (**Figure 37**B). Side chains of F671 and R773 are aligned perfectly in parallel to the plane of the purine ring in 3.5–4 Å distance indicating a planar stacking interaction. Besides the well known π-stacking between aromatic rings, also planar stacking interactions between arginines and aromatic rings have been reported before.[129] Additionally, two hydrogen bonds are formed between amines of the purine ring and the side chain hydroxyl-group of T648 and the backbone carbonyl-group of D647, respectively. Considering the exceptional good fit of the nucleobase in this pocket being stabilized by various forces (π-stacking and hydrogen-bonding), brings us to the conclusion that this is the predominant CoA binding pocket. Nevertheless, it is worth

mentioning, that formation of the binding pocket requires the accurate relative alignment of both subdomains and the presence of a CoA derivative. This could only be observed in chain D of the octanoyl-CoA soaked crystal and in chain D of the malonyl-CoA soaked crystal (5my0).
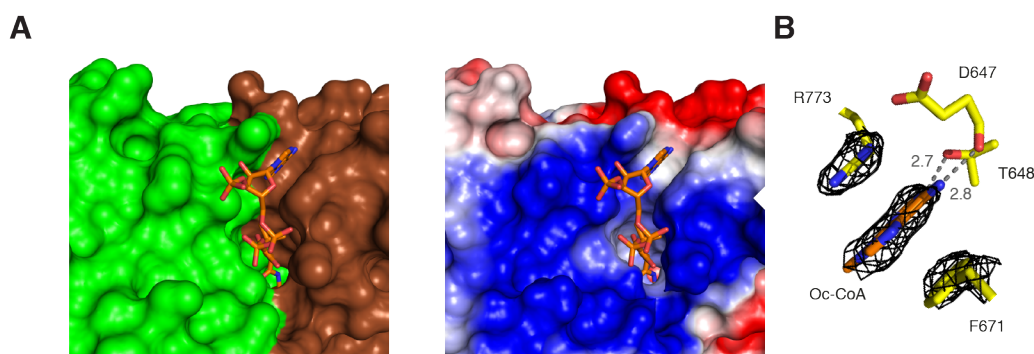
Figure 37: Deeper insight into the CoA binding pocket of chain D. (A) Investigation of the CoA binding pocket in chain D. Surfaces are coloured in either the subdomain colour code (left) or in electrostatic charges (right). (B) Detailed analysis of the nucleobase binding pocket. The nucleobase is coordinated by residues R773 and F671 via π-stacking interactions. Furthermore, two hydrogen bonds are formed between amines of the purine ring and the side chain hydroxyl-group of T648 or the backbone carbonyl-group of D647, respectively.

2.4.3.3. Further Analysis of Octanoyl-bound States of MAT

Basically, we observe three different octanoyl-moieties bound to the MAT domain, either in the acyl-enzyme state (chain A and D) or in the non-covalent octanoyl-CoA bound state. How can we decide or analyze, which of these orientations may be relevant during the enzyme catalyzed reaction and further, which enzyme conformation may reflect the conformation in the transition state complex best? Answering this question is of crucial importance to explain the unusual capability of MAT to transfer octanoyl-chains and may aid future rational mutagenesis approaches.

Stabilization of the negative charge occurring at the oxygen of the thioester group during the transition state by the enzyme is an essential feature for efficient catalysis of this enzyme class (see section 2.4.2). This oxyanion hole is created by the two backbone amide groups of M499 and L582, which have to be positioned in a sufficient distance and angle to the oxygen of the carbonyl group within the enzyme transition state complex. We exploited this aspect to identify the most relevant conformation of the octanoyl-chains in the active sites, which may similarly occur during enzyme-mediated catalysis. To this end, distances between the carbonyl groups of the active site esters or respective thioester and the backbone amide groups of M499 and L582 were measured (**Figure 38**). All

distances are longer than 3.5 Å and hence exceed the typical length of such hydrogen bonds. Especially, the oxygen of the carbonyl group of the bound octanoyl-moiety in chain D is up to 7.4 Å apart from the amide group of L582, which indicates that this conformation is likely not relevant during the progress of the reaction. This conclusion is supported by the fact that also the enzyme in chain D adopts an unusual conformation, as residue H683 from the catalytic dyad is rotated away from the catalytically meaningful position. Nevertheless, the measured distances for the enzyme-bound octanoyl moiety in chain A and for the non-covalently bound octanoyl-CoA in chain D are with 3.5–4 Å within a reasonable distance, not too far apart from typical hydrogen bond lengths. This means both orientations of the alkyl chains should be considered to identify the position of the octanoyl-moiety within the enzyme transition state complex, but a final conclusion certainly requires additional experiments.

To evaluate whether the observed conformation of the enzyme in chain A might occur during catalysis, one should further take the integrity of the catalytic dyad into considerations. The N_τ of the imidazole ring of H683 is with 2.7 Å in hydrogen bond distance to the active site serine and the N_π is 3.2, 3.5 and 3.0 Å away from backbone carbonyls of residues N738, L739 and S741. These distances compare very well to equivilant distances in chain D of the malonyl-CoA soaked structure and indicate that the catalytic dyad is intact.
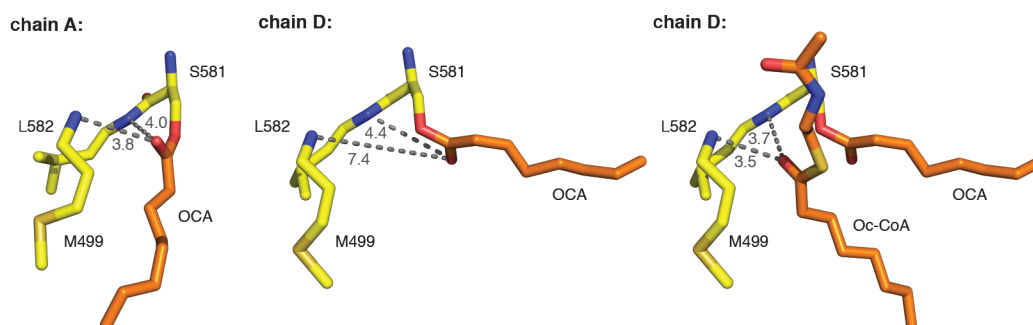
Figure 38: Analysis of the oxyanion hole in chain A and D. Hydrogen bond distances were measured between backbone amides of M499 and L582 and the carbonyl oxygen of the thioesters. Left and middle panel show the interaction of the covalently bound octanoyl-moiety and the right panel measures the corresponding distances to the non-covalently bound Oc-CoA in chain D.

Having crystal structures of murine MAT in different substrate bound states available and considering the previously solved structures of human and porcine MAT makes us wonder, whether one can draw a general structure-function relationship for the

MAT domain. Based on the assumption that certain plasticity of the active site to adopt individually to chemically diverse transition states is required for substrate promiscuity, it is important to identify the origin of conformational variability within the MAT domain.
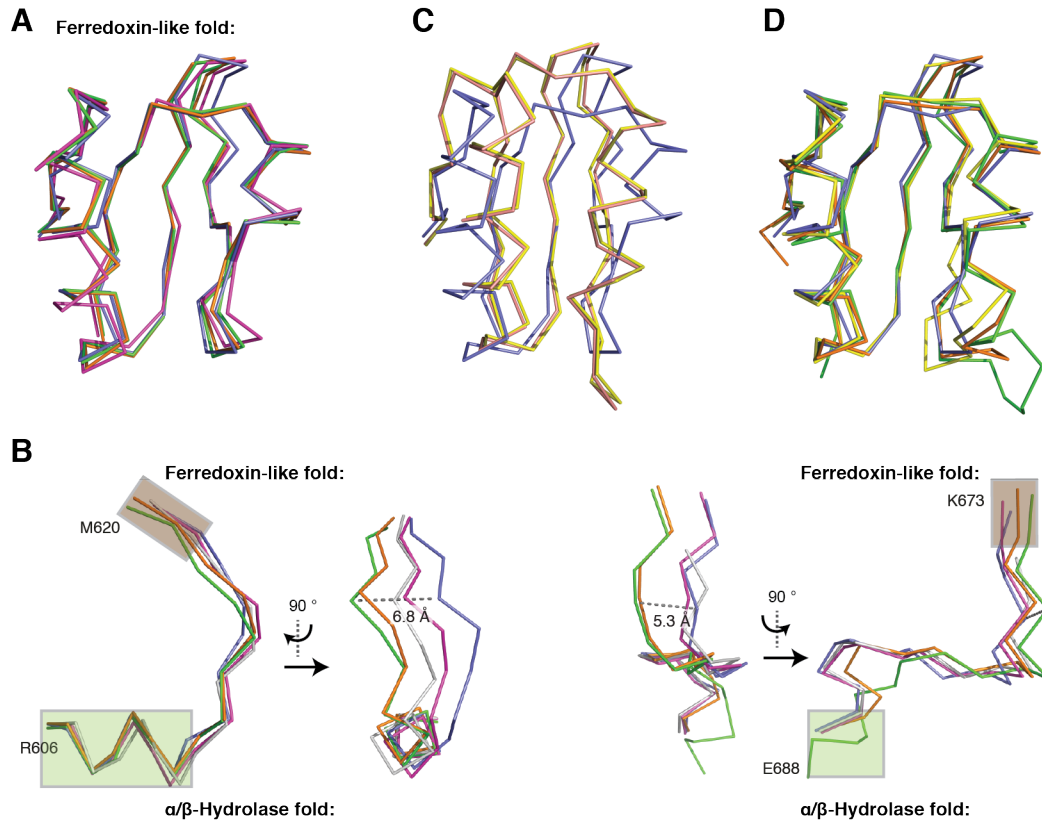
Figure 39: Analysis of the ferredoxin-like fold and the two linkers between the subdomains. (A) Ferredoxin-like fold based alignment (BB of residues 618-676) of chain C (grey) and D (green) of 6exg, chain D (orange) of 5my0 and chain A (purple) of porcine FAS (2gz9) with chain A (blue) of 6exg. Residue stretch 618-676 is shown in the ribbon depiction. (B) Analysis of linkers between the subdomains. Linker 1 (606-620, left panel) and linker 2 (673-688, right panel) are shown as ribbons after α/β-hydrolase fold based alignments. Both linker stretches are shown from two different perspectives rotated by 90°. Chain A (blue), chain C (grey) and chain D (green) from 6exg, chain D (orange) from 5my0 and human MAT chain A (purple) from 3hhd were used. Green and brown rectangles indicate secondary structure elements of the α/β-hydrolase- and ferredoxin-like-subdomain, respectively. (C) Ferredoxin-like fold based alignment of FabD (BB of residues 130-195) from malonyl-bound- (yellow; 2g2z) and apo-structure (orange; 1mla) with chain A (blue; 6exg). Corresponding residues are shown as ribbon. (D) Ferredoxin-like fold based alignment of M3$_{DEBS}$ chain B (yellow; BB of 687-748; 2qo3), M5$_{DEBS}$ chain A (orange; BB of 678-740; 2hg4) and AT0$_{AVES}$ chain A (green; BB of 155-221; 4rl1) with chain A (blue; 6exg). Corresponding residues are shown as ribbon.

We have already seen that the α/β-hydrolase subdomain is quite rigid and that different substrate bound states can be aligned very well with RMSDs less than 0.6 Å. Therefore, the rigidity of the ferredoxin-like subdomain was investigated next by an

ferredoxin-like fold based alignment (BB of residues 618–676) with chains A, C and D from the octanoyl-CoA soaked crystal (6exg), chain D from the malonyl-CoA soaked crystal (5my0) and chain A from the porcine FAS (2gz9) (**Figure 39A**). Again, all chains aligned quite well with RMSDs between 0.5–0.9 Å, at which the ferredoxin-like fold of the porcine FAS showed with 0.9 Å the major deviation. From the depiction in **Figure 39A**, it is clearly visible that the subdomain remains nearly unaltered throughout the aligned states, except some minor deviations in the loop regions. This was quite unexpected and led to the conclusion that most of the structural variability of MAT must originate from the positioning of both subdomains relative to each other and hence their linkage.

Indeed, the two linkers (linker 1: ca. 612–619 and linker 2: ca. 674–685) between the α/β-hydrolase- and the ferredoxin-like fold show a remarkable versatility. A α/β-hydrolase based alignment revisiting the same five chains used in the KS-based alignment, revealed a shift of up to 6.8 Å (linker 1) and 5.3 Å (linker 2) between corresponding residues (**Figure 39B**). Furthermore, residues 613-614 and 683-684 undergo significant changes in terms of main chain torsion angles as represented in the Ramachandran plot (**Figure 40**). In detail, the phi and psi angles of residues A613 and H614 of linker 1 change from the region of β-sheets to the right- and left-handed helical region, respectively, whereas respective angles of residues D612 and L615 remain nearly unaltered. Similarly, the phi and psi angles of residue H683 in linker 2 change from the left-handed helical region to the region of β-sheets, whereas the phi and psi angles of S684 change from the region of β-sheets to the right-handed helical region. Again main chain torsion angles of residues F682 and Y685 remain nearly unaltered.

At last, we were interested in the ferredoxin-like fold itself and how conserved this subdomain is between homologous structures. Again, chain A (6exg) served as the object of comparison and was structurally aligned with two subdomains from FabD (**Figure 39D**) and subdomains from modular PKSs: M3$_{DEBS}$, M5$_{DEBS}$ and AT0$_{AVES}$ (**Figure 39E**). FabD from the type II fatty acid synthase of *E. coli* was chosen for its exclusive specificity for malonyl-CoA. Similarly to murine MAT, binding of a substrate does not change the ferredoxin-like fold itself (compare malonyl-bound and apo-structure), but the bacterial subdomain shows significant structural differences to the murine subdomain (RMDS of 2.8–3.0 Å). Especially, the β-sheet is much more pronounced due to elongated β-strands and also α-helix 1 is extended substantially (twice as long). On the other hand, the representative ferredoxin-like folds from modular PKSs show a remarkable similarity to the murine subdomain (RMDS of 0.7–0.9 Å). Only AT0$_{AVES}$ possesses a slightly elongated loop between α-helix 1 and β-strand 2. It shall be mentioned that AT0$_{AVES}$ has a broad

tolerance for loading substrates, whereas ATs_{DEBS} have a rather strict specificity for methylmalony-CoA as far as known from literature.[69] Whether pronounced secondary structural elements in the ferredoxin-like fold are indeed necessary for a strict substrate specificity of an AT domain or whether they are just a consequence from a different evolution of FabD, does, however, require a broad systematic analysis of AT domains.
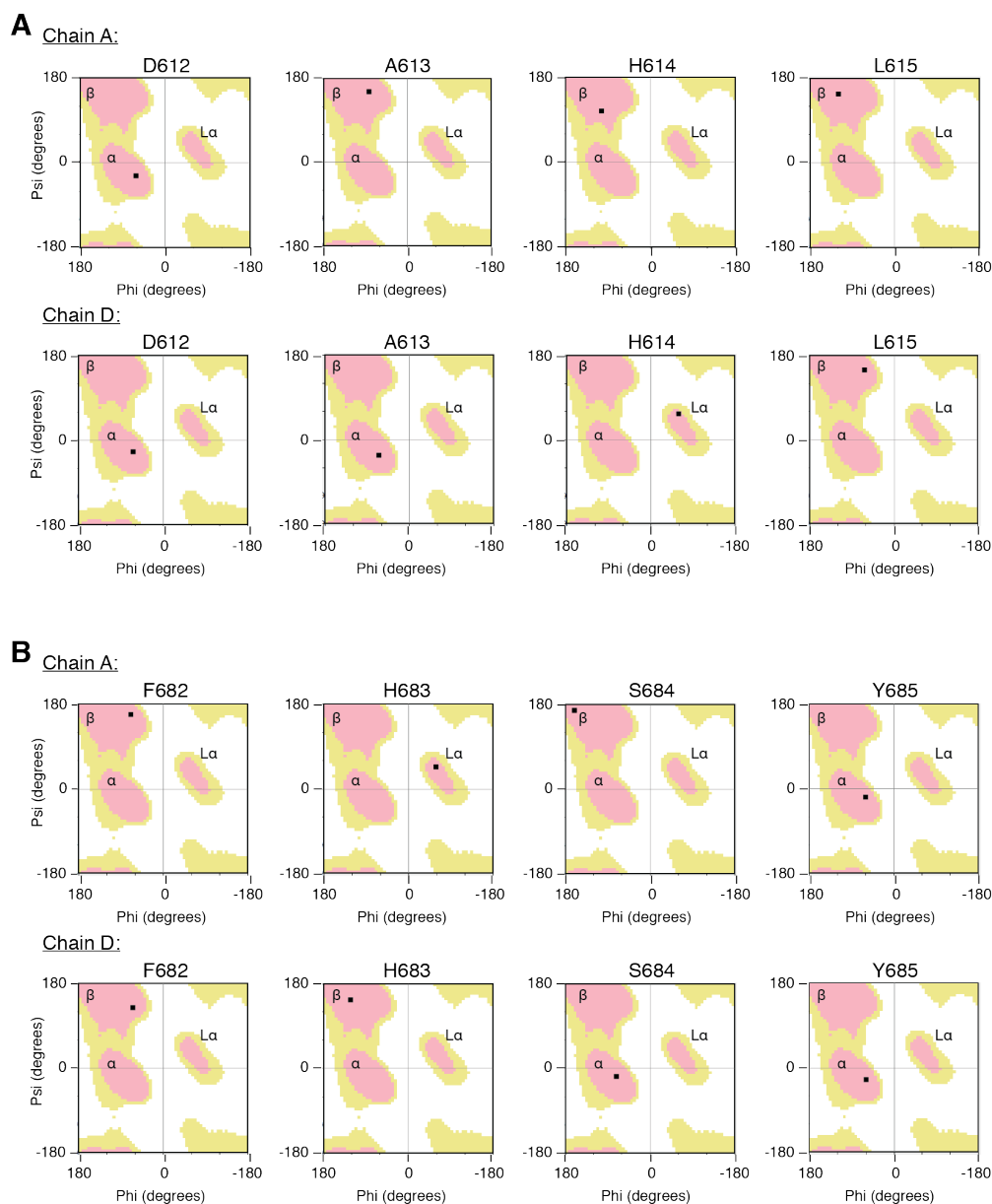
Figure 40: Ramachandran plots for crucial residues in both subdomain linkers. Significant changes in dihedral angles were seen for residues in linker 1 (A) 612-615 and linker 2 (B) 682-685 leading to changes to different allowed regions. Plots were created in coot. Used abbreviations indicate α – right-handed helical region, Lα – left-handed helical region and β – β-sheet region.

2.5 Rational Engineering of the Substrate Specificity of MAT

In the last two chapters, the structural and kinetic characterization of wild-type murine MAT has been reported. Utilizing our findings and gained experiences on this domain, we are aiming to rationally alter the substrate ambiguity, i.e. induce more substrate specificity into the MAT domain. It is generally believed that promiscuous precursor enzymes have acquired higher specificity and activity due to divergent evolution driven by selective pressure.[130] This theory has been confirmed in several studies by applying it to design enzymes with enhanced activity and more specificity in the laboratory.[131, 132] Hence, we are convinced that promiscuity of the MAT domain may also well serve as starting point for adaptive protein engineering intending to obtain an acyltransferase with higher specificity for a common substrate or even for unusual ones. The first step in such an approach is to identify residues (also called pasticity residues), which determine the specificity of the enzyme without disturbing overall catalytic efficiency. This shall further strengthen our molecular understanding of MAT-mediated catalysis as well as it shall further broaden the applicability of the domain.

2.5.1 Initial Mutant Screening

We have conducted a comprehensive mutant screening with 42 different variants aiming at identifying "hot spots" being responsible for the transfer of certain substrate groups (**Figure 41**). The initial approach based on ten positions, which are mainly located in the four sequence motifs determining substrate specificity that were introduced in section 1.4.2. Furthermore, we added position M499, which is a prominent residue of MATs from FASs, whereas specific ATs of most PKSs and also FabD are characterized by a conserved glutamine at this position. Based on previous structural information, also residues N738 and L739 were included in our considerations due to their vicinity to residue R606 in the acyl binding pocket.[64] Unfortunately, detailed structural information of murine MAT in different acyl bound states from our own laboratory (section 2.4.1 and 2.4.3) had not been available at the time when the mutagenesis study was started and other possible positions may arise based on new structural insights.
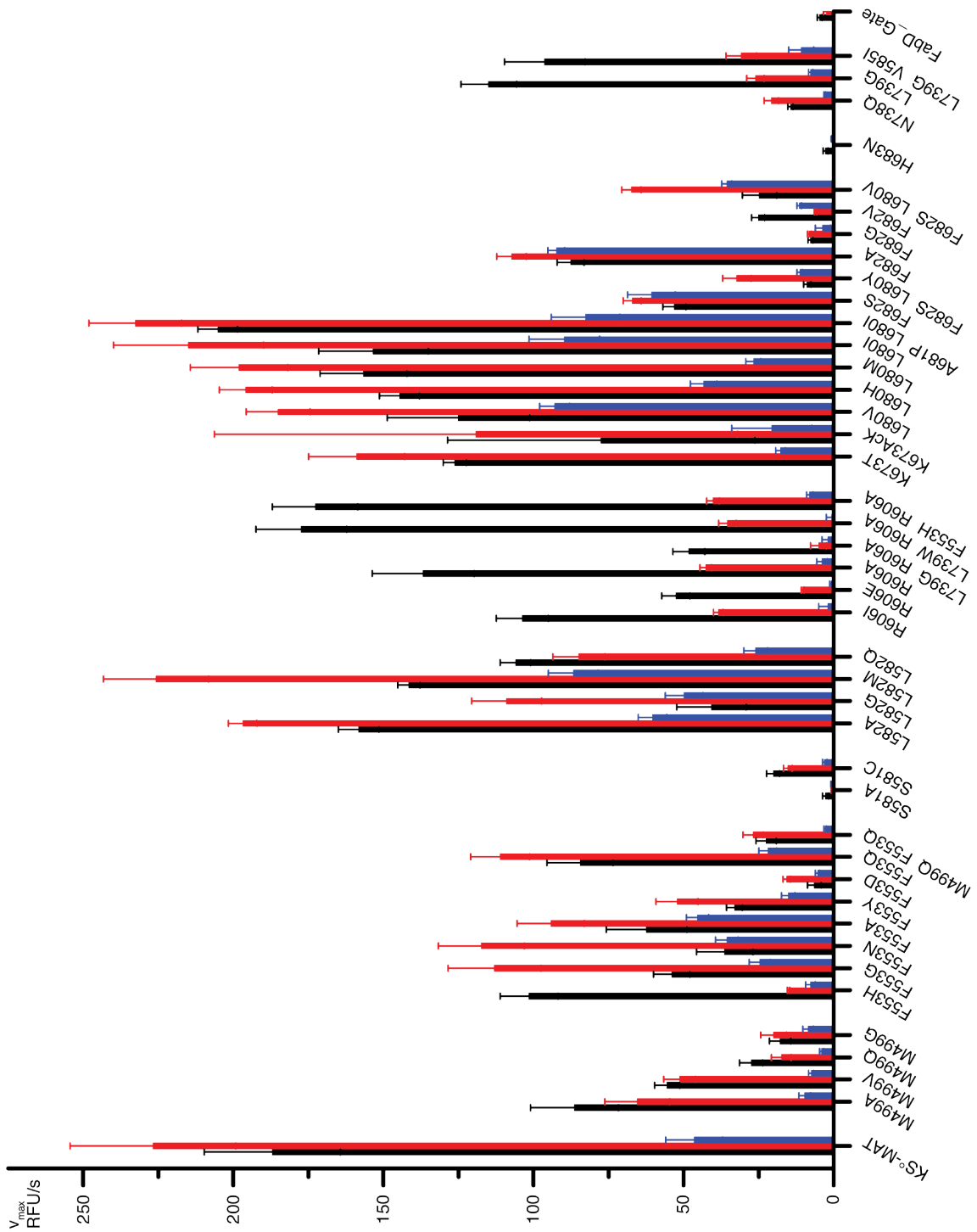
Figure 41: Comprehensive initial screening of MAT variants. Initial velocities for the substrates acetyl- (black bar), malonyl- (red bar) and methylmalonyl-CoA (blue bar) were determined with the αKGDH assay. Standard conditions were applied with fixed concentrations of substrates and ACP of 20 μM and 60 μM, respectively. Error bars reflect the standard deviation between three independent measurements using the same enzyme preparation.

The mutant screening was performed utilizing the KS-MAT didomain bearing the functional "knockout" C161G in the KS domain, similarly to the kinetic study on the wild-type MAT (section 2.3.2). All variants were separately expressed and purified, which was in parts performed by Dietmar Herzberg and Khanh Vu Huu within the scope of their master theses under my supervision,[110, 133] according to the established tandem affinity protocol. Utilizing the αKGDH assay (section 2.3.1.5), we ascertained values of the initial velocities for every variant with the three substrates, acetyl-, malonyl- and methylmalonyl-CoA at fixed concentrations of substrate (20 µM) and ACP (60 µM). The substrates were chosen to distinguish between generated loading- and extending-ATs plus to analyze the effect of the mutation on the type of extending-AT with respect of altering the preference for malonyl-CoA. We have established a quick screening set up allowing to measure initial velocities of seven mutants and three different substrate at once (see section 4.1.15).

In the following, some variants with significantly altered properties in comparison to the wild-type domain will be presented classified by their functionality. I) *Catalytically important residues without any impact on specificity.* Introducing mutations of the catalytic dyad, S581A and H683N (was preferred over H683A as this variant suffered from protein instability), essentially disrupt activity. Although a cysteine can also act as nucleophile, the mutation S581C reduces activity to 10 %. Similarly, changing of asparagine 738 to glutamine (N738Q) decreases activity to 10 %, which indicates the importance of proper positioning of this residue for backbone interaction with H683. Furthermore, glycine and valine at position 682 (F682G and F682V) next to the catalytic histidine led to a vast reduction of activity, although other residues at this position induced interesting properties as described below. Unexpectedly, all variants of M499 did not change the substrate specificity significantly, but diminished overall catalytic activity, which might be explained by the participation of the backbone carbonyl in forming the oxyanion whole. Finally, also exchanging the helix of the 'C-terminal region' motif with the corresponding helix of FabD failed due to stability issues of the mutant.

II) *Variants with acetyl-CoA specificity (loading-ATs).* The mutation R606A is well known from literature for abolishing the interaction of the guanidinium group with the carboxyl group of the malonyl moiety. Although the mutation induced a pronounced preference for acetyl moieties in murine MAT (a three times higher initial velocity compared to malonyl-CoA), the affect was less pronounced than reported for rat FAS. Instead of increasing the rate for acetyl transfer, it was reduced by circa 25 %.[51] Introducing residues isoleucine (often seen in loading-ATs) or glutamate (prospectively, negatively charged) at this position (R606I and R606E) further reduced the rate for acetyl

transfer and are not considered as promising. Though, two point mutations (F553H and L739W) in addition to R606A, indeed increased selectivity by slightly lowering the speed of malonyl transfer with simultaneously raising the rate of acetyl transfer almost to the wild-type value. Both mutations were chosen after the single point mutations F553H and L739G had attracted our attention with an unexpected high preference for acetyl moieties. The molecular basis for the behaviour of these variants remains unknown and requires further investigation.

III) *Variants with increased malonyl-CoA specificity (extender-ATs)*. The challenge to generate a specific variant of MAT for malonyl moieties is a matter of how to prevent the acceptance of acetyl groups. Unfortunately, there is no previous knowledge available from literature about how MAT manages acetyl transfer or rather how malonyl specific extender-ATs prevent the formation of an acetyl enzyme intermediate. An initial set of mutants was able to achieve initial promising results, and identified some variants with pronounced malonyl specificity, although a "real" malonyl specific variant of MAT, comparable to the highly specific FabD, has not yet been found. Most promising in this regard so far, seemed positions L582 and F553. Whereas variant L582M has a reduced rate of acetyl transfer by circa 25 % with an unaltered rate for malonyl moieties, there are three mutants L582G, F553G and F553N with an overall reduced activity (ca. 50 %), but a pronounced preference for malonyl moieties (up to 3.5 times faster malonyl than acetyl transfer).

IV) *Variants with increased methylmalonyl-CoA specificity (extender-ATs)*. We found the importance of sequence motifs 'YASH' and 'GHSXG' for determining the selectivity between elongating substrates confirmed in murine MAT. Several variants possessed an increased rate for the transfer of methylmalonyl moieties compared to the wild-type. Most prominently, mutations at position F682 to serine (F682S; see YASH motif) and even more to alanine (F682A) induced completely altered substrate specificities. Whereas, the rates of acetyl and malonyl transfers reduced two-fold, methylmalonyl moieties were transferred almost twice as fast making the mutant almost equally efficient for methylmalonyl-CoA. Additionally, the variant L680V shall be highlighted, as it has been inspired by avian FASs, which exclusively utilize valine at this position. This variant is characterized by a doubled transfer rate for methylmalonyl moieties with a concomitant reduction of the velocity for acetyl- and malonyl-CoA, which might indicate a distinct role of branched fatty acids in birds.

2.5.2 Further Characterization of Select Mutants

To verify preliminary results from the mutant screening and to gain further information about altered transfer kinetics including Michaelis constants, initial velocities were determined for select variants at different substrate concentrations. Utilizing the same set-up as in section 2.3.3 for the wild-type, apparent Michaelis and rate constants were obtained at a fixed concentration of the acceptor ACP at 60 μM. After initial K_m values had been approximated, final data were collected in the substrate range 0.2-5 × K_m and fit with the Hanes-Woolf function, which led to apparent kinetic parameters (Table 4 and Table 5).

Table 4: Kinetic analysis of the variant R606A at a fixed acceptor concentration of 60 μM ACP (n=6).

R606A variant	K_m^{app} (μM)	k_{cat}^{app} (s^{-1})	k_{cat}/K_m (10^6 × M^{-1} s^{-1})	Hydrolysis rate (10^{-3} × s^{-1})
Malonyl-CoA	18 ± 5	5 ± 1.7	0.3 ± 0.1	14.7 ± 2.3
Acetyl-CoA	3 ± 0.6	13 ± 2.1	4 ± 1.1	15.2 ± 2.1

The analysis was started with variant R606A, which is known from literature,[51] to confirm former findings with our experimental set up. Whereas hydrolysis rates remained nearly unaltered, the specificity constant for malonyl-CoA decreased significantly (as expected) being forty-fold smaller than the wild-type constant (k_{cat}/K_m of 0.3 ± 0.1 M^{-1} s^{-1}). This corresponds to a difference in transition state free energy of $\Delta\Delta G_{ES\ddagger}$ = 9.2 kJ mol^{-1} (2.3 kcal mol^{-1}), which is consistent with a H-bonding interaction. Contrary to the reported increase in transferring acetyl moieties for rat MAT, the murine variant was slightly slower than wild-type confirming the result from the screening. Interestingly, the addition of L739W indeed further increased specificity for acetyl groups, yielding an apparent specificity constant of 4.7 × 10^6 M^{-1} s^{-1} with an increased rate constant of 22 s^{-1}.

Table 5: Preliminary kinetic parameters of select variants at a fixed acceptor concentration of 60 µM ACP (n = 2).

R606A/L739W variant	$K_m{}^{app}$ (µM)	$k_{cat}{}^{app}$ (s^{-1})	k_{cat}/K_m ($10^6 \times$ M^{-1} s^{-1})
Malonyl-CoA	16	4.7	0.3
Acetyl-CoA	4.7	22	4.7
F553N variant			
Malonyl-CoA	2.8	11	3.9
Acetyl-CoA	2	2.9	1.4
L582M variant			
Malonyl-CoA	3.4	17	4.9
Acetyl-CoA	5.7	10	1.8
F682A variant			
Malonyl-CoA	2	6.4	3.2
Acetyl-CoA	3.4	6.2	1.8
Methylmalonyl-CoA	2.8	6.1	2.2

Replacement of F553 with asparagine yielded a malonyl specific variant with nearly unaltered K_ms, but five-fold decreased turnover rates. This variant prefers malonyl moieties by almost three times over acetyl groups judged by the corresponding specificity constants, which is an increase in specificity by two-fold compared to the wild-type properties (**Table 1**). A similar specificity was found in L582M, which is also characterized by an increased preference for malonyl moieties (almost three times, judged by k_{cat}/K_m values). This mutant increased in both K_m and k_{cat} values compared to the wild-type and to F553N, which can be seen as an advancement as natural evolution drives enzymes to higher kinetic parameters (see section 3.2).

The point mutation F682A has a drastic effect on transfer properties of murine MAT possessing equal kinetic parameters for all tested substrates. In detail, the respective specificity constants are decreased by four-fold (acetyl and malonyl moieties) and three-fold for methylmalonyl groups. Interestingly, the rate constants are essentially the same, which means a decrease of 2-3-fold for acetyl and malonyl moieties and an increase of 50 % for methylmalonyl groups compared to the wild-type.

Due to time limitations, velocity data on select variants does not yet cover large enough sample sets to satisfy standards for statistical analyses and has to be considered with care, accordingly.

Figure 42: Representative Hill fit for the transfer of octanoyl moieties of variant R606A/L739W at fixed ACP concentration (60 μM). Data points disobeyed standard Michaelis-Menten kinetics and were hence fit with the Hill equation (equation [12]). Derived kinetic parameters are given below the graph.
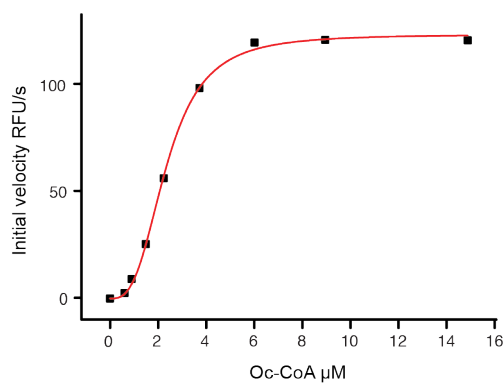
	$K^{\text{Oc-CoA}}$ M	k_{cat} s^{-1}	Hill coeff.
R606A/L739W	2.3×10^{-6}	4.1	3.0

For curiosity, the variant R606A/L739W was also tested with octanoyl-CoA as Rangan *et al.* reported an enormous increase in transferring longer carbon chains for the R606A variant. Such an acceleration could not be observed and the variant maintained a turnover rate comparable to the wild-type. Though, due to a shift of the K_m to higher substrate concentrations, it became visible, that the obtained velocity data did not obey Michealis-Menten kinetics. The shape of the curve was clearly sigmoidal (Figure 42), which suggested an allosteric or cooperative effect. Hence, data was fit with the Hill-equation (section 5.2.5; equation [12]) yielding a Hill coefficient of 3.0. This was very unexpected as we ruled out a cooperative effect between two MAT domains in a dimer due to distant active sites (133 Å between active site serines). Furthermore, a Hill coefficient greater than two cannot be obtained for a dimeric species and therefore this phenomenon can only be explained by an allosteric effect. This means that several molecules octanoyl-CoA (up to three) can bind to the MAT domain at allosteric sites effecting catalytic efficiency.

2.6 Site-specific Labelling of Multidomain Enzymes

Modifying proteins at specific sites can be a very powerful tool as the attachment of small organic molecules can shed light on various aspects on the molecular level. For example, the method facilitates to perform direct structure-function studies of physiological processes by specific crosslinking or to obtain spatial and time-resolved information about interaction partners by fluorecscence spectroscopy. The latter aspect inspired us to utilize site-specific labelling to analyze the interaction or relative spatial position of inherent domains in megaenzymes. It has been shown by several studies[31, 54, 134] that tremendous flexibility is an inherent feature of animal FAS and further that this

property is essential for catalytic activity. As this aspect forms also the basis for other classes of megaenzymes, as e.g. modular PKSs and NRPSs, and as currently hardly any information on protein dynamics can be found in literature, we decided to initiate a bold project in collaboration with Prof. Heilemann in 2014. Our goal in this approach was to use amber codon suppression to introduce handles site-specifically for the attachment of two fluorophores, which shall provide spatial and time-resolved information on individual enzymes in single-molecule fluorescence resonance energy transfer (FRET).

Amber codon suppression is a fantastic method, which has emerged to a very high level in the recent ten years. This technique expands the genetic code by reprogramming the rather rarely used amber (TAG) stop codon to encode new amino acids, which are not among the 20 commonly occurring proteinogenic ones. For this purpose, basically three compounds are required: a tRNA that recognizes the amber codon, the noncanonical amino acid of choice to encode and an aminoacyl-tRNA-synthetase (aaRS), which very specifically catalyzes the connection of the respective noncanonical amino acid with the respective tRNA. The pair of tRNA and aaRS is called orthogonal set, as it must not crosstalk with the endogenous machinery of the host cell. Typically, such orthogonal pairs originate from archaea and three distinct sets are commonly in use, namely from *Methanosarcina mazei (M. mazei), Methanosarcina bakeri* (*M. bakeri*) and *Methanocaldococcus jannaschii* (*M. jannaschii*).

We have comprised the project by cloning the two most promising vector systems in literature from the Schultz laboratory, namely pUltra and pEVOL (**Figure 43**).[135, 136] The practical work has been conducted with the collaboration of Christina Heil starting with her master thesis under my supervision,[137] which was extended into a PhD project. The application of this technique on mFAS was started from scratch to generate an own toolbox of materials and methods. In a project together with Christina Heil (with equally shared contributions), we therefore decided to clone a variety of different tRNA/aaRS pairs into both vector systems, to finally screen for the most sufficient orthogonal set. The cloning strategy and a comprehensive screening of the respective suppression efficiencies can be found in our shared manuscript (submitted, published at bioRχiv).[138] We focused on suppressor tRNA/aaRS pairs, which are specific for "clickable" amino acids, especially from the latest generation allowing copper-free click chemistry, and for acetyllysine, which is the product of the respective posttranslational modification found in animal FASs.
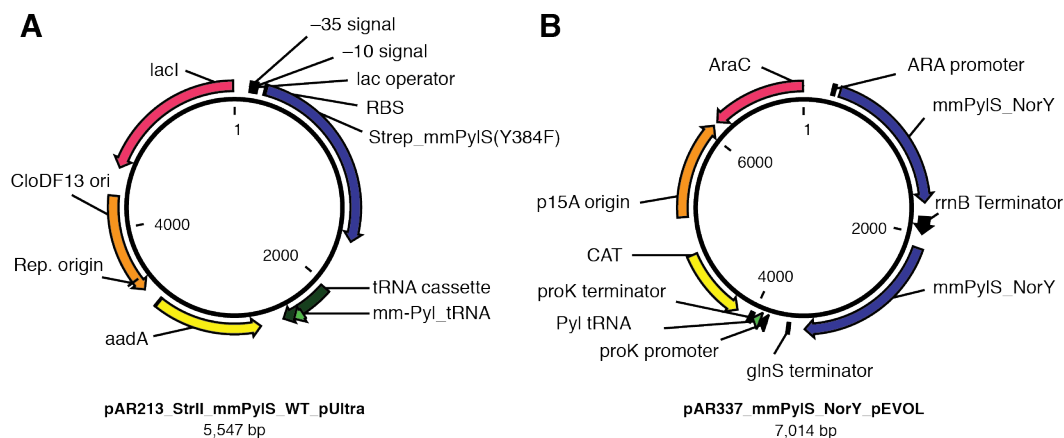
Figure 43: Representative vector maps of plasmids for amber codon suppression. Two vector systems, namely the pUltra- and pEVOL-system, were cloned utilizing information from the given references.[135, 136]

Besides plasmids encoding the orthogonal pairs, we had to find sufficient access to the modified amino acids of interest. At first, some amino acids were bought, but due to tremendous costs (one gram often costs between 200 and 1000 euro) or unavailability, we switched over to synthesize select compounds by ourselves. For this purpose, Jan Gajewski and me equipped an organic synthesis lab and established initial chemical routes with the help of three bachelor students Marija Trbljanic, Dietmar Herzberg and Bjarne Goebel. Especially the work of Bjarne Goebel shall be highlighted, who continued as a HIWI and master student,[139] and helped to establish the synthesis of seven noncanonical amino acids (**Figure 44**) in gram scale over the past years (section 4.2.5–4.2.23). Preparation of noncanonical amino acids in our own lab reduced the price significantly (ca. 15 € per gram for norbornene derivatives) and allowed the application in large-scale (liters) cell cultures.
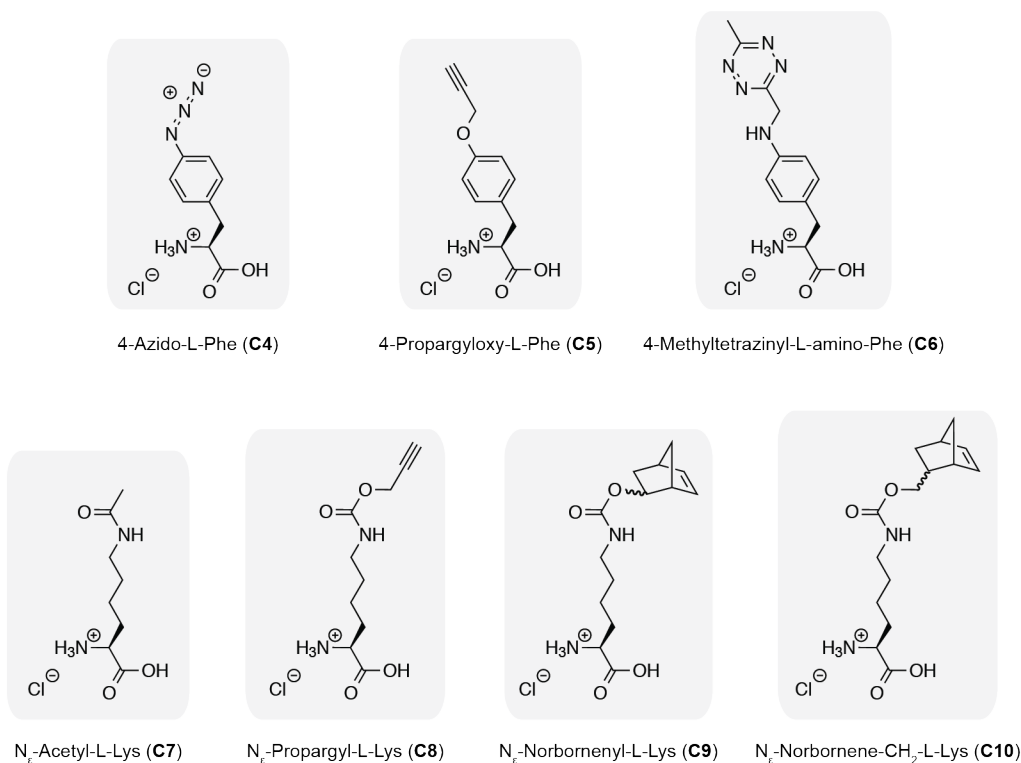
Figure 44: Compilation of noncanonical, mainly clickable amino acids, for which synthetic routes were established. All amino acids were synthesized as hydrochlorides. Common names are given below the chemical structures.

Comprehensive screening using an ACP-GFP fusion protein by Heil and students identified azido phenylalanine (**C4**) with suppressor pair tRNA/TyrRS (pAC^U_AzPhe) from *M. jannaschii* and norbornene lysine (**C10**) with suppressor pair tRNA/PylRS (pAC^U_NorLys) from *M. mazei* as the most sufficient systems. Although typical protein yields dropped significantly to 35% ± 8% (**C4**; pAC^U_AzPhe) and 18% ± 5% (C10; pAC^U_NorLys) of the wild-type, both approaches reproducibly provided sufficient amounts of modified proteins. By this means, we have achieved introduction of noncanonical amino acids into constructs: ACP, ACP-GFP, KS-MAT, processing part of mFAS, mFAS, PikAIII and DEBS 3.

As our constructs turned out to be very sensitive to copper, we focused on establishing copper-free click chemistry (**Figure 45A**). The norbornene moiety of residue **C10** reacted with tetrazine derivatives (commercially available Met-Tet-Cy5) in an inverse electron-demand Diels-Alder cycloaddition at mild conditions (RT and buffer of choice). The azide group of residue **C4** on the other hand was able to react with bicyclononyne derivatives (commercially available BCN-Atto647N) in a strain-promoted 1,3 dipolar Huisgen cycloaddition at standard conditions. Depending on the respective position of the noncanonical residue in the protein, which unexpectedly had an effect on click

efficiency, we could obtain a degree of labelling (determined by UV/Vis spectrometry) up to 37 % and 63 % using norbornene or azide mediated click chemistry, respectively. As a proof of principle, we expressed KS-MAT, PikAIII and DEBS 3 with norbornene residue (**C10**) and attached the fluorophor Met-Tet-Cy5, which is shown in **Figure 45B**.
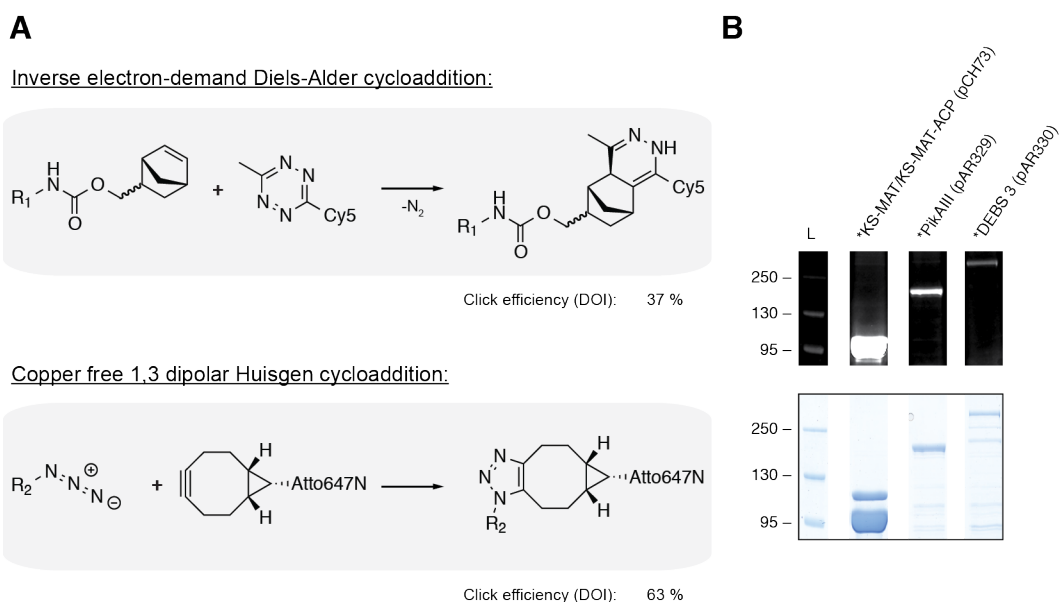
Figure 45: Copper free click chemistry. (A) Reaction schemes of utilized copper-free click reactions. Norbornene moieties react with tetrazine derivatives in an inverse electron-demand Diels-Alder cycloaddition, whereas azide groups react with bicyclononynyl moieties in 1,3 dipolar Huisgen cycloadditions. Rs indicate the attachment at the protein of interest. Typical click efficiencies are given below the reaction schemes, but optimization of the click conditions is still ongoing. (B) In-gel fluorescence and Commassie-stained SDS-PAGE of select multidomain constructs after click-reaction with Met-Tet-Cy5. * indicates the incorporation of **C10** via amber codon suppression. The construct *KS-MAT/KS-MAT-ACP refers to a heterodimer created *in vivo* by bicistronic expression.

In summary, we have successfully established the amber codon suppression system in our laboratory with synthetic routes for select noncanonical amino acids. In this way, we could introduce clickable residues into FAS and PKS constructs of up to 331 kDa and label them site-specifically with fluorophores. Although the degree of labelling should be further optimized, we can already generate double labelled samples utilizing the very efficient enzymatic attachment of fluorophores via Sfp at the ACP facilitating single-molecule FRET studies. Furthermore, the system can be adapted to attach spin labels, which can also provide spatial information via electron paramagnetic resonance (EPR) studies.

Chapter III: Discussion and Outlook

The essential biosynthesis of fatty acids in animals is accomplished by a megaenzyme, termed animal fatty acid synthase (FAS), which adopts a homodimeric, 'X-shaped' structure.[31] Besides highly conserved catalytic functionalities, also phylogenetic evidence has suggested that the animal *FASN* genes represent an evolutionary subset of type I *PKS* genes; i.e. the animal FAS is evolutionary nested within the type I PKS family. In fact, until the recent discovery of a pigment producing type I PKS (echinochrome PKS 2) in sea urchins, animal FAS used to be the only type I PKS occurring in animals and still remains the only ubiquitously found.[140] Being part of the primary metabolism, FASs are essential for their host in contrast to PKSs, which typically provide an advantage in a challenging environment. This special role has led to a distinct evolutionary pressure for animal FASs yielding a catalytically very efficient machinery. A FAS can produce palmitate in less than a second, whereas the type I modular PKS DEBS requires ~2 minutes despite the fact that more enzyme-catalyzed reactions are necessary for product formation.[141, 142] Furthermore, animal FAS fulfills different tasks in specialized tissues, e.g. synthesis of short FAs or branched FAs (section 1.3), which indicates the potential of the FAS fold for versatile usage.

In the following, the catalytic efficiency of FAS and the individual domains MAT and KS will be discussed, particularly with respect to the potential for biotechnological application. We will address the three stages of engineering animal FAS towards modular PKSs, defined in the aim of the thesis (section 1.7), which will be addressed in three separated sections. This may establish animal FAS as a practical model system to mechanistically study PKSs, which are often difficult to handle *in vitro*. Furthermore, it shall be discussed, which of our findings may actually be utilized in hybrid FAS/PKS systems to potentially create superior biosynthetic catalysts. Due to the fact that also appropriate enzyme assays needed to be established within the scope of this thesis to elucidate desired enzyme properties, compiled data will be critically evaluated regarding their reliability within these sections.

3.1   Establishment of Murine FAS as a Model System for PKSs

Despite the fact that animal FASs are well studied enzymes with also increasing relevance for widespread diseases in humans, like obesity, diabetes and cancer,[75, 76] heterologous expression in *E. coli* has not yet been established sufficiently. We were not able to reproduce data from the Wakil laboratory expressing human FAS with a N-terminal MBP fusion protein, which confirmed a statement of Pappenberger *et al.*, who also struggled in a more recent study.[64, 74] A possible explanation for the discrepancy might be, that Jayakumar *et al.* most likely utilized a fermenter (15 L cell culture was written in the manuscript) and did not state on any protein yields per litre cell culture. We can also not rule out, that trace amounts of active MBP-fused hFAS can be obtained from *E. coli*, but all attempts to improve expression failed. This included a comprehensive screening of culturing conditions, attaching different fusion proteins and co-expressing chaperones from bacteria as well as archaea. After eight months of trying, we had to accept that hFAS does not fulfil the criteria to serve as a model system (section 2.1) and we started to search for alternatives.

Very remarkably and unexpectedly though, murine FAS, which shares 81 % sequence identity with human FAS when aligned with the program BLAST,[143] turned out to perform very differently. The protein was sufficiently expressed in *E. coli* with typical yields of 1.5–2 mg purified protein out of 1 L cell culture without the necessity of gene codon optimization or the fusion of domains to increase solubility. The expression is very stable and reproducible and the applied tandem affinity purification strategy ensures the isolation of full-length protein with little appearance of nicked protein (**Figure 15**). After incubation at 37 °C as discovered in the Smith laboratory,[52] mFAS appeared almost exclusively in the active dimeric oligomeric state and showed very little tendency to aggregate. The enzyme is not modified posttranslationally by *E. coli* (contrary to the report of Jayakumar *et al.*), which facilitates to decide between the apo- and holo-enzyme by co-expressing Sfp yielding quantitative modification when desired (section 2.1.2). The opportunity to easily obtain apo-protein is advantageous as Sfp can also be used *in vitro* to attach different compounds to ACP, delivered as CoA-esters, including specific cross-linkers, fluorophores and natural or non-canonical substrates of FA synthesis.[83, 144] Inspired by the enzymatic attachment of one fluorophore at ACP, we have established the amber codon suppression technique to place a second fluorophore site-specifically in mFAS to facilitate FRET-studies (section 2.6). The labelling strategy has been successfully

worked out within this thesis and within the scope of three supervised master theses of Christina Heil, Maria Dell and Bjarne Goebel, where the first master thesis was expanded into a PhD thesis addressing the dynamics of mFAS.

As all parameters regarding protein quality seemed to be appropriate, we have analyzed the functionality of the heterologous expressed enzyme by the well-known NADPH consumption assay. Our values for mFAS were 564 ± 31 nmol/min/mg protein and 433 nmol/min/mg protein in the supervised master thesis of Aaron Himmler. These values agree very well with the value 462 nmol/min/mg protein for human FAS isolated from HepG2 cells,[145] but are only about a third of a value for chicken FAS (1478 nmol/min/mg protein protein) from liver in the same study. Typically, reported turnover rates for native FASs are in between 700-2500 nmol/min/mg protein at temperatures between 25–37 °C from different laboratories and isolated from different species.[42, 48] Besides relatively high fluctuations of these values in the literature, mFAS heterologously expressed in *E. coli* seems to be a factor 2-3 slower than enzymes isolated from tissue. How can this discrepancy be explained, when the protein seems to be integer regarding stability, oligomeric state and phosphopantetheinylation? It might indicate that further posttranslational modifications exist increasing catalytic activity, which are absent when expressed in *E. coli*. Indeed, several acetylation and phosphorylation sites were found in human FAS by proteomic analysis.[146, 147] Interestingly, two phosphoserines (S63 and S207) and three Nε-acetyllysines (K59, K70 and K298) were also identified in mFAS by similarity and independent studies, which are located in the KS domain. It would be the ideal target for regulatory modifications *in vivo* to adjust overall catalytic efficiency to the needs of the host. Unfortunately, this aspect could not be addressed within the scope of this thesis, but the identification of such activating modifications would be of high value for biosynthetic application and should be considered for future projects.

To gain a deeper insight into the working mode of mFAS, we have conducted an enzyme kinetic investigation according to a study of Cox *et al*..[44] Such analysis allows to extract the Michaelis constants for all three substrates and provides an overall k_{cat} value, which cannot be obtained experimentally due to the inhibitory effect among acetyl- and malonyl-CoA (section 1.3.1). All extracted kinetic parameters agreed relatively well between our analysis (section 2.1.2) and the study on FAS, isolated from chicken liver, except a factor of almost eight between the k_{cat} values (mFAS $k_{cat} = 3.1$ s^{-1} *vs* chicken FAS $k_{cat} = 23$ s^{-1}), which further strengthens the hypothesis of a less efficient KS domain due to heterologous expression in *E. coli*.[44] The obtained kinetic parameters can now be

inserted into equation [2] to simulate the behaviour of mFAS at different substrate concentrations, which has been nicely demonstrated within the master thesis of Aaron Himmler.[84]

Further, we have tested the capability of mFAS to be primed with different loading substrates than acetyl-CoA. Interestingly, using butyryl-CoA slightly increased the turnover rate by 7 %, which confirms previous reports on mammalian FASs.[42] Also other CoA-esters, which are activated intermediates of FA synthesis, were accepted (section 2.1.2; **Figure 16**), but phenylacetyl-CoA shall be highlighted (14 % turnover rate compared to acetyl-CoA usage), as this is a noncannonical substrate leading to phenyl derivatized FAs. In the absence of NADPH, mFAS produces TAL in two non-reductive condensation steps with subsequent cyclisation to form the lactone. We determined a TAL production rate of 23 ± 2.0 nmol/min/mg protein when acetyl- and malonyl-CoA were used.

Comparing these values to typical PKSs facilitates to estimate a biosynthetic potential of mFAS. A relatively similar enzyme in regards to the conceptional synthesis is the iterative PKS methylsalicylic acid synthase (MSAS) from *Penicillium patulum*. This enzyme produces 6-methylsalicyclic acid (MSA) in the presence of NADPH from acetyl- and malonyl-CoA and was characterized *in vitro* in the laboratory of Feodor Lynen.[148] To enable a good comparison values were converted into a consumption rate of malonyl-CoA per minute and mg protein. In terms of TAL production mFAS is almost twice as fast as MSAS (mFAS: 46.2 ± 0.7 *vs* MSAS: 25 nmol malonyl-CoA/min/mg;) and mFAS is slightly more efficient in FA synthesis than MSAS is in MSA production (mFAS: 282 ± 16 *vs* MSAS: 210 nmol malonyl-CoA/min/mg;). Although mFAS is slightly faster in our analysis, all values agree remarkably well, which is surprising, as the intermediate, which is cyclized to TAL, is a native intermediate during MSA synthesis, whereas it does not occur during FA synthesis. Hence, the ratio between TAL and FA production seems to be to high for mFAS indicating either a KS domain with altered properties (see above) or an influence of the protein concentration on catalytic efficiency (20 nM *vs* 700 nM for FA *vs* TAL production) regardless the stabilization by additives (section 2.1.2). Another possible explanation might be that this discrepancy indicates a decelerated reduction, dehydration, reduction sequence in heterologously expressed mFAS. Hammes stated that there is no clear rate determining step during each cycle of chain growth and that the β-ketoacyl reductase reaction is relatively slow in animal FA synthesis (although our data on indivudal KR activity suggest sufficient reduction rates ($k_{cat} = 12.0 \pm 0.4$ s^{-1}), at least when *trans*-1-decalone is used as substrate (**Figure 20**)).[141] A profound comparison to modular PKSs

is very difficult, as it is very challenging to obtain turnover rates *in vitro*. Nevertheless, Lowry *et al.* succeeded in reconstitution of the whole DEBS system *in vitro* and assigned a maximum turnover rate of 1.1 min^{-1} for the production of 6-deoxyerythronolide B, which is ten-times slower than FA production of mFAS (11 ± 0.6 min^{-1}).

In summary, within this thesis, a sufficient expression and purification strategy for an animal FAS in *E. coli* with high protein quality was established for the first time, which has laid the foundation to exploit this fold to study the PKS family.

3.2 Specific Activities of Individual Domains

Substrate specificity and catalytic efficiency of individual domains remain decisive in the multidomain architecture, although they might play a minor role compared to the type II system as interaction partners are constrained within the architecture. The most important domains in the PKS family are KS and AT domains as they choose and condense select substrates, which determines the carbon scaffold of a polyketide compound. Hence, we have decided to focus on murine KS and MAT domains, very early within this thesis. Especially due to their extraordinariness in regards to modular PKSs, the KS domain of FAS is characterized by a narrow substrate tolerance, but high catalytic efficiency. On the other hand MAT is bifunctional with accepting loading and extender units plus features very fast transfer rates. Both domains were characterized in detail structurally and functionally, exploiting excellent expression properties of the KS-MAT didomain in *E. coli* (section 2.2).

Enzymes can be studied in many different ways, all of which can provide important information about the mechanism of action. Nevertheless, a potential application requires an in-depth knowledge about the enzyme function and catalytic efficiency. In this regard, the classical approach to investigate enzyme kinetics is still the method of choice. Hence we began our anaylsis with searching for an appropriate assay to address enzyme functions most conveniently, sensitively and quantitatively. After struggling for a year, we found the answer in the αKGDH assay (section 2.2.1.5), which has been established very carefully. This assay allows a continuous fluorometric readout of NADH, which is generated by the conversion of liberated CoA by the coupled enzymatic system. The aspect of continuity provides superior quantity and quality of data points at defined time steps in comparison to end point measurements like HPLC analysis and does not require further chemical processing of samples. By this means we were able to achieve acurate results also at low substrate concentrations, which revealed

deviations from standard Michaelis-Menten kinetics for the MAT and KS domain (see section 3.2.1 and section 3.2.3). Nevertheless, the assay is practically challenging and needs some experience to obtain reliable results due to many relatively fast non-enzyme-mediated reactions in the background. Although most of these issues can be handled with proper controls, the relatively fast hydrolysis rate of succinyl-CoA remains an intrinsic restriction of the assay, which requires the use of very clean CoA-esters as substrates and limits all investigations to a narrow measuring window.

However, utilizing the αKGDH assay in combination with structural data on different acyl-enzyme states deepened our understanding of KS and MAT from animal FAS, which will be discussed in detail in the following sections.

3.2.1 Specificity of the MAT domain

It is well known from literature that MAT of animal FAS accepts both acetyl- and malonyl-CoA as substrates, which is already apparent from the cycle of FA synthesis (Figure 6). Further evidence existed, that MAT is way more tolerant towards a variety of substrates, as animal FA synthesis has also been conducted with substrates as propionyl-, butyryl-, acetoacetyl-, phenylacetyl- and methylmalonyl-CoA.[37, 38, 42, 47] Nevertheless, besides values for acetyl- and malonyl-CoA, no kinetic parameters for the individual MAT domain were available for such noncanonical substrates.[51] Consequently, we made use of the αKGDH assay to determine apparent kinetic constants for twelve different substrates (Table 1) and to obtain absolute kinetic parameters for acetyl- and malonyl-CoA as well as the acceptor ACP (Figure 1). This comprehensive analysis revealed a remarkable promiscuity of the MAT domain accepting all tested compounds as substrates. Furthermore, although malonyl-CoA remains the preferred substrate with a specificity constant of 12.2×10^6 M^{-1} s^{-1}, other CoA-esters were transferred only by a factor of 1.5-3 less efficiently with the exception of palmitoyl-CoA, which discriminates a factor of approximately twenty. Such a level of substrate ambiguity of MAT was unexpected, but led us to form the concept of utilizing the MAT fold as a universal AT for modular PKSs, which will be discussed in section 3.5.1.

Although it is common knowledge, that the ratio of the kinetic constants k_{cat}/K_m, which is a measure of enzyme–transition state complementarity, is most suited to define the catalytic efficiency of an enzyme,[58] in case of MAT, also substrate concentration and availability *in vitro* as well as *in vivo* have to be considered. Typically, substrate concentrations are 1-2 orders of magnitude higher than the K_ms of MAT *in vitro*, when FA synthesis is conducted. At such high concentrations of substrate, the reaction rate is

proportional to the turnover rate ($v = k_{cat}[E]_0$), because the enzyme is saturated with substrate meaning that the k_{cat} parameter becomes most important. Under these circumstances some differences in efficiency of MAT become more pronounced, e.g. malonyl moieties are transferred four times faster than methylmalonyl groups. As long as substrate concentrations are low, and hence below the Michaelis constant, the reaction rate obeys the rule $v = \frac{k_{cat}}{K_m}[E]_0[S]$, which has to be considered, when variants are evaluated regarding their efficiency as demonstrated in **Figure 46**. This will be discussed in section 3.2.2.
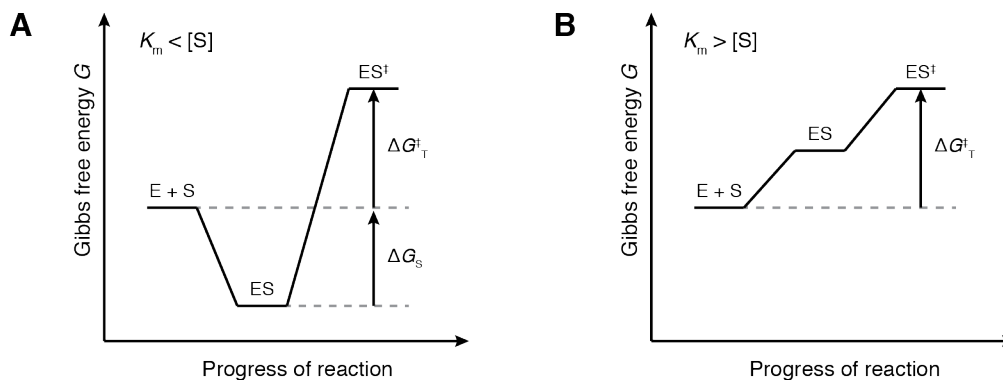
Figure 46: Two cases of enzyme evolution. In both cases the enzymes bind the transition state equally well, but in (A) the substrate is bound strongly, whereas in (B) the enzyme has evolved to bind the substrate weakly. Consequently, the activation energy in (A) is increased by the binding energy of the substrate in the Michaelis complex. Figure adapted from Fersht.[149]

How does the situation look *in vivo*? A general statement cannot be easily made due to a strong dependence of substrate concentrations on individual cell types or tissues and their respective metabolic state. Considering heterologous expression in *E. coli*, Takamura *et al.* reported concentrations of acetyl- and malonyl-CoA in the range of 20-600 µM and 4-90 µM, respectively.[150] This means, that typically substrate concentrations are indeed higher than the Michaelis constants leading to the proportional dependence of the reaction rate from k_{cat}. In specialized animal tissue, e.g. goose uropygial gland, sufficient methylmalonyl-CoA concentrations are guaranteed by an acetyl- and propionyl-CoA carboxylase with a subsequent decarboxylation of malonyl-CoA by specific decarboxylases.[151] Unfortunately, reliable *in vivo* concentrations of substrates are very difficult to find in literature, but such *in vivo* concentrations of substrates determine the required properties of engineered ATs, especially in regard to the introduction of noncanonical CoA-esters by feeding or by establishment of artificial pathways.[63]

Last but not least, we have also analyzed the MAT-mediated hydrolysis rate for all tested substrates, which is a well-known side reaction of transacylation (**Figure 8**). This was a very important aspect as increased hydrolysis rates for non-native substrates were postulated as a "proof reading" mechanism in ATs from PKSs determining their substrate specificities.[24, 118] We can definitely rule out such a "proof reading" role in MAT as enzyme-catalyzed hydrolysis rates appeared to be 10^2–10^3 orders of magnitude slower than transacylation rates. Nevertheless, we can conclude from our analysis, that it might be misleading to solely rely on hydrolysis kinetics as they are not necessarily proportional to transfer rates as e.g. crotonyl- and succinyl-CoA were hydrolyzed up to three times faster than acetyl- or malonyl-CoA. Although protection of the acyl enzyme state from nucleophilic attack by water remains an intrinsic property of every individual AT and might well be less pronounced in PKSs, a kinetic self-editing mechanism using CoA as a nucleophile seems much more likely. Considering high free CoA concentrations in the cell (1.4 mM in exponentially growing *E. coli*),[152] with relatively low K_ms in the single digit μM range, a very fast exchange frequency (probably even faster than the absolute turnover rate of 119 s^{-1} for malonyl-CoA) of native acyl moieties between CoA and the AT is guaranteed. Occasional acyl-enzyme states with less preferred substrates are than rather unloaded from the MAT domain than transferred to the acceptor ACP, which possesses a relatively high K_m (ca. 250 μM in murine FAS) and is predominantly occupied by polyketide intermediates. A kinetic self-editing mechanism of the whole animal FAS by CoA during FA synthesis has been already postulated in the 1980s making CoA essentially a co-substrate during synthesis.[153] It was shown that FA synthesis is inhibited when liberated CoA is withdrawn from the equilibrium *in vitro*, as it becomes impossible to unload acetyl moieties from the ACP domain during elongation cycles.

3.2.2   Comment on the Acceptance of Succinyl-CoA as a Substrate

Relatively early during carefully establishing the αKGDH assay in our laboratory, it became apparent that the MAT domain must be capable of transferring succinyl moieties donated from a CoA-ester. Indeed, succinyl-CoA, which is a C4-dicarbonic acid derived thioester that is well known for its central role in the citric acid cycle, turned out to be a relatively good substrate with a fast turnover rate (k_{cat} = 13.4 ± 0.6 s^{-1}), when transferred to ACP. The acceptance of succinyl-CoA was unexpected, since never mentioned in literature so far, and indicates a high plasticity of the active site of MAT being able to adopt to versatile functional groups, which will be discussed in the next section 3.2.2. Normally, succinyl-CoA occurs compartment in mitochondria, and might therefore not

be loaded to FAS under physiological conditions. But the capability of a MAT-mediated succinyl transfer to ACP (succinyl-CoA concentration in *E. coli* is 230 µM)[152] may well be biotechnologically interesting, in case the KS domain could be engineered to accept succinyl groups as starter units for a single reductive chain elongation with malonyl-CoA. This may hold for the prospect of FAS-mediated synthesis of adipic acid, which is the most important dicarboxylic acid from an industrial perspective being produced in a megaton scale per year due to the requirement as a precursor for nylon production.[154, 155]

3.2.1   Comment on the Directed Mutagenesis of MAT

Having revealed a general promiscuity of MAT for a versatile set of substrates, we were curious, whether one can specify the substrate spectrum by introducing single mutations in the active site. This is a converse approach to typical studies in literature, where specific ATs of PKSs are individually tried to be relaxed in their specificity by mutagenesis to accept other, often noncanonical substrates.[156, 157] It is generally believed, that during evolution enzymes with promiscuous function have acquired higher specificity and activity by selective pressure.[130] Hence, protein engineering in laboratories is commonly conducted with such enzymes, which are evolved to improved or altered catalytic functions.[131, 132] Inspired by this idea, we have performed a screening of 42 variants to test whether the MAT is a suited domain for a rational approach to adapt the specificity (section 2.5.1). The conclusion of our analysis is that indeed, already single mutations can alter the specificity dramatically, which shows the typical characteristic of plasticity (ability of proteins to alter their function with a small number of amino acid substitutions). Especially, generation of a specific AT for loading units can be obtained relatively easy by mutating arginine 606 to alanine. While this fact has been known since a study of Rangan *et al*. in 1997,[51] we have further refined the result by starting to investigate the effect of additional mutations. Furthermore, also the specificity for extender substrates can be increased, which has been shown for variant F553N and L582M having a pronounced preference for malonyl-CoA. Introducing a mutation in the 'YASH' motif (F682A) abolished the difference between malonyl- and methylmalonyl-CoA leading to a reduced turnover rate of acetyl- and malonyl-CoA and increased turnover rate of methylmalonyl-CoA. Hence, we have already identified five plasticity residues F553, L582, R606, L680 and F682, which are important to define the specificity of the domain and induce significant changes to the substrate spectrum when altered.

Aiming at evolving the MAT domain to accept loading *or* extender units, more plastic residues have to be identified in a next step to further prevent the transfer of

acetyl-moieties, which turned out to be more difficult than to abolish the transfer of malonyl moieties. This will further deepen our understanding of the molecular mechanism for substrate specificity and may enlarge our library of promising mutants. These shall be tested on noncanonical CoA-esters as provided by the laboratory of Tobias Erb to find sufficient AT domains for such compounds, which have not coevolved with ATs in nature and hence lack the existence of a natural efficient catalyst.[63].

3.2.2 Establishing a Model to Explain Substrate Ambiguity of the MAT domain

Having analyzed the substrate ambiguity in the MAT domain, the question remains of what the molecular basis for this behaviour is and what aspect distinguishes the MAT domain from typical specific ATs. To approach this question, one should start with summarizing observations and facts about specific, well-characterized ATs. Extender-ATs, like e.g. FabD or introduced extender ATs of DEBS or AVES (section 1.2), share one pronounced characteristic of possessing a conserved arginine residue at the bottom of the active site, which acts as a chelator for the carboxyl group of transferred elongating units. The guanidinium group of this residue is embedded within a H-bond network, which crosses through major parts of the α/β-hydrolase subdomain (**Figure 34**). This stabilization stretches the guanidinium group in a defined position allowing a bidentate complexation of extender units, which places the carbonyl group of the thioester readily for nucleophilic attack of the active site's serine. In the absence of this arginine residue catalytic activity in FabD of *E*. coli is dramatically reduced by 277-fold, which has recently been shown.[158] A closer look at crystal structures of such ATs (for a compilation of pdb codes, see section 2.4) reveals an open active site, comparable to the conformation of chain A in the octanoyl soaked crystal (**Figure 36B**), being created by the relative positioning of the two subdomains. Whether this cleft is persistent throughout all conformational states is unknown, but FabD has only been found yet in the open state independent of bound substrates (pdb: 1mla, 2g1h, 2g2o, 2g2y, 2g2z). Whereas, typical loading AT domains do not contain a respective arginine residue in the active site and neither do acetyltransferases in yeast FASs (pdb: 3hmj)[28] – instead, bulky hydrophobic residues like tryptophan (AT0$_{AVES}$) or isoleucine (fungal FAS) are present. Furthermore, revealed by structural information (pdb: 4rl1), the active site of AT0$_{AVES}$ is shaped as a hydrophobic tunnel and hence present in a closed conformation of the active site. The structural complementary of the enzyme to hydrophobic acyl chains may facilitate appropriate positioning of the substrate relative to the catalytic machinery, when no

functional groups are present that allow the positioning by hydrogen bonding or electrostatic forces.

We propose that MAT combines characteristics of both fields, namely the existence of the conserved arginine residue in combination with a closed, rather hydrophobic active site tunnel. In order to be able to transfer chemically and sterically diverse acyl moieties with similar efficiencies, the enzyme has to sample distinct conformations that allow stabilizing the different transition states. This requires a pronounced plasticity to fulfill the concept of conformational selectivity.[159] Indeed, we have found various relative positions of the two subdomains in different chains and different acyl bound states. Consequently, the active site appeared in different shapes (open and closed, see **Figure 36**) supporting the hypothesis. This can also be shown by anisotropic movements of the ferredoxin-like subdomain implying some kind of rotational opening and closing of the active site cleft (**Figure 47**).
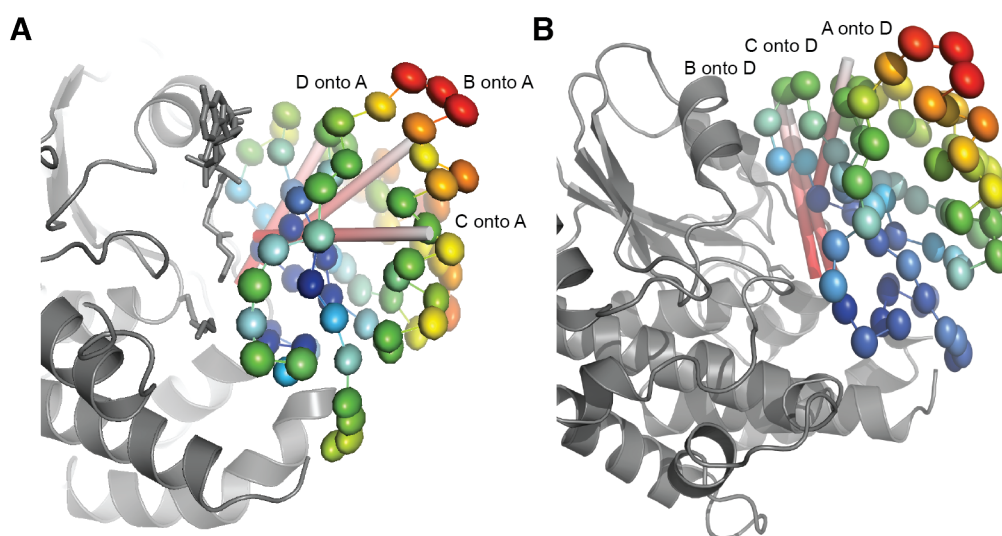
Figure 47: Anisotropic movement of C-alpha atoms in the ferredoxin-like subdomain. The anisotropic movement, as derived from the TLS tensors, is depicted by thermal ellipsoids colored by B-values (blue – low values; and red – high values) for the ferredoxin-like fold (616-684) region of chain A (A) and chain D (B). Rotation axes derived from superposition of ferredoxin-like fold (616-684) of chains B (A), C (B), and D (C) onto A (D) are shown in red-white cylindrical bars. For orientation, the rest of the KS-MAT chain is shown in grey.

Furthermore, we intended to find out, where the plasticity originates from or rather, which subdomains participate to this property. First, we have thought that the small ferredoxin-like subdomain is less stabilized and hence is itself very flexible. This, we could rule out due to the fact that independent of the active site appearance both subdomains align very well with only minor deviations and further the ferredoxin-like

domain of MAT resembles also homologous subdomains of specific ATs (**Figure 39**). Accordingly, only the linkers between the subdomains show remarkable flexibility, which has been demonstrated by structural superposition and by a torsion angle analysis. These linkers shape the acyl binding pocket and are hence very important for substrate specificity, which explains the impact of the 'YASH' motif in linker 2.

Although the conserved arginine residue is also present in MAT, we postulate that this residue plays a slightly different role in MAT. By comparing different crystal structures, we have found R606 in three different rotameric states dependent on the bound acyl moiety (**Figure S8**). Contrary to specific extender ATs, MAT does not possess an extended H-bond network to retain the guanidinium group in one position, mainly due to exchange of a conserved glutamine residue to phenylalanine (F553). Further, the catalytic efficiency of MAT is altered by mutating R606, but the transfer of malonyl moieties is by no means abolished. A possible role of R606 in MAT may be to compensate for the loss of a hydration shell, rather than placing extender units properly by complexation.

In summary, the MAT domain shows remarkable conformational plasticity mainly due to versatile linkage of the two subdomains, which might be an explanation for substrate promiscuity. This may allow the enzyme to adapt distinct conformational states to stabilze chemically diverse transition states originating from different acyl units. Appropriate positioning of the thioester group relative to the catalytic machinery may be accomplished by structural complementary of the enzyme and by binding the nucleobase of the CoA moiety within a pocket at the surface of the protein (**Figure 37**). The equilibrium of conformational states can be influenced by allosteric binding of unsual substrates as octanoyl-CoA, which was shown by a sigmoidal initial velocity curve (**Figure 42**) that can be fit with the Hill equation (equation [12]). Certainly, more experiments are required to confirm this model including crystal structures of select variants and further mutagenesis studies focusing on residues in the linker regions between the subdomains and the CoA binding pocket on the surface.

3.2.3   Specificity of the KS Domain

As the KS domain is the most important domain in FAS/PKS systems and still mechanistically not fully understood (section 1.5), we aimed at utilizing our knowledge about transthioesterification reactions to be able to also investigate such domains. Within the scope of this thesis, but in more detail within the master thesis of Aaron Himmler, we could demonstrate that the αKGDH assay can be used to analyze the first step in KS catalysis already reflecting the substrate specificity of the domain (**Figure 32**).

Furthermore, by carefully fitting the obtained velocity data, we could reveal a cooperative behaviour of the KS dimer, which has not been identified yet. Due to the fact that the acyl binding pockets of the active sites of the protomers merge via the dimeric interface, it might not be unexpected that loading of an acyl moiety in one active site influences kinetics of the other. It was confirmed by structural data that the domain is very rigid (in contrast to MAT) and only a short loop including the "gatekeeper" residue F395 moves while loading an acyl moiety. It remains to be shown to what extend altering of this residue may lead to increased tolerance for functional groups. As shown within this thesis, the set-up of a stable enzymatic assaying system is crucial for receiving reliable data that can analyze catalytic functions and guide engineering. On the basis of this example, I strongy recommend that before starting mutational studies on the KS domains, analytics need to be initially optimized. I see two major aspects for exploiting the αKGDH assays for analysis and design of the KS domain. First, other acyl donors than CoA-ester should be tested, as e.g. pantetheine esters, would provide better kinetic parameters, which are closer to the native ones. It has been initially tested that pantetheine is also accepted from the αKGDH complex as substrate, but whether kinetic parameters are sufficient (K_m at least in the single digit μM range, compare K_m for CoA of <0.1 μM)[160] to facilitate the utilization of this enzyme as a coupled system requires careful establishment. Furthermore, the option of using malonyl-ACP as an acceptor should be tested as this would allow to use the same set-up to measure condensation rates.

3.3   Modularity of the Animal FAS Fold

Within this thesis, we have addressed the question whether the animal FAS fold fulfils the requirements of a modular system. Such modularity is defined as a multi-component system that can be divided into smaller subsystems, which interact with each other and can be predictably interchanged with every present component retaining its individual functionality.[86] The inherent modularity of megasynthases is responsible for the versatility in product synthesis, and is considered as the key property that makes megasynthases highly relevant in biotechnological applications. In modular PKSs the absence or presence of modifying domains within a module controls the degree of processing in every elongating unit determining the functional groups present in the final product (section 1.2).

As animal FAS and iterative as well as modular PKSs are thought to share a common ancestor,[14] animal FAS can be seen as PKS equipped with a complete set of all

functional domains, which typically occur in PKSs. As such, the core fold of FAS, i.e. FAS without the TE domain, correlates with a fully reducing δ-module in modular PKSs. Acting on the assumption that the core fold of FAS exemplifies a common ancestory fold with PKSs, it shall have similar properties and as such show similar intrinsic modularity as PKS systems. We could demonstrate within this thesis and the corresponding master thesis of David Drexler that the FAS fold can be utilized to construct other partially reducing modules, which resemble β- and γ-modules in PKSs. All alterations were tolerated quite well yielding soluble protein with essentially unaffected tendency to aggregate (section 2.2.1). The fitness of the folds was judged by the yield of purified protein (this measure has been used in literature before, see references),[161, 162] behavior in SEC and stability of the constructs in thermal shift assays. Modularity can be seen manifested when deleting the ER domains, as it led to an increase of the monomeric fraction in SEC. The ER domain contributes with an area of 1600 Å^2 to the dimeric interface of FAS,[31] and hence a reduction of this dimerization area should shift the equilibrium from the dimeric to the monomeric state. As both oligomeric states are in a fine tuned equilibrium *in vitro*,[52] further experiments are required to quantify this effect, which could be accomplished by kinetic assembly or disassembly studies via HPLC-MALS analysis. It would be very interesting to find out, whether dimeric interfaces can be summed up in a modular fashion to determine the overall stability of a megasynthase. Remarkably, nature has already evolved a solution for modules of PKSs that do not contain ER domains, as small structural, dimeric domains, namely dimerization elements (DE) and docking domains (DD), are installed before and/or after the processing part to meet the demand.[163, 164]

In the next step of our analysis, we intended to investigate the integrity of the reducing PKS-like modules by functional assays. For this purpose, we decided to study partial KR activity as a measure for the modular character of the processing part, as this is the only component, which remains present in all reducing modules. Using a *trans*-1-decalone assay offered the opportunity of a very simple readout, which only involves the individual catalytic activity of the active site of the KR domain. The outcome of our analysis assigned functionality in all modular arrangements as long as the structural ΨME domain was present. However, obtained kinetic parameters revealed a strong impact of the DH domain on catalytic efficiency of the KR domain, which might be explained by a stabilizing effect of α-helix 989-999 on the Rossmann fold within KR. Furthermore, deletion of the TE domain (δ-module), which does not have a direct structural interface with the KR domain, unexpectedly increased the Michaelis constant of the *trans*-1-decalone reduction by 50 %, whereas deletion of the ER domain (γ-module) had no

further effect. As k_{cat} in the corresponding reactions was only slightly altered this might be explained by altered substrate binding affinity to the active site. A close look at the structure of the porcine FAS reveals that KR plays an extraordinary role in structuring the processing part in FAS being a docking site for all other modifying domains (**Figure 19**). To what extent this is also true for modular polyketide synthases is still under debate, but remains a key object of current research in the polyketide field. High impact publications on the structural organization of modular PKSs are still rare, and somehow contradictory, either due to indeed non-uniform architectures in different PKS systems or due to inadequate quality of protein samples.[70, 165-167]

Generation of an α-module can be reduced to the design of the linker that needs to be installed for covalently bridging an ACP to the condensing part. While this task seems to be straightforward, our analysis revealed that it is important where to attach the anchor point at the KS domain for the introduced linker. The idea behind the design of a first generation of α-modules was to utilize linker sequences from AVES starting right after the conserved motif in the post-AT linker (**Figure 48**), which could have led to short total linker lengths. These constructs led to lowered protein yields, decreased dimeric fractions and inactive protein, judged by a TAL assay. Interestingly, including further eleven amino acids of mFAS in the linker design, i.e. elongating the linker downstream of the post-AT motif by eleven more amino acids, led to a regain of activity. From this observation, we concluded that also this stretch of residues is important for normal KS activity in mFAS, which will be further discussed in section 3.5.2. For our purposes, the TAL production assay was a very powerful tool to demonstrate that all three domains possessed individual functionality in generated α-modules and revealed unexpected issues, when linkers were designed insufficiently. Furthermore, all the three domains have to interact with each other to produce TAL and hence this simple assay proved a functional interplay of all domains within the α-module.

A future goal will be to show a productive interplay of all domains present within our reducing PKS-like modules. On a simple level the TAL assay might again provide minimal information, as it proves that the ACP domain can efficiently reach the MAT and KS domain despite different structural organisations of the processing part. More informative would be to analyze the CoA-ester product spectrum, which should contain all functional groups generated by the present domains. For this purpose, we would utilize the ability of MAT to transfer a variety of acyl moieties (section 3.2.1) offering the opportunity to also unload polyketide-like intermediates. A similar aproach was conducted by Smith *et al*. who performed a comprehensive product analysis with rat FAS bearing a

variety of functional knockouts.[168] They are nicely demonstrating the outcome of a complex mixture of all kinds of compounds due to thermodynamic equilibriums in modifying reactions, i.e. domains like DH catalyze the back reaction when the product is not withdrawn from the equilibrium. Hence, this analysis would only give qualitative results, which would limit the significance for judging the effects of our engineering. Such modules are best studied within an assembly line of further downstream modules, which withdraw intermediate products of individual modules from the equilibrium by further elongation. The first attempt to create such folds beyond a singular module will be discussed in the next section.

Last but not least, to complete our analysis within the mFAS fold, modularity was also tested and confirmed within the condensing part. The MAT domain can be released from the KS-MAT didomain architecture retaining wild-type activity in two different arrangements. The construct LD-MAT resembles related ATs from *trans*-acting PKSs and the construct of a freestanding MAT domain with a C-terminal fusion protein mimics loading ATs in modular PKSs (**Figure S4**).

3.4 Bimodular Architecture and Impacts on Catalysis

Considering the core FAS fold as a δ-module in modular PKSs implies that it has to be possible to attach further modules either at the N- or at the C-terminus of FAS. Within this thesis, we have realized the simplest bimodular architecture using mFAS as a module 1 and attaching loading didomains containing AT and ACP domains (loading modules are termed module 0). Whereas we have failed to sufficiently express mFAS with "real" loading didomains from modular PKSs (**Figure 24**), it was possible to purify bimodular constructs utilizing artificial loading didomains built from the mFAS fold (**Figure S6**). Again, protein quality was confirmed by checking the tendency to aggregate, by peak patterns in SEC and by NADPH consumption assays. Interestingly, both bimodular constructs possessed decreased overall NADPH turnover rates compared to wild-type FAS, which we could not directly interpret. We started to exclude obvious reasons by confirming quantitative phosphopantetheinylation and by rechecking the protein samples by HPLC-SEC after activity assays had been conducted (**Figure S5**). Hence, the fused loading didomain either structurally compromised the mFAS fold in an unknown manner or the two additional domains can indeed functionally impact on the FA synthesis cycle. To exclude structural reasons we have tested functional knockouts of MAT0, ACP0 or both, which essentially retained catalytic efficiency of the

wild-type mFAS (**Figure 24**). Therefore, the effect has to originate from an acyl-ACP0, which can somehow functionally interact with the KS domain.

It remains illusive to date, which aspect causes the effect of an impaired FA synthesis. Basically there are two different possibilities how an acyl-ACP0 can interfere with FA production leading to a reduced overall rate. Witkowsky *et al.* stated that decarboxylation of malonyl-ACP does not take place in the native free KS domain, as the negatively charged carboxyl group of the malonyl moiety binds to the oxyanion hole within the active site.[169] When this scenario happens after a condensation step in the inherent FAS fold, the malonyl-ACP0 would be complexed at the KS domain acting as a competetive inhibitor for the acyl-loaded integral ACP domain. This would decrease the rate of every elongation cycle. The other possibility is that our loading module is actually capable of transferring acetyl units to the active site of the KS domain. When this occurs after a condensation step in the inherent FAS fold, the intermediate loaded on the integral ACP domain can only be processed, but not elongated any further. Dependent on the frequency of such events, butyryl or longer acyl moieties have to be unloaded from the integral ACP domain by the MAT or the TE domain. The TE domain is very specific for palmitoyl moieties and possesses slow hydrolysis rates for shorter acyl chains.[35] The MAT domain on the other hand has decreasing efficiencies for elongating acyl chains and the process of MAT-mediated unloading competes with the loading process of the native substrates malonyl- and acetyl-CoA. Hence, it is conceivable that unloading processes of intermediates from the inherent mFAS fold become rate determining for the overall FA synthesis. It is a future goal, to further investigate this aspect to shed light on the molecular mechanism behind the observation of reduced FA synthesis, which might be accomplished by analysing the FA product spectrum by HPLC-MS analysis. Unfortunately, these analytics have not been established in our laboratory yet and the establishment will require some time, as chemical esterification are required prior to the analysis. These constructs can have a great impact on kinetically analysing the transition from an iterative system to a modular one, which has only been shown for one iterative module of a PKS *in vivo* so far.[170]

3.5 Impact of Our Findings on the PKS Field

In this section two aspects of our findings shall be briefly discussed, which may directly affect current research in the PKS field. Protein engineering of modular PKSs turned out to be very challenging due to the requirement of detailed kinetic information about individual domains and in-depth knowledge about how different domains have to be assembled for functional cooperation.[162, 171] Until to date, very few AT domains of the vast amount of existing ATs in modular and iterative PKSs have been individually characterized. Hence, utilization of the one, well-characterized MAT fold and its variants may serve as versatile tool for PKS engineering in hybrid FAS/PKS systems. Furthermore, as a second aspect, we would like to emphasize the importance of the post-AT linker for the integrity of the KS domain of mFAS, which might question the postulated role in recent structural studies.[166, 172]

3.5.1 MAT as a Versatile Tool for PKSs Engineering

Our detailed analysis has demonstrated robustness and plasticity of the MAT fold, which are the two criteria for a promiscuous protein to be suited for an evolvability.[173] Robustness was shown, as the domain can easily be liberated from the KS-MAT didomain without losing catalytic efficiency (**Figure 30A**). Further the fold tolerates point mutations without suffering from significant destabilization (thermal shift assays with variants were performed within a master thesis).[133] In a mutagenesis study, we were able to alter the enzyme promiscuity by inducing substrate specificity with single substitutions of amino acids (**Figure 41**). By this means, five positions were identified, which are important for functional plasticity. Furthermore, the wild-type MAT domain exhibits some exciting properties, which might be beneficial for protein engineering. The domain is catalytically very efficient ($k_{cat}^{abs} = 119\ s^{-1}$ for malonyl-CoA) with relatively slow hydrolysis rates and last but not least, murine MAT is available in the heterologous host *E. coli* and, as such, easily amenable to *in vitro* engineering and mutational studies.

Given these facts, the MAT domain permits a versatile usage as a potential loading- and extending-AT in hybrid FAS/PKS systems, in the latter case even in *cis* or in *trans*. A single mutation (e.g. R606A) also functionally turns the domain into a typical loading-AT (although it shall be mentioned that such ATs possess all kinds of specificities, also for carboxyl groups as e.g. in the borrelidin synthase).[154] To incorporate designer acyl substrates, which lack naturally evolved AT domains, as e.g. provided by the Erb

laboratory,[63] the MAT domain may be a suitable promiscuous precursor enzyme to conduct directed evolution. Especially for this purpose, the low hydrolysis rate may increase the relevance for engineering, paving the way to energetically low costs of incorporation.

Nevertheless, further experiments are required to facilitate a direct application. These include the identification of plasticity residues that abolish the transfer of acetyl and related priming moieties. It is also important to increase the distinguishability between different extender substrates, although a limited substrate availability *in vivo* has to be taken into account, which may diminish the importance of very high substrate specificities. Last but not least, it has to be tested, how easy existing extender AT domains can be swapped in hybrid FAS/PKS KS$_{PKS}$-MAT didomains and whether this influences the catalytic efficiency of the integral KS$_{PKS}$ domain. A respective study to investigate the possibility of exchanging the AT domain within module 6 of DEBS was recently reported including effects of the engineering on protein stability and functionality.[162]

3.5.2 Linkage of ACP to the Condensing Part

Driven by the question of how long an "artificial" linker between the MAT and ACP domains has to be in a constructed α-module (**Figure 48A**), we have compared the influences of different anchor points for such linkers on the integrity of the KS domain (section 2.2.2). Although we have only considered attachments C-terminal to the conserved motif in the post-AT linker, we were surprised of how important the eleven following amino acid residues were for the stability and functionality of our constructs. This finding used to be an important aspect for a critical article about a recent structural study about pikAIII solved by cryoEM.[166, 167] In this report, the authors found a complete release of the post-AT linker from the KS surface and interpreted the resulting second active site entry as a new catalytically important feature of modular PKSs (**Figure 48B**).

Besides other inconsistencies in their postulated model (for details, see highlight article),[167] we think that this is a very critical aspect in their argumentation. It needs further verifications whether the property of murine KS-MAT didomain is exemplary for other KS-AT didomains in modular PKSs, but recent publications from the Keasling laboratory have confirmed the importance of post-AT linkers also in modular PKSs.[95, 162] Especially, the before mentioned report from 2017 clearly indicated the essential role of the C-terminal part of the post-AT linker leading to inactivity and a significant reduction in the yield of purified protein when exchanged.
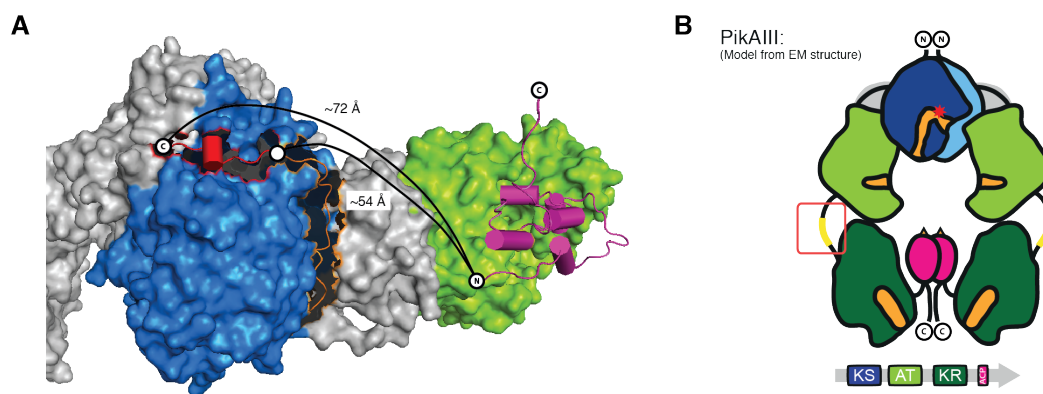
Figure 48: Linkage between ACP and MAT. (A) Different possibilities to link ACP to the condensing part. Surface depiction of KS-MAT (pdb: 5my0) and carton depiction of rat ACP (pdb: 2png), assembled at the active site entry of MAT. Positioning of the rat ACP was achieved by a structural superposition utilizing the crystal structure of the VinK-VinL complex (5czd).[65] The post-AT linker is shown as cartoon in orange (conserved motif from PKS) and red. Distances were measured from highlighted positions. (B) Model of the cryo-EM structure of PikAIII showing the release of the post-AT linker (yellow) from the KS-surface.[166]

3.6 Towards a Structural Investigation of a Part of DEBS by cryoEM

Inspired by the conformational flexibility of animal FAS,[54] which originates from a rather loose connection between the condensing and the processing part within a module, we intended to design a construct of DEBS amenable for structural studies by cryo electron microscopy (cryoEM). There is still a huge demand for structural information on modular PKSs, as the structural elucidation of such proteins is extraordinary difficult leading to a lack of knowledge about the overall architecture. In fact, only Dutta *et al*. managed to solve the structure of an entire module of a modular PKS by cryoEM,[166, 172] which led to a controversy due to inconsistencies with former reports (section 3.5.2).[167] To this end, we have been searching for an appropriate construct with sufficient protein quality for the past three years within a side project of my thesis to be able to conduct an analysis with cryoEM.

To circumvent the issue of tremendous flexibility within a module, assuming similar properties of modules of PKSs and the FAS fold, and to allow an investigation of module-module interactions, our construct is based on the processing part of module 4 and the condensing part of module 5 (M4proc-M5cond$_{DEBS}$). Our preliminary data looks promising (summarized in **Figure S9**) and work will be continued after this thesis in cooperation with Edoardo D'Imprima from the MPI for Biophysics (Department of Structural Biology, Prof. Werner Kühlbrandt and Dr. Janet Vonck).

3.7　Conclusion

Work within this thesis contributes significantly to the already broad knowledge about animal FASs and has made the murine FAS to a well-characterized member of this class. We have successfully established a reproducible and sufficient access to the protein in excellent quality and relatively high yield. In fact, our heterologous expression strategy of mFAS demonstrated the first sufficient expression of an animal FAS in *E. coli*. This adds FASs to our toolbox of manageable megaenzymes paving the way to exploit the catalytic efficiency in regards of microbial custom-compound biosynthesis (section 3.8).[174]

Especially, the acyltransferase domain MAT turned out to be a versatile tool to transfer a vast amount of chemically diverse compounds from CoA-esters to ACP offering the opportunity to incorporate those into fatty acids and potentially into polyketides. Our detailed structural and functional analysis, providing crystal structures of two different acyl-bound states and numbers of hydrolysis and transacylation rates for twelve exemplary CoA-esters, made the murine MAT to one of the best characterized AT domains. Further, we succeded in expressing the domain in different structural arrangements and in altering the substrate ambiguity by mutagenesis, which demonstrated robustness and plasticitity of the fold. While searching for appropriate analytics to investigate the MAT domain, the αKGDH assay turned out to be a well suited tool to obtain comprehensive enzyme kinetic data. Furthermore, initial experiments have demonstrated that the use of this assay can be expanded to also study other domains as KS or TE.

Protein engineering of mFAS experimentally confirmed the expected modular character of the fold, disclosing substantial structural variability in the animal FAS architecture. By exclusively utilizing native amino acid sequences, we have succeeded in rebuilding all four possible modules appearing in typical modular PKSs. Gained knowledge about the protein's behaviour with respective limitations (section 2.2.1) may assist with similar approaches in existing modular PKSs and may expand the biotechnological applicability of the animal FAS fold. We have further demonstrated functional integration of engineered loading didomains in bimodular constructs maintaining structural and functional integrity of the inherent FAS fold. These constructs mimic respective interactions of modules in modular PKSs providing an *in vitro* model system to kinetically study intermodular transfer between upstream ACPs and KS domains (section 3.8).

Hence, within this thesis animal FAS has been established as a model system to study PKSs and we commenced to engineer the FAS fold towards the design of modular PKSs. This aspect has been addressed on three levels of complexity from function of individual domains via organization of domains to form modules to the interplay of two modules in bimodular constructs.

3.8 Outlook

Our findings may be exploited to benefit further developments in future research in two aspects, i.e. utilizing the animal FAS fold to understand the working mode of PKSs on a molecular basis and to directly apply the unique properties of animal FAS for microbial custom-compound biosynthesis.

As the family of PKSs is diverse and modular PKSs are difficult to study *in vitro*, there is always a demand for appropriate model systems. In this regard, mFAS may serve as a very useful tool due to a good accessability via heterologous expression, excellent protein quality and comprehensive knowledge in the literature. Especially, reports from the 1980s-90s created a profound understanding of the molecular mechanism including a general equation (equation [2]) that describes enzyme kinetics of FA synthesis. Exploiting this knowledge, we suggest to use our bimodular constructs to study the transfer of intermediates between modules that exemplify respective arrangements in modular PKSs. The installation of an upstream loading module in the mFAS fold effects the function of the enzyme and demonstrates how the working mode of iterative modules may be altered by the overall architecture. Furthermore, mFAS facilitates a fine-tuned readout via FA product spectra, as every of the seven elongation cycles is directly apparent in the chain lengths of the produced fatty acids. With this system one could investigate influences of linker lengths, e.g. the ACP0-KS linker, effects of different ACP0-KS interfaces and the role of speed and specificity of the loading. It is not well understood why modules catalyze only one *or* several elongation steps in a modular PKS and respective studies are rare. The underlying mechanism of iterative *vs*. vectorial polyketide synthesis has been addressed *in vivo* by the Hertweck laboratory,[170] but our set-up would offer the opportunity for comprehensive studies *in vitro* including enzyme kinetics.

Structural information on enzymes contribute significantly to our understanding of their mode of action and are essential to aid protein engineering. Modular PKSs are difficult to handle *in vitro* and often suffer from instability and heterogeneity of sample preparations. These reasons have prevented to obtain high resolution data beyond

individual domains with only few exceptions. By assuming that modules of PKSs are structurally similar to the FAS fold and hence share equivalent properties, we concluded that flexibility may persist within a module and not necessarily between different modules of a modular PKS. Based on this assumption, we have designed our construct M4proc-M5cond$_{DEBS}$ and conducted first negative stain experiments (**Figure S9**). As preliminary data looks promising, we will put a strong focus on proceeding this project in collaboration with the Department of Structural Biology of the MPI for Biophysics. Further, our ability to label megaenzymes via amber codon suppression facilitates to attach fluorescent tags or spin labels site-specifically. Such samples may provide spatial and dynamic information, when investigated by FRET or EPR spectroscopy.

Besides providing basic principles about PKSs, the catalytic efficiency of mFAS may be directly exploited to synthesize relevant compounds in microorganisms. In this regard, mFAS allows a broader application than bacterial type I or yeast FAS due to their different overall architecture.[174] The open mFAS fold can tolerate domain deletions (shown within this thesis), and potentially substitutions relatively well, which demonstrates a high degree of architectural variability. Especially, the domains MAT and TE may serve as good targets for protein engineering, as structural interfaces with the residual fold are relatively small or non-existent. Swapping the TE domain does alter the termination step, which has already been shown for other thioesterases acting in *trans* or *cis*.[35, 175] Both approaches led to a reduction of the average chain length of FAs yielding short FAs, which are receiving increasing attention due to their potential usage as biofuels. Whereas other FAS systems (yeast FAS or type II systems) might be more suited for this approach, the core mFAS fold (our δ-module) would allow to attach different terminal modules, as e.g. sulfotransferase (ST)-TE domains or R (reductive) domains.[56] These give rise to completely different compound classes than fatty acids, namely terminal olefins and aldehydes or alcohols. Especially, utilizing R domains to produce 1-butanol seems to be an achievable goal, which is a very relevant compound with broad industrial application, e.g. as a very competitive biofuel for use in engines.[176] Both terminating domains can be found in PKSs and have already been used for protein engineering.[177, 178] Further, the finding that the MAT domain efficiently transfers succinyl moieties donated by succinyl-CoA may hold for the prospect of FAS-mediated synthesis of adipic acid.[154] For this purpose, mFAS must be modulated to accept the succinyl moiety for a single reductive chain elongation with malonyl-CoA. As the wildtype enzyme does not accept succinyl-CoA as a priming substrate (**Figure 16**), alternatively, the robust MAT domain could be introduced into a

respective PKS that already accepts carboxylic acids as starter units to exploit its excellent succinyl-transfer kinetics.[154]

Reuniting the families of FASs and PKSs to exploit the best catalytic properties from both fields, may offer new possibilities to create sufficient biocatalysts. Improving existing megaenzymes and altering their specificities aiming at derivatizing lead compounds may be one puzzle piece in challenging evolving resistancies of pathogens and may contribute to the economical use of microbial biosynthesis.

Chapter IV: Experimental Procedure

4.1 Material and Methods

Materials were purchased from Sigma (chemicals for synthesis, CoA-esters, β-NAD$^+$, NADH, NADPH, α-ketoglutarate dehydrogenase (αKGDH; porcine heart), α-ketoglutaric acid, thiamine pyrophosphate (TPP), and EDTA), Serva (BSA), New England Biolabs (NEB; restriction enzymes, Phusion High-Fidelity DNA Polymerase, plasmid pMAL-c5G), Clontech (In-Fusion HD Cloning Kit, Stellar Competent Cells), Carl Roth (isopropyl-$β$-D-thiogalactopyranoside (IPTG)), Qiagen (Ni-NTA affinity resin), Invitrogen (SDS-PAGE 4-12 % bis-tris gradient gels), Merck (plasmids pET22b and pETcoco-1), Merck Millipore (Amicon ultra centrifugal filters), Agilent (BL21gold(DE3)), IBA (Strep-Tactin columns; 5 mL), and GE Healthcare (Superdex 200 GL10/300 column, Superdex 200 HiLoad 16/60 and Superdex 200 HiLoad 26/60). The purity of CoA-esters was confirmed by HPLC-MS analysis before usage.

4.1.1 Preparation of Genomic DNA from *E. coli*

Genomic DNA from *E. coli* was obtained by following the instructions provided by eLABPROTOCOLS (https://www.elabprotocols.com/protocols). Although the procedure did not work very well, a DNA solution was gained that was sufficient to serve as a template for all respective PCRs within this thesis. In brief, a pre-culture (10 mL) of BL21gold(DE3) cells was grown and used for genomic DNA isolation. Cells were harvested by centrifugation (3,220 rcf for 10 min) and the cell pellets were resuspended in 500 µL elution buffer (Promega plasmid purification kit). The suspension was transferred to 2 mL tubes and 50 µL 10 % (v/v) SDS solution and 25 µL Proteinase K solution (20 mg mL^{-1}) were added. After incubation for 30 min at 55 °C (occasional mixing), 575 µL phenol/chloroform (1:1) were added and the viscous mixture was spun at 14,000 rcf for 10 min. The upper aqueous phase was transferred to a new tube, but the following described procedure failed. Instead, a white precipitate formed in the lower phase, which was extracted, washed with 70 % (v/v) ethanol and dissolved in elution buffer. This preparation contained sufficient amounts of genomic DNA and was stored at –20 °C.

4.1.2   Cloning of Plasmids

pVR01 (a pET22b(+) derivative) was used as a master plasmid for cloning the human FAS gene encoding a construct with an N-terminal StrepI and a C-terminal H(8)-tag. *FASN* genes from *Homo sapiens* and *Mus musculus* were purchased from Source BioScience (cDNA clones: IRAK110M20 and IRAVp968A0187D) as well as the murine thioesterase II gene (cDNA clone: IRAVp968A0864). Microbial genomic DNA was purchased from DSMZ: *Pyrococcus furiosus* (DSM3638), *Saccharopolyspora erythraea* (DSM40517), *Streptomyces avermitilis* (DSM46492), *Streptomyces venezuelae* (DSM41110) and *Streptomyces rapamycinicus* (DSM41530). Plasmids encoding chaperones from *E. coli* (DnaK, DnaJ, GrpE, GroEl and GroEs) were kindly provided by Prof. Hartl from the MPI Martinsried, whereas plasmids encoding DEBS 1, DEBS 2 and DEBS 3 (pBP130 and pBP144) were kindly provided by the Khosla laboratory in Stanford university. Plasmids encoding genes and tRNAs for amber codon suppression were either obtained from the Budisa laboratory: pJZ_mmPylS_pBCN (PylS of *Methanosarcina mazei*) or ordered at addgene: pSupAR-MbPylRS(DiZPK) (PylS of *Methanosarcina bakeri*) and pEVOL-pBpF (TyrS of *Methanocaldococcus jannaschii*). Fusion domains were cloned from commercially available plasmids.

We performed ligation independent cloning with the In-Fusion HD Cloning Kit (Clonetech) unless stated otherwise. All Primers mentioned in the following are listed in **Table S1** and information about all plasmids is summarized in **Table S2**. The sequences of all used plasmids were confirmed with the "dye terminator" method.

Generally, vectors were linearized with restriction enzymes or PCR and genes of interest, bearing complementary overhangs, were inserted by the In-Fusion HD Cloning Kit. In case inserts were assembled from different origins, all fragments were combined by megaprimer-PCRs to conduct In-Fusion cloning with three fragments at most. Templates for big PCR products (> 5000 bp) were always digested with Dpn1. Point mutations were generated with one "universal" reverse primer next to the position of interest and an overhang forward primer bearing the mutation. In case, a position was supposed to be randomized, we ordered forward primers with ambiguous nucleotides utilizing the standard code from IUPAC. When point mutations were introduced in large plasmids (> 8000 bp; pET22b derivatives), two separated PCRs were conducted using the primer pair of interest in combination with AR26/AR27.

In the following, the cloning of select constructs is described in detail, which was also part of a publication. The murine FAS gene was amplified by PCR with specific primers AR37 and AR38. By ligation independent cloning using 5' and 3' overhangs and

BamH1-HF/Xho1 digested pVR01, a pET22b derivative (termed pAR18) with an N-terminal StrepI and a C-terminal H(8)-tag was generated. The sequence of plasmid pAR18 was confirmed by sequencing. Plasmid pAR18 was used as a template for all following PCRs.

The KS-MAT encoding plasmid pAR69 was cloned by deletion of the respective sequence from pAR18 using primers AR163/AR148. Plasmid pAR70, encoding the KS knockout (C161G), was generated by PCR amplifying pAR69 with primers AR164/AR165 and transforming the purified PCR product into competent cells. Plasmids encoding MAT knockout (S581A) and MAT mutant (R606A) were cloned accordingly. Here, pAR70 was used as a template for PCRs using primer pairs AR310/AR301 and AR166/AR167 (pAR71 and pAR160). LD-MAT and MAT encoding plasmids were generated by amplifying with primers AR435/AR436 and AR512/AR513 and the purified PCR products were inserted into the BamH1-HF/Xho1 linearized vector pAR18 yielding pAR246 and pAR309, respectively.

Plasmid pAR327, encoding a C-terminal MBP fusion of MAT, was cloned by amplifying pAR309 with primers AR202/AR537. The MBP encoding sequence was amplified with primers AR539/AR540 using the vector pMAL-c5G as template. PCR products were assembled by ligation independent cloning and transformed in competent cells.

Plasmids encoding the bimodular constructs: LD-MAT-ACP-mFAS (pAR292) and MAT-ACP-mFAS (pAR340) were cloned in multiple steps: First, the BamH1 restriction site of pAR18 was exchanged to a Not1 restriction site by amplifying with primers AR467/AR468 and transforming into competent cells. The resulting plasmid (pAR264) was linearized by treatment with Not1-HF and Antarctic phosphatase and then used as backbone for both constructs in the ligation independent cloning reaction. The insert encoding LD-MAT-ACP was generated in two steps: ACP was linked to MAT by amplifying pAR18 with primers AR272/AR273 deleting the processing part. The resulting plasmid (pAR128) was used as a template for a PCR with primers AR494/AR496 to provide the insert for pAR292. The insert encoding MAT-ACP was cloned without intermediate transformation steps, MAT and ACP were amplified separately using pAR18 as a template and primers AR522/AR537 and AR560/AR561, respectively. Purified PCR products were combined by PCR (6 cycles without primers) and amplified with primers AR522/AR562 providing the insert for pAR340.

For purifying ACP, different expression plasmids were cloned. ACP was amplified with primers AR176/AR177 and the amplification product was inserted into a pAR18

derived vector pAR26, which was linearized with restriction enzymes BamH1-HF and Xho1, generating pAR75 (N-terminally StrepII-tagged). pAR26 was prepared from pAR18 with primer pair AR57/AR58. pAR75 was used as template for a PCR with primers AR202/AR227 to generate pAR100 (N-terminally strepII-tagged, C-terminally H(8) tagged).

The Sfp (*Bacillus subtilis* phosphopantetheinyl transferase) encoding co-expression plasmid pAR01H_SFP_bsub_pETcoco was prepared by linearizing pETcoco-1 with primers AR01H and AR02H by PCR. The Sfp encoding gene was ordered at GeneArt (ThermoFisher), and amplified with overhang primers AR03H and AR04H. Finally, the vector was assembled by ligation independent cloning.

Plasmid pAR352, encoding ACP and Sfp bicistronically, was generated by amplifying pAR100 with primers AR73/AR588. The Sfp-encoding sequence was amplified with primers AR81/574 using the vector pAR01H_SFP_bsub_pETcoco as a template. PCR products were assembled by ligation-independent cloning and transformed in competent cells.

4.1.1 Medium Scale Protein Expression with Enzymatic Lysis

Screening of expression conditions for human FAS constructs was conducted in medium scale cell cultures (50-100 mL). All plasmids were transformed into the respective chemically competent *E. coli* cell line and the transformants were grown overnight at 37 °C in 5 mL LB medium charged with the appropriate antibiotics and 1 % (w/v) glucose). 500-1000 µL of pre-cultures were used to inoculate 50-100 mL cultures and cells were grown under the respective conditions. Cells were harvested by centrifugation (3,220 rcf for 10 min) and the cell pellets were resuspended in 1 mL lysis buffer (200 mM potassium phosphate (pH 7.0), 10 % (v/v) glycerol, 25 % (m/v) sucrose, 10 mM EDTA and one tablet of PI cocktail (Roche) per 10 mL buffer). All following steps were performed on ice. After the suspension had been transferred to 2 mL tubes, 50 µL lysozyme solution (2 tips of small spatula in 0.5 mL H_2O) were added and the suspension was mixed for 1 min by a vortex mixer. Then, 60 µL 10 % (v/v) NP-40 were added to a final concentration of 0.5 % and mixed again. Subsequently, 6 µL 1 M $MgCl_2$ (final concentration was roughly 5 mM) and 50 µL DNase I solution (2 tips of small spatula in 0.5 mL H_2O) were added with interim mixing. The suspension was incubated for 1 h at 4 °C on a rotating wheel and cell debris was separated by centrifugation (20,000 rcf) for 1 h at 4 °C. 1 mL of the supernatant was transferred to a fresh tube and the residual liquid was carefully discarded. The cell pellet was resuspended in 0.5 mL 8 M urea

solution to dissolve protein aggregates and the cell debris was again separated by centrifugation (20,000 rcf) for 5 min at 4 °C. Samples from both solutions were taken (4 µL protein solution, 6 µL H_2O, 2 µL 6 × SDS loading buffer) and analyzed by SDS-PAGE.

4.1.2   Expression and Purification of FAS, KS-MAT and Truncated MAT

All plasmids were transformed into chemically competent *E. coli* BL21gold(DE3) cells. The transformants were grown overnight at 37 °C in 20 mL LB (100 µg mL^{-1} ampicillin (amp) and 1 % (w/v) glucose) medium. Pre-cultures were used to inoculate 1 L TB medium (100 µg mL^{-1} amp). Cultures were grown at 37 °C until they reached an optical density (OD$_{600}$) of 0.5–0.6. After cooling at 4 °C for 20 min, cultures were induced with 0.25 mM IPTG, and grown for additional 16 h at 20 °C and 180 rpm. Cells were harvested by centrifugation (4,000 rcf for 20 min). The cell pellets were resuspended in lysis buffer (50 mM sodium phosphate, 200 mM NaCl, 10 % (v/v) glycerol, 1 mM EDTA, 1 mM DTT, 20 mM imidazole (pH 7.6)) and lysed by French press. After centrifugation at 50,000 rcf for 30 min, the supernatant was mixed with 1 M MgCl$_2$ to a final concentration of 2 mM. The cytosol was transferred to Ni-NTA-columns and washed with 5 column volumes (CV) wash buffer (lysis buffer without EDTA). Bound protein was eluted with 2.5 CV elution buffer (50 mM sodium phosphate, 200 mM NaCl, 10 % (v/v) glycerol, 1 mM DTT, 300 mM imidazole (pH 7.6). The eluent was transferred to Strep-Tactin-columns, and washed with 5 CV strep-wash buffer (250 mM potassium phosphate, 10 % (v/v) glycerol, 1 mM EDTA, 1 mM DTT (pH 7.6)). Proteins were eluted with 2.5 CV elution buffer (strep-wash buffer containing 2.5 mM D-desthiobiotin). After concentration to 10–20 mg mL^{-1}, the proteins were frozen in liquid nitrogen and stored at –80 °C. Samples were thawn at 37 °C for 30–45 min and further polished by size-exclusion chromatography (SEC) using a Superdex 200 GL10/300 column equilibrated with the strep-wash buffer. Proteins were concentrated again to 10–20 mg mL^{-1} and stored frozen in aliquots using liquid nitrogen.

4.1.3   Expression and Purification of ACP

ACP for the activity assay was produced by co-transforming pAR100 with SFP_pETcoco into chemically competent *E. coli* BL21gold(DE3) cells. Overnight cultures were grown in LB (100 µg mL^{-1} ampicillin (amp), 36 µg mL^{-1} chloramphenicol (cam) and 1 % (w/v) glucose) at 37 °C. 2 L TB medium (100 µg mL^{-1} amp, 36 µg mL^{-1} cam) was inoculated and incubated at 37 °C until an optical cell density (OD$_{600}$) of 0.5–0.6 was

reached. After induction with 0.25 mM IPTG, cells were grown for 16 h at 37 °C. Cells were harvested by centrifugation, resuspended in lysis buffer and lysed by French press. After centrifugation (50,000 rcf for 30 min), the supernatant (supplemented with 2 mM MgCl$_2$) was transferred to Ni-NTA-columns and washed with 5 CV wash buffer. The protein was eluted with elution buffer (wash buffer containing 300 mM imidazole) and concentrated. Pooled fractions, were separated on a Superdex 200 HiLoad 16/60 or 26/60 SEC column equilibrated with buffer (50 mM sodium phosphate, 200 mM NaCl, 10 % (v/v) glycerol, 1 mM EDTA, 1 mM DTT). All fractions containing monomeric ACP were pooled and concentrated to 10–20 mg mL^{-1} (common yield was 20 mg purified protein per 1 L culture). Phosphopantetheinylation of ACP was confirmed by ESI mass spectrometry.

ACP, obtained from bicistronic co-expression with Sfp (pAR352), was expressed at 20 °C overnight and purified accordingly. BAP1 cells were transformed with pAR100 by electroporation and ACP was expressed at 20 and 37 °C over night.[113] Purified protein was analyzed accordingly (see **Figure S5**). After this analysis had been conducted ACP was always obtained from bicistronic co-expression with Sfp (pAR352).

4.1.4   **Expression and Purification of Active FAS Derivatives**

All constructs that need phosphopantetheinylation for activity were obtained by co-expressing with Sfp as described for ACP. In the later stages of this thesis plasmid SFP_pETcoco was exchanged by pAR357_SFP_pCDF-1b due to much better transformation efficiencies. Cell culture and protein purification were performed according to the protocol described for unmodified FAS, except that cell lysis and purification via Ni-NTA columns were performed in buffer (50 mM sodium phosphate, 450 mM NaCl, 20 % (v/v) glycerol, 1 mM DTT, (pH 7.6).

4.1.5   **Introduction of Noncanonical Amino Acids via Amber Codon Suppression**

For amber codon suppression the construct of interest, bearing an amber stop codon, was co-transformed with the desired suppression system (pUltra or pEVOL). The respective noncanonical amino acid was directly added to the main culture (2-4 mM, dissolved in 50 mL 0.5 M NaOH), when the cells were induced with IPTG. Cell culture and protein purification were performed according to the protocol described above (section 4.1.2).

4.1.6 HPLC-MS Analysis of ACPs

HPLC-MS analysis of ACP was performed using a Dionex UltiMate 3000 RSLC coupled to a Bruker micrOTOF-Q II equipped with an electrospray ionization source. Chromatographic separation was performed on a RP-C5 column (2.1 × 100 mm, particle size 3 µm, Discovery Bio WidePore) with a mobile-phase system consisting of solvent A (water with 0.1 % (v/v) formic acid) and solvent B (acetonitrile with 0.1 % (v/v) formic acid). The column was equilibrated with 95 % solvent A and 5 % solvent B at a flow rate of 0.4 mL min^{-1}. The concentration of acetonitrile was then linearly increased to 95 % over 14 min. Proteins were detected by monitoring the absorbance at 280 nm. ACP was found to elute at approximately 6.3 min. MS data was acquired in positive mode in the range from 100–1,500 *m/z* and later analyzed using DataAnalysis 4.0 software (Bruker Daltonik GmbH).

4.1.7 Determination of Protein Concentration

Protein concentrations were calculated from the absorbance at 280 nm, which was recorded on a NanoDrop 2000c (Thermo scientific). Extinction coefficients were calculated from the primary sequence without *N*-formylmethionine with CLC Main workbench (Qiagen). Absorbance of most important constructs 1 g L^{-1} at 280 nm (10 mm): 0.893 for mFAS; 1.053 for KS-MAT; 1.173 for LD-MAT; 1.498 for MAT$_{Ave}$-MBP; 0.907 for LD-MAT-ACP-mFAS; 0.905 for MAT-ACP-mFAS and 0.475 for ACP.

4.1.8 Occupancy of the MAT Domain in Thermodynamic Equilibrium

A protocol was established based on reports from the Smith laboratory using 6 % perchloric acid to stop the reaction and precipitate the protein.[108] Free thiol groups were quantified using the reagent by 7-Diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM), which was purchased from Sigma. Fluorescence was monitored with our Fusion Xpress gel documentation using the UVA 365 nm epi source with filter set F-535 Y2.

190 µL buffer (75 mM potassium phosphate (pH 7.4), 0.5 mM EDTA, 1 mM ascorbic acid) as negative control or buffer containing 2 µM acyltransferase were put into 1.5 mL tubes. 10 µL substrates (X-CoA; final concentration 10-20 µM; 5-10-fold excess) were added and the reaction mixture was incubated for 15-60 s at room temperature. The reaction was quenched by the addition of 50 µL 6 % (v/v) perchloric acid and the mixture was left on ice for 10 min to precipitate the protein. Aggregated protein was removed by centrifugation (20,000 rcf for 5 min) and 200 µL of the supernatant were carefully transferred to 96-well f-bottom microtiter plates (Greiner Bio-one). The solution

was neutralized with 50 µL NaOH solution (30 µL 1 M NaOH and 20 µL buffer) and 100 µL 60 µM CPM in buffer (very important to prepare freshly before usage) were added. After 15 min incubation at room temperature the fluorescence remained stable and was read out with our gel documentation.

A calibration was performed accordingly using the same procedure with free CoA in a final concentration range of 0.5–8 µM.

4.1.9 Activity Assay with 4-Methylumbelliferyl Acetate

Assays (100 µL volume) were run in 96-well f-bottom microtiter plates (Greiner Bio-one) and the fluorescence of 4-methylumbelliferone was monitored over a period of 20 min using a Tecan microplate reader at the following settings; excitation: 365 nm; emission: 455 nm; gain: 66.

Three solutions were prepared in assay buffer (75 mM potassium phosphate (pH 7), 0.5 mM EDTA, 1 mM ascorbic acid). Solution 1 contained 4-methylumbelliferyl acetate in 10-fold concentrated stocks in a range of 100-2000 µM. Solution 2 and 3 contained CoA (20 µM) and MAT (2.9 µM), respectively. The components were pipetted in order: 10 µL Sol 1, 50 µL Sol 2 and the reaction was initiated by the addition of 40 µL Sol 3. The final concentrations of all ingredients were 75 mM potassium phosphate, pH 7.0, 0.5 mM EDTA, 1 mM ascorbic acid, 10-200 µM 4-methylumbelliferyl acetate, 10 µM CoA and 1.15 µM MAT.

4.1.10 α-Ketoglutarate Dehydrogenase Coupled Activity Assay

The enzyme-coupled assay performed in this study was adapted from references.[69, 101] Briefly, assays were run in 96-well f-bottom microtiter plates (Greiner Bio-one) and NADH fluorescence was monitored using a ClarioStar microplate reader equipped with a dispenser (BMG labtech) at the following settings; excitation: 348-20 nm; emission: 476-20 nm; gain: 1900; focal height: 5.2 mm; flashes: 70; orbital averaging: 4 mm.

Similar to Dunn *et al.*, four different solutions were prepared in assay buffer (50 mM sodium phosphate, 10 % (v/v) glycerol, 1 mM DTT, 1 mM EDTA (pH 7.6), filtered and degased). Solution 1 (Sol 1) contained murine MAT in a 3.33-fold concentrated stock solution and supplemented with 0.1 mg mL^{-1} BSA. Solution 2 (Sol 2) contained 8 mM α-ketoglutaric acid, 1.6 mM NAD$^+$, 1.6 mM TPP and 60 mU/100 µL αKGDH, representing a 4-fold concentrated stock. Solution 3 (Sol 3) contained 4-fold concentrated CoA-esters, typically between 0.4–400 µM. Solution 4 (Sol 4) finally contained 5-fold concentrated

murine ACP, typically between 250-1500 µM. The components were pipetted in order: 30 µL Sol 1, 25 µL Sol 2 and 25 µL Sol 3, followed by mixing. The transfer reaction was initiated by 20 µL Sol 4, which was added by the dispenser. The final concentrations of all ingredients were 50 mM sodium phosphate, pH 7.6, 10 % (v/v) glycerol, 1 mM DTT, 1 mM EDTA, 2 mM α-ketoglutaric acid, 0.4 mM NAD$^+$, 0.4 mM TPP, 15 mU/100 µL αKGDH, 0.03 mg mL^{-1} BSA, 0.4–1 nM MAT, 50–300 µM ACP, 0.1–100 µM X-CoA (where X refers to the respective acyl-moiety of the assay). The background noise of the assay set-up was determined with assay buffer supplemented with 0.1 mg mL^{-1} BSA. Equidistant kinetic measurements were taken every 20 s for 5 min at 30 °C.

4.1.11 Hydrolysis Rate of MAT

The hydrolysis rate was determined as described previously with the variation that 30 µl Sol 1 (333 nM MAT or BSA-buffer) and 20 µL assay buffer were used instead of Sol 4. The final concentration of CoA-esters was 10 µM. All four solutions were pipetted manually and mixed. Each measurement was performed in technical triplicates.

4.1.12 Michaelis-Menten Kinetics of MAT

Determining the apparent Michaelis-Menten constant is an iterative process. Pre-experiments were initially performed to approach the approximate value of K_m. Eight data points were collected that cover substrate concentrations (Sol 3) of $0.2 \times K_m$; $0.3 \times K_m$; $0.5 \times K_m$; $0.75 \times K_m$; $1.25 \times K_m$; $2 \times K_m$; $3 \times K_m$; $5 \times K_m$. Every measurement was performed in technical triplicates, and additionally the corresponding background signal was recorded. Experiments were setup in a way such that changes in signal remained linear during the time ranges of measurement. While we generally recorded data to a consumption of the substrate up to 20 %, relatively low K_m at µM concentrations in combination with unfavorable signal-noise ratios made it necessary to extend ranges at low substrate concentrations up to a turnover of 50 %.

4.1.13 Analysis of Kinetic Data from the αKGDH Assay

For data analysis, the average of triplicates was calculated and the background subtracted. Relative fluorescence units were converted into concentrations (µM) using a calibration curve (**Figure 29**). All buffers were filtered, but occasional dust or air bubbles caused strong outliers, which were then omitted from further analysis. Quality of experimental observations was evaluated via standard deviations of triplicates. Data points from 40–140 s were fitted by linear regression to give the initial velocity of substrate

turnover. Michaelis-Menten constants and rate constants were determined with the Hanes-Woolf method (equation [10]). This linear regression method is known for its even weighting of data points, and turned out to well represent data.[58] Data quality was checked by also plotting via the Lineweaver-Burk- (equation [9]), Eadie-Hofstee-method (equation [11]) and non-linear regression. A representative data sheet is shown in **Figure S10**. Kinetic data were analyzed with Microsoft Excel.

Dunn et al. have defined the kinetic constants of MAT-mediated hydrolysis and transacylation to ACP in their supplementary information.[69] Based on the catalytic mechanism of MAT (**Figure 8**), we have introduced a constant for the off-rate of ACP (k_3), which had not been considered in the previous report. The analysis was mainly performed by Aaron Himmler and a derivation can be found in his master thesis.

4.1.14 Analysis of the Transthioesterification Reaction in the KS Domain

The protocol was accordinging to the MAT-mediated transfer reaction (section 4.1.10) with some adaptations. Briefly, assays were run in 96-well f-bottom microtiter plates (Greiner Bio-one) and NADH fluorescence was monitored using a ClarioStar microplate reader equipped with a dispenser (BMG labtech) at the following settings; excitation: 348-20 nm; emission: 476-20 nm; gain: 1200; focal height: 5.2 mm; flashes: 70; orbital averaging: 4 mm.

Four different solutions were prepared in the same assay buffer as in section 4.1.10. The final concentrations of all ingredients were 50 mM sodium phosphate, pH 7.6, 10 % (v/v) glycerol, 1 mM DTT, 1 mM EDTA, 2 mM α-ketoglutaric acid, 0.4 mM NAD$^+$, 0.4 mM TPP, 15 mU/100 µL αKGDH, 0.03 mg mL^{-1} BSA, 100–500 nM KS, 5–150 µM ACP, 25-500 µM X-CoA (where X refers to the respective acyl-moiety of the assay). The background noise of the assay set-up was determined with assay buffer supplemented with 0.1 mg mL^{-1} BSA. As relatively high concentrations of substrates are required due to high Michaelis constants, one has to particularly take care to use fresh, not hydrolyzed substrates. Equidistant kinetic measurements were taken every 20 s for 5 min at 25 °C.

4.1.15 Screening of MAT Variants

All settings were according to the protocol in section 4.1.10. Assays were run in 96-well f-bottom microtiter plates and pipetting was performed using the plate vertical. Every enzyme was measured in duplets (two columns) and every substrate (acetyl-, malonyl- and methylmalonyl-CoA) was added in a row. Thus, seven variants plus the respective background with all substrates were investigated in one set-up using two

plates, where the latter was prepared during the first measurement. All solutions were prepared in sufficient volumes for one run. Final data contains three independent measurements in duplets with the same batch of purified protein.

The final concentrations of all ingredients were 50 mM sodium phosphate, pH 7.6, 10 % (v/v) glycerol, 1 mM DTT, 1 mM EDTA, 2 mM α-ketoglutaric acid, 0.4 mM NAD$^+$, 0.4 mM TPP, 15 mU/100 µL αKGDH, 0.03 mg mL^{-1} BSA, 1–5 nM MAT, 60 µM ACP, 20 µM X-CoA (where X refers to the respective acyl-moiety of the assay).

4.1.16 Calculation of Transition State Energies

The energy of the transition state was calculated with:

$$\Delta G_{ES^{\ddagger}} = -RT\ln\left(\frac{k_{cat}}{K_m}\right) + RT\ln\left(\frac{k_B T}{h}\right) \qquad [\,5\,]$$

where k_B is the Boltzmann constant, T is the temperature in degrees Kelvin and h is Planck's constant.

The difference in transition state energies was calculated with experimentally measured values of the specificity constants (for example, see reference[179]):

$$\Delta\Delta G_{ES^{\ddagger}} = -RT\ln\left(\frac{(k_{cat}/K_m)_{Mal}}{(k_{cat}/K_m)_{Ac}}\right) \qquad [\,6\,]$$

where $(k_{cat}/K_m)_{Mal}$ and $(k_{cat}/K_m)_{Ac}$ refers to specificity constants of malonyl- and acetyl-transfer, respectively.

4.1.17 *In vitro* ACP Labeling with CoA 647

The degree of phosphopantetheinylation of constructs was quantified using a sensitive fluorescent assay.[83] Proteins at a concentration of 0.5 mg mL^{-1} were incubated in 50 mM sodium phosphate buffer (pH 7.6) containing 450 mM NaCl, 20 % (v/v) glycerol, 1 mM DTT, 10 mM MgCl$_2$, 1 µM CoA 647 and 10 µM purified Sfp-H6 for 1 h at 37 °C. In gel fluorescence was detected with a Fusion Xpress gel documentation system of Vilber equipped with a red excitation source (624 nm) and a F650/F695Y camera filter.

4.1.18 Overall Fatty Acid Synthase Activity

FAS activity was measured fluorometrically with a microplate reader by following the oxidation of NADPH at 25 °C in 50 mM potassium phosphate (pH 7.0), 75–100 µM acetyl-CoA, 100 µM malonyl-CoA and 40 µM NADPH. Alternatively, the absorbance at 340 nm was monitored with a NanoDrop (cuvette mode) using an extinction coefficient for NADPH of 6220 M^{-1}cm^{-1}. The enzyme was prepared in a 4-fold stock containing 20 %

(v/v) PEG 400 for stabilization, resulting in a final protein concentration of 20 nM and 5 % (v/v) PEG 400 in the assay. The reaction was initiated by the addition of malonyl-CoA. Every measurement was performed in technical triplicates and the corresponding background (without CoA-esters) was subtracted.

Microplate reader settings were: excitation: 348-20 nm; emission: 476-20 nm; gain: 1490; focal height: 5.7 mm; flashes: 17; orbital averaging: 4 mm.

4.1.19 Analysis of Kinetic Data from Overall Fatty Acid Synthesis

Data was analyzed according to a report of Hammes *et al.* using the software Origin.[44] To facilitate a global fit of the three data sets (varied substrate concentrations of NADPH, acetyl-CoA and malonyl-CoA) in origin, equation [2] had to be converted into:

$$v = s_{Ac} \frac{k_{cat}\,[E]_0}{1 + \frac{K_M}{[Mal-CoA]}\left(1 + \frac{x}{K_{A,i}}\right) + \frac{K_A}{x}\left(1 + \frac{[Mal-CoA]}{K_{M,i}}\right) + \frac{K_N}{[NADPH]}} \qquad [\,7\,]$$

$$+ s_{Mal} \frac{k_{cat}\,[E]_0}{1 + \frac{K_M}{x}\left(1 + \frac{[Ac-CoA]}{K_{A,i}}\right) + \frac{K_A}{[Ac-CoA]}\left(1 + \frac{x}{K_{M,i}}\right) + \frac{K_N}{[NADPH]}}$$

$$+ s_N \frac{k_{cat}\,[E]_0}{1 + \frac{K_M}{[Mal-CoA]}\left(1 + \frac{x}{K_{A,i}}\right) + \frac{K_A}{x}\left(1 + \frac{[Mal-CoA]}{K_{M,i}}\right) + \frac{K_N}{x}}$$

where parameters s_{Ac}, s_{Mal} and s_N were put to zero or one assigning the summand to the respective data set. As other substrate concentrations were kept constant, all concentrations reduced to simple numbers. Thereby, kinetic constants are defined as described by Hammes.[44]

4.1.20 Triacetic Acid Lactone (TAL) Production Assay

TAL production was monitored photometrically with a NanoDrop (cuvette mode) by observing the increase in absorbance at 298 nm. Two solutions were prepared in the assay buffer (50 mM potassium phosphate, pH 7.0, 10 % (v/v) glycerine, 1 mM DTT and 0.03-0.05 mg/mL BSA). Solution 1 contained the substrates acetyl- and malonyl-CoA (both 200 µM) and solution 2 contained the enzyme (1400 nM) in 2-fold concentrated stock solutions. 50 µL of both solutions (incubated at 25 °C) were mixed to final concentrations of 100 µM acetyl-CoA, 100 µM malonyl-CoA and 700 nM enzyme and the absorbance was monitored for 5 min. Absorption units were converted into concentrations by using the extinction coefficient for TAL of $\varepsilon = 2540\ M^{-1}\ cm^{-1}$.[180]

4.1.21 Crystallization of the KS-MAT Didomain

Initial crystallization conditions for KS-MAT were identified by carrying out screening trials using various commercial sparse-matrix crystallization screening kits (The SGC suite, The CSC suite, The JCSG+ suite, The JCSG Core I-IV Suites, The PACT Suite, The PEGs Suite and The PEGs II Suite (QIAGEN), Index HT (Hampton Research), Morpheus HT-96 and The PGA screen (Molecular Dimensions)) using a sitting-drop vapor-diffusion method technique on 96 well GreinerBioOne plates at 22 °C. Crystals were observed in the PEG Ion (condition D1) screen after about 2 days. These conditions were optimized with hanging drop vapor diffusion technique in 24-well plates using 2 µl protein solution (15-20 mg mL^{-1}) in the droplet mixed with 2 µl reservoir solution. Single-crystals were obtained at 0.2 M potassium-sodium tartrate, 25 % (w/v) PEG 3350 at 20 °C to sizes of about $75 \times 75 \times 75$ µm^3. Drops with the crystals were supplemented with 0.5 µl of 10 mM malonyl-CoA (Sigma-Aldrich) for up to 2 minutes and subsequently treated with a cryosolution containing 20 % (v/v) glycerol in the mother liquor. The crystal was then picked in a nylon fiber loop and vitrified into liquid nitrogen. Co-crystallization experiments gave micro-crystals that were unusable in any diffraction experiments.

4.1.22 Data Collection and Processing

Crystals were exposed to single wavelength X-radiation at the Swiss Light Source (X06SA), and maintained at 100 K, while data were recorded onto a detector (DECTRIS EIGER 16M). Data processing was performed with goeiger.com pipeline at X06SA within the *XDS*,[181] for indexing and integration, and *Aimless*[182] for scaling. All diffraction data are publicly available at https://zenodo.org/record/. Using the *CCP4i* suite, the calculations of Matthews coefficient[183] (V$_m$) for the cell gave values of 3.6 Å^3 Da^{-1} corresponding to a solvent content of 66 % for 4 molecules per asymmetric unit. The phase problem was solved using molecular replacement method with the program *Molrep*,[184] using the structural model of a monomer from the human FAS KS-MAT didomain (pdb accession code 3hhd).[64] After an initial rigid-body refinement, the model was subjected to repeated cycles of restrained refinement with *REFMAC*5.[185] Between the cycles of refinement, the model was built using *Coot*.[186] A sketch of malonyl serine molecule was drawn with the program *JLigand* and included in the refinement process.[187] Data collection and refinement statistics are given in **Table 2** and **Table 3**.

4.2   Chemical Synthesis

4.2.1   4-Methylumbelliferyl acetate (C1)[188]
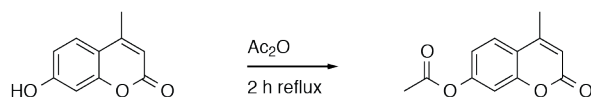
4-Methylumbelliferone (4.36 g, 25 mmol) was refluxed in 10 mL acetic anhydride for 2 h. Then the reaction mixture was poured into 30 mL ice-cold water to obtain the solid. After recrystallization by 95 % ethanol 4.8 g (89 % yield) of the product (**C1**) was obtained as white needles.

^{1}H-NMR (CDCl$_3$, 250 MHz) δ: 7.61 (d, ^3J = 8.5 Hz, 1H), 7.12-7.06 (m, 2H), 6.26 (d, ^4J = 1.2 Hz, 1H), 2.44 (d, ^4J = 1.2 Hz, 3H), 2.34 (s, 3H) ppm.
Spectrum is shown in **Figure S11**.

4.2.2   Meldrum's acid[189]
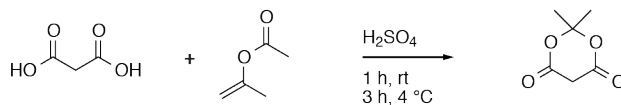
To a suspension of powdered malonic acid (10.4 g, 50 mmol, 1 eq) in isopropenyl acetate (5.5 g, 55 mmol, 1.1 eq) 0.1 mL concentrated sulfuric acid was added dropwise while stirring. After 45 min, most of the malonic acid dissolved and 8 mL acetone were added. The reaction mixture was allowed to stand in the refrigerator for 3 h and the resulting crystals were filtered by suction and washed with sufficient ice water. Recrystallization by dissolving with 20 ml acetone and 40 mL ice water yielded 5 g (35 %) of the product as white crystals.

^{1}H-NMR (CDCl$_3$, 400 MHz) δ: 3.63 (s, 2H), 1.79 (s, 6H) ppm.

4.2.3 4-Methylumbelliferyl malonate (C2)[190]
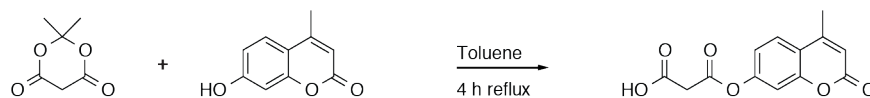
4-Methylumbelliferone (0.41 g, 2.3 mmol, 1 eq) and Meldrum's acid (0.5 g, 3.5 mmol, 1.5 eq) were heated in refluxing 15 mL toluene for 4 h. Upon cooling the product was precipitated and collected by filtration to yield 500 mg (82 %) of the product (C2) as brown crystals.

^{1}H-NMR (CD$_3$OD, 250 MHz) δ: 6.25 (d, ^3J = 9.4 Hz, 1H), 5.64-5.60 (m, 2H), 4.76 (d, ^4J = 1.2 Hz, 1H), 2.13 (s, 2H), 0.92 (d, ^4J = 1.2 Hz, 3H) ppm.
Spectrum is shown in Figure S12.

4.2.4 4-methyl-7-thioumbelliferone acetate (C3)
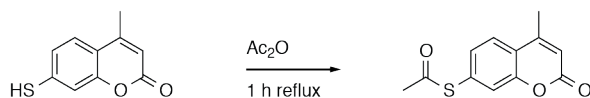
7-Mercapto-4-methylcoumarin (200 mg, 1 mmol) was refluxed in 5 mL acetic anhydride for 1 h. Then the reaction mixture was poured into 15 mL ice-cold water to obtain the solid. After recrystallization by 95 % ethanol, 244 mg (63 % yield) of the product (C3) was obtained as golden crystals.

^{1}H-NMR (CDCl$_3$, 250 MHz) δ: 7.64 (d, ^3J = 8.2 Hz, 1H), 7.42 (d, ^4J = 1.6 Hz, 1H), 7.34 (dd, ^3J = 8.2 Hz, ^4J = 1.7 Hz, 1H), 6.34 (s, ^4J = 1.1 Hz, 1H), 2.48 (s, 3H), 2.45 (d, ^4J = 1.1 Hz, 3H) ppm.
Spectrum is shown in Figure S13.

4.2.5 2-Amino-3-(4-azidophenyl)propanoic acid hydrochloride (AzPhe, C4)[191]
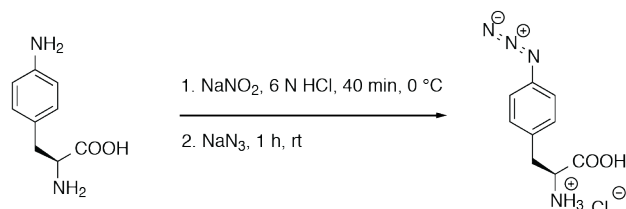
4-Aminophenylalanine (10.0 g, 55.4 mmol, 1 eq) was dissolved in 200 mL of an aqueous 6 N HCl solution and cooled down to 0 °C. A solution of $NaNO_2$ (5.0 g, 72.0 mmol, 1.3 eq) in 20 mL H_2O was prepared and slowly added. The reaction mixture was stirred for 40 min at 0 °C. Then, a solution of NaN_3 (5.4 g, 83.2 mmol, 1.5 eq) in 20 mL H_2O was prepared and added in one portion. The reaction mixture was kept at 0 °C for another 5 min and let warm to room temperature for 1 h. Evaporation of the solvent followed by column chromatography on silica gel (eluent DCM:MeOH = 7:3, r_f=0.25) gave 11.5 g (85 % yield) of the product (**C4**) as slightly yellow solid.

^{1}H-NMR (DMSO-d6, 250 MHz) δ 7.34 (d, ^3J = 8.4 Hz, 2H), 7.10 (d, ^3J = 8.4 Hz, 2H), 4.13 (dd, ^3J = 6.4 Hz, 6.3 Hz, 1H), 3.17-3.03 (m, 2H) ppm.
Spectrum is shown in **Figure S14**.

IR (KBr): Azide band at 2140 cm^{-1} matches with the literature.[192]

4.2.6 Methyl 2-((tert-butoxycarbonyl)amino)-3-(4-(ethynyloxy)phenyl)propanoate[193]
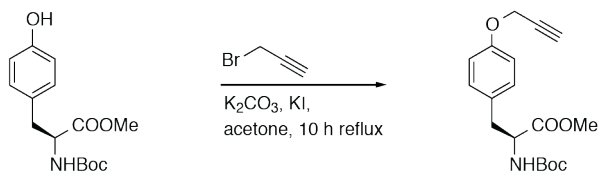
The protected tyrosine (13.1 g, 45.0 mmol, 1.0 eq), potassium carbonate (9.1 g, 67.5 mmol, 1.5 eq) and potassium iodide (0.7 g, 4.5 mmol, 0.1 eq) were put in a 250 mL flask and 150 mL dry acetone was added. Propargyl bromide (6 mL, 45 mmol, 1.0 eq) was added dropwise over 30 min and the mixture was kept overnight at room temperature without stirring. The mixture was refluxed at 150 °C for 10 h and rapidly cooled to room temperature afterwards. The mixture was filtrated two times and the solvent was removed under reduced pressure. The residue was dissolved in 100 mL ethyl acetate, washed with 30 mL H_2O twice and dried over $MgSO_4$. The solvent was removed under reduced pressure to yield 14.1 g (94 % yield) of the product as yellow oil.

^{1}H-NMR (CDCl$_3$, 250 MHz) δ 7.06 (d, ^3J = 8.75 Hz, 2H), 6.91 (d, ^3J = 8.75 Hz, 2H), 4.97-4.96 (m, 1H), 4.67 (d, ^3J = 2.50 Hz, 2H), 4.54-4.53 (m, 1H), 3.71 (s, 3H), 3.11-3.03 (m, 1H), 2.53-2.52 (m, 1H), 1.42 (s, 9H) ppm.

4.2.7 2-((Tert-butoxycarbonyl)amino)-3-(4-(ethynyloxy)phenyl)propanoic acid[194]
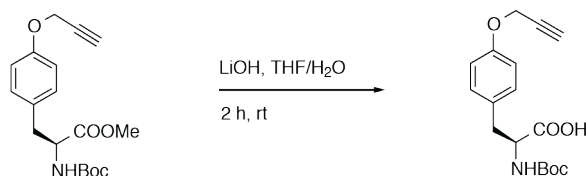
The double protected propargyl derivative (13.3 g, 40 mmol, 1 eq) was dissolved in 100 mL THF and 85 mL of an aqueous 1 M LiOH solution were added carefully. The yellow solution was stirred for 3 h at room temperature and the reaction was tracked via TLC (ethyl acetate:n-hexane = 1:1, r_f=0.8). The reaction was stopped by additions of 15 mL of a conc. HCl solution and the organic phase was separated. The aqueous phase was extracted three times with ethyl acetate (TLC control). The organic phases were combined, washed with brine and dried over MgSO$_4$. The solvent was removed under reduced pressure to yield 11.7 g (91 % yield) of the product as yellow oil.

^{1}H-NMR (CDCl$_3$, 250 MHz) δ 7.13 (d, ^3J = 8.51 Hz, 2H), 6.92 (d, ^3J = 8.70 Hz, 2H), 4.68 (d, ^3J = 2.40 Hz, 2H), 3.19-2.90 (m, 2H), 2.52 (t, ^3J = 2.40 Hz, 1H), 2.05 (s, 1H), 1.43 (s, 9H) ppm.

4.2.8 2-Amino-3-(4-(ethynyloxy)phenyl)propanoic acid hydrochloride (C5, PrPhe)[195]
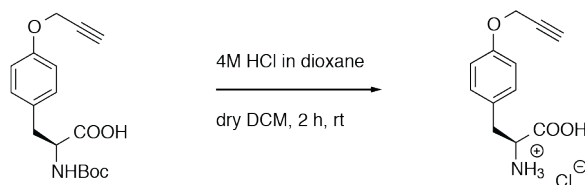
The carboxylic acid (11.7 g, 36.6 mmol, 1 eq) was dissolved in 25 mL dry DCM and 25 mL of a 4 M HCl solution in dioxane were added. The colourless suspension was stirred for 2.5 h at room temperature and the solvent was removed under reduced pressure. The residue was taken up with cold Et_2O, filtered and dried under vacuum. The product (**C5**) was a colourless solid that was obtained in a yield of 9.3 g (99 %yield).

^{1}H-NMR (DMSO-d6, 400 MHz) δ 8.40 (br s, 2H), 7.21 (d, $^3J = 8.56$ Hz, 2H), 6.94 (d, $^3J = 8.56$ Hz, 2H), 4.77 (d, $^3J = 2.32$ Hz, 2H), 4.11-4.08 (m, 1H), 3.57-3.56 (m, 1H), 3.08 (d, $^3J = 6.11$ Hz, 2H) ppm.
Spectrum is shown in **Figure S15**.

HRMS(MALDI +): *m/z* calc. for $C_{12}H_{13}NO_3$ [M+Na]$^+$ 242.07876, found 242.07914, Δm=0.00038, Δppm=1.6

4.2.9 Methylthiocarbohydrazide[195]
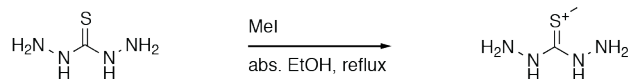
Thiocarbohydrazide (24.1 g, 0.23 mol, 1.0 eq) was suspended in 800 mL abs. ethanol and refluxed at 200 °C. Iodomethane (16 mL, 0.26 mol, 1.1 eq) was added and heating was continued for 1 h. The slightly yellow solution was filtered and then stirred for 1 h at room temperature. The flask was kept in the fridge overnight. Most of the solvent was decanted and the residue was removed under reduced pressure. The product was dried in vacuum to yield 33.3 g (59 % yield) of the product as colourless needles.

^{1}H-NMR (DMSO-d6, 250 MHz) δ 2.38 (s, 3H) ppm.

4.2.10 3-Methyl-6-(methylthio)-1,2,4,5-tetrazine[196]
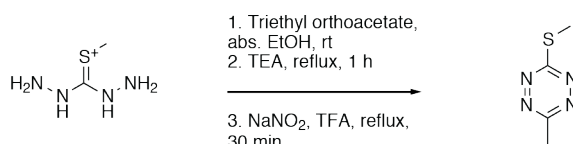
Methylthiocarbohydrazide (11.2 g, 45.1 mmol, 1.0 eq) was suspended in 300 mL abs. ethanol and stirred at room temperature. Triethyl orthoacetate (11.30 mL, 61.6 mmol, 1.4 eq) was added and stirring was continued for 30 min at room temperature. Triethylamine (6.29 mL, 45.1 mmol, 1 eq) was added (solution turned pink) and the mixture was refluxed for 60 min. $NaNO_2$ (6.4 g, 92.8 mmol, 2.1 eq) and TFA (3.48 mL, 45.1 mmol, 1.0 eq) were added and refluxing was continued for 30 min. 300 mL n-hexane was added and the solution was degassed with argon and stirred for 30 min at room temperature. 600 mL H_2O was added and the solution was extracted three times with 300 mL Et_2O each. The solution was dried over $MgSO_4$ and the solvent was removed under reduced pressure. The residue adsorbed on silica gel and purified via column chromatography (n-hexane:DCM=19:1). The respective fractions were combined and the

solvent was removed under reduced pressure to yield 1.5 g (23 % yield) of the product as red oil.

^{1}H-NMR (CDCl$_3$, 250 MHz) δ 2.97 (s, 3H), 2.72 (s, 3H) ppm.

4.2.11 2-((Tert-butoxycarbonyl)amino)-3-(4-((6-methyl-1,2,4,5-tetrazin-3-yl)amino)phenyl) propanoic acid[195]
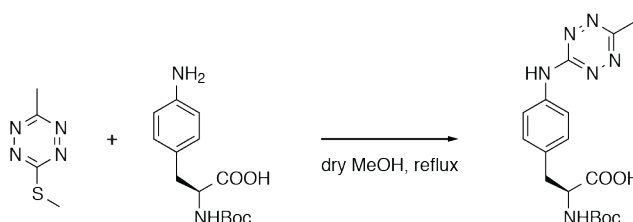
The tetrazine derivative (0.57 g, 4.0 mmol, 1.1 eq) and Boc-D-4-aminophenylalanine (1.02 g, 3.6 mmol, 1.0 eq) were suspended in 10 mL dry methanol and refluxed at 150 °C for 18 h. The solvent was removed under reduced pressure and the residue was adsorbed on silica gel and purified via column chromatography (CHCl$_3$ + 0.5 % up to 5.0 % methanol). The solvent was removed under reduced pressure to obtain 0.36 g (23 % yield) of the product as orange crystals.

^{1}H-NMR (DMSO-d6, 250 MHz) δ 12.60 (s, 1H), 10.62 (s, 1H), 7.63 (d, ^3J = 7.5 Hz, 2H), 7.25 (d, ^3J = 7.5 Hz, 2H), 4.08-4.05 (m, 1H), 3.04-2.79 (m, 2H), 2.78 (s, 3H), 1.34 (s, 9H) ppm.

4.2.12 2-Amino-3-(4-((6-methyl-1,2,4,5-tetrazin-3-yl)amino)phenyl)propanoic acid hydrochloride (C6, Tet-Phe)[195]
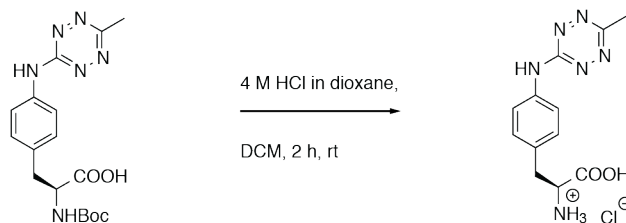
The Boc-protected tetrazine derivative (225 mg, 0.7 mmol, 1 eq) was dissolved in 8 mL dry DCM and 8 mL of a 4 M HCl solution in dioxane was added. The solution was stirred for 2 h at room temperature. The solvent was removed under reduced pressure and the residue was dissolved in DCM. The solvent was again removed under reduced pressure and dried in vacuum to yield 161 mg (87% yield) of the product as red crystals (**C6**).

^{1}**H-NMR** (DMSO-d6, 400 MHz) δ 10.70 (s, 1H), 8.41 (br s, 2H), 7.67 (d, ^3J = 8.56 Hz, 2H), 7.28 (d, ^3J = 8.56 Hz, 2H), 4.14 (br t, ^3J = 6.30 Hz, 1H), 3.11 (d, ^3J = 6.30 Hz, 2H), 2.78 (s, 3H) ppm.
Spectrum is shown in **Figure S16**.

HRMS(MALDI +): *m/z* calcd. for $C_{12}H_{14}N_6O_2$ [M+Na]$^+$ 297.10704, found 297.10681, Δm=0.00023, Δppm=0.8

4.2.13 6-Acetamido-2-((tert-butoxycarbonyl)amino)hexanoic acid
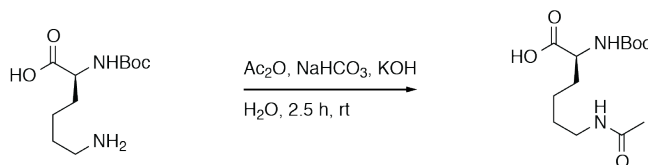
Boc-L-Lysine (3.06 g, 12.2 mmol, 1 eq) was dissolved in 60 mL H_2O and both sodium hydrogen carbonate (3.12 g, 36.6 mmol, 3 eq) and potassium hydroxide (0.85 g, 12.2 mmol, 1 eq) were added. Acetic anhydride (1.15 mL, 12.2 mmol, 1 eq) was added dropwise over 10 min and the solution was stirred for 2.5 h at room temperature. Then, the solution was dissolved in 60 mL ethyl acetate and brought to pH=2 by addition of a conc. HCl solution. The organic phase was separated and the aqueous phase was extracted three times with 30 mL ethyl acetate each. The organic phases were combined, washed with brine and dried over $MgSO_4$. The solvent was removed under reduced pressure to yield a slightly yellow oil. The oil was directly used in the next step without further purification.

^{1}H-NMR (DMSO-d6, 250 MHz) δ 12.11 (s, 2H), 7.79 (t, $^3J = 5.4$ Hz, 1H), 7.01 (d, $^3J = 8.2$ Hz, 1H), 3.89-3.77 (m, 1H), 3.00 (q, $^3J = 5.8$ Hz, 2H), 2.10 (s, 2H), 1.78 (s, 2H), 1.71-1.47 (m, 2H), 1.39 (s, 9H) ppm.

4.2.14 6-Acetamido-2-aminohexanoic acid hydrochloride (C7, AcLys)[195]
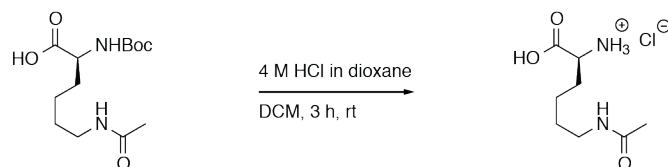
6-Acetylamino-Boc-L-lysine was dissolved a solution of 38 mL dioxane and 3 mL of a 4 M HCl solution in dioxane. The solution was stirred for 3 h at room temperature. The solvent was removed under reduced pressure and the residue was dissolved in 60 mL of an aqueous 1 M HCl solution. The solvent was removed azeotrope under reduced pressure by addition of four times 25 mL toluene to yield a yellow oil. After lyophilization for five times, the solvent was again removed azeotrope under reduced pressure and dried in vacuum to obtain 2.75 g (quantitative over two steps) of the product (C7) as off-white powder.

^{1}H-NMR (DMSO-d6, 250 MHz) δ 8.54-8.47 (m, 2H), 8.06-8.03 (m, 1H), 3.81 (q, ^3J = 5.5 Hz, 1H), 3.03-2.96 (m, 2H), 1.79 (s, 5H), 1.45-1.27 (m, 4H) ppm.

Spectrum is shown in **Figure S17**.

MS(ESI +): *m/z* calcd. for $C_8H_{16}N_2O_3$ [M+H]$^+$ 189.12, found 189.12

4.2.15 2-((Tert-butoxycarbonyl)amino)-6-(((prop-2-yn-1-yloxy)carbonyl)amino)hexanoic acid[197]
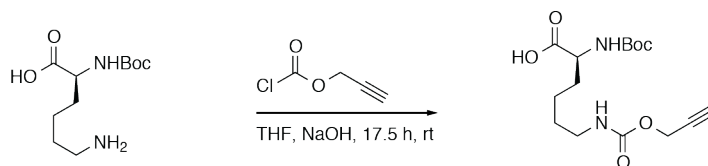
Boc-L-Lysine (3.0 g, 12.2 mmol, 1.3 eq) was added to a solution of 30 mL THF and 30 mL of an aqueous 1 M NaOH solution. The solution was cooled to 0 °C and propargyl chloroformate (0.96 mL, 9.7 mmol, 1 eq) was added dropwise over 10 min. The solution was stirred at room temperature for 17.5 h and then diluted with 60 mL ethyl acetate. Conc. HCl solution was added until the aqueous phase reached a pH of 1.5. The organic layer was separated and the aqueous phase was extracted three times with 30 mL ethyl acetate each. The organic phases were combined, washed with brine and dried over MgSO$_4$. The solvent was removed under reduced pressure to yield the product as yellow oil that was directly used in the next step without further purification.

^{1}H-NMR (DMSO-d6, 400 MHz): δ 12.38 (s, 1H), 7.30 (t, ^3J = 4.3 Hz, 1H), 7.00 (d, ^3J = 7.7 Hz, 1H), 4.80 (d, ^3J = 2.4 Hz, 1H), 4.59 (d, ^3J = 2.2 Hz, 2H), 3.86-3.77 (m, 1H), 3.46-3.43 (m, 1H), 2.96 (q, ^3J = 5.9 Hz, 2H), 1.38 (s, 9H), 1.68-1.46 (m, 2H), 1.33-1.24 (m, 2H) ppm.

4.2.16 2-Amino-6-(((prop-2-yn-1-yloxy)carbonyl)amino)hexanoic acid hydrochloride (C8, PrLys)[195]
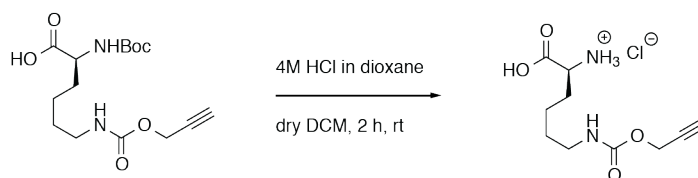
6-Propargylformylamino-Boc-L-lysine was dissolved in a solution of 30 mL dioxane and 2.4 mL of a 4 M HCl solution in dioxane. The solution was stirred for 4 h at room temperature. The solvent was removed under reduced pressure and the residue was dissolved in 80 mL of an aqueous 1 M HCl solution. The solvent was removed azeotrope under reduced pressure by addition of three times 25 mL toluene. The product (C8) was dried under vacuum to yield 1.7 g (55 % yield over two steps) of an off-white powder.

^{1}H-NMR (DMSO-d6, 400 MHz): δ 8.40 (br s, 2H), 7.36 (t, ^{3}J = 5.44 Hz, 1H), 4.60 (d, ^{3}J = 2.32 Hz, 2H), 3.85-3.84 (m, 1H), 3.48 (t, ^{3}J = 2.32 Hz, 1H), 2.98 (q, ^{3}J = 6.05 Hz, 2H), 1.79-1.78 (m, 2H), 1.41-1.38 (m, 3H) ppm.
Spectrum is shown in **Figure S18**.

HRMS(MALDI +): *m/z* calcd. for $C_{10}H_{16}N_2O_4$ [M+Na]$^+$ 251.10023, found 251.10030, Δm=0.00007, Δppm=0.3

4.2.17 Bicyclo[2.2.1]hept-5-en-2-ol[198]
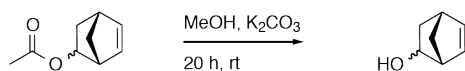
5-Norbornene-2yl acetate (10.0 g, 65.7 mmol, 1.0 eq) was added to a mixture of potassium carbonate (9.6 g, 69.7 mmol, 1.1 eq) in 100 mL dry methanol and the mixture was stirred for 20 h at room temperature. The solvent was removed under reduced pressure and the residue was dissolved in 300 mL H_2O. The aqueous solution was extracted five times with 100 mL ethyl acetate each. The organic phases were combined, washed with brine and dried over $MgSO_4$. The solvent was removed under reduced pressure to gain 6.2 g (86 % yield) of the product as off-white crystals.

^{1}H-NMR (CDCl₃, 400 MHz) δ 6.46-6.45 (m, 1H), 6.07-6.06 (m, 1H), 4.48-4.47 (m, 1H), 3.00 (s, 1H), 2.82 (s, 1H), 2.14-2.10 (m, 1H), 1.33-1.28 (m, 3H) ppm.

4.2.18 Bicyclo[2.2.1]hept-5-en-2-yl 2-(2,5-dioxopyrrolidin-1-yl)acetate[199]
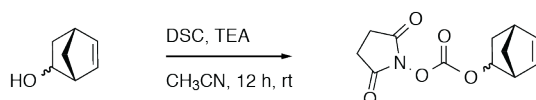
The norbornenol derivative (3.0 g, 27 mmol, 1 eq) was dissolved in 100 mL dry acetonitrile. Triethylamine (12 mL, 87 mmol, 3.2 eq) and *N,N'*-disuccinimidyl carbonate (11.3 g, 44 mmol, 1.2 eq) were added and the mixture was stirred for 12 h at room temperature. The solvent was removed under reduced pressure and the residue was dissolved in chloroform. The dissolved residue was adsorbed on silica gel and purified by column chromatography (DCM:Et₂O = 99.5:0.5, v/v) to gain 6.2 g (92 % yield) of the product as slightly yellow crystals.

^{1}H-NMR (CDCl₃, 250 MHz) δ 6.41-6.30 (m, 1H), 6.04-5.94 (m, 1H), 5.36 and 4.73 (m, 1H), 3.27 and 3.08 (m, 1H), 2.91 (s, 1H), 2.82 (s, 4H), 2.24-2.14 (m, 1H), 1.68-1.10 (m, 3H) ppm.

4.2.19 6-(((Bicyclo[2.2.1]hept-5-en-2-yloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino) hexanoic acid[199]
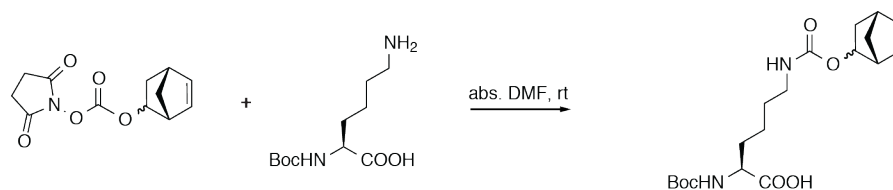
The activated norbornene (2.5 g, 10 mmol, 1 eq) was dissolved in 35 mL dry DMF and the Boc-protected lysine (3.2 g, 13 mmol, 1.3 eq) was added. The mixture was stirred over night at room temperature. The mixture was diluted with 300 mL H_2O and extracted three times with 150 mL ethyl acetate each. The organic phases were combined, washed three times with 150 mL H_2O each and once with 75 mL brine and dried over $MgSO_4$. The solvent was removed under reduced pressure, the residue was dissolved in DCM and the solvent was removed again under reduced pressure to yield the product as white foam. The foam was directly used in the next step.

^{1}H-NMR (CDCl$_3$, 400 MHz) δ 6.29-6.19 (m, 1H), 5.95-5.93 (m, 1H), 5.30-5.22 (m, 2H), 4.93-4.09 (m, 2H), 3.11 (br s, 2H), 2.80-2.79 (m, 1H), 2.09-2.08 (m, 1H), 1.82-1.30 (m, 15H), 0.88-0.87 (m, 1H) ppm.

4.2.20 2-Amino-6-(((bicyclo[2.2.1]hept-5-en-2-yloxy)carbonyl)amino)hexanoic acid hydrochloride (C9, NorK1)[195]
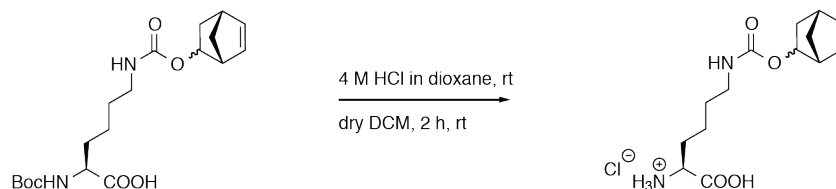
The Boc-protected norbornene lysine was dissolved in 25 mL dry DCM and 20 mL of a 4 M HCl solution in dioxane was added. The mixture was stirred for 1 h at room temperature and the solvent was removed under reduced pressure. The residue was diluted with Et_2O and filtrated. The solid was dried under vacuum to yield 3.1 g (97 % yield over two steps) of the product (**C9**) as colourless solid.

^{1}H-NMR (DMSO-d6, 400 MHz) δ 7.08-6.91 (m, 1H), 6.31-6.24 (m, 1H), 6.00-5.92 (m, 1H), 5.11-5.07 and 4.44-4.42 (m, 1H), 3.86-3.78 (m, 1H), 3.45-3.27 (m, 1H), 3.03 (s, 1H), 2.97-2.88 (m, 2H), 2.81-2.77 (m, 1H), 2.06-2.00 (m, 1H), 1.77-1.75 (m, 2H), 1.59-1.55 (m, 1H), 1.42-1.29 (m, 5H), 0.80-0.77 (m, 1H) ppm.
Spectrum is shown in **Figure S19**.

HRMS(MALDI +): *m/z* calc. for $C_{14}H_{22}N_2O_4$ [M+Na]$^+$ 305.14718, found 305.14741, Δm=0.00023, Δppm=0.8

4.2.21 Bicyclo[2.2.1]hept-5-en-2-ylmethyl   (2,5-dioxopyrrolidin-1-yl) carbonate[199]
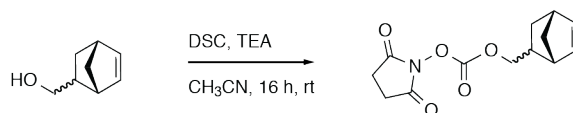
N,N'-Disuccinimidyl carbonate (DSC, 13.6 g, 53.6 mmol, 1.1 eq) was dissolved in 120 mL dry acetonitrile and 5-norbornene-2-methanol (5.84 mL, 48 mmol, 1 eq) was added. 24 mL triethylamine (TEA) were added and the mixture was stirred at room temperature overnight. The solvent was removed under reduced pressure and the residue was dissolved in dichloromethane (DCM). The dissolved residue was adsorbed on silica gel and purified by column chromatography (DCM:Et$_2$O = 99.5:0.5, v/v) to gain 10.2 g (80 % yield) of the product as colourless solid.

^{1}H-NMR (500 MHz, CDCl$_3$) δ 6.20 (dd, ^3J = 3.1 Hz, 5.7 Hz, 1H), 6.13-6.10 (m, 1H), 5.99 (dd, ^3J = 2.9 Hz, 5.7 Hz, 1H), 4.41 (dd, ^3J = 6.5 Hz, 6.4 Hz, 0.5H), 4.23 (m, 0.5H), 4.11 (dd, ^3J = 6.6 Hz, 10.3 Hz, 1H), 3.94 (dd, ^3J = 9.8 Hz, 10.1 Hz, 1H), 2.97 (s, 1H), 2.55-2.49 (m, 1H), 1.92-1.83 (m, 2H), 1.50 (dd, ^3J = 2.08 Hz, 8.4 Hz, 1H), 1.42-1.18 (m, 3H), 0.61-0.57 (m, 1H) ppm.

MS (ESI) *m/z*: [M+Na]$^+$ calc. for C$_{13}$H$_{15}$NO$_5$ 288.25, found 288.16.

4.2.22 6-(((bicyclo[2.2.1]hept-5-en-2-ylmethoxy)carbonyl)amino)-2-((tert-butoxycarbonyl) amino)hexanoic acid[199]
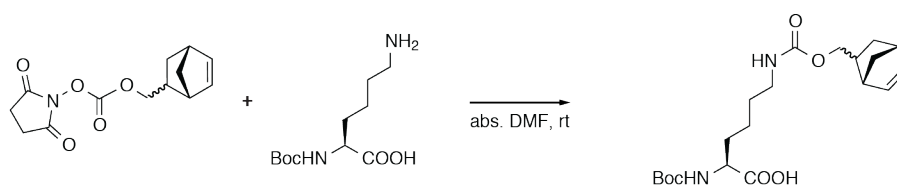
The activate norbornene ester (9.66 g, 36.4 mmol, 1 eq) was dissolved in 100 mL dry dimethylformamide (DMF) and BOC-protected lysine (9.88 g, 40.0 mmol, 1.1 eq) was added to the solution. The mixture was stirred at room temperature overnight. 400 mL

H_2O were added and the mixture was extracted with 500 mL ethyl acetate. The organic phase was washed with 150 mL H_2O and subsequently, the aqueous phase was extracted twice with 400 mL ethyl acetate. Organic phases were combined and again washed twice with 400 mL H_2O, 200 mL brine and dried over $MgSO_4$. The solvent was removed under reduced pressure and dried under vacuum to yield 14.37 g (99 % yield) of the slightly yellow foam.

^{1}H-NMR (500 MHz, $CDCl_3$) δ 6.17-12 (m, 1H), 6.11-6.06 (m, 1H), 5.95-5.94 (m, 1H), 5.29-5.27 (m, 1H), 4.36-4.26 (m, 1H), 3.20-3.17 (m, 3H), 2.98 (s, 1H), 2.89-2.85 (m, 1H), 1.92-1.80 (m, 3H), 1.57-1.51 (m, 3H), 1.48-1.39 (m, 14H) ppm.

MS (ESI) *m/z*: [M+H]$^+$ calc. for $C_{20}H_{32}N_2O_6$ 397.48, found 397.25.

4.2.23 2-Amino-6-(((bicyclo[2.2.1]hept-5-en-2-ylmethoxy)carbonyl)amino)hexanoic acid hydrochloride (C10, NorK2)[195]
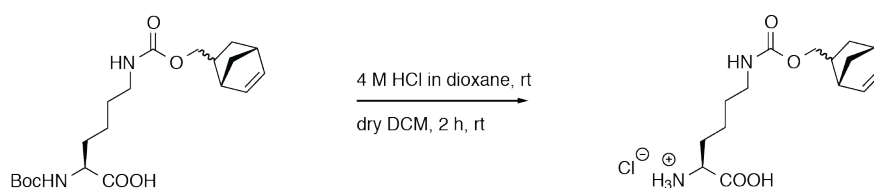
The BOC-protected intermediate (14.37 g, 36 mmol, 1 eq) was dissolved in 100 mL dry DCM and stirred at room temperature. 80 mL 4 N HCl in dioxane were added to the solution and the stirring was continued for 1 h. Evaporation under reduced pressure resulted in 10.42 g (87 % yield) of the non-natural amino acid NorK2 (**C10**) as colourless solid.

^{1}H-NMR (DSMO-d6, 250 MHz): The spectrum was difficult to interpret, but peaks matched with the literature.[200]

Spectrum is shown in **Figure S20**.

MS (ESI) *m/z*: [M+H]$^+$ calc. for $C_{15}H_{24}N_2O_4$ 297.36, found 297.20.

Chapter V: Appendix

5.1   Supplementary Figures
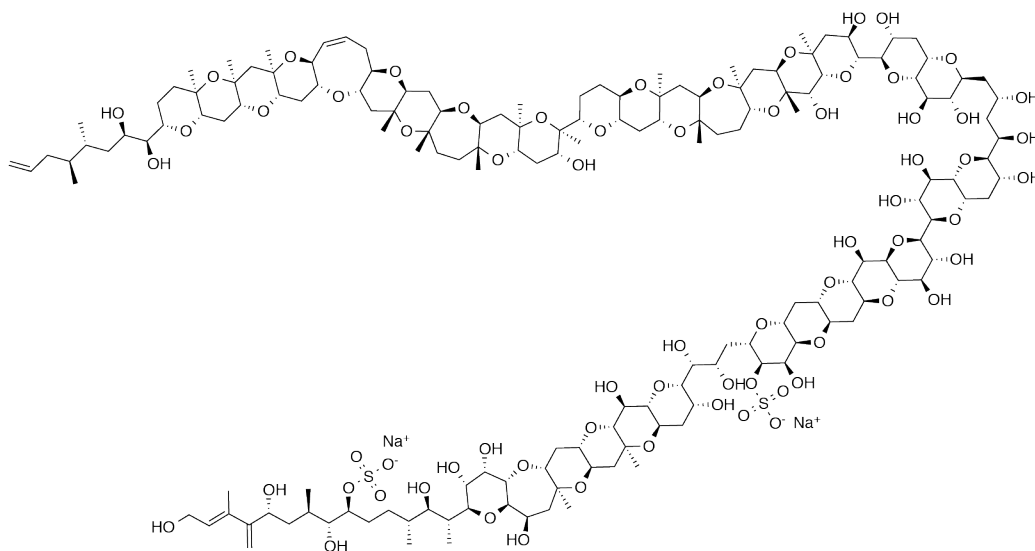
Figure S1: Chemical structure of Maitotoxin.

Multiple sequence alignment of FAS (fatty acid synthase) orthologues (FAS_MOUSE, FAS_RAT, FAS_HUMAN, FAS_PIG, FAS_CHICK) with a Conservation track shown beneath each block. Red stars (★) mark selected residues.

```
                 20         40         60         80        100
FAS_MOUSE MEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLKDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASL 100
FAS_RAT   MEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLKDLSKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAIVDGGINPASL 100
FAS_HUMAN MEEVVIAGMSGKLPESENLQEFWDNLIGGVDMVTDDDRRWKAGLYGLPRRSGKLKDLSRFDASFFGVHPKQAHTMDPQLRLLLEVTYEAIVDGGINPDSL 100
FAS_PIG   MEEVVIAGMSGKLPESENLEEFWANLIGGVDMVTADDRRWKAGLYGLPRRMGKLKDLSRFDASFFGVHSKQANTMDPQLRMLLEVTYEAIVDGGINPASL 100
FAS_CHICK MEDVVIAGIAGKLPESENLQEFWENLLNGVDMVTEDDRRWKPGIYGLPKRNGKLKDIKKFDASFFGVHPKQAHTMDPQLRLLLEVSYEAILDGGINPTAL 100
Conservation

                120        140        160 (★)    180        200
FAS_MOUSE RGTNTGVWVGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFFDFKGPSIALDTACSSSLLALQNAYQAIRSGECPAALVGGINLLLKPNTSVQF 200
FAS_RAT   RGTNTGVWVGVSGSEASEALSRDPETLLGYSMVGCQRAMMANRLSFFFDFKGPSIALDTACSSSLLALQNAYQAIRSGECPAAIVGGINLLLKPNTSVQF 200
FAS_HUMAN RGTHTGVWVGVSGSETSEALSRDPETLVGYSMVGCQRAMMANRLSFFFDFRGPSIALDTACSSSLMALQNAYQAIHSGQCPAAIVGGINVLLKPNTSVQF 200
FAS_PIG   RGTSTGVWVGVSSSDASEALSRDPETLVGYSMIGCQRAMMANLSFFFDFKGPSITIDTACSSSLLALQSAYQAIRGGECSAAVVGGLNVLLKPNSSLQF  200
FAS_CHICK RGTDTGVWVGASGSEALEALSQDPEELLGYSMTGCQRAMLANRISYFYDFTGPSLTIDTACSSSLMALENAYKARHGGCSAALVGGVNILLKPNTSVQF  200
Conservation

                220        240        260        280        300 (★)
FAS_MOUSE MKLGMLSPDGTCRSFDDSGSGYCRSEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGEVQEQLICSLYQPAGLAPESLEYIEAHGTGTKVG 300
FAS_RAT   MKLGMLSPDGTCRSFDDSGNGYCRAEAVVAVLLTKKSLARRVYATILNAGTNTDGCKEQGVTFPSGEAQEQLIRSLYQPGGVAPESLEYIEAHGTGTKVG 300
FAS_HUMAN LRLGMLSPEGTCKAFDTAGNGYCRSEGVVAVLLTKKSLARRVYATILNAGTNTDGFKEQGVTFPSGDIQEQLIRSLYQSAGVAPESFEYIEAHGTGTKVG 300
FAS_PIG   MKLGMLSQDGTCRSFDAEGTGYCRAEAVVAVLLTKKSLARRVYATILNAGTNTDGSKEQGVTFPSGDVQEQLIRSLYAPAGPDPESLEYIEAHGTGTKVG 300
FAS_CHICK MKLGMLSPDGACKAFDVSGNGYCRSEAVVVVLLTKKSMAKRVYATIVNAGSNTDGFKEQGVTFPSGEMQQQLVGSLYRECGIKPGDVEYVEAHGTGTKVG 300
Conservation

                320 (★)    340        360        380        400
FAS_MOUSE DPQELNGITRSLCAFRQAPLLIGSTKSNMGHPEPASGLAALTKVLLSLEHGVWAPNLHFHNPNPEIPALLDGRLQVVDRPLPVRGGNVGINSFGFGGSNV 400
FAS_RAT   DPQELNGITRSLCAFRQSPLLIGSTKSNMGHPEPASGLAALTKVLLSLENGVWAPNLHFHNPNPEIPALLDGRLQVVDRPLPVRGGIVGINSFGFGGANV 400
FAS_HUMAN DPQELNGITRALCATRQEPLLIGSTKSNMGHPEPASGLAALAKVLLSLEHGLWAPNLHFHSPNPEIPALLDGRLQVVDQPLPVRGGNVGINSFGFGGSNV 400
FAS_PIG   DPQELNGIVNALCATRREPLLIGSTKSNMGHPEPASGVAALIKVLLSLEHGVWAPNLHYHTPNPEIPALQDGRLQVVDRPLPIRGGNVGINSFGFGGSNV 400
FAS_CHICK DPQEVNGIVNVFCQCEREPLLIGSTKSNMGHPEPASGLAALAKVILSLEHGLWAPNLHFNDPNPDIPALHDGSLKVVCKPTPVKGGLVSINSFGFGGSNA 400
Conservation

                420        440        460        480        500
FAS_MOUSE HVILQPNTRQAPAPTAHAALPHLLHASGRTLEAVQDLLEQGRQHSQDLAFVSMLNDIAATPTAAMPFRGYTVLGVEGRVQEVQQVSTNKRPLWFICSGMG 500
FAS_RAT   HVILQPNTQQAPAPAPHAALPHLLHASGRTMEAVQGLLEQGRQHSQDLAFVSMLNDIAATPTAAMPFRGYTVLGVEGHVQAVQQVPASQRPLWFICSGMG 500
FAS_HUMAN HIILRPNTQPPPAPHAATLPRLLHASGRTPEAVQKLLEQGLRHSQDLAFLSMLNDIAAVPATAMPFRGYAVLGGEAGSQEVQQVPGSKRPVWFICSGMG  500
FAS_PIG   HVILQPNSRPAPPPAQHAALPRLLQASGRTLEAVQTLLEQGLRHSRDLAFVGMLNEIAAVSPVAMPFRGYAVLGGEAGSQEVQQVPGSKRPVWFICSGMG 500
FAS_CHICK HVILRPNEKKCQ-PQETCNLPRLVQVCGRTQEAVEILIEESRKHGGCSPFLSLLSDISAVPVSSMPYRGYTLVGTESDITEIQQVQASGRPLWYICSGMG 499
Conservation

                520        540        560        580 (★)    600
FAS_MOUSE TQWRGMGLSLMRLDSFRESILRSDEAVKPLGVKVSDLLLSTDERTFDDIVHAFVSLTAIQIALIDLLTSVGLKPDGIIGHSLGEVACGYADGCLSQREAV 600
FAS_RAT   TQWRGMGLSLMRLDSFRESILRSDEALKPLGVKVSDLLLSTDEHTFDDIVHSFVSLTAIQIALIDLLTSMGLKPDGIIGHSLGEVACGYADGCLSQREAV 600
FAS_HUMAN TQWRGMGLSLMRLDSFRDSILRSDEAVKPFGLKVSQLLLSTDESTFDDIVHSFVSLTAIQIALIDLLSCMGLRPDGIVGHSLGEVACGYADGCLSQEEAV 600
FAS_PIG   AQWQGMGLSLMRLDRFRDSILRSDQALKPLGLRVSDLLLSTDEAVLDDIVSSFVSLTSIQIALIDLLTSLGLQPDGIIGHSLGEVACGYADGCLTQEEAV 600
FAS_CHICK TQWKGMGLSLMLKLDLFRQSILRSDEALKSTGLKVSDLLLNADENTFDDTVHAFVGLAAIQIAQIDVLKAAGLQPDGILGHSVGELACGYADNSLSHEEAV 599
Conservation

                620        640        660        680 (★)    700
FAS_MOUSE LAAYWRGQCIKDAHLPPGSMAAVGLSWEECKQRCPAGVVPACHNSEDTVTISGPQAAVNEFVEQLKQEGVFAKEVRTGGLAFHSYFMEGIAPTLLQALKK 700
FAS_RAT   LAAYWRGQCIKDANLPAGSMAAVGLSWEECKQRCPPGVVPACHNSEDTVTISGPQAAVNEFVEQLKQEGVFAKEVRTGGLAFHSYFMEGIAPTLLQALKK 700
FAS_HUMAN LAAYWRGQCIKEAHLPPGAMAAVGLSWEECKQRCPPGVVPACHNSKDTVTISGPQAPVFEFVEQLKEGVFAKEVRTGGMAFHSYFMEAIAPPLLQELKK  700
FAS_PIG   LSSYWRGYCIKEANVLPGAMAAVGLSWEECKQRCPPGIVPACHNSKDTVTISGPQAAMSEFLQQLKREDVFVKEVRTGGIAFHSYFMESIAPTLLRQLRK 700
FAS_CHICK LAAYWRGRCVKEAKLPPGGMAAVGLTWEECKQRCPNVVPACHNSEDTVTVSGPLDSVSEFVTKLKKDGVFAKEVRRAGVAFHSYYMASIAPALLSALKK  699
Conservation

                720        740        760        780        800
FAS_MOUSE VIREPRPRSARWLSTSIPEAQWQSSLARTSSAEYNVNNLVSPVLFQEALWHIPEHAVVLEIAPHALLQAVLKRGVKSSCTIIPLMKRDHKDNLEFFLTNL 800
FAS_RAT   VIREPRPRSARWLSTSIPEAQWQSSLARTSSAEYNVNNLVSPVLFQEALWHVPEHAVVLEIAPHALLQAVLKRGVKPSCTIIPLMKRDHKDNLEFFLTNL 800
FAS_HUMAN VIREPKPRSARWLSTSIPEAQWHSSLARTSSAEYNVNNLVSPVLFQEALWHVPEHAVVLEIAPHALLQAVLKRGLKPSCTIIPLMKKDHRDNLEFFLAGI 800
FAS_PIG   VILDPKPRSKRWLSTSIPEAQWQGSLARTFSAEYSVNNLVSPVLFQEALQHVPAHAVVVEIAPHALLQAVLKRSLESSCTIIPLMKKDHRDNLEFFLSNV 800
FAS_CHICK VIPHPKPRSARWISTSIPESQWQSDLARNSSAEYHVNNLVNPVLFHEGLKHIPENAVVVEIAPHALLQAILRRTLKPTCTIIPLMKKDHKNNLEFFLTQT 799
Conservation

                820        840        860        880 (★)    900
FAS_MOUSE GKVHLTGINVNPNALFPPVEFPAPRGTPLISPHIKWDHSQTWDVPVAEDFPNGSSSSSATVYSIDASPESPDHYLVDHCIDGRVIFPGTGYLCLVWKTL 899
FAS_RAT   GKVHLTGIDINPNALFPPVEFPVPRGTPLISPHIKWDHSQTWDIPVAEDFPNGSSSSSATVYNIDASSESDHYLVDHCIDGRVLFPGTGYLYLVWKTL  899
FAS_HUMAN GRLHLSGIDANPNALFPPVEFPAPRGTPLISPLIKWDHSQTWDVPAAEDFPNGSGSPSAAIYNIDTSSESPDHYLVDHTLDGRVLFPATGYLSIVWKTL 899
FAS_PIG   GRLHLAGVSVNPNGLFPPVEFPAPRGTPLISPHXKWDHSQAWDVPSAADFPSGSSCSSVAVYKFDVSPESPDHYLVDHCIDGRVLFPGTGYLWLTWKTL 899
FAS_CHICK GKIHLTGINVLGNNLFPPVEYPVPVGTPLISPYIKWDHSQDWDVPKAEDFPSGSKGSASASVYNIDVSPDSPDHYLVGHCIDGRVLYPATGYLVLAWRTL 899
Conservation

                920        940        960        980       1,000
FAS_MOUSE ARSLGLSLEETPVVFENVSFHQATILPKTGTVALEVRLLEASHAFEVSD-TGNLIVSGKVYLWEDPNSKLFDHP--EVP-TPPESASVSRLTQGEVYKEL 995
FAS_RAT   ARSLSLSLEETPVVFENVTFHQATILPRTGTVPLEVRLLEASHAFEVSD-SGNLIVSGKVYQWEDPDSKLFDHP--EVP-IPAESESVSRLTQGEVYKEL 995
FAS_HUMAN ARALGLVEQLPVVFENVTFHQATILPKTGTVSLEVRLLEASHAFEVSE-NGNLVVSGKVYQWDDPDPRLFDHP--ESP-TPNPTEPLF-LAQAEVYKEL 994
FAS_PIG   ARALSQNLEETPVVFEDVTLHQATILPKTGTVSLEVRLLEASHAFEVSDSNGSLIASGKVYQWESPDPKLFDTR--AAV-DPADSTAEFRLSQGDVYKDL 996
FAS_CHICK ARSLGMVMEQTAVMFEEVTIHQATILPKKGSTQLEVRIMPASHSFEVS-GNGNLAVSGKISLLENDALKNFHNQLADFQ-SQANVTAKSGLLMEDVYQEL 997
Conservation

               1,020      1,040      1,060      1,080      1,100
FAS_MOUSE RLRGYDYGPQFQGICEATLEGEGGKLLWKDNWVTFMDTMLQVSILGSSQQSLQLPTRVTAIYIDPATHRQKVYRLKEDTQVADVTTSRCLGITVSGGIHI 1095
FAS_RAT   RLRGYDYGPHFQGVYEATLEGEGGKLLWKDNWVTFMDTMLQISILGFSKQSLQLPTRVTAIYIDPATHLQKVYMLEGDTQVADVTTSRCLGVTVSGGVYI 1095
FAS_HUMAN RLRGYDYGPHFQGILEASEGDSGRLLWKDNWVFMDTMLQMSILGSAKHGLYLPTRFTSIRIDPVTHRQKLYTLQDKAQVADVVVSRWLRVTVAGGVHI  1094
FAS_PIG   RLRGYDYGPFFQLVLESDLEGNRGRLQWNDSWVSFLDAMLHMSILAPGQLGLYLPTRFTSIRIDPVTHRQKLYTLQDTTQAADVVVDRNLNTVVAGGALF 1096
FAS_CHICK HLRGYNYGPTFQGVLECNSEGGSAGKILWNGNWVTFFLDTLLHLIVLAETGRSLRLPTRIRSVYIDPVLHQEQVYQYQDNVEAFDVVVDRCLDSLKAGGVQI 1097
Conservation

               1,120      1,140      1,160      1,180      1,200
FAS_MOUSE SRLQTTATSRRQQEQLVPTLEKFVFTPHMEAECLSESTALQKELQLCKGLARALQTKATQQGLKAAMLGQE--------DPPQHGLPRLLAAACQLQLNG 1187
FAS_RAT   SRLQTTATSRRQQEQLVPTLEKFVFTPHVEPECLSESAILQKELQLCKGLAKALQTKATQQGLKMTVPGLE--------DLPQHGGLPRLLAAACQLQLNG 1187
FAS_HUMAN SGLHTESAPRRQQEQQVPILEKFCFTPHTEEGCLSERAALQEELQLCKGLVQALQTTVTQQGLKMVYPGLD--GAQIPRDPSQQELPRLLSSAACRLQLNG 1192
FAS_PIG   LGAHSSVAPRRPQEHLKPILEKFCFTPHVESGCLAGNTALQEELQLCRGLAQALYPWQGLPGLD--QAQDPREAPQGSPLRLLAAACQLQLNG 1194
FAS_CHICK NGLHASVAPRRRQQERISPTLEKFSFVPYIESDCLSSSTQLHAYLEHCKGLIQKLQAKMALHGVKLVIHGLETKGAAAGSPPAQKGLQHILTEICRLELNG 1197
Conservation

               1,220      1,240      1,260      1,280      1,300
FAS_MOUSE NLQLELGEALAQERLLLPEDPLISGLLNSQALKACVDTALENLSTLKMKVAEVLAGEGHLYSRIPALLNTQPMLQLEYTATDRHPQALKDVQTKLQQHDV 1287
FAS_RAT   NLQLELGEVLARERLLLPEDPLISGLLNSQALKACIDTALENLSTLKMKVVEVLAGEGHLYSHISALLNTQPMLQLEYTATDRHPQALKDVQTKLQQHDV 1287
FAS_HUMAN NLQLELAQVLAGERPKLPEDPLLSGLLDSPALKACLDTAVENMPSLKMKVVEVLAGEGHLYSRIPGLLSPHPLLQLSYTATDRHPQALEAAQAELQQHDV 1292
FAS_PIG   NLQLELGQVLAQERPLLCDDPLLSGLLDAPALKACVDTALENMASPKMKVVEVLAGDGQLYSRIPALLNTQPVMDLDYTATDRNPQALEAAQAKLEQHV 1294
FAS_CHICK NPHSELEQIVTQEKMHLQDDPLLNGLLDSSELKTCLDVAKENTTSHRMKIVEALAGSGRLFSRVQSILNTQPLLQLDYIATDCTPETLSDNETELHDAGI 1297
Conservation
```

```
                 1,320          1,340          1,360          1,380          1,400
FAS_MOUSE  AQGGQWNPSDPAPSSLGALDLLVCNCALATLGDPALALDNMVAALKEGGFLLVHTVLKGHALGETLACLPS-EVQPAP-SLLSQEEWESLFSRKALHLVGL 1385
FAS_RAT    AQGGQWDPSGPAPTNLGALDLVVCNCALATLGDPALALDNMVAALKDGGFLLMHTVLKGHALGETLACLPS-EVQPGP-SFLSQEEWESLFSRKALHLVGL 1385
FAS_HUMAN  AQGGQWDPADPAPSALGSADLLVCNCAVAALGDPASALSNMVAALREGGFLLLHTLLRGHPLGDIVAFLTSTEPQYGQ-GILSQDAWESLFSRVSLRLVGL 1391
FAS_PIG    TQGGQWDPANPAPGSLGKADLLVCNCALATLGDPAVAVGNMAATLKEGGFLLHTLLAGHPLGEMVGFLTSPE-QGGR-HLLSQDQWESLFAGASLHLVAL 1392
FAS_CHICK  SFSQWDPSSLPSGNLTNADLAVCNCSTSVLGNTAEIISNLAAAVKEGGFVLLHTLLKEETLGEIVSFLTSPDLQQKH-SFLSQAQWEELFSKASLNLVAM 1396
Conservation

                 1,420          1,440          1,460          1,480          1,500
FAS_MOUSE  KRSFYGTALFLCRRAIPQEKPIFLSVEDTSFQWVDSLKSTLATS-SSQPVWLTAMDCPTSGVVGLVNCLRKEPGGHRIRCILLSNLSNTSHAPKLDPGSP 1484
FAS_RAT    KKSFYGTALFLCRRLSPQDKPIFLPVEDTSFQWVDSLKSILATS-SSQPVWLTAMNCPTSGVVGLVNCLRKEPGGHRIRCILLSNLSSTSHVPKLDPGSS 1484
FAS_HUMAN  KKSFYGSTLFLCRRPTPQDSPIFLPVDDTSFRWVESLKGILADEDSSRPVWLKAINCATSGVVGLVNCLRREPGGNRLRCVLLSNLSSTSHVPEVDPGSA 1491
FAS_PIG    KRSFYGSVLFLCRQQTPQDSPVFLSVEDTSFRWVDSLKDILADA-SSRPVWLMAVGCSTSGVVGMVNCLRKEPGGHRIRCVLVSNLSSTSPAPEMHPSSS 1491
FAS_CHICK  KRSFFGSVIFLCRRQSPAKAPILLPVDDTHYKWVDSLKEILADS-SEQPLWLTATNCGNSGILGMVNCLRLEAEGHRIRCVFVSNLSPSSTVPATSLSSL 1495
Conservation

                 1,520          1,540          1,560          1,580          1,600
FAS_MOUSE  ELQQVLKHDLVMNVYRDGAWGAFRHFQLEQDKPKEQTAHAFVNVLTRGDLASIRWVSSPLKHTQP-SSSGAQLCTVYYASLNFRDIMLATGKLSPDAIPG 1583
FAS_RAT    ELQKVLESDLVMNVYRDGAWGAFRHFQLEQDKPEEQTAHAFVNVLTRGDLASIRWVSSPLKHMQPPSSSGAQLCTVYYASLNFRDIMLATGKLSPDAIPG 1584
FAS_HUMAN  ELQKVLQGDLVMNVYRDGAWGAFRHFLLEEDKPEEPTAHAFVSTLTRGDLSSIRWVCSSLRHAQP-TCPGAQLCTVYYASLNFRDIMLATGKLSPDAIPG 1590
FAS_PIG    ELQKVLQGDLVMNVYRDGAWGAFRHFPLEQDRPEKQTEHAFVNVLSRGDLSSIRWVCSPLHYALPASCQD-RLCSVYYTSLNFRDVMLATGKLSPDSIPG 1590
FAS_CHICK  EMQKIIERDLVMNVYRDGKWGSFRHLPLQQAQPQELTECAYVNVLTRGDLSSLRWIVSPLRHFQT-TNPNVQLCKVYYASLNFWDIMLATGKLSPDAIPG 1594
Conservation

                 1,620          1,640          1,660          1,680          1,700
FAS_MOUSE  KWASRDCMLGMEFSGRDRCGRRVMGLVPAEGLATSVLLSSDFLWDVPSSWTLEEAASVPVVYTTAYYSLVVRGRIQRGETVLIHSGSGGVGQAAISIALS 1683
FAS_RAT    KWASRDCMLGMEFSGRDKCGRRVMGLVPAEGLATSVLLSPDFLWDVPSSWTLEEAASVPVVYTTAYYSLVVRGRIQHGETVLIHSGSGGVGQAAISIALS 1684
FAS_HUMAN  KWTSQDSLLGMEFSGRDASGKRVMGLVPAKGLATSVLLSPDFLWDVPSNWTLEEAASVPVYSTAYYALVVRGRVRPGETLLIHSGSGGVGQAAIAIALS 1690
FAS_PIG    KWLTRDCMLGMEFSGRDASGRRVMGMVPAEGLATSVLLLQHATWEVPSTWTLEEAASVPIVYTTAYYSLVVRGRMQPGESVLIHSGSGGVGQAAIAIALS 1690
FAS_CHICK  NWTLQQCMLGMEFSGRDLAGRRVMGLLPAKGLATVVDCDKRFLWEVPENWTLEEAASVPVVYATAYYALVVRGGMKKGESVLIHSGSGGVGQAAIAIALS 1694
Conservation

                 1,720          1,740          1,760          1,780          1,800
FAS_MOUSE  LGCRVFTTVGSAEKRAYLQARFPQLDDTSFANSRDTSFEQHVLLHTGGKGVDLVLNSLAEEKLQASVRCLAQHGRFLEIGKFDLSNNHPLGMAIFLKNVT 1783
FAS_RAT    LGCRVFTTVGSAEKRAYLQARFPQLDDTSFANSRDTSFEQHVLLHTGGKGVDLVLNSLAEEKLQASVRCLAQHGRFLEIGKFDLSNNHPLGMAIFLKNVT 1784
FAS_HUMAN  LGCRVFTTVGSAEKRAYLQARFPQLDSTSFANSRDTSFEQHVLWHTGGKGVDLVLNSLAEEKLQASVRCLATHGRFLEIGKFDLSQNHPLGMAIFLKNVT 1790
FAS_PIG    RGCRVFTTVGSAEKRAYLQARFPQLDETCFANSRDTSFEQHVLRHTAGKGVDLVLNSLAEEKLQASVRCLAQHGRFLEIGKFDLSNNHALGMAVFLKNVT 1790
FAS_CHICK  MGCRVFATVGSAEKREYLQARFPQLDANSFASSRNTTFQQHILRVTNGKGVSLVLNSLAEEKLQASLRCLAQHGRFLEIGKFDLSNNSQLGMALFLKNVA 1794
Conservation

                 1,820          1,840          1,860          1,880          1,900
FAS_MOUSE  FHGILLDALFEEANDSWREVAALLKAGIRDGVVKPLKCTVFPKAQVEDAFRYMAQGKHIGKVLVQVREEEPEAVLPGAQPTLISAISKTFCPAHKSYIIT 1883
FAS_RAT    FHGILLDALFEGANDSWREVAELLKAGIRDGVVKPLKCTVFPKAQVEDAFRYMAQGKHIGKVLVQVREEEPEAMLPGAQPTLISAISKTFCPEHKSYIIT 1884
FAS_HUMAN  FHGVLLDAFFNESSADWREVWALVQAGIRDGVVRPLKCTVFHGAQVEDAFRYMAQGKHIGKVVVQVREEEQGPAPRGLPPIALTGLSKTFCPPHKSYVIA 1890
FAS_PIG    FHGILLDSLFEEGGATWQEVSELLKAGIQEGVVQPLKCTVFPRTKVEAAFRYMAQGKHIGKVVIQVREEEQGPAPRGLPPIALTGLSKTFCPPHKSYVIT 1890
FAS_CHICK  FHGILLDSIFEEGNQEWEVVSELLTKGIKDGVVKPLRTTVFGKEEVEAAFRFMAQGKHIGKVMIKIQEEEKQYPLR-SEPVKLSAISRTSCPPTKSYIIT 1893
Conservation

                 1,920          1,940          1,960          1,980          2,000
FAS_MOUSE  GGLGGFGLELARWLVLRGAQRLVLTSRSGIRTGYQAKHIREWRRQGIQVLVSTSNVSSLEGARALIAEATKLGPVGGVFNLAMVLRDAMLENQTPELFQD 1983
FAS_RAT    GGLGGFGLELARWLVLRGAQRLVLTSRSGIRTGYQAKHVREWRRQGIHVLVSTSNVSSLEGARALIAEATKLGPVGGVFNLAMVLRDAMLENQTPEFFQD 1984
FAS_HUMAN  GGLGGFGLELAQWLIQRGVQKLVLTSRSGIRTGYQARVRRWRRQGVQVQVSTSNISSLEGARGLIAEAAQLGPVGGVFNLAVVLRDGLLENQTPEFFQD 1990
FAS_PIG    GGLGGFGLQLAQWLRLRGAQKLVLTSRSGIRTGYQARQVREWRRQGVQVLVSTSNASSLDGARSLITEATQLGPVGGVFNLAMVLRDAVLENQTPEFFQD 1990
FAS_CHICK  GGLGGFGLELAQWLIERGAQKLVLTSRSGIRTGYQAKCVREWKALGIQVLVSTSDVGTLEGTQLLIEEALKLGPVGGIFNLAVVLKDAMIENQTPELFWE 1993
Conservation

                 2,020          2,040          2,060          2,080          2,100
FAS_MOUSE  VNKPKYNGTLNLDRATREACPELDYFVAFSSVSCGRGNAGQTNYGFANSTMERICEQRRHDGLPGLAVQWGAIGDVGIVLEAMGTNDTVIGGTLPQRISS 2083
FAS_RAT    VNKPKYNGTLNLDRATREACPELDYFVAFSSVSCGRGNAGQSNYGFANSTMERICEQRRHDGLPGLAVQWGAIGDVGIILEAMGTNDTVVGGTLPQRISS 2084
FAS_HUMAN  VCKPKYSGTLNLDRVTREACPELDYFVVFSSVSCGRGNAGQSNYGFANSAMERICEKRRHEGLPGLAVQWGAIGDVGLVETMSTNDTIVSGTLPQRMAS 2090
FAS_PIG    VSKPKYSGTANLDRVTREACPELDYFVIFSSVSCGRGNAGQANYGFANSAMERICEKRRHDGLPGLAVQWGAIGDVGVVLETMGTNDTVIGGTLPQRIAS 2090
FAS_CHICK  VNKPKYSGTLHLDWVTRKKCPDLDYFVVFSSVSCGRGNAGQSNYGFANSAMERICEQRHHDGLPGLAVQWGAIGDVGI-LKAMGNREVVIGGTVLQQISS 2092
Conservation

                 2,120          2,140          2,160          2,180          2,200
FAS_MOUSE  CMEVLDLFLNQPHAVLSSFVLAEKKAVAHGDGDTQRDLVKAVAHILGIRDLAGINLDSTLADLGLDSLMGVEVRQILEREHDLVLPMREVRQLTLRKLQE 2183
FAS_RAT    CMEVLDLFLNQPHAVLSSFVLAEKKAVAHGDGEAQRDLVKAVAHILGIRDLAGINLDSSLADLGLDSLMGVEVRQILEREHDLVLPIREVRQLTLRKLQE 2184
FAS_HUMAN  CLEVLDLFLNQPHMVLSSFVLAEK-AAAYRDRDSQRDLVEAVAHILGIRDLAAVNLDSSLADLGLDSLMSVEVRQTLERELNLVLSVREVRQLTLRKLQE 2189
FAS_PIG    CLEVLDLFLNQPHVLSSFVLAEKKAAAAPRDGSSQKDLVKAVAHILGIRDVASINPDSTLVDLGLDSLMGVEVRQILEREHDLVLSMREVRQLSLRKLQE 2190
FAS_CHICK  CLEVLDMFLNQPHPVMSSFVLAEKVSVK-SEGGSQRDLVEAVAHILGVRDVSSLNAESSLADLGLDSLMGVEVRQTLERDYDIVMTMREIRLLTINKLRE 2191
Conservation

                 2,220          2,240          2,260          2,280          2,300
FAS_MOUSE  MSSKTDSATDTTAPKSRSD-TSLKQNQLNLSTLLVNPEGPTLTQLNSVQSSERPLFLVHPIEGSTTVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLAAYY 2282
FAS_RAT    MSSKAGSDTELAAPKSKND-TSLKQAQLNLSILLVNPEGPTLTRLNSVQSSERPLFLVHPIEGSTIVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLAAYY 2283
FAS_HUMAN  LSSKADEASELACPTPKEDGLAQQQTQLNLRSLLVNPEGPTLMRLNSVQSSERPLFLVHPIEGSTTVFHSLASRLSIPTYGLQCTRAAPLDSIHSLAAYY 2289
FAS_PIG    LSSKTSTDADPATPTSHEDSPVRQQATLNLSTLLVNPEGPTLTRLNSVQSSERPLFLVHPIEGSITVFHSLAAKLSIPTYGLQCTQAAPLDSIQSLASYY 2290
FAS_CHICK  LSSKTGTAEELK-PSQVLKTGPGEPPKLDLNNLLVNPEGPTITRLNEVQSTERPLFLVHPIEGSIAVFYTLASKLHMPCYGLQCTKAAPLDSIQSLASYY 2290
Conservation

                 2,320          2,340          2,360          2,380          2,400
FAS_MOUSE  IDCIKQVQPEGPYRIAGYSFGACVAFEMCSQLQAQGGPAPTHNNLFLFDGSHTYVLAYTQSYRAKMTPGCEAEAEAEALCFFIKQFLDVEHSKVLEALLP 2382
FAS_RAT    IDCIKQVQPEGPYRVAGYSFGACVAFEMCSQLQAQGGPAPAHNNLFLFDGSHTYVLAYTQSYRAKLTPGCEAEAEAEAICFFIKQFVDAEHSKVLEALLP 2383
FAS_HUMAN  IDCIRQVQPEGPYRVAGYSYGACVAFEMCSQLQAQGSPAPTHNSLFLFDGSPTYVLAYTQSYRAKLTPGCEAEAEETEAICFFVQQFTDMEHNRVLEALLP 2389
FAS_PIG    IECIRQVQPEGPYRIAGYSYGACVAFEMCSQLQAQGSATPGNHSLFLFDGSHSFVLAYTQSYVRAKMTPGCEAEAEAKAMYFFVQQFTDMEQGKVLEALIP 2390
FAS_CHICK  IDCMKQIQPEGPYRIAGYSFGACVAFEMCSQLQAQGGQNASHALNSLFLFDGSHSFVAAYTQSYRAKLTQGNEAALETEALCAFVQQFTGIEYNKLLEILLP 2390
Conservation

                 2,420          2,440          2,460          2,480          2,500
FAS_MOUSE  LKSLEDRVAASVDLITKSHHSLDRRELSFAAVSFYHKLRAADQYKPKAKYHGNVTLLRAKTGGTYGEDLGADYNLSQVCDGKVSVHIIEGDHRTLLEGSG 2482
FAS_RAT    LKSLEDRVAAAVDLITRSHQSLDRRDLSFAAVSFYYKLRAADQYKPKAKYHGNVILLRAKTGGTYGEDLGADYNLSQVCDGKVSVHIIEGDHRTLLEGSG 2483
FAS_HUMAN  LKGLEERVAAAVDLIITRSHQGLDRQELSFAARSFYYKLRAAEQYTPKAKYHGNVMLLRAKTGGAYGEDLGADYNLSQVCDGKVSVHVIEGDHRTLLEGSG 2489
FAS_PIG    LQGLEARVAATVDLITQSHAGLDRHALSFAARSFYYKLRAAENYWPQATYHGNVTLLRAKTGGAYGEDLGADYNLSQVCDGKVSVHVIEGDHRTLLEGSG 2489
FAS_CHICK  LEDLEARVNAAADLITQIHKNINREALSFAAASFYHKLKAADKYIPESKYHGNVTLMRAKTHNEYEEGLGGDYRLSEVCDGKVSVHIIEGDHRTLLEGDG 2490
Conservation

                 2,520
FAS_MOUSE  LESIINIIHSSLAEPRVSVREG 2504
FAS_RAT    LESIINIIHSSLAEPRVSVREG 2505
FAS_HUMAN  LESIISIIHSSLAEPRVSVREG 2511
FAS_PIG    LESILSIIHSCLAEPRVSVREG 2512
FAS_CHICK  VESIIGIIHGSLAEPRVSVREG 2512
Conservation
```

Figure S2: Sequence alignment of exemplary animal FASs. Uniprot accession codes: murine FAS: P19096; rat FAS: P12785; human FAS: P49327; porcine FAS: A5YV76 and chicken FAS: P12276. Primary sequences were aligned with CLC Main Workbench using the algorithm of Clustal Omega.[201] Domain boarders are given below the sequences using the typical color code and red stars indicate active site residues.
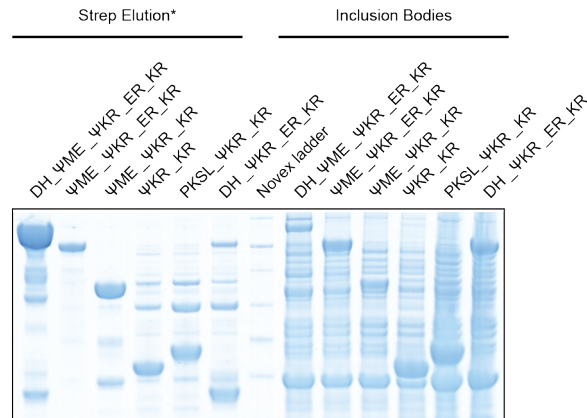
Figure S3: Commassie-stained SDS-PAGE of the processing part. Comparison of soluble and aggregated fractions of various truncated constructs. Asterisk indicates that exceptionally only a purification with a Strep-Tactin column was used explaining the relatively high degree of contamination.

```
AveAT0    ------MQRMDGGEEPRPAAGEVL--------GVADEADGGVVFVFPGQGPQWPGMGREL -46
MATload   MSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKGAGSSTNKRPLWFICSGMGTQWRGMGLSL -510
               *   .** .   ::.*..       . :.   . : *:  * * ** *** .*


AveAT0    LDASDVFRESVRACEAAFAPYVDWSVEQVLRDSPDAPGLDRVDVVQPTLFAVMISLAALW -106
MATload   MRL-DSFRESILRSDEAVKPLGVKVSDLLL--STDERTFDDIVHAFVSLTAIQIALIDLL -567
          :    * ****: .: *. *      : :*  * *   :*: .  :* *: *:*  *

                    ↓
AveAT0    RSQGVEPCAVLGHSLGEIAAAHVSGGLSLADAARVVTILWSQAQTTLAGTGALVSVAATPD -166
MATload   TSVGLKPDGIIGHSLGEVACGYADGCLSQREAV-LAAYWRGQCIKD--------AHLPPG -618
            * *::*  .:::*****:*:**..:..* **   :*. .:: *       .     . *.


AveAT0    ELLPRIAPWTED---NPARLAVAAVNGPRSTVVSGAREAVADLVADLTAAQVRTRMIP-V -222
MATload   SMAAVGLSWEECKQRCPAGVVPACHNSEDTVTISGPQAAVNEFVEQLKQEGVFAKEVRTG -678
          .:        * *    ** :. *. *.  :..:** : ** ::* :*.    * :: :

                    ↓
AveAT0    DVPAHSPLMYAIEERVVSGLLPI--TPRPSRIPFHSSV-TGGRL---DTRELDAAYWYRN -276
MATload   GLAFHSYFMEGIAPTLLQALKKVIREPRPRSARWLSTSIPEAQWQSSLARTSSAEYNVNN -738
          .:  ** :* .*   ::..* :   ***    : *:   .:     :* .* *  .*


AveAT0    MSSTVRFEPAARLLLQQGPKTFVEMSPHPVLTMGLQELAPDLGDTTGTADTVIMGTLRRG -336
MATload   LVSPVLFQEALWHI--PEHAVVLEIAPHALLQAVLKRG-------VKSSCTIIPLMKRDH -789
          : * * *: *   :         ..:*::** :*   *:.      . :: *:*      *


AveAT0    QGTLDHFLTSLAQLRGHGETSATT---VLS
MATload   KDNLEFFLTNLGKVHLTGIGGGSGGGSKIE
          :..*:.***.*.:::   *   ..:    :.
```

Figure S4: Sequence alignment of MAT with avermectin loading domain (AT0$_{AVES}$). The transferase fold of AT0$_{AVES}$ (pdb: 4rl1),[24] and its equivalent region in MAT is highlighted in green starting with β-sheet (AT0$_{AVES}$: 28-32, MAT: 492-496) and ending with α-helix (AT0$_{AVES}$: 339-352, MAT: 792-806). The N-terminal of the ACP0$_{AVES}$ domain (based on NMR structure of homologous ACP2 of DEBS (pdb: 2ju2),[96] and first three residues of the C-terminal MBP-fusion are highlighted in purple and blue respectively. Identical residues constitute the active site catalytic dyad (AT0$_{AVES}$: S120 and H227; MAT: S581 and H683) (black arrows). Residue R606 of MAT, necessary for transferring extender substrates, is absent in AT0$_{AVES}$ and was replaced by W145 (unfilled boxes) as identified by high resolution structures. Uniprot accession codes: murine FAS (*FASN*): P19096; AVES1 (*aveA1*): Q9S0R8. Sequence alignment was generated with the help of Clustal Omega within https://www.ebi.ac.uk/Tools/msa/clustalo.[201]
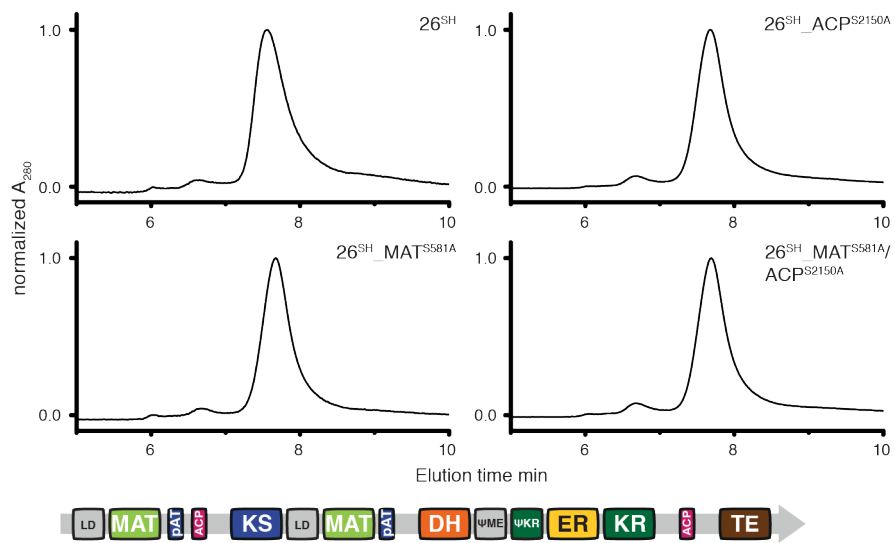
Figure S5: Analytical SEC of tested variants with HPLC to proof structural integrity of all samples. Three knockouts of construct 26^{SH} (LD-MAT-ACP-mFAS) with MAT^{S581A}, ACP^{S2150A} and MAT^{S581A} plus ACP^{S2150A} were investigated. Peaks correspond to a dimeric oligomeric state.
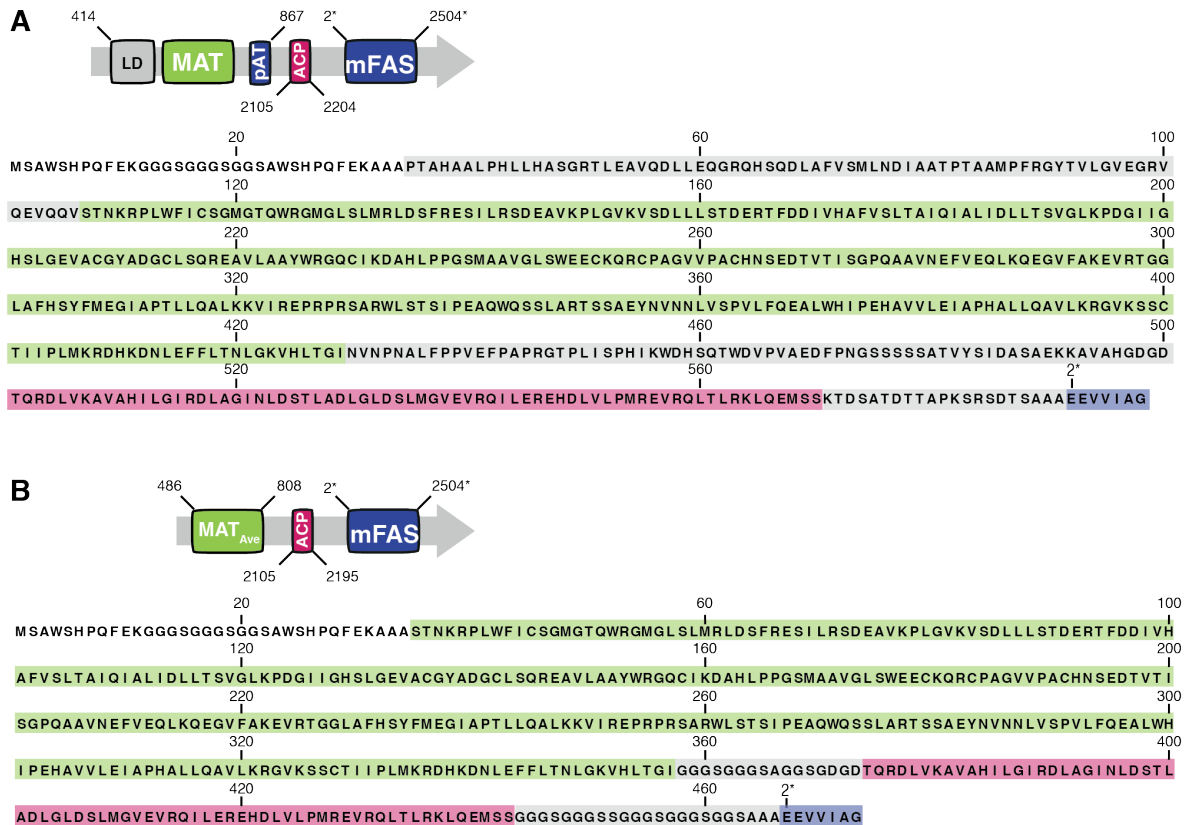
Figure S6: Organization of bimodular systems resembling loading-module–module1 assemblies. (A) Sequence of LD-MAT-ACP-mFAS (**26**). This construct was created by solely using murine FAS sequence. The linker between MAT and ACP was created by using the post-AT linker plus the first β-sheet of DH and connect it to the KR-ACP linker. The relative long linker was chosen as keeping the linker domain results in a long distance to the active site of MAT. The linkage between ACP and KS is created by the ACP-TE linker plus three alanines. (B) Sequence of MAT$_{Ave}$-ACP-mFAS (**29**). This construct contains the minimal MAT$_{Ave}$ fold resembling AT0$_{AVES}$. Linkers between MAT$_{Ave}$ and ACP, as well as the linker between ACP and KS are artificial GGS linkers with typical lengths of such organizations. As loading didomains are designed from mFAS, we have maintained the original residue numbering for the individual parts. The sequence of the integral mFAS is indicated with a star in superscript (*).
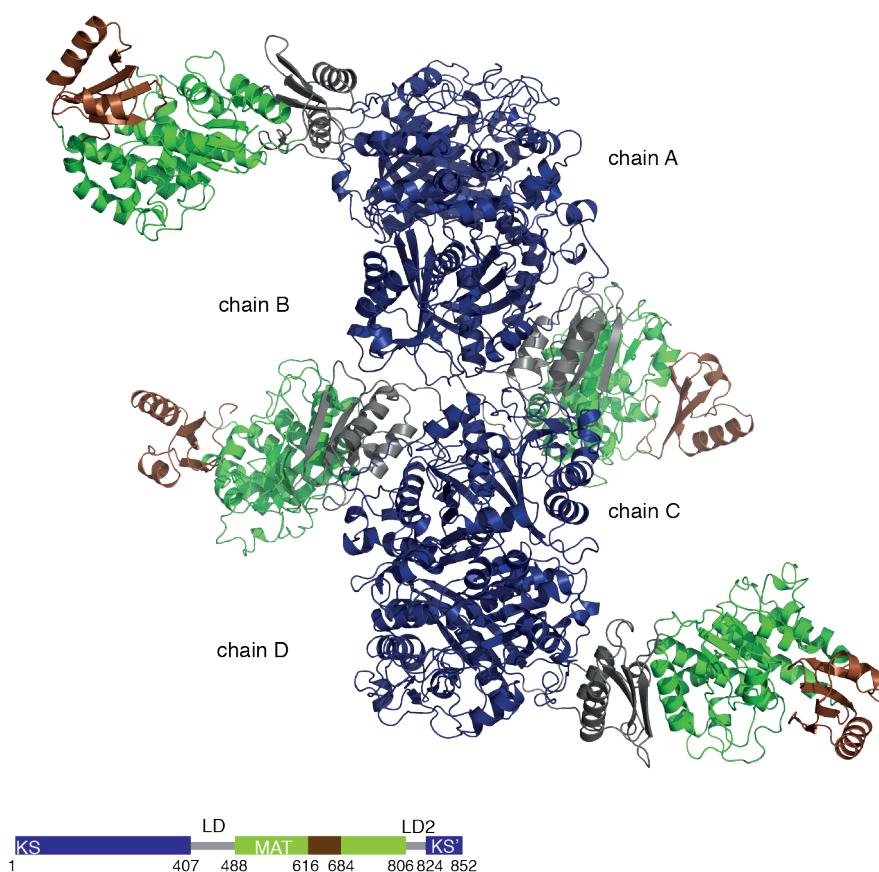
Figure S7: Arrangement of four polypeptide chains in the unit cell (pdb: 5my0). Domains and folds are color coded as depicted in the attached cartoon.
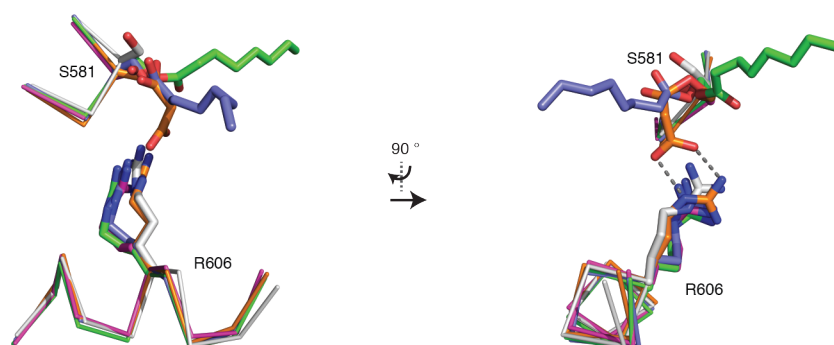
Figure S8: Rotameric states of R606 found in MAT. α/β-Hydrolase based superposition (BB of 488-615) of chain A (blue), chain C (white) and chain D (green) from Oc-CoA soaked crystals (6exg) with chain D (orange) of malonyl-CoA soaked crystals (5my0) and chain A (purple) of porcine FAS (2vz9). Important residues S581 and R606 are shown in sticks with covalent modifications of the serine represented also in sticks. For clarity residue stretch 580-583 and 601-610 are depicted as ribbons. R606 in octanoyl-bound active sites adopts the same conformation as R606 in the porcine FAS, whereas R606 in chain C (unbound) possesses the rotameric state of the unbound active sites previously found in human KS-MAT (3hhd).[64]

A

B

Constructs: pAR258b: StrI- and H(8)-tag
pAR355: GST- and H(8)-tag
pAR366: GST- and GST-H(8)-tag

C

Domain organization of construct 355:

covalently fused

D

E

F

G

H

I
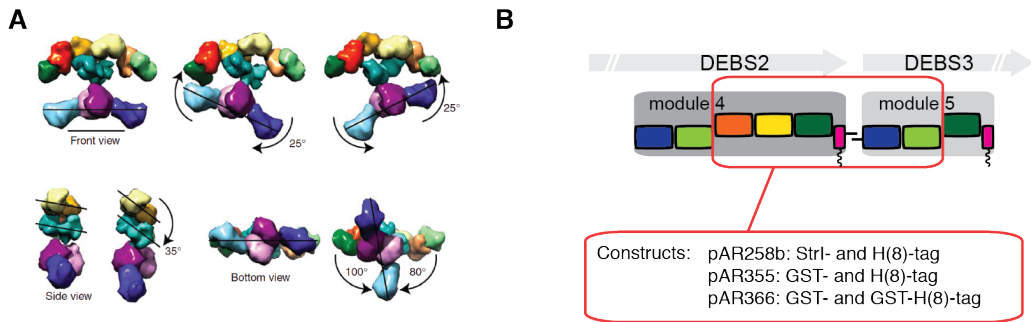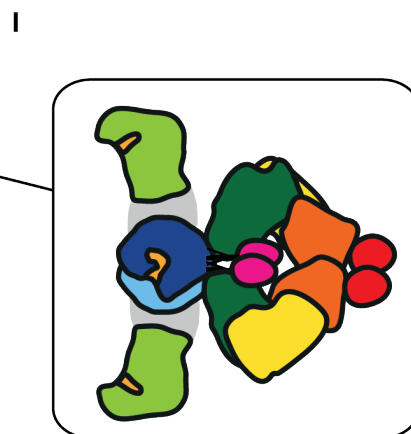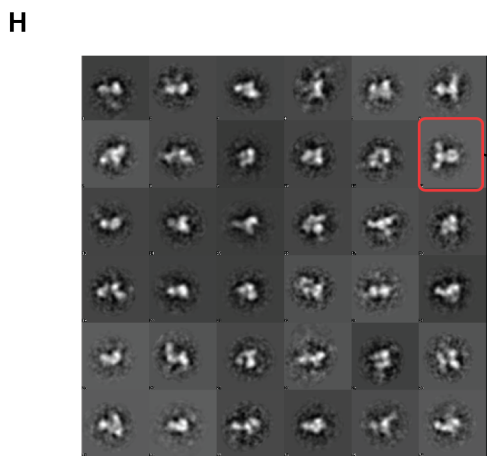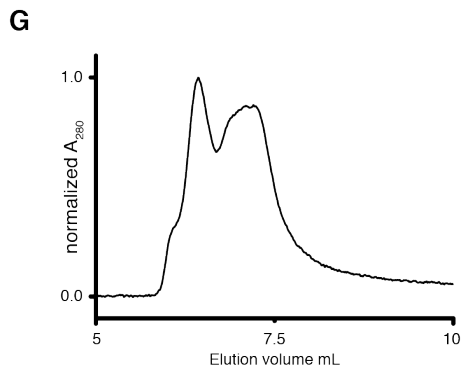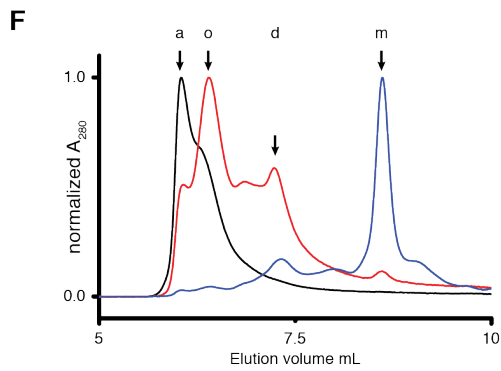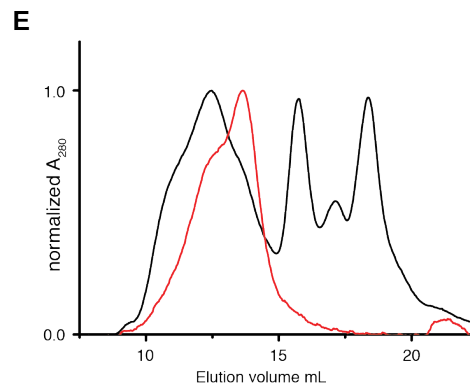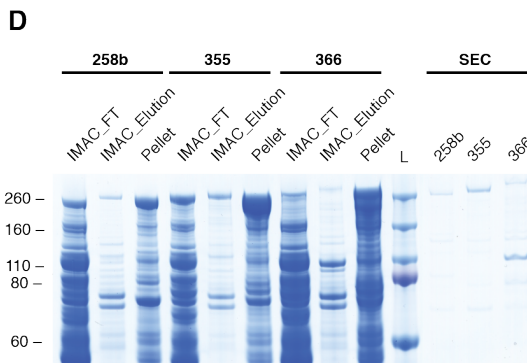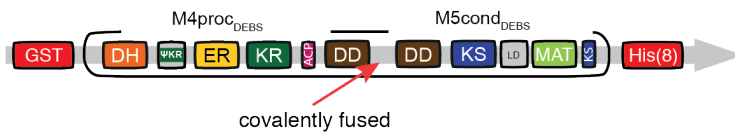
Figure S9: Preliminary data on structural interaction of module 4 and 5 in DEBS. (A) Tremendous flexibility found within the FAS fold structurally exemplifying module organization in PKSs. Data from Brignole *et al.*, who analyzed the rat FAS by single particle EM (negative stain).[54] (B) Overview on designed constructs of M4proc/M5cond_DEBS with different tag strategies. Based on the hypothesis that most of the flexibility originates within a module, designed constructs do not contain full modules to facilitate structural elucidation. (C) Domain organization of construct 355, which yielded the best protein quality. Docking domains (DD) of both parts were covalently fused as reported by Broadhurst *et al.*[163] (D) Representative preliminary purification strategy of three designed constructs. Coomassie-stained SDS-PAGE (NuPage 4-12 % bis-tris gradient gels) confirmed reasonable purity after IMAC plus SEC. (E) Preparative SEC (S6 10/300 column) of elution fraction from IMAC with construct 355 (black) with absorption normalized to the highest peak. At least four oligomeric species were found in the elution fraction. Red curve shows a second preparative SEC (S6 10/300 column) run after preliminary purification via SEC with a S6 16/70 column (old). (F) Analytical SEC with HPLC (Yarra 300x) to analyze different fractions from the preparative SEC run: Fr.1 (black), Fr. 2 (red) and Fr. 3 (blue). Arrows indicate found oligomeric states: m – monomeric, d – dimeric, o – higher oligomeric, a – aggregates. (G) Analytical SEC with HPLC (Yarra 300x) of the protein preparation, which was analyzed by negative stain. It shall be mentioned, that different oligomeric states were found due to insufficient purification or instability of the sample. (H) 2D classification of 700 particles from negative stain images using the software IMAGIC. Grid preparation, imaging and processing of the data was performed by Edoardo D'Imprima at the MPI of Biophysics, Frankfurt. (I) Postulated, possible arrangement of domains in construct 355. Cartoon depiction using the common color code shown in (C).
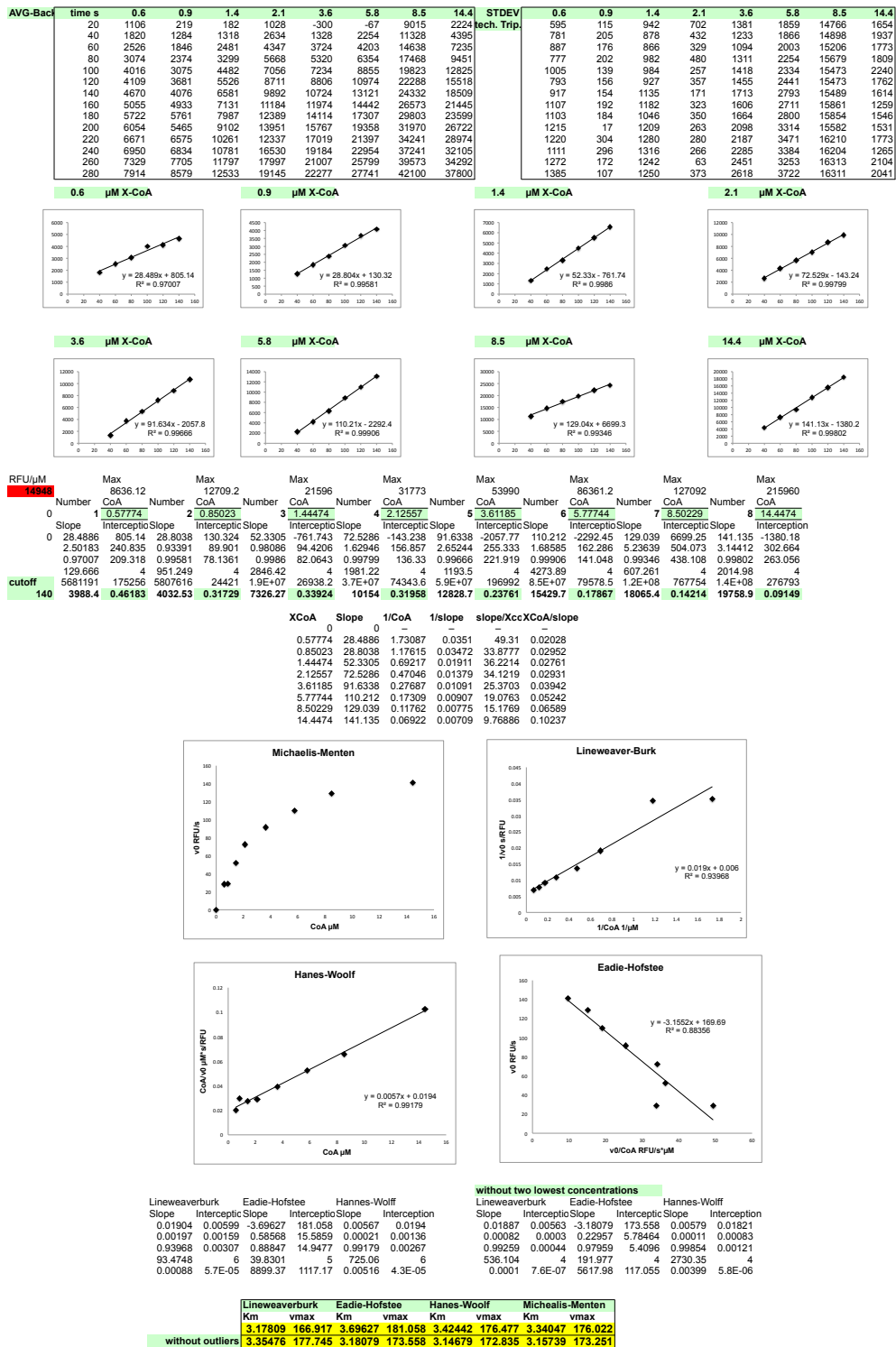
Figure S10: Representative analysis of one measurement in triplicates for specific transferase activity using acetyl-CoA. Initial velocities were determined for eight concentrations (0.2 K_m; 0.3 K_m; 0.5 K_m; 0.75 K_m; 1.25 K_m; 2 K_m; 3 K_m; 5 K_m) by linear regression of the data points from 40-140 s. Data was plotted by four methods: Michaelis-Menten (v_0 *vs.* [Ac-CoA]), Hanes-Woolf ([Ac-CoA]/v_0 *vs.* [Ac-CoA]), Lineweaver-Burk ($1/v_0$ *vs.* 1/[Ac-CoA]), Eadie-Hofstee (v_0 *vs.* v_0/([Ac-CoA]). Kinetic parameters were determined by linear regression with the help of the program Microsoft Excel and the sigmoidal Michaelis–Menten kinetics-fits with Origin.

5.2 Supplementary Equations

5.2.1 Michaelis-Menten Equation

$$v = \frac{V_{max}[S]}{K_m + [S]} \qquad\qquad V_{max} = k_{cat}[E]_0 \qquad\qquad [\,8\,]$$

5.2.2 Lineweaver-Burk Equation

$$\frac{1}{v} = \frac{K_m}{V_{max}} \frac{1}{[S]} + \frac{1}{V_{max}} \qquad\qquad [\,9\,]$$

5.2.3 Hanes-Woolf Equation

$$\frac{[S]}{v} = \frac{[S]}{V_{max}} + \frac{K_m}{V_{max}} \qquad\qquad [\,10\,]$$

5.2.4 Eadie-Hofstee Equation

$$v = V_{max} - K_m \frac{v}{[S]} \qquad\qquad [\,11\,]$$

5.2.5 Hill Equation

$$v = \frac{V_{max}[S]^h}{K' + [S]^h} \qquad\qquad [\,12\,]$$

Where the coefficient h is referred to as the Hill coefficient. The constant K' is different to the Michaelis constant and does no longer relate to the concentration of substrate required to attain half of the maximal velocity.

5.3 NMR Spectra of Compounds C1-C10
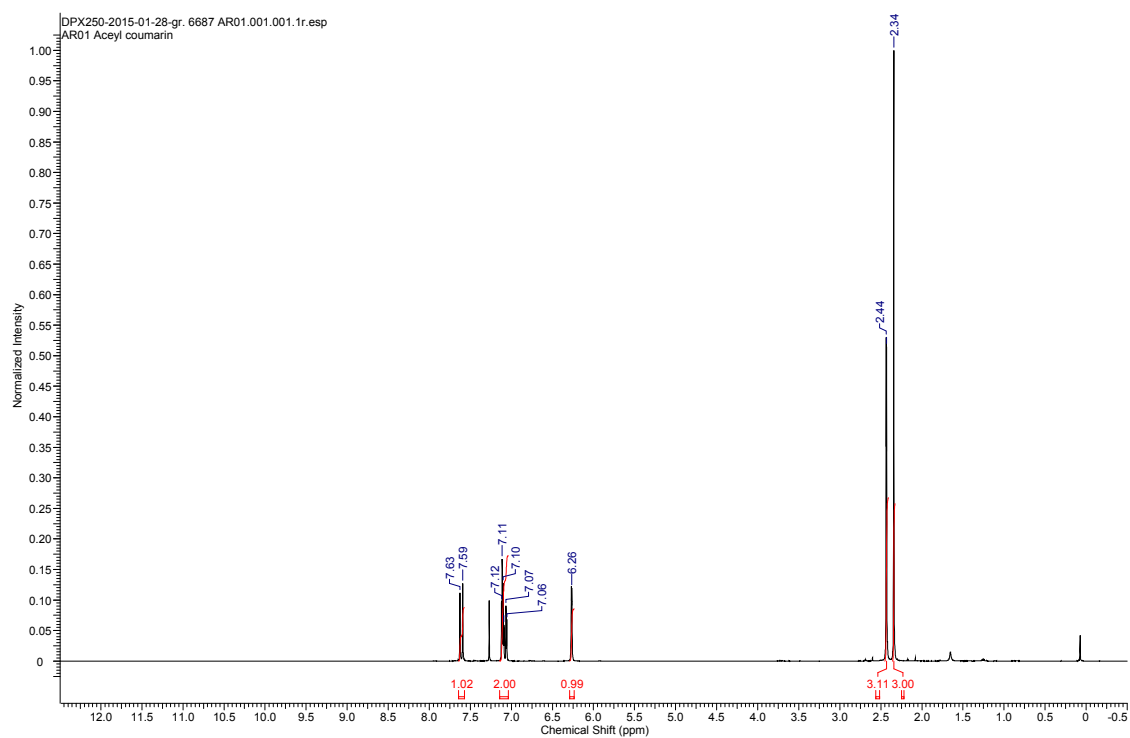
Figure S11: NMR spectrum of compound C1.
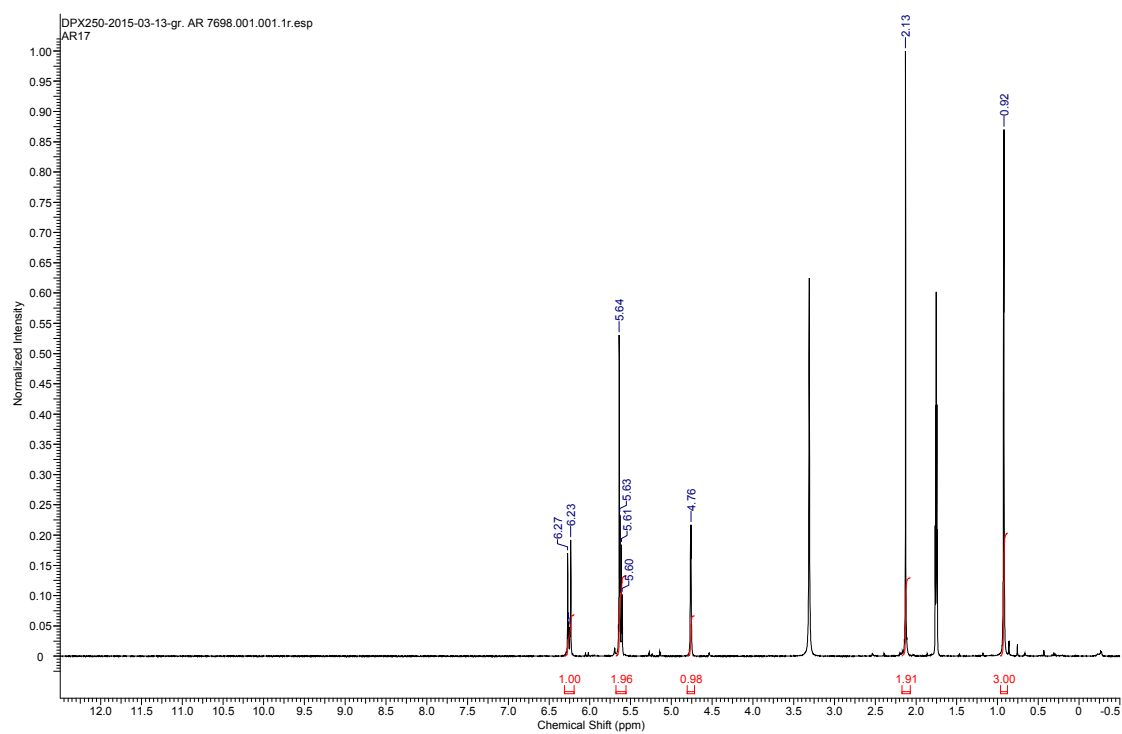
Figure S12: NMR spectrum of compound C2.
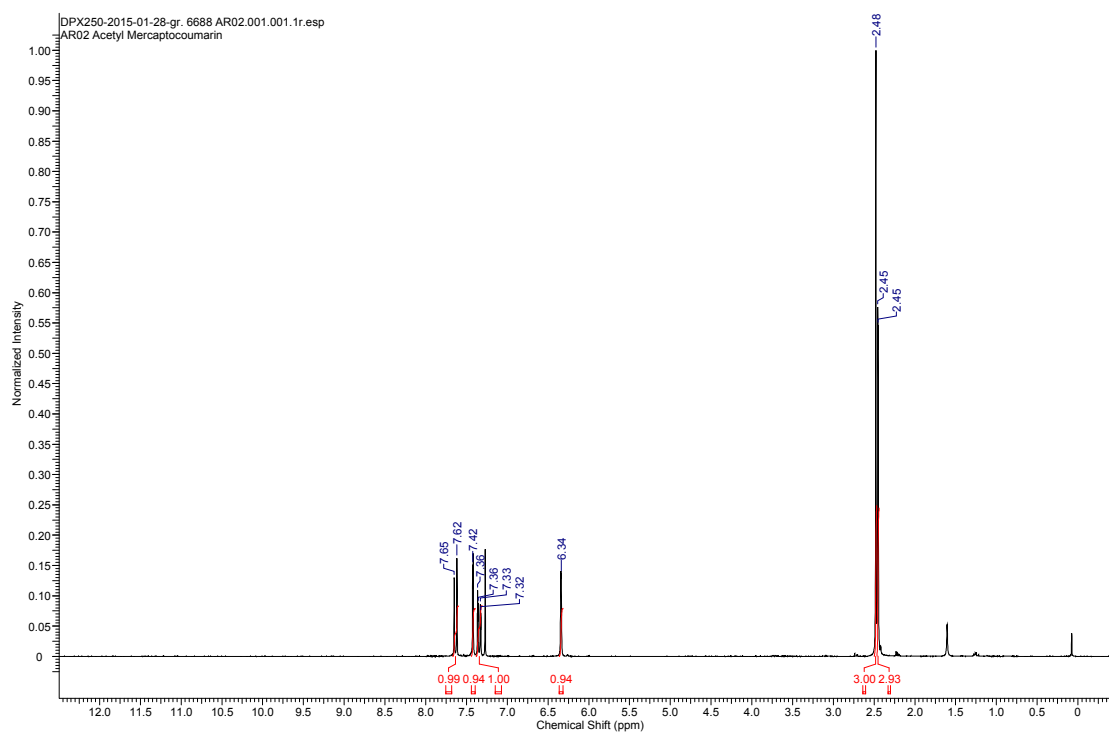
Figure S13: NMR spectrum of compound C3.
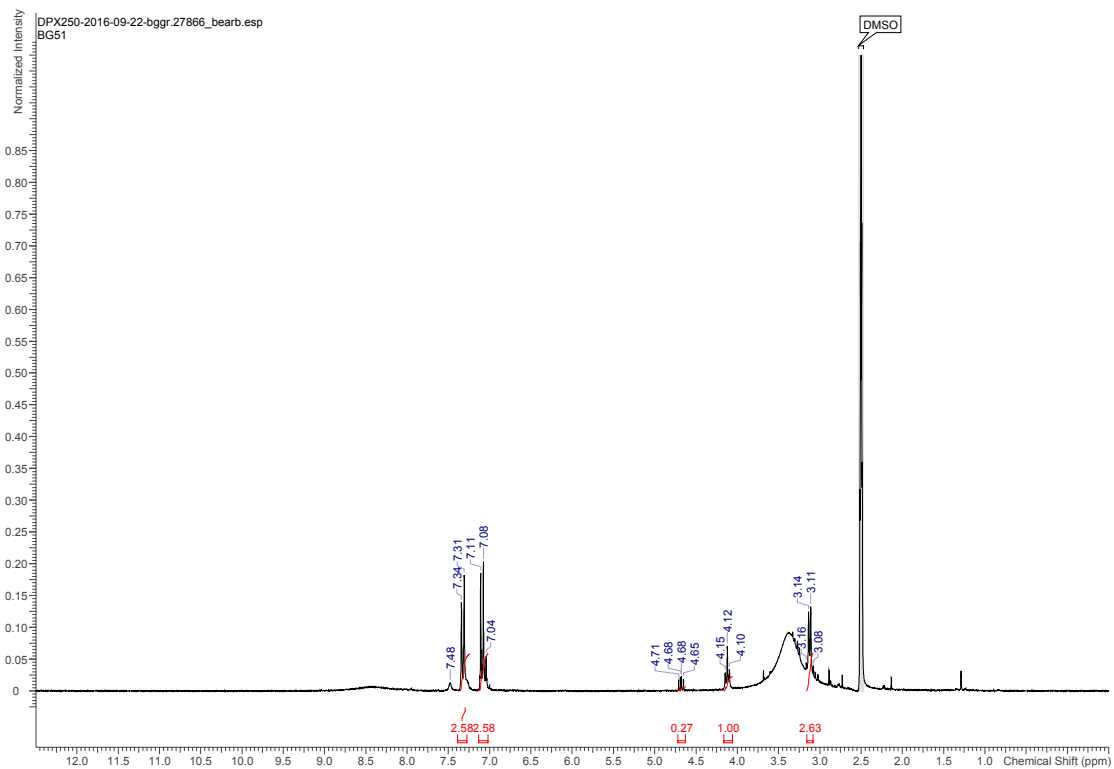
Figure S14: NMR spectrum of compound C4.
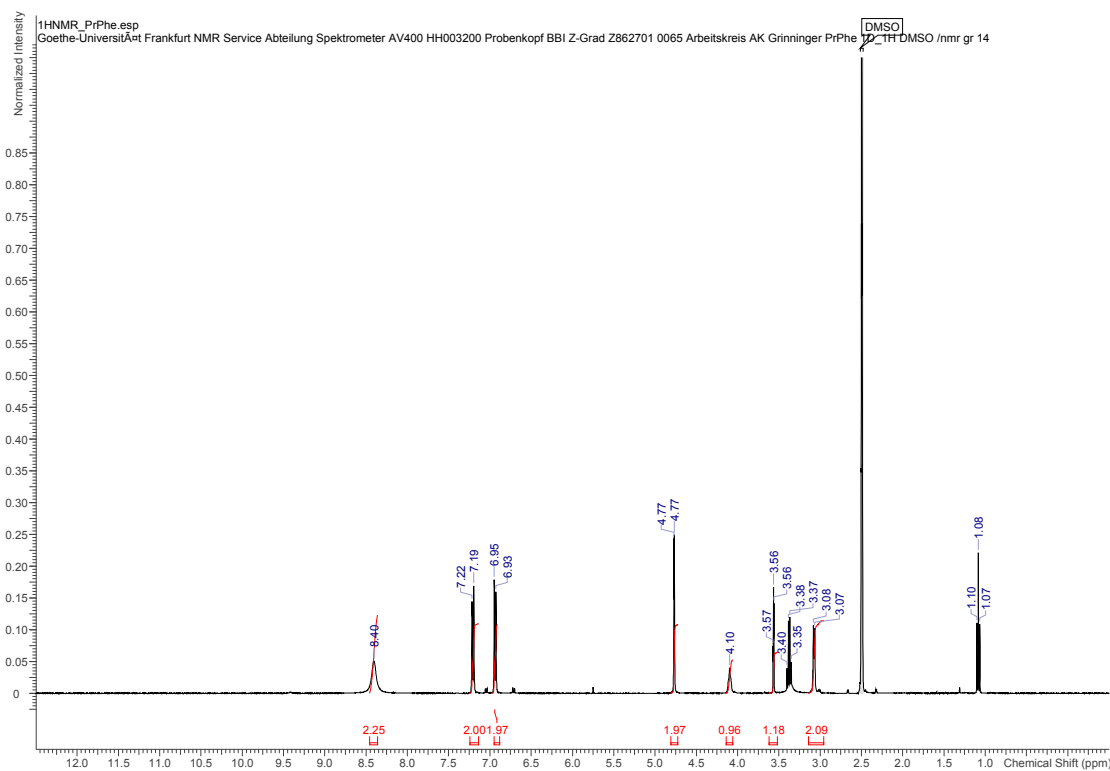
Figure S15: NMR spectrum of compound C5.
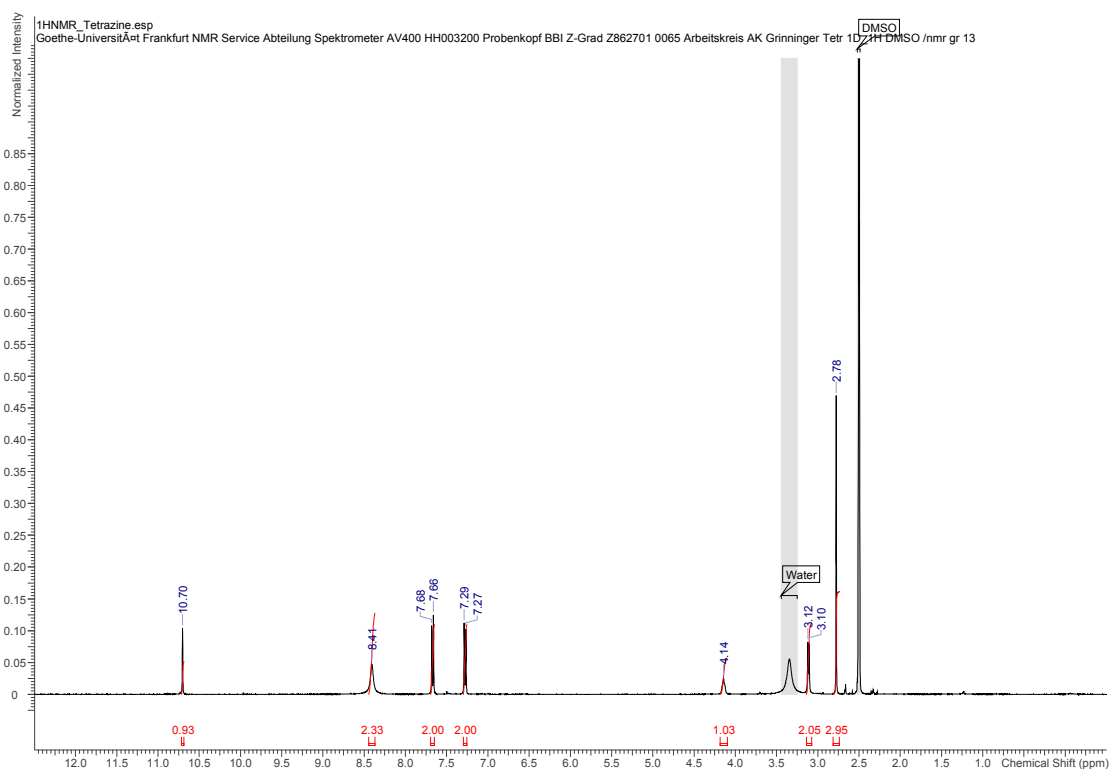
Figure S16: NMR spectrum of compound C6.

**Figure S17: NMR spectrum of compound C7.**


Figure S18: NMR spectrum of compound C8.

**Figure S19: NMR spectrum of compound C9.**

Figure S20: NMR spectrum of compound C10.

5.4 Primers and Plasmids

Table S1: List of primers.

Number	Name	Length	Orientation	T_m [°C]*	Sequence (5'3')
AR01	Strep_MBP_SLIC	54	forward	49.2	GGAGCCACCCGCAGTTCGAAAAAGGCGCCGGAAAAATCGAAGAAGGTAAACTGG
AR02	MBP_hFASE_SLIC	47	reverse	51.1	TCATACCTGCAATAACAACTTCTTCGGATCCGTAGTGTGCCGCACCC
AR03	Strep_TRX_SLIC	57	forward	52.8	GGAGCCACCCGCAGTTCGAAAAAGGCGCCGGAagcgataaaattattcacctgactg
AR04	TRX_hFASE_SLIC	49	reverse	54.9	TCATACCTGCAATAACAACTTCTTCGGATCCgccagaaccgaaccggc
AR05	Strep_SUMO3_SLIC	54	forward	53	GGAGCCACCCGCAGTTCGAAAAAGGCGCCGGAaatgaccacatcaacctgaagg
AR06	SUMO3_hFASE_SLIC	49	reverse	54.9	TCATACCTGCAATAACAACTTCTTCGGATCCggatccaccggtctgctg
AR07	MBP_seq_for	23	forward	58.8	GTCTGACCTTCCTGGTTGACCTG
AR08	MBP_seq_rev	23	reverse	58.8	CAGGTCAACCAGGAAGGTCAGAC
AR09	TRX_seq_for	24	forward	55.7	ccgttgcaaaactgaacatcgatc
AR10	TRX_seq_rev	24	reverse	55.7	gatcgatgttcagttttgcaacgg
AR11	SUMO3_seq_for	23	forward	57.1	ggcagatcagattcaggttcgac
AR12	SUMO3_seq_rev	23	reverse	57.1	gtcgaacctgaatctgatctgcc
AR13	hFASe_Δ11_TE_for	44	forward	57.1	GAAGCAAGCGAACTGGCATGTCAGACCCAACTGAATCTGCGTAG
AR14	hFASe_Δ11_TE_rev	24	reverse	57.4	ACATGCCAGTTCGCTTGCTTCATC
AR15	hFASe_Δ10f_TE_for	44	forward	56.7	GCAAGAACTGAGTAGCAAAGCACCGAAAGAAGATGGTCTGGCAC
AR16	hFASe_Δ10f_TE_rev	29	reverse	58.7	TGCTTTGCTACTCAGTTCTTGCAGTTTAC
AR17	hFASe_Δ20_TE_for	46	forward	59.1	GCAAGAACTGAGTAGCAAAGCACAACTGAATCTGCGTAGCCTGCTG
AR18	hFASe_aSTARTm_for	36	forward	55.3	GATCGGCGGCGTGGATATGGTGACCGATGATGATCG
AR19	hFASe_aSTARTm_rev	36	reverse	53.2	CACGCCGCCGATCAGATTATCCCAAAATTCTTGCAG
AR20	GATCCGGCTGCTAACA	32	forward	57.9	CATCACCACCATCACCACCATGAGCTACCGGC
AR21	MSAS_H6toH10_rev	33	reverse	58.3	ATGGTGGTGATGGTGGTGGTGCTCGAGTTTGGC
AR22	StrepIItoH8_ceffFAS_for	44	forward	60	CATCACCACCATCACCACCACCATGGCGTCGACGTGACCGAACC
AR23	pET22b_StrepIItoH8_rev	39	reverse	54	GTGATGGTGGTGATGGCTAGCCATATGTATATCCTTC
AR24	pET22b_SacB_for	41	forward	54.8	agaaggagatatacatatgtccacatatacctgccgttcac
AR25	SacB_pET22b_rev	40	reverse	56.7	ggtggtgctcgagcaatgccaataggatatcggcattttc
AR26	AMP_infusion_for	20	forward	55.9	GAG GAC CGA AGG AGC TAA CC
AR27	AMP_infusion_rev	20	reverse	55.9	GGT TAG CTC CTT CGG TCC TC
AR28	backbone_1800_rev	21	reverse	56.3	gaactctgtagcaccgcctac
AR29	backbone_2600_rev	20	reverse	53.8	cttgtctgtaagcggatgcc
AR30	backbone_3400_rev	20	reverse	55.9	ctgtcgcttgcggtattcgg
AR31	backbone_4200_rev	20	reverse	55.9	gagcatctggtcgcattggg
AR32	backbone_5000_rev	21	reverse	56.3	catcaccgatggggaagatcg
AR33	backbone_4500_for	16	forward	56.2	gcgacggcgcgtgcag
AR34	SacB_Seq_for	23	forward	55.3	gagagatcctcactacgtagaag
AR35	Strep_hFASh_infusion	43	forward	57.9	AGTTCGAAAAAGGCGCCGGATCCgaggaggtggtgattgccgg
AR36	hFASh_His_infusion	34	reverse	59.4	GGTGATGATGCTCGAGgccctcccgcacgctcac
AR37	Step_mFASm_infusion	41	forward	55.4	GTTCGAAAAAGGCGCCGGATCCgaggaggtggtgatagccg
AR38	mFASm_His_infusion	36	reverse	55.4	GGTGATGATGCTCGAGgccctcccgtacactcactc
AR39	MBP_hFASe_infusion_for	45	forward	54.8	CGTCGACGGATCCGAATTcGAAGAAGTTGTTATTGCAGGTATGAG
AR40	His_pMAL_infusion_rev	36	reverse	55.4	CGTTTTATTTGAAGCTtTCAGTGGTGGTGGTGGTGG
AR41	pMALtoStrII_hFASe_for	38	forward	53	CCCGCAGTTCGAAAAAGGATCCGAATTcGAAGAAGTTG
AR42	pMALtoPSTRII_rev	54	reverse	54	TTTTTCGAACTGCCGGTGGCTCCAGCTCAGCATAATCTATGGTCCTTGTTGGTC
AR43	T7term_deletion_for	23	forward	55.3	CTGAAAGGAGGAACTATATCCGG
AR44	T7term_deletion_rev	36	reverse	56.3	AGTTCCTCCTTTCAGGTTATTGCTCAGCGGTGGCAG
AR45	hFASh_seq_600for	18	forward	57.2	ccagcctgatggccctgc
AR46	hFASh_seq_900rev	19	reverse	55.4	gctgctcctggatatcccc
AR47	hFASh_seq_4500rev	20	reverse	55.9	cgaggtggcacagttgatgg
AR48	hFASh_seq_4200for	18	forward	54.9	ctgagccgcagtatggcc
AR49	IT1_StrI_for	26	forward	54.8	ACACATACATATGAGCGCTTGGAGCC
AR50	pET22b_IT1_rev	47	reverse	52.9	CTCATATGTATGTGTATCTCCTTCTTAAAGTTAAACAAAATTATTTC
AR51	IT2_StrI_for	26	forward	54.8	TGAGATACATATGAGCGCTTGGAGCC
AR52	pET22b_IT2_rev	47	reverse	52.9	CTCATATGTATCTCAATCTCCTTCTTAAAGTTAAACAAAATTATTTC
AR53	IT3_StrI_for	28	forward	54.8	CGAGACATACATATGAGCGCTTGGAGCC
AR54	pET22b_IT3_rev	49	reverse	52.9	CTCATATGTATGTCTCGATCTCCTTCTTAAAGTTAAACAAAATTATTTC
AR55	IT4_StrI_for	29	forward	53.7	CCCTTTAAGAAGGAGATATACATATGAGC
AR56	pET22b_IT4_rev	43	reverse	52.6	CTCCTTCTTAAAGGGAAACAAAATTATTTCTAGAGGGGAATTG
AR57	pET22b_STRItoII_for	33	forward	54.9	ATACATATGAGCGCTTGGAGCCACCCGCAGTTC
AR58	pET22b_STRItoII_rev	24	reverse	54	AGCGCTCATATGTATATCTCCTTC
AR59	siSD1_hFASe_for	41	forward	55.9	CTGGGAGGTGAACGTGGAGGTCCGGAAGTTCAGCAGGTTCC
AR60	siSD1_hFASe_rev	41	reverse	55.3	ACCTCCACGTTCACCTCCCAGAACTGCATAACCACGAAACG
AR61	siSD2_hFASe_for	36	forward	67.9m	GTGTTACCGTTGCCGGAGGTGTTCATATTAGCGGTC
AR62	siSD2_hFASe_rev	36	reverse	67.9m	GACCGCTAATATGAACACCTCCGGCAACGGTAACAC
AR63	siSD3_hFASe_for	32	forward	69.5m	CTGCGTCGTGAACCGGGAGGTAATCGTCTGCG
AR64	siSD3_hFASe_rev	32	reverse	69.5m	CGCAGACGATTACCTCCCGGTTCACGACGCAG
AR65	siSD4_hFASe_for	40	forward	56.3	CATTGCCGGAGGTCTGGGAGGTTTTGGTCTGGAACTGGCC
AR66	siSD4_hFASe_rev	35	reverse	56	TCCCAGACCTCCGGCAATGATATAGCTTTTGTGTG
AR67	siSD5_hFASe_for	37	forward	68.9m	GCGTGCCAAAACCGGAGGTGCATATGGTGAAGATCTG
AR68	siSD5_hFASe_rev	37	reverse	68.9m	CAGATCTTCACCATATGCACCTCCGGTTTTGGCACGC
AR69	STRIdel_MBP_for	44	forward	55.9	GGAGATATACATATGAAAATCGAAGAAGGTAAACTGGTAATCTG
AR70	STR1del_TRX_for	42	forward	55.2	GGAGATATACATATGagcgataaaattattcacctgactgac

*Basic melting temperatures (T_M) from OligoCalc, calculated for the annealing part of the primer.[202]

ᵐMelting temperature was calculated for the whole primer including the mismatch.

AR71	STR1del_SUMO_for	39	forward	55.7	GGAGATATACATATGaatgaccacatcaacctgaaggtg
AR72	pET22b_rev	30	reverse	53.4	CATATGTATATCTCCTTCTTAAAGTTAAAC
AR73	pETcoco-1_lin_for	22	forward	54.8	tgagatccggctgctaacaaag
AR74	pETcoco-1_lin_rev	30	reverse	54.8	catgcttatatctccttcttaaagttaaac
AR75	pETcoco_tig_for	41	forward	56.4	ggagatataagcatgCAAGTTTCAGTTGAAACCACTCAAGG
AR76	tig_pETcoco_rev	34	reverse	55.4	agcagccggatctcaCGCCTGCTGGTTCATCAGC
AR77	pETcoco_mycdodecin_for	39	forward	57.4	ggagatataagcatgAGCAATCACACCTACCGAGTGATC
AR78	mycdodecin_pETcoco_rev	34	reverse	55.4	agcagccggatctcaGGAATCCTCCAGGCGGAAG
AR79	pETcoco_hACP_for	35	forward	55.9	ggagatataagcatgCGTGATCGTGATAGCCAGCG
AR80	hACP_pETcoco_rev	33	reverse	57.2	agcagccggatctcaCGGGGTCGGACATGCCAG
AR81	pET22b_XbaI_for	31	forward	56.4	GGATAACAATTCCCCTCTAGAAATAATTTTG
AR82	tig_RBS_rev	43	reverse	55.4/58.6	CAAAATTATTTCTAGCGCGGAtcaCGCCTGCTGGTTCATCAGC
AR83	Dnak_RBS_rev	54	reverse	56.9	CAAAATTATTTCTAGCGCGGATCATTTTTTGTCTTTGACTTCTTCAAATTCAGC
AR84	pETcoco_DnaK_infusion_for	39	forward	55.2	ggagatataagcatgGGTAAAATAATTGGTATCGACCTG
AR85	Dnak_RBS_DnaJ_for	51	forward	54.5	GACAAAAAATGATCCGCGCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
AR86	DnaJ_pETcoco_infusion_rev	33	reverse	54.9	ggctggcttgccggccTCAGCGGGTCAGGTCGTC
AR87	MBP_hFASh_infusion_for	40	forward	57.9	TCGTCGACGGATCCGAATTcgaggaggtggtgattgccgg
AR88	pETcoco_pfuPre_for	40	forward	57.7	gaaggagatataagcatgCAAAACATTCCACCCCAAGTCC
AR89	pfuPre_pETcoco_rev	35	reverse	55.9	ggctggcttgccggccTCATCCAGCGGTTGGAGGTC
AR90	pfuPre_RBS_rev	41	reverse	55.9	CAAAATTATTTCTAGCGCGGATCATCCAGCGGTTGGAGGTC
AR91	Spacer_for	24	forward	54	AATTCGAGCTCGAACAACAACAAC
AR92	pMAL_rev	24	reverse	54	CATAATCTATGGTCCTTGTTGGTC
AR93	pMAL_tig_for	50	forward	56.4	GACCAACAAGGACCATAGATTATGCAAGTTTCAGTTGAAACCACTCAAGG
AR94	tig_spacer_infusion_rev	34	reverse	55.4	GTTCGAGCTCGAATTCGCCTGCTGGTTCATCAGC
AR95	pMAL_GST_for	45	forward	54	GACCAACAAGGACCATAGATTATGTCCCCTATACTAGGTTATTGG
AR96	GST_spacer_infusion_rev	35	reverse	53.8	GTTCGAGCTCGAATTTTTGGAGGATGGTCGCCAC
AR97	pETcoco_GroES_for	41	forward	54.8	gaaggagatataagcatgAATATTCGTCCATTGCATGATCG
AR98	GroES_RBS_GroEL_rev	58	reverse	54.4	CCTTAAATTCGTATGTTCAGTGTCGTGCGCGGATTACGCTTCAACAATTGCCAGAATG
AR99	RBS_GroEL_for	60	forward	54.4	CGACACTGAACATACGAATTTAAGGAATAAAGATAATGGCAGCTAAAGACGTAAAATTCG
AR100	GroEL_pETcoco_rev	36	reverse	54.9	gttagcagccggatctcaCATCATGCCGCCCATGCC
AR101	DnaJ_RBS_rev	46	reverse	54.9	CGCAGCATTCATCTCGCGTAAGCGCGGATCAGCGGGTCAGGTCGTC
AR102	RBS_GrpE_for	64	forward	54.8	CGCGAGATGAATGCTGCGAAAAAAACGCGGAGAAATTCATGAGTAGTAAAGAACAGAAAACGCC
AR103	GrpE_pETcoco_rev	42	reverse	55.7	gttagcagccggatctcaCGCTTTTGCTTTCGCTACAGTTAC
AR104	pETcoco_ClpB_for	45	forward	57.5	gaaggagatataagcatgCGTCTGGATCGTCTTACTAATAAATTC
AR105	RBS_pfuPreA_for	55	forward	54.1	GTTTAACTTTAAGAAGGAGATATACATATGGAAAACAATAAGGAATTGGAAAAGG
AR106	pfuPreA_pETcoco_rev	42	reverse	56.7	gttagcagccggatcTCACTTCTTAAGCTTGAAGCTCATTGC
AR107	pETcoco_PF1883_for	39	forward	54	gaaggagatataagcatgGTGAGGAGAATAAGAAGATGG
AR108	PF1883_pETcoco_rev	46	reverse	55.5	gttagcagccggatctcaTTCAACTTTAACTTCGAATCCTTCACTC
AR109	pETcoco_PfCPN_for	34	forward	55.4	gaaggagatataagcatgGCCCAGTTAGCAGGCC
AR110	PfCPN_pETcoco_rev	39	reverse	55.7	gttagcagccggatctcaGTCTAGATCACTGCTGAAGTC
AR111	pETcoco_IbpA_for	40	forward	54	gaaggagatataagcatgCGTAACTTTGATTTATCCCCGC
AR112	IbpA_RBS_rev	53	reverse	56.7	GCGAGTAAGTACCTGCAAATCCGAAGATTTCAGTTGATTTCGATACGGCGCGG
AR113	RBS_IbpB_for	23	forward	57.1	CGGATTTGCAGGTACTTACTCGC
AR114	IbpB_pETcoco_rev	40	reverse	57.1	gttagcagccggatctcaGCTATTTAACGCGGGACGTTCG
AR115	ClpB_pETcoco_rev	36	reverse	55.4	gttagcagccggatctcaCTGGACGGCGACAATCCG
AR116	Cam_infusion_for	22	forward	56.7	caccgtaacacgccacatcttg
AR117	Cam_infusion_rev	22	reverse	56.7	caagatgtggcgtgttaccggtg
AR118	pfuPreA_T7_rev	60	reverse	56.7	ccgctcacaattcccctatagtgagtcgtattaTCACTTCTTAAGCTTGAAGCTCATTGC
AR119	pETcoco_HtpG_for	43	forward	57	gaaggagatataagcATGAAAGGACAAGAAACTCGTGGTTTTC
AR120	HtpG_rev	22	reverse	56.7	TCAGGAAACCAGCAGCTGGTTC
AR121	HtpG_T7_rev	55	reverse	56.7	ccgctcacaattcccctatagtgagtcgtattaTCAGGAAACCAGCAGCTGGTTC
AR122	pfcPN_mut_for	73	forward		GCTCGGCAAAGGCAAGGGCAAAGGAGGTGGCAAGGGAGGAGGAAGCGGCGGCTTCAGCAGTGGCTAGGCTGA
AR123	pfcPN_mut_rev	36	reverse	(54.3)	CCTTTGCCCTTGCCTTTGCCGAGCTTGCTGGCAGCG
AR124	pAG425GAL_STRI_infusion_for	36	forward	55.9	gattctagaactagtgATGAGCGCTTGGAGCCATCC
AR125	H8_pAG425GAL_infusion_rev	36	reverse	55.4	ctaattacatgactcgaCTCAGTGGTGGTGGTGGTGGTG
AR126	pT7CFE1_STRI_infusion_for	36	forward	56.3	tatgccaccacccaTATGAGCGCTTGGAGCCATCC
AR127	H8_pT7CFE1_infusion_rev	35	reverse	55.4	gtgctcgagtgcggccCTCAGTGGTGGTGGTGGTG
AR128	pfCPN_His6_Δlid_for	50	forward	54.4	ccgcgcatcaccaccaccaccatgcgaaATGCTCAGAGAGATGGTCGAG
AR129	pfCPN_His6_Δlid_rev	55	reverse	54	cgctatggtggtggtggtgatgcgcggtttcCTCAAGTGCATCGTTAATTAGTGC
AR130	bT7_prom_pETcoco_seq_for	20	forward	55.9	gagatcgatctcgatcccgc
AR131	ClpB_T7_rev	53	reverse	55.9	ccgctcacaattccctatagtgagtcgtattatcaCTGGACGGCGACAATCC
AR132	STRI_DEBSM4_for	40	forward	59.4	GTTCGAAAAAGGCGCCGGATCCGCCGACGAGTCCGAGCCC
AR133	DEBSM4_H8_rev	35	reverse	61.9	GGTGATGATGCTGAGCGCCGGTGAGGCCGCGAAC
AR134	hFASe_ΔMAT_for	27	forward	57.2	CAGCCTagcCCGCGTGGTACACCGCTG
AR135	hFASe_ΔMAT_rev	30	reverse	55.4	CCACGCGGgctAGGCTGGGTATTCGGACGC
AR136	hFASe_KS_sol_for	43	forward	56.4	CGCTGagcAGTCCGgatATTAAATGGGATCATAGCTGGCATG
AR137	hFASe_KS_sol_rev	30	reverse	57.2	atcCGGACTgctCAGCGGTGTACCACGCGG
AR138	StrI_hMAT_infusion_for	33	forward	57.6	GTTCGAAAAAGGCGCCGCACCGGCACCGCATGC
AR139	hMAT_infusion_for	36	forward	57.6	AGAAGGAGATATACATATGGCACCGGCACCGCATGC
AR140	hMAT_pET22b_infusion_rev	39	reverse	56.3	GGTGATGATGCTCGATCAAGGCGGAAACAGTGCATTCGG
AR141	DEBSM4_2100rev	20	reverse	55.9	GCGAAGAACTCGTCCAGCTC
AR142	DEBSM4_1900for	18	forward	57.2	GAAGGTGGTGGCCCTGCG
AR143	DEBSM4_4000rev	19	reverse	57.6	GCACGAGACGTTGGGTGCC
AR144	DEBSM4_3900for	20	forward	57.9	GGAACCTGGGATGCTGCCTG
AR145	START_KS_MAT_for	18	forward	54.9	GGATCCgaggaggtggtg
AR146	START_KS_MAT_rev	18	reverse	54.9	caccacctcctcGGATCC
AR147	hDH_for	18	forward	54.9	tcaggttcccctcagcc
AR148	mMAT_hDH_rev	36	reverse	55.9	ctgagggggaacctgagccgtttgggaagtcctcag
AR149	mDH_for	20	forward	55.9	tccagctcctcctctgctac
AR150	hMAT_mDH_rev	35	reverse	55.4	cagaggaggagctggaaccgttggggaagtcctcg

AR151	FLAG_m/hFAS_for	40	forward	56.2	GACTACAAAGACGATGACGACAAGGGCGCCGGATCCgagg
AR152	pET22b_FLAG_rev	54	reverse	53.4	CTTGTCGTCATCGTCTTTGTAGTCCATATGTATATCTCCTTCTTAAAGTTAAAC
AR153	pfCPN_hCTT6A_for	35	forward	53.8	CAGCTCATAAGAGGTcttgttttggaccacggagc
AR154	hCCT6A_pfCPN_rev	45	reverse	54.8	CACAGCCTTGGGGTTgttacatttctcaataaaggtaaacttctc
AR155	hapical_pfCPN_for	22	forward	56.7	AACCCCAAGGCTGTGACAATCC
AR156	hapical_pfCPN_rev	21	reverse	54.4	ACCTCTTATGAGCTGGGTGTC
AR157	pETcoco_hNT_for	33	forward	54.9	gaaggagatataagcatggcggccggtgaagacc
AR158	hNT_pfCPN_rev	41	reverse	56	CTGCTGGGTGCTGAATtttgcatttcgtgaagcagcacattg
AR159	ΔpfNT_for	21	forward	54.4	ATTCAGCACCCAGCAGCTAAG
AR160	pfCPN_hCT_for	38	forward	54.8	GAGTGCTAGCGAGGCAgccaccaacattcttcttggttg
AR161	hCT_pETcoco_rev	36	reverse	53.5	agcagccggatctcaacctttcagagaagacattcc
AR162	ΔpfCT_rev	22	reverse	56.7	GCCTCGCTAGCACTCTTGATAG
AR163	hMAT_His_for	36	forward	53.2	caggttcccctcagccCATCATCACCACCACCACC
AR164	mKS(C161G)_for	34	forward	72.9m	gccctggacacagccGGCtcctccagcttgctgg
AR165	mKS(C161G)_rev	34	reverse	72.9m	ccagcaagctggaggaGCCggctgtgtccagggc
AR166	mMAT(R606A)_for	39	forward	70.8m	ctgtgcttgcagcttactggGCGgccagtgcatcaaag
AR167	mMAT(R606A)_rev	39	reverse	70.8m	ctttgatgcactggccCGCccagtaagctgcaagcacag
AR168	mMAT(M499Q)_for	34	forward	72.9m	gttcatctgctcagggCAGggcacgcagtggcgc
AR169	mMAT(M499Q)_rev	34	reverse	72.9m	gcgccactgcgtgccCTGccctgagcagatgaac
AR170	mMAT(F553Q)_for	36	forward	71.3m	gatgacatcgtgcatgccCAGgtgagcctcactgcc
AR171	mMAT(F553Q)_rev	36	reverse	71.3m	ggcagtgaggctcacCTGggcatgcacgatgtcatc
AR172	mMAT(CI609AM)_for	42	forward	75.2m	gcttactggcgaggccagGCGATGaaagatgcccacctcccg
AR173	mMAT(CI609AM)_rev	42	reverse	75.2m	cgggaggtgggcatctttCATCGCctggcctcgccagtaagc
AR174	mER(G1672V)_for	37	forward	73.3m	catccactcaggttcaGTGgtgtggggccaagcggcc
AR175	mER(G1672V)_rev	37	reverse	73.3m	ggccgcttggcccacaccCACtgaacctgagtggatg
AR176	STRII_mACP_for	30	forward	56.2	TCGAAAAAGGCGCCgggacggggacaccc
AR177	mACP_rev	35	reverse	54.3	GGTGATGATGCTCGATCAgggggctgtcgtgtcag
AR178	Strep_celFAS_infusion	46	forward	55.7	GTTCGAAAAAGGCGCCGGATCCGATAATACTGGAGAAGGATCGTCC
AR179	celFAS_His_infusion_rev	43	reverse	53.7	GGTGATGATGCTCGAGTTGAAGAATAATCCTGTTGATGTGTTC
AR180	2400for_celFAS_seq	23	forward	55.3	CTTGCTCTCCAGTGTTGTTCTAC
AR181	Strep_chiFAS_infusion_for	43	forward	55.9	AGTTCGAAAAAGGCGCCGGATCCGAAGACGTGGTGATCGCAGG
AR182	chiFAS_His_infusion_rev	39	reverse	56.7	GGTGATGATGCTCGAGACCTTCTCTGACACTGACACGTG
AR183	2400for_chiFAS_seq	22	forward	53	CTAACGCAGACTGGAAAGATTC
AR184	Strep_porFAS_infusion_for	41	forward	54.9	AGTTCGAAAAAGGCGCCGGATCCGAGGAGGTGGTGATCGCC
AR185	porFAS_His_infusion_rev	32	reverse	56.2	GGTGATGATGCTCGAGGCCCTCCCGCACGCTG
AR186	chiDH_for	20	forward	55.9	GGGTTCCAAAGGCTCTGCGTC
AR187	mMAT_chiDH_rev	35	reverse	55.9	AGAGCCTTTGGAACCgccgtttgggaagtcctcag
AR188	2400rev_chiFAS_seq	22	reverse	53	GAATCTTTCCAGTCTGCGTTAG
AR189	hMAT_rev	17	reverse	54.3	ggctgagggggaacctg
AR190	hKS(K436D)_for	31	forward	71	ctgaggccgtgcagGATctgctggagcaggg
AR191	hKS(K436D)_rev	31	reverse	71m	ccctgctccagcagATCctgcacggcctcag
AR192	hKS(L833H)_for	26	forward	53.2	ccccaCACatcaagtgggaccacagc
AR193	hKS(L833H)_rev	26	reverse	54.9	cttgatGTGtggggagatgagggag
AR194	mKS_lin_for	18	forward	54.9	ctgtggagttcccggctc
AR195	mKS_lin_rev	20	reverse	55.9	gtgttgggctggaggatgac
AR196	mKS/hMAT_for	30	forward	55.6	cctccagcccaacacacagccgccccccgc
AR197	hMAT/mKS_rev	35	reverse	55.9	ccgggaactccacaggtgggaacaaggcattgggg
AR198	hKS_lin_for	19	forward	55.4	cctgtggagttcccagctc
AR199	hKS_lin_rev	19	reverse	55.4	cgtgttgggcctcaggatg
AR200	hKS/mMAT_for	30	forward	55.6	ctgaggcccaacacgcggcaggcccctgcg
AR201	mMAT/hKS_rev	37	reverse	56.7	tgggaactccacaggtgggaacaaggcgttagggttg
AR202	LEHis_for	21	forward	56.3	CTCGAGCATCATCACCACCAC
AR203	ΔmTE_His_rev	36	reverse	56.3	GTGATGATGCTCGAGtgtcgtgtcagtagccgagtc
AR204	ΔmACPTE_His_rev	33	reverse	57.2	GTGATGATGCTCGAGtcccatgggccacagc
AR205	STRI_FabD_for	43	forward	53.5	TCGAAAAAGGCGCCGGATCCACGCAATTTGCATTTGTGTTCCC
AR206	FabD_H8_rev	31	reverse	56.7	GGTGATGATGCTCGAGTCGAGCGCCGCTGC
AR207	mMAT(K673T)_for	29	forward	>55.4	AGCTCGAGCGCCGCTGC
AR208	mMAT(K673T)_rev	35	reverse	>55.7	cgtacctcGGTggcaaacacaccttcttgctttag
AR209	mKS(LM203AA)_for	30	forward	56.7	gaagGCGggcGCGctcagcccggacggccac
AR210	mKS(LM203AA)_rev	32	reverse	54.8	gCGCgccCGCcttcatgaactgcacagaggtg
AR211	mKS(H293A)_for	27	forward	57.2	gaagccGCGggcacgggcaccaaggtg
AR212	mKS(H293A)_rev	32	reverse	55.7	gtgccCGCggcttcaatatactcaagcgactc
AR213	mKS(E333Q)_for	28	forward	57.6	caccctCAGcctgcctctgggcttgcag
AR214	mKS(E333Q)_rev	32	reverse	57.1	ggcaggCTGagggtgtcccagtgttggatttgg
AR215	mKS(M132G/Q136A)_for	35	forward	55.9	GGCgtgggctgcGCGgcgtgcaatgatggccaaccg
AR216	mKS(M132G/Q136A)_rev	34	reverse	55.4	CGCgcagcccacGCCgctgtagcccagaagcgtc
AR217	mFe_del_for	19	forward	57.6	gccccacattgctgcagg
AR218	mFe_del_rev	17	reverse	56.7	tccaggcgggaggtggg
AR219	eFe_infusion_for	35	forward	55.9	cacctcccgcctggaGCTATGGCGGCAATCATCGG
AR220	eFe_infusion_rev	37	reverse	56.7	cagcaatgtggggggcTGCTGGGTTTCATCAGCGCACAG
AR221	eInt_infusion_for	37	forward	55.4	ctaaagcaagaaggtgtgAAACGCGCGCTGCCGTTAC
AR222	mInt_del_rev	24	reverse	57.4	cacaccttcttgctttagctgctc
AR223	mMat_R606E_for	39	forward	69.7m	ctgtgcttgcagcttactggGAAgccagtgcatcaaag
AR224	mMat_R606E_rev	39	reverse	69.7m	ctttgatgcactggccTTCccagtaagctgcaagcacag
AR225	mMT_eGate_for	45	forward	55.9	GGGCAAAGTGCTTACTGGCCTGACGaagcgaggcgtgaagtccag
AR226	mMT_eGate_rev	43	reverse	57.6	AGTAAGCACTTTGCCCGGGCCGACctccagcaccacggcatgc
AR227	mACP_H8_rev	35	reverse	57.9	GTGATGATGCTCGAGgggggctgtcgtgtcagtag
AR228	STRI_Debs2M3(KS_MT)_for	40	forward	57.9	TCGAAAAAGGCGCCGGATCCGACCCGATCGCCATCGTCAG
AR229	Debs2M3(KS_MT)_H8_rev	33	reverse	61.5	GGTGATGATGCTCGAGGGAGGCGGCGGCGGGTG
AR230	mMat_V675P_for	36	forward	72.4m	gtgtgtttgccaaggagCCGcgaacaggaggcctgg

AR231	mMat_V675P_rev	36	reverse	72.4m	ccaggcctcctgttcgCGGctccttggcaaacacac
AR232	mMat_R676L_for	33	forward	69.4m	gtttgccaaggaggtaCTGacaggaggcctggc
AR233	mMat_R676L_rev	33	reverse	69.4m	gccaggcctcctgtCAGtacctccttggcaaac
AR234	mMAT_T677P_for	34	forward	72.9m	gccaaggaggtacgaCCGggaggcctggctttcc
AR235	mMAT_T677P_rev	34	reverse	72.9m	ggaaagccaggcctccCGGtcgtacctccttggc
AR236	mMAT_I549T_for	34	forward	68m	gcgcacctttgatgacACCgtgcatgcctttgtg
AR237	mMAT_I549T_rev	34	reverse	68m	cacaaaggcatgcacGGTgtcatcaaaggtgcgc
AR238	mMAT_F682S_for	35	forward	69.1m	gaacaggaggcctggctTCCcactcctacttcatg
AR239	mMAT_F682S_rev	35	reverse	69.1m	catgaagtaggagtgGGAagccaggcctcctgttc
AR240	mMAT_N738Q_for	36	forward	70.1m	ccgagtacaatgtcaacCAGctggtgagccctgtgc
AR241	mMAT_N738Q_rev	36	reverse	70.1m	gcacagggctcaccagCTGgttgacattgtactcgg
AR242	mMAT_V585Y_for	37	forward	71.1m	gggcactccttgggagagTATgcctgtggctatgcag
AR243	mMAT_V585Y_rev	37	reverse	71.1m	ctgcatagccacaggcATActctccaaggagtgccc
AR244	mMAT_R606I_for	39	forward	68.7m	ctgtgcttgcagcttactggATTggccagtgcatcaaag
AR245	mMAT_R606I_rev	39	reverse	68.7m	ctttgatgcactggccAATccagtaagctgcaagcacag
AR246	mMAT_N738I_for	36	forward	67.9m	ccgagtacaatgtcaacATTggtgagccctgtgc
AR247	mMAT_N738I_rev	36	reverse	67.9m	gcacagggctcaccagAATgttgacattgtactcgg
AR248	STRI_AVES2M3_for	42	forward	54.8	TCGAAAAAGGCGCCGGATCCCAATTGGCGAATGAAGCGAAGC
AR249	AVES2M3_H8_rev	30	reverse	55.6	GGTGATGATGCTCGAGGTCGGCGGGGACCG
AR250	STRI_AVES4M11_for	38	forward	54.9	TCGAAAAAGGCGCCGGATCCGTCGACGTTGACGAGCCG
AR251	AVES4M11_H8_rev	33	reverse	54.9	GGTGATGATGCTCGAGGTCGACATCCACACGCG
AR252	STRI_mTEII_for	43	forward	55.3	TCGAAAAAGGCGCCGGATCCGAGACAGCAGTCTATGCTAAGAG
AR253	mTEII_STOP_H8_rev	40	reverse	55.7	GGTGATGATGCTCGAGCTAGAAGTAGTCAAGTGAGGAGAG
AR254	mMAT_R606M_for	39	forward	68.7m	ctgtgcttgcagcttactggATGggccagtgcatcaaag
AR255	mMAT_R606M_rev	39	reverse	68.7m	ctttgatgcactggccCATccagtaagctgcaagcacag
AR256	mMAT_E668A_for	35	forward	69.1m	ggagcagctaaagcaaGCGggtgtgtttgccaagg
AR257	mMAT_E668A_rev	35	reverse	69.1m	ccttggcaaacacaccCGCttgctttagctgctcc
AR258	mMAT_L766E_for	34	forward	70.4m	gattgcgccccacgcaGAAttgcaggctgtcctg
AR259	mMAT_L766E_rev	34	reverse	70.4m	caggacagcctgcaaTTCtgcgtggggcgcaatc
AR260	mMAT_F661R_for	35	forward	67.9m	ggctgcagtgaatgaaCGTgtggagcagctaaagc
AR261	mMAT_F661R_rev	35	reverse	67.9m	gctttagctgctccacACGttcattcactgcagcc
AR262	mMAT_PP635EEAPG_for	32	forward	56.7	GAAGAAGCGCCGGGCggcgttggtgcctgctg
AR263	mMAT_PP635EEAPG_rev	37	reverse	56.7	GCCCCGGCGCTTCTTCgcagcgcgtttacattcctcc
AR264	mMAT2_I549T_for	28	forward	55.9	gatgacACCgtgcatgcctttgtgagcc
AR265	mMAT2_I549T_rev	31	reverse	54.4	catgcacGGTgtcatcaaaggtgcgctcatc
AR266	mKS_W_rev	20	reverse	55.9	ccaagtctgactgtggtccc
AR267	mACP_for	16	forward	56.2	ggggacggggacaccc
AR268	mMAT_AM3L_for	37	forward	54.9	ggaccacagtcagacttggCTCCAACCCACCCGGCAAG
AR269	AM3L_mACP_rev	35	reverse	55.4	gggtgtccccgtccccCACGAGGACATCACGCAGC
AR270	mMAT_AM11L_for	36	forward	54.3	ggaccacagtcagacttggCTGGATGTGGAGGGGGC
AR271	AM11L_mACP_rev	32	reverse	56.2	gggtgtccccgtccccGGCGAGGCAGTCGCGC
AR272	mMAT_mouseL_for	37	forward	56.3	cagcatcgacgccagtgcagagaagaaagctgtggcc
AR273	mouseL_mACP_rev	21	reverse	56.3	actggcgtcgatgctgtagac
AR274	mKS_F200X_for	30	forward	56.3	gtgcagMTGatgaagctgggcatgctcagc
AR275	mKS_F200X_rev	28	reverse	55.4	cttcatCAKctgcacagaggtgttcggc
AR276	mKS_F395A_for	32	forward	54.4	catttggcGCGgaggctccaatgttcatgtc
AR277	mKS_F395A_rev	28	reverse	54.4	ctccCGCgccaaatgagttgatgcccac
AR278	mKS_F395M_for	32	forward	54.4	catttggcATGgaggctccaatgttcatgtc
AR279	mKS_F395M_rev	28	reverse	54.4	ctccCATgccaaatgagttgatgcccac
AR280	mKS_A160M_for	29	forward	55.4	ggacacaATGtgctcctccagcttgctgg
AR281	mKS_A160M_rev	29	reverse	56.3	gagcaCATtgtgtcccagggcaatgcttgg
AR282	mMAT_M499G_for	26	forward	56.2	ctcagggGGTggcacgcagtggcgcg
AR283	mMAT_M499G_rev	29	reverse	56.3	gtgccACCccctgagcagatgaaccagag
AR284	mMAT_L739G/W_for	37	forward	55.4	tacaatgtcaacaacKGGgtgagccctgtgctcttcc
AR285	mMAT_L739G/W_rev	23	reverse	55.3	gttgttgacattgtactcggcag
AR286	mMAT_V585I_for	28	forward	55.9	gagagATTgcctgtggctatgcagatgg
AR287	mMAT_V585I_rev	28	reverse	54.9	cacaggcAATctctcccaaggagtgccc
AR288	mMAT_M499A/V_for	36	forward	56.2	ggttcatctgctcagggGYGggcacgcagtggcgcg
AR289	mMAT_M499X_rev	21	reverse	56.3	ccctgagcagatgaaccagag
AR290	mMAT_F553D/H/N/Y_for	37	forward	55.4	gacatcgtgcatgccNACgtgagcctcactgccatcc
AR291	mMAT_F553X_rev	23	reverse	57.1	ggcatgcacgatgtcatcaaagg
AR292	mMAT_F553A/G_for	37	forward	55.4	gacatcgtgcatgccGSCgtgagcctcactgccatcc
AR293	mMAT_L680Y_S_for	39	forward	54.4	gtacgaacaggaggcTATgctTCCcactcctacttcatg
AR294	mMAT_L680X_rev	19	reverse	55.4	gcctcctgttcgtacctcc
AR295	mMAT_L680M/V_for	40	forward	54.8	gtacgaacaggaggcRTGgctttccactcctacttcatgg
AR296	mMAT_L680H_for	40	forward	54.8	gtacgaacaggaggcCACgctttccactcctacttcatgg
AR297	mMAT_AM3L2_for	48	forward	57.2	gtcccggttgctgaggacttcccaaacggcCCGAGCGAAGGCCGTGAG
AR298	mMAT_AM11L2_for	47	forward	54.3	gtcccggttgctgaggacttcccaaacggcGTCTCTGGGCAGTGGGG
AR299	mKS_G_mouseL_rev	39	reverse	55.9	cctcagcaaccgggacatcccaagtctgactgtggtccc
AR300	mMAT_S580C_for	38	forward	55.9	ggcatcattgggcactGcttgggagaggttgcctgtgg
AR301	mMAT_S580X_rev	20	reverse	55.9	gtgcccaatgatgccgtcag
AR302	mKS_K151R_for	39	forward	55.4	ctttcttcttcgacttcCGTggaccaagcattgccctgg
AR303	mKS_K151X_rev	23	reverse	55.3	gaagtcgaagaagaagagagcc
AR304	mKS_K202R_for	36	forward	54.3	ctctgtgcagttcatgCGTctgggcatgctcagccc
AR305	mKS_K202X_rev	22	reverse	54.8	catgaactgcacagaggtgttc
AR306	mKS_V268E/A_for	40	forward	55.7	cccctctggagaagMGcaagaacaactcatctgctctctg
AR307	mKS_V268X_rev	23	reverse	55.3	ttctccagagggggaatgttacac
AR308	mKS_L166M_for	46	forward	54.4	gacacagcctgctcctccagcttgAtggcactacagaatgcctacc
AR309	mKS_I155T_rev	45	reverse	54.8	gagcaggctgtgtccagggcGGtgcttggtcctttgaagtcgaag
AR310	mMAT_S580A_for	38	forward	55.9	ggcatcattgggcacGccttgggagaggttgcctgtgg

AR311	ΔmER_for	35	forward	55.9	gagcaggacaagccctccaagaccttctgcccagc
AR312	ΔmER_rev	21	reverse	56.3	gggcttgtcctgctctaactg
AR313	ΔmDH_for	19	forward	55.4	tcacgcggcagcaagaac
AR314	ΔmDH_rev	33	reverse	54.9	ttgctgccgccgtgaagaggaggagctggagcc
AR315	mDH_H878A_for	36	forward	54.9	cactacctggtagacGCGtgcattgacggccgggtc
AR316	mDH_H878A_rev	20	reverse	55.9	gtctaccaggtagtggtcgg
AR317	mFASm_for	19	forward	55.4	gaggaggtggtgatagccg
AR318	mACP_mL_mFAS_rev	40	reverse	56.7	ggctatcaccacctcctcagacgtgtcactcctggacttg
AR319	mL_mACP_seq_for	22	forward	56.7	gacgccagtgcagagaagaaag
AR320	STRI_mTEI_for	43	forward	56.3	GTTCGAAAAAGGCGCCGGATCCaacctgagcacactgctggtg
AR321	mKS_G113M/I_for	36	forward	54.9	tgggtgggtgtgagtATKtcagaggcatccgaggcc
AR322	mKS_G113X_rev	19	reverse	55.4	actcacacccacccagacg
AR323	mKS_G113F/Y_for	36	forward	54.9	tgggtgggtgtgagtTWTtcagaggcatccgaggcc
AR324	mKS_G113L/W_for	36	forward	54.9	tgggtgggtgtgagtTKGtcagaggcatccgaggcc
AR325	mKS_S117H/Y_for	38	forward	55.9	agtggttcagaggcaYACgaggcccttagcagagatcc
AR326	mKS_S117X_rev	21	reverse	56.3	tgcctctgaaccactcacacc
AR327	mKS_M132F/Y_for	35	forward	56.7	cttctgggctacagcTWTgtgggctgccagccgtgc
AR328	mKS_M132X_rev	19	reverse	55.4	gctgtagcccagaagcgtc
AR329	mKS_M132W_for	35	forward	56.7	cttctgggctacagcTGGgtgggctgccagccgtgc
AR330	mKS_F200Y_for	39	forward	56.3	aacacctctgtgcagtAcatgaagctgggcatgctcagc
AR331	mKS_F200X_rev	19	reverse	55.4	ctgcacagaggtgttcggc
AR332	mTE_S2301A_for	39	forward	54.4	cgcatagctgggtacGCGtttggagcctgtgtagccttc
AR333	mTE_S2301X_rev	20	reverse	55.9	gtacccagctatgcggtagg
AR334	ΔmTE_STOP_His_rev	39	reverse	56.3	GTGATGATGCTCGAGTCAtgtcgtgtcagtagccgagtc
AR335	mKR_tylHTH_for	52	forward	55.4	GCGCCGGGTGCAGCAGAACTGGCGGAAGAACTGagacgccagggcatccaag
AR336	mKR_tylHTH_rev	51	reverse	55.9	GTTCTGCTGCACCCCGGCGCTTCCGGGCCACGtcgggaagtcagcacaagcc
AR337	mKR_tylmHanch_for	36	forward	57.2	CGTCGTctggcccggCGTctcgtgcttcgcggagcc
AR338	mKR_tylmHanch_rev	36	reverse	54.9	ACGccgggccagACGACGgccaaagccacctaggcc
AR339	m(KS_MAT)_rev	20	reverse	55.9	gccgtttgggaagtcctcag
AR340	mKS_G_AM3ACP_for	36	forward	57.2	gaggacttcccaaacggcCCGAGCGAAGGCCGTGAG
AR341	AM3ACP_mFAS_rev	35	reverse	59.1	ggctatcaccacctcctcGTCGGCGGGGACCGCAG
AR342	mKS_G_AM11ACP_for	37	forward	57.6	gaggacttcccaaacggcGTCTCTGGGCAGTGGGGTG
AR343	AM11ACP_mFAS_rev	38	reverse	57.9	ggctatcaccacctcctcGTCGACATCCACACGCGGTG
AR344	mACP_AM3L3_for	33	forward	57.2	caggaaatgtcctccGTCGGAGCGGACACAGGC
AR345	mACP_AM11L3_for	32	forward	56.7	caggaaatgtcctccAGCGGCGGAAGCCCGAC
AR346	mACP_rev	23	reverse	55.3	ggaggacatttcctgaagtttcc
AR347	pET22b_mTEII_for	40	forward	55.7	AGAAGGAGATATACATATGGAGACAGCAGTCTATGCTAAG
AR348	mTEII_H8_rev	38	reverse	54.8	GGTGATGATGCTCGAGGAAGTAGTCAAGTGAGGAGAGC
AR349	pET22b_mTEI_for	38	forward	56.7	AGAAGGAGATATACATATGtccaggagtgacacgtctc
AR350	mMAT_H683A_for	41	forward	55.7	ggaggcctggcttcGCctcctacttcatggaaggaattgc
AR351	mMAT_H683X_rev	20	reverse	55.9	gaaagccaggcctcctgttc
AR352	mKR_for	25	forward	54.4	accttggaaaagttcgttttcacac
AR353	PNG_AM4DE_for	33	forward	57.2	gacttcccaaacggcTCGGGTTCCGGTTCCGGG
AR354	AMDE_mKR_rev	33	reverse	54.9	gaactttccaaggtCCAGGTGTTGATGCGGGC
AR355	PNG_AM5DE_for	33	forward	54.9	gacttcccaaacggcCAAAGGCGTTCTCGCCGC
AR356	mKS(C161Q)_for	37	forward	55.4	gccctggacacagccCAGtcctccagcttgctggcac
AR357	mKS(C161X)_rev	20	reverse	57.9	ggctgtgtccagggcaatgc
AR358	mKS(R379Q/E)_for	37	forward	57.6	ctgcaggtggtcgatSAGcccctgcctgttcgtggtg
AR359	mKS(R379X)_rev	18	reverse	54.9	atcgaccacctgcagccg
AR360	mKS(SFA311ATE)_for	42	forward	54.4	ctgtgcgccACCcgccagGAAcctctgttaattggctccacc
AR361	mKS(SFA311ATE)_rev	40	reverse	54.8	gcgGGTggccgcacagGGCccgagtaatgccattcagttcc
AR362	ULTRAinPET_for	36	forward	55.4	CAATAGACATAAGCGGCCAGGTGGCACTTTTCGGGG
AR363	ULTRAinPET_rev	38	reverse	57.1	CAAAACCTTTCGCGGCCCATTCGCCAATCCGGATATAG
AR364	tac_mmPylS_for	34	forward	55.4	CAAGGACCATAGCATATGAGCGCTTGGAGCCACC
AR365	mmPylS_rrnB_rev	47	reverse	57	GTACGTAGTGTGCCGCACCTTACAGGTTGGTAGAAATCCCGTTATAG
AR366	proK_rev	21	reverse	56.3	AATGCGGGGCGCATCTTACTG
AR367	proK_mmtRNA_for	47	forward	56.4	CAGTAAGATGCGCCCCGCATTGGAAACCTGATCATGTAGATCGAATG
AR368	mmtRNA_terK_rev	35	reverse	57.9	GCTTTTCGAATTTGGTGGCGGAAACCCCGGGAATC
AR369	terK_for	24	forward	55.7	CCAAATTCGAAAAGCCTGCTCAAC
AR370	mmPylS_N346A_C348A_for	44	forward	55.9	GAGTTTACCATGCTGGCGTTCGCGCAGATGGGATCGGGATGCAC
AR371	mmPylS_N346X_rev	24	reverse	57.4	CAGCATGGTAAACTCTTCGAGGTG
AR372	mmPylS_F384Y_for	40	forward	54.8	GATTCCTGCATGGTCTATGGGGATACCCTTGATGTAATGC
AR373	mmPylS_F384X_rev	21	reverse	56.3	GACCATGCAGGAATCGCCTAC
AR374	mjTyrS_Y32L_for	46	forward	54.1	GATGAAAAATCTGCTctgATAGGTTTTGAACCAAGTGGTAAAATAC
AR375	mjTyrS_L65V_for	42	forward	55.2	GATATAATTATAgtgTTGGCTGATTTACACGCCTATTTAAAC
AR376	mjTyrS_L65V_rev	43	reverse	53.2	CAGCCAacacTATAATTATATCAAATCCAGCATTTTGTAAATC
AR377	mjTyrS_P107E_F108W_Q109M_for	57	forward	53.7	GCAAAATATGTTTATGGAAGTGAtggatgCTTGATAAGGATTATACACTGAATGTC
AR378	mjTyrS_D158G_I159A_for	53	forward	54.4	CTATCCAATAATGCAGGTTAATggcgcgCATTATTTAGGCGTTGATGTTGCAG
AR379	mjTyrS_Y32E_for	46	forward	54.1	GATGAAAAATCTGCTgaaATAGGTTTTGAACCAAGTGGTAAAATAC
AR380	mjTyrS_L65A_for	42	forward	55.2	GATATAATTATAgcgTTGGCTGATTTACACGCCTATTTAAAC
AR381	mjTyrS_L65A_rev	43	reverse	53.2	CAGCCAacgcTATAATTATATCAAATCCAGCATTTTGTAAATC
AR382	mjTyrS_P107E_F108P_Q109S_for	57	forward	53.7	GCAAAATATGTTTATGGAAGTgaaccgagcCTTGATAAGGATTATACACTGAATGTC
AR383	mjTyrS_D158G_L162G_for	57	forward	53.8	CTATCCAATAATGCAGGTTAATggcATTCATTATggcGGCGTTGATGTTGCAGTTGG
AR384	mmPylS_Y306A_for	39	forward	54.4	CTTGCTCCAAACCTTgcgAACTACCTGCGCAAGCTTGAC
AR385	mmPylS_Y306X_rev	21	reverse	54.4	AAGGTTTGGAGCAAGCATGGG
AR386	mmPylS_A306Y_for	39	forward	54.4	CTTGCTCCAAACCTTTACAACTACCTGCGCAAGCTTGAC
AR387	mACP_S2150A_for	37	forward	56.3	gacctcggcctggacGCGctcatgggtgtggaagttc
AR388	mACP_S2150X_rev	18	reverse	57.2	gtccagccgaggtctgc
AR389	ULTRAcassette_rev	40	reverse	55.4	GCCGCTTATGTCTATTGCTGGGGTCGTCGGTTCAGGGCAG
AR390	hKS_F256S_for	40	forward	56.3	caccaatacagatggcAGCaaggagcaaggcgtgaccttc

AR391	hKS_F256X_rev	20	reverse	55.9	gccatctgtattggtgccgg
AR392	mFAS_DHL1_for	35	forward	55.9	gaagacccgaactccgaccacccagaagtcccaac
AR393	mFAS_DHL1_rev	19	reverse	55.4	ggagttcgggtcttcccac
AR394	mFAS_DHL2_for	36	forward	55.9	gttattcgaccacccaacacccctgagtctgcatc
AR395	mFAS_DHL2_rev	22	reverse	54.8	tgggtggtcgaataacttggag
AR396	mFAS_DHL3_for	51	forward	56.7	gttattcgaccacccaacacccctgagtctgcatccgcctgacccaggg
AR397	mFAS_ERL1_for	35	forward	55.9	gaggaggagcctgaggctcagcccaccctgatttc
AR398	mFAS_ERL1_rev	18	reverse	57.2	ctcaggctcctcctcccg
AR399	mFAS_ERL2_for	42	forward	56.3	cttgtccaggtacggGGCGGCcccaccctgatttctgccatc
AR400	mFAS_ERL2_rev	22	reverse	56.7	ccgtacctggacaaggactttg
AR401	serMSAS_C181G_for	34	forward	57.2	GCGGTGGACACCGCGGGCTCTGCCTCGCTCACCG
AR402	serMSAS_C181X_rev	18	reverse	59.4	CGCGGTGTCCACCGCGAG
AR403	serMSAS_MAT_H8_rev	31	reverse	56.2	GGTGATGATGCTCGAGAGCGCCTGACGGGGC
AR404	mFAS_psKRL1_for	35	forward	55.4	tatcctgctgtccaactctcacgcccccaagttgg
AR405	mFAS_psKRL1_rev	21	reverse	56.3	gttggacagcaggatacaccg
AR406	mFAS_psKRL2_for	56	forward	55.4	tatcctgctgtccaactctCCGgcccccGCGACCAGCcctggctctccagagctac
AR407	mFAS_psMTL1_for	42	forward	57.2	cgggctctgcagaccGGCAGCGGCgaccctccacagcacggg
AR408	mFAS_psMTL1_rev	18	reverse	57.2	ggtctgcagagcccgtgc
AR409	mFAS_psMTdel_for	22	forward	54.8	ccacaggagaaacctatcttcc
AR410	mFAS_psMTdel_rev	35	reverse	55.9	aggtttctcctgtgggctctcagacaggcactcag
AR411	hFASh_DHL2_for	51	forward	56.7	ggatgaccctgaccccGCGctcttcgaccacccgccccaccccaaccccac
AR412	hFASh_DHL2_rev	19	reverse	57.6	ggggtcagggtcatcccac
AR413	hFASh_ERL1_for	48	forward	56.3	gcggaggagccggaggccCAGcccACCctgatgtcggccatctccaag
AR414	hFASh_ERL1_rev	18	reverse	57.2	ctccggctcctccgcaag
AR415	hFASh_psMTL1_for	47	forward	55.4	gcaggcactgcagaccacgGGCAGCGGCgaccctcacagcaggaac
AR416	hFASh_psMTL1_rev	19	reverse	57.6	ggtctgcagtgcctgcacc
AR417	hFASh_psMTL2_for	59	forward	55.4	gcaggcactgcagaccacgcagcaggggGCCcccgcccaggaccctcacagcaggaac
AR418	mmPylS_L301M_Y306L_L309A_for	46	forward	54.9	GATGGCTCCAAACCTTCTGAACTACGCGCGCAAGCTTGACAGGGCC
AR419	mmPylS_L301M_rev	36	reverse	56.7	AGGTTTGGAGCCATCATGGGTCTCAGGCAGAAGTTC
AR420	mmPylS_C348F_for	44	forward	55.9	GAGTTTACCATGCTGAACTTCTTTCAGATGGGATCGGGATGCAC
AR421	mMAT_H683N/D_for	42	forward	55.7	ggaggcctggctttcRACtcctacttcatggaaggaattgcc
AR422	StrI_mDH_for	36	forward	55.9	CGAAAAAGGCGCCGGATCCagctcctcctctgctac
AR423	StrI_mpsMT_for	38	forward	55.4	CGAAAAAGGCGCCGGATCCtcacggcggcagcaagaac
AR424	StrI_mpsKR_for	41	forward	54.8	CGAAAAAGGCGCCGGATCCccacaggagaaacctatcttcc
AR425	ΔmER_megaPCR_rev	27	reverse	56.3	cttggagggcttgtcctgctctaactg
AR426	StrI_DEBS1psKR1_for	38	forward	57.6	CGAAAAAGGCGCCGGATCCTCCACCGAGGTCGACGAGG
AR427	DEBS1psKR1_H8_rev	30	reverse	58.3	GGTGATGATGCTCGAGCGCAGCCCGCCGCG
AR428	StrepGS_rev	19	reverse	55.4	GGATCCGGCGCCTTTTTCG
AR429	m(LD_MAT)del_for	41	forward	55.9	cagcccaacacacggGGCGGCcgagggactcctctcatctc
AR430	m(LD_MAT)del_rev	19	reverse	57.6	ccgtgtgttgggctggagg
AR431	mMATdel_for	39	forward	57.4	caagaagtgcagcaaagctgcaccatcattcccttgatg
AR432	mMATdel_rev	22	reverse	56.7	ttgctgcacttcttggacacgg
AR433	STRI_mMAT_for	39	forward	56.3	GAAAAAGGCGCCGGATCCctctggttcatctgctcaggg
AR434	mMAT_H8_rev	34	reverse	57.6	GTGATGATGCTCGAGgcctcgcttcaggacagcc
AR435	STRI_mLD_for	34	forward	56.2	GAAAAAGGCGCCGGATCCgccccctgcgcccactg
AR436	mLD_H8_rev	36	reverse	56.3	GTGATGATGCTCGAGaggtgggaacaaggcgttagg
AR437	STRI_MBP_for	47	forward	53.7	GAAAAAGGCGCCGGATCCAAAATCGAAGAAGGTAAACTGGTAATCTG
AR438	MBP_mFASm_rev	42	reverse	56.4	ctatcaccacctcctcGTTATTGTTGTTGTTGTTCGAGCTCG
AR439	STRI_GST_for	48	forward	56.2	GAAAAAGGCGCCGGATCCTCCCCTATACTAGGTTATTGGAAAATTAAG
AR440	FLAG_mACP_for	33	forward	57.2	GACGATGACGACAAGgctgtggcccatggggac
AR441	STRII_DEBSII_for	49	forward	56.3	GAGCCACCCGCAGTTCGAAAAAGGCGCCGCTAGCACTGACAGCGAGAAG
AR442	pET22b_STRII_rev	54	reverse	53.4	GAACTGCGGGTGGCTCCAAGCGCTcatatgtatatctccttcttaaagttaaac
AR443	DEBS2_GGS_DEBS3_for	37	forward	55.9	GAGAGGACGGAGGTAGTAGCGGTGACAACGGCATGAC
AR444	DEBS2_GGS_DEBS3_rev	32	reverse	67.6	CTACCTCCGTCCTCTCCCCCGCCGAACCGCTG
AR445	pET22b_DEBSII_for	38	forward	55.9	gaaggagatatacatatgACTGACAGCGAGAAGGTGGC
AR446	DEBSII_for	21	forward	56.3	GCTAGCACTGACAGCGAGAAG
AR447	pET22b_msfGFP_For	33	forward	54.9	ggagatatacatatgAGCAAGGGCGAGGAGCTG
AR448	msfGFP_DEBSII_MPCR_rev	41	reverse	53.8	CTTCTCGCTGTCAGTGCTAGCCTTGTACAGCTCGTCCATGC
AR449	pUltra_MBP_for	45	forward	56.2	CAAGGACCCATAGCATATGAAAATCGAAGAAGGTAAACTGGTAATC
AR450	MBP_Linker_rev	26	reverse	56.4	GTTATTGTTGTTGTTGTTCGAGCTCG
AR451	MBP_mmPylS	45	forward	54.4	CAACAACAACAATAACGATAAAAAACCACTAAACACTCTGATATC
AR452	STRI_debsM4pro_for	36	forward	59.4	GAAAAAGGCGCCGGATCCGACGTCTCGGCGCTGGGC
AR453	debsM5con_H8_rev	32	reverse	59.1	GTGATGATGCTCGAGCCCGCCGGTGGGGATGG
AR454	DEBS2_GGS_DEBS3_MP_rev	38	reverse	67.6	CCGCTACTACCTCCGTCCTCTCCCCCGCCGAACCGCTG
AR455	STRI_rapsM1pro_for	37	forward	55.4	GAAAAAGGCGCCGGATCCGCCGATGGTCATCCACTGC
AR456	rapsM2con_H8_rev	32	reverse	56.7	GTGATGATGCTCGAGACCACCCGCAGTCGCCC
AR457	STRI_rapsM3pro_for	35	forward	54.3	GAAAAAGGCGCCGGATCCGCTGCGGGTGGTCATCC
AR458	rapsM4con_H8_rev	30	reverse	55.6	GTGATGATGCTCGAGCGCAGCCGCCCGGTC
AR459	STRI_rapsM7pro_for	35	forward	54.3	GAAAAAGGCGCCGGATCCGTTGCGGGCGGTCATCC
AR460	rapsM8con_H8_rev	30	reverse	55.6	GTGATGATGCTCGAGGCCCGCGGCCATCCG
AR461	STRI_rapsM13pro_for	35	forward	54.3	GAAAAAGGCGCCGGATCCGCTGCCGACGGTCATCC
AR462	rapsM14con_H8_rev	32	reverse	54.3	GTGATGATGCTCGAGAGCTGTGGCATCCGGGG
AR463	STRI_pikAM4pro_for	35	forward	54.3	GAAAAAGGCGCCGGATCCGACATCACCTCCGCCGG
AR464	pikAM5con_H8_rev	30	reverse	55.6	GTGATGATGCTCGAGGGCGGCCTGGTGGG
AR465	PIKA2_GGS_PIKA3_MP_for	40	forward	54.4	CTCCGACTTCGGAGGTTCAGCGAACAACGAAGACAAGCTC
AR466	PIKA2_GGS_PIKA3_MP_rev	30	reverse	54.9	CGCTGAACCTCCGAAGTCGGAGTCGCCCAG
AR467	Not1_mFAS_for	34	forward	55.4	GAAAAAGCGGCCGCAgaggaggtggtgatagccg
AR468	Not1_Strep1_rev	30	reverse	56.3	cTGCCGCCGCTTTTTCGAACTGCGGGTGGC
AR469	STRI_DEBS3_for	38	forward	55.9	GAAAAAGGCGCCGGATCCAGCGGTGACAACGGCATGAC
AR470	DEBS2_BamH1_DEBS3_rev	40	reverse	59.7	GTTGTCACCGCTGGATCCTCCGTCCTCTCCCCCGCCGAAC

AR471	STRI_PikAIII_for	39	forward	54.4	GAAAAAGGCGCCGGATCCGCGAACAACGAAGACAAGCTC
AR472	PikA2_BamH1_PikA3_rev	37	reverse	54.9	CGTTGTTCGCGGATCCTCCGAAGTCGGAGTCGCCCAG
AR473	DEBS3_H8_rev	34	reverse	54.9	GGTGATGATGCTCGAGCGAATTCCCTCCGCCCAG
AR474	pET22b_BamH1_DEBS3_for	41	forward	55.9	GGAGATATACATATGGGATCCAGCGGTGACAACGGCATGAC
AR475	STRI_DEBS2_for	38	forward	55.9	GAAAAAGGCGCCGGATCCACTGACAGCGAGAAGGTGGC
AR476	DEBS2_BamH1_DEBS3_rev	37	reverse	56.2	GTTGTCACCGCTGGATCCTCCGTCCTCTCCCCCGCCG
AR477	mmPylS_for	29	forward	54.4	GATAAAAAACCACTAAACACTCTGATATC
AR478	H6Sumo_for	37	forward	56.7	CAAGGACCATAGCATATGGGCAGCAGCCATCATCATC
AR479	hSumo1_mmPylS_rev	49	reverse	56.7	TAGTGGTTTTTTATCacctccATACACTCCACCGGTTTGTTCCTGGTAG
AR480	GST_rapsM13pro_for	34	forward	57.6	CATATGTCCATGGGCGCTGCCGACGGTCATCCAC
AR481	GST_GEKGAGS_for	36	forward	55.4	CATATGTCCATGGGCGGCGAAAAAGGCGCCGGATCC
AR482	MBP_Sumo1_for	43	forward	52.1	CAACAACAACAATAACGAAGGTGAATATATTAAACTCAAAGTC
AR483	pUltra_Sumo3_for	39	forward	54	CAAGGACCATAGCATatgaatgaccacatcaacctgaag
AR484	hSumo3_mmPylS_rev	45	reverse	54.9	TAGTGGTTTTTTATCacctccATACACTCCACCGGTctgctgctg
AR485	MBP_Sumo3_for	40	forward	55.7	CAACAACAACAATAACaatgaccacatcaacctgaaggtg
AR486	STRI_RAPSM14con_for	33	forward	56.7	AAAAGGCGCCGGATCCGAAGCTCCGGCGCCGAG
AR487	STRI_RAPSM13con_for	35	forward	54.3	GAAAAAGGCGCCGGATCCGTACGCCCGTCGGTGTC
AR488	RAPSM14_H8_rev	38	reverse	55.3	GTGATGATGCTCGAGTTCGTTCCTGGTCCTCCATTTTG
AR489	STRI_RAPSM12_psKR_for	31	forward	54.3	AAAGGCGCCGGATCCAACCCTGCCGAGCTGC
AR490	RAPSM12_KR_H8_rev	31	reverse	56.2	GTGATGATGCTCGAGTGCCAGGGCCGCCAGC
AR491	STRI_RAPSM12_X_for	31	forward	56.7	AAAGGCGCCGGATCCACGGACCGGGCGACTG
AR492	RAPSM13con_H8_rev	33	reverse	54.9	GTGATGATGCTCGAGTGGATGACCGTCGGCAGC
AR493	mDHstruc_QTV_for	43	forward	55.4	cactacctggtagacCAGAGCCGTGttccctggcactggctacc
AR494	Not1_mLD_for	36	forward	54.3	GTTCGAAAAAGCGGCCGCAcccactgcacacgctgc
AR495	postAT_Not1_rev	38	reverse	54.4	acctgctcTGCGGCCGCactgtggtcccacttgatgtg
AR496	mACP_Not1_rev	39	reverse	56.7	acctcctcTGCGGCCGCagacgtgtcactcctggacttg
AR497	STRI_RAPSM11con_for	37	forward	54.4	AAAAGGCGCCGGATCCCCTGAGCAAGACAAAGTGGTC
AR498	RAPSM11con_H8_rev	33	reverse	56.7	GGTGATGATGCTCGAGACCACCCGCAGTCGCCC
AR499	RAPSM11_KR_H8_rev	32	reverse	56.2	GGTGATGATGCTCGAGTGCCAGGGCGGCCAGC
AR500	RAPSM12con_H8_rev	33	reverse	56.7	GGTGATGATGCTCGAGAGTCGCCCGGTCCGTGC
AR501	STRI_RAPSM12con_for	31	forward	56.2	AAAAGGCGCCGGATCCGCCTTGGCTGCCCGG
AR502	mKR_K1988A_for	41	forward	54.8	ggatgtcaacaagcccGCGtacaatggcaccctgaacttg
AR503	mKR_K1988X_rev	22	reverse	56.7	gggctgttgacatcctggaag
AR504	mKR_K1988M_for	41	forward	54.8	ggatgtcaacaagcccATGtacaatggcaccctgaacttg
AR505	mKR_K1988AcK_for	41	forward	54.8	ggatgtcaacaagcccTAGtacaatggcaccctgaacttg
AR506	mKR_G1887X_rev	23	reverse	55.3	taggccaccagtgatgatgtaac
AR507	mKR_G1887A_for	37	forward	54.9	catcactggttggcctaGCGggctttggcctgagctg
AR508	Not1_AVES(AT0)_for	34	forward	54.3	GTTCGAAAAAGCGGCCCAGAGGATGGACGGCGGG
AR509	AVES(ACP0)_Not1_rev	32	reverse	56.2	cacctcctcTGCGGCCGCCGTGCCTCCGTGGC
AR510	Not1_DEBS(AT0)_for	37	forward	56.3	GTTCGAAAAAGCGGCCGACCTGTCAAAGCTCTCCGAC
AR511	DEBS(ACP0)_Not1_rev	34	reverse	57.2	cacctcctcTGCGGCCGGCGCCGCTTCGTTGGTC
AR512	STRI_mMAT2_for	37	forward	55.4	GAAAAAGGCGCCGGATCCtccaccaacaagcgcccac
AR513	mMAT2_H8_rev	34	reverse	55.4	GTGATGATGCTCGAGgatgcctgtgaggtgcacc
AR514	mMAT2_rev	19	reverse	55.4	gatgcctgtgaggtgcacc
AR515	mMAT2_mACP_Mega_for	37	forward	57.2	ggtgcacctcacaggcatcgctgtggcccatggggac
AR516	Not1_PikAI_for	36	forward	55.9	GTTCGAAAAAGCGGCCGGAATTACCAGGACCGGTGC
AR517	PikAI_Not1_rev	30	reverse	55.6	cacctcctcTGCGGCCACCGGCCGCCGCCAC
AR518	Not1_OlmAI_for	32	forward	56.2	GTTCGAAAAAGCGGCCTGCCGGCTGCCGGGTG
AR519	OlmAI_Not1_rev	24	reverse	57.9	cacctcctcTGCGGCGGCGGCCGG
AR520	Not1_PteAI_for	35	forward	55.4	GTTCGAAAAAGCGGCCGTCGAGTCCGAGCTCAACG
AR521	PteAI_Not1_rev	26	reverse	55.6	cacctcctcTGCGGCCACCGTCGCGG
AR522	Not1_mMAT_for	38	forward	55.4	GTTCGAAAAAGCGGCCGCAtccaccaacaagcgcccac
AR523	mACP2_Not1_rev	38	reverse		acctcctcTGCGGCCGCtgtcgtgtcagtagccgagtc
AR524	Not1_AVES(AT0)2_for	35	forward	55.4	GTTCGAAAAAGCGGCCATGGACGGCGGGGAAGAAC
AR525	AVES(ACP0)2_Not1_rev	35	reverse	56.7	cacctcctcTGCGGCCGGGTGGCCGCTGGGTGAGT
AR526	AR526_STRI_AVESACP0_for	35	forward	54.3	GAAAAAGGCGCCGGATCCGAGACGTCGGCGACCAC
AR527	AR527_AVESACP0_H8_rev	30	reverse	55.6	GTGATGATGCTCGAGTGCGGCCGCCGTGCC
AR528	AR528_STRI_AVESAT0_for	35	forward	54.3	GAAAAAGGCGCCGGATCCCAGAGGATGGACGGCGGG
AR529	AR529_AVESAT0_H8_rev	34	reverse	57.2	GGTGATGATGCTCGAGGGTGGTCGCCGACGTCTC
AR530	AR530_RAPSM14con_flag_rev	41	reverse		GTCTTTGTAGTCGGATCCGGCGCCAGCTGTGGCATCCGGGG
AR531	AR531_RAPS3_S5149A_for	41	forward		CAAGGATCTGGGCATGGATGCGCTGACGGCGATCGAACTCC
AR532	AR532_RAPS3_S5149X_rev	22	reverse	54.8	ATCCATGCCCAGATCCTTGAAC
AR533	AR533_RAPS3_Q5506TAG_for	34	forward	56.2	GCAGCGGGTGATCCGGTAGGCGCTGGCCAACGCC
AR534	AR534_RAPS3_Q5506X_rev	18	reverse	57.2	CCGGATCACCCGCTGCTG
AR535	AR535_RAPS3_A5509TAG_for	37	forward	56.2	GATCCGGCAGGCGCTGTAGAACGCCGGTCTCACCGCG
AR536	AR536_RAPS3_A5509X_rev	18	reverse	57.2	CAGCGCCTGCCGGATCAC
AR537	mMATstr_GGS_rev	43	reverse		CGATCCACCGCCAGAACCTCCACCgatgcctgtgaggtgcacc
AR538	GGS_His_for	38	forward		GTTCTGGCGGTGGATCGCTCGAGCATCATCACCACCAC
AR539	GGS_MalE_for	44	forward	53.7	GTTCTGGCGGTGGATCGAAAATCGAAGAAGGTAAACTGGTAATC
AR540	MalE_H8_rev	35	reverse	55.9	GTGATGATGCTCGAGAGTCTGCGCGTCTTTCAGGG
AR541	PikAIII_M5con_rev	15	reverse	55.6	GGCGGCGCTGGTGGG
AR542	PikAIII_M5c_M5p_for	32	forward	54.9	CCACCAGCGCCGCCGACGACTGGCGTTACCGC
AR543	PikAIII_M5p_H8_rev	33	reverse	57.2	GTGATGATGCTCGAGGGTGTTACGGGGGCCGAG
AR544	befPikA3KS5_ALAstop_for	39	forward	56.3	CAGTTCGAAAAAGGCtagGGATCCGCGAACAACGAAGAC
AR545	befDEBS3KS5_ALAstop_for	37	forward	55.4	CAGTTCGAAAAAGGCtagGGATCCAGCGGTGACAACG
AR546	STRI_DEBS3_ACP5_for	33	forward	58.3	GAAAAAGGCGCCGGATCCGGCGTGGCCGACGAGCCGC
AR547	DEBS3_ACP5_H8_rev	32	reverse	56.2	GGTGATGATGCTCGAGTGCGCCGACGAGCCGC
AR548	STRI_PikA3_ACP5_for	34	forward	56.2	GAAAAAGGCGCCGGATCCGGGCAGTCGTCGGCGC
AR549	PikAS3_ACP5_H8_rev	34	reverse	57.2	GGTGATGATGCTCGAGCGGTGCGAGGTACGCCTC
AR550	STRI_RAPS3_ACP14_for	37	forward	55.4	GAAAAAGGCGCCGGATCCCCCCTCCTATGCGGATGAGC

Table S2: List of plasmids.

Number	Important Construct	Name	Sequence confirmed
pAR001	MBP-hFAS	pAR01_StrepI_MBP_hFASe_H8_pET22b	✔
pAR002	TRX-hFAS	pAR02_StrepI_Trx_hFASe_H8_pET22b	✔
pAR003	Sumo-hFAS	pAR03_StrepI_SUMO3_hFASe_H8_pET22b	✔
		StrepI_hFASe_Δ11_TE_H8_pET22b	
		StrepI_hFASe_Δ10f_TE_H8_pET22b	
		StrepI_hFASe_Δ21_TE_H8_pET22b	
pAR004		StrepI_MBP_hFASe_Δ11_H8_pET22b	✔
pAR005		StrepI_MBP_hFASe_Δ10f_H8_pET22b	✔
pAR006		StrepI_MBP_hFASe_Δ21_H8_pET22b	✔
		StrepI_Trx_hFASe_Δ11_H8_pET22b	
pAR008		StrepI_Trx_hFASe_Δ10f_H8_pET22b	✔
pAR010	hFASe	pAR10_StrepI_hFASe(aSTARTm)_H8_pET22b	✔
pAR011	MBP-hFAS_aSm	pAR11_StrepI_MBP_hFASe(aSTARTm)_H8_pET22b	✔
pAR012	TRX-hFAS_aSm	pAR12_StrepI_Trx_hFASe(aSTARTm)_H8_pET22b	✔
pAR013	Sumo-hFAS_aSm	pAR13_StrepI_SUMO3_hFASe(aSTARTm)_H8_pET22b	✔
pAR014		pAR14_MSAS_H(10)_pET22(b)	✔
pAR015		pAR15_H8_Ceff_pET22(b)	✔
pAR016		pAR16_SacB_H(6)_pET22(b)	✔
pAR017	hFASh	pAR17_StrepI_hFASh_H8_pET22b	✔
pAR018	**1**; mFAS	pAR18_StrepI_mFASm_H8_pET22b	✔
pAR019	MBP-hFASe	pAR19_MBP_hFASe(aSm)_H8_pMAL-c5G	✔
pAR020		pAR20_STRII_hFASe(aSTARTm)_H8_pMAL-c5G	✔
pAR021		pAR21_StrepI_hFASe(aSTARTm)_H8_ΔT7t_pET22b	✔
pAR022		pAR22_IT1_StrepI_hFASe(aSTARTm)_H8_pET22b	✔
pAR023		pAR23_IT2_StrepI_hFASe(aSTARTm)_H8_pET22b	✔
pAR024		pAR24_IT3_StrepI_hFASe(aSTARTm)_H8_pET22b	✔
pAR025		pAR25_IT4_StrepI_hFASe(aSTARTm)_H8_pET22b	✔
pAR026	StrII-hFAS-H8	pAR26_StrepII_hFASe(aSTARTm)_H8_pET22b	✔
pAR027		pAR27_StrepI_hFASe(siSD_aSTARTm)_H8_pET22b	✔
pAR028		pAR28_MBP_hFASe(aSTARTm)_H8_pET22b	✔
pAR029		pAR29_Trx_hFASe_H8_pET22b	✔
pAR030		pAR30_SUMO3_hFASe_H8_pET22b	✔
pAR031		pAR31_Trx_hFASe_H8_pET32b	
pAR032		pAR33_mycdodecin_pETcoco	✔
pAR033		pAR34_hACP_pETcoco	✔
pAR034	DnaK-/DnaJ	pAR35_DnaK_RBS_DnaJ_pETcoco	
pAR035	TF+hFASe	pAR36_TF_RBS_StrepII_hFASe(aSTARTm)_H8_pET22b	✔
pAR036	TF(N+C)+hFASe	pAR36b_TF(N+C)_RBS_StrepII_hFASe(aSTARTm)_H8_pET22b	✔
pAR037		pAR37_DnaK_RBS_StrepII_hFASe(aSTARTm)_H8_pET22b	✔
pAR038		pAR38_MBP_hFASh_H8_pMAL-c5G	
pAR039		pAR39_pfuPre(b)_pETcoco	✔
pAR040		pAR40_pfuPre(b)_RBS_StrepII_hFASe(aSTARTm)_H8_pET22b	✔
pAR041		pAR41_TF_hFASe(aSm)_H8_pMAL-c5G	✔
pAR042		pAR42_GST_hFASe(aSm)_H8_pMAL-c5G	✔
pAR043	GroEL/GroES	pAR43_GroES_RBS_GroEL_pETcoco	✔
pAR044	DnaK-/DnaJ/GrpE	pAR44_DnaK_RBS_DnaJ_RBS_GrpE_pETcoco	✔
pAR045	prefoldin (pfuPfdB)	pAR45_pfuPre(b)_RBS_pfuPre(a)_pETcoco	✔
pAR046	PF1883	pAR46_PF1883_pETcoco	✔
pAR047	Cpn	pAR47_PfCPN_pETcoco	✔
pAR048	IbpA/IbpB	pAR48_IbpA_RBS_IbpB_pETcoco	✔
pAR049	ClpB	pAR49_ClpB_pETcoco	✔
pAR050		pAR50_PfuPreB_RBS_pfuPreA_T7_PfCPN_pETcoco	✔

pAR051		pAR51_ClpB_T7_DnaK_RBS_DnaJ_RBS_GrpE_pETcoco	✔
pAR052A		pAR52A_PfuPreB_RBS_pfuPreA_T7_PfCPN(MAmut)_pETcoco	
pAR052B		pAR52B_PfCPN(MAmut)_pETcoco	✔
pAR053A		pAR53A_STRI_hFASh_H8_pAG425GAL-ccdB	
pAR053B		pAR53B_STRI_hFASh_H8_pAG425GPD-ccdb	✔
pAR054A		pAR54A_STRI_mFASm_H8_pAG425GAL-ccdB	✔
pAR054B		pAR54B_STRI_mFASm_H8_pAG425GPD-ccdb	✔
pAR055		pAR55_STRI_hFASe_H8_pT7CFE1	
pAR056		pAR56_STRI_hFASh_H8_pT7CFE1	(✔)
pAR057		pAR57_PfCPN_H6Δlid_pETcoco	✔
pAR058		pAR58_StrepI_DEBS_Module_4_H8_pET22b	✔
pAR059		pAR59_MBP_hFASe(aSm)_ΔMAT_H8_pMAL-c5G	
pAR060		pAR60_StrepI_hMAT_pET22b	✔
pAR061		pAR61_hMAT_pET22b	✔
pAR062		pAR62_StrepI_m(KS_MAT)_hFASh_H8_pET22b	✔
pAR063		pAR63_StrepI_h(KS_MAT)_mFASm_H8_pET22b	✔
pAR064		pAR64_FLAG_mFASm_H8_pET22b	✔
pAR065		pAR65_PfCPN_hapical_pETcoco	
pAR066		pAR66_PfCPN(MAmut)_hapical_pETcoco	
pAR067		pAR67_hNT_PfCPN(MAmut)_hapical_pETcoco	
pAR068		pAR68_hNT_PfCPN_hapical_hCT_pETcoco	
pAR069	2; KS-MAT; condensing Part	pAR69_STRI_m(KS_MAT)_H8_pET22b	✔
pAR070	KS^C161G-MAT	pAR70_StrepI_m(KS(C161G)_MAT)_H8_pET22b	✔
pAR071		pAR71_StrepI_m(KS(C161G)_ATmut(R(606A))_H8_pET22b	✔
pAR072		pAR72_StrepI_m(KS(C161G)_MT(M499Q_F553Q)_H8_pET22b	✔
pAR073		pAR73_StrepI_m(KS(C161G)_MTmutplus_H8_pET22b	✔
pAR074		pAR74_StrepI_mFASm_ER(G1672V)_H8_pET22b	✔
pAR075		pAR75_StrepII_mACP_pET22b	✔
pAR076		pAR76_StrepI_celFAS_H8_pET22b	
pAR077		pAR77_StrepI_chiFAS_H8_pET22b	
pAR078		pAR78_StrepI_porFAS_H8_pET22b	
pAR079		pAR79_StrepI_ratFAS_H8_pET22b	
pAR080		pAR80_StrepI_maqFAS_H8_pET22b	
pAR081		pAR81_StrepI_m(KS_MAT)_chiFAS_H8_pET22b	✔
pAR082		pAR82_StrepI_h(KS_MAT)_H8_pET22b	✔
pAR083		pAR83_StrepI_h(KS_K436D_MAT)_H8_pET22b	✔
pAR084		pAR84_StrepI_h(KS_L833H_MAT)_H8_pET22b	✔
pAR085		pAR85_StrepI_h(KS_K436D_L833H_MAT)_H8_pET22b	✔
pAR086		pAR86_StrepI_mKS_hMAT_H8_pET22b	✔
pAR087		pAR87_StrepI_hKS_mMAT_H8_pET22b	✔
pAR088	14; mFAS_ΔTE	pAR88_StrepI_mFASm_ΔTE_H8_pET22b	✔
pAR089		pAR89_StrepI_mFASm_ΔACPTE_H8_pET22b	✔
pAR090		pAR90_StrepI_FabD_H8_pET22b	✔
pAR091		pAR91_StrepI_m(KS_K673T_MAT)_H8_pET22b	✔
pAR092		pAR92_StrepI_mFASm_LM203AA_H8_pET22b	✔
pAR093		pAR93_StrepI_mFASm_H293A_H8_pET22b	✔
pAR094		pAR94_StrepI_mFASm_E333Q_H8_pET22b	✔
pAR095		pAR95_StrepI_mFASm_M132G_Q136A_H8_pET22b	✔
pAR096		pAR96_StrepI_m(KS(C161G)_MTeFe)_H8_pET22b	✔
pAR097		pAR97_StrepI_m(KS(C161G)_MTeInt)_H8_pET22b	✔
pAR098		pAR98_StrepI_m(KS(C161G)_ATmut(R(606E))_H8_pET22b	✔
pAR099		pAR99_StrepI_m(KS(C161G)_MTeGate)_H8_pET22b	✔
pAR100	4; apo-ACP	pAR100_StrepII_mACP_H8_pET22b	✔
pAR101		pAR101_StrepI_DEBS2M3(KS_MMT)_H8_pET22b	✔
pAR102		pAR102_StrepI_m(KS(C161G)_MT(V675P))_H8_pET22b	✔
pAR103		pAR103_StrepI_m(KS(C161G)_MT(R676L))_H8_pET22b	✔
pAR104		pAR104_StrepI_m(KS(C161G)_MT(T677P))_H8_pET22b	✔
pAR105		pAR105_StrepI_m(KS(C161G)_MT(M499Q_F553Q_I549T))_H8_pET22b	✔
pAR106		pAR106_StrepI_m(KS(C161G)_MT(F682S))_H8_pET22b	✔
pAR107		pAR107_StrepI_m(KS(C161G)_MT(M499Q_F553Q_N738Q))_H8_pET22b	✔
pAR108		pAR108_StrepI_m(KS(C161G)_MT(M499Q_F553Q_V585Y)_H8_pET22b	✔
pAR109		pAR109_StrepI_m(KS(C161G)_ATmut(R606I)_H8_pET22b	✔
pAR110		pAR110_StrepI_m(KS(C161G)_ATmut(R606I_N738I)_H8_pET22b	✔

pAR111		pAR111_StrepI_m(KS(C161G)_ATmut(R606A_N738I)_H8_pET22b	✔
pAR112		pAR112_StrepI_m(KS(C161G)_MT(M499Q_I549T_F553Q_V585Y)_H8_pET22b	✔
pAR113		pAR113_StrepI_AVES2M3(Doc_KS_MT_ACP_L)_H8_pET22b	
pAR114		pAR114_StrepI_AVES4M11(KS_MMT_ACP_L)_H8_pET22b	
pAR115		pAR115_StrepI_mTEII_STOP_H8_pET22b	✔
pAR116		pAR116_StrepI_m(KS(C161G)_ATmut(R606M)_H8_pET22b	✔
pAR117		pAR117_StrepI_m(KS(C161G)_ATmut(E668A)_H8_pET22b	✔
pAR118		pAR118_StrepI_m(KS(C161G)_ATmut(L766E)_H8_pET22b	✔
pAR119		pAR119_StrepI_m(KS(C161G)_MT(F661R)_H8_pET22b	✔
pAR120		pAR120_StrepI_m(KS(C161G)_MT(PP635EEAPG)_H8_pET22b	✔
pAR121A		pAR121_StrepI_m(KS(C161G)_ATmut(R(606A)_I549T)_H8_pET22b	✔
pAR121B		pAR121B_StrepI_m(KS(C161G)_ATmut(I549T)_H8_pET22b	
pAR122		pAR122_StrepI_m(KS(C161G)_MT(N738Q))_H8_pET22b	✔
pAR123		pAR123_StrepI_m(KS(C161G)_MT(M499Q))_H8_pET22b	✔
pAR124		pAR124_StrepI_m(KS(C161G)_MT(F553Q))_H8_pET22b	✔
pAR125	**19**; KS-MAT-AM3L-ACP	pAR125_StrepI_m(KS_MAT)_AM3L_ACP_H8_pET22b	✔
pAR126	**20**; KS-MAT-AM11L-ACP	pAR126_StrepI_m(KS_MAT)_AM11L_ACP_H8_pET22b	✔
pAR127	**18**; KS-MAT-ACP	pAR127_StrepI_m(KS_MAT)_mouseL_ACP_H8_pET22b	✔
pAR128	**23**; KS-MAT-ACP-TE	pAR128_StrepI_m(KS_MAT)_mouseL_ACP_TE_H8_pET22b	✔
pAR129		pAR129_StrepI_mFASm_F200L_H8_pET22b	✔
pAR130		pAR130_StrepI_mFASm_F200M_H8_pET22b	
pAR131		pAR131_StrepI_mFASm_F395A_H8_pET22b	✔
pAR132		pAR132_StrepI_mFASm_F395M_H8_pET22b	✔
pAR133		pAR133_StrepI_mFASm_F395A_A160M_H8_pET22b	✔
pAR134		pAR134_StrepI_m(KS(C161G)_MT(M499G))_H8_pET22b	✔
pAR135		pAR135_StrepI_m(KS(C161G)_MT(L739G))_H8_pET22b	✔
pAR136A		pAR136_StrepI_m(KS(C161G)_ATmut(R606A_L739W))_H8_pET22b	✔
pAR136B		pAR136B_StrepI_m(KS(C161G)_ATmut(R606A_L739G))_H8_pET22b	✔
pAR137		pAR137_StrepI_m(KS(C161G)_MT(V585I_L739G))_H8_pET22b	✔
pAR138		pAR138_StrepI_m(KS(C161G)_MT(M499G_V585I_L739G))_H8_pET22b	✔
pAR139		pAR139_StrepI_m(KS(C161G)_MT(M499A))_H8_pET22b	✔
pAR140		pAR140_StrepI_m(KS(C161G)_MT(M499V))_H8_pET22b	✔
pAR141		pAR141_StrepI_m(KS(C161G)_MT(F553G))_H8_pET22b	✔
pAR142		pAR142_StrepI_m(KS(C161G)_MT(F553A))_H8_pET22b	✔
pAR143		pAR143_StrepI_m(KS(C161G)_MT(F553D))_H8_pET22b	✔
pAR144		pAR144_StrepI_m(KS(C161G)_MT(F553H))_H8_pET22b	✔
pAR145		pAR145_StrepI_m(KS(C161G)_MT(F553N))_H8_pET22b	✔
pAR146		pAR146_StrepI_m(KS(C161G)_MT(F553Y))_H8_pET22b	✔
pAR147		pAR147_StrepI_m(KS(C161G)_MT(L680Y_F682S))_H8_pET22b	✔
pAR148		pAR148_StrepI_m(KS(C161G)_MT(L680M))_H8_pET22b	✔
pAR149		pAR149_StrepI_m(KS(C161G)_MT(L680V))_H8_pET22b	✔
pAR150		pAR150_StrepI_m(KS(C161G)_MT(L680H))_H8_pET22b	✔
pAR151	**21**; KS-MAT-AM3L2-ACP	pAR151_StrepI_m(KS_MAT)_AM3L2_ACP_H8_pET22b	✔
pAR152	**22**; KS-MAT-AM11L2-ACP	pAR152_StrepI_m(KS_MAT)_AM11L2_ACP_H8_pET22b	✔
pAR153		pAR153_StrepI_m(KS(C161G)_MAT(S581C))_H8_pET22b	✔
pAR154		pAR154_StrepI_m(KS(K151R)_MAT)_mouseL_ACP_H8_pET22b	✔
pAR155		pAR155_StrepI_m(KS(K202R)_MAT)_mouseL_ACP_H8_pET22b	✔
pAR156		pAR156_StrepI_m(KS(V268E)_MAT)_mouseL_ACP_H8_pET22b	✔
pAR157		pAR157_StrepI_m(KS(V268A)_MAT)_mouseL_ACP_H8_pET22b	✔
pAR158		pAR158_StrepI_m(KS(L166M_I155T)_MAT)_mouseL_ACP_H8_pET22b	✔
pAR159	KS-MAT^{S581A}	pAR159_StrepI_m(KS_MAT(S581A))_H8_pET22b	✔
pAR160	KS^{C161G}-MAT^{S581A}	pAR160_StrepI_m(KS(C161G)_MAT(S581A))_H8_pET22b	✔
pAR161		pAR161_StrepI_m(KS(C161G)_MAT(S581A)_mouseL_ACP_H8_pET22b	✔
pAR162	**15**; mFAS_ΔER_ΔTE	pAR162_StrepI_mFASm_ΔER_ΔTE_H8_pET22b	✔
pAR163	**16**; mFAS_ΔDH_ΔER_ΔTE	pAR163_StrepI_mFASm_ΔDH_ΔER_ΔTE_H8_pET22b	✔
pAR164		pAR164_StrepI_mFASm_mER(G1672V)_ΔTE_H8_pET22b	✔
pAR165		pAR165_StrepI_mFASm_mDH(H878A)_mER(G1672V)_ΔTE_H8_pET22b	✔
pAR166		pAR166_StrepI_m(KS(C161G)_MAT_mouseL_ACP_H8_pET22b	✔
pAR167	**25**; KS-MAT-ACP-mFAS	pAR167_StrepI_m(KS_MAT_mL_ACP)_mL_mFASm_H8_pET22b	✔
pAR168	**24**; KS-MAT-ACP-m(KMA)	pAR168_StrepI_m(KS_MAT_mL_ACP)_mL_m(KS_MAT_mL_ACP)_H8_pET22b	✔
pAR169		pAR169_StrepI_mTEI_H8_pET22b	✔
pAR170		pAR170_StrepI_m(KS(G113M)_MAT(S581A))_H8_pET22b	✔

pAR171		pAR171_StrepI_m(KS(G113I)_MAT(S581A))_H8_pET22b	✔
pAR172		pAR172_StrepI_m(KS(G113F)_MAT(S581A))_H8_pET22b	✔
pAR173		pAR173_StrepI_m(KS(G113Y)_MAT(S581A))_H8_pET22b	✔
pAR174		pAR174_StrepI_m(KS(G113L)_MAT(S581A))_H8_pET22b	✔
pAR175		pAR175_StrepI_m(KS(G113W)_MAT(S581A))_H8_pET22b	✔
pAR176		pAR176_StrepI_m(KS(S117H)_MAT(S581A))_H8_pET22b	✔
pAR177		pAR177_StrepI_m(KS(S117Y)_MAT(S581A))_H8_pET22b	✔
pAR178		pAR178_StrepI_m(KS(M132F)_MAT(S581A))_H8_pET22b	✔
pAR179		pAR179_StrepI_m(KS(M132Y)_MAT(S581A))_H8_pET22b	✔
pAR180		pAR180_StrepI_m(KS(M132W)_MAT(S581A))_H8_pET22b	✔
pAR181		pAR181_StrepI_m(KS(F200Y)_MAT(S581A))_H8_pET22b	✔
pAR182		pAR182_StrepI_mFASm_mDH(H878A)_ER(G1672V)_H8_pET22b	✔
pAR183		pAR183_StrepI_mFASm_mTE(S2301A)_H8_pET22b	✔
pAR184		pAR184_StrepI_mFASm_ΔTE_STOP_H8_pET22b	✔
pAR185		pAR185_StrepI_mFASm_ΔDH_ΔER_H8_pET22b	✔
pAR186		pAR186_StrepI_mFASm_ΔDH_ΔER_tylHTH_H8_pET22b	✔
pAR187		pAR187_StrepI_mFASm_ΔDH_ΔER_tylmHanch_H8_pET22b	✔
pAR188		pAR188_StrepI_mFASm_ΔDH_ΔER_tylHTH_mHanch_H8_pET22b	✔
pAR189		pAR189_StrepI_m(KS_MAT)_AM3L2_ACP_AM3L3_m(KMA)_H8_pET22b	✔
pAR190		pAR190_StrepI_m(KS_MAT)_AM11L_ACP_AM11L3_M(KMA)_H8_pET22b	✔
pAR191		pAR191_StrepI_m(KS_MAT)_AM3ACP_m(KMA)_H8_pET22b	✔
pAR192		pAR192_StrepI_m(KS_MAT)_AM11ACP_m(KMA)_H8_pET22b	✔
pAR193		pAR193_mTEII_H8_pET22b	✔
pAR194	5 ; TE	pAR194_mTEI_H8_pET22b	✔
pAR195		pAR195_StrepI_m(KS(C161G)_AT(R606A)_mouseL_ACP_H8_pET22b	✔
pAR196		pAR196_StrepI_m(KS(C161G)_MT(M499G)_mouseL_ACP_H8_pET22b	✔
pAR197		pAR197_StrepI_m(KS(C161G)_MAT(H683A))_mouseL_ACP_H8_pET22b	✔
pAR198		pAR198_StrepI_m(KS(C161G)_MAT(S581A_H683A))_mouseL_ACP_H8_pET22b	✔
pAR199		pAR199_StrepI_m(KS_MAT)_AM4DE_KR_TE_H8_pET22b	✔
pAR200		pAR200_StrepI_m(KS_MAT)_AM5DE_KR_TE_H8_pET22b	✔
pAR201		pAR201_StrepI_m(KS_AT(R606A))_mouseL_ACP_H8_pET22b	✔
pAR202		pAR202_StrepI_m(KS_MMT(M499G))_mouseL_ACP_H8_pET22b	✔
pAR203		pAR203_StrepI_m(KS(C161Q)_MAT(S581A_H683A))_mouseL_ACP_H8_pET22b	
pAR204		pAR204_StrepI_m(KS(R379Q)_MAT)_H8_pET22b	✔
pAR205		pAR205_StrepI_m(KS(R379E)_MAT)_H8_pET22b	✔
pAR206		pAR206_StrepI_m(KS(SFA311ATE)_MAT)_H8_pET22b	✔
pAR207		pAR207_StrepI_m(KS_MAT)_mL_ACP(Q2182mut)_H8_pET22ultraPPF_R	✔
pAR208		pAR208_StrepI_mFAS_H8_pET22ultraPPF_R	
pAR209		pAR209_StrepI_m(KS_MAT)_mL_ACP_H8_pET22ultraPPF	
pAR210		pAR210_StrepI_mFAS_H8_pET22ultraPPF	
pAR211		pAR211_StrII_mmPylS_pBCN_Y306A_Y384F_pUltra	✔
pAR212		pAR212_StrII_mmPylS_aromatic_pUltra	✔
pAR213		pAR213_StrII_mmPylS_WT_pUltra	✔
pAR214		pAR214_mjTyrS_pCNF_D286R_pUltra	✔
pAR215		pAR215_mjTyrS_pTETF_D286R_pUltra	✔
pAR216		pAR216_StrepII_mACP_S2150A_msfGFP_H8_pET22b	✔
pAR217		pAR217_StrepI_h(KS(F256S)_MAT)_H8_pET22b	✔
pAR218		pAR218_StrepI_mFASm(DHL1)_H8_pET22b	✔
pAR219		pAR219_StrepI_mFASm(DHL2)_H8_pET22b	✔
pAR220		pAR220_StrepI_mFASm(DHL3)_H8_pET22b	✔
pAR221		pAR221_StrepI_mFASm(ERL1)_H8_pET22b	✔
pAR222		pAR222_StrepI_mFASm(ERL2)_H8_pET22b	✔
pAR223		pAR223_StrepI_m(KS(C161G)_MAT(S581A_H683A)_H8_pET22b	✔
pAR224		pAR224_StrepI_m(KS(C161G)_MAT(H683A))_H8_pET22b	✔
pAR225A		pAR225_StrepI_serMSAS(KS(C161G)_MAT_H8_pET22b	✔
pAR225B		pAR225b_StrepI_serMSAS(KS_MAT_H8_pET22b	✔
pAR226		pAR226_StrepI_mFASm(psKRL1)_H8_pET22b	✔
pAR227		pAR227_StrepI_mFASm(psKRL2)_H8_pET22b	✔
pAR228		pAR228_StrepI_mFASm(psMTL1)_H8_pET22b	✔
pAR229		pAR229_StrepI_mFASm_ΔpsMT_H8_pET22b	✔
pAR230		pAR230_StrepI_hFASh_shortLs_H8_pET22b	✔

pAR231		pAR231_StrepI_hFASh_shortLs2_H8_pET22b	
pAR232		pAR232_StrepI_hFASh_shortLs3_H8_pET22b	✔
pAR233		pAR233_StrII_mmPylS_AcKRS_pUltra	✔
pAR234		pAR234_StrepI_m(KS(C161G)_MAT(H683N))_H8_pET22b	✔
pAR235		pAR235_StrepI_m(KS(C161G)_MAT(H683D))_H8_pET22b	✔
pAR236	**3**; processing Part	pAR236_StrepI_m(DH_ps_ER_KR)_H8_pET22b	✔
pAR237	ΨME_ΨKR_ER_KR	pAR237_StrepI_m(ps_ER_KR)_H8_pET22b	✔
pAR238	ΨME_ΨKR_KR	pAR238_StrepI_m(ps_KR)_H8_pET22b	✔
pAR239	**17**; ΨKR_KR	pAR239_StrepI_m(psKR_KR)_H8_pET22b	✔
pAR240	PKSL_ΨKR_KR	pAR240_StrepI_m(PKSL_psKR_KR)_H8_pET22b	✔
pAR241	DH_ΨKR_ER_KR	pAR241_StrepI_m(DH_psKR_ER_KR)_H8_pET22b	✔
pAR242		pAR242_StrepI_DEBSI(psKR1_KR1)_H8_pET22b	✔
pAR243	**6**; KS$_{Extreme}$	pAR243_STRI_mKS_H8_pET22b	✔
pAR244	**7**; KS-LD	pAR244_STRI_m(KS_LD)_H8_pET22b	✔
pAR245	**10**; MAT$_{Extreme}$	pAR245_StrepI_mMAT_H8_pET22b	✔
pAR246	**8**; LD-MAT	pAR246_StrepI_m(LD_MAT)_H8_pET22b	✔
pAR247	**9**; LD-MAT-pAT	pAR247_STRI_m(LD_MAT_postATL)_H8_pET22b	✔
pAR248		pAR248_StrepI_MBP_mFASm_H8_pET22b	✔
pAR249		pAR249_StrepI_MBP_m(KS_MAT)_mouseL_ACP_TE_H8_pET22b	✔
pAR250		pAR250_StrepI_GST_mFASm_H8_pET22b	✔
pAR251		pAR251_StrepI_GST_m(KS_MAT)_mouseL_ACP_TE_H8_pET22b	✔
pAR252		pAR252_FLAG_mACP_H8_pET22b	✔
pAR253		pAR253_StrepI_m(KS(C161Q)_MAT(S581A_H683N))_H8_pET22b	
pAR254		pAR254_StrII_DEBS2_DEBS3_H6_pET22b	
pAR255		pAR255_DEBS2_DEBS3_H6_pET22b	
pAR256		pAR256_GFP_DEBS2_DEBS3_H6_pET22b	
pAR257		pAR257_MBP_mmPylS_WT_pUltra	✔
pAR258A		pAR258_StrepI_DEBS_M4pro_M5con_pET22b	
pAR258B	StrI-M4proc-M5cond$_{DEBS}$	pAR258B_StrepI_DEBS_M4pro_M5con_pET22b	✔
pAR259		pAR259_StrepI_RAPS_M1pro_M2con_pET22b	✔
pAR260		pAR260_StrepI_RAPS_M3pro_M4con_pET22b	
pAR261		pAR261_StrepI_RAPS_M7pro_M8con_pET22b	
pAR262		pAR262_StrepI_RAPS_M13pro_M14con_H8_pET22b1	✔
pAR263		pAR263_StrepI_PikA_M4pro_M5con_pET22b	✔
pAR264		pAR264_StrepI_Not1_mFASm_H8_pET22b	✔
pAR265		pAR265_StrepI_DEBS3M5(KS_MMT)_pET22b	✔
pAR266		pAR266_StrepI_PikA_M5con_pET22b	✔
pAR267		pAR267_StrepI_DEBS3_H8_pET22b	✔
pAR268		pAR268_DEBS3_H8_pET22b	✔
pAR269		pAR269_StrepI_DEBS2_DEBS3_H8_pET22b	✔
pAR270		pAR270_DEBS2_DEBS3_H8_pET22b	✔
pAR271		pAR271_H6_SUMO_mmPylS_WT_pUltra	✔
pAR272		pAR272_StrepI_DEBS_M4_M5con_pET22b	
pAR273		pAR273_MBP_mFASm_H8_pET22b	✔
pAR274		pAR274_MBP_m(KS_MAT)_mouseL_ACP_TE_H8_pET22b	✔
pAR275		pAR275_GST_RAPS_M13pro_M14con_H8_pMAL-c5G	
pAR276		pAR276_GST_GEK_RAPS_M13pro_M14con_H8_pMAL-c5G	(✔)
pAR277		pAR277_SUMO3_mmPylS_WT_pUltra	
pAR278		pAR278_MBP_SUMO_mmPylS_WT_pUltra	
pAR279		pAR279_MBP_SUMO3_mmPylS_WT_pUltra	
pAR280		pAR280_StrepI_RAPS_M14con_pET22b	✔
pAR281		pAR281_StrepI_RAPS_M13_M14con_pET22b	✔
pAR282		pAR282_StrepI_RAPS_M13pro_M14_pET22b	✔
pAR283		pAR283_StrepI_RAPS_M14_pET22b	✔
pAR284		pAR284_StrepI_RAPS_M12_psKR_KR_pET22b	✔
pAR285		pAR285_StrepI_RAPS_M12_DH°_psKR_KR_pET22b	✔
pAR286		pAR286_StrepI_RAPS_M13con_pET22b	✔
pAR287		pAR287_StrepI_RAPS_M12pro_M13con_pET22b	
pAR288		pAR288_StrepI_RAPS_M13_M14_pET22b	✔
pAR289		pAR289_StrepI_RAPS_M12pro_M13_M14_pET22b	
pAR290		pAR290_StrepI_mFASm_DH(Struc)_ΔER_ΔTE_H8_pET22b	

pAR291	**30**; LD-MAT-mFAS	pAR291_StrepI_m(LD_MAT)_mFASm_H8_pET22b	✔
pAR292	**26**; LD-MAT-ACP-mFAS	pAR292_StrepI_m(LD_MAT_mACP)_mFASm_H8_pET22b	✔
pAR293		pAR293_StrepI_RAPS_M11con_pET22b	✔
pAR294		pAR294_StrepI_RAPS_M11_DH°_psKR_KR_pET22b	✔
pAR295		pAR295_StrepI_RAPS_M11pro_M12con_pET22b	
pAR296		pAR296_StrepI_RAPS_M12con_pET22b	✔
pAR297		pAR297_StrepI_m(LD_MAT(R606A)_mACP)_mFASm_H8_pET22b	✔
pAR298		pAR298_StrepI_m(LD_MAT(S581A)_mACP)_mFASm_H8_pET22b	✔
pAR299		pAR299_StrepI_m(LD_MAT_mACP(S2150A))_mFASm_H8_pET22b	✔
pAR300		pAR300_StrepI_m(LD_MAT(S581A)_mACP(S2150A))_mFASm_H8_pET22b	✔
pAR301		pAR301_StrepI_m(DH_ps_ER_KR(K1988A))_H8_pET22b	✔
pAR302		pAR302_StrepI_m(DH_ps_ER_KR(K1988M))_H8_pET22b	✔
pAR303		pAR303_StrepI_m(DH_ps_ER_KR(K1988AcK))_H8_pET22b	✔
pAR304		pAR304_StrepI_m(DH_ps_ER_KR(G1887A))_H8_pET22b	✔
pAR305		pAR305_StrepI_m(DH_ps_ER(G1672V)_KR)_H8_pET22b	✔
pAR306	**27**; (AT0-ACP0)$_{AVES}$-mFAS	pAR306_StrepI_AVES(AT0_ACP0)_mFASm_H8_pET22b	✔
pAR307	**28**; (AT0-ACP0)$_{DEBS}$-mFAS	pAR307_StrepI_DEBS(AT0_ACP0)_mFASm_H8_pET22b	✔
pAR309	**11**; MAT$_{Ave}$	pAR309_StrepI_mMAT(Structure)_H8_pET22b	✔
pAR310	**13**; MAT$_{Ave}$-ACP	pAR310_StrepI_mMAT(Str)_L1_ACP_H8_pET22b	✔
pAR311		pAR311_StrepI_mMAT_mouseL_ACP_H8_pET22b	✔
pAR312		pAR312_StrepI_PikAI(KSq_AT0_ACP0)_mFASm_H8_pET22b	
pAR313		pAR313_StrepI_OlmAI(KSq_AT0_ACP0)_mFASm_H8_pET22b	
pAR314		pAR314_StrepI_PteA1(KSq_AT0_ACP0)_mFASm_H8_pET22b	
pAR315		pAR315_StrepI_m(MAT(Str)_mACP)_mFASm_H8_pET22b	✔
pAR316		pAR316_StrepI_AVES(AT0_ACP0)2_mFASm_H8_pET22b	✔
pAR317		pAR317_STRI_m(KS_MAT(R606A))_H8_pET22b	✔
pAR318		pAR318_StrepI_AVES(ACP0)_H8_pET22b	✔
pAR319		pAR319_StrepI_AVES(AT0)_H8_pET22b	✔
pAR320		pAR320_StrepI_RAPS_M13pro_M14con_flag_pET22b	
		Numbers pAR317-320 were utilized twice	
pAR317		pAR317_StrepI_mFASm_ΔER_ΔTE_DH(H878A)_H8_pET22b	✔
pAR318		pAR318_StrepI_mFASm_DHL2(Struc)_ΔER_ΔTE_H8_pET22b	✔
pAR319		pAR319_StrepI_mFASm_ΔER_ΔTE_DHL2(H878A)_H8_pET22b	✔
pAR320		pAR320_StrepI_mFASm_ΔER_H8_pET22b	✔
pAR321		pAR321_StrepI_RAPS_M13pro(S5149A)_M14con_flag_pET22b	✔
pAR322		pAR322_StrepI_RAPS_M13p_M14c_H8_pET22b	✔
pAR323		pAR323_StrepI_RAPS_M13p_M14c___H8_pET22b	✔
pAR324		pAR324_StrepI_RAPS_M13p_M14c*_H8_StrepI_RAPS_M13p(S5149A)_M14c_flag_pET22b	
pAR325		pAR325_StrepI_RAPS_M13p_M14c**_H8_StrepI_RAPS_M13p(S5149A)_M14c_flag_pET22b	
pAR326		pAR326_StrepI_mMAT(Structure)_GGGS_H8_pET22b	✔
pAR327	**12**; MAT$_{Ave}$-MBP	pAR327_StrepI_mMAT(Structure)_GGGS_MBP_H8_pET22b-1	✔
pAR328		pAR328_StrepI_PikAIII_M5_H8_pET22b	✔
pAR329		pAR329_StrepI_ALAmut_PikAIII_M5_pET22b	✔
pAR330		pAR330_StrepI_ALAmut_DEBS3_H8_pET22b	✔
pAR331		pAR331_StrepI_DEBS3(ACP5)_H8_pET22b	✔
pAR332		pAR332_StrepI_PikAIII(ACP5)_pET22b	✔
pAR333		pAR333_StrepI_RAPS(ACP14)_H8_pET22b	✔
pAR334		pAR334_StrII_RtFAS2_pET24b	
pAR335		pAR335_RtFAS2_StrII_pET24b	
pAR336		pAR336_StrepI_EcACP_H8_pET22b	✔
pAR337		pAR337_mmPylRS_NorY_pEVOL	✔
pAR338		pAR338_StrepI_mMAT(Str)_GGS_mACP(S2150A)_GGS_mFASm_H8_pET22b	
pAR339		pAR339_StrepI_mMAT(Str)(S581A)_GGS_mACP_GGS_mFASm_H8_pET22b	
pAR340	**29**; MAT$_{Ave}$-ACP-mFAS	pAR340_StrepI_mMAT(Str)_GGS_mACP_GGS_mFASm_H8_pET22b	✔
pAR341		pAR341_StrepI_mMAT(Str)(R606A)_GGS_mACP_GGS_mFASm_H8_pET22b	✔
pAR342		pAR342_StrepI_mMAT(Str)_GGS_mFASm_H8_pET22b	✔
pAR343		pAR343_StrepII_mACP_msfGFP_pET22b	✔
pAR344		pAR344_StrepI_m(KS(C161G)_MT(L680I)_H8_pET22b	✔
pAR345		pAR345_StrepI_m(KS(C161G)_MT(L680I)_(A680P)_H8_pET22b	✔
pAR346		pAR346_StrepI_m(KS(C161G)_MT(L582V))_H8_pET22b	✔
pAR347		pAR347_StrepI_m(KS(C161G)_MT(L582A))_H8_pET22b	✔
pAR348		pAR348_StrepI_m(KS(C161G)_MT(L582G))_H8_pET22b	✔
pAR349		pAR349_StrepI_m(KS(C161G)_MT(L582M))_H8_pET22b	✔
pAR350		pAR350_StrepI_m(KS(C161G)_MT(L582T))_H8_pET22b	✔

pAR351		pAR351_StrepI_m(KS(C161G)_MT(L582Q))_H8_pET22b	✔
pAR352A		pAR352_StrepII_mACP_H8_RBS_SFP_pET22b	
pAR352B	ACP	pAR352b_StrepII_mACP_H8_RBS_SFP_pET22b	✔
pAR353		pAR353_StrepI_mFASm_TEV_ACP_TE_H8_pET22b	
pAR354		pAR354_StrepI_mFASm_XA_ACP_TE_H8_pET22b	
pAR355	GST-M4proc-M5cond$_{DEBS}$	pAR355_GST_DEBS_M4pro_M5con_H8_pET22b	✔
pAR356		pAR356_CC_DEBS_M4pro_M5con_H8_pET22b	
pAR357	Sfp	pAR357_SFP_pCDF-1b	✔
pAR358		pAR358_StrepI_m(KS(C161G)_MAT(H580A))_H8_pET22b	✔
pAR359		pAR359_StrepI_m(KS(C161G)_MAT(H580N))_H8_pET22b	✔
pAR360		pAR360_StrepI_m(KS(C161G)_MAT(M620A))_H8_pET22b	✔
pAR361		pAR361_StrepI_m(KS(C161G)_MAT(C642A))_H8_pET22b	✔
pAR362		pAR362_StrepI_m(KS(C161G)_MAT(N644A))_H8_pET22b	✔
pAR363		pAR363_StrepI_m(KS(C161G)_MAT(N644S))_H8_pET22b	✔
pAR364		pAR364_StrepI_m(KS(C161G)_MAT(T650A))_H8_pET22b	✔
pAR365		pAR365_StrepI_m(KS(C161G)_MAT(T650V))_H8_pET22b	✔
pAR366	GST-M4proc-M5cond$_{DEBS}$-GST	pAR366_GST_DEBS_M4pro_M5con_GST_H8_pET22b	✔

5.5 Literature

[1] von Nussbaum, F., Brands, M., Hinzen, B., Weigand, S., and Häbich, D. (2006) Antibacterial Natural Products in Medicinal Chemistry—Exodus or Revival?, *Angew. Chem. Int. Ed. 45*, 5072-5129.

[2] Murata, M., Naoki, H., Matsunaga, S., Satake, M., and Yasumoto, T. (1994) Structure and Partial Stereochemical Assignments for Maitotoxin, the Most Toxic and Largest Natural Non-Biopolymer, *J. Am. Chem. Soc. 116*, 7098-7107.

[3] Krämer, K. (2015) Chemistry's toughest total synthesis challenge put on hold by lack of funds, In *Chemistry World*.

[4] Woodward, R. B., Logusch, E., Nambiar, K. P., Sakan, K., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., and Chen, C. H. (1981) Asymmetric total synthesis of erythromcin. 1. Synthesis of an erythronolide A secoacid derivative via asymmetric induction, *J. Am. Chem. Soc. 103*, 3210-3213.

[5] Woodward, R. B., Au-Yeung, B. W., Balaram, P., Browne, L. J., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., and Chen, C. H. (1981) Asymmetric total synthesis of erythromycin. 2. Synthesis of an erythronolide A lactone system, *J. Am. Chem. Soc. 103*, 3213-3215.

[6] Woodward, R. B., Logusch, E., Nambiar, K. P., Sakan, K., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., and Chen, C. H. (1981) Asymmetric total synthesis of erythromycin. 3. Total synthesis of erythromycin, *J. Am. Chem. Soc. 103*, 3215-3217.

[7] Rose, W., and Rybak, M. (2006) Tigecycline: first of a new class of antimicrobial agents, *Pharmacotherapy 26*, 1099–1110.

[8] Gennari, C., Vulpetti, A., Donghi, M., Mongelli, N., and Vanotti, E. (1996) Semisynthesis of Taxol: A Highly Enantio- and Diastereoselective Synthesis of the Side Chain and a New Method for Ester Formation at C13 Using Thioesters, *Angew. Chem. Int. Ed. 35*, 1723-1725.

[9] Boehm, M., Fuenfschilling, P. C., Krieger, M., Kuesters, E., and Struber, F. (2007) An Improved Manufacturing Process for the Antimalaria Drug Coartem. Part I, *Organic Process Research & Development 11*, 336-340.

[10] Dewick, P. M. (2009) *Medicinal Natural Products: A Biosynthetic Approach*, Vol. 3rd ed., Wiley.

[11] Ridley, C. P., Lee, H. Y., and Khosla, C. (2008) Evolution of polyketide synthases in bacteria, *Proceedings of the National Academy of Sciences 105*, 4595-4600.

[12] Castaldo, G., Zucko, J., Heidelberger, S., Vujaklija, D., Hranueli, D., Cullum, J., Wattana-Amorn, P., Crump, M. P., Crosby, J., and Long, P. F. (2008) Proposed Arrangement of Proteins Forming a Bacterial Type II Polyketide Synthase, *Chem. Biol. 15*, 1156-1165.

[13] Cox, R. J. (2007) Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets, *Org. Biomol. Chem. 5*, 2010.

[14] Keatinge-Clay, A. T. (2012) The structures of type I polyketide synthases, *Nat Prod Rep. 29*, 1050-1073.

[15] McGuire, J. M., Bunch, R. L., Anderson, R. C., Boaz, H. E., Flynn, E. H., Powell, H. M., and Smith, J. W. (1952) Ilotycin, a new antibiotic, *Antibiot. Chemother. 2*, 281-283.

[16] Heilman, F. R., Herrell, W. E., Wellmann, W. E., and Geraci, J. E. (1952) Some laboratory and clinical observations on a new antibiotic, erythromycin (Ilotycin), *Mayo Clin. Proc. 27*, 285-304.

[17] Washington II, J. A., and Wilson, W. R. (1985) Erythromycin: A Microbial and Clinical Perspective After 30 Years of Clinical Use (First of Two Parts), *Mayo Clin. Proc. 60*, 189-203.

[18] Washington II, J. A., and Wilson, W. R. (1985) Erythromycin: A Microbial and Clinical Perspective After 30 Years of Clinical Use (Second of Two Parts), *Mayo Clin. Proc. 60*, 271-278.

[19] Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., and Cane, D. E. (2007) Structure and Mechanism of the 6-Deoxyerythronolide B Synthase, *Annu. Rev. Biochem. 76*, 195-221.

[20] Bevitt, D. J., Cortés, J., Haydock, S. F., and Leadlay, P. F. (1992) 6-Deoxyerythronolide-B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme, *Eur. J. Biochem. 204*, 39-49.

[21] Roberts, G. A., Staunton, J., and Leadlay, P. F. (1993) 6-Deoxyerythronolide B synthase 3 from *Saccharopolyspora erythraea*: over-expression in *Escherichia coli*, purification and characterisation, *Biochem. Soc. Trans 21*, 32S.

[22] Ikeda, H., Nonomiya, T., Usami, M., Ohta, T., and Omura, S. (1999) Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. U.S.A. 96*, 9509-9514.

[23] Cane, D. E., Liang, T.-C., Kaplan, L. K., Nallin, M. K., Schulman, M. D., Hensens, O. D., Douglas, A. W., and Albers-Schönberg, G. (1983) Biosynthetic origin of the carbon skeleton and oxygen atoms of the avermectins, *J. Am. Chem. Soc. 105*, 4110-4112.

[24] Wang, F., Wang, Y., Ji, J., Zhou, Z., Yu, J., Zhu, H., Su, Z., Zhang, L., and Zheng, J. (2015) Structural and Functional Analysis of the Loading Acyltransferase from Avermectin Modular Polyketide Synthase, *ACS Chem. Biol. 10*, 1017-1025.

[25] Grininger, M. (2014) Perspectives on the evolution, assembly and conformational dynamics of fatty acid synthase type I (FAS I) systems, *Curr. Opin. Struct. Biol. 25*, 49-56.

[26] White, S. W., Zheng, J., Zhang, Y.-M., and Rock, C. O. (2005) The Structural Biology of Type II Fatty Acid Biosynthesis, *Annu Rev Biochem. 74*, 791-831.

[27] Maier, T., Leibundgut, M., Boehringer, D., and Ban, N. (2010) Structure and function of eukaryotic fatty acid synthases, *Q. Rev. Biophys. 43*, 373-422.

[28] Jenni, S., Leibundgut, M., Boehringer, D., Frick, C., Mikolasek, B., and Ban, N. (2007) Structure of fungal fatty acid synthase and implications for iterative substrate shuttling, *Science 316*, 254-261.

[29] Leibundgut, M., Jenni, S., Frick, C., and Ban, N. (2007) Structural Basis for Substrate Delivery by Acyl Carrier Protein in the Yeast Fatty Acid Synthase, *Science 316*, 288-290.

[30] Wakil, S. J. (1989) Fatty Acid Synthase, A Proficient Multifunctional Enzyme, *Biochemistry 28*, 4523-4530.

[31] Maier, T., Leibundgut, M., and Ban, N. (2008) The Crystal Structure of a Mammalian Fatty Acid Synthase, *Science 321*, 1315-1322.

[32] Jenke-Kodama, H., Sandmann, A., Müller, R., and Dittmann, E. (2005) Evolutionary Implications of Bacterial Polyketide Synthases, *Mol. Biol. Evol. 22*, 2027-2039.

[33] Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Pontén, F. (2015) Tissue-based map of the human proteome, *Science 347*, 1260419.

[34] Semenkovich, C. F., Coleman, T., and Fiedorek, F. T. (1995) Human fatty acid synthase mRNA: tissue distribution, genetic mapping, and kinetics of decay after glucose deprivation, *J. Lipid Res. 36*, 1507-1521.

[35] Libertini, L. J., and Smith, S. (1978) Purification and Properties of a Thioesterase from Lactating Rat Mammary Gland which Modifies the Product Specificity of Fatty Acid Synthetase, *J. Biol. Chem. 253*, 1393-1401.

[36] Yamazaki, T., Seyama, Y., Otsuka, H., Ogawa, H., and Yamakawa, T. (1981) Identification of Alkyldiacylglycerols Containing Saturated Methyl Branched Chains in the Harderian Gland of Guinea Pig, *J. Biochem. 89*, 683-691.

[37] Buckner, J. S., Kolattukudy, P. E., and Rogers, L. (1978) Synthesis of Multimethyl-Branched Fatty Acids by Avian and Mammalian Fatty Acid Synthetase and Its Regulation by Malonyl-CoA Decarboxylase in the Uropygial Gland, *Arch. Biochem. Biophys. 186*, 152-163.

[38] Smith, S., and Tsai, S.-C. (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases, *Nat. Prod. Rep. 24*, 1041-1072.

[39] Brady, R. O., and Gurin, S. (1952) Biosynthesis of fatty acids by cell-free or water-soluble enzyme systems, *J. Biol. Chem. 199*, 421-431.

[40] Dituri, F., and Gurin, S. (1953) Lipogenesis by homogenates or particle-free extracts of rat liver, *Arch. Biochem. Biophys. 43*, 231-232.

[41] Hsu, R. Y., Wasson, G., and Porter, J. W. (1965) The Purification and Properties of the Fatty Acid Synthetase of Pigeon Liver, *J. Biol. Chem. 240*, 3736-3746.

[42] Abdinejad, A., Fisher, A. M., and Kumar, S. (1981) Production and Utilization of Butyryl-CoA by Fatty Acid Synthetase from Mammalian Tissues, *Arch. Biochem. Biophys. 208*, 135-145.

[43] Katiyar, S. S., Cleland, W. W., and Porter, J. W. (1975) Fatty acid synthetase. A steady state kinetic analysis of the reaction catalyzed by the enzyme from pigeon liver, *J. Biol. Chem. 250*, 2709-2717.

[44] Cox, B. G., and Hammes, G. G. (1983) Steady-state kinetic study of fatty acid synthase from chicken liver, *Proc. Natl. Acad. Sci. U.S.A. 80*, 4233-4237.

[45] Cleland, W. W. (1963) The kinetics of enzyme-catalyzed reactions with two or more substrates or products, *Biochim. Biophys. Acta 67*, 104-137.

[46] Cleland, W. W. (1963) The kinetics of enzyme-catalyzed reactions with two or more substrates or products, *Biochim. Biophys. Acta 67*, 188-196.

[47] Smith, S., and Stern, A. (1983) The Effect of Aromatic CoA Esters on Fatty Acid Synthetase: Biosynthesis of ω-Phenyl Fatty Acids, *Arch. Biochem. Biophys. 222*, 259-265.

[48] Joshi, A. K., and Smith, S. (1993) Construction of a cDNA encoding the multifunctional animal fatty acid synthase and expression in *Spodoptera frugiperda* cells using baculoviral vectors, *Biochem. J. 296*, 143-149.

[49] Witkowski, A., Joshi, A. K., and Smith, S. (1997) Characterization of the Interthiol Acyltransferase Reaction Catalyzed by the β-Ketoacyl Synthase Domain of the Animal Fatty Acid Synthase, *Biochemistry 36*, 16338-16344.

[50] Witkowski, A., Joshi, A. K., and Smith, S. (2002) Mechanism of the β-Ketoacyl Synthase Reaction Catalyzed by the Animal Fatty Acid Synthase, *Biochemistry 41*, 10877-10887.

[51] Rangan, V. S., and Smith, S. (1997) 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.

[52] Rangan, V. S., Joshi, A. K., and Smith, S. (2001) Mapping the Functional Topology of the Animal Fatty Acid Synthase by Mutant Complementation *in Vitro*, *Biochemistry 40*, 10792-10799.

[53] Maier, T., Jenni, S., and Ban, N. (2006) Architecture of Mammalian Fatty Acid Synthase at 4.5 Å Resolution, *Science 311*, 1258-1262.

[54] Brignole, E. J., Smith, S., and Asturias, F. J. (2009) Conformational flexibility of metazoan fatty acid synthase enables catalysis, *Nat. Struct. Mol. Biol. 16*, 190-197.

[55] Zheng, J., Gay, D. C., Demeler, B., White, M. A., and Keatinge-Clay, A. T. (2012) Divergence of multimodular polyketide synthases revealed by a didomain structure, *Nat. Chem. Biol. 8*, 616-622.

[56] Barajas, J. F., Blake-Hedges, J. M., Bailey, C. B., Curran, S., and Keasling, J. D. (2017) Engineered polyketides: Synergy between protein and host level engineering, *Synth. Syst. Biotechnol. 2*, 147-166.

[57] Kage, H., Kreutzer, M. F., Wackler, B., Hoffmeister, D., and Nett, M. (2013) An Iterative Type I Polyketide Synthase Initiates the Biosynthesis of the Antimycoplasma Agent Micacocidin, *Chem. Biol. 20*, 764-771.

[58] Copeland, R. A. (2005) *Enzymes - A Practical Introduction to Structure, Mechanism and Data Analysis - 2nd Edition*, Weinheim, Germany: Wiley-VCH Verlag GmbH.

[59] Joshi, V. C., and Wakil, S. J. (1971) Studies on the mechanism of fatty acid synthesis. XXVI. Purification and properties of malonyl-coenzyme A–acyl carrier protein transacylase of *Escherichia coli*, *Arch. Biochem. Biophys. 143*, 493-505.

[60] Liou, G. F., Lau, J., Cane, D. E., and Khosla, C. (2003) Quantitative Analysis of Loading and Extender Acyltransferases of Modular Polyketide Synthases, *Biochemistry 42*, 200-207.

[61] Gregory, M. A., Petkovic, H., Lill, R. E., Moss, S. J., Wilkinson, B., Gaisser, S., Leadlay, P. F., and Sheridan, R. M. (2005) Mutasynthesis of Rapamycin Analogues through the Manipulation of a Gene Governing Starter Unit Biosynthesis, *Angew. Chem. Int. Ed. 44*, 4757-4760.

[62] Yadav, G., Gokhale, R. S., and Mohanty, D. (2003) Computational Approach for Prediction of Domain Organization and Substrate Specificity of Modular Polyketide Synthases, *J. Mol. Biol. 328*, 335-363.

[63] Peter, D. M., Schada von Borzyskowski, L., Kiefer, P., Christen, P., Vorholt, J. A., and Erb, T. J. (2015) Klassifizierung und Manipulation des synthetischen Potenzials carboxylierender Reduktasen aus dem Zentralmetabolismus und der Polyketid-Biosynthese, *Angew. Chem. 127*, 13659-13663.

[64] Pappenberger, G., Benz, J., Gsell, B., Hennig, M., Ruf, A., Stihle, M., Thoma, R., and Rudolph, M. G. (2010) Structure of the Human Fatty Acid Synthase KS–MAT Didomain as a Framework for Inhibitor Design, *J. Mol. Biol. 397*, 508-519.

[65] Miyanaga, A., Iwasawa, S., Shinohara, Y., Kudo, F., and Eguchi, T. (2016) Structure-based analysis of the molecular interactions between acyltransferase and acyl carrier protein in vicenistatin biosynthesis, *Proc. Natl. Acad. Sci. U. S. A. 113*, 1802-1807.

[66] Haydock, S. F., Aparicio, J. F., Molnár, I., Schwecke, T., Khaw, L. E., König, A., Marsden, A. F., Galloway, I. S., Staunton, J., and Leadlay, P. F. (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl) malonyl-CoA: acyl carrier protein transacylase domains in modular polyketide synthases, *FEBS Lett. 374*, 246-248.

[67] Lau, J., Fu, H., Cane, D. E., and Khosla, C. (1999) Dissecting the Role of Acyltransferase Domains of Modular Polyketide Synthases in the Choice and Stereochemical Fate of Extender Units, *Biochemistry 38*, 1643-1651.

[68] Arthur, C. J., Williams, C., Pottage, K., Płoskoń, E., Findlow, S. C., Burston, S. G., Simpson, T. J., Crump, M. P., and Crosby, J. (2009) Structure and Malonyl CoA-ACP Transacylase Binding of *Streptomyces coelicolor* Fatty Acid Synthase Acyl Carrier Protein, *ACS Chem. Biol. 4*, 625-636.

[69] Dunn, B. J., Cane, D. E., and Khosla, C. (2013) Mechanism and Specificity of an Acyltransferase Domain from a Modular Polyketide Synthase, *Biochemistry 52*, 1839-1841.

[70] Herbst, D. A., Jakob, R. P., Zähringer, F., and Maier, T. (2016) Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases, *Nature 531*, 533-537.

[71] Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) Tolerance and specificity of polyketide synthases, *Annu. Rev. Biochem. 68*, 219-253.

[72] Blaquiere, N., Shore, D. G., Rousseaux, S., and Fagnou, K. (2009) Decarboxylative Ketone Aldol Reactions: Development and Mechanistic Evaluation under Metal-Free Conditions, *J. Org. Chem. 74*, 6190-6198.

[73] Sari, D., Gupta, K., Raj, D. B. T. G., Aubert, A., Drncová, P., Garzoni, F., Fitzgerald, D., and Berger, I. (2016) The MultiBac Baculovirus/Insect Cell Expression Vector System for Producing Complex Protein Biologics, In *Advanced Technologies for Protein Complex Production and Characterization* (Vega, M. C., Ed.), pp 199-215, Springer International Publishing, Cham.

[74] Jayakumar, A., Huang, W. Y., Raetz, B., Chirala, S. S., and Wakil, S. J. (1996) Cloning and expression of the multifunctional human fatty acid synthase and its subdomains in *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A. 93*, 14509-14514.

[75] Ronnett, G. V., Kim, E.-K., Landree, L. E., and Tu, Y. (2005) Fatty acid metabolism as a target for obesity treatment, *Physiol. Behav. 85*, 25-35.

[76] Menendez, J. A., Vazquez-Martin, A., Ortega, F. J., and Fernandez-Real, J. M. (2009) Fatty Acid Synthase: Association with Insulin Resistance, Type 2 Diabetes, and Cancer, *Clin. Chem. 55*, 425-438.

[77] Baneyx, F., and Mujacic, M. (2004) Recombinant protein folding and misfolding in *Escherichia coli*, *Nat. Biotechnol. 22*, 1399-1408.

[78] Kolaj, O., Spada, S., Robin, S., and Wall, J. G. (2009) Use of folding modulators to improve heterologous protein production in *Escherichia coli*, *Microb Cell Fact 8*, 9.

[79] Yam, A. Y., Xia, Y., Lin, H.-T. J., Burlingame, A., Gerstein, M., and Frydman, J. (2008) Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies, *Nat. Struct. Mol. Biol. 15*, 1255-1262.

[80] Rüßmann, F., Stemp, M. J., Mönkemeyer, L., Etchells, S. A., Bracher, A., and Hartl, F. U. (2012) Folding of large multidomain proteins by partial encapsulation in the chaperonin TRiC/CCT, *Proc. Natl. Acad. Sci. U.S.A. 109*, 21208-21215.

[81] Peti, W., and Page, R. (2007) Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost, *Protein Expression Purif. 51*, 1-10.

[82] Luo, H., and Robb, F. T. (2011) A modulator domain controlling thermal stability in the Group II chaperonins of Archaea, *Arch. Biochem. Biophys. 512*, 111-118.

[83] Lee, K. K., Da Silva, N. A., and Kealey, J. T. (2009) Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*, *Anal. Biochem. 394*, 75-80.

[84] Himmler, A. (2017) Charakterisierung und Modifikation der Substratspezifität der Maus-β-Ketoacyl-Synthase.

[85] Clune, J., Mouret, J. B., and Lipson, H. (2013) The evolutionary origins of modularity, *Proc. R. Soc. Lond. B. Biol. Sci. 280*, 20122863.

[86] Khosla, C., Kapur, S., and Cane, D. E. (2009) Revisiting the modularity of modular polyketide synthases, *Curr. Opin. Chem. Biol. 13*, 135-143.

[87] Drexler, D. (2016) Charakterisierung katalytischer Eigenschaften der Ketoacylreduktase in Typ I Fettsäuresynthasen (FAS I).

[88] Beyer, D. (2014) Reaktionskontrolle in Fettsäuresynthasen aus Säugerzellen.

[89] Naggert, J., Witkowski, A., Wessa, B., and Smith, S. (1991) Expression in *Escherichia coli*, purification and characterization of two mammalian thioesterases involved in fatty acid synthesis, *Biochem. J. 273*, 787-790.

[90] Jacks, T. J., and Kircher, H. W. (1967) Fluorometric Assay for the Hydrolytic Activity of Lipase Using Fatty Acyl Esters of 4-Methylumbelliferone, *Anal. Biochem. 21*, 279-285.

[91] Zhang, W., Chakravarty, B., Zheng, F., Gu, Z., Wu, H., Mao, J., Wakil, S. J., and Quiocho, F. A. (2011) Crystal structure of FAS thioesterase domain with polyunsaturated fatty acyl adduct and inhibition by dihomo-γ-linolenic acid, *Proc. Natl. Acad. Sci. U.S.A. 108*, 15757-15762.

[92] Dutler, H., Coon, M. J., Kull, A., Vogel, H., Waldvogel, G., and Prelog, V. (1971) Fatty Acid Synthetase from Pig Liver. 1. Isolation of the Enzyme Complex and Characterization of the Component with Oxidoreductase Activity for Alicyclic Ketones, *Eur. J. Biochem. 22*, 203-212.

[93] Dutler, H., Kull, A., and Mislin, R. (1971) 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.

[94] Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies, *Anal. Biochem. 357*, 289-298.

[95] Yuzawa, S., Kapur, S., Cane, D. E., and Khosla, C. (2012) Role of a Conserved Arginine Residue in Linkers between the Ketosynthase and Acyltransferase Domains of Multimodular Polyketide Synthases, *Biochemistry 51*, 3708-3710.

[96] Alekseyev, V. Y., Liu, C. W., Cane, D. E., Puglisi, J. D., and Khosla, C. (2007) Solution structure and proposed domain-domain recognition interface of an acyl

carrier protein domain from a modular polyketide synthase, *Protein Sci. 16*, 2093-2107.

[97] Ploskoń, E., Arthur, C. J., Evans, S. E., Williams, C., Crosby, J., Simpson, T. J., and Crump, M. P. (2008) A Mammalian Type I Fatty Acid Synthase Acyl Carrier Protein Domain Does Not Sequester Acyl Chains, *J. Biol. Chem. 283*, 518-528.

[98] Lau, J., Cane, D. E., and Khosla, C. (2000) Substrate Specificity of the Loading Didomain of the Erythromycin Polyketide Synthase, *Biochemistry 39*, 10514-10520.

[99] Kubitschek, H. E. (1990) Cell volume increase in *Escherichia coli* after shifts to richer media, *J. Bacteriol. 172*, 94-101.

[100] Milo, R., and Phillips, R. (2015) *Cell Biology by the Numbers*, Garland Science.

[101] Molnos, J., Gardiner, R., Dale, G. E., and Lange, R. (2003) A continuous coupled enzyme assay for bacterial malonyl–CoA:acyl carrier protein transacylase (FabD), *Anal. Biochem. 319*, 171-176.

[102] Rangan, V. S., Serre, L., Witkowska, H. E., Bari, A., and Smith, S. (1997) Characterization of the malonyl-/acetyltransacylase domain of the multifunctional animal fatty acid synthase by expression in *Escherichia coli* and refolding *in vitro*, *Protein Eng. 10*, 561-566.

[103] Mikkelsen, J., Smith, S., Stern, A., and Knudsen, J. (1985) Stoichiometry of substrate binding to rat liver fatty acid synthetase, *Biochem. J. 230*, 435-440.

[104] Bonnett, S. A., Rath, C. M., Shareef, A.-R., Joels, J. R., Chemler, J. A., Håkansson, K., Reynolds, K., and Sherman, D. H. (2011) Acyl-CoA Subunit Selectivity in the Pikromycin Polyketide Synthase PikAIV: Steady-State Kinetics and Active-Site Occupancy Analysis by FTICR-MS, *Chem. Biol. 18*, 1075-1081.

[105] Chung, C. C., Ohwaki, K., Schneeweis, J. E., Stec, E., Varnerin, J. P., Goudreau, P. N., Chang, A., Cassaday, J., Yang, L., Yamakawa, T., Kornienko, O., Hodder, P., Inglese, J., Ferrer, M., Strulovici, B., Kusunoki, J., Tota, M. R., and Takagi, T. (2008) A Fluorescence-Based Thiol Quantification Assay for Ultra-High-Throughput Screening for Inhibitors of Coenzyme A Production, *Assay Drug Dev. Technol. 6*, 361-374.

[106] Sippel, T. O. (1981) New Fluorochromes for Thiols: Maleimide and Iodoacetamide Derivatives of a 3-Phenylcoumarin Fluorophore, *J. Histochem. Cytochem. 29*, 314-316.

[107] Gao, T., Yang, C., and Zheng, Y. G. (2012) Comparative studies of thiol-sensitive fluorogenic probes for HAT assays, *Anal. Bioanal. Chem. 405*, 1361-1371.

[108] Rangan, V. S., and Smith, S. (1996) Expression in *Escherichia coli* and Refolding of the Malonyl-/Acetyltransferase Domain of the Multifunctional Animal Fatty Acid Synthase, *J. Biol. Chem. 271*, 31749-31755.

[109] Lowry, B., Li, X., Robbins, T., Cane, D. E., and Khosla, C. (2016) A Turnstile Mechanism for the Controlled Growth of Biosynthetic Intermediates on Assembly Line Polyketide Synthases, *ACS Cent Sci 2*, 14-20.

[110] Vu Huu, K. (2016) Analyse der Substratspezifität von Transferasen.

[111] Richardson, R. D., and Smith, J. W. (2007) Novel antagonists of the thioesterase domain of human fatty acid synthase, *Mol. Cancer Ther. 6*, 2120-2126.

[112] Kim, Y., Tanner, K. G., and Denu, J. M. (2000) A Continuous, Nonradioactive Assay for Histone Acetyltransferases, *Anal. Biochem. 280*, 308-314.

[113] Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. coli*, *Science 291*, 1790-1792.

[114] Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V., and Stroh, J. G. (1999) Spontaneous α-*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.

[115] Rittner, A., Paithankar, K. S., Huu, K. V., and Grininger, M. (2018) 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.

[116] Misra, A., Sharma, S. K., Surolia, N., and Surolia, A. (2007) Self-Acylation Properties of Type II Fatty Acid Biosynthesis Acyl Carrier Protein, *Chem. Biol. 14*, 775-783.

[117] Seyama, Y., Otsuka, H., Kawaguchi, A., and Yamakawa, T. (1981) Fatty Acid Synthetase from the Harderian Gland of Guinea Pig: Biosynthesis of Methyl-Branched Fatty Acids, *J. Biochem. 90*, 789-798.

[118] Liew, C. W., Nilsson, M., Chen, M. W., Sun, H., Cornvik, T., Liang, Z. X., and Lescar, J. (2012) Crystal Structure of the Acyltransferase Domain of the Iterative Polyketide Synthase in Enediyne Biosynthesis, *J. Biol. Chem. 287*, 23203-23215.

[119] Tang, Y., Chen, A. Y., Kim, C.-Y., Cane, D. E., and Khosla, C. (2007) Structural and Mechanistic Analysis of Protein Interactions in Module 3 of the 6-Deoxyerythronolide B Synthase, *Chem. Biol. 14*, 931-943.

[120] Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase, *Proc. Natl. Acad. Sci. U.S.A. 103*, 11124-11129.

[121] Oefner, C., Schulz, H., D'Arcy, A., and Dale, G. E. (2006) Mapping the active site of *Escherichia coli* malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography, *Acta Crystallogr. D Biol. Crystallogr. 62*, 613-618.

[122] Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., and Larsen, S. (2001) Structures of β-Ketoacyl-Acyl Carrier Protein Synthase I Complexed with Fatty Acids Elucidate its Catalytic Machinery, *Structure 9*, 233-243.

[123] Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2256-2268.

[124] Afonine, P. V., Moriarty, N. W., Mustyakimov, M., Sobolev, O. V., Terwilliger, T. C., Turk, D., Urzhumtsev, A., and Adams, P. D. (2015) FEM: feature-enhanced map, *Acta Crystallogr. D Biol. Crystallogr. 71*, 646-666.

[125] Bunkoczi, G., Misquitta, S., Wu, X., Lee, W. H., Rojkova, A., Kochan, G., Kavanagh, K. L., Oppermann, U., and Smith, S. (2009) Structural Basis for Different Specificities of Acyltransferases Associated with the Human Cytosolic and Mitochondrial Fatty Acid Synthases, *Chem. Biol. 16*, 667-675.

[126] Nardini, M., and Dijkstra, B. W. (1999) α/β Hydrolase fold enzymes: the family keeps growing, *Curr. Opin. Struct. Biol. 9*, 732-737.

[127] Hol, W. G. (1985) Effects of the α-helix dipole upon the functioning and structure of proteins and peptides, *Adv. Biophys. 19*, 133-165.

[128] Witt, A. C., Lakshminarasimhan, M., Remington, B. C., Hasim, S., Pozharski, E., and Wilson, M. A. (2008) Cysteine pKa Depression by a Protonated Glutamic Acid in Human DJ-1, *Biochemistry 47*, 7430-7440.

[129] Flocco, M. M., and Mowbray, S. L. (1994) Planar stacking interactions of arginine and aromatic side-chains in proteins, *J. Mol. Biol. 235*, 709-717.

[130] O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities, *Chem. Biol. 6*, R91-R105.

[131] Yoshikuni, Y., Ferrin, T. E., and Keasling, J. D. (2006) Designed divergent evolution of enzyme function, *Nature 440*, 1078-1082.

[132] Mohamed, M. F., and Hollfelder, F. (2013) Efficient, crosswise catalytic promiscuity among enzymes that catalyze phosphoryl transfer, *Biochim. Biophys. Acta 1834*, 417-424.

[133] Herzberg, D. (2017) Analysis of transferase-activity of synthases.

[134] Yuan, Z. Y., and Hammes, G. G. (1986) Fluorescence Studies of Chicken Liver Fatty Acid Synthase: Segmental Flexibility and Distance Measurements., *J. Biol. Chem. 261*, 13643-13651.

[135] Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An Enhanced System for Unnatural Amino Acid Mutagenesis in *E. coli*, *J. Mol. Biol. 395*, 361-374.

[136] Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013) A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in *Escherichia coli*, *Biochemistry 52*, 1828-1837.

[137] Heil, C. (2014) Incorporation of Non-Natural Amino Acids in the Mammalian Fatty Acid Synthase.

[138] Heil, C. S., Rittner, A., Goebel, B., Beyer, D., and Grininger, M. (2018) Site-Specific Labelling of Multidomain Proteins by Amber Codon Suppression, *bioRχiv (*https://doi.org/10.1101/282525*)*.

[139] Goebel, B. (2017) Site Specific Labelling of Multidomain Enzymes via Amber Codon Suppression.

[140] Castoe, T. A., Stephens, T., Noonan, B. P., and Calestani, C. (2007) A novel group of type I polyketide synthases (PKS) in animals and the complex phylogenomics of PKSs, *Gene 392*, 47-58.

[141] Chang, S. I., and Hammes, G. G. (1990) Structure and Mechanism of Action of a Multifunctional Enzyme: Fatty Acid Synthase, *Acc. Chem. Res. 23*, 363-369.

[142] Pieper, R., Ebert-Khosla, S., Cane, D., and Khosla, C. (1996) Erythromycin Biosynthesis: Kinetic Studies on a Fully Active Modular Polyketide Synthase Using Natural and Unnatural Substrates, *Biochemistry 35*, 2054-2060.

[143] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic Local Alignment Search Tool, *J. Mol. Biol. 215*, 403-410.

[144] Worthington, A. S., Rivera, J., Heriberto, Alexander, M. D., and Burkart, M. D. (2006) Mechanism-Based Protein Cross-Linking Probes To Investigate Carrier Protein-Mediated Biosynthesis, *ACS Chem. Biol. 1*, 687-691.

[145] Jayakumar, A., Tai, M. H., Huang, W. Y., al-Feel, W., Hsu, M., Abu-Elheiga, L., Chirala, S. S., and Wakil, S. J. (1995) Human fatty acid synthase: properties and molecular cloning, *Proc. Natl. Acad. Sci. U.S.A. 92*, 8695-8699.

[146] Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions, *Science 325*, 834-840.

[147] Zhou, H., Di Palma, S., Preisinger, C., Peng, M., Polat, A., Heck, A., and Mohammed, S. (2103) Toward a Comprehensive Characterization of a Human Cancer Cell Phosphoproteome, *J. Proteome Res. 12*, 260-271.

[148] Dimroth, P., Walter, H., and Lynen, F. (1970) Biosynthesis of 6-methylsalicylic acid, *Eur. J. Biochem. 13*, 98-110.

[149] Fersht, A. R. (1999) *Structure and Mechanism in Protein Science*, W. H. Freeman and Company.

[150] Takamura, Y., and Nomura, G. (1988) Changes in the Intracellular Concentration of Acetyl-CoA and Malonyl-CoA in Relation to the Carbon and Energy Metabolism of *Escherichia coli* K12, *J. Gen. Microbiol. 134*, 2249-2253.

[151] Buckner, J. S., and Kolattukudy, P. E. (1975) Lipid Biosynthesis in Sebaceous Glands: Regulation of the Synthesis of *n*- and Branched Fatty Acids by Malonyl-Coenzyme A Decarboxylase, *Biochemistry 14*, 1768-1773.

[152] Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*, *Nat Meth 5*, 593-599.

[153] Soulié, J. M., Sheplock, G. J., Tian, W. X., and Hsu, R. Y. (1984) Transient kinetic studies of fatty acid synthetase. A kinetic self-editing mechanism for the loading of acetyl and malonyl residues and the role of coenzyme A, *J. Biol. Chem. 259*, 134-140.

[154] Hagen, A., Poust, S., Rond, T. d., Fortman, J. L., Katz, L., Petzold, C. J., and Keasling, J. D. (2016) Engineering a Polyketide Synthase for *In Vitro* Production of Adipic Acid, *ACS Synth. Biol. 5*, 21-27.

[155] Musser, M. T. (2005) *Adipic Acid*, Weinheim: Wiley-VCH, *Ullmann's Encyclopedia of Industrial Chemistry*.

[156] Bravo-Rodriguez, K., Klopries, S., Koopmans, K. R. M., Sundermann, U., Yahiaoui, S., Arens, J., Kushnir, S., Schulz, F., and Sanchez-Garcia, E. (2015) Substrate Flexibility of a Mutated Acyltransferase Domain and Implications for Polyketide Biosynthesis, *Chem. Biol. 22*, 1425-1430.

[157] Koryakina, I., Kasey, C., McArthur, J. B., Lowell, A. N., Chemler, J. A., Li, S., Hansen, D. A., Sherman, D. H., and Williams, G. J. (2016) Inversion of Extender Unit Selectivity in the Erythromycin Polyketide Synthase by Acyltransferase Domain Engineering, *ACS Chem. Biol. 12*, 114-123.

[158] Marcella, A. M., and Barb, A. W. (2017) The R117A variant of the *Escherichia coli* transacylase FabD synthesizes novel acyl-(acyl carrier proteins), *Appl. Microbiol. Biotechnol. 101*, 8431-8441.

[159] Hammes-Schiffer, S., and Benkovic, S. J. (2006) Relating Protein Motion to Catalysis, *Annu. Rev. Biochem. 75*, 519-541.

[160] Sanadi, D. R. (1969) KDH from pig heart, *Methods Enzymol. 13*, 52-55.

[161] Mayer, S., Rüdiger, S., Ang, H. C., Joerger, A. C., and Fersht, A. R. (2007) Correlation of Levels of Folded Recombinant p53 in *Escherichia coli* with Thermodynamic Stability *in Vitro*, *J. Mol. Biol. 372*, 268-276.

[162] Yuzawa, S., Deng, K., Wang, G., Baidoo, E. E. K., Northen, T. R., Adams, P. D., Katz, L., and Keasling, J. D. (2017) 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.

[163] Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., and Weissman, K. J. (2003) The Structure of Docking Domains in Modular Polyketide Synthases, *Chem. Biol. 10*, 723-731.

[164] Zheng, J., Fage, C. D., Demeler, B., Hoffman, D. W., and Keatinge-Clay, A. T. (2013) The Missing Linker: A Dimerization Motif Located within Polyketide Synthase Modules, *ACS Chem. Biol. 8*, 1263-1270.

[165] Edwards, A. L., Matsui, T., Weiss, T. M., and Khosla, C. (2014) Architectures of Whole-Module and Bimodular Proteins from the 6-Deoxyerythronolide B Synthase, *J. Mol. Biol. 426*, 2229-2245.

[166] Dutta, S., Whicher, J. R., Hansen, D. A., Hale, W. A., Chemler, J. A., Congdon, G. R., Narayan, A. R., Håkansson, K., Sherman, D. H., Smith, J. L., and Skiniotis, G. (2014) Structure of a modular polyketide synthase, *Nature 510*, 512-517.

[167] Rittner, A., and Grininger, M. (2014) Modular Polyketide Synthases (PKSs): A New Model Fits All?, *ChemBioChem 15*, 2489-2493.

[168] Witkowski, A., Joshi, A. K., and Smith, S. (2004) Characterization of the β-Carbon Processing Reactions of the Mammalian Cytosolic Fatty Acid Synthase: Role of the Central Core, *Biochemistry 43*, 10458-10466.

[169] Witkowski, A., Joshi, A. K., Lindqvist, Y., and Smith, S. (1999) Conversion of a β-Ketoacyl Synthase to a Malonyl Decarboxylase by Replacement of the Active-Site Cysteine with Glutamine, *Biochemistry 38*, 11643-11650.

[170] Busch, B., Ueberschaar, N., Behnken, S., Sugimoto, Y., Werneburg, M., Traitcheva, N., He, J., and Hertweck, C. (2013) Multifactorial Control of Iteration Events in a Modular Polyketide Assembly Line, *Angew. Chem. Int. Ed. 52*, 5285-5289.

[171] Hans, M., Hornung, A., Dziarnowski, A., Cane, D. E., and Khosla, C. (2003) Mechanistic Analysis of Acyl Transferase Domain Exchange in Polyketide Synthase Modules, *J. Am. Chem. Soc. 125*, 5366-5374.

[172] Whicher, J. R., Dutta, S., Hansen, D. A., Hale, W. A., Chemler, J. A., Dosey, A. M., Narayan, A. R. H., Håkansson, K., Sherman, D. H., Smith, J. L., and Skiniotis, G. (2014) Structural rearrangements of a polyketide synthase module during its catalytic cycle, *Nature 510*, 560-564.

[173] Khersonsky, O., and Tawfik, D. S. (2010) Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective, *Annu. Rev. Biochem. 79*, 471-505.

[174] Gajewski, J., Buelens, F., Serdjukow, S., Janßen, M., Cortina, N., Grubmüller, H., and Grininger, M. (2017) Engineering fatty acid synthases for directed polyketide production, *Nat. Chem. Biol. 13*, 363–365.

[175] Joshi, A. K., Witkowski, A., Berman, H. A., Zhang, L., and Smith, S. (2005) Effect of Modification of the Length and Flexibility of the Acyl Carrier Protein–Thioesterase Interdomain Linker on Functionality of the Animal Fatty Acid Synthase, *Biochemistry 44*, 4100-4107.

[176] Jin, C., Yao, M., Liu, H., Lee, C.-f. F., and Ji, J. (2011) Progress in the production and application of n-butanol as a biofuel, *Renew. Sust. Energy. Rev. 15*, 4080-4106.

[177] Gu, L., Wang, B., Kulkarni, A., Gehret, J. J., Lloyd, K. R., Gerwick, L., Gerwick, W. H., Wipf, P., Håkansson, K., Smith, J. L., and Sherman, D. H. (2009) Polyketide Decarboxylative Chain Termination Preceded by *O*-Sulfonation in Curacin A Biosynthesis, *J. Am. Chem. Soc. 131*, 16033-16035.

[178] Barajas, J. F., Phelan, R. M., Schaub, A. J., Kliewer, J. T., Kelly, P. J., Jackson, D. R., Luo, R., Keasling, J. D., and Tsai, S.-C. (2015) Comprehensive Structural and Biochemical Analysis of the Terminal Myxalamid Reductase Domain for the Engineered Production of Primary Alcohols, *Chem. Biol. 22*, 1018-1029.

[179] Wilkinson, A. J., Fersht, A. R., Blow, D. M., and Winter, G. (1983) Site-Directed Mutagenesis as a Probe of Enzyme Structure and Catalysis: Tyrosyl-tRNA Synthetase Cysteine-35 to Glycine-35 Mutation, *Biochemistry 22*, 3581-3586.

[180] Xie, D., Shao, Z., Achkar, J., Zha, W., Frost, J. W., and Zhao, H. (2006) Microbial Synthesis of Triacetic Acid Lactone, *Biotechnol. Bioeng. 93*, 727-736.

[181] Kabsch, W. (2010) *XDS*, *Acta Crystallogr. D Biol. Crystallogr. 66*, 125-132

[182] Evans, P. (2005) Scaling and assessment of data quality, *Acta Crystallogr. D Biol. Crystallogr. 62*, 72-82.

[183] Matthews, B. W. (1968) Solvent Content of Protein Crystals, *J. Mol. Biol. 33*, 491-497.

[184] Vagin, A., and Teplyakov, A. (1997) *MOLREP*: an Automated Program for Molecular Replacement, *J. Appl. Crystallogr. 30*, 1022-1025.

[185] Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) *REFMAC5* for the refinement of macromolecular crystal structures, *Acta Crystallogr. D Biol. Crystallogr. 67*, 355-367.

[186] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of *Coot*, *Acta Crystallogr. D Biol. Crystallogr. 66*, 486-501.

[187] Lebedev, A. A., Young, P., Isupov, M. N., Moroz, O. V., Vagin, A. A., and Murshudov, G. N. (2012) *JLigand*: a graphical tool for the *CCP*4 template-restraint library, *Acta Crystallogr. D Biol. Crystallogr. 68*, 431-440.

[188] Xi, G.-L., and Liu, Z.-Q. (2014) Coumarin moiety can enhance abilities of chalcones to inhibit DNA oxidation and to scavenge radicals, *Tetrahedron 70*, 8397-8404.

[189] Davidson, D., and Bernhard, S. A. (1948) The Structure of Meldrum's Supposed β-Lactonic Acid, *J. Am. Chem. Soc. 70*, 3426-3428.

[190] Ryu, Y., and Scott, A. I. (2003) Self-condensation of activated malonic acid half esters: a model for the decarboxylative Claisen condensation in polyketide biosynthesis, *Tetrahedron Lett. 44*, 7499-7502.

[191] Wieland, T., von Dungen, A., and Birr, C. (1971) Synthesis of an antitoxic antamanide variant using *p*-azido-phenylalanine in position 6, *Justus Liebigs Ann. Chem. 752*, 109-114.

[192] Lieber, E., Levering, D. R., and Patterson, L. (1951) Infrared Absorption Spectra of Compounds of High Nitrogen Content, *Anal. Chem. 23*, 1594-1604.

[193] Jain, S., and Reiser, O. (2008) Immobilization of Cobalt(II) Schiff Base Complexes on Polystyrene Resin and a Study of Their Catalytic Activity for the Aerobic Oxidation of Alcohols, *ChemSusChem 1*, 534-541.

[194] Borzilleri, R., Zhang, Y., Miller, M. M., and Seigal, B. A. (2014) Preparation of macrocyclic compounds for inhibition of inhibitors of apoptosis, In *PCT Int. Appl.*

[195] Seitchik, J. L., Peeler, J. C., Taylor, M. T., Blackman, M. L., Rhoads, T. W., Cooley, R. B., Refakis, C., Fox, J. M., and Mehl, R. A. (2012) Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific *in Vivo* Bioorthogonal Ligation with *trans*-Cyclooctenes, *J. Am. Chem. Soc. 134*, 2898-2901.

[196] Fields, S. C., Parker, M. H., and Erickson, W. R. (1994) A simple route to unsymmetrically substituted 1, 2, 4, 5-tetrazines, *J. Org. Chem. 59*, 8284-8287.

[197] Nguyen, D. P., Lusic, H., Neumann, H., Kapadnis, P. B., Deiters, A., and Chin, J. W. (2009) Genetic Encoding and Labeling of Aliphatic Azides and Alkynes in Recombinant Proteins *via* a Pyrrolysyl-tRNA Synthetase/tRNA_{CUA} Pair and Click Chemistry, *J. Am. Chem. Soc. 131*, 8720-8721.

[198] Khan, F. A., and Choudhury, S. (2010) An Efficient Synthesis of Substituted *meta*-Halophenols and Their Methyl Ethers: Insight into the Reaction Mechanism, *Eur. J. Org. Chem. 2010*, 2954-2970.

[199] Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J. W. (2012) Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction, *4*, 298-304.

[200] Plass, T., Milles, S., Koehler, C., Szymański, J., Mueller, R., Wießler, M., Schultz, C., and Lemke, E. A. (2012) Amino Acids for Diels-Alder Reactions in Living Cells, *Angew. Chem. Int. Ed. 51*, 4166-4170.

[201] Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol. 7*, 1-6.

[202] Kibbe, W. A. (2007) OligoCalc: an online oligonucleotide properties calculator, *Nucleic Acids Res. 35*, W43-W46.

5.6 Statement of personal contributions

All projects and strategies for the practical accomplishment within this thesis were conceived by myself based on the initial idea of Martin Grininger to utilize animal FASs as model systems to study modular PKSs.

Experiments presented within this thesis were mostly performed by myself or by students under my supervison, which was transparently marked at respective sites within the text and master theses were cited where appropriate. The only exception was the processing of X-ray crystallographic data with subsequent model building, which was performed by Karthik Paithankar. I participated in the process of structure elucidation until the exposure of crystals with X-rays at the synchrotron and from the refinement of the obtained model. Interpretation of structural data was again performed by myself with the assistance and guidance of Martin Grininger.

Writing of the thesis was exclusively conducted by myself, except sections that were similarly published in *"Characterization of the Polyspecific Transferase of Murine Type I Fatty Acid Synthase (FAS) and Implications for Polyketide Synthase (PKS) Engineering"*, which was marked at respective sites in the text. Karthik Paithankar and Martin Grininger participated significantly with editing and phrasing to the final version.

5.7 Non-scientific supplementary

5.7.1 Danksagung

Zuletzt, möchte ich mich besonders herzlich bei meiner Familie und meinen Freunden für die grandiose persönliche Unterstützung bedanken, insbesondere bei meinen Eltern für ein nicht endendes Interesse an Fettsäuresynthasen und verwandten Enzymen, für alle Ratschläge und für das Korrekturlesen der These. Ohne Euch wäre mir die Zeit mit Sicherheit schwerer gefallen. Liebe Lan-Na, was soll ich sagen, Du bist immer für mich da und erträgst ausdauernd meine wissenschaftliche, nerdige Ader. Ich bin Dir grenzenlos dankbar, dass Du die zugegebenermaßen nicht ideale Situation so lange toleriert hast und immer an meiner Seite warst.

5.7.2 Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich die vorgelegte Dissertation:

Engineering Animal Fatty Acid Synthase Towards Polyketide Synthases

selbständig angefertigt und mich nicht anderer Hilfsmittel als der in ihr angegebenen 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 zu haben, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Ort, Datum (Alexander Rittner)

5.7.3   Curriculum Vitae

Personal data:	

Name: **Alexander Rittner**

Address: Danneckerstr. 11, 60594 Frankfurt

E-mail: rittner@chemie.uni-frankfurt.de

Birth place/date: Bonn/25.02.1985

Current position:	

10/2012 – :   **Ph.D. student at the Buchmann Institute for Molecular Life Sciences of the Goethe University Frankfurt**
                    Supervisor: Prof. Dr. Martin Grininger, Institute of Organic Chemistry and Chemical Biology
                    Thesis title: "Engineering Animal Fatty Acid Synthase Towards Polyketide Synthases"

Education:	

10/2009-02/2012: **M.Sc. in Biochemistry at the Ludwig-Maximilians-University Munich**
                    Thesis title: "Characterisation of selected NTR1 variants for their Signalling Properties"
                    External Supervisor: Prof. Dr. Andreas Plückthun, Department of Biochemistry, University of Zurich

09/2007-10/2009: **B.Sc. in "Chemistry and Biochemistry" at the Ludwig-Maximilians-University Munich**
                    Thesis title: "Design eines funktionalisierten Salicylaldehyds zur Synthese 3,3'-verbrückter Bissalicylaldehyde"
                    Supervisor: Prof. Dr. Thomas Carell, Institute of Organic Chemitry

10/2005-09/2007: **Vordiploma's degree in "Biomedical Chemistry" at the Johannes Gutenberg-University Mainz**

School:	

08/1995-03/2004: **Graduation from Gymnasiums Nieder-Olm** (Abitur)

09/2001-12/2001: Stay abroad at the Clifton College in Bristol, England

Employments:

03/2010-02/2011: **Undergraduate research assistant (HiWi)** in the laboratory of Dr. Grininger at the Department of Membrane Biochemistry headed by Prof. Dr. Dieter Oesterhelt at the Max Planck Institute in Martinsried

05/2005-12/2011: **Biostatistic analysis of paternity cases** at the Institute of Legal Medicine in Kaiserslautern

05/2004-01/2005: **Community service** in the Pneumology department of the Universitätsklinikum Mainz

Scholarships:

06/2011-11/2011: **ERASMUS-Praktikum scholarship from the DAAD** (Deutscher Akademischer Austausch Dienst) for the master's thesis

Publications:

Schlinkmann, K. M., Hillenbrand, M., Rittner, A., Künz, M., Strohner, R. & Plückthun, A. (2012). Maximizing Detergent Stability and Functional Expression of a GPCR by Exhaustive Recombination and Evolution. *J. Mol. Biol. 422*, 414-428.

Rittner, A., and Grininger, M. (2014) Modular Polyketide Synthases (PKSs): A New Model Fits All?, *ChemBioChem 15*, 2489-2493.

Rittner, A., Paithankar, K. S., Huu, K. V., and Grininger, M. (2018) 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.

Heil, C. S., Rittner, A., Goebel, B., Beyer, D., and Grininger, M. (2018) Site-Specific Labelling of Multidomain Proteins by Amber Codon Suppression, *bioRχiv* (https://doi.org/10.1101/282525).