Enlarging the Product Spectrum of Fatty Acid 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
Jan Gajewski
aus Kirchheimbolanden

Frankfurt 2017
(D30)

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

Dekan: Prof. Dr. Michael Karras

1. Gutachter: Prof. Dr. Martin Grininger
2. Gutachter: Prof. Dr. Eckhard Boles

Datum der Disputation: 15.05.2017

Table of content

1 Deutsche Zusammenfassung

1.1 Hintergrund

In der Natur kommen Fettsäuren (FA) in allen eukaryotischen und bakteriellen Organismen vor. Die für die Herstellung verantwortlichen Enzyme werden als Fettsäuresynthasen (FAS) bezeichnet[1]. In den letzten Jahren erhielt die Erforschung von FAS eine erneute Aufmerksamkeitssteigerung, da ihr Potential für biotechnologische Anwendungen erkannt wurde[2]: Die durch FAS hergestellten FA können eine Schlüsselrolle in größeren Stoffwechselwegen spielen und weiter zu Methyl-/Ethylestern, Aldehyden, Fettalkoholen oder Alkanen umgesetzt werden[3]. Diese wiederum finden Anwendung als Lebensmittelzusätze, Emulgatoren und auch als Bestandteil von Biokraftstoffen.

Der Aufbau von FA folgt bei allen FAS dem gleichen grundlegenden Prinzip: Aus kleinen Vorläufermolekülen, Acetyl-CoA und Malonyl-CoA, werden lange Fettsäuren, typischerweise Palmitoyl-CoA (C_{16}-CoA) und Stearoyl-CoA (C_{18}-CoA), aufgebaut. Im ersten Schritt werden dabei Acetyl und Malonyl durch Transferasen auf das Enzym geladen, wo sie in der Ketoacylsynthase (KS) in einer Claisen-Kondensation miteinander verknüpft werden. Das daraus entstehende β-Ketoacyl-Zwischenprodukt bleibt kovalent an das zentrale Acyl-Carrier-Protein (ACP) gebunden, während die β-Keto-Gruppe durch eine Reihe von Reaktionsschritten modifiziert wird: Zunächst wird sie in der Ketoacylreduktase (KR) mithilfe von NADPH zum Hydroxyl reduziert. Im darauffolgenden Schritt wird in der Dehydratase (DH) Wasser abgespaltet. Nach einer letzten Reduktion in der Enoylreduktase (ER) liegt dann ein vollreduziertes Acyl-ACP vor. Dieses Acyl-ACP kann wieder in den Synthesekreislauf eingebracht werden und durch die Kondensation mit Malonyl in jedem Zyklus um weitere zwei Kohlenstoffe verlängert werden (genau wie Acetyl zuvor). Nach mehreren Durchläufen wird der Kreislauf abgebrochen, indem die vollreduzierte Acyl-Gruppe nicht wieder in die KS übertragen wird, sondern vom Enzym abgespaltet wird. Abhängig vom Typ der FAS, wird diese Abspaltung entweder durch Thioesterasen (zur freien FA) oder durch eine Malonyl-/Palmitoyl-Transferase (MPT; zum Acyl-CoA-Ester) katalysiert.

Je nachdem, ob die genannten katalytischen Domänen von FAS in einem Multienzymkomplex zusammengefasst sind oder sich auf separaten Proteinen befinden, werden FAS als Typ I bzw. Typ II klassifiziert. Typ I ist dabei typisch für

Eukaryoten, allerdings gibt es auch einige Bakterien (wie z.B. *Nocardia*, *Corynebacteria* und *Mycobacteria*), die solche Systeme besitzen[4]. Typ II Systeme findet man wiederum typischerweise in Prokaryoten, wobei die vorliegende Arbeit auf diese Klasse der FAS nur vereinzelt wieder eingeht.

Der biotechnologische Einsatz von FAS ist zur Zeit durch einige ihrer typischen Eigenschaften eingeschränkt, beispielsweise durch ihr natürliches Produktspektrum, das wie bereits dargelegt, typischerweise im Bereich der langen FA liegt. Für bestimmte Anwendungen sind jedoch andere Produkte gefragt: Als Beispiel kann hier die Anwendung in Benzin-ähnlichen Biokraftstoffen aufgeführt werden, für die kurzkettigere Produkte benötigt werden[5-6]. Für die biotechnologische Herstellung entsprechender FA ergibt sich daher ein Bedarf für ein Engineering von FAS.

Bisherige Studien zur Herstellung von kurzen FA stützen sich auf den Austausch ganzer Domänen und bedienen sich dabei bereits bekannter Substratspezifität von homologen Domänen. Auch erste Studien, die durch Austausch von einzelnen Aminosäuren die Herstellung von kurzen FA begünstigten, gibt es bereits: Diese beschränken sich jedoch auf Typ II FAS[7] oder die Produktion von kurzen FA als Nebenprodukte[8-9]. Durch die kompartimentierte Synthese in Typ I FAS und die daraus resultierenden vorteilhaften Syntheseleistung, besteht jedoch weiterhin Bedarf an der Aufklärung und der gezielten Veränderung von kettenlängenbestimmenden molekularen Vorgängen zur Herstellung von kurzen FA in Typ I FAS[5].

Die Verwendung der Erkenntnisse hinsichtlich der künstlichen Reaktionskontrolle in FAS sind dabei nicht auf diese Enzymklasse beschränkt. FAS werden aufgrund ihrer hohen Syntheseleistung und ihrer Größe zu den Megasynthasen gezählt, zu denen auch die Polyketidsynthasen und Nicht-Ribosomale-Peptid-Synthasen gehören. Mögliche Einblicke und Vorschläge zur Manipulation in FAS könnten als Beispiel für andere Megasynthasen dienen. Hervorzuheben ist an dieser Stelle nochmals die direkte Ähnlichkeit von FAS und Polyketidsynthasen, deren Synthese auf den gleichen grundlegenden Prinzipien basiert. Die Übertragbarkeit ist insbesondere aufgrund der großen Ähnlichkeit innerhalb der einzelnen Domänen gegeben.

1.2 Rationales Engineering von FAS

Der von uns gewählte Ansatz zum Engineering von FAS ist rationales Design. Ausgehend von Proteinstrukturdaten wurden Schlüsselprozesse mit Hilfe von gezielt gesetzten Mutationen gesteuert und so das Produktspektrum von FAS verändert. Der Fokus dieser Manipulationen lag auf der Kettenlängenkontrolle in FAS, genauer auf der gezielten Manipulation von FAS zur Herstellung von kurzkettigen Produkten anstelle der natürlich auftretenden langkettigen Produkte. Das von uns durchgeführte rationale Engineering wurde durch Arbeiten der Forschungsgruppe von Prof. Grubmüller (MPI für Biophysikalische Chemie, Göttingen) unterstützt. Deren Gegenstand waren die Molekular-Dynamik-Simulationen von Bindungsereignissen an einzelnen Domänen sowie die Entwicklung eines quantitativen Computer-Modells zur Simulation des Reaktions-Netzwerks von FAS.

Innerhalb der Gruppe der Typ I FAS, wurde die FAS aus *Saccharomyces cerevisiae* als Ausgangspunkt verwendet. Die ausgiebigen, verfügbaren Strukturdaten[10-13] und die weit zurückreichenden biochemischen Charakterisierungen[1] machen diese FAS zu einem geeigneten Kandidaten um als Grundlage für rationales Engineering zu dienen. Für die Aufklärung von Struktur-Funktionsbeziehungen war die Charakterisierung von FAS *in vitro* ein zentraler Bestandteil unserer Herangehensweise. Da der Zugang zu ausreichenden Mengen an *S. cerevisiae* FAS Protein in hoher Qualität jedoch beschränkt war, wurde die Typ I FAS aus *Corynebakterium ammoniagenes* als homologes Modellenzym gewählt. Diese Typ I FAS ist durch heterologe Expression in *E. coli* mit anschließender Affinitätschromatographie und Größenausschlusschromatographie in hohen Mengen und guter Qualität zugänglich[14]. Aufgrund ihrer hohen Homologie mit *S. cerevisiae* FAS[15], konnte letztere als gute Orientierung für das rationale Engineering dienen.

In einer ersten Studie stand die Modifikation der Kettenlängenkontrolle durch gezielte Veränderung von Substrat-Bindungsaffinitäten in den wichtigsten kettenlängenbestimmenden Domänen im Mittelpunkt. Zunächst wurde dafür die KS Domäne näher betrachtet, da sie in der Fettsäuresynthese die zentrale Reaktion der Kondensation katalysiert und somit direkt für die Verlängerung von Zwischenprodukten verantwortlich ist. In den *in vitro* Experimenten mit *C. ammoniagenes* FAS wurden insgesamt drei Mutationsstellen im KS Bindungskanal getestet. Die erste, G2599S, wurde in Inhibitor-Resistenz-Studien gefunden und dann mit einer erhöhten Produktion von kurzen FA in Verbindung gebracht[8-9]. Die zweite Mutation wurde durch Vergleich

von Strukturdaten identifiziert, die ein zentrales Methionin im KS Bindungskanal zeigen, das bei der Elongation von längeren Zwischenprodukten als zu überwindende Hürde angesehen wurde[10, 12, 16]. Ein Austausch zu dem noch größeren Tryptophan, M2600W, sollte diesen Effekt weiter verstärken. Die Kombination der beiden Mutationen im Konstrukt $FAS^{G2599S-M2600W}$ zeigte eine gesteigerte Synthese von kurzen Produkten. Der Anteil von C_8-CoA im Produktspektrum wurde so auf 18% erhöht[17-18].
In späteren *in vitro* Studien konnte die Produktspezifität von FAS abermals gesteigert werden: Die dritte Mutation in der KS Domäne konnte durch Sequenzvergleiche mit entfernt verwandten FAS, die in Naturstoffsynthesewegen mit der Herstellung von kurzen Fettsäuren in Verbindung gebracht wurden[19-21], identifiziert werden. Der Einbau in *C. ammoniagenes* FAS (dort L2628Y) führte dazu, dass C_8-CoA insgesamt 74% der detektierten CoA-Ester ausmachte.
Neben der KS Domäne wurden auch die Transferasen modifiziert. Auf Grundlage früherer Studien zur Substratspezifität von Transferasen konnte eine Mutation von Arginin zu Lysin identifiziert werden, die zu einem verringerten Laden von Malonyl führt[22-23]. An der äquivalenten Position in der MPT der *C. ammoniagenes* FAS (dort R1408K) führte diese Mutation in Verbindung mit den Mutationen G2599S und M2600W in der KS Domäne zu einer abermaligen Steigerung des C_8-CoA Anteils auf 47%, bei jedoch etwa gleichbleibender absoluter Ausbeute.
Auch in der Acetyltransferase (AT) wurde eine Mutation, I151A, getestet, die neben einer erhöhten Acetyltransferaseaktivität auch den Transfer von kurzen Acylketten zur Folge haben sollte[22-23]. Im Konstrukt $FAS^{G2599S-M2600W-I151A}$ wurde zwar die absolute Ausbeute gesteigert, in der Spezifität blieb diese hinter dem $FAS^{G2599S-M2600W-R1408K}$-Konstrukt zurück. Auch die Kombination von KS, MPT und AT Mutationen ($FAS^{G2599S-M2600W-I151A-R1408K}$) brachte keine weitere Verbesserung der Spezifität in der Produktion von kurzen Acyl-CoA-Estern.
Die molekulardynamischen Simulationen konnten zudem einen wichtigen Beitrag zum Verständnis der Wirkmechanismen der Mutationen liefern: Für die KS Domäne wurde für die mutierte Variante eine energetisch günstigere Bindung von C_8 berechnet, was jedoch unserer Theorie einer ungünstigeren Bindung durch die Einführung von großen Resten widersprach. Auf der anderen Seite, wurden unsere Annahmen zu den Reaktionsmechanismen für die Transferasen weitestgehend bestätigt: In der mutierten MPT wurde Malonyl weniger stark gebunden und für die mutierte AT Domäne konnte die mögliche Bindung von Acylketten wie C_8 gezeigt werden. Auch das kinetische

Computer-Modell ergab nach Berücksichtigung der Mutationen ähnliche Verteilungen wie in den *in vitro* Experimenten, wobei dies für die KS Mutationen erst mit einer zweiten Version erreicht werden konnte, die weniger strikte Vorgaben zu den veränderten kinetischen Vorgängen enthielt[17].

Neben der Modifikation von Substratbindungsaffinitäten, wurde auch die Interaktion von Domänen zur gezielten Veränderung von Synthesewegen auf FAS genutzt. Wieder diente die Kettenlängenkontrolle als beispielhaftes Ziel zur Quantifizierung unseres Engineerings. Durch die gezielte Veränderung der Oberfläche von katalytischen Domänen sollten dabei die Interaktionen von ACP und eben diesen beeinflusst werden und so die Reaktionsflüsse gesteuert werden.

Besonders erfolgreich war dieser Ansatz als die Interaktion von ACP und der KS Domäne modifiziert wurde. Als Ausgangssystem wurde *C. ammoniagenes* FAS verwendet, die bereits im KS Bindungskanal die Mutationen G2599S und M2600W besaß, da so Effekte leichter quantifizierbar waren. Auf der Oberfläche der KS Domäne wurden dann einzelne Reste ausgetauscht um die Wahrscheinlichkeit zu reduzieren, dass Acyl-ACP mit der KS interagiert. Eine Übertragung von Acyl-Resten in die KS für eine Elongation würde somit auch unwahrscheinlicher.

Für einige Konstrukte, z.B. N2557E, N2621D und D2622A (jeweils mit G2599S M2600W), wurde auch hier eine hohe spezifische Synthese von kurzen Acyl-CoA-Estern beobachtet, wobei der Produktanteil an C_8-CoA bis zu 77% ausmachte. Diese Studie zur gerichteten Synthese durch Engineering von Domänen-Domänen-Interaktionen ist als erste erfolgreiche Machbarkeitsstudie ihrer Art zu verstehen.

Einige kleinere Projekte dieser Promotion gingen über das reine Engineering der Kettenlängen in FAS hinaus und werden im Folgendem vorgestellt:

Wie bereits erwähnt, basieren die Synthesekonzepte von FAS und PKS auf den gleichen grundlegenden Prinzipien. FAS eignet sich deshalb auch als Modellprotein für die Erforschung von PKS Systemen, die oftmals schwer zugänglich sind. In einer Pilotstudie wurde ein gekoppelter Syntheseweg aus zwei in Reihe agierenden modifizierten FAS zur Herstellung eines einfachen Polyketids entwickelt. Als erste FAS wurden dabei einige der oben beschriebenen *C. ammoniagenes* FAS Konstrukte verwendet, welche C_8-CoA herstellen. Eine zweite modifizierte FAS nimmt C_8-CoA dann auf und verlängert es zweifach jedoch ohne Reduktionsschritt. Das so entstehende Triketid spaltet sich spontan zum Endprodukt 6-Heptyl-4-hydroxypran-2-on ab.

Zur weiteren Steigerung der Ausbeute wurden die entsprechenden *in vitro* Eintopfsynthesen auf zwei Arten optimiert: Zunächst wurden in der gekoppelten Synthese verschiedene Konstrukte der C_8-CoA-produzierenden FAS getestet. Nach Identifikation zweier vielversprechender FAS-Paare, wurden diese wiederum mit variierenden Konzentrationen von Acetyl-CoA und NADPH gemessen, wobei die Malonyl-CoA-Konzentration konstant gehalten wurde. Bezogen auf das jeweils limitierende Substrat, konnten so Ausbeuten von bis zu 35% erzielt werden. Diese guten Ergebnisse können als Maß für ein erfolgreiches Engineering gesehen werden, bei dem mit insgesamt nur fünf Mutationen verteilt auf zwei FAS ein Syntheseweg für die Herstellung eines einfachen Polyketids, 6-Heptyl-4-hydroxypran-2-on, geschaffen wurde[17].

Während alle bisherigen Studien dieser Promotionsarbeit Reaktionskontrolle in Rahmen von *in vitro* Experimenten untersuchten, wurde in einem Projekt weitergehend untersucht, ob die guten Ergebnisse zur Kettenlängenkontrolle auch in einem wesentlich komplexeren *in vivo* Umfeld zu ähnlichen Resultaten führen. In einer Kooperation mit der Forschungsgruppe von Prof. Boles des Fachbereichs der Biowissenschaften der Goethe Universität Frankfurt, wurden *Saccharomyces cerevisiae* Stämme zur Produktion von kurzen FA entwickelt. Dazu wurde ein Referenzstamm verwendet, dessen *de novo* FA-Synthese aus dem Genom entfernt wurde (FAS ist dort auf zwei Peptide und entsprechend auf zwei Gene verteilt). Die zuvor in den *in vitro* Studien identifizierten Mutationen in KS, MPT und AT Domänen wurden an äquivalenten Stellen in die eigene *S. cerevisiae* FAS eingeführt und diese wiederum auf Plasmiden in die Knockout-Stämme eingebracht. Unterschiede zwischen den Mutanten konnten so vergleichsweise einfach untersucht werden, wobei zu erwähnen ist, dass es sich nicht um ein optimiertes System zur Herstellung von kurzen FA handelte, da die verwendeten Plasmide nur in wenigen Kopien pro Zelle vorliegen[24], und zunächst auch die natürlichen Promotoren verwendet wurden)[25].

Die *in vivo* Ergebnisse können nichtsdestotrotz als ähnlich guter Erfolg wie die *in vitro* erhaltenen gewertet werden. Interessanterweise wurden bei den *S. cerevisiae* Fermentationen die kurzen Fettsäuren außerhalb der Zelle gefunden, was auf eine Hydrolyse der durch FAS hergestellte Acyl-CoA-Ester innerhalb der Zelle und anschließenden Export schließen lässt. Während potentielle Thioesterasen für die Hydrolyse für *S. cerevisiae* kürzlich beschrieben wurden[26-27], bleibt der Export bisher ungeklärt. Die höchsten Konzentrationen an kurzen FA lagen bei 118 mg/L YPD

Medium[28-29] und damit in vergleichbarer Höhe zu Ergebnissen bisheriger Studien (mit jedoch anderen, stark weitreichenderen Engineering-Strategien)[30].

1.3 Fazit

Die Synthese von FA aus kleinen Vorläufermolekülen durch FAS ist ein hochkomplexes Netzwerk von Reaktionen. Durch rationales Engineering von Schlüsselpositionen in KS, MPT und AT konnten Substratbindungsaffinitäten und folglich Reaktionsflüsse zur gezielten Herstellung von kurzen FA modifiziert werden, was in einer Publikation und einem Patent festgehalten wurde[17-18]. Darüber hinaus wurde in einer ersten Machbarkeitsstudie das Potential der Modifizierung von Domänen-Domänen-Interaktionen zur gerichteten Synthese am Beispiel der Herstellung von kurzen FA gezeigt.

Die so geschaffenen Grundlagen zur Kettenlängenkontrolle in FAS wurden um Studien erweitert, die angrenzende Themen behandelten. Zum einen wurden FAS in einen größeren, gekoppelten Syntheseweg für die *in vitro* Herstellung eines Polyketids eingebettet. Zum anderen wurde das Potential der zuvor identifizierten Mutationen zur Produktion von kurzen FA auch *in vivo* getestet, wo sie genau wie die *in vitro* Studien zu guten Ergebnissen (und einem weiteren Patent) führten[28-29].

Die Arbeiten dieser Promotion bieten somit wichtige Ansätze und biotechnologische Lösungen für Prozesse, in denen die Produktion von kurzen Acyl-CoA-Estern wichtiger Bestandteil ist.

2 Introduction

Fatty acids (FA) are molecules consisting of a polar head group and an aliphatic tail. They are ubiquitous in all known organisms throughout nature where they commonly serve as components of membranes, in energy storage, as signaling molecule or as a part of secondary pathways.

While FA and their biological synthesis have been the subject of research for decades[1], they have once more gained attention in the last couple of years due to the increased interest in their exploitation for biofuel production and their roles in green chemistry[2]. In organisms, FA are produced from small precursor molecules in a multistep synthesis by enzymes named fatty acid synthases (FAS).

2.1 Fatty acid synthases

Within the enzyme class of FAS, two types are distinguished according to their general organization: In type I FAS, all enzymatic centers are gathered on one multienzymatic complex. This form can be found in eukaryotes and few bacteria (such as *Corynebacterium, Rhodococcus, Nocardia* and *Mycobacterium*)[4]. In type II FAS which is typical for prokaryotes and mitochondria on the other hand, the enzymatic domains are all located on separate proteins [31]. As the focus of this work is on type I FAS, type II will not be described in detail, but references to type II FAS as homologue domains are made at certain points.

As far as the chemistry of FAS is concerned, it is essentially the same for both classes (Fig. 1, here specifically explained for fungal and bacterial type I FAS): The substrates are transferred onto the acyl carrier protein (ACP) by transferases. First the starter unit, acetyl, is loaded by the acetyl transferase (AT) from acetyl-CoA. In a similar mechanism, the elongation unit, malonyl, can be transferred by the malonyl/palmitoyl transferase (MPT). In a condensation reaction, the ketoacyl synthase (KS) domain then elongates acetyl (in later cycles acyl) with malonyl to form the β-ketoacyl-ACP intermediate. This intermediate is processed in a series of steps starting with the ketoacyl reductase (KR) where it is reduced with NADPH to β-hydroxyacyl-ACP. After dehydration in the dehydratase (DH), another reduction in the enoyl reductase (ER) follows. Here also, NADPH acts as the reduction equivalent, but electron transfer is mediated by FMN, a necessary cofactor[32].

The fully reduced acyl chain can then be loaded into the KS domain again for another cycle of elongation. After several repeats, the acyl chain is eventually released from the enzyme via the MPT domain as an acyl-CoA ester.

Differences exist in the product release step, e.g. in type II FAS the final product is often considered the acyl-ACP which can be cleft off as a free FA by a thioesterase (TE)[33]. TE are also used for product release by some type I FAS, i.e. the mammalian FAS[34].

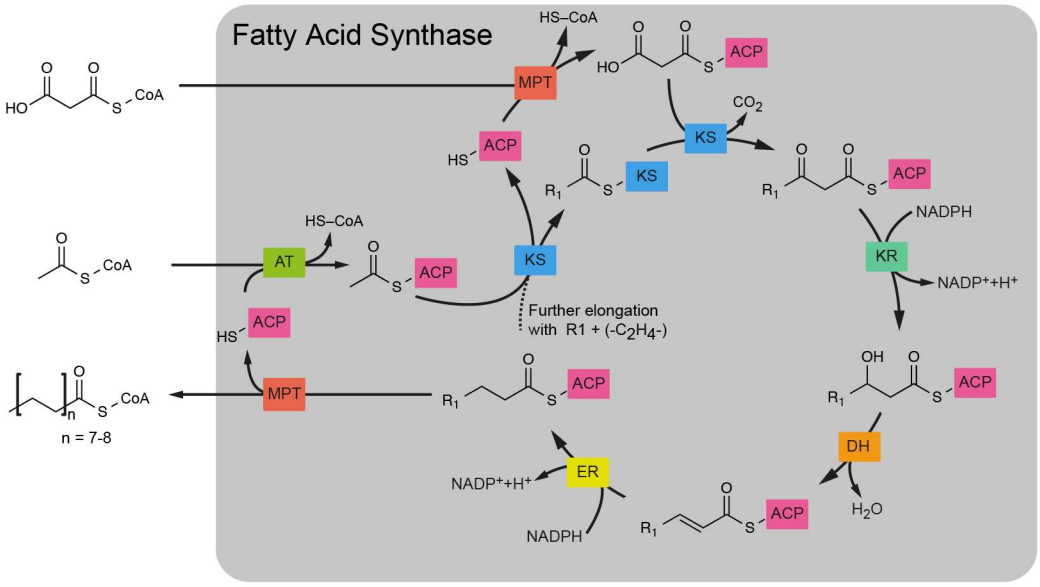


Figure 1: FA synthesis in type I FAS systems.
The FA synthesis cycle (here, shown specifically for yeast and bacterial type I FAS) is started with the loading of acetyl which is transferred onto ACP and then moved into the KS domain. ACP returns loaded with malonyl and elongates acetyl (in later cycles, acyl). After processing in the KR, the DH and in the ER, the fully reduced acyl chain can undergo more cycles of elongation until it is cleft off the enzyme by the MPT domain.

As an important aspect in synthesis, the organization of all enzymatic domains in one megasynthase complex are regarded an advantage for type I FAS. It can be easily assumed that the fixed proximity between catalytic sites evolved so that the high effective concentration of ACP can be exploited[35-36]. In type II FAS, the system relies on the abundance of ACP[37].

2.2 Structural aspects of type I FAS

Within the class of type I FAS, there is a further categorization in three sub-classes: fungal, bacterial, and mammalian FAS. The categorization is made based on

architectural differences with examples shown in Figure 2. Due to their exceptional big sizes and synthetic capacities, type I FAS are counted to the group of megasynthases.

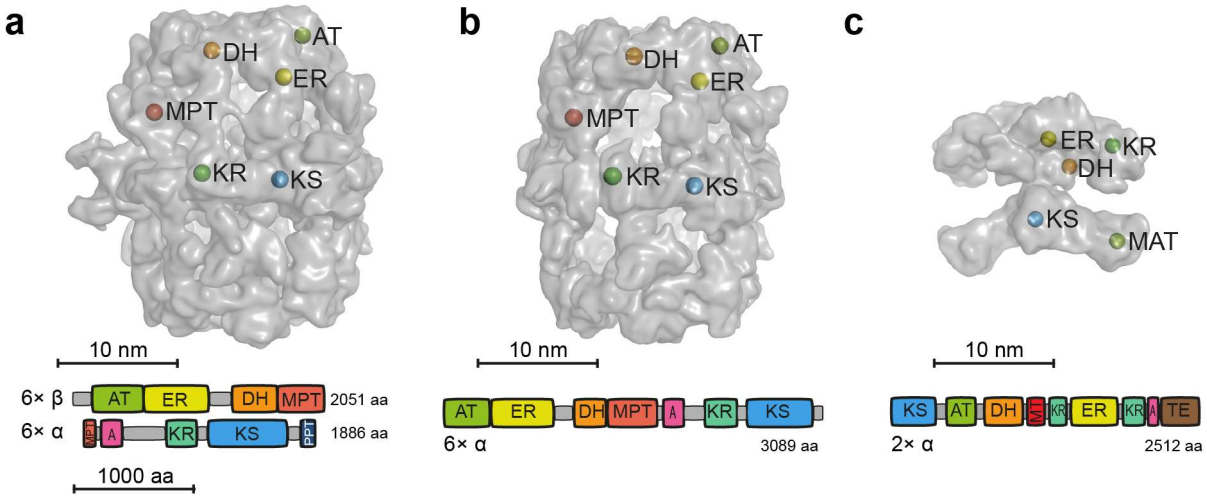


Figure 2: Classification of type I FAS.
The three classes of type I FAS are shown here in surface depiction. **(a)** Fungal FAS structure can be best described as a barrel. Although gene topological variants exist [38], one set of catalytic domains is typically distributed on two polypeptide chains. In *S. cerevisiae* FAS, several copies of each chain assemble to the $\alpha_6\beta_6$ heterododecamer. One set of catalytic domains is represented with spheres indicating the active centers (structure based on *Saccharomyces cerevisiae* FAS, protein data bank (PDB) code: 3HMJ[11]). **(b)** The bacterial FAS seems very similar to the fungal FAS. However, all catalytic domains are typically on one chain, leading to a homohexameric α_6 arrangement (structure shown here is from *Mycobacterium smegmatis*, PDB code: 3ZEN and 4B3Y[39]). **(c)** In contrast to the previous two, the architecture of the α_2 homodimer mammalian FAS is organized in an X-like shape (here, exemplary, the FAS from *Sus scrofa* (pig), PDB code: 2VZ8[40]). The TE domain is not resolved in the structure. Abbreviations: ketoacyl synthase (KS), acetyl transferase (AT), malonyl/palmitoyl transferase (MPT), enoyl reductase (ER), ketoacyl reductase (KR), dehydratase (DH) and malonyl/acetyl transferase (MAT), phosphopanteheine transferease (PPT), acyl carrier protein (A), pseudo methyl transferase (MT), thioesterase (TE).

Fungal FAS was the first solved in X-ray structure in full length[36] and has since been characterized by a whole set of structural studies with resolutions up to 3.1 Å[10-13]. The overall architecture of the $\alpha_6\beta_6$ heterododecamer is often described as a barrel, where the central wheel is formed by the six α chains. The two domes are each formed by three β chains.

The bacterial FAS shows high overall similarity to fungal FAS, which is visible at first sight from the available low resolution structural data[14, 39, 41]. There are, however, some differences such as 1) substantially fewer insertion elements, which extensively stabilize the barrel-shaped structure in fungal FAS, leading to a more

porous architecture and possibly higher overall flexibility[41]. 2) The catalytic domains are located on one polypeptide chain which resembles a β-/α-fusion of fungal FAS[15]. 3) The phosphopantetheine transferase (PPT) is not part of the enzyme but is located on a separate protein. In sum, the bacterial FAS has been regarded a minimal version of fungal FAS[38].

The mammalian FAS structure is in clear contrast to the two previously described type I FAS as can be seen in its 3.2 Å crystal structure[40]. The overall dimeric α_2 architecture can be described as X-shaped or as "gingerbread man"[2]. Especially its outstanding flexibility is worth mentioning at this point allowing swiveling and swinging motions[42].

2.2.1 Similarity to polyketide synthase (PKS) systems

An enzyme class which belongs to the megasynthase family and is related to FAS are the polyketide synthases (PKS). The products of PKS, the polyketides, have gained much attention due to their high potential as drugs with anti-cancer, anti-bacterial, anti-inflammatory or anti-cholesterol activity[43]. On the other hand, also non-desirable and even carcinogenic effects have been reported for polyketides, e.g. for Aflatoxin B_1, a polyketide produced by a series of *Aspergillus* strains[44].

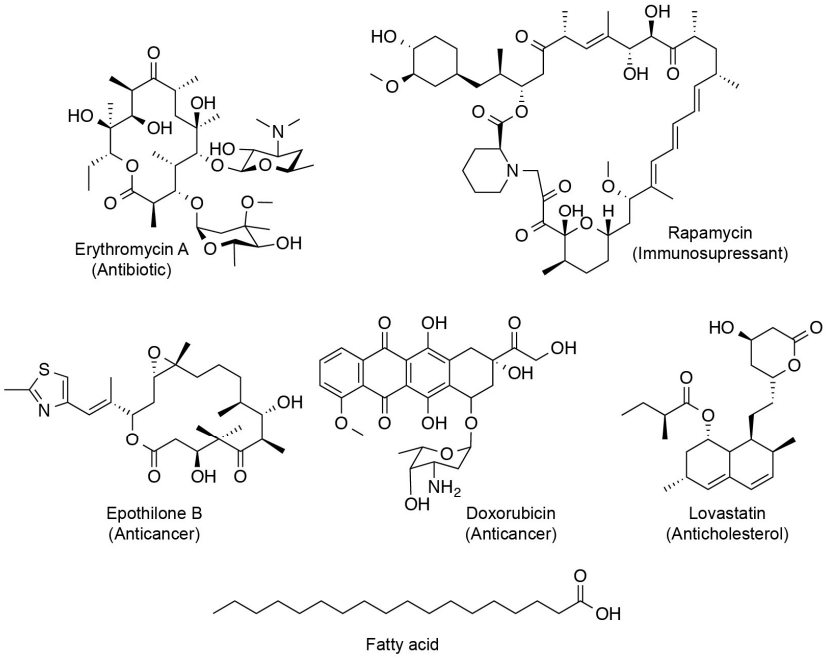


Figure 3: Examples of polyketides and FA.
While the structures of PK are very diverse[45-46], the production of PK and FA are based on the same chemical principles. Diversity in PK can derive from varied decorations which are either intrinsically formed by the PKS or by post-production processing.

In synthesis, PKS and FAS essentially share the same chemistry: carbon-carbon bonds are formed by a KS domain and the ketoacyl intermediate can be processed in its β-position. In contrast to FAS, processing in PK intermediates does not necessarily run through the full cycle of reduction and dehydration steps but they can be elongated before and thus, the functional group of the β-position remains. This varying degree of decoration is one of the main reasons for the wide variety of polyketides (Fig. 3)[43, 47]. In addition, polyketide synthesis can involve a number of further modifying enzymatic domains, both incorporated into PKS or standalone proteins, which are responsible for further polyketide modification during or after their synthesis. Examples are methyl transferases, oxidases, or cyclization domains[48]. Also the choice of starter units and elongation units is not as strict as in FAS and can include propionyl-CoA, methylmalonyl-CoA but also exotic compound such as benzoyl-CoA or cinnamoyl-CoA can be used in the synthesis by PKS[43].

For a broad classification, PKS are also distinguished by their domain organization (similar to FAS) with type I being specified as multienzymatic complex and type II having the enzymatic domains distributed on separate proteins. The additional group of type III PKS should only be mentioned for completeness (synthesis therein is different to type I and type II PKS as type III PKS do not use an ACP domain)[49].

There are more classifications into subcategories made according to certain synthesis parameters as for example their working mode and the degree of reduction in the products. The degree of reduction depends on the number of processing steps which are employed before an intermediate is subject to the next elongation. It can either be non-reductive, partially reductive or fully reductive[50]. The classification concerning the working mode distinguishes between two types: When all enzymatic centers are each used several times during synthesis, PKS are termed iterative. The second type of PKS, where intermediates are handed down from one set of catalytic domains (named module) to a next one after one round of elongation/processing, is called modular[51]. In this classification scheme, FAS could also be understood as a PKS, i.e. as a fully reductive iterative PKS[52].

As far as the structure of PKS is concerned there are also remarkable similarities to FAS. Within domains, the X-ray structures of KS, malonyl/acetyl transferases (MAT), KR, DH and TE, share substantial elements with their counterparts in FAS[51].

On a bigger scale, the architecture of full length PKS is still up to debate. Only few structures have been solved for the mentioned individual domains and small multidomain complexes[53-54], but no X-ray structure of a whole module could yet be solved[52]. Judging from the arrangement of the catalytic domains and the existing structural data, the mammalian FAS has served as a model for the overall structure[55] and it was hypothesized that PKS and mammalian FAS have a common ancestor[51]. Recently, an EM (electron microscopy) structure of a full length PKS was solved contradicting this theory[56], but its role as a model for the whole PKS field remains disputed[55, 57-58].

In general, the last decades of PKS research and their engineering were dominated by drawbacks. The hopes to easily harness the synthetic potential of PKS by swapping domains in a toolbox like fashion has had some success which, however, cannot be weighed up against the tremendous amount of times it failed[52]. Reasonable amounts of protein of PKS are neither easily obtainable nor easily handled and as a consequence PKS characterization has not developed to the same level as FAS. However, since FAS and PKS share the same principles in many aspects, using FAS to gain more information about PKS has been one way to overcome these problems.

2.3 Technological and industrial usage of FA

FA play a very important role in biotechnology and have recently again gained more attention due to their role as a platform chemical[2, 59]. The carboxyl (or in some cases the CoA ester) head group can be processed further to yield a variety of functional groups such as fatty acid ethyl esters (FAEE), fatty acid methyl esters (FAME), fatty alcohols, aldehydes, and alkanes as well as alkenes (Fig. 4)[3]. These compounds can be used in a wide range of products with the most prestigious one being biofuels. Since regular, fossil-derived fuels add to the greenhouse effect by releasing previously bound carbon to the atmosphere as CO_2, efforts to supplement or replace them by biofuels as a renewable energy source have been intensified in recent years. In one scenario, microorganisms are used for the conversion of biomass into organic compounds that can be used in combustion engines. In the best case, these biofuels share substantial characteristics with currently used fuels as far as corrosivity, flash point, energy density, and melting temperature are concerned[6]. Biofuels sharing the same properties with regular fuel are also termed “drop in” biofuels. They could easily be used with existing infrastructure, e.g. means of distribution, and already developed engines, which would keep costs at a minimum. The exact requirements on biofuels are depending on the type of combustion engine they are designated to power. In detail, three common classes are distinguished: gasoline, diesel, and kerosene. For the latter two, diesel and kerosene, usually long hydrocarbon chains are used (C_8–C_{21}; in C_n, n represents the carbons in the chain), while for gasoline short chain products are needed (C_4–C_8)[5-6]. While some proof of concept studies have shown the feasibility of using FA synthesis for biodiesel production (where long chains are needed)[60-62], short chain products as needed for biogasoline have lagged behind[5].

One of the main reasons for that is the missing availability of short chain FA. Naturally, the production of short FA does not occur at all or only in far less abundance than that of the typical long chains (C_{16}–C_{18})[63-64]. Natural production of short FA has only been reported for few exceptional cases e.g. in secondary metabolism[19-21], mammary glands[65], or as part of metabolic pathways in mitochondria[66]. For exploiting regular FA synthesis for the production of short chain products as needed for gasoline-like biofuels, engineering is required to impose a system of chain length control[59]. In biofuel research, chain length control has been

considered one of the biggest challenges (besides increasing yields and using lignocellulosic hydrolysates)[3]. Previous advancements in the field will be discussed later.

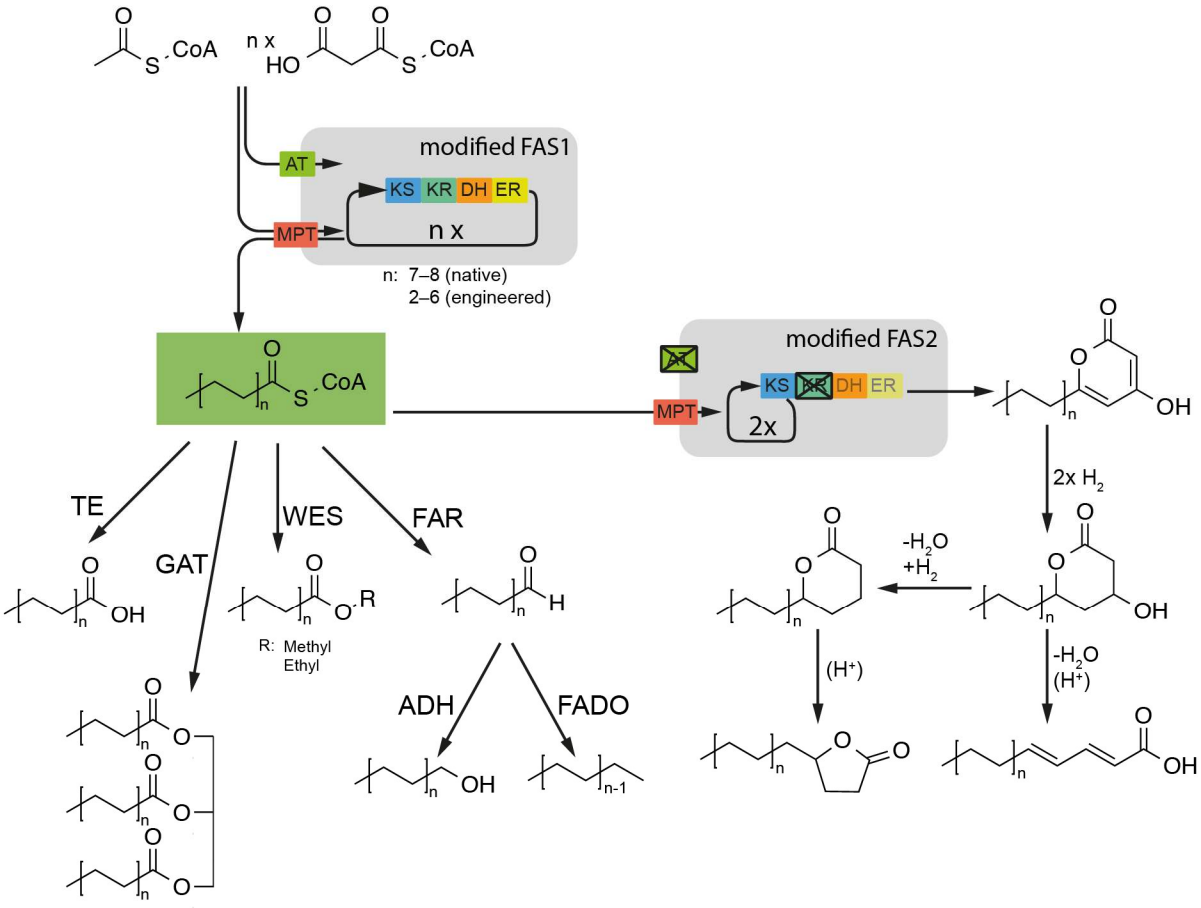


Figure 4: Fatty acid CoA esters as platform chemical.
The direct products of FAS, the fatty acid CoA esters, are synthesized from small substrate molecules, acetyl-CoA and malonyl-CoA. FA can be further processed to a variety of industrially relevant chemicals of which selected ones are shown here. Products are for instance free fatty acids, fatty acid methyl-/ethyl-esters (FAME/FAEE), aldehydes and even fatty alcohols and alkanes. In an additional synthesis route, the fatty acid CoA esters can be processed by a second FAS carrying modifications to yield the corresponding 6-alkyl-4-hydroxypyran-2-one. Starting from this compound a variety of additional substances are accessible synthetically[67]. The length of the acyl chain in the product is determined by how often it is elongated in FAS. Abbreviations: TE, thioesterase; GAT, glycerol acyl transferase; WES, wax ester synthase; FAR, fatty acyl-CoA reductase; ADH, alcohol dehydrogenase; FADO, fatty aldehyde deformylating oxygenase.

Besides the field of biofuels, there are other uses of short FA derivatives, some of them are directly connected to human consumption. As far as marketing is concerned, the labeling of compounds as “natural” is an important asset in that issue. According to the FDA (U.S. Food and Drug Administration) any compound can be

declared to be “natural”, when it is has been produced by a living organism or a live-derived source[68]. If biotechnologically obtained FA derived products do not include any chemical steps, they can meet this requirement[69].
Such products could for example derive from fatty alcohols and aldehydes, which are commonly used as soaps, detergents or additives to cosmetics[61]. Esters of short FA have also been used as flavoring compounds in various cases[70-71]. Furthermore, FA can be esterified to yield triacylglycerols, a compound class that is currently used as a food supplement to fight obesity. The intake of triacylglycerol esters of short FA has been connected to several beneficial effects which include the reduction of body fat and enhanced insulin sensitivity[68].
The development of these consumer oriented compounds can be considered the primary focus as selling prices of biofuels have to be very low which makes their production not as feasible in the short run[3].

FA can also be part of bigger synthesis pathways and thus, lead to an even greater diversity of final products. In naturally occurring pathways, FA are for example attached to bigger molecules, which has been reported for the synthesis of penicillin K[72]. In other cases, FA are not attached, but elongated again by other megasynthases that work together with FAS, for instance in the synthesis of Aflatoxin B_1[21] or in the synthesis of resorcylic acid lactones[73-74].
One example of an artificial pathway is also shown in Figure 4. The product of a first FAS is taken up by a second FAS and elongated again. Using this synthesis principle, small polyketides such as 6-alkyl-4-hydroxypyran-2-ones can be produced. These compounds share the same functional structure with other lactones like triacetic acid lactone (TAL), which is considered a platform chemical itself[67, 75]. Lactone conversion by chemical synthesis can provide access to a variety of additional molecules (see Fig. 4)[76]. Possible final products include for example δ-decalactone (derived from C_6-CoA) which is a widely spread flavoring compound (naturally present in sweet potato and other fruit, with a coconut-like or dairy fragrance)[77-78].

In regard to the biotechnological use, one peculiarity of fungal and bacterial type I FAS should be mentioned. As shown in the reaction scheme in Fig. 4, these FAS

have a beneficial characteristic: The final products of both classes of FAS are acyl-CoA esters and not free FA as in mammalian FAS or type II FAS systems. In acyl-CoA esters a high energetic level is preserved and therefore they can be processed further directly without previous activation[79].

2.4 Rational engineering of FAS to broaden the product spectrum

Whenever the natural product spectrum of FAS cannot be used directly (as previously discussed e.g. for biofuels in the gasoline range), efforts aim at adapting FA production in a way that it can meet the requirements. The main focus of this thesis was on chain length control in FA synthesis. The aim therein was twofold: 1) FAS was to be reprogrammed regarding its chain length distribution from long chain products towards short chain products. 2) The gained knowledge should be used to understand the underlying mechanisms of chain length control better.

Previous research on the production of short FA covered a whole range of strategies including expression or regulation of expression of certain proteins as well as protein engineering which was also our method of choice. More specifically, the focus of this dissertation was on the rational protein engineering which is characterized by the mutation of few, selected amino acid residues in order to change protein characteristics. This engineering method has certain advantages such as a smaller library size which can be tested in detail and the opportunity to directly map structure-function relationships[80]. Also, since there has not yet been a report on a method for efficient selection processes for short FA, directed evolution was ruled out since it could not have been used without laborious and very expensive screening. Competitive strategies for short FA production will be mentioned for comparison in this work as well, but will not be discussed extensively.

For the engineering of FAS and in particular the aspect of chain length control in FA synthesis (see Fig. 1), the key factors in the synthesis are: 1) the elongation step in the KS, 2) substrate loading and product unloading by the transferases by AT, MPT (or TE), and 3) intermediate shuttling the ACP. These domains will be discussed in detail in the following paragraphs. On the other hand, KR, DH and ER were neglected as they are only involved in modifying the intermediates. These synthesis steps were believed to be independent of chain length and it was assumed that they always would be conducted once an elongation took place.

2.4.1 KS

In FA synthesis, the KS is responsible for the formation of the new carbon-carbon bond and thus, for the actual elongation of an intermediate. In detail, an acyl (in the first cycle an acetyl) is first deposited in the KS domain where it covalently binds to the active site cysteine. This step is also named the “ping” step in the “ping pong” reaction scheme. It can be easily assumed, that this first step already is a key factor in chain length control, since substrate loading into the KS domain is a prerequisite for later elongation. Once, the acyl group is deposited in the KS domain, the free ACP needs to be loaded with malonyl by the MPT domain to proceed. The ACP returns to the KS domain where malonyl is first decarboxylated to yield the enolate intermediate which then elongates the acyl in a Claisen condensation to form the β-ketoacyl-ACP intermediate[81].

In early publications on FAS, it was already speculated that in fact chain length is determined when the fully reduced acyl chain is either loaded into the KS domain for another round of elongation or instead cleft off the enzyme by the MPT domain[82-83]. Up till today, it is still under debate which step within the KS reaction mechanism contributes most to chain length control. In related PKS systems, there are hints for the importance in chain length control of the loading step of acyl (“ping”)[84] as well as for the actual elongation step (“pong”)[85].

As far as the structure is concerned, X-ray structures of KS domains from type II FAS systems such as *E. coli* have also been solved numerous times with substrates or inhibitors bound[86-88]. Furthermore, structures for type I FAS KS domains are available among them also *S. cerevisiae* FAS[10, 12]. KS domains show a high over all homology as can be seen from the structural superposition of representatives from different organisms in Fig. 5[89].

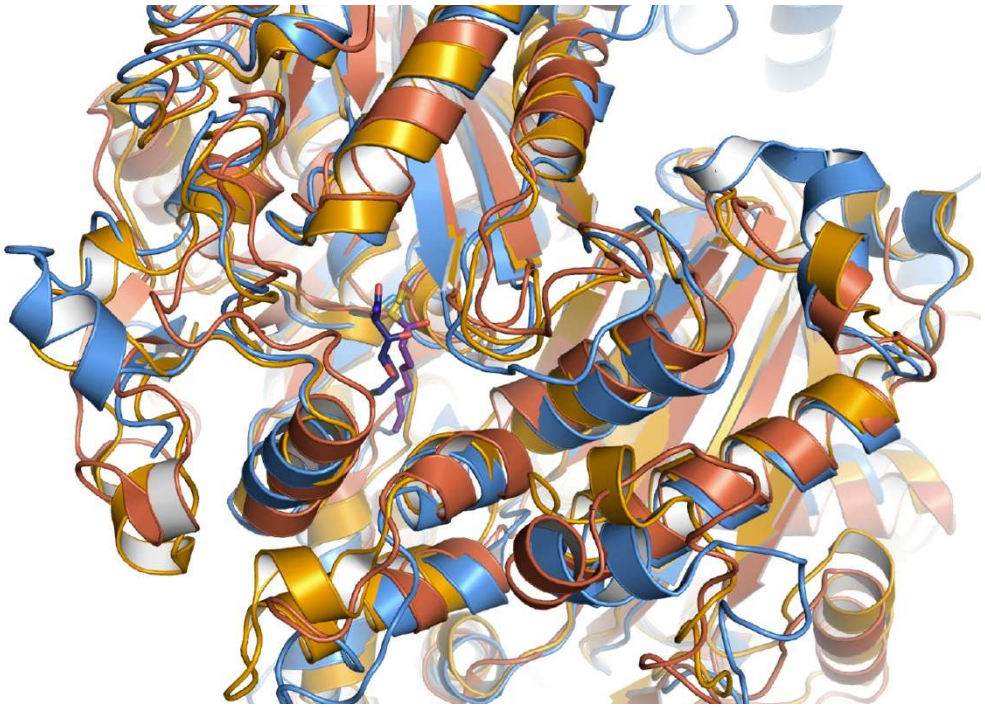


Figure 5: KS secondary structure comparison.
Representative KS domains from different backgrounds (mammalian, bacterial and fungal) were compared and displayed in cartoon depiction. The specific view shows one active center with ligands in stick representation. As a reference, the type I FAS *S. cerevisiae* KS domain (colored in blue, PDB code: 2VKZ[12]) with bound cerulenin (dark blue) was used and other KS domains were aligned using the "align" command of PyMOL (Version 1.7.0.3, Schrödinger, LLC). This algorithm first aligns based on sequence, followed by structural superposition (outliers are rejected during refinement with specific RMS; set to 3 Å). From type II FAS, *E. coli* FabB (orange, PDB: 2BUI[90]) aligns with RMSD of 1.5 Å to *S. cerevisiae* FAS (in 445 atoms) and has a hexanoyl bound in the active site (purple). As an example for type I mammalian FAS, the *Sus scrofa* (pig) KS domain (red, PDB code: 2VZ8[40]) was compared to *S. cerevisiae* in the same way (RMSD of 2.9 Å, 459 atoms).

As far as previous protein engineering efforts for short FA production are concerned, two publications should be mentioned explicitly: In the *E. coli* type II systems *FabF* (a KS derivate) was engineered and two big residues were introduced into the binding channel[7]. Also in type I FAS, a mutation was reported to enable *S. cerevisiae* FAS to produce short FA[8-9].

There are also other examples for short FA synthesis where KS domains were important but not further engineered: In one example on the *in vivo* production of FA in *E. coli*, short FA titers were increased when the described *FabF* derivative was used in combination with a metabolic dependent degradation of other KS in the cell[7, 91]. In a different study short FA synthesis was pushed by simply overexpressing *FabH*, the KS equivalent responsible for the starting condensation in that particular type II system[92].

In sum, these efforts have paved the way for further engineering but did not yet provide a well-regulated engineering to produce short FA by type I FAS. In several projects of this thesis, KS engineering was addressed. Three mutation sites were identified within the binding channel. They were first probed *in vitro* on a heterologously expressed *Corynebacterium ammoniagenes* FAS, where they showed very promising results. In the best construct the native system of chain length control was overwritten: C_8-CoA became the most dominant product making up 74% of the detected CoA esters while no production of long chain products could be detected anymore[18]. When the mutations were tested *in vivo* in *Saccharomyces cerevisiae* (baker's yeast) on its native FAS, the results were mostly confirmed with minor differences[28-29].

2.4.2 MPT

The MPT domain catalyzes the transfer of malonyl-CoA onto the enzyme. Mechanistically, it also follows a ping pong mechanism, wherein malonyl first covalently binds to the active site serine and is then transferred onto ACP yielding malonyl ACP. This process can, however work in both directions.

Since the MPT domain is responsible for the supply of the elongation unit, malonyl, it directly influences KS activity, where malonyl-ACP is needed for a productive elongation of an acyl chain. In early *in vitro* FAS experiments, it had already been established that the malonyl-/acetyl-CoA ratio was important[83, 93]. Lower relative concentrations of malonyl led to a higher amount of short FA.

There have also been previous engineering efforts on related FAS transferases that addressed substrate specificity. In the human malonyl/acetyl transferase (MAT) a conserved arginine was mutated to smaller residues such as lysine or alanine, which decreased the binding affinity of malonyl continuously[22]. In a later study, the crystal structure of the protein was determined and the arginine was identified to coordinate the free carboxyl group of malonyl[23].

Besides the loading of malonyl, the MPT domain is also responsible for the unloading of products. A fully reduced acyl chain is transferred from the ACP onto the MPT domain and then again on a free CoA. In this process, the MPT does not seem to distinguish between different lengths of products, but it has been shown to transfer a broad variety in lengths of fully reduced acyl chains[94-96].

Engineering of the MPT as a second major actor in chain length control, was believed to be hard to achieve[30]. However, the *in vitro* studies on *C. ammoniagenes* FAS eventually brought forward a powerful modification. Short chain yields increased from a share of 18% C_8-CoA to 47% C_8-CoA (out of the detected CoA esters) when this MPT mutation was introduced into a previously mutated FAS construct. When transferred into the *in vivo* environment of *S. cerevisiae*, an MPT modification by itself was even able to increase levels of C_8 to 80 mg/L, an incredible 137-fold increase over the wild type.

2.4.3 AT

The AT domain catalyzes the loading of acetyl from acetyl-CoA. As the transferases in FAS also share substantial similarities among each other, the structure and the mechanism are very similar to the MPT domain[36].

The previously mentioned influences of the malonyl-/acetyl-CoA ratio on chain length distribution[83, 93] have also sparked our interest in targeting the AT domain by engineering. The aim was to increase the availability of acetyl, the starter unit, and consequently push the production of short FA. Similar methods as described for the MPT domain were used which were encouraged by the studies on changing the substrate specificities in the human MAT[22-23]. Additionally, the mentioned publications also reported on allegedly opening up a new binding channel for short acyl chains such as octanoyl when certain mutations were introduced into their transferase. If indeed the AT domain's substrate spectrum was to be expanded to also transfer short acyl chains, it could act as a second exit pathway enhancing short chain product cleave off.

In the end, engineering on the AT domain has led to interesting results: In the *in vitro* testing of the *C. ammoniagenes* FAS, it could slightly increase short FA yields when introduced in constructs already carrying mutations, e.g. over all yields of C_8-CoA were 14 µM instead of 9 µM in the reference. In *S. cerevisiae* fermentations, an AT modification showed beneficial effects, e.g. when an AT mutation was introduced into a triple mutant FAS, short FA production was further increased 3-fold over the reference, making it the best performing strain in the particular study.

2.4.4 ACP

Besides the catalytic centers, the shuttling protein, the ACP, can be targeted for shifting product distributions. For general FAS activity, the ACP has to be posttranslationally modified by the attachment of phosphopantetheine so intermediates can bind to it as a thioester[11]. This covalent bond between intermediate and ACP is kept at any time during the synthesis cycle except when fully reduced acyl chain is waiting in the KS domain for further elongation. Thus, intermediates do not diffuse but are bound to the enzyme until they are released as CoA esters (or free FA as in the case of mammalian FAS).

The ACP itself is doubly tethered to the rest of the enzyme in bacterial and fungal FAS. Shuttling between the different catalytic sites follows a stochastic pattern[97]. Simulations also revealed that one ACP usually stays within its own set of catalytic domains, but could reach others as well[37].

Furthermore, the ACP was found to also interact with its bound intermediates itself. The growing acyl chains were observed to be sequestered within a hydrophobic pocket of the ACP, similar to a switchblade knife[10]. When in contact with a catalytic domain, the chain would then leave the hydrophobic pocket and find its way into the binding channel of the catalytic domain. This phenomenon, now commonly termed chain-flipping mechanism, was, however, mostly described for type II ACPs[98]. Previous engineering efforts even exploited this mechanism for short chain production by modifying the binding pocket to discriminate against longer chains[99]. In contrast, sequestered chains in the ACP have not been observed to the same extent in type I systems; NMR studies on *S. cerevisiae* ACP estimate sequestering at only 15% of the time[100].

ACP should in any case not only be understood as an attachment to the substrate. They themselves influence the reactions as studies on successful or failed partnering of ACPs and catalytic domains have indicated[101]. It has even been shown that when ACP is sequestering the intermediate in type II systems, the interaction of ACP with the catalytic domain precedes the loading of the intermediate into the binding pocket[102-103].

While these examples were primarily studied on type II FAS systems, it can be easily assumed that catalysis in type I systems is not either only dependent on substrates and on how substrates can be processed by the catalytic domains. Interactions of

substrate-ACP with catalytic sites influence reaction kinetics tremendously. This is of even higher importance when there are several competing options for an intermediate as for example when a fully reduced acyl chain is bound to the ACP: It can then either be elongated again in the KS or released via the MPT domain – a decision that directly influences chain length[82-83]. While previous mutation studies aimed at mapping these interaction sites[104-106], or enabling basic successful synthesis[107], modifying these interactions and exploiting them for the production of short FA was one project realized in this work.

In the conducted *in vitro* experiments on *C. ammoniagenes* FAS, single mutations on the surface of the KS domain proved very powerful and shifted the product spectrum in the test construct towards short chain products: C_8-CoA made up 77% (of the detected CoA esters) instead of 12% in the reference.

2.4.5 TE

In several FAS systems such as mammalian FAS or type II FAS, the product release step is not catalyzed by an MPT or a comparable transferase, but by a TE. TE hydrolyze either acyl-ACP or acyl-CoA (and in many cases both) to yield the free FA[108]. While in this thesis only MPT containing FAS systems were used, TE engineering plays such an important role that it should be presented for comparison. Previous efforts to increase the yield in free FA have heavily relied on the use of TE. In literature, there have been numerous reports on TE which have a high substrate specificity[109-110]. This substrate specificity has often also been used to produce short FA *in vivo*, e.g. in yeasts, such as *S. cerevisiae*[30] or *Yarrowia lipolytica*[64]. The probably most prominent representative of the TE enzyme class is tesA from *E. coli*. Natively located in the periplasm, the leaderless version ‘tesA can be expressed to stay in the cytosol where it led to tremendously increased titers of free FA [61]. In addition, one variant with a more specific affinity towards short acyl-esters, ‘tesA (L109P), has been found[92].

In general, producing short FA by exploiting TE substrate specificities can have some disadvantages due to low abundance of TE’s substrates. In *E. coli* for example, short chain ACP is not widely available and TE overexpression by itself is not very efficient[91]. Other disadvantages include that for many applications FA are further processed from their CoA ester form, which requires the free FA (as the TE produces

them) to be converted into CoA esters first. This process is catalyzed by fatty acid CoA synthetases (such as *fadD* of *E. coli*) at the cost of one ATP – a step, that can be skipped if products are released as CoA esters directly as in the case of MPT mediated product release[79]. Fungal and bacterial type I FAS are of clear advantage here.

2.4.1 Engineering FAS for the production of a polyketide

Since FAS and PKS systems essentially share the same chemistry and taking into account that extensive structural and biochemical data are available on FAS[10-13, 40], FAS is a suitable model protein from which conclusions can be drawn for the hardly obtainable PKS.

In similar fashion to previously mentioned naturally occurring systems, where FAS and PKS work hand in hand[21, 73-74], a two step pathway for the production of a simple polyketide was created. In addition to a module from the described *in vitro* studies on engineering *C. ammoniagenes* FAS for short FA production, a second FAS module was modified which takes up the short acyl-CoA esters which are produced by the first FAS, and then elongates them non-reductively to 6-heptyl-4-hydroxypyran-2-one (6-HHP) (see Fig. 4). Since 6-HHP shares the lactone structure with TAL, it is likely that 6-HHP can be processed in the same way to yield a variety of compounds[67, 75]. In our *in vitro* study of the one pot synthesis using the system of two modified FAS enzymes, we were able to gain a yield of up to 35% in reference to the limiting substrate over the whole course of the 14 catalyzed reaction steps (not even counting transferase catalyzed reactions)[17].

2.4.2 FAS as part of an *in vivo* metabolic pathway

The *in vitro* probing of FAS as performed during this work, should be seen as one step on the long way to its biotechnological use. From an economic point of view, the production of FA as final product or intermediate is only feasible in *in vivo* systems. While the use of microorganism such as bacteria, yeast or algae imply the use of cheap starting material, FA production requires further development to be profitable and therefore, a short introduction will be given here. The catalytic cycle of FAS is a complex network in itself and complexity is once more topped, when FA synthesis is examined imbedded into its *in vivo* metabolic network.

The initial study on transferring the mutations into an *in vivo* environment was conducted in cooperation with the group of Prof. Eckhard Boles from the biological faculty of the Goethe University Frankfurt (for a detailed listing on contributions, consult the appendix of this work). As mentioned above in the sections on specific catalytic domains, the first experiments in *S. cerevisiae* not only confirmed the *in vitro* data, but showed production up to 118 mg/L of short FA[28-29]. Even without any optimization of the *S. cerevisiae* strain, this result matched previous *in vivo* production in yeast[30].

To further increase FA titers in general, a series of metabolic engineering efforts were reported which dealt with upstream and downstream parts of pathways where FAS is involved. The limits are constantly pushed and free FA production (of long chain products) has been reported to be as high as 5.2 g/L in *E. coli*[111] and in *S. cerevisiae* up to 400 mg/L[60], 2.2 g/L[112] or more recently even 10.4 g/L[113]. For a complete list of measures to increase FA production in microorganisms, current reviews should be consulted, e.g. for *E. coli*[33, 114] or *S. cerevisiae*[115-116]. Unfortunately, the majority of studies, as well as current record holders, are only focused on the production of long FA. Still, some methods are likely to be beneficial for short FA production, too, since they are either 1) upstream to FAS and thus, anyway independent of chain length, or 2) downstream, but with a sufficiently unspecific substrate spectrum. A few selected examples will be introduced in more detail below.

Upstream of FAS, the acetyl-CoA carboxylase (ACC) is considered the first rate-limiting enzyme in FA synthesis. It produces malonyl-CoA by fixating a carbon dioxide group to an acetyl-CoA under the consumption of ATP[34]. Its activity is tightly controlled by phosphorylation which deactivates the enzyme[117]. Engineering towards a higher malonyl-CoA supply by developing a hyperactive ACC with a mutated phosphorylation site led to an increase in total lipid production[118]. Besides that, the expression of ACC is already down regulated by increased FA levels which can be overcome by exchanging the native promoter[60, 119]. The combination of higher malonyl-CoA availability with a FAS engineered to produce short FA could, on the one hand, lead to higher total FA titers, but on the other hand, shift product distribution towards longer FA as found for increased malonyl/acetyl ratios *in vitro*[83, 93].

As far as downstream or competing elements of the FA pathway are concerned, the counter process to FA synthesis, the beta-oxidation, can be considered the most important. In one of the many studies on beta-oxidation interruption, which aimed at ensuring that FA are not directly degraded again, key enzymes in the beta-oxidation have been knocked out and short FA production was increased up to 140-fold over the parent strain[120].

Regarding the downstream processes of FA synthesis, several options for their further conversion to free FA, esters, fatty alcohols, aldehydes or alkanes/alkenes have been discussed in a previous chapter (see Fig. 4).

Taken together, there are many possibilities of incorporating a modified FAS into reaction pathways in a plug-in like fashion. The engineering of FAS as a central piece in the synthesis chain is therefore a desirable goal.

3 Materials and methods

3.1 *In vitro* studies on *C. ammoniagenes* FAS

3.1.1 Initial cloning

FAS from *C. ammoniagenes* was initially cloned into a pET-22b(+)vector (Novagen, Merck Millipore, Darmstadt, Germany) from genomic DNA (DSM-20306, DSMZ, Braunschweig, Germany) in previous projects[14, 121]. An exemplary vector map is shown in Supplementary Fig. 1. It was constructed with an N-terminal strep twin tag with short linker in between the double tag motive (MSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKGAGS) or the shorter version, the strep II tag (MSAWSHPQFEKGAGS)[122]. Primers were designed to be used in Infusion cloning (15 bp overlap with vector ends at the 5' end, underlined):

fwd 5' AAAAGGCGCCGGATCCACTATTGGCATCTCTAACCACCGCCTGG 3' and
rev 5' GGTGATGATGCTCGAGCTGGTGGCTTGCCGTAGATCGCTTGC 3'.
Initial construction of the insert included approximately 270 bp after the actual stop codon since the gene was wrongly annotated in databases.

For later co-transformation with *C. ammoniagenes* FAS AcpS (for posttranslational addition of phosphopantetheine on the FAS), AcpS was cloned into a pETcoco vector (Novagen, Merck Millipore, Darmstadt Germany, also from a previous project[14]) and amplified (vector map shown in Supplementary Fig. 2). The following primers were used (the overlap to the vector is underlined, the parts binding the genomic DNA for the insert amplification are shown in bold):

fwd 5' AGAAGGAGATATAAGC**ATGCTCGACAACCGTGAAGCGATGAC** 3'
rev 5' TCGAGTGCGGCCTAGG**TTACCGCTGGTACCGCAGCAGG** 3'

3.1.2 Point mutations

For the introduction of point mutations, site directed mutagenesis was used and vectors were linearized via a PCR in which primers carried the mutation (for a better overview, primer are listed below in grouped tables, Table 1–4, and origin of plasmids is stated in Supplementary Table 1 and chapter "7 Statement of personal contributions") when they are not the own work but from the previous studies[14, 121]. After digestion with Dpn1 (New England Biolabs, Frankfurt, Germany), a preparative agarose gel (0.8%) was run and the DNA was extracted. The linearized vectors were used for transformation of competent *E. coli* cells (Stellar cells, Clontech, Mountain View, USA) according to the manufacture's protocol. All constructs were sequenced over the full length of the insert and its integration sites, if not stated otherwise.

Table 1: Primers for influencing substrate binding in the AT, MPT and KS domain.
The mutation site is indicated in bold, while the overlap of primers is underlined.

Mutation	Direction	Sequence (5' – 3' display)
I151A	fwd	CAGCTC**GCC**GGCGTCGCTATTTCTAAG
I151A	rev	GACGCC**GGC**GAGCTGCGCCAGCGCAATAAC
R1408K	fwd	GAAATCGTCTACGCC**AAG**GGTTTGACCATGCAC
R1408K	rev	GGCGTAGACGATTTCTACAACGGCTTCC
G2599S	fwd	GCTCGACCCAGGGCACG**AGT**ATGGGCGGCATGCAGTCG
G2599S	rev	CGACTGCATGCCGCCCAT**ACT**CGTGCCCTGGGTCGAGC
G2599S M2600W	fwd	CCCAGGGCACG**AGTTGG**GGCGGCATGCAGTCGATGCGC
G2599S M2600W	rev	GCGCATCGACTGCATGCCGCC**CCAACT**CGTGCCCTGGG
L2628Y	fwd	ATCTTGCAGGAAGCA**TAT**CCGAATGTCGTG
L2628	rev	TGCTTCCTGCAAGATGTCATTCGG

Table 2: Primers for mutations on the KS surface for influencing ACP–KS interaction.
The mutation site is indicated in bold typing, while the overlap between the primers is underlined.

Mutation	Direction	Sequence (5' – 3' display)
D2553A	fwd	CCGGCC**GCT**ATGATTGACAACCTCGACCGG
D2553A	rev	AATCAT**AGC**GGCCGGAATACCGTAGACACCGG
D2553N	fwd	CCGGCC**AAT**ATGATTGACAACCTCGACCGG
D2553N	rev	AATCAT**ATT**GGCCGGAATACCGTAGACACCGG
D2556A	fwd	CCGGCCGATATGATT**GCA**AACCTCGACCGGGTGGCCTTGTGG
N2557A	fwd	CCGGCCGATATGATTGAC**GCA**CTCGACCGGGTGGCCTTGTGG
N2557E	fwd	CCGGCCGATATGATTGAC**GAA**CTCGACCGGGTGGCCTTGTGG
D2556 N2557	rev	AATCATATCGGCCGGAATACCGTAGACACCGG
N2621A	fwd	GCGCAGCCGCGCCCG**GCT**GACATCTTGCAGGAAGCACTGCCG
N2621D	fwd	GCGCAGCCGCGCCCG**GAT**GACATCTTGCAGGAAGCACTGCCG
D2622A	fwd	GCGCAGCCGCGCCCGAAT**GCA**ATCTTGCAGGAAGCACTGCCG
N2621 D2622	rev	CGGGCGCGGCTGCGCCAG
A2696K	rev	GGCTTCGGCGATATG**AAA**GCCACGGCGGATTCCGGC
A2696	rev	CATATCGCCGAAGCCCGTAATGCC

Table 3: Primers for mutations on the surface of the MPT domain and in the AT domain.
The MPT surface mutations were designed to influence the ACP–MPT interaction and the mutation in the AT domain for broadening the presumed binding channel transferring short acyl chains. The mutation site is indicated in bold typing, while the overlap between the primers is underlined.

Mutation	Direction	Sequence (5' – 3' display)
K1435A	fwd	CGAAT**GCG**ATTGGCATCCGCGCCGAGGACG
K1435A	rev	GCCAAT**CGC**ATTCGGGCGCAGCGCTGCCAG
R1494A	fwd	GCATG**GCC**TCATTTATCATGATCCCGGGCATCGACG
R1494A	rev	TAAATGA**GGC**CATGCCCGGCGCGCGCGAC
Q1647A	fwd	CTCGGC**GCA**ACCCTGCGCTTGCCGCAATACGC
Q1647A	rev	CAGGGT**TGC**GCCGAGCATATTCGCAAGCGTCGG
V289A	fwd	GATGGCGCAGAGGCA**GCC**TTGACGGAGATCGTGAACTGGCC
L290A	rev	GATGGCGCAGAGGCAGTC**GCA**ACGGAGATCGTGAACTGGCC
V289A L290A	fwd	GATGGCGCAGAGGCA**GCCGCA**ACGGAGATCGTGAACTGGCC
V289 L290	rev	TGCCTCTGCGCCATCGTGAGCC

For the coupled one-pot synthesis of a PK using two FAS modules, the second FAS was converted into a simple PKS. The KR domain knockout (Y2227F) was introduced to block the regular synthesis cycle. Also, an AT knockout (S126A) was used to avoid the formation of side products.

Table 4: Primers for the module 2 FAS in the coupled pathway for the production of a PK.
The mutation site is indicated in bold typing, while the overlap between the primers is underlined.

Mutation	Direction	Sequence (5' – 3' display)
S126A	fwd	GTCGCACACATTGGCCAT**GCC**CAAGGCGCGCTTGCTAC
S126A	rev	GTAGCAAGCGCGCCTTG**GGC**ATGGCCAATGTGTGCGAC
Y2227F	fwd	CGGCGGCGACGGTGCC**TTT**GGTGAGTCCAAGGCTGCC
Y2227F	rev	GGCAGCCTTGGACTCACC**AAA**GGCACCGTCGCCGCCG

3.1.3 Expression

For the expression (previously established in a different project[14]), both the plasmid carrying the *C. ammoniagenes* FAS construct and the plasmid of the *C. ammoniagenes* FAS AcpS were co-transformed into *E. coli* competent cells (BL21 (DE3) gold, Agilent Technologies, Santa Clara, USA) according to the manufacture's protocol, plated on LB+1.5% agar (with 50 µg mL^{-1} ampicillin, 11 µg mL^{-1} chloramphenicol, 0.01% arabinose as final concentrations) and grown over night at 37 °C. Five colonies was picked and grown in a preculture (LB media with 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol and 0.01% arabinose) at 37 °C over night. The preculture was transferred into the main culture of TB media (also containing 100 µg mL^{-1} ampicillin, 34 µg mL^{-1} chloramphenicol and 0.01% arabinose). It was shaken at 37 °C with 180 rpm until an OD_{600} of roughly 0.8 was reached. The culture was then cooled to 20 °C and shaken at this temperature for 30 min before IPTG addition (final concentration 250 mM). Cultivation was continued over night at 20 °C. The media was then centrifuged at 4,000 rcf for 20 min. The wet cell pellet weight was typically around 20 g L^{-1} of TB culture. If the cells were not further processed directly, they were shock frozen in liquid N_2 and stored at -80 °C.

3.1.4 Protein purification

For protein purification previously established methods[14] were slightly changed: buffer W (100 mM Na_2HPO_4/NaH_2PO_4, pH 7.2, 100 mM NaCl, 1 mM EDTA) was added to the cell pellet to a total volume of 35 mL. After addition of 2 mg of DNAse I (AppliChem, Darmstadt, Germany) and protease inhibitor (complete EDTA-free, Roche, Mannheim, Germany), the cells were resuspended by stirring. French press was used to break open cells (16 000 psi). Lysates were then kept at 4 °C to avoid protein degradation. Following centrifugation (60,000 rcf for 60 min), the supernatant (approx. 30 mL) was transferred on a strep column (5 mL resin volume, IBA, Göttingen, Germany). After 8 column volumes (CV) of washing, the protein was eluted with buffer E (3 CV, same as buffer W but with 2.5 mM D-desthiobiotin). After checking fractions for impurities by SDS-PAGE (Supplementary Fig. 3a), promising fractions were concentrated in a centrifugal filter (100,000 nominal molecular weight limit, Amicon Ultra-4, Merck Millipore, Darmstadt Germany).

In a next step, the proteins were purified by size exclusion chromatography (column: Superose 6 10/300GL, GE Healthcare, buffer G: 100 mM Na_2HPO_4/NaH_2PO_4, pH 7.2, 100 mM NaCl) and were checked for the desired hexameric oligomers (Supplementary Fig. 3b). Those fractions were again pooled using a centrifugal filter to a concentration of 10 – 20 mg mL^{-1} of protein. Potential loss of FMN, known from literature[32], was compensated by the addition of a five fold molecular excess of FMN. Proteins were then mixed with glycerol (final concentration 50%) and stored at – 20 °C. Glycerol does not have an effect on product distributions or activity as shown in Supplementary Fig. 3c and Supplementary Fig. 4.

3.1.5 Activity assay of regular FAS

The activity assays were performed as described before[1] with buffer AB (400 mM KH_2PO_4/K_2HPO_4 pH 7.3, 3.5 mM DTT), 50 nmol acetyl-CoA, 30 nmol NADPH and 25 µg of the FAS protein. After initial recording of the absorption at 334 nm for some time, 60 nmol malonyl-CoA were added to start the reaction. The corrected rate of NADPH absorption (measured rate minus rate prior to malonyl-CoA addition) was calculated and the activity determined. Supplementary Figure 4 can be seen as an exemplary result.

3.1.6 Activity assay of module 2 FAS (lactone producing FAS)

The activity assay of module 2 was previously described in a different variation[121]. Here it was performed on a 60 µL scale including 30 µL of buffer AB (see above), 40 nmol C_8-CoA and 30 µg of the FAS protein. The reaction was incubated for two minutes at room temperature. After two minutes of monitoring the absorption at 298 nm (absorption of lactone ring of 6-HHP without interference of any CoA absorption)[123], the reaction was started with the addition of 50 nmol malonyl-CoA. For a typical result, see Supplementary Figure 5.

3.1.7 Product assay for FA product distribution

Preparation was on a 100 µL scale including 90 µL of a master mix (consisting of 50 µL of buffer AB, described above; 20 nmol acetyl-CoA, 100 nmol malonyl-CoA, 225 nmol NADPH all dissolved in water). The rest was filled with water and protein solution (final content 20 µg of the FAS per reaction). The assay was let to react at room temperature for 4 hours. Then, 440 µL of acetone (cooled in ice bath) and the internal standards (isoC_5-CoA and C_{17}-CoA, final concentration 50 µM; in C_n-CoA, n represents the number of carbons in the acyl chain) were added and the mix was vortexed for 20 s. The assays were stored at –20 °C over night to precipitate proteins and salts. The mixture was then centrifuged at 0 °C for 5 min at 20,000 rcf. Directly afterwards, 500 µL of the supernatant were transferred to a new vial and evaporated to dryness under reduced pressure using a SpeedVac. The samples were then redissolved in 40 µL water by sonification for 5 min. After another centrifugation step (5 min at 20,000 rcf), they were transferred into sample vials and measured via HPLC-UV-MS. For selected constructs, also biological repeats were recorded, which is described in more detail in Supplementary Note 1.

Recovery rates for this sample processing were examined using the internal standards at three different concentrations (20 µM, 50 µM, 100 µM) and four replicates each. For iC_5-CoA recovery rates were (74.5±0.9)% for 20 µM, (74.8±1.4)% for 50 µM, and (73.1±1.7)% for 100 µM. For C_{17}-CoA, data was obtained at (70.1±1.8)% for 20 µM, (75.0±2.0)% for 50 µM, and (72.4±1.9)% for 100 µM respectively.

3.1.8 Product assay for module 2 FAS (lactone production)

For quantifying 6-HHP production, enzymes were freshly prepared, since glycerol led to unidentifiable side products. The coupled one-pot assays were prepared on a 100 µL scale including 50 µL of buffer AB (see above), 40 nmol acetyl-CoA, 200 nmol malonyl-CoA, 225 nmol NADPH, 20 µg of the FAS protein FAS (module 1) and 30 µg of the FAS protein FAS (module 2).

When different substrate concentrations were tested, the following concentrations were used: acetyl-CoA in 10 nmol, 20 nmol, 30 nmol and 40 nmol and of NADPH 75 nmol, 125 nmol, 175 nmol and 225 nmol, malonyl-CoA was kept constant. In combination, a total of 16 data points were examined. Information on technical and biological repeats can be found in Supplementary Note 1.

Assays were allowed to react over night. As internal standard 6-octyl-4-hydroxypyran-2-one was added (for synthesis see Supplementary Note 2). The sample work up, then followed the same procedure as for the product assays for FA product distribution (see above): After precipitating the protein with acetone, the supernatant was separated and evaporated to dryness. The residue was redissolved in water and measured in an HPLC.

The sample work up procedure was also analyzed for the lactone products. Three different concentrations (10 µM, 50 µM, 100 µM) with four replicates for each concentration were tested and the following recovery rates were obtained: (74.7±5.1)% for 10 µM, (76.7±6.1)% for 50 µM, and (73.1±2.8)% for 100 µM.

3.1.9 HPLC quantification

The acyl-CoA esters were analyzed on a HPLC (Dionex Ultimate 3000 RSLC) coupled to an ESI-TOF mass spectrometer (micrOTOF-Q II, Bruker, Bremen, Germany). For chromatographic separation an RP-18 column (100 x 2.1 mm, particle size 1.7 µm, Waters Acquity BEH, Milford, USA) was used with a mobile-phase system of buffer A (water, 10 mM triethylamine/acidic acid buffer, adjusted to pH 9.0) and buffer B (acetonitrile). A multistep gradient at a flow rate of 0.25 mL min^{-1} was used. The starting condition of buffer B was at 7%, followed by a linear increase to 60% until 6 min, then to 70% until 9.5 min and finally to 90% until 10 min runtime. Data were acquired in negative mode in a scan range from 200–2000 *m/z.* For quantification, the UV trace at 260 nm was analyzed using DataAnalysis 4.0 software

(Bruker, Bremen, Germany). For calibration C_6-CoA, C_8-CoA, C_{10}-CoA, C_{12}-CoA, C_{14}-CoA, C_{16}-CoA, C_{18}-CoA as well as the internal standards iC_5-CoA and C_{17}-CoA were measured. Calibration points were measured three times each for 10 pmol, 25 pmol, 50 pmol, 100 pmol, 200 pmol, 1000 pmol and 2000 pmol. Included in the calibration range were only average signals that did not differ from their calculated value more than 20%.

Losses in the sample work up procedure were corrected when analyte concentrations were calculated in reference to the internal standards where iC_5-CoA served as a reference for shorter CoA esters (C_6-CoA, C_8-CoA, and C_{10}-CoA) and C_{17}-CoA for the longer ones (C_{12}-CoA, C_{14}-CoA, C_{16}-CoA, and C_{18}-CoA).

A typical result for a product assay HPLC chromatogram can be seen in Supplementary Figure 6.

For analyzing the coupled assay and its lactone products, the same LC-MS system (Dionex Ultimate 3000 RSLC coupled to a Bruker micrOTOF-Q II) was used. The chromatography column was, however, different (RP-18 column, 100 x 2.0 mm, particle size 2.5 µm, Phenomenex Luna HST, Torrance, USA). The used buffers were water and acetonitrile (each containing 0.1% formic acid). The gradient (flow rate of 0.40 mL min^{-1}) started with ACN at 5% until 2 min, followed by a linear increase to 95% until 8 min runtime. MS data acquired in positive mode in the range from 100–1500 *m/z* was analyzed for identification of the compounds using DataAnalysis 4.0 software (Bruker Daltonik GmbH). For quantification, the UV trace at 298 nm was analyzed.

Calibration was performed in reference to the internal standard (6-octyl-4-hydroxypyran-2-one) and gave a response factor of 0.9825 ± 0.0293 for the calibration range 10–400 µM. Concentrations of the internal standard itself can be calculated from the obtained equation $y = 7.467\ x - 9.078$ (with R^2 of 0.9996). Several calibration points (10 µM, 20 µM, 81 µM, 100 µM, 200 µM and 400 µM) were recorded with three replicates each, which all met the criteria of differing less than 15% (20% for the lower limit of quantification) from their calculated value.

For a typical result of a lactone chromatogram, please refer to Supplementary Figure 7.

3.1.1 Molecular dynamics simulations and kinetic model of FAS

Both, molecular dynamics simulations and the kinetic model of FAS were developed by our cooperation partner, the group of Prof. Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen. Their description was added as Supplementary Note 3 (it does not reflect own work, but was taken from the joint publication[17] and added for clarification).

3.2 In vivo studies of short FA production in *S. cerevisiae*

The i*n vivo* studies on S. cerevisiae were conducted in cooperation with the group of Eckhard Boles from the biology department. Individual contributions are listed here and in chapter “7 Statement of personal contributions”.

3.2.1 Description yeast strain

The initial *S. cerevisiae ⊿fas1 ⊿fas2* strain (BY.PK1238_1A_KO, haploid) was constructed by Peter Kötter and Manuel Fischer (University of Frankfurt, Germany). The gene reading frames of *FAS1* and *FAS2* were exchanged to a kanMX4 cassette, which created a knockout of the type I FAS system and geneticin resistance. The genotype is *Matα; ura3Δ0; his3Δ0; leu2Δ0; TRP1; lys2Δ0; MET15; fas1::uptag-kanMX4-downtag; fas2::uptag-kanMX4-downtag.*

3.2.2 Vector description

Centromeric pRS shuttle vectors, pRS313 (with *HIS3* auxotrophy marker) and pRS315 (*LEU2* marker) with single copy number were used[24]. The plasmids carrying the wildtype were provided by Manuel Fischer (for more information on origins of plasmids, see Supplementary Table 1 and chapter “7 Statement of personal contributions”). On pRS315, *FAS1* or its mutated versions were located, while pRS313 carried *FAS2* or mutations, respectively. The genes were under control of their native promoter (995 bp upstream for *FAS1* and 480 bp upstream for *FAS2*)[25]. As terminator sequences 295 bp and 258 bp downstream of the open reading frames were used. Initial cloning was conducted in *E. coli* with the Infusion HD cloning kit (Clontech, Mountain View, USA). *FAS1* and *FAS2* genes were copied from *S. cerevisiae* genomic DNA and set into pRS vectors using *Bam*HI and *Sal*I restriction sites. The chromosomal coordinates (with promoter and terminator sequences) in the strain S288C are for *FAS1* (*YKL182w*): Chr XI 99676-107121 and

for *FAS2* (YPL231w): ChrXVI 108172-114573. Exemplary vector maps are shown in Supplementary Figure 8 and Supplementary Figure 9.

3.2.3 Primers

For the introduction of mutations by site directed mutagenesis, the primers are listed below in Table 5. They were designed as complimentary pairs with the mutation at a central position (PCR was then performed covering the whole vector). As alternative, the overlap of the primers was approx. only 15 bp.

Table 5: Primers for the introduction of point mutations into *S. cerevisiae* FAS.
The mutation site is indicated in bold typing, while the overlap between the primers is underlined.

Mutation	Direction	Sequence (5' – 3' display)
I306A	fwd	TTCTTC**GCT**GGTGTTCGTTGTTACGAAGCATACCCAAACACTTCC
I306A	rev	ACACC**AGC**GAAGAATAATACAGTAATTGCTTTTCTTACGGAGACG
R1834K	fwd	AGTTGTGTTCTAC**AAA**GGTATGACTATGCAAGTTGCTGTTCC
R1834K	rev	CATAGTCATACC**TTT**GTAGAACACAACTTCAACTAAAGATTCGATAGAC
G1250S	fwd	TCTGGTTCT**TCT**ATGGGTGGTGTTTCTGCCTTACG
G1250S	rev	CAT**AGA**AGAACCAGAACAGTTACCAACCTCAGAAACATGTACG
G1250S M1251W	fwd	TCTGGTTCT**TCTTGG**GGTGGTGTTTCTGCCTTACG
G1250S M1251W	rev	**CCAAGA**AGAACCAGAACAGTTACCAACCTCAGAAACATGTACG
F1279Y	fwd	ATTTTACAAGAATCA**TAT**ATCAACACCATGTCCGC
F1279	rev	TGATTCTTGTAAAATATCATTTTGGACAGGC

3.2.4 Transformation

Transformations were conducted according to a modified lithium acetate protocol[124]. Plasmids of *FAS1* and *FAS2* were co-transformed using 1 µg of DNA for each construct. A pre-culture culture (5 mL) of the *Δfas1 Δfas2* strain was grown over night in YPD (1% yeast extract, 2% peptone, both produced by BD, Difco Laboratories, Sparks, USA; 2% dextrose, purchased from Roth, Karlsruhe, Germany) containing 200 µg/mL geneticin disulfate, free FA (myristic, palmitic and stearic acid, each 50 µg/mL) and 1% Tween20. The main culture (same composition as the pre-culture) was inoculated with the pre-culture and shaken at 30 °C and 200 rpm until

OD_{600} = 0.8 was reached. Then, 5 mL of this culture were centrifuged at 3000 rcf for 5 min followed by a washing step (resuspension of cells in 1 mL water and centrifugation). The cells were resuspended in lithium acetate solution (0.1 M) and let to rest for 5 min at room temperature. Following a short round of centrifugation (5000 rcf, 15 s), 240 µL PEG 1,500 solution (50%), 76 µL water, 36 µL lithium acetate solution (1.0 M), 5.0 µL single stranded DNA solution from salmon testis (10 mg/mL), and 2 µg of each plasmid were added. The mix was shortly vortexed and incubated for 30 min at 30 °C and again for 20 min at 42 °C. The mixture was centrifuged again (4000 rcf, 15 s, 4 °C) followed by another washing step (resuspension in 1 mL water and centrifugation at 4000 rcf for 15 s at 4 °C). Finally, 100 µL water were added and the cell suspension was plated on SCD –his –leu agar plates (also containing 200 µg/mL geneticin disulfate, free FA (myristic, palmitic and stearic acid, each 50 µg/mL) and 1% Tween20).

3.2.5 Cultures for product analysis

For product analyses, typically YPD medium (1% yeast extract, 2% peptone, both produced by BD, Difco Laboratories, Sparks, USA; 2% dextrose, purchased from Roth, Karlsruhe, Germany) was used to culture *S. cerevisiae* strains, if not clearly stated otherwise. Several colonies of the *S. cerevisiae* strains were always combined in a single pre-culture (5 ml YPD with 200 µg/mL geneticin disulfate, 50 mg/ml final concentration). The cultures were shaken at 200 rpm and 30 °C over night and the OD_{600} was noted. For the main culture, part of the pre-culture was transferred into fresh medium (50 ml YPD with 200 µg/mL geneticin disulfate, 50 mg/ml final concentration) to reach an OD_{600} of 0.1. The main cultures were then shaken for 48 h (200 rpm, 30 °C). The final OD_{600} was measured using a Ultrospec 2100 pro spectrophotometer (GE Healthcare, USA). Measurements were conducted by Renata Pavlovic (group of Prof. Eckhard Boles).

The particular samples with long FA supplementation contained additionally $C_{18:1}$ (final concentration of 1 mM) and Tergitol NP-40 (solution in water, 70%; final concentration 1%).

3.2.6 Sample processing

FA extraction was based a modified previously published protocol[30]: After centrifugation of the culture at 3,500 rcf for 15 min, the supernatant was divided into

10 ml portions where 0.2 mg of the internal standard, heptanoic acid (C_7; solved in methanol/chloroform 1:3), were added. The samples were acidified with 1 ml HCl (1 M), followed by the addition of 2.5 ml of a mixture of equal amounts of methanol and chloroform. After shaking the samples vigorously for 5 min, they were centrifuged at 3,500 rcf for 10 min. The chloroform phase was separated and last water were taken out. Under reduced pressure using a SpeedVac, the solvent were completely removed. The following methylation was conducted according to a previously published protocol[125].

Sample processing was conducted by both Jan Gajewski and Renata Pavlovic (group of Prof. Eckhard Boles).

3.2.7 Determination of free FA by gas chromatography (GC)

These measurements were conducted by Renata Pavlovic (group of Prof. Eckhard Boles). For the quantification of FA, the FA methyl esters were first taken up in hexane and then analyzed in a Perkin Elmer Clarus 400 gas chromatograph (Perkin Elmer, Rodgau, Germany) equipped with an Elite FFAP capillary column (30 m × 0.25 mm, film thickness: 0.25 µm; Perkin Elmer, Rodgau, Germany) and a flame ionization detector (Perkin Elmer, Rodgau, Germany). Helium was used as carrier gas. Of the sample, 1 µL was ken up into split injection (10 mL/min). Injector and detector temperatures were set to 200 °C and 250 °C, respectively. Separation followed the following temperature program: 50 °C for 5 min, followed by an increase of 10 °C/min to 120 °C (hold for 5 min), another increase of 15 °C/min to 180 °C (hold for 10 min), and a final increase of 20 °C/min to 220 °C for 37 min.

Three 10 mL aliquots of each culture were processed and analyzed separately. Between these three repeats, a standard deviation below 1 mg/L was typically reached (in 92% of the measured concentrations).

3.2.8 Metabolite analysis by HPLC

Also here, measurements were conducted by Renata Pavlovic (group of Prof. Eckhard Boles). Samples for glucose and ethanol quantification were prepared as followed: To 450 µL of the cell-free supernatant of the culture, 50 µL of 50% (w/v) 5-sulfosalicylic acid were added and vigorously shaken. After centrifugation (13,000 rcf, 5 min, 4 °C), the supernatant was separated and injected into an UHPLC+ system (Dionex UltiMate 3000, Thermo Scientific, Dreieich, Germany)

equipped with a HyperREZ XP Carbohydrate H+ 8 µm column (heated to 65 °C) and a refractive index detector (Thermo Shodex RI-101). As a buffer water with 5 mM sulfuric acid was used. The flow rate was set to 0.6 mL/min. For calibration, five standards (mixtures of D-glucose, ethanol, glycerol and acetate with concentrations of 0.05-2% (w/v)) were used.

4 Results

In this thesis, different aspects of manipulating reaction control in type I FAS were investigated. The strategy was to first test engineered FAS constructs in a low-complex environment *in vitro*. Therein, our research was focused on chain length control, but not exclusively. Promising modifications were then also tested in a more complex *in vivo* system where *S. cerevisiae* served as a model organism.

The engineering itself aimed at rationally planned, single mutations that were set in key catalytic domains. In the following subchapters, the results of the different engineering aspects of type I FAS will be presented.

The texts in this results section (and the corresponding Supplementary Information) were in parts prepared for manuscripts intended for publication (for individual contribution to the writing process, please consult the chapter “7 Statement of personal contributions”). Also, the very initial work is based on internal group projects. Its contributions are explained in detail in the mentioned statement.

4.1 Engineering binding channels for chain length control

For our first study on rational engineering of type I FAS and their *in vitro* testing (leading to the publication[17] and a patent[18]), a FAS was from *Corynebacterium ammoniagenes* (Uniprot Identifier: D5NXL2) was chosen. This FAS had some preferable characteristics: Expression can be done in *E. coli* to gain high amounts of active protein, and a quick and easy purification via a strep-tactin affinity chromatography and size exclusion chromatography was established which typically yielded 3–18 mg (protein)/L of TB culture (also see Supplementary Fig. 3a,c).

In the *C. ammoniagenes* FAS, the active centers are distributed on one polypeptide chain with 3016 amino acids with a total weight of 317 kDa[32, 126-127]. The protein further assembles to its active oligomeric state of a 1.9 MDa hexamer. The similarity to the *S. cerevisiae* FAS is striking as the *C. ammoniagenes* FAS can be seen as a fused *S. cerevisiae* FAS where its α-chain is attached to the end of the β-chain (see also Fig. 2)[15]. The longer sequences in *S. cerevisiae* FAS can be accounted to a bigger share of structural elements and the directly attached PPT domain, which are not present in bacterial type I FAS (and thus also not in *C. ammoniagenes* FAS)[38, 41]

For *S. cerevisiae* FAS there is also a series of structural data available[10-13]. The data was used as a basis for engineering *C. ammoniagenes* FAS which shares

substantial sequence identity with *S. cerevisiae* FAS, especially within individual domains e.g. 34.1% in the KS and 40.4% in the MPT (domain boarders set according to structure, PDB code: 2VKZ[12]). This procedure was also chosen since *S. cerevisiae* FAS could not be produced as easily for *in vitro* probing in the desired amounts.

For the initial *in vitro* study on manipulating chain length control in type I FAS on the example of *C. ammoniagenes* FAS, the engineering strategy was based on the following aspects:

As our starting point, we assumed that the kinetics of the domains KS, AT and MPT are the main determinants of FA chain length control. It has been reported before that the acyl-CoA output spectrum of FAS is sensitive to relative concentrations of the priming acetyl-CoA and the elongating malonyl-CoA[83]. Increased acetyl-CoA concentrations promote priming of synthesis, shifting the product spectrum towards shorter acyl-CoA esters, while increased malonyl-CoA concentrations promote elongation, shifting the spectrum towards longer products. Accordingly, we posited that we can influence chain length by modulating substrate binding affinities of AT and MPT. Further, the KS and MPT domain act directly on chain length control, insofar as the relative rates of the respective reactions determine whether the growing ACP bound acyl-chain is loaded in the KS, initiating another cycle of elongation, or loaded in the MPT, leading to product export. An early work has pointed out the interplay of these domains and suggested a simple numeric model in which chain length is determined by relative velocities of the condensation and the transferase reaction[83].

Our engineering efforts were supported by the group of Prof. Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen. They provided molecular dynamics simulations and a computational kinetic model of FAS in two versions (initial and a revised), which will be presented at the appropriate engineered domain.

4.1.1 Engineering the KS domain

As the MPT domain has been reported to be tolerant in binding of acyl chains of several chain lengths[95-96], we initially focused on achieving chain length control by engineering KS only. From the X-ray structural data of *S. cerevisiae* FAS, M1251, positioned centrally in the KS channel, has been proposed to act as a “gatekeeper”

to the KS binding channel by essentially populating two conformations; directed into the KS channel (closed state), leaving the outer part of the channel accessible for substrate binding, and rotated aside (open state), giving access also the distal portion of the KS channel (conformations as shown in Fig. 6)[12].

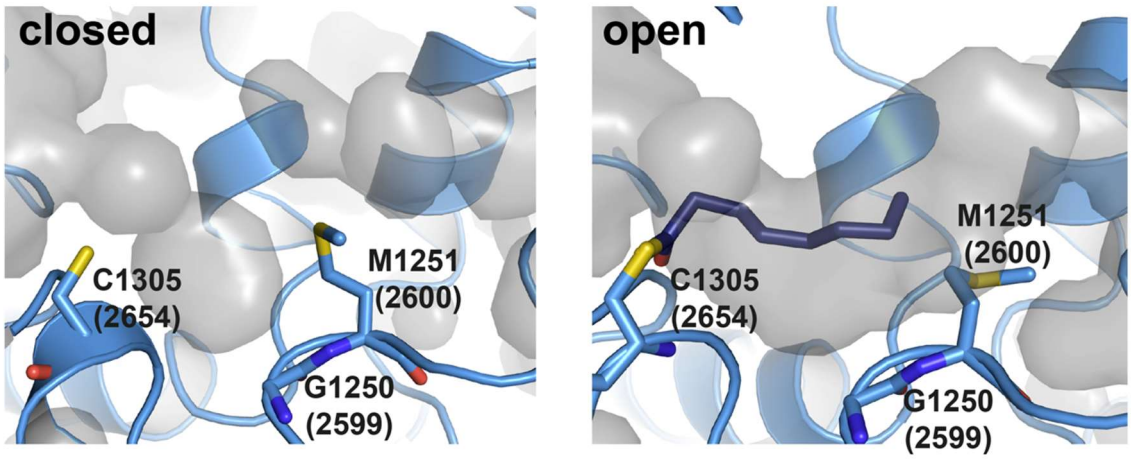


Figure 6: KS domain conformations.
In structural data of *S. cerevisiae* FAS an open and closed conformation was observed (*C. ammoniagenes* FAS numbering in parentheses). In the "closed" state (PDB code 2UV8)[10], M1251 is pointing into the KS channel, which makes only the outer part accessible (the binding channel is indicated as a gray cloud). The "open" state (PDB code 2VKZ)[12], is induced by the inhibitor cerulenin mimicking a bound acyl chain. Here, the central M1251 is pushed aside making the whole KS binding channel accessible. The neighboring mutation G1250S was believed to restrict M1251 further in its flexibility and freezing the "closed" state.

In human mitochondrial KS, homologous to the yeast FAS KS domain, the equivalent methionine was linked to a bimodal product distribution[16]. Presumably as the result of an energy penalty for rearrangement of the methionine, weak binding to human mitochondrial KS was found specifically for C_8-ACP, and was suggested to lead to increased production of C_6- as well as C_{10}- and C_{12}-products[128]. Direct evidence for the potential role of M1251 in modulating chain length control in yeast FAS was provided by *S. cerevisiae* strains carrying a G1250S mutation. Putatively on the same basis as for human mitochondrial KS, restricted conformational flexibility of the adjacent M1251 leads to increased production of short FA products[8-9]. Based on these previously reported findings, we sought to build upon the "gatekeeper" function of the M1251 equivalent M2600 of *C. ammoniagenes* FAS, and to this end constructed the *C. ammoniagenes* FAS mutants G2599S (FAS^{G2599S}) and G2599S-M2600W ($FAS^{G2599S\text{-}M2600W}$). In addition to the reported effect of the glycine to serine

mutation, we reasoned that a bulkier tryptophan residue in place of methionine would be expected to counteract more strongly the binding of acyl intermediates C_8 and longer, while still permitting the binding of C_6 and shorter. Consistent with this line of reasoning, our data show that these mutants indeed increased production of shorter FA (Fig. 7a,b and Supplementary Fig. 10).

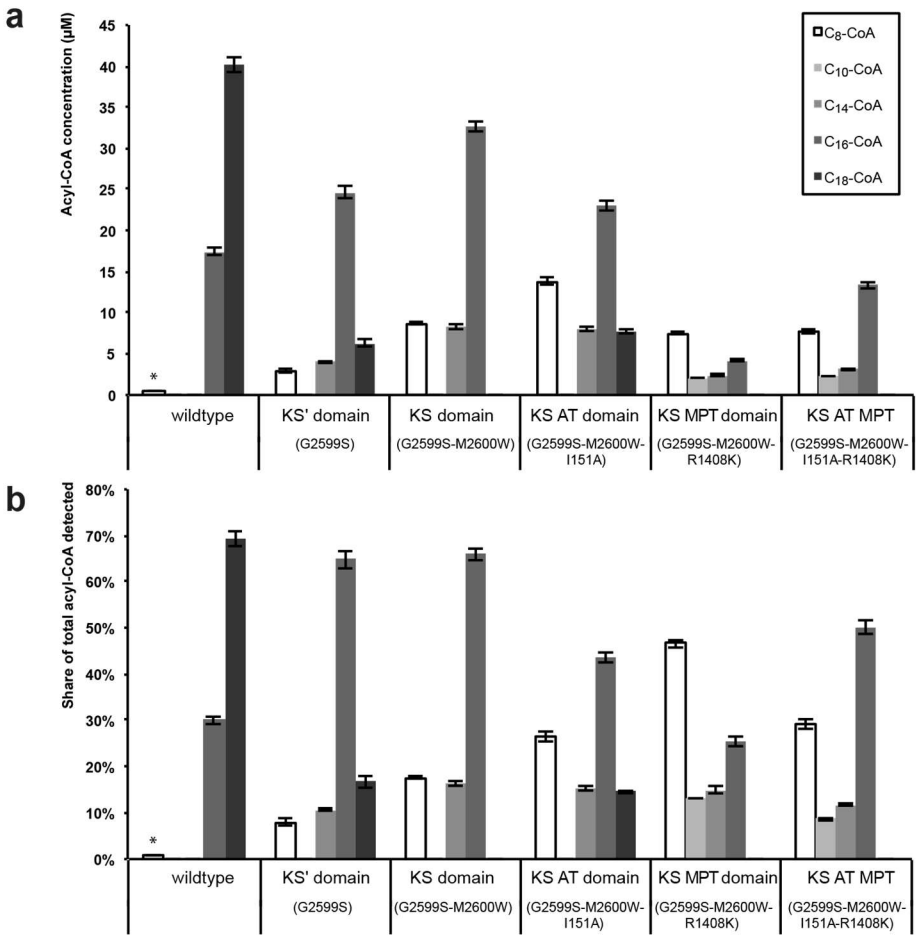


Figure 7: Product distributions of engineered FAS.
(a) Experimentally found concentrations of acyl-CoAs of specific chain lengths. FAS variants are abstracted by the domains carrying mutations while specific mutations are given in parentheses. Bars marked with (*) are based on values received by extrapolation (values outside of calibration range). **(b)** Percentage of molar concentrations of specific acyl-CoA in total acyl-CoAs found in the experiments.

Subsequently, we performed atomistic modeling and free energy calculations (Supplementary Note 3) on structural data of *S. cerevisiae* FAS (data for *C. ammoniagenes* FAS was not available). The equivalent combination to the *C. ammoniagenes* FAS$^{G2599S\text{-}M2600W}$ was evaluated in *S. cerevisiae* (where it corresponds to G1250S-M1251W) to evaluate the “gatekeeper” hypothesis, which predicts that the mutation would counteract C_8 binding (this work has been

conducted by the collaboration partner of the group of Prof. Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen). Against our expectations, free energy calculations indicated a pronounced effect in the opposite direction: the double KS mutation was calculated to enhance the binding of C_8, by 10.0 kJ/mol ± 4.4 kJ/mol (Fig. 8a), contradicting the "gatekeeper" rationale as formulated above. In support of this finding, a structural model of the $KS^{G1250S-M1251W}$:C_8 complex from atomistic molecular dynamics simulation (Fig. 8b) shows favorable close hydrophobic / van der Waals contacts of the acyl chain tail with the indole ring gatekeeping tryptophan, and no apparent displacement of its side chain.

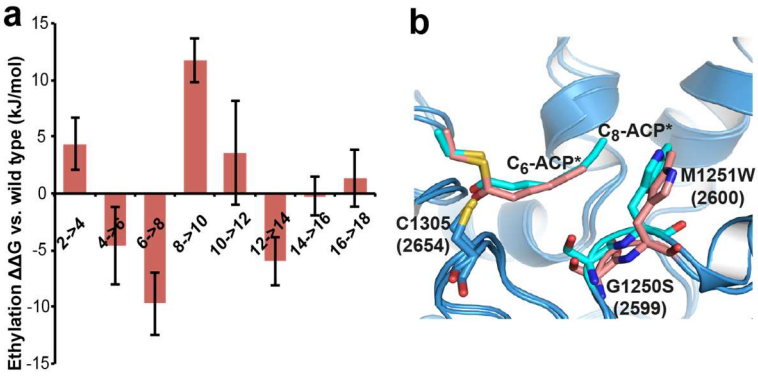


Figure 8: Results from molecular dynamics simulations of the KS domain.
(a) Binding free energy change for acyl chains C_4-C_{16} in the mutant $FAS^{G1250S-M1251W}$ with respect to wild type FAS, as calculated from molecular dynamics simulations. Alchemical free energy calculations estimating the free energy change associated with successive ethylation for both the wild type and mutant protein. Negative values indicate more favorable surroundings for the introduction of respective ethyl groups in the mutant protein compared to wild type. The key extension step C_6->C_8 (discussed in the main text) appears to be energetically favored in the mutant protein. **(b)** Snapshots in molecular dynamic simulations of C_6- and C_8-ACP binding to the modified KS domain. For molecular dynamics simulation, the modified KS domain of $FAS^{G2599S-M2600W}$ with bound acyl chains of different lengths was generated from the X-ray crystal structure of the *S. cerevisiae* FAS (PDB code 2UV8)[10]. Important residues are indicated in stick representation; active residue C1305 (C2654 in *C. ammoniagenes*), and mutated sites G1250S (G2599S) and M1251W (M2600W). The acyl-ACP are depicted in truncated forms (indicated by (*)) in the colors of the respective KS domain. Favorable interactions between the acyl-tail of octanoyl-ACP (C_8, peach) and the tryptophan are possible by their close proximity.

As FA synthesis in FAS is however a complex network of reactions, we didn't want examine isolated domains only and then assume effects on chain length control. Subtle modulation of several enzyme activities, their interplay as well as the relative availability of substrates dictate the final product distribution. To account for the expected complex interdependent effects and to guide the selection of mutations, protein engineering was combined with quantitative *in silico* modeling (programming

was conducted by the cooperation partner, the group of Prof. Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen; for further contribution to projects, see the statement at the end of this thesis, section “7 Statement of personal contributions”). Similarly as for the atomistic modeling described above, also the kinetic model was set-up with type I FAS of *Saccharomyces cerevisiae,* for which a wealth of biochemical and structural data have been reported, and which is highly homologous to *C. ammoniagenes* FAS in its active sites[4, 10, 12-13, 15].

The quantitative model represents a simplified canonical catalytic cycle consisting of 51 distinct reactions, which we explored by stochastic simulation (Fig. 9 and Supplementary Table 3)[129]. With only a few of the corresponding 102 forward and reverse rate constants directly available from experiment, we proceeded to restrict the parameter space of possible models by enforcing agreement with product spectra measured for yeast FAS, incorporating a further 30 experimental data points gathered under multiple sets of experimental conditions. From the large number of candidate parameter sets, an ensemble of the 1000 sets best matching the input data (hereafter referred to collectively as ‘the kinetic model’), was retained as a representative sample of the remaining parameter space, and used for the subsequent analysis and predictions.

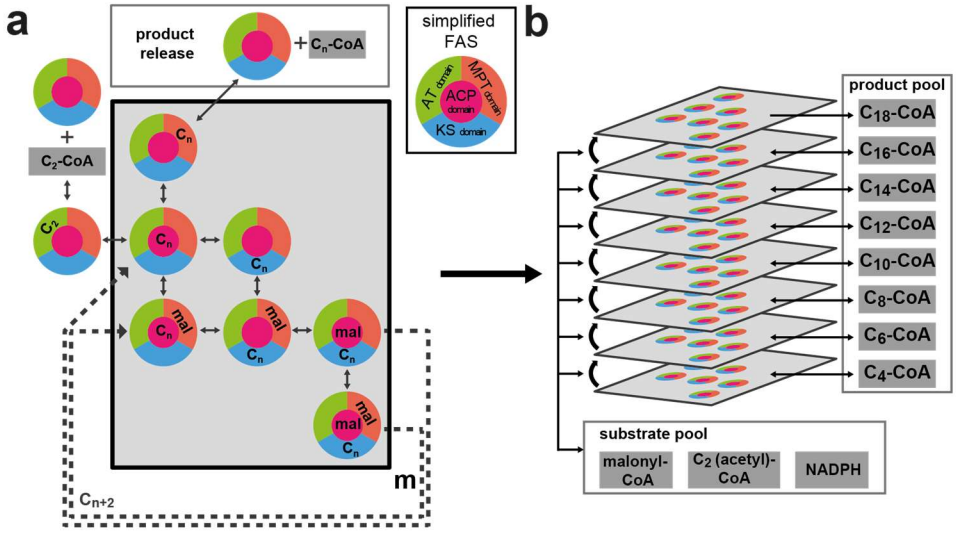


Figure 9: Development of a computational model for *in silico* FAS simulation.
(a) Schematic representation of the computational model as composed of a set of functionally equivalent layers. In the model, processing of an acyl intermediate of chain length C_n includes initiation, elongation and product release as well as ACP-mediated shuttling, while additional processing steps are neglected. Loading with acetyl-CoA (C_2-CoA) on the left is the first step, followed by stochastically changing loading states. When the ACP is occupied by malonyl (mal) and the KS

domain is loaded with an acyl chain (C_n), the latter can be elongated. A layer in the simulation starts where previously known occupancies reoccur (indicated by dashed lines) but with the elongated acyl chain (C_{n+2}). For clarity, not all possible permutations of site occupancy are explicitly shown. **(b)** Abstracted reaction space describing chain elongations from C_4 to C_{18}. FA synthesis is described by a stack of layers each representing the elongation of an acyl intermediate of chain length C_n.

We subsequently sought to represent the substantially more favorable KS:C_8 binding free energy that we found in our molecular dynamics simulations in the kinetic model. Specifically, we assumed that the substantially reduced free energy of the bound state would not yield a reduction in the forward rate of C_8 transfer to the mutant KS site, and tested combinations of kinetic coefficients selected in accordance with the calculated free energy change. No tested combination of parameters showed the experimentally observed property of increased C_8-CoA release, indicating that the computational model, as formulated, was incompatible with the experimental data from our engineering efforts.

After the results from the initial kinetic model did not reproduce our experimental findings, we revisited the mechanism of how a higher affinity of C_8 to the G2559S-M2600W-mutated KS domain could possible still lead to more short FA.
Underlying a first hypothesis, we note that a prominent side reaction of FAS/PKS protein family is the unproductive KS-mediated decarboxylation, without subsequent condensation of ACP-bound malonyl. For FAS, this side reaction has previously been reported as occurring at a negligible basal rate under normal conditions, but at a considerably higher rate when key catalytic residues are blocked through covalent modification[130]. We therefore hypothesized that an unusually strongly KS-bound C_8 intermediate could approximate a stalled state and induce decarboxylation, conceivably leading to elevated levels of C_8-CoA product since 1) C_8 is then more likely to be cleft off instead of undergoing elongation and 2) the decarboxylation of malonyl yields acetyl which itself is the starter unit, leading to more short FA *in vitro*[83, 93].
To test this alternate hypothesis *in silico*, we asked whether a kinetic model extended by the reaction KS:C_8 + ACP:malonyl → KS:C_8 + ACP:acetyl + CO_2 might account for the increase in C_8-CoA release. Although we were able to identify members of our model ensemble for which the introduction of the decarboxylation reaction

yielded an increase in C_8-CoA output, the minimum reaction rate for the decarboxylation reaction was at least an order of magnitude faster than the rate reported[44], thus appearing to contradict this interpretation.

We concluded our exploration of this hypothesis with further experimental measurements. *In vitro* reaction with ^{13}C-labelled acetyl-CoA (but unlabeled malonyl-CoA) did not show priming of FA synthesis with unlabeled acetyl, as would be expected in the case that significant quantities of unlabeled malonyl-CoA were undergoing decarboxylation in the KS-engineered FAS (Supplementary Fig. 11).

Following a second hypothesis of overly restrictive parameters, we eventually changed the model restrictions once again to gain a revised kinetic model: Relaxing our initial assumption that the calculated more favorable binding free energy for C_8 in the mutated KS substrate tunnel would not yield a reduction in the forward rate of the associated binding reaction, we identified 100 members of the 1000-member model ensemble for which a slowdown in the reaction ACP:C_8+KS → ACP+KS:C_8 by less than a factor of ten yielded an increase of at least 5% in C_8-CoA output. We note that in these model scenarios, the reduced forward rates for this reaction are compensated by greater slowdowns in the reverse rate of the same reaction, in order to enforce agreement with the calculated negative binding free energy change for the KS:C_8 complex (Fig. 10).

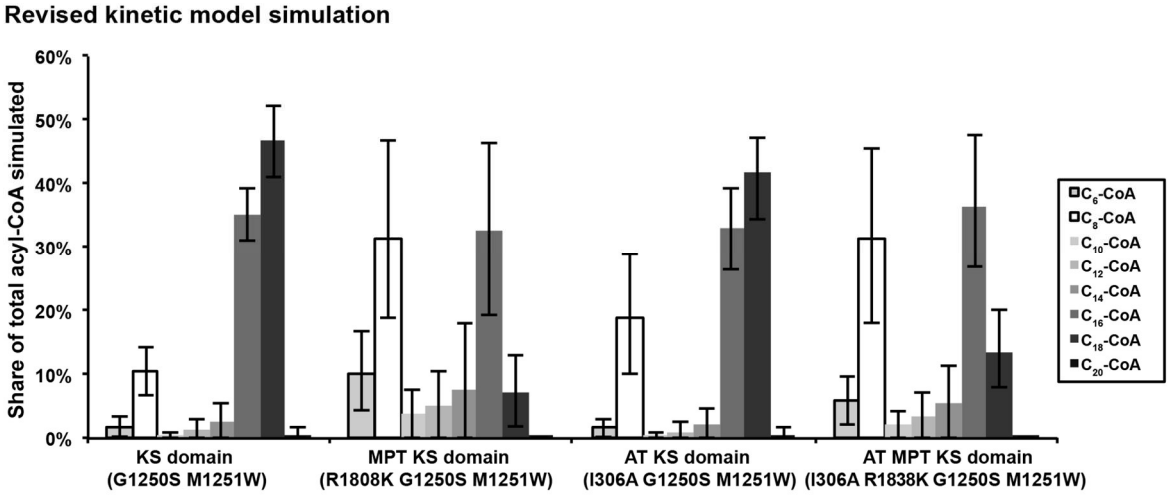


Figure 10: Computational output spectra of KS and combined transferase mutants.
FAS variants are abstracted by the domains carrying mutations. Percentages of molar concentrations of specific acyl-CoA in total acyl-CoAs found in the experiments are shown. Simulation of product spectra with the kinetic model for the KS domain mutant (G1250S M1251W, *S. cerevisiae* FAS

numbering), MPT KS domain mutant (R1834K G1250S M1251W, *S. cerevisiae* FAS), the AT KS domain mutant (I306A G1250S M1251W, *S. cerevisiae* FAS) and the AT MPT KS domain mutant (I306A R1808K G1250S M1251W, *S. cerevisiae* FAS).

We note that the agreement between experiment and model for the G1250S-M1251W combination arises by construction, since we selected only those parameter sets that matched the known C_8-enriched *C. ammoniagenes* product spectrum.

In further experiments, the enzyme activity, i.e. their maximum activity under substrate excess (v_{max}), was determined by quantifying the NADPH consumption of the reducing catalytic domains (KR and ER) as an indicator for overall FA synthesis[1]. Results for selected constructs are shown in Table 6 (mutations occurring in the table, that have not yet been presented, will be described later in their corresponding section).

Table 6: Enzymatic activities of FAS constructs.
Data were collected on different protein preparations (expression batches) as indicated. The sample means (SEM) are given with standard deviation of the average values of the three batches. When three or more values have been recorded for one batch, the standard deviation is also indicated. Activities are shown in mU/mg of protein (1 U was defined as the incorporation of 1 µmol of malonyl-CoA per minute). Activity was determined by monitoring NADPH consumption.

	Wild type	$FAS^{G2599S-M2600W}$	$FAS^{G2599S-M2600W-R1408K}$	$FAS^{G2599S-M2600W-I151A-R1408K}$
Batch 1	362 ± 55	158	61	43
2	210 ± 68	140		86
3	284 ± 28	109		
4	383 ± 31	51 ± 13		
SEM	310 ± 79	114	61	64

4.1.1 Engineering the transferase domains

In spite of the uncertainty as to the molecular basis underlying the results of our KS engineering, the experimental success of shifting the output spectra by engineering the KS-domain prompted further engineering for improving C_8-CoA production. As a second step of engineering, we focused on transferase domains AT and MPT. We first aimed at engineering the MPT domain towards higher C_8-CoA affinities, with the intention of facilitating its release. Eight candidate mutations, informed by molecular modeling of palmitoyl binding in the MPT domain of the *S. cerevisiae* FAS (molecular modeling performed by the group of Helmut Grubmüller), were generated at the equivalent positions in *C. ammoniagenes* FAS, with none of those constructs showing desired effects *in vitro* (Fig. 11a,b).

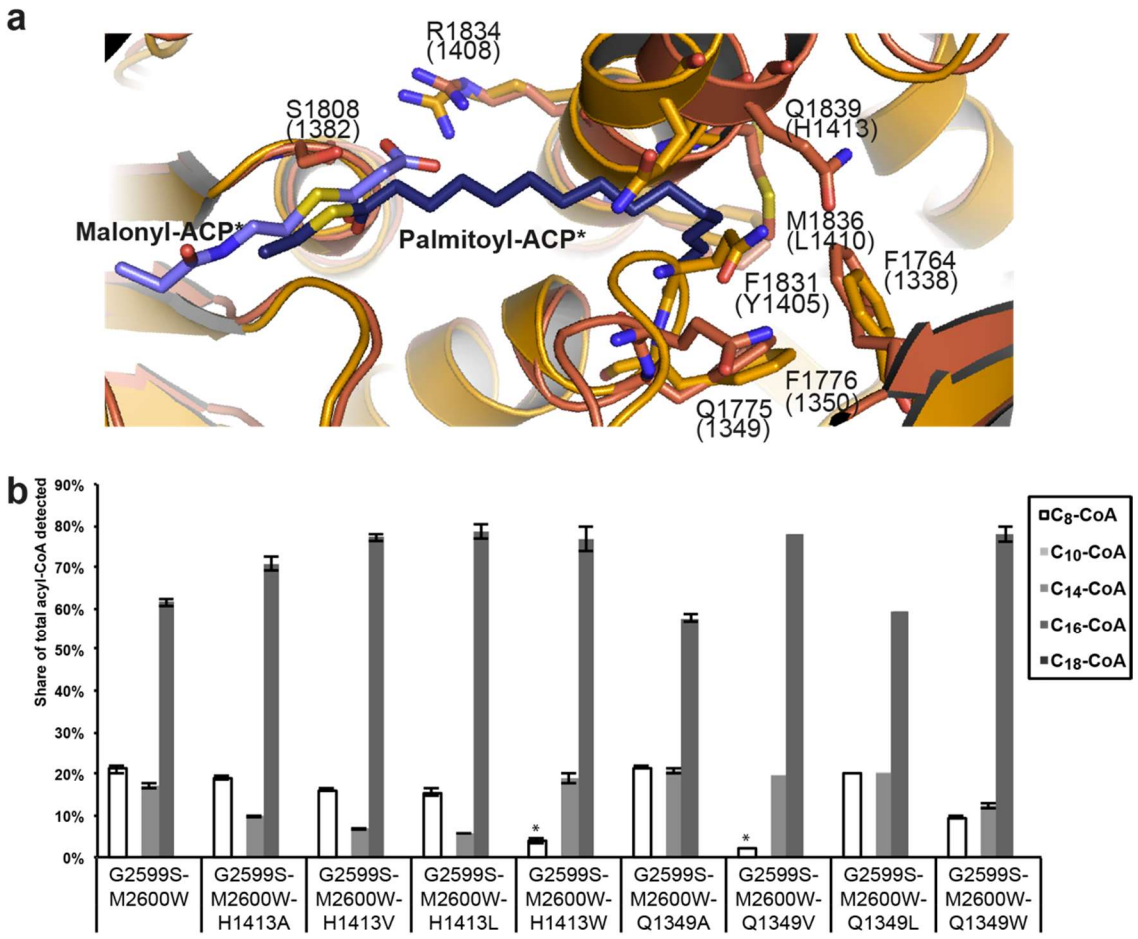


Figure 11: MPT domain engineering to enhance short FA cleave off.
(a) Molecular dynamics simulations of the *S. cerevisiae* FAS MPT domain (based on PDB code 2UV8[10]) were made for the binding of palmitoyl (C_{16}; ligand in dark blue, structure in red) and malonyl (ligand in light blue, structure in orange). Important residues are shown in stick representation, such as the active site serine (S1808, *S. cerevisiae* numbering; *C. ammoniagenes* numbering in the figure in brackets), the arginine involved in malonyl stabilization (R1834), the hydrophobic pocket (F1764, F1776, F1831, M1836) as postulated in literature[131] and two mutation sites that were tested to enhance short fatty acid release, Q1775 and Q1839 **(b)** Product spectra of further FAS variants mutated in MPT. Constructs were generated for higher C_8-CoA output by increasing its affinity towards the MPT after simulating how longer chains activate MPT transfer. Yields are displayed as percentage of molar concentration of acyl-CoA species in total acyl-CoAs detected, and reveal only inferior or similar C_8-CoA yield as the $FAS^{G2599S–M2600W}$ reference. Assays were analyzed using a different HPLC buffer system (buffer A: water with 0.05% triethylamine, buffer B: water/ACN 10:90 with 0.05% triethylamine).

In a complementary approach, we aimed to promote short acyl production by decreasing malonyl affinity[22-23]. An arginine was identified as being important for coordinating the free carboxyl group of malonyl (see Fig. 11a). The MPT domain mediates the dual functions of loading malonyl and releasing the acyl product by binding competitively to the same serine residue in the active site. Accordingly, a

decrease in malonyl affinity can be expected to decrease the rate of malonyl loading, and, as an additional desirable effect for our purpose, to lessen competition for acyl product release. Tested in conjunction with the KS domain mutations, the R1408K-modified *C. ammoniagenes* FAS ($FAS^{G2599S-M2600W-R1408K}$) yielded C_8-CoA corresponding to 47% of the total acyl-CoA products detected. Data suggest an increase in C_8-CoA and C_{10}-CoA at the expense of C_{16}-CoA (Fig. 7a,b). For mutations R1408A, R1408G and R1408Q FA synthetic activity was abolished which was confirmed HPLC measurements.

The biochemical experiments findings were also in close cooperation with the molecular dynamic modeling. Representing the equivalent R1834K mutation in the *S. cerevisiae* FAS-based theoretical model, we calculated weakened binding of malonyl-CoA in the MPT site by $\Delta\Delta G = +14.2 \pm 1.1$ kJ/mol. We sought to validate the initial kinetic model against this first transferase mutation as well. Consistent with our measurements on *C. ammoniagenes* FAS product spectra, the first kinetic model responded to the incorporation of the calculated effect with a broad shift towards shorter products (Supplementary Fig. 12). Intriguingly, this single perturbation led to a greater relative increase in the release of C_8-CoA than for the neighboring C_6 and C_{10} short chain products.

We also evaluated the MPT mutation in combination with the updated KS-mutated model (Fig. 10), since our representations of the former are based only on calculated thermodynamic quantities. Combining MPT mutation R1834K with the G1250S/M1251W double KS mutation, we observe increased C_8-CoA output, to 31±13%, compared to 46.6% observed experimentally. As with the MPT mutation in isolation, the model reproduced the general shift towards shorter products, including a considerable reduction in C_{18}-CoA release.

In a further step, we introduced the mutation I151A in the AT domain (additional to the KS domain mutations; $FAS^{G2599S-M2600W-I151A}$). This selection was guided by the same studies on the human MAT-transferase[22-23], where this mutation was suggestive of increased acetyl throughput, matching our aim of shifting the balance towards shorter products. Furthermore, this mutation was postulated to form a novel binding channel that can accommodate acyl chains with a length of up to C_{10},

potentially introducing a product-exporting acyl transferase activity for short-length acyl chains[22-23]. In measured *C. ammoniagenes* FAS product spectra (for $FAS^{G2599S-M2600W-I151A}$), the AT domain mutation I151A led to higher absolute and relative yields of C_8-CoA (see Fig. 7a,b). This increase came at the expense of the proportion of C_{16}-CoA; output of the longest observed product, C_{18}-CoA, was increased in comparison with the KS-mutated spectrum.

For the equivalent *S. cerevisiae* FAS mutation I306A in the theoretical model, a calculated ΔΔG value of -0.4 ± 1.6 kJ/mol indicates little to no effect on acetyl-CoA binding affinity. This speaks against the first hypothesis, that C_8-CoA output is enhanced by virtue of a favored loading of acetyl-CoA. We probed the second hypothesis, which proposes a novel short-acyl binding channel, with atomistic simulations of C_8-CoA in complex with equivalent *S. cerevisiae* mutant I306A. C_8-CoA could not be plausibly fitted in alignment with the catalytic residues in the wild type active site, but was accommodated in the I306A mutant with minor local rearrangement (Fig. 12). The resulting structure was used as a starting point for ten unbiased molecular dynamics simulations, each exceeding 400 ns in length. The simulations maintained a plausible pre-catalytic conformation, in alignment with the catalytic serine side chain and the oxyanion hole, in nine cases, suggesting a complex that is stable at least in the microsecond range.

Taken together, these results favor the hypothesis that I306A-mutated *S. cerevisiae* AT is a competent acyl transferase in this acyl chain length range as postulated in the publication[23]. Introducing the possibility of C_8 exit through the AT domain into the initial kinetic model with a forward rate of 3 s^{-1} for both the transfer of C_8 from ACP to AT domain and its exit as CoA ester, led to increased release of C_8-CoA product output in both the initial and the revised model (see Fig. 10 and Supplementary Fig. 12).

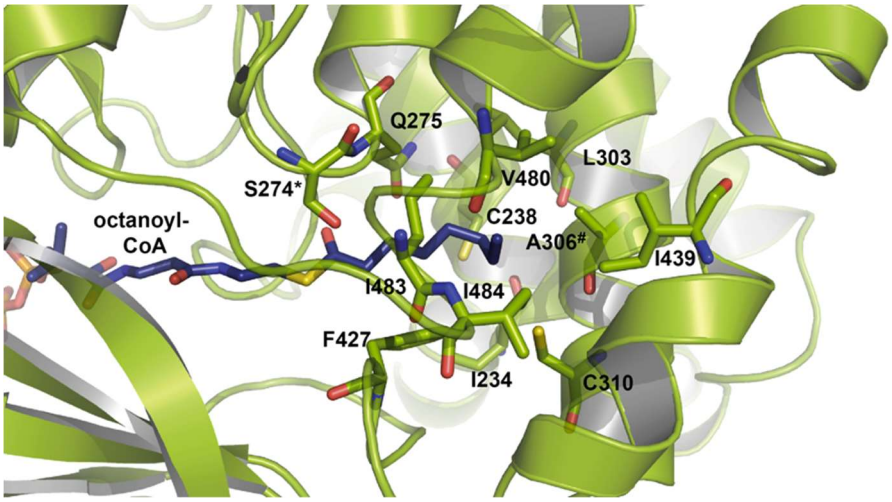


Figure 12: C_8-CoA binding in the AT-mutated I306A (I151A in *C. ammoniagenes* FAS).
C_8-CoA is shown as stick representation in blue with its nucleotide moiety extending to the left outside of the frame. Residues involved in forming the hydrophobic binding pocket (similar to the one described for the human MAT transferase[23]) are also shown as stick representation (I234, C238, Q275, L303, A306, C310, F427, I439, V480, I483, I484; all in *S. cerevisiae* numbering). For better clarity, cartoon representation of the secondary structures was set to 20% transparency. The active site S274 is marked with (*); the mutated A306 with (#).

C. ammoniagenes MPT mutation R1408K and AT mutation I151A, each of which individually brought significant increases in the relative yield of C_8-CoA, were further tested in combination (in addition to the KS domain mutations, yielding quadruply-mutated FAS$^{G2599S\text{-}M2600W\text{-}R1408K\text{-}I151A}$). The combination of these two beneficial mutations might be expected to be superior to each in isolation; however, their effects on relative C_8-CoA yield were not additive. Rather, the KS/MPT combination remained the highest-yielding for specific production of C_8-CoA (see Fig. 7a,b).
Incorporating the MPT and AT mutations simultaneously, the initial kinetic model reproduced the finding of non-additivity, with C_8-CoA output for the double MPT/AT mutant no higher than that for the MPT single mutant (see Supplementary Fig. 12). When introducing the I151A mutation in addition to the double KS mutation in the revised kinetic model, it yielded a prediction of modestly increased C_8-CoA release, at the cost of a modest decrease in the native $C16$-CoA and $C18$-CoA products, again in good agreement with the experimental finding (see Fig. 10). Applying all four mutations in combination, the notable experimental finding that the AT and MPT mutations, while each separately advantageous, do not in combination yield greater C_8-CoA production, is again reproduced in the model.

4.1.1 Additional engineering of the KS domain

In a subsequent study on KS engineering (independent of the mentioned publication[17]), another sensitive mutation site was found after FAS variants, that were apparently producing short FA naturally, were examined more closely. Particular FAS involved in the secondary metabolite pathway of Aflatoxin B_1, were presumed to have specificity for hexanoyl production[19-21]. In sequence alignments with other type I FAS from bacterial and fungal sources, a distinct mutation from leucine/phenylalanine to tyrosine was noted for the FAS from *Aspergillus nidulans* which had been directly linked to short FA production[19]. This mutation can also be seen in other *Aspergillus* FAS which have not yet been characterized in detail (Fig. 13).

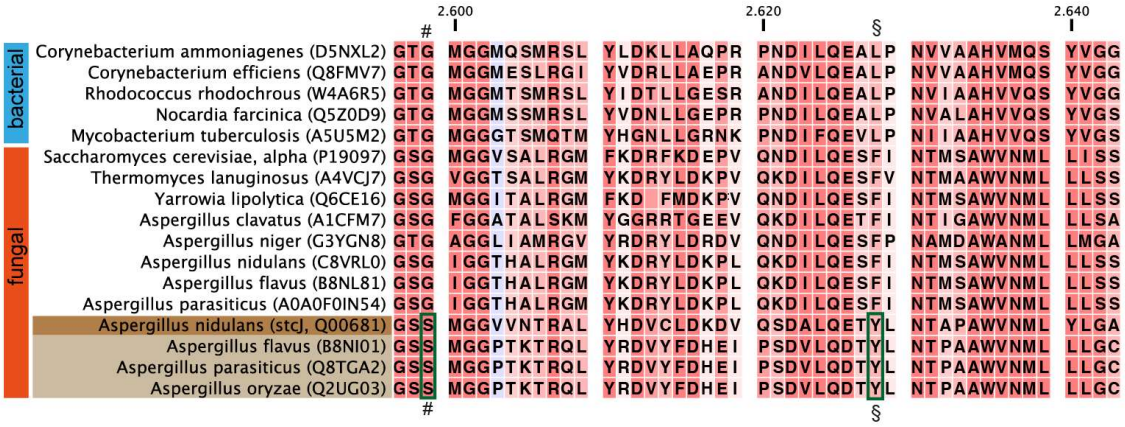


Figure 13: Sequence alignment of different type I FAS from bacteria and fungi.
Numbering is according to *C. ammoniagenes* FAS. In brackets, UniProt identifiers are listed for each FAS. Position 2628 (marked with §) is mostly conserved with big, hydrophilic residues. For *A. nidulans*, two FAS were reported, one for regular FA synthesis and one for short FA synthesis as part of a secondary metabolite pathway[19]. This particular FAS (gene *stcJ*, brown background) shows tyrosine at this position. The mutation can also be seen in other *Aspergillus* FAS that likely have the same product spectra (light brown background). A second typical mutation site was G2559S, which led to short FA production when introduced into other type I FAS[8-9].

We also incorporated the mutation into the bacterial FAS from *C. ammoniagenes*. Structural insight were gained by homology modeling with fungal FAS X-ray data[10, 12, 15, 132]. On that basis, we concluded that in L2628Y-mutated *C. ammoniagenes* FAS, the hydroxyphenyl moiety of tyrosine is putatively pointing directly into KS binding channel.

The product spectra analysis of the according construct FAS^{L2628Y} showed a clear shift towards shorter products with the C_8-CoA as the predominant species making up 74% of the detected CoA esters in that construct. Long chain production was decreased below the limit of detection in our set-up (Fig. 14 and Supplementary Fig. 13a). The overall activity of FAS^{L2628Y} was also determined (using the described method[1]) at a low value of 5.8 ± 3.5 mU/g (protein).

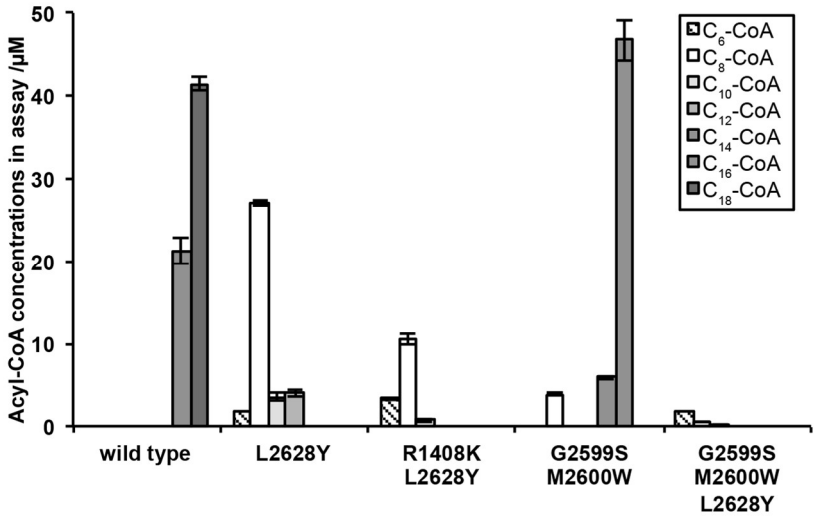


Figure 14: Product spectra of different FAS variants in comparison to the wild type.
By the introduction of the L2628Y mutation, the native function of the enzyme is changed from a long chain producing FAS to a short chain producing FAS. In combination with other mutations previously linked to changing the product spectra of FAS, the product spectra can be altered further. FA production can, however, also be heavily distorted or abolished as shown for the G2599S-M2600W-L2628Y construct in comparison to the previously described G2599S-M2600W construct. Error bars reflect the standard deviation of three assays using one enzyme purification.

Mutations that were previously described above, were also tested in combination with the novel L2628Y mutation with the aim of probing synergetic influences on the product spectra. The first mutation added to L2628Y, was R1408K that is located in the MPT domain active site. The $FAS^{R1408K-L2628Y}$ shows slightly higher share of C_6- and C_8-CoA (see Fig. 14). When combining the L2628Y with the AT mutation I151A, only minor modulations and no clear beneficial effect on the product spectra was observed (Supplementary Fig. 13b).

Also, the previously reported KS mutations G2599S and M2600W were probed in combination with L2628Y. Here it should be noted that the *Aspergillus* FAS that assumingly produce short FA carries a serine similar to the mutation G2599S. However, in our set-up, the corresponding *C. ammoniagenes* FAS construct,

FAS$^{G2599S-L2628Y}$, showed only minor product output (Supplementary Fig. 13b). There was no noteworthy production of acyl-CoA esters detected for the triple KS mutant, FAS$^{G2599S-M2600W-L2628Y}$. The recorded activities for the combinatorial FAS were similar to that of FAS^{L2628Y} in the low range, at around 1.8 – 11.9 mU/g (protein).

4.2 Modulation of domain-domain interactions

In all megasynthase systems, domain-domain interactions are of utmost importance. Particularly, the carrier protein ACP undergoes frequent domain-domain interaction during the synthetic cycle[133], where the surface recognition of the carrier protein and the enzymatic domains play a crucial role essentially determining the reaction pathway.

Our goal was to modulate the interactions between ACP and the domains KS and MPT for the production of short FA by rational design. The KS domain was chosen, since it is responsible for the actual condensation reaction of the acyl chain. In this process, the ACP has to interact twice with the KS domain; first for loading the acyl chain into the KS binding channel (ping step), and second for delivering malonyl, which after decarboxylation elongates the acyl chain (pong step) (see Fig. 1)[87]. In cleaving off the fully elongated acyl chains, the MPT domain is the direct counterpart to the KS domain in fungal and bacterial FAS. However, the MPT is also responsible for loading malonyl, the elongation unit, onto the ACP domain, and therefore acts as a key catalytic domain in dual manner on chain length control.

The mutations were set on the surface of the KS and MPT domain, while the ACP itself was left in its native form to ensure that measured effects did not derive from distortions somewhere else in the reaction cycle. By the introduction of mutations, we aimed to decrease the affinity between ACP and the targeted domains and thus, decreasing the probability of productive interactions.

When decreasing the ACP:KS affinity, we expected that acetyl or acyl, bound to ACP would less likely be loaded into the KS binding channel, and therefore less likely be elongated. This state of the acyl chain resting on the ACP could elevate the chances that short intermediates are already unloaded by the promiscuous MPT domain before their further elongation[94-96].

When decreasing the ACP:MPT affinity, we expected to mimic a lowered malonyl concentration by a suppressed import of malonyl into the fatty acid cycle (see Fig. 1). As mentioned before, a lowered ratio of malonyl-CoA to acetyl-CoA is known from *in vitro* assays to lead to short fatty acids[83, 93]. This effect might be counter-acted by a similarly suppressed product release. However, as product release only happens once per reaction cycle, while malonyl has to be loaded three times for octanoyl-CoA (as an example for short chains) or even eight times for stearoyl-CoA (a long chain) production, we speculated on the overweight of the impact from reduced malonyl loading.

As a particular challenge in our engineering efforts, we aimed at modulating the affinity of the ACP to the targeted domain in a balanced manner, in order to influence the product spectrum, while keeping at least a basal activity. This is different to conventional studies focusing mainly on the mapping of crucial residues in domain-domain interactions by inducing harsh effects, as, most often, the complete abolishment of the interactions[134].

As in the previous *in vitro* studies, *C. ammoniagenes* FAS (Uniprot Identifier: D5NXL2) was used. As before, we were able to exploit the atomic models of *S. cerevisiae* FAS collected in X-ray crystallographic studies[10, 12] for accurate homology modeling guiding our rational design of *C. ammoniagenes* FAS since they shares substantial sequence aspects[15].

As a template for protein engineering, we did not use the wild type, but instead a modified *C. ammoniagenes* FAS construct was chosen. This FAS construct, carrying two mutations in the KS domain (G2599S M2600W), has a bimodal product spectrum producing both short and long FA (see Fig. 7a,b). It therefore served as a sensitive reporter system for the quantification of effects of mutations influencing domain-domain interactions in either direction, which would be hard to do in the wild type that only produces short FA as a minor side product. All mutations referred to in this section of modulation of domain domain interactions, were therefore introduced into the template with the G2599S-M2600W background.

The individual mutation sites identified in this study were encouraged by previous studies on ACP-KS or ACP-MPT interactions of the separate proteins as occurring in type II FAS systems. Sequence alignments and structural comparisons then led to the candidate positions listed in Table 7. In particular, mutations sites on the KS

domain surface, such as D2553, D2556, and N2557, derived from a study analyzing the transient interaction FabB, a KS homologous protein in *E. coli*, and its ACP[135]. A2696 was identified from a crosslinking study on FabF, another KS homologue in *E. coli*[134]. Additionally, structural data of the *S. cerevisiae* FAS, which was crystallized with ACP docked to the KS domain[10], were included in our consideration; specifically contributing two novel, potentially crucial interaction sites N2621 and D2622 (Fig. 15a,b).

Table 7: Overview of examined surface mutations.
The mutation sites from this study are listed with the positions for *C. ammoniagenes* FAS (protein used for the presented *in vitro* study), for *S. cerevisiae* (from which the structure was used for rational planning) and for other proteins (which have previously been subject of ACP–domain interactions). For *S. cerevisiae* FAS, also the chain on which the mutation is located is listed in brackets.

	C. ammoniagenes	*S. cerevisiae*	Comments
KS domain	D2553	D1203 (α)	described mutation K63Q, FabB (*E. coli*)[135]
	D2556	S1206 (α)	described mutation K66Q, FabB (*E. coli*)[135]
	N2557	Q1207 (α)	similar to mutation K66Q, FabB (*E. coli*)[135]
	N2621	N1272 (α)	new in this study
	D2622	D1273 (α)	new in this study
	A2696	K1347 (α)	described mutation R206, FabF (*E. coli*) (0% binding in previous study)[134]
MPT domain	Q1647	R59 (α)	described mutation site R287, MCAT (*S. coelicolor*)[136]
	K1435	R1861 (β)	new in this study
	R1494	R1962 (β)	described mutation site R189, MCAT (*S. coelicolor*)[136]
AT domain	I151A	I306A (β)	described in previous chapter
	V289	I483 (β)	new in this study
	L290	I484 (β)	new in this study

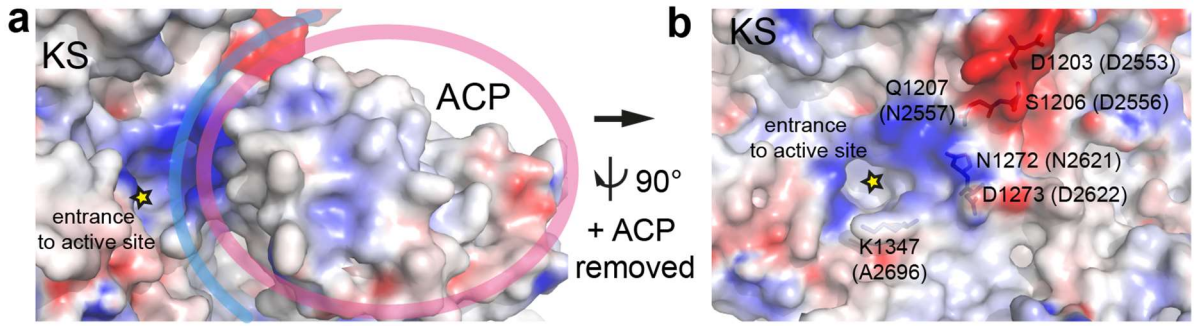


Figure 15: Structural basis for KS surface engineering.
Rational engineering of *C. ammoniagenes* FAS was based on the structure of the homologous protein from *S. cerevisiae* (PDB code: 2UV8[10]). **(a)** The ACP domain (on the right, circled in pink) is shown in its position locked to the KS domain (on the left). The surface depiction is colored according to its vacuum electrostatics (generated with PyMOL, Schrödinger, LLC; red, negatively charged; blue, positively charged). **(b)** The same structure as in (a) is displayed but rotated by 90° and with the ACP domain hidden for a view into the active site of the KS domain. Residues mutated in this study are shown in their native form in stick representation with *S. cerevisiae* and *C. ammoniagenes* (in brackets) numbering.

MPT engineering was based on a previous study on malonyl-CoA-ACP transferase (MCAT) from *S. coelicolor*[136], a related transferase from a type II FAS. Two potential sites in *C. ammoniagenes* FAS, R1494 and Q1647, were initially identified as candidate mutational sites. Given the general positive charge on the MPT surface around the entry to the active site (Fig. 16a), K1435 was chosen as another mutation site.

To reinstall an option for successful product release, the AT site was modified to also accept short acyl chains[22-23] and thus, potentially act as an alternative exit pathway for product release. The mutations I151A (previously described), V289A, and L290A (Fig. 16b) were introduced for this purpose (see Table 7).

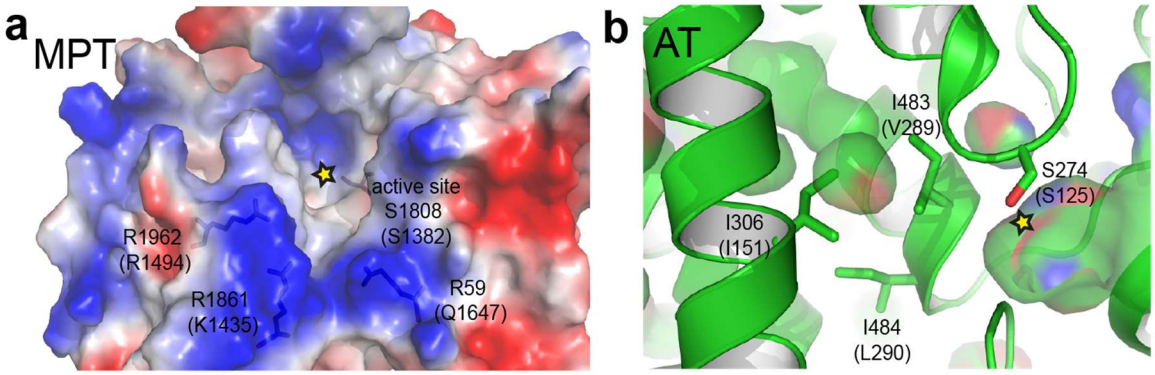


Figure 16: Structural basis for MPT surface and AT binding channel engineering.
As above, the *S. cerevisiae* FAS structure (PDB code: 2UV8[10]) was used as basis for engineering of *C. ammoniagenes* FAS. **(a)** The MPT domain surface representation is shown with vacuum electrostatics indicated (generated with PyMOL, Schrödinger, LLC; red, negatively charged; blue, positively charged). The active site and residues that were mutated are shown in their native form with *S. cerevisiae* and *C. ammoniagenes* (in brackets) numbering. **(b)** The AT binding site is shown in cartoon representation with the active site and the mutation sites in stick representation. The active site is loaded from the right (surface of entrance channel shown). With the introduction of the mutations, a binding channel presumably opens up extending to the top left.

The selected residues on the surface of the surface of the KS and the MPT domain were mutated to alanine and/or to residues with an opposite charge compared to the native one. These mutations were tested in the reporter FAS system (G2599S M2600W). *In vitro* product assays were performed and measured by HPLC-UV-MS. For several mutations in the KS domain, shifts towards shorter FA were observed. For three mutation sites, and especially for the three specific residue exchanges N2557E, N2621D and D2622A, beneficial effects were observed. C_8-CoA yield was increased by a factor of up to 3.6, making up 77% of all detected CoA esters (Fig. 17a and Supplementary Fig. 14).

For other mutation sites on the KS surface, no beneficial effects were observed. Interestingly, some constructs showed less short FA by shifting the bimodal output spectrum towards longer FA, reinstalling wild type like behavior in our template construct, opposite to what was previously anticipated (see Fig. 17a).

For the MPT-ACP interaction constructs, no significant increase in short chain FA output was observed (Fig. 17b). If at all, the construct comprising mutations I151A V289A L290A K1435A G2599S M2600W showed slightly higher amounts as the reference construct G2599S M2600W.

In addition to the product assays, also the activity was determined by spectroscopically recording NADPH consumption (Table 8)[1, 137]. Interestingly, the

constructs with the best yields in short FA were severely impaired in their activity, e.g. N2557E, N2621D, and D2622A only showing a fraction of the specific activity of the template construct. Accordingly, for mutants tested in this study, poor activities correlate with high short FA output. This correlation implies that just the severe decrease in ACP-KS affinity causing the observed drop in activity influences the product spectrum.

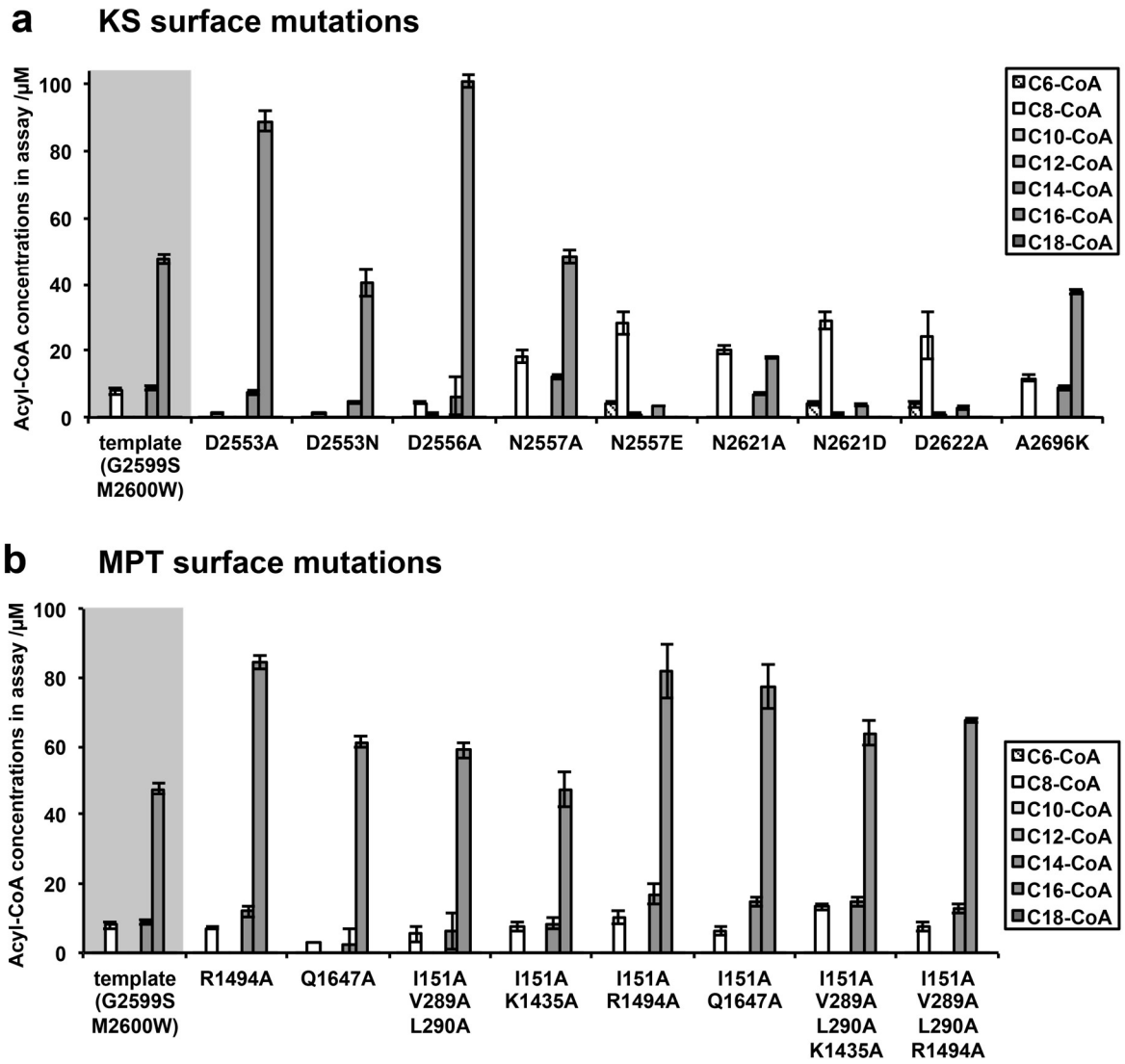


Figure 17: Product spectra of mutants with modulated ACP-domain interactions.
The product spectra were measured by HPLC-UV-MS and quantified at 260 nm. As a reference, the reporter FAS (G2599S M2600W) is listed, which also served as a template into which all surface mutations were introduced. **(a)** Product spectra of KS surface mutants modulating the ACP-KS interactions are shown. **(b)** Product spectra of MPT surface mutation (K1435A, R1494A, Q1647A) and AT binding channel mutations (I151A, V289A, L290A) are shown, also in combinations thereof. Error bars show the standard deviation of three assays for each construct.

Table 8: Activities of FAS constructs with surface mutations.
The activities of the FAS constructs was measured by NADPH consumption[1]. Here, 1 U was defined as the incorporation of 1 µmol of malonyl-CoA per minute. Wild type and the template are shown for comparison. All KS mutants as well as MPT/AT mutants also carry the template mutations (G2599S M2600W), which has not been stated in the figure explicitly. Error bars show the standard deviation from three measurements for each protein.

	Construct	Activity /mU/mg (protein)
	wild type	383 ± 31
Template	G2599S M2600W	51 ± 13
KS surface mutants	D2553A	106 ± 14
(all carry the GS MW	D2553N	44 ± 13
mutation)	D2556A	271 ± 27
	N2557A	77 ± 6
	N2557E	10 ± 6
	N2621A	31 ± 10
	N2621D	10 ± 3
	D2622A	7 ± 1
	A2696K	72 ± 1
MPT surface and AT channel mutants	R1494A	163 ± 14
(all carry the GS MW	Q1647A	77 ± 11
mutation)	I151A V289A	88 ± 14
	I151A L290A	217 ± 24
	I151A V289A L290A	229
	I151A K1435A	170 ± 28
	I151A R1494A	208 ± 48
	I151A Q1647A	133 ± 14
	I151A V289A L290A K1435A	173 ± 16
	I151A V289A L290A R1494A	116 ± 2

4.3 Engineering FAS for the production of a polyketide

Motivated by the detailed biochemical and structural characterization of fatty acid synthesis presented in these and earlier studies, as well as the established access to FAS proteins by recombinant methods[14, 41, 138], we used FAS as a model protein for expanding our understanding of the PKS/FAS superfamily of proteins (also included in the publication with early works on *in vitro* production of short FA)[17].

Using engineered FAS type I proteins, we aimed to set up an artificial biosynthetic pathway for the production of the simple polyketide that should be probed in a one pot synthesis *in vitro.* To this end, we used a first FAS construct (“module 1”) which

was optimized for the production of the short FA as described above. A second FAS construct ("module 2") was designed for further elongating C_8-CoAand terminating the synthesis by spontaneous lactonization yielding 6-heptyl-4-hydroxypyran-2-one (6-HHP) **1** (Fig. 18). For both modules, *Corynebacterium ammoniagenes* FAS (Uniprot Identifier: D5NXL2) was used as in our earlier studies.

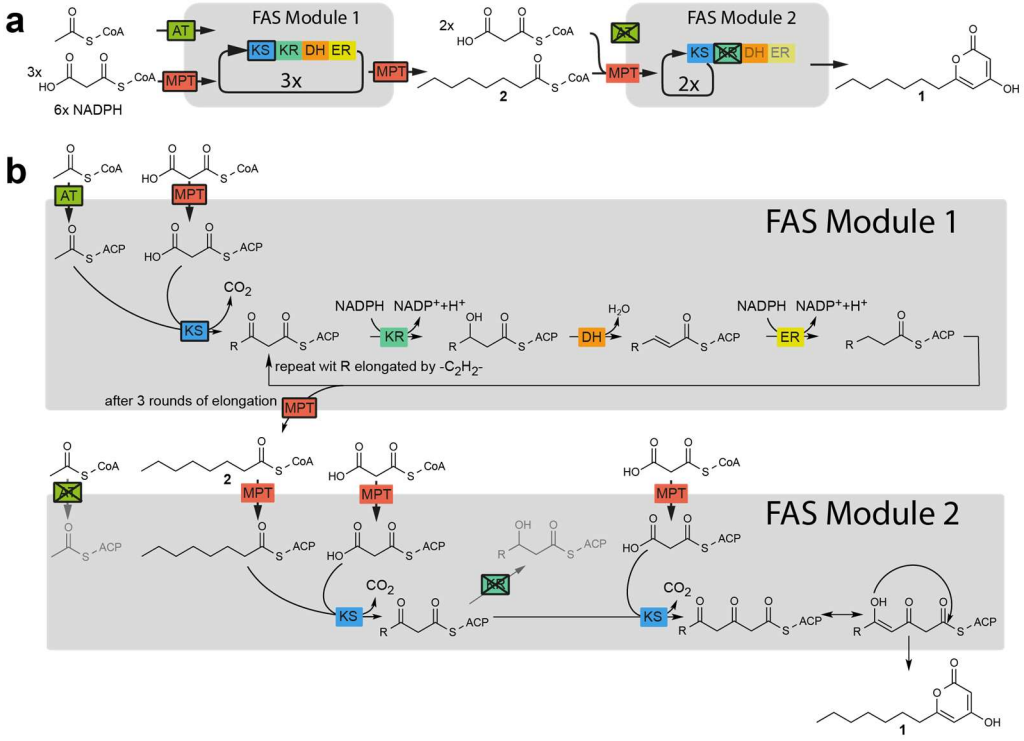


Figure 18: Engineering a pathway using FAS for the production of a simple PK.
(a) Overview of 6-HHP **1** synthesis from substrates acetyl-CoA, malonyl-CoA and the cofactor NADPH. Two engineered FAS, termed module 1 and module 2, work in sequence to first synthesize C_8-CoA **2** (module 1) and then the final 6-HHP **1** product (module 2). **(b)** Detailed representation of the individual steps involved in synthesis. Module 1 is modulated in FA chain length regulation. Mutations in domains KS, AT and MPT shift the FA product spectrum from the native products C_{16}- or C_{18}-CoA towards C_8-CoA **2**. Module 2 carries functional knock-outs in domains AT and KR to selectively accept and non-reductively elongate C_8-CoA **2**, respectively. The triketide dissociates from ACP by spontaneous lactonization to the final product **1**. In the scheme, engineered domains are highlighted by black frames as well as additionally crossed out when indicating a functional knock-out.

With several FAS mutants that could be employed as module 1, our engineering for this pathway required an additional FAS that had been turned into a simple PKS. For this, we produced a construct incorporating catalytic knockout mutations in the KR and AT sites. Mutation Y2227F in the KR domain impairs the proton transfer for neutralizing the hydroxyl anion in the reduction of the carbonyl group by NADPH[139],

thereby blocking the canonical reaction pathway that processes β-keto intermediates to saturated acyl chains. This mutation has previously been characterized as producing triacetic lactone (TAL)[123, 140]. With mutation S126A in the AT domain of module 2, we further blocked loading of acetyl-CoA, thereby restricting priming activity to those substrates that can be taken up through the less-specific MPT active site. This mutation enforces sequential synthesis, with module 2 processing the acyl-CoA products of module 1 (see Fig. 18b). Module 2 achieved about 5% of the specific activity of the wild type FAS of around 8–16 mU/mg (protein) compared to a typical value of approx. 300 mU/mg (protein) for the wild type. It should, however, be noted that while in both cases 1 U was defined as the incorporation of 1 µmol of malonyl-CoA per minute, activity of module 2 constructs were was measured by product formation of the lactone ring (increase in absorption at 298 nm) and not by NADPH consumption (as for the wild type).

Also a further developed construct for module 2 was tested: Guided by study on a type II KS domain, a mutation was found which led to an enhanced TAL production[141]. Incorporated at the equivalent position in the KS domain of *C. ammoniagenes* FAS, a construct for module 2 (FAS-M2$^{S126A-Y2227F-H2608A}$) was created. Testing, however, did not show an increased yield, but only 84% product formation of the regular FAS-M2$^{S126A-Y2227F}$, which stayed the first choice accordingly. In order to optimize yields of the overall reaction, we proceeded in two steps; first, several constructs of module 1 (producing short chain acyl-CoA esters) were tested with the same module 2 (FAS-M2$^{S126A-Y2227F}$) (Supplementary Fig. 15a). As a second measure, we optimized reactant and cofactor concentration for the two best performing combinations in the one pot reaction sequence (Fig. 19). The two combinations, FAS-M1$^{G2599S-M2600W-R1408K}$ → FAS-M2$^{S126A-Y2227F}$ and FAS-M1$^{G2599S-M2600W-I151A-R1408K}$ → FAS-M2$^{S126A-Y2227F}$, finally led to yields of the final polyketide product, 6-HHP, of up to 35% in a one pot synthesis under limiting concentrations of substrates acetyl-CoA and NADPH. Variation of the relative concentrations of the modules did not show a significant impact on product yield (see Supplementary Fig. 15b).

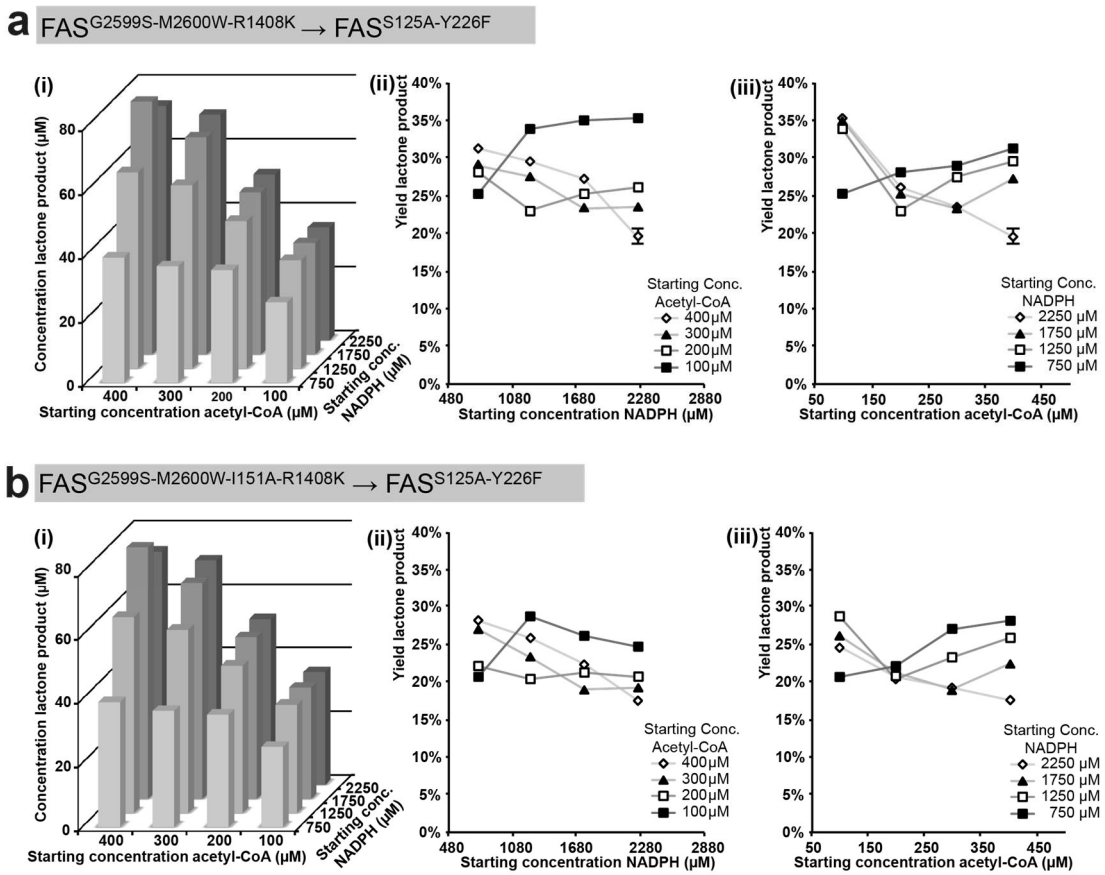


Figure 19: Output of the engineered reaction pathway at various substrate concentrations.
(a) Sequential synthesis with FAS-M1$^{G2599S-M2600W-R1408K}$ → FAS-M2$^{S125A-Y2227F}$. (i) 6-HHP molar concentrations recorded at different combinations of NADPH and acetyl-CoA starting concentrations. (ii/iii) Calculated from the final product concentrations of 6-HHP, the percentage yield in reference to the limiting substrate at each data point was calculated. The yields were then correlated to the starting concentration of NADPH and acetyl-CoA, respectively, and shown in the two panels. One sample was measured for each data point except for initial standard conditions (400 µM acetyl-CoA / 2250 µM NADPH) where four replicates were recorded as a representative for the low deviation within the same reaction conditions. Data points were connected for clarity (lines do not reflect experimental data) **(b)** Sequential synthesis with FAS-M1$^{G2599S-M2600W-I151A-R1408K}$ → FAS-M2$^{S125A-Y2227F}$. Presentation of data as in (a).

4.4 Production of short FA *in vivo* using *S. cerevisiae*

For the production of short FA in *S. cerevisiae* as an *in vivo* system, a *⊿fas1 ⊿fas2* strain was created. Two heterozygous strains, each carrying one deletion, were mated and then sporulated to gain the double knockout strain. Low copy vectors pRS315 and pRS313, with α- and β-chain encoding sequences *FAS2* and *FAS1* respectively, were transformed into the FAS-deficient strain. Both genes were provided under control of their natural promoters and terminators[25]. In this set-up, the plasmid-based FAS system was the only source for *de novo* FA synthesis. We favored this approach, as it is minimally invasive and allows comparison with wild type output spectra.

For the production of short FA, the cycle of *de novo* FA production was manipulated in all enzymatic centers at the equivalent positions as already performed in the *in vitro* study (with *C. ammoniagenes* FAS): Mutations were introduced individually or in combination in the KS domain (G1250S, M1251W, and F1279Y; all located on the α-chain), in the MPT domain (R1834K; located on the β-chain), and in the AT domain (I306A; located on the β-chain, too).

For the product analysis, a few additional aspects had to be considered as the *de novo* FA synthesis in the *in vivo* environment of *S. cerevisiae* is embedded into the metabolic network as shown in a simplified version in Fig. 20. Fungal FAS produce FA as CoA esters in the same fashion as bacterial FAS, which is different from mammalian FAS producing free FA. CoA esters were, however, not quantified in our set up. Instead, several acyl-CoA:ethanol acyl transferases were recently reported to act as TE for *S. cerevisiae* which hydrolyze short CoA esters into the free FA[26-27]. Short chain free FA will then be exported out of the cell, prompting us to analyze short FA of our strains in the culture media[30].

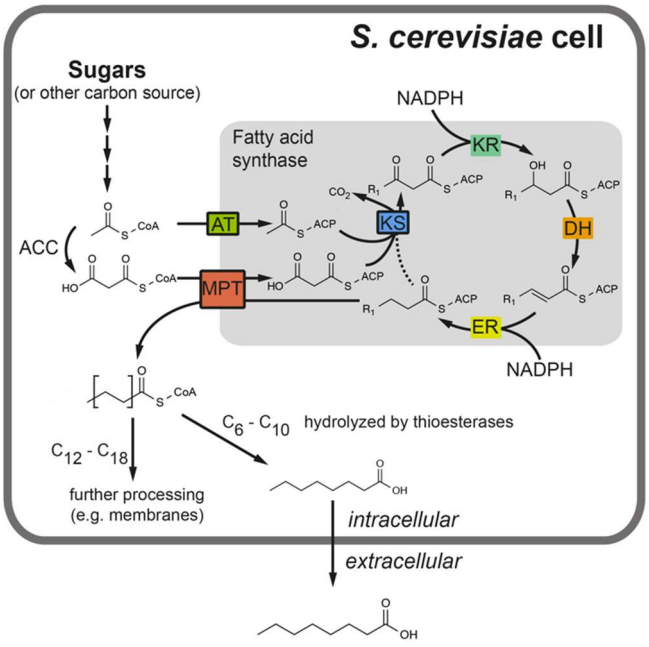


Figure 20: Fatty acid synthesis in *S. cerevisiae*.
The *de novo* FA production in *S. cerevisiae* is shown. FAS carries modifications in the domains KS, AT and MPT to produce short FA (instead of its native product, typically C_{16}- or C_{18}-CoA) from acetyl-CoA, malonyl-CoA and NADPH. *S. cerevisiae* FAS naturally produces CoA esters. Short free FA, appearing as hydrolyzed by unspecific TE (here shown exemplary for C_8), are then exported to the medium, from which they have been extracted for analysis.

4.4.1 Testing KS domain mutations *in vivo*

Based on the previous *in vitro* studies, we started also by introducing mutations into the KS domain. The first modification was G1250S. *S. cerevisiae* strains with this mutation were reported to increase levels of C_6 and its ester derivatives before[8]. This mutation had also been linked to enhanced resistance to cerulenin; an inhibitor that mimics KS-bound medium chain acyl compounds[12]. In comparison with the wild type, the G1250S strain showed a 9-fold increase in C_6 production (15.3 mg/L on average) (Fig. 21 and Supplementary Fig. 16).

As a second position in the KS domain, M1251 was mutated. Our intention was to enhance the previously assumed gatekeeper effect by replacing the thioether residue of M1251 with tryptophan as in the *in vitro* studies on the homologous bacterial FAS. Indeed, the G1250S-M1251W double-mutated *S. cerevisiae* strain showed increased levels of C_6 (with 19.9 mg/L, 12-fold increase over wild type) and C_8 (32.7 mg/L, 56-fold increase over wild type) (see Fig. 21).

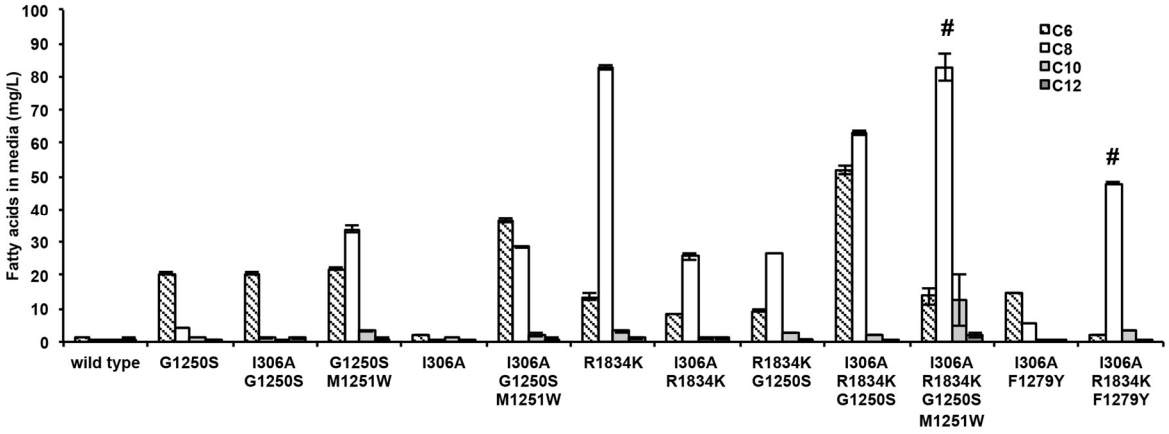


Figure 21: FA product spectra of YPD cultivations of *S. cerevisiae* strains with mutated FAS. For the measurements of the product spectra, cultures of *S. cerevisiae* were grown for 48 h at 30 °C, the media extracted and FA quantified via GC-FID. Error bars shown here reflect three technical repeats (for biological repeats, see Supplementary Fig. 16). The strain carrying the I306A-R1834K-G1250S-M11251W mutations and the strain with the I306A-R1834K-F1279Y mutations (both marked by #) only grew to approximately one third of the regular cell density as compared to the other strains.

The third promising KS mutation was also tested, which was initially identified in sequence alignments with FAS of reported C_6 producers from *Aspergillus*[21, 44]. Growth of G1250S-F1279Y double mutated strains, which resembles the naturally occurring mutant combination of the particular *Aspergillus* FAS (see Fig. 13) was strongly inhibited (Supplementary Table 16) and FA output spectra were hardly interpretable in chain length distribution. The effect of the F1279Y mutation was therefore solely tested in combination with transferase mutations as reported below.

4.4.2 Testing of the transferase mutations *in vivo*

The first transferase domain to be modified was the MPT domain. Guided by our *in vitro* studies, we mutated the equivalent position in *S. cerevisiae*, the central arginine R1834 that stabilizes the carboxyl group of malonyl. As the R1834K mutation lowers the affinity of malonyl, as reported for homologous proteins before[80,81], we expected a dual effect here, too: disfavoring malonyl uptake as well as favoring the competing acyl release.

The impact of the R1834K mutation in MPT domain on the product spectrum was dramatic, yielding production of mostly C_8 of 100 mg/L, corresponding to a 23-fold increase over wild type FAS (see Fig. 21).

The AT domain was next to be modified. From the previous studies on related transferases[22-23] and our previous *in vitro* experiments, we tested the corresponding mutation in our *in vivo* system (I306).

When installing the AT mutation I306A in *S. cerevisiae* FAS, the mutation, by itself, had no significant effect on short FA yield. However, in combination with other mutations, I306A drastically changed product distributions, as will be shown in the next section.

4.4.3 Combinations of mutations *in vivo*

Mutations were also tested in several combinations (see Fig. 21), leading to increased amounts in overall short chain FA as well as to the highly selective production of specific short FA. In general, recorded FA output spectra of strains carrying multiple mutated FAS implied a complex and partly non-intuitive behavior of individual mutational effects (Table 9).

The ambiguous impact of the I306A mutation (AT mutant) on *S. cerevisiae* FA production well highlights the complex molecular basis of chain length control. As mentioned above, the strain carrying solely I306A did not show an increased yield in short FA in comparison to wild type yeast. The insertion of the I306A mutation in the background of a KS double mutated strain, yielding a I306A-G1250S-M1250W-mutated strain, led to slightly increased levels of 66 mg/L short FA as compared to 57 mg/L (G1250S-M1250W). In the background of a R1834K-mutated strain, an unexpected negative effect of the I306A mutation on short chain FA production was found; the I306A-R1834K-mutated strain produced only 37 mg/L short FA in total, instead of 100 mg/L (R1834K).

In spite of a partly ambiguous impact on overall FA amounts, output spectra revealed specific influences of mutations on FA chain length. Table 9 provides an indication on the chain length that is favored by a particular mutation. Insertion of the mutations G1250S or I306A increases short FA output mostly by inducing C_6 FA production, while mutations M1251W or R1834K rather act on C_8 FA production levels, if short FA production is increasing at all.

Within tested strains, some combinations of mutations stand out as following: The highest total yield was achieved with the combination of I306A-R1834K-G1250S (AT MPT KS mutant), producing a total amount of 118 mg/L. For the specific production

of C_6, the double mutant I306A-G1250S (AT KS mutant) showed best results with 20 mg/L of C_6 accounting for 90% of the measured short FA (C_6 to C_{12}). A similar specificity, but for the production for C_8, was possible with the triple mutant I306A-R1834K-F1279Y (AT MPT KS mutant) with a titer of 48 mg/L for C_8 specifically, representing 89% of its short FA output.

4.4.4 Examining cell viability

For monitoring growth behavior, we generally measured wet cell pellet weights as well as, for selected strains, recorded OD_{600} at several time points (Supplementary Fig. 17a and Supplementary Table 5). For the growth in regular YPD, strains group into three classes according to their growth behavior; (i) strains with regular (wild type-like) growth rates, (ii) strains with reduced growth rates, and (iii) strain with extremely low or no detectable growth rates.

We assume that the growth correlates with the efficiency of the strains in producing short FA. Inhibited growth might either rely on the insufficient supply with long FA and/or the toxicity by short FA[142-143], both induced by a highly efficient short FA production. The observation of regular growth, when YPD medium was supplemented with oleic acid ($C_{18:1}$, 1 mM), suggests rather the limitation of long FA as affecting growth (see Supplementary Table 5).

Also, FA production of *S. cerevisiae* strains was measured when grown in supplemented YPD ($C_{18:1}$ at 1 mM). Strains generally showed lowered yields in short FA production with the only exception of the I306A-R1834K-G1250S-M1251W-mutated strain producing 131 mg/L instead of 104 mg/L of short FA (see Supplementary Fig. 18 and Fig. 21). Under growth in supplemented YPD ($C_{18:1}$ at 1 mM) also product spectra were different in the composition of FA. Intriguingly, we observed a shift towards shorter FA, as compared to growth in regular YPD, with the G1250S-M1251W-mutated strain as most prominent example. It is tempting to speculate that this effect might be based on inhibition of acetyl-CoA carboxylase Acc1[144], leading to lower malonyl-CoA levels and shorter FA[83, 93]. Also other more complex regulatory effects by long chain FA, as e.g. on the FA cycle or upstream enzymes, might be involved making the comparison with FA production data under normal non-supplemented medium difficult[144]. Moreover, the production of short FA by adding long FA is questionable for practical reasons.

Regarding a putative impact on growth by toxic short FA[142-143], we were interested whether a strong initial production of C_8 accounts for the inhibition of further growth. Product spectra of selected strains were measured after 12, 24 and 48 h; the latter constituting the regularly chosen measurement point (Supplementary Fig. 17b). While the I306A-R1834K-G1250S-M1251W-mutated strain reached its final cell density already after 24 h, different to e.g. the R1834K mutated-strain that reaches maximum cell density at a later time point (see Supplementary Fig. 17a), the FA production profile followed the generally observed course of maximum yield in short chain FA after 48 h. These data argue against an early growth limitation of some strains by high inhibitory concentration of short FA.

Glucose consumption and ethanol levels were recorded during the cultivations. All tested strains behaved very similar with glucose being entirely consumed after about 20 h, followed by a partial consumption of the produced ethanol (Supplementary Fig 18).

Table 9: Change of FA levels after introduction of FAS mutations.
To quantify how beneficial a mutation is, the results from one strain without the mutation (first line) is compared to the same strain with the mutation (second line). By dividing the second value by the first value, the increase a mutation adds to the yield (for C_6, C_8 and total short FA) is given as a factor ("x-fold increase"). Accordingly, values above 1 (in bold) are equivalent to a positive, higher yield; for values below 1, the yield dropped with the introduction of the mutation.

		C6 (mg/L)	C8 (mg/L)	total (C6 to C12) (mg/L)
G1250S (KS domain)	wild type	1,7	0,6	4,3
	G1250S	15,3	3,9	21,6
	X-fold increase	**9,1**	**6,7**	**5,0**
	I306A	1,2	0,5	3,6
	I306A G1250S	20,4	1,0	22,8
	X-fold increase	**17,0**	**2,2**	**6,4**
	R1834K	15,7	80,2	100,1
	R1834K G1250S	9,4	26,7	39,5
	X-fold increase	0,6	0,3	0,4
	I306A R1834K	10,1	25,1	37,4
	I306A R1834K G1250S	52,1	63,1	118,2
	X-fold increase	**5,2**	**2,5**	**3,2**
M1251W (KS domain)	G1250S	15,3	3,9	21,6
	G1250S M1251W	19,9	32,7	57,2
	X-fold increase	**1,3**	**8,3**	**2,6**
	I306A G1250S	20,4	1,0	22,8
	I306A G1250S M1251W	37,1	25,6	65,9
	X-fold increase	**1,8**	**25,2**	**2,9**
	I306A R1834K G1250S	52,1	63,1	118,2
	I306A R1834K G1250S M1251W	13,0	77,1	104,1
	X-fold increase	0,2	**1,2**	0,9
F1279Y (KS domain)	I306A	1,2	0,5	3,6
	I306A F1279Y	14,7	5,5	20,8
	X-fold increase	**12,3**	**11,9**	**5,8**
	I306A R1834K	10,1	25,1	37,4
	I306A R1834K F1279Y	2,2	47,9	53,9
	X-fold increase	0,2	**1,9**	**1,4**

Table 9: (continued)

		C6 (mg/L)	C8 (mg/L)	total (C6 to C12) (mg/L)
I306A (AT domain)	wild type	1,7	0,6	4,3
	I306A	1,2	0,5	3,6
	X-fold increase	0,7	0,8	0,8
	G1250S	15,3	3,9	21,6
	G1250S I306A	20,4	1,0	22,8
	X-fold increase	**1,3**	0,3	**1,1**
	G1250S M1251W	19,9	32,7	57,2
	G1250S M1251W I306A	37,1	25,6	65,9
	X-fold increase	**1,9**	0,8	**1,2**
	R1834K	15,7	80,2	100,1
	R1834K I306A	10,1	25,1	37,4
	X-fold increase	0,6	0,3	0,4
	R1834K G1250S	9,4	26,7	39,5
	R1834K G1250S I306A	52,1	63,1	118,2
	X-fold increase	**5,6**	**2,4**	**3,0**
R1834K (MPT domain)	wild type	1,7	0,6	4,3
	R1834K	15,7	80,2	100,1
	X-fold increase	**9,4**	**137,3**	**23,4**
	G1250S	15,3	3,9	21,6
	G1250S R1834K	9,4	26,7	39,5
	X-fold increase	0,6	**6,8**	**1,8**
	I306A G1250S	20,4	1,0	22,8
	I306A G1250S R1834K	52,1	63,1	118,2
	X-fold increase	**2,6**	**62,2**	**5,2**
	I306A	1,2	0,5	3,6
	I306A R1834K	10,1	25,1	37,4
	X-fold increase	**8,4**	**54,4**	**10,5**
	G1250S M1251W I306A	37,1	25,6	65,9
	G1250S M1251W I306A R1834K	13,0	77,1	104,1
	X-fold increase	0,4	**3,0**	**1,6**

5 Discussion

FA are ubiquitous compounds in organisms but also play an important role in biotechnology. The industrial and academic interest mostly stems from their role in production pathways for components of biofuels, but also from their generally versatile use for gaining access to a variety of other chemicals (see Fig. 4). In the FA *de novo* synthesis by FAS, understanding and manipulating the process can enable mankind to further adjust it to certain needs and gain access to new applications. The focus of this work was the rational manipulation of reaction aspects of FAS with an emphasis on chain length control.

Therein, our approach took on single reaction steps of FAS, mapped structure-function relationships and at the same time considered how they were embedded in the complex reaction network. The goal was the development of altered FAS modules which can be plugged into existing systems.

Possible usage includes the microbial production of gasoline-like biofuels that cannot be easily accessed since natural product distributions of FA synthesis are usually in the long chain range. As previously described, FA can also be part of bigger production pathways, where they are themselves further modified or bound to bigger compounds.

Furthermore, as FAS and PKS systems essentially share chemistry and high homology within catalytic domains, insights gained from FAS systems are likely to be applied to PKS systems.

In the following chapters, the results obtained during the course of this thesis will be discussed. As the focus of this work was on chain length control, there will be a clear emphasis on this topic. Different methods to reprogram type I FAS to produce short FA will be shown, both from initial *in vitro* experiments and the following *in vivo* studies. In some parts, engineering of FAS was extended further and a coupled FAS reaction pathway was established to produce a simple polyketide. First of all, the general approach and its classification in the biotechnological research field will be reviewed.

5.1 Settings for probing biotechnological engineering efforts

In order to address engineering objectives on FAS, our approach is placed in a specific setting. Generally, research for biotechnological production can cover a whole, continuous spectrum (Fig. 22): Starting from *in vitro* assays performed on individual domains, moving further towards more complex settings such as full constructs of multienzymatic complexes (if applicable) towards the reconstitution of complete pathways (more than one construct). Beyond that, *in vivo* studies represent another, higher level of complexity, ranging from small scale fermentations, over strain optimization to industrial scale fermentations. Not only the complexity of the systems increases when moving along the spectrum, also the cost per sample increases while accordingly the testable library size and the information of structure-function relationships typically decrease.

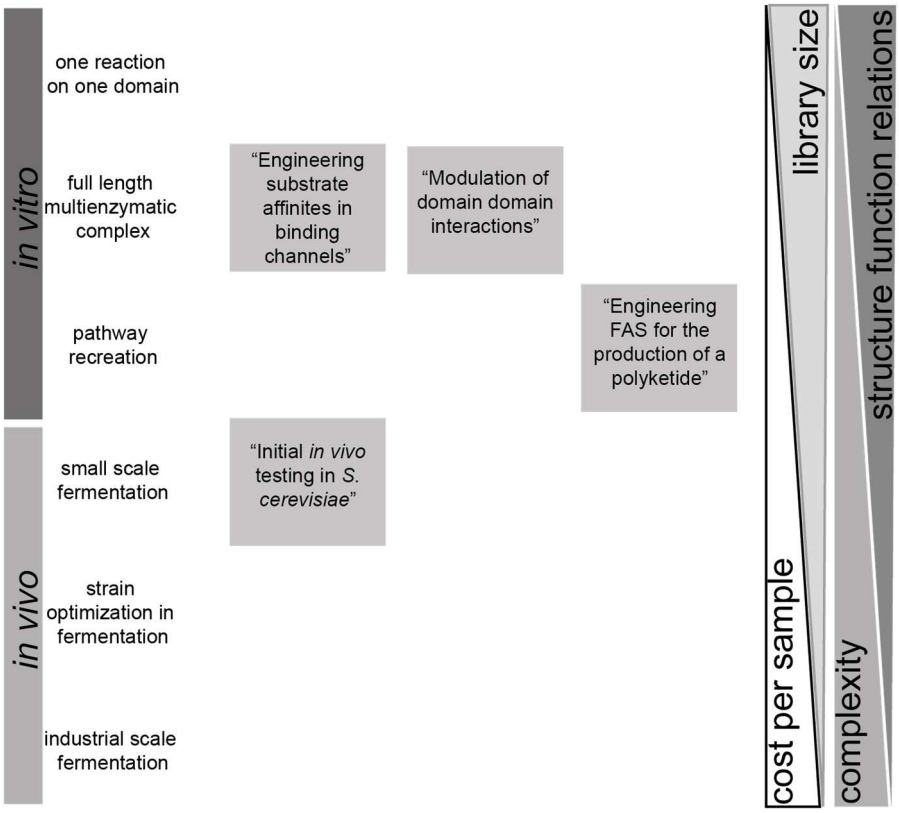


Figure 22: Environments for probing biotechnical engineering efforts.
The submitted publications are classified depending on the complexity of their testing environment. The spectrum starts with low complexity when only individual domains are studied and reaches up to high complex environments such as an industrial-scale fermentation. When moving to more complex environments, the cost and complexity increase while on the contrary, sample size and information on direct structure-function relationships decrease.

For our studies on rationally engineering FAS, initially an *in vitro* set up was chosen. The main reasons are that it enabled us to not only the characterize basic enzymatic functions such as activity and stability[1] but also control of reaction environments including substrate concentrations. Despite the fact, that individual catalytic domains were targeted with the rational engineering, testing was preferably done on the full-length constructs rather than on individual domains. It was judged that the interplay of all catalytic domains in this complex network can yield a better, more complete picture while direct conclusions could still be drawn on structure-function relationships. The results of the kinetic model of a FAS essentially supported this view as will be discussed later (chapter "5.2 Engineering for chain length control *in vitro*" for more information).

However, once engineering proved to be successful *in vitro*, additional studies were conducted in an *in vivo* environment, in *S. cerevisiae* to be more specific. The first step towards a possible application was taken; a goal that economically could never be achieved with an *in vitro* set up due to high costs of enzyme purification and expensive substrates. How our assumptions were confirmed and which new challenges derived from an *in vivo* set up will be discussed peripherally in a later chapter, "5.3.1 Moving into *in vivo* environments".

5.2 Engineering for chain length control *in vitro*

As starting point for engineering chain length control in *de novo* FA synthesis, a type I *C. ammoniagenes* FAS was probed *in vitro*[17]. Structural data on the homologous *S. cerevisiae* FAS served as a basis for protein design[10, 12-13, 15].

How the key catalytic reaction steps could be targeted was based on the following factors influencing chain length control: 1) The binding channel geometry allows certain chain lengths of intermediates to be bound and discriminates against others. This is especially important for the KS domain (in the elongation step) or the MPT domain (catalyzing the cleave off). 2) Substrate supply also determines if intermediates can be further elongated or if they are cleft off before they reach their full length. 3) Whether or not intermediates can be processed in a catalytic domain depends on the interactions of ACP with these domains.

All of these factors were addressed individually, but also in combination. They will be discussed in detail in the following paragraphs and first conclusions towards their interplay will be drawn supported by kinetic computer model calculations.

5.2.1 KS

The KS domain was intuitively the most obvious point to start engineering. It is involved in what was long before described as the decisive step in FA synthesis: A fully reduced acyl chain bound to the ACP and its decision between being loaded into the KS domain or released from the enzyme via the MPT domain[82-83]. As far as the KS mechanism is concerned, it can be divided into the ping step (the ac(et)yl transfer in the KS active site) and the pong step (the decarboxylation of malonyl which is followed by the condensation reaction) (Fig. 23)[87]. In between of the ping and pong step, the ACP has to be loaded with malonyl by the MPT domain.

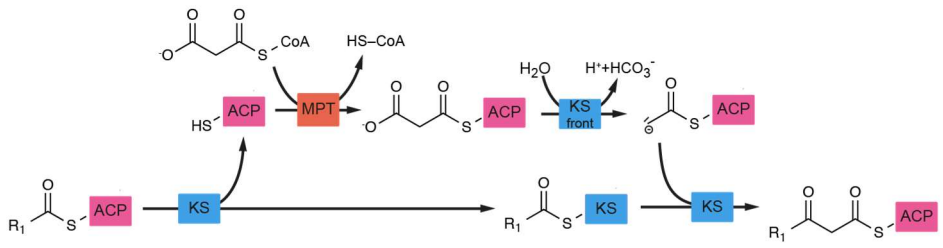


Figure 23: Reaction steps in KS mechanism.
First, the ac(et)yl is transferred into the KS domain where it then binds to the active site cysteine. In the following step, the unoccupied ACP has to be loaded with malonyl in the MPT domain. When the malonyl-ACP returns to the KS domain, malonyl is decarboxylated in the front part of the KS domain yielding a carbanion. The latter then reacts with the KS-bound acyl chain in a Claisen condensation. Thus, the bond of the acyl towards the KS domain is broken and the new ketoacyl-ACP intermediate is formed.

The engineering performed during this thesis initially focused on the modulation of the binding channel of the KS domain. In detail, three mutation sites were tested *in vitro* on *C. ammoniagenes* FAS. Their positions in the binding channel can be assumed from the KS structure of the homologous *S. cerevisiae* FAS (Fig. 24)[12, 15]. The introduction of the G2599S mutation increased levels of short FA approximately 7-fold (see Fig. 7a,b; under the assumption that the wild type yield was correctly extrapolated; it was outside of the calibration range), similar to the previous reports on the equivalent mutation in *S. cerevisiae* (G1250S)[8-9]. Looking at product

distribution, C_8-CoA is, however, only a side product. The majority of the detected acyl-CoA esters is C_{16}-CoA (64.8%).

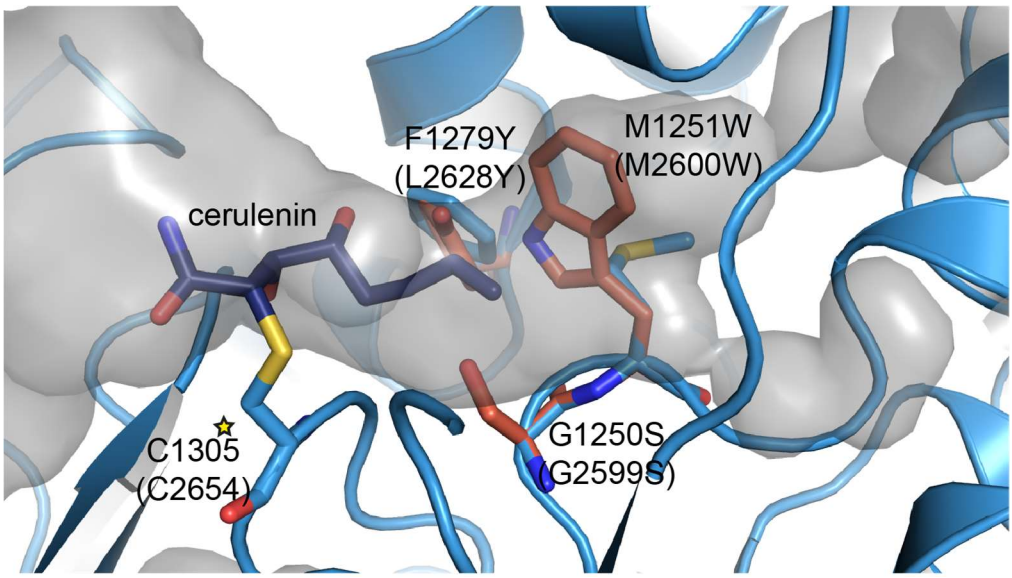


Figure 24: KS domain with mutated residues.
Here, on the example of the *S. cerevisiae* FAS, the KS domain is shown in cartoon depiction (based on PDB: 2VKZ[12]). The active site (C1305, marked with *) and the mutation sites, with their native residues (in blue and numbered) and mutations simulated by simple residue replacement (in red and numbered in brackets), are labeled explicitly. Mutated residues point directly into the binding channel (in gray) which in this structure is occupied by cerulenin, an inhibitor mimicking a short acyl chain.

The second mutation site M1251 was identified from structural data of *S. cerevisiae* FAS[10, 12]. In the past, a potential role of an equivalent methionine in a KS domain for chain length control was assumed[16]. Its mutation to bigger, bulkier residues was thought to introduce an efficient block into the KS binding channel, which would inhibit the elongation of intermediates longer than a certain chain length. Testing samples of G2599S-M2600W-mutated *C. ammoniagenes* FAS *in vitro* showed higher yields of short chain products: The concentration of C_8-CoA was once more increased 2.9-fold over the sole G2599S mutant. C_{16}-CoA still made up the majority of the detected CoA esters (66%) but C_{18}-CoA dropped below the limit of detection (see Fig. 7a,b).

For all *C. ammoniagenes* FAS constructs, their activity, i.e. their maximum activity under substrate excess (v_{max}), was determined by quantifying their NADPH consumption[1]. The results for G2599S and the G2599S-M2600W double mutant show a clear decrease compared to the wild type activity with typically around 50 - 150 mU/mg of protein in contrast to 310 mU/mg (for the wild type; 1 U was

defined as the incorporation of 1 µmol of malonyl-CoA per minute). The only change in these FAS constructs was the introduction of one or two mutations in the KS, which supports the hypothesis that the reaction step catalyzed by the KS domain has become the rate limiting step in FA synthesis. As no direct kinetic data was obtained for individual reaction steps on FAS, such a finding is susceptible to misinterpretation. Clearly, a modification in the KS channel does not necessarily only impact the KS reaction: a generally lowered stability of the enzyme could cause a drop in activity. In a broader picture, decreased activities of engineered enzyme are common[145].

In addition, a third mutation site was identified in sequence alignments where regular fungal and bacterial FAS were compared to FAS from *Aspergillus* which were supposedly specialized in short FA production as part of secondary metabolite pathways[19-21]. A tyrosine was found in a position where typically leucine or phenylalanine occurs. In the KS binding channel, the mutation is located at the opposite side to the previously described G2559S and M2600W mutations in the corresponding *S. cerevisiae* structure (see Fig. 24)[12]. The corresponding mutation in *C. ammoniagenes*, L2628Y (equivalent to F1279 in *S. cerevisiae* FAS), showed remarkable results when tested in *in vitro* assays (see Fig.15): The intrinsic system of chain length control was overwritten and no long chain products could be detected anymore. Instead C_8-CoA became the dominant product making up 74% of all detected CoA species.

It, however, can only be hypothesized why the *C. ammoniagenes* FAS^{L2628Y} produces C_8-CoA instead of a C_6 product which the *Aspergillus* FAS, that inspired this mutation, must produce for norsolorinic acid (the precursor of Aflatoxin B_1)[21]. Minor differences in the binding channel an in the positioning of the residues can be given as a possible reason here.

Interestingly, FAS^{L2628Y} showed only a limited activity of 5.8 ± 3.5 mU/mg (protein). In comparison to the other mutated FAS constructs (e.g. in other KS mutants and the ACP surface mutated constructs), this low activity correlates with specificity of short FA output. When indeed, the KS catalyzed reaction becomes the rate limiting step, the two competing reactions in chain length control, elongation and product release[82-83], are shifted towards the latter in the KS mutant construct since all other catalytic domains were left untouched.

On a different note, the alignments also revealed that the G2559S mutation in the KS channel naturally occurs at the equivalent position in the *Aspergillus* FAS. Even without the G2599S mutation, the L2628Y was sufficient in blocking the binding channel in *C. ammoniagenes* FAS. In fact, when the mutations were combined in *C. ammoniagenes* FAS, the double mutant G2599S-L2628Y only produced very little amounts of acyl-CoA esters. The same was found for the triple KS domain mutant G2599S-M2600W-L2628Y. Even though, the individual mutations performed well, when combined, the three mutations seemed to cause crowding in the KS channel and eventually most of the synthetic activity was abolished (see Fig. 7 and Fig. 14).

Besides the exceptional results that we obtained in these experiments, our analysis aimed at more detailed insights of the chain length mechanism. Even for FAS in their wild type form, the exact mechanisms of chain length control are still up to debate. In publications on KS domains of a PKS system, both ping and pong step were found to be dependent on chain length[84-85].
While by itself, the data connecting mutations and short FA output does not directly deliver the answer to how exactly the FAS reaction networks are changed, they can serve as a good basis for the discussion on how chain length control is influenced in the different reaction steps of FAS. The two possible steps in the KS reaction sequence, 1) the transfer of the acyl chain into the KS domain (ping) and 2) the decarboxylation reaction of malonyl-ACP followed by the actual condensation reaction of the enolate carbanion (pong, see Fig. 23), will therefore be discussed in detail in the following paragraphs.

5.2.1.1 Loading acyl chains into the KS domain (ping) as decisive step

Intuitively, introducing bigger, bulkier residues in the middle of the KS binding channel can easily lead to the assumption that they block the binding channel and thus, directly interfere in a gatekeeper like fashion with the loading of the acyl chain into the KS domain (ping step of the synthesis). For the mutations G2599S and M2600W it can, if at all, only be partially true, since the block is far from complete as long chain products are still the dominant species. In case chain length control really happens in the ping step of the KS reaction, the mode of action of G2599S and

M2600W can best be described as partially inhibiting the loading and as a consequence, elevating the chances of product cleave off by the MPT as the resting time of short chain on the ACP is increased.

Molecular dynamics simulations on the G2599S-M2600W-mutated equivalent in *S. cerevisiae* (there, G1250S-M1251W) pointed in a different direction at first. In fact, the simulations performed by the group of Prof. Grubmüller from the MPI in Göttingen indicated that short acyl chains like C_8 have a higher affinity in the KS binding channel after the introduction of the G2599S-M2600W mutation as compared to the wild type. This finding clearly contradicted our previous expectations of the mutations acting as a gatekeeper.
In the initial kinetic model, the higher affinity of C_8 to the KS domain also did not show any increased levels of short FA products. Only when the kinetic model was revisited and the newer version with fewer restrictions was implemented, a fraction of the tested parameter sets showed more FA as an output. In these cases, the G2599S-M2600W equivalent mutation was connected to a slower loading rate (and slower rates for its reaction in the other direction) despite the better final binding. As the revised model was adapted to some extent to reproduce the experimentally found results, its value for conclusions on the KS mechanism is limited. However, the model underlines the importance of KS loading in chain length control in the KS domain.

Previous publications on the role of methionine clearly favored the theory of a blocked loading when bulky, hydrophobic residues are located in the KS binding channel[7, 16]. While no final statement can be made for the effects of mutation equivalent to G2599S-M2600W on the ping step, inhibition of acyl transfer into the KS domain or slowing it down, can be seen as a likely mechanism in chain length control. How the mechanism works exactly, could also vary from one mutation to the other.
In the case of FAS^{L2628Y}, the mechanism was not studied as intensively and no molecular dynamics simulations were performed. From the disappearance of any long chain products, it can be concluded that FA exceeding a certain length cannot be elongated by the KS domain anymore.

5.2.1.2 Decarboxylation of malonyl-ACP and condensation reaction as the decisive step (pong step)

As a second possibility, it is also plausible that the ping step is indeed not the only or not even the most significant step for chain length control. Once the acyl chain is loaded into the KS domain, it still has to be elongated. A study on related KS domains from PKS systems for example suggests that substrate selectivity is higher in the pong step (elongation) than in the ping step (acylation)[85].

Even before the actual condensation reaction takes place, a changed binding of an intermediate in the KS binding channel could influence the decarboxylation reaction of the malonyl-ACP: The front end of the KS domain catalyzing the decarboxylation is sensitive to the loading state of the KS domain. If this was not the case, malonyl-ACP would continuously be decarboxylated and thus be inactivated even though there was no substrate bound in the KS domain. The process actually has been observed in FAS but only at very low basal rates[141].

If the binding of acyl chains to the KS was changed in a way that decarboxylation is inhibited or not favored anymore, no carbanion could form and thus, no elongation could happen. The acyl chain could in fact be loaded into the KS domain but without any elongation taking place. Since acyl chain transfer into and out of the KS domain is a reversible process, the acyl chain would eventually be transferred back on ACP and released from the enzyme. This theory is, however, difficult to be proven in type I systems: All enzymatic domains are located in one megasynthase complex and thus, the specific acylation of the KS domain is hard to quantify with ACP and transferases being intact (needed for loading of KS).

Following a successful acylation and decarboxylation of malonyl, the latter could be protonated to yield acetyl. Subsequently, malonyl is lost as an elongation unit[141]. It was reported that when certain inhibitors (in this particular case it was iodoacetamide) were bound in the KS active site in *S. cerevisiae* FAS, the decarboxylation and reaction to acetyl was increased substantially [130]. Applied to our case, it can be hypothesized that malonyl decarboxylation takes place but the positioning of the generated enolate carbanion and the acyl chain (bound to the KS)

could be distorted due to a changed binding geometry leading to the formation of acetyl instead of the β-keto intermediate. The exact angles in reactions can be considered as highly sensitive and minor distortion could lead to a diminished rate of successful condensation reactions.

If the criteria described above applied, there would be a lower rate of successful elongation reactions and at the same time, malonyl would be converted into acetyl which itself could serve as a starter unit eventually leading to more short FA *in vitro*[83, 93].

This theory of malonyl decarboxylation to yield acetyl without successful elongation steps was, however, judged to be unlikely after one key experiment: The G2599S-M2600W-mutated and the wild type *C. ammoniagenes* FAS were probed with ^{13}C-labeled acetyl-CoA. The product analysis did not show a noteworthy increase in products for the mutated FAS that could be traced back to decarboxylation of the unlabeled malonyl-CoA when compared to the wild type (see Supplementary Fig. 11).

Decarboxylation of malonyl without a successful condensation reaction can be ruled out as an important factor in chain length control, while both other options (slower or blocked acylation of the KS domain and inhibition of the decarboxylation reaction) seem more likely.

5.2.2 MPT

The MPT domain has a dual function in FAS: On the one hand, it is responsible for the transfer of malonyl, the elongation unit, and on the other hand, it also catalyzes the cleave off of products, typically C_{16}- and C_{18}-CoA. Early publications identified it as a key player in chain length control as it is the antagonist of the KS domain: The MPT releases products instead of letting them be elongated again in the KS domain[82-83]. Additionally, it has been known for a long time, that MPT does not show a distinct specificity in transfer as far as chain lengths are concerned[94-96].

Our first attempts to increase short FA output aimed at shifting the competition between elongation (in the KS domain) and product release (via the MPT domain) in favor of the latter by facilitating the transfer of short acyl chains by the MPT domain. As no structural data on MPT domain with an acyl chain bound to it have

been reported so far, molecular dynamics simulations based on the structure of an "empty" MPT in *S. cerevisiae* FAS were conducted by our cooperation partner, the group of Prof. Grubmüller from the MPI in Göttingen. The conformational changes induced in the MPT by binding a long acyl chain were studied. However, results remained speculative to some degree as the dynamic modeling of the acyl binding tunnel had to be based on a series of assumptions.

Our aim was to either stabilize this active transfer conformation or to modify the binding channel to accept shorter acyl chains more easily. Two glutamine residues (Q1775 and Q1839 in *S. cerevisiae* FAS) were identified and corresponding residues at said position were mutated to a variety of hydrophobic amino acids in *C. ammoniagenes* FAS (there, residues Q1349 and H1413) which already carried mutations in the KS domain. The constructs did not show any increase in short FA production *in vitro* (see Fig. 11). In the end, it can be assumed that this approach most likely contained too many points of uncertainty: 1) Binding of long acyl chains could not be elucidated without final doubts and 2) the conclusions drawn from the molecular dynamics simulations might not be easily transferrable towards the binding of short acyl chains.

Reactions catalyzed by the MPT were revisited and an inverse approach was followed, essentially elucidating an aspect in chain length control that had not been included in the historical view[82-83]: Cleave off of products is not only dependent on the specific affinity of acyl chains towards the MPT domain (where there is little discrimination[94-96]), but transfer is also heavily influenced by the availability of malonyl. Both possible ligands, malonyl and acyl chains, compete for the same active site. Thus, negatively influencing malonyl binding would shift the balance in competition for the active site in favor of product release. Furthermore, as a second consequence, a lowered malonyl affinity towards the MPT domain can be understood as if there was a lower malonyl concentration on the enzyme. This condition (mostly represented in lower malonyl-CoA to acetyl-CoA ratios) was already linked to increased production of short FA in previous *in vitro* studies[83, 93].

The engineering for lower malonyl affinity towards the MPT domain was encouraged by studies on related human malonyl/acetyl transferase[22-23]. Malonyl binding in transferases is highly conserved with an arginine coordinating the free carboxyl

group of malonyl (Fig. 25). The mentioned studies tested several mutations of the described arginine with decreasing malonyl affinity (Arg→Lys/Ala and Arg→Gln/Gly/Ala, respectively).

Applied to our case, engineering had to be carefully balanced as a decrease in malonyl transfer was likely to decrease total FAS activity and in the most extreme case abolish synthetic activity completely. When incorporated into the double KS mutant in the *C. ammoniagenes* FAS, only $FAS^{G2599S-M2600W-R1408K}$ showed activity and products in assay. While the output of C_8-CoA was not increased compared to the reference ($FAS^{G2599S-M2600W}$), specificity towards short chain products was tremendously increased: C_8-CoA made up 47% of the detected acyl-CoA esters while the share of long species (C_{14}- to C_{18}-CoA) dropped to 40%.

In all of the other mutations to alanine ($FAS^{G2599S-M2600W-R1408A}$), glycine ($FAS^{G2599S-M2600W-R1408G}$) and glutamine ($FAS^{G2599S-M2600W-R1408Q}$), no product were detected. The fact that only lysine had the ability to maintain synthetic activity led us to believe that a positive charge as provided by arginine or lysine is necessary to coordinate the carboxyl group.

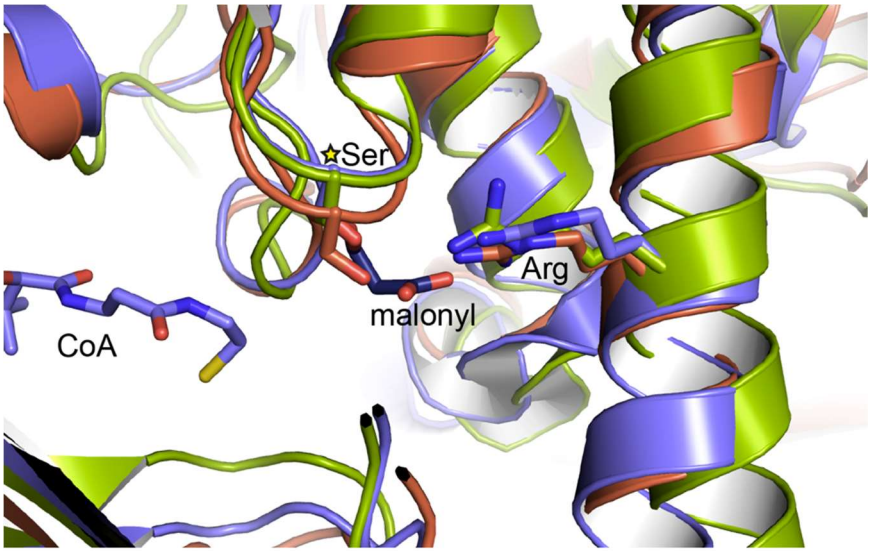


Figure 25: Malonyl binding in transferases.
Besides the active site serine (Ser, marked with *), a central element in malonyl binding transferases is an arginine (Arg) which coordinates the free carboxyl group of malonyl. Three relevant structures are shown in cartoon depiction: in red, the MPT domain of *S. cerevisiae* (PDB: 2UV8[10]); in blue, fabD, a malonyl transferase from *E. coli* structure with bound malonyl and CoA (PDB: 2G2Z[146]) and in green, the human malonyl/acetyl transferase that was used in the study to guide MPT engineering (PDB: 2JFD[23]).

For a better mechanistic insight, the R1408K mutation was further studied in molecular dynamics simulations on the equivalent position in the *S. cerevisiae* MPT structure (R1834K). While binding of long acyl chains was not influenced by the mutation, for malonyl a change in binding by $\Delta\Delta G = +14.2 \pm 1.1$ kJ/mol was observed. This weaker binding was in line with our expectation and incorporating it in the initial kinetic model (as a single mutation) and the revised kinetic model (in combination with the KS domain mutations) both led to an increase in short FA (see Fig. 12 and Fig. 10).

In sum, engineering the MPT domain for lower malonyl affinity and thereby shifting the competition of malonyl loading and product release in favor of the latter proved to be a successful approach for using the MPT domain in chain length control.

5.2.3 AT

As a second transferase, the AT domain was a target in our engineering strategy. While it might not be the first starting point for chain length control, several factors indicated that AT engineering can be harnessed for shifting the product spectrum towards shorter FA.

Substrate ratios of acetyl and malonyl have been shown to influence product distributions in type I FAS[83, 93]. Accordingly, more short FA would be produced if the AT domain was engineered for a better acetyl uptake. Generally, a successful engineering for improved catalytic activity seems unlikely as it can be assumed that yet, it is evolutionary optimized. However, there are reasons that speculatively speak against immediate selection pressure for highly optimized acetyl uptake: 1) Acetyl is needed only once during the synthesis. 2) When an acyl chain is deposited in the KS domain waiting for elongation, the free ACP has to be loaded with malonyl (and not acetyl) to ensure the successful continuity in the synthesis cycle. 3). Above that, intracellular levels of acetyl-CoA are typically higher than those of malonyl-CoA[147].

For our engineering of the AT domain, the studies on human MAT domains once again served as a starting point[22-23]. The authors not only observed that malonyl affinity could be lowered in their mutational studies but also that acetyl binding affinities were in fact increased. In addition, the introduced mutations allegedly even opened up a binding channel in MAT for bigger acyl chains e.g. for C_8.

Applied to our case, the engineering could have two beneficial consequences: 1) The elevated acetyl loading would act as if the effective acetyl concentration was increased and 2) a second way for product cleave off for short FA (besides MPT domain) would be installed. Both cases would lead to higher output of short chain products.

MAT and pure AT domains share high structural homology but also differ in important aspects: Since AT do not transfer malonyl, they lack the central arginine. However, mutating the corresponding position, an isoleucine, to a smaller residue such as alanine was still tested *in vitro* on *C. ammoniagenes* FAS (there, I151A). In combination with G2599S M2600W, the total yield of C_8-CoA was increased in comparison to the G2599S M2600W reference (see Fig. 7). Also, the share of C_8-CoA was higher than that of $FAS^{G2599S-M600W}$, but not as selective as the KS MPT construct $FAS^{G2599S-M600W-R1408K}$, e.g. the share of C_8-CoA was 26% compared to 47% (of the detected CoA esters), respectively.

Molecular dynamic studies using the *S. cerevisiae* structure were also conducted in order to gain a deeper insight into the molecular impact of the AT mutation. In a first assessment, an influence on acetyl binding to the AT domain was judged as unlikely since an interaction of the mutated residue and the bound acetyl would have to bridge a large 7 Å gap. However, further simulations were indicating a stabilization of short acyl chains, such as C_8, in the binding pocket as postulated by the earlier publications on the human MAT[22-23]. With this in mind, the contribution of the first possible effect, an increased acetyl binding, can be judged as negligible in comparison to the second one, the possibility to cleave off short acyl chains also via the AT.

The findings were confirmed in the initial kinetic model as well where short FA transfer via the AT was integrated. As a consequence, shorter FA were slightly increased. When probed in the revised model, where also the influence of the G2599S M2600W mutations was considered, the results of the *in vitro* data were fairly reproduced.

5.2.4 ACP domain-domain interactions

The previous engineering approaches targeted substrate binding affinities in binding channels. However, it has been long known that the synthesis on megasynthases is also highly dependent on domain–domain interactions, i.e. the interaction of ACP with the catalytic domains.

In our reasoning, the decision of whether or not an ACP bound intermediate is processed in a catalytic domain depends on 1) the intermediate itself (if it can be correctly positioned in the binding pocket), but 2) also on the interaction of ACP with said catalytic domain for bringing the intermediate close to the binding pocket. Furthermore, it has been illustrated for type II systems that the recognition of the ACP with the catalytic domain happens before the binding of an intermediate in the binding pocket[102-103]. On the other hand, the question if this two-step mechanism of ACP–domain recognition followed by chain flipping mechanism applies also to type I systems remains highly disputable since intermediates are far less sequestered in the ACP[100].

In a general sense, the recognition of ACP towards a catalytic domain can be judged as very important. In the end, numerous attempts of failed PKS engineering can be traced back to incompatible ACP–domain interactions. Studying these interactions has long been focused on the identification of important residues on both the ACP and the corresponding catalytic domain as was reported in various publications. In contrast, our approach aimed at changing these interactions to direct synthesis. Within the reaction cycle, the engineering therefore has to target points, where ACP faces the decision of interaction with one of at least two possible reaction partners. By modulating the interaction sites on the catalytic domains, the probability of one path or the other would be more likely as a consequence. To avoid interference in other parts of the reaction cycle, the ACP itself was left in its native form. The KS and the MPT were chosen as first engineering targets, as they are directly involved in chain length control, which will be discussed in more detail below.

Above all, the engineering had to be carefully balanced as excessively interfering in the recognition could possible abolish the interaction altogether – in our case, the synthesis would be stalled then, which had to be avoided.

In order to quantify possible effects in this proof of principle study, we chose to set the mutations in a *C. ammoniagenes* FAS already carrying the G2559S M2600W

mutations. This FAS reporter system set up enabled us to read out possible influences on chain length control more easily since it already had a bimodular product spectrum, where both short and long chain acyl-CoA esters were produced. Mutations modulating the ACP–domain interactions set in the G2599S M2600W background that would for example increase short chain production 2-fold, could more easily be detected than the equivalent 2-fold increase in the wild type background (where short chain products are only a hardly detectable side product). Also, a possible unexpected decrease in short chain production could simply be read out in the G2599S M2600W-mutated reporter FAS system.

5.2.4.1 KS-ACP

The interactions of the KS-ACP were chosen as it represents a decisive point in the fatty acid synthesis: As previously established, the decision on chain length is determined when a fully reduced acyl chain bound to the ACP is either loaded into the KS domain for another round of elongation or cleft off the enzyme by the MPT domain[82-83]. The mutations therefore aimed at decreasing the ACP's affinity towards the KS domain which could have several beneficial effects on short FA production: 1) A fully reduced acyl chain is less likely to be deposited in the KS for further elongation (ping step). The time it is additionally bound on the ACP elevates the chances of a premature product cleave off. 2) Even if the acyl chain was deposited in the KS domain and ready for elongation, a decreased affinity of ACP towards the KS could lower the chances of a malonyl-ACP returning to the KS domain (pong step). Since transfer reactions are reversible, malonyl could be cleft off the ACP again. The acyl chain could also be moved to the ACP again without being elongated and then be cleft off by the MPT domain.

A series of mutations on the surface of the KS domain was tested in the reporter FAS background (G2599S M2600W), with very promising results (see Fig. 17a). For five mutations, short acyl-CoA production was increased wherein the three mutations N2557E, N2621D and D2622A showed the highest amounts of short acyl-CoA esters with high specificity (e.g. shares of C_8-CoA of up to 77%) while long chain production decreased to a minimum.

Interestingly, the best performing constructs showed the lowest activities in the activity tests. It can be hypothesized that with the interference in the ACP-KS

interaction, the condensation reaction was indeed limited as planned. In all cases, at least a low activity and some products were detected which also indicates, that the interference in the ACP–KS interactions was not excessive. A combination of several mutations on the KS surface could have such an effect, but was not tested in the study.

In contrast to our expectations, we also found mutations on the surface of the KS domain, that led to a lower level of short FA (see Fig. 17a). In our line of reasoning, this would mean, that the probability of the condensation step was in fact not decreased but increased by the specific mutations on the KS surface. While the data truly points in this direction, some doubt remains as to if our mutations actually increased the affinity of the ACP towards the KS domain. An optimization of an ACP-domain interaction by the introduction of mutations can be judged as generally possible since there is a case where it has been reported for type II systems[148].

In sum, the results of this study can be regarded as the first proof of principle study that exploits the modulation of domain-domain interactions to direct synthesis towards specific products. The engineering approach is a subtle way of modifying an enzyme as the binding channels are not further changed (besides the preexisting mutations as in our reporter FAS for example). It is also worth to note, that in this study only the exchange of single residues on the surface of the KS domain was able to increase product shares of C_8-CoA from 12% in the G2599S M2600W reference to for example 77% in the FAS carrying an additional N2577E mutation.

5.2.4.2 MPT-ACP

Besides the KS domain, the interaction of the ACP with the MPT domain was also targeted in our approach of modulating domain-domain interactions. The MPT domain plays an important role in the condensation reaction as well: When an acyl chain is deposited in the KS domain, the free ACP is loaded with malonyl and then returns to the KS domain for the actual condensation reaction (see Fig. 1). It should be mentioned once again, that the ACP movement generally follows a stochastic pattern, though[37]. The supply of malonyl has a big influence on product spectra as shown before in our mutational studies on the MPT domain (see chapter 4.1.1 “Engineering the transferase domains”) or previous publication[83, 93]. Limiting this supply by interfering in the ACP-MPT interaction was essentially our goal.

Judging from the reaction cycle, a diminished interaction of ACP and MPT could however also influence the product release step negatively (see Fig. 1). Even though ACP statistically only has to interact with the MPT once for product release and several more times for fetching malonyl during the buildup of one acyl chain (3× in the synthesis of octanoyl-CoA and 7× for palmitoyl-CoA), we wanted to facilitate the product release step by engineering an alternative exit pathway for short acyl chains via the AT domain.

In the product spectra measurements, the combination of mutations on the surface of the MPT and in the AT domain did not show the desired higher output of short FA (see Fig. 17b). The differences in the product distributions in comparison to the reporter FAS reference (G2599S M2600W) were generally low and an interpretation was limited. It can be hypothesized that the rate by which ACP picks up malonyl from the MPT domain was not lowered enough to turn it into a bottleneck in FA synthesis. Either the introduced mutations on the MPT surface were not strong enough to interfere into the interaction, or product release was effected to a similar extent and not rescued by the alternative exit pathway in the mutated AT domain. Also, activities were still high, which could be taken as an additional indication that malonyl supply was not limited in a sufficient degree.

As the potential to direct synthesis had already been shown in the case of the ACP-KS interactions, the approach to modify the surface of the MPT should not be simply given up. In the future, more studies on the interaction of ACP and transferases can possibly provide better evidence of important residues for the interaction with ACP. Furthermore, a bigger library of potential mutation candidates could raise the chances of finding a successful mutation to also exploit the modulation of the ACP-MPT interactions for directing synthesis.

5.2.5 Interplay of domains and understanding network complexity

Our approach aimed at rational engineering of specific binding sites (for substrates and domains), but at the same time, their role in the reaction network needs to be understood. FA synthesis is a whole network of reactions that in case of type I FAS all happen on one enzymatic complex. The condensation, the two reduction steps, the dehydration, and all transfer reactions, highly depend on each other. In our mutational studies, we also saw that effects do not show a direct additive behavior.

For example, the introduction of the mutation in the AT domain (I151A) led to more short acyl-CoA esters when introduced into the background of the KS mutations yielding $FAS^{G2599S\text{-}M2600W\text{-}I151A}$, but when added to a construct carrying KS and MPT mutation (G2599S M2600W R1408K), it seemed to have the opposite effect in this new construct, $FAS^{G2599S\text{-}M2600W\text{-}R1408K\text{-}I151A}$ (see Fig. 7). In the end, chain length control and any other process on FAS are determined by their kinetic rates and the kinetic rates of competing processes.

The development of a computer model of FAS kinetics (in cooperation with the group of Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen) helped in that way (see Fig. 9). The analysis of different parameter sets for FAS kinetics shed light on interplay of processes that were not directly accessible otherwise. However, it should also be mentioned that the initial version of the kinetic model failed to explain the biochemically found product spectra of the KS mutations (G2599S M2600W in *C. ammoniagenes* FAS). When possible effects of these mutations were loosened, a revised model produced scenarios which matched our experimentally obtained data.

5.3 Engineering beyond chain length control

5.3.1 Engineering FAS for the production of a polyketide

Exceeding the previously described engineering for chain length control, a production pathway was established where two modified FAS work in sequence to produce a simple polyketide[17]. Therein, the first FAS, module 1, was constructed to mainly yield C_8-CoA. The engineering relied on the previously described modification of *C. ammoniagenes* FAS with mutations in the KS, AT and MPT domain or combinations thereof.

The second FAS, module 2, was constructed to take up C_8-CoA and elongate it further. A KR knockout, Y2227F, was introduced essentially blocking the proton transfer necessary for neutralizing the hydroxide anion after hydride transfer to the carbonyl group by NADPH[139]. This mutation has previously been characterized as producing triacetic lactone (TAL)[123, 140]. After the first elongation in the KS, the intermediate consequently could not be reduced but instead waited until it was elongated again in the KS, yielding the triketide. The product, 6-heptyl-4-hydroxypyran-2-one (6-HHP), then formed after a spontaneous lactonization and thereby also cleft itself off the enzyme. The second module also carried an AT knockout, S126A, to avoid acetyl to be loaded. Thus, the formation of TAL as a side product was to be avoided.

In our engineering, the native programming for production of long FA was overwritten by only five mutations distributed over two FAS modules, and a sequential pathway for the production of a simple polyketide was created. For testing the potential of this system, one-pot *in vitro* product assays were performed and optimized in two steps: In order to scan for the best production yield, several parameters such as different constructs for module 1, different relative amounts of module 1 and module 2, and finally a whole set of combinations of substrates were tested. The overall yield (in reference to the limiting substrate) was pushed to 35% (see Fig. 19). Considering that two FAS were involved with a total number of 14 enzymatic reaction steps (not counting transferase reactions), this result can be considered a success.

From the obtained results, some preliminary conclusions can be drawn: (1) When a reaction sequence exceeds a single module, the complexity of the reaction network increases to a point, where it cannot be easily understood anymore. (2) Some trends,

such as an increased acetyl-CoA level, previously connected to higher output of short FA[83, 93], were also beneficial in the coupled assays. (3) Extremely low acetyl-CoA concentrations, on the other hand, also led to high yields. Most likely this phenomenon can be linked to a lowered synthesis of short FA by module 1, and consequently a lowered usage of malonyl-CoA which then is still available in sufficient amounts for module 2.

Altogether, this study once more underlines the similarities of FAS and PKS and set an example for how FAS can be used for studying PK production. It should, however, also be clearly stated that the overall architecture of fungal and bacterial FAS differs tremendously from PKS, which are believed to resemble more the mammalian FAS structure[51]. Mammalian FAS therefore will serve as a better model enzyme when investigating PKS processes using FAS in the future.

5.3.1 Production of short FA *in vivo* using *S. cerevisiae*

As already discussed in chapter "5.1 Settings for probing biotechnological engineering efforts", probing enzyme engineering in an *in vitro* environment has a clear advantage of gaining information on structure-function relationships more easily. However, economic feasibility is most likely achieved when *in vivo* fermentation can be applied since costs of protein purification or pricy substrates are avoided.

For moving the production of short FA or derivatives thereof into an *in vivo* system, a cooperation with the group of Prof. Eckhard Boles from the biology department at the Goethe University was initiated where *S. cerevisiae* was chosen as a host[28-29]. In comparison to *E. coli*, the second model organism often used in oleo-chemical production, *S. cerevisiae* is thought to have generally lower lipid production but has advantages such as its tolerance of industrial fermentation conditions, high cell densities and immunity towards phage infestations[6, 149]. Furthermore, *S. cerevisiae* is well studied and expansive genetic tools are available[150]. The endogenous FAS system is also type I and was in fact used as a structural basis for engineering of the *C. ammoniagenes* FAS (see chapter "5.2 Engineering for chain length control *in vitro*").

The transfer into the *in vivo* system of *S. cerevisiae* adds a new complexity to our research since the *de novo* FA synthesis is embedded into a tightly regulated

metabolic network[34]. It was of high interest to see if the *in vitro* findings would hold up in an *in vivo* environment, despite the fact that short free FA have been shown to cause membrane stress and inhibit growth[143, 151].

Encouraged by the initial *in vitro* study on engineering *C. ammoniagenes* FAS for short FA production, mutations were transferred into the *S. cerevisiae* FAS. In comparison to the *C. ammoniagenes* FAS, the *S. cerevisiae* FAS consists of an α-chain and a β-chain, which closely resembles a split version of an *C. ammoniagenes* FAS[15]. The *S. cerevisiae* FAS correspondingly forms a $\alpha_6\beta_6$ heterododecamer as its functional oligomer. While its suitability for extensive *in vitro* studies was not given due to limited access to high amounts of protein at the time, the higher reported activities (2500 mU/mg[152-153] compared to 370 mU/mg for the *C. ammoniagenes* FAS[137]) and the general advantage of expressing native enzymes in their native host[119], led to the decision of modifying the endogenous *S. cerevisiae* FAS.

The different FAS were constructed by combining the known mutations from the *in vitro* studies and were then tested in a *Δfas1 Δfas2 S. cerevisiae* strain. By using low copy vectors with the α- and β-chain encoding sequences *FAS2* and *FAS1* and natural promoters and terminators[25], native FA synthesis was mimicked and represented the only source of *de novo* FA synthesis. Accordingly, the FAS systems were re-introduced on plasmids and had to fulfill a second task of providing sufficient long FA to ensure cell growth as well. Being far from optimized for highest titers, this system enabled a fairly direct read out of structure-function relationships.

The initial experiments did not find significantly increased levels of short FA in the *S. cerevisiae* cells. Short fatty acids were, however, found extracellularly as described in previous studies on short FA production in *S. cerevisiae*[30]. As *S. cerevisiae* FAS, in fact, produces acyl-CoA esters as their final product, only the recent publications on endogenous TE in *S. cerevisiae* were able to deliver an explanation for the high amounts of free FA that we found[26-27].

While in the wild type *S. cerevisiae* strain, short FA are only produced in low concentrations as a side product, the yields were increased in most cases when the mutated FAS were introduced into the cells. In several examples, cell growth was however severely inhibited, which can most likely be traced back to a lack in supply of long FA by the modified FAS. When supplemented with $C_{18:1}$ growth was restored in all modified strains.

The highest titers of approximately 118 mg/L were obtained for one construct carrying mutations in the AT, MPT and KS domain (I306A-R1834K-G1250S). In comparison to previous reports on short FA production in *S. cerevisiae* where 111 mg/L of mostly C_8 was obtained, these results are competitive[30]. In the already mentioned study, the approach was, however, different to our method. A human FAS with an exchanged TE was introduced in *S. cerevisiae* on a plasmid. The direct product of the human FAS was a free FA accordingly. The FA would have to be activated again before it could be processed further; a step that can be skipped if products are acyl-CoA ester as in our case and directly channeled into the next synthesis step[79]. Other differences include the use of a strong promoter, ADH2, instead of the native promoter employed in our study[25, 30, 154]. Without any optimization for highest general FA output, rational engineering was slightly superior and could easily be further improved e.g. after an exchange of the promoters to ADH2.

Besides the reached high titers for total short FA yields, some of our tested strains also showed a high specificity for certain chain lengths. For the production of C_6, the strain with mutations I306A-G1250S showed the highest specificity with 20 mg/L, making up a 90% share of all detected short FA. For C_8, the best specificity was obtained for the I306A-R1834K-F1279Y-mutated strain (48 mg/L, 89% share of C_8). How the individual mutations contribute to the modulation in chain length, was not visible at first sight. Effects seem to heavily influence each other in the *in vivo* environment and not always act in the same direction. As the most dominant example, the I306A mutation in the AT domain should be mentioned. When introduced as the only mutation in the wild type FAS, short FA titers seem not to have increased at all, while when adding it to an existing KS mutation (to generate the I306A-G1250S-M1250W-mutated strain) levels of short FA are slightly increased (66 mg/L short FA as compared to 57 mg/L before). In other cases, the I306A mutation even lowered short chain FA production, e.g. when the I306A R1834K-mutated strain produced 37 mg/L of short FA in total, instead of 100 mg/L for the sole R1834K mutation.

From the direct comparison of strains (see Fig. 21 and Table 9), some conclusions were drawn as to what influence individual mutations tend to have: For the mutations G1250S or I306A, levels were rather increased for C_6 FA production, while M1251W

or R1834K showed to add to higher titers of C_8 (if there was a beneficial effect of the mutation records at all).

These trends sparked speculations as to how individual mutations mechanistically act on chain length control. For the R1834K mutation, this was primarily linked to solely reducing the supply of malonyl while the rest of the FAS enzyme remains unchanged. It can be assumed that *S. cerevisiae* FAS would produce C_8 in its wild type form also, but a high supply of malonyl overwrites this programming of chain length control pushing chain length distributions towards the long chains.

For a second mutation, G1250S, located in the KS domain, its contribution to primarily short chain production could be attributed to its position in the KS binding channel (see Fig. 24). Located closer to the active center (C1305) than for example M1251W, it is likely that 1250S already influences the binding of shorter intermediates, thus leading to higher increases for C_6 products, rather than C_8 products (as the neighboring M1251W).

5.4 Implications on other megasynthases

The studies presented in this thesis were all conducted on FAS, but their impact is not limited to it. The knowledge gained from FAS can help to draw conclusions for other known megasynthase classes, especially for PKS. This transfer of knowledge can be understood intuitively, as PKS share the same basic principles in their chemical synthesis with FAS. In both systems, substrate specificities of binding channels play a crucial role. For both FAS and PKS, a high interest in transferase engineering to expand the possibly accepted building blocks has emerged[155].

Also, the interactions of carrier proteins with domains have become an important issue. Research on how these interactions work has been tremendously expanded in the previous years[104]. Our study on modulating these interactions to direct synthesis can be taken as an example for future engineering.

Most promising results directly concluded from our studies could be expected for iterative PKS. It should be noted that for the large group of modular PKS, their working mode is different since they hand down intermediates from one module to another, which catalyze only one synthesis cycle each. Engineering on these modular PKS is also feasible, though.

The presented research of this thesis should therefore be seen as addressing specific issues on FAS, e.g. chain length control, but also as providing a toolbox for engineering, which has to be adapted to the specific enzymes, of course.

5.5 Outlook

In the production of short FA through rationally engineering FAS, the presented research offers first successful solutions. Future research can incorporate these findings or push different aspects of the project further in several dimensions. On all levels of complexity of this biotechnological research (as shown in Fig. 22), further developments would be beneficial:

For low complex *in vitro* testing, the identification of individual reaction mechanisms on FAS, such as the exact molecular basis of the condensation step in the KS or the dynamics of product release via the MPT, is likely to enable insights into the native mechanisms of chain length control, which could then again be modified and exploited. Faster screening methods for testing mutations on individual domains could help identifying more possible mutation sites. The most important question would then be: Would these mutations hold up when synthesis was tested on full-length modules? When a good correlation can be shown, testing on individual domains would, however, lower the costs.

Research on full length FAS as it has been mostly conducted in our studies, could provide valuable further progress, too. Based on the existing mutation sites, further combinations could be tested. The libraries of constructs could be extended to larger sizes, which exceeded the scope of this work. Synergetic effects have already proven to play an important role, both *in vitro* and *in vivo* and as more combinations are possible, new powerful combinations could be found just by testing. Their analysis and comparison could possibly deliver valuable insights for the exact mechanisms of the existing mutations.

In more complex systems, promising but complex steps could possibly push the limits further. These steps include the analysis of metabolic fluxes and identification of potential bottlenecks in the synthesis, which would finally be followed by strain optimization. With probing modified FAS in the *in vivo* environment of *S. cerevisiae*, a first step in this direction was already taken. Judging from developments in strain

optimization of the research group of Prof. Boles, producing FA in gram scale is not an unrealistic target for the next years to come.

As far as *in vivo* systems are concerned, testing other organisms could be beneficial. Promising results can be expected from both bacterial or fungal systems: *Corynebacteria* have been heavily used in industry[156] and their optimization towards higher lipid production further underlines their potential as a host. As an example for fungal systems, *Yarrowia lypolitica* bears high potential due to their higher basal production of lipids in comparison to *S. cerevisiae*[157].

There have also been recent developments towards directed evolution approaches, which differ clearly from our rational engineering, but could bring forward interesting results for the production of short FA: While selection for short FA producing strains is still not possible, the development of a GPCR (G-protein coupled receptor) detecting short FA like C_8, this problem could possibly be overcome in the near future[158].

In sum, it can be expected that the next few years will bring further developments pushing the field of short FA production but also related issues as the development of biofuels in the gasoline range or similar biotechnologically produced compounds.

5.6 Conclusion

If the biotechnological production of chemicals can further replace or support regular synthetic chemistry, industry will be able to move away from fossil oils towards renewable sources. However, in many cases the much needed adaptation of biotechnological production systems is not yet developed to the necessary level.

For processes where short FA are needed, as for example in the microbial production of biofuels in the gasoline range[6], protein engineering had not yet delivered feasible solutions. In this thesis, several approaches to introduce chain length control on type I FAS were established and made available in a publication and two patents[17-18, 28-29]. Therein, engineering was focused on rational design based on available structural information.

First, the type I FAS from *C. ammoniagenes* was used as a model enzyme to probe modifications on FAS in a low complex *in vitro* environment in order to gain information about structure-function relationships. At this stage, engineering was conducted in several rounds, first addressing possible ways to alter product

distributions by changing substrate affinities through concise mutations in binding channels. Several FAS constructs were generated ranging from first successes, where short FA were produced as side products, to FAS where native chain length programming was overwritten and only short FA were produced.

Furthermore, another engineering target was addressed with the modification of domain-domain interactions on FAS. For its exploitation to direct synthesis, contact surfaces on catalytic domains were changed to interfere with ACP binding. This channeling of the kinetic process on the enzyme led to similar successes and short FA became the primary product.

The two approaches have proven to be potent tools to introduce systems of chain length control in FAS. This rational engineering has the big advantage that it is mostly minimally invasive and due to the high conservation of *de novo* FA synthesis, individual mutations could easily be used in other FAS (and their organisms) as well. Even heterologous expression of modified FAS genes is feasible.

Engineering was not only tested in a defined *in vitro* environment and but also in *S. cerevisiae* as an exemplary *in vivo* system. The results eventually confirmed the *in vitro* findings and proved that the chosen engineering could be transferred to more complex systems. Even before any optimization for highest output, the titers of short FA from *S. cerevisiae* fermentation matched previous reports with 118 mg/L[30].

In sum, this work covers several layers from basic research to preliminary applications. The presented modifications to create short FA producing FAS can be a key step in synthesis pathways and will likely enable a whole range of new succeeding research. It can be seen as a valuable contribution towards establishing novel ways for the production of chemicals from renewable sources.

6 Scientific supplementary information

The scientific supplementary information was also written in the course of the preparation of manuscripts. For more information on the individual contributions on the writing process, please refer to the section “7 Statement of personal contributions“.

6.1 Supplementary Notes

Supplementary Note 1: Validation of assay results (statistical coverage)

-Quantitative product analysis by HPLC:

We purchased standards for each acyl-CoA ester (purity checked in lab), and calibrated with standards providing the basis for quantifications. By using purchased compounds iC_5-CoA and C_{17}-CoA, as representatives of short and long chain acyl-CoA esters, we optimized the recovery rates of reaction work-up. Data showed high repeatability (see Supplementary Table 2 for details). iC_5-CoA and C_{17}-CoA were further used as internal standards in all assays ensuring precise quantification of losses in reaction work-up. In a similar manner we judged the repeatability of the coupled assay.

-Repeatability of reactions probed in biological and technical replicates:

For module 1, wild type FAS, G2599S and the G2599S-M2600W construct were prepared from several independent expressions. For each protein preparation (biological replicates; see Supplementary Fig. 12). The error bars in the graphs represent the standard deviation of these at four biologically independent protein purifications.

On the basis of this statistical evaluation and owing to challenges in protein preparations (we probed 16 different module 1 constructs), we did not further collect biological replicates on module 1 constructs. However, for each construct, we prepared at least three reactions followed by independent work-up and analytics. In two cases ($FAS1^{G2599S\text{-}M2600W\text{-}Q1349V}$ and $FAS1^{G2599S\text{-}M2600W\text{-}Q1349L}$), constructs could only be probed in a single assay each, since protein purification was unexpectedly low. Figure legends include outlines to statistics of depicted data.

For scanning trends, we chose to rely on single measurements owing to challenges in protein preparation. Furthermore, we probed substrate starting concentrations. To cover a wide range, we decide to perform single measurement; however, one substrate combination was tested again in four technical replicates, confirming our earlier observation of low variation in technical replicates. Figure legends include outlines to statistics of depicted data.

Supplementary Note 2: Quantification of CoA esters by HPLC

Initially, the recovery of the extraction procedure was tested with the internal standards (three different concentrations: 20 µM, 50 µM, 100 µM with four replicates). The obtained recovery rates for iC_5-CoA were (74.5±0.9)% for 20 µM, (74.8±1.4)% for 50 µM; (73.1±1.7)% for 100 µM respectively. For C17-CoA the rates were (70.1±1.8)% for 20 µM, (75.0±2.0)% for 50 µM; (72.4±1.9)% for 100 µM respectively.

The calibration coefficients and calibration ranges are listed for each CoA species in Supplementary Table 4. For the shorter acyl-CoAs (C_6-CoA, C_8-CoA, C_{10}-CoA), iC_5-CoA served as the internal standard to calculate loss from the work up procedure, for the longer acyl-CoA (C_{14}-CoA, C_{16}-CoA, C_{18}-CoA) C_{17}-CoA was used.

Concentrations for the calibration were 5 µM, 10 µM, 15 µM, 20 µM, 25 µM, 50 µM, 100 µM, 150 µM and 200 µM with three replicates each (used concentrations for the calibration are indicated by the calibration range). The signal of all points included in the calibration range differed from their calculated value no more than 15% (20% for the lower limit of quantification).

Supplementary Note 3: Chemical synthesis of the internal standard 6-octyl-4-hydroxypyran-2-one

For the synthesis of 6-octyl-4-hydroxypyran-2-one, a slightly modified protocol from reference [159] was used. 500 mg (3.96 mmol, 1 eq.) 4-hydroxy-6-methyl-2H-pyran-2-one were dried by suspending it three times in 20 ml toluene and removing the solvent under reduced pressure. Anhydrous THF (20 ml) was added to the residue and the resulting suspension was cooled to -78 °C. Then, 5.70 ml (9.12 mmol, 2.3 eq., 1.6 M solution in hexane) n-butyllithium was added dropwise over 15 min. The suspension was stirred at -78 °C for 4 h before 1.30 ml (1.79 g, 7.93 mmol, 2 eq.) 1-

iodoheptane were added dropwise. The reaction was slowly warmed to room temperature over night while being stirred. A 3 M HCl solution was used to quench the reaction by slowly adjusting the to pH 2. After separation of the phases, the aqueous phase was extracted with Et_2O (3x 30 ml). The combined organic phases were washed with brine and dried over Na_2SO_4. After filtration and evaporation of the solvent under reduced pressure, the residue was redissolved in acetone and concentrated onto silica gel. Purification by flash chromatography (silica gel, gradient of petrol ether/ethyl acetate 1:0, 9:1, 5:1, 4:1, 3:2) gave the desired product as a white powder (16.0 mg, 0.07 mmol, 2% yield). The product was stored at -20 °C until utilization. The purity was confirmed by HPLC-UV at 298 nm. Chemicals were obtained from Sigma-Aldrich.

MS (ESI) (*m/z*): $[M+H]^+$ calcd. for $C_{13}H_{20}O_3$ 225.1491; found 225.1510; ^{1}H NMR (400 MHz, DMSO): δ 5.91 (s, 1H), 5.16 (s, 1H), 2.41 (t, *J* = 7.5 Hz, 2H), 1.60-1.47 (m, 2H), 1.32-1.19 (m, 10H), 0.86 (t, *J* = 7.0 Hz, 3H)

Supplementary Note 4: Quantification of 6-HHP 1 by HPLC

For 6-HHP analysis in a sequential reaction, the recovery of the extraction procedure was tested on three different concentrations (10 µM, 50 µM, 100 µM) with four replicates for each concentration. The obtained rates were (74.7±5.1)% for 10 µM, (76.7±6.1)% for 50 µM; (73.1±2.8)% for 100 µM respectively.

Quantification was done in reference to the internal standard (6-octyl-4-hydroxypyran-2-one) using a response factor of 0.9825±0.0293 in the range of 10–400 µM. For the IS, the calibration equation was described by y = 7.467 x – 9.078 with an R^2 of 0.9996. Calibration points were 10 µM, 20 µM, 81 µM, 100 µM, 200 µM and 400 µM (with three replicates each), all meeting the criteria that all points included in the calibration range differed from their calculated value no more than 15% (20% for the lower limit of quantification)

Supplementary Note 5: Molecular dynamics simulations and kinetic model

The calculations of the molecular dynamics simulations and the kinetic model were conducted (and summarized) by our cooperation partner, the group of Prof. Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen. The descriptions are included in this thesis for completeness and are taken from the publication[17].

Calibration of model parameters

Our computational model of the reaction network of *S. cerevisiae* FAS represents 51 distinct reactions, enumerated in (supplementary Table 3). The core catalytic cycle was simplified by including only the canonical condensation and acetyl-, malonyl- and acyltransferase reactions, as catalyzed by the enzymatic activities of the KS, MPT and AT domains, and mediated through the ACP shuttle domain. The three catalytic activities responsible for the conversion of the post-condensation β-keto intermediate to a fully saturated acyl chain were not explicitly represented, under the assumption that these reactions proceed considerably faster than the overall time scale of the reaction. This assumption is justified by the observation that no unsaturated products are detected under normal conditions.

The 51 included reactions, each with a forward and reverse rate constant, yield a total of 102 model parameters. In the absence of an exhaustive experimental characterization of each of the individual reactions that comprise the network, we proceeded to narrow down the open parameter space in accordance with known properties of yeast FAS, as follows.

As a first step, we defined overall restrictions on the range of reaction rates to represent. Yeast FAS is known to process malonyl-CoA at a rate of about 2000 mU/mg FAS[152-153], which corresponds to 10-18 units consumed per second per synthetic unit, where a synthetic unit is defined as a grouping of one ACP shuttle domain and the single copy of each catalytic domain which is physically accessible to each ACP. Choosing a reference speed of 10/s for convenience, we assume that all relevant phenomena can be described by rates no more than three orders of magnitude slower or faster, yielding lower and upper bounds of $10^{-2}\ s^{-1}$ and $10^{4}\ s^{-1}$, thus representing a search range spanning six orders of magnitude.

A 'main sequence' of reactions, defined as essential steps, which must proceed forward at a certain rate to match the observed overall synthesis rate, was defined,

consisting of the following reactions, which were constrained to proceed at rates of no less than 10/s:

malonyl-CoA binding to the active site of MPT (given near-saturating concentrations of malonyl-CoA);
CoA release from MPT;
malonyl transfer from MPT to ACP;
transfer of acyl intermediates of 4 to 14-carbon length from ACP to KS;
condensation of ACP-bound malonyl with KS-bound acyl intermediate;
transfer of favored products C_{16} and C_{18} from ACP to MPT;
transfer of favored products C_{16} and C_{18} from MPT to CoA (product release).

Analogously, the reactions involved in acetyl-CoA priming (acetyl-CoA binding to AT, CoA release from AT, and acetyl transfer to ACP), which under saturating conditions occurs once for every 7 or 8 malonyl residues incorporated, were constrained to rates no slower than 1/s.
Acyl products in the range C_6-C_{14} can be generated in vitro in small but detectable quantities even under standard conditions[83]; a minimum rate for reactions required for these exit events was set 100-fold lower at 0.1/s.
The values of the Michaelis constant for substrates malonyl-CoA (8 μM) and acetyl-CoA (28 μM) have been reported[1]. Taken as a proxy for the binding affinity of these substrates for their respective canonical active sites, these values determine the ratio of the forward (k_1) and reverse (k_2) rates for the respective binding and unbinding reactions by the relationship $K_m=k_2/k_1$. For these reactions, the forward rate was picked randomly according to the general restrictions described above, and the reverse rate was calculated according to the experimentally determined ratio.
To sample relevant parameter sets from the large parameter space that remains after applying the constraints described above, a multi-level filtering process was implemented, to exclude the vast majority of candidate parameter sets, which were not descriptive of measured properties of yeast FAS described by Sumper et al.[83] Five sets of conditions were applied as filters, each corresponding to a scenario described by Sumper et al:

1) Yeast fatty acid synthase with iodoacetamide-induced condensation deficiency. Iodoacetamide-treated yeast FAS was shown to be deficient in the condensation reaction, with the gain of an ACP- and KS-dependent malonyl-CoA decarboxylase activity[83, 160], at a rate of 300 mU/mg under saturating conditions, corresponding to 2 s^{-1} units of malonyl-CoA per synthetic unit. Candidate parameter sets were filtered to exclude decarboxylation rates outside the range of 0.4-4 s^{-1}. This filter has the advantage of invoking only the small number of reactions involved in acetyl and malonyl import, export and transfer, bypassing the much larger number of chain length-dependent reactions which depend on the condensation reaction.
2) Standard reaction. Initial filtering was for an output of >80% C_{14} and longer acyl products.
3) High-acetyl-CoA condition, with 9:1 ratio of acetyl-CoA to malonyl-CoA.
4) Acetyl-CoA input throttled: acetyl-CoA is initially not present, and is generated at a known rate by a separate malonyl-CoA decarboxylase enzyme.
5) Malonyl-CoA input throttled: malonyl-CoA is initially not present, and is generated at a known rate by a separate acetyl-CoA carboxylase enzyme.

Having established (i) initial constraints for the parameter space based on known properties of the system, and (ii) a filtering procedure designed to identify parameter sets producing fatty acid product spectra matching those observed experimentally, we proceeded, using high-performance computing resources, to generate and evaluate large numbers of candidate parameter sets. For each candidate parameter set, the procedure is as follows:
populate the array of forward and reverse rate constants with randomly chosen values, chosen within the constraints defined above;
launch successive stochastic simulations of the FAS reaction network under the five filtering conditions given above;
reject and discard any parameter set when the output product spectrum fails to match any of the five filtering conditions.

In this step, we tested hundreds of millions of candidate parameter sets, and retained 31735 that matched, within a given tolerance, product output spectra for all five of the filter conditions outlined above.

The five-level filtering round was followed by a Monte Carlo optimization procedure. Candidate parameter sets were assigned a score *S* consisting of the root mean square deviation of the fractional representation of acyl chain lengths observed in the model (p_m) compared to those measured experimentally (p_e), with additional penalties for deviations for the rates of the iodoacetamide-blocked and 'standard' reactions 1) and 2) (r_1, r_2) from the reference values of 2 s^{-1} and 10 s^{-1}:

$$S = \sqrt{\frac{1}{9}\sum_{i=6,8,10,\ldots}^{22}(p_m - p_e)^2} + \frac{|r_1 - 2\mathrm{s}^{-1}| + |r_2 - 10\mathrm{s}^{-1}|}{100}$$

Parameters of the input models were then varied randomly, with any changes leading to a reduction in score being accepted.

We subjected each of the 31735 candidate parameter sets from the filtering stage to the described Monte Carlo optimization procedure. The best-scoring 1000 parameter sets were then chosen to compose the overall ensemble model for yeast FAS. The scores *S* for these 1000 sets ranged from 0.025 to 0.052.

Cross-validation of kinetic model against FAS mutations

Physical basis for representation of mutations in kinetic model

As described in the main text, we represented the effects of active site mutations in the MPT, KS and AT domains in the kinetic model, using free energy quantities calculated from atomistic simulation. These were calculated using truncated representations of the area surrounding each mutation, as described under Methods, necessarily excluding whole-protein scale effects in this megadalton complex, but accessing the relevant molecular recognition steps in atomic detail.

The relationship between the free energy change ΔG of a reaction, the equilibrium rate constant K, and the forward and reverse rate constants k_1 and k_2, is given by

$$\Delta G = -RT \ln K = -RT \ln \frac{k_1}{k_2}$$

From this relationship, a calculated change in the free energy difference of the reaction (ΔΔG) dictates a change in the ratio of the forward and reverse rate

constants k_1 and k_2. We accordingly sampled the range of possible thermodynamic outcomes consistent with the calculated change for each mutation by running tens of thousands of simulations of the kinetic network, each based on a single parameter set selected from the 1000-member kinetic model ensemble, with the respective rate constants affected by the mutation modified in accordance with the calculated ratio. To interpret the result, the product spectra of these tens of thousands of simulations of kinetic scenarios consistent with the calculated free energy change were aggregated and averaged, with population standard deviations (shown in the respective figures as error bars) giving an indication of the spread of likely outcomes of the mutation.

Cross-validation against mutations

Following model training against known experimental conditions for *S. cerevisiae*, we applied the framework described above to represent FAS mutations tested in our *C. ammoniagenes* experimental system. In so doing, we pursued the dual goals of cross-validating the model against untrained data, and of supporting our interpretation of the experimental results.

As presented in the main text, while was not initially possible to reconcile modeling results for the double G1250S/M1251W mutation to the KS domain with experimental output and the “gatekeeper” hypothesis that guided the design of the mutations, an updated conception of the “gatekeeper” rationale allowed us to suggest a mutually consistent interpretation of the model and results. Moreover, the transferase mutations were tested experimentally only in combination with the double KS mutation, thus complicating our interpretation of the transferase mutation results.

For this reason, we here divide the cross-validation against experimental data into two parts. In the first, we consider the effects of the tested transferase mutations in *S. cerevisiae* in isolation, without changes applied to the KS domain, and compare to the respective experimental effects in our *C. ammoniagenes* system for each of these mutations, with the caveat that the experimental results were obtained in combination with the double KS mutation. In the second part, we describe our

updated conception of the KS mutation, and again consider each transferase mutation in this updated context, yielding a more direct comparison with the experimental data, albeit subject to greater uncertainty.

Cross-validation without KS modification

For the AT domain mutation I151A (*C. ammoniagenes*) / I306A (*S. cerevisiae*), two hypotheses were considered as an explanation for increased C_8-CoA output (see Supplementary Fig. 12): increased acetyl throughput on the one hand, and the creation of a novel exit pathway for short-length acyl chains on the other. For *S. cerevisiae* mutation I306A, in the AT domain, we calculated a ΔΔG value of -0.4 ± 1.6 kJ/mol for the effect of the mutation on binding of acetyl-CoA, indicating little to no effect on binding affinity. This speaks against the first hypothesis, that C_8-CoA output is enhanced by virtue of a favored loading of acetyl-CoA.

We probed the second hypothesis, which proposes a novel short-acyl binding channel, with atomistic simulations of C_8-CoA in complex with *S. cerevisiae* mutant I306A. C_8-CoA could not be plausibly fitted in alignment with the catalytic residues in the wild type active site, but was accommodated in the I306A mutant with minor local rearrangement. The resulting structure was used as a starting point for ten unbiased molecular dynamics simulations, each exceeding 400 ns in length. The simulations maintained a plausible pre-catalytic conformation, in alignment with the catalytic serine side chain and the oxyanion hole, in nine cases, suggesting a complex that is stable at least in the microsecond range.

Additionally, we note that the first hypothesis, which suggested that an increase in acetyl availability as a result of the mutation would yield increased C_8-CoA production by virtue a general shift towards shorter products, as observed under conditions of high relative acetyl-CoA availability, would predict a reduction in the proportion of longer (C_{16}- and C_{18}-CoA esters) released. Conversely, were the relative availability of the AT active site to be reduced through competition with a non-canonical substrate in the form of short acyl chains, a relative reduction in acetyl availability could be expected, which could shift the balance towards longer products.

Taken together, these results favor the hypothesis that I306A-mutated *S. cerevisiae* AT is a competent acyltransferase in this acyl chain length range. Introducing the possibility of C_8 exit through the AT domain into the kinetic model with a forward rate of $3s^{-1}$ for both the transfer of C_8 from ACP to AT domain and its exit as CoA ester, led to increased release of C_8-CoA product at the expense of marginal reductions in C_{16}-CoA and C_{18}-CoA output. The rate of $3s^{-1}$ was chosen because, while we have no direct experimental or theoretical thermodynamic reference point, the hypothesis as formulated requires *a priori* a reaction rate on the same order as the overall synthesis reaction (ca. 10 s^{-1}).

C. ammoniagenes MPT mutation R1408K and AT mutation I151A, each of which individually brought significant increases in the relative yield of C_8-CoA, were further tested in combination (in addition to the KS domain mutations, yielding quadruply-mutated FAS1$^{G2599S\text{-}M2600W\text{-}R1408K\text{-}I151A}$). The combination of these two beneficial mutations might be expected to be superior to each in isolation; however, their effects on relative C_8-CoA yield were not additive. Rather, the *C. ammoniagenes* result showed the KS/MPT combination remained the highest-yielding for specific production of C_8-CoA (see Fig. 7).

Incorporating the MPT and AT mutations simultaneously, the kinetic model reproduced the finding of non-additivity, with C_8-CoA output for the double MPT/AT mutant no higher than that for the MPT single mutant (see Supplementary Fig. 12).

In summary, the representation of the tested transferase mutations in the kinetic model yielded correct reproduction of (i) the broad shift towards shorter products observed as a result of a malonyl throughput-restricting MPT mutation, (ii) the increased C_8-CoA yield as a result of the AT mutation, (iii) the non-additive nature of the MPT and AT mutations, and (iv) the slight shift towards longer products that accompanies the AT mutation.

KS-mediated chain length control revisited

In the light of the convincing cross-validation results from transferase engineering, we revisited the unexplained result we obtained in our attempt to determine the

molecular basis of KS-mediated chain length control. We identified two hypotheses, both potentially reconciling the apparent contradiction.

Underlying the first hypothesis, we note that a prominent side reaction of FAS/PKS protein family is the unproductive KS-mediated decarboxylation, without subsequent condensation. of ACP-bound malonyl. For FAS, this side reaction has previously been reported as occurring at a negligible basal rate under normal conditions, but at a considerably higher rate when key catalytic residues are blocked through covalent modification[130]. We therefore hypothesized that an unusually strongly KS-bound C_8 intermediate could approximate a stalled state and induce decarboxylation, conceivably leading to elevated levels of C_8-CoA product. To test this alternate hypothesis *in silico*, we asked whether a kinetic model extended by the reaction KS:C_8 + ACP:malonyl → KS:C_8 + ACP:acetyl + CO_2 might account for the increase in C_8-CoA release. Although we were able to identify members of our model ensemble for which the introduction of the decarboxylation reaction yielded an increase in C_8-CoA output, the minimum reaction rate for the decarboxylation reaction was at least an order of magnitude faster than the rate reported[130], thus appearing to contradict this interpretation. We concluded our exploration of this hypothesis with further experimental measurements. *In vitro* reaction with ^{13}C-labelled acetyl-CoA (but unlabeled malonyl-CoA) did not show priming of FA synthesis with unlabeled acetyl, as would be expected in the case that significant quantities of unlabeled malonyl-CoA were undergoing decarboxylation in the KS-engineered FAS (Supplementary Fig. 117).

Invoking a second hypothesis, we tested whether a less restrictive description of the influence of the double KS mutation on C_8 throughput might reconcile our findings. In our initial representation of the calculated enhanced binding affinity of the C_8:KS complex, we assumed that the calculated free energy change, considerably lower for the mutated C_8:KS state, would not yield a reduction in the forward rate of the reaction C_8:ACP + KS → ACP + C_8:KS. We relaxed this restriction to allow reductions in this reaction rate up to a factor of ten (while still fixing the ratio of forward and reverse rate constants in accordance with the calculated free energy difference). In doing so, we established that for 100 members of the 1000-member model ensemble, a slowdown in the reaction ACP:C_8+KS → ACP+KS:C_8 by less than a factor of ten yielded an increase of at least 5% in C_8-CoA output. We note

that in these model scenarios, the reduced forward rates for this reaction are compensated by greater slowdowns in the reverse rate of the same reaction, in order to enforce agreement with the calculated negative binding free energy change for the KS:C_8 complex.

We note that the agreement between experiment and model for the G1250S/M1251W combination arises by construction, since we selected only those parameter sets that matched the known C_8-enriched *C. ammoniagenes* product spectrum. Conversely, we may evaluate the transferase mutations in combination with the updated KS-mutated model, since our representations of the transferase mutations are based only on calculated thermodynamic quantities.

Combining MPT mutation R1834K with the G1250S/M1251W double KS mutation, we observe increased C_8-CoA output, to 31±13%, compared to 46.6% observed experimentally. As with the MPT mutation in isolation, the model reproduced the general shift towards shorter products, including a considerable reduction in C_{18}-CoA release.

Similarly, introducing the I151A mutation in addition to the double KS mutation yielded a prediction of modestly increased C_8-CoA release again in good agreement with the experimental finding.

Applying all four mutations in combination, the notable experimental finding that the AT and MPT mutations, while each separately advantageous, do not in combination yield greater C_8-CoA production, is again reproduced in the model.

Summarizing our combined theoretical and experimental exploration, we conclude that the studied KS double mutation does not represent a “gate” in the KS substrate tunnel, in the narrow sense of a steric barrier counteracting binding of long acyl chains as we initially conceived. Rather, our findings led us to interpret that the double mutation introduces a kinetic barrier that steers C_8-CoA away from KS-mediated elongation and thus towards release. A reevaluation of the cross-validation results for the transferase mutations did not contradict this implementation. The fact that the findings from this second round of cross-validation remain consistent with the experimental findings serves as a further confirmation of the usefulness of the model in representing the FAS catalytic cycle.

Atomistic modeling of MPT substrate

Unlike the KS domain active site, modeling of substrate interactions in the malonyl/palmitoyl transferase active site did not benefit from the availability of a co-crystal structure with an acyl chain analogue. Through knowledge of the catalytic mechanism of the acyltransferase activity and the crystal structure[31, 146], a putative pre-catalytic complex could in principle be constructed analogously to the approach described above for the active site of the KS domain, with distance restraints between the substrate carbonyl and the backbone amide hydrogens of Q1669 and L1809 representing the oxyanion hole, and from the side chain hydroxyl hydrogen of S1808 to the thioester carbon of the acyl intermediate. However, in the case of MPT, there is no clear pre-formed hydrophobic channel for longer acyl chains.

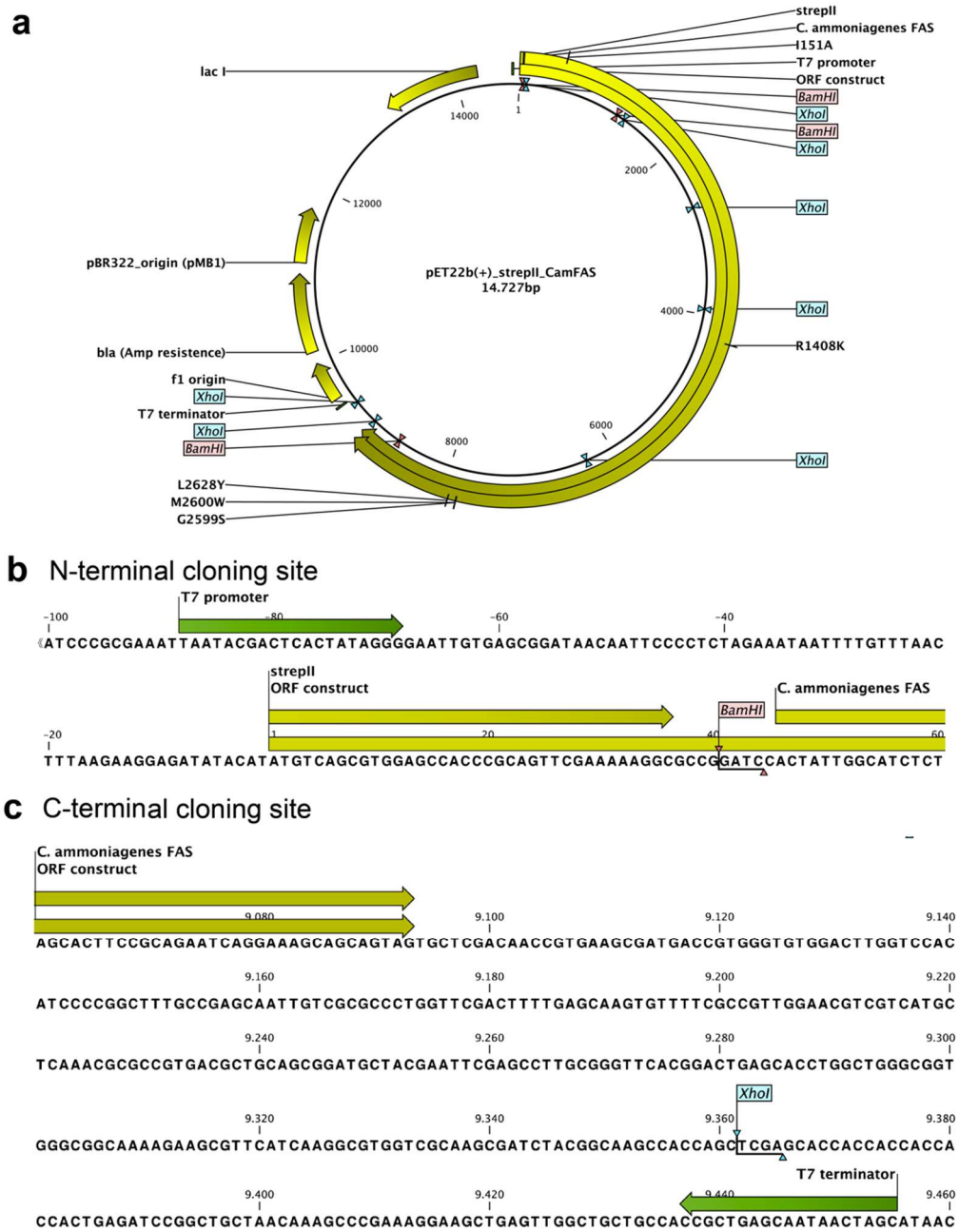
Evidence suggests that longer acyl chains will interact with patch of exposed hydrophobic residues[13]. However, initial modeling based on the crystal structure showed that a 16-18-carbon acyl chain is too short to bridge over a ridge formed by M1838, T1774 and neighboring residues, indicating that non-trivial rearrangement with respect to the crystal structure are required to accommodate the presumed long-chain substrate binding mode in MPT.

Using the palmitoylate (C_{16}) derivative as a probe substrate, molecular dynamics simulations were performed which made use of manual conformational selection to guide the exploration of the energy landscape in the direction of complexes fulfilling the experimentally determined structure-function hypothesis, without direct imposition of explicit biasing forces. Over multiple successive iterations, sets of 5-10 independent molecular dynamics simulations were launched, unbiased with the exception of the distance restraints maintaining a pre-catalytic conformation as described above, and allowed to run for up to 100 ns. The resulting conformations were then assessed individually, and a small subset of those qualitatively judged to most closely resemble a complex matching the experimental structure-function hypothesis were selected as input for a further iteration. After three such iterations, the molecular dynamics simulations arrived at a complex where both the putative pre-catalytic distance restraints and the interaction of the acyl chain with the hydrophobic patch were fulfilled.

The resulting structure notably differs from the crystal structure through the partial unwinding of the first turn of an alpha helix formed by residues 1775-1779, opening the ridge described above and allowing access to the hydrophobic patch. Conservation in the associated motif TQFTQP is high: 100% among the [set of sequences from database trimmed for redundancy] for F1776, and 81% for P1779, whose lack of a backbone amide for hydrogen bonding can be expected to lead to lower stability for this helix turn.

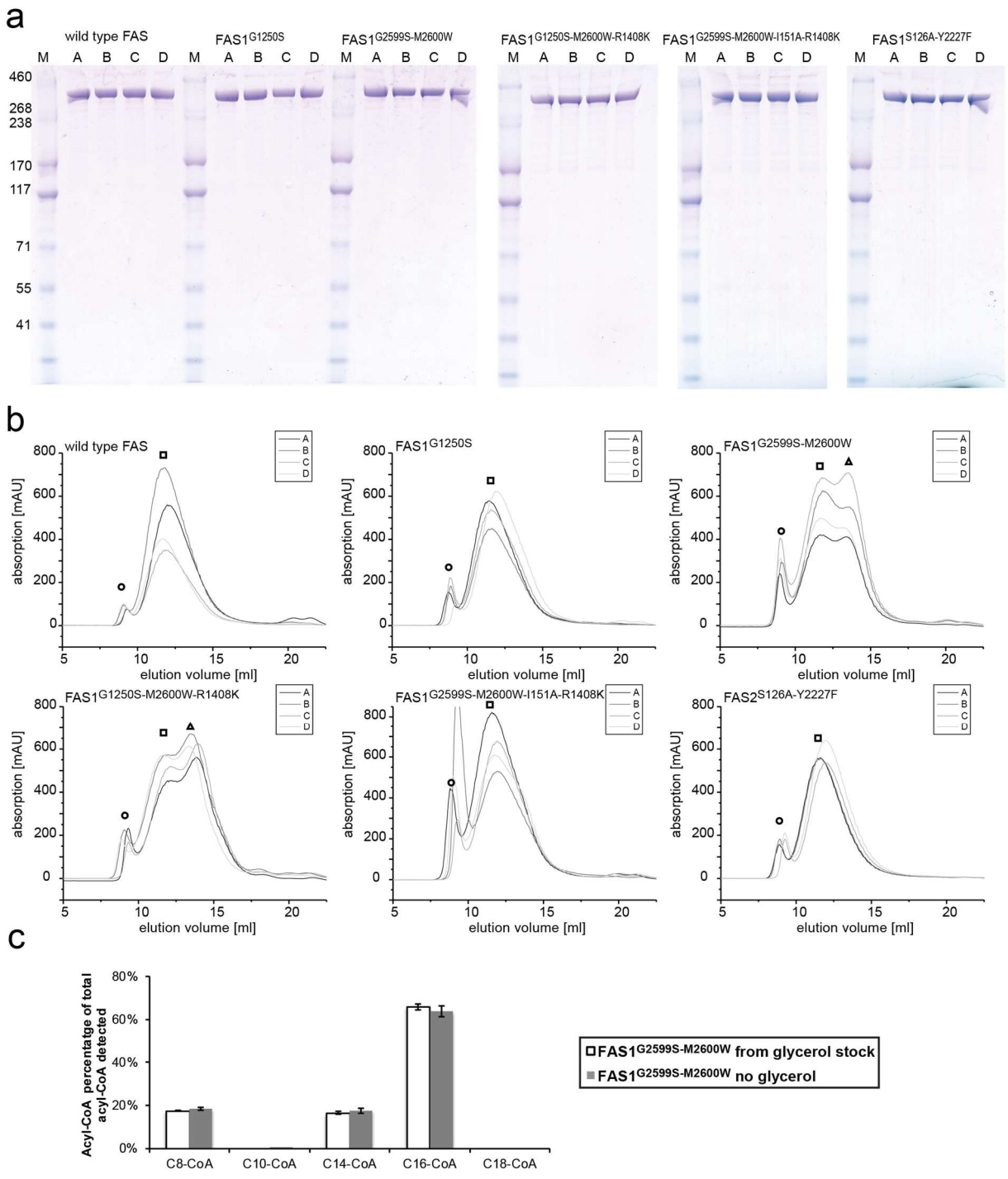
As this structure was derived using the C_{16} derivative as a probe, and since the associated structural perturbations are non-trivial and can be expected to vary as a function of substrate acyl chain length, we do not present calculated free energy differences for successive substrate methylations based on this model, as we judge that the atomistic simulation techniques applied in this work, short of intractably exhaustive sampling, would fail to capture the details of the energy landscape representing the non-trivial ‘open/closed’ transition.

6.2 Supplementary Figures

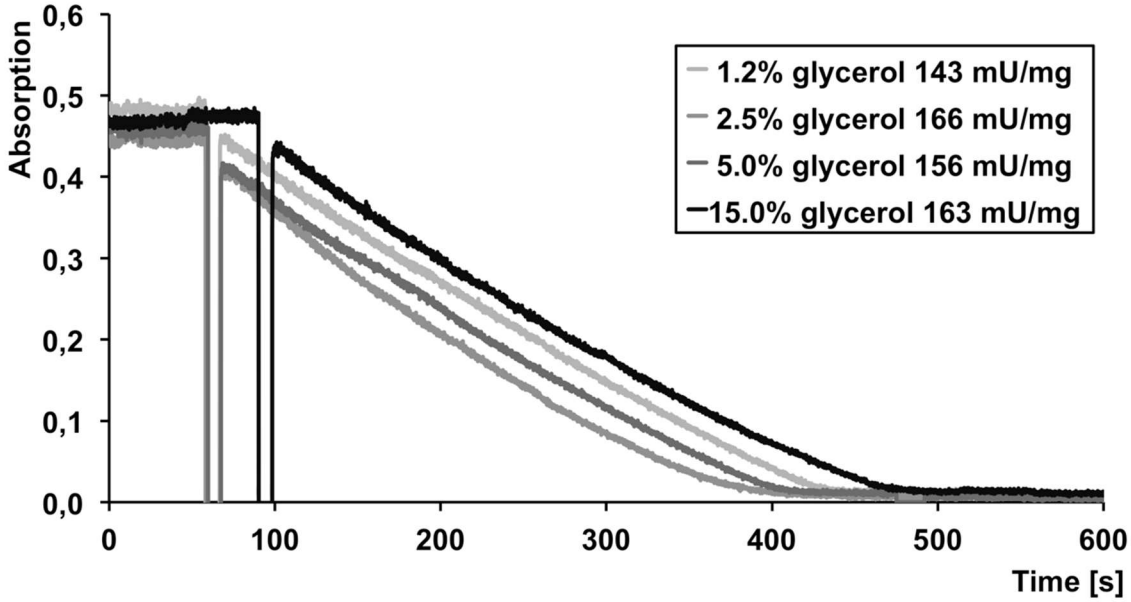


Supplementary Figure 1: *C. ammoniagenes* FAS expression vector used in the *in vitro* study. **(a)** Schematic view of the pET-22b(+)vector used for the expression of the *C. ammoniagenes* FAS gene in *E. coli*. The mutation sites of substrate affinity modulations are indicated. The mutation sites for the MPT and KS surface mutations were not included for clarity reasons. **(b)** The N-terminal part of the multiple cloning site is shown as sequence with strep II tag and the start of the insert of *C. ammoniagenes* FAS. **(c)** In the C-terminal part shown here, the actual stop codon of *C. ammoniagenes* FAS gene is located approx. 260 bp upstream to the XhoI restriction site that was used for cloning; the non-coding sequence is an artifact from a wrongly annotated gene in literature. Depiction in all panels was generated using CLC Main Workbench, Version 6.9.1, Waltham, USA. For origin of vectors, if they were not produced in the own work but in previous works[14, 121], consult Supplementary Table 1.

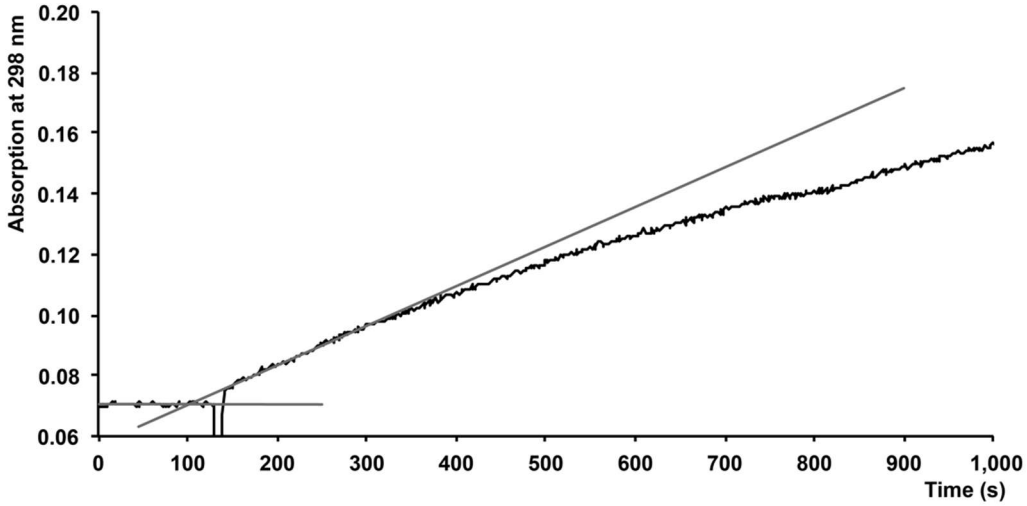
Supplementary Figure 2: AcpS expression vector (for study on *C. ammoniagenes* FAS).
(a) Schematic view of the pETcocol vector used for the expression of the AcpS gene (the phosphopantetheine transferase for activation of *C. ammoniagenes* FAS) in *E. coli*. **(b)** The sequence of the insert is shown here with promotor and terminator. Depiction in all panels was generated using CLC Main Workbench, Version 6.9.1, Waltham, USA.
For origin of vectors, if they were not produced in the own work but in previous works[14, 121], consult Supplementary Table 1.



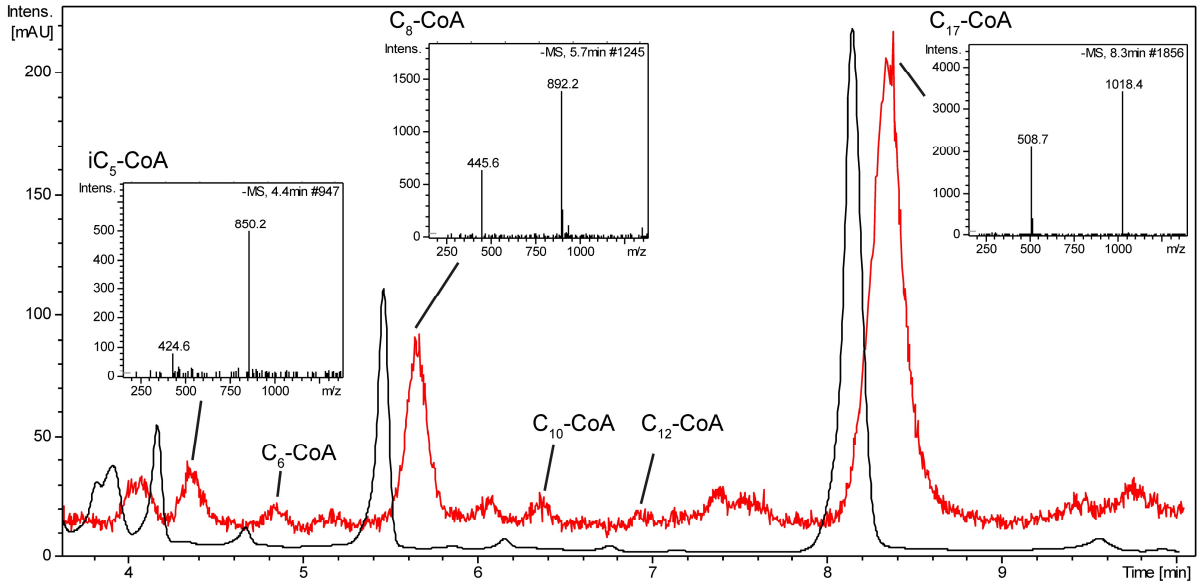
Supplementary Figure 3: Characterization of protein purification steps.
(a) SDS PAGE of purified FAS variants. Four biological repeats are shown for each construct. Proteins were purified by Melanie Janßen (also in b); the figure is taken from the joint publication[17]. **(b)** The size exclusion chromatography graphs show the purification and size distributions of the four biological repeats from a). Only fraction marked by the squares were used in assays (hexamers). The other fractions, marked by circle and triangle were assumed to be higher aggregates or lower oligomeric forms, respectively. **(c)** The Glycerol influence on product spectrum of FAS product spectrum of FAS$^{G2599\text{-}M2600W}$ as freshly prepared protein and as received from a glycerol stock were compared. For the sample with glycerol, its final glycerol content in the assay was 1.2%.



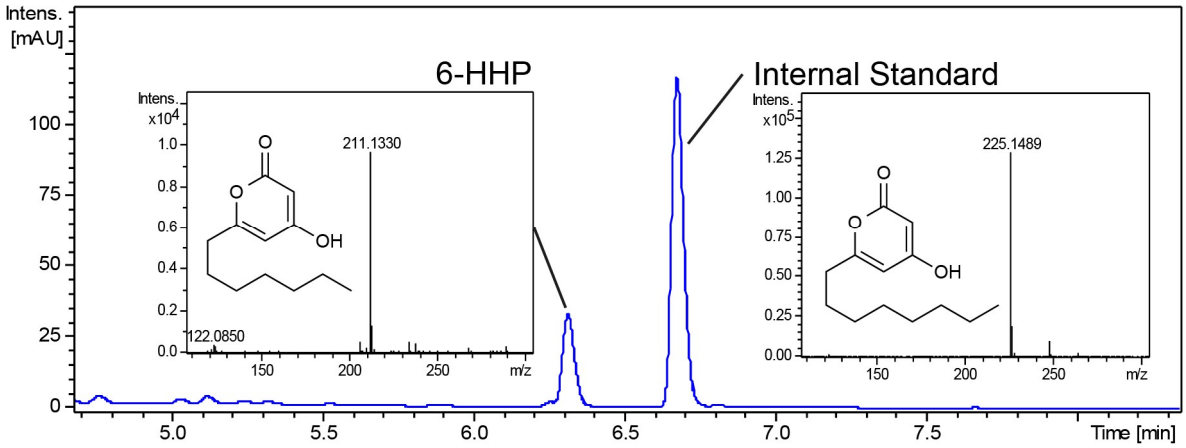
Supplementary Figure 4: Glycerol influence on the activity FAS in varying concentrations.
$FAS^{G2599\text{-}M2600W}$ was tested for activity with varying amounts of glycerol. NADPH decrease was monitored at 334 nm over time. After a period of constant absorption, the reaction is started by the addition of malonyl-CoA. Linear trends for both slopes, before and after the reaction, are used for the determination of enzymatic activity. A value of 6.0 mM^{-1} cm^{-1} was used as extinction coefficient $\varepsilon_{334\ nm}$[161]. The activities are listed in the legend inset.



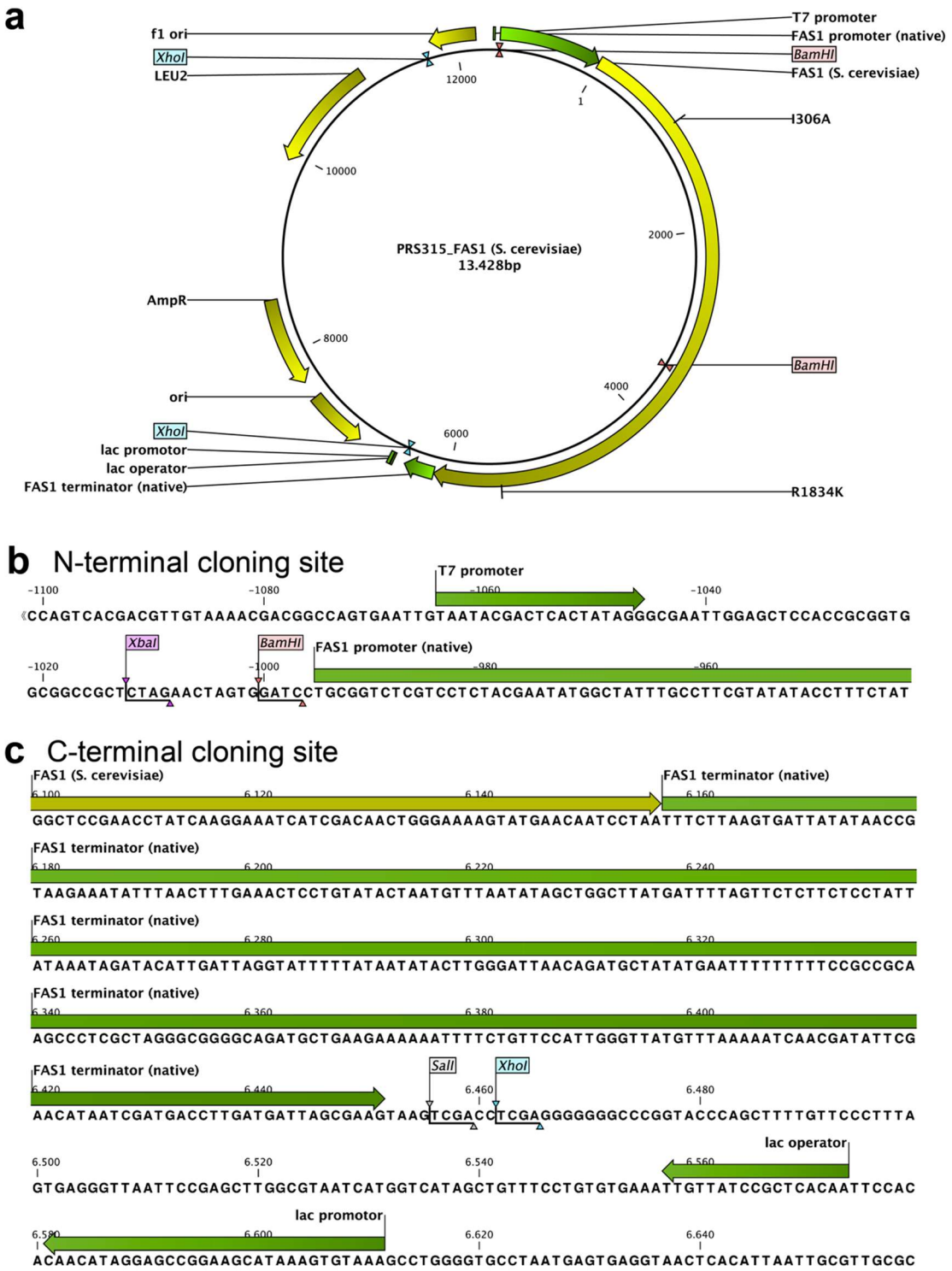
Supplementary Figure 5: Exemplary result of activity assay of module 2.
In this assay, the formation of the product 6-HHP was recorded at 298 nm over time. After a period of constant absorption, the reaction is started by the addition of malonyl-CoA. Linear trends for both slopes, before and after the reaction, are used for the determination of enzymatic activity[121].



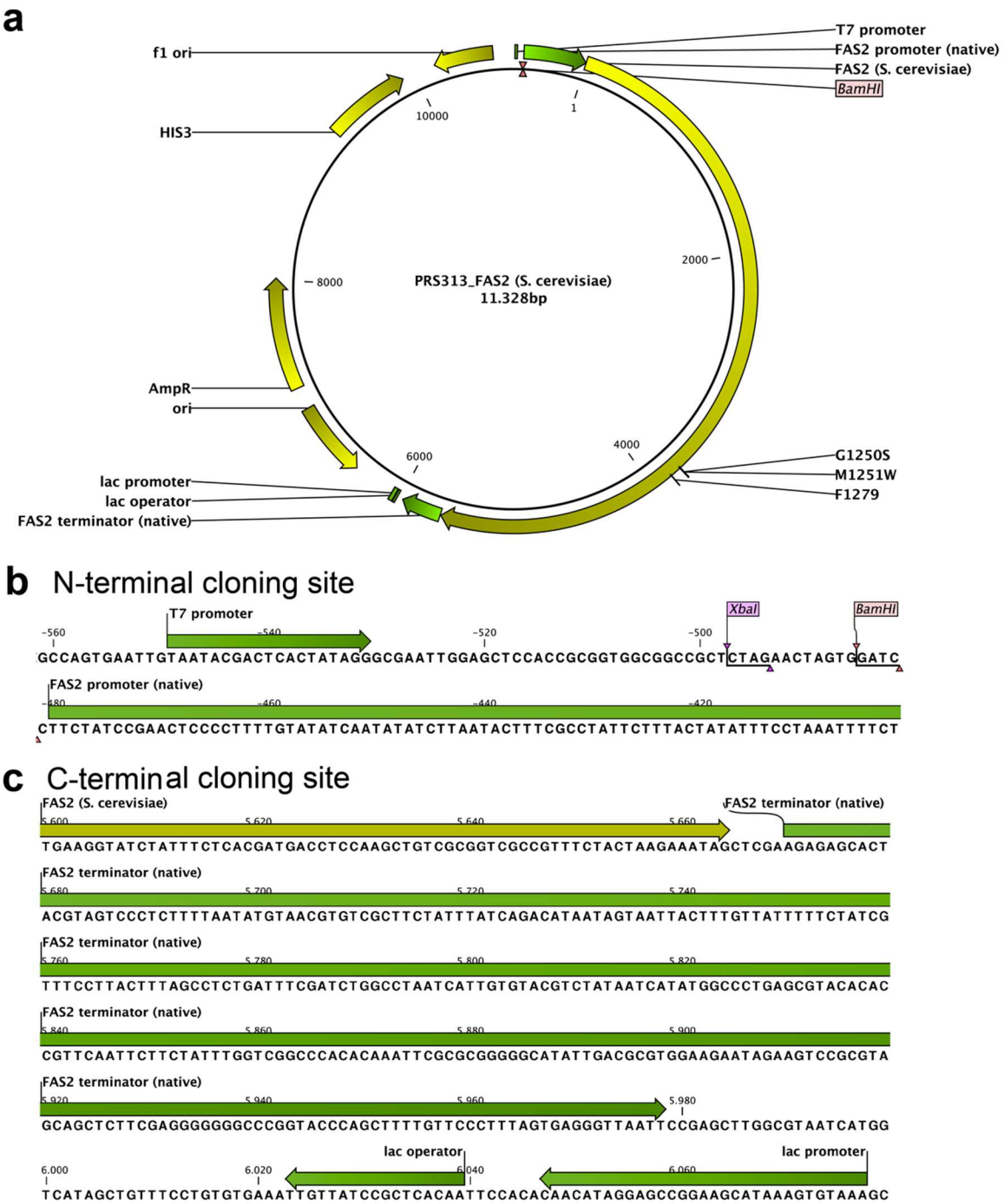
Supplementary Figure 6: Exemplary HPLC chromatogram for CoA ester analysis.
The HPLC chromatogram of one measurement of the FAS^{L2628Y} is displayed here with chromatogram traces for the UV absorption at 260 nm (black line) and the total ion count trace (red line). There is an offset of about 12 s between the two traces in this set up, which reflects the time for analytes to flow from the UV detector to the MS. Characteristic MS data is included as an example for the internal standards iC$_5$-CoA and C$_{17}$-CoA and the predominant product C$_8$-CoA. The UV trace was used for quantification (calibration parameters are listed in Supplementary Table 1).



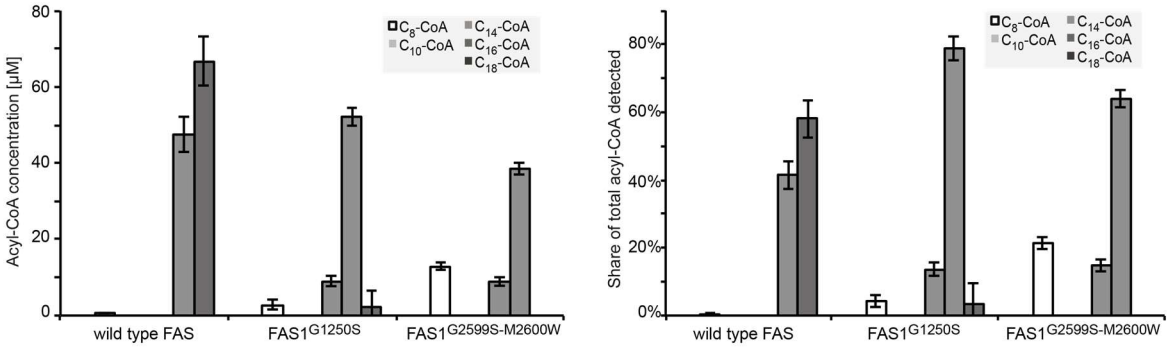
Supplementary Figure 7: Exemplary raw data of coupled assay.
Shown is the HPLC chromatogram at 298 nm from a coupled assay of FAS-M1$^{G2599S\text{-}M2600W\text{-}I151A\text{-}R1408K}$ → FAS-M2$^{S126A\text{-}Y2227F}$. For the product, 6-HHP, and the internal standard, the structures and mass spectra are shown.



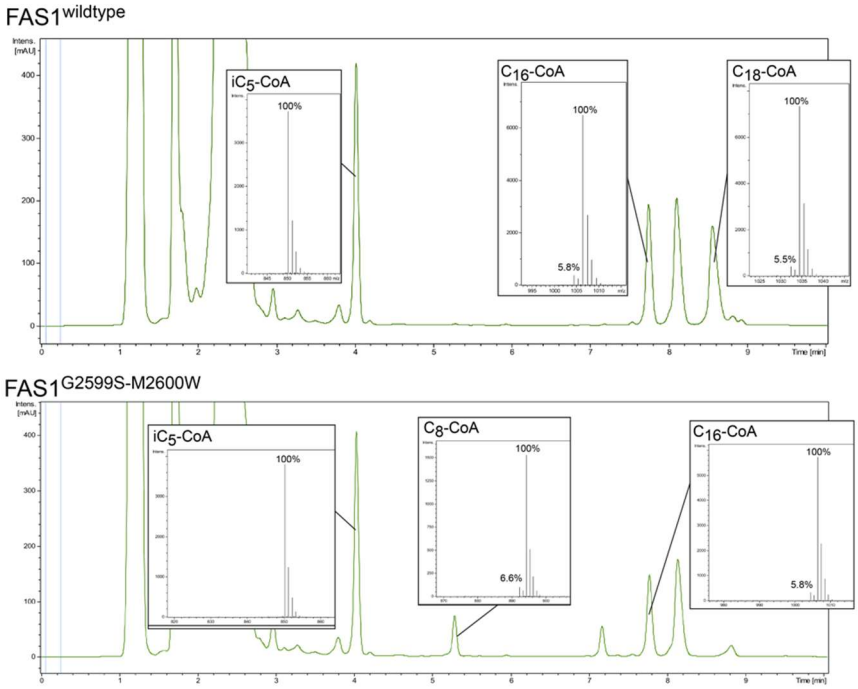
Supplementary Figure 8: *S. cerevisiae FAS1* expression vector used in the *in vivo* study.
(a) Schematic view of the PRS315 vector used for the expression of the *S. cerevisiae* FAS gene in its native host *S. cerevisiae*. The mutation sites are indicated. **(b)** The N-terminal part of the multiple cloning site is shown as sequence with the start of the native *FAS1* promoter. Numbering is in reference to the start codon of the *FAS2* open reading frame. **(c)** For the C-terminal, the transition of the S. cerevisiae FAS gene towards its natural terminator and into the vector is shown. Depiction in all panels was generated using CLC Main Workbench, Version 6.9.1, Waltham, USA. For origin of vectors, if they were not produced in the own work, consult Supplementary Table 1.



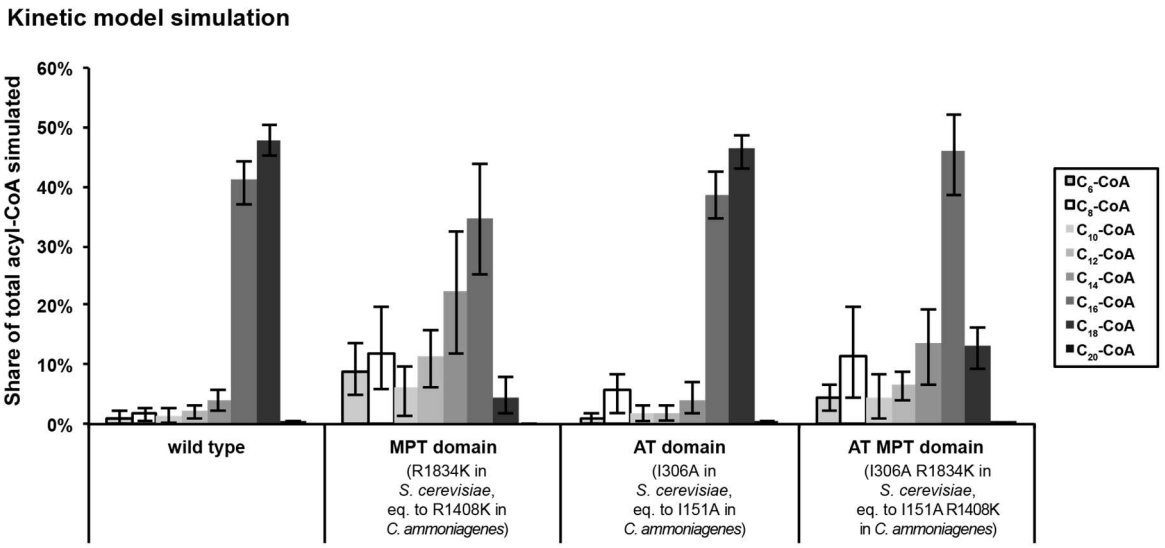
Supplementary Figure 9: *S. cerevisiae FAS2* expression vector used in the *in vivo* study.
(a) Schematic view of the PRS313 vector used for the expression of the *S. cerevisiae* FAS gene in its native host *S. cerevisiae*. The mutation sites are indicated. **(b)** The N-terminal part of the multiple cloning site is shown as sequence with the start of the native *FAS2* promoter. Numbering is in reference to the start codon of the *FAS2* open reading frame. **(c)** For the C-terminal, the transition of the S. cerevisiae FAS gene towards its natural terminator and into the vector is shown. Depiction in all panels was generated using CLC Main Workbench, Version 6.9.1, Waltham, USA. For origin of vectors, if they were not produced in the own work, consult Supplementary Table 1.



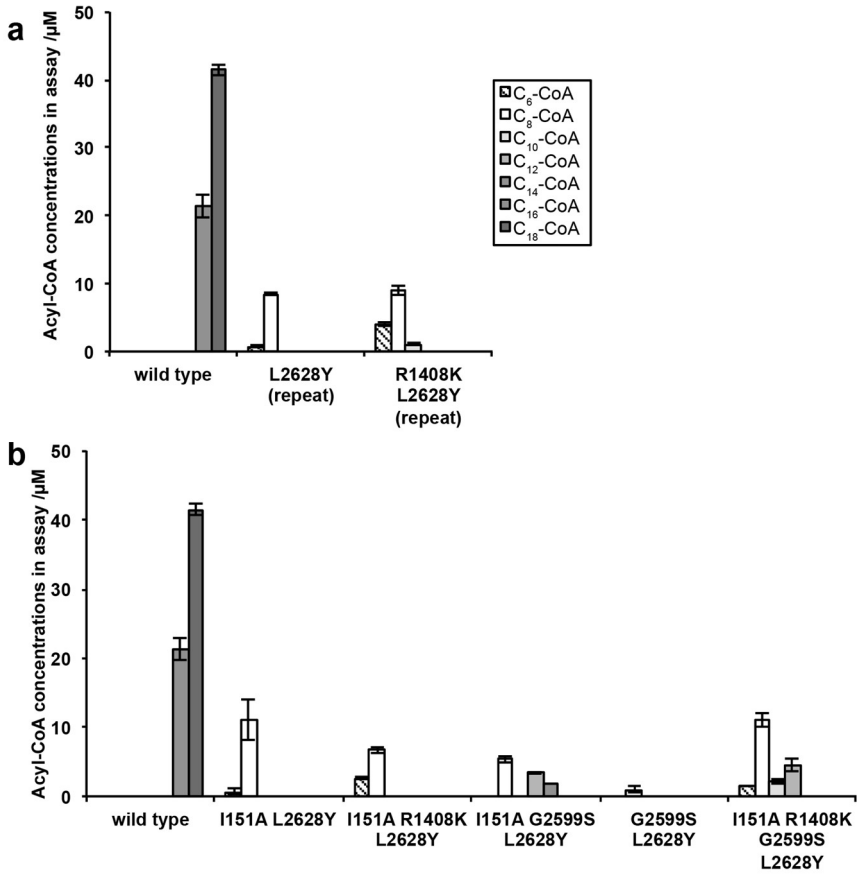
Supplementary Figure 10: Product distributions from additional FAS1 enzyme purifications. Biological repeats were prepared. Acyl-CoA concentrations and percentages of molar concentrations of specific acyl-CoA in total of acyl-CoAs produced by engineered FAS from four expressions for the wild type, FAS1^{G2599S} and FAS1$^{G2599S-M2600W}$. Compared to the concentrations reported in Figure 7, minor differences in the master mix of substrates most likely caused a slight shift to shorter FA. Proteins were purified by student Melanie Janßen; the figure was taken from the joint publication[17].



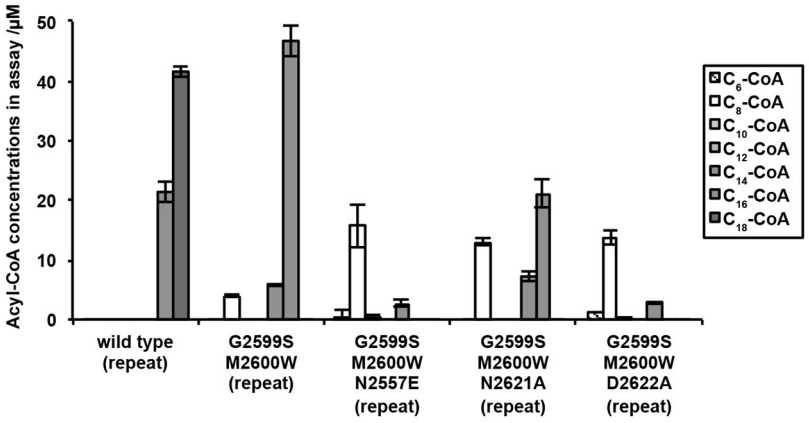
Supplementary Figure 11: Probing decarboxylation as a reason for shorter fatty acids. Shown are the HPLC chromatograms from product assays of FAS$^{wild\ type}$ and FAS$^{G2599S-M2600W}$ where ^{13}C-labelled acetyl-CoA was used as a starter. The exemplary mass spectra of the product peaks of C_8-CoA, C_{16}-CoA and C_{18}-CoA indicate that products mostly derive from the labeled ^{13}C-acetyl-CoA and have a molecular mass of M+2 in comparison to the expected ^{12}C-acetyl-CoA reactions (the intensities of these highest peaks were set to 100%). A small fraction, however, shows this regular mass M, which means, the acyl-CoA synthesis started from an unlabeled acetyl-CoA. Despite using only ^{13}C-labelled acetyl-CoA as a starter, it can be assumed that either a low basal decarboxylation activity[130] of the enzymes or an acetyl-CoA contamination of malonyl-CoA is the explanation. The fractions of unlabeled products in FAS$^{G2599S-M2600W}$ are, however, not significantly different from those found in wild type FAS. Mass spectra are also shown for iC_5-CoA, an internal standard, which shows no M+2 peak.



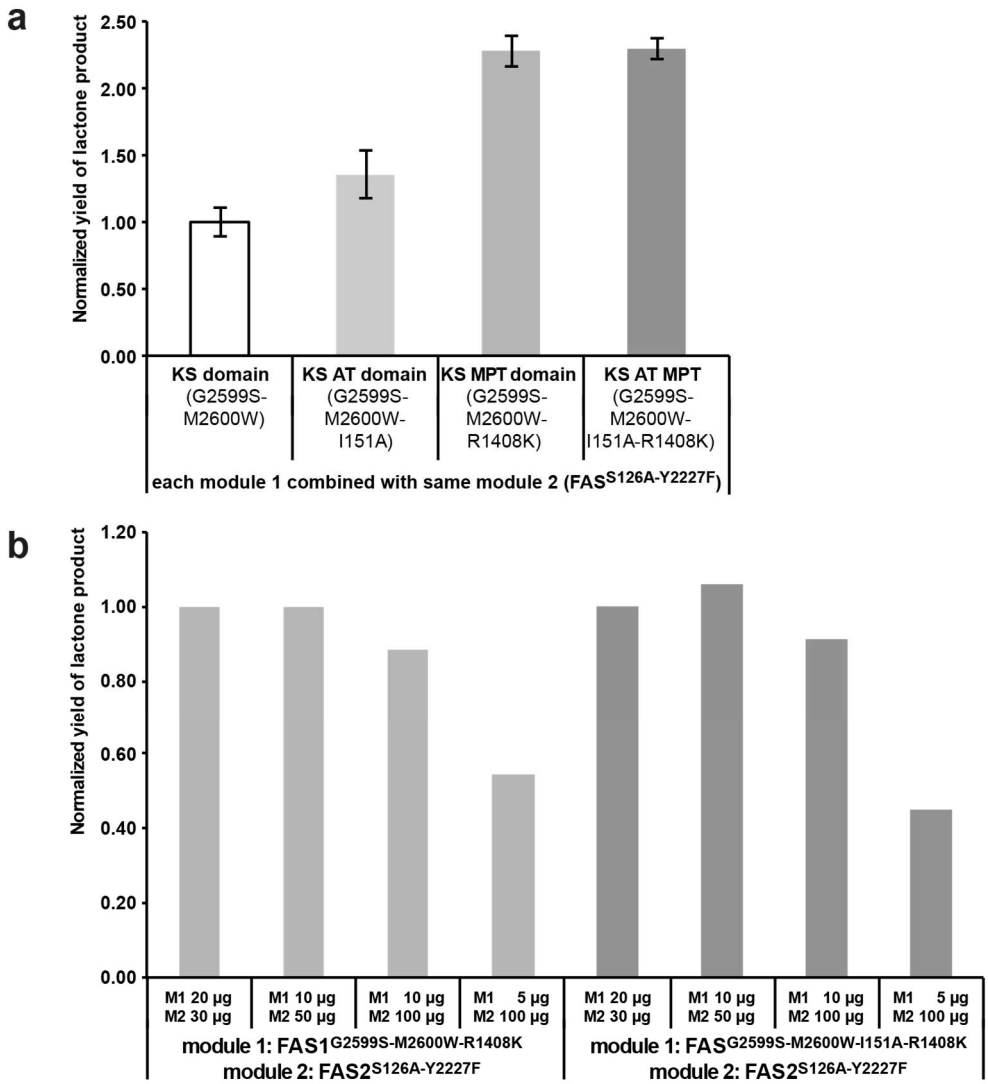
Supplementary Figure 12: Computational output spectra of engineered FAS.
Simulation of product spectra with the kinetic model for the wild type, MPT domain mutant (R1834K, *S. cerevisiae* FAS numbering), the AT domain mutant (I306A, *S. cerevisiae* FAS numbering) and the AT MPT double mutant (I306A-R1834K; combined changed parameters). Average output spectra from the kinetic model with at least 3000 model runs for each chart.



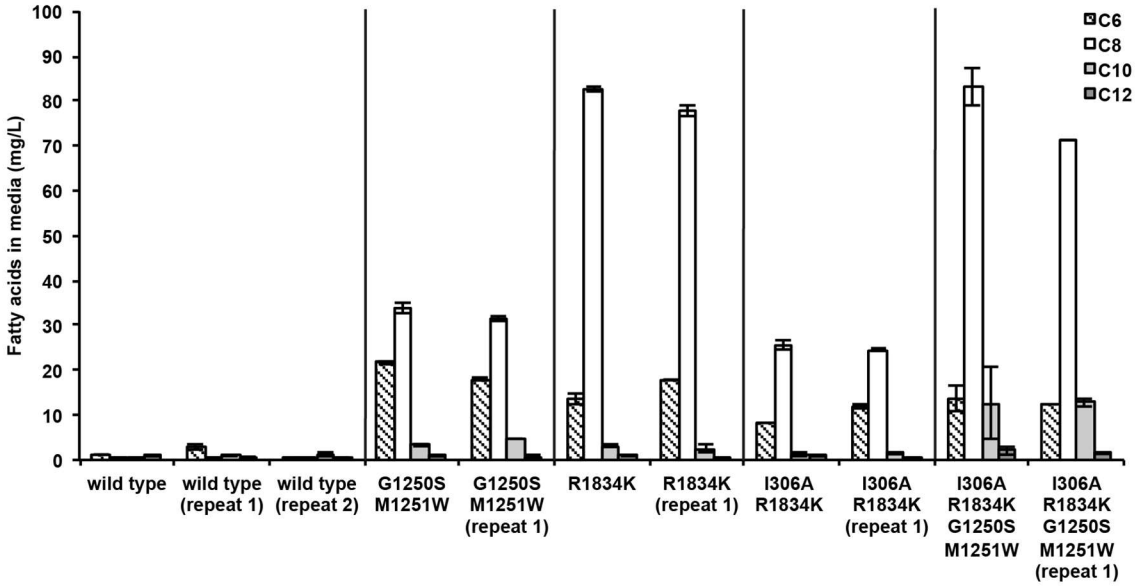
Supplementary Figure 13: Biological repeats and additional product spectra of combinations with L2628Y.
(a) For the most promising constructs, biological repeats were prepared, starting from separate transformations of cells. The results confirm the trends of the samples from the main text. Error bars reflect the standard deviation of three product assays. **(b)** Besides the constructs shown in the main text, additional FAS variants were tested that comprised the L2628Y mutation in combination with other mutations such as I151A (AT domain), R1408K (MPT domain), G2599S (KS domain) and M2600W (KS domain). Error bars reflect the standard deviation of three product assays.



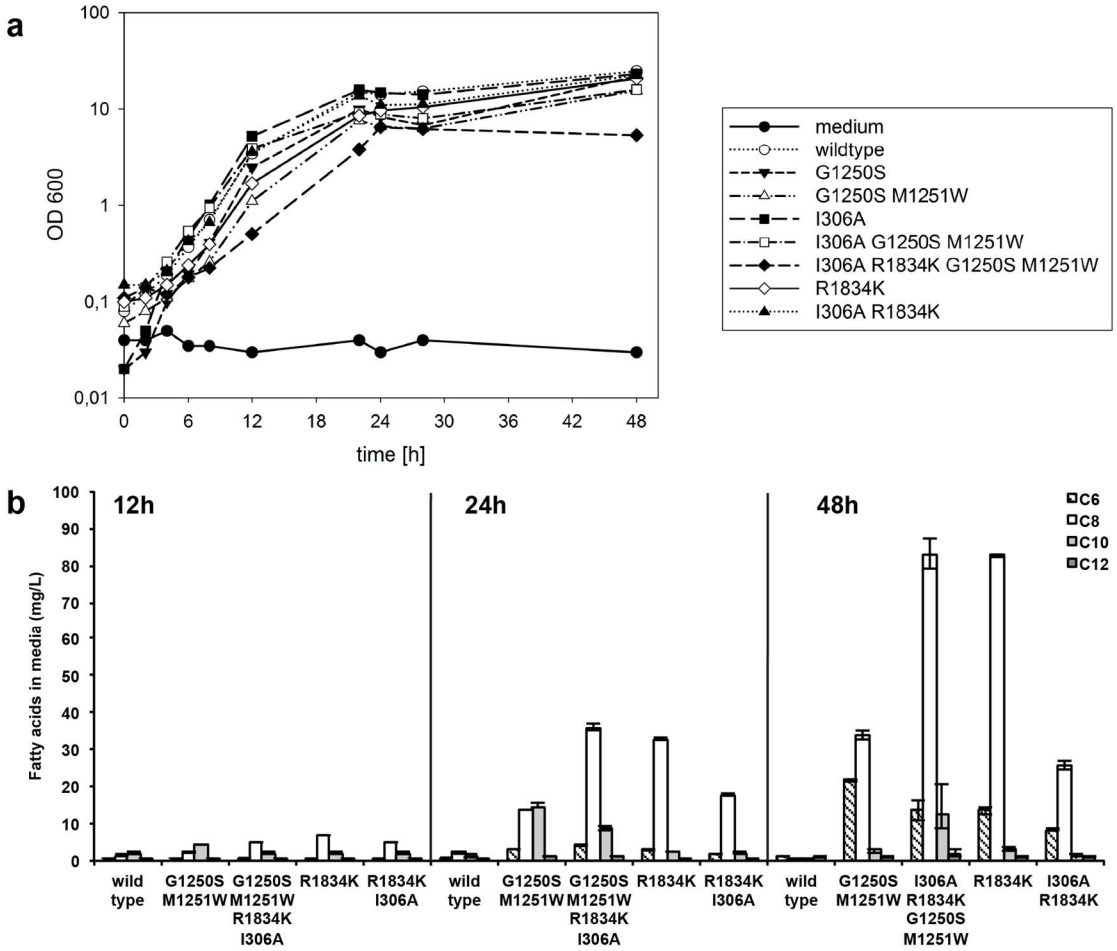
Supplementary Figure 14: Additional product spectra.
Besides the constructs shown in the main text, additional FAS variants were tested that comprised the L2628Y mutation in combination with other mutations such as I151A (AT domain), R1408K (MPT domain), G2599S (KS domain) and M2600W (KS domain). Error bars reflect the standard deviation of three product assays.



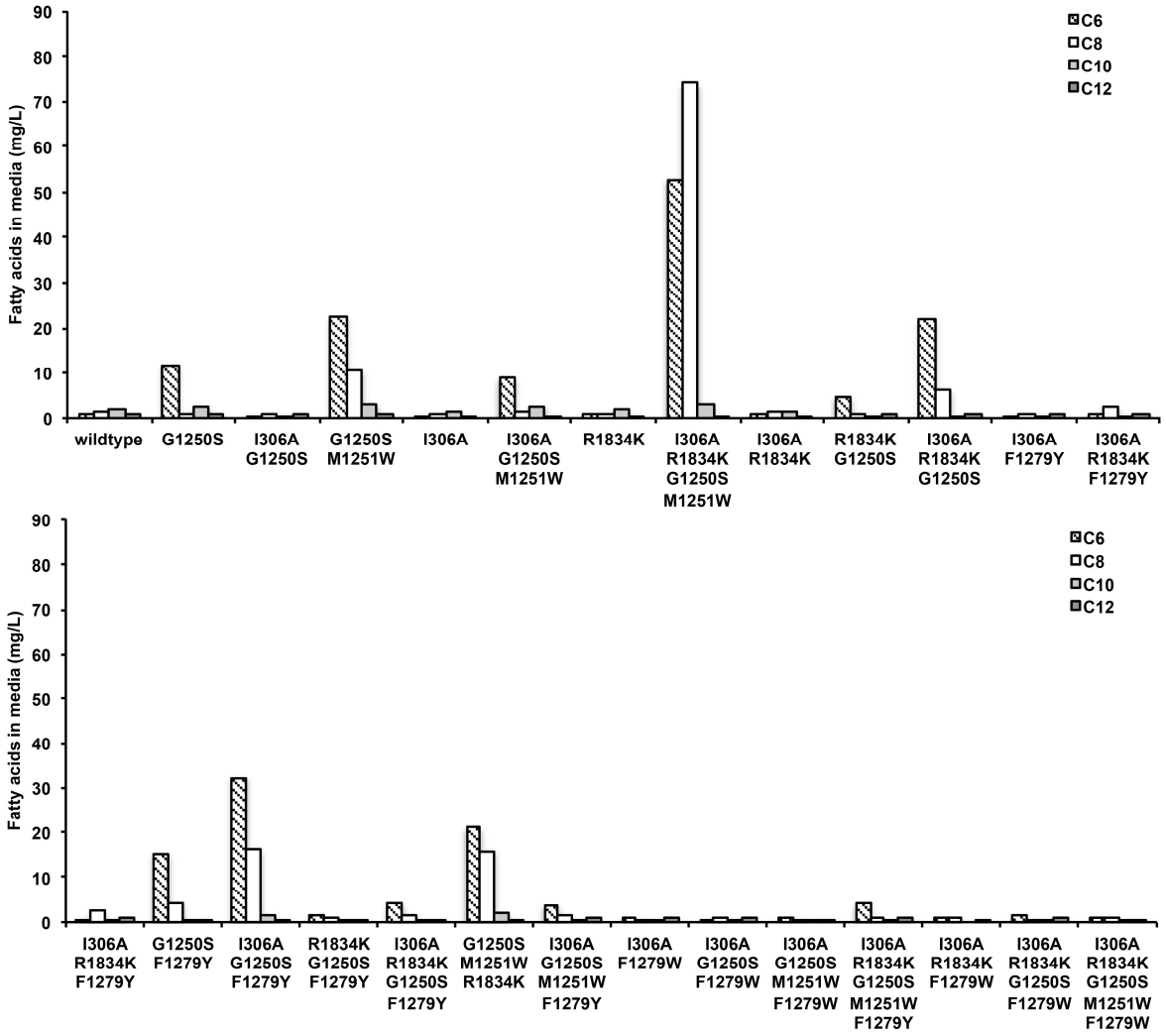
Supplementary Figure 15: Comparison of further variations in the sequential reaction.
(a) 6-HHP output with varied constructs used as module 1 ($FAS^{G2599S-M2600W}$, $FAS^{G2599S-M2600W-I151A}$, $FAS^{G2599S-M2600W-R1408K}$, $FAS^{G2599S-M2600W-I151A-R1408K}$) with the mutated domain shown combined with the same module 2 variant ($FAS^{S126A-Y2227F}$). Output was normalized to the coupled assay using $FAS^{G2599S-M2600W}$ as the module 1 construct. The error bars reflect the standard deviation of three separately prepared assays for each combination. **(b)** A quick scan (only single measurements) of 6-HHP output in correlation to decreasing ratios of module 1 to module 2 for the two combinations $FAS^{G2599S-M2600W-R1408K}$ → $FAS^{S126A-Y2227F}$ and $FAS^{G2599S-M2600W-I151A-R1408K}$ → $FAS^{S126A-Y2227F}$. The amounts of module 1 (M1) and module 2 (M2) reflect how much was used in one sample (standard 100 µL scale). Output for each module 1 / module 2 combination was normalized to its yield under standard conditions where 20 µg of module 1 and 30 µg of module 2 were used.



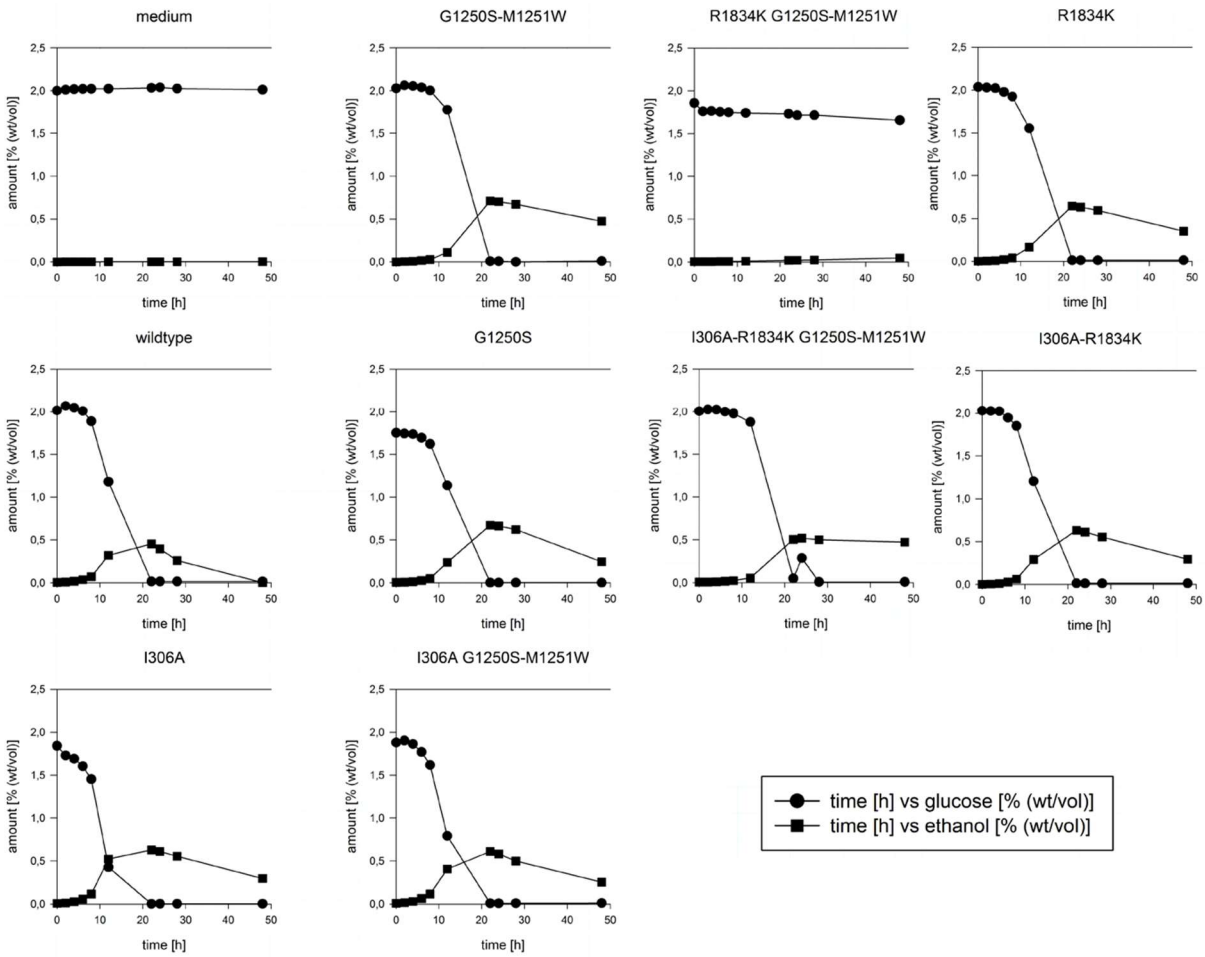
Supplementary Figure 16: Biological repeats of FA production in *S. cerevisiae*.
For the evaluation of biological repeatability, additional strains were grown. Their quantification is shown here in comparison to the samples presented in Figure 21. Error bars shown here reflect three technical repeats. Samples were grown by Renata Pavlovic.



Supplementary Figure 17: Growth properties of FAS engineered strains.
(a) Growth curves of engineered yeast strains (measured by Renata Pavlovic, group of Prof. Eckhard Boles)**.** For selected strains, the cell density was monitored at several time points (average of two measurements from one culture). **(b)** Product spectra at additional time points of selected strains. Besides the regular product spectra after 48 h, additional measurements were performed for selected strains after 12 h and 24 h. Error bars reflect the standard deviation of three technical repeats.



Supplementary Figure 18: Product spectra of FAS strains in YPD with oleic acid addition.
All strains were also grown in YPD supplemented with $C_{18:1}$ (1 mM). The resulting product spectra are significantly different from the experiments in regular YPD which might be explained by the inhibitory effects of $C_{18:1}$ on the whole FA biosynthesis pathway.



Supplementary Figure 19: Glucose and ethanol concentrations during yeast cultivations.
For selected strains, the supernatants of the medium were monitored at several time points during the 48 h cultivation (measured by Renata Pavlovic, group of Prof. Eckhard Boles). The amount of remaining glucose and produced ethanol in the fermentation medium was measured with HPLC. The shown data are results from one typical experiment.

6.3 Supplementary Tables

Supplementary Table 1: Plasmids for the *in vitro* study on chain length control in *C. ammoniagenes* FAS.
The most important plasmids is shown here with the mutation sites listed in the order as they appear on the plasmid. *C. ammoniagenes* FAS is shortened to CamFAS.

Project	Construct	Description in labjournal	Mutated domains	Origin (if not own work)
Engineering substrate affinities in binding channels	pET22b(+)_strepII_wtCamFAS	JAN129		Previous study[14]
	pET22b(+)_strepII_CamFAS_G2599S		KS	Previous master thesis[121]
	pET22b(+)_strepII_CamFAS_G2599S_M2600W		KS	Previous master thesis[121]
	pET22b(+)_strepII_CamFAS_L2628Y	JK37.7	KS	
	pET22b(+)_strepI_CamFAS_G2599S_L2628Y	JAN346.2	KS	
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_L2628Y	JK37.6	KS	
	pET22b(+)_strepII_CamFAS_R1408K	JAN232 C1	MPT	
	pET22b(+)_strepI_CamFAS_R1408K_G2598S_M2599W	JAN232 C2	MPT KS	
	pET22b(+)_strepII_CamFAS_I151A	JAN235 C2	AT	
	pET22b(+)_strepI_CamFAS_I151A_R1408K	JAN353.4B	AT MPT	
	pET22b(+)_strepI_CamFAS_I151A_G2599S	JAN331.2	AT KS	
	pET22b(+)_strepI-CamFAS_I151A_G2598S_M2599W	JAN232 C2	AT KS	
	pET22b(+)_strepI_CamFAS_I151A_L2628Y	JAN353.3	AT KS	
	pET22b(+)_strepI_CamFAS_I151A_G2599S_L2628Y	JAN334.13B	AT KS	
	pET22b(+)_strepI_CamFAS_I151A_R1408K_G2599S	BB6.1	AT MPT KS	
	pET22b(+)_strepI_CamFAS_I151A_R1408K_L2628Y	BB6.3	AT MPT KS	
	pET22b(+)_strepI_CamFAS_I151A_R1408K_G2599S_L2628Y	JAN353.12A	AT MPT KS	
	pET22b(+)_strepI-CamFAS_I151A_R1408K _G2598S_M2599W	JAN242.1	AT MPT KS	

Supplementary Table 1: continued

Project	Construct	Description in labjournal	Mutated domains
Modulating domain-domain interactions	pET22b(+)_strepI_CamFAS_G2598S_M2599W_D2553A	JAN354.3B	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_D2553N	JAN354.1	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_D2556A	JAN354.5G	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_N2557A	JAN354.6B	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_N2557E	JAN354.6A	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_N2621A	JAN354.8B	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_N2621D	MEL2.3B	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_D2622A	JAN354.11C	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_A2696K	JAN354.13A	KS
	pET22b(+)_strepI_CamFAS_G2598S_M2599W_ A2696N	JAN354.13C	KS
	pET22b(+)_strepI_CamFAS_ K1435A _G2598S_M2599W	SW IIIa	MPT KS
	pET22b(+)_strepI_CamFAS_ R1494A _G2598S_M2599W		MPT KS
	pET22b(+)_strepI_CamFAS_Q1647A_G2598S_M2599W	SW IIa	MPT KS
	pET22b(+)_strepI_CamFAS_I151A_L290A_G2598S_M2599W	SW VIIa	AT KS
	pET22b(+)_strepI_CamFAS_I151A_V289A_L290A_G2598S_M2599W	SW Xa	AT KS
	pET22b(+)_strepI_CamFAS_I151A_K1435A_G2598S_M2599W	JAN334.8B	AT MPT KS
	pET22b(+)_strepI_CamFAS_I151A_R1494A_G2598S_M2599W	JAN334.3A	AT MPT KS
	pET22b(+)_strepI_CamFAS_I151A_Q1647A_G2598S_M2599W	JAN334.1A	AT MPT KS
	pET22b(+)_strepI_CamFAS_I151A_V289A_L290A_K1435A_G2598S_M2599W	JAN334.9A	AT MPT KS
	pET22b(+)_strepI_CamFAS_I151A_V289A_L290A_R1494A_G2598S_M2599W	JAN334.5	AT MPT KS
Using FA for the production of a PK	pET22b(+)_strepII_CamFAS_S126A_Y2227F (described in earlier master thesis[121])		AT KR
	pET22b(+)_strepII_CamFAS_S126A_F2226Y_H2806A	JAN266 C3	AT KR KS
	pET22b(+)_strepII_CamFAS_S126A_Y2227F_R1408K	JAN242.4A	AT MPT KR
	pET22b(+)_strepII_CamFAS_S126A_Y2227F		AT KR

Supplementary Table 2: CoA calibration parameters.
Characteristics of one set of calibrations for CoA ester quantification is listed below. The UV trace at 260 nm was used for peak integration. Three replicates (independent dilutions) were used for each calibration point. Furthermore, all calibration points differed no more than 20% from their calculated value (otherwise, the calibration range had to be adjusted.

CoA ester species	Slope (m)	Intercept (c)	Correlation coefficient (R^2)	Calibration range
iC_5-CoA	2,459	0,455	0,9999	10 – 2000 pmol
C_{17}-CoA	2,563	-107,908	0,9989	100 – 2000 pmol
C_6-CoA	2,235	-3,279	0,9999	10 – 2000 pmol
C_8-CoA	2,492	1,271	1,0000	10 – 1000 pmol
C_{10}-CoA	1,862	10,894	0,9975	10 – 2000 pmol
C_{12}-CoA	2,305	-22,675	0,9989	25 – 2000 pmol
C_{14}-CoA	2,259	-18,476	0,9997	25 – 2000 pmol
C_{16}-CoA	2,392	-53,473	0,9999	50 – 2000 pmol
C_{18}-CoA	2,761	-111,587	0,9983	50 – 2000 pmol

Supplementary Table 3: Reactions included in the computational model of the FAS reaction network.

Kinetic parameters were assigned to forward and reverse reactions within constraints as described in the text. We defined a standard reaction rate under saturating conditions of 10 units malonyl-CoA consumed per second. “Priming” reactions are those that must occur once per product release cycle; reaction rates were chosen to ensure a throughput no slower than 1/s at standard substrate concentrations. “Main sequence” reactions are those that would result in a slower overall reaction rate than the experimentally observed ca. 10/s if they were to proceed any slower than this; therefore, reaction rates were chosen to ensure a throughput no slower than 10/s at standard substrate concentrations. The table was initially made by our collaboration partner, the group of Prof. Helmut Grubmüller and taken from the joint publication[17].

Reaction formula	Number of reactions	Class	Restrictions imposed
ACP:C_2 + AT^a <-> ACP + AT^a:C_2	**1**	Priming	>1/s; Michaelis constant: 28 µM
AT^a + AT^b + acetyl:CoA <-> AT^a:C_2 + AT^b:CoA	**1**	Priming	>1/s
AT^b:CoA <-> AT^b + CoA	**1**	Priming	>1/s
mal:CoA + MPT^a + MPTb <-> MPT^a:mal + MPT^b:CoA	**1**	Main sequence	>10/s; Michaelis constant: 8 µM
MPT^a:mal + ACP <-> MPT^a + ACP:mal	**1**	Main sequence	>10/s
MPT^b:CoA <-> MPT^b + CoA	**1**	Main sequence	>10/s
ACP:C_n + KS <-> ACP + KS:C_n	**11**	Main sequence for C_2-C_{18}	>10/s for C_2-C_{18}
ACP:C_n + MPT^a <-> ACP + MPT^a:C_n	**11**	Main sequence for C_{16} and C_{18}	>10/s for C_{16}-C_{18}
KS:C_n + ACP:mal <-> KS + ACP:C_{n+2} + CO_2	**11**	Main sequence for C_2-C_{18}	>10/s for C_2-C_{18}
MPT^a:C_n + MPT^b:CoA <-> MPT^a + MPT^b + C_n:CoA	**11**	Main sequence for C_{16} and C_{18}	>10/s for C_{16}-C_{18}

Supplementary Table 4: List of plasmids for the *in vivo* study on short chain *de novo* FA synthesis in *S. cerevisiae*.
A list of the important plasmids is shown here. *S. cerevisiae* FAS is not mentioned specifically in the name, but FAS1 and FAS2 are indicated.

Gene	Construct	Description in labjournal	Mutated domains
FAS1	PRS315_ScerFAS1wt (provided by Manuel Fischer)		
	PRS315_ScerFAS1_I306A	ASH C3	AT
	PRS315_ScerFAS1_R1834K	JAN303.1 C1	MPT
	PRS315_ScerFAS1_I306A_R1834K	JAN303.3 C1	AT MPT
FAS2	PRS313_ScerFAS2wt (provided by Manuel Fischer)		KS
	PRS313_ScerFAS2_G1250S	ASH C2	KS
	PRS313_ScerFAS2_G1250S_M1251W	ASH C3	KS
	PRS313_ScerFAS2_G1250S	JAN313.1	KS
	PRS313_ScerFAS2_G1250S_F1279Y	JAN313.2	KS
	PRS313_ScerFAS2_G1250S_M1251W_F1279Y	JAN313.3	KS
	PRS313_ScerFAS2_F1279W	JAN313.4	KS
	PRS313_ScerFAS2_G1250S_M1251W_F1279W	JAN313.6	KS

Supplementary Table 5: Cell densities (OD_{600} and wet pellet weight) after 48 h.
Just before further processing, the OD_{600} was measured for all samples, both when they were grown in regular YPD and in YPD supplemented with $C_{18:1}$ (1 mM). The measurements were conducted by Renata Pavlovic (Group of Prof. Eckhard Boles). For the growth in regular YPD, samples could be divided into three groups: regular growth (white background), reduced growth (light gray background) and very little/no growth (dark gray background). In YPD supplemented with $C_{18:1}$ all samples showed nearly the same densities. Also, the wet pellet weight was noted. It is, however, prone to errors since residual media that is stuck to the tube can make the results less reliable. For samples marked with (*), no main culture was inoculated as already the preculture showed no significant growth. OD measurements were conducted by Renata Pavlovic (group of Prof. Eckhard Boles).

	Samples in regular YPD		**Samples in YPD with $C_{18:1}$**	
	OD_{600}	wet cell pellet (g)	OD_{600}	wet cell pellet (g)
wild type	24,7	1,1	17,4	1,3
G1250S	21,9	1,1	19,2	1,4
G1250S M1251W	15,6	1,0	19,5	1,5
I306A	23,1	1,1	19,2	1,4
I306A G1250S M1251W	15,9	1,1	18,8	1,4
R1834K G1250S M1251W	3,5	0,2	22,0	1,5
I306A R1834K G1250S M1251W	5,4	0,5	20,4	1,7
R1834K	20,7	1,1	24,1	1,4
I306 R1834K	22,5	1,2	20,3	1,5
I306A G1250S	15,6	1,3	18,9	1,4
R1834K G1250S	10,2	0,9	19,6	1,4
I306A R1834K G1250S	14,0	1,1	18,3	1,3
G1250S F1279Y	0,1	0,3	17,5	1,4
I306A G1250S F1279Y	0,3	0,3	16,7	1,3
R1834K G1250S F1279Y	0,1	0,3	15,8	1,4
I306A R1834K G1250S F1279Y	0,1	0,3	17,0	1,5
I306A F1279Y	15,6	1,2	16,9	1,3
I306A G1250S M1251W F1279Y	- *	- *	18,2	1,4
I306A F1279W	- *	- *	17,7	1,3
I306A G1250S F1279W	- *	- *	21,8	1,4
I306A G1250S M1251W F1279W	- *	- *	18,4	1,3
I306A R1834K F1279Y	4,3	0,6	21,1	1,4
I306A R1834K G1250S M1251W F1279Y	- *	- *	20,5	1,3
I306A R1834K F1279W	- *	- *	23,6	1,4
I306A R1834K G1250S F1279W	- *	- *	15,8	1,4
I306A R1834K G1250S M1251W F1279W	- *	- *	21,1	1,3

7 Statement of personal contributions

7.1 Scientific work

This thesis developed within close collaborations. For better transparency, the individual contribution of other groups to the scientific work are stated here:

“Engineering substrate affinities in binding channels”

Cloning of the early constructs (please see Supplementary Table 1 for details) was conducted in previous projects in the MPI for Biochemistry, Munich by Andrej Matijczak, Sascha Serdjukov[121] und Mathias Enderle[14].

The group of Prof. Helmut Grubmüller from the MPI for Biophysical Chemistry in Göttingen conducted molecular dynamics simulations and programmed the kinetic model. The paragraphs describing such work were mainly written by them and taken from the joint publication[17]. Above that, effected paragraphs are clearly stated in the text.

Jan Gajewski planned and conducted the biochemical experiments, evaluated and interpreted results, which are shown here in this thesis. The experiments were supported by the work of several intern students (Valentin Hertz, Josefine Kahlstatt, Beata Berlin, Melanie Janßen) under supervision of Jan Gajewski.

“Modulation of domain-domain interactions”

Jan Gajewski planned and conducted all biochemical experiments, evaluated and interpreted results. The experiments were supported by the work of several intern students (Beata Berlin, Robin Schnieders, Melanie Janßen) under supervision of Jan Gajewski.

“Engineering FAS for the production of a polyketide”

Early works were based on a previous project from the MPI for Biochemistry in Munich[121]. However, Jan Gajewski planned and conducted all biochemical experiments, evaluated and interpreted results after the project was taken over from its initial state. The experiments were supported by the work of intern students (Maria Oliferovskaya and Melanie Janßen) under supervision of Jan Gajewski.

"Production of short FA *in vivo* using *S. cerevisiae*"

The group of Prof. Eckhard Boles from the biology department of the Goethe University Frankfurt was involved in planning and conducting of experiments and quantification of products.

In detail: Jan Gajewski planned and conducted the plasmids carrying the mutation, conducted the cloning of mutants, grew the cultures, extracted and processed the cultures. Furthermore, he evaluated and interpreted results as well as planned further experiments.

Renata Pavlovic also grew cultures, measured the OD, also extracted and processed the cultures (especially for biological repeats) and conducted all gas chromatographic as well as metabolites measurements.

The experiments were supported by the work of an intern student (Alexander Hirschhäuser) under supervision of Jan Gajewski.

The initial *S. cerevisiae* strain and the plasmids carrying the wildtype FAS were provided by Manuel Fischer.

7.2 Writing process

As mentioned, the **results section (and according Supplementary Information)** of this thesis is based on manuscripts intended for publication. The personal contribution of the author of this thesis towards the writing process of these manuscripts was as following:

“Engineering binding channels for chain length control” and “Engineering FAS for the production of a polyketide”

50% contributed personally – in detail, especially the paragraphs concerning the work of the group of Prof. Helmut Grubmüller (molecular dynamics simulations and programmed the kinetic model) were written by them. The parts (and also supplementary notes) are specifically marked. For the biochemical study and interpretation of the data, it can be stated that the text was written by Jan Gajewski and in the process of manuscript preparation heavily edited by Prof. Martin Grininger.

“Further KS engineering inspired by secondary metabolite production”

90% contributed personally – the parts were taken from a manuscript for publication. The text was written by Jan Gajewski and in the process of manuscript preparation edited by Prof. Martin Grininger.

“Modulating domain domain interaction for short FA production”

90% contributed personally – the parts were taken from a manuscript for publication. The text was written by Jan Gajewski and in the process of manuscript preparation edited by Prof. Martin Grininger.

“Production of short FA *in vivo* using *S. cerevisiae*”

60% contributed personally – the parts were taken from a manuscript for publication. The text was initially written by Jan Gajewski and in the process of manuscript preparation edited by Prof. Martin Grininger, Renata Pavlovic and Prof. Eckhard Boles.

8 Additional indexes

8.1 List of figures

8.2 List of tables

8.3 Abbreviations

µ	micro
6-HHP	6-heptyl-4-hydroxypyran-2-one
Å	Ångström
AT	acetyl transferase
ACC	acetyl-CoA carboxylase
ACP	Acyl-Carrier-Protein
AcpS	ACP synthase
ADH	alcohol dehydrogenase
bp	base pairs
C. ammoniagenes	*Corynebacterium ammoniagenes*
C_{16}-CoA	Palmitoyl-CoA
C_{18}-CoA	Stearoyl-CoA
CoA	coenzyme A
CV	column volume(s)
DH	dehydratase
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediamine tetraacetic acid
EM	electron microscopy
ER	enoyl reductase
ESI	electron spray ionization
FA	fatty acid(s)
FADO	fatty aldehyde deformylating oxygenase
FAEE	fatty acid ethyl esters
FAME	fatty acid methyl esters
FAR	fatty acyl-CoA reductase
FAS	fatty acid synthase(s)
FDA	U.S. Food and Drug Administration
Fig.	Figure
fwd	forward
GAT	glycerol acyl transferase
GPCR	G-protein coupled receptor

HPLC	high performance liquid chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
KR	ketoacyl reduktase
KS	ketoacyl synthase
L	liter
LB	lysogeny broth
m	milli
M1	module 1
M2	module 2
MAT	malonyl/acetyl transferase
MPT	malonyl/palmitoyl transferase
MT	pseudo methyl transferase
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance spectroscopy
PCR	polymerase chain reaction
PDB	protein data bank
PK	polyketide(s)
PKS	polyketide synthase(s)
PPT	phosphopanteheine transferease
psi	pound per square inch
rcf	relative centrifugal force
rev	reverse
RMSD	relative mean square deviation
rpm	rounds per minute
S. cerevisiae	*Saccharomyces cerevisiae*
TAL	triacetic acid lactone
TB	terrific broth
TE	thioesterase(s)
UV	ultra violet
WES	wax ester synthase

9 Literature

1. Lynen, F., [3] Yeast fatty acid synthase. In *Methods in Enzymology*, John, M. L., Ed. Academic Press: 1969; Vol. Volume 14, pp 17-33.

2. Beld, J.; Lee, D. J.; Burkart, M. D., Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Molecular BioSystems* **2015,** *11* (1), 38-59.

3. Lennen, R. M.; Pfleger, B. F., Microbial production of fatty acid-derived fuels and chemicals. *Current Opinion in Biotechnology* **2013,** *24* (6), 1044-1053.

4. Schweizer, E.; Hofmann, J., Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. *Microbiology and Molecular Biology Reviews* **2004,** *68* (3), 501-517.

5. Lee, S. Y.; Kim, H. M.; Cheon, S., Metabolic engineering for the production of hydrocarbon fuels. *Current Opinion in Biotechnology* **2015,** *33*, 15-22.

6. Peralta-Yahya, P. P.; Zhang, F.; del Cardayre, S. B.; Keasling, J. D., Microbial engineering for the production of advanced biofuels. *Nature* **2012,** *488* (7411), 320-328.

7. Val, D.; Banu, G.; Seshadri, K.; Lindqvist, Y.; Dehesh, K., Re-engineering ketoacyl synthase specificity. *Structure* **2000,** *8* (6), 565-566.

8. Aritomi, K.; Hirosawa, I.; Hoshida, H.; Shiigi, M.; Nishizawa, Y.; Kashiwagi, S.; Akada, R., Self-cloning Yeast Strains Containing Novel FAS2 Mutations Produce a Higher Amount of Ethyl Caproate in Japanese Sake TI. *Bioscience, Biotechnology, and Biochemistry* **2004,** *68* (1), 206-214.

9. Inokoshi, J.; Tomoda, H.; Hashimoto, H.; Watanabe, A.; Takeshima, H.; Ōmura, S., Cerulenin-resistant mutants of <i>Saccharomyces cerevisiae with an altered fatty acid synthase gene. *Molecular and General Genetics MGG* **1994,** *244* (1), 90-96.

10. Leibundgut, M.; Jenni, S.; Frick, C.; Ban, N., Structural Basis for Substrate Delivery by Acyl Carrier Protein in the Yeast Fatty Acid Synthase. *Science* **2007,** *316* (5822), 288-290.

11. Johansson, P.; Mulinacci, B.; Koestler, C.; Vollrath, R.; Oesterhelt, D.; Grininger, M., Multimeric Options for the Auto-Activation of the Saccharomyces cerevisiae FAS Type I Megasynthase. *Structure* **2009,** *17* (8), 1063-1074.

12. Johansson, P.; Wiltschi, B.; Kumari, P.; Kessler, B.; Vonrhein, C.; Vonck, J.; Oesterhelt, D.; Grininger, M., Inhibition of the fungal fatty acid synthase type I multienzyme complex. *Proceedings of the National Academy of Sciences* **2008,** *105* (35), 12803-12808.

13. Lomakin, I. B.; Xiong, Y.; Steitz, T. A., The Crystal Structure of Yeast Fatty Acid Synthase, a Cellular Machine with Eight Active Sites Working Together. *Cell* **2007,** *129* (2), 319-332.

14. Enderle, M.; McCarthy, A.; Paithankar, K. S.; Grininger, M., Crystallization and X-ray diffraction studies of a complete bacterial fatty-acid synthase type I. *Acta Crystallographica Section F* **2015,** *71* (11), 1401-1407.

15. Meurer, G.; Biermann, G.; Schütz, A.; Harth, S.; Schweizer, E., Molecular structure of the multifunctional fatty acid synthetase gene of Brevibacterium ammoniagenes: its sequence of catalytic domains is formally consistent with a head-to-tail fusion of the two yeast genes FAS1 and FAS2. *Molecular and General Genetics MGG* **1992,** *232* (1), 106-116.

16. Christensen, C. E.; Kragelund, B. B.; von Wettstein-Knowles, P.; Henriksen, A., Structure of the human β-ketoacyl [ACP] synthase from the mitochondrial type II fatty acid synthase. *Protein Science* **2007,** *16* (2), 261-272.

17. Gajewski, J.; Buelens, F.; Serdjukow, S.; Janßen, M.; Cortina, N. S.; Grubmüller, H.; Grininger, M., Engineering fatty acid synthases for directed polyketide production. *Nature Chemical Biology* **2017**, *accepted/in press*.

18. Gajewski, J.; Boles, E.; Grininger, M., Microbiological production of short fatty acids and uses thereof. *EP patent application 15 174 342.4 filed on June 29th,* **2015**.

19. Brown, D. W.; Adams, T. H.; Keller, N. P., Aspergillus has distinct fatty acid synthases for primary and secondary metabolism. *Proc Natl Acad Sci U S A* **1996,** *93* (25), 14873-7.

20. Watanabe, C. M. H.; Townsend, C. A., Initial Characterization of a Type I Fatty Acid Synthase and Polyketide Synthase Multienzyme Complex NorS in the Biosynthesis of Aflatoxin B1. *Chemistry & Biology* **2002,** *9* (9), 981-988.

21. Hitchman, T. S.; Schmidt, E. W.; Trail, F.; Rarick, M. D.; Linz, J. E.; Townsend, C. A., Hexanoate Synthase, a Specialized Type I Fatty Acid Synthase in Aflatoxin B1 Biosynthesis. *Bioorganic Chemistry* **2001,** *29* (5), 293-307.

22. Rangan, V. S.; Smith, S., Alteration of the Substrate Specificity of the Malonyl-CoA/Acetyl-CoA:Acyl Carrier ProteinS-Acyltransferase Domain of the

Multifunctional Fatty Acid Synthase by Mutation of a Single Arginine Residue. *Journal of Biological Chemistry* **1997,** *272* (18), 11975-11978.

23. Bunkoczi, G.; Misquitta, S.; Wu, X.; Lee, W. H.; Rojkova, A.; Kochan, G.; Kavanagh, K. L.; Oppermann, U.; Smith, S., Structural Basis for Different Specificities of Acyltransferases Associated with the Human Cytosolic and Mitochondrial Fatty Acid Synthases. *Chemistry & Biology* **2009,** *16* (6), 667-675.

24. Sikorski, R. S.; Hieter, P., A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces Cerevisiae. *Genetics* **1989,** *122* (1), 19-27.

25. Chirala, S. S., Coordinated regulation and inositol-mediated and fatty acid-mediated repression of fatty acid synthase genes in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** *89* (21), 10232-10236.

26. Saerens, S. M. G.; Verstrepen, K. J.; Van Laere, S. D. M.; Voet, A. R. D.; Van Dijck, P.; Delvaux, F. R.; Thevelein, J. M., The Saccharomyces cerevisiae EHT1 and EEB1 genes encode novel enzymes with medium-chain fatty acid ethyl ester synthesis and hydrolysis capacity. *Journal of Biological Chemistry* **2006,** *281* (7), 4446-4456.

27. Knight, M. J.; Bull, I. D.; Curnow, P., The yeast enzyme Eht1 is an octanoyl-CoA:ethanol acyltransferase that also functions as a thioesterase. *Yeast* **2014,** *31* (12), 463-74.

28. Gajewski, J.; Pavlovic, R.; Fischer, M.; Boles, E.; Grininger, M., Engineering fungal *de novo* fatty acid synthesis for short chain fatty acid production, *Nature Communications* **2017**, *accepted/in press*.

29. Gajewski, J.; Pavlovic, R.; Boles, E.; Grininger, M., Microbiological production of short fatty acids and uses thereof. *EP patent application 15 162 192.7 filed on April 1st,* **2015**.

30. Leber, C.; Da Silva, N. A., Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids. *Biotechnology and Bioengineering* **2014,** *111* (2), 347-358.

31. White, S. W.; Zheng, J.; Zhang, Y.-M.; Rock, C. O., The Structural Biology of Type II Fatty Acid Biosynthesis. *Annual Review of Biochemistry* **2005,** *74* (1), 791-831.

32. Morishima, N.; Ikai, A.; Noda, H.; Kawaguchi, A., Structure of bacterial fatty acid synthetase from brevibacterium ammoniagenes. *Biochimica et Biophysica*

Acta (BBA) - Protein Structure and Molecular Enzymology **1982,** *708* (3), 305-312.

33. Lennen, R. M.; Pfleger, B. F., Engineering Escherichia coli to synthesize free fatty acids. *Trends in Biotechnology* **2012,** *30* (0), 659-67.

34. Wakil, S. J.; Stoops, J. K.; Joshi, V. C., Fatty Acid Synthesis and its Regulation. *Annual Review of Biochemistry* **1983,** *52* (1), 537-579.

35. Lynen, F.; Engeser, H.; Foerster, E.-C.; Fox, J. L.; Hess, S.; Kresze, G.-B.; Schmitt, T.; Schreckenbach, T.; Siess, E.; Wieland, F.; Winnewisser, W., On the Structure of Fatty Acid Synthetase of Yeast. *European Journal of Biochemistry* **1980,** *112* (3), 431-442.

36. Jenni, S.; Leibundgut, M.; Maier, T.; Ban, N., Architecture of a Fungal Fatty Acid Synthase at 5 Å Resolution. *Science* **2006,** *311* (5765), 1263-1267.

37. Anselmi, C.; Grininger, M.; Gipson, P.; Faraldo-Gómez, J. D., Mechanism of Substrate Shuttling by the Acyl-Carrier Protein within the Fatty Acid Mega-Synthase. *Journal of the American Chemical Society* **2010,** *132* (35), 12357-12364.

38. Grininger, M., Perspectives on the evolution, assembly and conformational dynamics of fatty acid synthase type I (FAS I) systems. *Current Opinion in Structural Biology* **2014,** *25* (0), 49-56.

39. Boehringer, D.; Ban, N.; Leibundgut, M., 7.5-Å Cryo-EM Structure of the Mycobacterial Fatty Acid Synthase. *Journal of Molecular Biology* **2013,** *425* (0), 841-9.

40. Maier, T.; Leibundgut, M.; Ban, N., The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* **2008,** *321* (5894), 1315-1322.

41. Ciccarelli, L.; Connell, Sean R.; Enderle, M.; Mills, Deryck J.; Vonck, J.; Grininger, M., Structure and Conformational Variability of the Mycobacterium tuberculosis Fatty Acid Synthase Multienzyme Complex. *Structure* **2013,** *21* (7), 1251-1257.

42. Brignole, E. J.; Smith, S.; Asturias, F. J., Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat Struct Mol Biol* **2009,** *16* (2), 190-197.

43. Hertweck, C., The Biosynthetic Logic of Polyketide Diversity. *Angewandte Chemie International Edition* **2009,** *48* (26), 4688-4716.

44. Woloshuk, C. P.; Prieto, R., Genetic organization and function of the aflatoxin B1 biosynthetic genes. *FEMS Microbiology Letters* **1998,** *160* (2), 169-176.

45. Ladner, C. C.; Williams, G. J., Harnessing natural product assembly lines: structure, promiscuity, and engineering. *J Ind Microbiol Biotechnol* **2015,** *43* (2), 371-387.

46. Pickens, L. B.; Tang, Y.; Chooi, Y.-H., Metabolic Engineering for the Production of Natural Products. *Annual Review of Chemical and Biomolecular Engineering* **2011,** *2* (1), 211-236.

47. Khosla, C., Harnessing the Biosynthetic Potential of Modular Polyketide Synthases. *Chemical Reviews* **1997,** *97* (7), 2577-2590.

48. Walsh, C. T.; Fischbach, M. A., Natural Products Version 2.0: Connecting Genes to Molecules. *Journal of the American Chemical Society* **2010,** *132* (8), 2469-2493.

49. Shen, B., Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* **2003,** *7* (2), 285-295.

50. Cox, R. J., Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Organic & Biomolecular Chemistry* **2007,** *5* (13), 2010-2026.

51. Keatinge-Clay, A. T., The structures of type I polyketide synthases. *Natural Product Reports* **2012,** *29* (10), 1050-1073.

52. Weissman, K. J., Uncovering the structures of modular polyketide synthases. *Natural Product Reports* **2015,** *32* (3), 436-453.

53. Herbst, D. A.; Jakob, R. P.; Zähringer, F.; Maier, T., Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **2016,** *advance online publication*.

54. Zheng, J.; Taylor, C. A.; Piasecki, S. K.; Keatinge-Clay, A. T., Structural and Functional Analysis of A-Type Ketoreductases from the Amphotericin Modular Polyketide Synthase. *Structure* **2010,** *18* (8), 913-922.

55. Edwards, A. L.; Matsui, T.; Weiss, T. M.; Khosla, C., Architectures of Whole-Module and Bimodular Proteins from the 6-Deoxyerythronolide B Synthase. *Journal of Molecular Biology* **2014,** *426* (11), 2229-2245.

56. Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R. H.; Hakansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis,

G., Structure of a modular polyketide synthase. *Nature* **2014,** *510* (7506), 512-517.

57. Rittner, A.; Grininger, M., Modular Polyketide Synthases (PKSs): A New Model Fits All? *ChemBioChem* **2014,** *15* (17), 2489-2493.

58. Herbst, D. A.; Jakob, R. P.; Zähringer, F.; Maier, T., Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **2016,** *531* (7595), 533-537.

59. Pfleger, B. F.; Gossing, M.; Nielsen, J., Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabolic Engineering* **2015,** *29*, 1-11.

60. Runguphan, W.; Keasling, J. D., Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. *Metabolic Engineering* **2014,** *21* (0), 103-113.

61. Steen, E. J.; Kang, Y.; Bokinsky, G.; Hu, Z.; Schirmer, A.; McClure, A.; del Cardayre, S. B.; Keasling, J. D., Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **2010,** *463* (7280), 559-562.

62. Blatti, J. L.; Michaud, J.; Burkart, M. D., Engineering fatty acid biosynthesis in microalgae for sustainable biodiesel. *Current Opinion in Chemical Biology* **2013,** *17* (3), 496-505.

63. Harwood, J. L., Fatty Acid Metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology* **1988,** *39* (1), 101-138.

64. Rutter, C. D.; Zhang, S. Y.; Rao, C. V., Engineering Yarrowia lipolytica for production of medium-chain fatty acids. *Applied Microbiology and Biotechnology* **2015,** *99* (17), 7359-7368.

65. Grunnet, I.; Knudsen, J., Fatty-Acid Synthesis in Lactating-Goat Mammary Gland 1. Medium-Chain Fatty-Acid Synthesis. *European Journal of Biochemistry* **1979,** *95* (3), 497-502.

66. Hiltunen, J. K.; Chen, Z.; Haapalainen, A. M.; Wierenga, R. K.; Kastaniotis, A. J., Mitochondrial fatty acid synthesis – An adopted set of enzymes making a pathway of major importance for the cellular metabolism. *Progress in Lipid Research* **2010,** *49* (1), 27-45.

67. Chia, M.; Schwartz, T. J.; Shanks, B. H.; Dumesic, J. A., Triacetic acid lactone as a potential biorenewable platform chemical. *Green Chemistry* **2012,** *14* (7), 1850-1853.

68. Marten, B.; Pfeuffer, M.; Schrezenmeir, J., Medium-chain triglycerides. *International Dairy Journal* **2006,** *16* (11), 1374-1382.

69. Romero-Guido, C.; Belo, I.; Ta, T. M. N.; Cao-Hoang, L.; Alchihab, M.; Gomes, N.; Thonart, P.; Teixeira, J. A.; Destain, J.; Waché, Y., Biochemistry of lactone formation in yeast and fungi and its utilisation for the production of flavour and fragrance compounds. *Applied Microbiology and Biotechnology* **2011,** *89* (3), 535-547.

70. Mason, A. B.; Dufour, J. P., Alcohol acetyltransferases and the significance of ester synthesis in yeast. *Yeast* **2000,** *16* (14), 1287-1298.

71. Brault, G.; Shareck, F.; Hurtubise, Y.; Lepine, F.; Doucet, N., Short-Chain Flavor Ester Synthesis in Organic Media by an E-coli Whole-Cell Biocatalyst Expressing a Newly Characterized Heterologous Lipase. *PLoS ONE* **2014,** *9* (3).

72. Ferrero, M. A.; Reglero, A.; Martinvillacorta, J.; Fernandezcanon, J. M.; Luengo, J. M., Biosynthesis of benzylpenicillin (G), phenoxymethylpenicillin (V) and octanoylpenicillin (K) from glutathione S-derivatives. *Journal of Antibiotics* **1990,** *43* (6), 684-691.

73. Xu, Y.; Zhou, T.; Espinosa-Artiles, P.; Tang, Y.; Zhan, J.; Molnár, I., Insights into the Biosynthesis of 12-Membered Resorcylic Acid Lactones from Heterologous Production in Saccharomyces cerevisiae. *ACS Chemical Biology* **2014,** *9* (5), 1119-1127.

74. Xu, Y. Q.; Zhou, T.; Zhou, Z. F.; Su, S. Y.; Roberts, S. A.; Montfort, W. R.; Zeng, J.; Chen, M.; Zhang, W.; Lin, M.; Zhan, J. X.; Molnar, I., Rational reprogramming of fungal polyketide first-ring cyclization. *Proceedings of the National Academy of Sciences of the United States of America* **2013,** *110* (14), 5398-5403.

75. Kraus, G. A.; Basemann, K.; Guney, T., Selective pyrone functionalization: reductive alkylation of triacetic acid lactone. *Tetrahedron Letters* **2015,** *56* (23), 3494-3496.

76. Chia, M.; Haider, M. A.; Pollock, G.; Kraus, G. A.; Neurock, M.; Dumesic, J. A., Mechanistic Insights into Ring-Opening and Decarboxylation of 2-Pyrones in Liquid Water and Tetrahydrofuran. *Journal of the American Chemical Society* **2013,** *135* (15), 5699-5708.

77. Vandamme, E. J.; Soetaert, W., Bioflavours and fragrances via fermentation and biocatalysis. *Journal of Chemical Technology and Biotechnology* **2002,** *77* (12), 1323-1332.

78. Kang, W.-R.; Seo, M.-J.; An, J.-U.; Shin, K.-C.; Oh, D.-K., Production of δ-decalactone from linoleic acid via 13-hydroxy-9(Z)-octadecenoic acid intermediate by one-pot reaction using linoleate 13-hydratase and whole Yarrowia lipolytica cells. *Biotechnology Letters* **2016,** *38* (5), 817-823.

79. Haushalter, R. W.; Groff, D.; Deutsch, S.; The, L.; Chavkin, T. A.; Brunner, S. F.; Katz, L.; Keasling, J. D., Development of an orthogonal fatty acid biosynthesis system in E. coli for oleochemical production. *Metabolic Engineering* **2015,** *30* (0), 1-6.

80. Steiner, K.; Schwab, H., *Recent advances in rational approaches for enzyme engineering*. 2012.

81. Heath, R. J.; Rock, C. O., The Claisen condensation in biology. *Natural Product Reports* **2002,** *19* (5), 581-596.

82. Bloch, K.; Vance, D., Control Mechanisms in the Synthesis of Saturated Fatty Acids. *Annual Review of Biochemistry* **1977,** *46* (1), 263-298.

83. Sumper, M.; Riepertinger, C.; Lynen, F.; Oesterhelt, D., Die Synthese verschiedener Carbonsäuren durch den Multienzymkomplex der Fettsäuresynthese aus Hefe und die Erklärung ihrer Bildung. *European Journal of Biochemistry* **1969,** *10* (2), 377-387.

84. Jenner, M.; Frank, S.; Kampa, A.; Kohlhaas, C.; Pöplau, P.; Briggs, G. S.; Piel, J.; Oldham, N. J., Substrate Specificity in Ketosynthase Domains from trans-AT Polyketide Synthases. *Angewandte Chemie International Edition* **2013,** *52* (4), 1143-1147.

85. Jenner, M.; Afonso, J. P.; Bailey, H. R.; Frank, S.; Kampa, A.; Piel, J.; Oldham, N. J., Acyl-Chain Elongation Drives Ketosynthase Substrate Selectivity in trans-Acyltransferase Polyketide Synthases. *Angew. Chem.-Int. Edit.* **2015,** *54* (6), 1817-1821.

86. Gajiwala, K. S.; Margosiak, S.; Lu, J.; Cortez, J.; Su, Y.; Nie, Z.; Appelt, K., Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Letters* **2009,** *583* (17), 2939-2946.

87. Olsen, J. G.; Kadziola, A.; von Wettstein-Knowles, P.; Siggaard-Andersen, M.; Larsen, S., Structures of β-Ketoacyl-Acyl Carrier Protein Synthase I Complexed with Fatty Acids Elucidate its Catalytic Machinery. *Structure* **2001,** *9* (3), 233-243.

88. Moche, M.; Schneider, G.; Edwards, P.; Dehesh, K.; Lindqvist, Y., Structure of the Complex between the Antibiotic Cerulenin and Its Target, β-Ketoacyl-Acyl

Carrier Protein Synthase. *Journal of Biological Chemistry* **1999,** *274* (10), 6031-6034.

89. Chen, Y.; Kelly, E. E.; Masluk, R. P.; Nelson, C. L.; Cantu, D. C.; Reilly, P. J., Structural classification and properties of ketoacyl synthases. *Protein Science* **2011,** *20* (10), 1659-1667.

90. von Wettstein-Knowles, P.; Olsen, J. G.; McGuire, K. A.; Henriksen, A., Fatty acid synthesis. *FEBS Journal* **2006,** *273* (4), 695-710.

91. Torella, J. P.; Ford, T. J.; Kim, S. N.; Chen, A. M.; Way, J. C.; Silver, P. A., Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proceedings of the National Academy of Sciences* **2013,** *110* (28), 11290-11295.

92. Choi, Y. J.; Lee, S. Y., Microbial production of short-chain alkanes. *Nature* **2013,** *502* (7472), 571-574.

93. Kawaguchi, A.; Arai, K.; Seyama, Y.; Yamakawa, T.; Okuda, S., Substrate Control of Termination of Fatty Acid Biosynthesis by Fatty Acid Synthetase from Brevibacterium ammoniagenes. *The Journal of Biochemistry* **1980,** *88* (2), 303-306.

94. Schweizer, E.; Lerch, I.; Kroeplin-Rueff, L.; Lynen, F., Fatty Acyl Transferase. *European Journal of Biochemistry* **1970,** *15* (3), 472-482.

95. Engeser, H.; Hübner, K.; Straub, J.; Lynen, F., Identity of Malonyl and Palmitoyl Transferase of Fatty Acid Synthetase from Yeast. *European Journal of Biochemistry* **1979,** *101* (2), 407-412.

96. Pirson, W.; Schuhmann, L.; Lynen, F., The Specificity of Yeast Fatty-Acid Synthetase with Respect to the "Priming" Substrate. Decanoyl-CoA and Derivatives as "Primers" of Fatty-Acid Synthesis in vitro. *European Journal of Biochemistry* **1973,** *36* (1), 16-24.

97. Gipson, P.; Mills, D. J.; Wouts, R.; Grininger, M.; Vonck, J.; Kuhlbrandt, W., Direct structural insight into the substrate-shuttling mechanism of yeast fatty acid synthase by electron cryomicroscopy. *Proceedings of the National Academy of Sciences* **2010,** *107* (20), 9164-9169.

98. Cronan, J. E., The chain-flipping mechanism of ACP (acyl carrier protein)-dependent enzymes appears universal. *Biochemical Journal* **2014,** *460*, 157-163.

99. Liu, X.; Hicks, W. M.; Silver, P. A.; Way, J. C., Engineering acyl carrier protein to enhance production of shortened fatty acids. *Biotechnol. Biofuels* **2016,** *9* (1), 1-7.

100. Perez, D. R.; Leibundgut, M.; Wider, G., Interactions of the Acyl Chain with the Saccharomyces cerevisiae Acyl Carrier Protein. *Biochemistry* **2015,** *54* (13), 2205-2213.

101. Worthington, A. S.; Hur, G. H.; Meier, J. L.; Cheng, Q.; Moore, B. S.; Burkart, M. D., Probing the Compatibility of Type II Ketosynthase–Carrier Protein Partners. *ChemBioChem* **2008,** *9* (13), 2096-2103.

102. Beld, J.; Cang, H.; Burkart, M. D., Visualizing the Chain-Flipping Mechanism in Fatty-Acid Biosynthesis. *Angew. Chem.-Int. Edit.* **2014,** *53* (52), 14456-14461.

103. Ishikawa, F.; Haushalter, R. W.; Lee, D. J.; Finzel, K.; Burkart, M. D., Sulfonyl 3-Alkynyl Pantetheinamides as Mechanism-Based Cross-Linkers of Acyl Carrier Protein Dehydratase. *Journal of the American Chemical Society* **2013,** *135* (24), 8846-8849.

104. Finzel, K.; Lee, D. J.; Burkart, M. D., Using Modern Tools To Probe the Structure-Function Relationship of Fatty Acid Synthases. *ChemBioChem* **2015,** *16* (4), 528-547.

105. Worthington, A. S.; Burkart, M. D., One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Organic & Biomolecular Chemistry* **2006,** *4* (1), 44-46.

106. Marshall, C. G.; Burkart, M. D.; Meray, R. K.; Walsh, C. T., Carrier Protein Recognition in Siderophore-Producing Nonribosomal Peptide Synthetases. *Biochemistry* **2002,** *41* (26), 8429-8437.

107. Blatti, J. L.; Beld, J.; Behnke, C. A.; Mendez, M.; Mayfield, S. P.; Burkart, M. D., Manipulating Fatty Acid Biosynthesis in Microalgae for Biofuel through Protein-Protein Interactions. *PLoS ONE* **2012,** *7* (9), e42949.

108. Cantu, D. C.; Chen, Y.; Reilly, P. J., Thioesterases: A new perspective based on their primary and tertiary structures. *Protein Science : A Publication of the Protein Society* **2010,** *19* (7), 1281-1295.

109. Liu, T.; Vora, H.; Khosla, C., Quantitative analysis and engineering of fatty acid biosynthesis in E. coli. *Metabolic Engineering* **2010,** *12* (4), 378-386.

110. Voelker, T.; Worrell, A.; Anderson, L.; Bleibaum, J.; Fan, C.; Hawkins, D.; Radke, S.; Davies, H., Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. *Science* **1992,** *257* (5066), 72-74.

111. Zhang, F.; Ouellet, M.; Batth, T. S.; Adams, P. D.; Petzold, C. J.; Mukhopadhyay, A.; Keasling, J. D., Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **2012,** *14* (6), 653-660.

112. Leber, C.; Polson, B.; Fernandez-Moya, R.; Da Silva, N. A., Overproduction and secretion of free fatty acids through disrupted neutral lipid recycle in Saccharomyces cerevisiae. *Metabolic Engineering* **2015,** *28* (0), 54-62.

113. Zhou, Y. J.; Buijs, N. A.; Zhu, Z.; Qin, J.; Siewers, V.; Nielsen, J., Production of fatty acid-derived oleochemicals and biofuels by synthetic yeast cell factories. *Nature Communications* **2016,** *7*, 11709.

114. Tee, T. W.; Chowdhury, A.; Maranas, C. D.; Shanks, J. V., Systems metabolic engineering design: Fatty acid production as an emerging case study. *Biotechnology and Bioengineering* **2014,** *111* (5), 849-857.

115. Sheng, J.; Feng, X., Metabolic engineering of yeast to produce fatty acid-derived biofuels: bottlenecks and solutions. *Frontiers in Microbiology* **2015,** *6*.

116. Tang, X.; Lee, J.; Chen, W. N., Engineering the fatty acid metabolic pathway in Saccharomyces cerevisiae for advanced biofuel production. *Metabolic Engineering Communications* **2015,** *2*, 58-66.

117. Woods, A.; Munday, M. R.; Scott, J.; Yang, X.; Carlson, M.; Carling, D., Yeast SNF1 is functionally related to mammalian AMP-activated protein kinase and regulates acetyl-CoA carboxylase in vivo. *Journal of Biological Chemistry* **1994,** *269* (30), 19509-19515.

118. Hofbauer, Harald F.; Schopf, Florian H.; Schleifer, H.; Knittelfelder, Oskar L.; Pieber, B.; Rechberger, Gerald N.; Wolinski, H.; Gaspar, Maria L.; Kappe, C. O.; Stadlmann, J.; Mechtler, K.; Zenz, A.; Lohner, K.; Tehlivets, O.; Henry, Susan A.; Kohlwein, Sepp D., Regulation of Gene Expression through a Transcriptional Repressor that Senses Acyl-Chain Length in Membrane Phospholipids. *Developmental Cell* **2014,** *29* (6), 729-739.

119. d'Espaux, L.; Mendez-Perez, D.; Li, R.; Keasling, J. D., Synthetic biology for microbial production of lipid-based biofuels. *Current Opinion in Chemical Biology* **2015,** *29*, 58-65.

120. Leber, C.; Choi, J. W.; Polson, B.; Da Silva, N. A., Disrupted short chain specific β-oxidation and improved synthase expression increase synthesis of short chain fatty acids in Saccharomyces cerevisiae. *Biotechnology and Bioengineering* **2015**, n/a-n/a.

121. Serdjukov, S., Master Thesis "Efforts Towards a Directed Polyketide Synthesis". *Ludwig-Maximilians-University Munich/Max Planck Institute of Biochemistry, Martinsried* **2011**.

122. Schmidt, T. G. M.; Skerra, A., The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protocols* **2007,** *2* (6), 1528-1535.

123. Xie, D.; Shao, Z.; Achkar, J.; Zha, W.; Frost, J. W.; Zhao, H., Microbial synthesis of triacetic acid lactone. *Biotechnology and Bioengineering* **2006,** *93* (4), 727-736.

124. Schiestl, R.; Gietz, R. D., High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* **1989,** *16* (5-6), 339-346.

125. Ichihara, K. i.; Fukubayashi, Y., Preparation of fatty acid methyl esters for gas-liquid chromatography. *Journal of Lipid Research* **2010,** *51* (3), 635-640.

126. Stuible, H. P.; Wagner, C.; Andreou, I.; Huter, G.; Haselmann, J.; Schweizer, E., Identification and functional differentiation of two type I fatty acid synthases in Brevibacterium ammoniagenes. *Journal of Bacteriology* **1996,** *178* (16), 4787-4793.

127. Stuible, H.-P.; Meier, S.; Schweizer, E., Identification, Isolation and Biochemical Characterization of a Phosphopantetheine:Protein Transferase that Activates the Two Type-I Fatty Acid Synthases of Brevibacterium Ammoniagenes. *European Journal of Biochemistry* **1997,** *248* (2), 481-487.

128. Zhang, L.; Joshi, A. K.; Hofmann, J.; Schweizer, E.; Smith, S., Cloning, Expression, and Characterization of the Human Mitochondrial β-Ketoacyl Synthase: COMPLEMENTATION OF THE YEAST CEM1 KNOCK-OUT STRAIN. *Journal of Biological Chemistry* **2005,** *280* (13), 12422-12429.

129. Gillespie, D. T., Exact stochastic simulation of coupled chemical-reactions. *Journal of Physical Chemistry* **1977,** *81* (25), 2340-2361.

130. Kresze, G.-B.; Steber, L.; Oesterhelt, D.; Lynen, F., Reaction of Yeast Fatty Acid Synthetase with Iodoacetamide. *European Journal of Biochemistry* **1977,** *79* (1), 191-199.

131. Engeser, H.; Hübner, K.; Straub, J.; Lynen, F., Identity of Malonyl and Palmitoyl Transferase of Fatty Acid Synthetase from Yeast. *European Journal of Biochemistry* **1979,** *101* (2), 413-422.

132. Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolásek, B.; Ban, N., Structure of Fungal Fatty Acid Synthase and Implications for Iterative Substrate Shuttling. *Science* **2007,** *316* (5822), 254-261.

133. Crosby, J.; Crump, M. P., The structural role of the carrier protein - active controller or passive carrier. *Natural Product Reports* **2012,** *29* (10), 1111-1137.

134. Ye, Z.; Williams, G. J., Mapping a Ketosynthase:Acyl Carrier Protein Binding Interface via Unnatural Amino Acid-Mediated Photo-Cross-Linking. *Biochemistry* **2014,** *53* (48), 7494-7502.

135. Borgaro, J. G.; Chang, A.; Machutta, C. A.; Zhang, X.; Tonge, P. J., Substrate Recognition by beta-Ketoacyl-ACP Synthases. *Biochemistry* **2011,** *50* (49), 10678-10686.

136. Arthur, C. J.; Williams, C.; Pottage, K.; Płoskoń, E.; Findlow, S. C.; Burston, S. G.; Simpson, T. J.; Crump, M. P.; Crosby, J., Structure and Malonyl CoA-ACP Transacylase Binding of Streptomyces coelicolor Fatty Acid Synthase Acyl Carrier Protein. *ACS Chemical Biology* **2009,** *4* (8), 625-636.

137. Stuible, H.-P.; Meurer, G.; Schweizer, E., Heterologous Expression and Biochemical Characterization of two Functionally Different Type I Fatty Acid Synthases from Brevibacterium Ammoniagenes. *European Journal of Biochemistry* **1997,** *247* (1), 268-273.

138. Fischer, M.; Rhinow, D.; Zhu, Z.; Mills, D. J.; Zhao, Z. K.; Vonck, J.; Grininger, M., Cryo-EM structure of fatty acid synthase (FAS) from Rhodosporidium toruloides provides insights into the evolutionary development of fungal FAS. *Protein Science* **2015,** *24* (6), 987-995.

139. Keatinge-Clay, A. T.; Stroud, R. M., The Structure of a Ketoreductase Determines the Organization of the β-Carbon Processing Enzymes of Modular Polyketide Synthases. *Structure* **2006,** *14* (4), 737-748.

140. Zha, W.; Shao, Z.; Frost, J. W.; Zhao, H., Rational Pathway Engineering of Type I Fatty Acid Synthase Allows the Biosynthesis of Triacetic Acid Lactone from d-Glucose in Vivo. *Journal of the American Chemical Society* **2004,** *126* (14), 4534-4535.

141. Zhang, Y.-M.; Hurlbert, J.; White, S. W.; Rock, C. O., Roles of the Active Site Water, Histidine 303, and Phenylalanine 396 in the Catalytic Mechanism of the Elongation Condensing Enzyme of Streptococcus pneumoniae. *Journal of Biological Chemistry* **2006,** *281* (25), 17390-17399.

142. Borrull, A.; López-Martínez, G.; Poblet, M.; Cordero-Otero, R.; Rozès, N., New insights into the toxicity mechanism of octanoic and decanoic acids on Saccharomyces cerevisiae. *Yeast* **2015**, n/a-n/a.

143. Liu, P.; Chernyshov, A.; Najdi, T.; Fu, Y.; Dickerson, J.; Sandmeyer, S.; Jarboe, L., Membrane stress caused by octanoic acid in Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology* **2013,** *97* (7), 3239-3251.

144. Ogiwara, H.; Tanabe, T.; Nikawa, J.-i.; Numa, S., Inhibition of Rat-Liver Acetyl-Coenzyme-A Carboxylase by Palmitoyl-Coenzyme A. *European Journal of Biochemistry* **1978,** *89* (1), 33-41.

145. Gerlt, J. A.; Babbitt, P. C., Enzyme (re)design: lessons from natural evolution and computation. *Current Opinion in Chemical Biology* **2009,** *13* (1), 10-18.

146. Oefner, C.; Schulz, H.; D'Arcy, A.; Dale, G. E., Mapping the active site of Escherichia coli malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography. *Acta Crystallographica Section D* **2006,** *62* (6), 613-618.

147. Chohnan, S.; Furukawa, H.; Fujio, T.; Nishihara, H.; Takamura, Y., Changes in the size and composition of intracellular pools of nonesterified coenzyme A and coenzyme A thioesters in aerobic and facultatively anaerobic bacteria. *Applied and Environmental Microbiology* **1997,** *63* (2), 553-560.

148. Rafi, S.; Novichenok, P.; Kolappan, S.; Zhang, X.; Stratton, C. F.; Rawat, R.; Kisker, C.; Simmerling, C.; Tonge, P. J., Structure of Acyl Carrier Protein Bound to FabI, the FASII Enoyl Reductase from Escherichia coli. *Journal of Biological Chemistry* **2006,** *281* (51), 39285-39293.

149. Zhou, Y. J.; Buijs, N. A.; Siewers, V.; Nielsen, J., Fatty acid-derived biofuels and chemicals production in Saccharomyces cerevisiae. *Frontiers in Bioengineering and Biotechnology* **2014,** *2*.

150. Nielsen, J.; Larsson, C.; van Maris, A.; Pronk, J., Metabolic engineering of yeast for production of fuels and chemicals. *Current Opinion in Biotechnology* **2013,** *24* (3), 398-404.

151. Legras, J. L.; Erny, C.; Le Jeune, C.; Lollier, M.; Adolphe, Y.; Demuyter, C.; Delobel, P.; Blondin, B.; Karst, F., Activation of Two Different Resistance Mechanisms in Saccharomyces cerevisiae upon Exposure to Octanoic and Decanoic Acids. *Applied and Environmental Microbiology* **2010,** *76* (22), 7526-7535.

152. Wieland, F.; Renner, L.; VerfÜRth, C.; Lynen, F., Studies on the Multi-Enzyme Complex of Yeast Fatty-Acid Synthetase. *European Journal of Biochemistry* **1979,** *94* (1), 189-197.

153. Fichtlscherer, F.; Wellein, C.; Mittag, M.; Schweizer, E., A novel function of yeast fatty acid synthase. *European Journal of Biochemistry* **2000,** *267* (9), 2666-2671.

154. Michael Lee, K.; DaSilva, N. A., Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis. *Yeast* **2005,** *22* (6), 431-440.

155. Dunn, B. J.; Khosla, C., Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *Journal of The Royal Society Interface* **2013,** *10* (85).

156. Takeno, S.; Takasaki, M.; Urabayashi, A.; Mimura, A.; Muramatsu, T.; Mitsuhashi, S.; Ikeda, M., Development of Fatty Acid-Producing Corynebacterium glutamicum Strains. *Applied and Environmental Microbiology* **2013,** *79* (21), 6776-6783.

157. Beopoulos, A.; Cescut, J.; Haddouche, R.; Uribelarrea, J.-L.; Molina-Jouve, C.; Nicaud, J.-M., Yarrowia lipolytica as a model for bio-oil production. *Progress in Lipid Research* **2009,** *48* (6), 375-387.

158. Mukherjee, K.; Bhattacharyya, S.; Peralta-Yahya, P., GPCR-Based Chemical Biosensors for Medium-Chain Fatty Acids. *ACS Synthetic Biology* **2015**.

159. Zhang, X.; McLaughlin, M.; Muñoz, R. L. P.; Hsung, R. P.; Wang, J.; Swidorski, J., Syntheses of 2-Pyrones via Electrophilic Substitutions at C7 of 4-Hydroxy-6-methyl-2-pyrone through Mono- or Dianion Formation. *Synthesis* **2007,** *2007* (EFirst), 749-753.

160. Kresze, G.-B.; Steber, L.; Oesterhelt, D.; Lynen, F., Reaction of Yeast Fatty Acid Synthetase with Iodoacetamide. *European Journal of Biochemistry* **1977,** *79* (1), 181-190.

161. Ziegenhorn, J.; Senn, M.; Bücher, T., Molar absorptivities of beta-NADH and beta-NADPH. *Clinical Chemistry* **1976,** *22* (2), 151-60.

10 Non-scientific supplementary

10.1 Danksagung

Während meiner Promotion habe ich viel Unterstützung erhalten:

Als erstes möchte ich Martin Griniger danken für die sehr gute Betreuung und die Möglichkeit mich so zu entwickeln. Du hast mich als Mentor entscheidend geprägt und es hat mir auch ziemlich Spaß gemacht.

Lisa, danke, dass du dir so viele Geschichten über Fettsäuresynthasen angehört hast und einfach da warst.

Liebe Eltern, ohne das von euch gelegte Fundament wäre ich niemals bis hier gekommen.

Auch die Mitglieder unserer Arbeitsgruppe möchte ich hier erwähnen: Manuel, mit dem ich ab dem ersten Tag mein Büro teilte. Alex, danke für die guten Tipps und die Momente, wenn wir über Reaktionsmechanismen philosophiert haben. Thank you Karthik for always offering your support without any hesitation and for having high expectations. Ines, lieben Dank für deine gute Laune, obwohl wir gerne versucht haben so viel wie möglich an dich abzugeben. Kookie, thanks for all the cake but also your enthusiasm. Und natürlich Heike für alles Administrative. Die jüngeren Mitglieder unserer Gruppe möchte ich auch nicht vergessen – Christina, Florian und Maja, es war sehr schön mit euch zu arbeiten!

Als Mitglieder anderer Forschungsgruppen möchte ich die Kooperationspartner erwähnen: Zunächst Floris Buelens und Helmut Grubmüller, die wichtige wissenschaftliche Beiträge zum ersten Projekt geliefert haben.
Und natürlich möchte ich die sehr gute Zusammenarbeit mit Eckhard Boles und Renata Pavlovic erwähnen – vielen Dank für die super Abstimmung und die immer wieder sehr motivierenden gemeinsamen Überlegungen.

10.2 Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation 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.

Frankfurt, 30. Mai 17

gez. Jan Gajewski